



---

---

UNIVERSIDAD MICHOACANA DE  
SAN NICOLÁS DE HIDALGO

Facultad de Biología  
División de estudios de posgrado  
Programa Institucional de Maestría en Ciencias Biológicas

**Utilización de un sistema de tres híbridos en  
*Saccharomyces cerevisiae* para el análisis de  
interacciones de proteínas con la región líder del ARN  
del gen sacarosa fosfato sintasa de arroz (*sps1*)**

Tesis que presenta:  
Ernesto Vázquez Chimalhua.

Como requisito para obtener el grado de  
Maestro en Ciencias Biológicas.

Director de tesis.  
Dr. Miguel Martínez Trujillo.

Morelia, Michoacán, Diciembre de 2011.



## RESUMEN

El gen de la sacarosa fosfato sintasa, SPS (E.C. 2.3.1.14), cataliza la reacción entre la UDP-glucosa y la fructosa 6 fosfato para formar sacarosa 6 fosfato que es desfosforilada por la sacarosa 6 fosfato fosfatasa alcalina (SPP E.C. 3.1.3.00) para crear sacarosa. Éste carbohidrato es el principal componente de la mayoría de las plantas para transportar productos de la asimilación del carbono del tejido fuente al consumidor, probablemente debido a su alta solubilidad, baja reactividad y capacidad de almacenar energía, también la sacarosa puede ser sintetizada del almidón acumulado o ácidos grasos de organos de reserva como la semilla. El gen *sps1*, codificador para SPS de arroz índica, tiene un 5' UTR inusualmente largo de 368 nucleótidos, se ha visto que el líder de *sps1* es capaz de potenciar la traducción, es posible que dentro del líder haya elementos de secuencia/estructura que interactúen con proteínas que potencien la traducción. Para poder indentificar *in vivo* proteínas que interaccionen específicamente con el 5' líder de *sps1* se utilizó el sistema de tres híbridos, en este método se usó una biblioteca genómica de *A. thaliana*. Así se obtuvo que la proteína Spen codificada por el gen At4g12640 interacciona con el líder de *sps1*. Mediante la herramienta BLASTp de NCBI se descubrieron los ortólogos de arroz índica del gen At4g12640 de *A. thaliana*, OJ1381\_H04.27 y P0450E05.35, estos podrían funcionar sobre el líder de *sps1* como lo hizo la proteína *A. thaliana* en el sistema de tres híbridos. Análisis de motivos de secuencias dentro de los promotores del gen At4g12640 y sus ortólogos OJ1381\_H04.27 y P0450E05.35 con la base de datos PLACE revelaron que comparten elementos para expresión específica en semilla, tejido fotosintético y polen, compatible con la expresión específica registrada para *sps1* y *OsRRM* (homólogo de OJ1381\_H04.27 en arroz japónica) en los mismo tejidos. Para observar *in vivo* el patron de expresión de At4g12640, se decidió hacer construcciones que permitieran analizar la expresión de genes reporteros generada por diferentes versiones del promotor aislado. Amplificaciones de las versiones del promotor de -490 y -981 pb desde el atg de inicio de la traducción de At4g12640 se elaboraron para ser clonadas al vector binario pKGWFS7 y poder estudiarse en plantas. En este trabajo se concluyó la construcción con la primera versión del promotor, que será analizada por transformación de plantas en trabajos posteriores.

# ÍNDICE

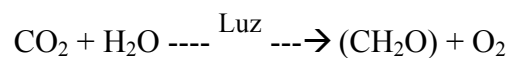
<b>I. Introducción.</b>	1
└ I.I. La fotosíntesis sucede en los Cloroplastos.	2
I.II. El Fotosistema II.	9
I.III. El complejo Citocromo B6f conecta a los fotosistemas I y II.	12
I.IV. Fotosistema I reduce NADP.	15
I.V. La ATP Sintasa y su mecanismo rotatorio.	18
I.VI. Fijación de CO <sub>2</sub> , fase oscura de la fotosíntesis y síntesis de sacarosa.	21
<b>II. Antecedentes.</b>	28
└ II.I. Características de SPS.	28
II.II. El gen de la sacarosa fosfato sintasa de <i>O. sativa</i> ssp. <i>indica</i> , <i>spsI</i> .	31
II.III. Método de tres híbridos en <i>S. cerevisiae</i> .	37
<b>III. Hipótesis.</b>	39
<b>IV. Objetivos.</b>	39
└ IV.I. Objetivo general.	39
IV.II. Objetivos particulares.	39
<b>V. Materiales y Métodos.</b>	40
└ V.I. Vectores de las moléculas en el sistema de tres híbridos.	40
V.II. <i>Saccharomyces cerevisiae</i> cepa L40-ura3.	42
V.III. Transformación secuencial de <i>S. cerevisiae</i> L40-ura3 con los plásmidos pHybLex/Zeo-MS2, pSPS-MS2 y la biblioteca de ADNc de <i>A. thaliana</i> y escrutinio para seleccionar la interacción.	43
V.IV. Ensayo de β-galactosidasa.	45
V.V. Extracción de ADN de levadura para recuperar los plásmidos de la biblioteca de ADNc de <i>Arabidopsis thaliana</i> .	45
V.VI. Selección de plásmidos de la biblioteca de ADNc extraídos de levadura.	46
V.VII. Secuenciación.	46
V.VIII. Identificación de las proteínas que generaron interacción en el sistema de tres híbridos.	46

V.IX. Identificación de las proteínas Spen de <i>A. thaliana</i> y arroz indica.	47
V.X. Alineación de la secuencia de aminoácidos entre las proteínas de arroz indica y <i>Arabidopsis thaliana</i> .	47
V.XI. Análisis de los promotores de los genes de <i>O. sativa</i> ssp. <i>indica</i> y de <i>A. thaliana</i> .	47
V.XII. Diseño de oligonucleótidos.	47
V.XIII. Extracción de ADN de <i>A. thaliana</i> y amplificaciones de las diferentes versiones del promotor del gen At4g12640.	48
V.XIV. Clonación de las diferentes versiones del promotor del gen At4g12640 al vector de entrada pCR8/GW/TOPO y secuenciación.	48
V.XV. Manejo de ADN plasmídico y transformación de células competentes.	49
V.XVI. Recombinación dirigida entre los sitios LR de los vectores de entrada y el vector destino pKGWFS7.	50
<b>VI. Resultados.</b>	51
└ VI.I. Análisis de la interacción del 5' UTR ARNm de <i>sps1</i> y proteínas codificadas en una biblioteca de ADNc de <i>A. thaliana</i> .	51
VI.II. Análisis funcional de la proteína codificada por el gen At4g12640.	56
VI.III. Aislamiento de diferentes versiones de tamaño del promotor del gen At4g12640 a partir del genoma de <i>A. thaliana</i> .	64
VI.IV. Discusión.	66
<b>VII. Bibliografía.</b>	70
<b>Anexo</b>	82

# I. INTRODUCCIÓN

Las plantas son la base de la pirámide alimentaria para los organismos heterótrofos, también son fuente de materia cruda como el almidón utilizado en las industrias o para la producción de farmacéuticos. El cultivo de plantas para la agricultura es un recurso esencial de la nutrición humana, el crecimiento de la población ha aumentado junto con la demanda de recursos naturales que en poco tiempo serán insuficientes, la inminente escasez de alimento global puede ser mitigada mediante la biotecnología. El conocimiento básico generado por las investigaciones no solo explica las funciones moleculares, bioquímicas, celulares o fisiológicas de las plantas sino también es una ciencia aplicada que aporta a la solución de grandes problemas económicos.

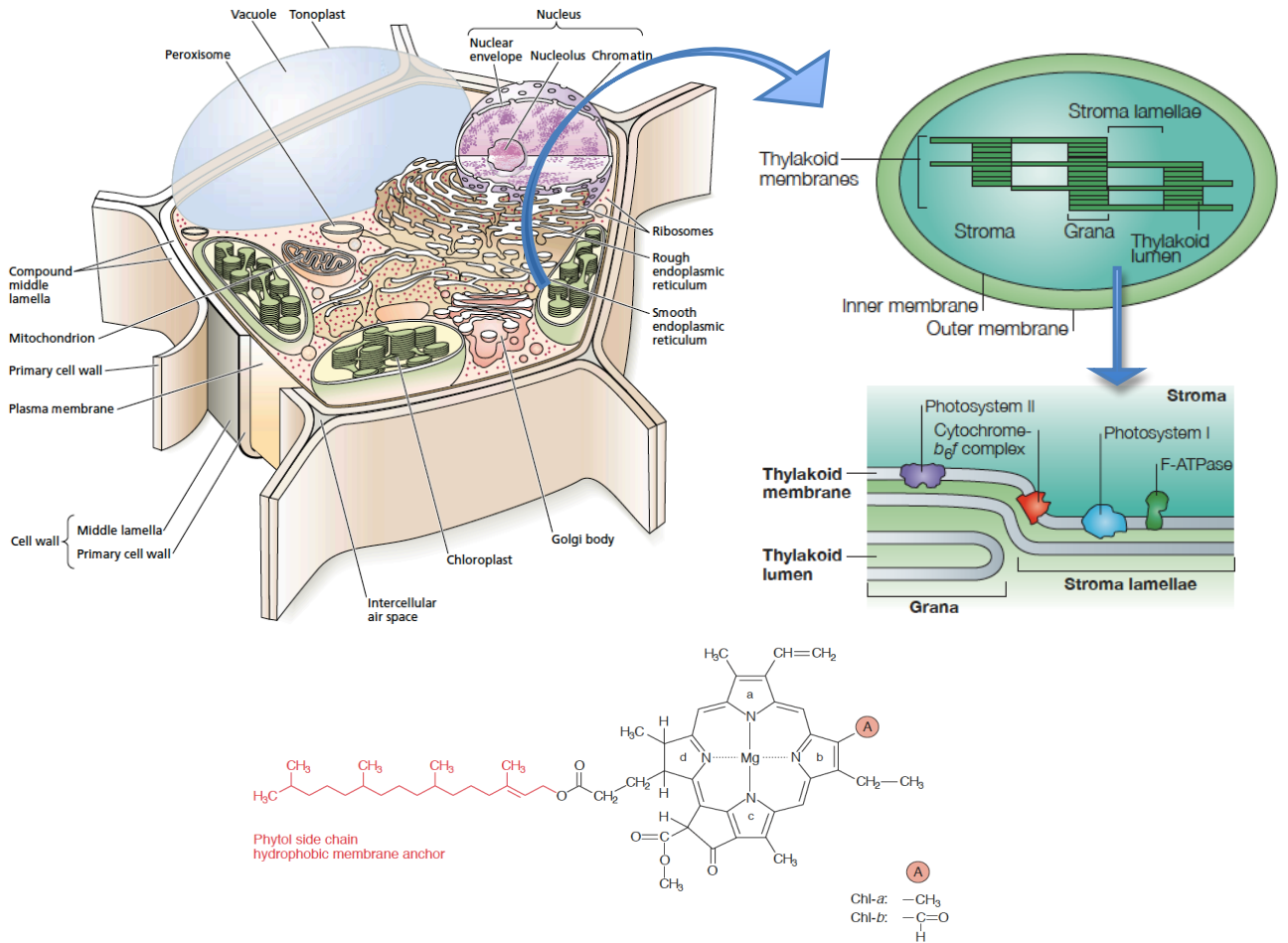
Las plantas fijan CO<sub>2</sub> y otras sustancias inorgánicas para convertirlas en compuestos orgánicos con los que forma el material útil en el crecimiento y propagación de la célula. La energía proveniente del sol es atrapada en un proceso llamado fotosíntesis, el dióxido de carbono y agua se combinan para elaborar carbohidratos (CH<sub>2</sub>O) y una molécula de oxígeno. Lo ilustra la siguiente fórmula de reacción (Taiz y Zeiger, 2002).



La fotosíntesis realizada por las plantas y cianobacterias crean la biomasa de la tierra, los animales dependemos del suplemento de carbohidratos, otros alimentos y sustancias como oxígeno o los depósitos de combustible fósil. La capacidad autótrofa de las plantas de usar CO<sub>2</sub> como única fuente de átomos de carbono requerida para la biosíntesis de almidón, sacarosa y celulosa, lípidos y proteínas, es contraria a la heterotrofia de los animales que no pueden manipular el CO<sub>2</sub> para producir su propia biomasa. Las plantas deben tener asombrosa flexibilidad metabólica capaz de adaptarse a las condiciones cambiantes del medio ambiente en el que están enraizadas y que les permiten utilizar diversas vías anabólicas que usan energía en forma de ATP y NADH o NADPH que es generado por las reacciones dependientes de la fotosíntesis para finalmente elaborar de su material celular. En las células fotosintéticas eucariotas el centro de reacción de la fotosíntesis tiene lugar en los cloroplastos (Nelson y Cox, 2008).

## **I.I. La fotosíntesis sucede en los Cloroplastos.**

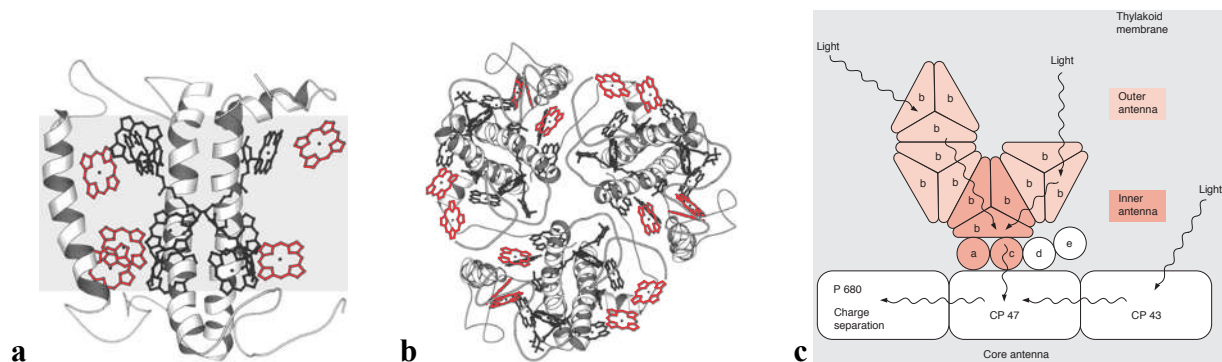
Los plástidos son una familia de organelos que solo existen en células de plantas, están delimitados por una doble membrana y contienen un pequeño genoma que codifica para algunas de sus propias proteínas. Estos se multiplican por fisión binaria y en la mayoría de los casos son heredados maternalmente a la naciente célula, es decir, que todos los plástidos descienden de los protoplástidos en la célula embrionaria. Durante la diferenciación celular en el desarrollo, los protoplástidos pueden convertirse, dependiendo su función en: Cromoplastos, leucoplastos, y cloroplastos (Heldt y Heldt, 2005). En las plantas superiores, el mesófilo es el tejido fotosintético más activo ya que tiene una gran cantidad de cloroplastos, aquí es donde ocurre tanto la absorción de la luz como la asimilación del carbono. Los cloroplastos son formados por la diferenciación de los protoplástidos, una célula madura del mesófilo puede contener hasta 50 cloroplastos, tienen forma de lente y pueden ajustar su posición dentro de la célula para recibir la mayor cantidad de luz posible, en las plantas superiores pueden llegar a medir de 3  $\mu\text{m}$  a 10  $\mu\text{m}$ . Además, tienen una doble envoltura de membranas, la membrana externa es permeable a pequeñas moléculas y iones mientras que la membrana interna encierra el estroma que contiene los tilacoides formados durante la diferenciación del cloroplasto por la invaginación de la membrana interna de la célula cloroplástica. Los tilacoides son vesículas cilíndricas rodeados de una membrana que encierran al lumen o espacio interno acuoso, diferenciado en dos dominios; uno es la agrupación en hileras de los sacos aplanados conjuntamente llamados grana, el otro dominio es la región de membrana que interconecta los grana se denomina lamela (Figura 1). Embebido en la membrana tilacoide se encuentran los pigmentos fotosintéticos y los complejos enzimáticos que llevan a cabo las reacciones luminosas y la síntesis de ATP y NADH (Figura 1). Las plantas superiores contienen tanto clorofila *a* como clorofila *b* para la captura de la luz absorbiendo entre 480 nm y 700 nm de longitud de onda, como otros cromóforos que también ayudan a la absorción de fotones. Los pigmentos de clorofila están formados por un anillo de tetrapirrol, también llamado porfirina, un átomo de magnesio (Mg) está colocado en el centro, parecido al grupo hemo encontrado en la hemoglobina, también en el cuarto anillo pirrol se une mediante enlace éster el fitol, que es una cadena de hidrocarburos con un doble enlace C-C y ramificaciones de grupos metilo ( $\text{CH}_3$ ) (Figura 1) (Heldt y Heldt, 2005; Stryer *et al.*, 2002).



**Figura 1. El cloroplasto de las plantas.** Es un organelo con doble membrana que encierra el estroma, análogo al citoplasma de la célula, los tilacoides son el tercer sistema de membranas interno donde ocurre la fotosíntesis, embebido en la membrana de los tilacoides, o lamela, se encuentran tanto los complejos proteicos como las clorofilas (*a* y *b*) necesarias para la absorción de luz y la síntesis de ATP.

Los fotosistemas localizados en la membrana tilacoide, son complejos organizados por grupos de **proteínas** que transfieren electrones unidos a las clorofilas *a* y *b*, así como pigmentos secundarios que ayudan a la absorción de luz a diferentes longitudes de onda como son los carotenos y xantofilas, todos los pigmentos pueden absorber fotones, sólo un par especial de moléculas de clorofila asociadas al centro de reacción son capaces de convertir la energía luminosa en energía química, el resto de pigmentos o moléculas antena forma complejos cosechadores de luz (LHC, Light-Harvesting Complex), se ha demostrado que el fotosistema I está relacionado con varias proteínas LHCs-I, similarmente el fotosistema II tiene diferentes LHCs-II que funcionan como antenas, los tipos de proteínas son LHC-IIa-d, sin embargo, el

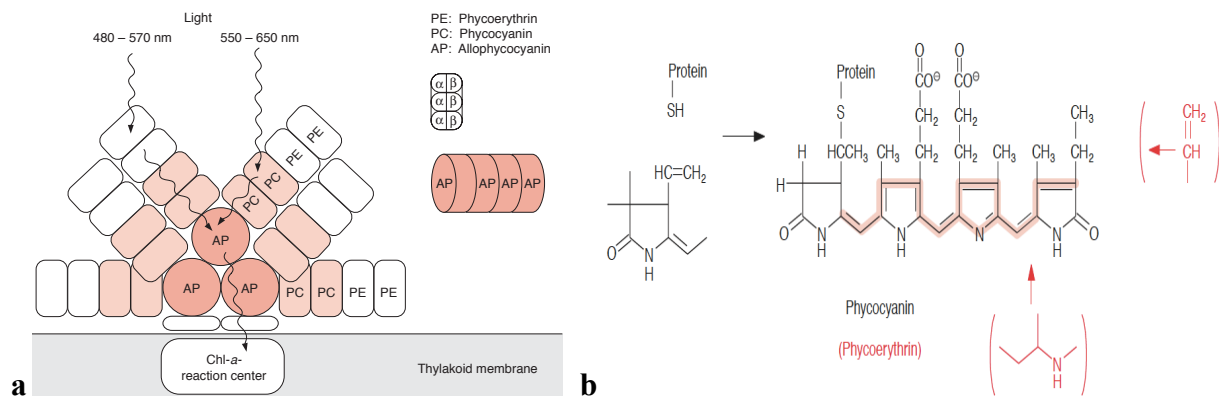
principal constituyente del fotosistema II es LHC-II*b*, el monómero es un péptido de 232 amino ácidos que forman tres segmentos  $\alpha$ -helicoidales transmembranales unidos a dos diferentes lípidos (1-Phosphatidylglycerol y 1-Digalactosyldiacylglycerol) junto con 7 moléculas de Chl-*a*, 5 moléculas de Chl-*b*, y 2 pigmentos secundarios de Luteína. LHC-II*b* existe en la membrana como trímeros dentro del mismo monómero, el excitón pasa de la clorofila *b* a la clorofila *a*, a su vez se transfiriere por contactos entre clorofilas *a* de los trímeros del complejo de antenas externo (Figura 2). Alrededor de 300 moléculas de clorofila están asociadas con cada centro de reacción, estas son las constituyentes de las antenas que funcionan para entregar electrones a los centros de reacción con los cuales están asociados, se ha pensado que la energía absorbida pasa entre los LHC hasta alcanzar el centro de reacción por transferencia de resonancia, la energía es absorbida y pasada a otra molécula por un mecanismo sin radiación (Heldt y Heldt, 2005; Taiz y Zeiger, 2002).



**Figura 2. Complejos cosechadores de luz. a).** Vista de lado del monómero LHC-II*b* en la membrana tilacoide, tres  $\alpha$ -helices expanden la membrana, clorofila *a* en negro y clorofila *b* en rojo, en el centro dos moléculas de luteína (en negro) funcionan como soporte. **b).** Vista de arriba del trímero de LHC-II desde el estroma. **c).** Esquema de los LHC en las antenas del fotosistema II de planta, LHC-II*a* es **a** y así sucesivamente. El complejo de antenas externo formado de LHC-II*b* captura excitones que son transferidos al complejo interno compuesto por algunas clorofilas que, mediante los monómeros LHC-II*a* y *c* se conectan a las proteínas CP47 y CP43.

Las algas rojas y cianobacterias poseen LHC que colectan la energía luminosa a longitudes muy bajas, esto es, ondas entre 480 nm y 650 nm. Los LHC están arreglados como partículas que van desde la parte superior de la membrana hasta cerca del centro de reacción, los llamados ficobilisomas son estructuras formadas por ficobilinas unidas covalentemente a

proteínas (ficobiliproteínas), mediante el grupo SH del residuo de cisteína de las proteínas y el grupo vinilo de la cadena de ficobilinas, ya que estas últimas son tetrapirroles de cadena abierta y estructuralmente están relacionadas a las clorofilas (Figura 3). Las proteínas ficocianina y aloficocianina se unen al cromóforo ficocianobilina, mientras que proteína ficoeritrina se une al pigmento ficoeritrobilina (Heldt y Heldt, 2005).



**Figura 3. Ficybilisomas y ficybilinas. a).** Esquema de la estructura de un ficybilisomas, vista de lado. Cada unidad constan de tres subunidades  $\alpha$  y  $\beta$ , cada una de estas puede tener de uno a cuatro ficybilinas como cromóforos. **b).** Estructura de las ficybilinas mostrando los anillos pirroles en cadena abierta, los cromóforos ficocianobilina y ficoeritrobilina se unen covalentemente vía enlace tioéter formado por la adición del grupo SH del residuo de cisteína de la proteína al grupo vinilo del cromóforo. Las conjugadas uniones dobles marcadas en rojo dan la característica de pigmento a la molécula.

En los fotosistemas, la mayoría de los pigmentos funcionan como antenas en el LHC absorbiendo fotones y excitando un electrón a un nivel de energía más alto, la energía ganada o disponible por la molécula de clorofila puede ser eliminada de formas alternativas:

1. La clorofila excitada puede re emitir el fotón y regresar a su estado energético basal, proceso conocido como **fluorescencia**, la longitud de onda emitida por la clorofila es mayor, es decir, de menor energía ya que una porción es liberada en forma de calor para entonces la clorofila fluérese en el rojo del espectro.
2. La clorofila excitada puede convertir la energía en calor directamente sin emitir fluorescencia, regresando al estado basal de energía.
3. La clorofila excitada puede transferir su energía a otra molécula de clorofila.

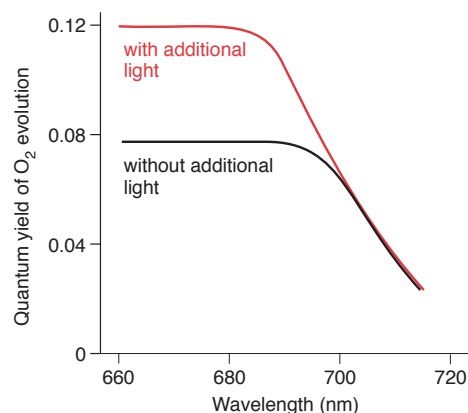
4. La energía del estado excitado de la clorofila puede ser usado para realizar reacciones químicas, proceso conocido como fotoquímico. Las reacciones fotoquímicas son de las más rápidas que se han conocido, tiene que serlo así ya que compite contra las otras tres posibilidades anteriores (Taiz y Zeiger, 2002).

El **excitón** es la energía absorbida por las moléculas de clorofila que es transferida a la clorofila vecina de los LHCs vía deslocalización de electrones promoviendo la elevación de un electrón a un orbital de energía más alto en las clorofilas, este electrón pasa por los LHCs hasta llegar al centro de reacción fotoquímico vía **Mecanismo Förster** (Heldt y Heldt, 2005). Los complejos que realizan las reacciones luminosas de la fotosíntesis en las plantas superiores no se distribuyen uniformemente por la membrana tilacoide (Figura 1), p.e. la localización del PSI (Photosystem I) es en la lamela, segregado el PSII (Photosystem II) que se haya en los dominios grana o región apilada de la membrana tilacoide, aquí encontramos en los márgenes de los grana también al complejo citocromo  $b_6 f$  (similar a citocromo  $bc_1$  mitocondrial de mamíferos, aves, levaduras y diferentes bacterias, indicando que evolucionaron de un ancestro común), y por último la F-ATPase puebla la parte de la lamela (Nelson y Ben, 2004).

Louis Duysens en 1952 encontró en membranas aisladas de la bacteria púrpura *Rhodospirillum rubrum* que al iluminarlas con pulsos cortos de luz, había una disminución en la absorción de luz a la longitud de onda específica de 890 nm, el efecto de “Blanqueamiento” o **bleaching** se debe al estado oxidado de la clorofila, es decir, a la pérdida de un electrón del centro de reacción fotoquímico, tal blanqueamiento también fue encontrado en la bacteria púrpura *Rhodobacter sphaeroides* pero a la longitud de onda de 870 nm. Años más tarde, utilizando tales técnicas Bessel Kok (USA) descubrió el mismo blanqueamiento a 700 nm correspondiente al fotosistema I en cloroplastos de plantas, paralelamente Horst Witt (Alemania) encontró un blanqueamiento del fotosistema II a la longitud de onda de 680 nm estudiando igualmente cloroplastos de plantas. En la bacteria sulfurosa verde el blanqueamiento es a la longitud de 840 nm, éste tiene fotosistema con dos rutas para el manejo del flujo de electrones, una es el flujo cíclico del electrón (usado por las bacterias púrpuras) y el otro en un flujo no cíclico. La estructura y función del fotosistema II en cloroplastos de

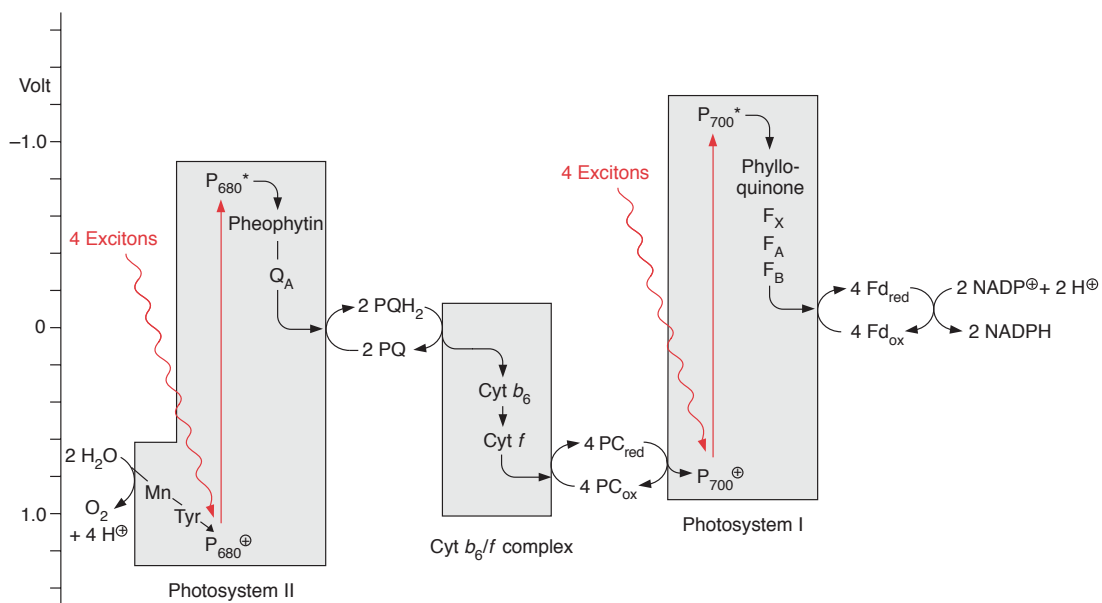
plantas corresponde al centro de reacción de la bacteria púrpura, mientras que fotosistema I de plantas en análogo al centro de reacción de la bacteria sulfurosa verde (Taiz y Zeiger 2002; Heldt y Heldt 2005; Nelson y Cox, 2008).

En la etapa tardía de 1950, Robert Emerson y colaboradores estudiaron la producción de O<sub>2</sub> por fotón absorbido o “**quantum yield**”, termino reciproco a “**quantum requirement**” que trata de los fotones absorbidos por molécula de O<sub>2</sub> producida, y revelaron que en el **espectro de acción** (que marca la respuesta producida de la luz, tal como la formación de oxígeno en función de la longitud de onda ) la producción de oxígeno cae drásticamente cuando la alga (*Chlorella*) es iluminada con longitudes de onda mayores a 680 nm, efecto conocido como **red drop**, indicando que la luz mas allá del rojo es ineficiente en el manejo de la fotosíntesis. Posteriormente Emerson y colaboradores resolvieron el problema al observar que al iluminar el alga con longitudes de onda mayor a 680 nm junto con luz roja de 680 nm la producción de oxígeno por fotón o quantum yield aumentaba drásticamente, en comparación a la suma de la producción de O<sub>2</sub> usando por separado luz roja y luz mas allá del rojo (Figura 4). Estas observaciones posteriormente ayudaron al descubrimiento de dos nuevos complejos fotosintéticos que funcionan en tándem, el fotosistema I y el fotosistema II (Nelson y Cox 2008; Heldt y Heldt, 2005; Taiz y Zeiger, 2002; Lodish *et al*, 2008).



**Figura 4. Funcionamiento en tándem de fotosistemas.** La producción de O<sub>2</sub> es mayor cuando la alga *Chlorella* es irradiada con luz adicional a 680 nm (luz roja) que la suma de la producción de O<sub>2</sub> por separado de luz roja y luz mas allá de rojo (línea negra). Llevando a la conclusión que la fotosíntesis es realizada por dos fotosistemas que trabajan en tándem pero con diferencias ligeras en la longitud de onda optima (de Emerson and Rabinowitch).

En 1960 Robert Hill y Fay Bendal postularon el **esquema Z** (Figura 5), llamado así por el potencial redox en forma de zigzag creado por el flujo de electrones a través de cada uno de los dos diferentes fotosistemas (I y II) localizados en la membrana tilacoide, cada fotosistema tiene su propio centro de reacción y LHC o antena diferente uno del otro, el fotosistema I es manejado a longitudes de onda de 700 nm o menos; el fotosistema II solo por longitudes de onda más cortas a 680 nm, los fotosistemas están conectados entre si por la **cadena transportadora de electrones** constituida por el complejo citocromo  $b_6f$  (complejo citocromo  $bc_1$  para bacterias) (Heldt y Heldt, 2005).



**Figura 5. Cadena transportadora de electrones.** Esquema en forma de Z del potencial redox por el flujo de electrones que pasa por el fotosistemas II, el citocromo  $b_6f$  y el fotosistema I para finalmente guardar la energía en productos estables, NADPH y ATP. Pulso cortos ( $5 \mu s$ ) de luz ( $<680 \text{ nm}$ ) activan el fotosistema II, e indica que la absorción de cuatro fotones (excitones) de luz son necesarios para generar una molécula de  $O_2$  a partir de  $2H_2O$ . Los electrones ganados en el desarrollo de  $O_2$  y  $4H^+$  son almacenados y utilizados en la formación de NADPH.

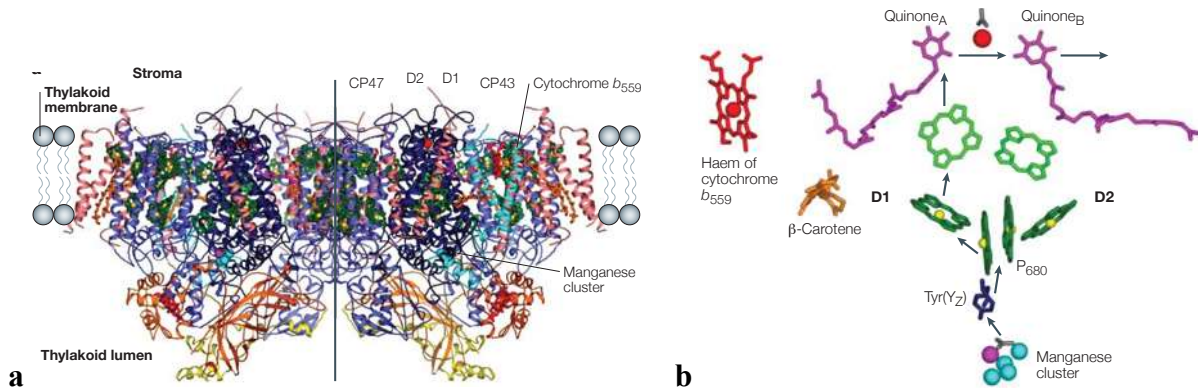
El aparato fotosintético de las plantas, algas y cianobacterias es más complejo que el único centro de reacción que poseen las bacterias, y parece que han evolucionado a través de la combinación de los fotocentros simples bacterianos. Los dos fotosistemas funcionan complementariamente que, junto con el complejo citocromo  $b_6f$  (Electron transport chain) canalizan los electrones hacia la formación de un NADPH y ATP generado por el gradiente de

protones. Hay una diferencia en el potencial eléctrico de 1.2 volts entre la oxidación del agua y la reducción del NADP (Figura 5), los dos fotosistemas juntos generan una diferencia en el potencial eléctrico de 3.45, así que, se pierde un tercio de energía (Heldt y Heldt, 2005).

## I.II. El Fotosistema II

Recientemente, Horst Witt y Wolfgang Saenger (ambos en Berlin) han revelado la estructura tridimensional del PS II con rayos X analizando cristales a 3.8 Å en la cianobacteria *Synechococcus elongatus*, mediante tal trabajo han concluido que los fotosistemas de plantas están contruidos por los mismos principios básico que en las bacterias. Con esto, y análisis en la secuencia genómica indican que los fotosistemas tienen un origen en común. El fotosistema II es un dímero de dos centros de reacción completos junto con sus respectivos complejos antena (Figura 6), el centro de reacción está constituido por dos proteínas similares de ~40 KDa, D1 y D2, de los cuales cada subunidad consiste de 5 helicasas transmembranales que, son homólogos a las proteínas L y M del centro de reacción de la bacteria púrpura. Flanqueando el centro de reacción, las subunidades CP43 y CP47 (Chlorophyll Protein) funcionan como proteínas cosechadoras de luz, cada una consiste de 6 helicasas transmembranales y unen 14 y 16 moléculas de clorofila respectivamente. La energía transferida del LHC hacia CP43/CP47 es mediada por proteínas cosechadoras de luz menores, CP24 (Lhcb6), CP26 (Lhcb5) y CP29 (Lhcb4). El fotosistema II cataliza la oxidación del agua y provee la mayor parte del oxígeno en la tierra, en el centro de reacción de éste fotosistema las proteínas D1 y D2 coordinan tanto el complejo productor de oxígeno de manganeso como los componentes del flujo de electrones, que son; un par especial de clorofilas *a*, en *Rp viridis* los anillos tetrapirroles están tan cercanos (0.3 nm) que sus orbitas se traslapan en estado excitado, la transferencia de electrones procede por el brazo L, en plantas el par especial está más alejado o débilmente unido que, probablemente sólo una molécula de clorofila reaccione con el excitón, cada clorofila del par especial está unida a las proteínas D1 y D2 formando dos brazos donde se unen cromóforos, en cada brazo se yace una molécula accesoria de clorofila *a* (en bacterias, como en plantas, la clorofila *a* accesoria del brazo M está en contacto con un β caroteno para abolir el estado triple de la clorofila en el centro de reacción), y una molécula de feofitina que, es una clorofila pero en lugar del átomo de magnesio en el centro hay dos

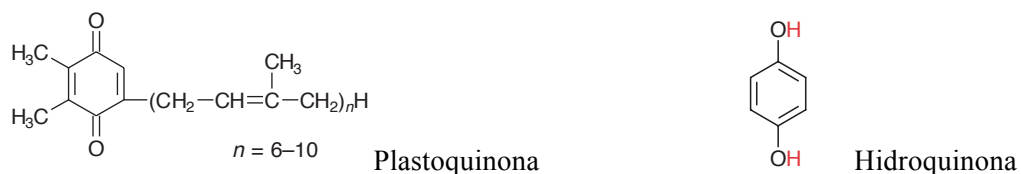
átomos de hidrógeno. La plastoquinona  $PQ_A$  se une fuertemente a la proteína D1 (en *Rp. viridis* el anillo de quinona  $Q_A$  va unido por puentes de hidrógeno e interacciones hidrofóbicas a la subunidad M), mientras que la plastoquinona  $PQ_B$  es móvil. Adicionalmente en el centro de reacción hay un átomo de hierro sin grupo hemo, entre  $P_{680}$  y el complejo de manganeso un aminoácido tirosina funciona como acarreador de electrones, se le conoce como  $Y_Z$ , y el citocromo  $b_{559}$  quizá esté involucrado en algún ciclo protector del PSII (Figura 6) (Ferreira *et al.*, 2004; Stryer *et al.*, 2002; Nelson y Cox, 2008)



**Figura 6. Vista perpendicular del dímero del PSII desde la membrana. a)** Subunidades D1 y D2 mostradas como estructura de cinta azul oscuro, CP43 y CP47 son las cintas moradas, subunidad citocromo- $b_{559}$  representado por la estructura en cinta azul claro y las subunidades extrínsecas por las cintas naranja y amarillo, las subunidades restantes son las cintas rosa claro, las clorofilas son mostradas en estructuras de bastones verde oscuro (Ion Mg como esfera amarilla), el complejo productor de oxígeno (grupo 4Mn) es simulado por las esferas azul claro y los iones de calcio por las esferas rosa oscuro, el grupo hemo es interpretado por bastones rojos y las esferas rojas son iones de hierro. Los bastones color naranja simula  $\beta$  caroteno y los bastones rosa oscuro son las quinonas. El eje de simetría del dímero está trazado por una línea. **b).** La flecha indica el flujo del electrón, Tirosina 161 ( $Y_Z$ ) en bastones azul oscuro y en bastones grises los iones bicarbonato, feofitina en verde claro y las clorofilas representadas por porfirinas simétricas. PDB ID 1S5L.

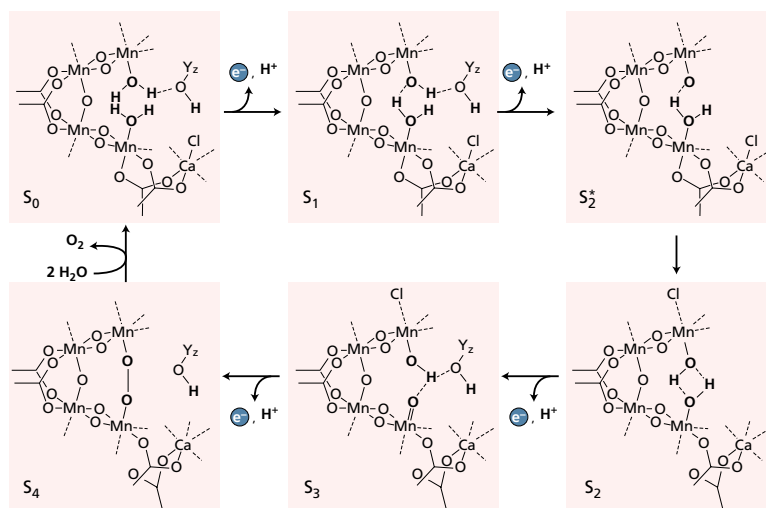
En el mismo centro de reacción se localizan los primeros **donadores de electrones** que llevan a cabo el primer proceso redox de la fotosíntesis, son el par especial de clorofilas  $a$  llamadas  $P_{680}$ , por la máxima longitud de onda a la que absorben luz, elevando un electrón a una orbita de mayor energía, tal molécula de clorofila  $P_{680}$  excitada transfiere el electrón al monómero de clorofila  $a$  accesoria e inmediatamente al **aceptor de electrones** feofitina quedando la clorofila del centro de reacción oxidado (carga positiva por la pérdida de un electrón) y el

aceptor del electrónes reducido (gana un electrón y la carga es negativa), en bacterias la diferencia de cargas ocurre en menos de 10 picosegundos,  $10^{-11}$  segundos. De la feofitina, el electrón viaja al aceptor (a  $10\text{\AA}$  de distancia en bacterias) inmovil  $\text{PQ}_A$ , suprimiendo así la *recombinación de cargas* que sucede cuando el electrón en la  $\text{Ph}^-$  regresa a  $\text{Chl}^+$  del centro de reacción para neutralizar la carga, en *R. viridis* la  $\text{Q}_A$  pasa el electrón al Fe sin hemo y posteriormente a la  $\text{Q}_B$  en  $6\mu\text{s}$ , es decir, reduce la móvil  $\text{PQ}_B$  a  $\text{PQ}_B^-$ , plastosemiquinona radical, la absorción de un segundo fotón o quantum por el centro de reacción mueve el segundo electrón a través de la cadena transportadora hasta llegar a reducir completamente a  $\text{PQ}_B^-$  en  $\text{PQ}_B^{2-}$  ya que, el sitio de unión de  $\text{PQ}_B$  yace del lado del estroma, toma dos protones del estroma para formar plastohidroquinona  $\text{PQH}_2$  que, es liberada del fotosistema II y considerada como el producto final (Nelson y Ben, 2004, Heldt y Heldt, 2005, Stryer *et al* 2002, Okamura *et al* 2000, Nelso y Cox, 2008).



La clorofila  $\text{P}_{680}^+$  oxidada tiene el potencial redox mas elevado dentro de los sistema biológicos ( $>1\text{V}$ ) (Figura 5), es un oxidante muy fuerte que extrae un electrón a la cercana tirosina ( $\text{Y}_Z$ ) ubicada en la proteína D1, la tirosina radical (Z) con déficit de un electrón oxida un ion manganeso de un grupo de 4 Mn iones, los cuales unen dos moléculas de agua como substrato y utiliza un ion calcio, un ion bicarbonato y un ion cloruro como cofactores. El evento anterior se repite dos veces para forma  $\text{PQH}_2$ , con cada transferencia individual de electrones el grupo 4 Mn se vuelve más oxidado, cuatro fotones absorbidos corresponden a una cuatro transferencia de electrones generando una carga de  $4+$  en el complejo productor de oxígeno (**oxigen-evolving complex**) oxidan  $2\text{H}_2\text{O} \rightarrow \text{O}_2 + 4\text{H}$  (Figura 7), formando oxígeno molecular y protones que son liberados en el lumen de la membrana tilacoide, actuando el complejo como una bomba de protones (Tommos y Babcock, 1999), además es el único complejo biológico conocido que produce  $\text{O}_2$ . El complejo productor de oxígeno está asociado con una proteína periférica membranal ( $M_r$  33,000) sobre el lado del lumen de la membrana tilacoide. La extracción de electrones puede ser impar, p.e. en bajas temperaturas,  $\text{P}_{680}^+$  bajo estas condiciones puede extraer  $e^-$  de su alrededor y causar daños, la estructura del PSII

sugiere una extracción alternativa de  $e^-$  donde se involucran el citocromo  $b_{559}$ , un carotenoides y una clorofila periférica (Figura 6) (Nelson y Ben, 2004; Heldt y Heldt, 2005; Stryer *et al.*, 2002; Okamura *et al.*, 2000; Nelson y Cox, 2008).

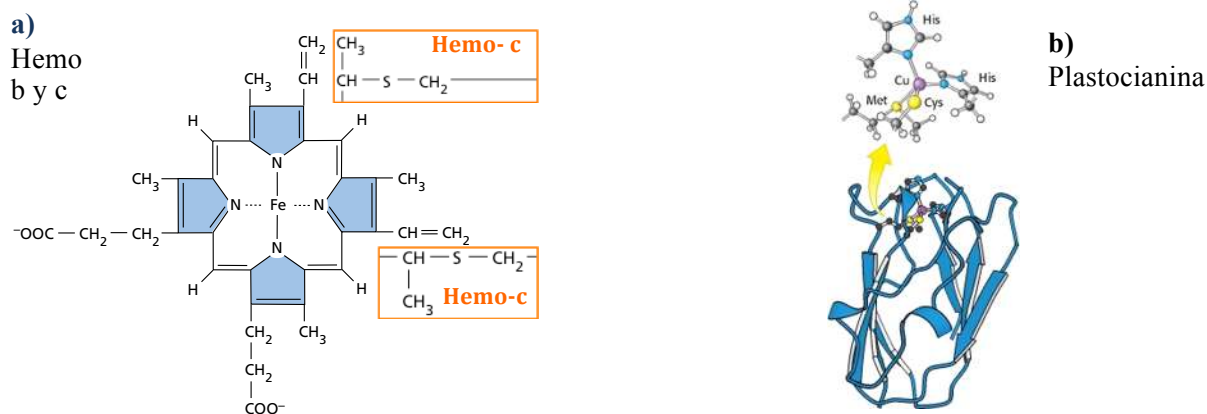


**Figura 7. Complejos productores de oxígeno.** Con experimentos realizados por Pierre Joliot y Bessel Kok se ha postulado que el complejo productor de oxígeno tiene diferentes estados de oxidación, conocidos como *estados S* y van del S<sub>0</sub> al S<sub>4</sub> hasta oxidar 2H<sub>2</sub>O. La tirosina es un aceptor de electrones entre en complejo de iones Mn y P<sub>680</sub>.

### I.III. El complejo Citocromo $B_6f$ conecta a los fotosistemas I y II.

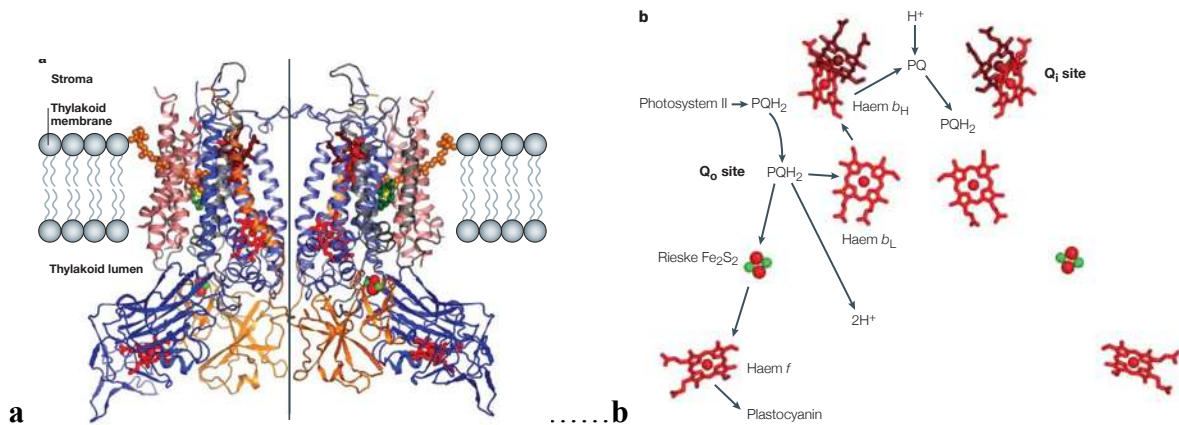
PQH<sub>2</sub> se disocia por la porción de hidrocarburos de la membrana hasta llegar al complejo Citocromo  $B_6f$  (Figura 5). Las estructuras de cristales del citocromo  $B_6f$  que se encuentra en la membrana tilacoide indican que funciona como homodímero, cada monómero consiste de cuatro subunidades: proteína citocromo- $b_6$  (23 KDa) integra dos grupos hemo  $b$ , proteína citocromo- $f$  (de foliar, 33 KDa) tiene un grupo hemo  $c$ , proteína hierro-azufre Rieske (Mr 20,000) y una subunidad IV (17 KDa), también cuatro pequeñas subunidades hidrofóbicas PetG, PerL, PetM, y Petn, dan al dímero un peso molecular de 217 KDa (Kurisu *et al.*, 2003; Stroebel *et al.*, 2003) y dos co-factores con función desconocida, una molécula de clorofila  $a$  y de  $\beta$  caroteno (Figura 9), conforman el complejo. Citocromo- $b_6$  expande la membrana como lo hace la subunidad IV, incluso ambos son homólogos en las mitades N-terminal y C-terminal, respectivamente, al complejo cit- $b_{Cl}$ , Los Citocromos son proteínas que unen de uno a dos

anillos tetrapirroles con átomos de Hierro en el centro, conocido como grupo hemo (en plantas el anillo tetrapirrol tiene un átomo de Mg), los citocromos funcionan como acarreadores de electrones, en comparación con las quinonas que transfieren dos  $e^-$  junto con dos protones. Estos están divididos en tres grupos *a*, *b*, y *c*, corresponde a los grupos hemos *a*, *b*, y *c*, en el cit-*b* el grupo hemo-*b* no está covalentemente unido a la proteína, en contraste con el grupo hemo-*c* unido covalentemente al cit-*c* mediante enlaces tioéter, donde se unen el SH<sub>2</sub> del residuo de cisteína y dos grupos vinilo del hemo-*b*, (Figura 8), el citocromo *c* del complejo citocromo *B<sub>6</sub>f* no está relacionado con el citocromo *c*<sub>1</sub> del complejo citocromo *b*<sub>C1</sub>, el restante grupo hemo *a* tiene una fila de tres unidades de isopreno unidas a un grupo vinilo del hemo *b*, la fila de isoprenos emula una ancla en la parte hidrofóbica de la membrana igual que en las quinonas. Acerca de la proteína hierro-azufre Rieske del complejo citocromo *B<sub>6</sub>f* es muy similar a la del complejo citocromo *b*<sub>C1</sub>, mientras que PetG, PerL, PetM, y Petn no tienen homólogos en el complejo *b*<sub>C1</sub>, el hemo *c* funcional del cit- *f* yace en la periferia de la membrana del lado luminal, del mismo lado está la proteína hierro-azufre Rieske que se ancla ligeramente a la membrana. El electrón pasa a través del complejo Citocromo *B<sub>6</sub>f* como la describe el **ciclo Q**, PQH<sub>2</sub> es oxidado al unirse al sitio Q<sub>0</sub> del lado luminal liberando dos protones al mismo lado, uno de los dos electrones va por el proceso cíclico que, incrementa (pmf) el número de protones bombeados transversalmente por la membrana, durante que el otro electrón sigue un ciclo lineal hasta llegar a la plastocianina. Durante el transporte lineal de electrones, PQH<sub>2</sub> es oxidado en el sitio Q<sub>0</sub> del lado del lumen por la proteína hierro-azufre Rieske, debido al potencial redox muy positivo de la proteína Rieske un electrón es quitado de la plastohidroquinona que se reduce a PQ<sup>-</sup>, a su vez libera dos protones al lado acuoso del lumen, este  $e^-$  es pasado al cit-*f* y enseguida a la plastocianina que, es una proteína (m m de 10.5 kDa) con un átomo de Cu (cobre), el cual está unido a un residuo de cisteína, uno de metionina, y dos de histidina de la misma proteína (Figura 8), la cual sucesivamente reduce las clorofilas P<sub>700</sub> del PSI (Carell *et al.*, 1997; Martinez *et al.*, 1994; Nelson y Ben, 2004).



**Figura 8. Citocromos y plastocianina. a).** Figura del grupo prostético del citocromo *b* y *c*. El hemo-*b* no está unido covalentemente al cit-*b* (azul), así el hemo-*c* (naranja) tiene similar estructura al hemo-*b*, pero está unido covalentemente a residuos de cisteína de la proteína cit-*c*.  $\text{Fe}^{2+}$  es el estado de oxidación en citocromos reducidos y  $\text{Fe}^{3+}$  el estado de oxidación en citocromos oxidados. **b).** Estructura de la plastocianina, un residuo de cisteína, uno de metionina y dos de histidina unen un átomo de Cu en un arreglo de tetraedro distorsionado que, cambia su estado de  $\text{Cu}^+$  a  $\text{Cu}^{2+}$ .

En el vía cíclica, la plastosemiquinona radical  $\text{PQ}^-$  producida es muy inestable y pasa el electrón entre la membrana al primer grupo hemo  $b_L$  del citocromo- $b_6$  y de ahí en el mismo cit- $b_6$  al segundo hemo  $b_H$  luego, reduce una oxidada plastoquinona en el sitio  $Q_1$  del lado del estroma, para producir un radical plastosemiquinona  $\text{PQ}^-$ , seguido de una segunda reducción toma dos protones del estroma para formar  $\text{PQH}_2$  que, se libera al pool de quinonas reducidas de la bicapa lipídica. En el sitio  $Q_0$ , cuatro protones son translocados al lumen por dos moléculas de  $\text{PQH}_2$  oxidadas, de las PQ hechas una retorna al PSII y la otra quinona se difunde por la capa de lípidos de la membrana para unirse a otro sitio  $Q_1$  donde es reducido por el grupo hemo  $b_H$  via  $\text{PQ}^-$  radical a  $\text{PQH}_2$ , la hidroquinona regenerada es oxidada por la proteína Rieske del lado del lumen (Figura 9), y así gradualmente (Carell *et al.*, 1997; Martinez *et al.*, 1994).

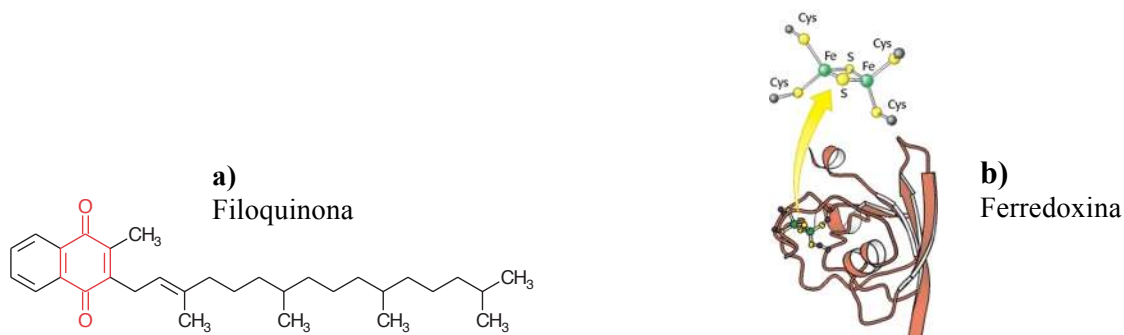


**Figura 9. Complejo citocromo-*B<sub>6</sub>f* de *Chlamydomonas reinhardtii* y los cofactores implicados en el transporte de electrones PDB ID 1Q90. a).** Vista perpendicular del complejo homodimérico Citocromo-*B<sub>6</sub>f* desde la membrana, el eje de simetría del dímero está resaltado por una línea. El citocromo-*b<sub>6</sub>* está mostrado con estructuras de listón moradas, la subunidad IV es la estructura de listón gris, las proteína hierro-azufre mostrada como listón naranja, y la subunidad citocromo-*f* es mostrada como estructura de listón azul oscuro. Las subunidades pequeñas PetG, PerL, PetM, y Petn son las estructuras de listón rosa ligero. Los grupos hemos-*b<sub>L</sub>*, *b<sub>H</sub>* y *f* son representados por estructuras de bastones rojos (las esferas rojas centrales simulan al ión de hierro) y los grupos hemo extras son los bastones rojo oscuro localizados entre el hemo-*b<sub>H</sub>* y el sitio Q<sub>1</sub>. Los bastones naranjas representan  $\beta$ -caroteno y las estructuras de bastones rosa oscuro representan las quinonas. **b).** Los cofactores posicionados como en la figura de a lado, y las flechas indican la transferencia de electrones en el ciclo Q de la oxireducción de la plastoquinona (PQ) y la plastocianina (Stroebel *et al.*, 2003).

#### I.IV. Fotosistema I reduce NADP

La soluble plastocianina reducida proveniente del citocromo *B<sub>6</sub>f* viaja a través del lumen hasta unirse a un sitio positivamente cargado del fotosistema I (PSI), dona un electrón y se difunde de regreso al citocromo *B<sub>6</sub>f*. El PSI está compuesto de dos partes; el centro de reacción y el complejo periférico cosechador de luz I (LHC-I), el PSI también contiene un par de clorofilas *a* (Chl<sub>a</sub>)<sub>2</sub> especiales que absorben la luz a 700 nm de longitud de onda como máximo, en contraste con el PSII, hay alrededor 100 clorofilas antenas ligadas intrínsecamente a dos grande proteínas (PsaA y PsaB) en el centro de reacción, el cual consta de 12-14 subunidades proteicas a saber; PsaA-PsaL, PsaN y PsaO. Los heterodímeros PsaA y PsaB como en D1 y D2 del PSII, fincan el corazón del PSI uniendo al par especial de chl-*a* P<sub>700</sub>, elceptor

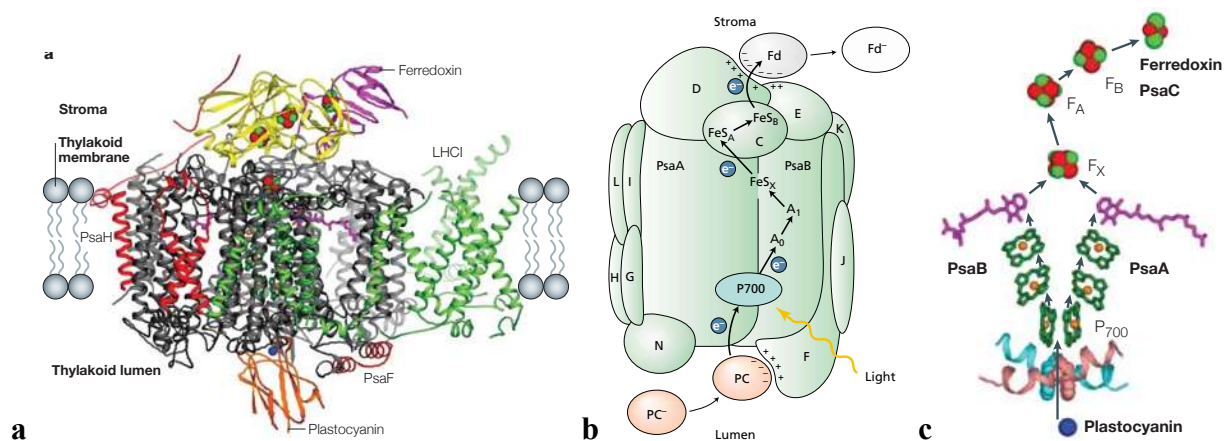
primario de electrón  $A_0$  (chl-*a*),  $A_1$  (filoquinona, figura 10) y una serie de tres proteínas membranales con un grupo de  $Fe_4-S_4$  conocidas como  $F_X$ ,  $F_A$  y  $F_B$  (unidos a PsaC, los dos últimos). La excitación por el fotón o el excitón en la  $P_{700}$  provoca una separación de cargas  $(Chla)_2^+$  que es luego reducida por la plastocianina al momento de empalmarse a PsaF, el electrón es transferido inmediatamente de  $P_{700}$  a  $A_0$  (monómero de Chl-*a*) y de aquí a una filoquinona  $A_1$  fuertemente unida (su función corresponde a  $Q_A$  del PSII), de la filoquinona radical el electrón va hacia el centro  $F_X$  y sucesivamente a los centros  $F_A$  y luego a  $F_B$ , en orden que, el electrón llega a la **Ferredoxina** (Figura 10 y 11), proteína  $Fe_2S_2$  de 11 KDa cercana al lado estromal de la membrana tilacoide, esta proteína acarrea solo un electrón a la vez y se encuentra ligada a la subunidad PsaD, PsaC y PsaE, la ferredoxina reduce a la flavoproteína asociada a la membrana tilacoidal **FNR** (ferredoxina  $NADP^+$  reductasa) aceptando un electrón a la ves, generando un flavino semiquinona, la proteína recibe un segundo electrón para formar  $FNRH_2$ , el cual transfiere dos electrones y un protón a  $NADP^+$  para crear NADPH, produciendo NADPH como producto final del transporte fotosintético no cíclico de electrones (Binda *et al.*, 1998; Xue *et al.*, 1998).



**Figura 10. Estructura de Filoquinona y Ferredoxina.** a) La filoquinona contiene la misma cadena fitol que la Clorofila *a* y su función corresponde al de la Quinona A ( $Q_A$ ) del PSII. La ferredoxina (b) es una proteína que contiene un grupo  $Fe_2S_2$  coordinado por un cuatro residuos de cisteína, el grupo  $Fe_2S_2$  puede acarrear electrones a la Ferredoxina  $NADP^+$  reductasa (FNR), así complementando el flujo no cíclico de electrones que empezó con la oxidación del  $H_2O$

El centro reacción del PSI en plantas usa un cinturón extrínseco de antenas cosechadoras de luz (LHCI) en el que la organización modular es de cuatro proteínas con clorofilas cosechadoras de luz (Lhca1-Lhca4), los cuatro Lhca se juntan en dos dímeros (Lhca1-4 y Lhca2-3) para formar una medialuna o cinturón que sobresale de la subunidad PsaF, la

asociación con el centro de reacción es asimétrico en el sentido que, solo el monómero Lhca1 a través de una unión hélice se une a PsaG, los restantes tres monómeros a causa que, no envuelven enlaces helice-helice interaccionan débilmente con el centro de reacción en los dominios lumenales y estromales y entre los mismos monómeros. De tal manera que, Lhca1 o el dímero Lhca1-Lhca4 actúa como ancla para el ensamblaje y estabilidad de los Lhca. Los dominios en el extremo C-terminales de PsaA y PsaB del PSI muestran el mismo arreglo helicoidal de D1 y D2 del PSII, similarmente, el extremo N-terminal de PsaA y PsaB es parecido en secuencia, estructura y pigmentos a las subunidades cosechadoras de luz intrínsecas CP47 y CP43 del PSII. Quizá, la característica más destacada de ambos fotosistemas son los componentes de que transfieren el electrón y su arreglo en dos brazos, en el cual, un par de cofactores está relacionado por un eje de simetría perpendicular a la membrana, esto deja claro que ambos fotosistemas provienen de un ancestro en común. Mientras que el PSII produce un potencial de oxidación muy fuerte en cualquier sistema biológico, el PSI se encarga de suplementar la célula con equivalentes oxidantes fuertes (Ben *et al.*, 2004; Ben *et al.*, 2003)



**Figura 11. Modelo del PSI y los cofactores involucrados en el transporte de electrones.** a) Vista de lado del PSI, las coordenadas estructurales del PSI fueron tomadas de PDB 1QZV, 1AG6 y 1A70. El centro de reacción comprende de 12 a 14 subunidades PsaA-PsaL, PsaN y PsaO. Las subunidades estromales son mostradas en estructuras de listón amarillas, las novedosas subunidades no encontradas en la contraparte de las cianobacterias están en listones rojos, entretanto que, las estructuras conservadas son los listones grises, el complejo LHCI (Lhca1-4) son los listones verdes. Los componentes del flujo de electrones del PSI son; en estructura de bastones verde oscuro las clorofilas con el ión Mg central como esfera naranja, las filoquinonas son los bastones morados y los grupos de  $Fe_4S_4$  son resaltados como esferas verdes y rojas, la plastocianina es simulada por estructura de listones

aranja con su ión cobre como esfera azul y la ferredoxina es el listón morado y su grupo  $\text{Fe}_2\text{S}_2$  son las esferas verdes y rojas. **b)** Dibujo de las subunidades del centro de reacción del PSI y los procesos en los que se involucran. **c)** Cofactores implicados en la conducción de electrones del PSI, posicionados como se indicó en la parte **a**, dos residuos de triptófano que pueden estar implicados en el paso del electrón de la plastocianina a  $\text{P}_{700}$  son presentados como listones rosa claro y azul claro dentro del contexto de su ambiente estructural secundario.

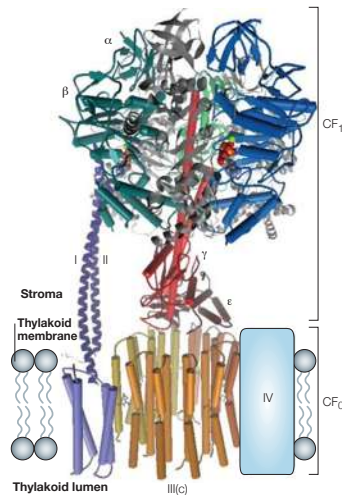
#### **I.V. La ATP Sintasa y su mecanismo rotatorio.**

Una parte de la energía luminosa capturada es almacenada como NADPH al viajar el electrón desde la molécula de agua. Sincronizadamente, los protones son bombeados al interior (lumen) de la membrana tilacoide conservando la energía en un potencial electroquímico, éste gradiente de protones provee la energía necesaria para la síntesis de ATP dependiente de luz a partir de ADP y un fosfato inorgánico ( $\text{P}_i$ ) por la ATPasa o ATP sintasa, evento conocido como **fotofosforilación**. Proceso descubierto por Daniel Arnon en 1954, en condiciones normales la fotofosforilación es llevada a cabo por un flujo de electrones, sin embargo, en condiciones controladas la fotofosforilación puede llevarse de manera independiente del flujo de electrones, conocido como flujo desacoplado. Ahora se sabe que, la fotofosforilación funciona vía mecanismo quimiosmótico, propuesto por Peter Mitchell en 1961 (Edinburgo), tal hipótesis menciona que el flujo de electrones acoplado a la fotofosforilación genera un gradiente electroquímico de protones, llamado Fuerza Protón Motriz (**PMF** por sus siglas en inglés o  $\Delta p$ ) que suplementa la energía para la síntesis de ATP. El principio básico de la **quimiosmosis** es la diferencia en la concentración de iones y la diferencia de potencial eléctrico a través de la membrana son fuentes de energía libre disponibles para célula, según la segunda ley de la termodinámica cualquier distribución de energía o materia no uniforme es fuente de energía. En 1966 André Jagendorf y colaboradores presentaron evidencia conclusiva que apoya la hipótesis quimiosmótica de la fotofosforilación, donde los cloroplastos sintetizan ATP en la oscuridad cuando se imponen a un gradiente electroquímico de pH a través de la membrana tilacoide, ellos suspendieron cloroplastos en un buffer ácido de pH 4 con el objetivo de acidificar el lumen tilacoidal, agregaron fosfato inorgánico ( $\text{P}_i$ ) y ADP a la suspensión y rápidamente la mezcla fue transferida a un buffer alcalino de pH 8, así generando una diferencia de pH de 4 unidades en el lumen ácido con respecto al exterior estromal,

provocando una ráfaga de síntesis de ATP acompañado con la desaparición del gradiente de pH a través de la membrana tilacoide (Taiz y Zeiger, 2002; Heldt y Heldt, 2005; Yoshida *et al.*, 2001).

De tal suerte que, la energía guardada en protones del proceso luminoso es útil para sintetizar ATP por la o ATPasa o complejo CF<sub>1</sub>-CF<sub>0</sub> (C es por cloroplasto y F de Factor). El complejo enzimático se encuentra tanto en la membrana interna mitocondrial, en la membrana plasmática en bacterias como en la membrana tilacoidal de cloroplastos, esta enzima se ha mantenido constante durante la evolución con  $\geq 60\%$  de su secuencia aminoacídica conservada en la subunidad catalítica  $\beta$ . La ATPasa es un complejo consistente de una parte integrada a la membrana tilacoide llamada F<sub>0</sub> que canaliza los protones pasantes por la membrana, unido a la porción estromal F<sub>1</sub> donde se haya el lugar catalítico encargado de la formación de ATP a partir de los substratos ADP y P<sub>i</sub>, utilizando la energía de PMF. Todas las F<sub>1</sub>-ATPasa están estructuradas por una parte hidrofílica CF<sub>1</sub> y CF<sub>0</sub> es la parte hidrofóbica. Las subunidades únicas menores  $\epsilon$ ,  $\gamma$  y  $\delta$  construyen el tallo central de F<sub>1</sub> conectado al cilindro hecho de alternadas tres copias de las subunidades mayores  $\alpha$  y  $\beta$  ( $\alpha_3\beta_3$ ) circundantes a la subunidad  $\gamma$ . El complejo hidrofóbico F<sub>0</sub> de cloroplastos, se conforma de las subunidades I, II, III, y IV en una posible disposición estequiométrica de 1:2:14:1 respectivamente, igual corresponde a las subunidades *a*, *b*, y *c* de mitocondria, mientras que la subunidad IV es similar en secuencia a la subunidad *a*. La subunidad IIIc contiene dos hélices transmembranales apareciendo en conjunto de 14 para formar un anillo que expande la membrana, se cree que, en la zona hidrofóbica de la membrana tilacoide la subunidad Ia se asocia con la superficie exterior del anillo de proteínas IIIc, además las dos subunidades F<sub>0</sub>IIb se asocian con F<sub>0</sub>Ia y F<sub>1</sub> $\delta$  como lo muestra la figura 12 (Boyer, 1989; Stock *et al.*, 2000) .

Análisis estructurales con cristalografía de rayos x confirman un arreglo del subcomplejo F<sub>1</sub> $\alpha\beta$  para formar una sola cámara con sitio de unión a un nucleótido de adenina, de esta manera se forman tres cámaras de reacción representando los diferentes momentos de la síntesis de ATP, una cámara contiene ADP ( $\beta_{DP}$ ), otra cámara lleva un análogo de ATP ( $\beta_{TP}$ ) (5' adenililimidodifosfato, AMP-PNP difiere en que los últimos dos residuos de fosfato están conectados



**Figura 12. Modelo de la ATPasa de cloroplasto.** Los dominios  $F_1$  y  $F_0$  también se conectan por un segundo tallo periférico sirviendo quizá como estabilizador al contrarrestar la tendencia de  $\alpha\beta$  de seguir el giro del tallo central  $F_1$ , erigido por las dos idénticas subunidades  $F_0IIb$  (morado), ambas interactúan en su dominio transmembranal (una  $\alpha$  helice cada uno) con la proteína de  $F_0Ia$  (morado) y la proteína  $\delta$  del dominio  $F_1$ , la subunidad  $\delta$  no es mostrada porque no se sabe su localización exacta, pero se piense que se conecta con las dos subunidades  $IIb$  de  $F_0$  y el subconjunto  $\alpha\beta$  de  $F_1$ . En rojo la proteína  $\gamma$  y en café la subunidad  $\epsilon$ , aunque no se conoce mucho de la subunidad  $IV$  ha sido agregada esquemáticamente, tanto el subconjunto  $\alpha\beta$  mostrado en verde, como el subconjunto  $\alpha\beta$  gris y azul simulan los diferentes estados de cambios en el sitio catalizador de ATP.

por un átomo N) y la restante cámara está vacante ( $\beta_E$ ). La subunidad  $\gamma$  es una estructura larga y enrollada en forma de espiral asimétrica que sobresale de entre  $F_1$  y se dobla hacia el lado de la cámara  $\alpha\beta$  con ADP ( $\beta_{DP}$ ). Según Boyer el movimiento rotatorio de  $\gamma$  fuerza al dominio del extremo c-terminal de la subunidad  $\beta$  a girar  $\sim 30^\circ$  alterando la estructura del sitio de unión al nucleótido de adenina de la subunidad  $\beta$ , estos cambios se alternan hasta completar un ciclo y liberar ATP, puede existir la acción contrario de hidrolizar el ATP induciendo la rotación de gama en el sentido opuesto, con esto se bombean protones en la dirección contraria. La rotación de  $\gamma$  y  $\epsilon$ , tallo central  $F_1$ , se conectada con la rotación del anillo de subunidades  $IIIc$  físicamente unidos al tallo de  $F_1$ , así formando el motor rotatorio de la ATPasa (Boyer, 1989; Stock *et al.*, 2000; Yoshida *et al.*, 2001).

Cada subunidad IIIc tiene dos  $\alpha$ -hélices que expanden la membrana, justamente se une el anillo en la parte hidrofóbica de la membrana la proteína Ia, teniendo consigo dos medios canales para protones que no expanden la membrana y por donde pueden pasar los protones pero no traspasar la membrana. Un medio canal esta del lado del lumen tilacoidal y el otro medio canal del lado del estroma. De los dos canales del prótido Ia cada cual interactúa con una proteína IIIc del anillo, en la mitad de la segunda hélice de alguna proteína IIIc se encuentra un residuo de aspartato (Asp 61), el Asp 61 dentro del canal luminal está en su forma ácida neutral teniendo contacto con la parte hidrofílica del lumen, en esta parte, el residuo Asp 61 es protonado para neutralizar su carga y ahora en anillo IIIc puede rotar pero solo en dirección de las manecillas del reloj, de esta manera el protón puede dar la vuelta al anillo y llegar al medio canal de la proteína Ia que esta conectado al lado estromal donde el protón se disocia del ácido aspártico y se mueve por el medio canal hacia el estroma, aquí la concentración de protones es menor y puede de esta manera ser restaurada. La rotación por el gradiente de protones del anillo de subunidades IIIc causa la rotación de la fuertemente unida proteína  $\gamma$ , a la vez el giro de  $\gamma$  origina un cambio en la conformación del sitio de unión al nucleótido de adenina en la subunidad  $\beta$ , el cual sucesivamente promueve la síntesis de ATP (Stryer *et al.*, 2002).

#### **I.VI. Fijación de CO<sub>2</sub>, fase oscura de la fotosíntesis y síntesis de sacarosa.**

Las plantas pueden sintetizar carbohidratos a partir de agua e incorporación de CO<sub>2</sub> usando el poder reductante de NADPH y la energía de ATP, generados por la oxidación de agua en las reacciones dependientes de luz de la fotosíntesis. Plantas y otros organismos autótrofos pueden usar CO<sub>2</sub> como única fuente de átomos de carbono para la producción de celulosa, almidón, proteínas, lípidos, ácidos nucleicos y otros compuestos orgánicos. Las reacciones que se llevan a cabo para la reducción de CO<sub>2</sub> son realizadas por enzimas en el estroma del cloroplasto, provistas de la energía del ATP y NADPH. Los organismos autótrofos fijan el CO<sub>2</sub> bajo el mismo mecanismo, entre 1946 y 1953 Melvin Calvin, Andrew Benson y James Bassham, en Berkeley, California, descubrieron el sistema por el cual se fija en CO<sub>2</sub>, las investigaciones fueron realizadas en la alga verde *Chlorella*, posteriores descubrimientos resolvieron que la fijación de CO<sub>2</sub> procede cíclicamente por lo que también es llamado **ciclo**

**de Calvin**, el ciclo de Calvin comprende tres estados, a saber:

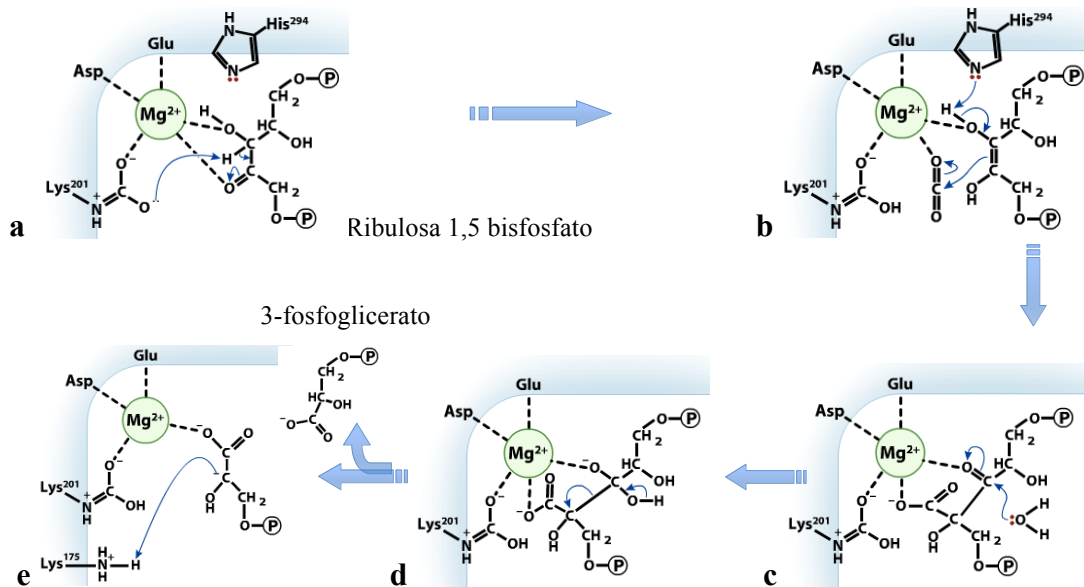
1er estado. *Carboxilación* del aceptor **ribulosa 1,5 bisfosfato** para dar origen a dos moléculas de 3-fosfoglicerato, primer intermediario estable.

2do estado. *Reducción* del intermediario 3-fosfoglicerato al carbohidrato gliceraldehido-3-fosfato (triosa fosfato).

3er estado. *Regeneración* del aceptor ribulosa 1,5 bisfosfato de las triosas fosfato formadas (Taiz y Zeiger, 2002).

El CO<sub>2</sub> es el compuesto más oxidado en la naturaleza (+4), el carbono del primer intermediario 3-fosfoglicerato es más reducido (+3), y la siguiente forma en gliceraldehido-3-fosfato es aún más reducido (+1), la reducción del carbono atmosférico facilita su incorporación a componentes orgánicos. La *carboxilación* sucede en el estroma, la condensación de CO<sub>2</sub> a la ribulosa 1-5 bisfosfato es realizada por la enzima **ribulosa 1,5 bisfosfato carboxilasa/oxigenasa** o **rubisco**. Existen dos tipos de rubisco, la forma I se encuentra en plantas vasculares, algas y cianobacterias, la forma II se presenta sólo en algunas bacterias y es más simple a causa que se forma de dos subunidades que semejan a dos subunidades grandes de la enzima de planta. La forma I de la enzima rubisco está elaborado de ocho subunidades grandes idénticas (M<sub>r</sub> 53,000; codificadas por el genoma del plástido) cada una contiene un sitio catalítico y ocho subunidades pequeñas idénticas (M<sub>r</sub> 14,000; codificados por el genoma nuclear) con función incierta, al parecer estabilizan el complejo enzimático pero no son necesarias para la carboxilación. Todas las subunidades grandes de rubisco contienen un residuo de lisina en el aminoácido 201 de ~470 aminoácidos, la rubisco activasa facilita la reacción del grupo amino ε de la lisina 201 con CO<sub>2</sub> (éste CO<sub>2</sub> es diferente al que se fija a ribulosa 1,5 bisfosfato) para formar **carbamato** (amida de ácido carbónico), incluso la rubisco requiere un cofactor para la activación, en esta caso Mg (magnesio) se une a la lisina carbamoilada y orienta los reactantes al sitio catalítico, la activación se debe a cambios en la conformación de la proteína (figura 13). La Ribulosa 1,5 bisfosfato se une al ión Mg mediante su grupo ceto del 2do carbón y el grupo hidroxilo adyacente del 3er carbón, complejo que fácilmente desprotona e isomeriza de ceto a enol a la ribulosa 1,5 bisfosfato produciendo un enediol, a su vez, el ión Mg polariza CO<sub>2</sub> atmosférico y lo fija con un ataque nucleofílico al carbono 2 del enediol generando 2-carboxi 3-ceto arabinitol 1,5 bisfosfato o β-cetoácido. Éste

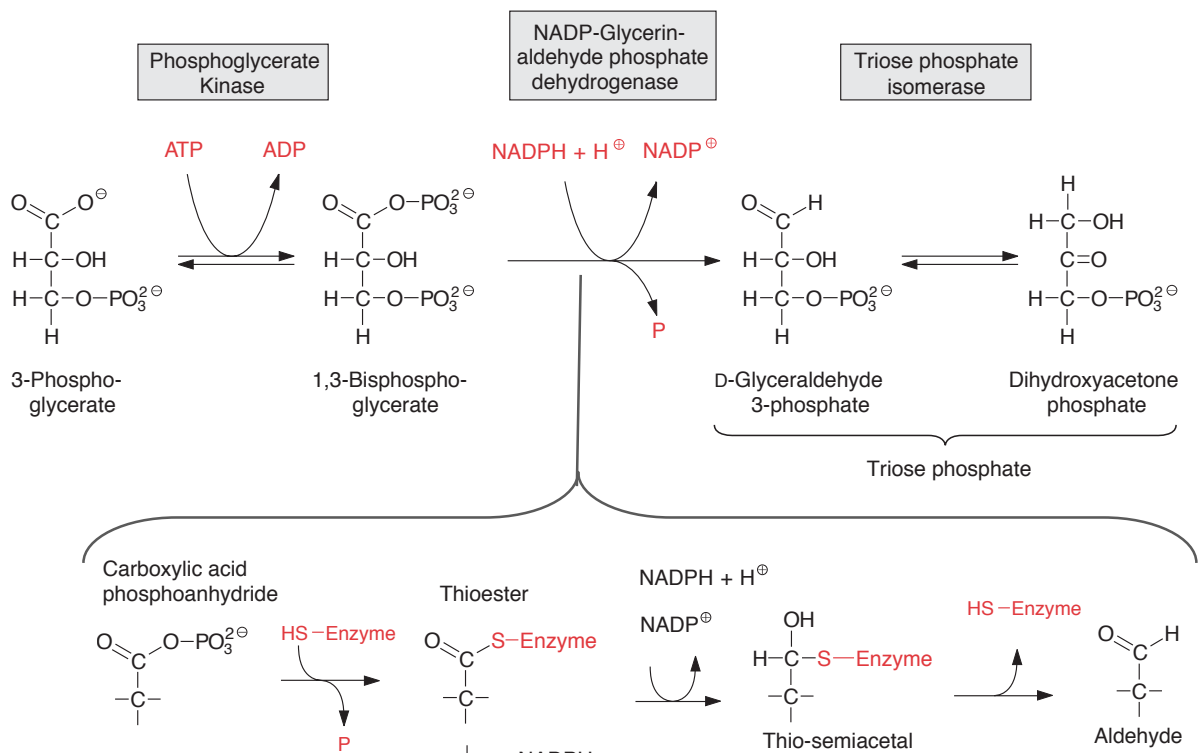
intermediario de 6 carbonos es hidrolizado enseguida al agregar una molécula de H<sub>2</sub>O en el carbono 3 y descompuesto en dos moléculas de 3 carbonos, **3-fosfoglicerato**, terminando el primer paso del ciclo, de fijar CO<sub>2</sub> (Heldt y Heldt, 2005).



**Figura 13. Fijación de CO<sub>2</sub> al aceptor ribulosa 1,5 bisfosfato, primer paso del ciclo de Calvin.** La fijación es realizada por la ribulosa 1,5 bisfosfato carboxilasa/oxigenasa, rubisco. El sitio activo de la rubisco yace en la subunidad larga donde se liga un ión Mg mediante los residuos de aspartato, glutamato y la lisina carbamoilada de la enzima, La ribulosa 1,5 bisfosfato se une al el ión Mg, aquí el aceptor dona un protón y es convertido en enediol, en este momento el CO<sub>2</sub> polarizado al ión Mg se fija al aceptor, construyendo un azúcar intermediario, una molécula de H<sub>2</sub>O hidroliza al azúcar ramificado en el C-3 rompiéndola en dos la molécula 3-fosfoglicerato. Uno es liberado al instante mientras el otro es protonado por el residuo de lisina 175 para su liberación.

El producto de la carboxilación y descomposición de ribulosa 1,5 bisfosfato a 3-fosfoglicerato pasa al segundo paso del ciclo, *la reducción*, al ser fosforilado por la 3-fosfoglicerato cinasa con el uso de ATP generado en las reacciones luminosas produciendo 1,3 bisfosfoglicerato, éste enseguida es reducido a gliceraldehido 3-fosfato y P<sub>i</sub> por la enzima NADP-gliceraldehido-fosfato-deshidrogenasa, primero cambiando el residuo fosfato del grupo carboxilo por un grupo SH del residuo de cisteína en el centro activo de la enzima y así formar un tioester, luego un NADPH de las reacción luminosa reduce al tioester para dar **gliceraldehido 3-fosfato** (triosa fosfato), la triosa fosfato isomeraza interconvierte de gliceraldehido 3-fosfato a

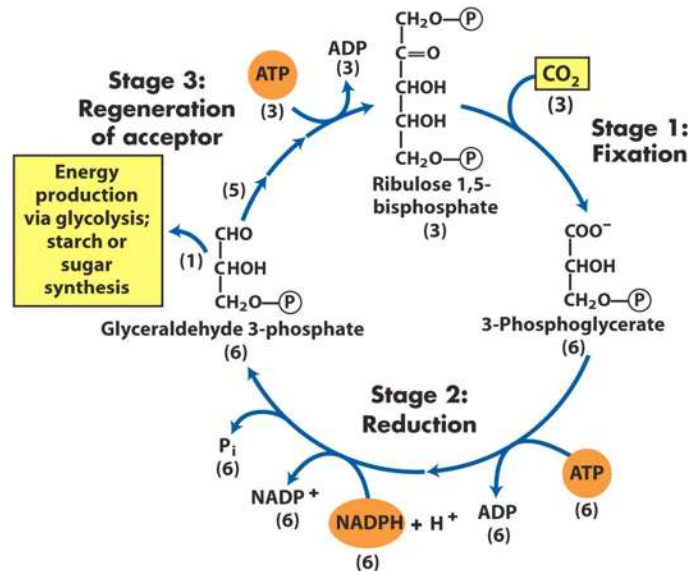
**dihidroxiacetona fosfato** (figura 14). La mayoría de las de las triosas fosfatos son recicladas para regenera al aceptor ribulosa 1,5 bisfosfato y comenzar de nuevo, el resto de las triosas fosfato pueden ser guardadas para futuro uso como almidón en el cloroplasto o para uso inmediato en el citosol tanto en formación de sacarosa para transportarse a tejidos no fotosintéticos como en la síntesis de hexosas para energía y material de construcción celular (Nelson y Cox, 2008).



**Figura 14. Triosas fosfato.** En la reducción, la enzima 3-fosfoglicerato cinasa transfiere un grupo fosforilo del ATP a 3-fosfoglicerato produciendo 1,3 bisfosfoglicerato, el anterior es reducido por la enzima gliceraldehído fosfato deshidrogenasa con el protón donado por NADPH catalizando a la triosa Gliceraldehído 3-fosfato que es sustrato de la enzima triosa fosfato isomerasa interconvirtiéndolo a la triosa dihidroxiacetona fosfato, la isomerasa mantiene a estas dos triosas en equilibrio.

En el tercer paso del ciclo es la *regeneración del aceptor ribulosa 1,5 bisfosfato* con los gliceraldehído 3-fosfato. De la fijación de tres moléculas de CO<sub>2</sub> con tres aceptores, se producen seis moléculas de triosas fosfato (dieciocho carbonos), cinco (quince carbonos) de seis triosas se reutilizan para regenerar tres ribulosa 1.5 bisfosfato, la ganancia neta es una triosa (tres carbonos) fosfato de seis que, será usada para formar almidón o sacarosa. La

reconstrucción del aceptor ribulosa 1,5 bisfosfato incluye la producción de varios azúcares intermediarios de cuatro, cinco, seis, y siete carbonos, el último paso en ésta serie de conversiones requiere un ATP per ribulosa 1,5 bisfosfato o un total de tres ATP (Nelson y Cox, 2008). De esta manera, se completa el ciclo de Calvin como lo ilustra la figura 15.



**Figura 15. Ciclo de Calvin.** Estequiométricamente, cada una de las tres moléculas de  $\text{CO}_2$  es fijada a una ribulosa 1,5 bisfosfato, eventualmente, producen seis moléculas de 3-fosfoglicerato, sólo un gliceraldehido 3-fosfato de tres carbonos es el aprovechamiento neto, mientras que las otras cinco partículas de triosas fosfato son recicladas en la generación del aceptor. Durante el ciclo, nueve ATP y seis NADPH provenientes de la fotosíntesis son consumidos.

La ribulosa 1,5-bisfosfato tiene actividad de oxigenasa, proceso nombrado **fotorespiración**, ya que el  $\text{CO}_2$  no es sustrato específico de la enzima, sino que, el oxígeno molecular ( $\text{O}_2$ ) compite por el sitio activo de la enzima tanto que, en uno de tres o cuatro ciclos de Calvin es fijado una molécula de  $\text{O}_2$  obteniendo un hipotético peróxido, intermediario que al agregar una molécula de agua se hidroliza en dos partes; una en gliceraldehido 3-fosfato y otra parte de carbonos llamada **2-fosfoglicolato**, un producto metabólicamente inútil. La fijación de  $\text{O}_2$  al aceptor no genera alguna ganancia neta de carbohidratos y el 2-fosfoglicolato entra al ciclo  $\text{C}_2$  para producir  $\text{CO}_2$ , implicando a 2 organelos más, peroxisomas y mitocondria. El 2-fosfoglicolato en el cloroplasto es rápidamente desfosforilado a glicolato por una fosfatasa específica, posteriormente es transportado al peroxisoma. El rescate del 2-fosfoglicolato es

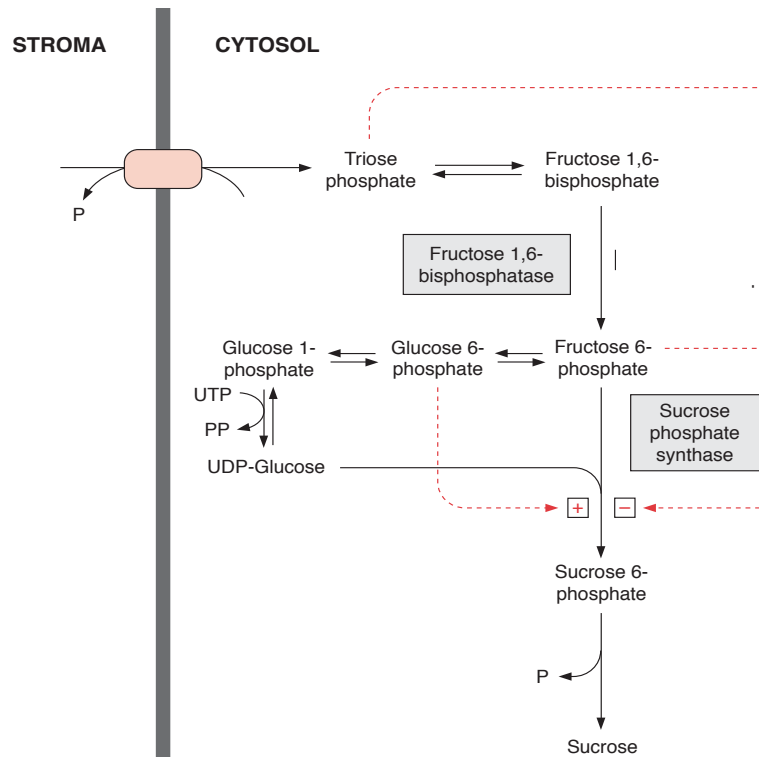
energéticamente costoso al haber más de diez reacciones enzimáticas entre los tres organelos (Stryer *et al.*, 2002).

En el estroma del cloroplasto de algunas de las plantas los carbohidratos son convertidos en almidón como reserva alimenticia temporal, en tanto que en la mayoría de las plantas las triosas fosfato preceden la síntesis de sacarosa, carbohidrato que es distribuido a otras partes de la planta como las raíces o algún lugar en desarrollo p.e. en la caña o el betabel. En la formación de almidón, gliceraldehído 3-fosfato y dihidroxiacetona fosfato se encuentran en equilibrio y son sustrato de la aldolasa fructosa 1,6 bisfosfato, la enzima da origen a *fructosa 1,6 bisfosfato*, en orden para ser desfosforilado por la fructosa 1,6 bisfosfatasa, liberando un  $P_i$  y una molécula de fructosa 6 fosfato, enseguida la última molécula es convertida en una glucosa 6 fosfato vía hexosa fosfato isomerasa, después la fosfoglucomutasa cambia el fosfato de la carbono 6 al 1, mutando la molécula a glucosa 1 fosfato, la hexosa 1 fosfato es activada condensándose con ATP vía ADP-glucosa pirofosforilasa a ADP-glucosa y liberando un pirofosfato que es inmediatamente hidrolizado a 2  $P_i$  por la pirofosfatasa. Sucesivamente, el residuo de glucosa de la ADP-glucosa es transferido vía *almidón sintasa* hacia la preexistente cadena de glucano creciente oxidando el grupo hidroxilo del no reducido carbono 4 terminal, resultando así en un enlace glicosídico  $\alpha$ 1-4. La amilosa consiste de aproximadamente 1000 moléculas de glucosa unidas, aún más grande es la amilopectina con  $10^4$  hasta  $10^5$  moléculas de glucosa unidas, además cada 20 o 25 residuos de glucosa hay ramificaciones con enlaces glicosídicos  $\alpha$ 1-6 a otros residuos de glucosa. Los enlaces  $\alpha$ 1-4 del almidón pueden ser rotos por la acción de la enzima  $\alpha$ -amilasa durante la germinación (Stryer *et al.*, 2002).

Por otra parte, algunas triosas fosfato en intercambio con fosfatos son exportadas por un transportador específico del estroma del cloroplasto al citosol celular, lugar en el que se sintetiza la sacarosa, disacárido consistente de glucosa y fructosa. La ruta metabólica en la que se forma la sacarosa es similar a la ruta en la síntesis de almidón, esto es con la utilización de isozimas (múltiples formas de enzimas con la misma función catalítica pero con diferente secuencia de aminoácidos, diferente afinidad por sustratos y otras propiedades), por la vía de la fructosa 1,6 bisfosfato y glucosa 1 fosfato con la diferencia que esta última es activada con UTP vía UDP-glucosa pirofosforilasa, enzima análoga a ADP-glucosa pirofosforilasa estromal

del cloroplasto, liberando UDP-glucosa y un pirofosfato que es hidrolizado a 2 P<sub>i</sub> (Heldt y Heldt, 2005).

La sacarosa 6 fosfato sintasa (SPS), figura 16, cataliza la transferencia del residuo glucosa de la Uridina Difosfato-glucosa a la fructosa 6 fosfato para elaborar sacarosa 6 fosfato enlazado mediante glucosa α1-β2 fructosa, incluso SPS integra como complejo a la sacarosa fosfato fosfatasa para desfosforilar la sacarosa 6 fosfato y crear sacarosa. La forma para transportar los carbohidratos es la sacarosa, p.e. en la producción de la celulosa la sacarosa sintasa (actúa reversamente a SPS) usa sacarosa citosólica y UDP para hacer fructosa y UDP-glucosa, la glucosa del último actúa como precursor en la síntesis de celulosa (Heldt y Heldt, 2005).



**Figura 16. Actividad de la sacarosa fosfato sintasa o SPS.** La enzima SPS favorece la reacción entre el carbono α1 de la glucosa de la UDP-glucosa y en carbono β2 de la fructosa, tanto la fructosa como la glucosa provienen de las triosas fosfato translocadas del estroma del cloroplasto al citosol celular, en esta parte de la célula las isozimas trabajan como en el estroma, condensando las triosas y formando fructosa 1,6 bisfosfato, así sucesivamente hasta formar glucosa 1 fosfato, la isozima UDP-glucosa pirofosforilasa condensa esta última hexosa con UTP en UDP-glucosa y libera PP<sub>i</sub>.

## II. ANTECEDENTES

EL gen sacarosa fosfato sintasa (SPS E.C. 2.3.1.14 ) actúa en el citoplasma de la célula sintetizando la reacción entre la UDP-glucosa y la fructosa 6 fosfato, tal glicosiltransferasa, posiciona el grupo glucosídico de la uridina difosfato glucosa a la F6P, formando así al disacárido sacarosa 6 fosfato, este es desfosforilado al instante por la enzima hidrolasa sacarosa 6 fosfato fosfatasa alcalina (SPP E.C. 3.1.3.00), con el cual forma complejo SPS, para crear sacarosa que puede ser transportada a los lugares no fotosintéticos de la planta y ser usado como fuente energética o ser precursor de la síntesis de compuestos como la celulosa. La actividad de SPS en hojas frecuentemente se correlaciona con los niveles de síntesis de sacarosa y exportación (Geigenberger *et al.*, 1999). La clonación de SPS, para su estudio, ha sido realizada en varias especies de plantas; primero fue la enzima de maíz (Worrel *et al.*, 1991), después de espinaca (Sonnewald *et al.*, 1993, Klein *et al.*, 1993), luego de papa (Sonnewald y Basner 1993), posteriormente de betabel (Hesse *et al.*, 1995) y de arroz (Valdez-Alarcón *et al.*, 1996).

### II.I. Características de SPS

Esta enzima en hojas de plantas puede ser regulada a diferentes niveles jerárquicos como en la regulación de la transcripción del gen, pero es principalmente controlada a nivel enzima, en oscuridad es inactivada o menos activa mediante la fosforilación covalente reversible en el grupo OH- del residuo de serina 158 vía sacarosa fosfato sintasa cinasa (SPSk) y para activarla es desfosforilada por la sacarosa fosfato sintasa fosfatasa (SPS-PP tipo 2A), la iluminación aumenta la diligencia de SPS-PP probablemente por la síntesis *de novo*; también la unión no covalente o alostérica del metabolito glucosa 6 fosfato (G6P) activa la enzima sacarosa fosfato sintasa e inhibe a SPSk. La afinidad de la enzima por los substratos UDP-G y F6P es modificada, SPS fosforilado muestra menos afinidad por los substratos como por su efector alostérico, SPS desfosforilado muestra mayor afinidad para el efector G6P como hacia los substratos y menor afinidad por Pi (inhibidor). La enzima SPS es aún mas compleja, se ha estudiado en tubérculos de papas y hojas de espinaca en condiciones hiperosmóticas y han resuelto que ésta enzima es activada al ser fosforilada en el grupo OH- del residuo de serina en la posición 424 por la cinasa SPSk dependiente de calcio activada por estrés osmótico, ésta fosforilación podría actuar antagónicamente a la fosforilación inhibitoria del residuo 158 de

serina conjuntamente con un decrecimiento del 90% en almidón y tres veces más sacarosa, relacionado con la acumulación del disacárido en reacción a estrés de agua y e incluso a las bajas temperaturas (Huber y Huber, 1992; Toroser y Huber, 1997; Geigenberger *et al.*, 1999).

En plantas superiores y en respuesta al estrés osmótico, la reducción del potencial de agua celular es controlado por la síntesis de osmolitos como el azúcar, facilitando la absorción de agua proveniente del ambiente y dirigiendo la repartición de carbono hacia la sacarosa (Toroser y Huber, 1997), aunando a lo anterior, se cree que SPS de planta regula osmóticamente la presión de la turgencia intracelular en eucariontes (Huynh *et al.*, 2004).

Al parecer, entre las especies discrepantes de SPS solo existen diferencias cuantitativas y no cualitativas, es decir que, la enzima específicamente actúa en la síntesis de sacarosa. Hay algunas desigualdades sobre el control de SPS, algunas especies presentan una marcada activación de SPS en presencia de luz (nombradas como clase I), mientras que la clase II es sensible a metabolitos bajo condiciones de saturación o limitación, por último SPS clase III no reacciona con la luz pero su actividad es influenciada por fosforilación (Huber y Huber, 1996).

Las variantes propiedades de SPS permiten distribuir el flujo de carbono o triosas P entre la síntesis de sacarosa y la de almidón en respuesta a señales de luz/oscuridad o cambios en la demanda de sacarosa. Actualmente, el rol en la repartición del carbono de SPS en hojas está bien entendido y ha sido ampliamente estudiado, sin embargo, falta comprender el papel de SPS en la distribución de carbono en la planta entera. En planta de tomate (*Lycopersicon esculentum*, dicotiledónea) fue expresado en gen SPS de maíz (*Zea mays*, monocotiledónea) encontrando que, la distribución de carbono fue mayor en tallo que en raíz, hubo por lo tanto, un incremento en el radio del peso de la masa seca del tallo correlacionado con el incremento de actividad de SPS en hojas y un decrecimiento en el radio del peso de la materia seca, no correlacionado con tal incremento de actividad SPS en raíz (Galtier *et al.*, 1993). Experimentos realizados con SPS de maíz en plantas transgénicas de tomate realizados por Laporte *et al.* (2001), demostraron que, actividad excesiva de SPS provoca un decrecimiento en el rango de síntesis de aminoácidos que podría explicar la respuesta no lineal del crecimiento de la planta, efecto del cambio de repartición de carbono entre el almidón,

sacarosa y compuestos iónicos (principalmente aminoácidos) y además, sólo una expresión de aproximadamente lo doble resulta en un crecimiento óptimo y producción de frutos.

En *Nicotiana tabacum* se sobre expresaron tres genes del metabolismo de la sacarosa, UDP-glucosa pirofosforilasa UGPasa, sacarosa fosfato sintasa SPS y sacarosa sintasa SuSy, creando inicialmente creando plantas con un único transgen, entre las cuales fueron sucesivamente cruzadas para engendrar líneas transgénicas portadoras con dos y tres transgenes; la manipulación de estos tres genes no tuvo efecto en la producción de celulosa, por su parte todas la líneas tuvieron incrementos en el crecimiento de la altura, esto es congruente con la sobre expresión individual de alguno de estos genes en experimentos previos realizados por Park *et al.*, (2008) y Coleman *et al.*, (2006) presentando un incremento del 15% en la altura del tabaco. Las plantas conteniendo dos transgenes exhibieron un incremento en la altura del 10% - 48%, y las plantas que sobre expresaban los tres genes SPS, UGPasa y SuSy tuvieron un ligero incremento en la altura con respecto a lo anterior, en general las líneas transgénicas exhibieron una cantidad reducida de carbohidratos en las hojas, probablemente por el transporte directo de los fotoasimilados al tallo y el rápido uso de estos para el aumento del crecimiento. En este mismo estudio se observó una alteración del número de internodos florales así como un retardo en el tiempo de floración relacionado con la altura de la planta, esto es, que la línea de planta más altas fue la última en florear, a causa que el almidón no se almacena en las plantas transgénicas es posible que esto sea una razón en el retraso de la floración. La floración ha sido asociada con el contenido de carbohidratos, la degradación del almidón y transporte de sacarosa hacia los meristemas del tallo y raíz (Coleman *et al.*, 2009).

SPS se ha registrado también en tallos de caña de azúcar (Zhu *et al.*, 1997), cotiledones heterotróficos de *Ricinus* (Geigenberger y Stitt, 1991), frutas de kiwi (MacRae *et al.*, 1992) y durante la maduración de bananas (Hubbard *et al.*, 1990). Lo anterior ha llevado a creer en algunas posibles funciones en tejidos no fotosintéticos; una es que, SPS podría re-sintetizar azúcar, como en tubérculos de papa, después que la sacarosa se escinda y viaje vía apoplástica o simplástica. Otra idea es que, SPS está involucrado en un ciclo regulatorio en el cual la sacarosa es simultáneamente degradada y re-sintetizada (Geigenberger y Stitt, 1993).

## II.II. El gen de la sacarosa fosfato sintasa de *O. sativa* ssp. *indica*, *sps1*.

En arroz (*Oriza sativa* sub. *Indica*) se caracterizó el gen *sps1* con 13 exones y 12 intrones, productor de una proteína de 1048 aa (116.4 kDa) que según por sus características bioquímicas se puede clasificar como clase I, asimismo, SPS de *A. thaliana* es clasificado como clase III. La comparación de la secuencia de aa de SPS de maíz y arroz, ambos de la clase I, reporta un 78.4% de identidad, en tanto que, la comparación de la porción N-terminal de arroz, maíz, espinaca y tomate, específicamente la región de 180 a 700 aa es notablemente conservada, al contrario de la parte C-terminal es menos conservada. Se determinó que la región 5' UTR de 368 nucleótidos, contiene algunas secuencias repetitivas involucradas en la formación de la estructuras secundarias. Aún más, mediante Southern blot se demostró que solo hay un gen codificante para SPS en el ADN genómico de arroz (Valdéz-Alarcón *et al.*, 1996)

En cuanto la expresión tejido específico del gen *sps1*, fue determinada transcripcionalmente por reacciones de RT-PCR in situ, detectando productos de la reacción RT-PCR del mRNA de *sps1* en células mesófilas de la hoja, en escutelo y aleurona de la semilla, ínterin no fueron detectados productos de RT-PCR en tejido vascular, interesantemente hay un gradiente de expresión del mRNA del gen *sps1* identificado por RT-PCR in situ en la hoja que cubre el tallo, siendo mayor la expresión en la parte más externa de la hoja que está más expuesta a la luz que gradualmente va disminuyendo hacia la parte más interna. Análisis por computadora del promotor del gen *sps1* hasta los 2,196 pb permitió identificar algunos motivos *cis* involucrados en la regulación de este gen, dos clases de elementos se encontraron en el promotor; uno es de la clase de elementos en respuesta a luz (LREs por light responsive elements), motivos regulados por luz y en tejido específico encontrados en genes de la fotosíntesis. La otra clase de motivos encontrados fueron los elementos en respuesta a giberelinas (GARE gibberellin-response elements) presentes también en los genes  $\alpha$ -amilasa, además, se encontró la secuencia *cis* principal de elementos de respuesta a baja temperatura. Para analizar el patrón de expresión de *sps1* tomaron 1.8 kb del promotor junto con el 5' UTR, el primer exón, el primer intrón y parte del segundo exón fusionados transcripcionalmente al gen de reportero GUS, creando un gen quimera *sps-gus1*, encontrando actividad de GUS en

células mesófilas y sin actividad en tejido vascular (xilema y floema) de la hoja, igualmente fue para la hoja que cubre el tallo pero con un patrón de mayor actividad GUS en la zona más externa de la hoja y menor actividad en la parte más interna o menos expuesta a la luz de la hoja. La presencia de secuencias LREs en el promotor de *sps1* proponen una regulación por luz, en las hojas de las plantas transgénicas *sps-gus1* la actividad de SPS es 65% mayor en plantas con 16 hrs de luz/8 hrs oscuridad en relación con la plantas en continua oscuridad, otras plantas fueron crecidas en las mismas condiciones para la detección de actividad GUS, no habiendo actividad de GUS en las hojas primarias de plantas en continua oscuridad, mientras que la actividad de GUS fue detectada en la hoja completa de plantas con luz. Tanto en hojas expandidas como en hojas que cubren el tallo, la distribución de la actividad de SPS es basipétala, lo que es, del ápice hacia la base. En contraste, durante el desarrollo de la hoja emergente sólo hay actividad GUS en el ápice, sugiriendo que *sps1* sigue un patrón de expresión transicional de tejido fuente a tejido consumidor. Conforme la hoja va desarrollándose de importador neto de azúcar, siendo la yema del brote y la hoja inmadura, la actividad de GUS va aumentando concomitante con la actividad de SPS durante en desarrollo de hojas expandidas o exportadoras de azúcar. Sobre el mismo gen quimera *sps-gus1* detectaron expresión de GUS en escutelo de plántulas de tres días de germinación como también durante la germinación de la semilla en la que el almidón, los lípidos o ácidos grasos son convertidos en el disacárido sacarosa que es transportado del endospermo al embrión a través del escutelo. Adicionalmente, en las plantas transgénicas *sps-gus1* se detecto actividad GUS en otro tejido, en el polen el promotor de *sps1* dirigió la expresión del gen quimera. En cuanto que no identificaron expresión de GUS en las raíces de los transgénicos, la tinción de GUS fue localizada en el polen de las inflorescencias inmaduras, pareciendo que *sps1* es expresado durante el desarrollo de polen (Chávez-Barcenás *et al.*, 2000).

El promotor de *sps1* no tiene caja TATA reconocible ni secuencias iniciadoras cerca del sitio de inicio de la transcripción, en su lugar, río arriba del sitio de inicio a la transcripción tiene algunas regiones ricas en G/C, secuencias degeneradas del lugar de unión a Sp1. Utilizando el sistema de bombardeo de partículas realizaron ensayos de expresión transitoria en discos de hojas de tabaco midiendo los radio de los focos de fluorescencia verde/azul del gen reportero GFP derivaron que, el promotor *sps1* es débil. Con el mismo sistema de bombardeo se

demonstró la expresión tejido específico dirigida por el promotor *sps1* en semillas de arroz germinantes, notablemente, delimitaron una parte hasta la posición -1830 del promotor con capacidad de expresión en ambas capas, escutelo y aleurona de la semilla, a partir de tal posición el promotor conduce una expresión exclusivamente en escutelo y no en aleurona de la semilla. Con plantas transgénicas de arroz se evidenció la expresión tejido específico controlada por el promotor *sps1*, análisis histoquímicos a semillas germinantes sugieren que desde -1857 del promotor hacia 5' se encuentran los elementos que suprimen la expresión en aleurona ya que sólo hay expresión en escutelo. Complementariamente, la expresión tejido específico se demostró al situar actividad GUS en el mesófilo y no detectar GUS en epidermis y tejido vascular de la hoja de plantas transgénicas. La expresión tejido específico también se identificó en granos de polen de la antera en desarrollo (Martínez-Trujillo *et al.*, 2004).

La parte mínima funcional del promotor de *sps1* se ha reducido a 22 pb en las proximidades de la posición -6 a +15 según el sitio de inicio de la transcripción, así descartando la intervención de las regiones ricas en G/C junto con el factor de unión Sp1 en los promotores de plantas, y eliminando la posibilidad de que actúen elementos DPE (Downstream Promoter Elements) por que la secuencia consenso no está presente, para esto, se identificó la secuencia TCACCC como iniciadora en base a que están superpuestos pirimidinas cumpliendo la regla de purina +1 y pirimidina -1 (Martínez-Trujillo *et al.*, 2008). Los promotores sin caja TATA se han visto en la mayoría de los genes nucleares de la fotosíntesis, apuntando a la presencia de elementos activadores en el promotor por luz (Nakamura *et al.*, 2002).

El 5' UTR o líder de *sps1* es inusualmente largo con 368 pb y se predice una estructura secundaria estable (con -106.6 Kcal/mol). En éste 5' UTR hay algunas secuencias TC como elementos potenciadores de la función del líder. El control génico del gen SPS tiene diferentes ordenes, ocurriendo cada uno progresivamente, siendo el principal de estos la regulación transcripcional. En el caso de *sps1*, la porción 5' UTR determina el nivel de expresión dirigido por el promotor, lo anterior fue determinado usando el método de bombardeo de partículas ensayado en disco de hojas de tabaco para la expresión transitoria de tres diferentes construcciones; una de estas es el promotor de -2 kb fusionado transcripcionalmente (hasta +20) al gen *uidA*, otra construcción, el promotor más 5' UTR, por último el promotor, el líder,

primer exón y parte del segundo (hasta +676) fusionados al gen *uidA*, obteniendo que, la remoción de los exones y el intrón no tiene cambio o efecto alguno en la acción de GUS con relación al promotor más el líder, en cambio el promotor solo, es decir, la eliminación del líder, confirió una disminución de hasta 10 veces en la intensidad del gen reportero. Sabiendo que la regulación de la expresión de GUS por el líder podría ser transcripcional o traduccional, se fusionó el líder hacia el extremo 5' del promotor truncado -90 y -46 del CaMV 35S para controlar la expresión de GUS, y no hubo ningún cambio en la expresión de GUS mostrando que no tiene efecto transcripcional. Por lo cual el 5' UTR fue insertado en dirección 3' del promotor entre éste y el gen *uidA*, los ensayos de expresión transitoria mostraron que el líder clonado sólo en su orientación nativa entre el promotor débil -46 de 35S CaMV y el gen *uidA* potencia la expresión de GUS hasta 10 veces, inclusive el 5' UTR clonado en medio de promotores fuertes, como el completo de 35S CaMV o CVMV, y *uidA* amplifica la expresión de GUS 30% más. Para estudiar SPS en condiciones fisiológicas dos vectores binarios fueron usados para crear plantas transgénicas de *A. thaliana*, uno contenía el promotor -46 de 35S CaMV::*uidA* y el otro -46::5'UTR-*uidA*, se elaboraron extractos de cada planta transgénica para cuantificar los niveles de GUS por análisis fluorométricos, se pudo medir 13 veces más actividad de GUS en la planta con el líder que en las líneas sin el líder; la estructura secundaria del 5' UTR podría intervenir en la regulación, para responder a esto, los niveles de ARNm de líder-GUS se evaluaron valiéndose de la técnica RT-PCR en dos pasos en las plantas transgénicas de *A. thaliana* con las construcciones con y sin líder; el ratio de los productos de las amplificaciones de PCR GUS provenientes del ARNm entre ambas líneas transgénicas reveló que los niveles de transcritos de GUS son semejantes y que el líder de *sps1* no altera la estabilidad del RNAm (Martínez-Trujillo *et al.*, 2003).

Con la intención de conocer como actúa el 5' UTR en la traducción, llevaron a cabo experimentos in vitro de construcciones génicas en la cuales el gen *uidA* se clonó al promotor T7 sin el líder y con el líder, usando un sistema de transcripción in vitro dependiente del promotor T7 su obtuvieron transcritos entrambos plásmidos, iguales cantidades de estos transcritos fueron sometidos a traducción in vitro usando extracto de germen de trigo con <sup>35</sup>S metionina, resultando en un incremento de 3 veces más incorporación de <sup>35</sup>S metionina en los productos de traducción de los transcritos de GUS con líder en comparación con transcritos sin

líder. Sugiriendo que el 5' UTR actúa como un potenciador traduccional, a través del cual están contenidos secuencias reiteradas o elementos estructurales responsables de la actividad potenciadora durante la traducción del ARNm (Martínez-Trujillo *et al.*, 2003).

En las plantas y animales, los 5' UTR son generalmente cortos de alrededor 20 a 100 nucleótidos (Joshi, 1997). Tales regiones líder son clasificadas conforme al mecanismo ejecutante de la siguiente manera: (a) presencia de uORFs cortos en la región líder, los cuales tienen un efecto inhibitorio durante la traducción, los uORFs falsos son traducidos y por lo tanto reducen el inicio de la traducción río abajo del codón de inicio (Fütterer y Hohn, 1996). (b) Potenciación de la traducción dirigido por el líder solo o coadyuvado con la región 3' trailer, como en el virus mosaico de tabaco (TMV con un líder de 68 bases) las repeticiones de secuencias en la región líder  $\Omega$  en donde una secuencia de 8 nucleótidos se haya repetida dos o tres veces al mismo tiempo que una región de 25 nt (CCA)<sub>n</sub> (Gallie y Walbot, 1992), en el gen nuclear *psaDb* codificador de la subunidad D del fotosistema I de *N. sylvestris* la estructura primaria del ARNm es responsable de la eficiencia traduccional, su 5' UTR (con 23 bases) aumentador de la traducción tiene secuencias redundantes de TC caracterizadas por controlar la actividad potenciadora del líder; las mismas secuencias repetitivas TC están en el líder del gen nuclear *FedA* de *A. thaliana* (Yamamoto *et al.*, 1995); ciertamente en el líder de *FedA* lleva la secuencia ACAAAA traslapada dos veces río arriba del ATG de inicio de la traducción (Caspar y Quail, 1993), repetición que además portan los genes de ferredoxina de espinaca (Wedel *et al.*, 1988) y de guisante, *Fed-1* (Dobres *et al.*, 1987), entre otros. Un caso de incremento en la traducción por colaboración de 5' con el 3' UTR es el del gen  $\alpha$ -amilasa de cebada (Gallie y Young, 1994). (c) Los potenciadores de la transcripción o elementos silenciadores han sido documentados en las regiones líderes de algunos genes como lo son, *PetH*, *PetE* y *PsaF* (codificadores para ferredoxina-NADP oxidoreductasa, plastocianina y subunidad III del fotosistema I, correspondientemente) que tienen secuencias reiterantes que incrementan la expresión cuantitativa transcripcional (Bolle *et al.*, 1994), por el contrario el líder del gen *LAT59* inhibe la acumulación del ARNm (Curie y McCormick, 1997). (d) La existencia de secuencias de elementos que median la estructura secundaria, estabilidad y vida media del ARNm han sido registradas, la abundancia del ARNm del gen *Fed-1* es regulado por luz, se evidenció un elemento repetido en el 5' UTR capaz de alterar la estabilidad del

ARNm y la capacidad de carga de poliribosoma, en luz el ARNm es estable cuando se asocia con el poliribosoma pero en oscuridad no se asocia con el poliribosoma y el ARNm es rápidamente degradado por un proceso que involucra estas secuencias (Dickey *et al.*, 1998), otros genes en los cuales el 5' UTR eleva la estabilidad y abundancia del ARNm son el *Lhcb1\*4* de guisante y un gen de heat-shock de *Daucus carota* (Anderson *et al.*, 1999).

Se ha propuesto que los líderes pueden actuar de manera pasiva al no presentar limitaciones a la subunidad ribosomal 40S durante la búsqueda del codón de inicio de traducción, o mediante un mecanismo activo que requiere un factor regulador actuante en *trans*. La redundancia funcional de secuencias repetitivas del 5' UTR podrían intervenir en la efectividad de la traducción estableciendo una afinidad mayor o un requerimiento menor de los componentes la maquinaria traduccional (Jobling y Gehrke, 1987; Timmer *et al.*, 1993; Wang *et al.*, 1997).

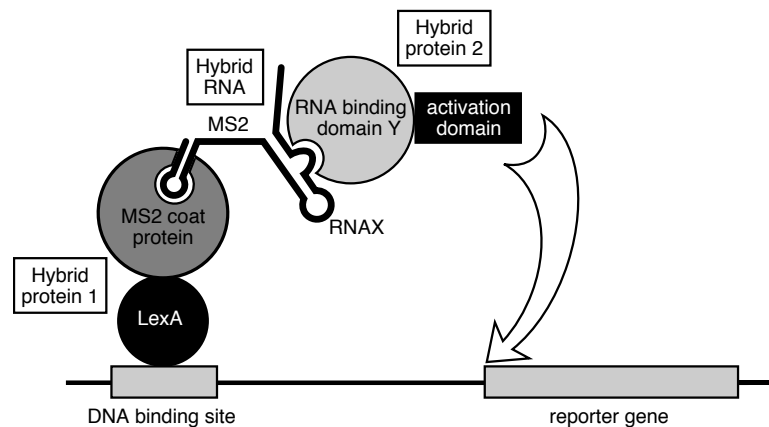
Las redundancias de elementos formados por secuencias consenso en la región líder del ARN genómico del TMV estimulan la traducción ensamblando específicamente una proteína de 102 kDa de unión a ARN en la región de poli(CAA) homólogo a HSP101 (Heat Shock Protein) (Wells *et al.*, 1998), la misma proteína también ha sido reconocida por unirse al iLRE (internal light-regulatory element) en el 5' UTR e intensificar la traducción del transcrito del único gen para ferredoxina *Fed-1* de guisante, secuencias ricas en CA similares al elemento poli(CAA) del líder  $\Omega$  del TMV han sido determinadas como el sitio de unión de la proteína HSP101 (Ling *et al.*, 2000). Comparablemente ambos genes el de *FedA* de *A. thaliana* junto con *psaDb* de *N. sylvestris* contienen secuencias profusas en TC (como la serie TCTCTCAA de *psaDb*), la substitución de tal caja de secuencias decrece la eficiencia de la traducción, entonces permitiendo especular que la traducción de los dos genes puede ser potenciados por el mismo mecanismo (Yamamoto *et al.*, 1995).

Análisis del líder de *sps1* exhibieron una estructura secundaria estable de -106.8 kcal/mol con varias configuraciones tallo y asa con efecto positivo en la traducción (Martínez-Trujillo *et al.*, 2002). Dentro del líder existen cinco regiones copiosas en TC, que forman parte de las estructuras tallo y asa del ARNm, como posibles sensores de la función potenciadora traduccional del líder de *sps1*. Deleciones hechas al 5' UTR han justificado una redundancia

funcional provocada por la distribución de los grupos de secuencias repetidas a causa que en las versiones truncadas del líder *sps1* hay al menos una repetición de TC (Martínez-Trujillo *et al.*, 2003). Por lo anterior, es posible suponer que alguna proteína específicamente actúe *trans* sobre el 5' UTR del ARNm de *sps1* y aumente la traducción.

### II.III. Método de tres híbridos en *S. cerevisiae*.

Para analizar los complejos ARN-proteína se ha usado un método genético molecular; el sistema de tres híbridos en levadura (*Saccharomyces cerevisiae*) toma ventaja de la unión específica de la proteína al ARN, la interacción puede ser detectada rápidamente en levadura de una manera independiente al rol biológico de la proteína y el ARN, el sistema faculta un monitoreo del crecimiento celular por auxotrofia aminoacídica, al momento de la activación del gen reportero nutricional *HIS3* (dictador de una proteína necesaria para la biosíntesis de histidina) o por el color de las colonias inducido por la actividad del gen reportero *LacZ* al metabolizar su substrato cromogénico (Bernstein *et al.*, 2002).



**Figura 1.** Diagrama del sistema de tres híbridos de levadura. Una proteína híbrida 1 contiene un dominio de unión a ADN (p.e. LexA) en el extremo N junto con un dominio de unión a ARN (p.e. proteína MS2 que se une y encapsida el ARN viral) en el extremo C terminal; El ARN quimera contiene dos concomitantes sitios de unión MS2 clonados junto a el ARNX de secuencia conocida e interés; en la proteína híbrida 2 el dominio de activación transcripcional (AD) Gal4 invariablemente va en el N terminal mientras que en el extremo C terminal va insertado el péptido Y de unión a ARN en búsqueda. Los genes reporteros *HIS3* y *LacZ* son regulados por los operadores LexA (cuatro repeticiones en el promotor de *HIS3* en tanto que ocho en el promotor de *LacZ*).

En el sistema de tres híbridos (Figura 1), la levadura contiene en su genoma los genes reporteros *HIS3* y *LacZ* bajo el control de múltiples copias de un sitio de unión del ADN de 17 nucleótidos, o UAS, upstream activating sequence (Putz *et al.*, 1996), secuencia reconocida por una proteína híbrida 1 hecha de un dominio de unión a ADN originaria de la proteína LexA de *E. coli* fusionado a la proteína capsídica de unión a ARN del bacteriófago MS2, el último es un pequeño polipéptido que se une específicamente como dímero a una secuencia conocida para tallo-asa de 21 nucleótidos del ARN MS2 (SenGupta *et al.*, 1996). La molécula quimera de ARN consta por una parte de dos copias en tándem del sitio de unión para la proteína MS2 unido a la secuencia ARN de interés conocida, X, en cual se asocia un polipéptido de unión a ARN en búsqueda, Y, el ARN híbrido es bifuncional a causa que contiene el sitio de unión para el péptido Y de unión a ARN. El polipéptido quimera 2 porta la proteína cognado Y de unión a ARN asociado al dominio de activación transcripcional Gal4 de levadura (Bernstein *et al.*, 2002).

El método fue diseñado para descubrir proteínas que se unen a una secuencia de ARN conocida, para la confirmación de interacciones ARN-proteína sospechadas o el análisis de mutaciones sobre la estructura específica de ARNs o proteínas interactuantes y para el descubrimiento de complejos multiproteicos. Cualitativamente la afinidad de la interacción, débil o fuerte, entre ARN-proteína se distingue por sus niveles de expresión de *HIS3* fenotípicamente y *LacZ* por ensayo colorimétrico. Ampliamente usado además para detectar proteínas cognado de unión a ARN provenientes de una biblioteca de ADNc que se unen a una secuencia específicas de ARN de interés. Una ventaja de la técnica es saber por medio de escrutinio cual clona de ADNc cognado codificó el péptido de unión a ARN e interés (Bernstein *et al.*, 2002; Hook *et al.*, 2005).

La existencia de un factor proteico que interaccione con el ARN de la región líder *sps1* es una posibilidad que no puede ser descartada a priori y que se analiza en este trabajo. La interacción de ARN con proteínas puede determinarse ya sea *in vitro* mediante el análisis de retardo de la migración del ARN en geles de poliacrilamida o bien en condiciones *in vivo*, para lo cual se utiliza el sistema de tres híbridos en la levadura *S. cerevisiae* (SenGupta *et al.* 1996).

### **III. HIPÓTESIS**

Existe una proteína que interacciona con la región líder en el ARN del gen *sps1*

### **IV. OBJETIVOS**

#### **IV.I. Objetivo general**

Determinar la interacción *in vivo* en el 5' UTR del ARN del gen sacarosa fosfato sintasa de arroz con proteínas.

#### **IV.II. Objetivos particulares**

Obtener secuencias de genes que codifiquen para proteínas que interaccionen con el 5' UTR del ARN del gen sacarosa fosfato sintasa de arroz.

Analizar tanto el promotor como la posible función biológica de la proteína encontrada mediante análisis bioinformático.

Aislar del genoma versiones de diferentes tamaños de la región promotora de la proteína encontrada y clonaras en vectores.

## V. MATERIALES Y MÉTODOS

### V.I. Vectores de las moléculas en el sistema de tres híbridos.

El vector pHybLex/Zeo-MS2 (Figura 1) codifica una proteína híbrida con un dominio de unión al ADN y un dominio de unión al ARN de MS2, que formará parte del RNA híbrido.

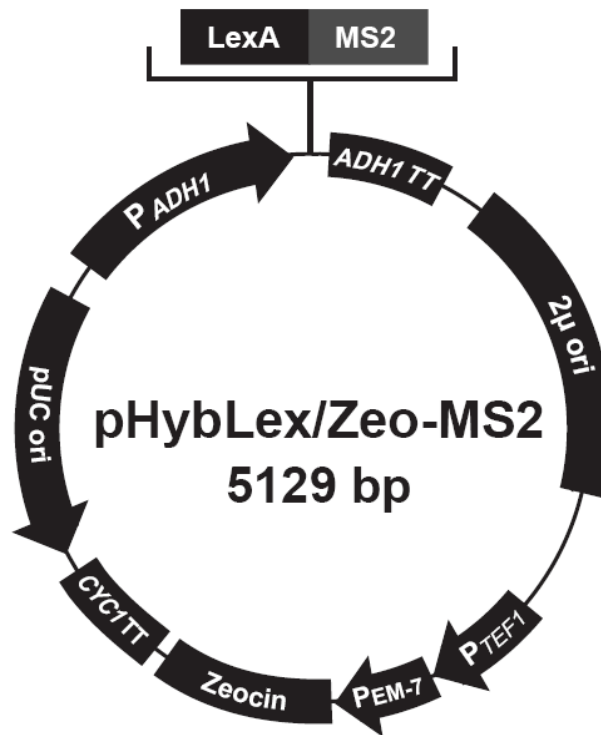


Figura 1. **pHybLex/Zeo-MS2**. Promotor ADH1: bases 1-399. Proteína de fusión LexA/MS2; ORF LexA: bases 420-1026, proteína de abrigo de MS2: bases 1044-1436. Región de término de transcripción de ADH1: bases 1508-1665. Origen de replicación para levadura 2μ: bases 1804-2685. Promotor TEF1: bases 3219-3627. Promotor EM-7: bases 3632-3698. Gen resistente a Zeocina: bases 3699-4073. Región de término de transcripción CYC1: bases 4074-4391. Origen de replicación para bacteria pUC: bases 4402-5075 (cadena complementaria)

pSPS-MS2. Construido a partir del vector pRH3' (Fig. 2) por López-Salmerón, 2007, produce el transcrito de ARN híbrido con la región líder *sps1* y el MS2 de unión al dominio de la proteína MS2. Formando un puente entre la proteína generada por pHybLex/Zeo-MS2 y alguna proteína codificada en la biblioteca de ADNc de *A. thaliana* que se una al ARN del líder *sps1*.

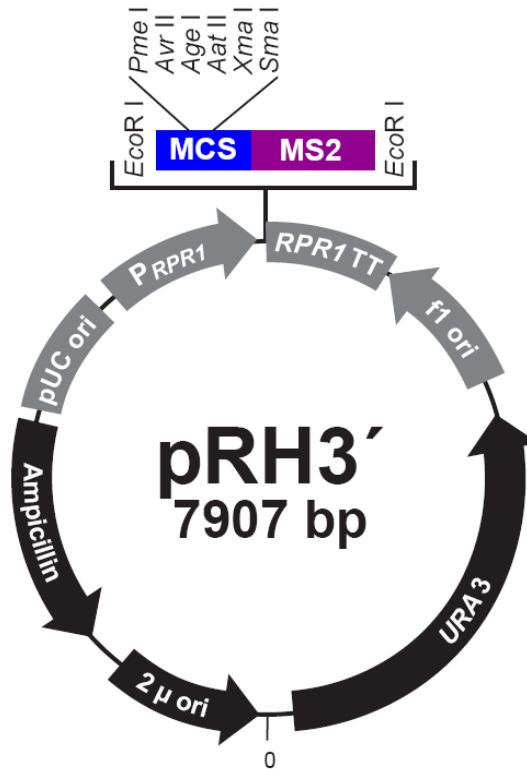


Figura 2. **Vector pRH3' genera el ARN híbrido, INVITROGEN.** Promotor RPR1: bases 3658-4065. Inicio de transcripción: 3981. Secuencia líder de ARNasa P: bases 3981-4065. Sitio de hibridación del cebador pRH Forward: 3937-3958. Sitio múltiple de clonación: bases 4094-4124. MS2 RNA: bases 4148-4220. Sitio de hibridación del cebador pRH Reverso: 4315-4337. Región de término de transcripción RPR1: 4247-4684. Origen f1: bases 6110-6565 (cadena complementaria). Gene URA3 bases 6615-7723 (cadena complementaria). Origen 2 μ: bases 534-1412 (cadena complementaria). Gene con resistencia a ampicilina: bases 1543-2395 (cadena complementaria). Origen pUC: bases 2548-3221.

En el vector pGAD10 (CLONTECH), Fig 3, está clonada la biblioteca genómica que contiene los ADNc de los genes de *Arabidopsis thaliana* fusionados a la región codificante del dominio de activación de la proteína GAL4, por lo que en la expresión se generan proteínas híbridas. La fuente de los ARN mensajeros fue de tejido verde vegetativo de una planta de 3 semanas de *A. thaliana* proporcionada por el Dr. Luis Herrera Estrella y el Dr. Gustavo Acevedo Hernández (CINVESTAV-IPN, Unidad Irapuato)

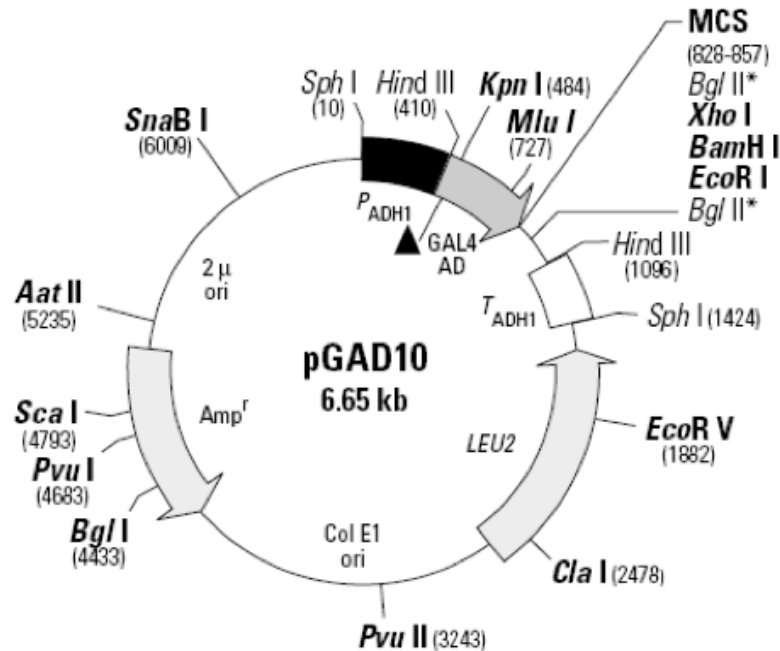


Figura 3. **Vector pGAD10**. Promotor truncado ADH1 de *S. cerevisiae*: bases 10-406. Codón de inicio (ATG) del polipéptido del dominio de activación: 422-424. Codón de alto 866-868, codones GAL4 768-881: 491-829, Señal de localización nuclear SV40: 452-472. Sitio múltiple de clonación: 828-857. Codones de alto de traducción: 859-861, 866-868 & 873-875. Señal de término de transcripción, fragmento terminador ADH1 de *S. cerevisiae*: 1096-1423. Secuencias codificadoras de LEU2: Codón de inicio (ATG):2631-2629; Codón de alto: 1539-1537. Origen de replicación del plásmido Col E1: 3448-4091. Gene de resistencia a ampicilina: Promotor: región -35: 5169-5164; región -10; 5146-5141. Inicio de transcripción: 5134, sitio de unión del ribosoma: 5111-5107. Secuencias codificantes de β-lactamasa: codón de inicio (ATG): 5099-5097; codón de alto: 4241-4239, Señal del péptido β-lactamasa: 5099-5031, Proteína madura de β-lactamasa: 5030-4242.

## V.II. *Saccharomyces cerevisiae* cepa L40-ura3.

Esta cepa contiene dos genes reporteros, lacZ e HIS3, cuya expresión es regulada por secuencias del operón LexA. La expresión de estos dos genes permite detectar la actividad de β-galactosidasa y seleccionar en medios carentes de histidina, respectivamente. La cepa requiere uracilo para su crecimiento. Se obtuvo a través de la empresa INVITROGEN. Algunas características importantes de L40-ura3 son:

Características de L40-ura3		
Cepa	Genotipo	Fenotipo
L40-ura3	MAT <i>a</i> , <i>ura3-52</i> , <i>leu2-3112</i> , <i>his3Δ200</i> , <i>trp1Δ1</i> , <i>ade2</i> , LYS2::( <i>LexA op</i> )4-HIS3, <i>ura3</i> ::( <i>LexA op</i> )8- <i>lacZ</i>	His <sup>-</sup> , Trp <sup>-</sup> , Ura <sup>-</sup> , Ade <sup>-</sup> , Leu <sup>-</sup>

Característica	Propósito
Expresión constitutiva de los genes dirigidos por el promotor <i>GAL1</i>	Permite la expresión de la proteína de interacción (“presa”) en un medio que contiene glucosa (YC o YPAD)
Marcador auxótrofo de histidina regulado por LexA	Reportero que permite la selección de interacciones positivas de RNA-proteína en un medio deficiente de histidina
Marcador auxótrofo <i>ura3</i>	Permite la selección de los plásmidos que contienen el gen <i>URA3</i> (Ej. pSPS-MS2)
Gen <i>LacZ</i> regulado por LexA	Reportero que permite la selección de interacciones positivas de RNA-proteína mediante el ensayo de la actividad de β-galactosidasa.
Marcador auxótrofo <i>trp1Δ1</i>	Permite la selección de plásmidos que contienen el gen <i>TRP1</i>

El medio YPD líquido (CLONTECH cat. 630409) o sólido (CLONTECH cat. 630410) es para crecimiento no selectivo y general de la cepa L40-ura3. Contiene 1% de extracto de levadura, 2% de peptona, 2% de glucosa. El medio sólido se complementa con 2% de agar bacteriológico.

El medio mínimo YC se utiliza para el crecimiento de levadura y es selectivo de acuerdo a la mezcla de aminoácidos adicionada. Se utilizó SD base CLONTECH cat. 630411 para líquido y cat. 630412 para sólido. A este medio se le agregan mezclas de aminoácidos requerida para la selección de la cepa con los plásmidos transformados de acuerdo a los marcadores. Se utilizaron los siguientes suplementos: -Ura (cat. 630416), -His/-Leu/-Ura (cat. 630423) y -Leu (cat. 630414).

### V.III. Transformación secuencial de *S. cerevisiae* L40-ura3 con los plásmidos pHybLex/Zeo-MS2, pSPS-MS2 y la biblioteca de ADNc de *A. thaliana* y escrutinio para seleccionar la interacción.

Se inoculó la cepa L40-ura3 en 10 ml. de medio YPD, y se incubó toda la noche a 30°C en agitación. Se diluyó el cultivo a una OD<sub>600</sub> de 0.4 en 50 ml de medio YPAD o YPD a una densidad de 5 x 10<sup>6</sup> células/ml. Se incubó este cultivo a 30°C en agitación a 200 rpm hasta una densidad de 2 x 10<sup>7</sup> células/ml (de 2 a 4 horas), posteriormente se realizó el siguiente

protocolo:

Concentración de las células por centrifugación a 2500 rpm por 5 minutos. Resuspensión de la pastilla con 40 ml de 1X TE, y centrifugación nuevamente (2500rpm x 5 min). Resuspensión de la pastilla en 2 ml de 1X LiAc/0.5X TE. Incubación a temperatura ambiente por 10 minutos. Preparación y adición de la mezcla de transformación (100 µg ADNTT (ADN acarreador de salmón INVITROGEN, 1 µg DNA del plásmido, 100 µl de células de levadura, 700 µl 1X LiAc/40% PEG-3350/1X TE). Incubación por 30 minutos a 30°C. Adición de 88 µl de DMSO. Incubación por 7 minutos a 42°C. Centrifugación a 14 000 rpm por 15 segundos. Resuspensión en 1 ml de 1X TE. Centrifugación a 14 000 rpm por 15 segundos. Eliminación del sobrenadante y resuspensión en 50-100µl de 1X TE. Plaqueo en el medio de selección deseado. Incubación a 30°C durante 48 horas.

La selección de la transformación con el plásmido pHybLex/Zeo-MS2 se realizó en medio YC con zeocina (200 µg/ml). A partir de esta levadura transformada se realizó la transformación con el plásmido pSPS-MS2 y se seleccionó la co-transformación en medio con zeocina (200 µg/ml) carente de uracilo.

El escrutinio de las interacciones de las moléculas híbridas se realizó de la siguiente manera. Inoculación de 100 ml de medio YC –U Zeo con una colonia de levadura (L40-Ura3) co-transformada con pSPS-MS2 y pHybLex/Zeo-MS2, a 30° C en agitación durante toda la noche hasta fase estacionaria. Transferencia del cultivo a 500 ml de medio YC –U Zeo fresco hasta alcanzar una OD<sub>600</sub> de 0.2 – 0.4. Incubación en agitación a 30° C por 4 hrs. Centrifugación a 2500 rpm por 5 min y retiro del sobrenadante. Resuspensión de la pastilla en 250 ml de TE 1X. Centrifugación a 2500 rpm por 5 min y retiro del sobrenadante. Resuspensión de la pastilla en 4 ml de AcLi/TE 0.5X. Incubación de las células a temperatura ambiente por 10 min. Transformación con 40 µl de ADN de la Biblioteca de ADNc (a una concentración de 1.9 µg/ml), 1 ml de ADN desnaturalizado de espera de salmón, 5 ml de células de levadura y 30 ml de AcLi 1X /40% PEG 3350 /TE 1x. Incubación de la solución a 30° C por 30 min. Adición de 3.5 ml de DMSO. Choque térmico a 42° C por 15 min en baño María. Incubación de la mezcla en hielo por 1 min. Centrifugación a 12000 rpm por 30s y decantado del sobrenadante. Resuspensión de la pastilla en 4 ml de TE 1x. Plaqueo de 50 µl por caja con

medio YC –ULH Zeo, hasta completar 100 cajas. Incubación a 30° C esperando el crecimiento de las primeras colonias. Otros plaqueos se realizaron en medio YC –ULH Zeo con el inhibidor 3-AT (3 amino 1, 2, 4, triazol) a una concentración de 20 mM, ya que la fuerza de la interacción se determina por los niveles de resistencia a este compuesto, que es un inhibidor competitivo del producto del gen HIS3 que se activa durante la interacción (Bernstein et al. 2002). Se tomaron las colonias que crecieron en los primeros 5 días y que tuvieron un tamaño al menos de un mm de ancho.

#### **V.IV. Ensayo de $\beta$ -galactosidasa.**

Se Agregó 1 g de agarosa (Sigma cat: A9539), a 100 ml (100mM) de buffer de fosfatos pH 7.0 (39 ml 0.1 M de KH<sub>2</sub>PO<sub>4</sub> y 61 ml 0.1 M de K<sub>2</sub>HPO<sub>4</sub>) y se disolvió calentando por 3-5 min en el horno de microondas. Se mantuvo la solución a 65° C. Por separado se agregaron 800  $\mu$ l de X-Gal/Dimetilformamida (25 mg/ml) en 8 ml de Dimetilformamida (Sigma) y se mezclaron con 12 ml de la agarosa de bajo punto de fusión disuelta en el buffer de fosfatos. Se incubó la solución a 65° C por 5 min y se colocó una ligera capa de la solución en las cajas de petri ensayadas, permitiendo de 5-10 min a que solidificara la agarosa. Se incubaron las cajas de petri en la oscuridad de 2-3 hrs o bien toda la noche a 30° C para que la colonias positivas a la actividad  $\beta$ -galactosidasa se tornaran de color azul.

#### **V.V. Extracción de ADN de levadura para recuperar los plásmidos de la biblioteca de ADNc de *Arabidopsis thaliana*.**

Inoculación de una colonia de levadura a 30° C en agitación en 10 ml de medio YC –Leu de 24 a 48 hrs. Centrifugación a 2000 rpm por 5 min. Eliminación del sobrenadante. Resuspensión de la pastilla en 1 ml de Sorbitol 1M, NaEDTA 0.1 M (pH 7.5). Adición de 50  $\mu$ l de solución de Liticasa (SIGMA) (4000 U/ml) e incubación por 1 h a 37° C. Centrifugación a 12000 rpm por 1 min. Eliminación del sobrenadante y resuspensión de las células en 1 ml de Tris-Cl 50 mM (pH 7.4) NaEDTA 20 mM. Adición de 100  $\mu$ l de SDS al 10%, mezclado e incubación a 65° C por 30 min. Adición de 400  $\mu$ l de acetato de potasio 5 M y transferencia a hielo por 1 h. Centrifugación a 12000 rpm por 5 min a 4° C. Transferencia del sobrenadante a

un tubo nuevo y adición de un volumen de isopropanol. Mezclado suave e incubación por 5 min a temperatura ambiente. Almacenamiento a -20° C toda la noche. Centrifugación por 5 min a 12000 rpm. Eliminación del sobrenadante y secado de la pastilla. Resuspensión de la pastilla en 100 µl de TE (pH 7.4).

#### **V.VI. Selección de plásmidos de la biblioteca de ADNc extraídos de la levadura.**

Las levaduras que mostraron crecimiento en medios YC-ULH Zeo 3-AT 20 mM incorporaron teóricamente los plásmidos pHybLex/Zeo-MS2, pSPS-MS2 y algún plásmido de la biblioteca de ADNc de *Arabidopsis thaliana*. Los dos últimos tienen resistencia a carbenicilina y el primero a zeocina. La mezcla de plásmidos extraída se utilizó para transformar *E. coli* DH5α y se hizo una selección en medio LB con carbenicilina (100 µg/µl). Se escogieron 10 colonias y se extrajo ADN por la técnica de miniprep. Se cortó el ADN con EcoRI y las mezclas se corrieron en geles de agarosa al 1%. Los vectores provenientes de la biblioteca de ADNc liberaron un fragmento de 6.65 Kb correspondientes al vector pGAD10. Los plásmidos de la biblioteca de ADNc recuperados se mandaron secuenciar utilizando el oligonucleótido 5' TAC CAC TAC AAT GGA TG 3'.

#### **V.VII. Secuenciación.**

Se utilizó el método de terminación de cadenas (Sanger et al. 1977), usando dideoxinucleótidos fluorescentes (ABI) y un secuenciador de DNA ABI PRISMTM 374. Se solicitaron los servicios del CINVESTAV-IPN Unidad Irapuato.

#### **V.VIII. Identificación de las proteínas que generaron interacción en el sistema de tres híbridos.**

Las secuencias de los ADNc de los plásmidos recuperados de la levadura se utilizaron para alinearse con secuencias de la base de datos del servidor del Nacional Center of Biotechnology Information (NCBI), mediante el uso de la herramienta BLASTn con la opción nucleotide collection (nr/nt). <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.

#### **V.IX. Identificación de las proteínas Spen de *A. thaliana* y arroz indica.**

La proteína del gen encontrado en el BLASTn, At4g12640, fue sometida a un análisis comparativo de las secuencias de aminoácidos para descubrir sus ortólogos de *O. sativa* ssp. *indica* usando la herramienta BLASTp con la opción Non-redundance protein sequences (nr) de NCBI. La posición, orientación y secuencia de los genes de arroz indica dentro del genoma fueron encontrados en la base de datos de ENA (European Nucleotide Archive) de EMBL-EBI <http://www.ebi.ac.uk/ena/>.

#### **V.X. Alineación de la secuencia de aminoácidos entre las proteínas de arroz indica y *Arabidopsis thaliana*.**

Las secuencias de aminoácido de las proteínas Spen de los genes At4g12640 y de arroz indica OJ1381\_H04.27 y P0450E05.35 fueron alineadas usando la herramienta BLAST2seq de NCBI

#### **V.XI. Análisis de los promotores de los genes de *O. sativa* ssp. *indica* y de *A. thaliana*.**

Las secuencias de nucleótidos de los promotores de los genes de arroz indica se obtuvieron al alinear la secuencia de los genes, obtenida de la base de datos ENA, contra la base de datos de NCBI en la herramienta BLASTn de *Oryza sativa* con la opción *O. sativa* (*indica* cultivar-group) WGS contigs. Mientras que la secuencia del promotor del gen de *A. thaliana* se encontró en la base de datos de NCBI.

Para conocer los elementos que regulan la expresión de estos genes se usó la base de datos de secuencias *cis* para plantas PLACE (<http://www.dna.affrc.go.jp/htdocs/PLACE/>).

#### **V.XII. Diseño de oligonucleótidos.**

Los oligonucleótidos fueron diseñados en la página de IDT (Integrated DNA Technologies) <http://www.idtdna.com/analyzer/applications/oligoanalyzer/>, además se les agregó en el extremo 5' la secuencia para corte con la enzima Hind III. Los oligos fueron:

Reverso 2

5'- AAGCTT TCT GCA TTC GTC AGA TCT ATC GCA -3'

Forward 500

5'- AAGCTT ACA TGA CGA GCA GAT CTA CGG AGA – 3'

Forward 1000

5' - AAGCTT ACC TCA CAG AAT TGA CAG ATC CA – 3'

### **V.XIII. Extracción de ADN de *A. thaliana* y amplificación de las diferentes versiones del promotor del gen At4g12640.**

El ADN de planta fue aislado usando el kit DNeasy Plant Mini (cat. 69104) tal bajo el mismo protocolo del fabricante. Este ADN se utilizó como templado para realizar PCRs con los oligonucleótidos diseñados empleando el kit para PCR de fermentas (EP0403), utilizando la fórmula de reacción del fabricante y el siguiente protocolo:

95° C 5' – por 35 ciclos; 95° C 1', 55° C 1', 68° C 1'– 68° C 10' y 4° C ∞'.

### **V.XIV. Clonación de las diferentes versiones del promotor del gen At4g12640 al vector de entrada pCR8/GW/TOPO y secuenciación.**

Las versiones amplificadas del promotor del gen At4g12640 por PCR se clonaron al vector de entrada Gateway pCR8/GW/TOPO, fig. 4, usando la mezcla de reacción sugerida por el mismo fabricante (cat. K2500-20, invitrogen). Para comprobar que los productos de PCR se hayan insertado en la orientación 5'-3' de acuerdo al vector, las clonaciones se mandaron secuenciar a el CINVESTAV-IPN Unidad Irapuato.

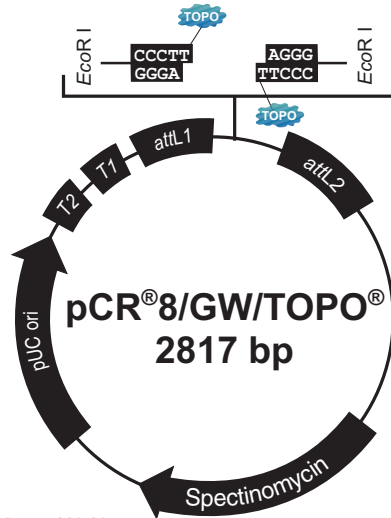


Figura 4. **Características del vector pCR8/GW/TOPO.** *rrnB* T2 secuencia de terminación de la transcripción: bases 268-295. *rrnB* T1 secuencia de terminación de la transcripción: bases 427-470. M13 sitio para el primer forward (-20): bases 537-552. *attL1*: bases 569-668. GW1 sitio de este primer: bases 607-631. TOPO sitio de reconocimiento 1: bases 678-682. TOPO sitio de reconocimiento 2: bases 683-687. *attL2*: bases 696-795. GW2 sitio de este primer: bases 733-757. T7 promotor/sitio del primer: 812-831 (c). M13 sitio para el primer reverso: bases 836-852. Espectinomicina promotor: bases 930-1063. Espectinomicina gen de resistencia (Spn): 1064-2074. pUC origen: bases 2141-2814. (c) = secuencia complementaria.

## V.XV. Manejo del ADN plasmídico y transformación de células competentes.

Para la replicación y obtención de ADN de los plásmidos se utilizó la cepa TOP10 de *E. coli* (INVITROGEN) y la cepa DH5 $\alpha$  (Sambrook et al. 1989). Para el aislamiento del ADN de los plásmidos bacterianos se utilizaron los métodos de Miniprep y Maxiprep (con polietilenglicol) descritos por Sambrook *et al.* (1989). Las condiciones de corte con endonucleasas de restricción se hicieron siguiendo las indicaciones del fabricante (INVITROGEN). Se transformaron células competentes de *E. coli* TOP10 o DH5 $\alpha$  con las mezclas de ligación, las cuales se dializaron previamente en membranas millipore con un tamaño de poro de 0.025  $\mu$ m, utilizando un electroporador Eppendorf 2510 a 1800 V.

El medio LB se utilizó para el crecimiento de *E. coli* contiene 1% de Bacto triptona, 0.05% de extracto de levadura y 1% de cloruro de sodio. Para medio sólido se adiciona 1.5 % de agar bacteriológico. En tanto que los geles de agarosa se elaboraron con 1% agarosa a 1X TAE.

## V.XVI. Recombinación dirigida entre los sitios LR de los vectores de entrada y el vector destino pKGWFS7.

Las versiones del promotor de At4g12640 clonadas en el vector de entrada pCR8/GW/TOPO se amplificaron por PCR con los oligos M13, llevando dentro de los productos los sitios de *attL1* y *attL2* que se recombinan con los sitios *attR1* y *attR2* correspondientemente del vector binario pKGWFS7 (fig. 5). La recombinación fue hecha con el Kit de invitrogen -LR clonase II enzyme Mix- (Cat. 11791-020) del sistema Gateway. Y se usó el producto de PCR como sustrato, a diferencia del protocolo del kit que sugiere usar a la construcción con el vector de entrada como sustrato.

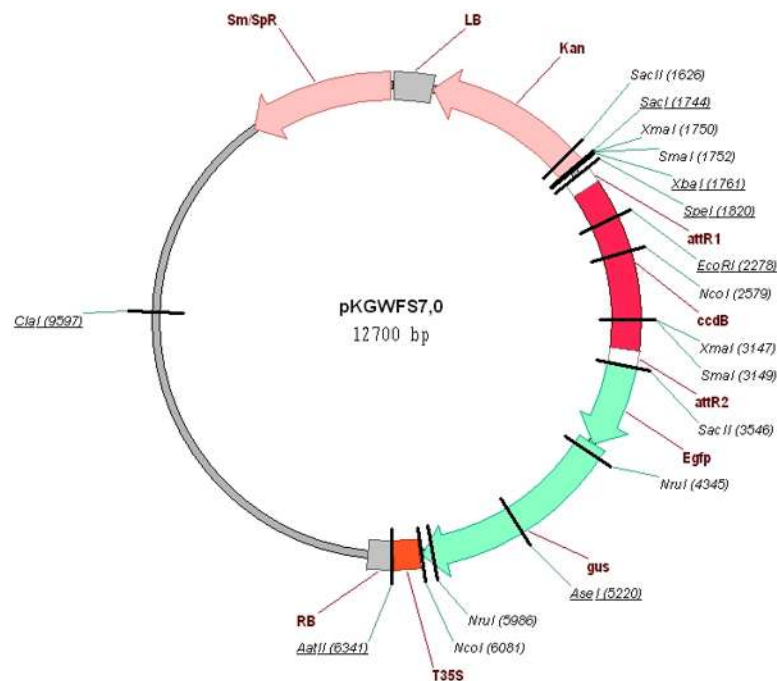


Figura 5. El vector binario pKGWFS7. Diseñado para el análisis de promotores regulando, en todos los casos, la expresión del gen reportero GUS-GFP. Este vector, además posee el gen *nptII* de resistencia a kanamicina bajo la regulación del promotor *nos*. Aparte posee el gen *ccdB*, que permite la selección negativa del mismo en *E. coli* DH5 $\alpha$  luego de la recombinación y transformación. Tanto el vector pKGWFS7,0 como pCR8/GW/TOPO poseen la resistencia al antibiótico espectinomocina.

## VI. RESULTADOS

Algunos escrutinios exitosos del sistema de tres híbridos han reportado diversas familias de proteínas de unión a ARN (Le *et al.*, 2000; Poleev *et al.*, 200; Xu *et al.*, 2001; Wood *et al.*, 2001). Para lograr el objetivo de identificar alguna proteína que interaccione específicamente con el 5' UTR del ARNm del gen *sps1* durante la traducción, se seleccionó el sistema de tres híbridos en *Saccharomyces cerevisiae* L40-*ura3*. La cepa se transformó introduciendo el plásmido pHybLex/Zeo-MS2 que expresa la proteína fusionada LexA-MS2 de unión a ADN y ARN correspondientemente, después se introdujo el vector pSPS-MS2 (López-Salmerón, 2007) que genera el ARN híbrido de interés. Posteriormente una biblioteca de ADNc de *A. thaliana* clonados en el vector pGAD10 fue introducida a la población de levaduras para realizar un escrutinio cuando el ARNX (5' UTR de *sps1*) interactúa con la proteína afin producida por el ADNc. Así, formando con las tres moléculas híbridas un puente transcripcional que activa el gen *HIS3* permitiendo que la levadura crezca en medio YC sin histidina, sin leucina, sin uracilo, con zeocina.

### VI.I. Análisis de la interacción del 5' UTR ARNm de *sps1* y proteínas codificadas en una biblioteca de ADNc de *A. thaliana*

Las células transformadas fueron cultivadas en 200 cajas de petri con medio sólido YC Zeo – HLU, después de cinco días crecieron aproximadamente  $6 \times 10^6$  colonias. Las cuales fueron cultivadas en 5 ml del mismo medio y sembradas en 100 cajas petri con medio sólido YC Zeo –HLU y 3-AT a 20 mM (3-amino 1,2,4- triazol, inhibidor competitivo de la enzima cifrada por *HIS3*) por 48 horas. En este lapso las primeras 200 colonias crecidas fueron escogidas para ser aisladas en cajas petri con YC Zeo –HLU 3-AT 20 mM. De ésta manera, se seleccionaron las primeras 33 cepas de levadura cepa L40-*ura3* con el mejor crecimiento en el medio de alta astringencia a las 48 horas (Cuadro 1). Se hicieron las extracciones de ADN total de cada una de las 33 cepas de levadura y se analizaron por electroforesis en geles de agarosa (Fig 1).

<b>Cuadro 1. Selección de colonias de levaduras y plásmidos</b>	
<b>Condiciones</b>	<b>Número</b>
Crecimiento en medio selectivo sin 3-AT a 120 h	≈ 6 000 000 colonias de levadura
Crecimiento en medio selectivo con 3-AT 20 mM a 48 h	200 colonias de levadura
Selección de colonias con mayor crecimiento en 3-AT 20 mM a 48 h	33 colonias de levadura
aislamiento de plásmidos cDNA's	15 plásmidos
Plásmidos con el DNAc en orientación correcta y adecuado posición para traducción	3 plásmidos
Plásmidos que confieren ensayo positivo de β-galactosidase	1 plásmido

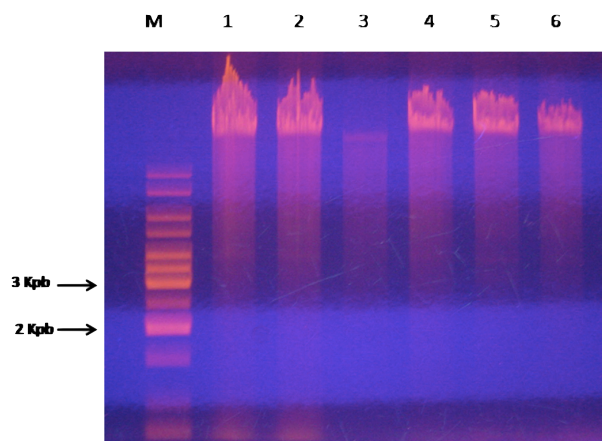


figura 1. **ADN total de seis colonias de levadura**; carril M, marcador de peso molecular; en la parte superior de los carriles 1-6 ADN cromosómico, mientras que en la parte media están los plásmidos introducidos. La mezcla de ADN fue posteriormente utilizada para transformar *E. coli* y así seleccionar vectores derivados de la biblioteca de ADNc.

Posteriormente, el ADN total de cada una de las 33 cepas de levadura conteniendo los plásmidos pSPS-MS2 que genera el ARN híbrido y el plásmido pGAD10 de la biblioteca genómica de ADNc de *A. thaliana*, fue utilizado para transformar a *E. coli* cepa DH5α y seleccionar en medio LB con carbenicilina, considerando que éstos vectores otorgan resistencia a este antibiótico, a excepción de pHybLex/Zeo-MS2. A las colonias de bacterias crecidas del medio selectivo se les extrajeron sus plásmidos y con cortes con la enzima *EcoR* I se pudo diferenciar cual plásmido había sido ingresado a la bacteria, en base a que en el vector

pSPS-MS2 un sitio para *EcoR* I flanquea el ARNX (5' UTR *sps1*) más el RNA MS2, liberando un fragmento de 440 pb y al vector de 7907 pb; mientras que en el vector pGAD10, el corte con *EcoR* I libera al vector de 6.65 kb y un fragmento de tamaño desconocido de ADNc. En la Fig. 2 se muestran los plásmidos de bacterias transformadas con el ADN total de una de las cepas de *S. cerevisiae* denominada 79; en el carril 7 la banda superior corresponde al vector pGAD10 de 6.65 kb y las otras dos bandas corresponden al ADNc con un tamaño aproximado de 2 kb y 700 pb. Con base en el tamaño del ADNc, éste fue uno de los 15 diferentes plásmidos considerados como candidatos para portar un ADNc que codificara para una proteína de unión a RNA, provenientes de cepas (Cuadro 1).

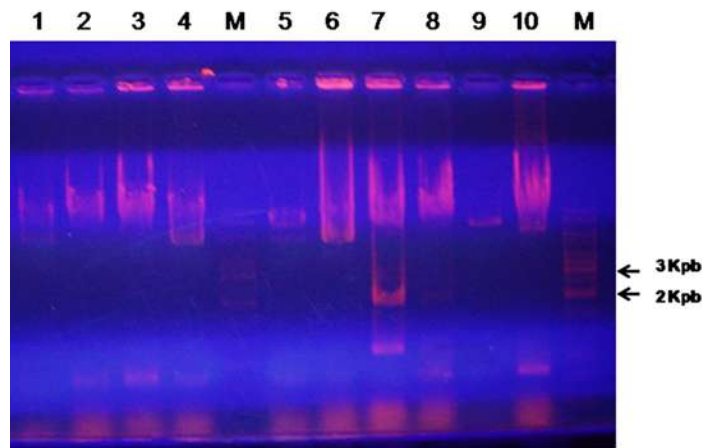


Figura 2. **Purificaciones de los plásmidos de *E. coli* transformados con ADN total de la cepa 79 de levadura.** El carril M corresponde al marcador molecular. Los demás carriles son plásmidos extraídos y cortados con la enzima *EcoR* I. Las escisiones que liberan pedazos de alrededor 440 pb iguales al tamaño del ARNX del vector pSPS-MS2 fueron descartados. En el carril 7 hay un banda adecuada a pGAD10 de 6.65 kb y otras dos de 2 kb y 700 pb consideradas como ADNc.

Los plásmidos pGAD10 seleccionados fueron purificados y secuenciados. Se obtuvieron 9 secuencias de ADNc, las cuales fueron alineadas con la base de datos de *A. thaliana* del National Center of Biotechnology information (NCBI) utilizando la herramienta BLASTn. En el cuadro 2 se presenta la información obtenida. Se determinó que cinco de los ADNc estaban clonadas en orientación inversa, y de las cuatro restantes, tres estaban en marco de lectura correcto, siendo los loci NM\_125810, NM\_117319 y NM\_117334.

Cuadro 2. Resultados del BLAST, de las secuencias de los ADNc elegidos, contra el genoma de *A. thaliana* en la base de datos de NCBI

Locus	ARN codificado	Observaciones de la clonación	Observaciones de la proteína
NM_117334	ARNm lineal de 2802 pb	Orientada correctamente en el vector pGAD10 y en el marco de lectura adecuado	Probable proteína
NM_125810	ARNm lineal de 480 pb	Orientada correctamente en el vector pGAD10 y en el marco de lectura adecuado	Proteína ribosomal S28; componente estructural del ribosoma
NM_117319	ARNm lineal de 786 pb	Orientada correctamente en el vector pGAD10 y en el marco de lectura adecuado	Inhibidor de proteasa/ almacenamiento de semillas/ transferencia de lípidos. De la familia LTP
NM_122009	ARNm lineal de 1049 pb	Orientada correctamente en el vector pGAD10. Con marco de lectura incorrecto para la traducción	Proteína nuclear relacionada con Ras
NM_128268	ARNm lineal de 2501 pb	Orientada incorrectamente en el vector pGAD10	Factor de transcripción SE
NM_120116	ARNm lineal de 1513 pb	Orientada incorrectamente en el vector pGAD10	De la familia F-box
NM_129704	ARNm lineal de 1157 pb	Orientada incorrectamente en el vector pGAD10	Probable proteína
NM_120952	ARNm lineal de 2300 pb	Orientada incorrectamente en el vector pGAD10	Probable proteína LepA de enlace-GTP
NM_117188	ARNm lineal de 737 pb	Orientada de manera incorrecta en el vector pGAD10	Factor de iniciación IF-1

El locus NM\_125810 corresponde al gen RPS28 que codifica la proteína S28 de 64 aminoácidos. Tal proteína funciona como constituyente estructural del ribosoma 40s y contiene un dominio de unión a ARN tipo S1 encontrado en una amplia variedad de proteínas asociadas a ARN. La proteína S28E, altamente conservada en arqueas y eucariontes, puede controlar la exportación del ARNr del núcleo, la biogénesis del ribosoma, la traducción el splicing del ARN y la cantidad del proceso de maduración del ARN, sin embargo, su mecanismo en el ribosoma es desconocida (Chang *et al.*, 2005).

El locus NM\_117319 codifica para una proteína de 182 aminoácidos bifuncional, como inhibidor de amilasa o como proteínas que transfieren lípidos, perteneciente a la subfamilia HPS (Hydrophobic Protein from Soybean), esta clase de proteínas defienden a la planta de

insectos y patógenos, transfieren lípidos entre membranas intracelulares y almacenan nutrientes en la semilla.

La proteína proveniente del ADNc de la cepa 79 de levadura y registrada en el locus NM\_117334 se identificó como el gen At4g12640 de *A. thaliana* con 4 exones y 3 intrones, su péptido se conforma por 823 aa (NP\_193001.2) y sin función conocida.

Para determinar que la interacción entre las proteínas codificadas por los ADNc fuera específica, se analizó la expresión del gen reportero *LacZ*, el cual se activa cuando interactúa el ARN y la proteína. En la figura 3 se presentan los resultados para la proteína 79 y la ribosomal S28. La interacción entre la proteína 79 y el ARN híbrido activó en gen *LacZ* además de proporcionar buen crecimiento de la cepa en el medio YC Zeo -HLU 3-AT 20 mM, en tanto que la proteína 43 no activó el gen reportero *LacZ* y el crecimiento en el medio selectivo fue mas lento. La proteína codificada por el locus MN\_117319 tampoco activó el gen reportero (no mostrado). Estos ensayos colorimétricos refuerzan las evidencias de la formación del puente transcripcional de las tres moléculas híbridas e interacción con la RNA polimerasa II en la cepa 79 con el plásmido que alberga el gen codificante para la proteína cognado.

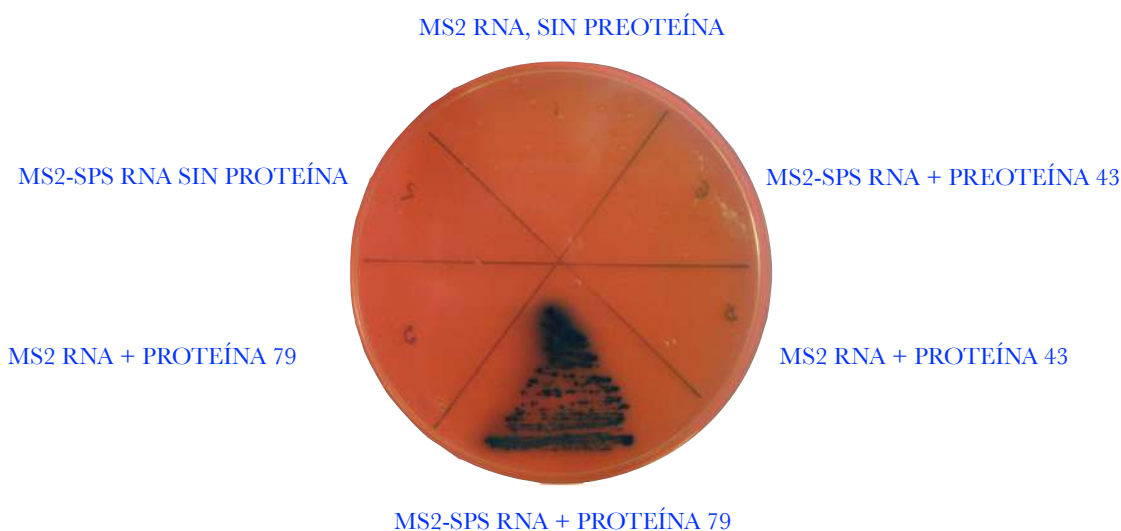


Figura 3. **Actividad positiva del gen *LacZ* en la cepa 79 de *S. cerevisiae*.** El medio YC Zeo –HLU contiene 3-AT a 20 mM y X-gal/dimetilformamida (2.5  $\mu$ g/ml). Solamente la cepa 79 activó el gen *LacZ* concomitante con el gen reportero *HIS3*.

## **VI.II. Análisis funcional de la proteína codificada por el gen At4g12640.**

La proteína NP\_193001.2 codificada por At4g12640 pertenece a la familia Spen (Split ends) caracterizada por un dominio RRM (RNA recognition motif) N-terminal y un dominio bien conservado C-terminal SPOC (Spen paralog and ortholog C-terminal). El dominio RRM es conocido como RBD (RNA binding domain) o RNP (Ribonucleoprotein domain), consta de 90 aa y tiene una estructura de cuatro laminas beta empaquetadas contra dos hélices alfa que interactúan con ARNss y ADNss como también con proteínas. RRM es encontrado en una variedad de proteínas hnRNPs (heterogeneous nuclear ribonucleoproteins) implicadas durante el splicing alternativo y proteínas componentes snRNPs (small nuclear ribonucleoproteins). En cuanto al dominio SPOC contiene alrededor de 165 residuos de amino ácidos y exhibe bajo grado de homología en la secuencia de aa de la región estructuralmente conservada (Shi-Yan *et al.*, 2007) que, consta de un núcleo de siete láminas beta formantes de un barril central y en algunos casos dos hélices pequeñas (Sánchez-Pulido *et al.*, (2004). SPOC está implicado en la señalización del desarrollo, facilitando la interacción proteína-proteína en el complejo de represión transcripcional. Sánchez-Pulido *et al.* (2004) utilizaron métodos bioinformáticos para identificar genes que codifican proteínas similares a las Spen en protozoarios y plantas, encontrando que las proteínas Spen podrían jugar un rol en la muerte celular programada y la generación de tumor, por lo tanto proveyendo una nueva función para el dominio SPOC. Las proteínas Spen participan en diferentes procesos biológicos como la diferenciación de la célula neuronal, sobrevivencia y orientación del axón (Chen y Rebay, 2000; Kuang *et al.*, 2000), regulación del ciclo celular (Lane *et al.*, 2000), represión en la identidad de la cabeza del tronco embrionario de *Drosophila* (Wiellette *et al.*, 1999), así como en la actuación en la señalización de la proteína Wingless (Wg) en los discos imaginales del ojo, alas y piernas de la mosca (Lin *et al.*, 2003). La proteína Spen humana SHARP (SMRT7HDAC1-associated represor proteína) ha sido identificada como componente del complejo de represión transcripcional tanto en el receptor nuclear como en las vías de señalización Notch/RBP-jkappa; por consiguiente las proteínas de la familia Spen parecen regular la transcripción mediante algunas vías de señalización (Sánchez-Pulido *et al.*, 2004).

En el mismo BLASTp de NCBI utilizado fueron halladas otras 40 proteínas Spen de plantas y 2 de animal (*Ailuropoda melanoleuca*) de una identidad mínima del 34% con respecto a la proteína del gen de *A. thaliana* At4g12640. Razonando que el ARNX de interés o cebo en el sistema de tres híbridos es de *Oriza sativa* subespecie *indica* y la biblioteca de ADNc de *A. thaliana*, se buscaron proteínas Spen de arroz con homología a la codificada por *A. thaliana*. Fueron identificadas dos proteínas Spen de arroz *indica*: una de 1005 aa también sin función conocida (EEC84301.1, etiqueta del locus OsI\_30787) con una identidad del 73% a la proteína de *A. thaliana*, junto con otra proteína hipotética con función desconocida (EEC84885.1, locus OsI\_32043) de 900 aa y 40% idéntica a la proteína de *A. thaliana*. El locus OsI\_30787 de *O. sativa* ssp. *indica* se localiza en la hebra reversa del cromosoma 9 y corresponde al gen OJ1381\_H04.27 conformado de tres exones y dos intrones, es homólogo (100%) al gen Os09g0298700 de *O. sativa* ssp. *japónica* nombrado *OsRRM* por Shi-Yan *et al.* (2007). El locus OsI\_32043 contiene al gen P0450E05.35 orientado en la hebra reversa del cromosoma 9, estructurado por seis exones y cinco intrones, el gen tiene un 100% homólogo en arroz ssp. *japónica* llamado Os09g0516300.

Particularmente, las proteínas NP\_193001.2 de *A. thaliana*, EEC84301 y EEC84885 ambas de *Oriza sativa* subespecie *indica*, llevan dos dominios RRM N-terminal y un dominio C-terminal SPOC. En la figura 4 se presenta la alineación de los dos dominios RRM N-terminal y el dominio SPOC alrededor de la parte media de los genes, realizado con la herramienta BLAST2seq de NCBI usando la proteína del gen At4g12640 contra las proteínas de arroz *indica* de los genes OsI\_30787 y OsI\_32043:

#### Primer RRM

```

OsI_32043 38-122      GGSGAPETNTLWVGNLPAQAAEDDVMAAFSPHGALDCVMAR
                   P + LWVGNLP E ++ F G L+ + +
At4g12640 15- 98      GRGRNPPSRHLWVGNLPHGILERELADRFLRFGELESIAFQ
                   G G PPSRHLWVG+L G+ +L++ FLR G++E ++
OsI_30787 41-124      GWGVAPPSRHLWVGSLSPGVAAADLSELFLRCGDVEGISRD

AGPRS YAFVLF RSVPEARAAALDALQGSKVKGS VVRLEFARPARA
G RSYAFV F +A AA+++LQG + G+ +R+EFA+ +
PG-RS YAFVNFNHDEDAFAAIESLQGFPLSGNPLRIEFAKAEKS
PG-RS+AFV F +EDA A + LQG L G P+RIEF+K +K
PG-RS FAFVTFAREEDAAVRELQGIHLRGAPIRIEFSKGDKG

```

## Segundo RRM

```
OSI_32043 220-303      VRNRESNPTNVLWVGFPNTAKINEEALRQAMAVHGAVTNTKVF  
+ +R + P+ VL++GFP  K+++ LR  + G +T  VFP  
at4g12640 142-225    MNDRNAEPSEVLYIGFPASLKVDDALLRNVFSFGEITKVTVFP  
R AEPSEVL+IGFP  LKVD+A L  FS FGE+ K+T FP  
OsI_30787 161-244    KRSRPAEPSEVLWIGFPVGLKVDEATLWEAFSPFGEVVKITTFP  
  
TRQYAFVEFATVGEASNAKKNLDGRLFNDRIQILFSNSE  
R YAFV+F + A AK++L G+LF + R+ I F+ SE  
GRS YAFVQFRNLMAACKAKESLQGKLFGNPRVHICFAKSE  
GR+YAFVQ+ + AAC+AKE+LQG LF NPRV ICF++S+  
GRTYAFVQYTTIAAACRAKETLQGNLFNPNRVSICFSRSD
```

## SPOC

```
VHRTSYLE-HFWRGSIAKGGSPVCRARCLPIRKGVEIPLPDVVNCSARTGLDMLAK  
++R S + W G+IAKGG+P+CRA+C P+ K +++ LP+ ++C+ARTGLDMLAK  
LNRPSLKD-WNWEGTIAKGGNPICRAKCFPVGKVMMDMLPEFLDCTARTGLDMLAK  
+ P LK+ W W+GTIAKGG PICRA+CFPVGKV++ MLPEFLDCTART L+ML+K  
YHEPLLKEEWKWDGTIAKGGTPICRARCFPVGKVLNFMMLPEFLDCTARTSLEMLSK  
  
HYRDASGFDIVFFLPDSEDDFVSYTEFLRYLGSKSRAGVVKVDGTTFLFLVPPSDFLRNV  
HY +S +VFF+P S+ D V Y EF+ YL K RA V K+D TTLFLVPPSDF V  
HY YQSSKAWVVFFVPGSDADIVFYDEFMHYLEEKQRAAVSKLDDTTTLFLVPPSDFSEKV  
HYYQ++ +WVFFVFP +DAD+ Y+EFM+YL +KQRAAV KL + ++LFLVPPSDFSE+V  
HY YQAASSWVVFFVPE NDADMAAYNEFMNYLGDKQRAAVCKLGERSSFLVPPSDFSEQV
```

```
LQVDGPERLYGVVLHIP OsI_32043 430-561  
L+V G + GV+L +  
LKVPGKLSISGVILRLE at4g12640 461-592  
L+VPGK+SISGVIL+ E  
LRVPGKVSISGVILKFE OsI_30787 487-619
```

Figura 4. Comparación de aminoácidos de los dominios RRM y SPOC del gen de *A. thaliana* At4g12640 contra los de sus ortólogos OsI\_30787 y OsI\_32043 de *O. sativa* ssp. *indica*. La proteína de *A. thaliana* comparte en sus dominios RRM1, RRM2 y SPOC una identidad de 49, 65 y 72 % correspondientemente, con la proteína de OsI\_30787. En tanto que, la misma proteína tiene una identidad de 38, 42 y 57 % en relación a los mismo dominios con la proteína del gen OsI\_32043.

Para conocer un poco más acerca de la posible función del gen At4g12640 como al mismo tiempo de los genes de arroz indica, se tomó de referencia al gen OJ1381\_H04.27 considerando que es 100% homólogo a *OsRRM* de arroz japónica reportado con actividad tejido específico en la regulación del desarrollo de las células especializadas del endospermo y polen. La proteína citada se localiza en el núcleo, lo que indica una función en la

transcripción del RNA o durante su procesamiento (Shi-Yan *et al.*,2007). Inclusive, una alineación con BLAST2seq de NCBI del promotor de OJ1381\_H04.27 contra el promotor del gen *OsRRM* corroboró la idea de que también los promotores de estos genes son homólogos. Aunque las proteínas ofrecen información sobre la posible función del gen, las regiones reguladoras de los genes aportan información sobre la regulación de la expresión. Se decidió analizar y comparar elementos reguladores conservados entre los promotores de arroz indica y *A. thaliana*.

PLACE (<http://www.dna.affrc.go.jp/htdocs/PLACE/> ) es una base de datos de motivos de secuencias nucleotídicas encontrados en elementos reguladores actuantes en *cis* del ADN de plantas (Higo *et al.*, 1999). Se analizaron 500 pb del promotor de los genes de arroz indica, Os09g0516300 y OJ1381\_H04.27, y de *A. thaliana* At4g12640, los cuales fueron sometidos a escaneo de señales. Los resultados ubicaron diferentes elementos de secuencias de unión a factores transcripcionales o de control en la expresión genética (anexo 1); entre los promotores de At4g12640, OJ1381\_H04.27 y P0450E05.35 hubo 6 elementos comunales, a saber; CAATBOX1 responsable de la actividad promotora tejido específico en semilla de guisante, CACTFTPPCA1 elemento *cis* para la expresión específica en mesófilo de *Flaveria trinervia*, DOFCOREZM sitio requerido para la unión de proteínas Dof en maíz, GTGANTG10 motivo encontrado en el promotor del gen *g10* de polen de tabaco, NODCON2GM uno de las dos secuencias consenso putativas de nodulación para soja (*Glycine max*), WRKY71OS en arroz es sitio de unión para el represor transcripcional de la ruta de señalización de giberelinas. A parte de los elementos *cis* anteriores, el gen At4g12640 con su ortólogo OJ1381\_H04.27 y *OsRRM* con expresión en endospermo y polen (Shi *et al.*,2007), comparten exclusivamente los motivos: ARR1AT o elemento de unión ARR1, GATABOX requerida para la expresión de alto nivel, regulada por luz y tejido específico, GT1CONSENSUS sitio de unión GT-1 en algunos genes regulados por luz, IBOX secuencia conservada corriente arriba de genes regulados por luz, POLLEN1LELAT52 uno de dos elementos reguladores co dependientes para la activación específica en polen del gen *Lat52* en tomate, ROOTMOTIFTAPOX1, por último, SEF4MOTIFGM7S sitio de unión para el factor SEF4.

La caja CAAT situada a -97, corriente arriba de la caja TATA en el gen *legA* de leguminosa, es responsable de la actividad tejido específico (Shirsat *et al.*, 1989). El gen codifica para la proteína mas importante de almacenamiento en la semilla. Estas proteínas son usadas como fuente de nitrógeno en plántulas en desarrollo, existen dos clases principales de proteínas de almacenamiento **en la semilla** de la leguminosa; 11 S leguminas y 7 S vicilinas, la síntesis de estas proteínas empieza en la mitad del desarrollo de la semilla hasta que comprende del 60%-80% de la proteína en la semilla madura (Shirsat *et al.*, 1989). La caja CAAT del gen de At4g12640 está en -81 del promotor, para P0450E05.35 de arroz se localiza a -150 del promotor y en OJ1381\_H04.27 va en las inmediaciones del inicio de la transcripción.

El factor transcripcional Dof se presenta únicamente en plantas y presentan un dominio de 52 residuos de aa altamente conservado de unión para la secuencia de reconocimiento AAAG; se ha propuesto que el objetivo del activador transcripcional Dof1 de maíz podría ser el promotor del gen fosfoenolpiruvato carboxilasa tipo C4 (C4PEPC), aún mas, la proteína Dof específica de endospermo de maíz, PBF, se une a la caja *cis* prolamina responsable de la expresión génica específica **para endospermo** del gen zeina. También, PBF puede interactuar con la proteína O2 (del tipo bZIP, basic leucine zipper) regulador de la expresión **en endospermo**, indicando que puede ser un regulador transcripcional para la expresión de genes en endospermo (Yanagisawa y Schmidt, 1999). En los promotores de los genes At4g12640 y OJ1381\_H04.27, la secuencia AAAG se encuentra en la misma posición -32 y a -70 del gen P0450E05.35, en el caso de C4PEPC se haya cerca del sitio de inicio de la transcripción. Aludiendo que, las proteínas de arroz y de thaliana podrían estar incluidas en una variedad de respuestas a señales y/o la expresión génica tejido específico en semilla. El gen  $\beta$ -conglucina subunidad- $\alpha'$  codifica para la subunidad  $\alpha'$  de la proteína de almacenamiento de semilla en soja  $\beta$ -conglucina compuesta de tres subunidades;  $\alpha$ ,  $\alpha'$  y  $\beta$ , las proteínas son expresadas por genes separados y expresados exclusivamente en diferentes tiempos de la maduración de la semilla, la secuencia RTTTTTR se ha encontrado en la posición -105 del gen  $\beta$ -conglucina subunidad- $\alpha'$ , con función potenciadora de la transcripción específica **en semilla**, éste elemento de secuencia une al factor nuclear SF4 que probablemente sea un proteína activadora de la transcripción (Allen *et al.*, 1989). Casualmente, la secuencia GTTTTTG en el gen OJ1381\_H04.27 de arroz está en -245 del promotor y en At4g12640 a -360 en el promotor.

El gen *OsWRKY71* sintetiza un represor transcripcional que puede inhibir la transducción de señales en respuesta a la hormona giberelina mediante la unión a la secuencia TGAC en genes de respuesta a la hormona giberelina (GA) como *Amy32b* ( $\alpha$ -amilasa) o  $\alpha$ -*Amy2* en células de **aleurona de la semilla de arroz**, al mismo tiempo se propone que OsWRKY71 interfiera con la interacción del activador DOF y GAMYB. La unión del represor OsWRKY71 a los elementos actuantes *cis* del promotor, podría atenuar la expresión génica al interactuar con los dominios de activación de los activadores transcripcionales y bloquear el contacto con maquinaria basal de transcripción o competir por los componentes de la maquinaria transcripcional (Zhang *et al.*, 2004).

El tetramero YACT es un componente clave del módulo Mem1 (mesophyll expression module) para la expresión específica **en mesófilo** del promotor del gen fosfoenolpiruvato carboxilasa, *ppcA1*, isoforma C4 de *Flaveria trinervia*, la secuencia T/CACT se encuentra en la parte distal del promotor esto decir a -1950 pb del promotortr (Gowik *et al.*, 2004). Encontrado tanto en los promotores de los genes de arroz como de thaliana, aunque no en la posición de *Flaveria trinervia*.

El elemento GATA en el promotor de genes *cab* es requerido para la actividad génica en **respuesta a luz** (Gilmartin *et al.*, 1990). Experimentos realizados en tabaco descubrieron al factor nuclear ASF-2 al unirse al motivo GATA, asimismo que per se no es suficiente para conferir respuesta a luz en genes *Cab* y *rbcS*, sino que incluso es probable que el factor ASF-2 medie una expresión hoja específica de igual forma ya que no hay actividad de unión en extractos de raíz (Lam y Chua, 1989). En los genes de *A. thaliana* At4g12640 y arroz indica OJ1381\_H04.27 existen varios de estos elementos GATA, habiendo uno en -91 y -95 de sus respectivos promotores, comúnmente encontrados hasta la posición -150 de los genes *cab*, así como en el promotor de 35S CaMV (Lam y Chua, 1989).

Una característica común de todos los elementos GT es una secuencia consenso amplia (G/T-A/T-GTG-R-W-AA-W-R-A/T) rica en los nucleótidos T y A más que G y T, las secuencias flanqueantes pueden ser diferentes. Existen proteínas GT-1 y GT-2 con motivos tri-helicoidales de unión a ADN únicos de plantas, los cuales reconocen elementos degenerados

de GT. GT-1 puede **estabilizar el complejo** TFIIA-TBP sobre el ADN a través de la interacción de TFIIA, funcionando como un co-activador que se une a GT-1 (Le Gourrierec *et al.*, 1999). Puede haber diferentes proteínas GT con diferentes funciones específicas, como para el gen de guisante *rbcS-3A* se registró que el sitio de unión identificado como Box II (GGTTAA) juega un rol importante en la transcripción regulada **por luz** (Green *et al.*, 1987). Otros elementos *cis* de unión a GT han sido implicados en la expresión tejido específico como en la expresión en polen del gen *lat52*, relacionados en la defensa (sitio de unión SBF-1 en el gen *chs15* de soja), y control del ciclo circadiano (sitio CGF en *cab2*) (Villan *et al.*, 1996). En el presente trabajo se encontró que en ambos genes de *A. thaliana* At4g12640 y *O. sativa* ssp. *índica* OJ1381\_H04.27 existe los elementos de unión al factor GT-1 como GAAAAA (GRWAAW). La caja GAAAAA en los genes Spen ortólogos OJ1381\_H04.27 y Atg12640 se encuentran repetidos y distribuidos, lo que sugiere una función redundante y reguladora de la transcripción por luz como la del gen de guisante *rbcS-3A* (Green *et al.*, 1987).

De forma similar, en los promotores de OJ1381\_H04.27 y At4g12640 que en el promotor del gen RBCS de *thaliana* y tomate el elemento I Box, GATAAG, es una secuencia necesaria y suficiente para la unión de GBF (G Box binding factor) distinta del motivo GT de unión para GT-1, encontrada corriente abajo o 3' de alguna de las secuencias redundantes para unión a GT-1; teorizando una acción recíproca física entre GBF y GT-1. Elevados niveles de actividad de unión de GBF fueron detectados en las hojas y el tallo mientras que una baja actividad se encontró en raíz (Giuliano *et al.*, 1988). La secuencia I box es funcional para la expresión regulada por **luz y tejido específico** del promotor de *rbcS-1A* de *A. thaliana*, específicamente es sitio de unión del factor GA-1 (Donald RGK y Cashmore AR, 1990). Implicando la significancia funcional de I Box en los genes de arroz y *A. thaliana* de que, diferentes proteínas de unión a ADN participan en la expresión regula por luz.

El gen OJ1381\_H04.27 contiene en la misma posición -97 del promotor que el gen *g10* de tabaco la caja GTGA, director de la expresión específica **en polen** del último gen y putativo homólogo del gen *lat56* de tomate. Un número de otros genes son coordinadamente expresados para la utilización de la pectina en el desarrollo del polen, como genes tipo pectato liasa *lat52*, *lat56* y *lat59*, con 54% de identidad en aminoácidos entre los últimos dos, y en su

promotor comparten la caja 56/59 desde -120 hasta -87 de sus promotores donde comparte la secuencia GTGA (Rogers *et al.*, 2001). La secuencia GTGA de los genes *g10* de tabaco y *lat56/59* de tomate es considerada con la misma función (Rogers *et al.*, 2001). El dato presente aquí revela la conservación de secuencias regulatorias a través de las especies para los genes homólogos, hemos descubierto que el motivo GTGA, presente en la caja LAT56/59 de los genes de tomate *lat56/59* como en el homólogo de tabaco *g10*, justamente se emplaza a -62 del promotor en At4g12640 y a -125 del promotor de P0450E05.35 de arroz indica, indicando conservación en la unión a un putativo factor transcripcional. El motivo de secuencia AGAAA se encuentra en At4g12640 a -86 pb del promotor, en cuanto al gen de arroz indica OJ1381\_H04.27 la secuencia está a -25 del promotor, consistentemente, el elemento AGAAA se halla en la posición -68 pb del promotor del gen *lat52* de tomate, y es uno de dos elementos *cis* co-dependientes reguladores de la transcripción específica en **polen** (Twell *et al.*, 1991). También en tomate, la secuencia es encontrada en el promotor del gen endo- $\beta$ -manasa llamado *LeMAN5*, el cual se expresa en polen, anteras y flores abiertas, pero no en raíz, tallo, hojas y silicuas (Filichkin *et al.*, 2004).

Las nodulinas podrían funcionar en, y estructurar la membrana peribacteroide de las células infectadas, la secuencia CTCTT está presente en las nodulinas 20,22,23 y 44 alrededor de la posición -130 pb y alrededor de la posición -120 pb de los genes *Lb* (Sandal *et al.*, 1987). Tal elemento es necesario para el control de la expresión nódulo específica (Stougaard *et al.*, 1990). Durante la simbiosis de soja (*Bradyrhizobium japonicum*) además de los genes leghemoglobin *Lb*, otros genes codificados por el hospedero son específicamente expresados **en el nódulo**, las nodulinas incluyen las formas nódulo específicas de uricasa, glutamina sintasa y sacarosa sintasa (Sandal *et al.*, 1987). En el promotor de At4g12640 CTCTT yace en el lugar -363 y en el promotor del gen de arroz OJ1381\_H04.27 yacen tres repeticiones de esa secuencia colocados en -177, -222 y -417 pb, en tanto que, en el otro gen de arroz P0450E05.35 hay 6 de repeticiones a -120, -137, -156, -205, -288 pb del promotor. Sugiriendo que estos elementos también estén implicados en la activación de estos genes.

En otra comparación, cuatro secuencias consenso más se hallaron entre los promotores de At4g12640 y P0450E05.35, siendo éstas; CCAATBOX1 caja localizada en el promotor de

genes heat sock que actúan en cooperación con HSEs para incrementar la actividad del promotor hs, OSE2ROOTNODULE una de las secuencias consenso OSE (organ-specific elements) característico de los promotores activados de células infectadas en los nódulos de raíz, SITEIIATCYTC elemento “site II” encontrado en el promotor de genes nucleares, SORLIP2AT secuencia representada en los promotores inducidos por luz en *A. thaliana*.

Mientras que, los genes de arroz indica, OJ1381\_H04.27 y P0450E05.35, en total tienen 9 sitios iguales en su promotor, los faltantes son; CURECORECR es el motivo principal del elemento en respuesta a cobre (CuRE) involucrado en la respuesta a oxígeno de estos genes de clamidomonas, EBOXBNNAPA regulador transcripcional de genes de almacenamiento en semilla, MYCCONSENSUSAT regula la transcripción de genes en el frío.

### VI.III. Aislamiento de diferentes versiones de tamaño del promotor del gen At4g12640 a partir del genoma de *A. thaliana*.

Debido a que el gen At4g12640 descubierto mediante el sistema de tres híbridos fue de una biblioteca genómica de *A. thaliana*, se decidió aislar la región del promotor y clonarla en un vector para posteriores análisis *in vivo*. Para lo anterior, se diseñaron oligonucleótidos que amplificaron la región 5' corriente arriba hasta -490 y -981 pb desde el atg de inicio de la traducción, la Fig. 5 muestra los productos de PCR.

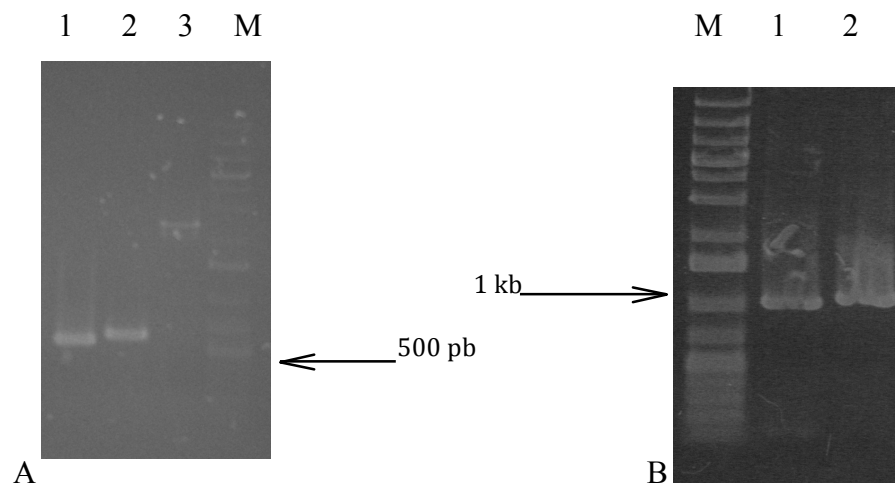


Figura 5. **Amplificación de la región promotora del gen At4g12640.** Se diseñaron oligonucleótidos para obtener el promotor de At4g12640 usando el ADN genómico de *A. thaliana* como templado. La banda de los carriles 2 en ambas fotos son los productos de PCR de interés; en la foto A amplificación de 588 pb; la foto B amplificación de 1.079 kb; M es el marcador molecular.

Los productos de PCR se clonaron en el vector de entrada Gateway pCR8/GW/TOPO, Fig. 6, el cual tiene los sitios *attL1* y *attL2*, que pueden recombinarse con los sitios *attR1* y *attR2* del vector binario destino que lleva los sitios LB y LR para que *Agrobacterium tumefaciens* transfiera el T-DNA durante la transformación de plantas. También fueron secuenciados los productos de PCR clonados para comprobar que se hayan colocado en el sentido correcto, de 5'-3' en relación al mismo sentido del vector.

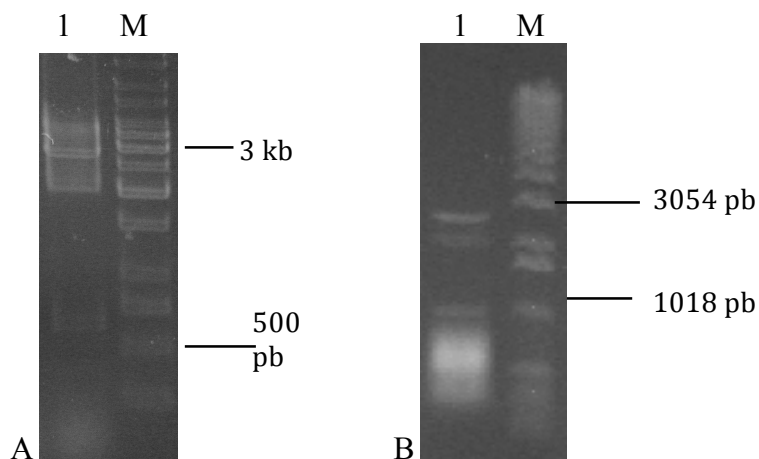


Figura 6. **Clonación de las versiones del promotor del gen At4g12640 en el vector de entrada pCR8/GW/TOPO.** Los fragmentos generados por PCR de las versiones del promotor se clonaron en el sitio *EcoR* I del vector de entrada pCR8/GW/TOPO. Los plásmidos fueron aislados de colonias transformadas y cortados con *EcoR* I para liberar en producto de PCR clonado. En la columna 1 en la foto A está el fragmento de 580 pb y en B de 1.079 kb, en ambas fotos la banda superior es el vector de entrada lineado de 2.817 kb; M es el marcador molecular.

La construcción con el vector pCR8/GW/TOPO se usó como templado para amplificar las versiones del promotor At4g12640 junto con los sitios de recombinación *attL1* y *attL2* con los primers M13, tales productos de PCR se usaron en la reacción de recombinación dirigida hacia el vector destino o binario pKGWFS7.0 (Fig 7) del sistema Gateway, el cual tiene los bordes

LB y LR para transferir la construcción con el gen reportero a plantas de *A. thaliana*. La reacción de recombinación se llevó acabo con el kit –LR Clonase II Enzyme Mix–.

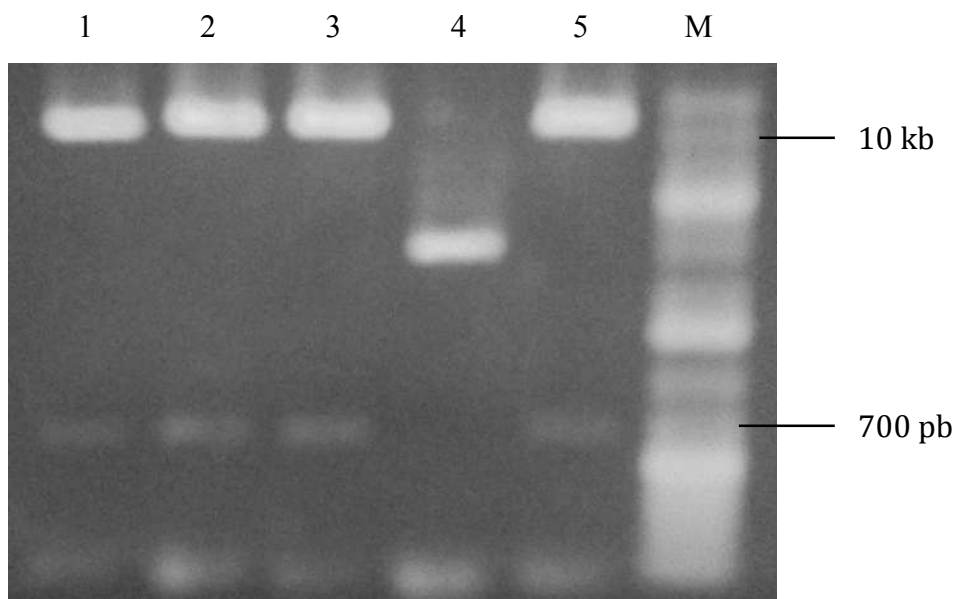


Figura 7. **Recombinación entre el vector de entrada y el vector destino pKGWFS7.** El fragmento del promotor hasta -490 pb de At4g12640 insertado en el vector pCR8/GW/TOPO se amplificó por PCR y se clonó al vector pKGWFS7 con el método de recombinación de sitios específicos del sistema Gateway. En los carriles 1-5, plasmidos extraídos de bacterias transformados con la construcción en el vector pKGWFS7 y cortados con *EcoR* I. Excepto en el carril 4, la banda media es el fragmento estudio de 588 pb liberado del vector pKGWFS7 que es la banda superior de 12700 pb. M es marcador molecular.

#### VI.IV. Discusión

El sistema de tres híbridos se usó para encontrar una proteína que interaccionara con la región líder del ARNm del gen *sps1* de arroz. Al carecer de una biblioteca genómica de arroz con fusiones de ADNc a un dominio de activación transcripcional, se utilizó una biblioteca de ADNc de *A. thaliana*. Fue posible encontrar una interacción específica del ARN líder de *sps1* con una proteína Spen codificada por el gen At4g12640. Considerando que la región líder tiene actividad potenciadora de la traducción tanto en sistemas *in vivo* en *A. thaliana* como *in vitro* en un sistema de extractos de germen de trigo (Martínez-Trujillo et al., 2003), se

consideró como una posibilidad que el efecto del líder fuese a través de la interacción de una proteína, como ha sido reportado para el líder del virus del mosaico del tabaco (Wells *et al.*, 1998). La interacción del ARN líder *sps1* con tal proteína abre muchas preguntas respecto a si es ésta la responsable de la potenciación de la traducción y si es así ¿de qué manera lleva acabo este proceso?.

Mediante la herramienta BLASTp de NCBI se encontraron dos proteínas Spen ortólogas de *O. sativa* ssp. *indica* en referencia a la proteína del gen At4g12640; una proteína con 73% de identidad codificada por el gen OJ1381\_H04.27 y la otra proteína 40% idéntica elaborada por el gen P0450E05.35. Sólo se sabe de estas proteínas que por sus características Spen, en la parte de sus dominios RRM podrían interaccionar proteína-ARN como lo hizo la codificada por At4g12640 con el 5' UTR de *sps1* en el sistema de tres híbridos del presente trabajo y por la parte del dominio SPOC interaccionar proteína-proteína como la hace la proteína humana Spen SHARP al unir a SMRT, un co-represor transcripcional que recluta a la histona desacetilasa (HDAC) (Lin *et al.*, 2003).

La expresión tejido específica en endospermo y polen de arroz japónica del gen Spen *OsRRM*, está programada por la distribución de los elementos de secuencia o motivos actuantes en *cis* de su promotor (Shi *et al.*, 2007). Ya que el gen *OsRRM* de arroz japónica es homólogo tanto en la secuencia codificante del gen como en promotor a OJ1381\_H04.27 de arroz índica, se realizó un análisis comparativo del promotor de éste último de índica con el otro gen P0450E05.35 de índica y el ortólogo en *A. thaliana* At4g12640.

En general, en los promotores de los genes At4g12640, OJ1381\_H04.27 y P0450E05.35 se hallaron elementos para expresión en tejido específico de semilla, así como secuencias *cis* de unión a activadores transcripcionales específicos para genes en endospermo (como el factor Dof y SF4), e incluso hay un elemento de unión para un represor transcripcional que se expresa en aleurona; éstos resultados de la base de datos PLACE, de manera congruente, concuerdan con la expresión específica en endospermo del gen *OsRRM* reportado por Shi *et al* (2007); justamente el gen *sps1* de arroz índica es capaz de dirigir una expresión tejido específico en escutelo y aleurona de la semilla (Chávez-Barcenas *et al*, 2000; Martínez-

Trujillo *et al.*, 2004). Entendiendo que los genes At4g12640, OJ1381\_H04.27 y P0450E05.35 podrían estar involucrados en la regulación del desarrollo embrionario durante la germinación. Del mismo modo como lo hizo en el sistema de tres híbridos con el 5' UTR de *sps1* de arroz índica, la proteína Spen codificada por el gen At4g12640 por su parte RRM N-terminal podría interactuar con el 5' UTR de SPS de *A. thaliana* en la traducción al momento de la germinación del embrión dentro de la semilla, dejando el dominio SPOC libre para interactuar con otra proteína y potenciar la traducción o conducir a algún otro evento de regulación. De manera similar, los genes de arroz índica OJ1381\_H04.27 y P0450E05.35 como igual sus homólogos de arroz japónica *OsRRM* y Os09g0516300 también podrían interaccionar respectivamente con el ARNm 5' UTR *sps1* y su homólogo *osSPS1* en arroz japónica durante la traducción. Es decir que, éstas proteínas pueden estar implicadas en la regulación post-transcripcional o procesamiento de RNA de genes necesarios para la germinación de la semilla.

Es interesante que, los genes At4g12640, OJ1381\_H04.27 y P0450E05.35 también tienen secuencias para expresión específica en polen y flor, además de otros elementos como de respuesta a luz y expresión específica en tejido mesófilo. Elementos *cis* similares se ubican en el promotor de *sps1* con expresión específica en polen y tejido mesófilo (Chávez-Barcenas *et al.*, 2000; Martínez-Trujillo *et al.*, 2004), aunado que *OsRRM* también contiene secuencias que le dan la capacidad de expresarse en polen (Shi *et al.*, 2007). Lo anterior hace que los elementos funcionales *cis* de los promotores confieran expresión de las proteínas Spen de los genes de *A. thaliana* y sus ortólogos de arroz índica en los mismos órganos especializados y estados específicos del desarrollo de la planta en que se expresa el gen *sps1* de arroz, lo que permitiría eventualmente la interacción del líder *sps1* y la proteína Spen.

Las proteínas Spen codificadas por el gen de *A. thaliana* como sus ortólogos de arroz índica puede que estén relacionadas en el procesamiento transcripcional y post-transcripcional, como también en la estabilidad del ARNm, degradación y regulación de la traducción de genes específicos expresados en la fotosíntesis de tejido mesófilo, floración y desarrollo del polen. Por ejemplo. La proteína nuclear Spen en *A. thaliana* FPA participa en el control del tiempo de floración, regulando la escisión alternativa del RNA y poliadenilación; es probable que

FPA mediante la selección del sitio poli(A) proximal en el *FPA* pre-RNAm auto-regule negativamente la expresión, resultando en la pérdida del codón de alto de RNA y codificando para solo el primer RRM de FPA. La función de FPA es redundante con FCA, otra proteína específica de unión a ARN en planta que controla el tiempo de floración y silenciamiento de ARN, físicamente interactúa con FY para controlar la expresión de *FLC*. FY es un factor conservado de procesamiento del ARN 3' (Hornyk *et al.* 2010). Por lo tanto, la versatilidad de las proteínas Spen sobre la unión de los dominios RRM y SPOC hacia sus cognados es grande ya que participan en diferentes mecanismo moleculares requeridos en un espacio-tiempo especial del desarrollo de la planta, mismamente lo demuestra la presencia de proteínas Spen en desiguales etapas del desarrollo de la planta y estructuras reproductivas, como lo son la semilla y polen.

Se decidió demostrar *in vivo* que el gen descubierto At4g12640 es capaz de dirigir una expresión tejido específico en endospermo, polen y tejido mesófilo, tal como lo sugieren los motivos del promotor encontrados por la base de datos PLACE. Por lo que, se hicieron dos construcciones usando productos de PCR en dirección 5' de -490 pb y -981 pb desde el atg de inicio de la traducción de At4g12640. En este trabajo se clonó la amplificación del promotor hasta -490 después del atg de inicio de la traducción, en el vector destino pKGWFS7 del sistema Gateway. Con la construcción anterior se transformarán plantas de *A. thaliana* vía *Agrobacterium tumefaciens*, lo cual se está realizando en el laboratorio para un trabajo posterior de tesis. Teorizando que los genes reporteros se expresarán en la planta, acorde a los motivos del promotor hallados en los análisis informáticos.

## VII. BIBLIOGRAFÍA

Allen RD, Bernier F, Lessard PA, Beachy RN. (1989). Nuclear factors interact with a soybean beta-conglycinin enhancer. *Plant Cell* 1, 623-631.

Anderson M.B., K. Folta, K.M. Warpeha, J. Gibbons, J. Gao, L.S. Kaufman. (1999). Blue light-directed destabilization of the pea Lhcb1\*4 transcript depends on sequences within the 5' untranslated region, *Plant Cell* 11, 1579-1589.

Ben-Shem, A., Frolow, F. & Nelson, N. (2003). The crystal structure of plant photosystem I. *Nature* 426, 630-635.

Ben-Shem, A., Frolow, F. & Nelson, N. (2004). Light harvesting by plant photosystem I. *Photosynthesis Res.* 81, 239–250.

Bernstein David S., Natascha Buter, Craig Stumpf, and Marvin Wickens. (2002). Analyzing mRNA–protein complexes using a yeast three-hybrid system. *Methods* 26, 123–141.

Binda, C., Coda, A., Aliverti, A., Zanetti, G. & Mattevi, A. (1998). Structure of the mutant E92K of [2Fe-2S] ferredoxin I from *Spinacia oleracea* at 1.7 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* 54, 1353–1358.

Bolle C., S. Sopory, T. Lübbersedt, R.G. Herrmann, R. Oelmu. (1994). Segments encoding 5'-untranslated leaders of genes for thylakoid proteins contains cis-elements essential for transcription. *Plant J.* 6, 513-523.

Boyer D. Paul. (1989). A perspective of the binding change mechanism for ATP sintesis. *FASEB journal* 3, 2164-2178.

Boyer, D. Paul. (1997). The ATP synthase, a splendid molecular machine. *Annu. Rev. Biochem.* 66, 717–749.

Carrell, C. J., Zhang, H., Cramer, W. A. & Smith, J. L. (1997). Biological identity and diversity in photosynthesis and respiration: structure of the lumen-side domain of the chloroplast Rieske protein. *Structure* 5, 1613–1625.

Caspar Timothy and Peter H. Quail. (1993). Promoter and leader regions involved in the expresión of the Arabidopsis ferredoxin A gene. *The Plant Journal* 3 (1), 161-174

Chang, I. F., Szick-Miranda, K., Pan, S., J. Bailey- Serres. (2005). Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes. *Plant Physiology* 137 (3), 848-62.

Chávez-Bárceñas, A.T., Váldez-Alarcón, J.J., Martínez-Trujillo, M., Chen, L., Xoconostle-Cazárez, B., Lucas, W.J., and Herrera-Estrella, L. (2000). Tissue-Specific and Developmental Pattern of Expression of the Rice *sps1* Gene. *Plant Physiology* 124, 641-653.

Chen F, Rebay I. (2000). Split ends, a new component of the Drosophila EGF receptor pathway, regulates development of midline glial cells. *Curr Biol* 10 (2), 943-946.

Coleman HD, Ellis DD, Gilbert M, Mansfield SD. (2006). Up-regulation of sucrose synthase and UDP-glucose pyrophosphorylase impacts plant growth and metabolism. *Plant Biotechnol J* 4, 87–101.

Curie C., S. McCormick. (1997). A strong inhibitor of gene expression in the 5' untranslated region of the pollen-specific *LAT59* gene of tomato, *Plant Cell* 9, 2025-2036.

Dickey L.F., M.E. Petracek, T.T. Nguyen, E.R. Hansen, W.F. Thompson. (1998). Light regulated Fed-1 mRNA requires an element in the 5' untranslated region and correlates with differential poly-ribosome association. *Plant Cell* 10, 475-484.

Dobres, M., Elliot, R., Watson, J. and Thompson, W. (1987). A phytochrome regulated transcript encodes ferredoxin I. *Plant Mol. Biol.* 8, 53-59.

Donald RGK, Cashmore AR. (1990). Mutation of either G box or I box sequences profoundly affects expression from the Arabidopsis *rbcS*-1A promoter. EMBO J 9, 1717-1726.

Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J. & Iwata, S. (2004). Architecture of the photosynthetic oxygen-evolving center. Science 303, 1831–1838.

Filichkin SA, Leonard JM, Monteros A, Liu PP, Nonogaki H. (2004). A novel endo-beta-mannanase gene in tomato *LeMAN5* is associated with anther and pollen development. Plant Physiol. 134, 1080-1087.

Fütterer J. and T. Hohn. (1996). Translation in plants-rules and exceptions, Plant Mol. Biol. 32 158 -189.

Gallie D.R., V. Walbot. (1992). Identification of the motifs within the tobacco mosaic virus 5' leader responsible for enhancing translation. Nucleic Acids Res. 20, 4631-4638.

Gallie D.R., T.E. Young. (1994). The regulation of gene expression in transformed maize aleurone and endosperm protoplasts. Analysis of promoter activity, intron enhancement, and mRNA untranslated region on expression, Plant Physiol. 106, 929-939.

Galtier Nathalie, Christine H. Foyer, Joan Huber, Toni A. Voelker, and Steven C. Huber. (1993). Effects of Elevated Sucrose-Phosphate Synthase Activity on Photosynthesis, Assimilate Partitioning and Growth in Tomato (*Lycopersicon esculentum* var UC82B). Plant Physiol 101, 535-543.

Geigenberger, P. and Stitt, M. (1991). A 'futile' cycle of sucrose synthesis and degradation is involved in regulating partitioning between sucrose, starch and respiration in cotyledons of germinating *Ricinus communis* L. seedlings when phloem transport is inhibited. Plant journal, 185, 81-90.

Geigenberger, P. and Stitt, M. (1993). Sucrose synthase catalyses a readily reversible reaction in vivo in developing potato tubers and other plant tissues. *Plant journal*, 189, 329-339.

Geigenberger Peter, Ralph Reimholz, Uta Deiting, Uwe Sonnewald and Mark Stitt. (1999). Decreased expression of sucrose phosphate synthase strongly inhibits the water stress-induced synthesis of sucrose in growing potato tubers. *Plant Journal*, 19(2), 119-129.

Gilmartin PM, Sarokin L, Memelink J, Chua N-H. (1990). Molecular light switches for plant genes. *Plant Cell* 2, 369-378.

Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR. (1988). An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proc Natl Acad Sci USA* 85, 7089-7093.

Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Westhoff P. (2004). *cis*-Regulatory elements for mesophyll-specific gene expression in the C4 plant *Flaveria trinervia*, the promoter of the C4 phosphoenolpyruvate carboxylase gene. *Plant Cell*. 16, 1077-1090.

Green, P.J., Kay, S.A., and Chua, N.-H. (1987). Sequence-specific interactions of a pea nuclear factor with light-responsive elements upstream of the *rbcS-3A* gene. *EM60 J*. 6, 2543-2549.

Heather D. Coleman, Leigh Beamish, Anya Reid, Ji-Young Park, Shawn D. Mansfield. (2009). Altered sucrose metabolism impacts plant biomass production and flower development. *Transgenic Res* DOI 10.1007/s11248-009-9309-5

Heldt Hans-Walter y Heldt Fiona. (2005). *Plant Biochemistry*, Tercera Edición, Elsevier.

Hesse H, Sonnewald U, Willmitzer L. (1995). Cloning and expression analysis of sucrose-phosphate synthase from sugar beet (*Beta vulgaris*). *Mol. Gen. Genet.* 247: 515–20

Higo, K., Y. Ugawa, M. Iwamoto and T. Korenaga. (1999). Plant cis-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Research* 27 (1), 297-300.

Hook Brad, Bernstein David Zhang Beilin and Wickens Marvin, (2005), RNA–protein interactions in the yeast three-hybrid system: Affinity, sensitivity, and enhanced library screening, *RNA* 11, 227-233

Hornyik Csaba, Lionel C. Terzi, and Gordon G. Simpson. (2010). The Spen Family Protein FPA Controls Alternative Cleavage and Polyadenylation of RNA. *Developmental Cell* 18, 203–213.

Hubbard, N.L., Pharr, M.D. and Huber, S.C. (1990). Role of SPS in sucrose biosynthesis in ripening bananas and its relationship to the respiratory climacteric. *Plant Physiol.* 94, 201-208.

Huber Steven C. and Joan L.Huber. (1992). Role of Sucrose-Phosphate Synthase in Sucrose Metabolism in Leaves. *Plant Physiol* 99, 1275-1278.

Huber Steven C. and Joan L. Huber. (1996). Role and Regulation of Sucrose-Phosphate Synthase in Higher Plants. *Plant Physiol* 47, 431-444.

Huynh Frederick, Tien-Chye Tan, Kunchithapadam Swaminathan and Bharat K. C. Pate. (2004). Expression, purification and preliminary crystallographic analysis of sucrose phosphate synthase (SPS) from *Halothermothrix orenii*. *Acta cryst.* F61, 116-117.

Jobling S.A., Gehrke L. (1987) Enhanced translation of chimaeric messenger RNAs containing a plant viral untranslated leader sequence. *Nature* 325, 622-625.

Joshi C.P. (1987). An inspection of the domain between putative TATA box and translation start site in 79 plant genes, *Nucleic Acids Res.* 15, 6643-6653.

Klein RR, Crafts-Brandner SJ, Salvucci ME. (1993). Cloning and developmental expression of the sucrose-phosphate-synthase gene from spinach. *Planta* 190, 498–510

Kuang B, Wu SC, Shin Y, Luo L, Kolodziej P. (2000). Split ends encodes large nuclear proteins that regulate neuronal cell fate and axon extension in the *Drosophila* embryo. *Development* 127, 1517-1529.

Kurisu, G., Zhang, H., Smith, J. L. & Cramer, W. A. (2003). Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science* 302, 1009–1014.

Lam E, Chua NH. (1989). ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in cab promoters. *Plant Cell* 1, 1147-1156.

Lane ME, Elend M, Heidmann D, Herr A, Marzodko S, Herzig A, Lehner CF. (2000). A screen for modifiers of cyclin E function in *Drosophila melanogaster* identifies Cdk2 mutations, revealing the insignificance of putative phosphorylation sites in Cdk2. *Genetics*, 155, 233-244.

Laporte Marianne M., Julie A. Galagan, Amy L. Prash, Peter J. Vanderveer, David T. Hanson, Christine K. Shewmaker, Thomas D. Sharkey. (2001). Promoter strength and tissue specificity effects on growth of tomato plants transformed with maize sucrose-phosphate synthase. *Planta* 212, 817-822.

Le Gourrierec J, Li YF, Zhou DX. (1999). Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA-TBP-TATA complex. *Plant J* 18, 663-668.

Le S.-Y., R. Sternglanz, C.W. Greider. (2000). Identification of two RNA-binding proteins associated with human telomerase RNA. *Mol. Biol. Cell* 11, 999–1010.

Lin HV, Doroquez DB, Cho S, Chen F, Rebay I, Cadigan KM. (2003). Splits ends is a tissue/promoter specific regulator of Wingless signaling. *Development*, 130, 3125-3135.

Ling J., D.R. Wells, R.T. Tanguay, L.F. Dickey, W.F. Thompson, D. Gallie. (2000). Heat shock protein HSP101 binds to the Fed-1 internal light regulatory element and mediates its high translational activity, *Plant Cell* 12, 1213-1227.

López Salmerón Vadir. 2007. Establecimiento de un sistema de tres híbridos en levadura para buscar la interacción del RNA de la región líder del gen *sps1* de arroz con proteínas de *Arabidopsis thaliana*. Tesis de Licenciatura. Facultad de biología. Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Micoacán. 139 pp.

MacRae, E., Quick, W.P., Benker, C. and Stitt, M. (1992). Carbohydrate metabolism during the ripening of Kiwi fruit. *Plant journal* 188, 314-323.

Martinez, S. E., Huang, D., Szczepaniak, A., Cramer, W. A. & Smith, J. L. (1994). Crystal structure of chloroplast cytochrome f reveals a novel cytochrome fold and unexpected heme ligation. *Structure* 2, 95–105.

Martínez-Trujillo M. (2002). Análisis funcional de regiones involucradas en la expresión del gen de Sacarosa-Fosfato Sintasa de arroz (*Oryza sativa* L.). Tesis de Doctorado. CINVESTAV-IPN. Irapuato, Guanajuato. 112 pp.

Martínez-Trujillo, M., Limones-Briones, V., Chávez-Bárcenas, A.T., and Herrera-Estrella, L. (2003). Functional analysis of the 5' untranslated region of the sucrose phosphate synthase rice gene (*sps1*). *Plant Science* 165, 9-20.

Martínez-Trujillo, M., Chávez Bárcenas, A.T., Limones-Briones, V., Simpson, L., and Herrera-Estrella, L. (2004). Functional analysis of the promoter of the rice sucrose phosphate synthase gene (*sps1*). *Plant Science* 166, 131-140.

Martínez-Trujillo Miguel, Gamaliel Valdivia-Rojas, Gloria Solís Guzmán, José Luis Cabrera-Ponce. (2008). A 22-bp sequence of the core promoter from the Indica rice sucrose-phosphate synthase gene (*sps1*) is sufficient to confer basal transcription activity. *African Journal of*

Nakamura M, Tsunoda T, Obokata J. (2002). Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. *Plant J.* 29, 1-10.

Nelson Nathan and Ben-Shem Adam. (2004). The complex architecture of oxygenic photosynthesis, *Nature Reviews, Molecular Cell Biology* 5, 1 - 12.

Nelson David L. y Michael M. Cox. (2008). *Lehninger Principles of biochemistry*, Quinta edición, W.H. Freeman and Company

Okamura, M. Y., Paddock, M. L., Graige, M. S., and Feher, G. (2000). Proton and electron transfer in bacterial reaction centers. *Biochim. Biophys. Acta* 1458, 148–163.

Park J-Y, Canam T, Kang KY, Ellis DD, Mansfield SD. (2008). Overexpression of an arabidopsis family a sucrose phosphate synthase (SPS) gene alters plant growth and fibre development. *Transgenic Res* 17, 181–192.

Poleev A., A. Hartmann, S. Stamm. (2000). A trans-acting factor, isolated by the three-hybrid system, that influences alternative splicing of the amyloid precursor proteína minigen. *Eur. J. Biochem.* 267, 4002–4010.

Putz U., P. Skehel, D. Kuhl, (1996), A three hibrids sistem for the analysis and detection of RNA-protein interactions *Nucleic Acids Res.* 24, 4838–4840.

Rogers HJ, Bate N, Combe J, Sullivan J, Sweetman J, Swan C, Lonsdale DM, Twell D. (2001). Functional analysis of cis-regulatory elements within the promoter of the tobacco late pollen gene g10. *Plant Mol Biol* 45, 577-585.

Sandal NN, Bojsen K, Marcker KA. (1987). A small family of nodule specific genes from soybean. *Nucleic Acids Res.* 15, 1507-1519.

Sánchez-Pulido Luis, Ana M Rojas, Karel H van Wely, Carlos Martinez-A and Alfonso Valencia. (2004). SPOC: A widely distributed domain associated with cancer, apoptosis and transcription. *BMC Bioinformatics* 5: 91.

SenGupta D.J., B. Zhang, B. Kraemer, P. Pochart, S. Fields, M. Wickens. (1996). A three hybrid system to detect RNA-protein in vivo, *Proc. Natl. Acad. Sci. USA* 93, 8496–8501.

Shi-Yan Chen, Zong-Yang Wang, Xiu-Ling Cai. (2007). OsRRM, a Spen-like rice gene expressed specifically in the endosperma. *Cell Research* 17, 713-721.

Shirsat A, Wilford N, Croy R, Boulter D. (1989). Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. *Mol Gen Genet* 215, 326-331.

Sonnewald U, Basner A. 1993. EMBL Data Library, Accession No. S34172

Sonnewald U, Quick WP, MacRae E, Krause KP, Stitt M. (1993). Purification, cloning and expression of spinach leaf sucrose-phosphate synthase in *E. coli*. *Planta* 189, 174–81

Stock Daniela, Gibbons Clyde, Arechaga Ignacio, Andrew GW Leslie and John E Walker. (2000). The rotary mechanism of ATP synthase. *Elsevier* 10, 672-679.

Stougaard J, Jorgensen JE, Christensen T, Kuhle A, Marcker KA. (1990). Interdependence and nodule specificity of cis-acting regulatory elements in the soybean leghemoglobin *lbc3* and *N23* gene promoters. *Mol Gen Genet.* 220, 353-360.

Stroebel, D., Choquet, Y., Popot, J. L. & Picot, D. (2003). An atypical haem in the cytochrome b6f complex. *Nature* 426, 413–418.

Stryer Luber, Jhon L. Tymoczko y Jeremy M. Berg. (2002), *Biochemistry*, W. H. Freeman and Company

Taiz Lincoln y Zeiger Eduardo, 2002, Plant Fisiology, Tercera Edition, Sinauer Associates, Inc, Publisher.

Timmer R.T., Benkowski L.A., Schodin D., Lax S.R., Metz A.M., Ravel J.M., Browning K.S. (1993). The 5' and 3' untranslated regions of satellite tobacco necrosis virus RNA affect translational efficiency and dependence on a 5' cap structure. *Journal of Biological Chemistry* 268, 9504-9510.

Tommos, C., and Babcock, G. T. (1999). Oxygen production in nature: A light-driven metalloradical enzyme process. *Acc. Chem. Res.* 37, 18–25.

Toroser Dikran and Steven C. Huber. (1997). Protein Phosphorylation as a Mechanism for Osmotic-Stress Activation of Sucrose-Phosphate Synthase in Spinach leaves. *Plant Physiol* 114, 947-955.

Twell D, J Yamaguchi, R A Wing, et al. (1991). Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes Dev.* 5, 496-507.

Valdéz-Alarcón, J.J., Fernando, M., Salerno, G., Jiménez-Moraila, B. and Herrera-Estrella, L. (1996). Characterization of a rice sucrose-phosphate synthase-encoding gene. *Gene* 170, 217-222.

Villain P, Mache R, Zhou DX. (1996). The mechanism of GT element-mediated cell type-specific transcriptional control. *J Biol Chem* 271, 32593-32598.

Wang S., Browning K.S., Miller W.A. (1997). A viral sequence in the 3'-untranslated region mimics a 5' cap in facilitating translation of uncapped mRNA. *EMBO Journal* 16, 4107-4116.

Wedel N., Bartling, D. and Hernnann, R. (1988). Analysis of cDNA clones encoding the entire ferredoxin I precursor poly-peptide from spinach. *Bot. Acta* 101, 295-300.

Wells D.R., R.L. Tanguay, H. Le, D.R. Gallie. (1998). HSP101 functions as a specific translational regulatory protein whose activity is regulated by nutrient status, *Genes Dev.* 12, 3236-3251.

Wiellette EL, Harding KW, Mace KA, Ronshaugen MR, Wang FY, McGinnis W. (1999). *Spn* encodes an RNP motif protein that interacts with Hox pathways to repress the development of head-like sclerites in the *Drosophila* trunk. *Development* 126, 5373-5385.

Wood J., R.M. Frederickson, S. Fields, A.H. Patel. (2001). Hepatitis C virus 3'X Region Interacts with Human ribosomal Proteins. *J. Virol.* 75, 1348–1358.

Worrell AC, Bruneau J-M, Summerfelt K, Boersig M, Voelker TA. (1991). Expression of a maize sucrose phosphate synthase in tomato alters leaf carbohydrate partitioning. *Plant Cell* 3:1121–30

Xu P., L. Wen, G. Benegal, X. Wang, G.A. Buck. (2001). Identification of a spliced leader RNA binding protein from *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* 112, 39–49.

Xue Y., Okvist, M., Hansson, O. & Young, S. (1998). Crystal structure of spinach plastocyanin at 1.7 Å resolution. *Protein Sci.* 7, 2099–2105.

Yamamoto Y., H. Tsuji, J. Obokata. (1995). 5'-leader of a photosystem I gene in *Nicotiana sylvestris*, *psaDb*, contains a translational enhancer, *J. Biol. Chem.* 270, 12466-12470.

Yanagisawa S, Schmidt RJ. (1999). Diversity and similarity among recognition sequences of Dof transcription factors. *Plant J* 17, 209-214.

Yoshida Masasuke, Eiro Muneyuki and Toru Hisabori. (2001). ATP synthase — A marvellous Rotary Engine of the cell. *Nature review molcellbio* 2, 669-677.

Zhang ZL, Xie Z, Zou X, Casaretto J, Ho TH, Shen QJ. (2004). A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. *Plant Physiol.* 134, 1500-1513.

Zhu, Y.J., Komor, E. and Moore, P.H. (1997). Sucrose accumulation in the sugarcane stem is regulated by the difference between the activities of soluble acid invertase and sucrose phosphate synthase. *Plant Physiol.* 115, 609-616.

# ANEXO

## At4g12640

Web Signal Scan Program

Database Searched: PLACE

This is the sequence you submitted

>../tmp/sigscan//signalseq.12284 [Unknown form], 500 bases, CB734C5B  
checksum.

```
GAAACATGACGAGCAGATCTACGGAGAGGAGTGTGAGTTTCTGGGGTTTG
AGGAGTGTGTGAGGTTGAAGATGGAGATAGATACGATTTGCGATAGATGT
TAGACGGAGTTTTGGTTGGGGATTTGAAGGTTTTTGCTCTTGGAGTTGAA
GCCATTGATGGTGTAAAGCTAGGGTTTTCTGGGTGAAGAATGAAGAGTGA
ATGAAGAGGAAGAAGAAGACGATGATGATGATGAAGAAGGAGGTAGAGAG
AGAAAGTTGGCGGGAAAAGGTGGCGGACACTTCTTCTTCTTCGCATCAGA
CAAGAGTTTGCAGAAAAGGGTTTAACTTTGGGCCGAATATTGGGCTTAT
TAACCCAACTCATCACTTTAGGTTTATCCAAATTAATCATAACATTTTGG
GCCGATTTGATAAGAAATCAATAATATAGGTCCTTTTGTGAAAGAAAAAA
AAAAAACGAAGAGCACAAAGTGAAATGGGGAATTGTTTTTTGTAGAGCC
```

RESULTS OF YOUR SIGNAL SCAN SEARCH REQUEST

../tmp/sigscan//signaldone.12284: 500 base pairs  
Signal database file: user.dat

(+) = Current Strand  
(-) = Opposite Strand

```
1      GAAACATGACGAGCAGATCTACGGAGAGGAGTGTGAGTTTCTGGGGTTTG
      (+)ASF1MOTIFCAMV S000024
      (+)WRKY71OS S000447
      (-)CACTFTPPCA1 S000449
      (+)GTGANTG10 S000378
      (+)EECCRCAH1 S000494
      (-)POLLEN1LELAT52 S000245

51     AGGAGTGTGTGAGGTTGAAGATGGAGATAGATACGATTTGCGATAGATGT
      (-)CACTFTPPCA1 S000449
      (+)GTGANTG10 S000378
      (+)GATABOX S000039
      (+)GATABOX S000039
      (+)ARR1AT S000454
      (+)EECCRCAH1 S000494
      (+)PRECONSCRHSP70A S000506
      (+)PRECONSCRHSP70A S000506
      (+)GATABOX S000039

101    TAGACGGAGTTTTGGTTGGGGATTTGAAGGTTTTTGCTCTTGGAGTTGAA
      (-)REALPHALGLHCB21 S000362
      (-)MYBPZM S000179
      (+)ARR1AT S000454
      (+)SEF4MOTIFGM7S S000103
```

```

(+ )NODCON2GM S000462
(+ )OSE2ROOTNODULE S000468
(- )CAREOSREP1 S000421

151 GCCATTGATGGTGTAAGCTAGGGTTTTCTGGGTGAAGAATGAAGAGTGA
(- )CAATBOX1 S000028
(- )UP2ATMSD S000472
(- )GT1CONSENSUS S000198
(- )GT1GMSCAM4 S000453
(- )POLLEN1LELAT52 S000245
(+ )GTGANTG10 S000378
(- )INRNTPSADB S000395
(- )NODCON2GM S000462
(- )OSE2ROOTNODULE
S000468
(- )INRNTPSADB S000395
(- )CACTFTPPCA1
S000449
(+ )GTGANTG10 S000378

201 ATGAAGAGGAAGAAGAAGACGATGATGATGATGAAGAAGGAGGTAGAGAG
(- )NODCON2GM S000462
(- )OSE2ROOTNODULE S000468
(- )CTRMCAV35S S000460

251 AGAAAGTTGGCGGGAAAAGGTGGCGGACACTTCTTCTTCTCGCATCAGA
(+ )POLLEN1LELAT52 S000245
(+ )DOFCOREZM S000265
(- )E2FAT S000417
(+ )E2F1OSPCNA S000396
(- )E2FANTRNR S000366
(- )E2FCONSENSUS S000476
(+ )GT1CONSENSUS S000198
(- )PYRIMIDINEBOXOSRAMY1A S000259
(+ )DOFCOREZM S000265
(- )SORLIPI1AT S000482
(+ )CACTFTPPCA1 S000449

301 CAAGAGTTTGCAGAGAAAAGGGTTTAACTTTGGGCCGAATATTGGGCTTAT
(- )NODCON2GM S000462
(- )OSE2ROOTNODULE S000468
(+ )POLLEN1LELAT52 S000245
(- )PYRIMIDINEBOXOSRAMY1A S000259
(+ )DOFCOREZM S000265
(+ )TBOXATGAPB S000383
(- )DOFCOREZM S000265
(+ )SITEIIATCYTC S000474
(+ )SORLIP2AT S000483
(- )-10PEHVPSBD S000392
(- )ROOTMOTIFTAPOX1 S000098
(+ )ROOTMOTIFTAPOX1 S000098
(- )CAATBOX1 S000028
(- )CCAATBOX1 S000030
(+ )SITEIIATCYTC S000474
(- )GT1CORE
S000125

```

351 TAACCCAACATCATCACTTTAGGTTTATCCAAATTAATCATAACATTTTGG  
 (+)SEF3MOTIFGM S000115  
 (+)CAREOSREP1 S000421  
 (+)PREATPRODH S000450  
 (+)ANAERO3CONSENSUS S000479  
 (-)GTGANTG10 S000378  
 (+)CACTFTPPCA1 S000449  
 (+)NTBBF1ARROLB S000273  
 (-)DOFCOREZM S000265  
 (-)TAAAGSTKST1 S000387  
 (-)GT1CONSENSUS S000198  
 (+)SREATMSD S000470  
 (-)IBOXCORE S000199  
 (+)TATCCAOSAMY S000403  
 (-)GATABOX S000039  
 (-)MYBST1 S000180  
 (+)CIACADIANLELHC S000252  
 (-)ARR1AT S000454  
 (-)UP1ATMSD S000471  
 (+)SITEIIATCYTC  
 S000474  
 (+)SORLIP2AT  
 S000483  
 401 GCCGATTTGATAAGAAATCAATAATATAGGTCCTTTTGTGAAAGAAAAA  
 (+)ARR1AT S000454  
 (+)GATABOX S000039  
 (+)IBOX S000124  
 (+)IBOXCORE S000199  
 (+)IBOXCORENT S000424  
 (+)POLLEN1LELAT52 S000245  
 (-)ARR1AT S000454  
 (+)CAATBOX1 S000028  
 (+)POLASIG3 S000088  
 (-)ROOTMOTIFTAPOX1 S000098  
 (+)PYRIMIDINEBOXOSRAMY1A S000259  
 (-)DOFCOREZM S000265  
 (+)GTGANTG10 S000378  
 (+)DOFCOREZM S000265  
 (+)POLLEN1LELAT52  
 S000245  
 (+)GT1CONSENSUS S000198  
 (+)GT1GMSCAM4 S000453  
 (-)MARTBOX S000067  
 (-)MARTBOX S000067  
 (-)MARTBOX S000067  
 (-)MARTBOX S000067  
 451 AAAAAACGAAGAGCACAAAGTGAAATGGGGAATTGTTTTTTGTAGAGCC  
 (-)NODCON2GM S000462  
 (-)OSE2ROOTNODULE S000468  
 (-)TBOXATGAPB S000383  
 (+)DOFCOREZM S000265  
 (-)INRNTPSADB S000395  
 (-)CACTFTPPCA1 S000449

(+)GTGANTG10 S000378  
(-)CAATBOX1 S000028

501

-----  
o If you use this program in published research, please cite:  
- Higo, K., Y. Ugawa, M. Iwamoto and T. Korenaga (1999) Plant cis-acting regulatory DNA elements (PLACE) database:1999. Nucleic Acids Research Vol.27 No.1 pp. 297-300.  
- Prestridge, D.S. (1991) SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements. CABIOS 7, 203-206.

### OsI\_30787

Web Signal Scan Program

Database Searched: PLACE

This is the sequence you submitted  
>../tmp/sigscan//signalseq.11808 [Unknown form], 502 bases, 3BF0E985 checksum.

```
ggatgaacaaattccagtcgctccgatgctgccaatcgatggccatgat
ttggtaccgcatgccctcaaggatggtaaaaaatctctttctataatatt
agcatgcatgcgcttaactccaaaatatatggatggatatattctaagtt
ttataaatatgatagatagaaatggtgaaatattgtggaaatgggtgttg
ttttagtattcaataaatcgactctcttcatctcaaaacataaaaagt
ttttgttgatgtgacacattctagtagtactacgaatctggataaaactagta
tgtgtttcatgtaatcaaaactctttatattttgagatagatggagtata
gatatgatattgtaagcttataagatttgatatgcattaattattgtcca
gtgataagccatgcttgcatccattttctcggcaaacgaaaaaaaaacacg
gaagggaaaaatagtaaagggcagcaagaaaaaaatggaaaaacaattg
ca
```

RESULTS OF YOUR SIGNAL SCAN SEARCH REQUEST

../tmp/sigscan//sigaldone.11808: 502 base pairs  
Signal database file: user.dat

(+) = Current Strand  
(-) = Opposite Strand

```
1      GGTATGAACAAATTCCAGTCGCTCCGATGCTGCCAAATCGATGGCCATGAT
                                         (-)ARR1AT S000454
                                           (+)ARR1AT S000454
```

51 TTGGTACCGCATGCCCTCAAGGATGGTAAAAATCTCTTTCTATAATATT  
(+ )CURECORECR S000493  
(- )CURECORECR S000493  
(- )BOXLCORED PAL S000492  
(+ )S1FBOXSORPS1L21 S000223  
(+ )GT1CONSENSUS S000198  
(+ )CCA1ATLHCB1 S000149  
(- )ARR1AT S000454  
(+ )NODCON2GM S000462  
(+ )OSE2ROOTNODULE S000468  
(- )DOFCOREZM S000265  
(- )POLLEN1LELAT52 S000245  
(- )BOXIINTPATPB S000296  
(- )ROOTMOTIFTAPOX1  
S000098 (+ )ROOTMOTIFTAPOX1  
S000098 (+ )CPBCSPOR S000491

101 AGCATGCATGCGCTTAACTCCAAAATATATGGATGGATATATTCTAAGTT  
(+ )RYREPEATBNNAPA S000264  
(+ )RYREPEATGMGY2 S000105  
(+ )RYREPEATLEGUMINBOX S000100  
(+ )RYREPEATVFLEB4 S000102  
(- )RYREPEATVFLEB4 S000102  
(- )RYREPEATGMGY2 S000105  
(- )RYREPEATLEGUMINBOX S000100  
(- )RYREPEATBNNAPA S000264  
(- )ROOTMOTIFTAPOX1 S000098  
(- )AMYBOX2 S000021  
(- )TATCCAYMOTIFOSRAMY3D S000256  
(- )TATCCAOSAMY S000403  
(+ )MYBST1 S000180  
(+ )GATABOX S000039  
(+ )ROOTMOTIFTAPOX1 S000098  
(+ )-10PEHVPSBD S000392

151 TTATAAATATGATAGATAGAAATGTTGAAATATTGTGGAAATGGTGTGG  
(+ )TATABOX2 S000109  
(- )SEF1MOTIF S000006  
(- )ROOTMOTIFTAPOX1 S000098  
(+ )GATABOX S000039  
(+ )GATABOX S000039  
(+ )BOXIINTPATPB S000296  
(+ )POLLEN1LELAT52 S000245  
(- )RAV1AAT S000314  
(- )ROOTMOTIFTAPOX1 S000098  
(+ )ROOTMOTIFTAPOX1 S000098  
(- )CAATBOX1 S000028  
(+ )GT1CONSENSUS S000198  
(- )CANBNNAPA S000148  
(- )RAV1AAT S000314  
(- )REALPHALGLHCB21  
S000362 (- )MYB1AT S000408

201 TTTGTAGTATTCAATAAATCGTACTCTCTTCATCTCAAAACATAAAAAGT  
 (-)CACTFTPPCA1 S000449  
 (+)CAATBOX1 S000028  
 (+)POLASIG1 S000080  
 (-)ARR1AT S000454  
 (+)CURECORECR S000493  
 (-)CURECORECR S000493  
 (+)CACTFTPPCA1 S000449  
 (+)NODCON2GM S000462  
 (+)OSE2ROOTNODULE S000468  
 (-)MARARS S000064  
 (+)DOFCOREZM S000265  
 (+)SEF4MOTIFGM7S  
  
 S000103  
  
 251 TTTTGTGGAATGTGACACATTCTAGTACTACGAATCTGGATAAACTAGTA  
 (-)RAV1AAT S000314  
 (-)TGTCACACMCUCUMISIN S000422  
 (+)GTGANTG10 S000378  
 (+)WRKY71OS S000447  
 (-)BIHD1OS S000498  
 (-)CACTFTPPCA1 S000449  
 (+)CURECORECR S000493  
 (-)CURECORECR S000493  
 (+)CACTFTPPCA1 S000449  
 (-)ARR1AT S000454  
 (-)TATCCAOSAMY S000403  
 (+)MYBST1 S000180  
 (-)SREATMSD S000470  
 (+)GATABOX S000039  
 (+)GT1CONSENSUS S000198  
 (+)IBOXCORE S000199  
 (-)CACTFTPPCA1  
  
 S000449  
  
 301 TGTGTTTCATGTAATCAAAACTCTTTATATTTTGAGATAGATGGAGTATA  
 (-)ARR1AT S000454  
 (+)NODCON2GM S000462  
 (+)OSE2ROOTNODULE S000468  
 (-)DOFCOREZM S000265  
 (-)TAAAGSTKST1 S000387  
 (+)ROOTMOTIFTAPOX1 S000098  
 (+)GATABOX S000039  
 (-)CACTFTPPCA1 S000449  
  
 351 GATATGATATTGTAAGCTTATAAGATTTGATATGCATTAATTATTGTCCA  
 (+)GATABOX S000039  
 (+)GATABOX S000039  
 (+)ROOTMOTIFTAPOX1 S000098  
 (-)CAATBOX1 S000028  
 (+)ARR1AT S000454  
 (+)GATABOX S000039  
 (-)POLASIG3 S000088  
 (-)CAATBOX1 S000028

(-)CACTFTPPCA1

S000449

401 GTGATAAGCCATGCTTGCATCCATTTTCTCGGCAAACGAAAAAAAAACACG  
 (+)GTGANTG10 S000378  
 (+)GATABOX S000039  
 (+)IBOX S000124  
 (+)IBOXCORE S000199  
 (-)GT1CONSENSUS S000198  
 (-)POLLEN1LELAT52 S000245  
 (+)GT1CONSENSUS S000198  
 (+)GT1GMSCAM4 S000453

451 GAAGGGAAAAATAGTAAAGGGCAGCAAGAAAAAAATGGAAAAACAATTG  
 (+)GT1CONSENSUS S000198  
 (+)GT1CONSENSUS S000198  
 (+)GT1GMSCAM4 S000453  
 (-)SP8BFIBSP8BIB S000184  
 (-)CACTFTPPCA1 S000449  
 (+)TAAAGSTKST1 S000387  
 (+)DOFCOREZM S000265  
 (+)POLLEN1LELAT52 S000245  
 (+)GT1CONSENSUS S000198  
 (+)GT1GMSCAM4 S000453  
 (+)GT1CONSENSUS S000198  
 (+)GT1CONSENSUS S000198  
 (+)GT1GMSCAM4 S000453  
 (+)CAATBOX1 S000028  
 (+)EBOXBNNAPA S000144  
 (+)MYCCONSENSUSAT

S000407

(-)EBOXBNNAPA S000144  
 (-)MYCCONSENSUSAT

S000407

(-)CAATBOX1 S000028

501 CA

-----

o If you use this program in published research, please cite:

- Higo, K., Y. Ugawa, M. Iwamoto and T. Korenaga (1999) Plant cis-acting regulatory DNA elements (PLACE) database:1999. Nucleic Acids Research Vol.27 No.1 pp. 297-300.
- Prestridge, D.S. (1991) SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements. CABIOS 7, 203-206.

## OsI\_32043

Web Signal Scan Program

Database Searched: PLACE

This is the sequence you submitted

>../tmp/sigscan//signalseq.11602 [Unknown form], 506 bases, 4C4B5FCA  
checksum.

```
ttgtgactagtgtgagagcaagtttaaatagtgtagccaactactatctcaaa  
atcacctatagtcacataatagccaattcatataatagttatatataaa  
catatagtacactattaatacatgaatacatatcatacaaacattgtgtt  
ttggagtccgtgctataactggctataaatttatagcccgtgcttttct  
ctctactcttttatcttctcgacatgtgtttatcttctcttatagtctgc  
tacagtacctgcttgccagtgatgctgttcctggccctctctttccacag  
aacggcctataaggatcaaaaagccctatgccagtttctcttgtttcaat  
ctgggcctcttcccttgcttttactcttctcaagcccggcctgctaaa  
tgccgtgttgctcccgccacgtaaccggaagcatcgctcgaaccctcct  
aggccgctgctgccccacacgtcatatactcagcccgatcccggcctt  
gtctca
```

### RESULTS OF YOUR SIGNAL SCAN SEARCH REQUEST

../tmp/sigscan//signaldone.11602: 506 base pairs  
Signal database file: user.dat

(+) = Current Strand  
(-) = Opposite Strand

```
1      TTGTGACTAGTGAGAGCAAGTTTAATAGTGTAGCCAACACTACTATCTCAAA  
      (+)GTGANTG10 S000378  
      (+)WBOXHVIS01 S000442  
      (+)WBOXNTERF3 S000457  
      (+)WRKY71OS S000447  
      (-)CACTFTPPCA1 S000449  
      (+)GTGANTG10 S000378  
      (-)CACTFTPPCA1 S000449  
      (+)CACTFTPPCA1 S000449  
      (-)GATABOX S000039  
      (-)ARR1AT S000454  
  
51     ATCACCTATAGTCAACATAATAGCCAATTCATATAATAGTTATATATAAA  
      (-)GTGANTG10 S000378  
      (-)WBOXHVIS01 S000442  
      (-)WBOXNTERF3 S000457  
      (-)WBOXATNPR1 S000390  
      (-)WRKY71OS S000447  
      (+)RAV1AAT S000314  
      (+)CCAATBOX1 S000030  
      (+)CAATBOX1 S000028  
      (+)CARGCW8GAT S000431  
      (-)CARGCW8GAT S000431
```

(-)TATABOX4 S000111  
(+)TATABOX4 S000111  
(-)TATAPVTRNALEU S000340

101 CATATAGTACACTATTAATACATGAATACATATCATAACAAACATTGTGTT  
(-)CACTFTPPCA1 S000449  
(+)CURECORECR S000493  
(-)CURECORECR S000493  
(+)CACTFTPPCA1 S000449  
(+)TATABOX3 S000110  
(-)TATABOX3 S000110  
(-)-10PEHVPSBD S000392  
(-)GATABOX S000039  
(-)CAATBOX1 S000028

151 TTGGAGTCCGTGCTATAACTGGCTATAAATTTATAGCCCGCTGCTTTTCT  
(+)MYB2AT S000177  
(+)MYB2CONSENSUSAT S000409  
(-)MYBCORE S000176  
(+)TATABOX2 S000109  
(-)TATABOX2 S000109  
(-)BS1EGCCR S000352  
(-)ANAERO2CONSENSUS S000478  
(-)DOFCOREZM S000265  
(-)POLLEN1LELAT52

S000245

201 CTCTACTCTTTTATCTTCTCGACATGTGTTTATCTTCTTATAGTCTGC  
(+)CACTFTPPCA1 S000449  
(+)NODCON2GM S000462  
(+)OSE2ROOTNODULE S000468  
(-)DOFCOREZM S000265  
(-)GT1CONSENSUS S000198  
(-)IBOXCORE S000199  
(-)GATABOX S000039  
(+)EBOXBNNAPA S000144  
(+)MYCATERD1 S000413  
(+)MYCCONSENSUSAT S000407  
(-)DPBFCOREDCDC3 S000292  
(-)EBOXBNNAPA S000144  
(-)MYCATRD22 S000174  
(-)MYCCONSENSUSAT S000407  
(-)GT1CONSENSUS S000198  
(-)IBOXCORE S000199  
(-)GATABOX S000039  
(+)NODCON2GM S000462  
(+)OSE2ROOTNODULE S000468

251 TACAGTACCTGCTTGCCAGTGATGCTGTTCCCTGGCCCTCTTTCCACAG  
(-)CACTFTPPCA1 S000449  
(+)CURECORECR S000493  
(-)CURECORECR S000493  
(-)CACTFTPPCA1 S000449  
(+)GTGANTG10 S000378  
(-)SORLIP2AT S000483  
(+)NODCON2GM S000462

```

(+ )OSE2ROOTNODULE S000468
(- )DOFCOREZM S000265
(- )SV40COREENHAN S000123

301 AACGGCCTATAAGGATCAAAAAGCCCTATGCCAGTTTCTCTTGTTCAT
(+ )MYBCOREATCYCB1 S000502
(+ )DOFCOREZM S000265
(- )POLLEN1LELAT52 S000245
(+ )NODCON2GM S000462
(+ )OSE2ROOTNODULE S000468
(+ )INRNTPSADB S000395
(+ )CAATBOX1 S000028
(- )ARR1AT S000454

351 CTGGGCCTCTTCCCTTGTCTTTTACTCTTCTCAAGCCCGCCTGCTAAA
(+ )SITEIIATCYTC S000474
(+ )SORLIP2AT S000483
(+ )NODCON2GM S000462
(+ )OSE2ROOTNODULE S000468
(- )DOFCOREZM S000265
(+ )CACTFTPPCA1 S000449
(+ )NODCON2GM S000462
(+ )OSE2ROOTNODULE S000468

401 TGCCGTGTTGCTCCCGCCACGTAACCGAAAGCATCGCTCGAACCCCTCCT
(+ )UPRMOTIFIIAT S000426
(- )RAV1AAT S000314
(+ )SORLIP1AT S000482
(- )ABREATCONSENSUS S000406
(- )ACGTABREMOTIFA2OSEM S000394
(- )BOXIIPCCHS S000229
(- )ABRELATERD1 S000414
(+ )ACGTATERD1 S000415
(- )ACGTATERD1 S000415
(- )MYBCORE S000176
(+ )DOFCOREZM S000265

451 AGGCCGCCGTGCCCCACACGTCATATACTCAGCCCCGATCCCGCCCGGCTT
(+ )GCCCORE S000430
(- )ABRELATERD1 S000414
(+ )ACGTATERD1 S000415
(+ )HEXMOTIFTAH3H4 S000053
(- )ACGTATERD1 S000415
(- )TGACGTVMAMY S000377
(- )ASF1MOTIFCAMV S000024
(- )WRKY71OS S000447
(+ )CACTFTPPCA1 S000449
(+ )SEBFCONSSTPR10A
S000391
(+ )ARFAT S000270

501 GTCTCA
(- )SURECOREATSULTR11 S000499

```

