

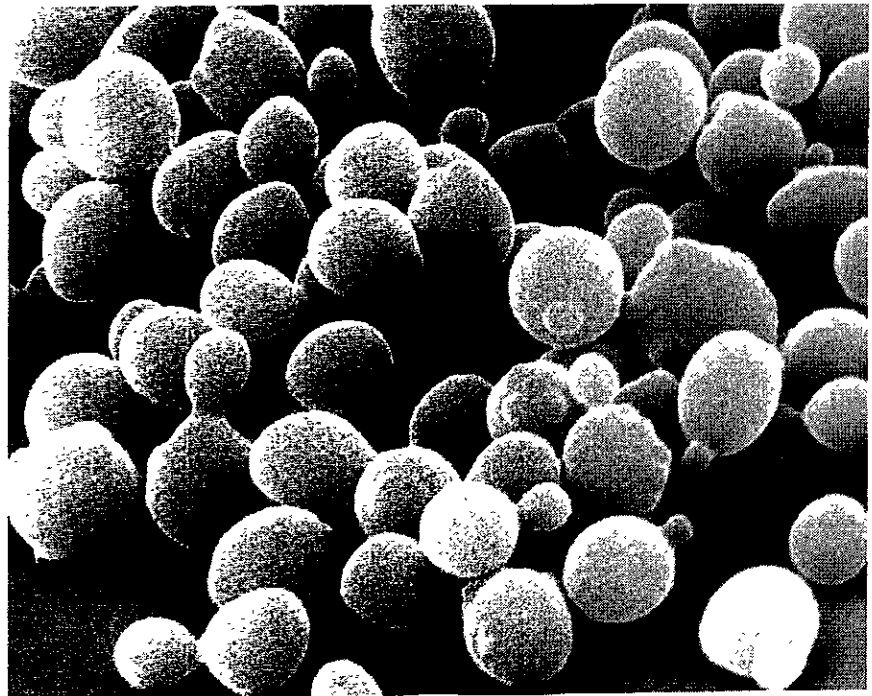
03008  
2  
Lej



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

INSTITUTO DE BIOTECNOLOGIA

CARACTERIZACION DE GENES INVOLUCRADOS EN LA RESPUESTA AL ESTRES OSMOTICO



T E S I S

QUE PARA OBTENER EL GRADO DE DOCTOR EN BIOTECNOLOGIA *Biomedicas* PRESENTA Adriana Garay Arroyo

Cuernavaca, Morelos, Julio de 1999

TESIS CON FALLA DE ORIGEN

274800



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

A Mauricio/Sofía

A Sofía/Mauricio

## Agradecimientos

A Mario, por todo lo bueno, por el amor que hemos construido en nuestra maravillosa vida juntos, porque hicimos a Mauricio y a Sofía y eso es lo mejor que me ha pasado en la vida.

A mis padres, porque gracias a la educación y el cariño que recibí se ha realizado este trabajo.

A mi mamá, por el apoyo incondicional y porque a lo largo del tiempo se ha convertido en una de mis mejores amigas.

A Lalo, mi favorito, porque me dió una lección de vida que todavía duele.

A Alejandra, por la confianza que me ha tenido desde que nos conocimos, por el apoyo en todo momento, por ser tan académica y, al mismo tiempo, tan buena amiga.

A mi mejor amigo, Chema, quien se convirtió en parte fundamental de mi vida cotidiana con sus comentarios y sus risas.

A Roberto, por ser mi buen amigo y por compartir mucho mas que las levaduras.

A mis amigas, Elena, Consuelo, María, Lucy, Nieves, Martha, Irma, Roxy y Helena, por todo lo que hemos pasado juntas y por ser tan necesarias en mi vida.

A Claudia, Lucero y Bere por todos los momentos de risa y por las muy buenas platicas esporádicas.

A mis compañeros del laboratorio, Roxy, Pancho, Oscar, Blanca, Liz, Pepe, Marina, Sonia, Isadora, Esperanza, Abraham, Emmanuel, Sergio, Jordi, Larissa, LuzMa, Jose Luis, Magda, y Ma. Jesús, quienes hacen del laboratorio un lugar muy agradable para trabajar.

A los miembros de mi comite tutorial, Dra. Alejandra A. Covarrubias Robles, Dra. Alicia González Manjarréz, Dr. Enrique Morett por aguantar todo el proceso de este trabajo y enriquecerlo con sus comentarios.

A los miembros del jurado Dr. Victor Bustamante, Dr. Gustavo Pedraza, Dr. David Romero, Dr. Jorge Luis Folch, Dra. Susana López, Dr. Jesús Aguirre y Dra. Alejandra Covarrubias por la revisión y las sugerencias hechas a esta tesis.

Resumen.....	1
--------------	---

## Capítulo I

### a) Introducción

i) El potencial hídrico y su relación con el estrés osmótico.....	3
ii) Adaptación fisiológica	
ii-a) Compuestos osmocompatibles.....	5
ii-b) La levadura <i>S.cerevisiae</i> como modelo para el estudio de la respuesta a estrés osmótico en organismos eucariotes.....	6
ii-c) Fisiología de <i>S.cerevisiae</i> en condiciones de estrés osmótico.....	6
iii) Vías de transducción de señales en la respuesta a estrés osmótico	
iii-a) La vía de HOG.....	12
- Elemento presente en los promotores de genes regulados por estrés osmótico: STRE.....	15
iii-b) La vía de AMPc-PKA.....	16
iii-c) La vía de regulación mediada por Ssn6p-Tup1p-Sok1p.....	18
b) Hipótesis.....	20
c) Objetivos.....	20
d) Artículo	
e) Resultados anexos	
i) Hibridación tipo Southern.....	21
ii) Cinética de expresión de los genes GRE bajo diferentes concentraciones de sorbitol.....	22
f) Conclusiones y discusión	
i) GRE1 forma parte de las hidrófilinas que son proteínas altamente hidrófilas.....	25
ii) GRE2 forma parte de la superfamilia (DHE/DFR) de mamíferos 3 $\beta$ -deshidrogenasa hidroxisteroide/plantas dihidroflavonol-4-reductasa (DFR).....	25
iii) GRE3 tiene alta homología con xilasas reductasa descritas en diferentes hongos.....	29
iv) Regulación transcripcional de los genes <i>GRE</i> .....	31

## Capítulo II

### **a) Introducción**

i) Proteínas abundantes de la embriogénesis tardía (proteínas LEA).....	35
a) Grupo 1 (Familia D-19).....	35
b) Grupo 2 (Familia D-11).....	36
c) Grupo 3 (Familia D-7).....	36
d) Grupo 4 (Familia D-113).....	38
e) Grupo 6 (PvLEA-18).....	38
f) Grupo 5 (Proteínas LEAs atípicas) .....	39
ii) Posibles funciones de las proteínas LEA.....	41

<b>b) Hipótesis.....</b>	<b>43</b>
--------------------------	-----------

<b>c) Objetivos.....</b>	<b>43</b>
--------------------------	-----------

### **d) Artículo**

<b>e) Conclusiones y discusión.....</b>	<b>44</b>
-----------------------------------------	-----------

<b>Bibliografía.....</b>	<b>46</b>
--------------------------	-----------

## Resumen

Las múltiples ventajas que presenta la levadura *Saccharomyces cerevisiae* como modelo experimental han permitido en los últimos años avanzar enormemente en el conocimiento de los mecanismos moleculares involucrados en la respuesta a diferentes tipos de estrés ambiental como son el estrés por calor, el oxidativo y, en particular, el estrés osmótico. En esta levadura se han caracterizado tres diferentes mecanismos para responder a condiciones de estrés osmótico: la vía de HOG (High Osmolarity Glycerol), la vía de PKA-AMPc y la vía mediada por el complejo represor Ssn6p-Tup1p-Sko1p. Estas vías parecen actuar de manera independiente aunque, al parecer, existe una gran interacción entre ellas lo cual le permite al organismo contender eficientemente con el estrés que se presente y el resultado es una respuesta compleja. Muchos de los elementos descritos en *S. cerevisiae* se han descrito en plantas y mamíferos, por lo que la información recabada para este organismo servirá para el estudio de esta respuesta en eucariotes superiores.

El trabajo que presento en esta tesis está dividido en dos partes principales.

En la primera parte hago una descripción de los patrones de expresión de tres genes de la levadura *S. cerevisiae* que no habían sido descritos previamente como genes de respuesta a estrés y a los cuales se denominaron genes *GRE* (Genes de Respuesta a Estrés). *GRE1* codifica para una proteína que no muestra similitud con ninguna proteína reportada hasta ahora en los bancos de datos; sin embargo, tiene características fisicoquímicas que se parecen a las descritas para las proteínas LEA (Late Embriogenesis Abundant) de plantas. *GRE2* codifica para una proteína que tiene similitud con enzimas de plantas que participan en la biosíntesis de metabolitos secundarios: 24-25% con dihidroflavonol 4-deshidrogenasa y 36% con cinamoil-CoA reductasa. *GRE3* tiene similitud con enzimas que participan en la síntesis de osmolitos, presenta un 64-65% de identidad con reductasas de xilosas de diferentes hongos (Garay-Arroyo y Covarrubias, 1999). Los tres genes están regulados positivamente por la vía de HOG y por los factores transcripcionales Msn2p y Msn4p y negativamente por AMPc. Tomando en cuenta los patrones de expresión de los genes *GRE* bajo condiciones de estrés osmótico y la semejanza (a nivel de secuencia o de características estructurales) que presentan los productos de estos genes con proteínas de plantas (caracterizados como parte de la respuesta al estrés hídrico), podemos sugerir la importancia de cada uno de estos genes en la respuesta de diferentes organismos a condiciones de estrés osmótico.

En la segunda parte del trabajo describo el hallazgo de proteínas tipo LEA codificadas en los genomas de *Escherichia coli* y de *S. cerevisiae*. Estas proteínas no

presentan similitud con las proteínas LEA a nivel de secuencia de aminoácidos pero tienen propiedades fisicoquímicas similares (alta hidrofiliidad y más del 8% de glicinas en su composición de aminoácidos). El análisis de la expresión de los genes que codifican para estas proteínas mostró que un gran porcentaje (100% para el caso de *E. coli* y 70% para el caso de *S. cerevisiae*) se induce bajo condiciones de estrés osmótico. Esta última correlación entre inducción y características fisicoquímicas refuerza la idea de que este tipo de proteínas debe ser de importancia general en la respuesta adaptativa a condiciones de déficit hídrico.



# Capítulo 1

## Introducción

### El potencial hídrico y su relación con el estrés osmótico.

El agua es el compuesto más abundante en los organismos vivos, constituye del 70 al 90% del peso de la mayor parte de estos y juega un papel central en la vida ya que es el solvente por excelencia de muchos componentes celulares y participa en la estructuración de proteínas, ácidos nucleicos y membranas. Por otro lado, el contenido hídrico intracelular determina el turgor y, por lo tanto, el tamaño y la forma de la célula y del organismo completo. La disponibilidad del agua es un factor que limita los procesos biológicos ya que las actividades metabólicas de los organismos requieren de una concentración hídrica óptima por lo que, en ambientes con poca agua, se han seleccionado mecanismos especializados que le permiten a un organismo contener con situaciones extremas. En algunos casos estos mecanismos están dirigidos a incrementar la eficiencia de uso del agua de tal forma que, en muy bajas concentraciones de ésta, sea posible mantener en funcionamiento procesos vitales y/o asegurar la progenie. De cualquier manera, la gran mayoría de las especies vive en ambientes intermedios, en donde la disponibilidad de agua varía temporal y espacialmente, por lo que se han desarrollado mecanismos para contener con ligeros cambios en el potencial hídrico.

El potencial hídrico ( $\psi_w$ ) es una medida cuantitativa de la energía libre asociada con el agua; el agua difunde de una región de alta a otra de baja actividad (ésta es la energía libre que lleva a cabo la difusión). Este potencial, en células que poseen pared celular, se define como:

$$(\psi_w) = P - \pi$$

esta ecuación indica que el  $\psi_w$  está afectado, principalmente, por dos factores: la presión de turgencia ( $P$ ) y la presión osmótica ( $\pi$ ). La reducción en el valor de  $\psi_w$  causado por la interacción del agua con los materiales insolubles y las paredes, se define como potencial de matriz, que no aparece en la ecuación por haberse comprobado empíricamente que su contribución al valor de  $\psi_w$  es despreciable (Taiz y Zeiger, 1991).

La presencia de la pared celular en bacterias, hongos y plantas es una característica distintiva de estos grupos de organismos, limita el tamaño de la célula y previene la ruptura de la membrana plasmática cuando se incorpora agua del medio. Únicamente en

organismos con pared celular existe lo que se conoce como presión de turgencia (P); ésta se define como la presión hidrostática generada por la membrana celular en contra de la pared celular para equilibrar las diferencias entre la presión interna y la externa.

La presión osmótica está definida por la concentración de solutos de una solución y es un valor positivo; sin embargo, se resta de la ecuación ya que la interacción del agua con los solutos tiene un efecto negativo sobre el  $\psi_w$ . Es importante considerar, para medidas del potencial osmótico, el tipo de soluto que se utilice cuando se trabaje con concentraciones osmóticas altas, ya que en función del soluto varía el potencial hídrico. La Fig. 1 muestra que las medidas de  $\psi_w$  generadas con la misma concentración de NaCl o de glicerol son diferentes. Esto quiere decir que, únicamente tomando en cuenta esta característica, para un organismo es más favorable acumular glicerol ya que con una menor concentración (en relación a la que tendría que tener acumulando NaCl) obtendrá el mismo valor de  $\psi_w$ . Este dato explica, de manera parcial, por qué evolutivamente se han seleccionado una serie de compuestos para osmorregular (estos compuestos osmocompatibles tienen otras propiedades que se analizarán más adelante).

Cuando dos soluciones con diferentes potenciales hídricos están separadas por una membrana semipermeable, el agua fluye a favor de su gradiente hasta que se igualan las presiones hídricas. Si un organismo se somete a un ambiente donde el potencial hídrico externo es más alto (una condición de estrés hipo-osmótico), el agua entra a la célula y ésta se hincha hasta que la presión de turgencia equilibra las diferencias de presión entre el interior y el exterior. Simultáneamente a estos cambios en la presión de turgencia, la presión osmótica (el segundo componente que define al  $\psi_w$ ) también cambia debido a la dilución en la concentración intracelular de solutos. El organismo necesita osmorregular (equilibrar las presiones hídricas modulando los niveles de osmolitos) y parte de la respuesta celular involucra la inducción de procesos que aceleran el catabolismo de solutos o que provocan su expulsión. Por el contrario, cuando el potencial hídrico externo es más bajo (estrés hiperosmótico), el agua sale de la célula y la presión de turgencia disminuye. En este caso, se requiere aumentar la presión osmótica interna para equilibrar las presiones hídricas (osmorregular). Este proceso se lleva a cabo ya sea por síntesis intracelular y/o por incorporación de solutos del medio.

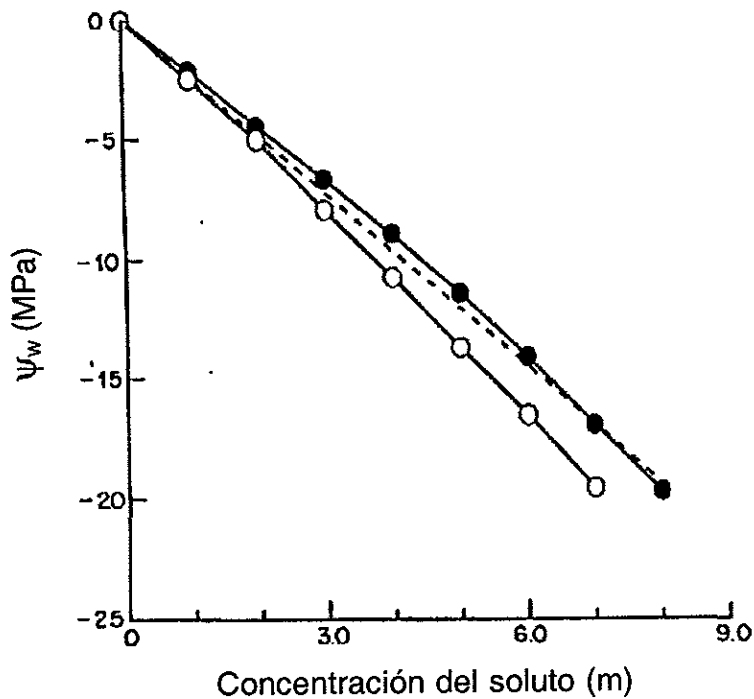


Figura 1. Potencial de agua de una solución a 20°C en relación a la molalidad del soluto. Las soluciones se ajustaron con NaCl (●) o con glicerol (○). (---) Representa una solución ideal. Modificado de Blomberg, 1988.

## Adaptación fisiológica.

### a) Compuestos osmocompatibles

La osmorregulación se puede definir como el proceso encargado de mantener constantes el volumen celular y la presión de turgencia cuando existen cambios en el potencial hídrico del medio. Este proceso involucra la síntesis o degradación de compuestos de bajo peso molecular conocidos como osmolitos compatibles que se acumulan para modular la presión osmótica intracelular. Los osmolitos compatibles se encuentran presentes en todos los organismos vivos y se postula que existió un proceso de evolución convergente en la selección del tipo de sustancias que se acumulan. Hay cuatro tipos de sustancias que preferencialmente se acumulan en condiciones de estrés hiperosmótico: (1) polioles y azúcares, (2) aminoácidos y derivados, (3) urea y metilaminas y (4) iones inorgánicos (Yancey, et al., 1982).

Se ha visto que los tres primeros grupos de osmolitos se acumulan, bajo condiciones de estrés osmótico, en prácticamente todos los organismos estudiados; por ejemplo, en microorganismos se favorece la acumulación de polioles (glicerol, sorbitol,

arabitol, etc..) y aminoácidos (glutamato, prolina y betaina); en plantas se acumulan, entre otras cosas, azúcares como la trehalosa y, en algunos peces y anfibios, se acumula preferencialmente urea (Yancey, et al., 1982). Sin embargo, se sabe que los iones inorgánicos son poco utilizados como compuestos osmorreguladores ya que altas concentraciones de iones intracelulares tienen efectos desestabilizadores. A la fecha sólo se ha encontrado un organismo, *Halobacterium*, que acumula  $K^+$  en condiciones de estrés osmótico, como compuesto osmocompatible. En esta bacteria se han seleccionado proteínas cuya composición presenta un mayor contenido de aminoácidos polares (Asp y Glu), lo cual evita interacciones hidrofóbicas y les permite funcionar en altas concentraciones iónicas (Yancey et al., 1982). La adaptación de este organismo ha sido tan extrema que esta bacteria solo puede vivir en potenciales hídricos negativos (ver Tabla 1); lo anterior parece indicar que, para colonizar nuevos nichos, se han seleccionado preferencialmente aquellos organismos que presentan una composición "normal" de aminoácidos en las proteínas y que osmorregulan usando cualquiera de las otras sustancias compatibles (ver Yancey et al., 1982).

#### **b) La levadura *S.cerevisiae* como modelo para el estudio de la respuesta a estrés osmótico en organismos eucariotes.**

La síntesis y degradación de los osmolitos constituye una parte de la respuesta de los organismos a condiciones de baja disponibilidad de agua, ya que un fenómeno tan complejo involucra cambios a muchos niveles y en diferentes grados. Por lo tanto, este estudio se facilita en modelos experimentales que poseen un ciclo de vida corto, un sistema genético sencillo, una fisiología caracterizada y cualidades que permitan un análisis molecular detallado, etc.. Este es el caso de la levadura *S. cerevisiae* que presenta, además, la ventaja adicional de que se conoce la secuencia completa de su genoma y existe un sistema accesible que permite adquirir los datos sobre el análisis inicial de esta información. Por las ventajas que sistemas como éste ofrecen, el grado de conocimiento acerca de los diferentes procesos avanza a gran velocidad; específicamente la respuesta de la levadura a estrés hiperosmótico ha sido ampliamente caracterizada, lo cual permite incursionar en los diferentes mecanismos involucrados en ella; así como discernir entre aquellos mecanismos que comparten más de una señal disparadora.

#### **c) Fisiología de *S. cerevisiae* en condiciones de estrés osmótico.**

La levadura *S. cerevisiae* es un organismo poco tolerante a bajos potenciales

hídricos (Tabla 1); sin embargo, se sabe que en condiciones naturales vive en las plantas y en las frutas que éstas producen, en donde ocurren cambios drásticos y constantes en la osmolaridad del medio. Por ejemplo, se sabe que durante las primeras etapas de la fermentación de la uva, la levadura está expuesta a una solución con más del 20% de azúcares (una condición con un potencial hídrico muy bajo), mientras que al final de este proceso metabólico se encuentra en una solución sin azúcar y con un 10% de etanol (Hohmann, 1997). Este organismo ha tenido que desarrollar estrategias adaptativas de regulación osmótica para mantener su actividad metabólica interna y poder vivir en estas condiciones cambiantes.

Tabla 1. Rango de tolerancia de varios microorganismos a diferentes potenciales hídricos

Organismo	Potencial Hídrico Bares			
	0	-200	-400	-600
<b>Bacterias</b>				
<i>Escherichia coli</i>	—			
<i>Staphylococcus aureus</i>	—————			
<i>Lactobacillus plantarum</i>	—————			
<i>Halobacterium salinarum</i>	—————			
<b>Hongos</b>				
<i>Saccharomyces cerevisiae</i>	—————			
<i>Saccharomyces rouxii</i>	—————			
<i>Debaromyces hanseii</i>	—————			
<i>Aspergillus niger</i>	—————			
<i>Xeromyces bisporus</i>	—————			
<b>Algas</b>				
<i>Dunaliella tertiolecta</i>	—————			
<i>Dunaliella viridis</i>	—————			

(1 Bar= 10MPa) Tomado de Blomberg, 1988

La adaptación de *S. cerevisiae* a estos cambios osmóticos se puede analizar a diferentes niveles: desde cambios permanentes como la presencia de pared celular, hasta cambios reversibles como son los siguientes:

(1) La imposición de un estrés con 7% de NaCl induce en la levadura una disminución de volumen en un 30-35% (Hohmann, 1997). Mientras más alta es la concentración de NaCl

aplicada en el medio, menor es el tamaño de las células de levaduras.

(2) Otra modificación que ocurre, es un engrosamiento de la pared celular y proyecciones de la misma hacia el citoplasma. También se han reportado cambios en la composición de la pared en estas condiciones de estrés (Morris et al., 1983).

(3) Tinciones con rodamina-faloídina muestran que el citoesqueleto se desorganiza en respuesta a un estrés hiperosmótico (Chowdhury et al., 1992). En estas condiciones de estrés, ensamblar el citoesqueleto le toma a la levadura, aproximadamente, dos horas (Hohmann, 1997). Este tiempo correlaciona con la recuperación del crecimiento sugiriendo que estos dos fenómenos están coordinados y que el ensamblaje del citoesqueleto es necesario para que la célula recupere su tasa de crecimiento.

(4) Como consecuencia del desensamblaje del citoesqueleto se pierde la polaridad en las nuevas gemaciones y se producen pseudohifas (Brewster y Gustin, 1994) y

(5) cambios en las vías metabólicas necesarios para contender con la nueva situación ambiental.

Muchos de los cambios adaptativos arriba descritos y otros, involucran la transcripción de nuevos genes. Los RNA mensajeros que se han descrito hasta el momento cuya transcripción se incrementa cuando el organismo se somete a una condición de estrés osmótico se pueden clasificar en diferentes grupos: aquéllos que tienen que ver con metabolismo de carbono, aquéllos que se caracterizaron por participar en un tipo específico de estrés: estrés por calor, estrés oxidativo, etc.. (ver Tabla 2). De este análisis, así como del hecho de que muchos genes además de incrementar su transcripción en condiciones de estrés osmótico lo hacen también en respuesta a otros tipos de estrés, ha sido posible hipotetizar que muchos de los productos de éstos genes deben de participar en la respuesta a una situación de estrés en general (ver Tabla 2). Asimismo, se ha visto que una serie de transcritos regulados por estrés osmótico (*GPD1* (glicerol 3-fosfato deshidrogenasa), *HXK1* (hexocinasa), *GLK1* (glucocinasa)), están involucrados en las primeras etapas de la glucólisis (ver Tabla 2) lo cual sugiere que el metabolismo de la levadura se está ajustando a las nuevas condiciones de crecimiento.

Además del estudio a nivel de transcripción, se han hecho análisis de cambios en el patrón de acumulación de proteínas, analizados por electroforesis bidimensional de levaduras crecidas en estrés osmótico. Esto ha permitido identificar, aproximadamente, 200 proteínas cuya expresión se incrementa bajo estas condiciones de estrés. De estas 200 proteínas, 20 incrementan su expresión significativamente en condiciones de estrés moderado (0.7M NaCl) (Blomberg, 1997) y sólo cinco de éstas son proteínas identificadas:

GPD1, GPP2 (glicerol 3-fosfatasa), CTT1 (catalasa citosólica), TPS2 (trehalosa fosfatasa) y HSP104. De estos datos se puede desprender que no se ha caracterizado ni la sexta parte de todas las proteínas inducidas en estrés osmótico por lo que nuestra comprensión del fenómeno todavía es muy incompleta.

1

Tabla 2. Caracterización de genes de respuesta a estrés osmótico

Gen	HOG	PKA	STRE	MSN2/4	SSN6 -TUP1	Otros estreses	Referencias.	
Metabolismo de carbono	<i>HXT1</i>	Sí	No				33	
	<i>ENO1</i>					FE	56	
	<i>GPD1</i>	Sí	No	Sí	No	Sí	C	17,33,48,62
	<i>GPP2</i>	Sí	No-Sí	Sí	Sí		C	17,33
	<i>GPP1</i>	No	Sí	Sí			C	33
	<i>HXK1</i>			Sí	Sí		C.F	17,33
	<i>GLK1</i>	Sí	Sí	Sí	Sí		C.F	33
	<i>GLO1</i>	Sí		Sí	Sí		I	37,69
	<i>DAK1</i>							
	<i>GAC1</i>			Sí			O.pH	61,62
	<i>TPS1</i>			Sí	Sí		C	17,62
	<i>TPS2</i>			Sí			C.FE	62
	<i>CTT1</i>	Sí	Sí	Sí		Sí	O.pH.FE	17,48,61,62,71
	<i>HSP12</i>		Sí	Sí	Sí	Sí	C.FE	33,62,71
	<i>HSP26</i>			Sí	Sí	Sí	FE	17,62,71
	<i>HSP104</i>		Sí	Sí	Sí	Sí	C.O.pH	17,61,62
	<i>DDR2</i>	Sí		Sí	Sí		C.O.FE.DD	61,62,69,71
	<i>DDR48</i>		No	No	Sí	Sí	DD	33,48,52,62
	<i>HAL1</i>	No	Sí	Sí*		Sí		48,62
	<i>ENA1</i>	Sí	Sí	Sí		Sí	C.F.pH.FE	33,48,49,62
<i>ALD2</i>	Sí	Sí	Sí*	Sí		FE	17,52,62	
<i>GRE1</i>	Sí	Sí	Sí	Sí		C.I.FE	30	
<i>GRE2</i>	Sí	Sí	No	Sí		C.O.I.FE	30	
<i>GRE3</i>	Sí	Sí	No	Sí		C.O.I.FE	30	
<i>PAI3</i>	Sí	Sí	Sí			FE	52,62	
<i>SIP18</i>	Sí	Sí	Sí			FE	52,62	
<i>UBI4</i>		Sí	Sí			C.O.pH	17,61,62,69	
<i>PTP2</i>	Sí		Sí				39	
<i>PTP3</i>	Sí		No				39	
<i>CYC7</i>			Sí				34	

C: Calor  
F: Frío  
O: Oxidativo  
I: Iónico  
FE: Fase Estacionaria  
DD: Daño a DNA  
\* Secuencia no funcional

Otro aspecto muy importantes para entender el proceso de osmorregulación y la acumulación de solutos en *S. cerevisiae* es la compartimentalización celular. En hongos y plantas, la vacuola es un organelo en el cual se acumulan iones y otro tipo de sustancias tóxicas permitiendo, de esta manera, disminuir la concentración citoplásmica y proteger macromoléculas y estructuras intracelulares importantes. En *S. cerevisiae* se han encontrado mutantes hiperosmosensibles afectadas ya sea en la biogénesis vacuolar o bien en la secreción adecuada de proteínas vacuolares al periplasma. Esto indica que en este organismo, la vacuola juega un papel importante en el proceso de osmorregulación (Latterich y Watson, 1991).

Se han postulado tres diferentes fases en la respuesta de la levadura a condiciones

de estrés hiperosmótico. La primera es un proceso puramente fisicoquímico que en la levadura es de, aproximadamente, un minuto (Hohmann, 1997) en el cual la célula pierde agua pasivamente para equilibrar las presiones hídricas. La segunda fase de la respuesta a este tipo de estrés es la adaptación fisiológica que se completa cuando el organismo osmorregula y, la tercer fase es la recuperación del crecimiento, que se dá una vez que la levadura se ha adaptado al nuevo ambiente.

La segunda fase, la adaptación fisiológica a un estrés hiperosmótico, involucra la acumulación de glicerol, que es el osmolito compatible por excelencia en este organismo (Blomberg y Adler 1989). El glicerol se produce por la vía de la glucólisis y se puede utilizar como fuente de carbono (ver Fig. 2).

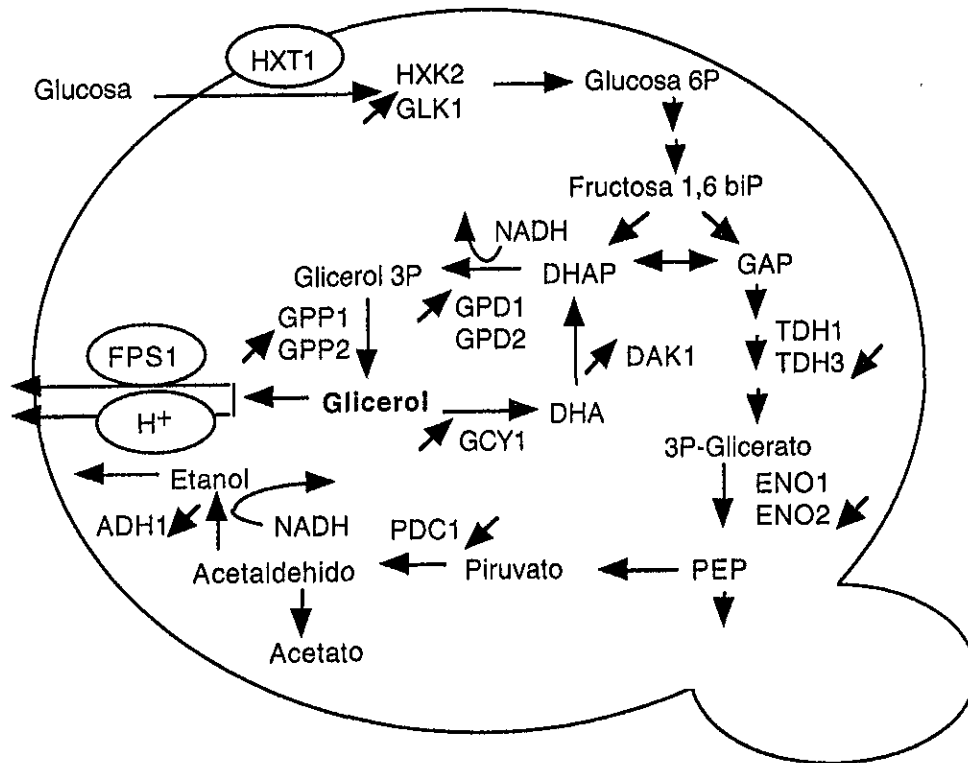


Figura 2. Vía de producción y de eliminación del glicerol. Modificado de Norbeck y Blomberg, 1997. Las flechas a la izquierda de las proteínas indican un incremento y las flechas a la derecha indican una disminución en la concentración y/o actividad de las mismas .

Entre los ajustes metabólicos que favorecen la producción de glicerol en condiciones de estrés osmótico se encuentran los siguientes:

- (1) Un incremento en la transcripción de los genes *GPD1* y *GPP2*; estos genes codifican



para enzimas que catalizan los últimos pasos en la vía de síntesis de glicerol (ver Fig. 2).

(2) Un aumento en la actividad de la fosfofructocinasa (es una enzima de la glucólisis que cataliza una reacción irreversible).

(3) Inhibición de la producción de etanol; probablemente debida a que la ADH1 utiliza el mismo cofactor que la GPD1 (NADH).

(4) Una disminución en la transcripción de los genes que codifican para las enzimas TDH3 (gliceraldehído 3-fosfato deshidrogenasa) y ENO2 (Enolasa). Esto es relevante para la producción de glicerol ya que, al no existir estas enzimas en cantidades normales, no se utiliza gliceraldehído-3-fosfato (GAP) y, por lo tanto, se aumenta la poza de fosfato de dihidroxiacetona (DHAP; precursor del glicerol) (ver Fig. 2).

(5) Baja en la expresión de la piruvato descarboxilasa (PDC1) y de la alcohol deshidrogenasa (ADH1) y una elevación en la actividad de la aldehído deshidrogenasa (ALD6). Estos cambios también son importantes para la producción de glicerol ya que se incrementa la concentración de NADH (Blomberg, 1997).

La acumulación de glicerol es producto, principalmente, de dos eventos : un incremento en su producción y una inhibición de su salida (Albertyn et al., 1994; Ansell et al., 1997; Blomberg y Adler 1989; Luyten et al., 1995).

La enzima responsable de la síntesis de glicerol en *S. cerevisiae* es la GPD. El transcrito de la GPD1 se induce en condiciones de estrés hiperosmótico y una interrupción del gen produce un fenotipo de osmosensibilidad (Albertyn et al., 1994; Larsson et al., 1993). Se ha descrito otra isoforma de este gen (GPD2) cuyo transcrito no se induce en estas condiciones de estrés y cuya interrupción no produce un fenotipo de osmosensibilidad. Sin embargo, la doble interrupción en *GPD1* y *GPD2* da lugar a una levadura altamente osmosensible que produce cantidades despreciables de glicerol (Ansell et al., 1997). El último paso en la síntesis de glicerol (ver Fig. 2) involucra la defosforilación del glicerol-3-fosfato a glicerol, esta reacción es catalizada por las fosfatasa conocidas como GPP1 y GPP2. La *GPP2* se regula transcripcionalmente por estrés hiperosmótico y por la vía de HOG, mientras que la *GPP1* se expresa constitutivamente y se ha demostrado que codifica para una proteína que, *in vitro*, funciona como fosfatasa (Norbeck et al., 1996). La interrupción de cualquiera de estos dos genes, *GPP1* y *GPP2*, o de ambos, no genera ningún fenotipo de sensibilidad a estrés osmótico (Hirayama et al., 1995), lo cual sugiere la existencia de una fosfatasa adicional involucrada en la producción de glicerol.

En cuanto al transporte de glicerol se refiere, se han descrito dos transportadores:

uno que funciona por difusión facilitada (FPS1) y otro por difusión activa como simporter con H<sup>+</sup> (Sutherland et al., 1997).

FPS1 es un miembro de un grupo de proteínas transmembranales formadoras de canales conocidas como MIP (por Membrane Intrinsic Protein); este transportador no se ha podido ubicar en ninguno de los cinco diferentes grupos filogenéticos en que se han clasificado a los miembros de esta familia, ya que el tamaño de la proteína es extremadamente grande en comparación con las otras (669 aminoácidos con respecto a 230 aminoácidos que, normalmente, tienen las 70 proteínas que se han descrito dentro de esta familia). Sin embargo, en cuanto a su función se refiere, se podría ubicar en el primer grupo, ya que se ha demostrado *in vivo* que transporta glicerol y se ha probado que el transportador de *E. coli* (GlpF) es capaz de complementar la función de transporte de una mutante en FPS1 (Luyten et al., 1995). Se sabe que el transporte mediado por FPS1 es el causante de la salida del 80% del glicerol en *S. cerevisiae* y, su apertura y cerrado responde, en cuestión de minutos, al cambio en el potencial hídrico intracelular (Luyten et al., 1995). Cuando la levadura se enfrenta a un estrés osmótico, el transportador se cierra permitiendo la acumulación intracelular de glicerol y, cuando hay un estrés hipo-osmótico, se abre para permitir la salida del mismo. Hasta el momento, la mutante de este gen es la única que se conoce que presenta un fenotipo de sensibilidad tanto en alta como en baja osmolaridad (Hohmann, 1997). Aparentemente, *FPS1* no está regulado por la vía de HOG (Luyten et al., 1995) pero la actividad enzimática está modulada por otra cinasa (PKC) que se sabe participa en la respuesta de la levadura a condiciones de estrés hipo-osmótico (Varela y Mager, 1996).

Por otro lado, el transportador activo, que funciona como simporter con H<sup>+</sup>, parece no estar implicado en el ajuste osmótico y se reprime cuando se crece a la levadura en glucosa como fuente de carbono; por lo que se ha propuesto que juega un papel en el cambio de un metabolismo glucolítico a uno gluconeogénico (Sutherland et al., 1997).

## **Vías de transducción de señales en la respuesta a estrés osmótico.**

### **1.- La vía de HOG**

En los últimos años, se ha publicado una gran cantidad de información acerca de las vías de transducción de señales que dan lugar a la respuesta a estrés osmótico en *S. cerevisiae*. Como se mencionó anteriormente, el estrés osmótico induce la síntesis de glicerol a través de la enzima GPD. La caracterización de mutantes incapaces de sintetizar glicerol ante una situación de estrés osmótico llevó a la identificación de la vía denominada

HOG por High Osmolarity Glycerol (Brewster et al., 1993). Esta vía regula la expresión del gen que codifica para la GPD1 y, de esta manera, induce la síntesis y consecuente acumulación de glicerol. Los datos acumulados a la fecha han permitido establecer un modelo en el cual se muestran algunos de los elementos involucrados en esta vía de transducción de señales (ver Fig. 3A). Se han descrito dos osmosensores en la membrana de las células: SLN1, que se inactiva en alta osmolaridad, y SHO1, que se activa por estrés hiperosmótico. Estos osmosensores actúan sobre una vía de MAP cinasas que termina en la fosforilación de la proteína HOG1 que, a su vez, presenta actividad de cinasa. Las MAP cinasas son módulos de señalamiento intracelular compuestos de tres cinasas que actúan secuencialmente; estas cinasas transmiten una señal extracelular al interior de la célula para dar una respuesta celular apropiada. En la vía de HOG, MAPKKK corresponde a SSK2/22 que fosforila a la cinasa MAPKK (que, en ésta vía, es PBS2) que, a su vez, fosforila a la cinasa MAPK (que, en ésta vía, es HOG1). La cinasa HOG1, se fosforila en un residuo de tirosina (esto sucede un minuto después de aplicar un estrés moderado: 0.3M NaCl (Brewster, et al.,1993)) y es dependiente de la cinasa PBS2 (Schüller et al., 1994). Esta vía regula a dos factores transcripcionales, Msn2p y Msn4p (es posible que HOG1 sea el responsable de la fosforilación de estos dos factores), que son los responsables de activar muchos genes de respuesta al estrés (Boy-Marcotte et al., 1998; Treger et al., 1998).

El osmosensor SLN1 forma un sistema de dos componentes con YPD1-SSK1; como se mencionó anteriormente, SLN1 se inactiva en presencia de condiciones de estrés hiperosmótico y no fosforila a SSK1 que, de esta manera, es activo y enciende la vía de HOG (ver Fig. 3B). Se sabía que las mutaciones en los genes *HOG1* y *PBS2* son osmosensibles; sin embargo no se encontró este fenotipo de osmosensibilidad en las mutantes de los genes *SSK1/SSK2* y *SSK1/SSK22* (ver Fig. 3A). Por esta razón se sugirió la existencia de un segundo osmosensor que regulara la vía de HOG a partir de PBS2. A este segundo osmosensor se le ha denominado SHO1 y, recientemente, se ha demostrado que esta proteína es capaz de formar un complejo con PBS2, HOG1 y STE11 (Maeda et al., 1994; Posas y Saito, 1997). STE11 es una MAPKKK que funciona en la vía de respuesta a feromonas; sin embargo, se sabe que cuando forma el complejo con las cinasas de la vía HOG, no interacciona con la otra vía permitiendo, de esta manera, especificidad en la respuesta. Se ha demostrado que no hay un cruce entre las dos vías (Posas y Saito,1997) y que PBS2 funciona como una proteína reclutadora para ambas.

También se ha reportado que *SLN1* es un gen esencial para la levadura y se ha propuesto que esta letalidad puede deberse al hecho de que se acumulen altas

concentraciones de los productos regulados por la vía de HOG ó a que metabolitos esenciales se desvíen para la producción de estos compuestos. Durante la búsqueda de mutantes supresoras de la mutación en *SLN1*, se identificaron cuatro diferentes fosfatasa, dos de tirosina PTP2 y PTP3 y dos de serina-treonina PTC1 y PTC3. Se sabe que los productos de PTP2 y PTP3 controlan la desfosforilación del producto de *HOG1* (Maeda et al., 1994), que son genes regulados transcripcionalmente en condiciones de estrés osmótico y a través de la vía de HOG (Jacoby et al., 1997); esto último sugiere que esta vía tiene un mecanismo intrínseco de inactivación. No se sabe cual es el sustrato de las otras dos fosfatasa (PTC1 y PTC3) pero se piensa que actúan sobre las últimas dos cinasa de la vía de HOG (PBS2 y HOG1) (Maeda et al., 1994).

Se sabía que el producto de HOG1 tiene que estar fosforilado para ser activo; recientemente se ha demostrado que es necesaria esta fosforilación para que la proteína se localice intranuclearmente. Este hecho hace pensar en que la actividad de fosforilación mediada por la proteína HOG1 es nuclear (Ferrigno et al., 1998).

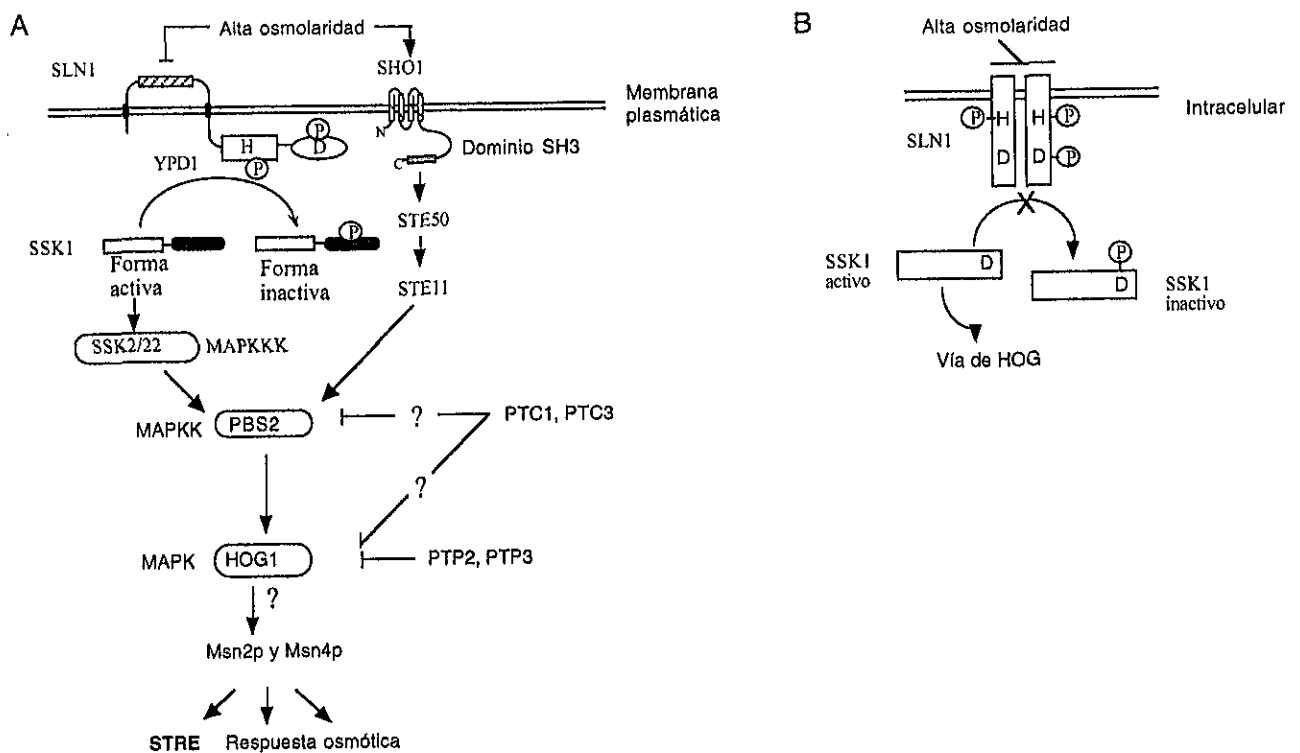


Figura 3. Vía de transducción de señales "HOG"

No se sabe cuales son los factores transcripcionales regulados por la vía de HOG;

sin embargo, si se conocen dos factores transcripcionales, Msn2p y Msn4p, que contienen dos dedos de Zinc  $^{2+}$  en su parte carboxilo terminal, involucrados en la respuesta a estrés en general. Se ha visto que cepas que carecen de Msn2p y Msn4p son sensibles a diferentes condiciones de estrés (como el osmótico) y no acumulan muchos de los transcritos de respuesta a estos diferentes tipos de estrés. Por otro lado, se demostró que estos factores transcripcionales se unen a una región del DNA conocida como STRE (ver más adelante) (Martínez-Pastor et al., 1996; Schmitt y McEntee, 1996). Específicamente, es posible pensar en la existencia de otros factores transcripcionales, además de Msn2p y Msn4p, que regulen la respuesta a estrés osmótico. Esta hipótesis se basa en el hecho de que la doble mutante *msn2-msn4* crece igual que la cepa silvestre en 0.8M de NaCl y que se encontró actividad de la fusión STRE a un gen reportero (ver siguiente sección) en la triple mutante *pbs2-msn2-msn4* bajo estas condiciones de estrés (Martínez-Pastor et al., 1996); esto último sugiere que esta respuesta es independiente tanto de los factores transcripcionales Msn2p-Msn4p como de la vía de HOG.

#### **Elemento presente en los promotores de genes regulados por estrés osmótico: STRE.**

La gran mayoría de los genes regulados por la vía de HOG tienen una secuencia en sus regiones promotoras conocida como STRE (STress Response Element; CCCCT o AGGGG). Este elemento es suficiente para inducir la expresión de un gen reportero ante diferentes situaciones de estrés, tales como estrés oxidativo, estrés osmótico, falta de nutrientes, crecimiento en bajo pH, etc.. (Schüller et al., 1994). Este hecho ha llevado a varios investigadores a proponer que el elemento STRE puede funcionar como un punto integrador de la respuesta a estrés en la levadura *S. cerevisiae* (Ruis y Schüller, 1995) y explicar, por lo menos parcialmente, la inducción de transcritos semejantes ante diferentes condiciones de estrés. A pesar de existir este cruce en la inducción de los genes en respuesta a diferentes condiciones de estrés, hay cierta especificidad en cada una de los diferentes condiciones de estrés; por ejemplo, se sabe que la fosforilación y, por lo tanto, la activación de HOG1, se da únicamente en respuesta a estrés osmótico (Schüller et al., 1994).

Como ya mencioné anteriormente, se han caracterizado dos factores transcripcionales Msn2p y Msn4p que, *in vitro*, se unen al elemento STRE. Por otro lado, mediante mapeo fino (footprinting), ha sido posible definir que la ocupación del elemento STRE presente en las regiones promotoras de los genes *CTT1* y *HSP104* se incrementa

en células en estrés y depende de la presencia de los dos factores transcripcionales Msn2p y Msn4p (Görner et al., 1998). Se demostró que ambos factores se acumulan en el núcleo bajo condiciones de estrés (calor, estrés osmótico, etanol, sorbato), siendo esta localización rápida, reversible e independiente de la síntesis de proteínas. Por otro lado, también se ha demostrado que la localización citoplásmica de estos factores está asociada a una alta actividad de la proteína cinasa A dependiente de AMPc (PKA) (Görner et al., 1998). Estos datos apuntan fuertemente a que la regulación de la respuesta a condiciones de estrés, mediada por estos dos factores transcripcionales, está dada por la localización citoplásmica-nuclear de los mismos y que esta localización depende, en gran medida, de la actividad de PKA.

Se han encontrado genes regulados por estrés osmótico y por la vía de HOG que no presentan STRE en sus regiones promotoras: *HOR7* (Hirayama et al., 1995) y *DDR48* (Miralles y Serrano, 1995) y genes como *HAL1*, *ALD2* y *ENA1* que presentan la secuencia STRE pero no es funcional (Alepuz et al., 1997; Márquez y Serrano, 1998; Miralles y Serrano, 1995). En algunos casos se propone la existencia de otra secuencia no identificada que funciona como STRE (Miralles y Serrano, 1995) y en otros, se argumenta que el elemento conocido como PDS (Post Diauxic Shift; A/TAGGGA o TCCCTT/A) es el mediador de la respuesta a estrés (Boy-Marcotte et al., 1998; Garay-Arroyo y Covarrubias, 1999). El elemento PDS, descrito a principios de los noventa, se activa tanto por estrés por calor como en las fases tardías de crecimiento de una levadura crecida en glucosa como fuente de carbono (Boorstein y Craig 1990a; Boorstein y Craig 1990b). La secuencia consenso que lo define es muy similar a la secuencia del elemento STRE y se ha observado, mediante experimentos de retardamiento en gel, que el elemento PDS es capaz de competir *in vitro* con el elemento STRE (Marchler et al., 1993), por lo que se ha considerado que puede ser reconocido por los mismos factores transcripcionales (Boy-Marcotte et al., 1998; Garay-Arroyo y Covarrubias, 1999).

## **2.- La vía de AMPc-PKA**

Existe una correlación directa entre la velocidad de crecimiento y la sensibilidad a estrés osmótico. Cuando la levadura se crece en glucosa como fuente de carbono (crecimiento rápido-fermentativo) los niveles de AMPc se incrementan y, por lo tanto, la actividad de la proteína cinasa dependiente de AMPc (PKA) también aumenta. Una alta actividad de esta cinasa provoca la represión transcripcional de muchos genes de respuesta a estrés; lo inverso también se cumple, una baja actividad de esta cinasa dá

como resultado la inducción transcripcional de estos genes. La cinasa PKA está compuesta por dos subunidades, la reguladora (inhibitoria) BCY1 y la catalítica, la cual a su vez está formada por los productos de tres genes homólogos *TPK1*, *TPK2* y *TPK3*. El AMPc modula la actividad de esta cinasa, ya que se une a la subunidad reguladora BCY1 y libera la parte catalítica que, una vez libre, se activa (ver Fig. 4).

Los efectos de PKA sobre los genes de estrés están mediados, al menos en parte, por el papel que juega esta cinasa en la localización citoplásmica de los factores transcripcionales Msn2p y Msn4p (ver apartado anterior; Görner *et al.*, 1998). Estos datos sugieren que el AMPc pudiera ser un mediador de la respuesta a estrés ya que, de esta manera, es posible coordinar la respuesta metabólica y la protectora. Así pues, genes que poseen en sus regiones promotoras el elemento STRE, ó el elemento PDS, responden a cambios metabólicos durante el crecimiento y/o a cambios ambientales bajo condiciones de estrés modulados por los niveles de AMPc y, por lo tanto, de Msn2p y Msn4p activo (en el núcleo).

Actualmente se desconoce el mecanismo por el cual la levadura aumenta los niveles de AMPc cuando crece en glucosa como fuente de carbono, pero se ha propuesto que la despolarización de la membrana y/o la acidificación intracelular pueden ser los disparadores de la señal (Thevelein, 1994). Una vez que se internaliza la señal aumentan los niveles de AMPc que están determinados por dos procesos opuestos: por síntesis dependiente de la adenilato ciclasa (CYR1 o CDC35) e hidrólisis mediada por dos fosfodiesterasas; PDE1 y PDE2, una de baja y otra de alta afinidad, respectivamente.

La actividad de la adenilato ciclasa (CYR1) está regulada por las proteínas RAS que son activas cuando están unidas a GTP; esta unión está regulada positivamente por la proteína (CDC25), que estimula el intercambio GDP-GTP y, negativamente, por las proteínas IRA1-IRA2, que estimulan el intercambio GTP-GDP (ver Fig. 4).

Recientemente se ha reportado que el gen que codifica para la proteína PDE2 se regula por los factores transcripcionales Msn2p y Msn4p (Treger *et al.*, 1998), lo cual quiere decir que la transcripción de este gen está regulada indirectamente por la actividad de PKA que, a su vez, está regulada por la concentración de AMPc que, depende de la actividad de estas fosfodiesterasas.

Por otro lado, se ha demostrado que existe una correlación inversa entre los niveles de AMPc (y por lo tanto de PKA activo) y la concentración de NaCl presente en el medio de crecimiento (aún en levaduras crecidas en glucosa como fuente de carbono); se demostró que el efecto es más pronunciado en respuesta a concentraciones más altas de NaCl

(Márquez y Serrano, 1996).

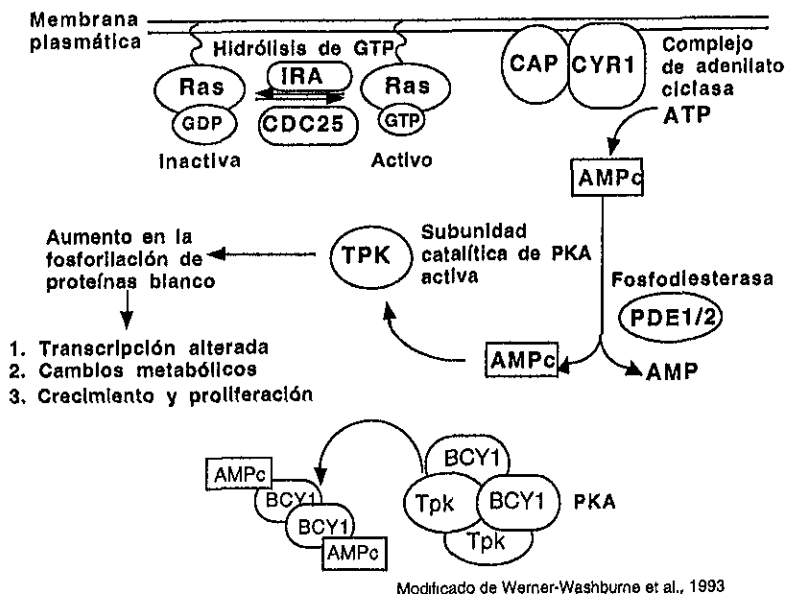


Figura 4. Vía de regulación de la proteína cinasa-A (PKA) dependiente de AMPc.

Además de BCY1, recientemente se han caracterizado dos proteínas que unen AMPc. Estas proteínas se encuentran ancladas en la membrana por medio de un residuo fosfatidilinositol y se ha propuesto que se liberan por una fosfolipasa C cuando el organismo crece en fuentes de carbono fermentables (Thevelein, 1994). No se sabe si alguna de estas dos proteínas participa en los fenómenos mediados por AMPc hasta ahora descritos.

### 3.- La vía de regulación mediada por el complejo represor Ssn6p-Tup1p-Sko1p.

Recientemente se ha identificado una secuencia en la región promotora del gen *HAL1* que funciona como modulador negativo en la respuesta de la levadura a condiciones de estrés osmótico. Esta región es suficiente para reprimir la transcripción y conferir respuesta a estrés osmótico al promotor del gen *CYC1*. Los datos reportados sugieren que esta secuencia es el blanco de un complejo de proteínas llamado Ssn6p-Tup1p-Prp que se ha caracterizado como parte de un sistema de represión bajo otros tipos de estrés (Márquez *et al.*, 1998). Actualmente se sabe que, en el caso específico de estrés



osmótico, la Prp es un factor transcripcional conocido como Sko1p, que contienen dos dedos de Zinc  $^{2+}$  en su parte carboxilo terminal, que se une a una secuencia llamada CRE (TGACGTTT). Este elemento es similar a una secuencia reportada en mamíferos de respuesta a AMPc (Proft y Serrano, 1999). Se ha sugerido que el complejo Ssn6p-Tup1p-Sko1p está unido, en condiciones normales de crecimiento, a una secuencia del promotor de *HAL1* y que este complejo se libera bajo condiciones de estrés osmótico. El sistema de Ssn6p-Tup1p-Sko1p parece estar involucrado en el control de la transcripción de genes tanto dependientes (*CTT1*, *ENA1*, *ALD2* y *SIP18*) como independientes de la vía de HOG (*DDR48* y *HAL1*). Esto indica que la transcripción de muchos de los genes activados por la vía de HOG se reprimen por la unión de este complejo y, por lo tanto, para su expresión óptima no sólo se requeriría de la activación por la vía de HOG sino también de la des-represión del complejo Ssn6p-Tup1p-Sko1p. Parecería entonces que, la contribución de la vía de HOG en la activación de muchos de los genes de respuesta a estrés, esta enmascarada por la liberación del complejo Ssn6p-Tup1p-Sko1p. Esta idea se ve apoyada por el hecho de que el fenotipo y control transcripcional observado en la mutante de *hog1* se puede suprimir con la mutante en *ssn6-tup1* y con la mutante en *sko1* (Márquez et al., 1998; Proft y Serrano, 1999).

## I. Hipótesis

A través de caracterizar genes que se regulen por condiciones de estrés osmótico podremos entender un poco más el proceso que le permite a la levadura adaptarse a vivir bajo estas condiciones de estrés.

## II. Objetivo general

Aislar y caracterizar genes de levadura inducidos por condiciones de estrés osmótico.

Prácticamente todo el trabajo que se conoce hasta el momento acerca de la respuesta de la levadura *S. cerevisiae* a condiciones de estrés osmótico se ha obtenido gracias a la facilidad metodológica que presenta este organismo para obtener mutantes y poder caracterizarlas genéticamente. Sin embargo, una gran cantidad de genes tienen homólogos funcionales (ésto queda claro con la posibilidad de analizar la secuencia completa del genoma de éste organismo) por lo que no es posible obtener mutantes con fenotipo en estos genes y *caracterizar el fenómeno completamente. Una alternativa, que complementa a la estrategia de obtención de mutantes, para estudiar la respuesta de la levadura a una cierta condición particular de estrés es la obtención de genes inducidos bajo estas condiciones de estrés. Esta última estrategia es la que decidimos abordar en este trabajo.*

## III. Objetivos particulares

1. Aislar genes cuya transcripción se incremente en condiciones de estrés osmótico. Secuenciar estos productos, deducir la secuencia de aminoácidos y comparar con los bancos de datos para buscar homologías con otras proteínas.
2. Caracterizar la expresión de los transcritos bajo diferentes condiciones de estrés (iónico, oxidativo, calor, etc..) mediante experimentos tipo northern.
3. Determinar el número de copias de los genes *GRE* en el genoma de la cepa RS58 de levadura a través de experimentos tipo Southern .
4. Llevar a cabo mutaciones específicas, en cada uno de los genes aislados, por medio de una disrupción génica. Caracterizar el fenotipo de las cepas mutantes en diferentes medios de crecimiento.

YEAST

*Yeast* 14, 000–000 (1999)

## Three Genes Whose Expression is Induced by Stress in *Saccharomyces cerevisiae*

ADRIANA GARAY-ARROYO<sup>1</sup> AND ALEJANDRA A. COVARRUBIAS<sup>1\*</sup>

<sup>1</sup>*Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apdo. Postal 510-3, Cuernavaca, Morelos 62250, México*

In this work we report the isolation and characterization of three genes induced by different stress conditions in the yeast *Saccharomyces cerevisiae*. These genes, named *GRE1*, *GRE2* and *GRE3*, were identified by the differential display technique using total RNAs obtained from yeast grown under hyperosmotic conditions. Northern analysis of RNA obtained from different growth conditions shows that their corresponding transcripts accumulate not only in response to osmotic stress but also to ionic, oxidative and heat stress. Analysis of the deduced amino acid sequences indicated that *GRE1*, *GRE2* and *GRE3* correspond to ORFs YPL223C, YOL151W and YHR104W, respectively. Additionally, it suggested that *GRE1* encodes a hydrophilic polypeptide that it is not homologous to any known protein but has features resembling the late embryogenesis abundant (LEA) proteins characterized in higher plants; *GRE2* encodes a putative reductase with similarity to plant dihydroflavonol-4-reductases; and *GRE3* codifies for a keto-aldose reductase highly related to fungal xylose-reductases. The three genes are induced in the late growth phases in agreement with the presence of PDS elements in their promoter regions. The three of them are under the control of the HOG pathway, even though *GRE1* and *GRE2* promoter regions do not present the consensus core STRE sequence. In addition, *GRE1* and *GRE3* are regulated negatively by the cAMP-PKA transduction pathway and positively by the transcriptional factors Msn2p and Msn4p. Gene disruptions of the *GRE* genes did not show a phenotype in any of the tested stress conditions. Copyright © 1999 John Wiley & Sons, Ltd.

**KEY WORDS** — stress response; HOG pathway; LEA proteins; aldose reductase; dihydroflavonol-4-reductase; *GRE* genes

### INTRODUCTION

Organisms have mechanisms that allow them to adapt to constant fluctuations in their living environment. These mechanisms, which involve the perception of the changing environment and the transduction of the stress signal, result in the modulation of pre-existing proteins and the transcriptional induction of a set of genes whose products counteract the hazardous consequences of the imposed stress.

Correspondence to: A. A. Covarrubias, Departamento de Biología Molecular de Plantas, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apdo. Postal 510-3, Cuernavaca, Morelos 62250, México. Tel: 52 (73) 29-1643; fax: 52 (73) 13-9988; e-mail: crobles@ibt.unam.mx; adriana@ibt.unam.mx.

Contract/grant sponsor: Consejo Nacional de Ciencia y Tecnología, México; Contract/grant numbers: 0131P-N9506, 26242-N.

Contract/grant sponsor: European Community; Contract/grant number: C11\*CT94-0082.

The versatility of the yeast *Saccharomyces cerevisiae* as an experimental model has allowed the study of different stress responses, such as heat shock, osmotic and oxidative stresses, nutrient starvation, etc. (Mager and Ferreira, 1993; Serrano, 1996). Among other contributions, these studies have provided evidence that suggest the existence of a general stress response system that can be activated by various factors and, upon activation, protect cells against a number of conditions that endanger their survival (Mager and Varela, 1993; Parsell and Lindquist, 1994).

In *S. cerevisiae* the existence of at least four different DNA elements that respond to stress signals and participate in the transcriptional regulation of different genes has been demonstrated. Among the best characterized is the heat shock element (HSE), which interacts with the so called heat shock transcription factor (HSF) (Parsell and Lindquist, 1994). Other elements participate in the

response to multiple stress conditions, such as the PDS (post-diauxic shift) element that mediates gene activation during diauxic growth and stationary phase, and under heat shock stress (Boorstein and Craig, 1990a, 1990b; Werner-Washburne *et al.*, 1993). The stress response element (STRE) constitutes a sequence that is activated by various stress conditions, such as nitrogen starvation, osmotic and oxidative stress, low external pH, weak organic acids, and ethanol (Belazzi *et al.*, 1991; Marchler *et al.*, 1993; reviewed in Ruis and Schüller, 1995). Although the transcription factor binding to STRE has not yet been unambiguously identified, recent observations show that the transcription activators Msn2p and Msn4p are not only required for a high level of expression of genes activated via STRE, but also bind specifically to these elements *in vitro* (Martínez-Pastor *et al.*, 1996; Schmitt and McEntee, 1996). Finally, the AP-1 responsive elements (ARE) have been identified in the promoter region of genes typically involved in the response to oxidative stress (Gounalaki and Thireos, 1994; Ruis and Schüller, 1995).

Increase in the external osmolarity induces a water deficit situation, which reduces growth and viability of yeast cells as a consequence of the loss of an osmotic gradient across the membrane. A water deficit can be reached during adverse environmental conditions such as freezing, salt stress or desiccation. Certain developmental stages, such as spore formation, also implicate dehydration. One of the first responses, maintained throughout evolution, that was characterized in organisms living under hyperosmotic conditions was the synthesis of osmolytes or compatible solutes (Yancey *et al.*, 1982). In *S. cerevisiae*, glycerol is the major osmolyte synthesized under hyperosmotic stress and is responsible for most of the osmotic adjustment of this organism (Albertyn *et al.*, 1994; Blomberg and Adler, 1989). The enzyme responsible for the production of this osmolyte is glycerol-3-phosphate dehydrogenase, NAD<sup>+</sup> (GPD). The *GPD1* gene is induced under osmotic stress and the lack of its product renders the yeast osmosensitive (Albertyn *et al.*, 1994). A second gene encoding a GPD isozyme, *GPD2*, has been described, as being important for glycerol production, although it is not induced by osmotic stress conditions (Ansell *et al.*, 1997).

In *S. cerevisiae*, the signal transduction pathway that ends in the production of glycerol has been widely characterized and constitutes an example of

a two-component regulatory system in eukaryotes (Maeda *et al.*, 1994). This transduction pathway consists of a MAP kinase cascade (MAPKKK, Ssk2/22p; MAPKK, Pbs2p; MAPK, Hog1p; Brewster *et al.*, 1993; Maeda *et al.*, 1994) that results in the induction of several genes, including *GPD1* (Albertyn *et al.*, 1994; Schüller *et al.*, 1994). The kinase cascade is activated by the product of two different genes that encode osmosensors on the membrane of yeast cells: *SLN1*, which forms the two-component system with *YPD1-SSK1*, and *SHO1*, which acts on *PBS2* through *STE11* (Maeda *et al.*, 1994; Maeda *et al.*, 1995; Posas *et al.*, 1996; Posas and Saito, 1997). This regulatory system is known as the high osmolarity glycerol (HOG) pathway as it was initially characterized as the regulatory mechanism for glycerol production during yeast growth in high osmolarity medium (Brewster *et al.*, 1993).

In this work we report the isolation and characterization of three genes, *GRE1*, *GRE2* and *GRE3*, whose expression is induced by osmotic stress. We show that these genes also respond to other adverse situations: oxidative, heat, and ionic stresses and during the normal development of yeast, in the diauxic and postdiauxic-stationary growth stages. Furthermore, we report that the three *GRE* genes are positively regulated by the HOG pathway. Two of them, *GRE1* and *GRE3*, are positively regulated by the transcriptional factors Msn2p–Msn4p and negatively regulated by the cAMP–PKA transduction pathway.

## MATERIAL AND METHODS

### *Yeast strains, media and growth conditions*

The *S. cerevisiae* strains used in this work are described in Table 1. Cells were routinely grown on YPD medium [1% yeast extract (Difco), 2% Bacto-peptone (Difco), and 2% glucose] or SD medium [2% glucose, 0.7% Yeast Nitrogen Base without amino acids (Difco), and 0.05 M MES (Sigma), pH 6.0] supplemented with adenine (30 µg/ml), uracil (30 µg/ml), histidine (30 µg/ml) and leucine (100 µg/ml) at 30°C, unless otherwise indicated. Solid media contained 2% agar (Difco). For all the shock experiments, cells were grown in YPD medium to mid-logarithmic phase (OD<sub>660</sub> of 0.5) and solutions of sorbitol, hydrogen peroxide, sodium chloride or lithium chloride were added to the final concentrations indicated. For the heat shock experiments, yeast cultures grown in YPD

GENES INDUCED BY STRESS IN YEAST

Table 1. Strains of *S. cerevisiae* used in this study.

Strain	Genotype	Source or reference*
RS58	MAT $\alpha$ <i>ura3, leu2, his3, ade2</i>	Gaxiola <i>et al.</i> , 1992
<i>hog1</i>		RS58 <i>hog 1::URA3</i>
<i>gre1</i>	RS58 <i>gre 1::URA3</i>	
<i>gre2</i>	RS58 <i>gre 2::URA3</i>	
<i>gre3</i>	RS58 <i>gre 3::URA3</i>	
<i>bcy1</i>		RS58 <i>bcy1::URA3</i>
W303	MAT $\alpha$ , SUC2, <i>ade2, can1, his3, leu2, trp1, ura3</i>	Martínez-Pastor <i>et al.</i> , 1996
<i>msn2-<del>msn4</del></i>	W303 <i>msn2-3::HIS3</i> and <i>msn4-1::TRP1</i>	Martínez-Pastor <i>et al.</i> , 1996

\*Source is this study, unless otherwise indicated.

were shifted from 30°C to 37°C. Cells were maintained under the stress conditions during the periods indicated. For the batch culture experiment, stationary-phase YPD yeast cultures were used to inoculate YPD medium containing 2 M sorbitol or, for the control conditions, YPD. Aliquots were removed to determine the cell density and cells were harvested at different times, as indicated.

*Yeast transformation*

We followed the Li-acetate method for transformation of yeast, as described initially by Ito *et al.* (1983) and modified by Lauermaun (1991).

*Differential display method*

The 'differential display' of specific transcripts is based on the PCR amplification of cDNAs (Liang and Pardee, 1992) obtained from total RNA (1 µg) purified from yeast cultures grown to an OD<sub>660</sub> of 0.5 in YPD medium or in YPD supplemented with 2 M sorbitol. An oligo dT(15) with a GC at its 3' end was used for the cDNA synthesis. The cDNA obtained was the template for the subsequent PCR that was carried out with two different decamers, 5'-GGC GGT TGT C-3' for *GRE1* and *GRE2*, and 5'-CTG GGC AAC T-3' for *GRE3*, in combination with the oligo dT(15)-GC. The procedure used was that of Liang and Pardee (1995). The accumulation pattern of the transcripts corresponding to the selected amplified cDNAs was confirmed by using the fragments eluted from the DNA sequencing gel as probes in Northern blot experiments. DNA fragments corresponding to the osmotic

regulated transcripts were cloned in pKS vectors (Stratagene) and subjected to further sequence analysis.

*Isolation of the full-length genes*

The DNA sequences obtained from the PCR fragment clones were used to obtain the complete sequence of the three *GRE* genes from the *S. cerevisiae* genome database (Stanford University, USA). To obtain full-length sequences of the *GRE* genes, we carried out PCR amplifications using specific primers designed according to the known corresponding sequences.

The primers used were: for *GRE1*, 5'-GCC AAA CAC TAC CGC ATA AAA GC-3' and 5'-CAA ACT GAC ATT AAT AAA AAC G-3'; for *GRE2*, 5'-GAT AGC AGT ATC ACA CGC CC-3' and 5'-CAG CTT TTC TTG TGA TGT GTTGC-3'; and for *GRE3*, 5'-ATA AAT CGT AAA GGA AAA TTG G-3' and 5'-GGA AGC TGC TAT TAA TAT CCC G-3'. The PCR reactions using genomic DNA as template were as follows: 35 cycles of 30 s at 94°C, 30 s at 55°C and 45 s at 72°C. The resulting PCR products were cloned in *EcoRV* digested pKS plasmids. The clones obtained were further used as probes for the corresponding Northern and Southern blot experiments, as described below.

*Constructions of yeast strains carrying the disrupted genes*

Disruptions of the three *GRE* genes were constructed using a *URA3-HindIII* cassette (obtained from plasmid Yep24) as selective marker. To get the *GRE1* gene-disruption, the pKS-*GRE1* clone

was digested with *Bam*HI, ends were filled-in with Klenow and ligated to the refilled *URA3-Hind*III cassette. In the case of the *GRE2* gene, the *URA3* cassette was digested with *Hind*III, refilled, and inserted into an *Eco*RV internal site. The *GRE3* gene was disrupted by inserting the *URA3-Hind*III cassette into the internal *Hind*III site of the pKS-*GRE3* clone. Linear fragments containing the *GRE* gene disruptions were introduced into the RS58 strain by transformation. The desired mutants were selected on SD medium without uracil and the disruptions were confirmed by Southern blot experiments. The *HOG1* and *BCY1* disruptions were obtained by transforming RS58 with independent linear fragments containing the disrupted genes. The mutated *HOG1* and *BCY1* genes were kindly given by S. Hohmann and R. Gaxiola, respectively. The disruptions were confirmed, in the case of *HOG1*, by testing their osmotic sensitivity and, at the regulatory level, looking at the expression of *GPD1* and *CTT1* genes under osmotic stress (data not shown). In the case of the *BCY1* gene, transformants were tested by their glycogen content using the iodine staining procedure (Chester, 1968). The strains W303 and W303 (*msn2-msn4*) were kindly given by Dr F. Estruch.

#### *Nucleic acids purification*

Total RNA was extracted as indicated by Collart and Oliviero (1995), except that a final step was included in which total RNA was precipitated with 2 M LiCl to eliminate transfer RNA. The RNA was resuspended in one volume of water and one volume of 4 M LiCl and incubated for 2 h at 4°C. RNAs were then centrifuged at 16 000 × *g* for 5 min and washed twice with 70% (v/v) ethanol. Finally, the resulting RNA was resuspended in water and the concentration determined spectrophotometrically. Genomic DNA was obtained as described by Hoffman (1995).

#### *Northern and Southern blot analysis*

For Northern blot experiments, 10 µg of total RNA was separated by electrophoresis in formaldehyde-1.2% agarose gel and blotted to Hybond membranes following standard protocols (Hybond-Amersham). The mRNA levels were quantified using the NIH Image 1.60b7. The resulting values were normalized with those obtained from the 25S rRNA hybridization signal; those numbers represent an arbitrary unit which

was obtained by subtracting the intensity value obtained for each transcript for a loading correction factor according to the intensities of the 25S rRNA band. For Southern blot hybridization experiments, 20 µg of *Eco*RI-digested genomic DNA was used per lane. DNA was subjected to electrophoresis in agarose gels and blotted to nylon-N<sup>+</sup> membranes (Hybond-Amersham). In all hybridization experiments, the membranes were probed with fragments carrying the full-length *GRE* genes. Probes were labelled with a commercial random primer kit (Dupont) using [ $\alpha$ -<sup>32</sup>P] dCTP (3000 Ci/mmol; Amersham). High-stringency conditions were used for all hybridizations and washes.

#### *DNA sequencing*

DNA sequencing was carried out following the Sequenase procedure (US Biochemicals). Both strand sequences from pKS-derived clones were obtained using 'reverse' and 'universal' primers.

## RESULTS

### *Isolation and characterization of three genes induced by osmotic stress*

Three genes whose expression is induced by osmotic stress (2 M sorbitol) were identified by differential display RT-PCR. To confirm the osmotic induction of these transcripts, we used the selected amplified fragments for Northern blot experiments (data not shown). Three of the tested fragments that showed transcript accumulation in response to osmotic stress were selected for further cloning and sequencing. The DNA sequences obtained were compared against the *S. cerevisiae* genome database (Stanford University). Three different genes were identified and their complete nucleotide sequences were compared against databases, showing that they are genes not identified previously as responsive to hyperosmotic conditions in yeast. Since the three genes were responsive not only to osmotic stress but also to additional stress conditions, we called them '*GRE*' for '*Genes de Respuesta a Estres*' (stress responsive genes, Spanish). The deduced amino acid sequences indicated that *GRE1* encodes a highly hydrophilic protein with an amino acid composition similar to that observed for proteins of the LEA (Late Embryogenesis Abundant) family from plants (Dure, 1993; Ingram and Bartels, 1996; Figure 1). *GRE2* encodes a protein 24–25%

GENES INDUCED BY STRESS IN YEAST

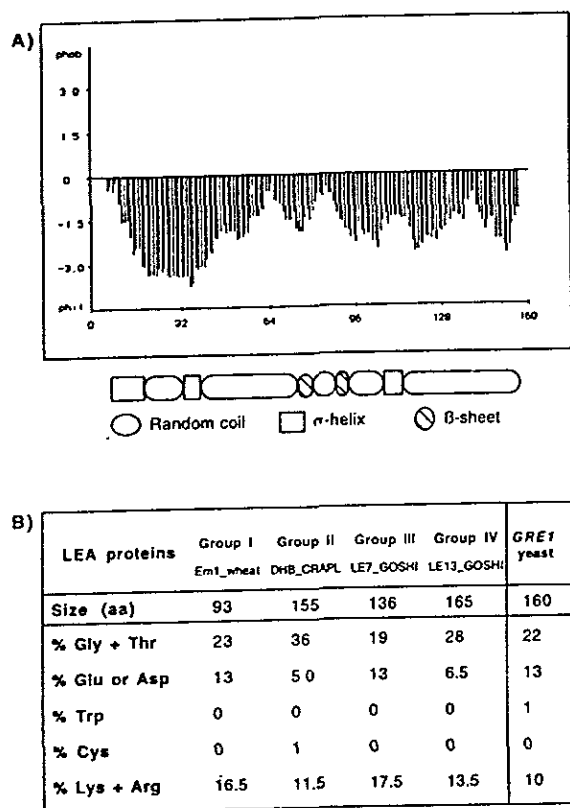


Figure 1. Structural characteristics of the *GRE1* gene product. (A) Hydropathic profile and predicted secondary structure of the protein encoded by the *GRE1* gene. This analysis was made using the Wisconsin Sequence Analysis Package. (B) Comparison between the amino acid composition of the *GRE1* gene product and that corresponding to a representative member of the different groups of the LEA proteins from plants. The table in this figure shows those amino acids which are distinctive in LEA proteins (Dure, 1993).

identical to dihydroflavonol-4-reductases from different plants (Casamayor *et al.*, 1995; Sparvoli *et al.*, 1994), and *GRE3* encodes a putative ketoaldose reductase that shows high identity with xylose reductases from different yeasts (Mewes *et al.*, 1997): 69% with that isolated from *Kluyveromyces lactis* (Billard *et al.*, 1995), and 58% with the one from *Pichia stipitis* (Takuma *et al.*, 1991). In addition, *GRE3* shows significant homology ( $\approx 34\%$  identity) with aldose reductases from other organisms (Mewes *et al.*, 1997). The tetrapeptide IPKS, which is considered to be the active site for both aldehyde and aldose reductases (Bartels *et al.*, 1991), is present in this protein. The Accession Nos for the *GRE* genes on the yeast genome database at

Stanford University are: YPL223C for *GRE1*, YOL151W for *GRE2* and YHR104W for *GRE3*.

Genomic Southern blot analysis indicates that strain RS58 contains a single copy of *GRE1* and *GRE3*, whereas three additional homologous sequences were detected for *GRE2* (data not shown). These data are in agreement with the information obtained from the yeast genome database, which indicates the presence of single *GRE1* and *GRE3* genes, as well as the presence of three genes with  $\geq 60\%$  homology to *GRE2*. Additionally, we carried out a comparative analysis (BLAST) of the deduced amino acid sequences, which indicated that the products of the three *GRE2* homologous genes correspond to other putative reductases with an identity of 55–60% to the *GRE2* protein (YDR541C, YGL157W, YGL039W). Five additional putative aldose reductases, 30–40% identical to *GRE3*, were also found. However, none of the additional aldose reductases seem to be a xylose reductase, based on their identities with *bona fide* xylose reductases. No protein homologous to the *GRE1* product was found.

*Accumulation patterns of the GRE transcripts during growth under stressed and non-stressed conditions*

Northern blot experiments using total RNA obtained from yeast cultures grown under osmotic stress and harvested at logarithmic phase ( $OD_{660}$  of 0.5), confirmed that all *GRE* genes are clearly responsive to hyperosmotic conditions (Figure 2A). The sizes of the *GRE* transcripts were determined to be, approximately 0.5 kb for *GRE1* and 1.2 kb for *GRE2* and *GRE3*. In a typical Northern hybridization, we detected a 25-fold induction for *GRE1*, 14-fold induction for *GRE2* and 38-fold induction for *GRE3* in response to high osmolarity in the growth medium (Figure 2A).

To observe the *GRE* transcript accumulation under stressed and non-stressed conditions, we took aliquots from cultures grown on YPD and YPD+2M sorbitol at different times during growth to stationary phase, as indicated in Figure 2B. The resulting data indicate that although transcript accumulation varied with osmotic stress, it was also affected by the growth stage (Figure 2C). When yeast cells were grown in a hyperosmotic environment, the highest accumulation was detected during the log phase, gradually decreasing as the culture reached the stationary phase. The

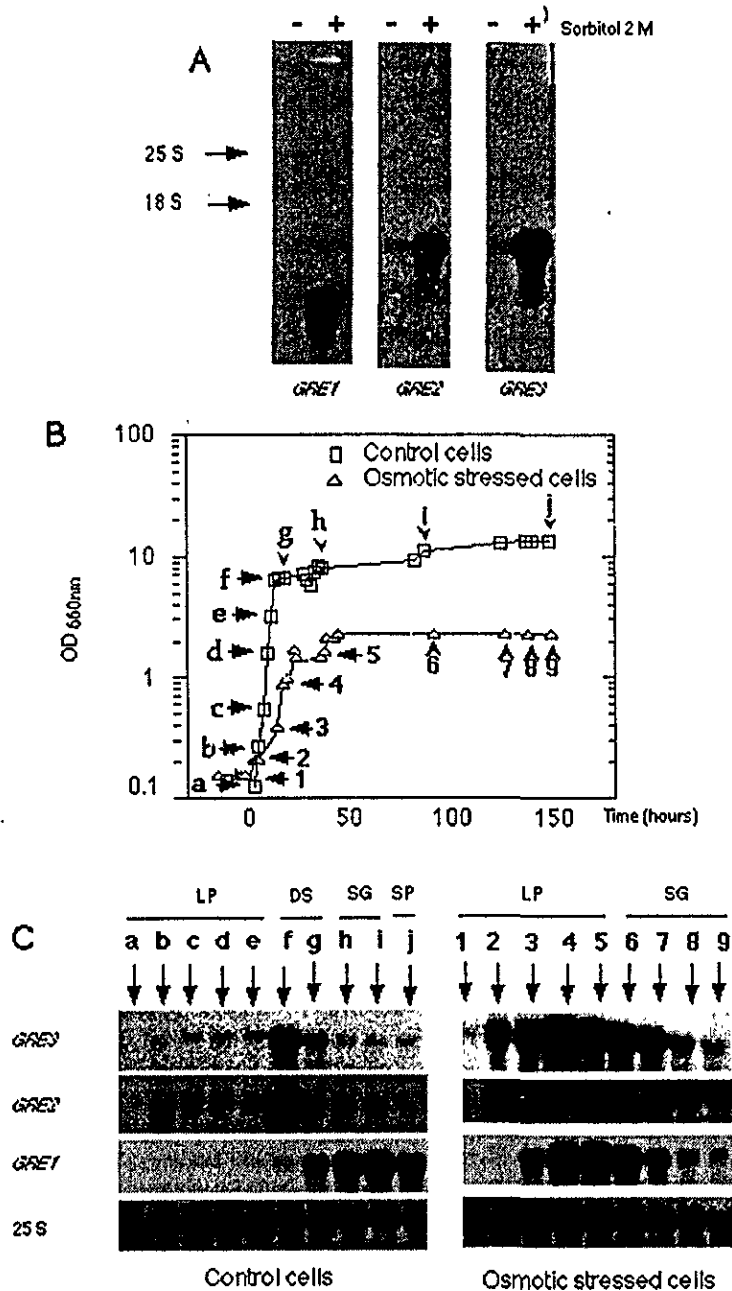


Figure 2. *GRE* transcript accumulation patterns. (A) Accumulation of the *GRE* transcripts under osmotic stress. Total RNA (10 µg/lane) obtained from yeast cells (RS58) that had been grown to an  $OD_{660}$  0.5 on YPD (-) or YPD plus 2 M sorbitol (+) was loaded on a 1.2% agarose-formaldehyde gel. After electrophoresis, bands were blotted on nylon membranes and hybridized with [ $\alpha$ - $^{32}$ P] dCTP-labelled full-length cDNAs, as indicated in Material and Methods. The position of the ribosomal RNAs (25S and 18S) is shown at the left of the figure. (B) Growth curves of yeast cultures grown in YPD medium (control cells) and YPD+2 M sorbitol (osmotic stressed cells). Aliquots were removed to determine the cell density at the times indicated by arrows. Each number and letter indicates an RNA extraction point and the corresponding lanes in the following panel. (C) Effect of the growth phase on the accumulation of the *GRE* transcripts. Cells were grown as explained in panel (B) and harvested, at the times indicated by the arrows, for RNA extraction. Total RNA was purified and loaded (10 µg/lane) on a 1.2% agarose-formaldehyde gel. After electrophoresis, bands were blotted on nylon membranes and hybridized with [ $\alpha$ - $^{32}$ P] dCTP-labelled full-length cDNAs. We used the 25S ribosomal RNA as a reference of the amount of RNA loaded: LP, logarithmic phase; DS, diauxic shift; SG, slow growth; SP, stationary phase.



GENES INDUCED BY STRESS IN YEAST

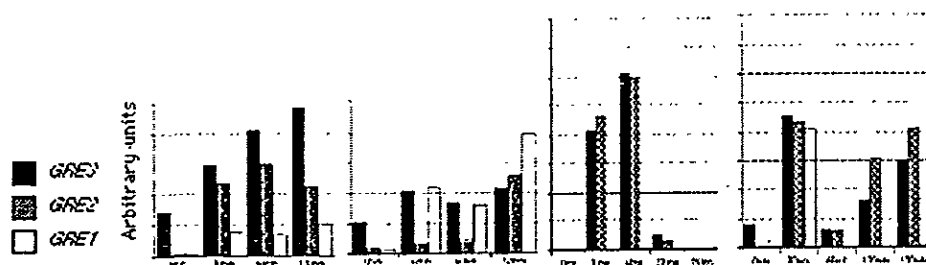
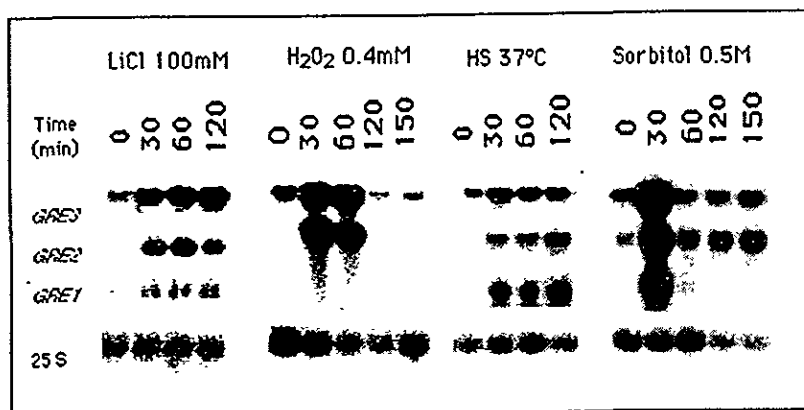


Figure 3. Time course of the *GRE* transcript accumulation under heat shock, ionic, oxidative and osmotic stress. Cells were harvested at the times indicated and total RNA was extracted from yeast cultures grown to an  $OD_{660}$  of 0.5. At this point, the different stresses were imposed by adding sorbitol, LiCl or  $H_2O_2$  in solution to final concentrations of 0.5 M, 100 mM and 0.4 mM, respectively (see Materials and Methods). For the heat shock experiment, the cultures were shifted from 30°C to 37°C for the indicated periods. RNAs were loaded (10  $\mu$ signg/lane), blotted and hybridized with [ $\alpha$ - $^{32}P$ ] dCTP-labelled full-length cDNAs. The 25S ribosomal RNA was used as loading reference. The mRNA levels were quantified using the NIH Image 1.60b7. The resulting values were normalized with those obtained from the 25S rRNA hybridization signal, as indicated in Materials and Methods.

effect is more evident during growth in YPD without sorbitol. In this case, *GRE2* and *GRE3* accumulated to high levels during the diauxic shift, whilst *GRE1* transcript was expressed during the postdiauxic and stationary phases.

*GRE transcript accumulation during osmotic, ionic, oxidative and heat shock conditions*

The fact that *GRE2* and *GRE3* gene expression is induced in the diauxic shift phase, when the yeast switch to respiratory metabolism and changes in the redox status occur (Werner-Washburne *et al.*, 1993), suggested that their products could be involved in other stress responses. The same was implied for the *GRE1* gene, whose expression is induced in control cells after the postdiauxic phase and during the stationary phase, where other stresses (e.g. nutrient starvation) could be sensed. The responsiveness of the *GRE* genes

to other stresses was tested by looking at their transcript accumulation in Northern hybridization experiments, using total RNA obtained from cells stressed by osmotic, ionic, oxidative and heat shock. In these experiments, yeast cultures grown in YPD to an  $OD_{660}$  of 0.5 were subjected to stressful environments for different times.

Ionic shock was imposed by adding LiCl to a final concentration of 100 mM. This concentration was selected in order to preferentially induce ionic toxicity rather than osmotic stress (Gläser *et al.*, 1993). For the three transcripts, accumulation started to be detected at 30 min after LiCl addition. For *GRE1* the accumulation remained at the same level during the period of the stress treatment; for *GRE2* there was a further increase at 60 min and then the transcript accumulation started to decrease; and for *GRE3*, there was an increase in the accumulation of the transcript

throughout the stress treatment (Figure 3). As in the case of ionic shock, accumulation of all GRE transcripts was detected after 30 min of heat shock. For *GRE1* and *GRE2*, the highest transcript accumulation occurs after 120 min, whereas the *GRE3* transcript remained at the same level at all times tested (Figure 3).

A different expression pattern was observed for the oxidative and osmotic shock treatments. When a final concentration of 0.4 mM H<sub>2</sub>O<sub>2</sub> was added to the cultures, the *GRE2* and *GRE3* mRNAs showed their highest accumulation at 60 min of treatment. After 120 min, the accumulation of both transcripts decline (Figure 3). In contrast, *GRE1* mRNA was not detected under H<sub>2</sub>O<sub>2</sub> stress. In the case of osmotic shock, imposed by treatment with 0.5 M sorbitol (low osmolarity), all GRE transcripts showed a peak of induction at 30 min; afterwards, at 60 min, a severe decline in the level of the transcripts occurs (Figure 3). Interestingly, we have observed that the induction time of the three GRE transcripts depends on the concentration of the osmolytes in the medium; i.e. low concentrations (0.5 M sorbitol or 0.3 M NaCl) induce a faster accumulation, whereas high concentrations (1.5 M sorbitol or 1 M NaCl) result in a slower induction (see Figure 4A, B; A. Garay-Arroyo, unpublished results).

#### *Regulation of the GRE genes by the HOG and the cAMP-protein kinase A pathways*

The HOG pathway is the best characterized regulatory system for osmotic responsive genes in *S. cerevisiae*; mutations in this MAP kinase pathway render cells osmosensitive. To investigate whether the GRE genes were under the control of the HOG pathway, a strain containing a *HOG1* disruption was constructed as indicated in Materials and Methods (Table 1). The analysis was carried out by looking at the GRE transcripts accumulation in wild-type and *hog1* strains subjected to osmotic stress. As shown in Figure 4A, none of the GRE genes was induced in the *hog1* strain in response to osmotic stress, indicating that *HOG1* positively regulates their expression. This result is in agreement with the finding of one STRE in the *GRE3* promoter region and of related sequences in the other GRE promoters (Figure 5). These homologous sequences correspond to the so-called 'PDS element' known to be involved in the regulation of genes expressed during the diauxic or stationary growth phases (Figures 2, 5).

Additional regulatory pathways have been reported to regulate the osmotic response in yeast through the STRE. The cAMP-protein kinase A pathway has been shown to negatively control the expression of osmotic-responsive genes. A mutation in the gene encoding the regulatory subunit of protein kinase A (*BCY1*) results in a high cAMP-independent protein kinase A activity and, as a consequence, in repression of several genes induced by osmotic stress (reviewed in Ruis and Schüller, 1995). To look at the effect of this mutation on the expression of the GRE genes, a *bcy1* strain was constructed using an *URA3* disrupted *BCY1* gene (Table 1). Both, the wild-type and the *bcy1* strains were subjected to osmotic stress and RNAs were isolated at different times, as indicated in Figure 4B. The data show that *GRE1* and *GRE3* gene expression is modulated by cAMP levels. In contrast, *GRE2* expression was unaffected when compared with its expression in the wild-type strain (Figure 4B). Again, these data are consistent with the presence of various STRE and PDS elements in the *GRE1* and *GRE3* promoter region (Figure 5). The lack of effect in the case of *GRE2* expression could be explained by the presence of a unique PDS element or the absence of a typical STRE (Figure 5).

Recent reports by Martínez-Pastor *et al.* (1996) and Schmitt and McEntee (1996), showed that the MSN2p-MSN4p transcriptional factors are able to bind *in vitro* to the STRE. Additionally, it is known that the double mutant *msn2-*msn4** is sensitive to osmotic, ionic and oxidative stress conditions. To see whether these transcriptional factors were able to regulate the induction of the GRE genes, we carried out a Northern analysis of wild-type and mutant strains under stressed and non-stressed conditions (Figure 4C). The results suggest that *GRE1* and *GRE3* expression is regulated by these transcription factors; however, as in the case of the *bcy1* strain, the *GRE2* transcript levels were unaffected. The presence of a typical STRE in the case of the *GRE3* promoter region correlates with the transcript accumulation in strain *msn2-*msn4**. This is not the case for the *GRE1* gene, where no consensus STRE was found; instead, six PDS elements were detected. A possible explanation for the *GRE1* responsiveness to the absence of Msn2p-Msn4p transcriptional factors could be related to the high homology between the STRE and PDS element, as well as to the presence of more than one PDS sequence. For *GRE2*, the absence of a STRE and the presence of

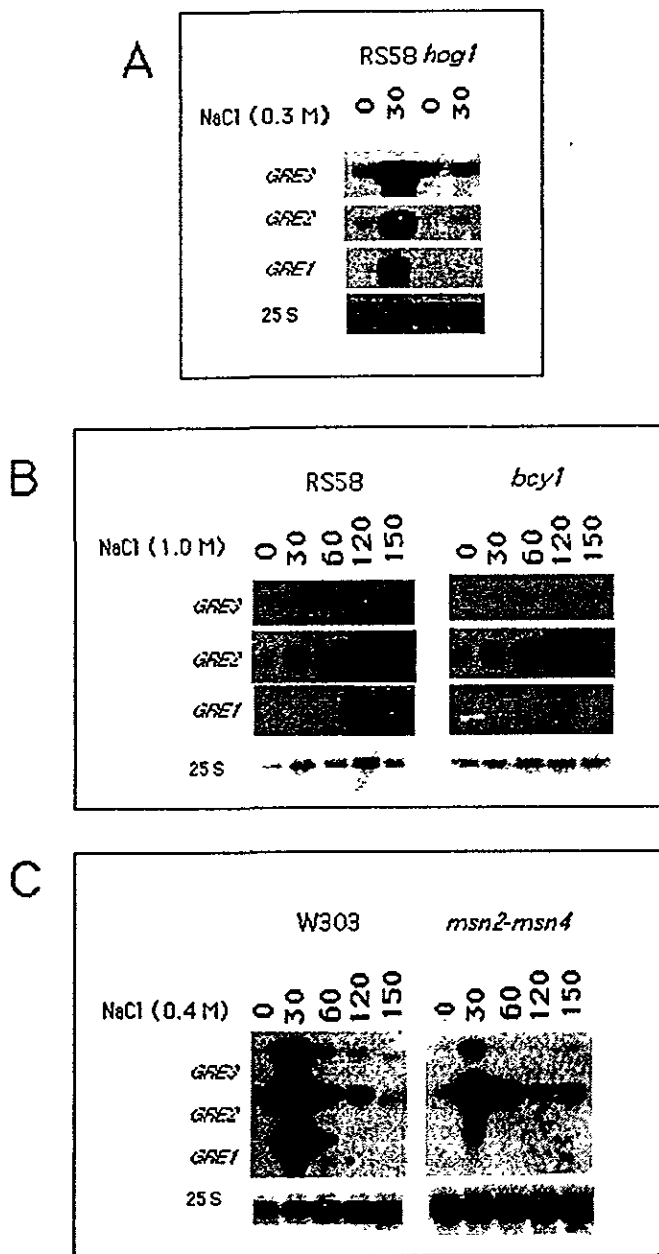


Figure 4. Effect of mutations in *HOG1*, *BCY1* and *MSN2-MSN4* genes on the accumulation of the *GRE* transcripts under osmotic stress. (A) Wild-type and *hog1* strains were subjected to 0.3 M NaCl treatment for 30 min. (B) Total RNA was prepared from wild-type (RS58) and *bcy1* strains after incubation during the indicated periods on YPD (time 0) and YPD + 1.0 M NaCl (30, 60, 120, 150 min). (C) Northern analysis of total RNA prepared from wild-type (RS58) and *msn2-msn4* strains after incubation during the indicated periods on YPD (time 0) and YPD with 0.4 M NaCl (30, 60, 120, 150 min). Northern experiments were performed with 10 µg of total RNA obtained from yeast cultures grown to an OD<sub>660</sub> of 0.5. [ $\alpha$ -<sup>32</sup>P] dCTP-labelled full-length cDNAs were used as hybridization probes in all Northern experiments. 25S ribosomal RNA was used as a loading reference.

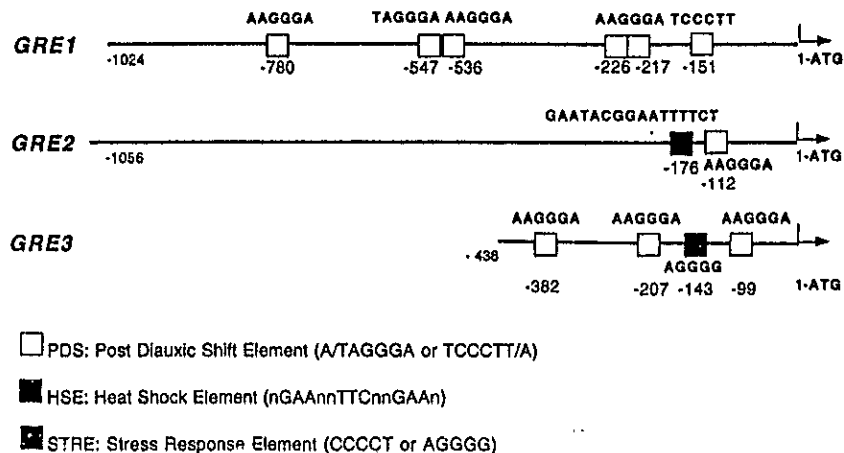


Figure 5. Putative control elements contained in the promoter regions of the *GRE* genes. The complete sequence of the *GRE* genes promoter regions were obtained from the Stanford University yeast database. The complete intergenic regions were analysed for the presence of the different sequences, using FIND PATTERNS from GCG software package. The sequence of the regulatory elements identified are shown, as well as their precise locations.

a unique PDS sequence could be a possible explanation for its lack of response in a *msn2-*msn4** background.

#### Disruption of the *GRE* genes

To address the role of the products of the *GRE* genes in environmental stress responses, we constructed strains carrying disrupted *GRE* genes. *GRE1*, *GRE2* and *GRE3* were disrupted by inserting the *URA3* gene into their coding regions. Each disruption was introduced into the RS58 genome by recombination and selected by complementation of uracil auxotrophy. The positive candidates were analysed by Southern blot to confirm the presence of the disrupted *GRE* genes (data not shown). Phenotypic analyses were done by growing yeast cells in YPD solid medium containing sorbitol (1.3, 1.5 and 2 M), NaCl (1.0, 1.3 and 1.5 M), LiCl (0.1 and 0.2 M), or H<sub>2</sub>O<sub>2</sub> (5 and 10 mM). The results indicate that the *GRE* genes are dispensable for growth, at least under the osmotic, ionic and oxidative stress conditions tested (data not shown).

#### DISCUSSION

In *S. cerevisiae*, a general stress response system induced by various environmental stimuli has been shown to be responsible for the protection of yeast cells against a number of conditions that endanger their survival (Mager and Varela, 1993; Parsell and

Lindquist, 1994; Trollmo *et al.*, 1988; Varela *et al.*, 1992). Specific responses can also be detected for each stress situation. In the case of an osmotic stress, an osmosensitive phenotype has been shown in mutants defective in very diverse functions, such as vacuole sorting and biogenesis (Latterich and Watson, 1991), cytoskeleton structure (Chowdhury *et al.*, 1992), osmolyte production (Albertyn *et al.*, 1994; Larsson *et al.*, 1993), etc. These data suggest that a complex response is needed to adapt to media with a low water activity.

To contribute to the understanding of the molecular events leading to an osmoregulatory response, we have isolated and characterized three genes that are induced in response to osmotic stress conditions. The analysis of their amino acid sequences revealed the involvement, in this response, of an unidentified hydrophilic protein (*GRE1*), a putative dihydroflavonol reductase (*GRE2*), and a putative xylose reductase (*GRE3*). The analysis of the *GRE1* amino acid composition suggested a functional relationship with the so-called LEA proteins from plants. As in the case of LEA proteins, *GRE1* is highly hydrophilic and presents a high content of charged amino acids and glycine, which allows us to predict that it is a non-globular (random-coiled) and highly soluble protein (Dure, 1993; Roberts *et al.*, 1993; Figure 1). Additionally, LEA proteins and *GRE1* are induced by water-deficit environments, either externally imposed or during development

(Figures 2, 3; Dure, 1993; Ingram and Bartels, 1996). Interestingly, by searching the yeast genome databank, we have found additional analogous proteins (Garay-Arroyo *et al.*, in preparation). The finding of analogous proteins in yeasts and plants whose transcripts are induced under osmotic stress raises the possibility of a wider occurrence in nature, as well as a general role during this response. The lack of phenotype in the *GRE1* gene disruption correlates with the idea that these analogous proteins carry out similar functions during the yeast response to stress.

*GRE2* encodes a protein similar to dihydroflavonol-4-reductases (DFR) from plants. It has been shown that in higher plants, these enzymes are involved in the synthesis of secondary compounds, anthocyanins and flavonoids, using as co-factor NADP<sup>+</sup> (Kristiansen and Rohde, 1991). In addition, its induction by water stress has been reported in *Vigna unguiculata* (Iuchi *et al.*, 1996). In *S. cerevisiae*, it has been suggested that these enzymes may be involved in the synthesis of certain co-factors (Mewes *et al.*, 1997). The fact that the expression of this gene is positively affected by oxidative stress and during the diauxic shift, where a metabolic change from fermentation to respiration occurs (Werner-Washburne *et al.*, 1993), suggests that its function might be related to the balance of the NADP<sup>+</sup>/NADPH pools, which could also be adversely modified under hyperosmotic conditions. Since the *S. cerevisiae* genome contains three additional *GRE2* homologous genes, the absence of phenotype of the *GRE2* disruption under the conditions tested is not surprising.

The third gene described, *GRE3*, encodes a protein homologous to fungal xylose reductases (XR). This enzyme is a member of the family of aldo-keto reductases that uses NADPH as co-factor and catalyses the reduction of carbonyl metabolites (sugars) to produce polyols (Mahler and Cordes, 1971). In this particular case, xylitol is the product of the reaction. In different organisms, polyols have been shown to be important as osmoprotectants and during osmotic adjustment since they contribute to maintain an adequate osmotic potential without interfering with macromolecules (Yancey *et al.*, 1982). In *S. cerevisiae*, glycerol is the major osmolyte synthesized during osmotic stress conditions (Blomberg and Adler, 1989) but additional polyols have been reported for other yeasts (Brown and Edgley, 1980). In the yeast *Pichia stipitis*, xylitol produced by XR is the

substrate of a xylose dehydrogenase (XD), which converts it into xylulose that enters into the pentose phosphate pathway (Verduyn *et al.*, 1985). In *S. cerevisiae*, D-xylose is the substrate of XR and is not a good carbon source unless is provided with additional sugars, ethanol or glycerol. Under these conditions, 95% of the xylose used is converted to xylitol, which accumulates, suggesting that the limiting step for the xylulose production is the xylose dehydrogenase-catalysed reaction (Van Zyl *et al.*, 1989). Since the disruption of *GRE3*, the unique XR gene in *S. cerevisiae*, does not present a detectable phenotype under stress conditions, it is difficult to define its precise function. However, it cannot be discounted that xylitol could be acting as an osmolyte in a particular growth condition, and/or that the enzyme activity participates as a modulator of the NADP<sup>+</sup>/NADPH pools during stress (Hahn-Hagerdal *et al.*, 1996).

The results reported here show that the three *GRE* genes are induced in response not only to hyperosmotic conditions but also to other stress situations (Figure 3). This suggests that they may be under the control of a more general stress response network. Additionally, their induction, at the growth late stages, suggests that there is a cross-talk between the regulatory pathways controlling growth and those regulating the stress response. A similar pattern of expression has been observed for other stress genes (Schüller *et al.*, 1994). Analysis of the *GRE* gene promoter regions, together with their expression patterns in wild-type and mutant strains, reveals that they are under the control of the HOG pathway, as are many other genes responsive to osmotic stress. However, only one of the three *GRE* genes, *GRE3*, has an orthodox structure of the STRE in its promoter. In the case of *GRE1* and *GRE2*, we did not detect a consensus core STRE sequence; instead, we found six PDS elements in *GRE1* and one in *GRE2* promoter regions. This observation allowed us to hypothesize that the HOG pathway is modulating the expression of these genes, either through a different sequence, as it occurs with other genes (Miralles and Serrano, 1995) or through the PDS elements. This last possibility is supported by the fact that PDS elements are able to compete with STRE sequences *in vitro* (Marchler *et al.*, 1993). According to the role of the Msn2–Msn4 transcriptional factors as mediators of STRE-dependent regulation, we have found that *GRE3* is positively regulated by these factors, because its transcript level decreases in the *msn2–msn4* mu-

tant. The same effect was observed for *GRE1*, which does not contain typical STRE sequences. This observation supports the idea that PDS elements present in the *GRE1* promoter can be recognized as STRE, maybe functioning in a co-operative mode. Apparently, *GRE2* did not respond to Msn2p-Msn4p absence, which could be explained by the lack of STRE sequences and the presence of only one PDS element. Experiments using promoter-*lacZ* fusions and variants of these elements are in progress in order to prove the functionality of these putative promoter elements.

The observation that the *GRE* genes are induced not only under osmotic stress but also during the late growth stages suggested the participation of cAMP-PKA pathway in the regulation of these genes. The *GRE1* and *GRE3* transcript accumulation patterns obtained using a *bcy1* strain indicate that PKA negatively regulates their expression (Figure 4B). The participation of the PDS element in regulation by cAMP-PKA has been shown using PDS-*lacZ* fusions (Boorstein and Craig, 1990a). The presence of PDS elements in the promoter region of both genes strongly suggests their participation in the observed regulation by cAMP-PKA. Furthermore, these results correlate with the PDS role during late growth stages and support the idea of considering cAMP as a major component of the intracellular signal of sub-optimal nutrient conditions (Boorstein and Craig, 1990a). However, additional sequences might be required for the response to cAMP, since the *GRE2* gene is able to respond to the diauxic shift (Figure 2B) but not to cAMP-PKA regulation (Figure 4B), even though its promoter region contains one PDS element. A more extensive analysis using different mutants affected in cAMP levels is needed to validate these hypotheses. The function of the *GRE* genes during the yeast stress response could not be defined, given the lack of phenotype in the disruptions that were analysed. One explanation for this result is the functional redundancy found for the three GRE proteins, which complicates analysis regarding their participation in adaptation to stress conditions.

#### ACKNOWLEDGEMENTS

We thank M. Werner-Washburne, M. Rocha, J. L. Reyes, J. Folch and W. Hansberg for critical reading of the manuscript. We also thank R.M. Solórzano for technical assistance, P. Gaytán and

E. López for the synthesis of the oligonucleotides used in this study, and J. Folch for advice on genetic analysis. This work was supported by grants from Consejo Nacional de Ciencia y Tecnología (México) 0131P-N9506 and 26242-N, and from European Communities CII\*CT94-0082.

#### REFERENCES

- ✓ Albertyn, J., Hohmann, S., Thevelein, J. M. and Prior, B. A. (1994). *GPD1*, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *Saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathway. *Mol. Cell. Biol.* **14**, 4135-4144.
- ✓ Ansell, R., Granath, K., Hohmann, S., Thevelein, J. M. and Adler, L. (1997). The two isozymes for yeast NAD-dependent glycerol 3-phosphate dehydrogenase encoded by *GPD1* and *GPD2* have distinct roles in osmoadaptation and redox regulation. *EMBO J.* **16**, 2179-2187.
- ✓ Bartels, D., Engelhardt, K., Roncarati, R., Schneider, K., Rotter, M. and Salamini, F. (1991). An ABA and GA modulated gene expressed in the barley embryo encodes an aldose reductase related protein. *EMBO J.* **10**, 1037-1043.
- ✓ Belazzi, T., Wagner, A., Wieser, R., Schanz, M., Adam, G., Hartig, A. and Ruis, H. (1991). Negative regulation of transcription of the *Saccharomyces cerevisiae* catalase T (*CTT1*) gene by cAMP is mediated by a positive control element. *EMBO J.* **10**, 585-592.
- ✓ Billard, P., Menart, S., Fleer, R. and Bolotin-Fukuhara, M. (1995). Isolation and characterization of the gene encoding xylose reductase from *Kluyveromyces lactis*. *Gene* **162**, 93-97.
- ✓ Blomberg, A. and Adler, L. (1989). Roles of glycerol and glycerol-3-phosphate dehydrogenase (NAD<sup>+</sup>) in acquired osmotolerance of *Saccharomyces cerevisiae*. *J. Bacteriol.* **171**, 1087-1092.
- ✓ Boorstein, W. R. and Craig, E. A. (1990a). Regulation of a yeast *HSP70* gene by a cAMP responsive transcriptional control element. *EMBO J.* **9**, 2543-2553.
- ✓ Boorstein, W. R. and Craig, E. A. (1990b). Transcriptional regulation of *SSA3*, an *HSP70* gene from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10**, 3262-3267.
- ✓ Brewster, J. L., de Valoir, T., Dwyer, N. D., Winter, E. and Gustin, M. C. (1993). An osmosensing signal transduction pathway in yeast. *Science* **259**, 1760-1763.
- ✓ Brown, A. D. and Edgley, M. (1980). Osmoregulation in yeast. In Rains, W. D., Valentine, R. C. and Hollaender, A. (Eds), *Genetic Engineering of Osmoregulation*. Plenum, Washington, D.C., pp. 75-90.
- ✓ Casamayor, A., Aldea, M., Casas, C., Herrero, E., Gamo, F. J., Lafuente, M. J., Gancedo, C. and Arifio,

- J. (1995). DNA sequence analysis of a 13 kbp fragment of the left arm of the yeast chromosome XV containing seven new open reading frames. *Yeast* 11, 1281-1288.
- Collart, M. and Oliviero, S. (1995). Preparation of yeast RNA. In Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (Eds), *Current Protocols in Molecular Biology*. Wiley, New York, pp. 13.12.1-13.12.4.
- Chester, V. (1968). Heritable glycogen-storage deficiency in yeast and its induction by ultra-violet light. *J. Gen. Microbiol.* 51, 49-56.
- Chowdhury, S., Smith, K. W. and Gustin, M. C. (1992). Osmotic stress and the yeast cytoskeleton: phenotype-specific suppression of an actin mutation. *J. Cell Biol.* 118, 561-571.
- Dure, L. III. (1993). Structural motifs in LEA proteins. In Close, T. J. and Bray, E. A. (Eds), *Plant Responses to Cellular Dehydration during Environmental Stress*. American Society of Plant Physiologist, Riverside, pp. 91-103.
- Gläser, H. U., Thomas, D., Gaxiola, R., Montrichard, F., Surdin-Kerjan, Y. and Serrano, R. (1993). Salt tolerance and methionine biosynthesis in *Saccharomyces cerevisiae* involve a putative phosphatase gene. *EMBO J.* 12, 3105-3110.
- Gounalaki, N. and Thireos, G. (1994). Yap1, a yeast transcriptional activator that mediates multidrug resistance, regulates the metabolic stress response. *EMBO J.* 13, 4036-4041.
- Hahn-Hagerdal, B., Hallborn, J., Jeppsson, H., Meinander, N., Walfridsson, M., Ojamo, H., Penttila, M. and Zimmermann, F. K. (1996). Redox balances in recombinant *Saccharomyces cerevisiae*. *Ann. N.Y. Acad. Sci.* 782, 286-296.
- Hoffman, C. (1995). Preparation of yeast DNA, RNA, and proteins. In Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (Eds), *Current Protocols in Molecular Biology*. Wiley, New York, pp. 13.11.1-13.11.3.
- Ingram, J. and Bartels, D. (1996). The molecular basis of dehydration tolerance in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 47, 377-403.
- Ito, H., Fukuda, K. and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* 153, 163-168.
- Iuchi, S., Yamaguchi-Shinozaki, K., Urao, T., Terao, T. and Shinozaki, K. (1996). Novel drought-inducible genes in the highly drought-tolerant cowpea: cloning of cDNAs and analysis of the expression of the corresponding genes. *Plant Cell Physiol.* 37, 1073-1082.
- Kristiansen, K. N. and Rohde, W. (1991). Structure of the *Hordeum vulgare* gene encoding dihydroflavonol-4-reductase and molecular analysis of *ant18* mutants blocked in flavonoid synthesis. *Mol. Gen. Genet.* 230, 49-59.
- Larsson, K., Ansell, R., Eriksson, P. and Adler, L. (1993). A gene encoding sn-glycerol-3-phosphate dehydrogenase (NAD<sup>+</sup>) complements an osmosensitive mutant of *Saccharomyces cerevisiae*. *Mol. Microbiol.* 10, 1101-1111.
- Latterich, M. and Watson, M. D. (1991). Isolation and characterization of osmosensitive vacuolar mutants of *Saccharomyces cerevisiae*. *Mol. Microbiol.* 5, 2417-2426.
- Lauermann, V. (1991). Ethanol improves the transformation efficiency of intact yeast cells. *Curr. Genet.* 20, 1-3.
- Liang, P. and Pardee, A. B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257, 967-971.
- Liang, P. and Pardee, A. B. (1995). Differential display of mRNA by PCR. In Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (Eds), *Current Protocols in Molecular Biology*. Wiley, New York, pp. 15.8.1-15.8.8.
- Maeda, T., Takekawa, M. and Saito, H. (1995). Activation of yeast PBS2 MAPKK by MAPKKs or by binding of an SH3-containing osmosensor. *Science* 269, 554-558.
- Maeda, T., Wurgler-Murphy, S. M. and Saito, H. (1994). A two-component system that regulates an osmosensing MAP kinase cascade in yeast. *Nature* 369, 242-245.
- Mager, W. H. and Ferreira, M. (1993). Stress response of yeast. *Biochem. J.* 290, 1-13.
- Mager, W. H. and Varela, J. C. S. (1993). Osmostress response of the yeast *Saccharomyces*. *Mol. Microbiol.* 10, 253-258.
- Mahler, H. R. and Cordes, E. H. (1971). *Biological Chemistry*. Harper & Row, New York.
- Marchler, G., Schüller, C., Adam, G. and Ruis, H. (1993). A *Saccharomyces cerevisiae* UAS element controlled by protein kinase A activates transcription in response to a variety of stress conditions. *EMBO J.* 12, 1997-2003.
- Martínez-Pastor, M. T., Marchler, G., Schüller, C., Marchler-Bauer, A., Ruis, H. and Estruch, F. (1996). The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress-response element (STRE). *EMBO J.* 15, 227-2235.
- Mewes, H. W., Albermann, K., Heumann, K., Liebl, S. and Pfeiffer, F. (1997). MIPS: a database for protein sequences, homology data and yeast genome information. *Nucleic Acids Res.* 25, 28-30.
- Miralles, V. J. and Serrano, R. (1995). A genomic locus in *Saccharomyces cerevisiae* with four genes up-regulated by osmotic stress. *Mol. Microbiol.* 17, 653-662.
- Parsell, D. A. and Lindquist, S. (1994). Heat shock proteins and stress tolerance. In Morimoto, R. I., Tissières, A. and Georgopoulos, C. (Eds), *The Biology*

- of Heat Shock Proteins and Molecular Chaperones*. Cold Spring Harbor Laboratory Press, New York, pp. 457-494.
- ✓Posas, F. and Saito, H. (1997). Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: scaffold role of Pbs2 MAPKK. *Science* **276**, 1702-1705.
- ✓Posas, F., Wurgler-Murphy, S. M., Maeda, T., Witten, E. A., Cam Thai, T. and Saito, H. (1996). Yeast HOG1 MAP kinase cascade is regulated by a multi-step phosphorelay mechanism in the *SLN1-YPD1-SSK1* 'two component' osmosensor. *Cell* **86**, 865-875.
- ✓Roberts, J. K., DeSimone, N. A., Lingle, W. L. and Dure, L. III. (1993). Cellular concentrations and uniformity of cell-type accumulation of two Lea proteins in cotton embryos. *Plant Cell* **5**, 769-780.
- ✓Ruis, H. and Schüller, C. (1995). Stress signaling in yeast. *BioEssays* **17**, 959-965.
- ✓Schmitt, A. P. and McEntee, K. (1996). Msn2, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **93**, 5777-5782.
- ✓Schüller, C., Brewster, J. L., Alexander, M. R., Gustin, M. C. and Ruis, H. (1994). The HOG pathway controls osmotic regulation of transcription via the stress response element (STRE) of the *Saccharomyces cerevisiae* CTT1 gene. *EMBO J.* **13**, 4382-4389.
- ✓Serrano, R. (1996). Salt tolerance in plants and microorganisms: toxicity targets and defense responses. *Int. Rev. Cytol.* **165**, 1-52.
- ✓Sparvoli, F., Martin, C., Scienza, A., Gavazzi, G. and Tonelli, C. (1994). Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). *Plant Mol. Biol.* **24**, 743-755.
- ✓Takuma, S., Nakashima, N., Tantirungkij, M., Kinoshita, S., Okada, H., Seki, T. and Yoshida, T. (1991). Isolation of xylose reductase gene of *Pichia stipitis* and its expression in *Saccharomyces cerevisiae*. *Appl. Biochem. Biotechnol.* **28-29**, 327-340.
- ✓Trollmo, C., André, L., Blomberg, A. and Adler, L. (1988). Physiological overlap between osmotolerance and thermotolerance in *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* **56**, 321-326.
- ✓Van Zyl, C., Prior, B. A., Kilian, S. G. and Kock, J. L. (1989). D-xylose utilization by *Saccharomyces cerevisiae*. *J. Gen. Microbiol.* **135**, 2791-2798.
- ✓Varela, J. C. S., Van Beekvelt, C. A., Planta, R. J. and Mager, W. H. (1992). Osmostress-induced changes in yeast gene expression. *Mol. Microbiol.* **6**, 2183-2190.
- ✓Verduyn, C., Van Kleef, R., Frank, J., Schreuder, H., Van Dijken, J. P. and Scheffers, A. (1985). Properties of the NAD(P)H-dependent xylose reductase from the xylose-fermenting yeast *Pichia stipitis*. *Biochem. J.* **226**, 669-677.
- ✓Werner-Washburne, M., Braun, E., Johnston, G. C. and Singer, R. A. (1993). Stationary phase in the yeast *Saccharomyces cerevisiae*. *Microbiol. Rev.* **57**, 383-401.
- ✓Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. and Somero, G. N. (1982). Living with water stress: evolution of osmolyte systems. *Science* **217**, 1214-1222.



## Resultados anexos

Prácticamente todos los resultados de la caracterización de los tres genes *GRE* se encuentran descritos en el artículo de la revista *Yeast* que se anexa. Sin embargo, existen algunos datos adicionales que arrojaron más información sobre estos genes y que no se muestran en el artículo.

### Hibridación tipo Southern.

Como parte de la caracterización funcional de los genes *GRE* era necesario hacer una mutación dirigida, por medio de una interrupción génica, de cada uno ellos en el genoma de *S. cerevisiae* para probar si la cepa resultante era sensible a diferentes condiciones de estrés. Para ello fue necesario conocer el número de copias de cada uno de los tres genes en el genoma de la cepa RS58 (ver Tabla 1; Garay-Arroyo y Covarrubias, 1999); por lo que se realizó un experimento tipo Southern genómico con DNA total de *S. cerevisiae* (ver Material y Métodos en Garay-Arroyo y Covarrubias, 1999.) digerido con diferentes enzimas: *EcoR1*, *EcoRV*, *HindIII* y *HincII*.

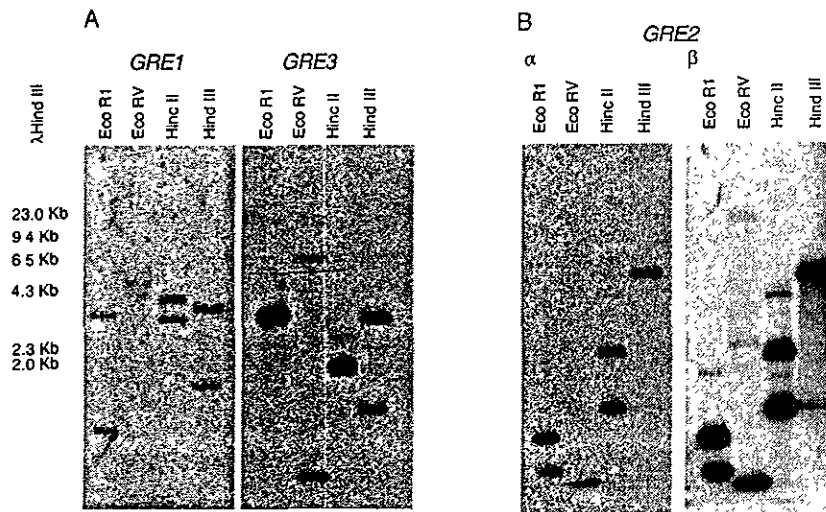


Fig. 1. Southern genómico. Se cargaron 20 $\mu$ g de DNA total de *S. cerevisiae* digeridos con las enzimas *EcoR1*, *EcoRV*, *HincII* y *HindIII*. Se usó el fragmento correspondiente al gen completo de cada uno de los genes *GRE* como sonda. La hibridación y los lavados se hicieron a alta astringencia. Los pesos moleculares están indicados a la izquierda de la figura.

Como se puede ver en la Fig. 1A, se encontró únicamente una copia de *GRE1*; (este gen presenta un sitio interno para las enzimas *EcoR1*, *HincII* y *HindIII*) y una copia de *GRE3* (que también presenta un sitio interno para las enzimas *HincII* y *HindIII* y dos sitios internos para *EcoRV*). En el caso de *GRE2* fué posible detectar más de una copia del gen en el genoma de la cepa RS58 cuando se sobreexpuso la membrana de hibridación (comparar los patrones  $\alpha$  y  $\beta$  de la Fig. 1B). También para este gen se sabe, por los patrones de restricción, que tiene un sitio interno para la enzima *EcoRI*, y tres sitios de restricción para las enzimas *EcoRV* y *HincII*; sin embargo, para éstas últimas enzimas no se detectan tres bandas ya que uno de los fragmentos resultantes es menor de 100pb (ver Fig. 1B $\beta$ ). Lo anterior concuerda con lo reportado para los tres genes *GRE* en la base de datos del genoma de levadura que se encuentra en la Universidad de Stanford .

### **Cinética de expresión de los genes *GRE* bajo diferentes concentraciones de sorbitol.**

En *S. cerevisiae* se han definido condiciones de alta y baja osmolaridad para los cuales existen sensores y vías de trasducción de señales diferentes (Márquez y Serrano, 1996). Con la finalidad de definir el comportamiento de los genes *GRE* en respuesta a condiciones de alta y baja osmolaridad, se realizó un experimento para ver la cinética temporal de inducción bajo diferentes concentraciones de sorbitol. Como se puede ver en la Fig. 2, a medida que se aumenta la concentración del osmolito, se incrementa el tiempo de inducción de los tres genes *GRE*. De hecho, se puede ver el mismo tipo de comportamiento cuando se trata de los genes *GPD1* y *CTT1*; (ver Fig. 2). Es interesante que los cinco genes reportados presenten un pico de inducción en los mismos tiempos para todas las concentraciones probadas. Una posible explicación de este fenómeno es que a bajos potenciales osmóticos externos (0.5 M sorbitol) la levadura se adapta rápidamente al medio y, por esa razón no requiere de mantener la transcripción de los diferentes genes por mucho tiempo; de hecho, se ha visto que a estas mismas concentraciones osmóticas, la fosforilación de HOG1 tiene un pico a los 15-20 minutos de aplicado el estrés y luego regresa a su fosforilación basal (Brewster et al., 1993). La regulación temporal de la expresión de los genes *GRE*, *CTT1* y *GPD1* en condiciones de baja osmolaridad depende de la

expresión de HOG1 (Fig. 4A, Garay-Arroyo y Covarrubias, 1999; Hirayama et al., 1995; Shüller et al., 1994); esta expresión se puede correlacionar con el tiempo de fosforilación máxima que presenta esta cinasa.

Por otro lado, conforme se aumenta la osmolaridad y el estrés resulta más severo pareciera que la levadura se tarda en inducir los diferentes genes de respuesta ya que la fase de adaptación es más larga. De cualquier manera, a 1.0 M de sorbitol la respuesta sigue siendo transitoria en los tiempos en los que se midió, ya que después de los 60 min. baja la transcripción de los cinco genes (ver Fig. 2). En 1.5 M de sorbitol, el tiempo de inducción de los genes es aún mayor y, no sabemos si la respuesta es transitoria o no porque no se hicieron mediciones a tiempos más largos; sin embargo, suponemos que debería bajar la expresión de los diferentes transcritos dado el comportamiento que muestran los genes a 0.5 M y a 1.0 M de sorbitol. Sería interesante saber si las proteínas GRE presentan el mismo tipo de comportamiento y el estudiar por qué la inducción de éstas es temporal, si la condición de estrés se mantiene.

Es importante notar que los experimentos tipo northern realizados con las diferentes concentraciones de sorbitol se llevaron a cabo con levaduras crecidas hasta fase logarítmica momento en que se aplicó el estrés (experimentos de choque). Este comportamiento no se observa cuando el transcrito es cuantificado a lo largo de la curva de crecimiento en donde la condición de estrés se mantiene constante; en este último caso, los genes se expresan tanto en fase lag como en exponencial y la expresión baja cuando el cultivo llega a fase estacionaria (Fig. 1; Garay-Arroyo y Covarrubias, 1999).

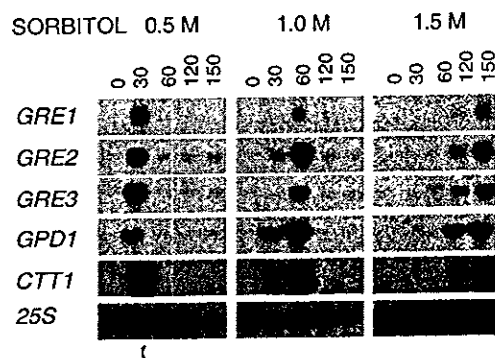


Fig. 2. Expresión de los genes *GRE* bajo diferentes condiciones de estrés osmótico. Los cultivos de *S. cerevisiae* se crecieron hasta una densidad de 0.5 (600nm) y el estrés se impuso con una solución concentrada de sorbitol hasta llegar a la Molaridad que esta indicada en la parte superior de la figura. Las células se cosechan a los tiempos indicados después de aplicado el estrés y se corren 10µg de RNA total. Como sonda se utilizaron los fragmentos completos de los tres genes *GRE* y de *GPD1*, y un fragmento interno *EcoRI* del gen *CTT1*. Como control de carga se utilizó el gen ribosomal 25S.

En cuanto a la regulación de la expresión de los genes en alta osmolaridad (1.0 M y 1.5M de sorbitol) se puede hipotetizar que: 1) la inducción es independiente de HOG1 y, por lo tanto, otra vía o molécula inductora es la responsable y 2) la cinética de fosforilación de HOG1 es diferente en altas concentraciones. Estos datos se podrían corroborar experimentalmente de dos maneras, por un lado midiendo la fosforilación de HOG1 en cultivos sometidos a estrés hiperosmótico con altas concentraciones del osmolito y, por otro lado, viendo la expresión de diferentes genes en el fondo genético *hog1* a estas mismas concentraciones. Está reportado que la expresión del gen *ENA1* está mediada por dos diferentes vías: a bajas concentraciones funciona la vía de HOG y a altas concentraciones funciona la vía de calcineurina (Márquez et al., 1996). Sin embargo, esto no ocurre con *GPD1* en donde se observa que HOG1 controla la expresión a alta y baja osmolaridad pero con diferente cinética; lo anterior quiere decir que, en una mutante *hog1*, se observa el transcrito del gen a diferentes tiempos comparado con la inducción que se observa en la cepa silvestre; lo anterior hablaría de una independencia de HOG1 en la inducción de este gen en esos tiempos (Hohmann, 1997). Hohmann (1977) propone que la vía de HOG podría estar funcionando en las primeras etapas de la respuesta a estrés osmótico y que luego se requeriría de otras vías para completar el proceso. Esto mismo (aunque podría ser por diferentes vías) podría estar ocurriendo con los genes *GRE*.

## Conclusiones y discusión

1.- Se identificaron tres genes (*GRE1*, *GRE2* y *GRE3*), no caracterizados previamente, que se inducen en diferentes condiciones de estrés (osmótico, iónico, oxidativo, calor, fase estacionaria). Los tres genes *GRE* se regulan positivamente por la vía de HOG y por los factores transcripcionales Msn2p y Msn4p y negativamente por AMPc. No se ha podido determinar el papel de los genes *GRE* durante la respuesta a estrés debido a que la interrupción de cada uno de ellos no resulta en un fenotipo de sensibilidad en las diferentes condiciones de estrés probadas.

Los genes *GRE* responden no sólo a condiciones de estrés osmótico sino a una variedad de estreses como son iónico, oxidativo, por calor y por desarrollo; es debido a esta característica que los denominamos *GRE*, por Genes de Respuesta a Estrés.

***GRE1* forma parte de un grupo de proteínas llamadas hidrofílicas que se caracterizan por ser altamente hidrofílicas.**

*GRE1* es un gen que codifica para una proteína que no tiene homología a nivel de secuencia primaria con ninguna proteína hasta ahora reportada en los bancos de datos. Sin embargo, recientemente reportamos una nueva familia de proteínas con características fisicoquímicas similares a las reportadas para las proteínas LEA y a las encontradas en *GRE1*, a las cuales denominamos hidrofílicas. La identificación y caracterización de los genes que codifican para estas proteínas en los genoma de *S. cerevisiae* y de *E. coli* se encuentran en el artículo titulado "Highly hydrophilic proteins are common during conditions of water deficit in prokaryotes and eucaryotes" que se anexa en ésta tesis.

***GRE2* forma parte de la superfamilia (DHE/DFR) representada por las enzimas de mamíferos 3 $\beta$ -deshidrogenasa hidroxisteroide y las de plantas dihidroflavonol-4-reductasa (DFR).**

Por medio de una comparación con los bancos de datos y por datos bibliográficos (Casamayor, et al., 1995) se encontró que las enzimas que más alta homología presentan con *GRE2* son las llamadas DFR de plantas. El porcentaje de identidad encontrado (24-25%) es bajo, por lo que no es muy seguro que estas dos enzimas usen el mismo sustrato ni catalicen una reacción similar. Sin embargo, gracias a la homología entre *GRE2* y las

DFR se encontró que GRE2 forma parte de una superfamilia de proteínas (tanto de animales como de bacterias y plantas) en donde también están incluidas las DFR (Baker et. al., 1990). Para saber las relaciones que guardaban los diferentes miembros de la familia con GRE2 se trazó un árbol filogenético (ver Fig. 1) en donde se puede ver que esta proteína se parece más a sus homólogos de levadura (YGL039w, YGL157w y YDR541c) que a cualquier otra enzima. Desgraciadamente, sólo para GRE2 se tiene algo de información acerca de la expresión y regulación del gen ya que los tres genes restantes sólo han sido identificados como un “marco abierto de lectura” en el proyecto de secuenciación del genoma, por lo que no se sabe si se expresan, ni en que condiciones lo hacen.

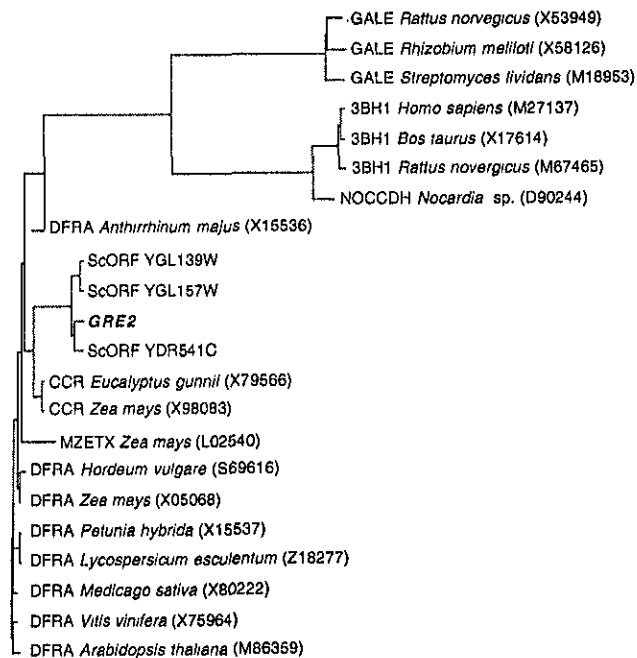


Fig. 3. Árbol filogenético de GRE2 y secuencias relacionadas. El árbol se construyó usando las secuencias depositadas en los bancos de datos y los números de accesos están indicados después de las especies a las cuales pertenece cada proteína. Los aminoácidos se alinearon usando el programa Pileup; y las distancias se obtuvieron por el método de Kimura usando el programa de Distances. Finalmente, el árbol se obtuvo con el programa Growtree. Todos los programas son del paquete de GCG, Wisconsin, Madison.

UDP galactosa-4-epimerasa (GALE); 3-β- hidroxil-5ene-esteroide-deshidrogenasa (3BH1); NAD(P)-dependiente Colesterol deshidrogenasa (NOCDDH); Dihidroflavonol 4-reductasa (DFRA); NADH HC-toxina reductasa (MZETX); Cinnamoyl CoA- reductasa (CCR); ScORF: *S. cerevisiae* ORF.

También se puede ver en la Fig. 1 que hay una enzima de plantas que participa en la biosíntesis de ligninas: la CCR (Cinamoil-CoA-oxidoreductasa dependiente de NADP<sup>+</sup>) que está evolutivamente más cercana al bloque donde se encuentra GRE2 que las

realización de un alineamiento entre GRE2 y las dos enzimas CCR lo cual mostró un 36% de identidad a nivel de aminoácidos entre estas dos grupos de enzimas. Estos datos confirman que, en efecto, GRE2 es mas parecida a las CCR que a las DFR.

Desafortunadamente esta homología no aporta ninguna información adicional a la posible función de GRE2 ya que la enzima CCR también participa (igual que las DFR) en el metabolismo secundario de las plantas y no se ha reportado ninguna actividad similar en levadura. De acuerdo a Baker y Blasco (1992) la UDP-4-galactosa-epimerasa pudiera ser el ancestro de las enzimas de la superfamilia (DHE/DFR) por lo que se podría pensar que la función de los genes de levadura pudiera estar más relacionado a metabolismo de azúcares o de ácidos nucleicos (ver distancias en la Fig. 1). Por esta razón se realizó un alineamiento entre el posible antecesor y GRE2 donde se observó que la homología a nivel de aminoácidos es muy baja, únicamente 36% de identidad en la región de unión a NADPH, por lo que no es muy factible que participen en la misma reacción enzimática. De cualquier manera sería interesante saber si la función de los genes de levadura está más relacionada a metabolismo de azúcares o de ácidos nucleicos que a la biosíntesis de metabolitos secundarios.

Otro dato adicional de la posible función de GRE2 está dado por el hecho de que el transcrito se induce en condiciones de estrés oxidativo y cuando la levadura cambia de metabolismo fermentativo a uno respiratorio (Figs. 2 y 3; Garay-Arroyo y Covarrubias, 1999), por lo que la enzima podría estar involucrada en modular el potencial redox de la célula. Se sabe que el cofactor de muchos miembros de la superfamilia (DHE/DFR) es NADPH así que el papel de GRE2 podría ser el de modulador de las pozas de NADPH bajo diferentes condiciones ambientales. En la Fig. 2 se muestra un alineamiento en donde se muestra el sitio de unión al cofactor NADPH, que se conserva en GRE2 y en varios miembros de la superfamilia de (DHE/DFR).

ygl1039w	~~~~~	~~~~~	~~~MTTEKTV	VFVSGATGFI	ALHVVDLLK
ygl1157w	~~~~~	~~~~~	~~~MTTDTT	VFVSGATGFI	ALHIMNDLLK
gre2	~~~~~	~~~~~	~~~~~ms	vfvsgangfi	aghivdl11k
ydr541c	~~~~~	~~~~~	~~~~~MSNT	VLVSGASGFI	ALHILSOLLK
X79566	~~~~~	~~~~~mp	vdalpgsgqt	vcvtgaggfi	aswlvkl11e
X98083	mtvvdavvss	tdagapaaaa	apvpagngqt	vcvtgaagyi	aswlvkl11e
S69616	~~~~~	~~~~~	~~~mdgnkgp	vvtgasgfv	gswlvk11q
X05068	~~~~~	~~~~~	meggesakgt	vlvtgasgfv	gswlvk11q
X80222	~~~~~	~~~~~	~~~~~	~~~~~	gswlvml1e
X75964	~~~~~	~~~~~	~~~msgqset	vcvtgasgfi	gswlvml1e
M86359	~~~~~	~~~~~	~~~mvsqket	vcvtgasgfi	gswlvml1e
X15537	~~~~~	~~~~~	mplhlrcsat	vcvtgaagfi	gswlvml1e
Z18277	~~~~~	~maseahavv	dahsppkttt	vvtggagfi	gswlvml1e
X15536	~~~~~	~msptslnts	setappsstt	vcvtgaagfi	gswlvml1e
L02540	~~~~~	~~~~~	maekesngvr	vcvtggagfi	gswlvml1e

ygl1039w	<b>TGYKVI</b> GSGR	SQEKNDGLLK	KFKSNPNLSM	EIVEDIAAPN	AFDKVFQKHG
ygl1157w	<b>AGYTFV</b> GSGR	SQEKNDGLLK	KFNNNPKLSM	EIVEDIAAPN	AFDEVFKKHG
gre2	edykvigsar	sqekaenlte	afgnnpkfsm	evvpdiskld	afdhvfqkhg
ydr541c	<b>QDYKVI</b> GTVR	SHEKEAKLLR	QFQHNPNTL	EIVPDISHPN	AFDKVLQKRG
X79566	<b>rgytr</b> atvr	npdd.pkng	lrel.egase	rltl.ykgdl	mdygsleeai
X98083	<b>kgytr</b> atvr	npdd.pknah	lral.dgaae	rlil.ckadl	ldydai crav
S69616	<b>agytv</b> atvr	dpanvektkp	l1el.pgake	rlsi.wkadl	sedgsfneai
X05068	<b>agytv</b> atvr	dpanvgktp	lmdl.pgate	rlsi.wkadl	aeegsfddai
X80222	<b>pgymv</b> atvr	dpenlkkvsp	l1el.pgaks	klsi.wkadl	geegsfdeai
X75964	<b>rrlt</b> atvr	dptnvkkvkh	l1dl.pkaet	hltl.wkadl	adegsfdeai
M86359	<b>rgyfv</b> atvr	dpgnlkkvqh	l1dl.pnakt	qltl.wkadl	seegsyddai
X15537	<b>rgynv</b> atvr	dpenkkvkh	l1el.pkadt	nl1l.lkadl	tveg sfdeai
Z18277	<b>rgynv</b> atvr	dpenqkkvkh	l1el.pkadt	nl1l.wkadl	aveg sfdeai
X15536	<b>rgytv</b> atvr	dpgnmkkvkh	l1el.pkadt	nl1l.wkadm	tveg sfdeai
L02540	<b>kgytv</b> atvr	ntgdeakagl	lrrlvpgaae	rlrl.fqadl	fdaatfapai

Fig. 2. Alineamiento del probable sitio de unión a NAD(P) localizado en la región amino-terminal de diferentes miembros de la superfamilia (DHE/DFR). Se indican los números de acceso de cada una de las proteínas a la izquierda de la figura. Las proteínas se obtuvieron de las secuencias depositadas en los bancos de datos. Los aminoácidos se alinearon usando el programa Pileup del paquete de GCG, Wisconsin, Madison. En negro están indicados los aminoácidos idénticos por lo menos en diez de las secuencias utilizadas para llevar a cabo el alineamiento.

**GRE3 tiene alta homología con xilosas reductasa descritas en diferentes hongos.**

GRE3 es miembro de una familia de aldo-ceto reductasas que presentan una



secuencia conservada (IPKS), que es el sitio activo tanto en aldehído como en aldosa-reductasas (ver Fig. 3). Gran parte de los miembros de esta familia catalizan la siguiente reacción:

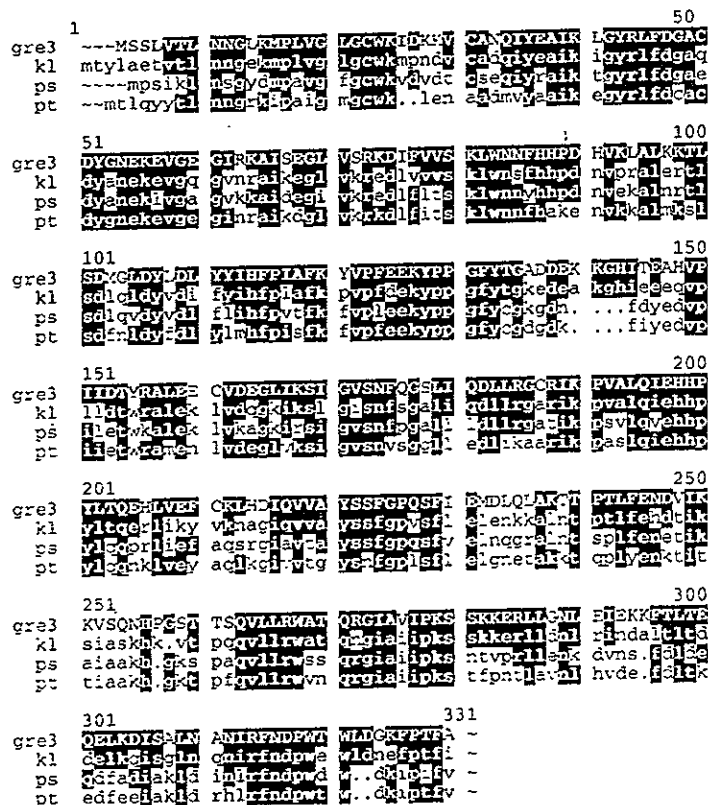
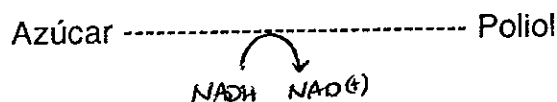


Fig.3. Alineamiento múltiple de diferentes xilosas reductasas. Las proteínas se obtuvieron de las secuencias depositadas en los bancos de datos. Los aminoácidos se alinearon usando el programa Pileup del paquete de GCG, Wisconsin, Madison. KL (*Kluyveromyces lactis*; Billard et al., 1995), PS (*Pichia stipitis*; Takuma et al., 1991) y PT (*Pachysolen tannophilus*; Bolen et al., 1996). En negro se encuentran indicados todos los aminoácidos idénticos por lo menos en dos de las cuatro secuencias utilizadas.

Esta proteína presenta una alta homología (65-66%, ver Fig. 3) con reductasas de xilosa (XR) descritas en diferentes hongos (Billard et al., 1995; Bolen et al., 1996; Takuma et al., 1991). La XR es la enzima que cataliza la primera reacción de la vía de síntesis de

xilulosa, que puede ser utilizada como fuente de carbono por *S. cerevisiae*:



El producto de la reacción de la enzima XR es xilitol y se requiere de una xilitol deshidrogenasa (XD) para producir xilulosa (ver esquema anterior) la cual puede entrar al ciclo de las pentosas. Sin embargo, desde hace tiempo se sabía que la levadura *S. cerevisiae* es incapaz de crecer en xilosa como única fuente de carbono, aunque sí puede utilizar xilulosa (Van Zyl et al., 1989). Por lo anterior se pensaba que la XR no existía en *S. cerevisiae* y se llevaron a cabo algunos experimentos introduciendo la enzima heteróloga XR de *Pichia stipitis* en *S. cerevisiae* (Takuma et al., 1991) para probar que la falta de crecimiento de esta levadura se debía a la ausencia de la enzima. En esta levadura transformada tampoco se observó crecimiento en xilosa como única fuente de carbono por lo que se sugirió que la enzima limitante podría ser la siguiente enzima de la vía de síntesis de xilulosa, la XD y, que se necesitaba tener una levadura con ambas enzimas heterólogas para que creciera en xilosa como fuente de carbono. Sin embargo, esta hipótesis no ha sido probada ya que no se ha construido esta doble transformante conteniendo tanto la XR como la XD. Actualmente, con la disponibilidad de toda la secuencia del genoma y, por los datos que se obtuvieron sobre *GRE3*, se sabe no sólo que ambas enzimas están presentes en el genoma de la levadura (existe un "marco abierto de lectura" que tiene entre un 50-60% de identidad con la enzima XD, datos no mostrados) sino que, al menos *GRE3* se expresa bajo diferentes condiciones de crecimiento. A la fecha se desconoce si la enzima XD en la levadura es funcional y en que condiciones se expresa. Por lo anterior se puede postular que *GRE3* podría catalizar una reacción para producir un osmolito compatible (que bien pudiera ser xilitol) bajo ciertas condiciones de crecimiento usando diferentes sustratos como se ha postulado para otras reductasas de xilosa (Bolen et al., 1985).

Por otro lado, prácticamente todas las reductasas de xilosa reportadas utilizan NADPH como cofactor (*Pichia stipitis* también puede utilizar NADH) así que, como en el caso de *GRE2*, también es posible proponer que *GRE3* juega un papel en modular las pozas de su cofactor NADPH bajo ciertas condiciones de estrés y/o ciertas condiciones de crecimiento.

Es interesante notar que muchos de los cambios tanto en las actividades de las proteínas como en la inducción de genes que se regulan por condiciones de estrés osmótico se comparten con aquellas que se han observado en condiciones de estrés oxidativo (comparar Fig. 2 de la introducción de éste capítulo y Fig. 6 de Godon et al.,

1998). Uno podría hipotetizar que el factor común en estos dos tipos de estrés es que una condición de estrés osmótico induce primariamente un estrés oxidativo. Esta hipótesis sería compatible con la posible función de GRE2 y GRE3 como moduladores de las pozas de NADPH (NADH) y, por tanto, del potencial redox.

### **Regulación transcripcional de los genes *GRE*.**

En cuanto a la regulación de estos genes se refiere se puede concluir lo siguiente: presentan una regulación semejante a lo reportado para otros genes de respuesta a estrés, ya que se regulan positivamente por de la vía de HOG y los genes *GRE1* y *GRE3* son regulados negativamente por AMPc-PKA. Sin embargo, en el caso de *GRE2* no encontramos regulación negativa mediada por la cinasa PKA, ya que no cambian sus niveles transcripcionales en la mutante *bcy1* (comparados con el control). Este dato sugirió que los niveles de AMPc no estaban regulando la transcripción de este gen. Sin embargo, en estudios recientes (Benítez, comunicación personal), usando el promotor de *GRE2* fusionado con un gen reportero (*lacZ*) se encontró que el promotor del gen *GRE2* responde a los niveles de AMPc, de tal forma que una mutante que carece de la *CDC25* y, por tanto presenta niveles bajos de AMPc (OL520-1; Flattery-O'Brien et al., 1997), posee una alta actividad de  $\beta$  galactosidasa; este efecto se revierte por la adición de AMPc. Los datos anteriores indican que la mutante *bcy1* genera resultados que se interpretan incorrectamente (*GPP2* (Hirayama et al., 1995; Boy-Marcotte et al., 1998)) y sugieren que el AMPc puede estar regulando la expresión transcripcional por medio de otras proteínas diferentes a la cinasa PKA.

Adicionalmente, se demostró que los factores transcripcionales Msn2p y Msn4p que se sabe que, *in vitro*, se unen al elemento STRE, regulan positivamente a *GRE1* y *GRE3*. En el caso de *GRE2*, cuando el análisis de expresión se lleva a cabo utilizando como sonda el gen completo, no es posible observar efecto en el fondo genético *msn2* y *msn4*; sin embargo, cuando se usó como sonda la región 3' no traducida del gen, es posible detectar que *GRE2*, al igual que *GRE1* y *GRE3*, se regulan positivamente por los factores transcripcionales Msn2p y Msn4p (datos no mostrados). Estos datos van de acuerdo con la existencia de más de un gen homólogo de *GRE2* en *S. cerevisiae*, de tal forma que cuando se utiliza la sonda completa es posible detectar los diferentes transcritos similares a *GRE2*, en tanto que cuando se usa la región 3' no traducida, específica de *GRE2*, sólo se detecta el transcrito de éste gen. Estos datos correlacionan con datos recientes obtenidos en el laboratorio (E. Benítez, comunicación personal) en donde el promotor de *GRE2* esta

fusionado al gen *lacZ*, y no se observa actividad del gen reportero en la mutante *msn2* y *msn4*; lo anterior indica que estos dos factores transcripcionales funcionan como reguladores positivos de *GRE2*. Por otro lado, es posible concluir que alguno (s) de los tres genes homólogos a *GRE2* no se regula por estos factores transcripcionales.

Recientemente se ha demostrado una interacción entre los factores transcripcionales Msn2p y Msn4p y PKA. Cuando hay condiciones de estrés en el medio de crecimiento, los factores transcripcionales Msn2p y Msn4p se encuentran, preferencialmente, en el núcleo de las levaduras, activando genes de respuesta a estrés. Cuando hay condiciones óptimas de crecimiento y glucosa como fuente de carbono se incrementa la actividad de la proteína cinasa dependiente de AMPc (PKA) y los factores Msn2p y Msn4p se encuentran, principalmente, en el citoplasma (Görner et al., 1998). Esto correlaciona con los datos de los northern de los genes *GRE 1* y *GRE3* en los diferentes fondos genéticos. Ambos genes se regulan positivamente por los factores Msn2p y Msn4p y negativamente por PKA. Adicionalmente, y apoyando esta observación, se ha visto que genes que se inducen en la fase diaúxica (cuando la levadura cambia de un metabolismo fermentativo a uno respiratorio) se regulan positivamente por los factores transcripcionales Msn2p y Msn4p y negativamente por AMPc (Boy-Marcotte et al., 1998), lo cual concuerda con lo encontrado para los genes *GRE2* y *GRE3* (ver Fig. 2 Garay-Arroyo y Covarrubias, 1999).

El análisis de las regiones promotoras de los diferentes genes permitió la identificación de diferentes cajas o elementos de regulación que ayudaron a explicar el por qué de la inducción que observamos en los diferentes fondos genéticos (ver figura 5; Garay-Arroyo y Covarrubias, 1999). Se ha demostrado que la secuencia STRE es un elemento que confiere respuesta a una gran variedad de estreses cuando se fusiona con un gen reportero (Schüller et al., 1994). Como se puede ver en la Fig 5 del artículo de la revista *Yeast*, el gen *GRE1* no presenta ningún elemento STRE pero tiene 6 elementos PDS; *GRE2* tiene una caja HSE y un elemento PDS y *GRE3* presenta un elemento STRE y tres elementos PDS. Nosotros proponemos que este último elemento, PDS, podría funcionar como el elemento regulador en muchas condiciones de estrés y/o que hay un elemento alternativo que no se ha identificado y que funciona como el responsable de la activación. De hecho, y apoyando la primera hipótesis, se ha demostrado que el elemento PDS compite con el elemento STRE en geles de retardamiento (Marchler et al., 1993) y se ha propuesto que los factores transcripcionales Msn2p y Msn4p pudieran estar uniéndose a esta secuencia (Boy-Marcotte et al., 1998). Sería interesante demostrar si la participación, en condiciones de estrés osmótico, de un elemento STRE es equivalente a

aquella que presenta un elemento PDS y, si los elementos PDS pudieran estar actuando sinérgicamente entre sí o en conjunto con el elemento STRE, ésta última hipótesis de sinergismo está comprobada para los elementos PDS-HSE (Kobayashi y MacEntee, 1993) y para diferentes elementos STRE (Varela et al., 1995).

No entendemos aún el papel de estos genes durante la respuesta a estrés ya que la interrupción de ninguno de ellos resulta en un fenotipo de sensibilidad, al menos en las condiciones de estrés estudiadas hasta ahora. Es posible que la causa de no haber detectado ningún fenotipo se deba no sólo a la redundancia genética sino también a la redundancia funcional. De ahí la importancia de generar mutantes múltiples que permitan profundizar en este análisis. Por otro lado, sería necesario explorar la función de los productos de estos genes, por enfoques bioquímicos que permitan comprobar la identidad sugerida por el análisis de su secuencia.

Por último, presento un esquema en donde se resumen los diferentes factores que se han descrito hasta el momento y juegan un papel en modular la expresión de los genes bajo condiciones de estrés osmótico. Esta respuesta involucra, principalmente, tres vías de regulación; una de ellas actúa de manera positiva, la vía de HOG, y dos ejercen un efecto negativo, la regulación ejercida por la cinasa PKA y la actividad represora del complejo Ssn6p-Tup1p-Sko1p. (ver Fig. 4A y 4B). Como se puede ver en la Fig. 4A, en condiciones normales de crecimiento (en glucosa como fuente de carbono) la vía de HOG está apagada; los niveles de AMPc se encuentran altos (por lo que la actividad de la cinasa PKA también está alta) y los factores transcripcionales Msn2p y Msn4p se localizan en el citoplasma. Además, el complejo represor Ssn6p-Tup1p-Sko1p se encuentra unido a su secuencia en las regiones promotoras reprimiendo la expresión del gen. Cuando el organismo se somete a condiciones de estrés osmótico (ver Fig. 4B), la vía de HOG se activa, la cinasa HOG1 se fosforila (se activa) y se va al núcleo; los niveles de AMPc (y por lo tanto la actividad de PKA) disminuyen y los factores transcripcionales Msn2p y Msn4p se van al núcleo. Paralelamente a estos cambios, el complejo represor Ssn6p-Tup1p-Sko1p se libera de los promotores y deja de reprimir. Todos estos cambios actúan simultáneamente para que la levadura se adapte a las nuevas condiciones de crecimiento.

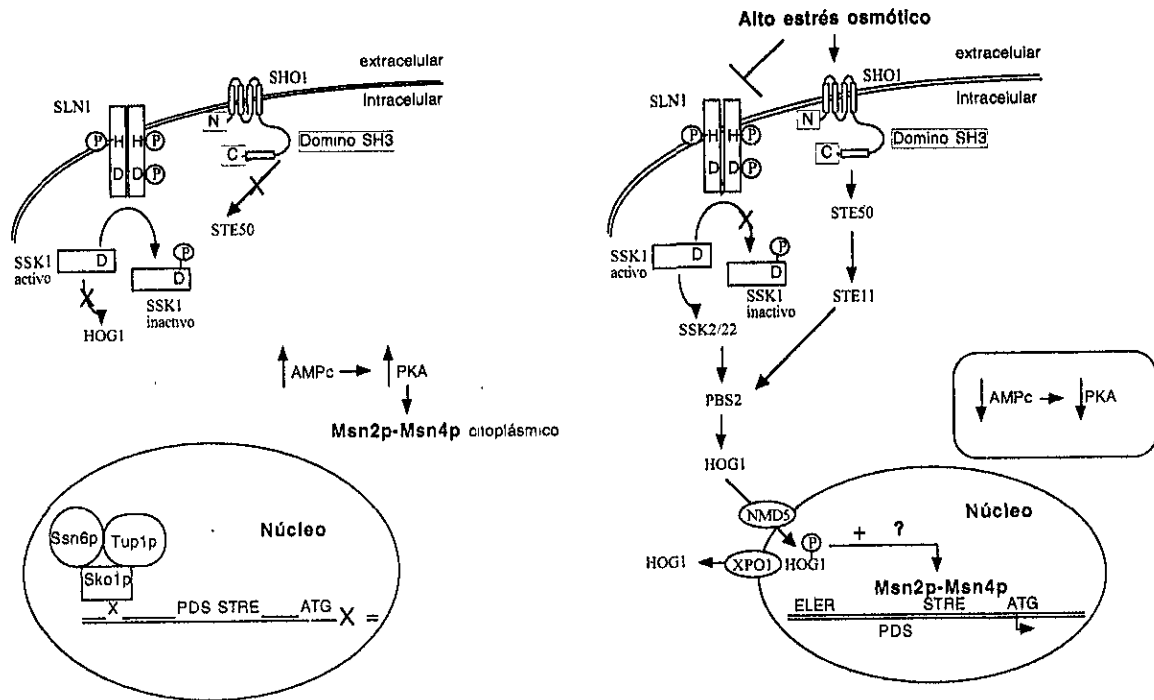


Figura 4A. Modelo de una célula en condiciones óptimas de crecimiento. Figura 4B. Modelo de una célula crecida en condiciones de estrés osmótico

Podemos concluir, por la información que se conoce hasta el momento de la respuesta de la levadura *S. cerevisiae* ante condiciones de estrés hiperosmótico, que existe una íntima interacción entre la respuesta a estrés y la modulación metabólica, que este tipo de estrés induce una gran variedad de proteínas y transcritos con funciones diversas, lo cual indica que la respuesta es compleja, y que se requiere de esta diversidad para contender con una situación tan adversa como es la limitación de agua. Aún queda por determinar las interacciones entre estos genes y los factores que modulan su actividad; si existen condiciones en las que algunos genes sean más importantes que otros y cuáles son estas condiciones.

## Capítulo 2

### Introducción

#### **i) Proteínas abundantes de la embriogénesis tardía (proteínas LEA).**

Las proteínas LEA son proteínas ubicuas en plantas superiores que se agrupan debido a su inducción en las últimas etapas de la embriogénesis tardía, justo antes del inicio de la desecación de la semilla. Estas proteínas se acumulan en respuesta a estrés osmótico en los tejidos vegetativos o por aplicación exógena de ABA (hormona ácido abscísico) en embriones inmaduros o en tejidos vegetativos no sujetos a estrés (Revisado en Dure III, 1993). En diferentes casos se ha demostrado que, en condiciones de estrés osmótico, la inducción de sus genes está mediada por la hormona ABA. Por lo anterior, se puede decir que el inductor común de la expresión de estas proteínas, es la pérdida de agua de la célula.

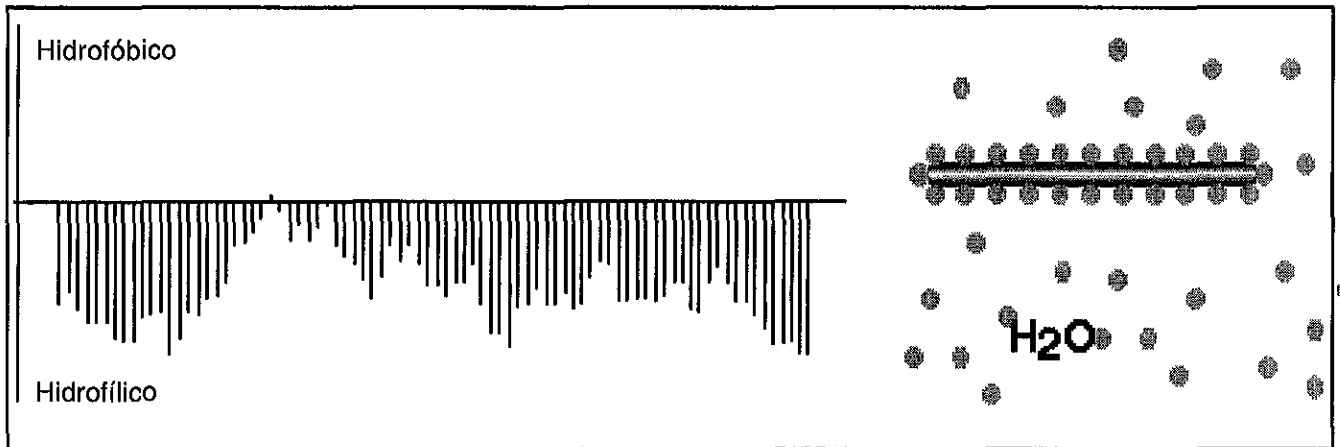
Las proteínas LEA son altamente hidrofílicas, tienen un alto contenido de glicinas y de aminoácidos cargados y, en su mayoría, carecen de cisteínas y triptófanos. Se predice que algunas de ellas no poseen estructura globular y las que tienen algún grado de estructuración forman  $\alpha$ -hélices anfifílicas (Revisado en Dure III, 1993). Estas proteínas se agrupan en cinco familias de acuerdo a la homología en su secuencia de aminoácidos; sin embargo, se han encontrado algunas proteínas que no se pueden clasificar en ninguno de estos grupos, por lo que se les considera proteínas LEA atípicas y se han clasificado en un grupo aparte (5) (Colmenero-Flores et al., 1999; Dure III, 1993; Galau et al., 1993).

Se sabe poco sobre la localización de las proteínas LEA, algunos miembros de las familias 3 y 4 están presentes en todos los tipos celulares del embrión (Roberts et al., 1993, Goday et al., 1994, Close et al., 1993a) y, parece que todas las proteínas LEA están presentes en el núcleo y citosol de tejidos meristemáticos, vasculares y provasculares de la parte vegetativa de plantas sujetas a estrés (Close et al., 1993a; Goday et al., 1994, Niogret et al., 1996) .

#### **a) Grupo 1 (familia D-19).**

Una característica muy particular de esta familia es que los miembros presentan una alta homología entre sí (65% de identidad y 80% de similitud a nivel de aminoácidos). Son proteínas pequeñas (entre 90-110 aa) y no presentan estructura secundaria ya que tienen un alto contenido de glicinas (16-21%) distribuidas a lo largo de toda la proteína. El grupo de McCubbin et al. (1985) encontró que la proteína *Em* de trigo se hidrata

considerablemente más que proteínas globulares y que esta propiedad puede deberse a que tiene un 70% de estructura en "random coil" en condiciones fisiológicas. Esta propiedad de estar en "random coil", parece correlacionar con la capacidad de ser altamente hidratable. Se ha propuesto que esta característica ayuda a prevenir o combatir la desecación.



**Fig. 1:** Estructura y perfil hidropático de una proteína LEA del grupo 1. Como se puede ver la proteína es altamente hidrofílica y su forma es extendida, por lo que tiene gran capacidad para interactuar con el agua (círculos)

### b) Grupo 2 o dehidrinas (familia D-11).

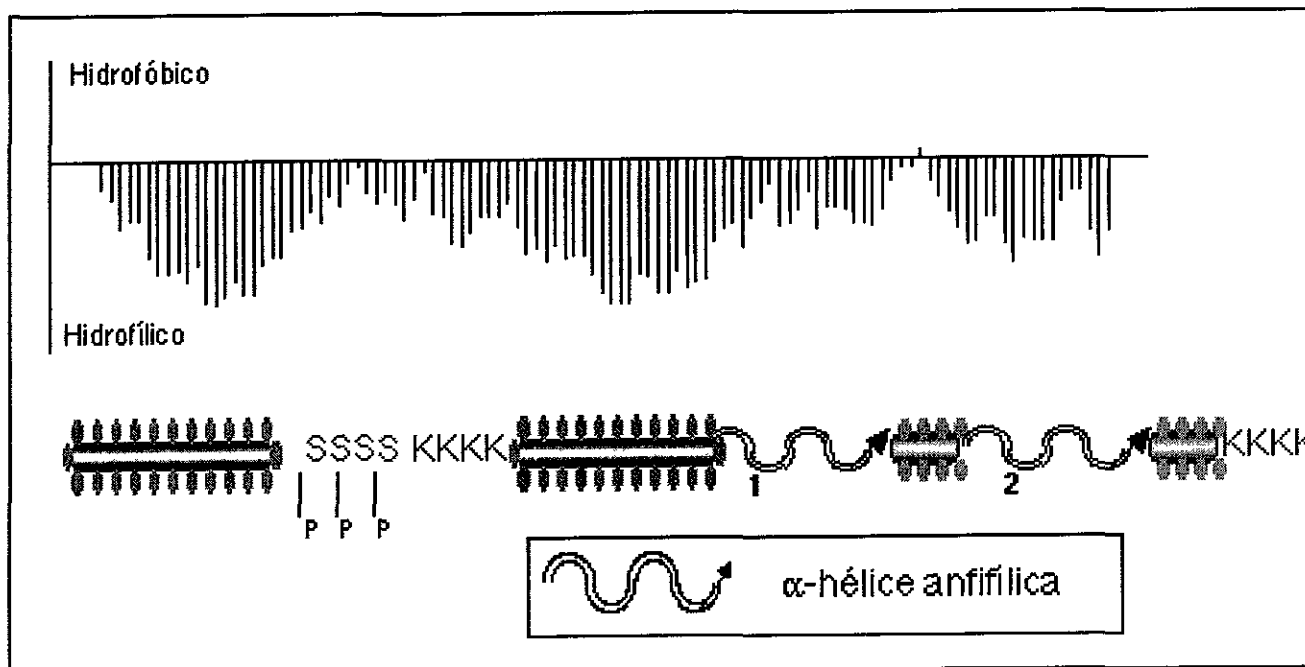
Son también proteínas muy hidrofílicas que poseen, en el carboxilo terminal, una región de 14 aa (EKKGIMDKIKEKLP) que puede estar hasta tres veces en tandem; esta región forma una  $\alpha$ -hélice anfifílica. La proteína presenta una serie de 6-9 serinas contiguas fosforilables seguidas por una región rica en aminoácidos cargados positivamente (principalmente lisinas). Goday et al. (1994) demostraron que las proteínas encargadas de acarrear a otras proteínas al núcleo, reconocen este motivo. Las dehidrinas representan alrededor del 0.5% de la proteína total soluble de un embrión (Dure III, 1993).

### c) Grupo 3 (familia D-7).

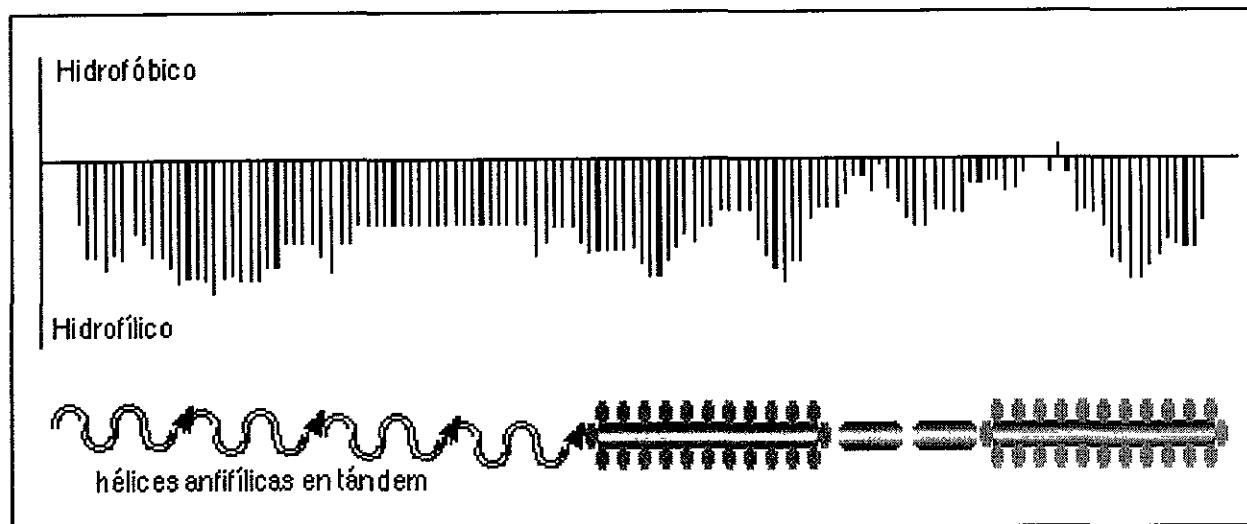
Los miembros de esta familia presentan tamaños muy variables debido a la presencia de 11 aminoácidos (TAQAAKEKAGE) conservados y repetidos en tandem (entre 5 en algodón y 13 en *Brassica*). Como en el caso de las dehidrinas, estos motivos están muy conservados y constituyen  $\alpha$ -hélices anfifílicas que pueden formar un homodímero. Se postula que este dímero puede secuestrar iones. Se encuentran en embriones de algodón en una concentración estimada de 0.34 mM, lo que supone que constituye el 4% de la



proteína citosólica, no organelar (Revisado en Dure III, 1993).



**Fig. 2:** Perfil hidropático y esquema de la estructura de una proteína LEA del grupo 2 o dehidrina. El tracto de serinas está indicado con una S y está seguido de aminoácidos cargados positivamente (K). Las barras gruesas esquematizan las regiones predichas como "random coil" y los óvalos representan moléculas de H<sub>2</sub>O

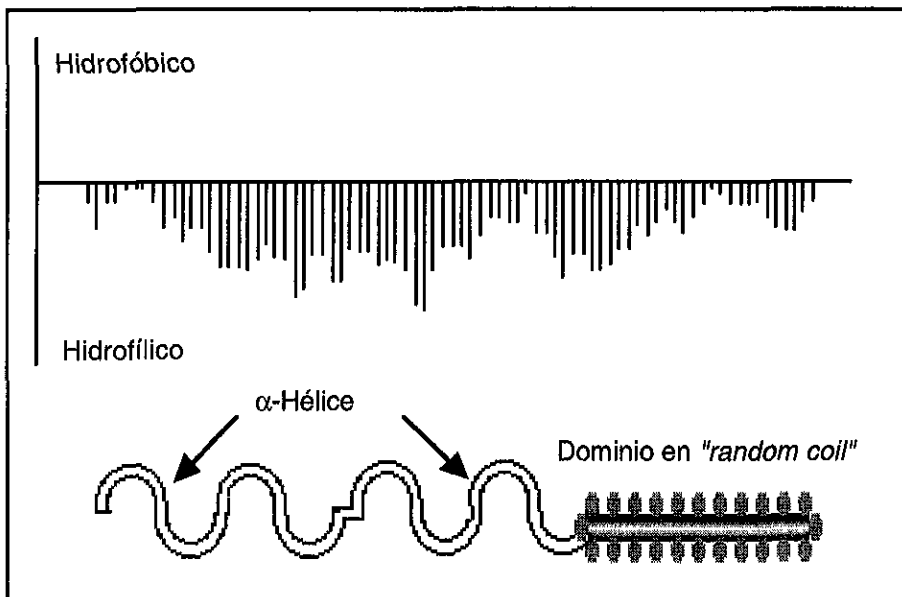


**Fig. 3:** Perfil hidropático y esquema de la estructura de una LEA del grupo 3

**d) Grupo 4 (Familia D-113).**

Poseen un rango muy variable de tamaño (88-175 aminoácidos), son ricas en

glicinas y treoninas y, a diferencia de las otras familias, son muy ricas en alaninas (11-20%). Según las predicciones, poseen siempre un dominio muy conservado de unos 60-80 aminoácidos con estructura prácticamente ininterrumpida de  $\alpha$ -hélice, mientras que el resto de la proteína varía en tamaño y su secuencia está poco conservada. Aún cuando, en esta última región, no hay homología a nivel de aminoácidos, hay similitud fisicoquímica y se mantiene la estructura en "random coil". Estas proteínas también llegan a alcanzar una concentración 0.3 mM en el embrión antes de que empiece la fase de desecación. (Revisado en Dure III, 1993).

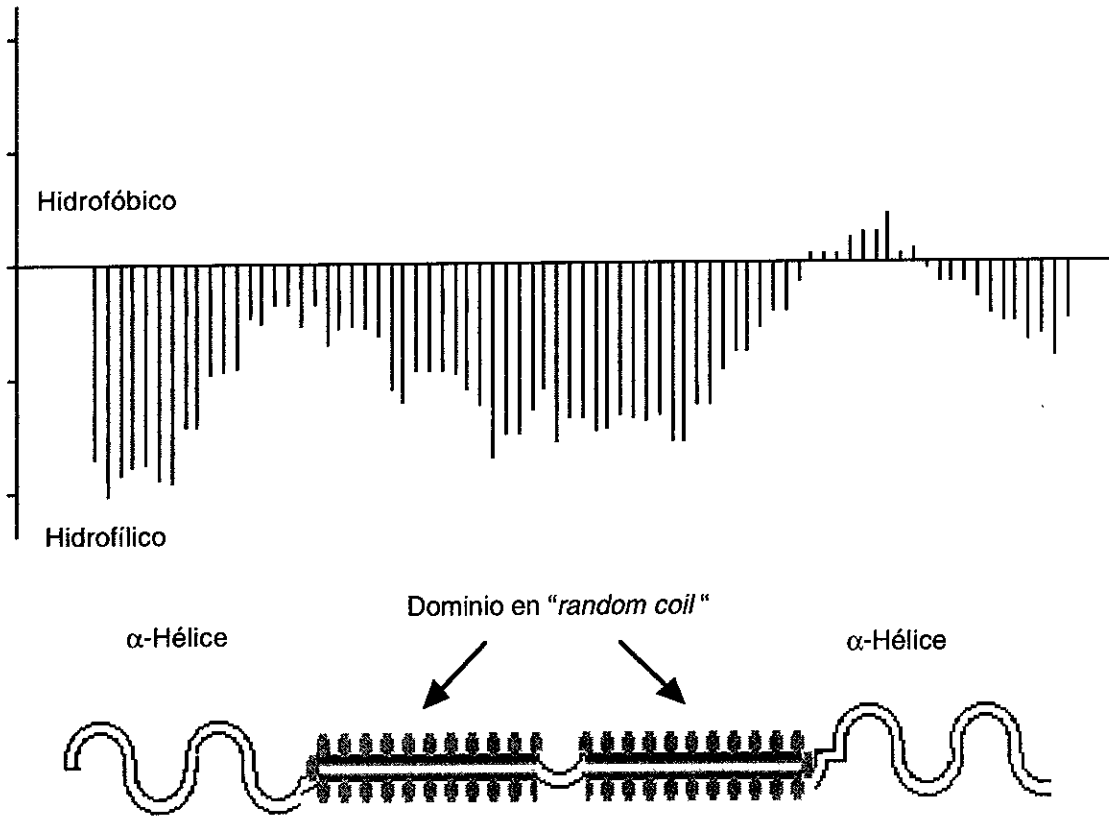


**Fig. 4:** Perfil hidropático y esquema de la estructura de una proteína LEA del grupo 4.

#### e) Grupo 6 (PvLEA-18).

Esta proteína se caracterizó en *Phaseolus vulgaris* (cv Negro Jamapa) y se constituyó como una nueva familia, ya que se han encontrado genes homólogos y/o proteínas relacionadas tanto en otras dicotiledóneas como *Arabidopsis* como en monocotiledóneas como *Zea mays* (maíz). Además, se han encontrado proteínas relacionadas en diferentes variedades de frijol como son var. cacahuete y var. flor de mayo (Colmenero-Flores et al., 1999). En general, son proteínas de bajo peso molecular (aprox. 10 KDa) con alto contenido de estructura en "random coil" (Fig. 5). Se ha visto que ésta proteína se acumula en semilla seca (con una concentración aproximada de 0.51  $\mu$ M en embriones de 36h) y en respuesta a situaciones de déficit hídrico. También es capaz de

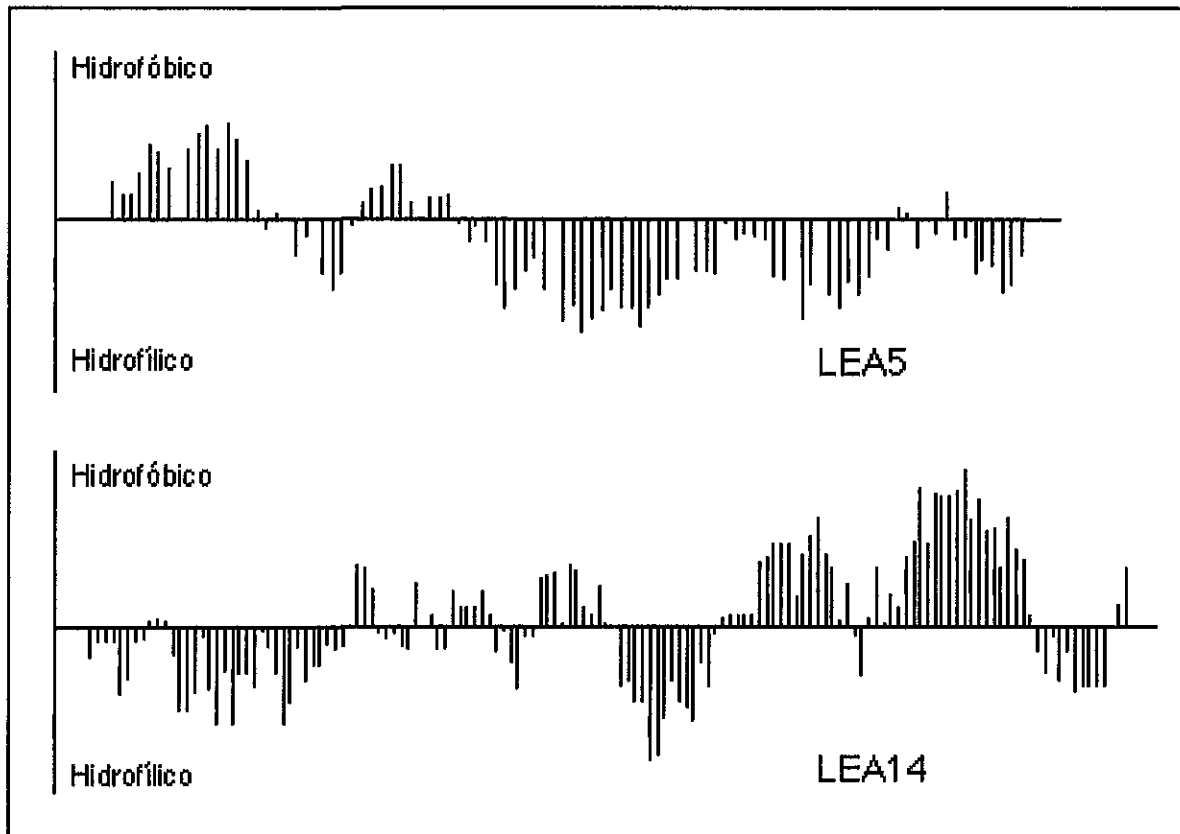
responder durante el desarrollo de la planta, encontrándose en aquellos tejidos en donde los potenciales osmóticos son más negativos, como el cilindro vascular y la epidermis; o bien, en formas especializadas como el polen que es una estructura de resistencia que se somete a un proceso de deshidratación severo.



**Fig. 5:** Perfil hidropático y esquema de la estructura de una proteína LEA de la familia 6 (PvLEA-18).

### Grupo 5 (LEAs atípicas).

Las llamamos atípicas ya que no pertenecen a ninguna de las 5 familias descritas y no presentan el mismo grado de hidrofiliidad que predomina en estos grupos, ni tienen la misma composición de aminoácidos que caracteriza a las proteínas LEA hidrofílicas. Se puede decir que son proteínas de embriogénesis tardía que cumplen funciones diferentes a las "típicas" proteínas LEA hidrofílicas.



**Fig. 6:** Perfil hidropático de las proteínas LEA atípicas 5 y 14. Notese el alto contenido de regiones hidrofóbicas.

Una representante de este grupo es la proteína LEA5 de algodón que tiene un peso molecular de 11.4 KDa, que presenta dominios hidrofóbicos en la región amino terminal que pudieran constituir una región transmembranal o una señal de tránsito al retículo endoplásmico (Jackson et al., 1997). El resto de la proteína es hidrofílica y no es particularmente rica en glicinas, pero sí en serinas (casi 15%). La proteína LEA14, también de algodón, es otra proteína LEA atípica ya que es moderadamente hidrofóbica en casi toda su extensión. Es una proteína de 16.4 KDa con bajo contenido en glicinas y que también se acumula abundantemente en la hoja, en condiciones de deshidratación (Galau et al., 1993). Se han encontrado en otras especies proteínas homólogas a las proteínas LEA 5 y LEA 14 de algodón que también se acumulan en embriogénesis tardía y que se inducen por otros procesos diferentes a la deshidratación. En el caso de la proteína LEA5 hay homología con proteínas cuyos genes son inducidos por auxina (Yamamoto et al., 1992), por ácido giberélico, por choque térmico, por metales pesados y por tuberización (Jackson et al., 1997). La proteína LEA14 es casi idéntica a la proteína Lemmi9 de tomate, que se induce en las agallas generadas por la infección de un nematodo (Van der Eycken

et al., 1996) (Fig. 6).

## ii) Posibles funciones de las proteínas LEA.

La función biológica de las proteínas LEA se desconoce. Su alta hidrofiliidad, la elevada concentración en la que llegan a acumularse en el embrión y la variabilidad en los tamaños de los polipéptidos dentro de una misma familia hacen muy poco probable que sean enzimas o que tengan una función estructural. Por otra parte, se ha hipotetizado que su conformación abierta y su gran hidrofiliidad les proporciona una gran capacidad para interaccionar con el agua y crear ambientes hidratados en donde se protegerían macromoléculas y otras estructuras celulares de los efectos de la sequía. Además, en condiciones de deshidratación severa, podrían evitar la cristalización de componentes celulares (Ingram y Bartels, 1996; McCubbin y Kay, 1985).

A las proteínas LEA se les han asignado varias funciones hipotéticas:

- 1) Retención de agua. Se han determinado las propiedades hidrodinámicas de la proteína de maíz *Em1*, una proteína LEA del grupo I, y se observó que un 70% de la proteína es "random coil" (lo cual confirmaría lo predicho para una proteína tan hidrofílica) y, por lo tanto, altamente hidratable (McCubbin y Kay, 1985). Esta propiedad física debería facilitar la interacción con el agua en células que sufren estrés osmótico, e impedir la deshidratación completa de la semilla. Walters et al. (1997) encontraron que, en semillas de trigo, las proteínas que permanecen solubles tras el hervor (principalmente proteínas LEA) poseen azúcares fuertemente unidos y que dichos complejos tienen propiedades de hidratación especiales. Ellos proponen que estas proteínas sirven como amortiguadores de la hidratación.
- 2) Funcionar como osmolitos compatibles. Las LEAs podrían también estabilizar un sistema a través del mecanismo de exclusión molecular que se le ha adjudicado a los osmolitos compatibles. Por ejemplo, las proteínas LEA del grupo 1 y 4 podrían estar implicadas en la solvatación de estructuras citosólicas a través de sus amplios dominios de estructura abierta (Baker et al., 1988). Estableciendo un paralelismo con los azúcares solubles, las regiones abiertas de las proteínas LEA exponen numerosos grupos hidroxilo que tienen una gran capacidad de solvatación. Estas regiones de estructura aleatoria les podría permitir unirse a la superficie de estructuras celulares, sustituyendo al agua en sus interacciones a través de puentes de hidrógeno, de forma equivalente al mecanismo que se le ha adjudicado a los disacáridos (como la trehalosa) como estabilizadores en condiciones de deshidratación extrema.
- 3) Funcionar como chaperonas de proteínas desnaturalizadas (Close et al., 1996). Este

autor propone que los segmentos polares de las dehidrinas podría formar puentes de hidrógeno con los grupos polares de las macromoléculas y, de esta manera, prevenir la coagulación de las mismas manteniéndolas en solución.

4) Atrapar iones. En el caso de las proteínas LEA del grupo 4, se ha encontrado que una región importante de la  $\alpha$ -hélice gigante posee una distribución preferencial de cargas (+) y (-) hacia lados opuestos de la hélice, lo que podría generar campos eléctricos que atraparían iones (Imai et al., 1996). Las proteínas LEA del grupo 3 poseen residuos en sus  $\alpha$ -hélices anfifílicas (gln3, glu7, lys8 y glu11) que, de estar formando puentes salinos con otra hélice anfipática, podrían pasar a formar puentes salinos con iones del medio cuando dichos iones rebasen una concentración crítica durante la deshidratación (Dure III, 1993).

Por otro lado, se han llevado a cabo una serie de experimentos, tanto *in vivo* como *in vitro*, para tratar de entender la función de estas proteínas: (1) un miembro del grupo 2 de espinaca es capaz de proteger *in-vitro* a la enzima lactato deshidrogenasa frente a inactivación por congelamiento (Kazuoka y Oeda, 1994); (2) la sobreexpresión de una proteína LEA del grupo 3 de cebada, HVA1, aumenta la resistencia del arroz a la deshidratación y a la salinidad (Xu DP et al., 1996); (3) se reportó un aumento en la resistencia de la levadura *S. cerevisiae* a la salinidad y al congelamiento cuando se sobreexpresa una proteína LEA del grupo 4, LE25 (Imai et al., 1996). Sin embargo, aún cuando estas proteínas confieran resistencia a la desecación, se desconocen los mecanismos moleculares por los cuales lo logran.

## **I. Hipótesis**

Las características fisicoquímicas comunes a las proteínas LEA de plantas, como son el hecho de tener un alto porcentaje de glicinas (arriba del 8%) y una hidrofiliidad alta ( $>1$ ), también están presentes en proteínas involucradas en la respuesta a déficit hídrico de diferentes organismos, tanto procariotes como eucariotes.

## **II. Objetivo general**

Identificar proteínas que presenten un alto contenido de glicinas así como una elevada hidrofiliidad en otros organismos diferentes a plantas y analizar su expresión bajo condiciones de déficit hídrico.

## **III. Objetivos particulares**

1. Desarrollar un algoritmo que considere como parámetros de búsqueda un alto contenido de Gly (arriba del 8%) y una alta hidrofiliidad ( $>1$ ).
2. Realizar la búsqueda utilizando como modelos experimentales, en un inicio, los genomas de *E. coli* y de *S. cerevisiae*.
3. Hacer experimentos tipo northern blot para caracterizar la expresión de los transcritos de estas proteínas bajo condiciones de déficit hídrico.

# **Highly hydrophilic proteins are common during water deficit in different organisms.**

**A. Garay-Arroyo\*<sup>†</sup>, J. M. Colmenero-Flores\*<sup>†</sup>, A. Garciarrubio<sup>§</sup> and A. A. Covarrubias<sup>†¶</sup>**

<sup>†</sup>Departamento de Biología Molecular de Plantas and <sup>§</sup>Departamento de Bioestructura y Reconocimiento Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apdo. Postal 510-3, 62250 Cuernavaca, Mor. MEXICO.

Classification: Biological Science (Biochemistry)

Manuscript information: 28 pages.

Abstract word count: 249

Word and characters count in the complete text (including Table and Figure legends): 6,085 and 32,445.

\* These authors made an equal contribution to this work

¶ Corresponding author



**ABSTRACT.** Many organisms have to contend against water limited environments, in some cases these conditions can be extreme such as the dry seeds of higher plants. Very hydrophilic proteins known as Late Embryogenesis Abundant Proteins (LEAPs) have been found to accumulate not only in mature seeds but also during periods of water deficit in the plant vegetative organs. In spite of the lack of a significant homology among the sequences in the different groups of LEAPs, two unifying features can be distinguished, their extreme hydrophilicity and high percentage of glycine residues. Based on these structural features, we devised a powerful search criteria to screen for this kind of proteins in the available data banks. After demonstrating that the criterion adequately selects the LEAPs from among the plant subset of SwissProt database, a methodic screening was carried out in other taxon-specific subsets. From this search, a limited number of proteins with these structural characteristics were found. We called them "**hydrophilins**". Here, we report that within the genomes of *Escherichia coli* and *Saccharomyces cerevisiae* only 5 and 12 proteins, respectively, met our criteria. This finding suggests that hydrophilins are not very common: they usually constitute less than 1% of all proteins. We also demonstrate that most of the genes encoding these proteins in *E. coli* and *S. cerevisiae* are induced by osmotic stress. Together, these data indicate that this physicochemical criteria is a good predictor of participation in osmotic adaptation, and that the same solution has evolved in different taxonomic groups.

## Introduction

Very hydrophilic proteins, known as Late Embryogenesis Abundant Proteins (LEAPs) have been found to accumulate under conditions of extreme desiccation in higher plants. These proteins accumulate to high levels during the last stage of seed formation (when a natural desiccation of the seed tissues take place) and during periods of water deficit in vegetative organs (5). This set of proteins have been grouped on the basis of the presence of regions of significant homology with at least one protein of another species. This has allowed to distinguish, at least, six classes of LEAPs (11, 20). Although significant homology has not been detected between the members of the different classes, an unifying and outstanding feature of these proteins is their high hydrophilicity, and their high percentage of glycines and charged amino acids (4, 13, 14). Their amino acid sequences allow to predict that they exist principally as random coils (13). This property has been confirmed in few cases with purified proteins (23, 24), and it is supported by the fact that this type of proteins do not coagulate upon heating (7, 8). Most of the LEA protein families have been identified in a wide range of different plant species, to the extent that they could be considered ubiquitous in plants (9, 20). Finally, it has been shown that, at least, one of these protein families, the so called dehydrins, is present in a range of photosynthetic organisms including lower plants, algae and cyanobacteria (7, 9).

Two important questions are, how do these structural and physicochemical characteristics relate to the largely unknown function of LEAPs, and if they represent a solution to a plant specific problem, or to a more general one?. Recently, proteins with similar characteristics have been described in other organisms. That is the case of *HSP12* (26) and *GRE1* (17)

from *S. cerevisiae*, whose amino acid composition shows a high content of charged and hydrophilic amino acids distributed throughout their primary structure. In both cases, it has been shown that their corresponding transcripts accumulate in response to hyperosmotic conditions (17, 32). The *CON6* encoded protein of *Neurospora crassa*, also exhibit hydrophilic and highly charged repeated domains. The *CON6* polypeptide is present at high levels at the late stages of conidiation and in mature conidia, stages that involve a nutrient deprivation and a developmentally imposed (30).

Since the LEAPs of different classes have no evident sequence homology, it is unlikely that the structural and physicochemical characteristics they share are due to common descent. A more plausible hypothesis is that those characteristics allow LEAPs to carry on their specific functions under partial dehydration. If so, we should expect to find outside the plant kingdom that proteins with similar characteristics are used to face low water situations. We have chosen the name hydrophilins to group the proteins which present, as the LEAPs, extreme hydrophilicity and high percentage of glycine residues. Our aim in this work has been to explore how widespread, outside the plant LEAPs, is the association between these characteristics and the participation in the response to hyperosmotic conditions.

Conceptually, there were three stages in our strategy: first, with the help of an *ad hoc* program, we devised by trial and error a simple but general criterion to search the sequence databases for proteins having physicochemical resemblance to LEAPs. Then, using that criterion, proteins were fetched from several taxon specific databases. Finally, we investigated, either from the literature or by performing northern blot experiments whether the transcripts of the selected genes accumulate under water deficit. After

finding isolated examples in several species, we tested exhaustively *E. coli* and *S. cerevisiae*, two species whose genomes have been completely sequenced. A surprisingly high proportion of the proteins chosen by the physicochemical criterion also satisfied the expression requirement. Thus, we conclude that hydrophilins exist in several kingdoms including plant, bacteria and fungi, where they represent a similar solution to a prevalent problem of life under low water conditions.

## MATERIALS AND METHODS

**Strains, media and growth conditions.** The strain RS58 of *S. cerevisiae* was used in this work (19). Yeast cells were routinely grown on YPD medium (1% yeast extract, 2% Bacto peptone, and 2% glucose) at 30°C. The stress treatment involved a 30 min osmotic shock imposed by adding a final concentration of 0.5 M sorbitol to a yeast culture grown up to 0.5 OD<sub>660</sub>. For *E. coli*, most of the experiments were carried out using strain DH5 $\alpha$ . *E. coli* cells were routinely grown on LB medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl) at 37°C. As in the case of yeast cells, an osmotic shock was imposed to bacterial cultures grown up to 0.5 OD<sub>590</sub>. NaCl was added to a final concentration of 0.5 M and cells were harvested after 10, 20 and 60 min of treatment.

**DNA constructions.** All hydrophilins full length genes were obtained by PCR amplification. The primers used to obtain *S. cerevisiae* hydrophilin gene probes were: S65242-5 (5'-CTC GAG GCC AAA CAC TAC CGC ATA AAA GC-3'); S65242-3 (5'-CTC GAG CAA ACT GAC ATT AAT AAA AAC G-3'); STF2\_YEAST-5 (5'-AAT CAA CAG TAA CAA ACC GC-3') STF2\_YEAST -3 (5'-CCT TTT ATC TTG AAA CTG TGC-3'); YBM6\_YEAST-5 (5'-ATG TCT GCT AAC GAT TAC TAC GGC GG-3') YBM6\_YEAST-3 (5'-GAA AGG CAA CTG CAA AGA ACA GAA GGG G-3'); YFBO\_YEAST-5 (5'-TTC GTC CAC CTG ACC GAG CCG GC-3') YFBO\_YEAST-3 (5'-CCG TCA AAT CCA TCA TCA AAT CC-3'); IF1A\_YEAST-5 (5'-ATG GGT AAG AAA AAC ACT AAA GGT GG-3') IF1A\_YEAST-3 (5'-GAT ATT GAG GAC ACT TGT TCG GCC-3'); YN8D\_YEAST-5 (5'-GCA TAT AAG GAA CCA AGG GAA AAG C-3') YN8D\_YEAST-3 (5'-TTC TCA TTA TTA TAT ACG AAT ACG-3'); YDL-5 (5'-GGA AAA AGA CCC AAC AGT GCA GCC GG-3') YDL-3 (5'-GGA TGA ATT CCT GCA GGA ACA GGC-3') YJ04-5 (5'-CAA GAA GAA GAT GTT AAG GAG

GG-3') YJ04-3 (5'-CGT TAC AAC CCC GAA CCC GTG CG-3'); YJS4-5 (5'-GCC GAT AGT GCA CTG GAA CC-3') YJS4-3 (5'-CAA TCC TTC TTA ACT GTA TCC C-3'); YNT0-5 (5'-ATG AAG TTC TCT TCT GTT ACT GC-3') YNT0-3 (5'-CAG GAT GAA GGG GGT AGG GGG GG-3'), RL44-5 (5'-TCC TGA ATC GTA GGG CAA ATC C-3') RL44-3 (5'-AAT TGC AAA GCT TGA CCC TTT TGC-3') and MTC-5 (5'-GAT GCG TCT TTT CCG CTG AAC CG-3') MTC-3 (5'-GAT AGG TTT TCA GCA GCG GGT ACC-3'). To obtain the *E. coli* hydrophilin probes, the primers used were: YCIG-5 (5'-ATG CAA CGC AAC GGG AGA AAT AAC-3') YCIG-3 (5'-GCG AGT TTT GCC AGT GCC CGG G-3'); YJBJ-5 (5'-CTG AAT AAA TAA TCC TAA GCC-3) YJBJ-3 (5'-GAA ACC GCG TTA AGA GGG CG-3'); PRTL-5 (5'-GAA CTT CCC AAC GAA TCC GC-3') PRTL-3 (5'-CGA TGG TGG TGG GGG AAG G-3'); RMF-5 (5'-CCG TCT CCG GTA TGT TGC C-3') RMF-3 (5'-CAG ACA GAG TAT CAA AAG GCG-3') and YHDL-5 (5'-ATG AGT CGA TAT CAG CAT ACT AAA GG-3') YHDL-3 (5'-GCT GAA AGA GCA GAA TAA CCG C-3'). PCR amplifications were carried out using chromosomal DNA as template. The purified PCR products were used as probes for all the northern blot experiments.

**Northern blot analysis.** In *E. coli* and *S. cerevisiae* the RNA was extracted following standard procedures (10) with some modifications. To remove the remaining transfer RNA, an additional precipitation step using LiCl to a final concentration of 2.0M was included. After centrifugation, RNA pellet was washed twice with ethanol 70% and resuspended in water (17). For all northern blot experiments, 10 µg of total RNA was separated by electrophoresis in formaldehyde-1.2% agarose gel and blotted onto nylon N<sup>+</sup> (Amersham). All hybridizations were carried out at high stringency conditions (65°C, PSE) (6).

**Computational Methods.** To search the protein databases, an *ad hoc* program written in Perl was used. For each sequence, the program records the amino acid composition and performs a hydrophilicity analysis based on the algorithm of Kyte and Doolittle (21), using a window size of 12 residues. The program's output is a list, sorted according to a given score, of the sequences that passed a selection criteria. The selection criteria and the scoring function can be given separately to the program at runtime. Throughout this work the selection criterion was an average hydrophilicity above 1 and at least 8% Gly; the score was defined as  $100 \times (\%Gly - 8) \times (\text{average hydrophilicity} - 1)$ . After trying different combinations of several structural and physicochemical features common to LEAPs, this criterion was chosen because it is simple and it adequately discriminates between LEAPs and other plant proteins. The score function, was chosen to give zero at the threshold values, and to grow larger as a protein becomes more distinguishable from the bulk of the non-LEAPs. Different sequence set were used in this study. To develop the selection criteria and asses its adequacy, we used a control set which includes almost all known LEAPs, 56 sequences in total. All other sets were derived from the SwissProt database (Release 36). Sequences recorded as fragments and those shorter than 30 residues were not used. To find new hydrophilins we used different taxon specific subsets of the SwissProt database. Classification into taxons was made by the Fetch program of the Wisconsin Package (Version 9.1, Genetics Computer Group (GCG), Madison, Wisc.). The groups considered were plant, bacteria (excluding *E. coli*), fungi (excluding *S cerevisiae*), *E. coli* and *S cerevisiae*. In the case of *E. coli* and *S cerevisiae*, searches were performed over hypothetically translated ORFs of the corresponding complete genomes (*Escherichia coli* K-12 MG1655 complete genome, version M52, Genbank

U00096; *Sacharomyces* Genome Database, Release 4.1). Three state secondary structure prediction was performed by the PHD sec service of the PredictProtein Server at <http://www.emblheidelberg.de/predictprotein>. The computer program and the sequences in the control set are available from the authors.



## RESULTS

**Identification of proteins with common characteristics to the LEAPs in other organisms different to plants.** Probably, arising from diverse origins, LEAPs do not share a universal sequence pattern. However, they have several distinguishing physicochemical features, mainly, high glycine content and large proportion of charged residues (13). Also, LEAPs present distinctive hydropathy profiles, which helps to recognize them among the rest of the proteins (see Table 1). Our first goal was to devise a simple but powerful selection criteria, based on the features described above, that would allow us to retrieve non-plant LEA-like proteins from available databases. The search criteria was established by searching a control set, which included most known LEAP sequences and the plant subset of the SwissProt database, with an *ad hoc* program (see Materials and Methods) that only considered glycine content and average hydrophilicity, the two straight forward properties which resulted in a high selectivity for LEAPs (see Fig. 1). Thus, in real searches only proteins having average hydrophilicity above 1 and, at least, 8% glycine content were retained. We designated this group of proteins as hydrophilins (Fig. 1). In addition to these fixed thresholds, the algorithm assigned a score to the searched proteins which allowed us to sort them by their significance. The score, defined as  $(\%Gly - 8) \times (\text{average hydrophilicity} - 1)$ , is arbitrary but has the property of being zero at the threshold values, and becoming larger as a protein moves away from the bulk of the non-LEAPs (see Fig. 1).

In the case of plant proteins, most of the known LEAPs (52 out of 57) passed the search criteria and were classified as hydrophilins, with high scores. As expected, there was a group of LEAPs that clearly separated from the rest since they do not present a high Gly content and are not specially

hydrophilic. They are considered atypical and belong to group V, such as LEA5 and LEA14 from cotton (16). For this reason, they were excluded from the control set of LEPs, as can be seen in Fig. 1. Another indication of the specificity of the search criteria was that among the total proteins analyzed from the plant set (5792) only 72 were selected, of which 52 are known LEAPs (Fig. 1). Additional hydrophilins were identified: two LEAPs not previously considered, two of the so-called COR proteins, a flocculant-active protein (MO2), a seed anti-microbial peptide (MBP1) and several ribosomal proteins (14). Interestingly, some of these proteins have been associated to water deficit situations. That is the case of the alfalfa COR410 wheat CORA proteins, which not only present homology with some LEAPs but also it is known that their genes are induced during cold and drought stress (12, 22). On the other hand, MBP1 and MO2 proteins are found in the desiccated seed (15, 18). Regarding the ribosomal proteins, although their structural characteristics could be also related to their association with ribosomes, little information is available on their expression patterns under stress conditions (p10851; p50888; p49195; p26871; p27083; p50892; p27070; p48857; p49213; p48130; p51190; p49164; p49689; p28520). When the same study was performed with the fungi and bacteria protein subsets (excluding *S. cerevisiae* and *E. coli*, see below), again a discrete group of proteins were classified as hydrophilins. In fungi, only 12 out of 2264 sequences passed the selection criterion. This group was formed by two proteins, *CON6* and *CON10*, involved in conidiation (a developmental stage where a dehydration process occurs) (28, 36), a mitochondrial (2), six ribosomal proteins (p52809; p27076; p31028; p31027; p31866; p27075) and three hypothetical proteins. In bacteria, 14 candidates out of 12911 sequences were selected. In the hydrophilins group three proteins involved in *B. subtilis* sporulation (GSIB,

SSPF and COTT) were included; one of this, GSIB, has also been found to be induced in response to different stress conditions (3, 31, 33). Also, 11 additional hydrophilins from different bacterial species were identified: 7 were hypothetical, 2 are classified as single-stranded DNA-binding proteins (p40947; p44409) and 2 as ribosomal proteins (p52829; p46386).

**Systematical search of hydrophilins in two complete genomes: *E. coli* and *S. cerevisiae*.** To reinforce the analysis described above, we studied the genomes of two of the best characterized microorganisms which, however, are phylogenetically very distant: *E. coli* and *S. cerevisiae*. For both of them, the complete genomic sequences are available. This was important because the study of complete genomes eliminates many sources of bias in the composition of the databases, specially the fact that genes tend to be isolated and studied by virtue of being related to other known genes. In *E. coli*, only five candidates were found when the selection criterion was tried against the 4289 proteins of release M52 of the complete genome, whereas the genome of *S. cerevisiae* codes for 6200 proteins, 12 of which qualified (see Table 1 and Fig. 1).

Many of the identified proteins corresponded to ORFs that had not been characterized before and whose amino acid sequences did not show significative homology among them or with other proteins in the data banks; while some others have already been described, in some cases with a particular function. In *E. coli* a specific function of 2 of the 5 selected hydrophilins has been characterized. RMF (Ribosome Modulation Factor) is a protein associated with the 100S ribosome dimers and which start to accumulate in *E. coli* cells upon growth transition from exponential to stationary phase (34, 37). PRTL has been described as a small and basic

polypeptide which is referred as the P-protein since it presents some homology with the trout protamine, however its function is still unknown (1).

In yeast, from the 12 selected hydrophilins, the genes from 3 of them have been reported to be induced by osmotic stress, that is the case of *HSP12*, *SIP18*, and *GRE1* whose function is not known. On the other hand, a functional characterization has been reported for *STF2*, *IF1A* and *RL44*. *STF2* encodes a 15 kD polypeptide that has been suggested to help stabilize the complex between F1F0-ATPase and the inhibitor protein upon cessation of phosphorylation keeping the enzyme inactive during de-energization of mitochondrial membranes (38). *IF1A* has been described to encode an eucaryotic initiation factor 1A which seems to be required for maximal rate of protein biosynthesis (35). *RL44* (YP2 $\alpha$ ) encodes for one of four ribosomal acidic phosphoproteins characterized in *S. cerevisiae* which have been involved in the regulation of the activity of the 60S ribosome subunit. It has been shown that these proteins are not required for cell viability but regulate the pattern of expression in yeast (27).

**Prediction of the hydrophilins secondary structure.** One of the striking characteristics of LEAPs is that they are predicted to be highly unstructured, existing principally as random coils. This has been confirmed in few cases by circular dichroism (23, 24). Accordingly, when the selected *E. coli* and *S. cerevisiae* hydrophilins were subjected to secondary structure prediction analysis, the results indicated that in every case they tend to adopt a coil configuration over much of their length (50-80%) (Fig. 2). In contrast, in a control group made of 89 randomly chosen *E. coli* proteins, only 6 proteins (7 %), had over 50% predicted coil regions. Besides their extreme hydrophilicity, LEAPs also have characteristic hydrophilicity profiles. As shown in Table 1, the hydrophilicity profiles of the *E. coli* and *S. cerevisiae*

candidates, and of other proteins discussed in this work, resemble those of LEAPs.

**Hydrophilins expression patterns under osmotic stress conditions.** The high correlation between the expression pattern under water limitation and the physicochemical characteristics present in different hydrophilins (including LEAPs) suggested that this could be conserved in most of them. To test this hypothesis, we carried out northern blot analysis using as probes the amplified sequences corresponding to the hydrophilins obtained from *E. coli* and *S. cerevisiae* (see section above). These probes were hybridized against total RNA obtained from *E. coli* or *S. cerevisiae* grown in optimal conditions or under hyper-osmotic stress (see Materials and Methods). The probes corresponding to *SIP18* and *HSP12* were not included in this analysis since they already have been found to be inducible by high osmoticum by other groups (25, 32). The results in Fig. 3, showed that all *E. coli* hydrophilin transcripts accumulate in response to osmotic stress. The highest accumulation was for YCIG\_ECOLI and YJBJ\_ECOLI transcripts after 20 min of hyperosmotic treatment. In both cases, we detected additional transcripts of higher molecular weight responsive to the imposed stress conditions. This observation suggested that YCIG\_ECOLI and YJBJ\_ECOLI genes are part of operons being transcribed as polycistronic mRNAs. A significant transcript accumulation in response to osmotic stress was also detected for the PRTL\_ECOLI gene whose highest accumulation occurred after 10 min of treatment (Fig. 3). Regarding the RMF\_ECOLI transcript, it was also able to respond to water deficit, however in contrast with the other hydrophilin genes, the mRNA accumulation was detected until 60 min of stress treatment (Fig. 3). The accumulation pattern corresponding

to YHDL\_ECOLI transcript was not included since its levels were undetectable by this analysis.

The analysis of the expression patterns of the 12 *S. cerevisiae* hydrophilin transcripts showed that only 8 accumulated in response to hyperosmotic conditions: S65242; STF2\_YEAST, SI18\_YEAST, YBM6\_YEAST, HS12\_YEAST, YJ04\_YEAST; YJS4\_YEAST; YNT0\_YEAST. (see Fig. 4). As indicated above, we only pursue the study of those yeast hydrophilins whose osmotic stress response has not been reported before. The results in Fig. 4 show that, as in the case of *E. coli*, while the accumulation levels of the yeast hydrophilin transcripts vary among them, their inducibility by water limitation seems to be a common characteristic. Interestingly, a correlation was found between the level of induction and the score of the *E. coli* and *S. cerevisiae* hydrophilins. In yeast, for example, the clearest induction was for *GRE1*, the protein with the highest score, while the 4 candidates that did not accumulate (YFBO\_YEAST; IF1A\_YEAST; S67772; RL44\_YEAST), all belong to the lower half of the range. This result reinforces the suitability of the search criteria to select new hydrophilins as they appear in databases.

## DISCUSSION

Many genes and/or proteins have been found to be responsive to water deficit conditions in different organisms. In higher plants the group of the so-called LEAPs present the common characteristic of being induced when a loss of water from the cell occurs (during normal desiccation of seeds, or by environmental dehydration during water, salt, and cold stress). In this work we focused our attention on the structural unifying aspects of most of these proteins: their extreme hydrophilicity and high content of Gly residues. The methodic screening that we carried out indicated, first, that proteins with these properties can be clearly distinguished from the bulk of proteins of an organism, and second, that these structural features are not only present in the LEAPs but rather they are dispersed among proteins from different organisms. This was evident from the analysis of the *E. coli* and *S. cerevisiae* genomes which showed only 5 and 12 proteins, respectively, with these characteristics. Importantly, the criterion we used in this analysis, which does not imply sequence homology and is based in two simple characteristics, is highly selective. Systematically, we found that approximately 1% of the proteins passed this criterion. From the complete OWL database, only 1,800 out of 250,000 sequences (0.72%) were selected. Of the two characteristics, the hydrophilicity threshold is by far the most selective, eliminating 19 out of 20 proteins, compared to the glycine threshold which merely eliminates 4 of every 5 proteins.

An additional conclusion from this work is the remarkable correlation, at least in plants, *E. coli* and *S. cerevisiae*, between expression under water deficit and the physicochemical characteristics of hydrophilins. The strongest correlation was found for *E. coli* where the genes encoding 4 of the 5 proteins selected as hydrophilins are induced by osmotic treatment (one

transcript was undetectable in our conditions). In the case of *S. cerevisiae*, there were some exceptions: 4 of the 12 selected hydrophilin transcripts did not accumulate after osmotic shock. Reassurance comes from the observation that on average the exceptions present lower scores than the osmotic induced hydrophilins, suggesting that they are not true hydrophilins or that they become induced under different conditions (see Fig. 4).

By addressing the question of how many of the well studied proteins in *E. coli* are known to be related to a hyperosmotic condition, the probability to which a randomly selected protein happens to be inducible by water deficit was estimated. In the *E. coli* genome, 1477 ORFs are said to be over 95% identical to SwissProt entries. Of these entries (which constitute a non-redundant set), 880 have some annotation under "Function" or "Induction", 26 of which suggest some relationship to dehydration or hyperosmosis. Thus, approximately, 3% or less of the proteins in any group are expected to be induced by osmotic shock. In contrast, the data in this work indicate that all the *E. coli* hydrophilin genes are induced under these conditions, whereas in *S. cerevisiae* 70% of the selected genes show to be responsive. Therefore, we could conclude that the association found between the selected physicochemical properties and the induction by osmotic shock cannot be given by chance.

In agreement with the transcript accumulation patterns, the promoters for some hydrophilin genes contain *cis*-elements that have been shown to participate in the activation by stress conditions. In particular, in the promoters of the yeast hydrophilin genes that were found to be induced by osmotic stress, at least one putative stress response element (STRE) (29) was identified (data not shown).



Our criterion also selected many proteins from other groups besides plants, *E. coli* and *S. cerevisiae*. Only in few cases, we found information to suggest that these candidates are induced by water deficit. This has to be considered as an incomplete list. Future releases of the SwissProt database, with more entries, and information on the function or inducibility of their sequences, will undoubtedly provide many more examples. This is true even of the plant group. The case is different for *E. coli* and *S. cerevisiae* for which their lists should be complete, since their genomes have been fully analyzed, at least in this regard.

Metazoa were not analyzed in this work because, with few exceptions, the response to dehydration has not been studied in this group. However, a preliminar examination indicates that SwissProt database has 193 metazoan proteins selected by our algorithm (data not shown). This set of proteins is very heterogeneous according to their described function. It can be distinguished toxins, DNA binding proteins, ribosomal proteins and some extracellular proteins. At present, a functional correlation with dehydration has not been established in this group.

Even though no significative homology was found among most of the different hydrophilins, common functional features could be distinguished. That is the case of those which seem to interact with nucleic acids such as protamine, the single-stranded DNA binding proteins or the ribosomal proteins that interact with RNA. Thus, it could be argued that their structural properties (presence of charged amino acids) are easily associated with their ability to bind nucleic acids, this would be the same for all DNA or RNA binding proteins; however, not all the proteins with this functional characteristic were classified as hydrophilins. For example, in *E. coli*, only one ribosomal protein, the RMF, was included in the hydrophilins group.

Interestingly, the RMF protein shows an unusual behaviour since it starts to accumulate in *E. coli* cells upon growth transition from exponential to stationary phase and, in contrast to other ribosome-components or associated proteins, the *rmf* expression depends, inversely, on growth rate (37), suggesting its participation in the bacterial response to stress conditions. Accordingly, we have found that the RMF protein plays a role in the *E. coli* adaptative response to osmotic stress because a strain lacking the RMF protein (37) showed to be osmosensitive under hyper-osmotic treatments (Garay-Arroyo *et al.*, unpublished results). Given that a significative proportion of the hydrophilins from different organisms are ribosomal proteins (organellar and cytoplasmic), it could be hipotesized that some of them were selected as part of an adaptative mechanism that help to maintain a functional translational apparatus under water limited environments.

As indicated above, the hydrophilins constitute an heterogeneous group of proteins, where, at least in the case of plant hydrophilins (mostly LEAPs), six different families can be identified based on the presence of regions of significant homology with at least another protein of the same or different species (11, 20). The analysis of the hydrophilins protein sequences has also showed that some of them could be grouped in the class 1 of LEAPs since homologous repetitive motifs are conserved, not only in all the members of this group, but also in the GSIB protein from *B. subtilis* (33) and in the *CON10* and the *YCIG\_ECOLI* proteins from *N. crassa* and *E. coli*, respectively (see Tables 1). The presence of conserved motifs among the proteins in each family suggested that they play similar functions, and probably different to those carried out by hydrophilins in other groups. Therefore, those common physicochemical properties in hydrophilins could be the result of the selection of a general solution in proteins that could have

different functions. For instance, as stabilizers of either cytosolic structures (e.g. membranes) or as macromolecular complexes (transcriptosome, ribosomes, etc).

The high content of Gly residues in most LEAPs has allowed to propose that these proteins are found in a random coiled structure (13, 14). This is consistent with experimental data obtained for some LEAPs (Em1, Dsp16 and G50) which showed a high percentage of random coiled moieties, suggesting that these proteins may be largely unfolded in a native state (23, 24). In the case of hydrophilins, the analysis of their predicted three-structure also suggests that characteristically they exist as lossed-structured proteins (see Fig. 2), in contrast with most of the proteins in an organism, which might be related to their hypothetical role in helping to maintain a minimum cellular water for survival.

If we consider that the sequence of a protein determines its function only indirectly, by guaranteeing that the protein will have the adequate structural and physicochemical characteristics, it is possible that exceptional characteristics could serve to search for the function when the proteins have no sequence similarity, due either to convergence or to very distant ancestry. In our case this was a fruitful strategy. By showing that their physicochemical characteristics are associated to a common expression pattern, our results suggest that hydrophilins represent a common adaptation to osmotic stress in the most diverse taxons.

## **ACKNOWLEDGMENTS**

We thank B. Barkla, M. Rocha, J.L. Folch, F. Campos and E. Alvarez-Buylla for critical reading of the manuscript, R. M. Solórzano for technical assistance and P. Gaytán and E. López for the synthesis of oligonucleotides. This work was supported by grants from Consejo Nacional de Ciencia y Tecnología (México) 0131P-N9506 and 26242-N to A.A.C.

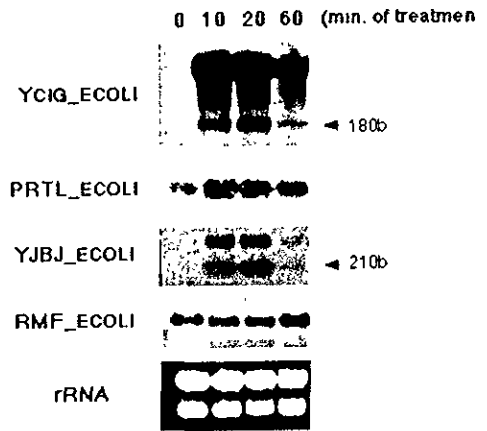
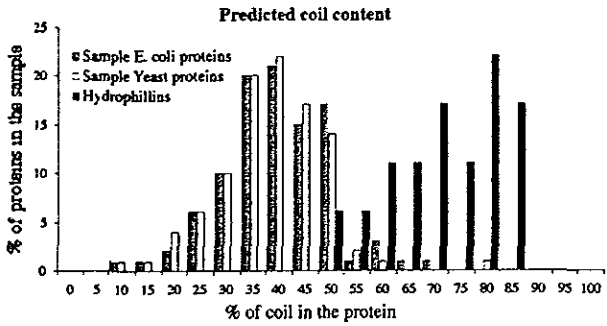
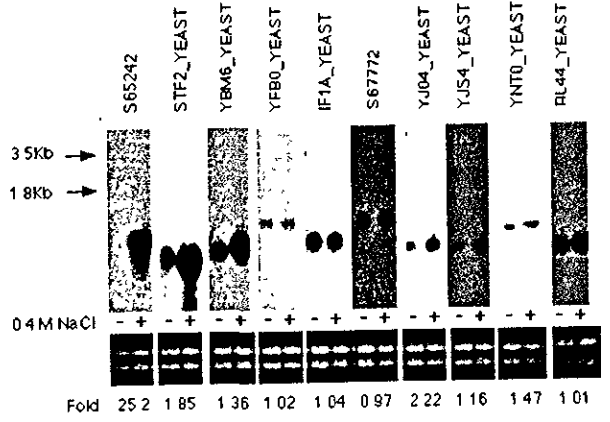
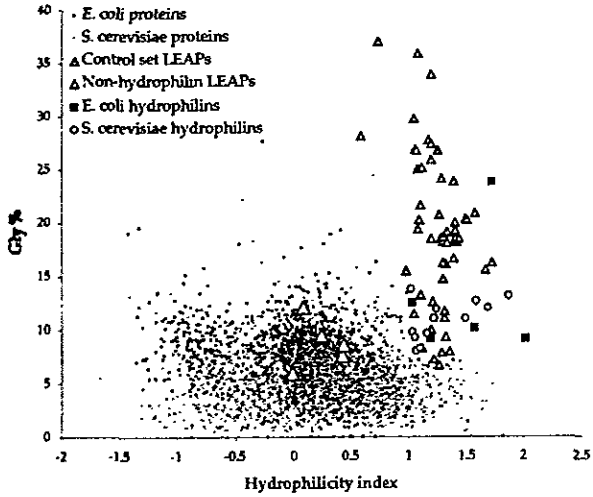


Table 1. Hydrophilins from different organisms

		Hydropathy profile	N.L.I.	%Hy (I)	%Coil (I)	Score (I)	Ref.	
BACTERIA	<i>E. coli</i>	YCG_ECOLI	1.72	23.3	88.1	3280	This work	
		PRTL_ECOLI	2.01	8.8	72.7	1032	1	
		YBJ_ECOLI	1.57	10.0	66.5	792	This work	
		YHDL_ECOLI	1.04	12.0	48.0	497	This work	
		RMF_ECOLI	1.19	8.9	62.7	263	34, 37	
	<i>B. subtilis</i>	GSIB_BACSU	1.80	15.0	49.0	1890	33	
		COTT_BACSU	1.10	11.0	81.0	438	3	
	FUNGI	<i>S. cerevisiae</i>	S65242	1.87	13.7	83.3	1652	17
			STF2_YEAST	1.68	12.0	86.7	1232	38
			SIP19_YEAST	1.58	12.6	77.2	1206	26
YBM6_YEAST			1.49	10.8	82.0	838	This work	
HS12_YEAST			1.23	11.8	60.6	675	32	
YFBO_YEAST			1.02	14.0	72.0	609	This work	
<i>S. cerevisiae</i>		IF1A_YEAST	1.22	11.0	64.7	554	35	
		S67772	1.24	9.3	58.7	358	This work	
		YJ04_YEAST	1.16	8.52	62.0	319	This work	
		YJS4_YEAST	1.03	8.52	56.0	205	This work	
		YNT0_YEAST	1.05	9.26	78.9	180	This work	
		RL44_YEAST	1.11	9.0	69.0	161	27	
		<i>N. crassa</i>	CON8_NEUCR	1.32	10.0	74.0	483	36
			CONX_NEUCR	1.17	20.0	59.0	1522	28, 30
		PLANTS	<i>L. esculentum</i>	LE19_GOSHI	1.29	15.0	55.0	1103
DH11_GOSHI	1.32			13.0	80.0	835	13	
LE7_GOSHI	1.30			8.1	24.0	256	13	
LE13_GOSHI	1.32			18.2	69.0	1614	13	
PL1EA-18	1.38			12.8	76.0	977	11	

25 aa

Table 2. Primers used to obtain *E. coli* and *S. cerevisiae* hydrophilin gene probes:

Name	Sequence (5' to 3')
YCG_ECOLI	ATG CAA CGC AAC GGG AGA AAT AAC GGG AGT TTT GCC AGT GCC CGG G
YBJ_ECOLI	CTG AAT AAA TAA TCC TAA GCC GAA ACC GCG TTA AGA GGG CG GAA CTT CCC AAC GAA TCC GC
PRTL_ECOLI	CGA TGG TGG TGG GGG AAG G CCG TCT CCG GTA TGT TGC C
RMF_ECOLI	CAG ACA GAG TAT CAA AAG GCG ATG AGT CGA TAT CAG CAT ACT AAA GG GCT GAA AGA GCA GAA TAA CCG C
S65242	CTC GAG GCC AAA CAC TAC CGC ATA AAA GC CTC GAG CAA ACT GAC ATT AAT AAA AAC G AAT CAA CAG TAA CAA ACC GC CCT TTT ATC TTG AAA CTG TGC
STF2_YEAST	ATG TGT GCT AAC GAT TAC TAC GGC GG GAA AGG CAA CTG CAA AGA ACA GAA GGG G
YFBO_YEAST	TTC GTC CAC CTG ACC GAG CCG GC CCG TCA AAT CCA TCA TCA AAT CC
IF1A_YEAST	ATG GGT AAG AAA AAC ACT AAA GGT GG GAT ATT GAG GAC ACT TGT TCG GCC
YN8D_YEAST	GCA TAT AAG GAA CCA AGG GAA AAG C TTC TCA TTA TTA TAT ACG AAT ACG
YDL_YEAST	GGA AAA AGA CCC AAC AGT GCA GCC GG GGA TGA ATT CCT GCA GGA ACA GGC
YJ04_YEAST	CAA GAA GAA GAT GTT AAG GAG GG CGT TAC AAC CCC GAA CCC GTG CG
YJS4_YEAST	GCC GAT AGT GCA CTG GAA CC CAA TCC TTC TTA ACT GTA TCC C
YNT0_YEAST	ATG AAG TTC TCT TCT GTT ACT GC CAG GAT GAA GGG GGT AGG GGG GG
RL44_YEAST	TCC TGA ATC GTA GGG CAA ATC C AAT TGC AAA GCT TGA CCC TTT TGC

## REFERENCES

1. Altman, S., Model, P., Dixon, G. H. & Wosnick, N. A. (1981) *Cell* **26**, 299-304.
2. Amegadzie, B. Y., Zitomer, R. S. & Hollenberg, C. P. (1990) *Yeast* **6**, 429-440.
3. Aronson, A. I., Song, H. Y. & Bourne, N. (1989) *Mol. Microbiol.* **3**, 437-444.
4. Baker, J., Steele, C. & Dure III, L. (1988) *Plant Mol. Biol.* **11**, 277-291.
5. Bray, E. A. (1997) *Trends Plant Sci.* **2**, 48-54.
6. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci.* **81**, 1991-1995.
7. Close, T. J. & Lambert, P. J. (1993) *Plant Physiol.* **101**, 773-779.
8. Close, T. J. (1996) *Physiol. Plant* **97**, 795-803.
9. Close, T. J. (1997) *Physiologia plantarum* **100**, 291-296.
10. Collart, M. & Olivieron, S. (1995) in *Current Protocols*, eds. Vol. pp.
11. Colmenero-Flores, J. M., Campos, F., Garcarrubio, A. & Covarrubias, A. A. (1997) *Plant Mol. Biol.* **35**, 393-405.
12. Danyluk, J., Houde, M., Rassart, E. & Sarhan, F. (1994) *FEBS Lett* **344**, 20-24.
13. Dure III, L. (1993a) in *Plant responses to cellular dehydration during environmental stress.*, eds. Close, T. J. & Bray, E. A. (Am. Soc. Plant Physiol., Rockville, MD) vol. 10, pp. 91-103.
14. Dure III, L. (1993b) *Plant J.* **3**, 363-69.
15. Duvick, J. P., Rood, T., Rao, A. G. & Marshak, D. R. (1992) *J. Biol. Chem.*

- 267**, 18814-18820.
16. Galau, G. A., Wang, H. Y. & Hughes, D. W. (1993) *Plant Physiol.* **101**, 695-696.
  17. Garay-Arroyo, A. & Covarrubias, A. A. (1999) *Yeast*
  18. Gassenschmidt, U., Jany, K. D., Tauscher, B. & Niebergall, H. (1995) *Biochim. Biophys. Acta* **1243**, 477-481.
  19. Gaxiola, R., Larrinoa, I. F., Villalba, J. M. & Serrano, R. (1992) *EMBO J* **11**, 3157-3164.
  20. Ingram, J. & Bartels, D. (1996) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 377-403.
  21. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105-132.
  22. Laberge, S., Castonguay, Y. & Vezina, L. P. (1993) *Plant Physiol.* **101**, 1411-1412.
  23. Lisse, T., Bartels, D., Kalbitzer, H. R. & Jaenicke, R. (1996) *Biol. Chem.* **377**, 555-561.
  24. McCubbin, W. D. & Kay, C. M. (1985) *Can. J. Biochem.* **63**, 803-810.
  25. Miralles, V. J. & Serrano, R. (1995) *Mol. Microbiol.* **17**, 653-662.
  26. Mtwisha, L., Brandt, W., McCready, S. & Lindsey, G. G. (1998) *Plant Mol Biol.* **37**, 513-521.
  27. Remacha, M., Jimenez-Diaz, A., Bermejo, B., Rodriguez-Gabriel, M. A., Guarinos, E. & Ballesta, J. P. G. (1995) *Molecular and Cellular Biology* **15**, 4754-4762.
  28. Roberts, A. N., Berlin, V., Hager, K. M. & Yanofsky, C. (1988) *Mol Cell Biol* **8**, 2411-2418.



29. Ruis, H. & Schüller, C. (1995) *BioEssays* **17**, 959-965.
30. Springer, M. L. & Yanofsky, C. (1992) *Genes Dev* **6**, 1052-7.
31. Stephens, M. A., Lang, N., Sandman, K. & Losick, R. (1984) *J. Mol. Biol.* **176**, 333-348.
32. Varela, J. C., van Beekvelt, C., Planta, R. J. & Mager, W. H. (1992) *Mol. Microbiol.* **6**, 2183-2190.
33. Völker, U., Egelmann, S., Maul, B., Riethdorf, S., Völker, A., Schmid, R., Mach, H. & Hecker, M. (1994) *Microbiology* **140**, 741-752.
34. Wada, A., Igarashi, K., Yoshimura, S., Aimoto, S. & Ishihama, A. (1995) *Biochem Biophys Res Commun* **214**, 410-417.
35. Wei, C. L., Kainuma, M. & Hershey, W. B. (1995) *J. Biol. Chem.* **270**, 22788-22794.
36. White, B. T. & Yanofsky, C. (1993) *Dev. Biol.* **160**, 254-264.
37. Yamagishi, M., Matsushima, H., Wada, A., Sakagami, M., Fujita, N. & Ishihama, A. (1993) *EMBO J* **12**, 625-630.
38. Yoshida, Y., Sato, T., Hashimoto, T., Ichikawa, N., Nakai, S., Yoshikawa, H., Imamoto, F. & Tagawa, K. (1990) *FEBS Lett.* **192**, 49-53.

## TABLE AND FIGURE LEGENDS

**Table 1.** Characteristics of the hydrophilins found in different organisms. The protein hydrophobic profiles were obtained using MacMolly's Translate package. **(1)** H.I.: Hydrophobic Index (see Materials and Methods, ref. 21). **(2)** %Gly: Glycine content expressed as percentage of amino acid content; **(3)** %Coil: proportion of the polypeptide predicted to be structured as a random coil; **(4)** Index of significance: corresponds to an arbitrary number that was chosen to give zero at the threshold values and which increases as a protein becomes more distinguishable from the bulk of the non-LEA proteins (see text for details).

**Figure 1.** Plot of the hydrophilicity and percentage of Gly content of some relevant protein groups. This analysis included data obtained for all *E. coli* and *S. cerevisiae* proteins as well as for LEA proteins in the control set. For technical limitations of the graphing program, only 4,000 of the approximately 6,200 proteins in *S. cerevisiae* are represented. The LEA proteins which were not classified as hydrophilins have also been incorporated. The hydrophilicity index was obtained as described in Materials and Methods (ref. 21).

**Figure 2.** Distribution of 89 randomly chosen *E. coli* and *S. cerevisiae* proteins, respectively, and the 17 *E. coli* and *S. cerevisiae* hydrophilins, according to their predicted contents of random coil structures. Three state secondary structure prediction was performed by the PHD sec service of the PredictProtein Server.

**Figure 3.** Time course of the *E. coli* hydrophilins transcript accumulation in response to osmotic stress. *E. coli* cultures were grown to an OD<sub>590nm</sub> of 0.5 and at this point the stress was imposed by adding NaCl to a final concentration of 0.4 M (see Materials and Methods for details). Cells were harvested at the times indicated and total RNA was extracted. RNA (10 µg/lane) was electrophoretically separated, blotted and hybridized against the corresponding [ $\alpha$ -<sup>32</sup>P] dCTP-labeled probes. High-stringency conditions were used for all hybridizations and washes. Ethidium bromide stained ribosomal RNAs are shown at the bottom of the figure as loading reference.

**Figure 4.** *S. cerevisiae* hydrophilin transcript accumulation in response to osmotic stress. *S. cerevisiae* cultures were grown to an OD<sub>660nm</sub> of 0.5 and at this point the stress was imposed by the addition of sorbitol to a final concentration of 0.5 M (see Materials and Methods for details). Cells were harvested after 30 min of treatment and total RNA was extracted. RNAs (10 µg/lane) were electrophoretically separated, blotted and hybridized against the corresponding [ $\alpha$ -<sup>32</sup>P]dCTP-labeled probes. High-stringency conditions were used for all hybridizations and washes. Ethidium bromide stained ribosomal RNAs are shown at the bottom of the figure as loading reference. The numbers below show the quantification of transcript accumulation when compared to the transcript levels obtained under control conditions. The quantification was done using the NIH Image 1.60b7.

## Conclusiones y discusión

**1.- Se identificaron proteínas tipo LEA que nosotros denominamos hidrofillinas tanto en un organismo procarlote, *E. coli*, como en un organismo eucariote, *S. cerevisiae*. La gran mayoría de los transcritos que codifican para estas proteínas se inducen en condiciones de estrés osmótico (100% para el caso de *E. coli* y 70% para el caso de *S. cerevisiae*) lo cual sugiere la participación de este tipo de proteínas en estas condiciones de estrés. Es importante puntualizar la alta correlación que existe, al menos en *E. coli*, *S. cerevisiae* y plantas entre la expresión bajo condiciones de baja disponibilidad de agua y las características fisicoquímicas de las hidrofillinas**

Hasta la fecha, se han caracterizado muchos genes y/o proteínas como parte de la respuesta de los organismos a condiciones de déficit hídrico. En particular, un grupo muy abundante de proteínas inducidas por condiciones de baja disponibilidad de agua son las llamadas proteínas LEA de plantas superiores, que se acumulan principalmente en semilla. Hasta el momento sólo se han reportado cuatro proteínas tipo LEA presentes en otros organismos. En tres de los casos se encontró similitud a nivel de secuencia primaria: (1) una proteína de cianobacteria que se acumula por estrés osmótico y reacciona con un anticuerpo que reconoce a las dehidrinas de plantas (Close y Lammers, 1993b); (2) se reportó similitud entre unas regiones internas de proteínas de LEAs del grupo 1 y una proteína de *Bacillus subtilis* (GsiB) que se sabe se induce bajo diferentes condiciones de estrés (incluido el estrés osmótico) (Stacy y Aalen, 1998); y (3) la HSP12 de *S. cerevisiae*. Mtwisha et al., 1998 discuten las propiedades de alta hidrofiliidad y pocos residuos hidrofóbicos que comparte la HSP12 con las proteínas LEA. El último ejemplo proviene de nuestro trabajo, en donde se caracteriza una proteína, deducida de la secuencia de nucleótidos, en base a sus características fisicoquímicas: (4) La proteína GRE1 presenta las características estructurales de las proteínas LEA y está codificada por un transcrito que se induce en condiciones de estrés osmótico (Garay-Arroyo y Covarrubias, 1999).

En este trabajo se demostró que el criterio de búsqueda fundamentado en características fisicoquímicas resultó ser altamente selectivo: sólo se encontraron 5 proteínas en *E. coli* y 12 en *S. cerevisiae*. A este nuevo grupo de proteínas se les denominó hidrofillinas por los patrones tan hidrofílicos que presentan (ver Tabla 1 artículo

anexo). Dentro del grupo de las hidrofílicas de levadura se encuentran incluidas las proteínas S65242 (GRE1), STF2\_YEAST, SI18\_YEAST, YBM6\_YEAST, HS12\_YEAST (HSP12), YFB0\_YEAST, IF1A\_YEAST, S67772, YJO4\_YEAST, YJS4\_YEAST, YNT0\_YEAST y RL44\_YEAST (ver Tabla 1). El grupo de hidrofílicas de *E. coli* está conformado por cinco miembros: YCIG\_ECOLI, PRTL\_ECOLI, YJBJ\_ECOLI, YHDL\_ECOLI y RMF\_ECOLI. Además, se demostró que gran parte de los transcritos que codifican para estas proteínas están inducidos por condiciones de estrés osmótico, lo cual sustenta la asociación encontrada entre características fisicoquímicas y respuesta a déficit hídrico.

En base a que las hidrofílicas no presentan identidad a nivel de secuencia de aminoácidos, y que muchas de ellas tienen asignada alguna función particular (en el caso de *E. coli* PRTL y RMF; en el caso de *S. cerevisiae* STF2 y IF1A), creemos que podrían estar funcionando modularmente. Es decir, estas proteínas mantienen una región hidrofílica que es la que les permite estar activas bajo condiciones de baja disponibilidad de agua pero tienen una región específica en donde se localiza algún módulo funcional. Por ejemplo, la RMF es una proteína que se asocia con los ribosomas por medio del módulo funcional, pero su parte hidrofílica le permitiría realizar sus funciones bajo condiciones de déficit hídrico.

Actualmente, se están realizando diversos ensayos, tanto *in vitro* como *in vivo*, para tratar de entender la función que llevan a cabo las hidrofílicas. Los ensayos *in vitro* se están enfocando en probar las propiedades protectoras de estas proteínas sobre la permanencia de la actividad enzimática en ambientes con baja disponibilidad de agua. Los ensayos *in vivo* se están llevando a cabo de dos maneras; por un lado haciendo mutaciones de cada una de las hidrofílicas de *E. coli* y buscando un fenotipo de sensibilidad y, por otro lado, sobreexpresando los genes (tanto en *E. coli* como en *S. cerevisiae*) y tratando de encontrar alguna condición de estrés en donde la sobreexpresión de estas proteínas otorgue un fenotipo de resistencia.

## **Bibliografia.**

1. **Albertyn J, Hohmann S, Thevelein JM, Prior BA** (1994) GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *Saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathway. *Mol. Cel. Biol.* **14**: 4135-4144
2. **Alepuz PM, Cunningham KW, Estruch F** (1997) Glucose repression affects ion homeostasis in yeast through the regulation of the stress-activated ENA1 gene. *Mol. Microbiol.* **26**: 91-8
3. **Ansell R, Granath K, Hohmann S, Thevelein J, Adler L** (1997) The two isoenzymes for yeast NAD<sup>+</sup>-dependent glycerol 3-phosphate dehydrogenase encoded by *GPD1* and *GPD2* have distinct roles in osmoadaptation and redox regulation. *EMBO J.* **16**: 2179-2187
4. **Baker J, Steele C, Dure III L** (1988) Sequence and characterization of 6 Lea proteins and their genes from cotton. *Plant. Mol. Biol.* **11**: 277-291
5. **Baker ME, Blasco R** (1992) Expansion of the mammalian 3 $\beta$ -hydroxysteroid/plant dihydroflavonol reductase superfamily to include a bacterial cholesterol dehydrogenase, a bacterial UDP-galactose 4-epimerase, and open reading frames in vaccinia virus and fish lymphocystis disease virus. *FEBS Lett.* **301**: 89-93
6. **Baker ME, Luu-The V, Simard J, Labrie F** (1990) A common ancestor for mammalian 3 $\beta$ -hydroxysteroid dehydrogenase and plant dihydroflavonol reductase. *Biochem. J.* **269**: 558-559
7. **Bartels D, Engelhardt K, Roncarati R, Schneider K, Rotter M, Salamini F** (1991) An ABA and GA modulated gene expressed in the barley embryo encodes an aldose reductase related protein. *EMBO J.* **10**(5): 1037-1043
8. **Billard P, Menart S, Fler R, Bolotin-Fukuhara M** (1995) Isolation and characterization of the gene encoding xylose reductase from *Kluyveromyces lactis*. *Gene* **162**(1): 93-97
9. **Blomberg A.** (1988). Osmoregulation and osmotolerance in yeast. Thesis.

10. **Blomberg A** (1997) Osmoresponsive proteins and functional assessment strategies in *Saccharomyces cerevisiae*. *Electrophoresis* **18**: 1429-1440
11. **Blomberg A, Adler L** (1989) Roles of glycerol and glycerol-3-phosphate dehydrogenase (NAD<sup>+</sup>) in acquired osmotolerance of *Saccharomyces cerevisiae*. *J. Bacteriol.* **171**: 1087-1092
12. **Blomberg A, Lennart A** (1992) Physiology of osmotolerance in fungi. *Advances in Microbial Physiol.* **33**: 145-212
13. **Bolen PL, Detroy RW** (1985) Induction of NADPH-linked D-xylose reductase and NAD-linked xylitol dehydrogenase activities in *Pachysolen tannophilus* by D-xylose, L-arabinose or D- galactose. *Biotech. and Bioeng.* **27**: 302-307
14. **Bolen PL, Hayman GT, Shepherd HS** (1996) Sequence and analysis of an aldose (xylose) reductase gene from the xylose-fermenting yeast *Pachysolen tannophilus*. *Yeast* **12**:1367-75
15. **Boorstein WR, Craig EA** (1990a) Regulation of a yeast HSP70 gene by a cAMP responsive transcriptional control element. *EMBO J.* **9**: 2543-2553
16. **Boorstein WR, Craig EA** (1990b) Transcriptional regulation of *SSA3*, an HSP70 gene from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10**: 3262-3267
17. **Boy-Marcotte E, Perrot M, Bussereau F, Boucherie H, Jaquet M** (1998) Msn2p and Msn4p control a large number of genes induced at the diauxic transition which are repressed by cyclic AMP in *Saccharomyces cerevisiae*. *J. Bacteriol.* **180**(5): 1044-1052
18. **Brewster JL, de Valoir T, Dwyer ND, Winter E, Gustin MC** (1993) An osmosensing signal transduction pathway in yeast. *Science* **259**: 1760-1763
19. **Brewster JL, Gustin MC** (1994) Positioning of cell growth and division after osmotic stress requires a MAP kinase pathway. *Yeast* **10**: 425-439
20. **Casamayor A, Khalid H, Balcells L, Aldea M, Casas C, Herrero E, Arino J** (1996) Sequence analysis of a 13.4 kbp fragment from the left arm of chromosome XV reveals a malate dehydrogenase gene, a putative Ser/Thr protein kinase, the ribosomal L25 gene and four new open reading

frames. *Yeast* **12**(10B): 1013-20

21. **Chowdhury S, Smith KW, Gustin MC** (1992) Osmotic stress and the yeast cytoskeleton: phenotype-specific suppression of an actin mutation. *The J. Cell Biol.* **118**: 561-571
22. **Close TJ** (1996) Dehydrins: emergence of a biochemical role of a family of plant dehydration proteins. *Physiol. Plant.* **97**: 19-27
23. **Close TJ, Fenton RD, Yang A, Asghar R, DeMason DA** (1993a) Dehydrin: the protein. *In* TJ Close, EA Bray, ed. *Plant responses to cellular dehydration during environmental stress.* Am. Soc. Plant Physiol., Rockville, MD.
24. **Close TJ, Lammers PJ** (1993b) An osmotic stress protein from cyanobacteria is immunologically related to plant dehydrins. *Plant Mol. Biol.* **25**: 479-492
25. **Colmenero-Flores JM, Moreno LP, Smith CE, Covarrubias AA** (1999) Pvlea-18, a member of a new late-embryogenesis-abundant protein family that accumulates during water stress and in the growing regions of well-irrigated bean seedlings. *Plant Physiol* **120**(1): 93-104
26. **Dure III L** (1993) A repeating 11-mer amino acid motif and plant dessication. *Plant J.* **3**: 363-369
27. **Ferrigno P, Posas F, Koepf D, Saito H, Silver PA** (1998) Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin beta homologs NMD5 and XPO1. *EMBO J.* **17**: 5606-14
28. **Flattery-O'Brien JA, Grant CM, Dawes IW** (1997) Stationary-phase regulation of the *Saccharomyces cerevisiae* SOD2 gene is dependent on additive effects of HAP2/3/4/5- and STRE-binding elements. *Mol. Microbiol.* **23**: 303-12
29. **Galau GA, Wang HY, Hughes DW** (1993) Cotton Lea5 and Lea14 encode atypical late embryogenesis-abundant proteins. *Plant Physiol.* **101**: 695-6
30. **Garay-Arroyo A, Covarrubias AA** (1999) Three genes whose expression is induced by different stress conditions in *S. cerevisiae*. *Yeast* (en prensa).



31. **Goday A, Jensen AB, Cullianez-Macia FA, Mar Alba M, Figueras M, Serratos J, Torrent M, Pages M** (1994) The maize abscisic acid-responsive protein Rab17 is located in the nucleus and interacts with nuclear localization signals. *Plant Cell* **6**: 351-360
32. **Godon Ch, Lagniel G, Lee J, Buhler JM, Kieffer S, Perrot M, Boucheire H, Toledano MB, Labarre J** (1998) The H<sub>2</sub>O<sub>2</sub> stimulon in *Saccharomyces cerevisiae*. *J Biol. Chem.* **273**: 22480-22489
33. **Görner W, Durchschlag E, Martinez-Pastor MT, Estruch F, Ammerer G, Hamilton B, Ruis H, Schuller C** (1998) Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. *Genes Dev.* **12**: 586-97
34. **Hirayama T, Maeda T, Saito H, Shinozaki K** (1995) Cloning and characterization of seven cDNAs for hyperosmolarity-responsive (HOR) genes of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **249**: 127-138
35. **Hohmann S** (1997) Shaping up: The response of yeast to osmotic stress. *In* S Hohmann, WH Mager, ed. *Yeast stress responses*. Landes Company, Texas, USA.
36. **Imai R, Chang L, Ohta A, Bray EA, Takagi M** (1996) A lea-class gene of tomato confers salt and freezing tolerance when expressed in *Saccharomyces cerevisiae*. *Gene* **170**: 243-248
37. **Ingram J, Bartels D** (1996) The molecular basis of dehydration tolerance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**: 377-403
38. **Inoue Y, Tsujimoto Y, Kimura A** (1998) Expression of the glyoxalase I gene of *Saccharomyces cerevisiae* is regulated by high osmolarity glycerol mitogen-activated protein kinase pathway in osmotic stress response. *J. Biol. Chem.* 1998 **273**: 2977-83
39. **Jackson S, Gascon J, Carrera E, Monte E, Prat S** (1997) Cloning and expression analysis of a gene that shows developmental regulation upon tuberization in potato. *Plant Mol. Biol.* **33**: 169-174
40. **Jacoby T, Flanagan H, Faykin A, Seto AG, Mattison C, Ota I** (1997) Two protein-tyrosine phosphatases inactivate the osmotic stress response

pathway in yeast by targeting the mitogen-activated protein kinase, Hog1. J. Biol. Chem. **272**: 17749-55

41. **Kazuoka T, Oeda K** (1994) Purification and characterization of COR85-oligomeric complex from cold acclimated spinach. Plant Cell Physiol. **35**: 601-611

42. **Kobayashi N, McEntee K** (1993) Identification of *cis* and *trans* components of a novel heat shock stress regulatory pathway in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **13**: 248-256

43. **Larsson K, Ansell R, Eriksson P, Adler L** (1993) A gene encoding sn-glycerol-3-phosphate dehydrogenase (NAD<sup>+</sup>) complements an osmosensitive mutant of *Saccharomyces cerevisiae*. Mol. Microbiol. **10**: 1101-1111

44. **Latterich M, Watson MD** (1991) Isolation and characterization of osmosensitive vacuolar mutants of *Saccharomyces cerevisiae*. Mol. Microbiol. **5**: 2417-2426

45. **Latterich M, Watson MD** (1993) Evidence for a dual osmoregulatory mechanism in the yeast *Saccharomyces cerevisiae*. Biochem. Biophys. Res. Com. **191**: 1111-1117

46. **Luyten K, Albertyn J, Fourie Skibbe W, Prior BA, Ramos J, Thevelein JM, Hohmann S** (1995) Fps1, a yeast membrane of the MIP family of channel proteins, is a facilitator for glycerol uptake and efflux and is inactive under osmotic stress. EMBO J. **14**: 1360-1371

47. **Maeda T, Wurgler-Murphy SM, Saito H** (1994) A two-component system that regulates and osmosensing MAP kinase cascade in yeast. Nature **369**: 242-245

48. **Marchler G, Schüller C, Adam G, Ruis H** (1993) A *Saccharomyces cerevisiae* UAS element controlled by protein kinase A activates transcription in response to a variety of stress conditions. EMBO J. **12**: 1997-2003

49. **Márquez JA, Pascual-Ahuir A, Proft M, Serrano R** (1998) The Ssn6-Tup1 repressor complex of *Saccharomyces cerevisiae* is involved in the osmotic induction of HOG-dependent and -independent genes. EMBO J. **17**: 2543-2553

50. **Márquez JA, Serrano R** (1996) Multiple transduction pathways regulate the sodium-extrusion gene PMR2/ENA1 during salt stress in yeast. *FEBS Lett.* **382**: 89-92
51. **Martínez-Pastor MT, Marchler G, Schüller C, Marchler-Bauer A, Ruis H, Estruch F** (1996) The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress-response element (STRE). *EMBO J.* **15**: 227-2235
52. **McCubbin WD, Kay CM** (1985) Hydrodynamics and optical properties of the wheat Em protein. *Can. J. Biochem.* **63**: 803-810
53. **Miralles VJ, Serrano R** (1995) A genomic locus in *Saccharomyces cerevisiae* with four genes up-regulated by osmotic stress. *Mol. Microbiol.* **17**: 653-662
54. **Morris GJ, Winters L, Coulson GE, Clarke KJ** (1986) Effect of osmotic stress on the ultrastructure and viability of the yeast *Saccharomyces cerevisiae*. *J. Gen. Microbiol.* **11**: 2023-34
55. **Mtwisha L, Brandt W, McCreedy S, Lindsey GG** (1998) HSP12 is a LEA-like protein in *Saccharomyces cerevisiae*. *Plant Mol. Biol.* **37**: 513-521
56. **Niogret MF, Cullianez-Macia FA, Goday A, Mar Alba M, Pages M** (1996) Expression and cellular localization of rab28 mRNA and Rab28 protein during maize embryogenesis. *Plant J.* **9**: 549-557
57. **Norbeck J, Blomberg A** (1997) Metabolic and regulatory changes associated with growth of *Saccharomyces cerevisiae* in 1.4M NaCl. *J. Biol. Chem.* **272**: 5544-5554
58. **Norbeck J, Pahlmann AK, Akhtar N, Blomberg A, Adler L** (1996) Purification and characterization of two isoenzymes of DL-glycerol 3-phosphatase from *Saccharomyces cerevisiae*. Identification of the corresponding *GPP1* and *GPP2* genes and evidence for osmotic regulation of Gpp2p expression by the osmosensing MAP kinase signal transduction pathway. *J. Biol. Chem.* **271**: 13875-13881
59. **Posas F, Saito H** (1997) Osmotic activation of the HOG MAPK pathway via Ste11p MAPKKK: Scaffold role of Pbs2p MAPKK. *Science* **276**: 1702-1705

60. **Proft M, Serrano R** (1999) Repressors and upstream repressing sequences of the stress-regulated *ENA1* gene in *Saccharomyces cerevisiae*: bZIP proteins Sko1p confers HOG-dependent osmotic regulation. *Mol. Cell. Biol.* **19**: 537-546
61. **Roberts JK, DeSimone NA, Lingle WL, Dure III L** (1993) Cellular concentrations and uniformity of cell-type accumulation of two lea proteins in cotton embryos. *Plant Cell.* **5**: 769-780
62. **Ruis H, Schüller C** (1995) Stress signaling in yeast. *BioEssays* **17**: 959-965
63. **Schmitt AP, McEntee K** (1996) Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **93**: 5777-5782
64. **Schüller C, Brewster JL, Alexander MR, Gustin MC, Ruis H** (1994) The HOG pathway controls osmotic regulation of transcription via the stress response element (STRE) of the *Saccharomyces cerevisiae* CTT1 gene. *EMBO J.* **13**: 4382-4389
65. **Serrano R** (1996) Salt tolerance in plants and microorganisms: toxicity targets and defense responses. *Int. Rev. Cytol.* **165**: 1-52
66. **Shen Q, Chen CN, Brands A, Pan SM, Ho TD** (1997) Developmental regulation and potential function of a barley stress-induced gene whose homologs are present in diverse organisms including human and yeast. The American Society of Plant Physiologists. **Abstract number**: 234
67. **Stacy RA, Aalen RB** (1998) Identification of sequence homology between the internal hydrophilic repeated motifs of group 1 late-embryogenesis-abundant proteins in plants and hydrophilic repeats of the general stress protein GsiB of *Bacillus subtilis*. *Planta* **206**: 476-478
68. **Sutherland FC, Lages F, Lucas C, Luyten K, Albertyn J, Hohmann S, Prior BA, Killian SG** (1997) Characteristics of Fps1-dependent and -independent glycerol transport in *Saccharomyces cerevisiae*. *J. Bacteriol.* **179**: 7790-7795
69. **Taiz L, Zeiger E** (1991) *Plant Physiology*. Editorial: The Benjamin/Cummings Publishing Company, Inc. 565 pp.

70. **Takuma S, Nakashima N, Tantirungkij M, Kinoshita S, Okada H, Seki T, Yoshida T** (1991) Isolation of xylose reductase gene of *Pichia stipitis* and its expression in *Saccharomyces cerevisiae*. *Appl. Biochem. Biotechnol.* **28-29**: 327-40
71. **Thevelein JM** (1994) Signal transduction in yeast. *Yeast* **10**: 1753-1790
72. **Treger JM, Schmitt AP, Simon JR, McEntee K** (1998) Transcriptional factor mutations reveal regulatory complexities of heat shock and newly identified stress genes in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **273**: 26875-9
73. **Van der WE, de Almeida JE, Inzé D, Van Montagu M, Gheysen G** (1996) A molecular study of root-knot nematode-induced feeding sites. *Plant J.* **9**: 45-54
74. **Van zyl C, Prior BA, Kilian SG, Kock JLF** (1989) D-xylose utilization by *Saccharomyces cerevisiae*. *J. Gen. Microbiol.* **135**: 2791-2798
75. **Varela JCS, Mager WH** (1996) Response of *Saccharomyces cerevisiae* to changes in external osmolarity. *Microbiol.* **142**: 721-731
76. **Varela JCS, Praekelt UM, Meacock PA, Planta RJ, Mager WH** (1995) The *Saccharomyces cerevisiae* *HSP12* gene is activated by the high-osmolarity glycerol pathway and negatively regulated by protein kinase A. *Mol. Cell. Biol.* **15**: 6232-6245
77. **Walters C, Walker-Simmons MK, Ried JL** (1997) Heat-soluble proteins extracted from wheat seed embryos have tightly bound sugars and unusual hydration properties. The American Society of Plant Physiologists **Abstract number**: 503
78. **Werner-Washburne M, Braun E, Johnston GC, Singer RA** (1993) Stationary phase in the yeast *Saccharomyces cerevisiae*. *Microbiol. Rev.* **57**: 383-401
79. **Xu D, Duan X, Wang B, Hong B, David Ho T-H, Wu R** (1996) Expression of a Late Embryogenesis Abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiol.* **110**: 249-257

80. **Yamamoto KT, Mori H, Imaseki H** (1992) Novel mRNA sequences induced by indole-3-acetic acid in sections of elongating hypocotyls of mung bean (*Vigna radiata*). *Plant Cell Physiol.* **33**: 13-20
81. **Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN** (1982) Living with water stress: evolution of osmolyte systems. *Science* **217**:(1214-1222)