



# UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## Maestría y Doctorado en Ciencias Bioquímicas

Caracterización de la sensibilidad estructural al ambiente osmótico de regiones intrínsecamente desordenadas *in vivo*

TESIS

QUE PARA OPTAR POR EL GRADO DE:  
Maestro en Ciencias

PRESENTA:  
Constanza Enriquez Toledo

TUTOR PRINCIPAL:  
Dr. César Luis Cuevas Velázquez  
[Facultad de Química](#)

MIEMBROS DEL COMITÉ TUTOR:  
Dr. Alejandro Sosa Peinado  
[Facultad de Medicina](#)

Dr. Takuya Nishigaki  
[Instituto de Biotecnología](#)

Ciudad de México. Diciembre, 2022



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

## **AGRADECIMIENTOS ACADÉMICOS**

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el donativo de Ciencia de Frontera CONACYT con la clave CF-2019/252952 al Dr. César Luis Cuevas Velázquez.

A la Universidad Nacional Autónoma de México y al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la Beca Nacional para Estudios de Posgrado (CVU 1083636) que recibí. Así como el apoyo del “Programa de Apoyo a los Estudios del Posgrado” (PAEP) para asistir a congresos.

Agradezco a los miembros del comité tutor por las valiosas sugerencias, comentarios: Dr. Alejandro Sosa Peinado de la Facultad de Medicina y el Dr. Takuya Nishigaki del Instituto de Biotecnología.

Agradezco a los miembros del jurado por el tiempo para revisar y discutir el presente trabajo de tesis: Dr. Daniel Alejandro Fernández Velasco, Dra. Rosana Sánchez López, Dr. Luis Cardenas Torres, Dr. Enrique García Hernández y Dr. José Luis Reyes Taboada.

A la Facultad de Química, Departamento de Bioquímica, Conjunto E de la Universidad Nacional Autónoma de México.

Este proyecto se realizó en el laboratorio 103 del Conjunto E de la Facultad de Química, bajo la dirección del Dr. César Luis Cuevas Velázquez.

## Índice

Abreviaciones	4
Resumen	6
Abstract	7
I. Introducción	8
1.1 Proteínas y regiones intrínsecamente desordenadas	8
1.2 La sensibilidad estructural de las regiones intrínsecamente desordenadas al ambiente fisicoquímico	9
1.3 Transferencia de energía de resonancia de fluorescencia (FRET)	10
1.4. Osmolaridad intracelular	14
1.5 Estrés hiperosmótico	14
1.6 <i>Saccharomyces cerevisiae</i>	15
1.6.1 <i>Saccharomyces cerevisiae</i> como organismo modelo	15
1.6.2 Respuesta al estrés osmótico en <i>Saccharomyces cerevisiae</i>	17
II. Antecedentes	19
III. Hipótesis	21
IV. Objetivo general	21
V. Objetivos particulares	21
VI. Materiales y métodos	22
6.1 Caracterización <i>in silico</i> de las secuencias desordenadas	22
6.2 Síntesis de las IDRs	23
6.3 Expresión de las IDRs en la levadura	24
6.4 Ensayos FRET	25
VII. Resultados	26
7.1 Construcción de la base de datos con las secuencias de IDRs candidatas	26
7.2 Identificación de regiones intrínsecamente desordenadas con el potencial de formar hélice $\alpha$	27
7.3 Establecimiento del parámetro de dualidad y calibración	28
7.4 Selección de tres regiones intrínsecamente desordenadas con dualidad	29
7.4.1 La proteína de unión a elementos sensible a AMP cíclico 1 (CREB1)	31
7.4.2 La proteína componente esencial del pirenoide 1 (EPYC1)	32

7.4.3 El factor de transcripción U1-type domain-containing protein (U1TDCP)	33
7.5 Análisis de la dualidad de regiones intrínsecamente desordenadas seleccionadas	34
7.6 Caracterización de las propiedades fisicoquímicas de las regiones intrínsecamente desordenadas con CIDER	35
7.7 Expresión de las regiones intrínsecamente desordenadas de CREB1, EPYC1 y U1TDCP en levadura	38
7.8 Determinación de la sensibilidad estructural de las regiones intrínsecamente desordenadas de CREB1, EPYC1 y U1TDCP	41
VIII. Discusión	45
IX. Conclusión	48
IX. Perspectivas	49
IX. Referencias	50
X. Anexos metodológicos	57
10.1 Protocolo de transformación por choque térmico de levadura BJ5465	57
10.2 Protocolo de ensamblaje Gibson Assembly Master Mix	58
10.3 Condiciones de PCR en levadura para amplificación de construcciones IDR-CREB1, IDR-EPYC1, IDR-U1TDCP.	58
10.4 Protocolo para determinación de la sensibilidad estructural mediante FRET <i>in vivo</i>	58

## **Abreviaciones**

ABP: Proteína de unión a arabinosa. Del inglés *Arabinose-Binding Protein*.

CREB1: Proteína de unión a elementos sensibles a AMP cíclico. Del inglés *cAMP Response Element-Binding Protein 1*.

CIDER: *Classification of Intrinsically Disordered Ensemble Regions*.

DC: Dicroísmo circular.

DxAm: Emisión del aceptor tras la excitación del donador.

DxAm: Emisión del donador tras la excitación del aceptor.

EPYC1: Proteína componente esencial del pirenoide 1. Del inglés *Essential Pyrenoid Component 1*.

FRET: Transferencia de energía de resonancia de fluorescencia. Del inglés *Fluorescence Resonance Energy Transfer*.

FP: Proteínas fluorescentes. Del inglés *Fluorescent Proteins*.

HOG: Glicerol de alta osmolaridad. Del inglés *High Osmolarity Glycerol*.

IDPs: Proteínas intrínsecamente desordenadas. Del inglés *Intrinsically Disordered Proteins*.

IDRs: Regiones intrínsecamente desordenadas. Del inglés *Intrinsically Disordered Regions*.

LEA: Proteínas abundantes en la embriogénesis tardía. Del inglés *Late Embryogenesis Abundant*.

LLPS: Separación de fase líquido-líquido. Del inglés *Liquid-liquid Phase Separation*.

MAPK: Proteína cinasa activada por mitógenos. Del inglés *Mitogen Activated Protein Kinase*.

NFAT5: Factor de transcripción de células T 5. Del inglés *Nuclear Factor of Activated T-cells 5*.

NMR: Resonancia magnética nuclear.

ORF: Marco de lectura abierto. Del inglés *Open Reading Frame*.

PCR: Reacción en cadena de la polimerasa. Del inglés *Polymerase Chain Reaction*.

PEG: Polietilenglicol.

PM: Marcador de peso molecular.

PONDR: Del inglés *Predictor Of Natural Disordered Regions*.

rpm: Revoluciones por minuto.

SAXS: Dispersión de rayos X por ángulo pequeño.

SED1: Del inglés *Sensor Expressing Disordered protein 1*.

FT: Factor de transcripción.

U1TDCP: Proteína que contiene dominio tipo U1. Del inglés *U1-type Domain-Containing Protein*.

## Resumen

Las regiones intrínsecamente desordenadas (IDRs) son ubicuas en los proteomas de todos los dominios de la vida y desempeñan papeles esenciales en la función celular. A diferencia de las proteínas globulares, las IDRs no presentan una estructura tridimensional bien definida y se comportan como un ensamble de conformaciones altamente dinámicas y cambiantes. Debido a estas características las IDRs se exponen ampliamente a la composición del medio, siendo susceptibles a presentar sensibilidad estructural a los cambios en el ambiente fisicoquímico circundante.

En algunos casos, las IDRs pueden plegarse a conformaciones de tipo hélice  $\alpha$  en respuesta a los cambios en el ambiente fisicoquímico o bien a la interacción con otras proteínas. A pesar de que la sensibilidad estructural al ambiente podría desempeñar un papel relevante en las funciones de las IDRs, solamente un número limitado de ellas han sido estudiadas *in vitro* y en menor medida caracterizadas en un contexto celular.

En este trabajo se realizó la caracterización de la sensibilidad estructural de tres IDRs en respuesta a cambios en el ambiente osmótico de las células vivas de la levadura *Saccharomyces cerevisiae*. Se estudió a la proteína de unión al elemento sensible al AMP cíclico 1 (CREB1) de rata (*Rattus norvegicus*), la proteína componente esencial del pirenoide 1 (EPYC1) de alga (*Chlamydomonas reinhardtii*) y la proteína que contiene dominio tipo U1 (U1TDCP) de planta (*Arabidopsis thaliana*). Las tres IDRs se predijeron como intrínsecamente desordenadas, sin embargo, también mostraron niveles significativos de predicción de hélice  $\alpha$ . Mediante el sistema de transferencia de energía de resonancia de Förster (FRET) se estudió el efecto del incremento en la osmolaridad del medio extracelular sobre la transición a estados compactos de las IDRs *in vivo*, a través de los tratamientos de choque hiperosmótico inducido por el aumento en las concentraciones de cloruro de sodio (NaCl).

Las IDRs CREB1, EPYC1 y U1TDCP, clonadas en el sistema FRET y expresadas en las células de *Saccharomyces cerevisiae* mostraron un incremento en la razón FRET *in vivo* en función de la osmolaridad inducida por el incremento en la concentración extracelular de NaCl. Los resultados indican que la estructura de las tres IDRs estudiadas se compacta ante los efectos del incremento en la osmolaridad inducida por NaCl, lo que proporciona nueva evidencia sobre la sensibilidad estructural al ambiente fisicoquímico *in vivo* de las IDRs de diferentes organismos.

## Abstract

Intrinsically disordered regions (IDRs) are ubiquitous in all proteomes of life and play essential roles in cell function. In contrast with globular proteins, IDRs do not present a tridimensional structure and they behave like an ensemble of changing highly dynamic conformations. Due to these features, IDRs are exposed to the media composition and possibly susceptible to structural sensitivity to changes in the chemical composition.

In some cases, the IDRs can fold in more ordered conformations like alpha helix conformation in response to changes in the physicochemical environment or by interactions with interaction partners. Despite the structural sensitivity to the environment, a feature of IDRs that could play a crucial role in their function, only a limited number of IDRs have been studied *in vitro* and a reduced number of them have been characterized in a cellular context.

This work characterized the structural sensitivity of IDRs of cAMP Response Element-Binding Protein 1 (CREB1) of the rat *Rattus norvegicus*, IDR of Essential Pyrenoid Component 1 (EPYC1) of algae *Chlamydomonas reinhardtii*, and IDR *U1-type Domain-Containing Protein* (U1TDCP) of plant *Arabidopsis Thaliana*, in response to changes in the osmotic environment of the yeast *Saccharomyces cerevisiae* living cells.

All IDRs were predicted as disordered and at the same time, they show high prediction for alpha-helix formation. Using a Foster Resonance Energy Transfer (FRET) system we were able to analyze the extended to a compact state transition of IDR *in vivo* following the increase of osmolarity caused by the increment of sodium chloride concentrations. The three IDRs analyzed using the FRET system displayed an increment of FRET ratio *in vivo* due to the treatments with sodium chloride. These results indicate that the global dimensions of the IDR were reduced in response to osmolarity increase due to NaCl. These results provide evidence of structural sensitivity to the physicochemical environment *in vivo* of IDRs from different organisms.

## I. Introducción

### 1.1 Proteínas y regiones intrínsecamente desordenadas

Las proteínas intrínsecamente desordenadas (IDPs, del inglés *Intrinsically Disordered Proteins*) evidencian que la adquisición de una estructura tridimensional marginalmente estable no es un prerrequisito para la funcionalidad de una proteína. Las IDPs, a diferencia de las proteínas globulares, son capaces de llevar a cabo una o más funciones a pesar de carecer de una estructura terciaria o secundaria estable en su estado nativo (van der Lee et al., 2014). Las IDPs pueden ser completamente desordenadas o contener segmentos desordenados denominados como regiones intrínsecamente desordenadas (IDRs, del inglés *Intrinsically Disordered Regions*). Las IDRs son segmentos de polipéptidos  $\geq 30$  aminoácidos desprovistos total o parcialmente de una estructura tridimensional única y en su lugar presentan un rango de distintas conformaciones en equilibrio dinámico.

La identificación del desorden estructural presente en las IDRs permitió el desarrollo de múltiples algoritmos bioinformáticos con la capacidad de predecir el desorden a partir de la secuencia de aminoácidos (Tompa 2012). Basados en estos predictores, se estima que el desorden estructural es abundante en los proteomas de todos los dominios de la vida. Se ha predicho que el 2.0% de las proteínas de arqueas, el 4.2% de las proteínas de eubacterias contienen regiones desordenadas. Y de forma considerablemente mayor, el 33.0% del total de las proteínas presentes en el proteoma de los eucariotas es desordenado (Ward et al., 2004). Particularmente en la levadura *Saccharomyces cerevisiae*, del 50 al 60% de las proteínas contienen al menos una región desordenada (Tompa et al., 2006). Y en el humano el desorden estructural representa aproximadamente del 30% al 43.6% del proteoma (Dyson y Wright 2015; Piovesan et al., 2021).

En los últimos veinte años y tras múltiples esfuerzos por develar la función del desorden estructural, se ha demostrado que las IDRs pueden mediar múltiples procesos celulares desde la regulación de la transcripción, la traducción, las redes de señalización celular y la formación de organelos sin membrana (Uversky 2017; Soranno 2020). A diferencia de las proteínas plegadas, donde las interacciones proteína-proteína se logran a través de interfaces estables y definidas, la plasticidad estructural de las IDRs les permite interactuar con múltiples ligandos a las que se unen con baja afinidad, pero conservando una alta especificidad (Dyson y Wright 2005; Oldfield y Dunker 2014). Durante esta interacción las IDRs pueden

modificar su preferencia conformacional transitando al orden o también a una conformación más extendida después de unirse a sus blancos (Tompa y Fuxreiter, 2008). Aunque esta respuesta es variada, la flexibilidad estructural de las IDR's ante ciertos elementos de interacción o el ambiente fisicoquímico subyace a la secuencia de aminoácidos (Das et al., 2015).

Las secuencias de las IDR's presentan una alta proporción en residuos cargados o polares (Arg, Pro, Gln, Ser, Glu y Lys) y de aminoácidos pequeños (Gly, Ala) que promueven el desorden estructural (Dyson y Wright, 2005). Y una baja abundancia o ausencia de aminoácidos hidrofóbicos (Ile, Leu, Val) y voluminosos (Trp, Tyr, Phe) promotores del orden estructural (Uversky, 2014).

Como se mencionó con anterioridad, las IDR's pueden adquirir una estructura más ordenada. Este comportamiento es desencadenado generalmente por dos componentes, la presencia de otros elementos de interacción y el ambiente (Uversky y Dunker, 2010). Algunas IDR's adquieren una forma más estable tras la unión a un ligando, el costo de la entropía necesaria para el plegamiento lo paga la entalpía de unión a su ligando (Dyson y Wright, 2005). Por otro lado, las IDR's poseen un área superficial extensa que las pone en contacto con el medio circundante siendo altamente expuestas a los cambios en las condiciones del ambiente fisicoquímico como la osmolaridad, el pH y el hacinamiento molecular, los cuales pueden inducir cambios conformacionales en las regiones intrínsecamente desordenadas (Uversky 2009; Wicky et al., 2017; Cuevas-Velazquez et al., 2021; Gutierrez et al., 2022).

## **1.2 La sensibilidad estructural de las regiones intrínsecamente desordenadas al ambiente fisicoquímico**

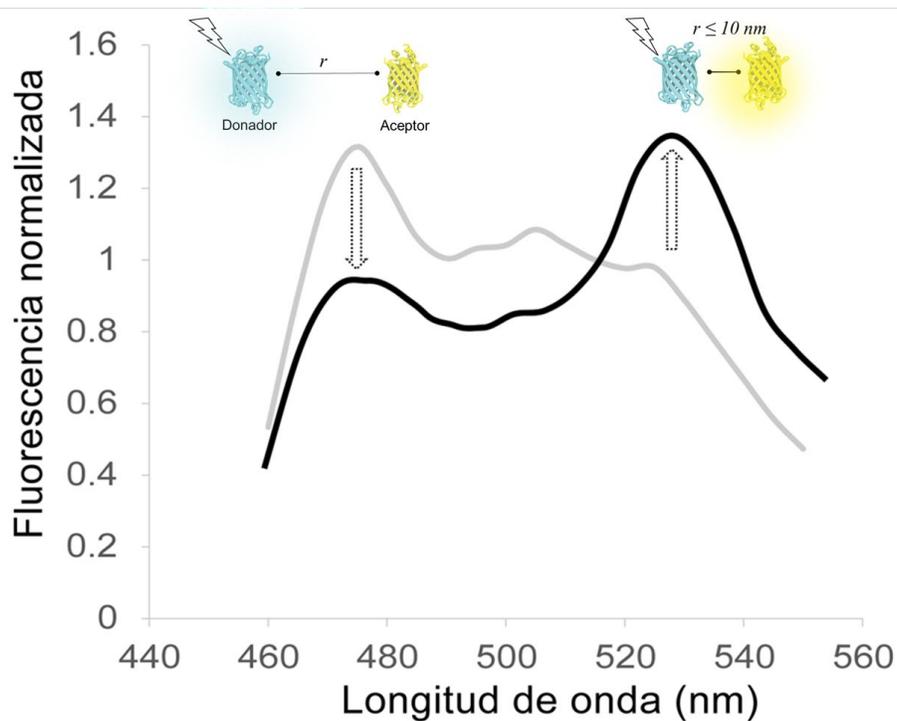
La sensibilidad estructural al ambiente fisicoquímico es una propiedad de muchas IDR's (Holehouse y Sukenink 2020; Moses et al., 2020). Las alteraciones en las propiedades de la solución que rodea a las IDR's pueden generar alteraciones en su radio de giro y por ende en sus dimensiones (Wicky et al., 2017). Estas transiciones conformacionales de las IDR's causadas por los cambios en el ambiente fisicoquímico pueden explicarse, en parte, por la secuencia de aminoácidos y la identidad del solvente. Las interacciones residuo-residuo y residuo-solvente que se establecen dependen del contenido de carga electrostática y su distribución a lo largo de la IDR (Das y Pappu en 2013; Holehouse et al., 2017). Adicionalmente, se ha observado que la baja hidrofobicidad, y el contenido de prolina influyen en el estado conformacional de

las IDR's (Taneja y Holehouse 2021). Por ello, la sensibilidad estructural de las IDR's, demostrada *in vitro* y estimada *in silico* es diversa (Moses et al., 2020).

La sensibilidad estructural de las IDR's a las fluctuaciones en el ambiente físicoquímico ha sido caracterizada mayormente *in vitro*. A través de estos ensayos se ha demostrado que las IDR's pueden compactarse ante el cambio en ciertas propiedades del ambiente físicoquímico, por ejemplo, el incremento en la osmolaridad del medio. Condición en la que experimentan cambios transitorios hacia una estructura secundaria, como la formación de hélices  $\alpha$  (Cuevas-Velazquez et al., 2016). No obstante, el contexto en el que se ha estudiado la sensibilidad estructural de las IDR's está alejado de representar a las condiciones presentes dentro de las células. Debido a su flexibilidad inherente, las IDR's también pueden presentar una sensibilidad estructural al entorno celular, donde hasta el 40% del volumen total celular está ocupado por otras macromoléculas. En estas condiciones, los efectos del volumen excluido llevan a ciertas IDR's a compactarse en presencia de altas concentraciones de agentes hacinantes (Fonin et al., 2018). Una de las técnicas más utilizadas para estimar el cambio estructural de las IDR's en el entorno celular en respuesta a cambios en el ambiente físicoquímico es la transferencia de energía de resonancia de fluorescencia (FRET) (Soranno et al., 2014; Basak et al., 2021).

### **1.3 Transferencia de energía de resonancia de fluorescencia (FRET)**

El fenómeno de transferencia de energía de resonancia de fluorescencia (FRET) entre una molécula donadora fluorescente de alta energía y una molécula aceptora fluorescente de menor energía, se refiere a la transferencia de energía no radiante y dependiente de la distancia ( $r$ ) de 1 a 10 nm que ocurre tras la excitación del donador hacia el aceptor (Campbell 2009). La transferencia de energía resulta en la disminución del espectro de emisión de fluorescencia que corresponde al donador tras su excitación ( $DxDm$ ) al mismo tiempo que el espectro de fluorescencia de aceptor aumenta ( $DxAm$ ) resultado del acercamiento entre ambas moléculas (**Figura 1**).



**Figura 1.** Espectro de emisión de fluorescencia tras la transferencia de energía tipo FRET. El espectro de emisión de fluorescencia del donador (línea gris continua) disminuye al mismo tiempo que el espectro de emisión de fluorescencia del aceptor (línea negra continua) aumenta.

El sistema FRET ha sido utilizado para estudiar la sensibilidad estructural de las IDRs (Moses et al., 2020). Estos diseños consisten en una proteína fluorescente donadora y una proteína fluorescente aceptora unidas a cada extremo de la IDR. En este caso, la reducción de la distancia inicial entre las proteínas fluorescentes ocasionada por el cambio estructural de la IDR genera variaciones en la lectura de fluorescencia observable, que se cuantifica como el cambio en la eficiencia de la transferencia de energía.

La eficiencia de FRET es inversamente proporcional a la sexta potencia de la distancia entre dos fluoróforos (Lakowicz 2006). Una eficiencia FRET alta significa que un mayor porcentaje de energía se transfiere desde el donador hacia el aceptor, debido al acortamiento de la distancia entre ambos componentes.

La ecuación matemática que describe el cambio en la eficiencia FRET ( $\Delta E$ ) en función de la distancia entre los fluoróforos en el estado inicial y final es:

$$E = \frac{1}{\left(1 + \frac{R^6}{R_0^6}\right)} \quad (1)$$

En la ecuación 1, R es la distancia entre la proteína fluorescente donadora y aceptora.  $R_0$  se refiere a la distancia en la cual ocurre el 50% de eficiencia FRET, llamada en lo consiguiente como “distancia de Förster”, la cual está definida por:

$$R_0 = 0.02108 \left( \kappa^2 \varphi_D n^{-4} \left( \int_0^\infty f_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda \right) \right)^{\frac{1}{6}} \text{ (en nm)} \quad (2)$$

Donde  $\kappa^2$  es el factor de orientación inter-dipolo de un fluoróforo, el cual se asume que vale 2/3 (del valor promedio de todas las orientaciones posibles de un dipolo antes de la transferencia de energía).  $n$  es el índice de refracción del medio circundante del fluoróforo.  $\Phi_D$  es la eficiencia cuántica del donador en ausencia del aceptor.  $f_D$  es la intensidad de fluorescencia corregida del donador en la longitud de onda  $\lambda$ .  $\varepsilon_A$  es el coeficiente de extinción molar del aceptor ( $M^{-1} \cdot cm^{-1}$ ) en la longitud de onda  $\lambda$ .

En la ecuación 2, la integral representa la superposición entre el espectro de emisión del donador y el espectro de excitación del aceptor. De los tres términos restantes, el factor de orientación difiere entre el estado inicial de una IDR y su cambio estructural posterior, así como el índice de refracción que puede variar dependiendo del contexto. En consecuencia, podemos agrupar y definir la constante C para los términos que no se modifican y sustituirla en la ecuación para eficiencia FRET inicial:

$$E_{inicial} = \frac{1}{1 + \frac{R^6}{C n_{inicial}^{-4} \kappa_{inicial}^5}} \quad (3)$$

De tal forma que la eficiencia FRET final ( $E_{final}$ ) se puede expresar en función de  $R_{final}$  (la distancia final entre el donador y aceptor),  $n^{-4}$  y  $\kappa^2$ . Reescribiendo la ecuación a:

$$\Delta E = (E_{final} - E_{inicial}) = \left( \frac{1}{1 + \frac{R_{final}^6}{C n_{final}^{-4} \kappa_{final}^2}} - \frac{1}{1 + \frac{R_{inicial}^6}{C n_{inicial}^{-4} \kappa_{inicial}^2}} \right) \quad (4)$$

En la ecuación 4, seis variables determinan la magnitud de la eficiencia FRET total ( $\Delta E$ ) que son:  $R_{inicial}$ ,  $R_{final}$ ,  $n_{inicial}^{-4}$ ,  $n_{final}^{-4}$ ,  $\kappa_{inicial}^2$  y  $\kappa_{final}^2$ . De tal forma que  $\Delta E$  tendrá un valor distinto a cero si la estructura de la IDR se modifica reduciendo la distancia entre los fluoróforos. Teóricamente, cuanto más se reduzca la distancia entre el donador y aceptor asociado al cambio estructural del estado inicial y final de la IDR, más grande será el rango de eficiencia FRET en el que se opera. Es decir, a mayor  $\Delta E$  mayor la sensibilidad estructural de la IDR.

Otro factor importante en el sistema FRET es el rango dinámico. El rango dinámico es crucial para la detección de eventos celulares con alta sensibilidad y se obtiene con la siguiente ecuación:

$$\frac{(E_{max} - E_{min})}{(E_{min})} \quad (5)$$

donde se divide la eficiencia FRET máxima ( $E_{max}$ ) menos la eficiencia FRET mínima ( $E_{min}$ ) entre  $E_{min}$ . La eficiencia FRET y el rango dinámico están relacionados por una curva sigmoide con la pendiente mayor en su punto medio ( $E=0.5$  y  $R=R_0$ ). Motivo por el cual  $\Delta E$  debe ser lo suficientemente grande para proporcionar una buena señal en el rango dinámico con un cambio de razón FRET mayor a la señal proveniente del fondo (Campbell et al., 2009).

Experimentalmente, este cambio en las distancias del donador y el aceptor se obtiene midiendo la intensidad de fluorescencia del donador en presencia y en ausencia del aceptor obteniendo un cociente:

$$E = 1 - \frac{F_{DA}}{F_D} \quad (6)$$

Otra forma de estimar la eficiencia de FRET es midiendo el tiempo de vida media del donador en presencia del aceptor ( $\tau_{DA}$ ) y en ausencia del aceptor ( $\tau_D$ ).

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \quad (7)$$

#### **1.4. Osmolaridad intracelular**

Las fluctuaciones osmóticas intracelulares son una de las perturbaciones fisicoquímicas más frecuentes que experimentan las células (Li et al., 2020). En ausencia de estresores externos, la actividad metabólica de la célula induce grandes cambios en la concentración de diferentes metabolitos que alteran la osmolaridad intracelular. Las perturbaciones osmóticas intracelulares severas se desencadenan fácilmente por condiciones de estrés inducidas por el ambiente, donde la osmolaridad fuera de las células cambia dramáticamente. Por un lado, un efecto notorio de la baja osmolaridad es el aumento del volumen celular y por otro lado, el de la alta osmolaridad es la disminución del volumen celular asociado a una fina regulación de la función celular (Miermont et al., 2013).

La homeostasis de los líquidos en los organismos se conoce como osmorregulación y controla la presión osmótica intracelular para evitar la pérdida de agua, ya que su disponibilidad es un requisito fundamental para una amplia gama de funciones celulares. Los organismos han desarrollado diversas estrategias para enfrentar el estrés hiperosmótico generado por las altas concentraciones de solutos orgánicos e inorgánicos (Hosseiniyan Khatibi et al., 2019) y por la disminución en el contenido de agua en el exterior de la célula. Una de las adaptaciones que preservan el equilibrio osmótico celular es la síntesis y acumulación de osmolitos orgánicos, también denominados solutos compatibles. Existen cuatro principales grupos de osmolitos: los carbohidratos, los polioles (por ejemplo, sorbitol, inositoles, glicerol, etc.), los aminoácidos y las metilaminas (Lang et al., 1998).

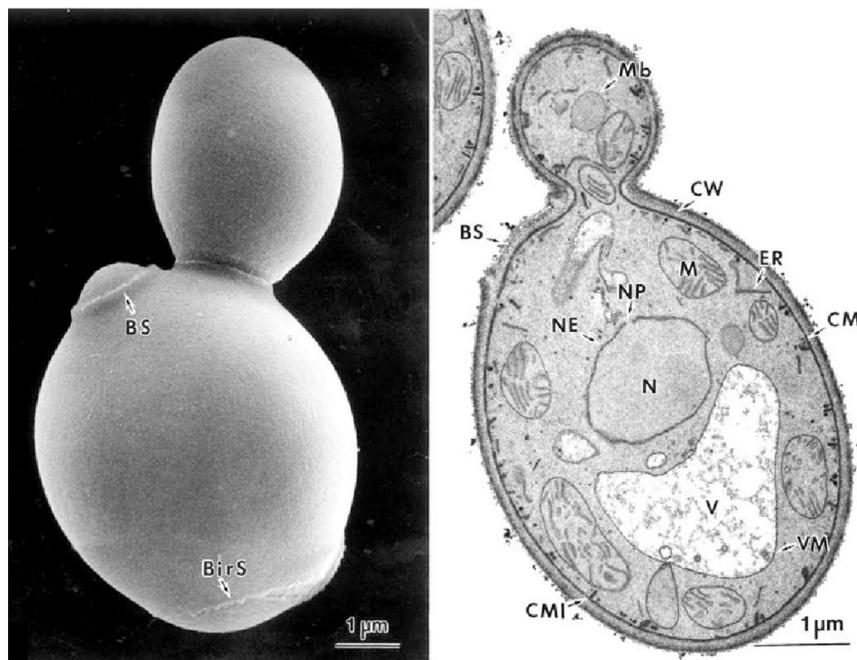
#### **1.5 Estrés hiperosmótico**

El choque hiperosmótico provoca la salida pasiva de agua fuera de la célula. Esto da como resultado un colapso inmediato del volumen celular y un aumento concomitante de la concentración de solutos. La reducción del volumen celular asociada al estrés hiperosmótico induce el incremento del hacinamiento macromolecular y la viscosidad intracelular que a su vez alteran una gran variedad de procesos celulares como, la estructura del citoesqueleto, la remodelación de la cromatina, la detención del ciclo celular y la apoptosis. La respuesta al estrés osmótico ha sido estudiado principalmente a través del organismo modelo *S. cerevisiae* (Stadmler et al., 2017).

## 1.6 *Saccharomyces cerevisiae*

### 1.6.1 *Saccharomyces cerevisiae* como organismo modelo

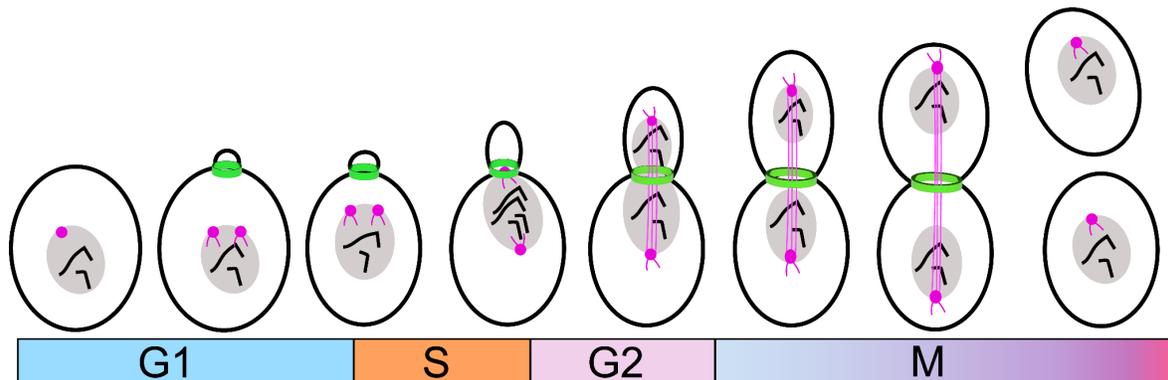
Las células de *S. cerevisiae* tienen una forma elipsoidal. El tamaño de las células es variable, generalmente el ancho se encuentra entre 1-7  $\mu\text{m}$ , incluso las células individuales de una cepa pura pueden mostrar heterogeneidad morfológica (**Figura 2**). El genoma de *S. cerevisiae* es de aproximadamente 12 Mb, organizado en 16 cromosomas y con más de 6000 genes que codifican para proteínas (*Saccharomyces Genome Database*, SGD). Los genes de la levadura tienen pocos intrones y las regiones intergénicas son muy cortas. La redundancia genética en el genoma de la levadura es baja. Las levaduras se cultivan en medio líquido y se aíslan como colonias de células individuales en medios sólidos. El tiempo de generación es muy corto (alrededor de 90 minutos), por lo que grandes poblaciones de individuos se pueden cultivar y analizar rápidamente. Como organismo modelo, muchos de los genes de la levadura están relacionados evolutiva y funcionalmente con los genes de eucariotas superiores (Burgess et al., 2017).



**Figura 2. Características generales de *S. cerevisiae*.** BS, cicatriz de gemación; BirS, cicatriz de nacimiento; CM, membrana celular; CMI, invaginación de la membrana celular; CW, pared celular; ER, Retículo endoplásmico; M, mitocondria; Mb, microcuerpo; N, núcleo NP, poro nuclear; NE, envoltura nuclear; V, vacuola; VM, membrana de la vacuola (Tomado de Osumi 2012).

El ciclo de vida de la levadura se alterna entre células haploides y diploides. Ambos tipos celulares se pueden reproducir de manera asexual por gemación. Sin embargo, la reproducción sexual ocurre únicamente en células haploides de apareamiento opuesto entre células de tipo sexual  $a$  y  $\alpha$ . Cuando las células tipo  $a$  y tipo  $\alpha$  se aparean, fusionan sus núcleos generando una sola célula diploide. En condiciones de inanición, las células diploides ingresan al programa de esporulación que desencadena la meiosis con la reducción concomitante del genoma diploide para producir cuatro esporas haploides. Las cuatro esporas de una meiosis única se mantienen juntas como una tétrada, rodeadas por una pared gruesa llamada ascus. Cuando las condiciones son favorables, las esporas haploides germinan y se fusionan para generar nuevamente células diploides (Burgess et al, 2017).

El estado vegetativo de la levadura ocurre por gemación, a medida que avanza el ciclo celular la morfología celular cambia. Cuando la levadura aparece como una célula redonda está en la fase G1 del ciclo celular (Howell y Lew 2012). Una pequeña yema emerge de un extremo cuando las células entran a la fase S, Durante esta etapa, los cromosomas se duplican, produciendo pares de cromátidas hermanas. En la fase G2, la yema crece y los pares de cromátidas hermanas permanecen unidas entre sí. Durante la mitosis, los cromosomas se condensan y las cromátidas hermanas se segregan entre la célula madre y la yema. Finalmente, la citocinesis deja núcleos genéticamente idénticos en dos células G1, una célula madre y una célula hija (**Figura 3**).



**Figura 3.** El ciclo de vida vegetativo de *Saccharomyces cerevisiae*. Se distinguen en la fase G1, el establecimiento de la polarización de la célula. En la fase S el brote de la yema y la duplicación de los cromosomas. En la fase G2 el crecimiento de la yema y el comienzo de la segregación de los cromosomas. Finalmente, en la fase M se observa la distribución de las cromátidas entre la célula madre y la célula hija seguido de separación celular por citocinesis (Modificado de Howell y Lew 2012).

### 1.6.2 Respuesta al estrés osmótico en *Saccharomyces cerevisiae*

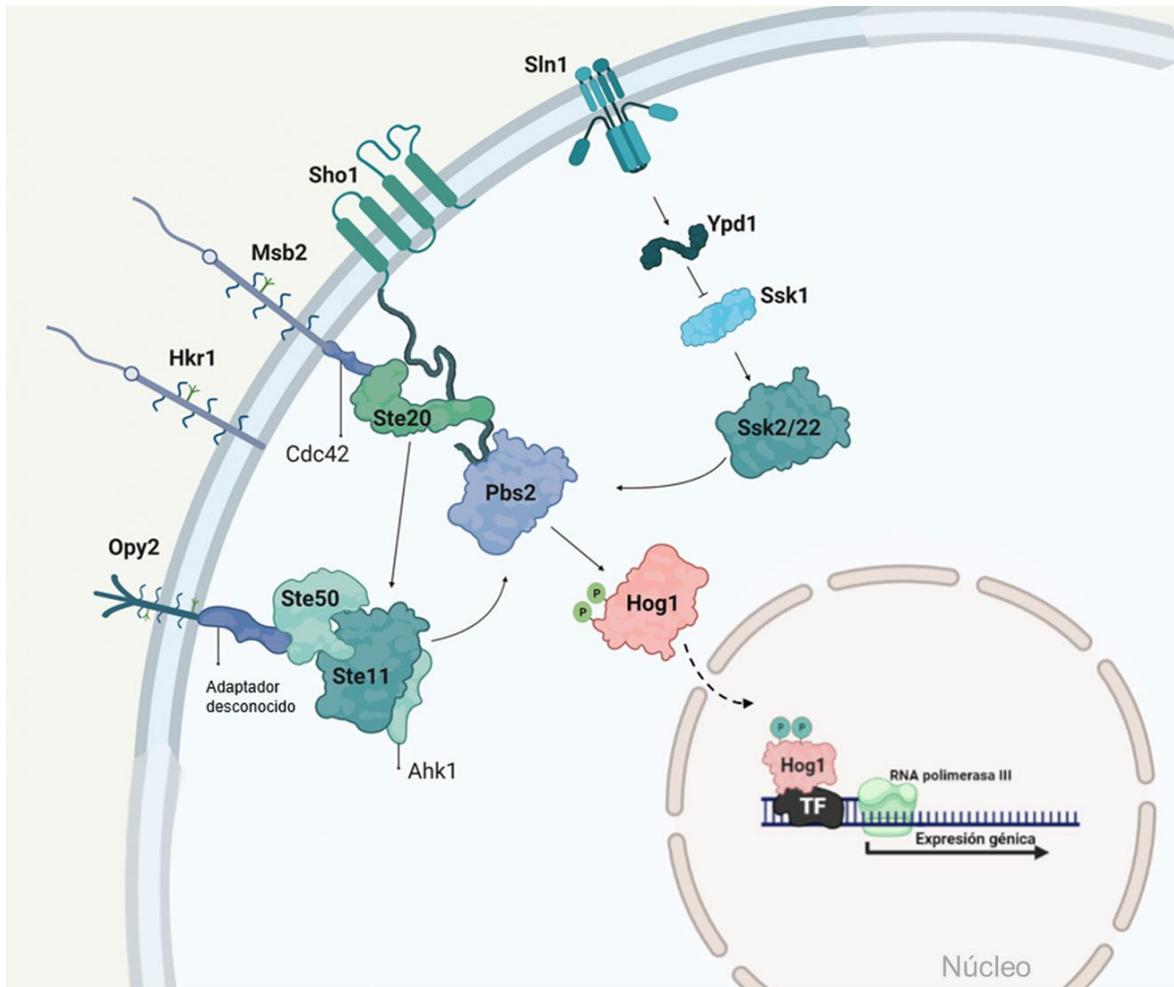
Para asegurar la adaptación osmótica al ser sometida a un ambiente hiperosmótico, *S. cerevisiae* produce una rápida acumulación de glicerol a través de la vía de señalización de glicerol de alta osmolaridad (HOG, del inglés *High Osmolarity Glycerol Pathway*). La vía HOG involucra la cascada de proteínas cinasas activadas por mitógenos (MAPK, del inglés Mitogen Activated Protein Kinase). La cascada consiste en las MAPKKK Ssk2/Ssk22/Ste11, la MAPKK Pbs2 y la MAPK Hog1. En respuesta al choque hiperosmótico la cascada de MAPK puede activarse a través de dos ramas de entrada (Hersen et al., 2008).

La primera rama consiste en un sistema de transferencia de grupos fosfato integrado por la histidina cinasa Sln1 localizada en la membrana plasmática, la proteína de fosfotransferencia de histidina Ypd1, y la proteína reguladora de respuesta Ssk1 (**Figura 4**). En ausencia de estrés hiperosmótico el sistema Sln1-Ypd1-Ssk1 es activo resultando en el estado fosforilado de Ssk1. El estrés hiperosmótico inactiva Sln1, lo que resulta en la acumulación de Ssk1 no fosforilado, que a su vez estimula cambios conformacionales en la cinasa Ssk2/Ssk22 provocando su autofosforilación. Una vez fosforilado, Ssk2/Ssk22 activa a Pbs2 que activa por fosforilación a Hog1. Posteriormente Hog1 fosforilado entra en el núcleo para activar la expresión de genes involucrados en la producción de glicerol (Hersen et al., 2008). Dos de los genes cuya expresión es estimulada codifican para las enzimas Gpd1 y Gpp2, que catalizan la conversión de dihidroxiacetona fosfato (DHAP) a través de glicerol-3-fosfato (G3P) en glicerol (Yaakoub et al., 2021).

La segunda ruta para activar la síntesis de glicerol en respuesta al estrés hiperosmótico involucra al osmosensor transmembranal Sho1 que activa a Pbs2 a través de la cinasa Ste11 (Tatebayashi et al. 2015). Sho1 sufre cambios estructurales que le permiten reclutar cinasas como Ste11, Ste20 y Pbs2, además de varias proteínas más como Ste50 (una proteína adaptadora para Ste11), Opy2 (un ancla transmembranal para Ste50), Cdc42 (una GTPasa citosólica y activador de Ste20) y la proteína Ahk1. La activación de esta rama culmina con la fosforilación de Pbs2 por Ste11 y la consecuente fosforilación de Hog1 (Tatebayashi et al. 2006; Nishimura et al. 2016; Takayama et al. 2019).

La activación de Hog1 conduce a su localización nuclear y la respuesta transcripcional (**Figura 4**). La activación de Hog1 y la estimulación transcripcional dependiente de Hog1 son procesos transitorios. Las proteínas fosfatasas citosólicas y nucleares actúan como reguladores negativos de la vía. La acumulación de glicerol también se controla mediante el cierre rápido de Fps1,

que es un canal de glicerol regulado por osmolaridad, lo que aumenta la presión osmótica interna (Klipp et al., 2005).

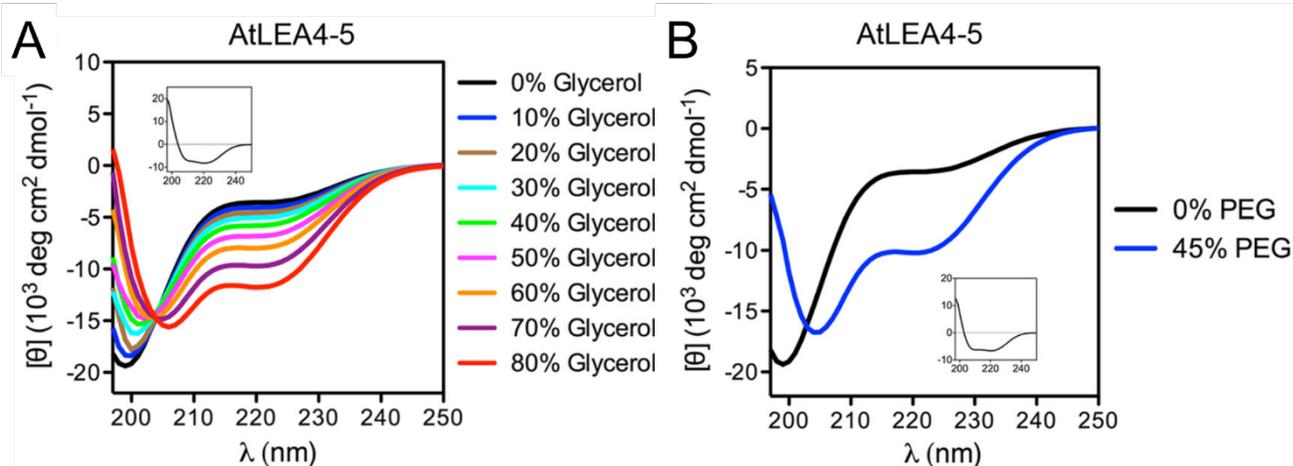


**Figura 4.** La vía HOG en *S. cerevisiae*. La vía HOG de *S. cerevisiae* implica la cascada de cinasas que consta de Hog1 MAPK, Pbs2 MAPKK y Ssk2/Ssk22/Ste11 MAPKKK. Esta cascada tiene dos ramas de activación que convergen en Pbs2; la rama de Sln1 y la rama Sho1. La rama Sln1 consiste en la histidina cinasa Sln1, la proteína de fosfotransferencia de histidina Ypd1, y la proteína reguladora de respuesta Ssk1 y la MAPKKK Ssk2/Ssk22. La rama Sho1 consta de la proteína transmembrana Sho1 junto con la MAPKKK Ste11, la quinada Ste20, Ste50 (proteína adaptadora para Ste11), Opy2 (ancla transmembranal de Ste50), Cdc42 (GTPasa citosólica y activadora de Ste20) y la proteína Ahk1 que asocia a Sho1, Ste11 y Pbs2. Se muestran también representados los posibles osmosensores: Msb2 y Hkr1 (Modificada de Yaakoub et al., 2021).

## II. Antecedentes

El desorden estructural en las proteínas es una característica presente en todos los dominios de la vida, particularmente abundante en los eucariotas, donde desempeña papeles clave en múltiples vías celulares (Dyson y Wright, 2005; Peng et al., 2015). Trabajos previos han demostrado la importancia del cambio estructural de las regiones intrínsecamente desordenadas (IDRs, del inglés *Intrinsically Disordered Regions*) para realizar sus funciones (Robustelli et al., 2020). Sin embargo, aún falta conocer cuáles son los elementos que determinan el cambio estructural desde un estado desordenado hacia estados compactos o más extendidos y si la regulación en los grados de sensibilidad estructural tiene un efecto en las funciones de las IDRs. Entre esos candidatos están los cambios en el ambiente fisicoquímico que son capaces de alterar las preferencias conformacionales de las IDRs. Los enfoques utilizados para abordar este cuestionamiento han sido estudiados principalmente de forma *in vitro* (Holehouse y Sukenink, 2020; Moses et al., 2020), proporcionando valiosa información de la sensibilidad estructural de las IDRs. Sin embargo, poco se sabe si las condiciones controladas en las que se ha estudiado a las IDRs reflejan su comportamiento dentro del contexto celular.

AtLEA4-5, es una proteína intrínsecamente desordenada del grupo cuatro de las proteínas LEA (del inglés *Late Embryogenesis Abundant*) de la planta *Arabidopsis thaliana* (*Arabidopsis*) cuya sensibilidad estructural a los cambios en el ambiente fisicoquímico ha sido demostrada *in vitro* e *in vivo*. Inicialmente, AtLEA4-5 fue caracterizada *in vitro* por ensayos de difracción circular (DC) donde se sometió a la proteína a condiciones de alta osmolaridad y de alto hacinamiento molecular con tratamientos de glicerol y polietilenglicol, respectivamente (Cuevas-Velazquez et al., 2016). En estos experimentos, observaron que AtLEA4-5 era capaz de transitar de un estado desordenado a una conformación de hélice  $\alpha$  en función del incremento en la osmolaridad (**Figura 5A**) y el hacinamiento molecular (**Figura 5B**). De esta forma, se evidenció la sensibilidad estructural *in vitro* de AtLEA4-5 ante el cambio en la osmolaridad del medio, resultando en la adquisición de una conformación de hélice  $\alpha$  (Cuevas-Velazquez et al., 2016).



**Figura 5.** Ensayos de dicroísmo circular UV lejano de AtLEA4-5 en diferentes condiciones de gliceroles y PEG. **A.** Espectros de DC UV lejano de AtLEA4-5 en condiciones de glicerol al 0% (negro), 10% (azul), 20% (marrón), 30% (cian), 40% (verde), 50% (magenta), 60% (naranja), 70% (púrpura) y 80% (rojo). **B.** AtLEA4-5 en solución (línea negra) y en una mezcla de PEG/agua de 45% PEG 5000 (línea azul) (Tomado de Cuevas-Velazquez et al., 2016).

Posteriormente, Cuevas-Velázquez y colaboradores (2021), implementaron un sistema FRET con el cual se estudió la sensibilidad estructural de AtLEA4-5 ante choques hiperosmóticos *in vivo* en las células de múltiples organismos. El sistema fue integrado por un par de proteínas fluorescentes (FP). La proteína fluorescente mCerulean3 como donante y la proteína fluorescente Citrina como aceptor fusionadas a los extremos amino y carboxilo terminal de AtLEA4-5 respectivamente. De esta manera, se estudió la compactación de AtLEA4-5 ante los tratamientos de choque hiperosmótico inducido por las concentraciones crecientes de NaCl, KCl y sorbitol. Así como la sensibilidad estructural ante el incremento en el hacinamiento macromolecular. Estos datos sugieren que la conformación de las IDRs con capacidad de plegarse en hélice  $\alpha$  podría estar modulada por la alta osmolaridad y el hacinamiento molecular *in vivo*.

Debido a que la sensibilidad a los cambios ambientales de las IDRs podría estar conservada en los dominios de la vida, nos cuestionamos si otras IDRs con el potencial de formar hélices  $\alpha$  presentan sensibilidad estructural ante los cambios en el ambiente osmótico intracelular. Para contestar a esta pregunta, en el presente trabajo se estudió la sensibilidad estructural de tres IDRs mediante la técnica de FRET en células vivas de *Saccharomyces cerevisiae*. Estudiar la sensibilidad estructural de las IDRs nos permitirá conocer si el ambiente osmótico intracelular regula la conformación de las IDRs, lo cual podría estar íntimamente ligado a su función molecular.

### **III. Hipótesis**

Las variaciones en el ambiente osmótico intracelular inducen la compactación de regiones intrínsecamente desordenadas con el potencial de formar hélices  $\alpha$  en células vivas de *Saccharomyces cerevisiae*.

### **IV. Objetivo general**

Caracterizar el cambio conformacional de regiones intrínsecamente desordenadas con el potencial de formar hélices  $\alpha$  en respuesta a choques hiperosmóticos *in vivo* en *Saccharomyces cerevisiae*.

### **V. Objetivos particulares**

1. Seleccionar y caracterizar regiones intrínsecamente desordenadas con el potencial de formar hélices  $\alpha$ .
2. Clonar los genes que codifican para las regiones intrínsecamente desordenadas del primer objetivo en el sistema FRET de *Saccharomyces cerevisiae*.
3. Expresar las construcciones en *Saccharomyces cerevisiae*.
4. Caracterizar los cambios en los niveles de FRET en respuesta a choques hiperosmóticos en células vivas de *Saccharomyces cerevisiae*.

## VI. Materiales y métodos

Se realizó una búsqueda de las regiones intrínsecamente desordenadas provenientes de diferentes organismos eucariotas. La búsqueda se dirigió hacia aquellas IDRs con la capacidad reportada en la literatura de adquirir una estructura secundaria de hélice  $\alpha$ , así como desorden estructural. Se buscaron secuencias de una longitud de 60 a 100 aminoácidos. Para ello se consultó la literatura y la base de datos de proteínas intrínsecamente desordenadas de MobiDB (<https://mobidb.org/>). La base de datos de MobiDB permitió buscar las anotaciones de las IDRs curadas de la literatura, además de que incluye la predicción de desorden estructural, así como los límites de las regiones intrínsecamente desordenadas, lo que facilitó la localización de la posición inicial y final de las IDRs. Las secuencias de aminoácidos de las IDRs candidatas se extrajeron de la base de datos de Uniprot ([www.uniprot.org](http://www.uniprot.org)).

### 6.1 Caracterización *in silico* de las secuencias desordenadas

De la búsqueda de secuencias desordenadas se procedió a caracterizar el porcentaje de desorden por residuo de aminoácido y el porcentaje de desorden total de las secuencias de aminoácidos de las IDRs candidatas, a través del predictor de regiones desordenadas naturales (PONDR, del inglés *Prediction Of Natural Disordered Regions*) con el algoritmo VL-XT (<http://www.pondr.com/>). Este predictor integra tres redes neuronales (Romero et al., 2000) y es capaz de clasificar cada residuo dentro de una secuencia como ordenada o desordenada utilizando un rango de puntaje. Un valor arriba o igual de 0.5 indica que el aminoácido se encuentra en una región desordenada y por debajo de 0.5 en una región ordenada.

La tendencia a la formación de estructura secundaria tipo hélice  $\alpha$  por residuo de las IDRs candidatas se evaluó con el predictor Agadir ([agadir.crg.es](http://agadir.crg.es)). También se determinó la composición de aminoácidos de las secuencias, así como la masa molecular, el punto isoeléctrico, el coeficiente de extinción molar, el índice de inestabilidad, el índice alifático y el promedio general de hidropatía con la herramienta ProtParam del servidor ExPASy.

A través de la herramienta computacional CIDER (del inglés *Classification of Intrinsically Disordered Ensemble Regions*) (Holehouse et al., 2017) se realizó el análisis de una gama de las propiedades fisicoquímicas codificadas en las secuencias de las regiones intrínsecamente desordenadas candidatas. CIDER cuenta con una versión de servidor web de acceso libre a través de <http://pappulab.wustl.edu/CIDER>, que toma las secuencias de IDRs en formato

fasta como entradas para realizar el cálculo de una serie de parámetros y diagramas específicos de las secuencias (Holehouse et al., 2017).

## 6.2 Síntesis de las IDRs

Se solicitó la síntesis del marco de lectura abierto (ORF, del inglés *Open Reading Frame*) de tres IDRs seleccionadas a la compañía Integrated DNA Technologies (IDT, USA). Los fragmentos de DNA solicitados contuvieron a las regiones codificantes de las IDRs más seis pares de bases pertenecientes al sitio de restricción de ScaI hacia el 5' y seis pares de bases hacia el 3' de BglII, más nueve pares de bases homólogas a las secuencias de las proteínas fluorescentes del sistema FRET contenido en el plásmido pDR38 (**Figura 4**). Posteriormente la compañía IDT clonó los fragmentos de DNA de cada IDR en el plásmido pIDTSMART-AMP con resistencia a ampicilina.

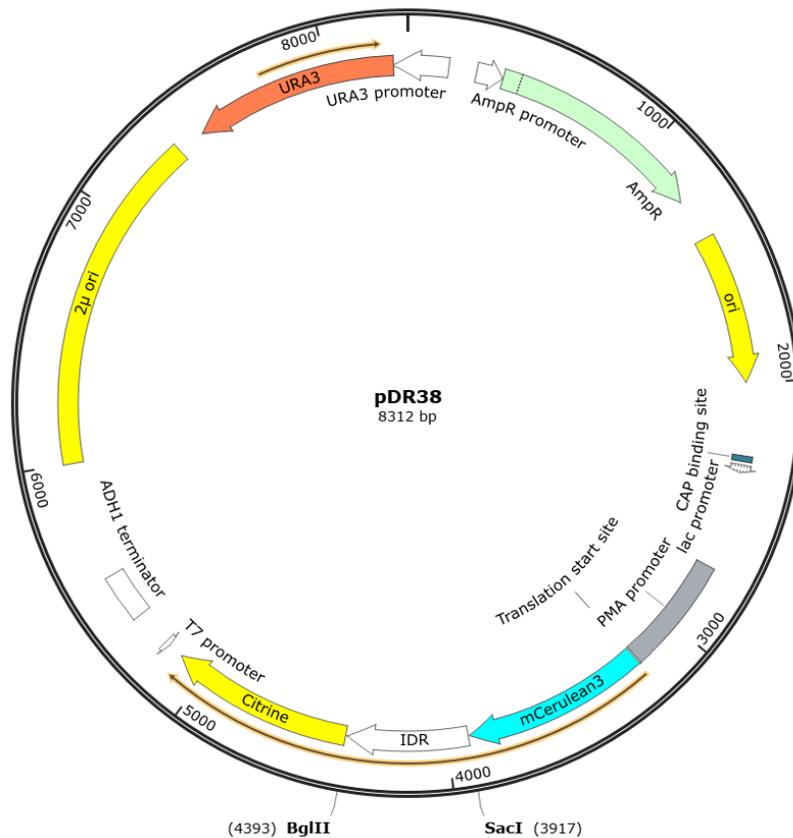
Las secuencias sintetizadas se clonaron individualmente por el método de ensamblaje de Gibson. Para ello primero se amplificaron cada una de las secuencias codificantes de las IDRs por PCR empleando la DNA Taq polimerasa de arranque en caliente Platinum II de alta fidelidad, usando oligonucleótidos específicos para cada construcción (ver Tabla 2). Se realizó la purificación de la banda del tamaño esperado con el kit de extracción y purificación de banda Zymoclean™ Gel DNA Recovery de la marca Zymo Research (USA). Posteriormente se obtuvo al plásmido pDR38 libre de la región codificante de AtLEA4-5 con una restricción doble con las enzimas ScaI y BglII. La reacción de ensamblaje de Gibson procedió entonces en condiciones isotérmicas (ver Anexos).

**Tabla 2.** Oligonucleótidos utilizados en este proyecto

Número	Descripción	Secuencia (5'-->3')
CET015	Forward para amplificar IDR-CREB1	GGA TGA ACT GTA TAA AGA GCT CCA AAT TTC TAC TAT T
CET016	Reverso para amplificar IDR-CREB1	CCT CAC CTT TTG ACA CCA TAG ATC TAG TTT CTT CTT CAG A
CET017	Forward para amplificar IDR-EPYC1	GGA TGA ACT GTA TAA AGA GCT CGT TTC TCC AAC TAG G
CET018	Reverso para amplificar IDR-EPYC1	CCT CAC CTT TTG ACA CCA TAG ATC TAG CTT TTT TAG AAG C
CET019	Reverso para amplificar IDR-U1TDCP	GGA TGA ACT GTA TAA AGA GCT CAG AGC TTT GGG TAT G
CET020	Reverso para amplificar IDR-U1TDCP	CCT CAC CTT TTG ACA CCA TAG ATC TAG ATT TTT TAG AAG AAC C

Las construcciones finales contaron con el marco de lectura abierto (ORF) de cada una de las proteínas fluorescentes posicionados a cada lado de la región codificante de la IDR seleccionada. En estas construcciones se utilizó como

aceptor a la proteína fluorescente Citrina y como donador a mCerulean3, en el plásmido pDR38 (**Figura 4**) para su expresión en la cepa de levadura BJ5465.



**Figura 4.** Plásmido pDR38 para expresión en levadura, contiene sistema FRET conformado de las proteínas fluorescentes mCerulean3 y Citrina. Obtenido del trabajo de Cuevas-Velazquez et al., 2021.

### 6.3 Expresión de las IDRs en la levadura

Las construcciones se transformaron en la cepa de *S. cerevisiae* BJ5465, deficiente en las proteasas (Pep4 y Prb1), con el protocolo de choque térmico (ver anexos). La selección de células transformadas se realizó por auxotrofia en el medio de selección carente de uracilo (SD-ura) (ver anexos), solo las colonias transformadas con la construcción correspondiente crecieron. Posteriormente, se verificó la transformación mediante la emisión de la fluorescencia en placa. Para ello, se incidió con luz azul de una longitud de onda de 465 nm sobre las colonias de levadura crecidas en medio de selección usando un transiluminador E4000 SmartBlue™ (Accuris Instruments, USA) y se confirmó por PCR de colonia de levadura con oligonucleótidos y condiciones de amplificación específicos para cada construcción (ver anexos).

## 6.4 Ensayos FRET

Se cultivó durante la noche el inóculo con la colonia de levadura que contuvo la construcción correspondiente en 3 mL de medio de selección líquido a 200 rpm a 30°C hasta alcanzar un rango de densidad óptica de 1 a 2 OD<sub>600</sub>. Se colectó 1 mL de cultivo por centrifugación a 12,000 rpm respectivamente y se realizó un proceso de lavado con buffer MES 50 mM, pH 6 dos veces. Posteriormente se resuspendieron en 1 mL del mismo buffer y se colocaron en reservorios de 25 mL. Posteriormente, se evaluó la sensibilidad estructural de las tres construcciones de IDRs expresadas de forma independiente en las células levadura a través de tratamientos con soluciones hipertónicas.

Los ensayos FRET se montaron en placas de 96 pocillos negras con fondo transparente, para ello, se mezclaron 50 µL de las levaduras con 150 µL de las soluciones de NaCl, según fue el caso. Las concentraciones finales de las soluciones utilizadas en el ensayo de choque hiperosmótico fueron 0 mM, 200 mM, 400 mM, 600 mM, 800 mM, 1000 mM y 1500 mM de NaCl. Se realizaron 3 réplicas técnicas para cada uno de los tratamientos. Se utilizó NaCl para estos ensayos debido a que es el osmolito que indujo un mayor cambio en el cociente FRET de AtLEA4-5 (Cuevas-Velazquez et al., 2021).

Inmediatamente después de la aplicación del tratamiento con NaCl, se midió el espectro de emisión de fluorescencia usando el lector de placas CLARIOStar (BMG LABTECH). La longitud de onda de excitación fue de 433 nm y la emisión se registró de 460 a 550 nm. El ancho de banda utilizado en la excitación y emisión fue de 10 nm, la ganancia se ajustó manualmente para cada construcción. En las mediciones se incluyó a la proteína AtLEA4-5 como referencia de una IDR y la proteína de unión a arabinosa (ABP) como referencia de una proteína globular. Se realizaron 3 réplicas independientes de todas las mediciones.

## VII. Resultados

### 7.1 Construcción de la base de datos con las secuencias de IDRs candidatas

Resultado de la búsqueda intensiva de secuencias de IDRs, se obtuvieron 50 secuencias de IDRs candidatas (Tabla 1). Estas secuencias se recopilieron para la construcción de una base de datos (<https://docs.google.com/spreadsheets/d/1yj0QPnwMA4zrCzZL97gRdbGcJoAfCFb1dWluOfM-s18/edit?usp=sharing>) donde se anotó el organismo al que pertenecía cada IDR, la longitud de aminoácidos, así como el identificador de Uniprot, MobiDB, y las referencias de las publicaciones científicas que las reportaron. La construcción de esta base de información permitió la posterior caracterización *in silico* de las secuencias de IDRs candidatas.

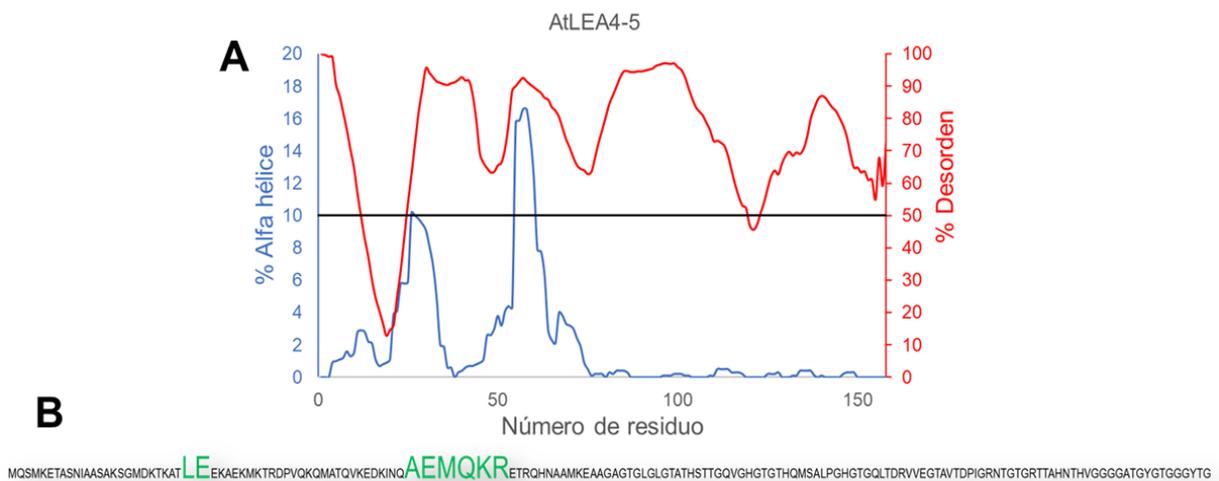
**Tabla 1.** Lista de 50 IDRs candidatas

IDR	Organismo	Nombre de la proteína	UniProt	MoBI	Tamaño (aa)
IDR-001	<i>Schizosaccharomyces pombe</i>	26S proteasome complex subunit rpn15	O14140	O14140	71
IDR-002	<i>Homo sapiens</i>	Breast cancer type 1 susceptibility protein	P38398	P38398	100
IDR-003	<i>Homo sapiens</i>	Cellular tumor antigen p53	P04637	P04637	74
IDR-004	<i>Homo sapiens</i>	Death domain-associated protein 6	Q9UER7	Q9UER7	74
IDR-005	<i>Mus musculus</i>	Bcl-2-binding component 3	Q99ML1	Q99ML1	54
IDR-006	<i>Rattus norvegicus</i>	Sequestosome-1	O08623	O08623	100
IDR-007	<i>Homo sapiens</i>	Cyclin-dependent kinase inhibitor 1B	P46527	P46527	81
IDR-008	<i>Homo sapiens</i>	Splicing factor 3B subunit 1	O75533	O75533	100
IDR-009	<i>Homo sapiens</i>	Polyglutamine-binding protein 1	O60828	O60828	50
IDR-010	<i>Saccharomyces cerevisiae</i>	Pre-mRNA-splicing factor NTR2	P36118	P36118	100
IDR-011	<i>Saccharomyces cerevisiae</i>	Nucleoporin NUP2	P32499	P32499	100
IDR-012	<i>Mus musculus</i>	RWD domain-containing protein 1	Q9CQK7	Q9CQK7	100
IDR-013	<i>Mus musculus</i>	Protein phosphatase inhibitor 2	Q9DCL8	Q9DCL8	50
IDR-014	<i>Mus musculus</i>	POU domain class 2-associating factor 1	Q64693	Q64693	65
IDR-015	<i>Homo sapiens</i>	Achaete-scute homolog 1	P50553	P50553	90
IDR-016	<i>Homo sapiens</i>	Achaete-scute homolog 1	P50553	P50553	58
IDR-017	<i>Arabidopsis thaliana</i>	Transcription factor TCP8	Q9C518	Q9C518	100
IDR-018	<i>Arabidopsis thaliana</i>	Transcription factor TCP8	Q9C518	Q9C518	100
IDR-019	<i>Arabidopsis thaliana</i>	Transcription factor HY5	O24646	O24646	77
IDR-020	<i>Arabidopsis thaliana</i>	BRI1 kinase inhibitor 1	Q9FMZ0	Q9FMZ0	100
IDR-021	<i>Rattus norvegicus</i>	Cyclic AMP-responsive element-binding protein 1	P15337	P15337	60
IDR-022	<i>Proteus hauseri</i>	Modification methylase Pvull	P11409	P11409	58
IDR-023	<i>Saccharomyces cerevisiae</i>	General control protein GCN4	P03069	P03069	50
IDR-024	<i>Escherichia coli</i>	Protein GrpE	P09372	P09372	50
IDR-025	<i>Saccharomyces cerevisiae</i>	Acetolactate synthase catalytic subunit, mitochondrial	P07342	P07342	50
IDR-026	<i>Saccharomyces cerevisiae</i>	Acetyl-CoA carboxylase	Q00955	Q00955	53
IDR-027	<i>Homo sapiens</i>	Wee1-like protein kinase	P30291	P30291	50
IDR-028	<i>Rattus norvegicus</i>	Vesicle-associated membrane protein 2	P63045	P63045	88
IDR-029	<i>Drosophila melanogaster</i>	FACT complex subunit Ssrp1	Q05344	Q05344	68
IDR-030	<i>Drosophila melanogaster</i>	FACT complex subunit Ssrp1	Q05344	Q05344	50
IDR-031	<i>Drosophila melanogaster</i>	FACT complex subunit Ssrp1	Q05344	Q05344	99
IDR-032	<i>Bos taurus</i>	Cyclic nucleotide-gated cation channel beta-1	Q28181-4	Q28181	50
IDR-033	<i>Streptococcus gordonii</i>	Amylase-binding protein AbpA	A8A273	A8A273	80
IDR-034	<i>Arabidopsis thaliana</i>	U1-type domain-containing protein	AT3G05760	AT3G05760	94
IDR-035	<i>Homo sapiens</i>	Insulinoma-associated protein 1	Q01101	Q01101	50
IDR-036	<i>Saccharomyces cerevisiae</i>	Transcriptional regulatory protein ASH1	P34233	P34233	84
IDR-037	<i>Rattus norvegicus</i>	Glial cell line-derived neurotrophic factor	Q07731	Q07731	50
IDR-038	<i>Homo sapiens</i>	Alpha-synuclein	P37840	P37840	50
IDR-039	<i>Rattus norvegicus</i>	Bcl-2-like protein 1	P53563	P53563	50
IDR-040	<i>Saccharomyces cerevisiae</i>	Actin-related protein 4	P80428	P80428	50
IDR-041	<i>Homo sapiens</i>	Mitochondrial intermembrane space import and assembly protein 40	Q8N4Q1	Q8N4Q1	50
IDR-042	<i>Homo sapiens</i>	Regulator of nonsense transcripts 2	Q9HAU5	Q9HAU5	50
IDR-043	<i>Homo sapiens</i>	Glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein 1	Q8IV16	Q8IV16	50
IDR-044	<i>Rattus norvegicus</i>	High mobility group protein B1	P63159	P63159	74
IDR-045	<i>Aquifex aeolicus</i>	50S ribosomal protein L20	O67086	O67086	59
IDR-046	<i>Homo sapiens</i>	UDP-glucose 6-dehydrogenase	O60701	O60701	50
IDR-047	<i>Chlamydomonas reinhardtii</i>	EPYC1	-	-	63
IDR-048	<i>Homo sapiens</i>	Nuclear receptor coactivator 3	Q9Y6Q9	Q9Y6Q9	71
IDR-049	<i>Mus musculus</i>	Histone lysine acetyltransferase CREBBP	P45481	P45481	96
IDR-050	<i>Homo sapiens</i>	Tumor suppressor ARF	Q8N726	Q8N726	73

## 7.2 Identificación de regiones intrínsecamente desordenadas con el potencial de formar hélice $\alpha$

Se desarrolló un parámetro para predecir la capacidad a experimentar cambios conformacionales en respuesta a los efectos del cambio en la osmolaridad de las IDR candidatas. Para ello, se consideraron las características de la secuencia de AtLEA4-5.

AtLEA4-5 consta de 158 aminoácidos, con una alta predicción de desorden total del 89.87 % y un porcentaje de hélice  $\alpha$  total de 2.11 %. La predicción de AtLEA4-5 contiene dos hélices  $\alpha$  con un puntaje de más del 10%. Los resultados de predicción por residuo indican que ciertas regiones de la IDR cuentan con residuos predichos como desordenados y a su vez con predicción de formación de hélice  $\alpha$  (Figura 5).



**Figura 5.** Predicción de desorden y hélice  $\alpha$  de AtLEA4-5. **A.** La proteína AtLEA4-5 presenta dos picos con alta predicción de desorden y con alta predicción de hélice  $\alpha$ . **B.** El primer pico corresponde a dos aminoácidos con predicción simultánea de desorden y de hélice  $\alpha$  y el segundo pico a seis aminoácidos con estas características.

Como resultado de este análisis se observó la presencia de dos regiones que presentaron un alto porcentaje de predicción de desorden y moderado de hélice  $\alpha$ , la primera con dos aminoácidos y la segunda con seis aminoácidos. De esta forma, se obtuvo sistemáticamente la predicción de desorden y hélice  $\alpha$  características en las secuencias de IDRs seleccionadas (ver métodos) y se calcularon los valores de las predicciones tomando como referencia el carácter dual observado en los residuos de AtLEA4-5, con el fin de predecir la capacidad de las secuencias a experimentar cambios conformacionales en respuesta del incremento en la osmolaridad. La obtención de las predicciones antes

mencionadas permitió desarrollar un nuevo parámetro al que se le llamó “parámetro de dualidad”.

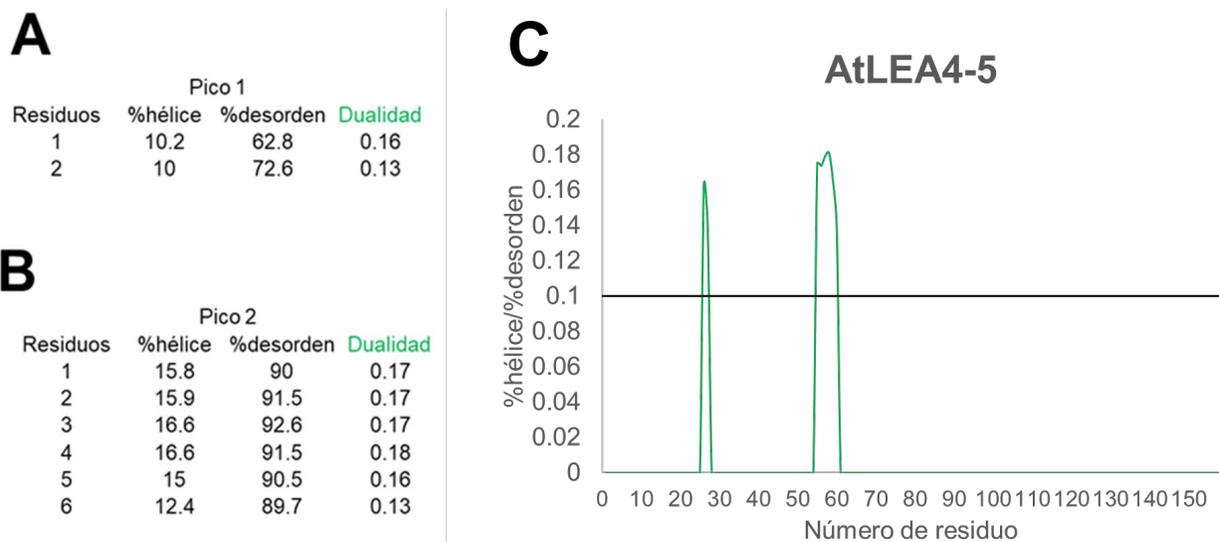
### 7.3 Establecimiento del parámetro de dualidad y calibración

La dualidad fue definida como el cociente del porcentaje de predicción de hélice  $\alpha$  entre el porcentaje de predicción de desorden por aminoácido:

$$D = \frac{\text{Porcentaje}_{\text{alfa hélice}}}{\text{Porcentaje}_{\text{desorden}}}$$

El parámetro dual consideró únicamente a los aminoácidos que cumplieran con ambas predicciones de ser desordenados y de formar hélice  $\alpha$ . Para calibrar la dualidad se estableció el corte de porcentaje de hélice  $\alpha$  de  $\geq 10\%$  como límite mínimo para el análisis, debido a que es el valor más bajo de hélice  $\alpha$  presente en el primer pico dual de AtLEA4-5, por debajo de este valor umbral los valores fueron descartados en la predicción de la dualidad. Con respecto al desorden, el límite mínimo se estableció en seis residuos consecutivos predichos como desordenados con el predictor PONDR (ver sección 6.1 de métodos) tomando como base el segundo pico dual de AtLEA4-5, los residuos por debajo de este rango no fueron considerados para el cálculo de la dualidad.

Para establecer el umbral de dualidad se tomaron ciertas consideraciones, la primera fue la predicción del comportamiento dual ideal, es decir, residuos que sean desordenados y cuenten con  $\geq 10\%$  de hélice  $\alpha$  determinarían el límite superior particular de cada secuencia. La segunda consideración fue que la propiedad dual debía ser propia de IDRs altamente desordenadas, por ello el límite de dualidad inferior se estableció en 0.1, buscando aquellas IDRs con un porcentaje de desorden cercano o igual al 100% (**Figura 6**).



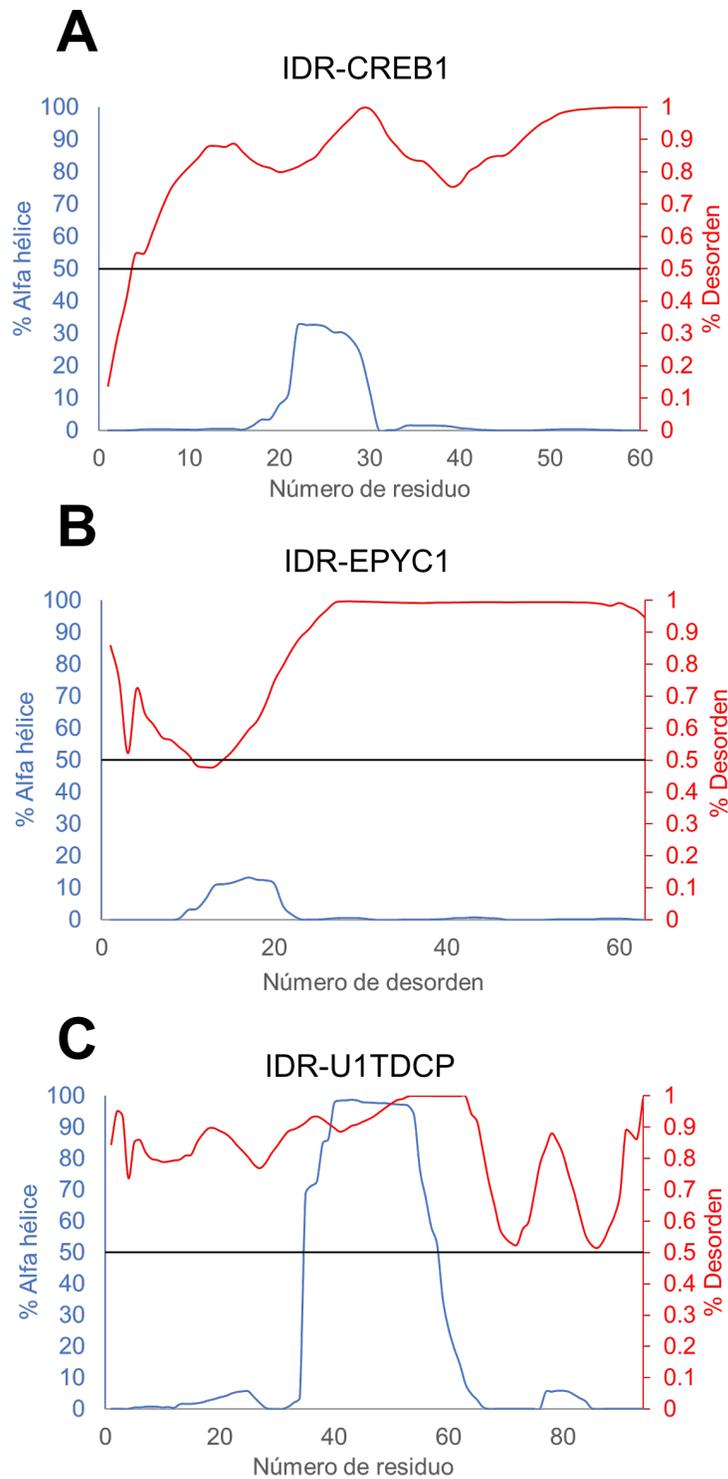
**Figura 6.** Calibración del parámetro dual de acuerdo con el análisis de AtLEA4-5. **A.** Se muestra la predicción de hélice  $\alpha$  y desorden para el primer pico que cuenta con dos residuos duales. **B.** Predicción de hélice  $\alpha$  y desorden para el primer pico que cuenta con seis residuos duales. **C.** Representación del parámetro de dualidad a lo largo de la secuencia de AtLEA4-5. Se muestra el límite inferior de dualidad, el cual está establecido para residuos con niveles de 100% de desorden.

#### 7.4 Selección de tres regiones intrínsecamente desordenadas con dualidad

Se analizó el nivel de desorden y la tendencia a la formación de hélice  $\alpha$  de las secuencias de cada una de las IDRs candidatas (**Figura 7**). Posterior a este análisis se obtuvo el parámetro de dualidad de todas las secuencias candidatas para la selección final de tres secuencias de IDR con residuos con predicción dual provenientes de tres organismos modelo distintos, como se muestra en la Tabla 3.

**Tabla 3.** IDRs seleccionadas de organismos diferentes

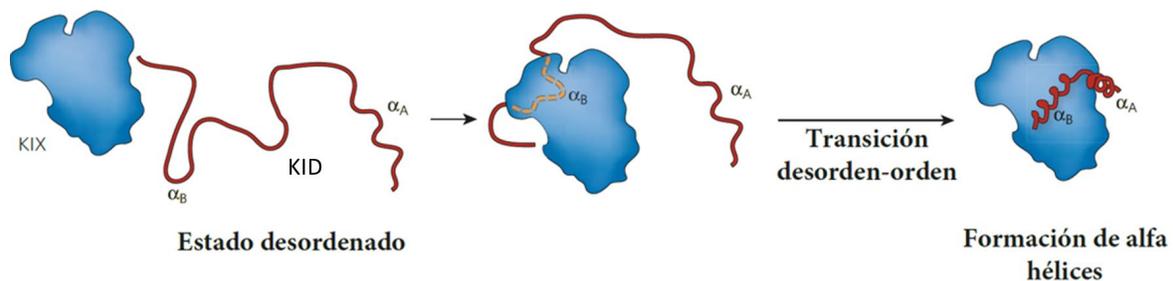
Organismo	Nombre de la proteína	UniProt	Longitud de la IDR (aa)	No. residuos duales
<i>Rattus norvegicus</i>	Cyclic AMP-responsive element-binding protein 1 (CREB1)	P15337	60	10
<i>Chlamydomonas reinhardtii</i>	Essential Pyrenoid Component 1 (EPYC1)	-	63	7
<i>Arabidopsis thaliana</i>	U1-type domain-containing protein (U1TDCP)	AT3G057	60	28



**Figura 7.** Predicción de desorden (rojo) y hélice  $\alpha$  (azul) de IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP. **A.** IDR-CREB1 presenta un solo pico con alta predicción de desorden y con predicción de hélice  $\alpha$ . **B.** IDR-EPYC1 presenta un solo pico con ambas características de desorden y hélice  $\alpha$ . **C.** IDR-U1TDCP presenta un pico con alta predicción simultánea de desorden y de hélice  $\alpha$ .

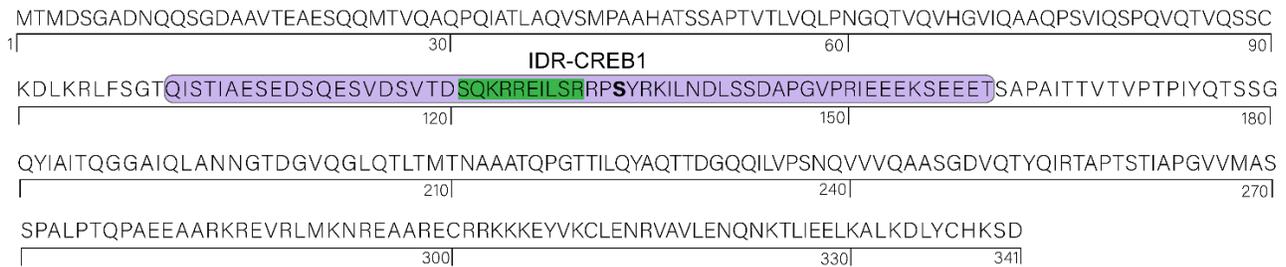
### 7.4.1 La proteína de unión a elementos sensible a AMP cíclico 1 (CREB1)

La proteína CREB1 es una proteína intrínsecamente desordenada con la capacidad de transitar a un estado ordenado de hélice  $\alpha$  tras su fosforilación y la posterior unión con su compañero de interacción. CREB1 contiene al dominio inducible por cinasa (KID, del inglés *kinase inducible domain*) que en su estado no fosforilado es desordenado en solución (Radhakrishnan et al., 1998). Sin embargo, tras la fosforilación en el residuo de serina 133 (S133), se promueve la asociación con su compañero de interacción, el dominio KIX (del inglés *kinase-inducible domain (KID) interacting domain*) de la proteína de unión a CREB1 (CBP, del inglés *CREB binding protein*). Experimentando la transición al estado ordenado (**Figura 8**). El cambio de estado conformacional es necesario para llevar a cabo la función de CREB1 como activador transcripcional inducible al unirse al elemento de respuesta de cAMP (CRE, del inglés *cAMP response element*) del DNA, una secuencia presente en muchos promotores virales y celulares (Liu et al., 2020).



**Figura 8.** Mecanismo de transición desorden a orden de CREB1 fosforilado en el dominio KID tras la interacción con el dominio KIX de la proteína CBP, con la consecuente formación de dos hélices  $\alpha$ ;  $\alpha_B$  y  $\alpha_A$ . Modificado de Eliezer y Palmer, 2007.

El dominio KID de *Rattus norvegicus* consta de 60 residuos de longitud (residuos 101 al 160) y se seleccionó para su análisis *in vivo* porque tiene dos regiones con propensión a la formación de hélices  $\alpha$  en solución (residuos 120-129; 134-144) que, al unirse a su compañero de interacción se estabiliza a una conformación de hélice-giro-hélice centrada alrededor del residuo S133 (**Figura 9**). No se seleccionaron las regiones que flanquean a las hélices ya que, de acuerdo con la literatura, estos residuos no aumentan la propensión a la formación de hélice  $\alpha$  del dominio de KID (Radhakrishnan et al., 1997; Radhakrishnan et al., 1998).



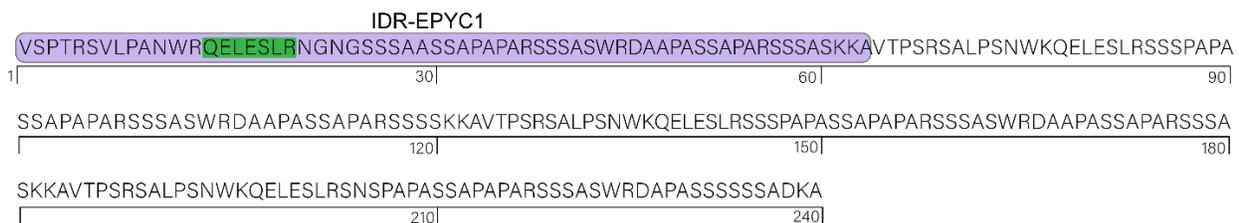
**Figura 9.** Secuencia de aminoácidos de la proteína CREB1 de *Rattus norvegicus*. En morado se muestra al dominio KID (residuos 101 al 160) que es la región intrínsecamente desordenada evaluada en el presente trabajo, en verde los residuos con predicción dual.

#### 7.4.2 La proteína componente esencial del pirenoide 1 (EPYC1)

EPYC1 es una proteína intrínsecamente desordenada de *Chlamydomonas reinhardtii*, la cual permite el establecimiento y mantenimiento de una estructura fundamental para la fijación del dióxido de carbono en algas, el pirenoide. El pirenoide es una estructura esférica que se encuentra en el estroma de los cloroplastos y que utiliza a EPYC1 para formar una matriz proteica que concentra a enzimas como Rubisco y a su chaperona la Rubisco activasa (Mackinder et al., 2016).

La proteína EPYC1 consiste en cuatro repeticiones casi idénticas de 60 aminoácidos, flanqueadas por el amino terminal y carboxilo terminal. Estas repeticiones cuentan con predicción de desorden y un dominio con predicción de hélice  $\alpha$  (Mackinder et al., 2016) (**Figura 10**). Además, EPYC1 facilita la formación y organización del pirenoide a través de la separación de fase líquido-líquido (LLPS) (Freeman Rosenzweig et al., 2017).

Se seleccionó la primera repetición (residuos 52-114) para su análisis *in vivo* porque tiene predicción de desorden y de hélice  $\alpha$ . Además de estar involucrada en el fenómeno de LLPS, el cual podría estar regulado por cambios en el ambiente fisicoquímico (Mackinder et al., 2016).



**Figura 10.** Secuencia de aminoácidos de la proteína EPYC1 de *Chlamydomonas reinhardtii*. En morado se muestra la región intrínsecamente desordenada que se evaluó en este trabajo, en verde los residuos con predicción dual.

### 7.4.3 El factor de transcripción U1-type domain-containing protein (U1TDCP)

Los análisis globales de la predicción del desorden estructural en los factores de transcripción (FT) han revelado que cerca del 80% de los FT de los eucariontes contienen una o más IDR (Liu et al., 2006). Para la planta modelo *Arabidopsis thaliana* (*Arabidopsis*) este porcentaje asciende al 91% (Salladini et al., 2020).

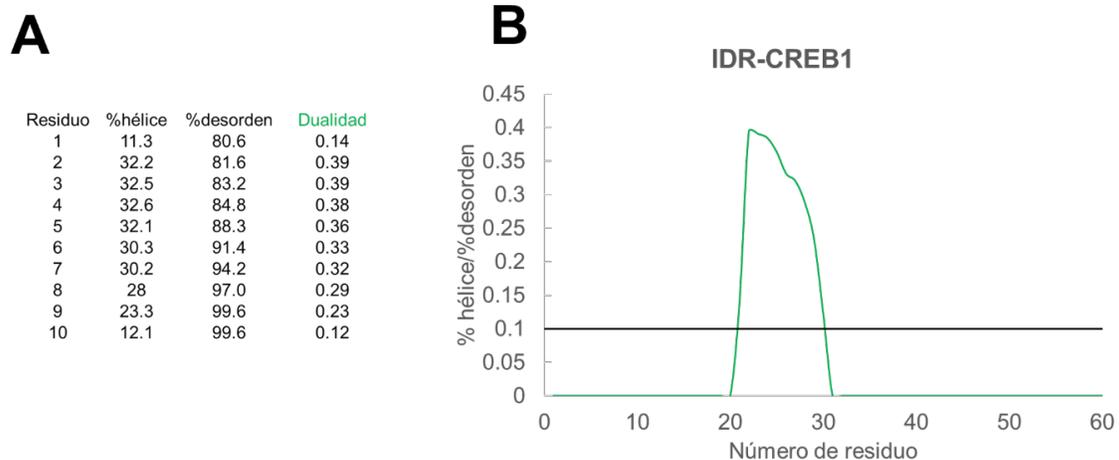
Se seleccionó la IDR más larga del factor de transcripción U1-type domain-containing protein (U1TDCP) de *Arabidopsis* porque además de su alta predicción de desorden pertenece a la familia de dedos de zinc C2H2, esta familia de FT juega un papel importante en la resistencia al estrés abiótico en las plantas (**Figura 11**) (Han et al., 2020).



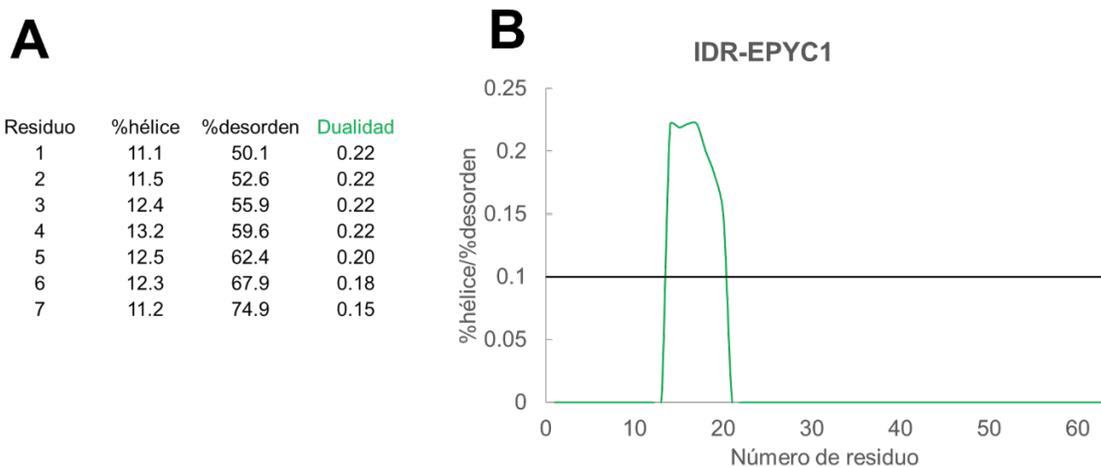
**Figura 11.** Secuencia de aminoácidos de la proteína U1TDCP de *Arabidopsis*. En morado se muestra la región intrínsecamente desordenada que se evaluó en este trabajo, en verde los residuos con predicción dual.

## 7.5 Análisis de la dualidad de regiones intrínsecamente desordenadas seleccionadas

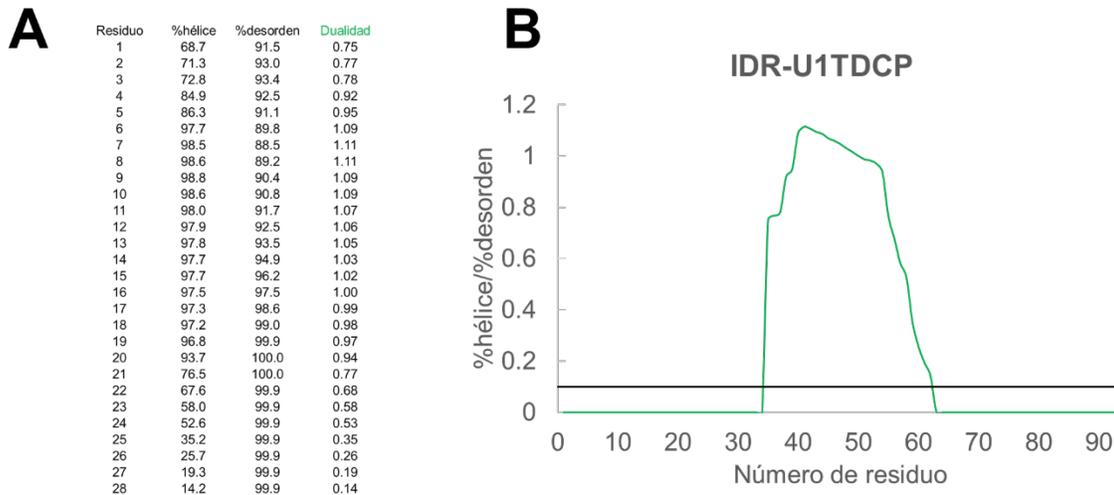
IDR-U1TDCP del factor transcripcional de Arabidopsis contó con la cobertura más alta de dualidad, representado por el 29.7% de los residuos duales del total de la secuencia, situada en la región intermedia de la IDR (residuos 35 al 62) así como los valores más altos de predicción dual 0.82 (**Figura 14**). IDR-CREB1 cuenta con el 16.6% de la secuencia con residuos duales los cuales alcanzaron un valor dual promedio de 0.29 en la región intermedia (residuos 21 al 30) (**Figura 12**). Finalmente, IDR-EPYC1 contó con una cobertura de dualidad del 11% con un valor promedio de 0.20 en los residuos 14 al 20 (**Figura 13**).



**Figura 12.** Representación del parámetro de dualidad a lo largo de la secuencia de IDR-CREB1. **A.** Se muestran los porcentajes de la predicción de hélice  $\alpha$  y desorden para los 10 residuos duales. **B.** Se representa la dualidad presente en los residuos 21 al 30 de IDR-CREB1.



**Figura 13.** Representación del parámetro de dualidad a lo largo de la secuencia de IDR-EPYC1. **A.** Se muestran los porcentajes de la predicción de hélice  $\alpha$  y desorden para los 7 residuos duales. **B.** Se representa la dualidad presente en los residuos 14 al 20 de IDR-EPYC1.



**Figura 14.** Representación del parámetro de dualidad a lo largo de la secuencia de IDR-U1TDCP. **A.** Se muestran los porcentajes de la predicción de hélice  $\alpha$  y desorden para los 28 residuos duales. **B.** Se representa la dualidad presente en los residuos 35 al 62 de IDR-U1TDCP.

## 7.6 Caracterización de las propiedades fisicoquímicas de las regiones intrínsecamente desordenadas con CIDER

Con la finalidad de obtener una referencia cuantitativa de lo que se podría esperar en el comportamiento de las IDRs seleccionadas tomando como referencia a las características de AtLEA4-5, se empleó la herramienta CIDER para analizar las secuencias de aminoácidos de las IDRs seleccionadas.

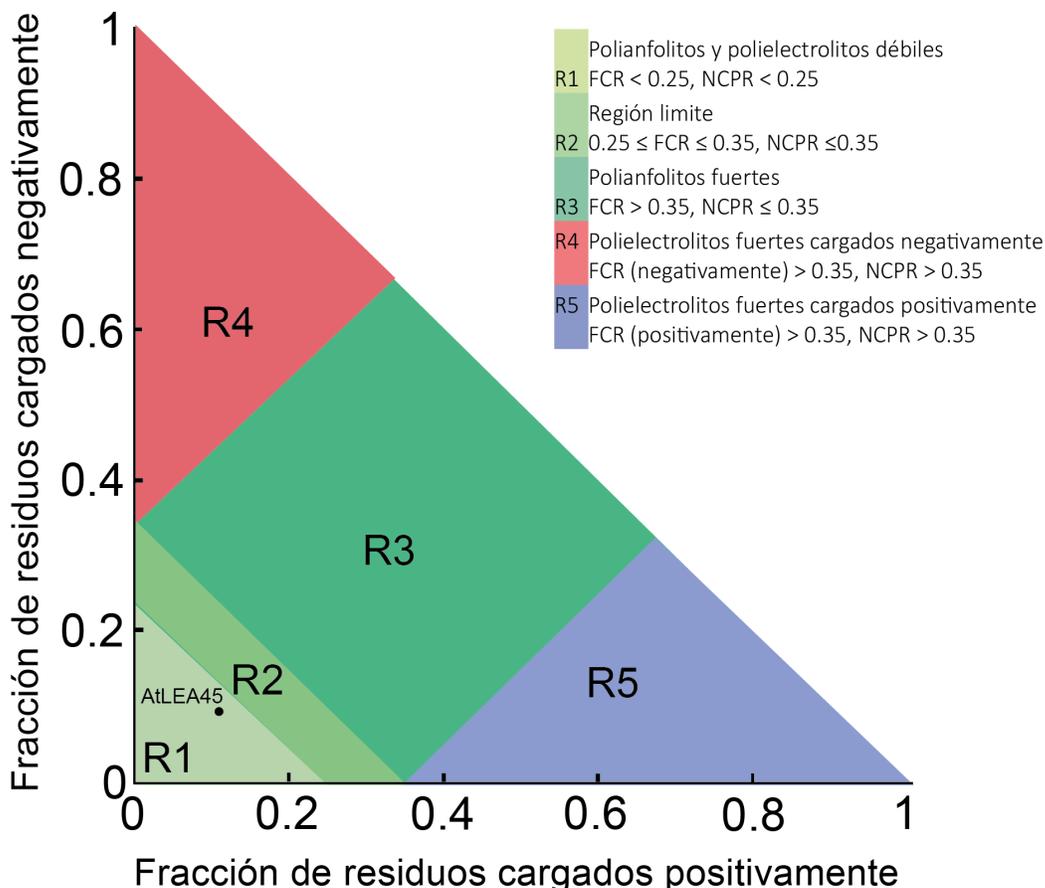
CIDER es una herramienta computacional diseñada para analizar de forma rápida una amplia gama de propiedades fisicoquímicas codificadas en la secuencia de las IDPs. En la **Figura 15** se muestra el resumen de los parámetros calculados con CIDER tras el análisis de la secuencia de AtLEA4-5, incluyen la longitud de la secuencia, el patrón de carga (cuantificado por  $\kappa$ ), la fracción de residuos cargados (FCR), la carga neta por residuo (NCPR), la hidropatía neta y la fracción de residuos que promueven el desorden.

ID	Length	K	FCR	NCPR	Hydropathy	Disorder promoting	Plot region
AtLEA45	158	0.061	0.203	0.025	3.685	0.810	1

**Figura 15.** Parámetros de secuencia calculados por CIDER. (ID) ID de secuencia como se define la entrada. (Length) Número de residuos de la secuencia. (K) El patrón de carga (cuantificado por  $\kappa$ ). (FCR) Fracción de residuos cargados. (NCPR) La carga neta por residuo. (Hydropathy) La hidropatía neta. (Disorder promoting) La fracción de residuos que promueven el desorden. (Plot region) La posición en el diagrama de estados.

Basado en los parámetros calculados, CIDER es capaz de obtener una clasificación de las diferentes clases conformacionales de las IDRs usando los valores de las fracciones de residuos cargados positiva y negativamente. El diagrama de estados (**Figura 16**), posee una clasificación de cinco clases conformacionales donde es posible agrupar a las IDRs para generar las predicciones de las propiedades conformacionales de una secuencia de dadas.

## Diagrama de estados

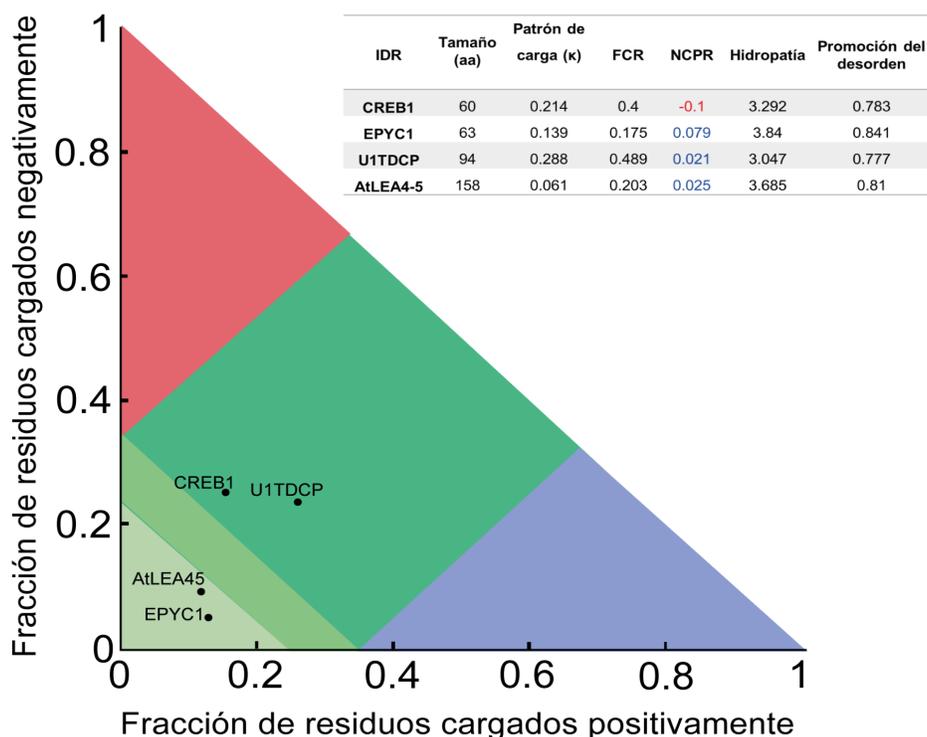


**Figura 16.** Diagrama de estados para la clasificación de las secuencias de proteínas intrínsecamente desordenadas. La secuencia representada, AtLEA4-5, se coloca en el diagrama en función de la fracción de residuos cargado positiva y negativamente. El diagrama de estado se divide en cinco *regiones* (R1 a la R5) (Das y Pappu en 2013).

La *región R1* corresponde a polianfolitos débiles o polielectrolitos débiles que forman glóbulos fundidos que tienden a compactarse. La *región R3* corresponde a polianfolitos fuertes que adquieren conformaciones no globulares en forma de espiral o en forma de horquilla. Por otro lado, la *región R2* corresponde a aquellas conformaciones que tienden a transitar entre los tipos de

conformaciones adoptadas por las secuencias que se clasifican dentro de las regiones 1 y 3. Se ha observado que las proteínas que caen dentro de la región 2 existen en dos o más ensambles de estados conformacionales a las que se les ha denominado proteínas tipo Janus (Madhurima et al., 2021). Y finalmente las secuencias que son designadas a las *regiones R4 y R5* son polielectrolitos fuertes con  $FCR > 0.35$  y  $NCPR > 0.3$ . Se espera que estas secuencias adquieran conformaciones en forma de espiral.

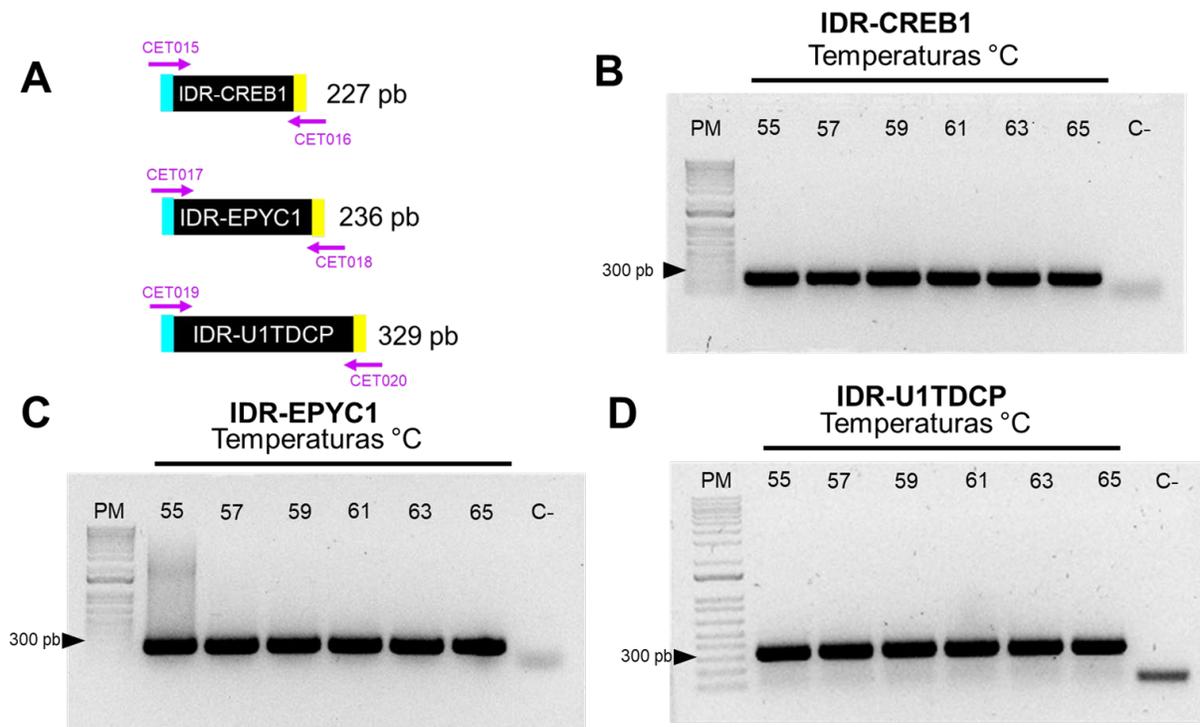
El análisis de las IDR con el servidor de CIDER proporcionó las distribuciones de las secuencias en el diagrama de estados. La IDR de EPYC1 al igual que AtLEA4-5 cae dentro de la clasificación de polianfolitos débiles, mientras que las IDRs de CREB1 y U1TDCP se encuentran en la clasificación de polianfolitos fuertes (**Figura 17**). Cabe señalar que en el análisis de las secuencias con CIDER también se deben tener ciertas reservas ya que esta clasificación no debe tomarse como una prueba directa del comportamiento que tendrá la proteína, es decir, los límites de las diferentes regiones del diagrama no deben tomarse como umbrales estrictos que establecen comportamientos radicalmente diferentes a cada lado de la línea divisoria.



**Figura 17.** Determinación de la posición de las IDRs seleccionadas en el diagrama de estados de CIDER.

## 7.7 Expresión de las regiones intrínsecamente desordenadas de CREB1, EPYC1 y U1TDCP en levadura

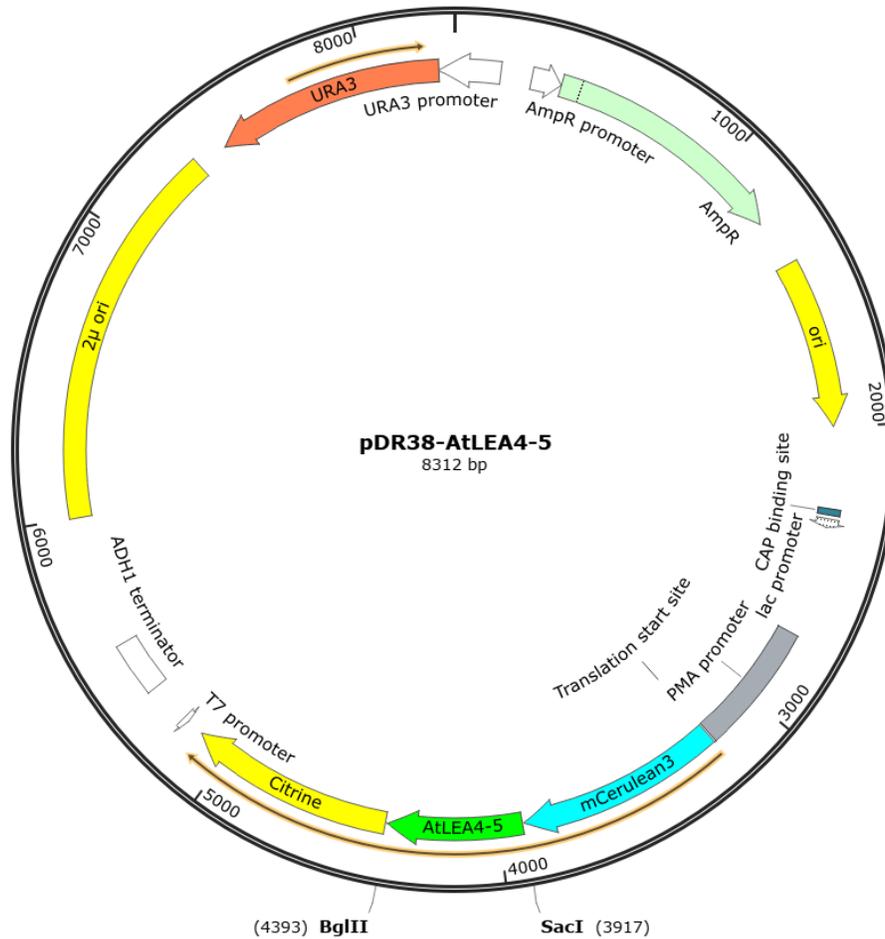
Los plásmidos sintetizados por la compañía IDT que contenían a las IDR de interés se analizaron por PCR con gradientes de temperatura de alineamiento para determinar la mejor condición de amplificación con los oligonucleótidos específicos para cada secuencia. Debido a que los resultados de PCR mostraron una cantidad semejante de producto de amplificación en todas las temperaturas de alineamiento, se determinó 63°C como la temperatura de alineamiento (**Figura 18**). Posteriormente se realizó PCR con Taq polimerasa Platinum II Hot-Start. El producto de PCR (con el peso esperado para cada IDR) se separó en un gel de agarosa 1% y se purificó utilizando el kit de extracción y purificación de banda de Zymo para su posterior clonación.



**Figura 18.** A) Esquema de los ORF sintetizados que codifican para las diferentes IDR. Se indican los oligonucleótidos utilizados para la amplificación desde el plásmido pIDTSMART-AMP y utilizados para su posterior clonación en pDR38. Gel de agarosa 1% de PCR con gradiente de temperatura para CREB1 (B), EPYC1 (C) y U1TDCP (D). PM: marcador de peso molecular de 1 Kb; C-: control negativo. La flecha negra indica la banda de un peso de 300 pb.

A continuación, se realizó la doble restricción del plásmido pDR38-AtLEA4-5 (**Figura 19**) con las enzimas de restricción *SacI* y *BglII* para obtener un plásmido

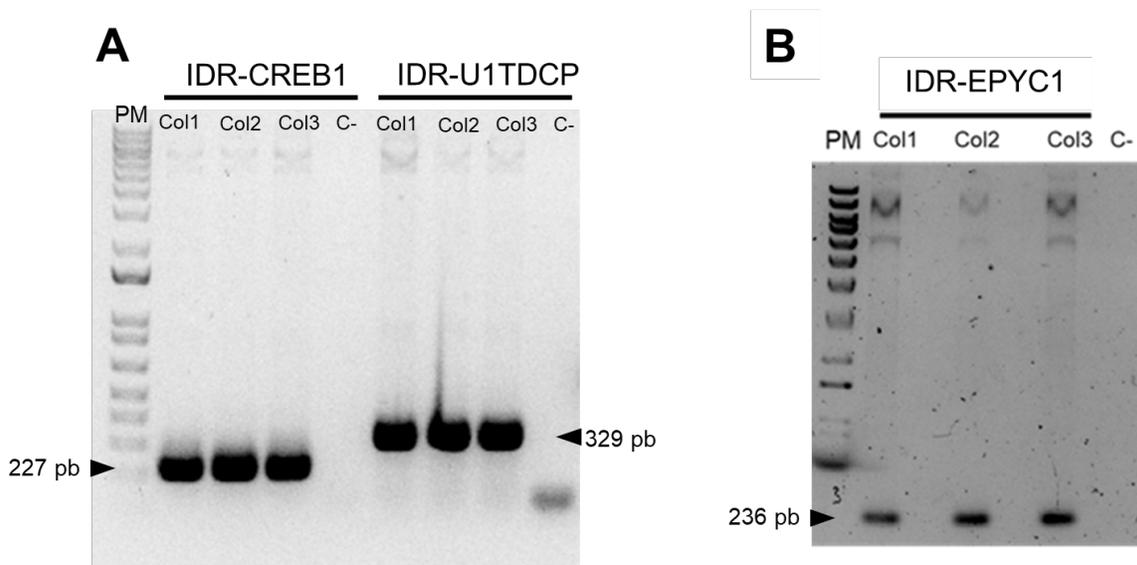
linearizado que contiene el sistema FRET sin la región codificante de la proteína AtLEA4-5. El resultado de la digestión se separó en un gel de agarosa 1% y el plásmido digerido se purificó. Consistentemente con lo esperado, observamos durante la purificación, la banda correspondiente a la AtLEA4-5.



**Figura 19.** Mapa del plásmido pDR38-AtLEA4-5. Se indican los sitios de restricción *Bgl*III y *Sac*I.

Obtenidos los fragmentos de las IDR y el plásmido pDRF38-AtLEA4-5 libre de la región codificante para AtLEA4-5, se procedió a clonar cada IDR en el sistema FRET utilizando el método de Gibson Assembly; el cual se basa en la actividad de las exonucleasas 5, una polimerasa y una ligasa que ensamblan el DNA de forma direccional. El producto de la reacción de ensamblaje de las IDRs dentro del plásmido pDRF38 se transformó en células competentes de *E. coli* One Shot® Mach1™. Cinco colonias candidatas por cada construcción se validaron por PCR de colonia. Posteriormente se extrajo el DNA plasmídico de tres clonas

transformadas con pDR38-IDR-CREB1, pDR38-IDR-EPYC1 e pDR38-IDR-U1TDCP respectivamente. Asimismo, se realizó PCR sobre el DNA plasmídico extraído, el tamaño de los productos de PCR correspondió al tamaño esperado para cada una de las construcciones (**Figura 20**). Posteriormente se secuenció el DNA plasmídico extraído de las tres clonas de cada construcción, para confirmar el correcto ensamblaje de las IDRs en el sistema FRET y la ausencia de mutaciones. Para ello, se solicitó la secuenciación por el método de Sanger en la unidad de síntesis y secuenciación del Instituto de Biotecnología, UNAM. Finalmente se seleccionó la clona que mostró la integridad de la secuencia codificante de interés para cada una de las construcciones, para pDR38-IDR-CREB1 clona 1, pDR38-IDR-EPYC1 clona 3, e pDR38-IDR-U1TDCP clona 1.

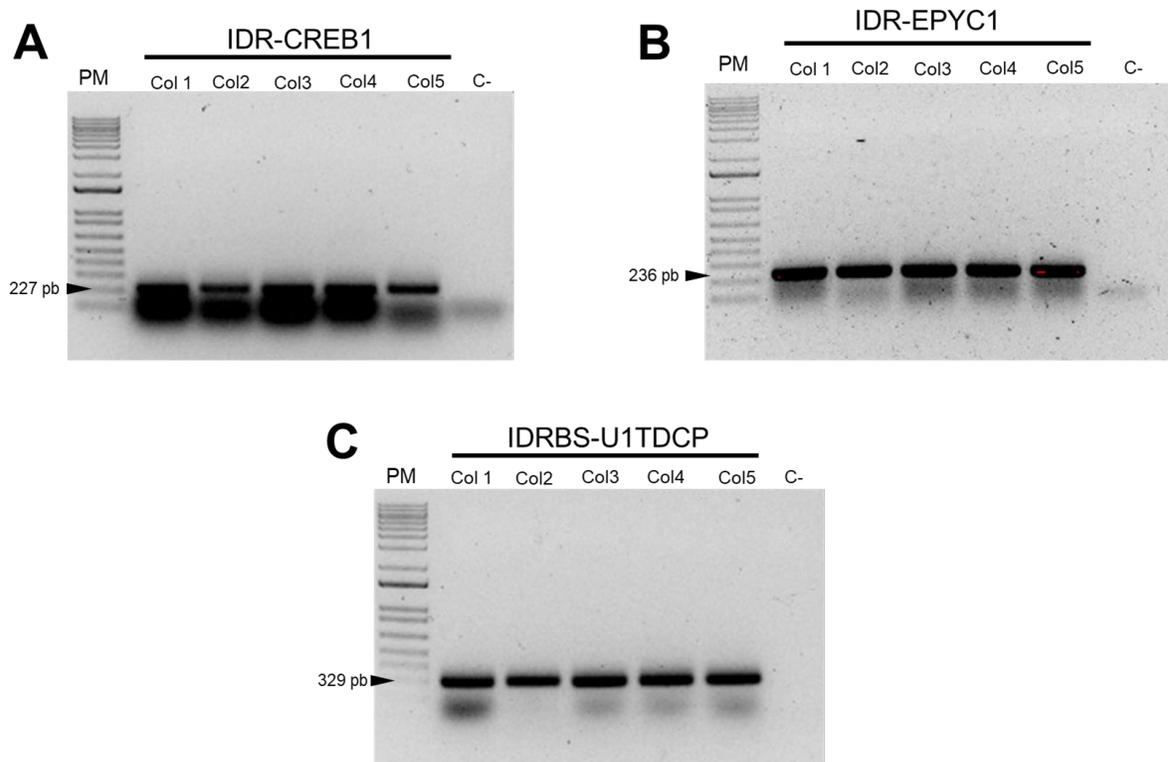


**Figura 20.** A) PCR sobre DNA plasmídico extraído de tres colonias independientes que portan las construcciones IDR-CREB1, IDR-U1TDCP dentro del sistema FRET. Se observan los productos correspondientes al tamaño esperado de 227 y 329 pb respectivamente. B) PCR sobre DNA plasmídico de tres colonias independientes que portan la construcción IDR-EPYC1 dentro del sistema FRET. Se observan los productos correspondientes al tamaño esperado de 236 pb. Gel de agarosa 1%. PM: marcador de peso molecular de 1 Kb; C-: control negativo.

En el siguiente paso las células de levadura se transformaron con los DNAs plasmídicos corroborados por secuenciación de cada construcción. Posteriormente se realizó PCR sobre las colonias de levadura. El tamaño de los productos de PCR correspondió al tamaño esperado de cada una de las construcciones. En el caso de la construcción de pDR38-IDR-CREB1, se observó amplificación del producto esperado de 227 pb en las cinco colonias de levaduras probadas. De igual manera para las colonias de levadura transformadas con

pDR38-IDR-EPYC1 e pDR38-IDR-U1TDCP se obtuvieron productos de 236 pb y 329 bp respectivamente (**Figura 21**).

Por último, se seleccionó una sola colonia de levadura por construcción para realizar los ensayos de sensibilidad estructural. Para pDR38-IDR-CREB1 se seleccionó la colonia cinco, para pDR38-IDR-EPYC1 la colonia dos y para pDR38-IDR-U1TDCP se seleccionó la colonia dos.



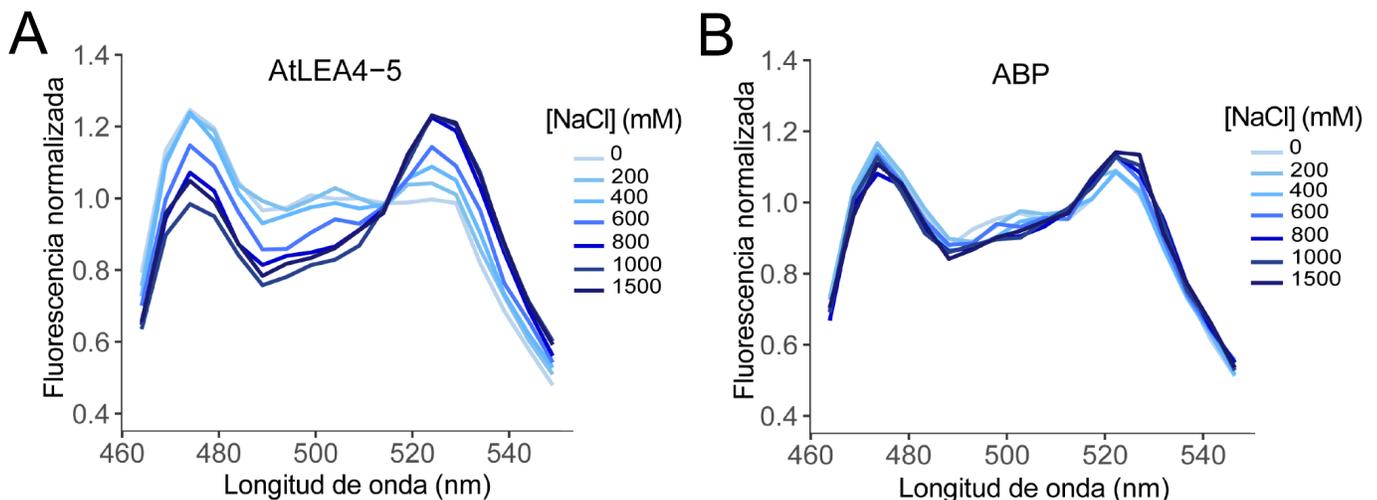
**Figura 21.** PCR sobre colonias de *S. cerevisiae* cepa BJ5465 transformadas con las construcciones IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP dentro del sistema FRET. A) Se observan los productos correspondientes al tamaño esperado de 227 pb de IDR-CREB1 en las cinco colonias. B) Se observa la amplificación del producto esperado de 236 pb para las cinco colonias transformadas con IDR-EPYC1. C) Se observa la amplificación en las cinco colonias para el producto del tamaño esperado de 329 pb. El control negativo no amplificó en ninguno de los casos. Gel de agarosa 1%. PM: marcador de peso molecular de 1 Kb; C-: control negativo.

### 7.8 Determinación de la sensibilidad estructural de las regiones intrínsecamente desordenadas de CREB1, EPYC1 y U1TDCP

Con las construcciones pDR38-IDR-CREB1, pDR38-IDR-EPYC1 e pDR38-IDR-U1TDCP confirmadas en clonas de levadura se procedió a determinar la

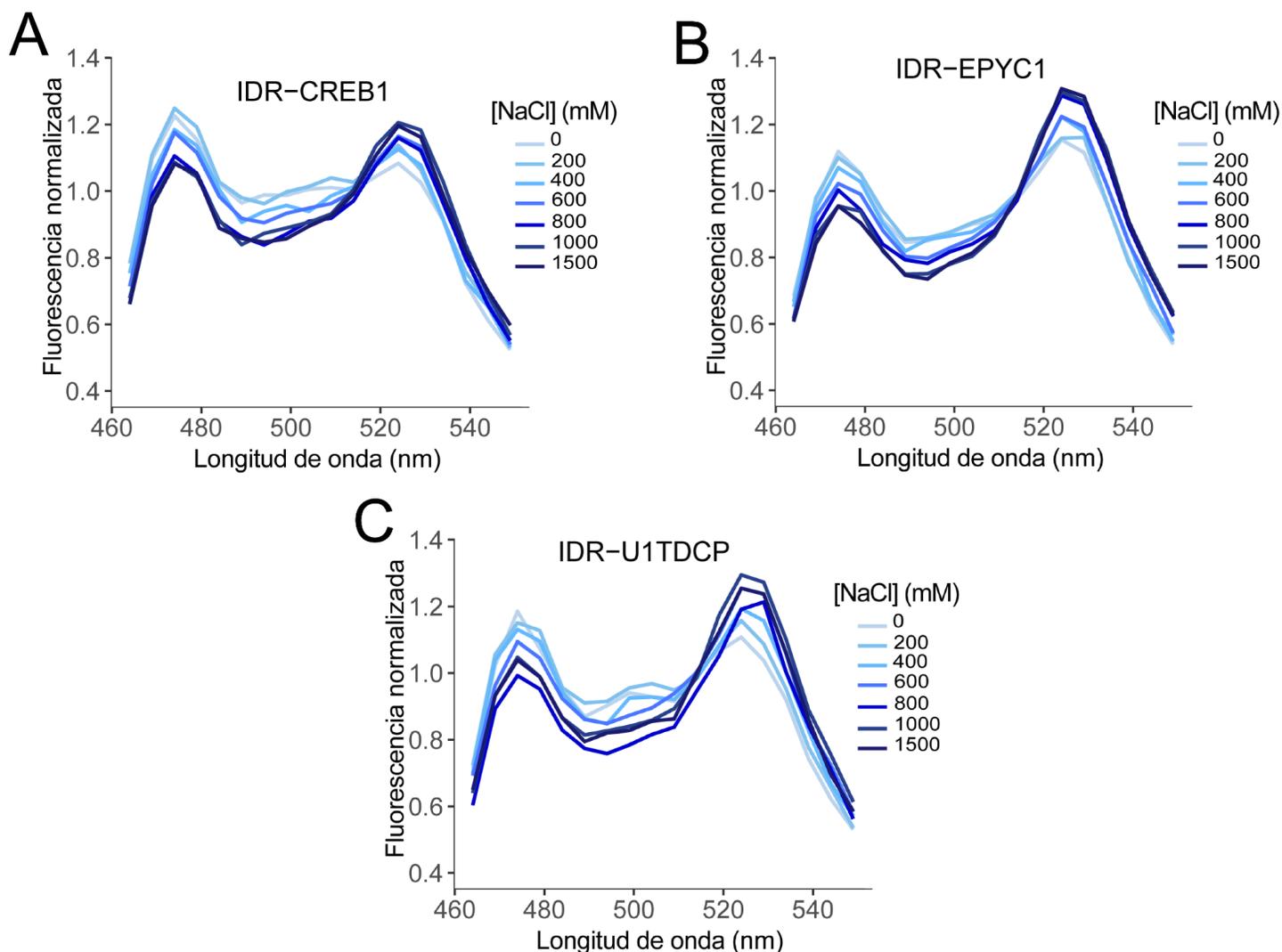
sensibilidad estructural de las IDRs a cambios en el ambiente osmótico *in vivo*. Se realizaron mediciones del espectro de emisión de fluorescencia de las células de levadura que expresan las construcciones sometidas a tratamientos de estrés hiperosmótico con NaCl. Las referencias utilizadas fueron la construcción pDR38-AtLEA4-5, que tiene altos niveles de sensibilidad, y la construcción pDR38-ABP que consistió en la proteína globular, llamada proteína de unión a arabinosa (ABP, *arabinose binding protein*), la cual posee poca sensibilidad estructural (Cuevas-Velazquez et al., 2021).

El espectro de emisión de fluorescencia es normalizado dividiendo los valores de emisión de fluorescencia relativa obtenidos entre el valor a 515 nm. De esta forma se observa en la construcción pDR38-AtLEA4-5, el pico de emisión de fluorescencia del donador (mCerulean3, 475 nm) disminuye al incrementar las concentraciones de NaCl, con el concomitante incremento del pico de emisión de fluorescencia del aceptor (Citrina, 525 nm) (**Figura 22A**). Este tipo de comportamiento se espera en un sistema de transferencia de energía tipo FRET reportado previamente por Cuevas-Velázquez y colaboradores (2021). En contraste y de acuerdo con lo esperado, la ABP mostró cambios sutiles en respuesta al choque hiperosmótico (**Figura 22B**). Estas observaciones sugieren que el cambio estructural de la IDR AtLEA4-5 induce el acercamiento espacial del donador con respecto al aceptor, incrementando la transferencia de energía en respuesta al choque hiperosmótico inducido con NaCl *in vivo*, consistente con lo reportado previamente (Cuevas-Velazquez et al., 2021).



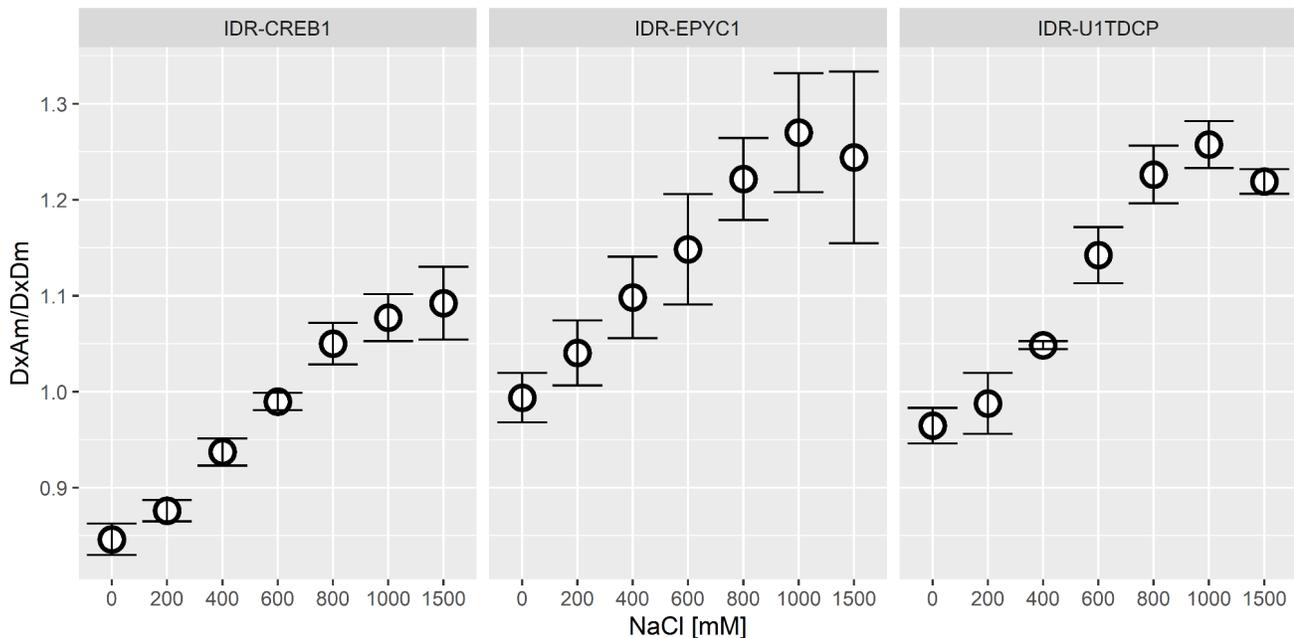
**Figura 22.** Espectros de emisión de fluorescencia normalizada a 515 nm de células de levadura vivas que expresan la construcción AtLEA4-5 (izquierda) y la construcción ABP (derecha), tratadas con 0 mM, 200 mM, 400 mM, 600 mM, 800 mM, 1000 mM y 1500 mM de NaCl.

Tras demostrar que los tratamientos probados indujeron una respuesta de transferencia de energía en la construcción de referencia pDR38-AtLEA4-5, y no en pDR38-ABP, se procedió a someter a las células de levadura que expresaron pDR38-IDR-CREB1, pDR38-IDR-EPYC1 e pDR38-IDR-U1TDCP al choque hiperosmótico inducido por el incremento en la concentración de NaCl. Las tres construcciones mostraron un comportamiento tipo FRET, al mostrar un aumento en la intensidad de fluorescencia normalizada del aceptor tras la excitación del donador, con la concomitante disminución en la intensidad de fluorescencia normalizada del donador (Figura 23).



**Figura 23.** Espectros de emisión de fluorescencia normalizada a 515 nm de células de levadura vivas tratadas con 0 mM, 200 mM, 400 mM, 600 mM, 800 mM, 1000 mM y 1500 mM de NaCl, que expresan la construcción (A) IDR-CREB1, (B) IDR-EPYC1, (C) IDR-U1TDCP.

Con la finalidad de cuantificar los niveles de sensibilidad estructural de las tres IDR estudiadas, se cuantificó el cambio en la razón de FRET, el cual se obtuvo dividiendo la emisión del aceptor a 525 nm entre la emisión del donador a 475 nm (DxAm/DxDm). Estas mediciones permitieron determinar que la sensibilidad estructural al choque hiperosmótico *in vivo* de las IDRs; CREB1, EPYC1 y U1TDCP fue significativamente mayor ( $p < 0.001$ ) a la sensibilidad estructural de la referencia ABP (**Figura 24**), como se observó en los espectros de emisión de fluorescencia (**Figura 22 y Figura 23**). En conjunto, estos datos muestran que variaciones en el ambiente osmótico intracelular de *S. cerevisiae* inducen una compactación en la estructura de las IDRs de CREB1, EPYC1 y U1TDCP.



**Figura 24.** Razón de FRET (DxAm/DxDm) de las células de levadura vivas que expresan a las IDRs estudiadas. Las células se trataron con 0 mM, 200 mM, 400 mM, 600 mM, 800 mM, 1000 mM y 1500 mM de NaCl. Cada condición se midió por triplicado y se realizaron cuatro mediciones independientes para cada construcción. Se muestran las medias  $\pm$  desviación estándar de 3 experimentos independientes.

## VIII. Discusión

Las IDRs se presentan en un ensamble de conformaciones que pueden transitar hacia una estructura secundaria más ordenada en respuesta al ambiente fisicoquímico. El trabajo experimental que sustenta esta aseveración ha sido realizado principalmente *in vitro* (Holehouse y Sukenink, 2020; Moses et al., 2020; Cuevas-Velazquez et al., 2021), proporcionando valiosa información del comportamiento estructural de las IDRs bajo condiciones controladas. Sin embargo, poco se sabe sobre si el comportamiento descrito *in vitro* refleja el comportamiento de las IDRs dentro del contexto celular. Para contribuir en la respuesta de este cuestionamiento en este trabajo reportamos el desarrollo del parámetro de dualidad, capaz de predecir la sensibilidad estructural de las IDRs ante los cambios en el ambiente osmótico intracelular. Así como la caracterización de la sensibilidad estructural de tres IDRs; CREB1, EPYC1 y U1TDCP con el potencial de formar hélice  $\alpha$  ante tratamientos por choque osmótico en células vivas de levadura utilizando el sistema FRET.

El desarrollo del parámetro de dualidad nos permitió seleccionar tres IDRs (IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP) para evaluar la sensibilidad estructural *in vivo* ante los tratamientos de choque hiperosmótico en células de levadura. Nuestros datos muestran que las IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP presentan sensibilidad estructural a diferentes concentraciones de NaCl sugiriendo que la predicción dual, que considera a la formación de estructuras tipo hélice  $\alpha$  y desorden, es capaz de predecir la transición estructural de un estado extendido a uno compacto de las regiones intrínsecamente desordenadas en respuesta al cambio en el ambiente osmótico intracelular. Resulta interesante que la IDR-U1TDCP cuya predicción de dualidad es mayor, sin embargo, presentó un comportamiento similar al de IDR-EPYC1 de menor dualidad, indicando que nuestro parámetro propuesto no es suficiente para predecir el rango de sensibilidad estructural de las IDRs ante los efectos del aumento en la osmolaridad *in vivo*. Estas observaciones sugieren que existen otras propiedades además de la propensión de formación de hélice  $\alpha$  y la alta predicción al desorden en las IDRs de CREB1, EPYC1 y U1TDCP que dirigen la compactación de las secuencias durante la alteración en el ambiente osmótico intracelular.

Por otro lado, el análisis de las propiedades de las secuencias de IDRs analizadas con CIDER reflejan la complejidad de predecir la compactación de las IDRs en respuesta al cambio en el ambiente fisicoquímico. El comportamiento conformacional de las IDR estudiadas de acuerdo con su clasificación en el diagrama de estados de CIDER no fue el esperado. En el caso de la IDR-EPYC1 al ser un polianfolito débil se esperaba que tendiera a ser más compacto con

respecto a las IDR-CREB1 e IDR-U1TDCP que se clasificaron como polianfolitos fuertes y de acuerdo con otros trabajos las secuencias que se clasifican dentro de la R3 tienden a adquirir estructuras desplegadas (Ginell et al., 2020). Sin embargo, nuestros datos *in vivo* indican que tanto IDR-EPYC1 como IDR-U1TDCP poseen una compactación similar en respuesta a los tratamientos de choque hiperosmótico inducido por NaCl.

Los ensayos de sensibilidad estructural en las células vivas de levadura mostraron cambios en el cociente de la señal de FRET en respuesta a los tratamientos por choque osmótico en las células expresoras de las construcciones de pDR38-IDR-CREB1, pDR38-IDR-EPYC1 e pDR38-IDR-U1TDCP en comparación con las células expresoras de la proteína de referencia ABP (reportada previamente como poco sensible al tratamiento por choque osmótico *in vivo* (Cuevas-Velazquez et al., 2021).

En primer lugar, nuestras observaciones indican que *in vivo* y en ausencia de choque osmótico, las proteínas fluorescentes de las construcciones de pDR38-IDR-CREB1, pDR38-IDR-EPYC1 e pDR38-IDR-U1TDCP están lo suficientemente cercanas para transferir la energía desde el donador hacia el aceptor, dando como resultado una señal de FRET basal en comparación con los espectros de fluorescencia de las células expresoras de la pDR38-AtLEA4-5 donde no se observa este comportamiento. Es posible que las diferencias en las longitudes de las IDRs: IDR-CREB1 (60 aa), IDR-EPYC1 (63 aa) e IDR-U1TDCP (94 aa) comparado con los 158 residuos que conforman a la AtLEA4-5 podrían influir en la señal inicial de FRET basal en ausencia de tratamiento.

En segundo lugar, la magnitud del cambio en el cociente FRET difiere entre las construcciones IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP debido posiblemente a que las IDRs poseen diferentes estados desordenados iniciales, influidos por la topología y el acomodo de los aminoácidos, así como la interacción con moléculas en el medio intracelular. Cabe mencionar que, de las IDRs estudiadas en el presente trabajo, solo se ha demostrado el estado desordenado *in vitro* de la IDR-CREB1 (Radhakrishnan et al., 1998), en contraste con la IDR-EPYC1 y la IDR-U1TDCP cuyos antecedentes de desorden estructural se sustenta en predicciones *in silico* (Freeman Rosenzweig et al., 2017). El nivel de cambio en el cociente de FRET observado en la IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP podría estar influido por otros efectos asociados a los tratamientos de choque osmótico. Uno de los efectos resultantes del cambio en la osmolaridad intracelular es el incremento del hacinamiento molecular dentro de la célula, lo que puede promover la compactación de las IDRs o incluso no alterar su estructura conformacional (Fonin et al., 2018). Debido a que algunas IDR tienen preferencias estructurales

que pueden ser susceptibles incluso a los cambios leves en el entorno intracelular, mientras que otras IDR pueden mostrar una notable resistencia estructural (Moses et al., 2021).

Un hallazgo importante de este trabajo es el rango de sensibilidad estructural de las IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP ante los tratamientos de choque hiperosmótico. Mientras que el cociente de FRET de la IDR-CREB1 aumenta de forma sutil ante las concentraciones crecientes de NaCl, la IDR-EPYC1 y la IDR-U1TDCP presentan un cambio significativo desde 0.5 M de NaCl a 1 M de NaCl. Ante 1.5 M de NaCl la IDR-U1TDCP comienza a disminuir el cociente FRET sugiriendo el acotamiento de su rango de sensibilidad estructural comparado con la IDR-CREB1 y la IDR-EPYC1, las cuales de acuerdo con nuestras observaciones podrían extender su sensibilidad estructural hacia tratamientos con mayores concentraciones de NaCl.

Finalmente resulta interesante la posibilidad de explotar la diversidad de la sensibilidad estructural observada en IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP para el diseño de biosensores activados por el ambiente osmótico que puedan ser expresados de forma heteróloga en organismos eucariotas. El uso de biosensores basados en IDRs nos permitiría contestar preguntas fundamentales de la biología celular, como por ejemplo ¿cómo es que las perturbaciones en el ambiente repercuten en los diferentes aspectos fisicoquímicos de la biología intracelular?

## **IX. Conclusión**

Las IDR's CREB1, EPYC1 y U1TDCP presentan niveles de sensibilidad estructural diferencial ante las variaciones en el ambiente osmótico intracelular de la levadura por tratamientos de NaCl. El sistema FRET sugiere que la transición desde el estado inicial extendido de las IDR's a uno compacto aminora las distancias entre los extremos de las construcciones, lo cual podría estar influenciado por el tamaño de la IDR, el acomodo de los residuos y la topología que adoptan las construcciones ante los tratamientos. Además de otros factores, como la interacción con otras moléculas intracelulares y los efectos concomitantes del choque hiperosmótico como el hacinamiento molecular.

Por otro lado, la propensión de formación de hélice  $\alpha$  y el desorden no es suficiente para predecir la sensibilidad estructural de las IDR's ante los cambios en el ambiente, sugiriendo que existen otras características particulares en las secuencias que determinan la respuesta al ambiente osmótico intracelular.

Nuestro estudio apunta a que es necesario refinar el parámetro tomando en cuenta otras propiedades de las secuencias para predecir la compactación de las IDR's. Como ejemplo, tomar en cuenta el número y extensión de las regiones con predicción dual, así como el posible acomodo de las hélices  $\alpha$  y las topologías que adoptan las IDR's en un contexto celular. Adicionalmente sería conveniente analizar la composición de aminoácidos de las regiones duales e integrar en el parámetro de dualidad el patrón de cargas, la fracción de residuos cargados, la carga neta por residuo y la hidropatía (resultados obtenidos con CIDER) para hacer más robusto el parámetro de dualidad.

## **IX. Perspectivas**

Evaluar de forma experimental el desorden estructural de la IDR-EPYC1 y la IDR-U1TDCP, así como la capacidad de adquisición de estructura secundaria de hélice  $\alpha$  inducida por agentes como glicerol o polietilen glicol, a través de ensayos *in vitro*.

Evaluar si el efecto en la compactación de las IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP es dependiente del tipo de osmolito utilizado en los ensayos de choque osmótico.

Evaluar la transición a estados conformacionales compactos, ya sean hélices  $\alpha$  u otras estructuras secundarias como hojas  $\beta$  que podrían conducir a estados más compactos de las IDRs.

Analizar la sensibilidad estructural de las IDRs en un contexto de proteína nativa.

Explotar los diferentes rangos de sensibilidad estructural demostrada de las IDR-CREB1, IDR-EPYC1 e IDR-U1TDCP para la generación de biosensores fluorescentes codificados genéticamente que nos permitan monitorear cambios el ambiente intracelular en tiempo real y de forma no destructiva en múltiples organismos.

## IX. Referencias

- Basak, S., Saikia, N., Dougherty, L., Guo, Z., Wu, F., Mindlin, F., Lary, J. W., Cole, J. L., Ding, F., & Bowen, M. E. (2021). Probing Interdomain Linkers and Protein Supertertiary Structure In Vitro and in Live Cells with Fluorescent Protein Resonance Energy Transfer. *Journal of Molecular Biology*, 433(5), 166793. doi.org/10.1016/j.jmb.2020.166793
- Belott, C., Janis, B. & Menze, M. A. (2020). Liquid-liquid phase separation promotes animal desiccation tolerance. *Proceedings of the National Academy of Sciences*, 117(44), 27676-27684. doi.org/10.1073/pnas.2014463117
- Bourque, C. W. (2008). Central mechanisms of osmosensation and systemic osmoregulation. *Nature Reviews Neuroscience*, 9(7), 519–531. doi-org.pbidi.unam.mx:2443/10.1038/nrn2400
- Burgess, S. M., Powers, T., & Mell, J. C. (2017). Budding Yeast *Saccharomyces cerevisiae* as a Model Genetic Organism. *eLS*, 1-12. doi.org/10.1002/9780470015902.a0000821.pub2
- Campbell, R. E. (2009). Fluorescent-Protein-Based Biosensors: Modulation of Energy Transfer as a Design Principle. *Analytical Chemistry*, 81(15). 5972-9. doi: 10.1021/ac802613w.
- Cuevas-Velazquez, C. L., Covarrubias, A. A. (2011). Las proteínas desordenadas y su función: Una nueva forma de ver la estructura de las proteínas y la respuesta de las plantas al estrés. *Tip Revista Especializada en Ciencias Químico-Biológicas*, 14(2), 97-105.
- Cuevas-Velazquez, C. L., Saab-Rincón, G., Reyes, J. L., & Covarrubias, A. A. (2016). The unstructured n-terminal region of Arabidopsis group 4 late embryogenesis abundant (LEA) proteins is required for folding and for chaperone-like activity under water deficit. *Journal of Biological Chemistry*, 291(20), 10893–10903. doi.org/10.1074/jbc.M116.720318
- Cuevas-Velazquez, C. L., Vellosillo, T., Guadalupe, K., Schmidt, H. B., Yu, F., Moses, D., Brophy, J. A. N., Cosio-Acosta, D., Das, A., Wang, L., Jones, A. M., Covarrubias, A. A., Sukenik, S. & Dinneny, J. R. (2021). Intrinsically disordered protein biosensor tracks the physical-chemical effects of osmotic stress on cells. *Nature Communications*, 12(1). doi.org/10.1038/s41467-021-25736-8
- Das, R. K., Ruff, K. M. & Pappu, R. V. (2015, junio). Relating sequence encoded information to form and function of intrinsically disordered proteins. *Current Opinion in Structural Biology*, 32, 102-112. doi.org/10.1016/j.sbi.2015.03.008
- Dyson, H. J., & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology*, 6(3) 197–208. doi.org/10.1038/nrm1589
- Eliezer, D. & Palmer, A. G. (2007) Proteins hunt and gather. *Nature*. *Nature*, 447(7147), 920-921. doi.org/10.1038/447920a

Ellis, R. J. (2001). Macromolecular crowding: An important but neglected aspect of the intracellular environment. *Current Opinion in Structural Biology* 11 (1),114–119. doi.org/10.1016/S0959-440X(00)00172-X

Freeman Rosenzweig, E. S., Xu, B., Kuhn Cuellar, L., Martinez-Sanchez, A., Schaffer, M., Strauss, M., Cartwright, H. N., Ronceray, P., Pitzko, J. M., Förster, F., Wingreen, N. S., Engel, B. D., Mackinder, L. C. M., & Jonikas, M. C. (2017). The Eukaryotic CO<sub>2</sub>-Concentrating Organelle Is Liquid-like and Exhibits Dynamic Reorganization. *Cell*, 171(1), 148–162. doi-org.pbidi.unam.mx:2443/10.1016/j.cell.2017.08.008

Fonin, A. V., Darling, A. L., Kuznetsova, I. M., Turoverov, K. K., & Uversky, V. N. (2018). Intrinsically disordered proteins in crowded milieu: when chaos prevails within the cellular gumbo. *Cellular and molecular life sciences*, 75(21), 3907–3929. doi.org/10.1007/s00018-018-2894-9

Gianni, S., & Jemth, P. (2019). Affinity versus specificity in coupled binding and folding reactions. *Protein Engineering, Design and Selection*, 32(8), 355–357. doi.org/10.1093/protein/gzz020

Ginell, G. M. & Holehouse, A. S. (2020). Analyzing the Sequences of Intrinsically Disordered Regions with CIDER and localCIDER. *Methods in Molecular Biology*, 103-126. https://doi.org/10.1007/978-1-0716-0524-0\_5

Greenwald, E. C., Mehta, S., & Zhang, J. (2018). Genetically encoded fluorescent biosensors illuminate the spatiotemporal regulation of signaling networks. *Chemical Reviews*, 118(24), 11707–11794. doi.org/10.1021/acs.chemrev.8b00333

Guillén-Boixet, J., Kopach, A., Holehouse, A. S., Wittmann, S., Jahnel, M., Schlüßler, R., Kim, K., Trussina, I. R. E. A., Wang, J., Mateju, D., Poser, I., Maharana, S., Ruer-Gruß, M., Richter, D., Zhang, X., Chang, Y.-T., Guck, J., Honigmann, A., Mahamid, J., Franzmann, T. M. (2020). RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. *Cell*, 181(2), 346–361. doi.org/10.1016/j.cell.2020.03.049

Gutierrez, J. I., Brittingham, G. P., Karadeniz, Y., Tran, K. D., Dutta, A., Holehouse, A. S., Peterson, C. L. & Holt, L. J. (2022). SWI/SNF senses carbon starvation with a pH-sensitive low-complexity sequence. *eLife*, 11. doi.org/10.7554/elife.70344

Hashimoto, T., Horikawa, D. D., Saito, Y., Kuwahara, H., Kozuka-Hata, H., Shin-I, T., Minakuchi, Y., Ohishi, K., Motoyama, A., Aizu, T., Enomoto, A., Kondo, K., Tanaka, S., Hara, Y., Koshikawa, S., Sagara, H., Miura, T., Yokobori, S. I., Miyagawa, K., . . . Kunieda, T. (2016). Extremotolerant tardigrade genome and improved radiotolerance of human cultured cells by tardigrade-unique protein. *Nature Communications*, 7(1). https://doi.org/10.1038/ncomms12808

Hersen, P., McClean, M. N., Mahadevan, L., & Ramanathan, S. (2008). Signal Processing by the HOG MAP Kinase Pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 105(20), 7165–7170. doi.org/10.1073/pnas.0710770105 Hohmann,

- S. (2009). Control of high osmolarity signalling in the yeast *Saccharomyces cerevisiae*. *FEBS Letters*, 583 (24), 4025–4029. doi.org/10.1016/j.febslet.2009.10.069
- Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. & Pappu, R. V. (2017). CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophysical Journal*, 112(1), 16-21. https://doi.org/10.1016/j.bpj.2016.11.3200
- Holehouse, A. S., & Sukenik, S. (2020). Controlling Structural Bias in Intrinsically Disordered Proteins Using Solution Space Scanning. *Journal of Chemical Theory and Computation*, 16(3), 1794–1805. doi.org/10.1021/acs.jctc.9b00604
- Hosseiniyan Khatibi, S. M., Zununi Vahed, F., Sharifi, S., Ardalan, M., Mohajel Shoja, M. & Zununi Vahed, S. (2019). Osmolytes resist against harsh osmolarity: Something old something new. *Biochimie*, 158, 156-164. https://doi.org/10.1016/j.biochi.2019.01.002
- Howell, A. S., & Lew, D. J. (2012). Morphogenesis and the cell cycle. *Genetics*, 190(1), 51–77. doi.org/10.1534/genetics.111.128314
- Ibraheem, A. & Campbell, R. E. (2010). Designs and applications of fluorescent protein-based biosensors. *Current Opinion in Chemical Biology*, 14(1), 30-36. https://doi.org/10.1016/j.cbpa.2009.09.033
- Klipp, E., Nordlander, B., Krüger, R., Gennemark, P. & Hohmann, S. (2005). Integrative model of the response of yeast to osmotic shock. *Nature Biotechnology*, 23(8), 975-982. https://doi.org/10.1038/nbt1114
- Kovacs D., Kalmar E., Torok Z., & Tompa P. (2008). Chaperone Activity of ERD10 and ERD14, Two Disordered Stress-Related Plant Proteins. *Plant Physiology*, 147(1), 381–390. doi-org.pbidi.unam.mx:2443/10.1104/pp.108.118208
- Kulkarni, P., & Uversky, V. N. (2019). Intrinsically Disordered Proteins in Chronic Diseases. *Biomolecules*, 9(4), 147. doi.org/10.3390/biom9040147
- Lakowicz, J. R. (2006). Principles of fluorescence spectroscopy. In *Principles of Fluorescence Spectroscopy*. Springer. doi.org/10.1007/978-0-387-46312-4
- Lang, F., Busch, G. L., Ritter, M., Völkl, H., Waldegger, S., Gulbins, E., & Häussinger, D. (1998). Functional significance of cell volume regulatory mechanisms. *Physiological Reviews*, 78 (1), 247–306. doi.org/10.1152/physrev.1998.78.1.247
- Li, Y., Konstantopoulos, K., Zhao, R., Mori, Y., & Sun, S. X. (2020). The importance of water and hydraulic pressure in cell dynamics. *Journal of cell science* 133(20). doi.org/10.1242/jcs.240341

- Liu, N., Guo, Y., Ning, S. & Duan, M. (2020). Phosphorylation regulates the binding of intrinsically disordered proteins via a flexible conformation selection mechanism. *Communications Chemistry*, 3(1). <https://doi.org/10.1038/s42004-020-00370-5>
- Liu, J., Perumal, N. B., Oldfield, C. J., Su, E. W., Uversky, V. N., & Dunker, A. K. (2006). Intrinsic disorder in transcription factors. *Biochemistry*, 45(22), 6873–6888. [doi.org/10.1021/bi0602718](https://doi.org/10.1021/bi0602718)
- Madhurima, K., Nandi, B., & Sekhar, A. (2021). Metamorphic proteins: the Janus proteins of structural biology. *Open Biology*, 11(4), 210012. [doi.org/10.1098/rsob.210012](https://doi.org/10.1098/rsob.210012)
- Mackinder, L. C. M., Chen, V. K., Rosenzweig, E. S. F., Pallesen, L., Reeves, G., Itakura, A., Jonikas, M. C., Meyer, M. T., Mitchell, M. C., Caspari, O., Griffiths, H., Mettler-Altmann, T., Sommer, F., Mühlhaus, T., Schroda, M., Stitt, M., Roth, R., Goodenough, U., & Geimer, S. (2016). A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proceedings of the National Academy of Sciences of the United States of America*, 113(21), 5958–5963. [doi.org/10.1073/pnas.1522866113](https://doi.org/10.1073/pnas.1522866113)
- Miermont, A., Waharte, F., Hu, S., McClean, M. N., Bottani, S., Léon, S. & Hersen, P. (2013). Severe osmotic compression triggers a slowdown of intracellular signaling, which can be explained by molecular crowding. *Proceedings of the National Academy of Sciences*, 110(14), 5725–5730. [doi.org/10.1073/pnas.1215367110](https://doi.org/10.1073/pnas.1215367110)
- Mika, J. T., & Poolman, B. (2011). Macromolecule diffusion and confinement in prokaryotic cells. In *Current Opinion in Biotechnology*, 22(1), 117–126. [doi.org/10.1016/j.copbio.2010.09.009](https://doi.org/10.1016/j.copbio.2010.09.009)
- Moses, D., Yu, F., Ginell, G. M., Shamoan, N. M., Koenig, P. S., Holehouse, A. S., & Sukenik, S. (2020). Probing the hidden sensitivity of intrinsically disordered proteins to their chemical environment. *bioRxiv* (p. 2020.08.17.252478). [bioRxiv. doi.org/10.1101/2020.08.17.252478](https://doi.org/10.1101/2020.08.17.252478)
- Moses, D., Yu, F., Ginell, G. M., Shamoan, N. M., Koenig, P. S., Holehouse, A. S., & Sukenik, S. (2020). Revealing the Hidden Sensitivity of Intrinsically Disordered Proteins to their Chemical Environment. *Journal of Physical Chemistry Letters*, 11(23), 10131–10136. [doi.org/10.1021/acs.jpcclett.0c02822](https://doi.org/10.1021/acs.jpcclett.0c02822)
- Moses, D., Guadalupe, K., Yu, F., Flores, E., Perez, A., McAnelley, R., Shamoan, N., Cuevas-Zepeda, E., Merg, A., Martin, E., Holehouse, A., & Sukenik, S. (2021). Hidden Structure in Disordered Proteins is Adaptive to Intracellular Changes. *bioRxiv*; 2021. DOI: [10.1101/2021.11.24.469609](https://doi.org/10.1101/2021.11.24.469609).
- Nishimura, A., Yamamoto, K., Oyama, M., Kozuka-Hata, H., Saito, H., & Tatebayashi, K. (2016). Scaffold Protein Ahk1, Which Associates with Hkr1, Sho1, Ste11, and Pbs2, Inhibits Cross Talk Signaling from the Hkr1 Osmosensor to the Kss1 Mitogen-Activated Protein Kinase. *Molecular and cellular biology*, 36(7), 1109–1123. [doi.org/10.1128/MCB.01017-15](https://doi.org/10.1128/MCB.01017-15)

Oldfield, C. J. & Dunker, A. K. (2014). Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annual Review of Biochemistry*, 83(1), 553-584. doi.org/10.1146/annurev-biochem-072711-164947

Olvera-Carrillo, Y., Campos, F., Reyes, J. L., Garcarrubio, A., & Covarrubias, A. A. (2010). Functional analysis of the group 4 late embryogenesis abundant proteins reveals their relevance in the adaptive response during water deficit in arabidopsis. *Plant Physiology*, 154(1), 373–390. doi.org/10.1104/pp.110.158964

Osumi, M. (2012). Visualization of yeast cells by electron microscopy. *Microscopy*, 61(6), 343-365. doi.org/10.1093/jmicro/dfs082

Peng, Z., Yan, J., Fan, X., Mizianty, M. J., Xue, B., Wang, K., Hu, G., Uversky, V. N. & Kurgan, L. (2014). Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. *Cellular and Molecular Life Sciences*, 72(1), 137-151. doi.org/10.1007/s00018-014-1661-9

Piovesan, D., Necci, M., Escobedo, N., Monzon, A. M., Hatos, A., Mičetić, I., Quaglia, F., Paladin, L., Ramasamy, P., Dosztányi, Z., Vranken, W. F., Davey, N., Parisi, G., Fuxreiter, M. & Tosatto, S. E. (2020, 25 noviembre). MobiDB: intrinsically disordered proteins in 2021. *Nucleic Acids Research*, 49(D1), D361-D367. doi.org/10.1093/nar/gkaa1058

Pittas, T., Zuo, W. & Boersma, A. J. (2021). Engineering crowding sensitivity into protein linkers. *Linkers in Biomacromolecules*, 51-81. doi.org/10.1016/bs.mie.2020.09.007

Radhakrishnan, I., Pérez-Alvarado, G. C., Dyson, H. & Wright, P. E. (1998, 3 julio). Conformational preferences in the Ser133-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Letters*, 430(3), 317-322. doi.org/10.1016/s0014-5793(98)00680-2

Radhakrishnan, I., Pérez-Alvarado, G. C., Parker, D., Dyson, H., Montminy, M. R. & Wright, P. E. (1997). Solution Structure of the KIX Domain of CBP Bound to the Transactivation Domain of CREB: A Model for Activator:Coactivator Interactions. *Cell*, 91(6), 741-752. doi.org/10.1016/s0092-8674(00)80463-8

Robustelli, P., Piana, S. & Shaw, D. E. (2020). Mechanism of Coupled Folding-upon-Binding of an Intrinsically Disordered Protein. *Journal of the American Chemical Society*, 142(25), 11092-11101. doi.org/10.1021/jacs.0c03217

*Saccharomyces Genome Database*. www.yeastgenome.org/

Salladini, E., Jørgensen, M., Theisen, F. F., & Skriver, K. (2020). Intrinsic Disorder in Plant Transcription Factor Systems: Functional Implications. *International journal of molecular sciences*, 21(24), 9755. doi.org/10.3390/ijms21249755

- Soranno, A., Koenig, I., Borgia, M. B., Hofmann, H., Zosel, F., Nettels, D., & Schuler, B. (2014). Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proceedings of the National Academy of Sciences of the United States of America*, 111(13), 4874–4879. doi.org/10.1016/j.bpj.2013.11.1572
- Stadtmiller, S. S., Gorenek-Benitez, A. H., Guseman, A. J., & Pielak, G. J. (2017). Osmotic Shock Induced Protein Destabilization in Living Cells and Its Reversal by Glycine Betaine. *Journal of Molecular Biology*, 429(8), 1155–1161. doi.org/10.1016/j.jmb.2017.03.001
- Takayama, T., Yamamoto, K., Saito, H., & Tatebayashi, K. (2019). Interaction between the transmembrane domains of Sho1 and Opy2 enhances the signaling efficiency of the Hog1 MAP kinase cascade in *Saccharomyces cerevisiae*. *PLoS one*, 14(1), e0211380. doi.org/10.1371/journal.pone.0211380
- Tatebayashi, K., Yamamoto, K., Nagoya, M., Takayama, T., Nishimura, A., Sakurai, M., Momma, T., & Saito, H. (2015). Osmosensing and scaffolding functions of the oligomeric four-transmembrane domain osmosensor Sho1. *Nature communications*, 6, 6975. doi.org/10.1038/ncomms7975
- Tatebayashi, K., Yamamoto, K., Tanaka, K., Tomida, T., Maruoka, T., Kasukawa, E., & Saito, H. (2006). Adaptor functions of Cdc42, Ste50, and Sho1 in the yeast osmoregulatory HOG MAPK pathway. *The EMBO journal*, 25(13), 3033–3044. doi.org/10.1038/sj.emboj.7601192
- Theillet, F. X., Binolfi, A., Frembgen-Kesner, T., Hingorani, K., Sarkar, M., Kyne, C., Li, C., Crowley, P. B., Gierasch, L., Pielak, G. J., Elcock, A. H., Gershenson, A. & Selenko, P. (2014, 5 junio). Physicochemical Properties of Cells and Their Effects on Intrinsically Disordered Proteins (IDPs). *Chemical Reviews*, 114(13), 6661-6714. doi.org/10.1021/cr400695p
- Tompa, P., Dosztányi, Z., & Simon, I. (2006). Prevalent structural disorder in *E. coli* and *S. cerevisiae* proteomes. *Journal of Proteome Research*, 5(8), 1996–2000. doi.org/10.1021/pr0600881
- Tompa, P., & Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends in Biochemical Sciences*, 33(1), 2–8. doi.org/10.1016/j.tibs.2007.10.003
- Tompa, P. Intrinsically disordered proteins: a 10-year recap. (2012). *Trends Biochemical Sciences*, 37(12), 509-16. doi: 10.1016/j.tibs.2012.08.004.
- Uversky, V. N. (2009). Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, Counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein Journal*, 28(7–8), 305–325. doi.org/10.1007/s10930-009-9201-4

- Uversky, V. N. & Dunker, A. K. (2010). Understanding protein non-folding. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1804(6), 1231-1264. doi.org/10.1016/j.bbapap.2010.01.017
- Uversky, V. N. (2017). Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Current Opinion in Structural Biology*, 44, 18-30. doi.org/10.1016/j.sbi.2016.10.015
- Uversky, V. N. (2014). *Intrinsically Disordered Proteins*. (SpringerBriefs in Molecular Science) (2014th ed.). Springer. van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D. T., Kim, P. M., Kriwacki, R. W., Oldfield, C. J., Pappu, R. V., Tompa, P., Uversky, V. N., Wright, P. E. & Babu, M. M. (2014). Classification of Intrinsically Disordered Regions and Proteins. *Chemical Reviews*, 114(13), 6589-6631. doi.org/10.1021/cr400525m
- Walia, A., Waadt, R., & Jones, A. M. (2018). Genetically Encoded Biosensors in Plants: Pathways to Discovery. *Annual Review of Plant Biology*, 69(1), 497–524. doi.org/10.1146/annurev-arplant-042817-040104
- Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., & Jones, D. T. (2004). Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *Journal of Molecular Biology*, 337(3), 635–645. doi.org/10.1016/j.jmb.2004.02.002
- Wright, P. E. & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, 16(1), 18-29. doi.org/10.1038/nrm3920
- Yaakoub, H., Sanchez, N. S., Ongay-Larios, L., Courdavault, V., Calenda, A., Bouchara, J. P., Coria, R. & Papon, N. (2021). The high osmolarity glycerol (HOG) pathway in fungi. *Critical Reviews in Microbiology*, 1-39. doi.org/10.1080/1040841x.2021.2011834
- Yamaguchi A, Tanaka S, Yamaguchi S, Kuwahara H, Takamura C, Imajoh-Ohmi S, et al. (2012). Two Novel Heat-Soluble Protein Families Abundantly Expressed in an Anhydrobiotic Tardigrade. *PLoS ONE* 7(8): e44209. doi.org/10.1371/journal.pone.0044209
- Zhou, X., Mehta, S., & Zhang, J. (2020). Genetically Encodable Fluorescent and Bioluminescent Biosensors Light Up Signaling Networks. In *Trends in Biochemical Sciences* 45(10), 889–905. doi.org/10.1016/j.tibs.2020.06.001
- Zimmerman, S. B., & Trach, S. O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*. *Journal of Molecular Biology*, 222(3), 599–620. doi.org/10.1016/0022-2836(91)90499-V
- Wicky Basile I. M., Shammas Sarah L., & Clarke Jane. (2017). Affinity of IDPs to their targets is modulated by ion-specific changes in kinetics and residual structure. *Proceedings of the National Academy of Sciences of the United States of America*, 114(37), 9882–9887. doi.org/10.1073/pnas.1705105114

## X. Anexos metodológicos

### 10.1 Protocolo de transformación por choque térmico de levadura BJ5465

#### Reactivos

- Buffer Lazy bones:  
40% PEG 3350  
Acetato de litio 0.1 M  
Buffer Tris (pH 7.5) 10 mM  
EDTA 1 mM  
Pesar los componentes, agregar el agua e incubar a 60°C hasta que el PEG se disuelva. Cuando se prepare este buffer no se agita. El PEG debe permanecer intacto sino la levadura morirá. Se esteriliza por filtración.
- DNA de esperma de Salmón (2 mg/mL)
- Buffer Tris (pH 7.5) 10 mM

#### Procedimiento

1. Inocular 1 mL de medio YPD con la cepa BJ5465.
2. Crecer toda la noche a 30 °C con agitación constante (200 rpm).
3. Colectar 1 mL del inóculo en tubo eppendorf y centrifugar a 12,000 rpm por 1 minuto. Descartar el sobrenadante.
4. Resuspender las células en 1 mL de Buffer Tris 10 mM pH 7.5.
5. Centrifugar a 12,000 rpm por 1 min. Descartar el sobrenadante.
6. Resuspender las células en 500 uL de Buffer Lazy Bones.
7. Añadir 25 uL de DNA de esperma de Salmón (2 mg/mL). El DNA de esperma de salmón debe ser previamente hervido a 100°C por 5 minutos y posteriormente expuesto a hielo por 5 min.
8. Añadir 500 ng de DNA plasmídico de interés.
9. Vortexear por 1 minuto.
10. Incubar a temperatura ambiente durante 1 hora.
11. Someter a un choque térmico a 42°C por 12 minutos e inmediatamente incubar en hielo por 12 minutos.
12. Centrifugar a 12,000 rpm por 1 min. Descartar el sobrenadante.
13. Resuspender las células en 1 mL de Tris 10 mM pH 7.5
14. Plaquear 200 uL en medio SD-ura (o cualquier medio de interés) con perlas estériles. El crecimiento en cajas de SD-ura es de 2 a 3 días a 30°C.

## 10.2 Protocolo de ensamblaje Gibson Assembly Master Mix

Reactivos	Volumen (uL)
Gibson.....	5
Plásmido abierto.....	2
Inserto.....	1
Agua libre de nucleasas.....	2
Volumen total de la reacción.....	10

Incubar en el termociclador a 50°C por 15 min.

Transformar 2 uL del resultado del ensamblaje Gibson a una alícuota de *E. coli* One Shot Mach 1. Plaquear 200 uL y el resto en cajas de medio LB + Amp y crecer O.N. 37 °C

## 10.3 Condiciones de PCR en levadura para amplificación de construcciones IDR-CREB1, IDR-EPYC1, IDR-U1TDCP.

94 °C	10 min
94 °C	30 segundos
63 °C	30 segundos
72 °C	45 segundos
72 °C	5 minutos

Por 30 ciclos.

## 10.4 Protocolo para determinación de la sensibilidad estructural mediante FRET *in vivo*

- 1) Preparación de inóculos de levadura transformada con las construcciones IDR-CREB1, IDR-EPYC1, IDR-U1TDCP.

En un tubo añadir 2.7 mL de agua destilada estéril más 0.3 mL de SD-ura 10X. Inocular con parte de la colonia en placa que expresa a la construcción de interés. Crecer durante toda la noche a 30°C y agitación constante.

- 2) Preparación de células de levadura para someter a tratamiento de choque hiperosmótico con NaCl.
  - Tomar 1 mL de inóculo y depositar en tubo eppendorf.
  - Centrifugar a 12 000 rpm por 1 min. Descartar el sobrenadante.
  - Resuspender con 1 mL de MES 50 mM pH 6.
  - Centrifugar a 12 000 rpm por 1 min. Descartar el sobrenadante.
  - Resuspender con 1 mL de MES 50 mM pH 6.
  - Centrifugar a 12000 rpm por 1 min. Descartar el sobrenadante.

→ Resuspender con 1 mL de MES 50 mM pH 6. Usar las células para el ensayo.

3) Establecimiento de las condiciones del lector de placas CLARIOStar

- No. de puntos de escaneo de longitud de onda: 91
- Longitud de onda de excitación [nm]: 433
- Ancho de banda de excitación [nm]: 10
- Longitud de onda de emisión [nm]: 460.0 -> 550.0
- Stepwidth [nm]: 1.0
- Ancho de banda de emisión [nm]: 10
- Ganancia: 2183
- Ganancia obtenida por: Introducido manualmente
- Altura focal obtenida por: enfoque automático
- Altura focal [mm]: 3.9
- Longitud de onda utilizada para el ajuste de ganancia [nm]: 490