

# LAS PROTEÍNAS DESORDENADAS Y SU FUNCIÓN: UNA NUEVA FORMA DE VER LA ESTRUCTURA DE LAS PROTEÍNAS Y LA RESPUESTA DE LAS PLANTAS AL ESTRÉS

**César Luis Cuevas-Velázquez y Alejandra A. Covarrubias-Robles\***

Departamento de Biología Molecular de Plantas, Instituto de Biotecnología,  
Universidad Nacional Autónoma de México, Apdo. Postal 510-3, C.P. 62250  
Cuernavaca, Mor., México. E-mail: \*crobles@ibt.unam.mx

## RESUMEN

El dogma que relaciona la función de una proteína con una estructura tridimensional definida ha sido desafiado durante los últimos años por el descubrimiento y caracterización de las proteínas conocidas como proteínas no estructuradas o desordenadas. Estas proteínas poseen una elevada flexibilidad estructural la cual les permite adoptar estructuras diferentes y, por tanto, reconocer ligandos diversos conservando la especificidad en el reconocimiento de los mismos. A las proteínas de este tipo, altamente hidrofílicas y que se acumulan ante condiciones de déficit hídrico (sequía, salinidad, congelamiento) se les ha denominado hidrofílicas. En plantas, las hidrofílicas mejor caracterizadas son las proteínas LEA (del inglés Late Embryogenesis Abundant) que se acumulan abundantemente en la semilla seca y en tejidos vegetativos cuando las plantas se exponen a condiciones de limitación de agua. Evidencia reciente ha demostrado que las proteínas LEA se requieren para que las plantas toleren y se adapten a condiciones de baja disponibilidad de agua. Esta revisión describe los datos más relevantes que asocian las características fisicoquímicas de estas proteínas con su flexibilidad estructural y cómo se afecta ésta por las condiciones ambientales; así como, aquéllos relacionados con sus posibles funciones en la célula vegetal ante situaciones de limitación de agua.

**Palabras Clave:** Déficit hídrico, estrés ambiental, hidrofílicas, proteínas intrínsecamente no estructuradas (PINES), proteínas LEA.

## ABSTRACT

The dogma that relates the function of a protein with a defined three-dimensional structure has been challenged in recent years by the discovery and characterization of a set of proteins known as unstructured or disordered proteins. These proteins have a high structural flexibility which allows them to adopt different structures and therefore recognize different ligands while retaining their specificity. Proteins of this type, which are highly hydrophilic and accumulate under water deficit (drought, salinity, freezing) have been recently characterized and named hydrophilins. In plants, the best characterized hydrophilins are the LEA proteins (for Late Embryogenesis Abundant), which accumulate abundantly in the dry seed and in vegetative tissues when plants are exposed to water-limited environments. Recent evidence has shown that LEA proteins are required for plants to tolerate and adapt to conditions of low-water availability. This review describes the most relevant data regarding their structural flexibility and how this is affected by environmental conditions. Also, it addresses information related to their possible functions in plant cells that are exposed to water deficit.

**Key Words:** Water deficit, environmental stress, hydrophilins, intrinsically unstructured proteins (IUP), LEA proteins.

## INTRODUCCIÓN

Las proteínas son las biomoléculas más versátiles y diversas de la célula. Están presentes en todos los procesos biológicos, indicativo de la capacidad que tienen para desarrollar un amplio número de funciones. ¿Qué hay detrás de la versatilidad funcional de estas moléculas?, la respuesta es: su estructura tridimensional. La estructura de una proteína es la forma espacial más estable que adopta en cierto ambiente, la cual depende directamente de su secuencia de aminoácidos. A partir del descubrimiento de las primeras estructuras cristalográficas comenzó a establecerse la idea de que la función de cada proteína depende en gran medida de su estructura tridimensional, a tal grado que se postuló que “*para que una proteína sea funcional, debe poseer una estructura tridimensional bien definida*”<sup>1</sup>, idea que por muchos años ha prevalecido en muchas áreas de la ciencia. El paradigma de estructura-función fue totalmente aceptado por la comunidad científica durante muchos años. Los conceptos de estructuras secundarias como hélices alfa y láminas beta fueron ampliamente utilizados y se aceptaron como las unidades estructurales fundamentales de las proteínas. Sin embargo, a lo largo del tiempo se han encontrado proteínas a las cuales no se les ha podido asignar, con las metodologías especializadas, alguna estructura secundaria conocida, lo que generó la noción de que existen proteínas que carecen de una estructura estable; es decir, que poseen una “estructura o plegamiento azaroso” (en inglés “random coil”)<sup>2</sup>. Inicialmente, este tipo de plegamiento flexible, azaroso o sin estructura se encontraba solamente en ciertas regiones de las proteínas que sólo se consideraban como conectores entre regiones con estructuras definidas o dominios

funcionales. A medida que ha aumentado el número de estructuras cristalográficas descritas, también ha aumentado la cantidad de proteínas con regiones con plegamientos azarosos, que varían en longitud y en número en una misma proteína. Por otro lado, el descubrimiento de nuevas proteínas y la caracterización de algunas de las ya descritas ha revelado la existencia de múltiples proteínas que, en la mayor parte de su extensión, presentan una estructura flexible o un plegamiento azaroso. Esto ha sido un reto para el concepto clásico de estructura-función, ya que de acuerdo a este concepto, estas proteínas no tendrían función alguna; sin embargo, ya se había demostrado que proteínas en las que abunda este tipo de plegamiento están involucradas en diferentes vías de señalización, algunas otras funcionan como factores transcripcionales, o bien son proteínas abundantes en algunos tejidos de diferentes organismos. Este cúmulo de evidencias, junto con el hecho de que son proteínas comunes en todas las especies, ha llevado a abrir una nueva sección en el capítulo de estructura-función de las proteínas, en el que se incluyen a este nuevo grupo de proteínas a las que se les ha denominado como “proteínas intrínsecamente desordenadas” (PIDs o IDPs, del inglés Intrinsically Disordered Proteins) o “proteínas no estructuradas” (PINEs o IUPs, del inglés Intrinsically Unstructured Proteins)<sup>2</sup>.

### ¿SIN ESTRUCTURA NO HAY FUNCIÓN?

Las PIDs o PINEs son proteínas en las que abundan las regiones intrínsecamente desordenadas (RID o IDR, del inglés Intrinsically Disordered Regions) y que, por tanto, carecen total o parcialmente de una estructura tridimensional bien definida (Fig. 1). Como se

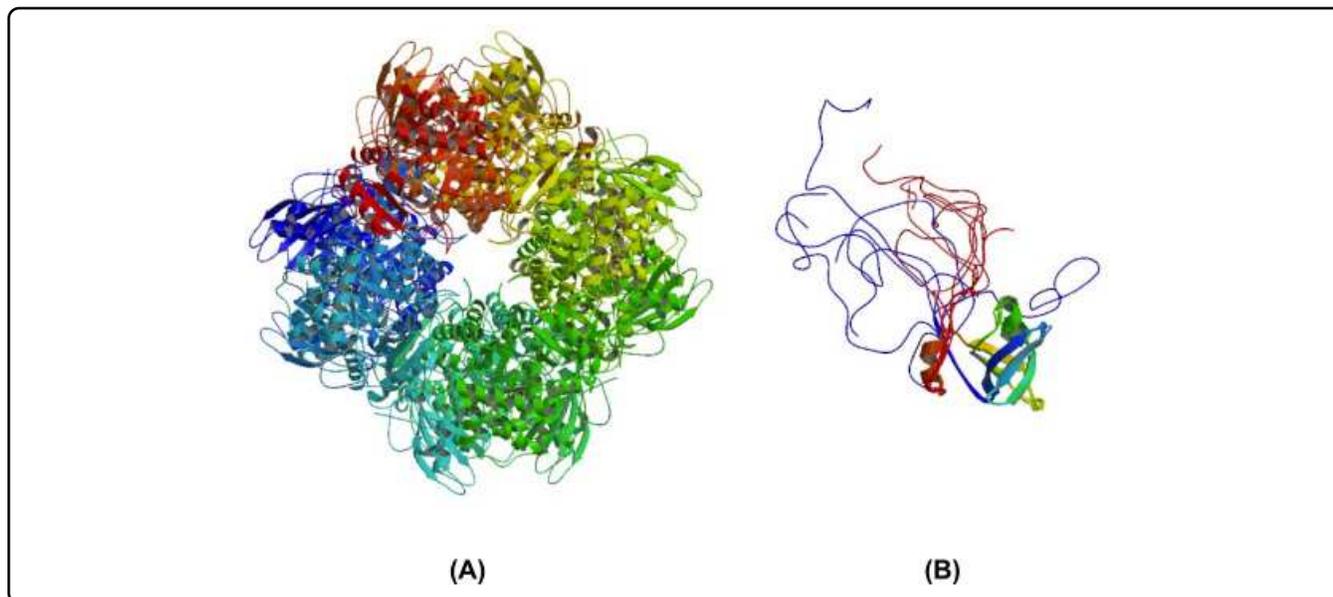


Figura 1. Estructuras de dos proteínas distintas: (A) la ribulosa 1,5-bifosfato-carboxilasa-oxigenasa (rubisco) de espinaca (PDB:1RCX) es una proteína globular; (B) el factor de inicio de la traducción (eIF1A) de humano (PDB:1D7Q), en el que la mayor parte de su estructura es desordenada. Las líneas en (B) representan las posibles conformaciones que adquiere dicha zona de la proteína.

menciona arriba, la literatura especializada contiene cada vez más evidencia que demuestra que, contrario a lo establecido en el paradigma de estructura-función, las PINEs son capaces de llevar a cabo una o más funciones. De hecho, son proteínas que desempeñan funciones muy importantes en la célula, por ejemplo, entre las PINEs y/o proteínas con RIDs podemos encontrar receptores membranales, proteínas de andamiaje, proteínas del citoesqueleto, factores transcripcionales y receptores nucleares de hormonas, entre otras<sup>3</sup>. Su importancia resalta aún más ya que mutaciones en algunas de estas proteínas, las cuales comprometen su función cambiando su ambiente estructural, están asociadas a enfermedades tan serias como el cáncer y desórdenes neurodegenerativos<sup>4</sup>. El advenimiento de los métodos masivos para obtener información genómica y la aplicación de la informática predictiva ha permitido en los últimos años obtener información que indica que las PINEs comprenden una alta proporción de las proteínas codificadas en un genoma eucarionte, alrededor del 30% se predicen como proteínas parcial o totalmente desordenadas<sup>5</sup>. Esto ha llevado a múltiples preguntas de relevancia biológica referentes a la relación estructura-función en una proteína y al significado del desorden o flexibilidad estructural en estas moléculas. Preguntas como ¿por qué se ha

favorecido el desorden estructural en las proteínas a lo largo de la evolución? ¿representa alguna ventaja evolutiva para una proteína el poseer diferentes niveles de desorden? ¿qué tipo de funciones se encuentran asociadas a estas características fisicoquímicas?

En general, las PIDs o PINEs son proteínas que presentan una composición de aminoácidos particular, son proteínas de baja complejidad que tienen una baja abundancia o carecen de aminoácidos hidrofóbicos y/o voluminosos (Val, Ile, Met, Phe, Trp, Tyr) que son la base de las proteínas globulares, pues promueven un plegamiento espontáneo y, por otro lado, poseen una alta proporción de aminoácidos cargados o polares (Gln, Ser, Pro, Glu, Lys, Arg) y en algunos casos tienen una alta abundancia de ciertos aminoácidos pequeños (Gly, Ala)<sup>6</sup>. Muchas de las regiones ricas en estos aminoácidos han sido reconocidas desde hace varios años como fundamentales para ciertas funciones en proteínas regulatorias, como lo son los factores transcripcionales, en los cuales estas regiones son la base para la interacción con los ácidos nucleicos y, por tanto, para su función como activadores o represores de la expresión genética<sup>7</sup>. Estudios recientes muestran que frecuentemente las PINEs contienen

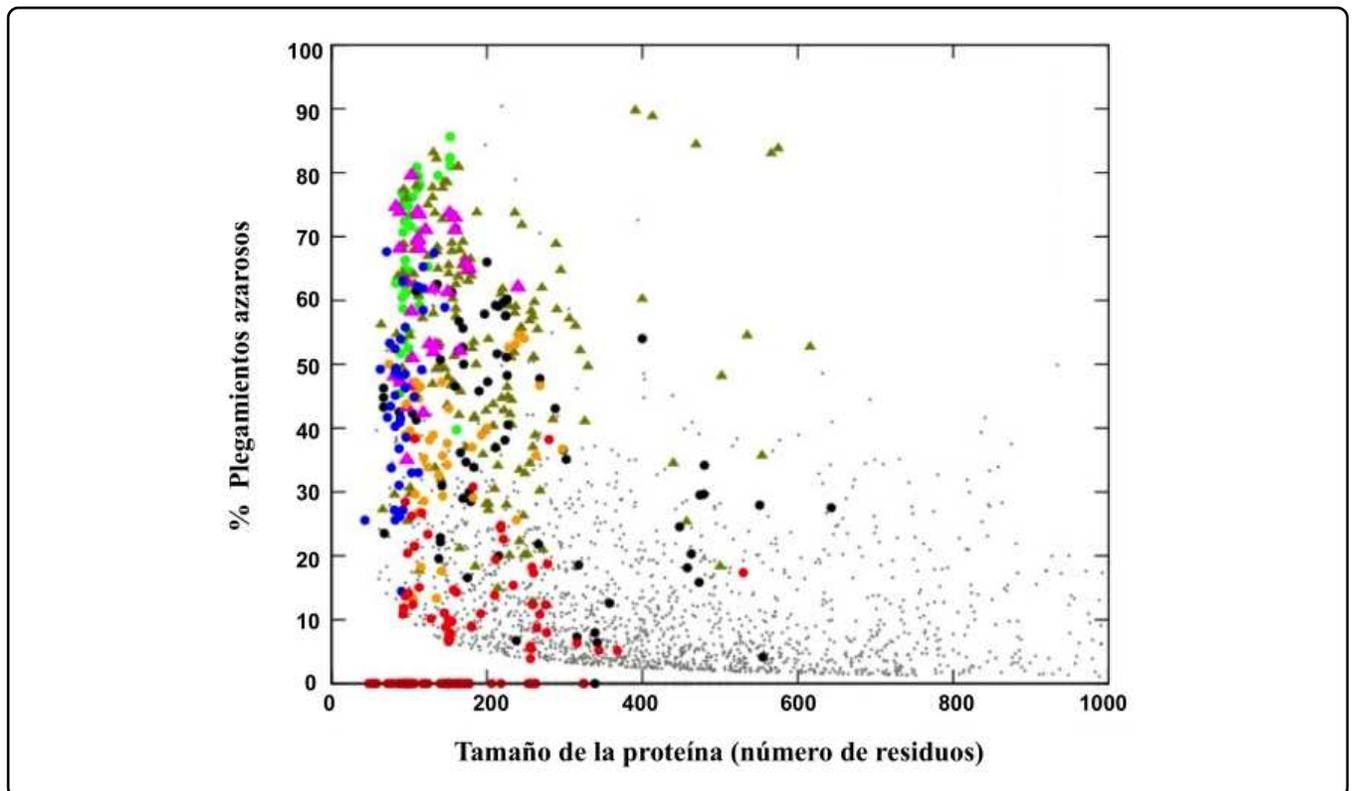


Figura 2. Representación esquemática del rango de plegamientos azarosos (*random coil*) existente en las proteínas LEA. Los porcentajes de plegamientos azarosos y el número de residuos en las proteínas se muestra en los ejes de las ordenadas y de las abscisas, respectivamente. Las proteínas LEA están representadas por círculos y triángulos llenos, en tanto que los puntos más pequeños representan proteínas diferentes a las proteínas LEA descritas para las células vegetales. La mayoría de las proteínas en las células vegetales presenta porcentajes de plegamiento azaroso iguales o menores a 0 y no se muestran en esta figura (Modificada de Battaglia *et al.*<sup>8</sup>).

regiones repetidas, las cuales se ha propuesto que pudieron haber evolucionado por expansión<sup>9</sup>. Es su composición de aminoácidos y características como éstas últimas las responsables de que la proteína sea incapaz de obtener una estructura estable en solución acuosa. Tenemos que considerar que todo tipo de estructura en las macromoléculas se ha caracterizado primordialmente en solución acuosa, de tal forma que la descripción del orden o flexibilidad estructural de una proteína corresponde generalmente a la estructura más estable que llega a adquirir en solución acuosa. Como se mencionó antes, hay que recordar que entre las PINEs las hay con diferentes grados de desorden, de tal forma que entre ellas existen proteínas que tienen múltiples dominios los cuales están conectados por regiones desordenadas o flexibles y que en cierto punto les permiten obtener una estructura compacta y más ordenada; otras que poseen amplias regiones intrínsecamente desordenadas (glóbulos fundidos o en inglés 'molten globules') y otras ordenadas, y las que son mayoritariamente desplegadas, heterogéneas o desordenadas (Fig. 2). Es probable que esta flexibilidad estructural que se reconoce en solución acuosa rara vez se dé bajo condiciones fisiológicas en la célula, ya que las condiciones del microambiente podrían promover la interacción entre ciertos aminoácidos que en solución acuosa no ocurre, de tal forma que las proteínas podrían tender a adquirir cierto grado de estructura secundaria en determinadas regiones.

Con esta información ahora regresemos a la pregunta de ¿por qué se ha favorecido el desorden estructural en las proteínas a lo largo de la evolución? Reflexiones sobre esta pregunta ha llevado a proponer varias hipótesis; quizás una de las más favorecidas sea la que considera que las PINEs son proteínas más flexibles o maleables, de tal forma que pueden adoptar diferentes estructuras funcionales que les permitirían reconocer ligandos diversos. Es posible incluso pensar que este tipo de proteínas puede presentar un mayor número de interfases intermoleculares en un momento determinado que, en el caso de las proteínas ordenadas o globulares, implicaría que éstas fuesen mucho más grandes. Esto probablemente representó una restricción evolutiva pues el aumentar la capacidad funcional de las proteínas ordenadas o globulares por un aumento en su tamaño llevaría a un mayor amontonamiento molecular en la célula, probablemente desventajoso para algunas funciones pues disminuiría el agua disponible, entre otros efectos. Por otro lado, cabe considerar que la mayor flexibilidad estructural de las PINEs no sólo abre la posibilidad de presentar más de una interfase intermolecular, sino también de exhibir más de una estructura funcional dependiendo de las condiciones microambientales y de la presencia de posibles ligandos a través de los cuales se reflejaría su función. Por ejemplo, de acuerdo a la situación metabólica de la célula es posible que se modifique el pH intracelular, el contenido de ciertos iones, el estado oxidante, el agua disponible, etc.; lo cual podría favorecer una u otra estructura y, por tanto, una u otra función. Dependiendo del estado de desarrollo de un organismo podrían estar presentes

ciertas moléculas y otras no, lo que favorecería la interacción con ciertas moléculas en un estadio determinado y con otras en uno diferente, y, por tanto, en cada caso realizaría una función distinta. Sin duda un escenario como éste sería ventajoso para muchos organismos pues no sería necesario incrementar su capacidad de almacenaje genético (genomas más grandes) y resultaría en la adquisición de respuestas adaptativas más rápidas y, quizás, más exitosas. Hipótesis como éstas se ven favorecidas por observaciones como las que resultan del análisis de la secuencia de genomas de organismos complejos. Por ejemplo, la secuencia del genoma humano ha mostrado que nuestro material genético codifica para aproximadamente 25,000 proteínas, el mismo número aproximado de proteínas codificadas por el genoma de la pequeña hierba *Arabidopsis thaliana*; esto ha llevado a una paradoja aún no resuelta relacionada con la falta de correlación entre la complejidad de un organismo y el número de genes que codifican para proteínas. Una explicación parcial a esta paradoja pudiera ser que la complejidad funcional de un organismo no necesariamente se resuelve con tener más genes que codifiquen proteínas diferentes, sino algunos genes que codifiquen proteínas multifuncionales (también llamadas 'moonlighting' en inglés) como lo serían las PINEs.

Algunas de las ideas que se describen arriba han sido comprobadas experimentalmente de tal manera que ahora sabemos que las PINEs tienen la capacidad de fluctuar entre diferentes estados conformacionales en una escala de tiempo muy corta (del orden de nano a micro-segundos)<sup>10</sup>. Esta alta flexibilidad por unidad de tiempo les da la habilidad de interactuar con un abanico de moléculas distintas con las que se unen con baja afinidad pero conservando una alta especificidad<sup>10</sup>, lo cual representa una gran ventaja desde el punto de vista biológico y evolutivo. Ante esta posibilidad, las PINEs serían capaces incluso de interactuar específicamente con la misma molécula pero de formas distintas (Fig. 3). Así pues, algunas PINEs tienen la capacidad de funcionar como interruptores (*switches*); es decir, la interacción de una PINE, en un estado conformacional dado, con una molécula blanco determinada promueve que una vía de señalización se encuentre activa; sin embargo, si el estado conformacional de la PINE cambia (modificando la dinámica de la interacción con su blanco), la vía se inactiva (Fig. 3). Éste es el caso de las p21 y p27, las cuales pueden tener tanto un efecto inhibitor, de hecho conocidas como proteínas inhibitoras de las quinasa dependientes de ciclinas (CDKs), como activador sobre las mismas CDKs, lo cual se reportó posteriormente<sup>11</sup>. Otros ejemplos similares se han descrito en la literatura con proteínas como la DHPR (receptor de la dihidropiridina, detector de voltaje)<sup>12</sup>, la CFTR (regulador de la conductancia transmembranal en la fibrosis cística)<sup>13</sup>; o bien, el caso de la calpastatina, un inhibidor de la proteasa activada por calcio, calpaina. La calpastatina favorece la activación de la calpaina facilitando su unión a calcio; sin embargo, también la inhibe a través de unirse por un dominio específico<sup>14</sup> (para mayor información ver Dyson y Wright<sup>5</sup>).

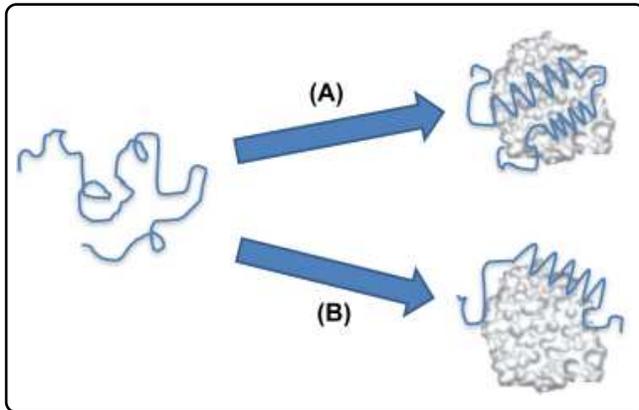


Figura 3. Las PINEs son capaces de realizar más de una función, gracias a su flexibilidad. Una PINE puede unir un ligando con una conformación dada y realizar una función específica (A); sin embargo, la misma PINE puede interactuar con el mismo ligando (o con otro distinto) a través de una conformación diferente y, por tanto, realizar una función completamente distinta (B). Las PINEs, por lo tanto, son capaces de realizar más de una función, por lo que se constituyen como un grupo de proteínas multifuncionales (que en inglés se refieren como “moonlighting”).

La multifuncionalidad que se ha observado en algunas de las PINEs es el resultado de modificaciones postraduccionales que favorecen ciertos cambios conformacionales, lo que les permite presentar una interfase diferente y con ello modular el reconocimiento a sus moléculas blanco (Fig. 4). Éste es el caso de modificaciones por fosforilación en diferentes PINEs, como sucede con CFTR (regulador de la conductancia transmembranal en la fibrosis quística), el cual ante una señal es fosforilado, lo que induce un cambio en su conformación y en el modo de unión de esta proteína a sus blancos<sup>15</sup>. Otro ejemplo bien caracterizado ocurre en CBP (CREB Binding Protein), cuyo dominio KIX interactúa con el dominio KID de CREB. Esta interacción requiere de la fosforilación del dominio KID, la cual habilita la alta afinidad por el dominio KIX. En este caso, tanto la forma fosforilada como la no fosforilada del dominio KID se encuentran no estructuradas en solución; sin embargo, la KID fosforilada se pliega al unirse al dominio KIX, de esta manera la flexibilidad intrínseca del dominio KID de CBP, ante la modificación por fosforilación, facilita la interacción con el dominio KIX de CREB<sup>6</sup>.

La flexibilidad estructural de las PINEs les permite unirse de forma específica a ligandos diversos, a la fecha se han descrito como moléculas interactoras de las PINEs a otras proteínas, ARN, ADN, membranas e incluso moléculas pequeñas como iones o azúcares<sup>16</sup>. Así pues, la función de las PINEs es también diversa, sus funciones incluyen regulación transcripcional y traduccional, transducción de señales celulares, fosforilación de proteínas, almacenamiento de moléculas pequeñas, chaperonas de proteínas y de ARN, regulación del ensamblaje de grandes complejos multiproteicos y del ribosoma, entre otras<sup>16</sup>.

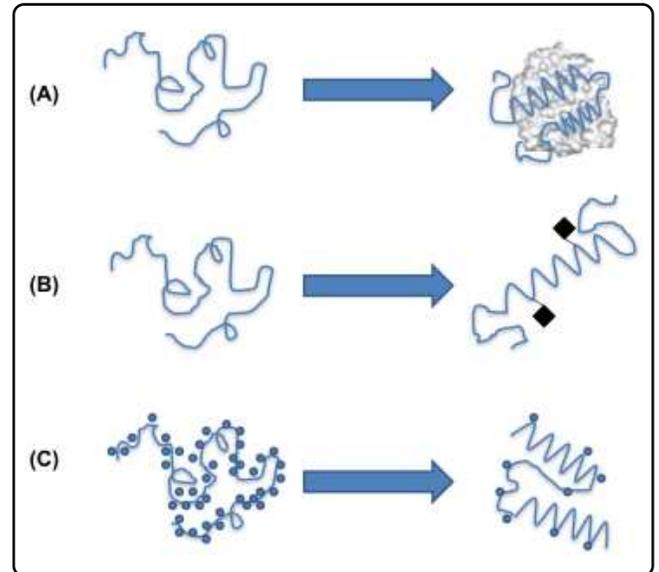


Figura 4. El plegamiento de las PINEs puede ser inducido por diferentes moléculas o condiciones ambientales en la célula. La interacción con los aminoácidos de un ligando (A) puede inducir en la PINE una conformación más estable. (B) Las modificaciones postraduccionales son otro factor que influye en el plegamiento de una PINE, el cual puede acompañarse además de la interacción con su(s) ligando(s). (C) La disponibilidad de agua también puede ser capaz de modificar la estructura de las PINEs; al cambiar la cantidad de agua disponible en el ambiente, la interacción de las PINEs con el agua se modificaría, lo que tiene como consecuencia un cambio en su conformación. Los rombos representan alguna modificación postraduccional (fosforilación, glicosilación, etc); los círculos representan moléculas de agua.

La plasticidad intrínseca de las PINEs no significa necesariamente que sean incapaces de adquirir una estructura más estable. Como se mencionaba antes, es posible que los cambios en su microambiente favorezcan una mayor estabilidad o un mayor orden estructural. Además, es posible que la interacción o unión con un ligando no sólo facilite un cambio conformacional sino que también lo estabilice (Fig. 4). Esta posibilidad ha sido demostrada para algunas PINEs y, en algunos casos, se ha estudiado con detalle. Algunos reportes recientes indican que los cambios estructurales que llevan a una conformación ordenada en una PINE se dan de manera acoplada, concomitantemente, con la unión a su ligando<sup>6</sup>. El costo de la entropía necesaria para plegar a una proteína desplegada lo paga la entalpía generada por la unión a su ligando. Evidencias como ésta han llevado a considerar que las PINEs adquieren la estructura necesaria para llevar a cabo su función al unirse a sus moléculas blanco. Sin embargo, aún queda abierta la pregunta de cómo es que las PINEs son capaces de reconocer a sus ligandos, ¿acaso se da en ausencia de un orden estructural determinado? o ¿es necesario cierto orden para presentar las interfases adecuadas en el momento apropiado? ¿es el ambiente el

responsable de proporcionar las condiciones fisicoquímicas necesarias para facilitar o inducir tales cambios conformacionales o todo se reduce al azar? Éstas son preguntas que aún quedan por dilucidar.

### LAS PINES Y LA RESPUESTA AL ESTRÉS

Hasta el momento ha sido demostrado que las PINEs son proteínas importantes en la célula y que son determinantes para el adecuado cumplimiento de las funciones celulares basales. Más recientemente, se ha encontrado que las PINEs también participan en la respuesta de los organismos a distintos estímulos o condiciones desfavorables para la célula. Un ejemplo clásico de las proteínas no estructuradas o flexibles que participan en la respuesta a condiciones adversas son las "hidrofilinas". Las hidrofilinas se definen como aquellas proteínas que tienen un índice de hidrofilicidad mayor a 1 y un contenido de glicina y de otros aminoácidos pequeños de más del 6%<sup>17</sup>. Además de estar presentes en todos los dominios taxonómicos, se ha encontrado que su acumulación está asociada notablemente con la exposición de las células o de los organismos a condiciones de limitación de agua. Llama la atención que en los diferentes *phyla* en donde se han encontrado, están acumuladas en respuesta a una baja disponibilidad de agua en el ambiente, dígase arqueas, bacterias, hongos, plantas, insectos, artrópodos y otros animales<sup>17</sup>. También se acumulan en aquellas células que transitan por etapas de desarrollo que involucran un déficit hídrico y en estructuras u órganos que pasan por estados de deshidratación severa<sup>17</sup>.

El hecho de que las hidrofilinas compartan la gran mayoría de sus características fisicoquímicas con las descritas para las PINEs y, de que se haya predicho por análisis *in silico* un alto porcentaje de desorden en todas ellas, permite incluirlas, al menos hipotéticamente, en el grupo de las PINEs o PIDs; sin embargo, a la fecha sólo de muy pocas hidrofilinas conocemos datos experimentales sobre sus características estructurales.

La hipótesis de que la presencia de las hidrofilinas confiera una ventaja a los diferentes organismos para contender con la limitación de agua parece evidente; sin embargo, aún quedan por responder preguntas como: ¿cuál es el mecanismo de acción de las hidrofilinas a nivel molecular? ¿cuál es la importancia de poseer una estructura flexible en este tipo de proteínas? y ¿por qué las características de alta hidrofilicidad y desorden estructural se han mantenido en estas proteínas a lo largo de la evolución en los diferentes organismos y se han relacionado íntimamente con situaciones de déficit hídrico? Éstas son interrogantes que han sido abordadas recientemente y que aún no se entienden del todo.

### LA RESPUESTA DE LAS PLANTAS A UN DÉFICIT HÍDRICO REQUIERE DE CIERTO NIVEL DE DESORDEN

Un caso muy interesante que surge del estudio de las hidrofilinas en las plantas es el de un grupo de proteínas muy amplio y

abundante, conocido como proteínas LEA (por sus siglas en inglés Late Embryogenesis Abundant). Estas proteínas se encontraron por primera vez, hace ya casi 30 años, por ser proteínas abundantes en la etapa tardía de la embriogénesis de la semilla, etapa en la cual ocurre una desecación natural muy alta<sup>18</sup>. Posteriormente se encontró que, además de acumularse en la semilla, se acumulan abundantemente en tejidos vegetativos sometidos a déficit hídrico (sequía, salinidad y congelamiento)<sup>19</sup>. La correlación encontrada entre la acumulación de las proteínas LEA y condiciones adversas de limitación de agua sugirió que estas proteínas juegan un papel fundamental en la respuesta y en la tolerancia de las plantas a estas condiciones. Esta hipótesis se ha visto reforzada por el hecho de que la gran mayoría de las proteínas LEA se acumulan en respuesta a tratamientos con ácido abscísico (ABA)<sup>20</sup>, conocido como la hormona del estrés, por ser una de las hormonas responsables de desencadenar una de las cascadas de señalización que dan lugar a la expresión de proteínas de respuesta que conllevan a la adaptación y/o tolerancia al estrés.

Desde el descubrimiento de las proteínas LEA, numerosos esfuerzos se han enfocado en determinar si son realmente una estrategia adaptativa utilizada por las plantas. En especial, estudios de genética funcional utilizando la planta modelo *Arabidopsis thaliana* han demostrado que al anular la expresión de 1, 2 ó 3 diferentes proteínas LEA de un mismo grupo, causa una deficiencia en la resistencia a sequía o a estrés osmótico, en comparación con las plantas silvestres<sup>21</sup>. Esto se ha visto también para proteínas LEA de otros grupos y, se ha encontrado que al complementar las mutantes con las proteínas correspondientes, el fenotipo se restablece. Experimentos como éstos han demostrado que las proteínas LEA son fundamentales para que las plantas puedan contender con situaciones de déficit hídrico.

Las proteínas LEA se han clasificado en al menos siete grupos, con base en la similitud en su secuencia de aminoácidos, de tal forma que cada grupo se distingue por ciertos motivos consenso distintivos<sup>8</sup>. En estudios previos utilizando difracción circular se ha demostrado que algunas proteínas LEA presentan estructuras con un alto nivel de desorden<sup>22</sup>. Sin embargo, algunos de estos estudios también demuestran que al someterlas a un tratamiento con compuestos que compiten por el agua cercana a las proteínas, como lo es el trifluoroetanol (TFE), estas proteínas tienden a adquirir cierto nivel de estructura, en su mayoría estructuras tipo alfa hélices u hojas beta. Estos datos indican que las condiciones de disponibilidad de agua son capaces de inducir cambios conformacionales en estas proteínas que las lleven a adquirir un mayor orden estructural. Consistente con estos datos, se encontró que el amontonamiento molecular inducido por la combinación de PEG (poli-etilen-glicol) y glicerol también es capaz de promover cambios conformacionales en estas proteínas. Estos resultados sugieren que las condiciones ambientales son capaces de modificar la organización estructural de las PINEs y,

en particular, de las proteínas LEA, lo cual adiciona un elemento más a las hipótesis propuestas dirigidas a entender el funcionamiento de este tipo de proteínas. Así, nosotros proponemos que las proteínas LEA presentan diferentes organizaciones estructurales dependiendo de la cantidad de agua disponible y/o del amontonamiento molecular en la célula, lo que a su vez podría modular el reconocimiento de diferentes ligandos, en concordancia con las hipótesis previamente propuestas para las PINes. Estas proteínas son altamente flexibles en solución acuosa, condición poco frecuente en la célula; ante una disminución en la disponibilidad de agua y, ante el consecuente incremento en la concentración intracelular de las macromoléculas (amontonamiento molecular), las proteínas LEA tenderán a adquirir una conformación más definida que les permitiría reconocer algún o algunos ligandos y de esta forma estabilizarse (Fig. 5).

Esta hipótesis se ha combinado con el modelo de trabajo al que hemos llegado, de acuerdo con los datos obtenidos a lo largo de varios años, los cuales muestran que las proteínas LEA son capaces de prevenir los cambios estructurales que se promueven por los efectos de la limitación de agua en proteínas/enzimas reporteras (como la lactato deshidrogenasa, LDH), que pudieran llegar a su desnaturalización y consecuente agregación, y que tienen como consecuencia una desactivación de éstas. Este efecto protector es evidente en ensayos *in vitro* en los que al deshidratar de manera progresiva a la enzima reportera, ésta

pierde paulatinamente su actividad; sin embargo, si antes de iniciar la deshidratación se le agrega una proteína LEA, en una relación molar 1:1, la actividad de la enzima reportera se mantiene en un mayor nivel que cuando no es protegida por la adición de la proteína LEA<sup>23</sup>. El hecho de que estas proteínas sean capaces de proteger a la enzima reportera en una relación molar 1:1 es indicativo de que uno de los posibles mecanismos de protección sea a través de una interacción directa proteína LEA-proteína blanco. Adicionalmente, datos provenientes de este tipo de ensayos nos han indicado que la capacidad protectora de las proteínas LEA de los diferentes grupos es distinta, algunas presentan una protección de casi 100% aún bajo condiciones de pérdida de agua mayores al 98%, aún mejor que la misma trehalosa, un disacárido con propiedades protectoras demostradas bajo diferentes condiciones<sup>23</sup> (Fig. 5). Sin embargo, otras proteínas LEA no muestran protección en estos ensayos. Es probable que las limitaciones de estos ensayos *in vitro*; entre otros, el uso restringido de enzimas reporteras, susceptibles a estos tratamientos, que no son los ligandos naturales, no permitan detectar su capacidad protectora. También es posible que su función esté dirigida a proteger otro tipo de moléculas, como ácidos nucleicos (ADN y/o ARN); o bien, algunas estructuras celulares como podrían ser membranas o citoesqueleto. Esto nos obligará a buscar alternativas experimentales que nos permitan explorar otros posibles blancos o funciones. Cabe mencionar que también hemos observado protección de enzimas reporteras por proteínas LEA en ensayos

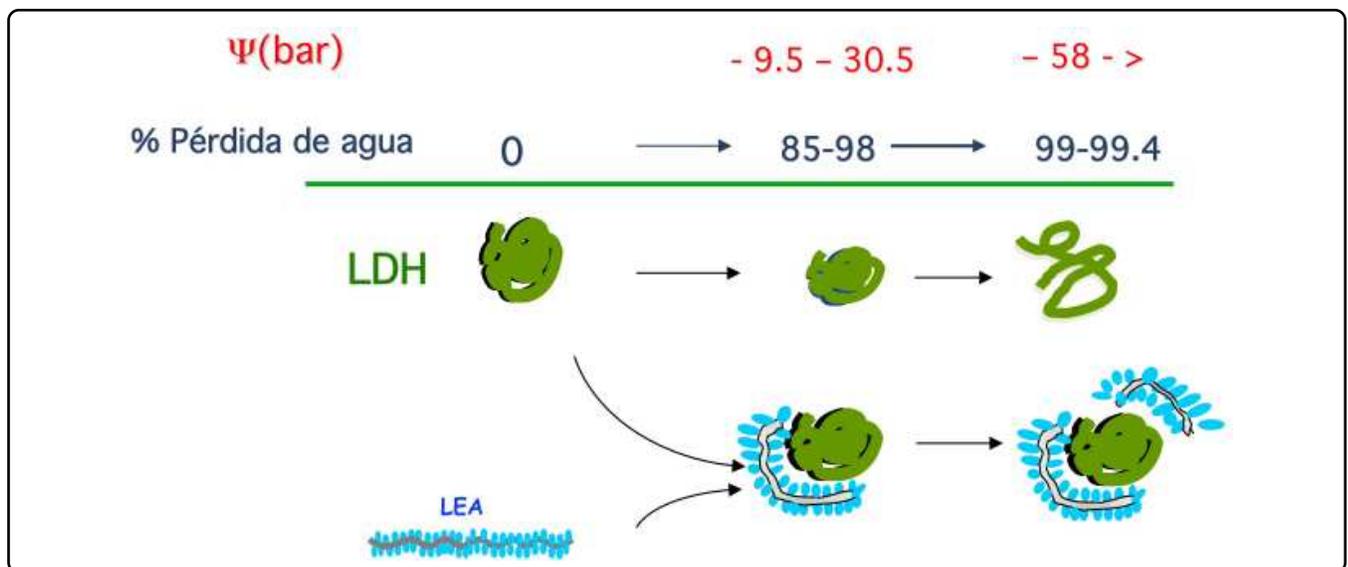


Figura 5. El modelo de acción de algunas de las proteínas LEA surge a partir de los ensayos *in vitro* descritos por Reyes et al.<sup>23,24</sup> (ver texto). Cuando las células sufren pérdida de agua, estas condiciones pueden promover cambios conformacionales en las proteínas de mantenimiento, provocando una disminución en sus actividades catalíticas o una menor eficiencia en sus funciones; condiciones más severas, pueden llegar a inducir una pérdida total de su estructura o desnaturalización y, con ello una pérdida también total de su actividad. En condiciones en las que existe una cantidad óptima de agua, las proteínas LEA se encontrarían mayoritariamente desplegadas. Cuando la cantidad de agua disponible disminuye, se podría modificar la conformación de las proteínas LEA hacia un estado "activo", en el que ahora serían capaces de interactuar con las proteínas de mantenimiento (u otros ligandos o blancos), previniendo así la pérdida de su estructura y, por lo tanto, de su función.

de congelamiento-descongelamiento<sup>24</sup>, otra condición que impone un déficit hídrico. Por otro lado, la aplicación de diferentes ensayos *in vitro* nos han mostrado que aunque estas proteínas pueden prevenir la desactivación de enzimas reporteras causada por los efectos de la deshidratación parcial, las proteínas LEA no son capaces de proveer tal protección ante temperaturas elevadas, lo cual indica que su función está más bien dirigida a contender con situaciones de baja disponibilidad de agua. Interesantemente, chaperonas moleculares como los son las proteínas de choque térmico de bajo peso molecular (sHSPs, del inglés small Heat Shock Proteins) no son capaces de prevenir la pérdida de actividad de la enzima reportera, ni su posterior desnaturalización, ante situaciones de deshidratación severa, lo cual indica que las sHSPs son incapaces de ejercer su función renaturalizadora en estas condiciones de estrés<sup>23</sup> y, por otro lado, refuerza la importancia funcional de las características fisicoquímicas de las proteínas LEA ante condiciones de déficit hídrico.

A pesar de las limitaciones que presentan estos ensayos *in vitro*, nos han permitido detectar una función para algunas de estas proteínas, y se presentan como una herramienta para explorar algunas otras preguntas en relación a las características estructurales de las proteínas LEA. Entre otras, utilizando proteínas mutantes en estos motivos, generadas *in vitro*, podríamos contestar cuál es el significado funcional de los motivos conservados en alguna de las familias o grupos de estas proteínas, cómo contribuyen estos motivos en la función protectora de estas proteínas, en estos ensayos, y cuáles son los aminoácidos relevantes para ello. La información que obtengamos será de utilidad para predecir la relevancia de posibles cambios conformacionales en estas condiciones. Aún más importante, será la posibilidad de probar estas mutantes *in vivo*, transformando plantas con estas variantes y evaluando los fenotipos que se obtengan, lo cual nos permitirá extrapolar nuestros resultados *in vitro* a su función en la planta.

La falta de similitud con otras proteínas de la célula, junto con su desorden o flexibilidad estructural, han sido obstáculos que han frenado un mayor avance en el conocimiento de la función de estas proteínas. Sin embargo, estas mismas características hacen de estas proteínas un modelo no sólo enigmático, sino también atractivo para su estudio. Las proteínas LEA se podrían considerar como un buen modelo para estudiar la función de las PINEs en plantas y, sobre todo, para entender el por qué proteínas como éstas han sido seleccionadas a lo largo de la evolución como una solución general a un problema común en los diferentes organismos como lo es la limitación de agua.

#### AGRADECIMIENTOS

Este trabajo fue apoyado parcialmente por DGAPA (IN222309). César Luis Cuevas es apoyado con una beca de doctorado por el CONACyT.

#### REFERENCIAS

1. Stryer, L. Biochemistry (Freeman and Co, New York, 1995). 1064 págs.
2. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**, 527-533 (2002).
3. Ward, J.J., *et al.* Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **337**, 635-645 (2004).
4. Iakoucheva, L.M., Brown, C.J., Lawson, J.D., Obradovic, Z. & Dunker, A.K. Intrinsic disorder in cell-signalling and cancer-associated proteins. *J. Mol. Biol.* **323**, 573-584 (2002).
5. Schad, E., Tompa, P. & Hegyi, H. The relationship between proteome size, structural disorder and organism complexity. *Genome Biol.* **12** (2011).
6. Dyson, H.J. & Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197-208 (2005).
7. Mitchell, P.J. & Tijan, R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378 (1989).
8. Battaglia, M., *et al.* The enigmatic LEA proteins and other hydrophilins. *Plant Physiol.* **148**, 6-24 (2008).
9. Tompa, P. Intrinsically unstructured proteins evolve by repeat expansion. *Bioessays* **25**, 847-855 (2003).
10. Tsvetkov, P., Reuven, N. & Shaul, Y. The nanny model for IDPs. *Nat. Chem. Biol.* **5**, 778-781 (2009).
11. Cheng, M., *et al.* The p21(Cip) and p27(Kip) CDK "inhibitors" are essential activators of cyclin D-dependent kinases in murine fibroblasts. *EMBO J.* **18**, 1571-1583 (1999).
12. Haarmann, C.S., *et al.* The random-coil 'C' fragment of the dihydropyridine receptor II-III loop can activate or inhibit native skeletal ryanodine receptors. *Biochem. J.* **372**, 305-316 (2003).
13. Ma, J. Stimulatory and inhibitory functions of the R domain on CFTR chloride channel. *News Physiol. Sci.* **15**, 154-158 (2000).
14. Tompa, P., Mucsi, Z., Orosz, G. & Friedrich, P. Calpastinin subdomains A and C are activators of calpain. *J. Biol. Chem.* **277**, 9022-9026 (2002).
15. Schwiebert, E.M., *et al.* CFTR is a conductance regulator as well as a chloride channel. *Physiol. Rev.* **79**, S145-S166 (1999).
16. Tompa, P. & Csermely, P. The role of structural disorder in the function of RNA and protein chaperones. *FASEB J.* **18**, 1169-1175 (2004).
17. Garay-Arroyo, A., Colmenero-Flores, J.M., Garciarrubio, A. & Covarrubias, A.A. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *J. Biol. Chem.* **275**, 5668-5674 (2000).
18. Dure, L. & Chlan, C. Developmental biochemistry of cotton seed embryogenesis and germination. XII. Purification and properties of principal storage proteins. *Plant Physiol.* **68**, 180-186 (1981).
19. Bies-Ethève, N., *et al.* Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* **67**, 107-124 (2008).
20. Zimmermann, P., Hirsch-Hoffmann, M., Henning, L. & Gruissem, W. GENEVESTIGATOR: *Arabidopsis* microarray database and analysis toolbox. *Plant Physiol.* **136**, 2621-2632 (2004).
21. Olvera-Carrillo, Y., Campos, F., Reyes, J.L., Garciarrubio, A. & Covarrubias, A.A. Functional analysis of the group 4 late

- embryogenesis abundant proteins reveals their relevance in the adaptive response during water deficit in *Arabidopsis*. *Plant Physiol.* **154**, 373-390 (2010).
22. Soulages, J.L., Kim, K., Walters, C. & Cushman, J.C. Temperature-induced extended helix/random coil transitions in a group I late embryogenesis-abundant protein from soybean. *Plant Physiol.* **128**, 822-832 (2002).
23. Reyes, J.L., *et al.* Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects *in vitro*. *Plant Cell Environ.* **28**, 709-718 (2005).
24. Reyes, J.L., *et al.* Functional dissection of hydrophilins during *in vitro* freeze protection. *Plant Cell Environ.* **31**, 1781-1790 (2007).