



UNIVERSIDAD DEL PAPALOAPAN

INGENIERÍA EN BIOTECNOLOGÍA
CAMPUS TUXTEPEC

CARACTERIZACIÓN FUNCIONAL Y ESTRUCTURAL DEL
DOMINIO N-TERMINAL DEL FACTOR DE TRANSCRIPCIÓN
SUB1A-1 DE ARROZ (*Oryza sativa*) EXPRESADO EN
Arabidopsis thaliana.

T E S I S

QUE PARA OBTENER EL GRADO DE:

Ingeniero en Biotecnología

PRESENTA:

WALTER JOSUÉ HERNÁNDEZ SANTOS

DIRECTOR:

DR. JULIÁN MARIO PEÑA CASTRO

CO-DIRECTOR:

DR. FRANCISCO NOÉ MENDOZA AMBROSIO



TUXTEPEC, OAXACA

DICIEMBRE, 2016



UNIVERSIDAD DEL PAPALOAPAN CAMPUS TUXTEPEC

ACTA DE REVISIÓN DE TESIS

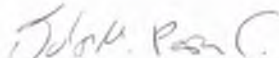
En la Ciudad de San Juan Bautista Tuxtepec, Oaxaca, el día 22 de Noviembre del 2016 a las **12:00 hr.** se reunieron en la sala de juntas del Instituto de Biotecnología, los miembros de la comisión revisora de tesis designada por la jefatura de la carrera de **Ingeniería en Biotecnología** de la Universidad del Papaloapan, con la finalidad de examinar la tesis titulada **Caracterización del dominio N-terminal del factor de transcripción SUB1A-1 de arroz (*Oryza sativa*)** presentado por el alumno **Walter Josué Hernández Santos**, con número de matrícula **11090013**, aspirante al grado de **Licenciatura**

Después de intercambiar opiniones, los miembros de la comisión manifestaron **QUE LA TESIS SATISFACE LOS REQUISITOS SEÑALADOS POR LAS DISPOSICIONES REGLAMENTARIAS VIGENTES, OTORGANDO SU APROBACIÓN PARA QUE EL ASPIRANTE PUEDA PROCEDER CON EL PROCESO DE TITULACIÓN**


Tuxtepec, Oax., a 22 de Noviembre del 2016.

ATENTAMENTE LA COMISIÓN REVISORA

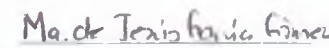
Asesor de Tesis


Dr. Julián Mario Peña Castro
Profesor- Investigador


Co-Asesor de Tesis


Dr. Francisco Noé Mendoza Ambrosio
Profesor- Investigador

Presidente


Dra. Maria de Jesus Garcia Gómez
Profesor- Investigador

Secretario


Dra. Ariana Arlene Huerta Heredia
Profesor-Investigador

Vocal


Dr. Ramon Garduño Juárez
Profesor- Investigador

A mi familia

El presente trabajo se desarrolló en los laboratorios de Biotecnología Vegetal, Biología Molecular y Química Teórica de la Universidad del Papaloapan, Campus Tuxtepec bajo la asesoría del Dr. Julián Mario Peña Castro del Instituto de Biotecnología y Co-Asesoría del Dr. Francisco Noé Mendoza Ambrosio del Instituto de Química Aplicada, con el apoyo del Consejo Nacional de Ciencia y Tecnología a través de la beca 22838 y los proyectos CONACyT - Secretaría de Educación Pública de México para Jóvenes Investigadores Ciencia Básica 152642 asignado a Julián Mario Peña-Castro y 169619 a Blanca Estela Barrera-Figueroa.

Partes de este trabajo fueron presentadas en los siguientes eventos:

Walter Josué Hernández Santos, Julián Mario Peña Castro y Francisco Noé Mendoza Ambrosio. *Creación de un método basado en dinámica molecular para elucidar la estructura de la proteína SUB1A-1 de arroz.* Cuernavaca, Morelos, 27 al 31 de julio de 2015. V Taller de dinámica molecular: Algoritmos, Análisis y Aplicaciones en Programas Paralelos. Instituto de Ciencias Físicas de la Universidad Nacional Autónoma de México.

Walter Josué Hernández-Santos, Francisco Noé Mendoza-Ambrosio y Julián Mario Peña-Castro. *Theoretical and experimental studies of the structure of the SUB1A-1 protein, responsible for flooding tolerance in rice.* Oaxaca, Oaxaca, México 3 -7 de Noviembre de 2015. The 4th International Workshop Frontiers in Protein Folding, Evolution and Function and 5o Congreso de la Rama de Fisicoquímica, Estructura y Diseño de Proteínas de la Sociedad Mexicana de Bioquímica,

AGRADECIMIENTOS

A los integrantes del comité tutorial que evaluaron esta tesis: Dra. María de Jesús García Gómez, Dra. Ariana Arlene Huerta Heredia, Dr. Andrés Aguirre Cruz y Dra. Laura Patricia Ramirez Coutiño de la Universidad del Papaloapan y Dr. Ramón Garduño Juárez del Instituto de Ciencias Físicas de la Universidad Nacional Autónoma de México

A la Dra. Jacqueline Capataz Tafúr del Laboratorio de Cultivo de Células Vegetales, al Dr. Guillermo Ramírez Galicia del Laboratorio de Química Teórica, y al Dr. Ramón Garduño Juárez por las facilidades otorgadas para el uso de equipos y espacios a su cargo.

Al personal de laboratorios QFB. Leticia Saldaña e Ing. Juan Hernández por las facilidades otorgadas para el uso de reactivos y materiales a su cargo.

A la Dra. Ana Karín Navarro Martínez y a la Lic. Yesenia Barrientos Arenal por el apoyo otorgado en el proceso administrativo de esta tesis.

Al M.C. José Luis Velasco Bolom del Instituto de Ciencias Físicas por sus atinados consejos en el campo de la biofísica.

A Osiris, por su invaluable e incondicional apoyo sumados a la valiosa compañía en estos años.

A Erick Manzano e Ivonne Azamar, por su amistad y apoyo durante todo este tiempo.

Al Dr. Francisco Noé Mendoza Ambrosio por fungir como Co-Asesor de este proyecto y apoyarme siempre.

Finalmente al Dr. Julián Mario Peña Castro, asesor de esta tesis, por ser un dedicado profesor, asesor, consejero y amigo desde el primer día.

Mi más sincera gratitud

WALTER JOSUÉ HERNÁNDEZ SANTOS

CONTENIDO

1. Introducción	2
1.1. La inundación	2
1.1.1. Escape y Quiescencia	3
1.2. Los genes <i>SUB1</i>	5
1.3. El factor de transcripción <i>SUB1A-1</i>	6
1.4. Mecanismo molecular de acción de los ERFs	8
1.5. Los dominios conservados en los ERFs	11
1.5.1. El dominio N-terminal	11
1.5.2. El dominio AP2	12
1.6. Modelo computacional	15
1.7. Avances biotecnológicos con <i>SUB1A</i>	16
1.8. <i>Arabidopsis thaliana</i>	18
2. Justificación	20
2.1. Hipótesis	22
2.2. Objetivos	22
2.2.1. Objetivo General	22
2.2.2. Objetivos Específicos	22
3. Estrategia Experimental	23
4. Metodología	24
4.1. Material Vegetal	24
4.1.1. Condiciones de Crecimiento	25
4.2. Análisis de resistencia en inundación	26
4.3. Monitoreo del retraso en la floración	26
4.4. Determinación de la expresión de <i>SUB1A-1</i> y <i>SUB1C-1</i>	27
4.4.1. Extracción de RNA	27
4.4.2. Determinación de la concentración de ARN extraído	28
4.4.3. Verificación de ausencia de DNA	28
4.4.4. Síntesis de cDNA	29
4.4.5. Verificación de ADN sintetizado	30

4.4.6. Amplificación de los transcritos de <i>SUB1A-1</i> y <i>SUB1C-1</i>	30
4.5. Medición cualitativa de la retención de carbohidratos	30
4.6. Estrategia Computacional	31
4.6.1. Análisis bioinformático	31
4.6.2. Uso del servidor I-Tasser	31
4.6.3. dinámica molecular	32
4.6.4. Acoplamiento Molecular	32
5. Resultados y Discusiones	34
5.1. Efecto del cassette de expresión sobre el fenotipo	34
5.2. Efecto del cassette de expresión sobre la tolerancia a la inundación	36
5.3. Efecto del transcrito sobre la floración	37
5.4. Efecto del transgén sobre la retención de carbohidratos	40
5.5. Expresión de <i>SUB1A-1</i> y <i>SUB1C-1</i>	42
5.6. Obtención de un modelo teórico para la proteína SUB1A-1	42
5.7. Refinamiento estructural por dinámica molecular	43
5.7.1. Minimización de energía	44
5.7.2. Equilibrio NVT	44
5.7.3. Equilibrio NPT	45
5.7.4. Recocido simulado (GSA)	46
5.7.5. Dinámica de producción en NPT	47
5.8. Validación del modelo de la proteína SUB1A-1	50
5.8.1. Análisis de ángulos diedros	51
5.9. Acoplamiento del modelo de proteína con un segmento de ADN afín	53
5.10. Localización estructural de dominios	53
5.10.1. Posición del dominio AP2	53
5.10.2. Posición del dominio N-terminal	54
6. Conclusiones y Perspectivas	58
6.1. Conclusiones	58
6.2. Perspectivas	59
A. Anexos	A1
A.1. Secuencias de SUB1A	A1
A.1.1. Nucleótidos	A1
A.1.2. Aminoácidos	A1
A.2. Configuraciones de ciclos para PCR	A2
A.3. Primers utilizados para PCR	A2
A.4. Composición de las soluciones utilizadas	A3
A.4.1. Formaldehído - Buffer de fosfatos pH 6.6	A3

A.4.2. Solución de yodo Lugol	A4
A.4.3. Gel de agarosa al 1.6% w/v	A4
A.4.4. Gel de agarosa al 2% w/v	A4
A.4.5. EtBr/TAE 1X	A4
A.5. Algoritmo de regresión lineal en Fortran 2008 para calcular el LT_{50}	A5
A.6. Configuraciones de las diferentes etapas de simulación	A6

Referencias Citadas

Ref 1

ÍNDICE DE FIGURAS

1.1. Metabolismo de carbohidratos en condiciones normales y de hipoxia	4
1.2. Composición de la región en el locus <i>SUB1</i> del cultivar FR13A de arroz tolerante . .	6
1.3. Mecanismos de acción de las proteínas ERF de <i>Arabidopsis</i> en condiciones norma- les y de hipoxia	9
1.4. Mecanismo de acción molecular de <i>SUB1A</i> durante el estrés por inundación en arroz.	10
1.5. Dominios conservados en la familia de los ERFs	11
1.6. Comparativo de las secuencias de los ERFs del locus SUB1 en la región AP2 . . .	12
1.7. El dominio de unión AP2 y su modelo de interacción con DNA	14
1.8. Supervivencia de plantas transformadas con <i>SUB1A-1</i> y <i>SUB1C-1</i>	17
1.9. Fenotipo de las plantas Col-0 transformadas con <i>SUB1A</i> y <i>SUB1C</i>	18
1.10. Etapas de crecimiento de Col-0 en fotoperiodo largo	19
4.1. Cassettes de expresión que se utilizaron para transformar a las líneas de <i>Arabidop- sis thaliana</i> que se emplearon en este trabajo.	24
4.2. Selección de las plantas transformadas en medio MS con el herbicida Basta	25
5.1. Fenotipo de las líneas de <i>Arabidopsis thaliana</i> analizadas en este trabajo.	35
5.2. Supervivencia de plantas de <i>Arabidopsis</i> a la inundación	38
5.3. Valores promedio de los tiempos de floración de las líneas utilizadas. 3 experimen- tos	39
5.4. Tinción con yodo de plántulas de 14 días de edad al final de la noche	41
5.5. Expresión de los genes <i>SUB1A-1</i> y <i>SUB1C-1</i>	42
5.6. Modelos arrojados por I-TASSER	43
5.7. Evolución de la energía potencial durante el proceso de minimización de la proteína SUB1A-1.	44
5.8. Evolución de la presión del sistema durante la simulación de la proteína SUB1A-1 en NVT.	45
5.9. Movimiento del backbone de la proteína SUB1A-1 durante 80 <i>ns</i> de simulación en NPT	46
5.10. Temperatura y RMSD del sistema durante la simulación en GSA	47
5.11. Evolución de la Energía potencial y la densidad del sistema durante la simulación GSA	48

5.12. Movimiento del esqueleto carbonado durante todas las etapas de simulación	49
5.13. Progreso de la definición de estructura secundaria en la simulación	50
5.14. Proteína SUB1A-1 antes y después del refinamiento estructural	51
5.15. Mapa de ramachandran para el modelo final de la proteína <i>SUB1A-1</i>	52
5.16. Acoplamiento del modelo de SUB1A-1 obtenido y una secuencia de ADN afín	54
5.17. Posición de los dominios en el modelo de SUB1A-1	55
5.18. Modelo final de la posición del N-terminal	56
5.19. Modelo final de la proteína SUB1A-1 unida al marcador FLAG	57

ÍNDICE DE TABLAS

4.1. Líneas transformadas seleccionadas para la extracción de RNA	27
4.2. Concentraciones calculadas de ARN en cada muestra	29
4.3. Volúmenes de los reactivos utilizados para la síntesis de la primera cadena	29
5.1. Tiempos letales 50 (LT ₅₀) de las líneas analizadas durante el estrés por inundación	37
5.2. Dominios encontrados en la secuencia de SUB1A con ayuda del paquete MEME . .	53
A.1. PCR con GoTaq Green	A2
A.2. Valores utilizados en la configuración de la minimización de energía	A6
A.3. Valores utilizados en la configuración de la simulación NVT	A6
A.4. Valores utilizados en la configuración de la simulación NPT	A7
A.5. Parámetros utilizados en la etapa GSA	A7

LISTA DE ABREVIATURAS

Abreviatura	Significado
ABA	Ácido absícico
ADH1	Alcohol deshidrogenasa-1
ADN	Acido desoxirribonucleico
AM	Acoplamiento molecular
AP2	Dominio de unión a ADN Apétala
ARN	Acido ribonucleico
ARNm	Acido ribonucleico mensajero
ARNt	Acido ribonucleico de transferencia
ATP	Trifosfato de adenosina
BR	Brasinosteroides
CSA	Acido cistein-sulfónico
dH ₂ O	Agua destilada
deH ₂ O	Agua destilada estéril
diH ₂ O	Agua desionizada
ERFs	Factores de respuesta al etileno
EtBr	Bromuro de Etidio
DRX	Difracción de rayos X
GA	Acido giberélico
HRE	Factores de respuesta a hipoxia
GSA	Recocido simulado
LOES	Mecanismo de escape
LOQS	Mecanismo de quiescencia
LT ₅₀	Tiempo letal 50
kb	Kilobases
MAS	Selección asistida por marcadores
MAPK	Proteín cinasas activadas por mitógeno
MD	dinámica molecular
ns	Nanosegundos
pb	pares de bases
PCO	Cistein oxidasas vegetales
PDC	Piruvato descarboxilasa
p/p	Relación peso-peso
QTL	Análisis cuantitativo de locus

Abreviatura	Significado
RMN	Resonancia magnética nuclear
RNS	Especies reactivas de nitrógeno
ROS	Especies reactivas de oxígeno
SNP	Polimorfismo de nucleótido simple
TCA	Ciclo de los ácidos tricarbónicos
UTR	Región no traducida
μ s	microsegundo

RESUMEN

Caracterización funcional y estructural del dominio N-terminal del factor de transcripción SUB1A-1 de arroz (*Oryza sativa*) expresado en *Arabidopsis thaliana*.

Actualmente, la humanidad enfrenta un nuevo paradigma compuesto por el aumento en la población, la demanda de alimentos que se genera y el incremento en eventos climatológicos productos del calentamiento global. Por ello, resulta de vital importancia generar tecnologías que permitan obtener cultivos resistentes a dichos eventos con la menor pérdida posible de rendimientos y características nutritivas y sensoriales.

En el campo de la biotecnología vegetal se han identificado a los ERFs del locus *SUBMERGENCE1* como los elementos genéticos responsables de la capacidad de sobrevivir a la inundación en los cultivares tolerantes de arroz. Sin embargo, esta capacidad de responder al estrés por inundación, es una capacidad endémica propia de sólo algunas variedades de este cereal.

Los primeros intentos de generación de plantas mediante la transformación con el ERF *SUB1A-1* no lograron generar plantas con mayor tolerancia a la inundación, pero sí permitieron trasladar algunas otras funciones correspondientes a este factor de transcripción como retención de carbohidratos, atraso en la floración y menor elongación celular.

La hipótesis de este trabajo es que *SUB1A* posee un mecanismo de regulación que depende de la protección ó exposición de su dominio N-terminal, por lo que al estar comprometido por la adición de elementos, como al marcador inmunológico FLAG, producirá fenotipos de diferente intensidad de resistencia a la inundación.

Para demostrar lo anterior, en este trabajo se analizaron plantas transformadas de *Arabidopsis thaliana* con combinaciones de 2 promotores (uno inducible y otro constitutivo) con el gen *SUB1A-1* y su homólogo de menor intensidad *SUB1C-1* y la presencia de un marcador FLAG protegiendo o exponiendo al dominio N-terminal. Adicionalmente, se realizó un proceso secuencial de modelado computacional para obtener un modelo tridimensional teórico de la proteína SUB1A.

De este modo, se concluyó que la intensidad de la resistencia a la inundación y el fenotipo observado en las plantas modificadas está directamente relacionado con el orden de los elementos utilizados en el cassette de expresión. Con el modelo teórico en tres dimensiones de SUB1A se analizaron los dominios presentes en este factor de transcripción y permite proponer un mecanismo de modulación estereoespacial que involucra al dominio N-terminal de la proteína SUB1A.

1. INTRODUCCIÓN

1.1. La inundación

La inundación vegetal se presenta cuando el agua asociada a fenómenos meteorológicos cubre totalmente al organismo vegetal (Tamang y Fukao, 2015). Cuando la acumulación de agua no excede el nivel de la parte aérea, es decir, no cubre totalmente al organismo, recibe el nombre de anegamiento (Shiono y Yamada, 2014). Tanto la inundación como el anegamiento generan estrés en el organismo vegetal.

Después de la sequía, la inundación es el segundo estrés abiótico más perjudicial para los cultivos y organismos vegetales. El valor de las pérdidas causadas por esta condición resulta complicado de calcular, sin embargo, en los Estados Unidos se estima que un promedio de 3 mil millones de dólares por año se pagaron por las empresas aseguradoras a los agricultores en el periodo 2011-2014 (Bailey-Serres *y cols.*, 2012b). En Asia, las pérdidas económicas se estiman en cerca de 1,500 millones de dólares (Xu *y cols.*, 2006). En México, las principales áreas afectadas corresponden a zonas costeras con actividad agrícola (Manson *y cols.*, 2009), en las que se calculó en 2014 una pérdida promedio del 35 % de la producción anual (Peña-Castro, 2014). Se estima que en los próximos años se incrementará el riesgo de eventos meteorológicos como huracanes y tormentas en todo el planeta como consecuencia del cambio climático, lo que incrementará el riesgo de inundaciones (Hirabayashi *y cols.*, 2013).

Durante la inundación, el agua disminuye la difusión de gases, tanto de la planta al aire (ej. etileno) como a la inversa (CO_2 y O_2), y también reduce la captación de luz y la capacidad fotosintética de las plantas (Voeselek y Sasidharan, 2013).

El conjunto de estos fenómenos colocan al organismo vegetal en una situación de crisis energética. Al disminuir el oxígeno, se crean condiciones de hipoxia, que se define como la disponibilidad de oxígeno por debajo de la concentración normal o normóxica de 21 % p/p (Voeselek y Bailey-Serres, 2015). La reducción en la asimilación de luz y la hipoxia, provocan el cese de producción de ATP por medio de la ruta aerobia glucólisis - ciclo de los ácidos tricarbónicos - cadena de transporte de electrones (Branco-Price *y cols.*, 2008).

Adicionalmente, la acumulación de agua provoca la proliferación de microorganismos y aumenta la probabilidad de que estos entren en contacto con el tejido vegetal (Hsu *y cols.*, 2013; Voesenek *y Bailey-Serres*, 2015). La inundación también modifica las capacidades electroquímicas del suelo, provocando con ello que las concentraciones de iones y metales como el magnesio (Mg), hierro (Fe), boro (B), sodio (Na) y aluminio (Al) aumente considerablemente hasta volverse tóxicos para las plantas (Bailey-Serres *y Voesenek*, 2008; Setter *y cols.*, 2009).

La crisis energética del organismo se enfrenta por medio del uso del metabolismo fermentativo. El proceso de cambio de metabolismo inicia con la desviación del flujo de electrones hacia el lactato como producto final y cuya acumulación reduce el pH citoplasmático. Lo anterior tiene como consecuencia que se active a la enzima piruvato descarboxilasa (PDC) (Mithran *y cols.*, 2014) y al mismo tiempo se inhiba la lactato deshidrogenasa (LDH) para que ahora el nuevo aceptor de electrones sea el etanol, una molécula menos tóxica a través de la enzima alcohol deshidrogenasa (ADH). Las enzimas LDH, ADH y PDC son suficientes energéticamente para convertirse en el principal mecanismo de recirculación de NAD⁺ y mantener a la glucólisis como principal ruta de obtención de ATP (Figura 1.1).

Sin embargo, como consecuencia del tiempo de adaptación a esta vía metabólica y a las oscilaciones diurnas, los tejidos vegetales se ven más propensos al daño oxidativo provocado por las especies reactivas de oxígeno (ROS; *Reactive Oxygen Species*) y nitrógeno (RNS; *Reactive Nitrogen Species*) que se forman, debido principalmente al funcionamiento errático de la cadena mitocondrial de transporte de electrones que no cuenta con oxígeno, su receptor final de electrones (Gupta *y cols.*, 2009; Blokhina *y cols.*, 2003). Todo lo anterior tiene como consecuencia, que los tejidos sufran necrosis, clorosis, retraso en el crecimiento, y en casos severos, la muerte completa del organismo vegetal (Peña-Castro, 2014).

1.1.1. Escape y Quiescencia

La búsqueda de mecanismos naturales de respuesta y tolerancia a la inundación condujo a la investigación, en primer instancia, de las plantas acuáticas. Se encontró que estas implementan como mecanismo de adaptación evolutiva el desarrollo de aerénquimas (estructuras especializadas tubulares huecas rellenas de aire) y raíces adventicias (Voesenek *y cols.*, 2006). Desde el punto de vista de la importancia agroindustrial, estos mecanismos son deficientes o inexistentes en los cultivares modernos de arroz (*Oryza sativa*), maíz (*Zea mays*), caña de azúcar (*Saccharum sp.*) y trigo (*Triticum sp.*). Aunado a la distancia genética existente entre estas plantas, es muy difícil el traslado interespecie de estos mecanismos de tolerancia.

poros. Sin embargo, esta estrategia de supervivencia resulta ineficiente cuando el nivel de la inundación (fenómeno que para la planta es imposible de determinar) es muy elevado, ya que los carbohidratos disponibles, pocas veces son suficientes para estimular niveles de elongación que permitan superar altos niveles de agua. De tal manera, su importancia como mecanismo de tolerancia está restringido a ciertos nichos ecológicos (Pucciariello y Perata, 2013).

El segundo mecanismo se conoce como *quiescencia* (LOQS; *Low-Oxygen Quiescence Syndrome*, del latín *quiescere*; estar quieto) y utiliza, a diferencia del escape, una estrategia intensiva de conservación de energía (Bailey-Serres y Voeselek, 2008; Pucciariello y Perata, 2013). Para ello, se reprime la actividad metabólica para consumir una cantidad discreta de energía y conservar las funciones celulares esenciales. Así, de manera simultánea se detiene el crecimiento y la floración, y se induce el uso del metabolismo fermentativo para poder sobrevivir por periodos relativamente prolongados de tiempo bajo el agua. Este mecanismo se descubrió en plantas de arroz, en las que se observó que la quiescencia tiene las ventajas de permitir que el metabolismo vegetal se restablezca una vez que el estrés ha terminado con menos efectos negativos resultantes (Xu y cols., 2006). El principal gen responsable de esta respuesta es *SUB1A-1* y se detallará en la Sección 1.3.

A pesar de que sus efectos parecen totalmente contrarios ambos mecanismos tienen en común que no involucran el uso de estructuras biológicas especializadas, sino que promueven la reducción o aumento del metabolismo vegetal mediante ajustes moleculares (Bailey-Serres y Voeselek, 2008).

1.2. Los genes *SUB1*

Se han encontrado variedades de plantas que exhiben diferentes niveles de resistencia al estrés por inundación, de entre todas ellas, el arroz (*Oryza sativa*) es probablemente la más estudiada dado que generalmente su modo de producción y su ecosistema natural incorporan constantemente niveles variables de agua (Bailey-Serres y cols., 2010). Un cultivar en específico de la variedad *indica*, FR13A (IR40931-26) resulta altamente tolerante por el mecanismo de quiescencia y puede sobrevivir hasta 1 semana más (14 días en total) que otras variedades bajo inundación completa y renovar su crecimiento una vez que el estrés desaparece (Vergara y Mazaredo, 1975; Xu y Mackill, 1996).

En FR13A se encontró por análisis cuantitativo de locus (QTL; *Quantitative Trait Locus*) y clonaciones secuenciales, que la resistencia reside en los genes del locus denominado *SUBMERGENCE 1* (*SUB1*) que tiene una longitud de 182 kb y se encuentra ubicado cerca del cen-

trómero del cromosoma 9. En el locus *SUB1* están presentes 3 genes: *SUB1A*, *SUB1B* y *SUB1C* (Figura 1.2). Los tres genes *SUB1* codifican para factores de transcripción de respuesta al etileno (ERFs; *Ethylene Response Factors*). Adicionalmente, en el locus *SUB1* se han encontrado otros 10 genes por análisis de secuencia, de 4 de ellos se han reportado transcritos, mientras que los 6 restantes se encuentran actualmente catalogados como proteínas hipotéticas (Xu y cols., 2006). No obstante, los estudios realizados por (Fukao y cols., 2009) revelan que, con base en la variabilidad genética, *SUB1C* es el gen ancestral del locus *SUB1* a partir del cual, se derivaron evolutivamente los genes *SUB1A*, *SUB1B* y *SUB1C* (Fukao y cols., 2009; Pucciariello y Perata, 2013).

Por su similitud de secuencia, sobretodo en su segmento de unión al DNA, los genes del locus *SUB1*, se han clasificado como parte de la familia VII de los ERFs (Nakano y cols., 2006), una familia multigénica de proteínas vegetales que se han asociado con respuestas reguladoras de múltiples estreses como los generados por insectos, sequía, frío y enfermedades en múltiples sistemas vegetales (Gutterson y Reuber, 2004; McGrath y cols., 2005; Nakano y cols., 2006).

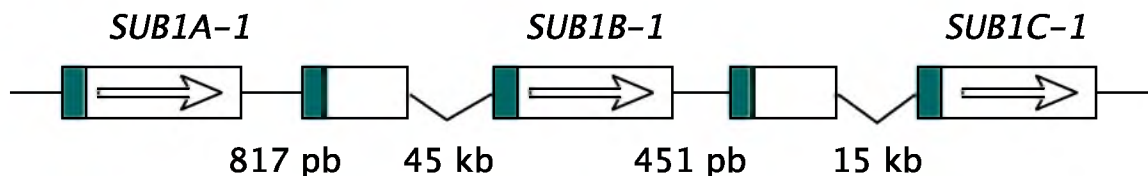


Figura 1.2: Composición de la región en el locus *SUB1* del cultivar FR13A de arroz tolerante que contiene a los ERFs. Las cajas blancas representan las regiones codificantes, en verde se muestran las regiones promotoras, las flechas indican el sentido de la transcripción. Adaptado de Xu y cols. (2006)

1.3. El factor de transcripción *SUB1A-1*

SUB1B y *SUB1C*, están presentes en la región *SUB1* de todas las variedades de arroz analizadas hasta la fecha, sin embargo, la presencia del gen *SUB1A* es variable y se ha asociado a las variedades tolerantes (Xu y cols., 2006). Existen 2 alelos para este gen (Li y cols., 2011), el primero de ellos, *SUB1A-1* es reconocido como un alelo específico de tolerancia (Fukao y cols., 2006), mientras que el segundo alelo, *SUB1A-2*, no se asocia aún a la tolerancia (Pucciariello y Perata, 2013). La principal diferencia entre estos dos alelos es un polimorfismo de nucleótido simple (SNP; *Single Nucleotide Polymorphism*) en la posición 556 que codifica a una prolina en vez de

una serina en la posición 186 de la proteína (Parnell *y cols.*, 2005). Hipotéticamente, esto impide el reconocimiento de quinasas activadas por mitógeno (*Mitogen-Activated Protein Kinase*; MAPK), lo que podría hacer que exista diferencia en la tolerancia a la inundación. Por otra parte, se han encontrado nueve alelos para *SUB1B* y siete para *SUB1C*, sin embargo, con excepción de la regulación negativa del alelo *SUB1C-1*, a la fecha ninguno de estos se ha asociado positivamente a la resistencia a la inundación.

SUB1A está constituido por 1312 nucleótidos (Apéndice A.1.1), de los cuales 149 corresponden a la región 5'-UTR, 846 al marco abierto de lectura (*Open Reading Frame*; ORF) que codifica para una proteína de 281 aminoácidos (aproximadamente 30 kDa, Apéndice A.1.2) y finalmente una región 3'-UTR de 317 nucleótidos .

La sobreexpresión de *SUB1A-1* en variedades intolerantes de arroz confiere un aumento en la tolerancia a la inundación, la disminución de la transcripción de *SUB1C* y un aumento en la expresión de los mensajeros que codifican para la alcohol deshidrogenasa 1 (ADH1), una enzima característica del metabolismo fermentativo (Agarwal *y cols.*, 2007). Lo anterior sugiere que *SUB1A-1* es un elemento relevante en la resistencia a la inundación, se ha determinado que *SUB1A-1* es responsable del 70 % de la variabilidad en respuesta a la inundación (Niroula *y cols.*, 2012).

La participación del resto del locus *SUB1* parece significativa también, estudios hechos por Xu *y cols.* (2006) revelan que durante la inundación, en la variedad tolerante aumentan los niveles de concentración de los mRNA de *SUB1A* y *SUB1C*, los cuales disminuyen cuando el estrés desaparece. Sin embargo, en variedades intolerantes, la inducción de *SUB1C* es más elevada que en la variedad tolerante. Por otro lado, los transcritos de *SUB1B* aumentan débilmente durante la inundación. En escrutinios de un número elevado de variedades de arroz se logró asociar, gracias a los cultivares carentes de *SUB1A*, que *SUB1C* es responsable de una respuesta primitiva de quiescencia (Niroula *y cols.*, 2012). La sobreexpresión de *SUB1C* en *Arabidopsis* también apoya la anterior hipótesis (Peña-Castro *y cols.*, 2011).

Además del arroz, se han buscado elementos responsables de la respuesta al estrés por inundación en otros organismos, como maíz (Campbell *y cols.*, 2015), *Brachypodium distachyon* (Rivera-Contreras *y cols.*, 2016a) y *Arabidopsis*. En este último, se encontró un total de 5 genes ERFs inducibles por hipoxia que pertenecen a la misma clasificación que *SUB1A-1*, 3 de ellos, At1g53910, At3g14230 y At3g16770, tienen alta homología (al igual que *SUB1A*) con una proteína nativa de *Arabidopsis*, APETALA2, por lo que se denominaron RAP (*RELATIVE TO APETALA*) y se identifican generalmente como RAP2.12, RAP2.2 y RAP2.3 respectivamente. Licausi *y cols.* (2010), demostraron que además de los RAPs existen otros factores de respuesta a hipoxia tipo

ERF llamados *HYPOXIA RESPONSIVE ELEMENTS* (HRE) en *Arabidopsis* que cuando se interrumpen por T-ADN (Knock out) en el genoma de *Arabidopsis*, la sensibilidad a la hipoxia es mayor que en plantas silvestres.

Recientemente, se ha encontrado que el estrés por hipoxia desencadena la síntesis de miRNAs (porciones de ARN de entre 13 y 21 pb cuya función es generalmente la de silenciamiento específico) pero se cree que estos juegan un papel menor en la respuesta al estrés por hipoxia (Licausi *y cols.*, 2011b).

1.4. Mecanismo molecular de acción de los ERFs

La investigación de los mecanismos de acción en el estrés por inundación se ha centrado en los RAPs de la planta modelo *Arabidopsis*. Es en este contexto que se sabe que la síntesis de estos ERFs y sus ortólogos se controla directamente por la concentración de etileno y de oxígeno presentes en el entorno celular. En condiciones normóxicas, los gases son capaces de difundirse libremente en la atmósfera, por lo que el etileno se encuentra en bajas concentraciones en la periferia celular lo que resulta en una baja transcripción de estos ERFs. Al haber oxígeno presente en el citoplasma vegetal, las pocas proteínas que se sintetizan se oxidan en la cisteína-2 del motivo N-terminal MCGGIL que marca a la proteína para su futura degradación proteosomal por medio del complejo enzimático N-terminal (Licausi *y cols.*, 2011a; Gibbs *y cols.*, 2011; Weits *y cols.*, 2014, Sección 1.5.1). Aquellas copias de la proteína que no se oxidan, por lo general, se secuestran por una proteína de membrana del tipo unión a acetil coenzima A (ACBP; *ACYL-COA-BINDING PROTEIN*), lo que mantiene una transcripción mínima de los genes centrales de hipoxia (GCH) por parte de los ERFs [Figura 1.3](#).

Ante un evento como la inundación, el etileno es incapaz de difundirse a la atmósfera debido a la ausencia al agua acumulada alrededor del organismo y al encontrarse en altas concentraciones, promueve fuertemente la transcripción de los ERFs. Las proteínas sintetizadas no se oxidan en la cisteína del dominio MCGG debido a la ausencia de oxígeno y por lo tanto no se degradan en el proteosoma. Cuando la concentración de las proteínas ERFs alcanza niveles adecuados, se introducen al núcleo celular gracias a la presencia en de una secuencia de localización nuclear ([E/A]****[R/K]**[K/S]; [Sección 1.5.2](#)) e inicia fuertemente la transcripción de los GCH (revisado en [Peña-Castro, 2014](#)).

Sin embargo, y a pesar de ser altamente homólogo a los RAPs y pertenecer a la misma familia de ERFs, *SUB1A* parece tener un mecanismo diferente. [Gibbs *y cols.* \(2011\)](#) demostraron, por medio de la manipulación de elementos de la ruta N-terminal, que a pesar de poseer el dominio de

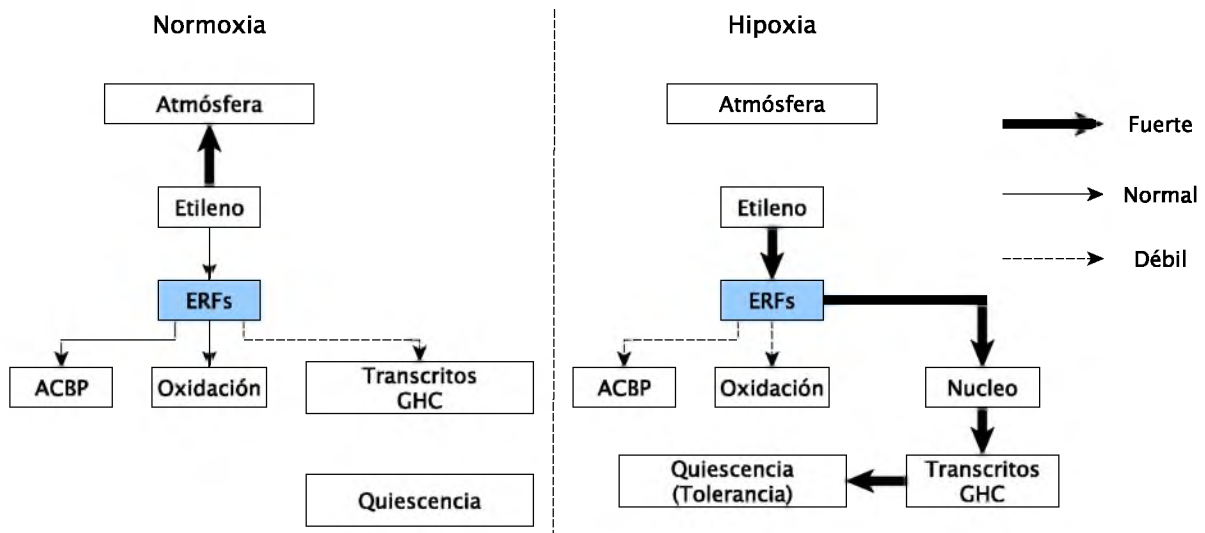


Figura 1.3: Mecanismos de acción de las proteínas ERF de *Arabidopsis* en condiciones normales y de hipoxia. Figura generada con base en Bailey-Serres y Voesenek (2008); Bailey-Serres y Colmer (2014)

reconocimiento N-terminal, *SUB1A* no se degrada por esta ruta, sin embargo Peña-Castro *y cols.* (2011) demostraron que *SUB1A-1* es susceptible de degradación proteosomal, al probar, mediante inmunoprecipitación de extractos de plántulas, que la concentración detectable de *SUB1A* aumenta cuando se trata a las plantas con el inhibidor MG132 del proteosoma y que se reduce en ausencia de dicho inhibidor. Lo anterior sugiere que *SUB1A* se degrada por el proteosoma sin ser procesado por la ruta N-terminal por medio de un mecanismo aún no determinado (Gibbs *y cols.*, 2011).

Jung *y cols.* (2010) demostraron que en el transcriptoma movilizado por *SUB1A*, se encuentran múltiples elementos que afectan el metabolismo en general, incluidos otros factores de transcripción, que incluyen a otros ERFs que serían responsables de la modulación del aumento en la sensibilidad al ácido abscísico (ABA; *abscisic acid*) y represión del ácido giberélico (GA; *gibberellic acid*).

SUB1A-1 actúa directamente sobre la homeostasis energética del organismo aumentando la sensibilidad al ABA, una hormona que se ha asociado a la respuesta al estrés por sequía (Fukao *y cols.*, 2011), por medio de 2 proteínas conocidas como *SLENDER RICE 1* (SLR1) y *SLENDER RICE 1-LIKE 1* (SLRL1), que reprimen la producción del GA, el principal antagonista del ABA (Fukao y Bailey-Serres, 2008; Fukao *y cols.*, 2011). El GA también es responsable, entre otras

cosas, de la movilización de carbohidratos y de la estimulación de la elongación. De tal forma, *SUB1A* mantiene la regulación hormonal del mecanismo quiescente. En presencia de *SUB1A*, el uso de las reservas de carbohidratos se restringen y se utilizan por vías anaerobias de producción de ATP (Figura 1.4).

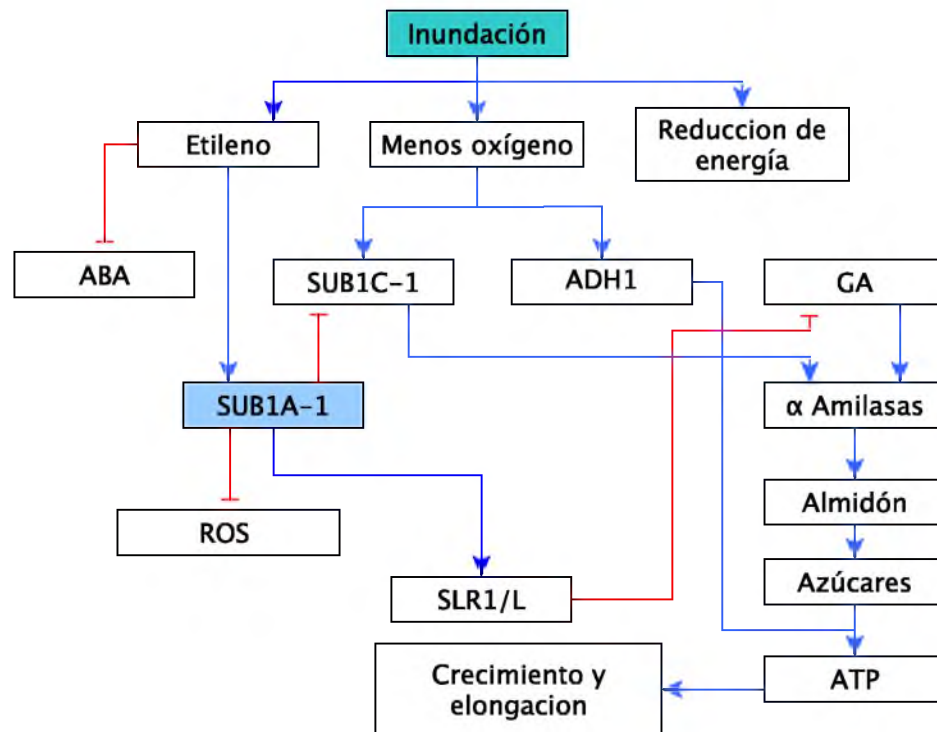


Figura 1.4: Mecanismo de acción molecular de *SUB1A* durante el estrés por inundación en arroz. Figura generada a partir de Bailey-Serres y Voeselek (2008); Fukao y Bailey-Serres (2008); Jung *et cols.* (2010); Bailey-Serres y Voeselek (2010); Bailey-Serres y Colmer (2014)

Avances recientes en el entendimiento de los fenómenos que ocurren durante y después de la inundación revelan que cuando la inundación cede, el organismo vegetal sufre estrés oxidativo al tener que readaptarse al entorno aerobio (Tamang y Fukao, 2015) y también de estrés energético-luminoso cuando el fotosistema recibe luz abruptamente. Se ha demostrado que las plantas son más propensas a sufrir desecación de hojas en las primeras horas de haber terminado el evento de inundación, y que generalmente, enfrentan una falta de nutrientes por el efecto del agua sobre la dinámica de iones en el suelo (Setter *et cols.*, 2010).

1.5. Los dominios conservados en los ERFs

Las proteínas ERFs pertenecen a una megafamilia de proteínas denominada AP2/ERF, que está formada por cuatro subfamilias; AP2, RAV, ERF y DREB. Cada subfamilia se ha asociado a diversos mecanismos de sobrevivencia en las plantas y les da la versatilidad necesaria al ser organismos sésiles que han evolucionado para controlar redes metabólicas de adaptación y poder enfrentar los cambios de su entorno y sobrevivir (Mizoi *y cols.*, 2012).

Las proteínas ERFs se identifican por poseer en su secuencia al menos un dominio APETA-LA. Además poseen una cantidad variable de dominios altamente conservados evolutivamente que se han determinado por análisis bioinformático de secuencia (Nakano *y cols.*, 2006; Zamora-Hernandez, 2016). Recientemente se demostró que tales dominios responden a una amplio conjunto de funciones, desde sensores de oxígeno hasta activador o desactivador transcripcional en respuesta a múltiples estímulos (Figura 1.5). En el caso del grupo VII de los ERFs, se han obtenido estas secuencias gracias al análisis de la planta modelo *Arabidopsis*. Nakano *y cols.* (2006) encontraron que pueden contener hasta 10 dominios, siendo los más estudiados el N-terminal (Sección 1.5.1) y el de unión a ADN (Sección 1.5.2).

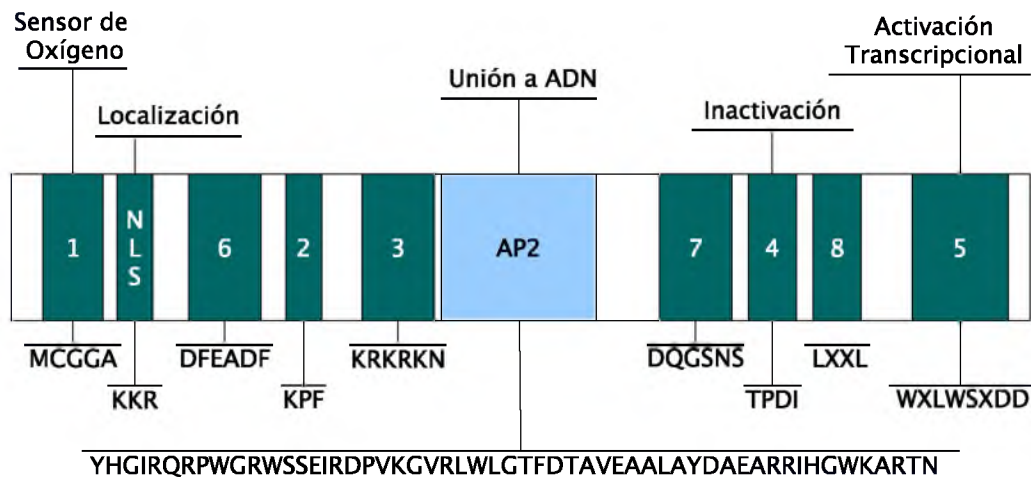


Figura 1.5: Dominios conservados en la familia de los ERFs, Adaptado de Zamora-Hernandez (2016)

1.5.1. El dominio N-terminal

Por su importancia en el mecanismo de acción de los ERFs, uno de los dominios más estudiados es el dominio N-terminal, el cual consiste en los residuos conservados MCGG(A/I)L, y que se

ha asociado a la señalización para la degradación proteosomal a través del mecanismo conocido como la ruta N-terminal (Bachmair *y cols.*, 1986; Graciet *y cols.*, 2010; Weits *y cols.*, 2014). En condiciones normales, esta ruta controla un amplio conjunto de funciones asociadas al desarrollo de brotes y hojas (Graciet *y cols.*, 2009). Sin embargo, en condiciones hipóxicas, la cisteína ubicada en la posición 2 del dominio N-terminal es oxidada por enzimas de la familia cisteína oxidasas (PCO; *PLANT CYSTEINE OXIDASES*) para formar ácido cistein-sulfónico (CSA; *cystein sulphonic acid*). Durante este paso, la proteína madura pierde la metionina N-terminal por efecto de proteínas de la familia metionil-aminopeptidasas (Bradshaw *y cols.*, 1998). Lo anterior coloca a la cisteína como el residuo N-terminal, lo que la hace reconocible por las proteínas PCO. El CSA actúa como un sitio de reconocimiento para la maquinaria de degradación proteosomal. Se sabe que la síntesis de enzimas de la familia PCO es regulada por los propios genes ERFs, por lo que cuando los ERFs se sintetizan en altas cantidades se reprime la síntesis de PCOs y se crea un bucle de autorregulación (Weits *y cols.*, 2014).

1.5.2. El dominio AP2

Los ERFs poseen en común una copia de la región conservada APETALA2 (AP2) (Xu *y cols.*, 2006), llamada así por la proteína del mismo nombre encontrada en *Arabidopsis* y que se asoció a la floración (Jofuku *y cols.*, 1994) y que es un dominio de 56 residuos de aminoácidos (Weigel, 1995). Su longitud es variable según la proteína y el organismo en el que se encuentre (Shigyo *y cols.*, 2006). A pesar de ser un grupo altamente conservado a lo largo de las proteínas ERFs, existen diferencias discretas en el dominio AP2 de cada ERF. En el locus *SUB1*, el dominio AP2 de *SUB1B* y *SUB1C* tiene los 2 aminoácidos característicos del subgrupo B2 de la familia de los ERFs (distintivos del grupo VII); alanina y ácido aspártico ubicados en las posiciones 14 y 19 respectivamente (asteriscos en la Figura 1.6), mientras que en *SUB1A* dicho dominio incorpora una serina en la posición 15.

	1	11	21	31	41	51	
	### # #	# #*	# *	# #	#		#
SUB1A-1	YEYHGIRQRP	WGRWSSEIRD	PVKGVRLWLG	TFDTAVEAAL	AYDAEARRIH	GWKARTNF	
SUB1B-1	YPYRGVRQRP	WGRWASEIRD	PVKGARVWLG	TFDTAVEAAR	AYDAEARRIH	GHKARTNF	
SUB1C-1	YPYRGIRQRP	WGRWASEIRD	PVKGIRVWLG	TFDTAEGAAR	AYDDEVRRYI	GGNAKTNF	

Figura 1.6: Comparativo de las secuencias de los ERFs del locus *SUB1* en la región AP2. Figura generada a partir de Xu *y cols.* (2006)

Dada la alta homología entre los dominios AP2 de los ERFs, se puede asumir que en todos

ellos la estructura tridimensional es semejante a la región original de la proteína APETALA 2. [Allen y cols. \(1998\)](#) analizaron por resonancia magnética nuclear (RMN) la estructura de dicha región y reportaron sus coordenadas atómicas (PDB ID: 2GCC). Se pudo identificar que el dominio AP2 posee característicamente 3 hojas β antiparalelas (rojo en la [Figura 1.7a](#)) conformadas por los segmentos Val149-Arg152, Lys156-Asp163 y Ala169-Phe176 respectivamente y una α -hélice (azul en la [Figura 1.7a](#)) de 16 residuos (Thr178-Arg194).

El análisis muestra que la estructura de AP2 se estabiliza por contactos hidrofóbicos en las cadenas laterales que se forman por los pares de residuos Tyr146-Val149, Phe157-Ala159, Ile161-Val171, Leu173-Phe176, Ala179-Ala182, Ala183-Ala185, Tyr186-Ala189, Ala190-Ala198 y Leu200-Phe202 (líneas azules en la [Figura 1.7c](#)).

La posición de la α -hélice, casi paralela a la segunda hoja β en dirección N-terminal a C-terminal, parece estar determinada por la interacción de los múltiples residuos de alanina presentes con los residuos hidrofóbicos de las hojas β ; Phe157, Ile161, Val171 y Phe176 (líneas rojas en la [Figura 1.7c](#)), lo cual provee a la estructura de estabilidad por rigidez. El bucle N-terminal His145-Gly148 parece tener mayor estabilidad que el C-terminal Gly195-Pro203 ([Allen y cols., 1998](#)).

[Allen y cols. \(1998\)](#) determinaron que las 3 hojas β son las responsables directas de la interacción con un segmento específico de ADN mediante unión no covalente tipo B-DNA. El análisis sugiere que los residuos de AP2 responsables de dicha unión de carácter no covalente hidrofóbico son Arg147, Gly 148, Arg150, Arg152, Trp154, Lys156, Arg162, Arg170, Trp172, Thr175 y Tyr186 (asteriscos en la [Figura 1.7c](#)). Con esta información se generó el modelo de unión que se muestra en la [Figura 1.7b](#) (PDB ID: 1GCC). Se determinó en general el dominio AP2, y sus homólogos reconocen las secuencias de ADN semejantes a la secuencia base AGCCGCC.

A pesar de la importancia biológica de *SUB1A*, es remarcable la ausencia de estudios estructurales de esta proteína y los mecanismos de acción de sus dominios. La mayoría de las investigaciones realizadas a la fecha en ERFs utilizan como modelo de estudio a las proteínas RAPs, especialmente a RAP2.12 pues es el principal elemento de respuesta de *Arabidopsis* a la inundación ([Papdi y cols., 2015](#); [Kosmacz y cols., 2015](#)). Entre estos estudios se encuentran los ensayos de transactivación con los que [Bui y cols. \(2015\)](#) determinaron que la capacidad de activación de la transcripción de RAP2.12 reside en un dominio conservado en el extremo C-terminal (LWS9), e identificaron que la función del motivo interno CMVII-4 (NSFGCSFSQ**TP) definido por [Nakano y cols. \(2006\)](#), es inhibir la transcripción.

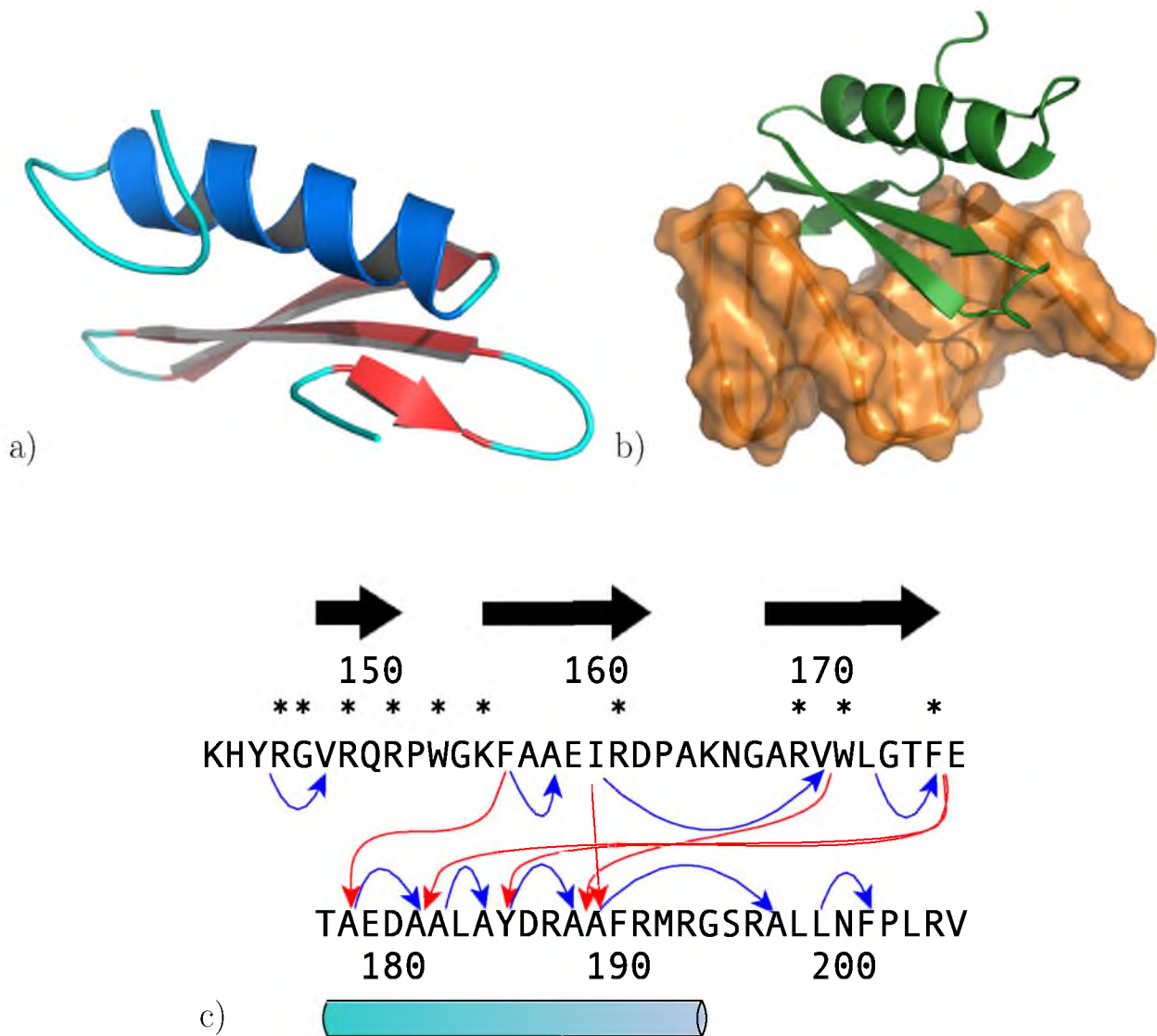


Figura 1.7: El dominio de unión AP2 y su modelo de interacción con DNA. a) Sección AP2 obtenida por RMN (PDB ID: 2GCC), en rojo se muestran las hojas β , en azul, la α -hélice. b) Modelo de interacción de AP2 con ADN (PDB ID: 1GCC), en verde el dominio AP2, en naranja, una porción de ADN de secuencia afín. c) Fuerzas no covalentes (polares) que mantienen la estructura, las flechas negras representan a las hojas β , el cilindro azul representa la α -hélice, las flechas azules señalan a los pares hidrofóbicos, las flechas rojas a los residuos estabilizadores del α -hélice, los residuos responsables de la unión a ADN se señalan con un asterisco. Figura generada a partir de [Allen y cols. \(1998\)](#)

El conocimiento actual de la función de tales dominios se logró sin el uso de un archivo de coordenadas atómicas. Se desconocen las razones por las que no existe tal información en las

bases de datos. Al no existir tal información, no es posible el análisis tridimensional de la proteína y sus dominios, lo que permitiría determinar teóricamente las posiciones y mecanismos de los mismos.

1.6. Modelo computacional

Una alternativa que ofrece una solución a la falta de un archivo de coordenadas atómicas, implica el uso de métodos teóricos basados en parámetros biofísicos. La dinámica molecular (MD; *Molecular Dynamics*) se utiliza para obtener una predicción teórica-computacional de un sistema químico mediante el procesamiento secuencial de algoritmos matemáticos como el algoritmo de Verlet (1967, 1968) y su versión actualizada, *Velocity Verlet* (Swope y cols., 1982). La MD se fundamenta en la segunda ley de Newton y produce una estimación a nivel atómico de un sistema químico. Gracias a la MD es posible simular una porción representativa del sistema problema y su evolución a través del tiempo hasta que el sistema adopte su conformación de menor energía. Gracias a la MD se han podido determinar las estructuras más probables de péptidos y proteínas pequeñas a partir de simulaciones de plegado de proteínas (Zhou y cols., 2011; Zhang y Ma, 2010, 2012; Liu y cols., 2014). Sin embargo cuando se aplica únicamente MD, la capacidad de cómputo requerida para simular el plegado de las proteínas se incrementa exponencialmente conforme aumenta la longitud de la cadena, por lo que esta vía de solución puede ser poco práctica o incluso imposible si no se cuenta con la capacidad de cómputo necesaria.

Para resolver este problema, se han desarrollado estrategias como el recocido simulado (GSA; *Generalized Simulated Annealing*) que permite reducir considerablemente el tiempo de cómputo al someter a la proteína a un incremento de temperatura seguido de un descenso de la misma para inducir la adopción de la estructura de mayor orden y menor energía posible (Agostini y cols., 2006). Adicionalmente, el incremento de la cantidad de estructuras depositadas en las bases de datos como el *Protein Data Bank* permitieron el desarrollo de modelos de predicción de estructura secundaria basados en homología (McDougal y cols., 2014). Generalmente los algoritmos de estos modelos se alojan en servidores web y suponen que segmentos de secuencias proteicas iguales poseen una estructura tridimensional semejante (Xia y Wei, 2014). La combinación de estos métodos en un proceso completo, podría permitir la obtención de un modelo de la estructura de una proteína completa en un periodo de tiempo razonable (Hao y cols., 2015).

Cuando se dispone de las estructuras atómicas de las partículas que se piensa interactúan, se pueden tratar computacionalmente con métodos como el acoplamiento molecular (AM), que utiliza la forma tridimensional de los modelos moleculares para predecir, en base a propiedades

físicas de los mismos, si dos o más moléculas tienen la posibilidad de interactuar entre sí para formar un complejo estable. El AM permite además, predecir el mecanismo y la conformación que adoptan las moléculas cuando se comienza a formar, o a deformar, un complejo estable. Se ha demostrado que el AM se puede utilizar para determinar los mecanismos de unión de enzimas con sus sustratos (Deeb *y cols.*, 2010). Se puede hipotetizar que también se puede emplear este método para analizar la manera en la que un factor de transcripción se acopla con su secuencia nucleotídica blanco (Prajapati *y cols.*, 2013).

1.7. Avances biotecnológicos con *SUB1A*

Con técnicas de biotecnología vegetal convencional (cruzas y retrocruzas) y aprovechando la heredabilidad mendeliana de caracteres en el arroz, Xu *y cols.* (2004) lograron la introgresión de la capacidad de tolerancia a la inundación desde la variedad tolerante FR13A (*indica*) a la variedad sensible de arroz M202 (*japonica*). Más tarde, Xu *y cols.* (2006) apoyados en métodos de selección asistida por marcadores lograron hacer una introgresión en la variedad *indica* Swarna. Las plantas Swarna-SUB1 mostraron altos niveles de resistencia a la inundación. Actualmente, esta nueva variedad se cultiva en las Filipinas (Singh *y cols.*, 2013) sin cambios en los rendimientos de producción debidos a la selección (Swarna: 6.3 ± 0.1 t ha, Swarna-SUB1: 6.4 ± 0.1 t ha).

Las nuevas líneas demostraron una productividad superior en condiciones de inundación de hasta 3.8 toneladas por hectárea contra 1.6 toneladas de las variedades silvestres y sin pérdida de las características sensoriales o de calidad del grano (Septiningsih *y cols.*, 2009; Bailey-Serres *y Voeselek, 2010*). Este paso fue una mejora importante debido a que el cultivar FR13A posee características indeseables de sabor (Neeraja *y cols.*, 2007).

Estas investigaciones demostraron que el locus SUB1 es un factor genético de la tolerancia que se puede transferir exitosamente entre variedades. Estas nuevas variedades se generaron sin el uso de técnicas de ADN recombinante, lo cual hizo posible la implementación inmediata de estos cultivos en siembra a campo abierto principalmente en zonas de Asia (Septiningsih *y cols.*, 2009; Dar *y cols.*, 2013; Singh *y cols.*, 2009), región en la que se encuentran dos terceras partes de los cultivos mundiales de arroz, que enfrentan constantemente niveles variables de inundación, principalmente la India, Tailandia y Bangladesh (Bailey-Serres *y cols.*, 2010).

En un primer intento por trasladar la resistencia a la inundación a plantas diferentes al arroz, Peña-Castro *y cols.* (2011) sobreexpresaron ectópicamente a *SUB1A-1* y *SUB1C-1* en plantas de *Arabidopsis* variedad Columbia (*Col-0*) usando el promotor constitutivo del virus del mosaico de la coliflor CaMV 35S (*Cauliflower Mosaic Virus*) combinado con marcación inmunológica tipo FLAG

protegiendo el extremo N-terminal. Sin embargo, contrario a lo que se esperaba, los individuos transformados resultaron altamente sensibles a la inundación como se muestra en la [Figura 1.8](#). Por otra parte, este estudio demostró que la inhibición de la floración es parte del mecanismo de quiescencia y está regulado por *SUB1A-1* y que el gen *SUB1C* modifica el fenotipo de las plantas haciendo que sus hojas sean más pequeñas y se enrollen ([Figura 1.9](#)).

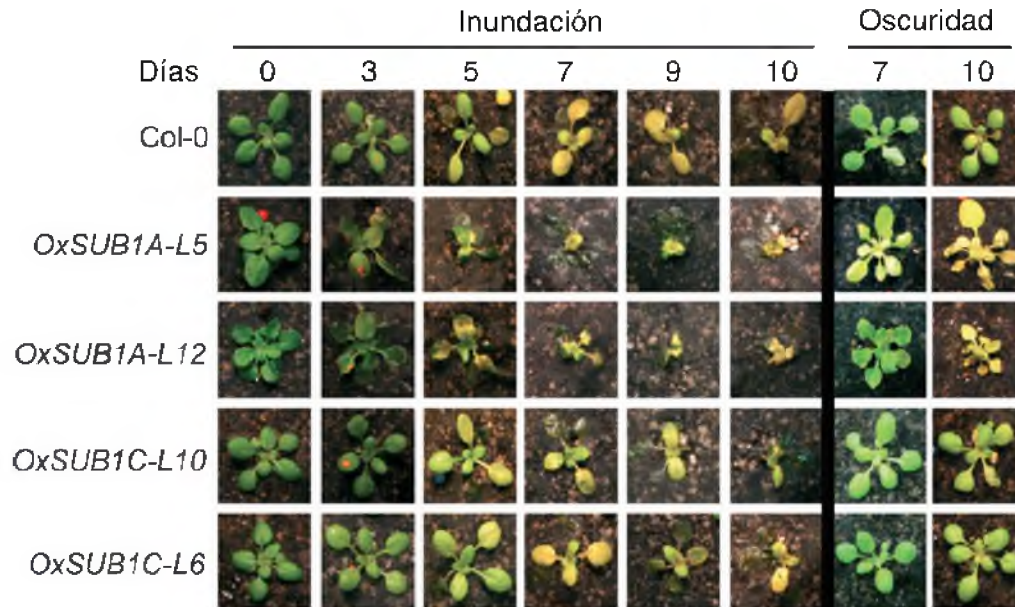


Figura 1.8: Sobrevivencia de plantas transformadas con *SUB1A-1* y *SUB1C-1* bloqueados en el extremo N-terminal con el marcador FLAG por [Peña-Castro y cols. \(2011\)](#). Se indica sobreexpresión constitutiva (Ox), la identidad del gen SUB1 y la línea probada (L), los números indican progenie a partir de una misma planta transformada.

[Núñez-López y cols. \(2015\)](#) observaron la capacidad de retención de carbohidratos en plantas de *Arabidopsis* que sobreexpresaron a *SUB1A-1*, especialmente almidón, y propusieron utilizar a *SUB1A-1* como una estrategia para mejorar la obtención de un incremento en la cantidad de azúcares fermentables por medio de sacarificación, con el objeto de lograr su posterior aplicación en la producción de bioetanol como alternativa energética. No obstante, este estudio demostró que la presencia de *SUB1A-1* retrasa el crecimiento, especialmente en etapas posteriores a la floración.

A pesar de que se ha logrado dotar a variedades de arroz intolerantes con la capacidad de tolerar la inundación utilizando al gen *SUB1A-1* ([Xu y cols., 2004, 2006](#); [Bailey-Serres y cols., 2010](#)), la transformación de una especie diferente supone un importante reto ya que los organismos vegetales, aun cuando pertenezcan a las familias cercanas evolutivamente, modulan su crecimiento



Figura 1.9: Fenotipo de las plantas Col-0 transformadas con *SUB1A* y *SUB1C* por Peña-Castro *y cols.* (2011)

y sus respuestas ante estreses de diferentes maneras y utilizan rutas metabólicas y metabolitos diferentes (Jackson y Ismail, 2015). Además hay otras posibilidades, lo cual complica la inserción de una región codificante en el genoma cuyo producto además, puede ser reconocido como una invasión y ser atacada por la maquinaria de defensa inmunológica del organismo huésped, modificada postraduccionalmente y que afecte el nivel de expresión (Butaye *y cols.*, 2005).

Prajapati *y cols.* (2013) obtuvieron una imagen tridimensional del factor de transcripción SUB1A-1 utilizando homología empírica computacional, sin embargo el resultado fue una estructura poco resuelta (< 40% del total de la secuencia). Adicionalmente, el modelo final supone un bloqueo espacial en la región AP2, por lo que se puede predecir que este modelo se puede refinar más con la adición de nuevos pasos de análisis que incluyan nuevas herramientas como la MD.

1.8. *Arabidopsis thaliana*

En el presente trabajo, se utilizó a la planta modelo *Arabidopsis thaliana* como el modelo de estudio para probar la hipótesis. Lo anterior debido a que se ha utilizado extensamente en estudios de inundación y pruebas de concepto de la transferencia de factores génicos de tolerancia (Lee *y cols.*, 2011; Vashisht *y cols.*, 2011; Peña-Castro *y cols.*, 2011; Weits *y cols.*, 2014). Adicionalmente, *Arabidopsis thaliana* es una planta dicotiledónea del género de la familia *Brassicaceae* que carece de importancia agrícola pero que se usa ampliamente como organismo modelo debido a la simplicidad de su genoma (aproximadamente 145 Mpb, 5 cromosomas, totalmente secuenciada).

do), su ciclo de vida corto de 12 semanas, su abundante producción de semillas y la facilidad de cultivo (Weigel y Glazebrook, 2002). Un ecotipo en particular de esta planta, Columbia-0 (Col-0) es ampliamente utilizado en estudios de biología molecular y sus etapas de crecimiento están muy bien caracterizadas (Boyes *y cols.*, 2001, Figura 1.10).

Al ser una planta originaria de las Montañas Harz en Alemania y otras zonas al noreste de Europa (The *Arabidopsis* Information Resource, 2016), es sensible a los cambios en la calidad y duración de la luz por lo que también se utiliza como modelo de estudio en problemas que involucran variaciones en el fotoperiodo.

Algunas otras características que hacen de *Arabidopsis* una planta modelo por excelencia son su alta capacidad productiva, pues una sola planta produce hasta 5000 semillas.

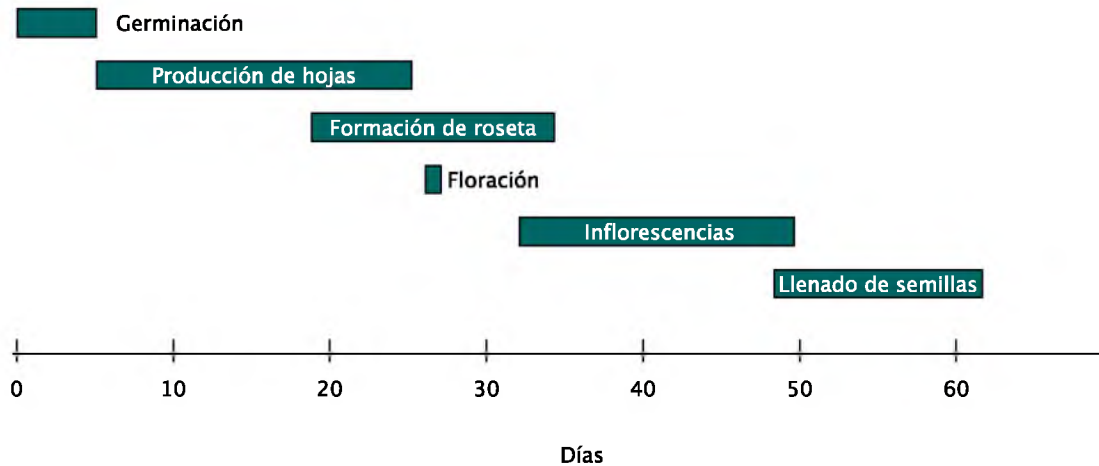


Figura 1.10: Etapas de crecimiento de Col-0 en fotoperiodo largo, Adaptado de Boyes *y cols.* (2001)

2. JUSTIFICACIÓN

Por la posición geográfica de México, una gran parte de su territorio se encuentra cercano a regiones costeras entre los océanos Pacífico y Atlántico, por lo que es un país propenso a sufrir eventos climáticos como lluvias y huracanes, que se traducen en elevados riesgos de inundaciones principalmente en estas zonas costeras. Se espera que, como consecuencia del cambio climático la secuencia de estos eventos aumente (Bailey-Serres *y cols.*, 2012a).

Por estas razones, resulta primordial la implementación de tecnologías que permitan obtener cultivos tolerantes a la inundación, con la menor pérdida de propiedades sensoriales, nutritivas y de rendimientos productivos, y sin riesgos para el consumidor, garantizando así la alimentación humana.

Desde que los ERFs se identificaron por bioinformática en *Arabidopsis* y arroz, se determinó que su función podría ser muy importante para definir la respuesta vegetal al estrés por inundación y que estaría definida por los múltiples dominios que poseen. Recientemente, esto se ha abordado de forma experimental confirmando que los dominios poseen funciones relevantes para la percepción de oxígeno, unión a DNA, inhibición o promoción de la transcripción (Nakano *y cols.*, 2006; Bui *y cols.*, 2015; Papdi *y cols.*, 2015). Sin embargo, la mayor parte de este conocimiento se ha obtenido usando a la planta modelo *Arabidopsis* y el gen *SUB1A-1* de arroz, a pesar de su importancia biotecnológica moderna, no ha sido abordado de la misma forma. El análisis de su secuencia proteica indica que al igual que los demás ERFs, está dotado varios de estos dominios conservados evolutivamente (Jung *y cols.*, 2010).

Por lo tanto, para poderse usar eficientemente como factor genético para transferir características de importancia biotecnológica como la tolerancia a la inundación y la conservación de almidón, debe pasar por un análisis estructural de dominios.

La evidencia experimental al momento usando la transferencia directa de *SUB1A*, ha logrado sólo conocimientos parciales, pero que sugieren que algunos de sus dominios en efecto están conservando su función molecular. Con esta información, se puede suponer, por ejemplo, que el cassette de expresión utilizado por Peña-Castro *y cols.* (2011) comprometió espacialmente a alguno de estos dominios, principalmente el dominio N-terminal causando una baja supervivencia

en inundación; Núñez-López *y cols.* (2015) utilizaron esta característica para lograr la efectiva retención de carbohidratos.

Adicionalmente, el análisis de expresión sugiere que el uso de un promotor constitutivo como el CaMV 35S en el cassette de expresión podría generar una severa crisis en el organismo vegetal al promover la quiescencia desde edades tempranas del desarrollo, lo que provoca que cuando el estrés se presente, las plantas sean incapaces de aumentar su tolerancia a niveles mayores de estrés y perezcan. El uso de promotores inducibles de activación conjunta con el estrés, como los presentes en el metabolismo fermentativo, podría promover el mecanismo LOQS sólo en las condiciones necesarias.

No se dispone en la actualidad de un archivo de coordenadas atómicas de SUB1A-1 o cualquiera de sus análogos, con lo que se podría determinar cómo afecta la transformación antes mencionada, o cualquier otra, a alguna región en la proteína SUB1A-1 madura. Este problema puede solucionarse por medio de simulación por MD, que ha demostrado ser efectiva como vía para la obtención de una imagen de un estado fisicoquímicamente posible para un sistema complejo como las proteínas. Así, se podría predecir la posición tridimensional de los dominios estructurales del factor SUB1A-1 y, al tratarse biocomputacionalmente (por medio de acoplamientos moleculares), esta información podría brindar mayor entendimiento acerca del mecanismo de acción de este factor de transcripción.

El entendimiento en los fenómenos involucrados en la exposición o protección del dominio N-terminal del factor de transcripción SUB1A-1 de arroz podría conducir al desarrollo de un nuevo tren de investigación biotecnológica que permita eventualmente obtener cultivos tolerantes a la inundación.

2.1. Hipótesis

La exposición al entorno del dominio N-terminal de la proteína SUB1A-1 de arroz producirá fenotipos de diferente intensidad de resistencia a la inundación en plantas transgénicas de *Arabidopsis thaliana*.

2.2. Objetivos

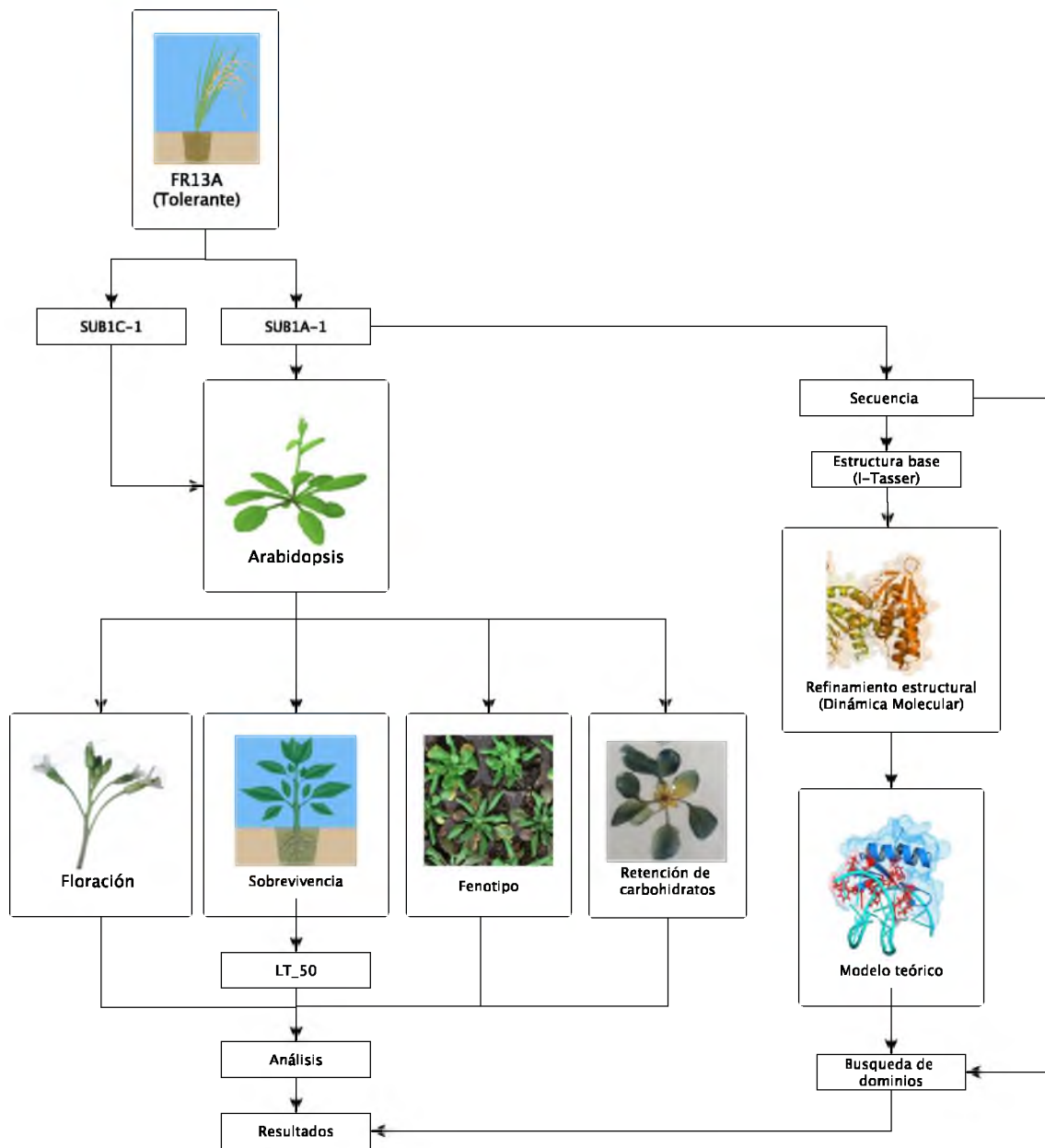
2.2.1. Objetivo General

Conocer los fenotipos generados por la exposición y el bloqueo físico del dominio N-terminal de la proteína SUB1A-1 y su homólogo SUB1C, así como sus diferentes niveles de resistencia a la inundación en plantas transgénicas de *Arabidopsis thaliana*.

2.2.2. Objetivos Específicos

1. Predecir por dinámica molecular la localización de los dominios AP2, N-terminal, C-terminal y MCGG asociados al mecanismo de acción del factor de transcripción SUB1A-1.
2. Evaluar el nivel de tolerancia a la inundación de plantas transgénicas transformadas con *SUB1A-1* con el extremo N-terminal bloqueado y libre.
3. Cuantificar el nivel de expresión de *SUB1A-1* en las plantas transformadas.
4. Evaluar la conservación de características fenotípicas asociadas a la acción de SUB1A-1 (floración, conservación de almidón y arquitectura).

3. ESTRATEGIA EXPERIMENTAL



4. METODOLOGÍA

4.1. Material Vegetal

Se utilizó a la planta *Arabidopsis thaliana* var. Col-0 tipo silvestre y sus líneas transformadas con los genes de los factores de transcripción SUB1A-1 y SUB1C-1 de arroz en cassettes de expresión que lo combinan con el promotor de la ADH de *Arabidopsis* (pADH, inducible por estrés por inundación) y pCaMV 35S (p35S, promotor constitutivo del virus del mosaico de la coliflor). Para el promotor p35S se colocó un marcador de selección inmunológica tipo FLAG bloqueando el extremo N-terminal o el C-terminal (Figura 4.1).

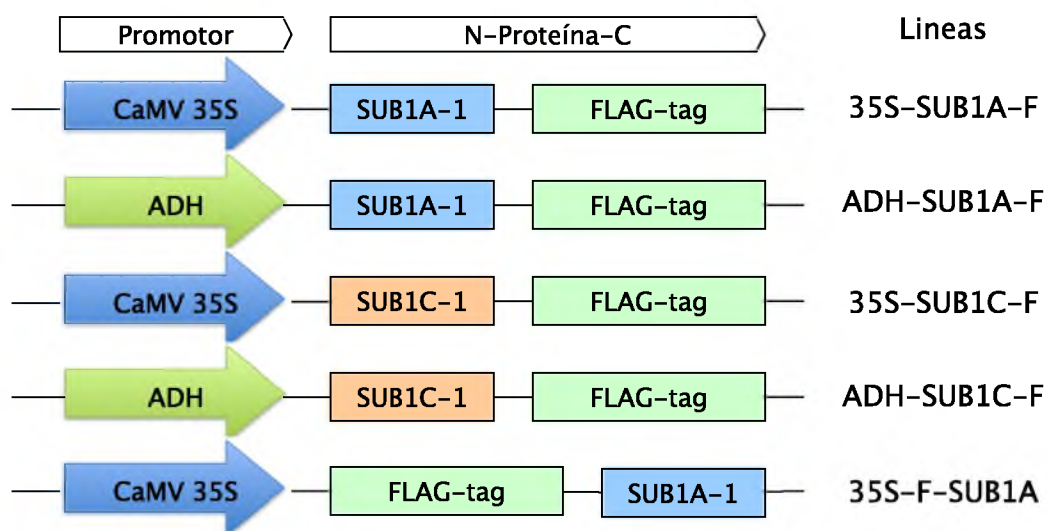


Figura 4.1: Cassettes de expresión que se utilizaron para transformar a las líneas de *Arabidopsis thaliana* que se emplearon en este trabajo.

Las construcciones quiméricas de ADN y la transformación genética se realizaron previamente en la University of California at Riverside por el Dr. Julián Mario Peña-Castro y la selección en placas de BASTA (Figura 4.2) se realizó por la alumna de servicio social Adriana Prado Fuentes

en el Laboratorio de Biotecnología Vegetal de la Universidad del Papaloapan.

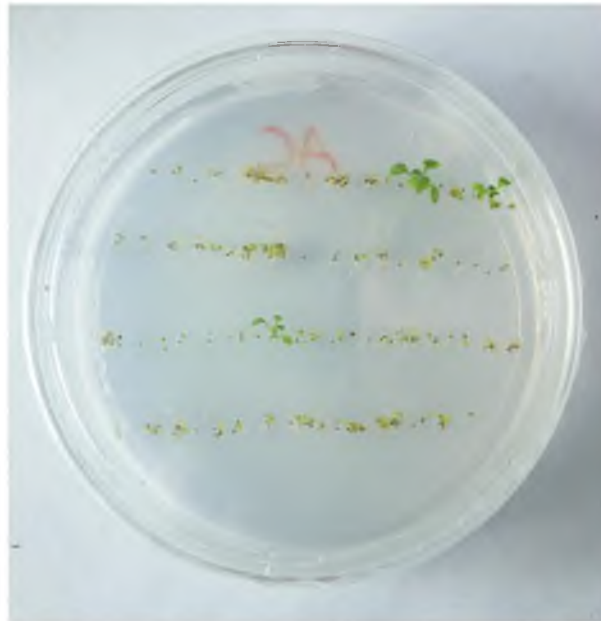


Figura 4.2: Selección de las plantas transformadas en medio MS con el herbicida Basta (50 mg/ml). Se observan las plántulas resistentes de color verde y las sensibles (no transformadas) de color amarillo. Se muestra una placa de selección para la construcción ADH-SUB1C-F.

Para apreciar las diferencias en los organismos provocadas por las variaciones en los cassettes de expresión, se analizaron las líneas provenientes de 2 eventos de transformación para cada combinación descrita en la [Figura 4.1](#). Se colocó un evento con fenotipo fuerte y otro moderado. Las semillas de dichas líneas están actualmente disponibles en el Laboratorio de Biotecnología Vegetal de la Universidad del Papaloapan bajo la custodia del Dr. Julián Mario Peña Castro bajo utilización confinada de organismos genéticamente modificados (Registro SAGARPA 001-2016).

4.1.1. Condiciones de Crecimiento

Las semillas de las plantas transformadas y silvestres se desinfectaron con un lavado en etanol al 70 % por 5 min. seguido de otro con hipoclorito de sodio en mezcla 1:1 con agua destilada estéril (deH_2O) por 3 min. Posteriormente, se incubaron por 4 días a 4 °C, para detener el proceso de dormancia y sincronizar la germinación. Las semillas se sembraron en placas Petri con 40 ml de medio MS 0.5X (sacarosa 1 %, sales Murashige-Skoog 0.5X, pH 5.7, agar 1 %) y se colocaron verticalmente bajo fotoperiodo corto (8h luz/16h oscuridad) por 7 días. Después, las semillas se trasladaron a sustrato sólido (mezcla Sunshine #3 - perlita en relación 3:1) y se añadió fertilizante NPK 12:12:17 al 3 % p/p. Las plántulas permanecieron en condiciones controladas en cámara de

crecimiento automatizada (Convicon Adaptis CMP6010) a 23 °C con una humedad relativa del 60% bajo régimen lumínico de 150 $\mu\text{moles m}^{-2}\text{s}^{-1}$ en fotoperiodo corto (8h luz / 16h oscuridad).

Las plantas se regaron por goteo a diario durante los primeros 7 días después de haber sido trasladadas a sustrato. Posteriormente, se regaron por inmersión en agua potable cada 2 días. Para reducir efectos diferenciales de las condiciones de crecimiento, las plantas se cambiaron aleatoriamente de posición dentro de la cámara después de cada riego.

4.2. Análisis de resistencia en inundación

Se desarrolló un análisis del desempeño de estas líneas durante condiciones controladas de inundación, para ello se siguió el protocolo de [Rivera-Contreras y cols. \(2016b\)](#) con las siguientes modificaciones: 15 individuos de 28 días de edad se sometieron a inundación completa dos horas antes del final del día bajo influencia del fotoperiodo durante 2, 4, 6 y 8 días en cuarto de crecimiento vegetal a 23 °C. Después de este periodo los individuos se sometieron a 7 días de tratamiento de recuperación en cámara automatizada (Convicon Adaptis CMP6010) en las condiciones descritas anteriormente. Al final de esta etapa se registró el número de hojas y porcentaje de individuos sobrevivientes para calcular el Tiempo Letal 50 (LT_{50}), usando un algoritmo de regresión lineal construido en base al método simple de regresión lineal ([Chapra y Canale \(2006\)](#), [Apéndice A.5](#)). Este experimento se realizó por triplicado para fines del análisis estadístico y se analizó con una prueba simple de T de Student de una sola vía con $\alpha = 0.05$, todas las líneas se probaron simultáneamente.

4.3. Monitoreo del retraso en la floración

Como indicador del efecto de los cassettes de expresión sobre los procesos de floración de la planta, se desarrolló un experimento en el que 15 individuos crecieron en condiciones controladas ([Sección 4.1.1](#)) y se monitoreó el tiempo de floración. Este experimento se realizó por triplicado para fines del análisis estadístico, el cual consistió de una prueba simple de T-Student con $\alpha < 0.05$.

4.4. Determinación de la expresión de *SUB1A-1* y *SUB1C-1*

Cinco individuos de cada una de las líneas transformadas de la [Tabla 4.1](#), así como de la línea control Col-0 de 14 días de edad se sometieron a dos días de inundación completa en condiciones de fotoperiodo largo (16h luz, 8h oscuridad). Se cosecharon muestras de la parte aérea al final del día y se colocaron en envoltorios de aluminio en donde se congelaron por inmersión en nitrógeno líquido (NL). Posteriormente, las muestras se mantuvieron a -80 °C en ultracongelador (Thermo Scientific 9000) hasta su uso.

Tabla 4.1: Líneas transformadas seleccionadas para la extracción de RNA

Promotor	SUB1A-1			
	Control	Inundación	Clave	Fenotipo
ADH	ADH-SUB1A-F-1 (AA1)	ADH-SUB1A-F-1 (AA1)	AA-10-1	Fuerte
	ADH-SUB1A-F-2 (AA7)	ADH-SUB1A-F-2 (AA7)	AA-18-7	Débil
35S	35S-SUB1A-F-1 (A4)		3A-3-14	Fuerte
	35S-SUB1A-F-2 (A3)		3A-9-15	Débil
	35S-F-SUB1A-1 (L5)		L5	Débil
	35S-F-SUB1A-2 (L12)		L12	Fuerte
SUB1C-1				
ADH	ADH-SUB1C-F-1 (AC6)	ADH-SUB1C-F-1 (AC6)	AC-4-6	Fuerte
	ADH-SUB1C-F-2 (AC20)	ADH-SUB1C-F-2 (AC20)	AC-9-20	Débil
35S	35S-SUB1C-F-2 (C9)		3C-9-9	Fuerte
	35S-SUB1C-F-1 (C6)		3C-6-10	Débil

4.4.1. Extracción de RNA

Las muestras se trituraron con NL en mortero estéril, en seguida se colocaron en microtubos precongelados y se almacenaron a -80 °C por 24 h. Las muestras se procesaron con el kit de extracción *Direct-zol ARN MiniPrep* (Zymo Research #R2050). Para ello, se agregaron 500 μ l del reactivo *TRI-Reagent* por cada 50 mg de tejido (aproximadamente 0.1 ml de tejido triturado contenidos en un microtubo). La mezcla resultante se homogeneizó en vortex (Genie 2, Scientific Industries) durante 3 min, posteriormente se centrifugó a 12,000 rpm (Pico 17, Thermo Scientific) durante 2 min a temperatura ambiente y el sobrenadante se transfirió a un microtubo de 1.5 ml en donde se agregaron 500 μ l de etanol puro y se mezcló vigorosamente en vortex. La mezcla se transfirió a la columna incluida en el kit acoplada a un tubo colector, se centrifugó a 12,000

rpm por 1 minuto y se descartó el sobrenadante, se añadieron 400 μ l de solución buffer de lavado (incluida en el kit) y se centrifugó por un minuto para después descartar el sobrenadante.

Para remover el ADN de la muestra, se mezclaron en un microtubo nuevo 5 μ l de DNAsa I (Zymo Research E 1011-A 250 U en 4 ml) con 75 μ l de buffer de digestión de ADN (ambos incluidos en el kit) y se mezclaron gentilmente para después añadirse a la columna e incubarse a temperatura ambiente por 15 minutos.

Posteriormente, la columna se lavó 2 veces con 400 μ l de buffer de prelavado y una con 700 μ l de buffer de lavado, ambos incluidos en el kit, y se centrifugó por 1 min a 12,000 rpm. Finalmente, la columna se transfirió a un nuevo tubo y se añadieron 25 μ l de agua libre de RNAsas y DNAsas. Se centrifugó por 2 minutos a 13,000 rpm para eluir el ARN que se almacenó a -80 °C.

4.4.2. Determinación de la concentración de ARN extraído

La concentración del ARN se midió por absorbancia en la región UV-Vis por medio de un nanoespectrofotómetro (NanoDrop 2000, Thermo Scientific). Se colocó 1 μ l de cada extracción en el equipo y se leyó a longitudes de onda de 230, 260 y 280 nm, procurando que las relaciones de absorbancia 280/260 y 260/230 fueran cercanas a 2. Los resultados se analizaron con el software NanoDrop 2000 obteniendo las concentraciones de la [Tabla 4.2](#). Posteriormente, se realizaron las diluciones correspondientes para obtener 12 μ l de cada muestra a la misma concentración (15.5 ng/ μ l), tomando como referencia aquella con la menor concentración de ARN extraído.

4.4.3. Verificación de ausencia de DNA

Para verificar la remoción de ADN de las muestras, se realizó una PCR utilizando a la enzima *GoTaq Green Master Mix* (Promega #M7122) y primers para amplificación del gen *TUBULIN2*: (*TUB2*); (GenBank: M84700.1, [Wenkel y cols., 2006](#); [Apéndice A.3](#)). A cada volumen de reacción de 19.5 μ l se le adicionó 0.5 μ l de la dilución del ARN que le correspondiera. Se utilizó el programa de la [Tabla A.1](#) ([Apéndice A.2](#)) en un termociclador automático (Techne, Tc-512).

Los productos de PCR se separaron por electroforesis en gel de agarosa/TAE al 1.6 % w/v ([Apéndice A.4.3](#)). Se utilizó un marcador de peso molecular *GeneRuler ADN Ladder Mix* (Thermo Scientific #SM0331 [Apéndice A.2](#)). Los geles se tiñeron con solución de EtBr /TAE 1X ([Apéndice A.4.5](#)) y se visualizaron con ayuda de un fotodocumentador (Syngene, InGenius LHR).

Tabla 4.2: Concentraciones calculadas de ARN en cada muestra

Línea	Condición	Concentración (ng/ μ l)
35S-SUB1A-F-2	Control	141.1
35S-SUB1A-F-2	Inundación	63.7
35S-SUB1A-F-3	Control	18.4
ADH-SUB1A-F-1	Control	32.8
ADH-SUB1A-F-1	Inundación	52.5
ADH-SUB1A-F-2	Control	72.2
ADH-SUB1A-F-2	Inundación	71.5
ADH-SUB1A-F-3	Inundación	47.0
35S-SUB1C-F-1	Control	22.7
35S-SUB1C-F-2	Control	54.3
ADH-SUB1C-1	Control	17.8
ADH-SUB1C-1	Inundación	21.1
ADH-SUB1C-2	Control	15.5
ADH-SUB1C-2	Inundación	35.0
ADH-SUB1C-2.2	Inundación	22
35S-F-SUB1A	Control	76.2
X0	Control	54.8
X0	Inundación	110.1

4.4.4. Síntesis de cDNA

Una vez que se verificó la correcta extracción de RNA, se procedió a realizar la síntesis de la primera cadena de ADN complementario (cDNA) por medio de RT-PCR de un sólo ciclo con ayuda del kit *RevertAid H minus First Strand cDNA Synthesis* (Thermo Scientific #K1632). En microtubos de 0.2 ml se colocaron 5.5 μ l de dilución de ARN de cada muestra, se añadieron 0.5 μ l de oligo (dT)₁₈ provisto con el kit, y posteriormente se añadieron 4 μ l de la mezcla de los componentes del kit descrita en la [Tabla 4.3](#). La mezcla se incubó por 60 min a 42 °C en un termociclador automático (Techne, Tc-512).

Tabla 4.3: Volúmenes de los reactivos utilizados para la síntesis de cADN(en μ l)

Componente	volumen
Buffer de reacción 5X	2
Inhibidor de ARNsa <i>RiboLock</i>	0.5
Mezcla de dNTPs 10 mM	1
Transcriptasa reversa	0.5
volumen total	4

4.4.5. Verificación de ADN sintetizado

La verificación de la correcta síntesis del cADN a partir del ARN se realizó por PCR utilizando a la enzima *GoTaq Green Master Mix* y primers para la amplificación del gen *TUB2*. Se utilizó gADN de *Arabidopsis* como control positivo y como control negativo se utilizó la reacción sin nucleótido molde. La PCR siguió el programa detallado en la [Tabla A.1 \(Apéndice A.2\)](#) en un termociclador automático. Posteriormente, los productos de PCR se visualizaron en un gel de agarosa/TAE al 2 % w/v y se tiñeron con solución EtBr/TAE 1X ([Apéndice A.4.4](#)).

4.4.6. Amplificación de los transcritos de *SUB1A-1* y *SUB1C-1*

Una vez que se verificó la integridad del cDNA, se procedió a utilizarlo como molde para realizar una PCR utilizando a la enzima *GoTaq Green Master Mix* con los primers para la amplificación de los genes *SUB1A-1* y *SUB1C-1* ([Apéndice A.3](#)). La PCR siguió el programa de la [Tabla A.1 \(Apéndice A.2\)](#) en un termociclador automático (Tecne Tc-512). Posteriormente, los productos de PCR se separaron por electroforesis en gel de agarosa al 2 % w/v en TAE 1X ([Apéndice A.4.4](#)). La imagen de la separación de los productos de PCR en el gel se visualizó con ayuda de un fotodocumentador.

4.5. Medición cualitativa de la retención de carbohidratos

Para observar cualitativamente el efecto de los cassettes de expresión sobre la dinámica de retención de carbohidratos en forma de almidón, se realizó una tinción con yodo ([Núñez-López y cols., 2015](#)). Muestras de la parte aérea de cada línea transformada, así como de la línea control Col-0 de 15 días de edad que crecieron en las condiciones descritas en la sección [Sección 4.1.1](#) se cosecharon al final de la noche y se infiltraron al vacío con una solución de formaldehído al 3.7 % en buffer de fosfatos a pH de 6.6 ([Apéndice A.4.1](#)). Posteriormente, las muestras se sumergieron en solución de etanol al 80 % a 80 °C y se calentaron en baño maría a 80 °C por 30 min para remover la clorofila, se enjuagaron con diH₂O para rehidatarlas.

Las muestras se tiñeron por inmersión en solución de Lugol ([Apéndice A.4.2](#)) bajo estricta protección luminosa por 30 minutos. Las muestras se enjuagaron con abundante diH₂O hasta dejar de apreciarse coloración en el agua de enjuague. Las muestras se observaron con ayuda de la luz blanca producida por un transiluminador.

4.6. Estrategia Computacional

Como estrategia para el análisis de la conformación tridimensional de SUB1A-1 y la posición espacial de sus dominios conservados se desarrolló un tren de modelado teórico de esta proteína.

4.6.1. Análisis bioinformático

Se obtuvo la secuencia del gen *SUB1A-1* de arroz de la base de datos NCBI bajo el número de acceso DQ011598.1_4912..5747 (GI: 93504064, Xu y cols., 2006; Fukao y cols., 2009). Posteriormente se tradujo a proteína por medio del algoritmo *Expasy Translate Tool* (Gasteiger y cols., 2003, web.expasy.org/translate/) y se comparó con la reportada previamente (gb: AAZ06207.1).

Se obtuvo la secuencia consenso del dominio AP2 para la familia VII de los ERFs (Nakano y cols., 2006; Fukao y cols., 2006; Peña-Castro y cols., 2011), y se verificó por multialineamientos de secuencia con el algoritmo MUSCLE (Edgar, 2004, <http://www.ebi.ac.uk/Tools/msa/muscle/>) en el paquete de software Unipro UGENE (Okonechnikov y cols., 2012, <http://ugene.net>). Una vez que se obtuvieron los residuos conservados, se alinearon a la secuencia de *SUB1A* para determinar la posición y longitud del dominio AP2.

Se utilizó la suite bioinformática MEME (Bailey y cols., 2009; <http://meme-suite.org/>) para la elucidación de las posiciones de los dominios reportados por Nakano y cols. (2006) presentes en la proteína SUB1A, se aplicó el criterio de discriminación al software. Una vez que se localizaron los dominios presentes, se buscó la posición específica por multialineamientos de secuencia con ayuda del algoritmo Muscle.

4.6.2. Uso del servidor I-Tasser

La secuencia de la proteína de SUB1A-1 (gb: AAZ06207.1, Apéndice A.1.2) se envió al servidor de búsqueda de estructura secundaria por homología I-Tasser sin restricciones (Roy y cols., 2010, zhanglab.ccmb.med.umich.edu/I-TASSER). Se usó al dominio AP2 como "semilla" debido a su alta estabilidad. Los modelos arrojados por este servidor se analizaron por medio del C-score (prefiriendo valores cercanos a 0) para determinar el modelo más confiable.

Adicionalmente, se evaluó el coeficiente de desviación cuadrática media o RMS de los alineamientos estructurales de los modelos producidos por I-Tasser y el archivo de coordenadas de la región AP2 (PDB ID: 2GCC). En este punto se verificó que la región de anclaje a ADN no se

encontrara bloqueada, y que la porción ordenada del modelo fuera la mayor posible.

Se seleccionó como mejor modelo aquel con un buen C-score, un RMSD pequeño, la mayor porción ordenada y el menor bloqueo a la región de anclaje.

4.6.3. dinámica molecular

El mejor modelo del paso anterior se refinó por DM, por medio del paquete de software GRO-MACS (Bekker *y cols.*, 1993; van der Spoel *y cols.*, 2005) en su versión 5.1. Se utilizó el algoritmo de descenso rápido para la minimización de energía, posteriormente, se realizó una simulación por DM de todos los átomos utilizando agua como disolvente y equilibrando las cargas de la proteína con NaCl, para después procesarse con el campo de fuerza CHARMM 27 en estaciones de trabajo híbridas CPU/GPU (NVIDIA 700 Cuda engine) en cinco etapas descritas a continuación:

- 855 picosegundos (ps) de minimización de energía
- 200 nanosegundos (ns) de simulación por MD en equilibrio constante de concentración, volumen y temperatura (NVT) a 300 K.
- 80 *ns* de equilibrio a concentración, presión y temperatura constantes (NPT) a 298 K.
- 20 *ns* de SA de 21 pasos con intervalo de incremento/decremento de 30 K.
- MD de producción NPT por el tiempo necesario para alcanzar una conformación estable. Para esto se tomó como criterio decisivo el aporte del radio de giro total (rGXYZ), el coeficiente de desviación cuadrática media (RMSD) y la evolución de estructura secundaria. Al final de esta etapa se dispondrá de un modelo putativo de SUB1A-1.

4.6.4. Acoplamiento Molecular

Finalmente, el modelo putativo de la estructura de SUB1A-1 se sometió a una sesión de Acoplamiento Molecular rígido y flexible mediante los paquetes de software Vina v1.1.2 de AutoDock y CLC Drug Discovery Workbench v2.5 de Quiagen. Se calcularon en esta etapa los potenciales de interacción de la proteína SUB1A-1 y su energía de acoplamiento con una región afín a su sitio de interacción (PDB ID: 1GCC:BxC) para validar el modelo propuesto. Lo anterior para demostrar que la estructura reproduce las condiciones naturales, en este punto se realizó el análisis de los dominios.

Una vez que se validó el acoplamiento estructural, se localizaron espacialmente los dominios de la proteína SUB1A-1 encontrados por análisis de secuencia y se determinó su posición y se

determinó el efecto del bloqueo espacial de la proteína por el uso del marcaje inmunológico.

5. RESULTADOS Y DISCUSIONES

5.1. Efecto del cassette de expresión sobre el fenotipo

La transformación con los cassettes de expresión empleados ([Figura 4.1](#)) tuvo efectos diferentes sobre el fenotipo de *Arabidopsis*. Para estudiar lo anterior, se seleccionaron como controles para realizar las comparaciones a las plantas tipo silvestre (Col-0) y la línea 35S-F-SUB1A previamente reportada ([Peña-Castro y cols., 2011](#)). 35S-F-SUB1A, posee una protección de secuencia FLAG (DYKDDDDK) en el extremo N-terminal, se caracterizó por tener hojas pequeñas y ensanchadas, con peciolo cortos, un color verde claro, floración tardía, clorosis temprana, y tamaño menor respecto a Col-0 ([Figura 5.1](#))

En las líneas transformadas con el cassette 35S-SUB1A-F, que deja al extremo N-terminal desprotegido, las hojas también se redujeron en tamaño, adoptaron una forma ensanchada, y presentaron clorosis; pero el fenotipo fue menos pronunciado que en las líneas 35S-F-SUB1A. Incluso, se pudieron recuperar líneas con un fenotipo más moderado (35S-SUB1A-F2) que no se pudieron encontrar cuando el extremo N-terminal se bloqueó ([Peña-Castro y cols., 2011](#)). Todas las líneas tipo 35S-SUB1A-F fueron fértiles, mientras que las líneas 35S-F-SUB1A son estériles, en su mayoría (sólo una fértil se recuperó en un escrutinio de más de 300 individuos). Estos resultados indican que al proteger al extremo N-terminal, en este caso con una secuencia FLAG, se fomentaría una mayor estabilidad de la proteína SUB1A-1 que se observa como un fenotipo más intenso que el que se observa cuando la secuencia FLAG se coloca en el extremo C.

En el caso de las líneas pADH (ADH-SUB1A-F) el fenotipo se mantuvo casi sin cambios con respecto a Col-0, la excepción fue la presencia de un color morado en las hojas maduras. De esta forma, el cambio de promotor, permitió eliminar los efectos negativos sobre el crecimiento de la planta que se han observado al usar una expresión constitutiva.

Cuando el gen *SUB1C* se expresó de forma constitutiva con una protección FLAG en el extremo N-terminal, el fenotipo con respecto a *SUB1A* fue muy diferente. *SUB1C* provocó el enrollamiento de las hojas y una casi total esterilidad de las plantas ([Peña-Castro y cols., 2011](#); [Figura 1.9](#)), esto último fue la razón por la cual no se incluyeron en este trabajo dichas líneas trans-

génicas, pues su descendencia se ha agotado. Sin embargo, al retirar la protección N-terminal (35S-SUB1C-F), el fenotipo de esterilidad desapareció y el enrollamiento de hojas, se presentó, pero disminuido.

Por otra parte, las líneas ADH-SUB1C-F mantuvieron hojas del mismo tamaño que la línea control Col-0, sin embargo las hojas maduras presentaron un color morado y aparecieron en ellas bordes dentados, y menor enrollamiento. La altura del tallo floral de estas líneas fue indistinguible de Col-0. Lo anterior sugiere que la inserción de un promotor constitutivo resulta dañino para la planta ya que no tiene la posibilidad de detener los efectos del mecanismo quiescente aun cuando el estrés no este presente. Por otra parte, cuando el promotor promueve la transcripción de los genes *SUB1* de forma condicional, el efecto sobre la fisiología es más discreto. La intensidad en el efecto fisiológico observado, también se relaciona directamente con el tiempo que el mecanismo quiescente está activo, por lo que la inserción de un promotor inducible parece ser la opción biotecnológica más adecuada si lo que se desea es obtener la menor cantidad de modificaciones fisiológicas en el organismo.

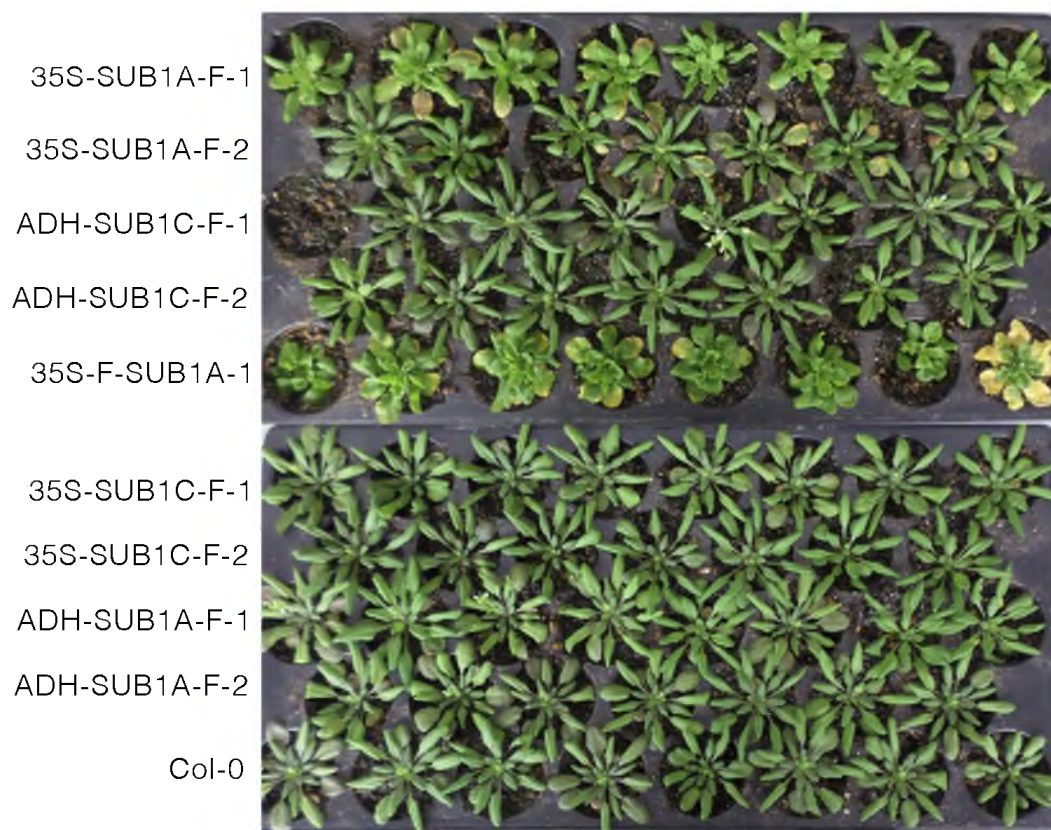


Figura 5.1: Fenotipo de las líneas de *Arabidopsis thaliana* analizadas en este trabajo.

5.2. Efecto del cassette de expresión sobre la tolerancia a la inundación

Además de las modificaciones en el fenotipo, las líneas transformadas sufrieron modificaciones en la capacidad de sobrevivencia al estrés por inundación como resultado directo de la manipulación genética recibida. Después de someterse a inundación por periodos de 2, 4, 6 y 8 días y tras una etapa de recuperación de 7 días, se monitoreó el porcentaje de supervivencia y se expresó en forma de tiempo letal 50 (LT_{50}).

De igual modo que con la modificación fenotípica, la resistencia a la inundación resultó en un evento dependiente del arreglo de los elementos que integraron el cassette de expresión, en especial la protección N-terminal. El conocimiento previo indica que la expresión constitutiva de *SUB1A*, contrario a lo que se deseaba, ocasionó un fenotipo sensible a la inundación. Estos resultados se confirmaron en este trabajo pues las líneas 35S-F-SUB1A fueron las de menor tolerancia al estrés, por el marcador FLAG bloqueando el extremo N-terminal. La remoción de la protección no mejoró la condición, pues fueron las líneas que siguieron en sensibilidad al estrés. Finalmente, la mayor sobrevivencia al estrés por inundación observada perteneció a las líneas cuyos cassettes están compuestos por el promotor inducible ADH y el transgén *SUB1C*; todavía con diferencias significativas con respecto a Col-0. De tal manera, ninguna construcción fue capaz de transferir la capacidad de tolerancia que se observa en el arroz, a *Arabidopsis* (Figura 5.2).

Todo lo anterior muestra que el orden de los elementos en el cassette de transformación utilizado en cada una de las líneas tiene un efecto sobre la capacidad de supervivencia de los individuos (Tabla 5.1). Se pudo notar que la resistencia a inundación otorgada a los individuos de las líneas ADH-SUB1C es la más cercana a la de la línea control, puesto que en ambos casos el valor de LT_{50} es cercano a 9 (ADH-SUB1A-F-2) o superior (ADH-SUB1a-F-1).

Debido a que *SUB1C* no ha sido asociado a la tolerancia al estrés por inundación (Niroula y cols., 2012), se puede hipotetizar que el fenotipo Col-0, que se ha usado extensamente en la investigación, es un fenotipo tolerante a la inundación de forma natural. Así, puede ser difícil observar fenotipos con tolerancia mejorada en este fondo

Vashisht y cols. (2011) realizaron un escrutinio de 86 variedades de *Arabidopsis* en inundación y encontraron una variabilidad en la LT_{50} desde 4 hasta 11 días. Col-0 tuvo un registro de 9 días, colocándose en la parte superior de la tabla. Los ecotipos más sensibles como Cvi-0, Tanz 0 Bay-0, pueden ser mejores candidatos para probar nuevas estrategias biotecnológicas.

Tabla 5.1: Tiempos letales 50 (LT_{50}) de las líneas analizadas durante el estrés por inundación

Línea	LT_{50}	
Control (Col-0)	9.3 ± 0.3	a
35S-F- <i>SUB1A</i> -1	4.8 ± 0.5	b
35S-F- <i>SUB1A</i> -2	4.4 ± 0.3	b
35S- <i>SUB1A</i> -F-1	6.1 ± 0.6	c
35S- <i>SUB1A</i> -F-2	4.7 ± 0.4	b
35S- <i>SUB1A</i> -F-3	5.0 ± 0.5	b
ADH- <i>SUB1A</i> -F-1	6.6 ± 0.3	c
ADH- <i>SUB1A</i> -F-2	6.8 ± 0.2	c
ADH- <i>SUB1A</i> -F-3	5.6 ± 0.2	c
35S- <i>SUB1C</i> -F-1	7.9 ± 0.1	d
35S- <i>SUB1C</i> -F-2	6.6 ± 0.2	c
ADH- <i>SUB1C</i> -F-1	13.9 ± 0.5	e
ADH- <i>SUB1C</i> -F-2	8.6 ± 0.3	d

Letras diferentes indican diferencias estadísticamente significativas, prueba T de Student con $\alpha < 0.05$.

Los datos son el resultados de tres experimentos independientes con $n=15$.

5.3. Efecto del transcrito sobre la floración

Durante el estudio de las líneas 35S-F-*SUB1A* se observó que el cambio fenotípico más notorio fue el atraso en la floración. Al profundizar más, se descubrió que las plantas que expresan a *SUB1A*, tanto en *Arabidopsis* como en arroz, reprimen la expresión de los genes esenciales de floración *CONSTANS (CO)* y *FLOWERING LOCUS T (FT)*. El resultado de este mecanismo sería la inhibición temporal de la floración para ahorrar recursos energéticos durante la inundación dentro del esquema de la quiescencia (Peña-Castro y cols., 2011).

En esta tesis, se observó, como resultado de la inserción de las diferentes combinaciones de cassettes de expresión en las líneas transformadas, un desplazamiento en los periodos de floración de las líneas transformadas con respecto a Col-0.(Figura 5.3).

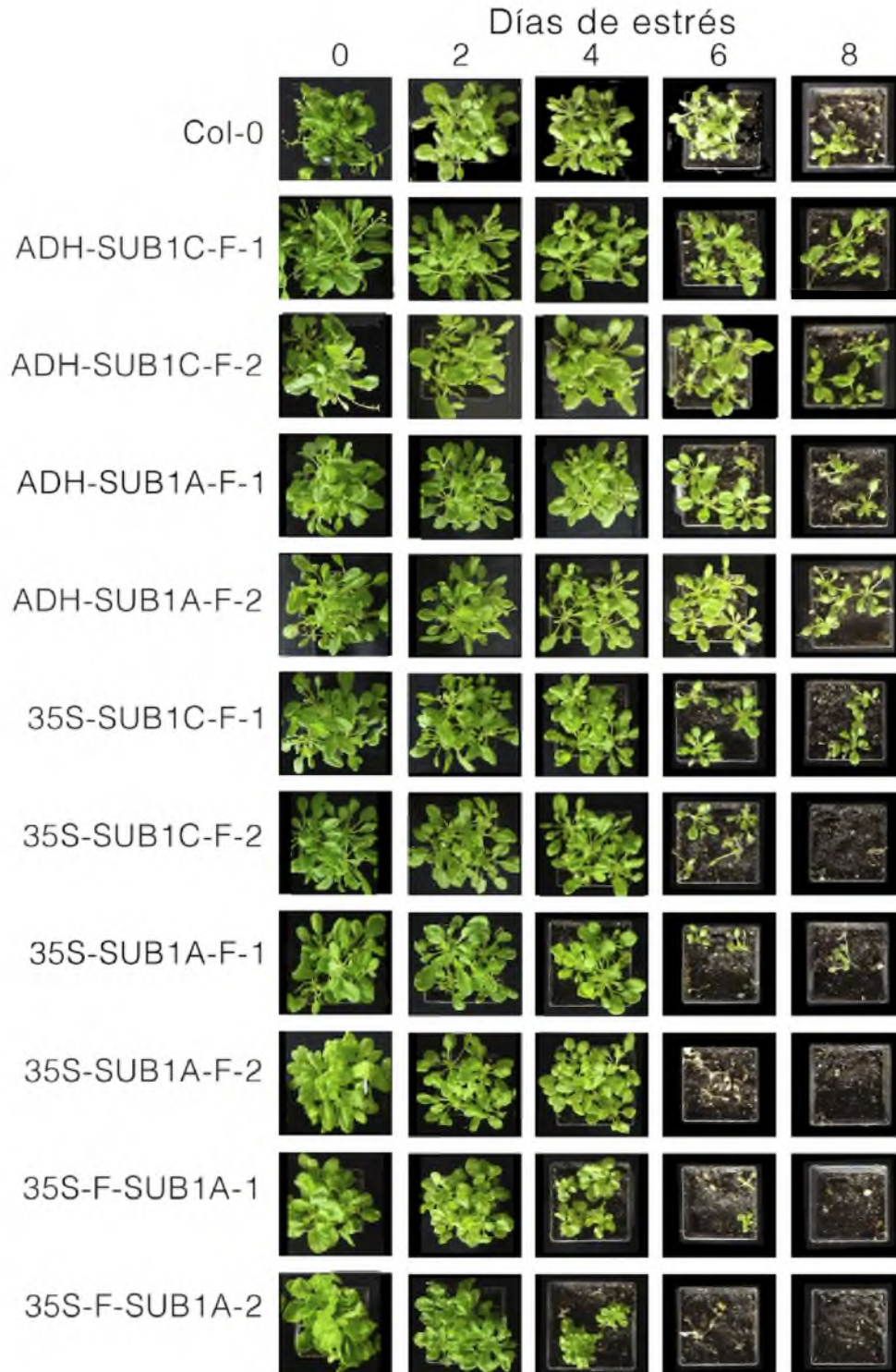


Figura 5.2: Supervivencia de plantas de *Arabidopsis* a la inundación a los 7 días posteriores al estrés aplicado durante los tiempos indicados.

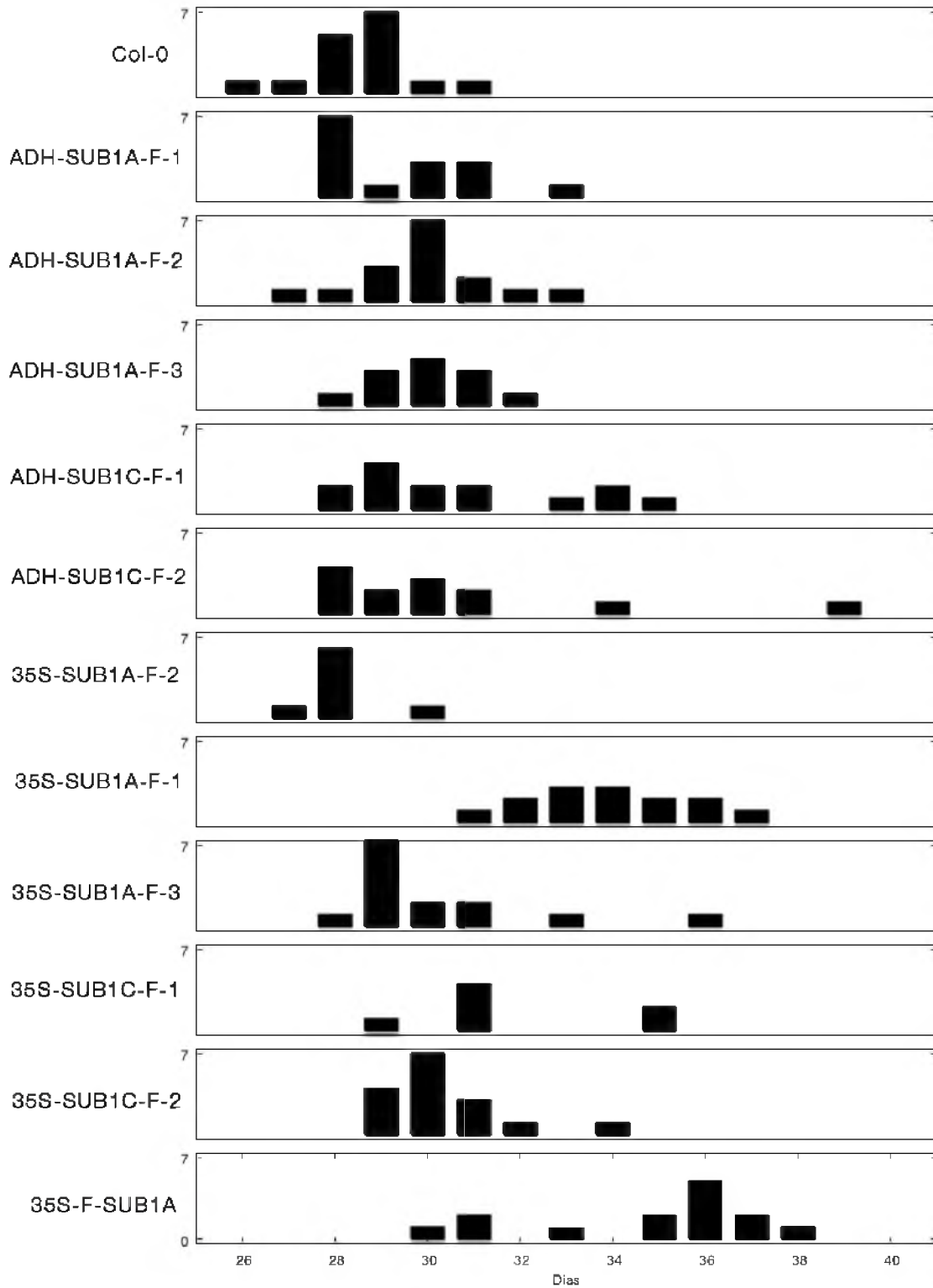


Figura 5.3: Valores promedio de los tiempos de floración de las líneas utilizadas. 3 experimentos

Primero, la floración de la línea 35S-F-SUB1A se atrasó hasta los 35 - 38 días, estos son 9 días después que Col-0. En las líneas sin protección en el N-terminal 35S-SUB1A-F, el intervalo de floración fue menor que el de la línea protegida. Las líneas provistas del promotor inducible pADH también incrementaron el intervalo de tiempo a floración, pero casi indistinguible de Col-0.

Este experimento añade evidencia experimental a la conclusión de que la combinación y el orden de los elementos en el cassette de expresión es responsable directo en la intensidad de las variables de respuesta (floración, resistencia a la inundación y fenotipo) pues, cuando la intensidad del cassette se espera sea baja, como con pADH o sin protección en el extremo N-terminal, los efectos observados son menores en comparación con la línea 35S-F-SUB1A.

5.4. Efecto del transgén sobre la retención de carbohidratos

En trabajos anteriores, se observó que las plantas que expresan a *SUB1A* con el extremo N-terminal protegido, retienen más almidón que las plantas silvestres en días largos y etapas juveniles (Núñez-López y cols., 2015).

Para registrar el efecto de los cassettes de expresión sobre esta característica, tres plántulas de cada línea de 15 días de edad cultivadas en fotoperiodo largo y corto, se cosecharon al final de la noche y se tiñeron en solución de Lugol. En este experimento se pudo comprobar que la línea 35S-F-SUB1A tuvo más hojas con coloración azul, y que las líneas que expresaron a *SUB1C* constitutivamente tenían menos, respecto a Col-0 (Figura 5.4).

La remoción de la protección FLAG no disminuyó la coloración en las líneas *SUB1A*, lo que indica que esta característica se puede lograr con ambos tipos de construcciones en día largo. No fue así cuando *SUB1A* estuvo bajo el control de pADH donde se perdió la capacidad de acumular almidón.

Cuando *SUB1C* estuvo bajo el control de pADH, se observó un efecto discrepante entre las dos líneas probadas, una acumuló más que Col-0 (2) mientras que otra se mantuvo igual (1). Para discernir esta diferencia, se propone recurrir a métodos cuantitativos como los de trabajos anteriores.

Cuando este experimento se realizó en días cortos, se observó que la capacidad de retener almidón de la línea 35S-F-SUB1A y la capacidad de perderlo de las líneas 35S-SUB1C-F, se mantuvo. Ante el cambio de régimen luminoso, las líneas pADH-SUB1A perdieron la capacidad de retener almidón que sí se observó en días largos (Figura 5.4).

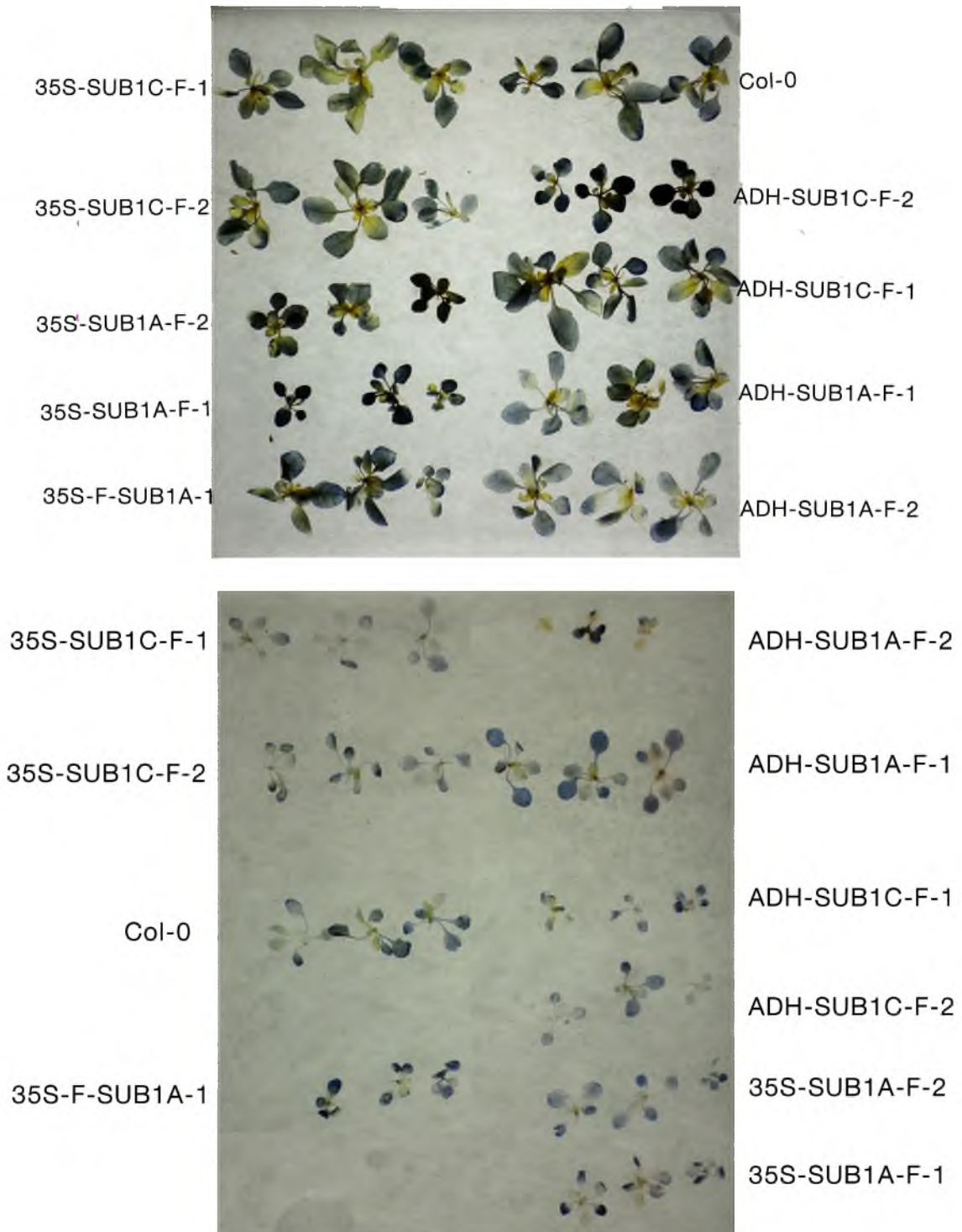


Figura 5.4: Tinción con yodo de plántulas de 14 días de edad al final de la noche. Arriba, fotoperiodo largo (16h luz, 8h oscuridad); Abajo, fotoperiodo corto (8h luz, 16h oscuridad). El almidón se observa de color azul

5.5. Expresión de *SUB1A-1* y *SUB1C-1*

Para observar la expresión de los transgenes en las plantas estudiadas, se seleccionó una línea para cada gen que estuviera bajo el control del promotor constitutivo fuerte 35S. Para *SUB1A* fue la línea 35S-SUB1A-F-1 y para *SUB1C* fue 35S-SUB1C-F-1. Como se observa en la [Figura 5.5](#), se detectó una banda de intensidad diferente a la observada en el control Col-0 no transformado. De esta manera, se comprobó que las construcciones realizadas sí tienen la capacidad de expresar los genes de interés en *Arabidopsis*.

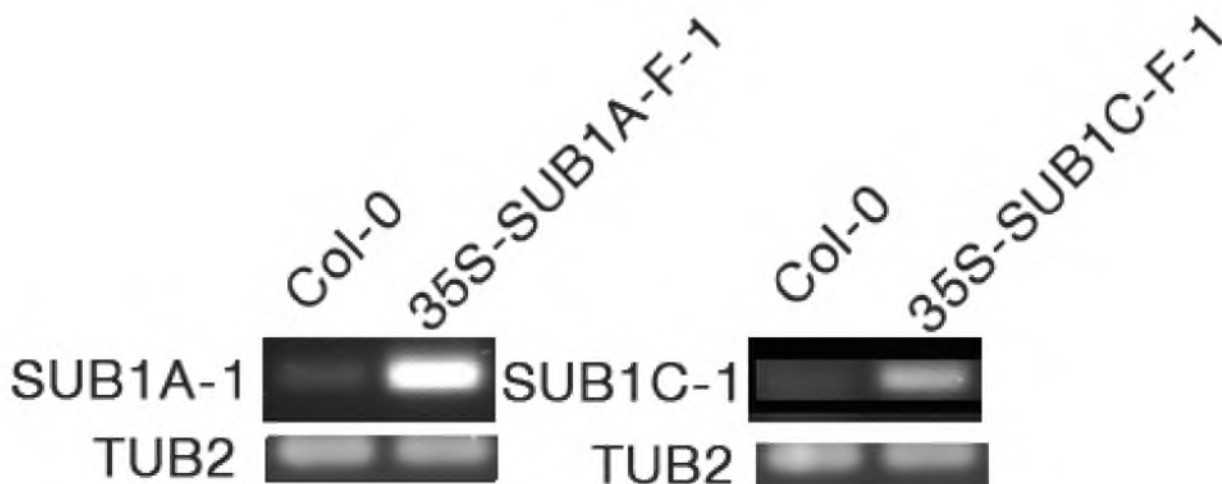


Figura 5.5: Expresión de los genes *SUB1A-1* y *SUB1C-1*. Se observan geles de agarosa al 2% teñidos con EtBr con la separación electroforética de RT-PCR de los genes y cADNs indicados

5.6. Obtención de un modelo teórico para la proteína SUB1A-1

Para observar la accesibilidad del dominio N-terminal en la proteína SUB1A-1, se realizó una simulación computacional por medio de dinámica molecular. La secuencia de la proteína SUB1A-1 se sometió al servidor I-TASSER sin restricciones. El servidor arrojó 5 modelos que se muestran en la [Figura 5.6](#), los cuales se analizaron alineando espacialmente con el archivo de cristalografía del dominio AP2 (PDB:1GCC). Se puede observar que en los casos de los modelos 1, 3 y 5, AP2 se encuentra bloqueado, por lo que una simulación realizada a partir de cualquiera de estos modelos no representa la realidad biológica del factor de transcripción. Se seleccionó al modelo 2 por tener a AP2 disponible, un valor de C-Score alto (-4.00), un RMS alto y la menor porción

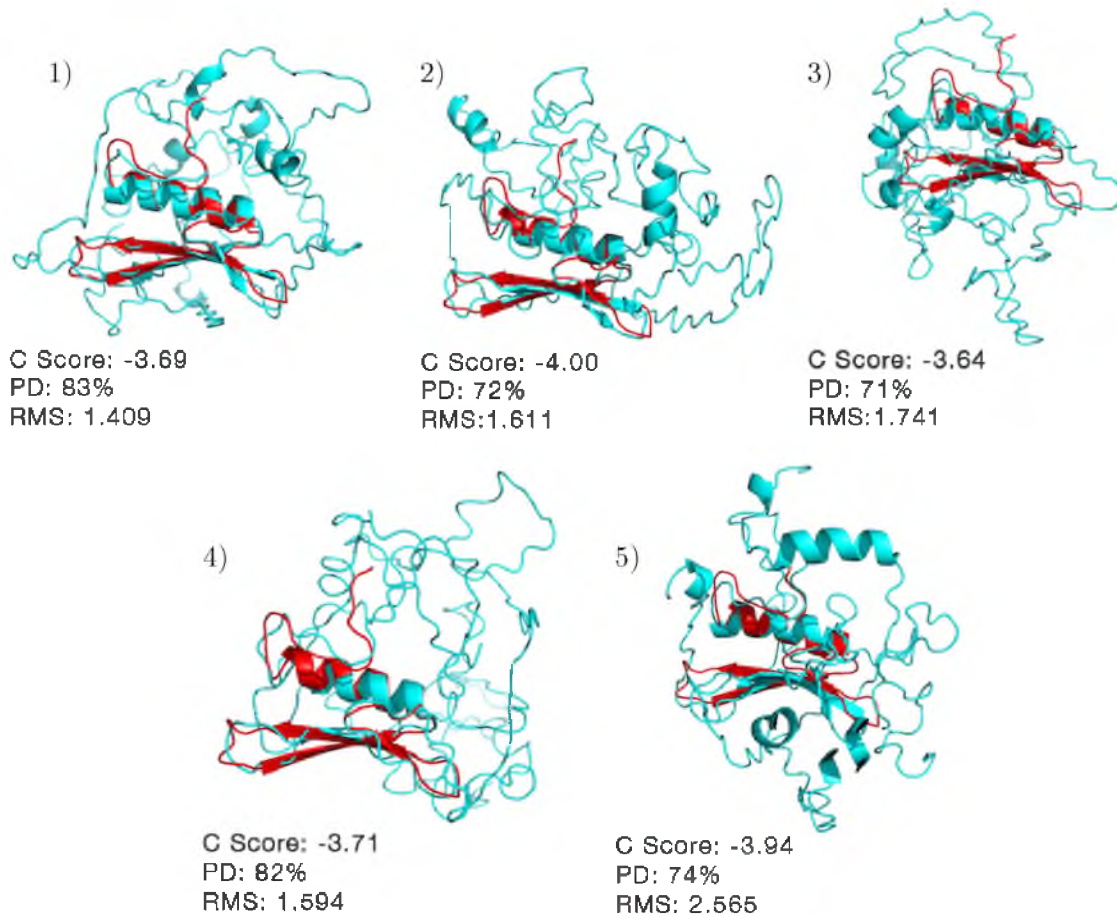


Figura 5.6: Modelos de la proteína SUB1A-1 arrojados por el servidor I-TASSER (azul) alineados con el archivo de cristalografía de dominio AP2 (rojo)

desordenada.

5.7. Refinamiento estructural por dinámica molecular

Una vez que se determinó que el modelo 2 resultó el más adecuado para ser refinado por MD, se determinó por cálculo de distancias que la celda de simulación cúbica óptima es de 9.496 nm de arista con el centro de masa de la proteína en las coordenadas 4.748,4.748,4.748. Posteriormente, la celda se llenó con 25,572 moléculas agua utilizando el modelo rígido simple tip3p (Jorgensen *y cols.*, 1983). La carga neta de la proteína (-12) se equilibró por adición de NaCl, para

ello, se sustituyeron 528 moléculas de agua con 270 átomos de sodio y 258 átomos de cloro para obtener una carga neta de cero y una concentración salina de 0.05 mMol. Así, se representaron las condiciones fisiológicas del citoplasma vegetal, el pH se supuso cercano a la neutralidad en 6.7.

5.7.1. Minimización de energía

Una vez que se equilibraron las cargas de la proteína, se realizó una simulación de todos los átomos sobre la celda de simulación en el campo de fuerza CHARMM 27 (Karplus, 1983). Inicialmente, el sistema se minimizó energéticamente con el algoritmo de Descenso Rápido con las condiciones listadas en la Tabla A.2 (Apéndice A.6) y convergió en un valor de energía aceptable de 826.83 KJ/mol/nm después de 855 pasos (855 ps), en este punto el valor de la energía potencial fue de -1.4006482×10^6 KJ/mol/nm (Figura 5.7).

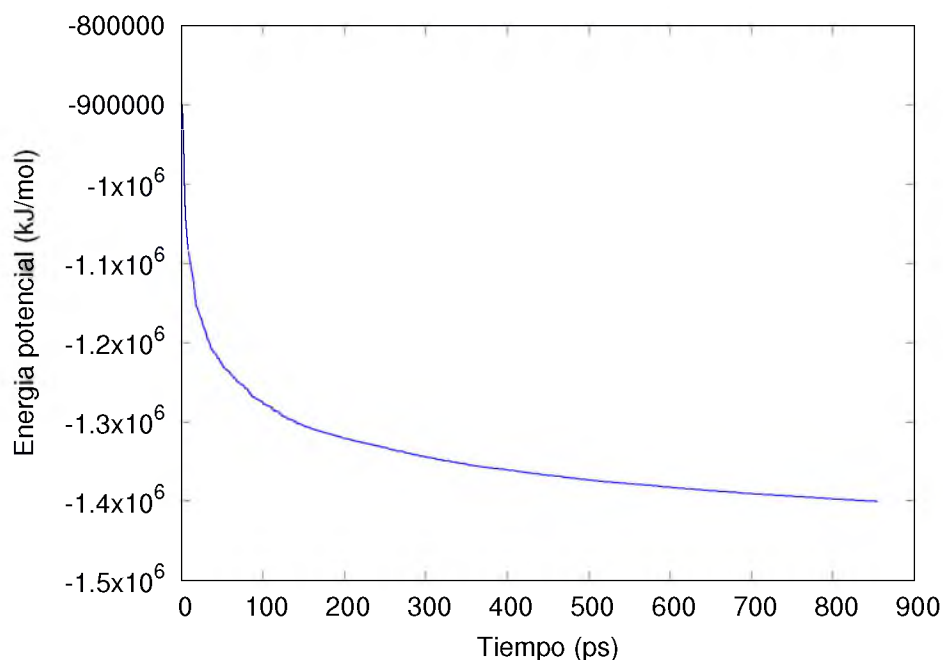


Figura 5.7: Evolución de la energía potencial durante el proceso de minimización de la proteína SUB1A-1.

5.7.2. Equilibrio NVT

Una vez que la celda de simulación que contenía al modelo 2 alcanzó la estructura de menor energía, se realizó sobre ella una ronda de simulación de 200 ps en el ensamble NVT (número

de partículas, volumen y temperatura constantes) con el fin de permitir que el sistema adoptara la conformación de menor energía sin el efecto del barostato (algoritmo de restricción en el aumento o disminución de la presión en la celda de simulación).

Para esta etapa, la temperatura inicial de la celda se configuró a 300 K y para su control se utilizó la modificación de V-rescale al termostato de Berendsen (Bussi *y cols.*, 2007). El volumen de simulación se mantuvo fijo en 856.2924 nm³. El total de átomos simulados fue de 79858, de los cuales, 4198 son de la proteína (281 residuos de amino ácidos), 75132 son de moléculas de agua, 270 átomos de sodio y 258 átomos de cloro. Los parámetros utilizados para esta etapa se detallan en la [Tabla A.3](#) (Apéndice A.6).

Después de 8 ps, la presión del sistema tomó un valor estable y se mantuvo en valores cercanos a -1000 bar ([Figura 5.8](#)) por lo que no se requirió de mayor tiempo de simulación para esta etapa.

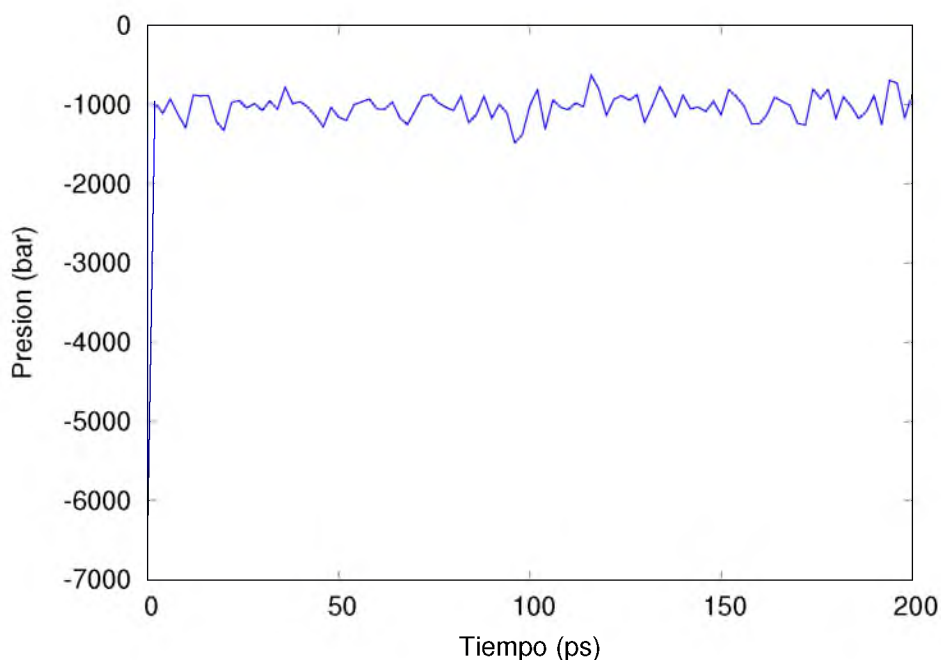


Figura 5.8: Evolución de la presión del sistema durante la simulación de la proteína SUB1A-1 en NVT.

5.7.3. Equilibrio NPT

Una vez que la celda se encontraba en su estado de menor energía y después de alcanzar el equilibrio de presión en el ensamble NVT, se realizó sobre ella una simulación de 80,000 ps (80 ns) en el ensamble NPT (número de partículas, presión y temperatura constantes) con los

parámetros descritos en la [Tabla A.4 \(Apéndice A.6\)](#). Esta etapa de la simulación permitió obtener un estado estable de la proteína que mostraba indicios de plegado en forma de cambios en el radio de giro del esqueleto de la proteína o *backbone* indicando que su posición estaba cambiando por efecto de la evolución del sistema como se puede apreciar en la [Figura 5.9](#). Sin embargo, el valor absoluto máximo de este radio de giro después de 80 ns resultó en tan solo 0.01 nm, mismo que sugería más tiempo de simulación para alcanzar el óptimo adecuado. A pesar de ello, la pendiente promedio de esta propiedad se mantenía positiva de acuerdo a lo reportado en la bibliografía ([Papdi y cols., 2015](#))

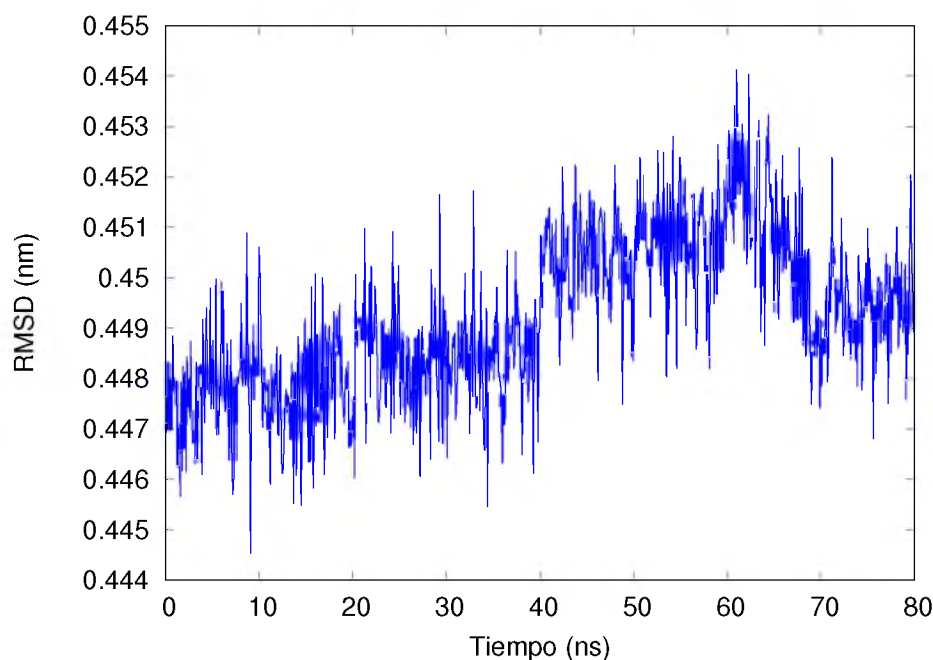


Figura 5.9: Movimiento del backbone de la proteína SUB1A-1 durante 80 ns de simulación en NPT

5.7.4. Recocido simulado (GSA)

Para reducir el tiempo total de cómputo posterior al periodo de simulación NPT, se realizó una simulación con el método de recocido simulado o GSA durante 20 ns divididos en 21 etapas de aproximadamente 952 ps cada una. Durante las primeras 8 etapas se incrementó la temperatura del sistema hasta un valor de 510 K en intervalos de 30 K con el fin de dotarlo de energía cinética (aproximadamente 340 000 kJ/mol) y así inducir los movimientos en la arquitectura de los grupos R de los residuos de aminoácidos y posteriormente del esqueleto carbonado. Los intervalos pequeños de tiempo impidieron la descomposición de los enlaces de la proteína. Los 7

intervalos posteriores llevaron este valor nuevamente a los 300 K para forzar a esta estructura altamente energizada a adoptar la forma de menor energía y de mayor orden. Finalmente, los 6 últimos intervalos permitieron a la estructura estabilizarse a 300 K (trazo rojo en la [Figura 5.10](#)). Los parámetros de esta etapa se detallan en la [Tabla A.5](#) ([Apéndice A.6](#)).

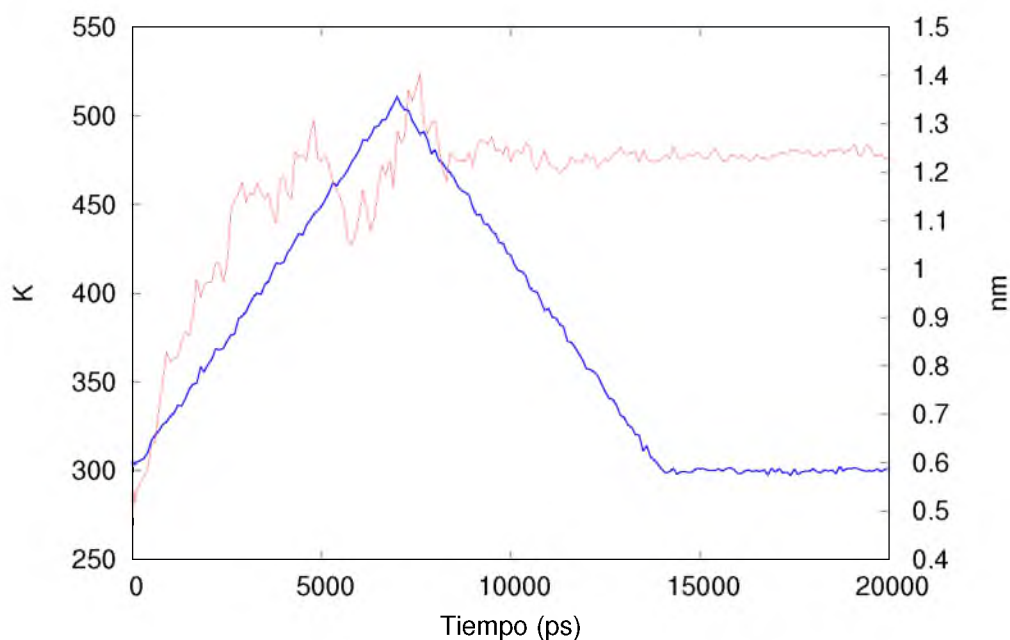


Figura 5.10: Temperatura (trazo azul, eje izquierdo) y RMSD (trazo rojo, eje derecho) del sistema de la proteína SUB1A-1 durante la simulación en GSA

Al finalizar la simulación GSA, la estructura adoptó una configuración espacial más ordenada que en el paso anterior. Adicionalmente, el valor absoluto del indicador de plegado o RMSD, con valor de 0.95 nm (trazo rojo en la [Figura 5.10](#)), fue 90 veces mayor que el obtenido en la simulación NPT previa.

Tal como se esperaba, la energía potencial y la densidad del sistema se comportaron de manera recíproca ([Figura 5.11](#)). Esto es un indicador de la estabilidad de la proteína durante los incrementos y decrementos en la energía cinética a los que se sometió.

5.7.5. Dinámica de producción en NPT

Aún cuando la simulación en GSA permitió modificar el valor de RMSD más rápido que la simulación previa NPT, se puede suponer que el tiempo tan corto de simulación en esta etapa podría representar un estado no estable de la molécula. Para descartar esta posibilidad, después de la simulación en GSA se realizó una simulación de producción con las mismas condiciones

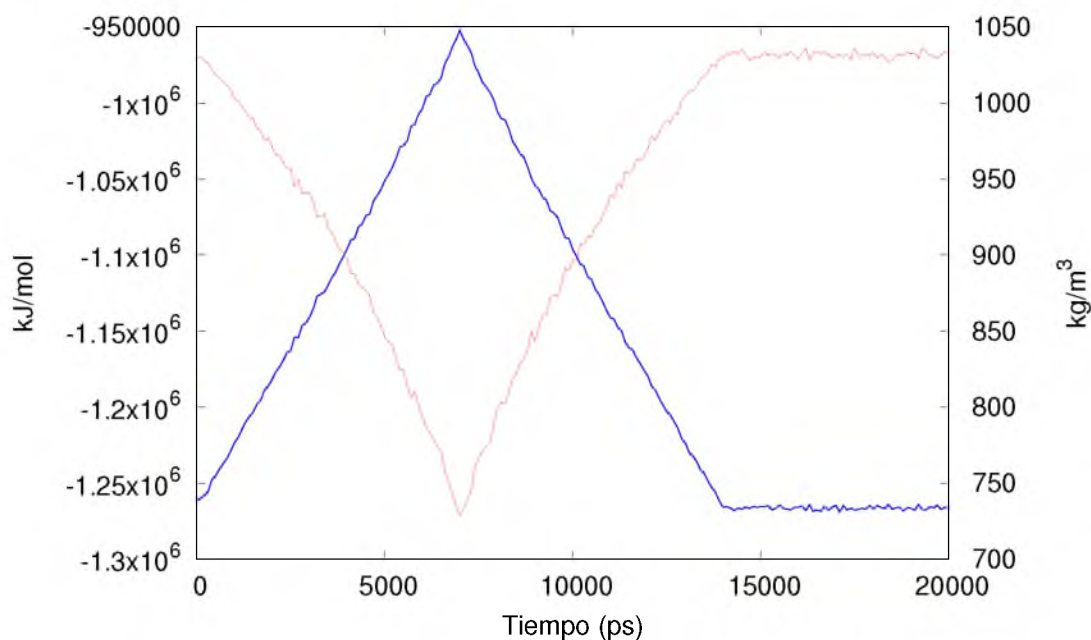


Figura 5.11: Evolución de la Energía potencial (trazo azul, eje izquierdo) y la densidad (trazo rojo, eje derecho) del sistema de la proteína SUB1A-1 durante la simulación GSA

que la sección NPT previa, con el fin de dar al sistema la cantidad de tiempo suficiente para adoptar, si así fuese, la estructura más relajada de la proteína o en caso de que la etapa GSA hubiese permitido alcanzar la estructura más estable, aumentar el valor del RMSD para obtener una imagen del sistema aún más definida.

En esta etapa se utilizaron 2 parámetros de control, el primero de ellos fue el valor del RMSD, que debía adoptar valores estables por al menos 40 ns, y el segundo, el diccionario de estructuras secundarias (DSSP; *Define Secondary Structure of Proteins*) que determina mediante el análisis de los ángulos del *backbone* si un residuo pertenece a una estructura tridimensional definida de α -hélices, hojas- β , 3-helix, giros, puentes o bucles.

Para llevar a cabo este objetivo, se realizaron etapas de simulación en NPT de 20 ns cada una. Al final de cada etapa, se monitoreó la evolución del RMSD y DSSP y cuando alguno de ellos no fue estable, se realizó una etapa posterior. Esta estrategia produjo resultados positivos después de 55 etapas de simulación, logrando un tiempo total de simulación (desde la minimización de energía) de 1,221,085 ps (1.22 μ s), lo cual supone una mejora considerable en comparación de los actuales métodos de plegado de proteínas que requieren en promedio de 7 μ s para obtener valores estables de los parámetros enlistados y que se limitan a un sistema de no más de 20 residuos.

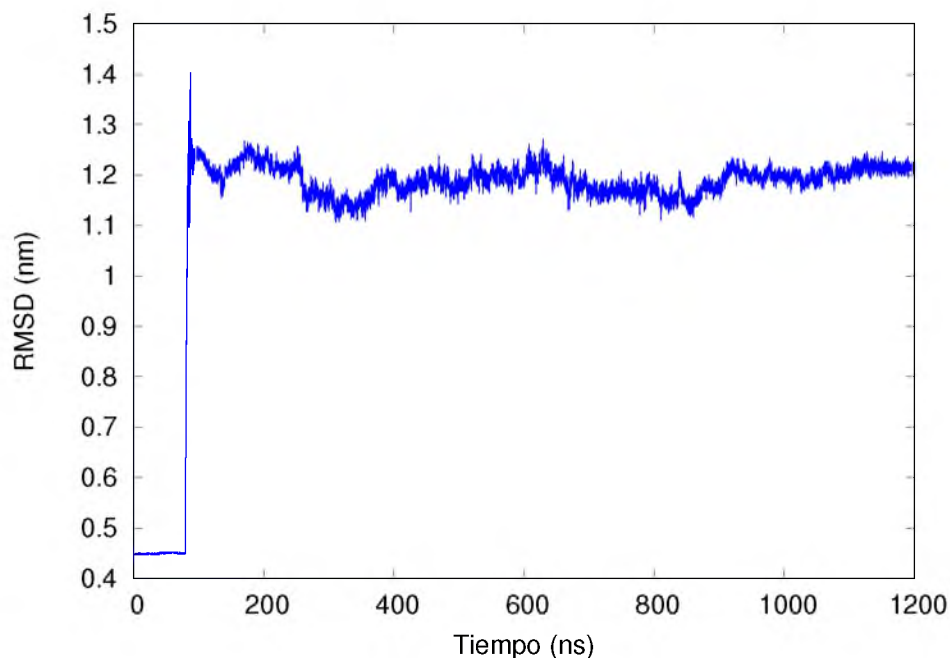


Figura 5.12: Movimiento del esqueleto carbonado durante todas las etapas de simulación

Al finalizar esta segunda etapa en NPT, el esqueleto carbonado de la estructura tuvo un RMSD con valor de 1.21 nm con respecto al modelo inicial arrojado por el servidor I-TASSER. Este valor se encontró estable a partir de un tiempo de simulación de $1.1\mu s$ (Figura 5.12). Esto supone que el proceso aquí descrito logró producir cambios estructurales en la proteína SUB1A-1. Sin embargo, el RMSD no aporta información acerca de cuáles son dichos cambios, es por ello que se recurrió al algoritmo DSSP, cuya imagen global se presenta en la Figura 5.13. El eje de las ordenadas indica la suma de residuos que se añaden a una de las estructuras definidas, por lo que es de suponerse que cuando estos valores se tornen estables, la estructura dejará de sufrir modificaciones y por lo tanto se habrá alcanzado una configuración del sistema válida para las condiciones fisiológicas establecidas al inicio de la simulación.

Esta imagen también permite observar que, en etapas tempranas de la simulación, la estructura de la proteína se encontraba en su mayoría conformada por *loops* o *coils* (trazo negro), lo cual concuerda con el alto valor de la porción desordenada. Adicionalmente, la Figura 5.13 muestra que las hojas beta se encontraban estables desde etapas tempranas de la simulación, por la decisión de usar al segmento AP2 como semilla de homología.

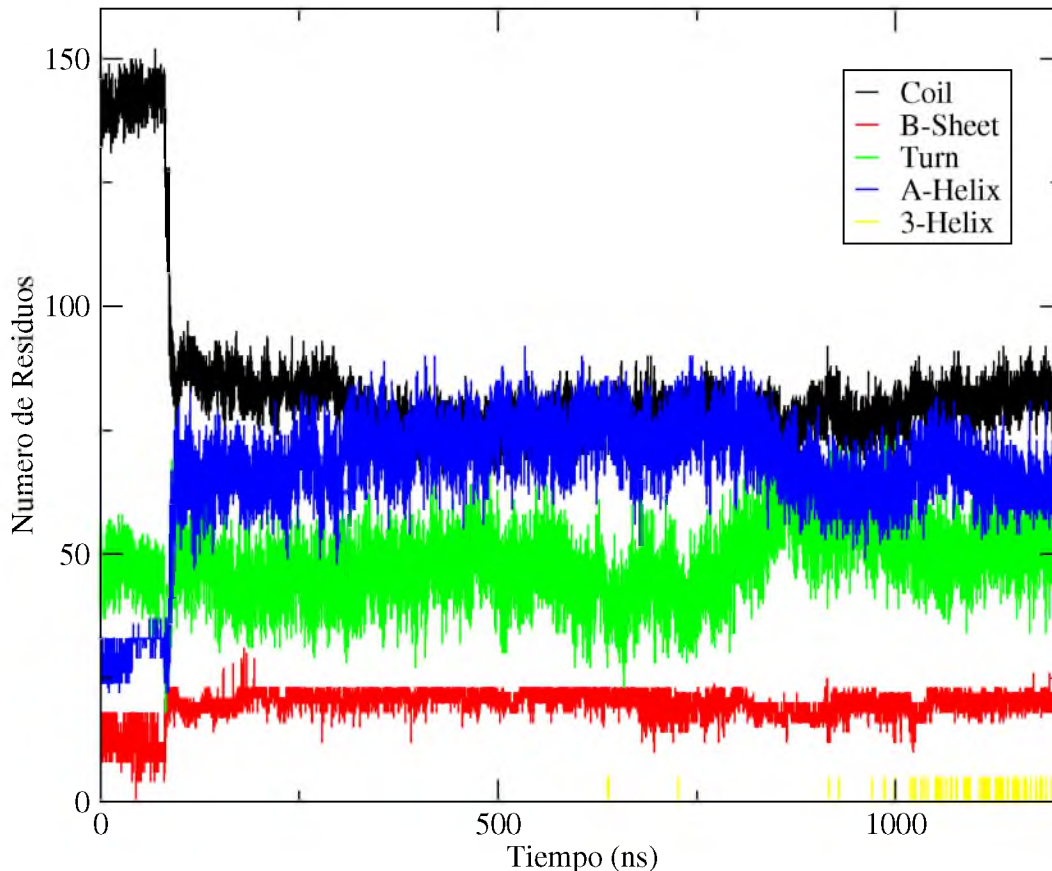


Figura 5.13: Progreso de la definición de estructura secundaria en la simulación

5.8. Validación del modelo de la proteína SUB1A-1

Una vez alcanzados valores estables de RMSD y DSSP para la proteína simulada, se obtuvo la última imagen del sistema y se convirtió a un archivo de estructura legible y se minimizó por el algoritmo de descenso rápido por 500 ps para eliminar efectos estocásticos remanentes de la simulación dependiente del tiempo, resultando en la imagen final de la proteína (Figura 5.14b). El modelo se analizó con la herramienta `d_ssp` del paquete de software GROMACS concluyendo que la porción desordenada se redujo del 73% hasta el 43%.

Para confirmar la calidad estructural, el modelo se trató bajo los parámetros de calidad a los que típicamente se someten las estructuras obtenidas experimentalmente por medio de RMN o difracción de rayos X (DRX) del servidor de validación del PDB. Se corroboró que este modelo cumple con los parámetros propios de los archivos de coordenadas obtenidos por métodos experimentales y puede ser usado en estudios de acoplamiento molecular.

A partir del refinamiento de la estructura por MD se pueden enunciar que, a pesar de haber logrado obtener una imagen posible del estado químico de la proteína SUB1A-1, el resultado arrojó una molécula significativamente más organizada que al inicio pero no definida en su totalidad, por lo que se requiere de mayor tiempo de análisis para obtener una estructura plegada al 80% (es decir con < 20% de porción desordenada).

Sin embargo, y considerando que el análisis del diccionario de estructuras secundarias muestra que la estructura alcanza el equilibrio de conformación luego de 1 μ s de simulación por MD, se puede pensar que para lograr valores de porción ordenada mayores al valor aquí reportado de 57% se requerirá de tiempos de procesamiento significativamente mayores.

Por otra parte, aun cuando el tiempo de simulación total con la metodología propuesta en este trabajo es insuficiente para obtener un modelo plegado en mas del 80% de su secuencia, sí supone una mejora con respecto al uso individual de otros (DM, AM y GSA) para la obtención de un sistema proteico. Por lo que este modelo se encuentra dentro de un resultado aceptable y útil (Karplus y McCammon, 2002).

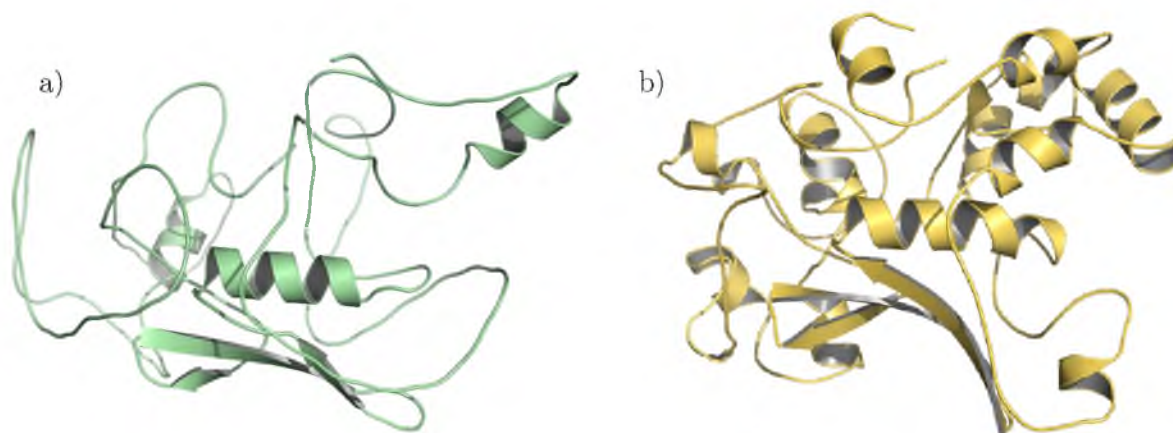


Figura 5.14: Vista de la proteína SUB1A-1 antes del refinamiento estructural (verde, izquierda) y modelo al final del tratamiento computacional (amarillo, derecha).

5.8.1. Análisis de ángulos diedros

En un primer análisis del modelo obtenido, se monitorearon los ángulos diedros de los enlaces peptídicos de la proteína final. Para ello, se obtuvo el mapa de Ramachandran para el sistema mediante la herramienta `g_rama` del paquete de software GROMACS y se verificó que no existieran puntos en las zonas no admitidas. Como se puede apreciar en la [Figura 5.15](#), la proteína se encontró dentro de los límites permitidos, corroborando que el modelo obtenido posee las

características estructurales de una proteína natural.

A partir de la última imagen del sistema, se generó un archivo de coordenadas en formato PDB que se procesó y sincronizó con el paquete *pdbeditor* (Lee y Kim, 2009) 2009) para que cumpliera con los formatos oficiales del Protein Data Bank. El modelo se encuentra disponible por medio de petición al correo electrónico julianpc@unpa.edu.mx.

El modelo final de SUB1A-1 que se obtuvo, tiene mejores parámetros computacionales que el propuesto por Prajapati y cols. (2013) al considerar la realidad biológica del ERF SUB1A-1, y complementar los análisis realizados con el C Score, verificación por PDB y análisis de ángulos diedros presentados en forma de mapa de Ramachandran.

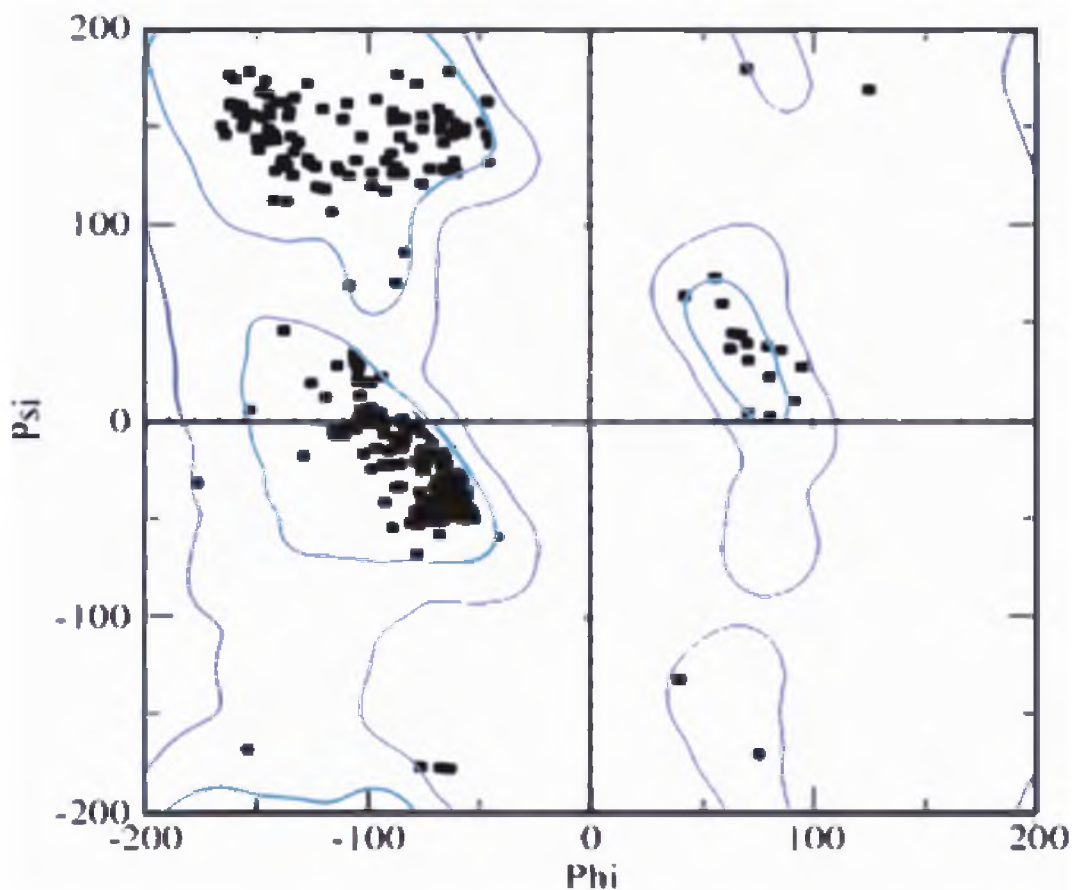


Figura 5.15: Mapa de Ramachandran para el modelo final de la proteína *SUB1A-1*. En cuadros sólidos en negro se representan los puntos obtenidos por medio del algoritmo *g_rama*; en trazo azul, zonas de alto valor estructural; en trazo púrpura, zonas de valor estructural medio-alto.

5.9. Acoplamiento del modelo de proteína con un segmento de ADN afin

El modelo de la proteína SUB1A se acopló a la secuencia de ADN TAGCCGCCAGCGCTGGC-GGCTA (PDBID: 1GCCxB:C) por medio de los paquetes de software Vina y CLC DrugDiscovery. Sin embargo, el tamaño del sistema superó las capacidades de ambos paquetes. Así, se optó por el uso del servidor especializado NPdock (<http://http://genesilico.pl/NPDock/>, Tuszynska y cols., 2015) que se alimentó de forma separada con los 2 archivos, el de SUB1A y el de la secuencia de DNA. Con ellos, se realizó una corrida sin restricciones de 5000 pasos, refinados por el método estadístico Monte Carlo y arrojó un PDB con el modelo ilustrado en la [Figura 5.16](#). Se observa cómo el dominio AP2 es la zona que interactúa con el DNA, tal y como se reportó previamente en la literatura a partir de estructuras determinadas experimentalmente ([Read y cols., 2016](#)).

Dada la alta homología de los dominios AP2 de los ERFs se puede hipotetizar que el resto de los genes *SUB1* son también elementos funcionales, al menos a nivel de este dominio.

5.10. Localización estructural de dominios

Una vez obtenida la imagen tridimensional del sistema, se usó el paquete PyMOL para la visualización de los dominios encontrados con el paquete bioinformático MEME ([Tabla 5.2](#)). Una vista general de la posición espacial de estos dominios se muestra en la [Figura 5.17](#).

Tabla 5.2: Dominios encontrados en la secuencia de SUB1A con ayuda del paquete MEME

Dominio	Secuencia	Residuos	Función putativa
N-terminal	MCGG(A/I)L	1:7	Sensor de oxígeno
CMVII-5 (LWS)	LWS(I/L/Y)DE*N	58:66	Promoción de transcripción
CMVII-3	[E/A]****[R/K]**[K/S]	38:43	localización nuclear
AP2	Apéndice A.1.1	121:179	unión a DNA
CMVII-6	E D*R**D*GDG**E*D	15:26	inhibición de transcripción
CMVII-8	Y****AYD*L***F	227:239	desconocido

5.10.1. Posición del dominio AP2

La proteína SUB1A-1 se analizó por visualización con el paquete de software PyMOL (Schrodinger inc, [DeLano, 2002](#)). Con el paquete MEME Suite se ubicaron en ella las posiciones de los

6 dominios presentes en la proteína SUB1A-1 característicos de los ERFs (Tabla 5.2).

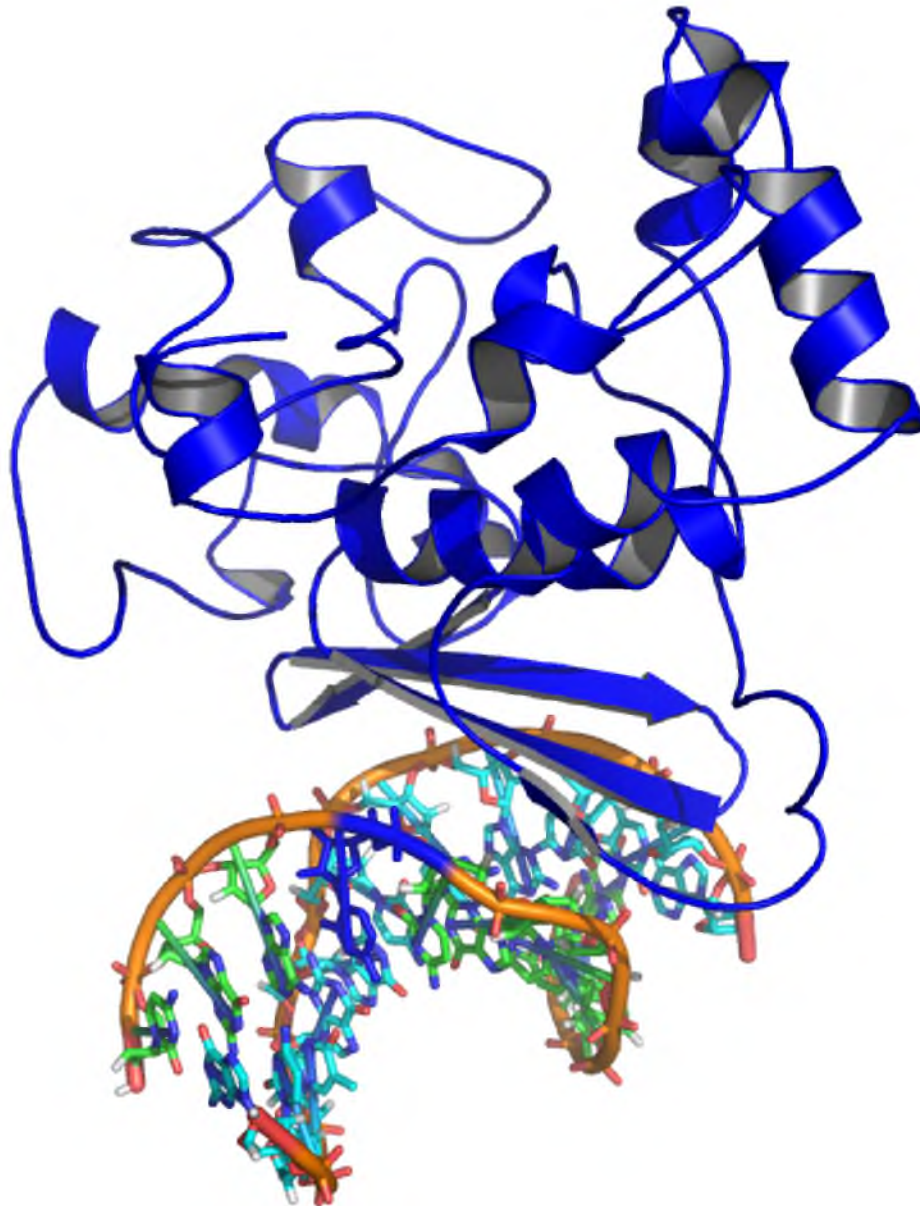


Figura 5.16: Acoplamiento del modelo de SUB1A-1 obtenido y una secuencia de ADN afín

5.10.2. Posición del dominio N-terminal

Después de analizar el plegado de la proteína obtenido por MD, se puede proponer que el dominio N-terminal de la proteína SUB1A-1 se encuentra proyectado desde el centro de la estructura

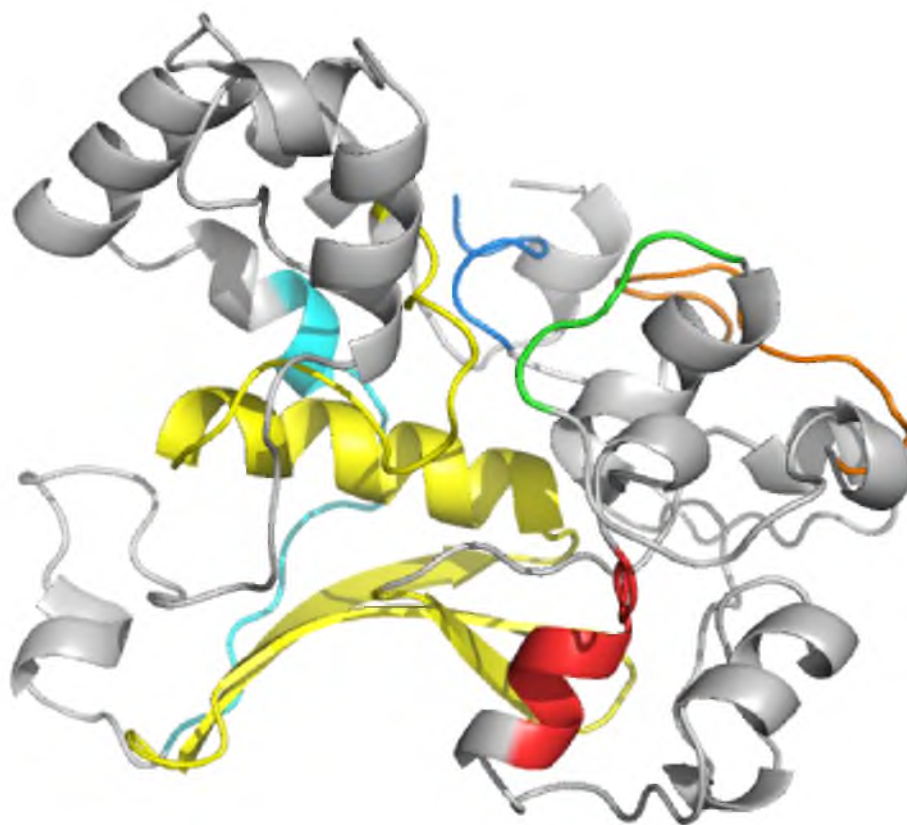


Figura 5.17: Posición de los dominios en el modelo de SUB1A-1. En amarillo, el dominio AP2; en azul el dominio N-terminal; en rojo, el dominio LWS; en verde, el dominio CMVII-3; en naranja el dominio CMVII-6; en cian el dominio CMVII-8.

entre dos grupos voluminosos formados por la estructura que adoptan las regiones adyacentes al dominio AP2 formadas por aproximadamente 30 aminoácidos [Figura 5.18](#). En base a la flexibilidad aparente de estas regiones, se puede proponer que este factor de transcripción sufre modificaciones estructurales una vez que la estructura del complejo SUB1A-ADN se estabiliza para distribuir la energía potencial de unión entre las cadenas laterales, lo cual, podría suponer un sistema de regulación estereoespacial que expondría o protegería a este dominio de la oxidación N-terminal. Si el dominio de oxidación requiere de este sistema de activación/desactivación para promover la transcripción de los GCH es de suponerse que cualquier adición espacial como la de un marcador inmunológico tipo FLAG ([Figura 5.19](#)) propiciaría el bloqueo de este sistema, impidiendo con ello la regulación de la proteína SUB1A-1 y por ende, explicaría la mayor intensidad de los fenotipos cuando ese dominio está protegido. La falta de tolerancia a la inundación que se observa en los individuos transformados con estos cassettes, en especial aquellos de las líneas 35S-F-SUB1A, son el reflejo de la carencia de este sistema de regulación.

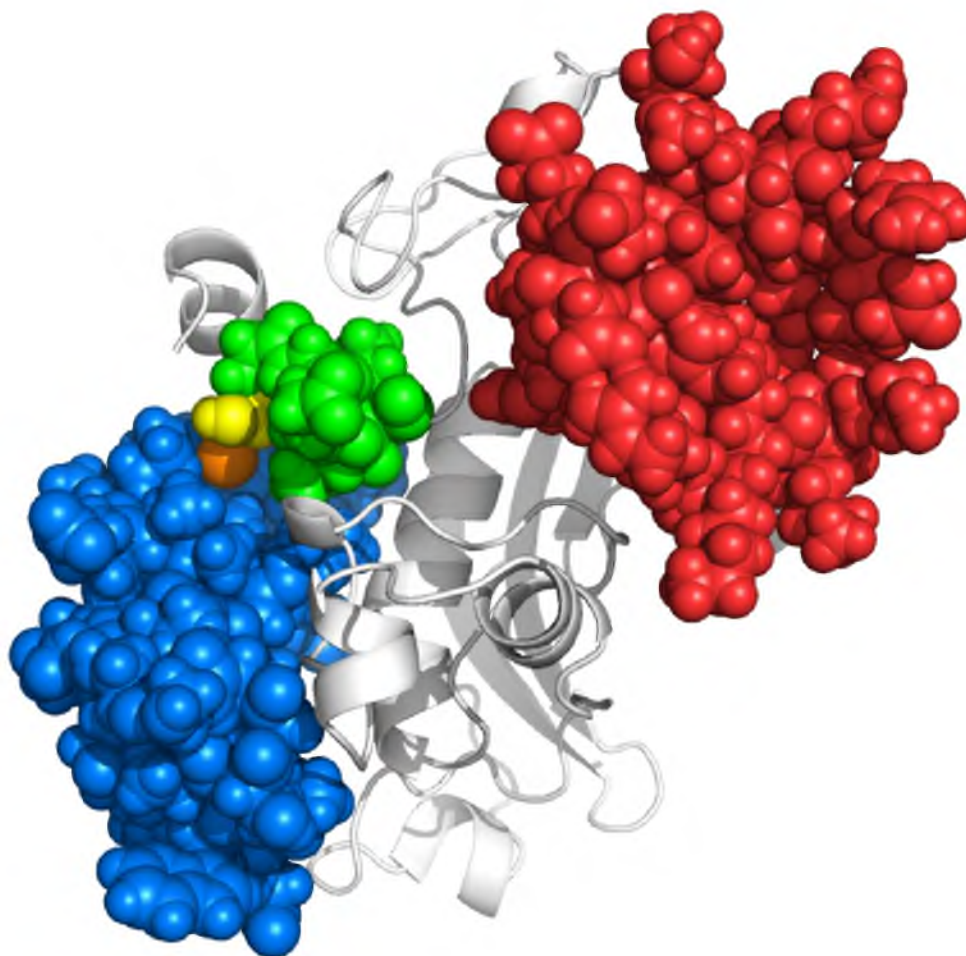


Figura 5.18: Posición del N-terminal en Modelo final de la proteína SUB1A-1 sin la metionina-1; en rojo y azul los dominios abultados derecho (residuos 26:61) e izquierdo (residuos 234:272); en verde el dominio N-terminal (residuos 2:7); en amarillo, la cisteína 2, en naranja el grupo sulfidrido.

Los factores de transcripción requieren de estos mecanismos estereoespaciales de control para poder interactuar con las proteínas que regulan sus actividades. Aunque para SUB1A-1 se han determinado varias proteínas de identidad desconocida como interactuantes, aún no se conoce su importancia biológica (Seo *y cols.*, 2011). Para los ERFs de *Arabidopsis*, si se han demostrado varias interacciones proteína-proteína relevantes como el caso de las PCO, ACBP1 y de HYPOXIA RELATED ELEMENT 1 (HRE1), todos responsables de bucles autoregulatorios (Gibbs *y cols.*, 2014; Weits *y cols.*, 2014; Lee *y cols.*, 2011). Incluso, Bui *y cols.* (2015) determinaron que en la secuencia de los ERFs RAP2.12 y RAP2.2 existen dominios que realizan funciones contradictorias de promoción y represión de la expresión [Figura 1.5](#).

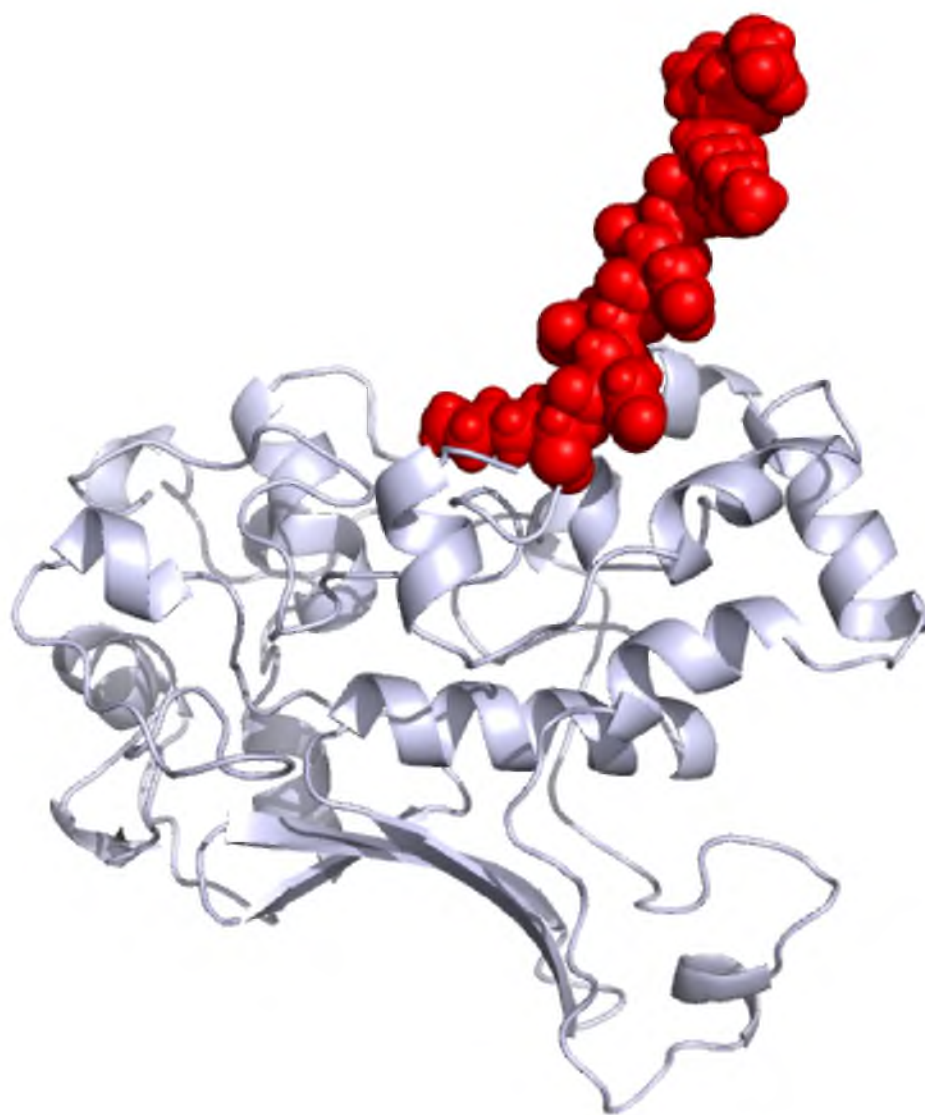


Figura 5.19: Modelo final de la proteína SUB1A-1 (plateado) unida al marcador FLAG (rojo)

6. CONCLUSIONES Y PERSPECTIVAS

6.1. Conclusiones

Mediante el estudio de líneas transgénicas de *Arabidopsis thaliana* transformadas con cassettes de expresión que poseen al dominio N-terminal MCGGIL con o sin protección de una secuencia FLAG, se pudo demostrar que este dominio es funcional. Al estar protegido, se observaron fenotipos de floración tardía, hojas redondeadas, senescencia temprana y acumulación de almidón más intensos que cuando se encontró expuesto y susceptible a regulación, probablemente por el mecanismo de degradación proteosómica de la regla N-terminal. Estas observaciones confirmaron la hipótesis inicial del trabajo.

No obstante lo anterior, no se pudo obtener un fenotipo de tolerancia aumentada al estrés por inundación con ninguna combinación de gen *SUB1*, promotor inducible por inundación o constitutivo, ni funcionalidad del dominio N-terminal. Desde el punto de vista biotecnológico, sí se logró recuperar la fertilidad y el crecimiento normal de las líneas transgénicas mediante la exposición del dominio N-terminal al entorno.

Con el uso de herramientas computacionales, se logró obtener un modelo teórico de la proteína SUB1A-1 compatible con los parámetros característicos de estructuras proteicas determinadas experimentalmente. El análisis de este modelo permitió identificar en la estructura de SUB1A-1 a varios dominios de los ERFs de la subfamilia VII. De especial importancia, el dominio N-terminal se encontró expuesto en la superficie de la proteína entre dos bloques de aminoácidos compuestos por bucles y α -hélices de aparente movilidad elevada que sugieren la presencia de un mecanismo de regulación estereoespacial. Proponemos que involucrará a las proteínas del mecanismo proteolítico N-terminal, como ATE, PRT, PCO y el proteosoma.

Los resultados de los experimentos de fisiología vegetal y química computacional aquí presentados, pueden ser utilizados en el futuro para diseñar mejores estrategias biotecnológicas con el fin de obtener plantas con mayor tolerancia a la inundación, mayor acumulación de almidón y menores efectos indeseados en el desarrollo vegetal.

6.2. Perspectivas

Con la información aquí presentada, se propone que los futuros estudios biotecnológicos que pretendan transferir la característica de tolerancia a la inundación conferida por SUB1A-1 que se ha observado en el arroz, deben realizarse en un fondo genético diferente a Col-0, en especial, alguno menos tolerante de forma natural como un paso previo para su implementación en cultivos comerciales.

Estos estudios, también deberán ampliar la cantidad de promotores probados para modificar la etapa de desarrollo, la intensidad y la localización de la expresión del gen *SUB1A-1* para lograr plantas con características biotecnológicas de interés. Por ejemplo, para el caso de la inundación se pueden probar promotores sensibles a menor intensidad luminosa capaces de detectar el advenimiento de una tormenta o promotores activados por químicos. Para el caso de la acumulación de almidón, promotores que respondan a etapas del desarrollo no juveniles o específicos de hojas maduras.

La demostración experimental y teórica de que la exposición del dominio N-terminal de SUB1A-1 incide en el fenotipo de la planta, muy probablemente por la estabilidad de la proteína, indica que el cambio de promotores será efectivo sólo si siempre se utiliza una versión no susceptible a regulación proteolítica. Esto puede ser por medio de bloqueo con otros amino ácidos (como el aquí probado), mutagénesis sitio-específico de la cisteína 2, o la remoción total del dominio.

El tren de análisis biocomputacional propuesto en esta tesis, se puede utilizar para diseñar simulaciones complejas de interacción de SUB1A-1 (u otros ERFs) con otras moléculas, como el oxígeno, reguladores proteicos y secuencias de DNA. También se puede usar para proponer estructuras de proteínas, que al igual que SUB1A, sólo se conoce una porción de su estructura determinada experimentalmente.

A. ANEXOS

A.1. Secuencias de SUB1A

A.1.1. Nucleótidos

```
ATGTGTGGAGGAGAAGTGATCCCCGCCGACATGCCGGCGGGCCGTTACGCCACGCCACGGCGACGGC
GAGACATGGGTTGACAGAAAGAGGAGGAACAAGAAGAAGAGGAAGCGCGGCCGACGAAGAATGGGAG
GCCGCCTTCCAGGAGTTCATGGCTGCTGACGACGACGACGACGGCGGGCGGACTCGTGTTAAGTAGTAAA
TCTTTGGTGTGAGGTACCAGGTGAAAATGATGCAGGCCGGGGCGCCGCCACCATGTCCATGCCG
CTGGACCCCGTGACCGAGGAGGCCGAGCCGGCGGTGGCTGAGAAGCCTCGCCGGCGCCGGCCGAGGCGG
AGCTACGAGTACCACGGCATCCGGCAGCGGCCGTGGGGCGGTGGTCTCGGAGATCCGCGACCCCGTC
AAGGGCGTCCGCCTCTGGCTCGGCACCTTCGACACCGCCGTCGAAGCCGCGCTCGCCTACGACGCCGAG
GCCCCCGCATCCACGGCTGAAAGCCCGGACAACTTCCCACCCGCGATCTTTCTTCGCCGCCGCCG
CCGTCGCAGCCGCTCTGCTTCTTGCTCAACGACAACGGCCTCATACAATCGGAGAAGCGCCGACCGAC
GACGCCCGCTCGACGTCGACGTCGACGACGGAGGCGTCCGGCGACGGCGCATACAACCTGGAGTGCTGC
TCGGACGACGTGATGGACAGCCTCCTCGCCGGCTACGACGTGGCCAGCGGCGACGACATATGGACATGG
ACATCTGGAGCCTCCTCCACCTCTGTTAACCAAGAGATCAAGACCCCATCGATCCACAAAAACATATCA
TATGCAGGGGAAGCCTGA
```

A.1.2. Aminoácidos

La región en color azul corresponde al dominio AP2

```
MCGGEVIPAD MPAAPFTPRH GDGETWVDRK RRNKKRKRK ADEEWEAAFQ EFMAADDDDD
GGGLVLSSKS LVLRSPEGND AGRGAAATMS MPLDPVTEEA EPAVAEPKRR RRPRESYEYH
GIRQRPWGRW SSEIRDPVKG VRLWLGTFDT AVEAALAYDA EARRIHGWKA RTNFPPADLS
SPPPPSQPLC FLLNDNGLIT IGEAPTDDAA STSTSTTEAS GDARIQLECC SDDVMDSLLA
GYDVASGDDI WTWTSGASST SVNQEIKTPS IHQNISYAGE A
```

A.2. Configuraciones de ciclos para PCR

Tabla A.1: PCR con GoTaq Green

Paso	Temperatura	Duración	Repeticiones
Desnaturalización inicial	95	3 min	1
Desnaturalización	95	30s	
Alineamiento	55	30s	35
Extensión	72	1 min	
Elongación final	72	2 min	1

Marcador de peso molecular

Se utilizó el marcador de peso molecular *GeneRuler ADN Ladder Mix* (Thermo Scientific # SM0331), con la proporción recomendada por el fabricante de 1 μ l de Ladder de ADN con 1 μ l de buffer de carga 6X *Loading Dye Solution* y 4 μ l de ddH₂O para un volumen de carga de 6 μ l.

A.3. Primers utilizados para PCR

Los primers utilizados se diseñaron en el laboratorio de Biotecnología vegetal de la Universidad del Papaloapan y se sintetizaron por la empresa Macrogen

Set de primers para la amplificación del gen de *TUBULIN2* de *Arabidopsis thaliana* (TUB2)

Se diseñaron con seguro contra contaminación por ADN genómico al amplificar una región de 667 pb cuando existen intrones y 166 pb de la secuencia madura ((Wenkel y cols., 2006)).

Primer	Secuencia	T _m (C)	%GC
Directo	5'-GAGCCTTACAACGCTACTCTGTCTGTC-3'	69.8	51.8
Reverso	5'-CACCAGACATAGTAGCAGAAATCAAG-3'	64.7	42.3

Set de primers para la amplificación de *SUB1A*

Los primers utilizados para la amplificación de los genes *SUB1A* y *SUB1C* se obtuvieron de acuerdo a lo reportado por Fukao *y cols.* (2009). El tamaño del amplicón se calculó en 101 pb con ayuda del algoritmo *PCR test* de la Sequence Manipulation Suite (Stothard, 2000), disponible en: <http://www.bioinformatics.org/sms2/index.html>.

Primer	Secuencia	T_m (°C)	%GC
Directo	5'-GACAAACTTCCCACCCGC-3'	53	61.11
Reverso	5'-GCGCTTCTCCGATTGTGATG-3'	54	55.00

Set de primers para la amplificación de *SUB1C*

El tamaño del amplicón se calculó en 90 pb con ayuda del algoritmo *PCR test* de Sequence Manipulation Suite (Stothard, 2000)

Primer	Secuencia	T_m (°C)	%GC
Directo	5'-ATACTCATCGAGTGCTGCTCCGAC-3'	59	54.17
Reverso	5'-TTAGCTCCAGAAGCGCATGTC-3'	52	50.00

A.4. Composición de las soluciones utilizadas

A.4.1. Formaldehído - Buffer de fosfatos pH 6.6

Se siguió el protocolo de la Universidad de California en Berkeley. Se prepararon 100 ml de soluciones stock como sigue:

Solución 1: Stock 1M de Fosfato de potasio monobásico (KH_2PO_4)

Se añadieron 13.06 g de fosfato de potasio monobásico (Sigma Aldrich) a 100 ml de deH_2O , se mezcló vigorosamente con ayuda de un vortex hasta lograr la disolución completa.

Solución 2: Stock 1M de Fosfato de potasio dibásico (K_2HPO_4)

Se añadieron 17.42 g de fosfato de potasio dibásico (J.T. Baker) a 100 ml de deH_2O , se mezcló vigorosamente con ayuda de un vortex hasta lograr la disolución completa.

Posteriormente, se mezclaron 7.62 ml de la solución 1 con 12.38 ml de la solución 2, esta mezcla se añadió a 180 ml de deH₂O. Finalmente, se añadieron a la mezcla 7.612 ml de formaldehído el solución al 37%.

A.4.2. Solución de yodo Lugol

En un volumen de 300 ml se añadieron:

- 1 g de Yodo sublimado (I₂, Química Meyer) en perlas previamente maceradas en mortero
- 2 g de Youro de potasio granulado (KI; J. T. Baker)

La mezcla final (KI 2% w/v, I₂ 1% w/v) se agitó por 30 min con ayuda de un agitador magnético en un contenedor cubierto con papel aluminio.

A.4.3. Gel de agarosa al 1.6% w/v

Se añadieron 200 ml de buffer TAE 1X a 3.2 g de agarosa (Ibi Scientific, IB70040) y se fundieron completamente con ayuda de un horno de microondas (Daewoo KR 16) en intervalos de 30 s de operación. Ya fundida, la solución fue vertida en los moldes apropiados previamente nivelados, se dejó reposar durante 20 min para su solidificación.

A.4.4. Gel de agarosa al 2% w/v

Se añadieron 200 ml de buffer TAE 1X a 4 g de agarosa (Ibi Scientific, IB70040)) y se fundieron completamente con ayuda de un horno de microondas (Daewoo KR 16) en intervalos de 30 s de operación. Ya fundida, la solución fue vertida en los moldes apropiados previamente nivelados, se dejó reposar durante 20 min para su solidificación.

A.4.5. EtBr/TAE 1X

A 20 ml de solución TAE 50X se le añadieron a 980 ml de dH₂Oy se mezclaron por agitación manual. Posteriormente se añadieron 40 μl de EtBr de un stock de 10 mg/ml w/v y se cubrió de la luz con envoltorio de papel aluminio.

A.5. Algoritmo de regresión lineal en Fortran 2008 para calcular el LT_{50}

```

PROGRAM regresionlineal
  IMPLICIT NONE
  REAL :: m, a, b, r, sum_x, sum_y, x_prom, y_prom, sum_xy, sum_xc, sum_yc
  REAL :: ra, rb, rc, rd, p3, p2, p1
  INTEGER, PARAMETER :: n_max=100
  INTEGER :: n, i, st1, st2, st3, st4, st5, st6, st7
  REAL, ALLOCATABLE, DIMENSION(:)::x,s,sc,scc, y, yr, xr
  sum_xy=0;
  sum_xc=0;
  sum_yc=0;
  sum_x=0
  write (*,*) 'Leer cantidad de puntos a analizar'
  read(*,*) n
  DO i=1,n
    write(*,*) 'Leer coordenadas (x,y) del punto', i
    read(*,*) x(i), y(i)
    sum_x=sum_x+x(i)
    sum_y=sum_y+y(i)
  END DO
  x_prom=sum_x/n
  y_prom=sum_y/n
  DO i=1,n
    s(i)=x(i)*y(i)
    sum_xy=sum_xy+s(i)
    sc(i)=x(i)**2
    sum_xc=sum_xc+sc(i)
    scc(i)=y(i)**2
    sum_yc=sum_yc+scc(i)
  END DO
  m=(n*sum_xy-sum_x*sum_y)/(n*sum_xc-(sum_x**2));

```

```

b=y_prom-m*x_prom;
ra=((n*sum_xy)-(sum_x*sum_y));
rb=((n*sum_xc)-(sum_x**2))**0.5;
rc=((n*sum_yc)-(sum_y**2))**0.5;
rd=rb*rc;
r=(ra/rd)**2;
write(*,*) 'la_pendiente_es_', m
write(*,*) 'la_ordenada_al_origen_es:', b
write(*,*) 'el_valor_de_r^2_de_la_regresion_es:', r, '%'
END PROGRAM regresionlineal

```

A.6. Configuraciones de las diferentes etapas de simulación

Tabla A.2: Valores utilizados en la configuración de la minimización de energía

Campo	Valor	Unidades	Comentario
Valor máximo de energía	1000	<i>KJ/mol/nm</i>	
Valor de paso de iteración	0.01	<i>KJ/mol/nm</i>	
Número de pasos	50000		valor máximo
Frecuencia de lista de vecinos	10	pasos	
Método de lista de vecinos	grid		
Radio de corte	1.2	nm	
Radio de corte coulombico	1.2	nm	
Radio de Van der Waals	1.2	nm	

Tabla A.3: Valores utilizados en la configuración de la simulación NVT

Campo	Valor	Unidades	Comentario
Integrador	Leapfrog		
Valor de paso de tiempo	2	fs	
Frecuencia de lista de vecinos	10	pasos	
Radio de corte	1.4	nm	
Radio de Van der Waals	1.4	nm	
Radio de Coulomb	1.4	nm	
Temperatura	300	K	

Tabla A.4: Valores utilizados en la configuración de la simulación NPT

Campo	Valor	Unidades	Comentario
Integrador	Leapfrog		
Valor de paso de tiempo	2	fs	
Frecuencia de lista de vecinos	10	pasos	
Radio de corte	1.4	nm	
Radio de Van der Waals	1.4	nm	
Radio de Coulomb	1.4	nm	
Temperatura	300	K	
Presión	1	bar	
Compresibilidad	$4.5 \cdot 10^{-5}$	bar^{-1}	

Tabla A.5: Parámetros utilizados en la etapa GSA

Campo	Valor	Unidades
Integrador	Leapfrog	
Valor de paso de tiempo	2	fs
Frecuencia de lista de vecinos	10	pasos
Radio de corte	1.4	nm
Radio de Van der Waals	1.4	nm
Radio de Coulomb	1.4	nm
Presión	1	bar
Temperaturas (incremento)	300 330 360 390 420 450 480 510	
(Reducción)	480 450 420 390 360 330 300	
(Estabilización)	300 300 300 300 300 300	

REFERENCIAS CITADAS

- Agarwal, S., Kapoor, A., Lakshmi, O. S., y Grover, A. (2007). Production and phenotypic analysis of rice transgenics with altered levels of pyruvate decarboxylase and alcohol dehydrogenase proteins. *Plant Physiology and Biochemistry*, 45(9):637–46. DOI: [10.1016/j.plaphy.2007.07.008](https://doi.org/10.1016/j.plaphy.2007.07.008).
- Agostini, F. P., Soares-Pinto, D. D. O., Moret, M. A., Osthoff, C., y Pascutti, P. G. (2006). Generalized simulated annealing applied to protein folding studies. *Journal of Computational Chemistry*, 27(11):1142–1155. DOI: [10.1002/jcc.20428](https://doi.org/10.1002/jcc.20428).
- Allen, M. D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., y Suzuki, M. (1998). A novel mode of DNA recognition by a β -sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *The EMBO Journal*, 17(18):5484–5496. DOI: [10.1093/emboj/17.18.5484](https://doi.org/10.1093/emboj/17.18.5484).
- Bachmair, A., Finley, D., y Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science*, 234(4773):179–186. DOI: [10.1126/science.3018930](https://doi.org/10.1126/science.3018930).
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., y Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. page gkp335. DOI: [10.1093/nar/gkp335](https://doi.org/10.1093/nar/gkp335).
- Bailey-Serres, J. y Colmer, T. (2014). Plant tolerance of flooding stress - recent advances. *Plant, Cell & environment*, 37(10):2211–2215. DOI: [10.1111/pce.12420](https://doi.org/10.1111/pce.12420).
- Bailey-Serres, J., Fukao, T., Gibbs, D. J., Holdsworth, M. J., Lee, S. C., Licausi, F., Perata, P., Voesenek, L. A. C. J., y van Dongen, J. T. (2012a). Making sense of low oxygen sensing. *Trends in Plant Science*, 17(3):129–38. DOI: [10.1016/j.tplants.2011.12.004](https://doi.org/10.1016/j.tplants.2011.12.004).
- Bailey-Serres, J., Fukao, T., Ronald, P., Ismail, A., Heuer, S., y Mackill, D. (2010). Submergence tolerant rice: SUB1's journey from landrace to modern cultivar. *Rice*, 3(2-3):138–47. DOI: [10.1007/s12284-010-9048-5](https://doi.org/10.1007/s12284-010-9048-5).
- Bailey-Serres, J., Lee, S., y Brinton, E. (2012b). Waterproofing crops: Effective flooding survival strategies. *Plant Physiology*, 160(4):1698–1709. DOI: [10.1104/pp.112.208173](https://doi.org/10.1104/pp.112.208173).
- Bailey-Serres, J. y Voesenek, L. (2008). Flooding stress: acclimations and genetic diversity. *Annu. Rev. Plant Biol.*, 59:313–339. DOI: [10.1146/annurev.arplant.59.032607.092752](https://doi.org/10.1146/annurev.arplant.59.032607.092752).
- Bailey-Serres, J. y Voesenek, L. A. C. J. (2010). Life in the balance: a signaling network

- controlling survival of flooding. *Current Opinion in Plant Biology*, 13(5):489–94. DOI: [10.1016/j.pbi.2010.08.002](https://doi.org/10.1016/j.pbi.2010.08.002).
- Banti, V., Giuntoli, B., Gonzali, S., Loreti, E., Magneschi, L., Novi, G., Paparelli, E., Parlanti, S., Pucciariello, C., Santaniello, A., y Perata, P. (2013). Low oxygen response mechanisms in green organisms. *International Journal of Molecular Sciences*, 14(3):4734–61. DOI: [10.3390/ijms14034734](https://doi.org/10.3390/ijms14034734).
- Bekker, H., Berendsen, H. J. C., Dijkstra, E. J., Achterop, S., van Drunen, R., van der Spoel, D., Sijbers, A., Keegstra, H., Reitsma, B., y Renardus, M. K. R. (1993). Gromacs: A parallel computer for molecular dynamics simulations. *Physics Computing*.
- Blokhina, O., Virolainen, E., y Fagerstedt, K. V. (2003). Antioxidants, oxidative damage and oxygen deprivation stress: a review. *Annals of Botany*, 91(2):179–194. DOI: [10.1093/aob/mcf118](https://doi.org/10.1093/aob/mcf118).
- Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., y Görlach, J. (2001). Growth stage–based phenotypic analysis of Arabidopsis a model for high throughput functional genomics in plants. *The Plant Cell*, 13(7):1499–1510. DOI: [10.1105/TPC.010011](https://doi.org/10.1105/TPC.010011).
- Bradshaw, R. A., Brickey, W. W., y Walker, K. W. (1998). N-terminal processing: the methionine aminopeptidase and N α -acetyl transferase families. *Trends in Biochemical Sciences*, 23(7):263–267. DOI: [10.1016/S0968-0004\(98\)01227-4](https://doi.org/10.1016/S0968-0004(98)01227-4).
- Branco-Price, C., Kaiser, K. A., Jang, C. J., Larive, C. K., y Bailey-Serres, J. (2008). Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in Arabidopsis thaliana. *The Plant Journal*, 56(5):743–755. DOI: [10.1111/j.1365-313X.2008.03642.x](https://doi.org/10.1111/j.1365-313X.2008.03642.x).
- Bui, L. T., Giuntoli, B., Kosmacz, M., Parlanti, S., y Licausi, F. (2015). Constitutively expressed ERF-VII transcription factors redundantly activate the core anaerobic response in Arabidopsis thaliana. *Plant Science*, 236:37–43. DOI: [10.1016/j.plantsci.2015.03.008](https://doi.org/10.1016/j.plantsci.2015.03.008).
- Bussi, G., Donadio, D., y Parrinello, M. (2007). Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101.
- Butaye, K. M., Cammue, B. P., Delauré, S. L., y De Bolle, M. F. (2005). Approaches to minimize variation of transgene expression in plants. *Molecular Breeding*, 16(1):79–91. DOI: [10.1007/s11032-005-4929-9](https://doi.org/10.1007/s11032-005-4929-9).
- Campbell, M. T., Proctor, C. A., Dou, Y., Schmitz, A. J., Phansak, P., Kruger, G. R., Zhang, C., y Walia, H. (2015). Genetic and molecular characterization of submergence response identifies Subtol6 as a major submergence tolerance locus in maize. *PloS One*, 10(3):e0120385. DOI: [10.1371/journal.pone.0120385](https://doi.org/10.1371/journal.pone.0120385).

- Chapra, S. C. y Canale, R. P. (2006). *Numerical methods for engineers*. McGraw-Hill New York, 5th edición.
- Dar, M. H., de Janvry, A., Emerick, K., Raitzer, D., y Sadoulet, E. (2013). Flood-tolerant rice reduces yield variability and raises expected yield, differentially benefitting socially disadvantaged groups. *Scientific Reports*, 3. DOI: [10.1038/srep03315](https://doi.org/10.1038/srep03315).
- Deeb, O., Rosales-Hernández, M. C., Gómez-Castro, C., Garduño-Juárez, R., y Correa-Basurto, J. (2010). Exploration of human serum albumin binding sites by docking and molecular dynamics flexible ligand–protein interactions. *Biopolymers*, 93(2):161–170. DOI: [10.1002/bip.21314](https://doi.org/10.1002/bip.21314).
- DeLano, W. L. (2002). The PyMOL molecular graphics system.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5):1792–1797. DOI: [10.1093/nar/gkh340](https://doi.org/10.1093/nar/gkh340).
- Fukao, T. y Bailey-Serres, J. (2008). Submergence tolerance conferred by Sub1A is mediated by SLR1 and SLRL1 restriction of gibberellin responses in rice. *Proceedings of the National Academy of Sciences*, 105(43):16814–16819. DOI: [10.1073/pnas.0807821105](https://doi.org/10.1073/pnas.0807821105).
- Fukao, T., Harris, T., y Bailey-Serres, J. (2009). Evolutionary analysis of the Sub1 gene cluster that confers submergence tolerance to domesticated rice. *Annals of Botany*, 103(2):143–150. DOI: [10.1093/aob/mcn172](https://doi.org/10.1093/aob/mcn172).
- Fukao, T., Xu, K., Ronald, P. C., y Bailey-Serres, J. (2006). A variable cluster of ethylene response factor–like genes regulates metabolic and developmental acclimation responses to submergence in rice. *The Plant Cell*, 18(8):2021–2034. DOI: [10.1105/tpc.106.043000](https://doi.org/10.1105/tpc.106.043000).
- Fukao, T., Yeung, E., y Bailey-Serres, J. (2011). The submergence tolerance regulator SUB1A mediates crosstalk between submergence and drought tolerance in rice. *The Plant Cell*, 23(1):412–427. DOI: [10.1105/tpc.110.080325](https://doi.org/10.1105/tpc.110.080325).
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., y Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research*, 31(13):3784–3788. DOI: [10.1093/nar/gkg563](https://doi.org/10.1093/nar/gkg563).
- Gibbs, D. J., Lee, S. C., Isa, N. M., Gramuglia, S., Fukao, T., Bassel, G. W., Correia, C. S., Corbinau, F., Theodoulou, F. L., Bailey-Serres, J., y cols. (2011). Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. *Nature*, 479(7373):415–418. DOI: [10.1038/nature10534](https://doi.org/10.1038/nature10534).
- Gibbs, D. J., Md Isa, N., Movahedi, M., Lozano-Juste, J., Mendiando, G. M., Berckhan, S., Marín-de la Rosa, N., Vicente Conde, J., Sousa Correia, C., Pearce, S. P., Bassel, G. W., Hamali, B., Talloji, P., Tomé, D. F. A., Coego, A., Beynon, J., Alabadí, D., Bachmair, A., León, J., Gray, J. E., Theodoulou, F. L., y Holdsworth, M. J. (2014). Nitric oxide sensing in plants is mediated by

- proteolytic control of group VII ERF transcription factors. *Molecular Cell*, 53(3):369–79. DOI: [10.1016/j.molcel.2013.12.020](https://doi.org/10.1016/j.molcel.2013.12.020).
- Graciet, E., Mesiti, F., y Wellmer, F. (2010). Structure and evolutionary conservation of the plant N-end rule pathway. *The Plant Journal*, 61(5):741–751. DOI: [10.1111/j.1365-313X.2009.04099.x](https://doi.org/10.1111/j.1365-313X.2009.04099.x).
- Graciet, E., Walter, F., Maoiléidigh, D. Ó., Pollmann, S., Meyerowitz, E. M., Varshavsky, A., y Wellmer, F. (2009). The N-end rule pathway controls multiple functions during Arabidopsis shoot and leaf development. *Proceedings of the National Academy of Sciences*, 106(32):13618–13623. DOI: [10.1073/pnas.0906404106](https://doi.org/10.1073/pnas.0906404106).
- Gupta, K. J., Zabalza, A., y Van Dongen, J. T. (2009). Regulation of respiration when the oxygen availability changes. *Physiologia Plantarum*, 137(4):383–391. DOI: [10.1111/j.1399-3054.2009.01253.x](https://doi.org/10.1111/j.1399-3054.2009.01253.x).
- Gutterson, N. y Reuber, T. L. (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. *Current Opinion in Plant Biology*, 7(4):465–471. DOI: [doi:10.1016/j.pbi.2004.04.007](https://doi.org/10.1016/j.pbi.2004.04.007).
- Hao, G.-F., Xu, W.-F., Yang, S.-G., y Yang, G.-F. (2015). Multiple Simulated Annealing-Molecular Dynamics (MSA-MD) for Conformational Space Search of Peptide and Miniprotein. *Scientific Reports*, 5. DOI: [10.1038/srep15568](https://doi.org/10.1038/srep15568).
- Hattori, Y., Nagai, K., Furukawa, S., Song, X.-J., Kawano, R., Sakakibara, H., Wu, J., Matsumoto, T., Yoshimura, A., Kitano, H., y cols. (2009). The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. *Nature*, 460(7258):1026–1030. DOI: [10.1038/nature08258](https://doi.org/10.1038/nature08258).
- Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., y Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, 3(9):816–821. DOI: [10.1038/nclimate1911](https://doi.org/10.1038/nclimate1911).
- Hsu, F.-C., Chou, M.-Y., Chou, S.-J., Li, Y.-R., Peng, H.-P., y Shih, M.-C. (2013). Submergence confers immunity mediated by the WRKY22 transcription factor in Arabidopsis. *The Plant Cell*, 25(7):2699–2713. DOI: [10.1105/tpc.113.114447](https://doi.org/10.1105/tpc.113.114447).
- Jackson, M. B. y Ismail, A. M. (2015). Introduction to the Special Issue: Electrons, water and rice fields: plant response and adaptation to flooding and submergence stress. *AoB Plants*, 7(1):plv078. DOI: [10.1093/aobpla/plv078](https://doi.org/10.1093/aobpla/plv078).
- Jofuku, K. D., der Boer, B. G., Van Montagu, M., y Okamoto, J. K. (1994). Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell*, 6(9):1211–25.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., y Klein, M. L. (1983). Compa-

- parison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935. DOI: [10.1063/1.445869](https://doi.org/10.1063/1.445869).
- Jung, K.-H., Seo, Y.-S., Walia, H., Cao, P., Fukao, T., Canlas, P. E., Amonpant, F., Bailey-Serres, J., y Ronald, P. C. (2010). The submergence tolerance regulator Sub1A mediates stress-responsive expression of AP2/ERF transcription factors. *Plant Physiology*, 152(3):1674–1692. DOI: [10.1104/pp.109.152157](https://doi.org/10.1104/pp.109.152157).
- Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4:187217. DOI: [10.1002/jcc.540040211](https://doi.org/10.1002/jcc.540040211).
- Karplus, M. y McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural & Molecular Biology*, 9(9):646–652.
- Kosmacz, M., Parlanti, S., Schwarzländer, M., Kragler, F., Licausi, F., y Van Dongen, J. T. (2015). The stability and nuclear localization of the transcription factor RAP2. 12 are dynamically regulated by oxygen concentration. *Plant, Cell & Environment*, 38(6):1094–1103. DOI: [10.1111/pce.12493](https://doi.org/10.1111/pce.12493).
- Lee, J. y Kim, S.-H. (2009). PDB Editor: a user-friendly Java-based Protein Data Bank file editor with a GUI. *Acta Crystallographica Section D: Biological Crystallography*, 65(4):399–402. DOI: [10.1107/S090744490900451X](https://doi.org/10.1107/S090744490900451X).
- Lee, S. C., Mustroph, A., Sasidharan, R., Vashisht, D., Pedersen, O., Oosumi, T., Voesenek, L. A., y Bailey-Serres, J. (2011). Molecular characterization of the submergence response of the *Arabidopsis thaliana* ecotype Columbia. *New Phytologist*, 190(2):457–471. DOI: [10.1111/j.1469-8137.2010.03590.x](https://doi.org/10.1111/j.1469-8137.2010.03590.x).
- Li, Z.-X., Septiningsih, E., Quilloy-Mercado, S., McNally, K., y Mackill, D. (2011). Identification of SUB1A alleles from wild rice *Oryza rufipogon* Griff. *Genetic Resources and Crop Evolution*, 58(8):1237–1242. DOI: [10.1007/s10722-010-9657-2](https://doi.org/10.1007/s10722-010-9657-2).
- Licausi, F., Kosmacz, M., Weits, D. A., Giuntoli, B., Giorgi, F. M., Voesenek, L. A. C. J., Perata, P., y van Dongen, J. T. (2011a). Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. *Nature*, 479(7373):419–22. DOI: [10.1038/nature10536](https://doi.org/10.1038/nature10536).
- Licausi, F., van Dongen, J. T., Giuntoli, B., Novi, G., Santaniello, A., Geigenberger, P., y Perata, P. (2010). HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana*. *The Plant Journal*, 62(2):302–15. DOI: [10.1111/j.1365-313X.2010.04149.x](https://doi.org/10.1111/j.1365-313X.2010.04149.x).
- Licausi, F., Weits, D. A., Pant, B. D., Scheible, W.-R., Geigenberger, P., y van Dongen, J. T. (2011b). Hypoxia responsive gene expression is mediated by various subsets of transcription factors and

- miRNAs that are determined by the actual oxygen availability. *New Phytologist*, 190(2):442–456. DOI: [10.1111/j.1469-8137.2010.03451.x](https://doi.org/10.1111/j.1469-8137.2010.03451.x).
- Liu, Y., Prigozhin, M. B., Schulten, K., y Gruebele, M. (2014). Observation of complete pressure-jump protein refolding in molecular dynamics simulation and experiment. *Journal of the American Chemical Society*, 136(11):4265–4272. DOI: [10.1021/ja412639u](https://doi.org/10.1021/ja412639u).
- Manson, R., Jardel Peláez, E., Jiménez Espinosa, M., y cols. (2009). Perturbaciones y desastres naturales: impactos sobre las ecorregiones, la biodiversidad y el bienestar socioeconómico, en: Capital natural de México, vol. II: Estado de conservación y tendencias de cambio. *Conabio, Mexico*, 1(1):131–184.
- McDougal, O. M., Cornia, N., Sambasivarao, S., Remm, A., Mallory, C., Oxford, J. T., Maupin, C. M., y Andersen, T. (2014). Homology modeling and molecular docking for the science curriculum. *Biochemistry and Molecular Biology Education*, 42(2):179–182. DOI: [10.1002/bmb.20767](https://doi.org/10.1002/bmb.20767).
- McGrath, K. C., Dombrecht, B., Manners, J. M., Schenk, P. M., Edgar, C. I., Maclean, D. J., Scheible, W.-R., Udvardi, M. K., y Kazan, K. (2005). Repressor-and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. *Plant Physiology*, 139(2):949–959. DOI: [10.1104/pp.105.068544](https://doi.org/10.1104/pp.105.068544).
- Mithran, M., Paparelli, E., Novi, G., Perata, P., y Loreti, E. (2014). Analysis of the role of the pyruvate decarboxylase gene family in Arabidopsis thaliana under low-oxygen conditions. *Plant Biology*, 16(1):28–34. DOI: [10.1111/plb.12005](https://doi.org/10.1111/plb.12005).
- Mizoi, J., Shinozaki, K., y Yamaguchi-Shinozaki, K. (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1819(2):86–96. DOI: [10.1016/j.bbagr.2011.08.004](https://doi.org/10.1016/j.bbagr.2011.08.004).
- Nakano, T., Suzuki, K., Fujimura, T., y Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiology*, 140(2):411–32. DOI: [10.1104/pp.105.073783](https://doi.org/10.1104/pp.105.073783).
- Neeraja, C., Maghirang-Rodriguez, R., Pamplona, A., Heuer, S., Collard, B., Septiningsih, E., Vergara, G., Sanchez, D., Xu, K., Ismail, A., y cols. (2007). A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. *Theoretical and Applied Genetics*, 115(6):767–776. DOI: [10.1007/s00122-007-0607-0](https://doi.org/10.1007/s00122-007-0607-0).
- Niroula, R. K., Pucciariello, C., Ho, V. T., Novi, G., Fukao, T., y Perata, P. (2012). SUB1A-dependent and-independent mechanisms are involved in the flooding tolerance of wild rice species. *The Plant Journal*, 72(2):282–293. DOI: [10.1111/j.1365-313X.2012.05078.x](https://doi.org/10.1111/j.1365-313X.2012.05078.x).
- Núñez-López, L., Aguirre-Cruz, A., Barrera-Figueroa, B. E., y Peña-Castro, J. M. (2015). Impro-

- vement of enzymatic saccharification yield in *Arabidopsis thaliana* by ectopic expression of the rice SUB1A-1 transcription factor. *Peer Journal*, PeerJ 3:e817. DOI: [10.7717/peerj.817](https://doi.org/10.7717/peerj.817).
- Okonechnikov, K., Golosova, O., Fursov, M., y cols. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8):1166–1167. DOI: [10.1093/bioinformatics/bts091](https://doi.org/10.1093/bioinformatics/bts091).
- Papdi, C., Pérez-Salamó, I., Joseph, M. P., Giuntoli, B., Bögre, L., Koncz, C., y Szabados, L. (2015). The low oxygen, oxidative and osmotic stress responses synergistically act through the ethylene response factor VII genes RAP2. 12, RAP2. 2 and RAP2. 3. *The Plant Journal*, 82(5):772–784. DOI: [10.1111/tpj.12848](https://doi.org/10.1111/tpj.12848).
- Parnell, S. C., Marotti, L. A., Kiang, L., Torres, M. P., Borchers, C. H., y Dohlman, H. G. (2005). Phosphorylation of the RGS protein Sst2 by the MAP kinase Fus3 and use of Sst2 as a model to analyze determinants of substrate sequence specificity. *Biochemistry*, 44(22):8159–8166. DOI: [10.1021/bi0503091](https://doi.org/10.1021/bi0503091).
- Peña-Castro, J. M. (2014). Respuesta molecular de las plantas ante el estrés por inundación: lecciones aprendidas del gen SUB1A. *Revista Fitotecnia Mexicana*, 37(4):325–337.
- Peña-Castro, J. M., van Zanten, M., Lee, S. C., Patel, M. R., Voeselek, L. A., Fukao, T., y Bailey-Serres, J. (2011). Expression of rice SUB1A and SUB1C transcription factors in *Arabidopsis* uncovers flowering inhibition as a submergence tolerance mechanism. *The Plant Journal*, 67(3):434–446. DOI: [10.1111/j.1365-313X.2011.04605.x](https://doi.org/10.1111/j.1365-313X.2011.04605.x).
- Prajapati, G. K., Kashyap, N., Kumar, A., y Pandey, D. M. (2013). Identification of GCC Box in the Promoter region of Ubiquinol Cytochrome C Chaperone Gene using Molecular Beacon Probe and its *in silico* Protein-DNA interaction study in Rice (*Oryza sativa L.*). *International Journal of Computational Bioinformatics and In Silico Modeling*, 2(5):213–222.
- Pucciariello, C. y Perata, P. (2013). Quiescence in rice submergence tolerance: an evolutionary hypothesis. *Trends in Plant Science*, 18(7):377–381. DOI: [10.1016/j.tplants.2013.04.007](https://doi.org/10.1016/j.tplants.2013.04.007).
- Read, R. J., Adams, P. D., Arendall, W. Bryan, I., Brunger, A. T., Emsley, P., Joosten, R. P., Kleywegt, G. J., Krissinel, E. B., Lütke, T., Otwinowski, Z., Perrakis, A., Richardson, J. S., Sheffler, W. H., Smith, J. L., Tickle, I. J., Vriend, G., y Zwart, P. H. (2016). A New Generation of Crystallographic Validation Tools for the Protein Data Bank. *Structure*, 19(10):1395–1412. DOI: [10.1016/j.str.2011.08.006](https://doi.org/10.1016/j.str.2011.08.006).
- Rivera-Contreras, I., Zamora-Hernández, T., Huerta-Heredia, A., Capataz-Tafur, J., Barrera-Figueroa, B., Juntawong, P., y Peña-Castro, J. (2016a). Transcriptomic analysis of submergence-tolerant and sensitive *Brachypodium distachyon* ecotypes reveals oxidative stress as a major tolerance factor. *Scientific reports*, 6:27686. DOI: [10.1038/srep27686](https://doi.org/10.1038/srep27686).
- Rivera-Contreras, I., Zamora-Hernández, T., Huerta-Heredia, A., Capataz-Tafur, J., Barrera-

- Figueroa, B., Juntawong, P., y Peña-Castro, J. (2016b). Transcriptomic analysis of submergence-tolerant and sensitive *Brachypodium distachyon* ecotypes reveals oxidative stress as a major tolerance factor. *Scientific reports*, 6:27686. DOI: [10.1038/srep27686](https://doi.org/10.1038/srep27686).
- Roy, A., Kucukural, A., y Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols*, 5(4):725–738. DOI: [10.1038/nprot.2010.5](https://doi.org/10.1038/nprot.2010.5).
- Seo, Y.-S., Chern, M., Bartley, L. E., Han, M., Jung, K.-H., Lee, I., Walia, H., Richter, T., Xu, X., Cao, P., Bai, W., Ramanan, R., Amonpant, F., Arul, L., Canlas, P. E., Ruan, R., and Xuwei Chen, C.-J. P., Hwang, S., Jeon, J.-S., y Ronald, P. C. (2011). Towards Establishment of a Rice Stress Response Interactome. *PLoS Genetics*, 7(4):e1002020. DOI: [10.1371/journal.pgen.1002020](https://doi.org/10.1371/journal.pgen.1002020).
- Septiningsih, E. M., Pamplona, A. M., Sanchez, D. L., Neeraja, C. N., Vergara, G. V., Heuer, S., Ismail, A. M., y Mackill, D. J. (2009). Development of submergence-tolerant rice cultivars: the Sub1 locus and beyond. *Annals of Botany*, 103(2):151–60. DOI: [10.1093/aob/mcn206](https://doi.org/10.1093/aob/mcn206).
- Setter, T., Waters, I., Sharma, S., Singh, K., Kulshreshtha, N., Yaduvanshi, N., Ram, P., Singh, B., Rane, J., McDonald, G., y cols. (2009). Review of wheat improvement for waterlogging tolerance in Australia and India: the importance of anaerobiosis and element toxicities associated with different soils. *Annals of Botany*, 103(2):221–235. DOI: [10.1093/aob/mcn137](https://doi.org/10.1093/aob/mcn137).
- Setter, T. L., Bhekasut, P., y Greenway, H. (2010). Desiccation of leaves after de-submergence is one cause for intolerance to complete submergence of the rice cultivar IR 42. *Functional Plant Biology*, 37(11):1096–1104. DOI: [10.1071/FP10025](https://doi.org/10.1071/FP10025).
- Shigyo, M., Hasebe, M., y Ito, M. (2006). Molecular evolution of the AP2 subfamily. *Gene*, 366(1):156–65. DOI: [10.1016/j.gene.2005.08.009](https://doi.org/10.1016/j.gene.2005.08.009).
- Shiono, K. y Yamada, S. (2014). Waterlogging tolerance and capacity for oxygen transport in *Brachypodium distachyon* (Bd21). *Plant Root*, 8(0):5–12. DOI: [10.3117/plantroot.8.5](https://doi.org/10.3117/plantroot.8.5).
- Singh, S., Mackill, D. J., y Ismail, A. M. (2009). Responses of SUB1 rice introgression lines to submergence in the field: yield and grain quality. *Field Crops Research*, 113(1):12–23. DOI: [10.1016/j.fcr.2009.04.003](https://doi.org/10.1016/j.fcr.2009.04.003).
- Singh, U., Dar, M., Singh, S., Zaidi, N., Bari, M., Mackill, D., Collard, B., Singh, V., Singh, J., Reddy, J., y cols. (2013). *Field performance, dissemination, impact and tracking of submergence tolerant (Sub1) rice varieties in South Asia.*, volume 1. SABRAO.
- Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques*, 28(6):1102–1104.
- Swope, W. C., Andersen, H. C., Berens, P. H., y Wilson, K. R. (1982). A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of mo-

- lecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649. DOI: [10.1063/1.442716](https://doi.org/10.1063/1.442716).
- Tamang, B. G. y Fukao, T. (2015). Plant Adaptation to Multiple Stresses during Submergence and Following Desubmergence. *International Journal of Molecular Sciences*, 16(12):30164–30180. DOI: [10.3390/ijms161226226](https://doi.org/10.3390/ijms161226226).
- The Arabidopsis Information Resource (2016). About Arabidopsis, consultado en <https://www.arabidopsis.org/portals/education/aboutarabidopsis.jsp> en Julio de 2016.
- Tuszynska, I., Magnus, M., Jonak, K., Dawson, W., y Bujnicki, J. M. (2015). NPDock: a web server for protein–nucleic acid docking. page gkv493. DOI: [10.1093/nar/gkv493](https://doi.org/10.1093/nar/gkv493).
- van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., y Berendsen, H. J. C. (2005). GROMACS: Fast, Flexible and Free. *Journal of Computational Chemistry*, 26(1):1701–1718.
- Vashisht, D., Hesselink, A., Pierik, R., Ammerlaan, J., Bailey-Serres, J., Visser, E., Pedersen, O., Van Zanten, M., Vreugdenhil, D., Jamar, D., y cols. (2011). Natural variation of submergence tolerance among Arabidopsis thaliana accessions. *New Phytologist*, 190(2):299–310. DOI: [10.1111/j.1469-8137.2010.03552.x](https://doi.org/10.1111/j.1469-8137.2010.03552.x).
- Vergara, B. y Mazaredo, A. (1975). Screening for resistance to submergence under greenhouse conditions. In *6. Scientific Meeting of the Crop Science Society of the Philippines. Bacolod City (Philippines). 8 May 1975*.
- Verlet, L. (1967). Computer experiments on classical fluids I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159(1):98:103. DOI: [10.1103/PhysRev.159.98](https://doi.org/10.1103/PhysRev.159.98).
- Verlet, L. (1968). Computer experiments on classical fluids. II. Equilibrium correlation functions. *Physical Review*, 165(1):201:214. DOI: [10.1103/PhysRev.165.201](https://doi.org/10.1103/PhysRev.165.201).
- Voesenek, L. y Bailey-Serres, J. (2015). Flood adaptive traits and processes: An overview. *New Phytologist*, 206(1):57–73. DOI: [10.1111/nph.13209](https://doi.org/10.1111/nph.13209).
- Voesenek, L., Colmer, T., Pierik, R., Millenaar, F., y Peeters, A. (2006). How plants cope with complete submergence. *New phytologist*, 170(2):213–226. DOI: [10.1111/j.1469-8137.2006.01692.x](https://doi.org/10.1111/j.1469-8137.2006.01692.x).
- Voesenek, L. y Sasidharan, R. (2013). Ethylene–and oxygen signalling–drive plant survival during flooding. *Plant Biology*, 15(3):426–435. DOI: [10.1111/plb.12014](https://doi.org/10.1111/plb.12014).
- Weigel, D. (1995). The APETALA2 domain is related to a novel type of DNA binding domain. *The Plant Cell*, 7(4):388.
- Weigel, D. y Glazebrook, J. (2002). Arabidopsis. *A Laboratory Manual*, 165.
- Weits, D. A., Giuntoli, B., Kosmacz, M., Parlanti, S., Hubberten, H.-M., Riegler, H., Hoefgen, R., Perata, P., van Dongen, J. T., y Licausi, F. (2014). Plant cysteine oxidases control the oxygen-

- dependent branch of the N-end-rule pathway. *Nature communications*, 5. DOI: [10.1038/ncomms4425](https://doi.org/10.1038/ncomms4425).
- Wenkel, S., Turck, F., Singer, K., Gissot, L., Le Gourrierc, J., Samach, A., y Coupland, G. (2006). CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis. *The Plant Cell*, 18(11):2971–2984. DOI: [10.1105/tpc.106](https://doi.org/10.1105/tpc.106).
- Xia, K. y Wei, G.-W. (2014). Persistent homology analysis of protein structure, flexibility, and folding. *International Journal for Numerical Methods in Biomedical Engineering*, 30(8):814–844. DOI: [10.1002/cnm.2655](https://doi.org/10.1002/cnm.2655).
- Xu, K., Deb, R., y Mackill, D. J. (2004). A microsatellite marker and a codominant PCR-based marker for marker-assisted selection of submergence tolerance in rice. *Crop science*, 44(1):248–253. DOI: [doi:10.2135/cropsci2004.2480](https://doi.org/10.2135/cropsci2004.2480).
- Xu, K. y Mackill, D. J. (1996). A major locus for submergence tolerance mapped on rice chromosome 9. *Molecular Breeding*, 2(3):219–224. DOI: [10.1007/BF00564199](https://doi.org/10.1007/BF00564199).
- Xu, K., Xu, X., Fukao, T., Canlas, P., Maghirang-Rodriguez, R., Heuer, S., Ismail, A., Bailey-Serres, J., Ronald, P., y Mackill, D. (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature*, 442(7103):705–708. DOI: [10.1038/nature04920](https://doi.org/10.1038/nature04920).
- Zamora-Hernandez, T. (2016). *Identificación y estudio de la expresión de los Factores de Respuesta al Etileno del Grupo VII de Brachypodium distachyon que participan en la respuesta al estrés por inundación*. Tesis, Universidad del Papaloapan.
- Zhang, C. y Ma, J. (2010). Enhanced sampling and applications in protein folding in explicit solvent. *The Journal of Chemical Physics*, 132(24):244101. DOI: [10.1063/1.3435332](https://doi.org/10.1063/1.3435332).
- Zhang, C. y Ma, J. (2012). Folding helical proteins in explicit solvent using dihedral-biased tempering. *Proceedings of the National Academy of Sciences*, 109(21):8139–8144. DOI: [10.1073/pnas.1112143109](https://doi.org/10.1073/pnas.1112143109).
- Zhou, Y., Duan, Y., Yang, Y., Faraggi, E., y Lei, H. (2011). Trends in template/fragment-free protein structure prediction. *Theoretical Chemistry Accounts*, 128(1):3–16. DOI: [10.1007/s00214-010-0799-2](https://doi.org/10.1007/s00214-010-0799-2).