

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA



**José Pedro Fonseca**

**“Vias de sinalização de estresses em plantas: análise da região promotora do gene *NIMIN-1* de *Arabidopsis thaliana* e da proteína ScCBL1 de cana-de-açúcar (*Saccharum spp.*)”**

Tese apresentada ao Instituto de Biologia para obtenção do Título de Doutor em Genética e Biologia Molecular, na área de Genética Vegetal e Melhoramento.

Este exemplar corresponde à versão final da tese defendida pelo(a) candidato (a)

**JOSÉ PEDRO FONSECA**

*[Handwritten signature]*

aprovada pela Comissão Julgadora

Orientador: Prof. Dr. Marcelo Menossi

Co-Orientador: Prof. Dr. Ricardo Aparicio

Campinas, 2010

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP

F733v

Fonseca, Jose Pedro

Vias de sinalização de estresses em plantas: análise da região promotora do gene NIMIN-1 de *Arabidopsis thaliana* e da proteína ScCBL1 de cana-de-açúcar (*Saccharum spp.*) / Jose Pedro Fonseca. – Campinas, SP: [s.n.], 2010.

Orientadores: Marcelo Menossi Teixeira, Ricardo Aparício.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Gene NIMIN-1. 2. TGA. 3. *Arabidopsis thaliana*. 4. Cana-de-açúcar – Melhoramento genético. 5. Proteína CBL. 6. Transdução de sinal celular. I. Teixeira, Marcelo Menossi. II. Aparício, Ricardo. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

(rcdt/ib)

**Título em inglês:** Signal transductional pathways under biotic and abiotic stress in plants: functional analysis of NIMIN-1 promoter region in *Arabidopsis* and characterization of a calcium binding protein (ScCBL1) in sugar cane (*Saccharum spp.*).

**Palavras-chave em inglês:** NIMIN-1 gene; TGA; *Arabidopsis thaliana*; Sugarcane – Genetic enhancement; CBL protein; Cellular signal transduction.

**Área de concentração:** Genética Vegetal e Melhoramento.

**Titulação:** Doutor em Genética e Biologia Molecular.

**Banca examinadora:** Marcelo Menossi Teixeira, Luis Eduardo Aranha Camargo, Helaine Carrer, Márcio José da Silva, Gonçalo Apolinário de Souza Filho.

**Data da defesa:** 28/04/2010.

**Programa de Pós-Graduação:** Genética e Biologia Molecular.

Campinas, 28 de Abril de 2010

**BANCA EXAMINADORA**

Prof. Dr .Marcelo Menossi (Orientador)



Assinatura

Prof. Dr. Luis Eduardo Aranha Camargo



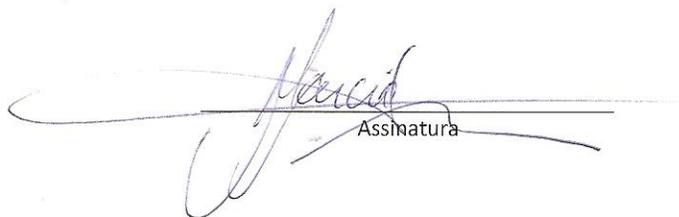
Assinatura

Profa. Dra. Helaine Carrer



Assinatura

Prof. Dr. Márcio José da Silva



Assinatura

Prof. Dr. Gonçalo Apolinário de Souza Filho



Assinatura

Prof. Dr. Jorg Kobarg

Assinatura

Prof. Dr. Carlos Colombo

Assinatura

Prof. Dr. Fabio Papes

Assinatura

## AGRADECIMENTOS

Meu muito obrigado ao Prof. Dr. Marcelo Menossi pela orientação, apoio e paciência;

Ao meu orientador do doutorado sanduiche Prof. Dr. Cristopher D. Town e todo o grupo de genômica de plantas do John Craig Venter Institute nos EUA.

Ao meu co-orientador Ricardo Aparício do Instituto de Química da UNICAMP pela co-orientação e coordenação da parte estrutural do trabalho.

Ao Prof. Dr. Marcelo Menossi, Prof. Dr. Luis Eduardo Aranha Camargo, Prof. Dr. Gonçalo Guimarães Pereira pelas sugestões e críticas na banca prévia.

Ao Prof. Dr. Marcelo Menossi, Prof. Dr. Luis Eduardo Aranha Camargo, Profa. Dra. Helaine Carrer, Prof. Dr. Márcio José da Silva e Prof. Dr. Gonçalo Apolinário de Souza Filho pela participação na banca.

A todos os amigos do LGF pelo companheirismo do dia a dia.

Ao RoboLab do LNLS pela colaboração com as análises cristalográficas da ScCBL1.

A estrutura oferecida pelo LNLS.

Ao laboratório do Prof. Dr. Hiroshi por gentilmente ceder o uso do sonicador sempre que necessário.

A minha família: minha mãe, meu pai, meu irmão e minha avó que sempre me apoiaram.

A minha querida Carolina; esta tese é dedicada a você.

Aos grandes amigos de Unicamp Marcela, Vera, José Eustáquio, Marcos e Bruno pela solidariedade e amizade.

## ÍNDICE

<b>RESUMO</b>	1
<b>ABSTRACT</b>	4
<b>INTRODUÇÃO GERAL</b>	7
Vias de sinalização em resposta estresse em plantas	7
<b>CAPÍTULO I - Análise funcional da região promotora do gene <i>NIMIN-1</i> em <i>Arabidopsis thaliana</i></b>	9
<b>INTRODUÇÃO</b>	10
Resistência Sistêmica Adquirida (SAR)	10
Fatores de transcrição bZIP em Arabidopsis	12
Fatores de transcrição TGA	13
Hibridização do sitio de ligação de TGA2 a microarranjos em Arabidopsis	16
<i>NIMIN-1</i>	17
<b>OBJETIVOS</b>	19
<b>MATERIAIS E MÉTODOS</b>	20
Material Vegetal	20
Clonagem Gateway das regiões promotoras de cinco genes de Arabidopsis	20
Transformação de <i>Agrobacterium</i> por “triparental mating”	25
Transformação estável de Arabidopsis	26
Ensaio de expressão transiente por agroinfiltração	26
Extração de RNA e síntese de cDNA	27
PCR quantitativo em tempo real	28
<b>RESULTADOS E DISCUSSÃO</b>	30

Padrão de expressão de quatro genes candidatos	30
Análise funcional da região promotora de <i>NIMIN-1</i>	32
<b>CONCLUSÕES</b>	40
<b>REFERÊNCIAS</b>	41
<b>CAPÍTULO II - caracterização da proteína ScCBL1 em cana-de-açúcar (<i>Saccharum spp.</i>).</b>	47
<b>INTRODUÇÃO</b>	48
O papel do cálcio na resposta a estresses bióticos e abióticos	48
Proteínas sensoras de cálcio em plantas	49
Proteínas tipo Calcineurina B (CBL) em plantas	49
Modificação pós-traducional das CBL's	52
Proteômica estrutural	53
Dicroísmo circular (CD) de proteínas	53
Espalhamento dinâmico de luz (“Dynamic Light Scattering”, DLS)	56
Cristalografia de Proteínas	56
Espalhamento de Raios-X a Baixos Ângulos (SAXS)	60
<b>OBJETIVOS</b>	65
<b>MATERIAIS E MÉTODOS</b>	66
Clonagem e transformação de <i>Escherichia coli</i>	66
Ensaio de expressão	67
Deteção da proteína recombinante	68
Purificação por cromatografia de afinidade em resina de níquel (Ni-NTA)	69
Espalhamento dinâmico de luz (DLS)	69

Espectroscopia de Dicroísmo Circular (CD)	70
Espalhamento de Raios-X a Baixos Ângulos (SAXS)	71
Ensaio de Cristalização de proteínas	72
<b>RESULTADOS E DISCUSSÃO</b>	73
Expressão, detecção por Western blott e purificação das proteínas ScCBL1, ScCBL3 e ScCBL6	73
Caracterização da estrutura secundária de ScCBL1 por dicroísmo circular (CD)	80
Caracterização por espalhamento dinâmico de luz (DLS)	81
Ensaio de cristalização	83
Caracterização estrutural por espalhamento de raios-x a baixos ângulos (SAXS)	85
<b>CONCLUSÕES</b>	98
<b>CONCLUSÃO GERAL E PERSPECTIVAS</b>	99
<b>REFERÊNCIAS</b>	102
<b>ANEXOS</b>	108
Anexo I	108
Anexo II	109
Anexo III	110
Anexo IV	113
Anexo V	114

## LISTA DE ABREVIATURAS

AGI	Código de identificação de genes em <i>Arabidopsis Thaliana</i>
BSA	Albumina de soro bovino
cDNA	DNA complementar
CDS	Sequência codante de um gene
ChIP	Imunoprecipitação da cromatina
ChIP-chip	Hibridização de alvos imunoprecipitados por ChIP a microarranjos de DNA
Ct	Número de ciclos em uma reação de PCR em tempo real que o sinal de fluorescência de uma amostra leva para ser detectado
DNA	Ácido desoxirribonucléico
DPI	Dias pós-infiltração
dsDNA	DNA dupla-fita
DTT	1,4-ditiotreitol
Genes PR	Genes relacionados a patogenicidade ou defesa em plantas
GFP	Proteína Verde Fluorescente
GUS	$\beta$ -glucuronidase
HR	Resposta Hipersensível, mecanismo de defesa em plantas
INA	Ácido 2,6 dicloroisonicotínico
IPTG	isopropil-tio- $\beta$ -D-galactopiranosídeo
kDa	KiloDaltons
LB	Reação Gateway LB (clonagem no vetor de entrada)
LR	Reação Gateway LR (clonagem no vetor destino)

ORF	Open reading frame (Sequência nucleotídica compreendida entre um códon de início ATG e um códon de terminação)
qRT-PCR	PCR quantitativo em tempo real
RNA	Ácido ribonucléico
SDS-PAGE	Eletroforese em gel de poliacrilamida com duodecil-sulfato de sódio
SA	Ácido Salicílico
SAR	Resistência Sistêmica Adquirida
TBS	Tampão Tris salino
$\Delta\Delta Ct$	Método de cálculo da expressão relativa de um gene baseado na normalização do gene de interesse em relação a um gene invariante.

## RESUMO

Estresses bióticos e abióticos como a seca, salinidade e ataque por patógenos são responsáveis por perdas significantes em culturas de grãos ao redor do mundo. Diversos genes são regulados em resposta a esses fatores e podem ser ativados ou reprimidos para gerar uma resposta específica na planta de maneira a gerar uma resposta de defesa que atenua os efeitos do estresse e promoção de tolerância pela planta. É importante entendermos o funcionamento desses mecanismos moleculares, e dos genes e proteínas envolvidas nestas vias de sinalização para um melhor conhecimento de como estas vias de transdução operam em plantas, bem como no desenvolvimento de variedades de plantas tolerantes.

No capítulo I deste trabalho nós descrevemos a análise funcional de um motivo de ligação do fator TGA localizado na região promotora do gene *NIMIN-1* que é altamente induzido por ácido salicílico (SA) durante defesa de plantas (estresse biótico). Fatores TGA desempenham um papel chave na defesa de plantas através da interação com elementos presentes na região promotora de genes de defesa para induzir a sua expressão. O ácido salicílico (SA) é um fito-hormônio que induz a expressão do gene que codifica a proteína NIMIN-1. Essa proteína interage com a proteína NPR1/NIM1, reguladora da resistência sistêmica adquirida (SAR). Neste trabalho foi investigado se um motivo de ligação do fator TGA2 “TGACG”, localizado na região promotora imediatamente anterior ao sítio de iniciação da transcrição de *NIMIN-1*, é necessário a indução de *NIMIN-1* por ácido salicílico. Uma versão mutagenizada do promotor do gene *NIMIN-1* foi gerada por mutação sítio-dirigida. Nós geramos construções T-DNA nas quais tanto a região promotora nativa quanto a mutagenizada foram fusionadas aos repórteres proteína verde fluorescente (GFP) e

beta-glucuronidase (GUS). Foram geradas plantas transgênicas e a expressão de GFP sob o controle da região promotora de *NIMIN-1* foi observada em raízes, no pecíolo e folhas de *Arabidopsis*. A atividade dirigida pelo promotor mutagenizado e o selvagem foi quantificada por PCR quantitativo em tempo real. Tanto a construção contendo o promotor de *NIMIN-1* como a cópia endógena de *NIMIN-1* foram induzidas por SA. Em contraste, a construção promotora mutagenizada foi bem menos sensível a SA que o promotor nativo de *NIMIN-1*. Esse dado indica que o motivo de ligação do fator TGA2 analisado está diretamente envolvido na modulação da expressão de *NIMIN-1* induzida por SA em *Arabidopsis*.

No capítulo II nós descrevemos a caracterização da proteína ScCBL1 de cana-de-çúcar que apresenta elevada identidade com membros da família de proteínas sensoras de cálcio do tipo calcineurina B (CBL) em plantas. Experimentos de duplo-híbrido realizados em nosso grupo mostram que a proteína ScCBL1 interage com uma proteína quinase (ScCIPK8). Trabalhos prévios também desenvolvidos em nosso laboratório demonstraram que ScCIPK8 está envolvida no metabolismo do carboidrato e é diferencialmente expressa em resposta a ABA. O gene *ScCBL1* foi clonado e a proteína codificada expressa e purificada a partir do extrato solúvel por cromatografia de afinidade usando resina Ni-NTA e a proteína eluída foi usada para estudos estruturais. Análises por espectrometria por difração circular (CD) mostraram que a proteína ScCBL1 é composta predominantemente por  $\alpha$ -hélices, em concordância com programas de predição da sequência de aminoácido desta proteína. Experimentos de espalhamento de raio-x a baixos ângulos (SAXS) indicaram que as amostras obtidas da ScCBL1 estavam homogêneas e monodispersas em solução e que ocorre uma mudança em seu estado oligomérico quando adicionado o agente

reduzidor DTT, ocorrendo uma diminuição na intensidade de espalhamento ( $I(0)$ ) a uma ordem de 1,56 para a mesma concentração, acompanhado de uma diminuição de 10 Å em seu raio de giro. As análises por SAXS indicaram que a proteína ScCBL1 é pentamérica em seu estado nativo e um trímero quando adicionado DTT. Análises por SAXS também indicaram que a proteína ScCBL1 está enovelada em solução, apresentando estrutura terciária estável. A modelagem de baixa resolução *ab initio*, juntamente com a função de distribuição das distâncias  $P(r)$  indicaram que a proteína possui um formato alongado (prolato). Ensaio inicial de cristalização a partir de amostras monodispersas da proteína ScCBL1, confirmadas por DLS, foram feitas e um monocristal simétrico de aproximadamente 0.05 mm de diâmetro obtido além de outros sinais promissores para um refinamento das condições de cristalização de ScCBL1 para determinação de sua estrutura tridimensional a uma alta resolução. Esses dados contribuem para uma caracterização inicial da estrutura da proteína ScCBL1.

## ABSTRACT

Many biotic and abiotic stresses such as drought, salinity and pathogen attack are responsible for major crop losses around the world. Several genes are regulated in response to these factors to counteract the stress effects and promote plant tolerance. Understanding the molecular mechanisms as well as the roles of genes and proteins involved in stress signaling pathways is key to a better understanding of plant tolerance.

We report in chapter I the functional analysis of a TGA binding factor located in the promoter region of *NIMIN-1* that is highly induced by SA during plant defense against pathogen attack (biotic stress). TGA factors play a key role in plant defense by binding to the promoter region of defense genes, inducing their expression. Salicylic acid (SA) induces the expression of the gene encoding NIMIN-1, which interacts with NPR1/NIM1, a key regulator of systemic acquired resistance. In this work we investigated whether the TGA2-binding motif TGACG located upstream of the *NIMIN-1* gene is necessary for SA induction of *NIMIN-1* expression. A mutated version of the *NIMIN-1* promoter was created by site-directed mutagenesis. We generated T-DNA constructs in which native *NIMIN-1* and mutated promoters were fused to green fluorescent protein and  $\beta$ -glucuronidase reporters. We produced transgenic *Arabidopsis* plants and observed *NIMIN-1* promoter-driven green fluorescent protein expression in the roots, petiole and leaves. Constructs were also agroinfiltrated into the leaves to evaluate gene expression of mutagenized and native promoters driving expression of GUS using quantitative real-time RT-PCR. We characterized the normal gene response to SA and compared it to the response of the mutant version of the *NIMIN-1* promoter. Both the native *NIMIN-1* construct and an

endogenous copy of *NIMIN-1* were induced by SA. However, the mutated promoter construct was much less sensitive to SA than the native *NIMIN-1* promoter, indicating that this TGA2-binding motif is directly involved in the modulation of SA-induced *NIMIN-1* expression in *Arabidopsis*.

In chapter II we describe the characterization of a sugarcane ScCBL1 protein which displays high sequence identity to the calcium binding protein family calcineurin B-like (CBL) from plants. Using the two-hybrid system our group has shown that ScCBL1 binds to a protein kinase (ScCIPK8). Previous work done in our laboratory also showed that ScCIPK8 is involved in carbohydrate metabolism and that it is differentially expressed in response to ABA. ScCBL1 was cloned, expressed and purified by one round of affinity chromatography using a Ni-NTA resin and the purified eluted protein was used for structural analysis. Spectroscopic analysis by circular dichroism (CD) showed that ScCBL1 is mainly composed of  $\alpha$ -helices agreeing with secondary structure prediction by PredictProtein server. Small Angle X-Ray Scattering (SAXS) analysis showed that ScCBL1 sample is homogeneous and monodisperse in solution and that the protein undergoes an oligomeric change when DTT is added, with a decrease in scattering intensity ( $I(0)$ ) by a factor of 1,56 for samples with the same concentration, together with a decrease in the radius of gyration of 10 Å. SAXS experiments also showed that ScCBL1 is pentameric in its native state and the protein undergoes a change in its oligomeric state to a trimer when DTT is added. SAXS experiments also showed the protein is folded in solution and *ab initio* modeling of ScCBL1 protein envelope together with the pair-distribution function  $P(r)$  indicates that the protein has a rod-like, elongated shape. An initial crystallization screening with ScCBL1 monodisperse samples (confirmed by DLS

experiments) was carried out and some crystallization signals were obtained, including a single crystal of around 0.05 mm in length. These data shed light on the structural features of ScCBL1.

## **I. INTRODUÇÃO GERAL**

O termo estresse em seu sentido biológico foi definido como “uma força ou condição adversa que inibe o funcionamento normal e bem estar de um sistema biológico como, por exemplo, uma planta” (Jones e Jones., 1989). Estresses abióticos tais como frio, calor, salinidade, seca e escassez de nutrientes no solo, bem como de estresses bióticos, como o ataque de vírus, bactéria, fungo, insetos e roedores provocam perdas na produção de grãos mundial causando uma diminuição na produção do rendimento de grandes culturas em mais de 50% (Mahajan e Tuteja., 2005; Bray et al., 2000; Flowers., 2004). Em um cenário no qual a demanda por alimentos é cada vez maior, é desejável não só um aumento na produção de grãos como também o desenvolvimento de novas variedades de plantas tolerantes a diversos tipos destes estresses bióticos e abióticos. Desta forma seriam evitadas perdas de produtividade e aumentariam as garantias da oferta de alimentos para uma população global em crescimento.

Portanto, um entendimento dos mecanismos moleculares envolvidos nestas vias de sinalização em resposta a diversos tipos de estresse, bem como os genes envolvidos é uma primeira etapa para o melhoramento genético de variedades tolerantes.

### **I.1. Vias de sinalização em resposta a estresse em plantas**

Plantas são constantemente desafiadas por diversos tipos de estresses bióticos e abióticos, e por serem organismos sésseis, desenvolveram uma complexa rede de sinalização ao longo da evolução para adaptarem-se a todos estes fatores. De maneira geral o sinal de estresse é primeiramente detectado na membrana por receptores e então através

de uma via de transdução de sinal (“signal transduction”) na célula ativa a transcrição de genes responsivos ao estresse específico. Os mecanismos desta rede de sinalização foram caracterizados previamente e consistem, tipicamente, na percepção, transmissão e processamento de um sinal para geração de uma resposta por parte da planta através da interação de fito-hormônios, mensageiros secundários, fatores de transcrição (FTs), genes e proteínas (Pandey., 2008).

No presente trabalho foi feita a análise funcional da região de ligação do fator TGA2 “TGACG” na região promotora do gene *NIMIN-1* em *Arabidopsis*, envolvido na via de sinalização de defesa de plantas, especificamente em SAR (resistência sistêmica adquirida) (Capítulo I). No capítulo II foi feita a caracterização da proteína sensora de cálcio ScCBL1 em cana-de-açúcar (*Saccharum spp.*), relacionada a via de sinalização em resposta a estresse por seca.

**CAPÍTULO I - Análise funcional da região promotora do gene  
*NIMIN-1* em *Arabidopsis thaliana***

# 1. INTRODUÇÃO

## 1.1. Resistência sistêmica adquirida (SAR)

As plantas desenvolveram sofisticados mecanismos moleculares de defesa contra patógenos ao longo da evolução. Um destes mecanismos de defesa é conhecido como resposta hipersensível (HR) que consiste na expressão de genes de defesa pela planta, acompanhada pela codificação de genes de avirulência pelo patógeno (Baker et al., 1997; Staskawicz et al., 1995). A expressão de genes HR pela planta leva a uma rápida resposta local via morte celular programada e consequente confinamento do patógeno nos locais iniciais de infecção (Glocova et al., 2005). Outro mecanismo de defesa em plantas é conhecido como Resistência Sistêmica Adquirida (“Systemic Acquired Resistance” ou SAR do inglês). A SAR pode ser definida como um sistema geral de defesa da planta quando atacada por patógeno, que se desenvolve em regiões distais, não infectadas da planta, tornando estas regiões menos sensíveis a novos ataques (Ross., 1961; Glazebrook et al., 1996; Durrant e Dong., 2004; Ryals et al., 1996). Tanto em *Arabidopsis* como em tabaco o estabelecimento de reações de defesa pela planta é acompanhado por um aumento da biosíntese de ácido salicílico (SA). Portanto a via de sinalização relacionada à SAR tem início com a síntese do fito-hormônio SA, tanto no tecido infectado como não infectado da planta, como resposta a infecção pelo patógeno e culmina com a síntese de proteínas relacionadas à patogenicidade (PR). De fato, SA é necessário para indução dos genes de defesa (“pathogenesis-related” ou genes PR) e resistência ao patógeno durante a SAR (Zhang et al., 2003).

A via de transdução bioquímica em SAR ocorre da seguinte forma: a acumulação de ácido salicílico ativa uma proteína citosólica intermediária NPR1, que por sua vez interage com fatores de transcrição TGA, que se ligam à região promotora e regulam a expressão de genes de defesa (Johnson et al., 2003) (Figura 1).

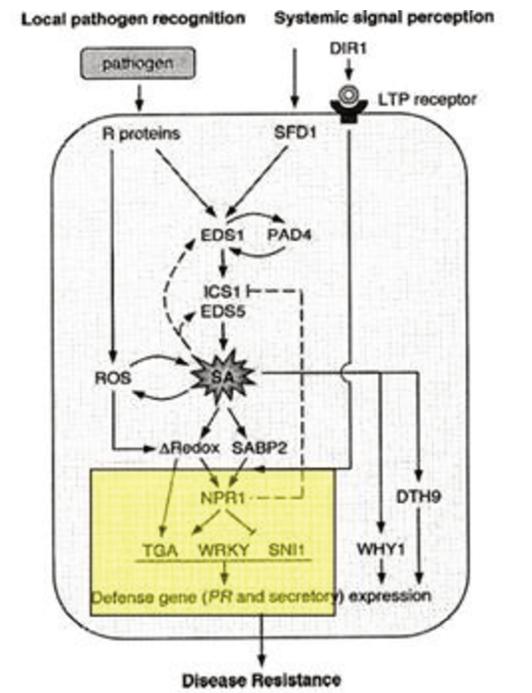


Figura 1: Esquema da via da Resistência Sistêmica Adquirida. A via de sinalização de defesa ativada na planta que tem início com a acumulação de SA em resposta a estresse xenobiótico, infecção ou defesa a patógeno (Durrant et al., 2004). Ocorre então a ativação da proteína NPR1, que ativa proteínas que se ligam às regiões promotoras, ativando genes efetores.

A proteína NPR1 desempenha um papel central em SAR através de mudanças em seu estado celular quando é reduzida, passando de um estado oligomérico para uma forma monomérica ativa (Mou et al., 2003). Esta forma monomérica migra do citoplasma para o núcleo onde se ligará especificamente ao fator TGA2 e outros fatores de transcrição do tipo TGA, que por sua vez irão se ligar e regular a expressão de genes de defesa (Johnson et al., 2003). Plantas transgênicas superexpressando NPR1 apresentaram maior resistência ao

patógenos bacteriano *Pseudomonas syringae* (linhagem ES4326) em *Arabidopsis* (Cao et al., 1998). De fato, já foi demonstrado que tratamento de plantas com ácido salicílico leva a indução de genes de defesa (genes PR) e conseqüentemente estabelecimento de SAR (Cao et al., 1998; Ward et al., 1991). Similarmente já foi proposto que os fatores TGA também interagem e regulam a expressão de outros fatores de transcrição (por exemplo, os fatores WRKY) que limitariam a taxa de expressão de genes de defesa num mecanismo de regulação negativa (Johnson et al., 2003).

## **1.2. Fatores de transcrição bZIP em Arabidopsis**

Fatores de transcrição participam ativamente nos mais diversos processos biológicos em plantas. Os fatores de regulação de transcrição do tipo “basic leucine zipper” (bZIP) foram descritos em todos os eucariotos. O seu domínio de ligação ao DNA é formado por uma região rica em aminoácidos básicos que fazem ligação com as bases do DNA e por um zíper de leucinas constituído por repetições de aminoácidos hidrofóbicos (principalmente leucinas) a cada sete aminoácidos e que causa a dimerização (Figura 2). Estudos cristalográficos do fator bZIP GCN4 de levedura ligado a sua seqüência de DNA alvo revelaram que o domínio bZIP forma uma  $\alpha$ -hélice. Os dois zíperes de leucinas interagem produzindo uma estrutura “coiled-coil” e as regiões básicas do dímero encaixam-se no sulco maior da seqüência de DNA alvo (Ellenberg et al., 1992; König e Richmond., 1993; Hurst., 1996).

Análises genéticas, moleculares e bioquímicas indicam que os fatores bZIP são reguladores importantes de processos específicos de angiospermas (plantas com flores) como a fotomorfogênese (Osterlund et al., 2000); desenvolvimento de órgãos (Walsh et al., 1997; Chuang et al., 1999; Schlögl et al., 2008); mecanismos de defesa (Niggeweg et al.,

2000; Zhang et al., 1999; Despres et al., 2000); via de sinalização de hormônios e da sacarose (Rook et al., 1998; Finkelstein e Lynch., 2000; Schlögl et al., 2008) e resposta a luz ( Schindler et al., 1992) entre outros.

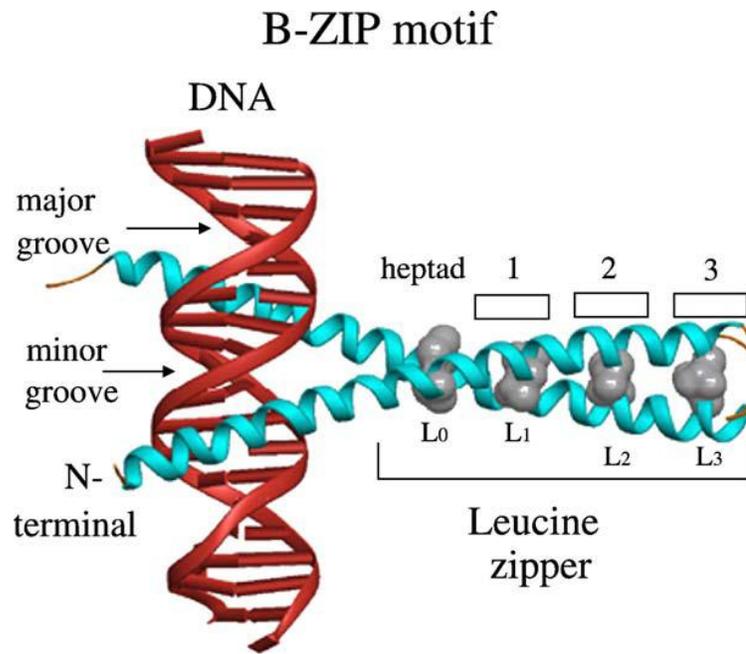


Figura 2: Estrutura de raios-X do dímero B-ZIP GCN4 ligado a dupla-fita de DNA em levedura (Ellenberg et al., 1992).

Com a disponibilização das seqüências genômicas de *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000), verificou-se que aproximadamente 5,9% de todos os seus genes (~28.000) codificam para aproximadamente 1500 fatores de transcrição (The Arabidopsis Genome Initiative., 2000; Riechmann et al., 2000; Yamada et al., 2003). Sessenta e sete proteínas bZIP foram identificadas em uma anotação recente do genoma de *A. thaliana* (Deppmann et al., 2004).

### 1.3. Fatores de transcrição TGA

Neste trabalho é abordado mais especificamente um fator de transcrição bZIP do grupo D, que compreende os fatores TGA. De acordo com a classificação de Jakoby et al.

(2002), fatores de transcrição bZIP do grupo D participam em dois processos biológicos: defesa contra patógenos e desenvolvimento vegetal (Figura 3). Seu envolvimento em defesa provém de trabalho com fatores TGA em tabaco e *Arabidopsis* (Katagiri et al., 1989; Lam et al., 1989).

Os Fatores TGA foram primeiramente descritos em tabaco pela sua capacidade em ligar-se especificamente ao elemento as-1 localizado na região promotora do vírus mosaico da couve flor CaMV 35S e promover transcrição (Katagiri et al., 1989; Lam et al., 1989). O envolvimento dos fatores TGA na via de sinalização de ácido salicílico foi confirmado pela presença de sítios de ligação dos fatores TGA na região promotora do gene PR-1 e pelo papel desses sítios na expressão do gene *PR-1* modulada por ácido salicílico (Lebel et al., 1998). A evidência genética do papel dos fatores TGA em SAR foi demonstrada por Zhang et al (2003) na qual a indução do gene de defesa PR-1 pelo INA (análogo a SA) e resistência ao patógeno *Peronospora parasitica* foi abolida no mutante triplo “knockout” TGA 2,5,6, abolindo portanto a SAR e em trabalho usando o sistema de duplo híbrido em levedura que identificou interação com a proteína NPR1 (Zhang et al., 1999; Despres et al., 2000).

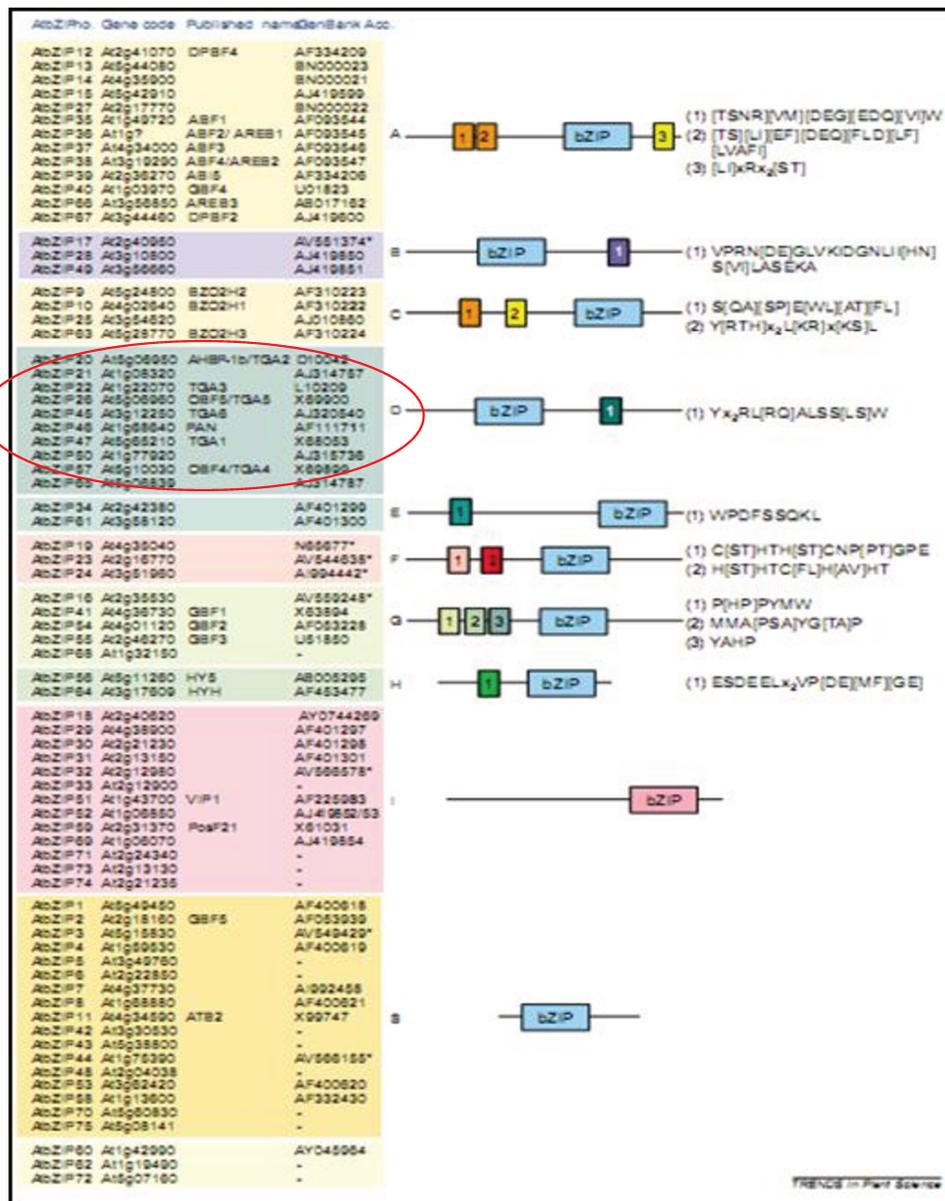


Figura 3: Os fatores TGA em Arabidopsis são uma subfamília conservada de fatores de transcrição bZIP do grupo D, composta por dez membros que estão envolvidos tanto em defesa da planta como desenvolvimento vegetal. Essa classificação foi baseada por similaridade de sequência da região básica e a presença de motivos conservados entre os fatores bZIP de Arabidopsis. Os domínios bZIP aparecem em azul e os domínios conservados característicos de cada subfamília numerados (adaptado de Jakoby et al., 2002).

A indução da proteína PR-1 não foi bloqueada no mutante TGA6 ou no mutante duplo (TGA2,5), revelando o papel redundante dos fatores TGA2,5 e 6. Experimentos “in

*vitro*” demonstraram que o motivo ‘TGACG’ é suficiente para a ligação do fator TGA (Lam et al., 1989). Em *Arabidopsis* a família TGA é composta por dez membros e foi demonstrada a ligação *in vivo* de dois destes (TGA2 e TGA3) à região promotora do gene *PR-1* através de imunoprecipitação da cromatina (ChIP) (Johnson et al., 2003).

#### **1.4. Hibridação do sítio de ligação de TGA2 a microarranjos em Arabidopsis**

A tecnologia ChIP permite o enriquecimento das regiões de DNA alvo as quais se ligam os fatores de transcrição (interação DNA - proteína) com a utilização de anticorpos específicos contra a proteína que está sendo analisada. Este método permite identificar a associação de proteínas a uma região do genoma *in vivo* através da fixação da proteína ligada ao DNA pelo formaldeído. A proteína é então imunoprecipitada com anticorpos específicos, e após a ligação do DNA com a proteína ser desfeito, o DNA imunoprecipitado pode ser amplificado por PCR ou usado para hibridação contra o genoma de *Arabidopsis* (ChIP-chip). Esta última técnica permite avaliar em quais regiões do genoma inteiro o fator de transcrição específico é ligado (Vinson et al., 2006). A identificação de mais de cem motivos de ligação de fatores de transcrição foi feita na levedura *Saccharomyces cerevisiae* através da hibridização da cromatina imunoprecipitada a “arrays” de genoma inteiro em uma técnica mais comumente referida como ChIP-chip (Lee et al., 2002).

Esforços recentes para identificar sítios de ligação para o fator TGA2 em *Arabidopsis* utilizaram o método ChIP-chip para a hibridização da cromatina imunoprecipitada de TGA2 a “arrays” de genoma inteiro ou contendo somente regiões promotoras em *Arabidopsis*. Este último “array” definido como regiões de 2 Kb localizadas

anteriormente ao códon de iniciação para todos os genes anotados na época do trabalho no “TIGR Release 4” (Thibaud-Nissen et al., 2006).

Um total de 51 regiões do genoma foram encontradas “enriquecidas” e a análise com o programa de busca por motivos SIFT (Hudson e Quail, 2003) revelou que cinquenta e cinco por cento destas regiões continham o motivo de ligação dos fatores TGA (TGACG). Para identificar os genes que são provavelmente alvos diretos dos fatores TGA somente os genes localizados a uma distância de 3 Kb de uma região de ligação identificada por ChIP-chip do fator TGA2 foram selecionados. Cento e dezessete genes foram selecionados de acordo com este critério (10% de um total de 1297 genes), mas somente vinte e quatro destes genes contém o sítio de ligação para o fator TGA2 e eram simultaneamente induzidos por ácido salicílico.

Diversos sítios de ligação ao fator TGA2 foram identificados pelo método de ChIP-chip descrito anteriormente, mas somente aqueles que continham o motivo de ligação de TGA2, localizados na região promotora próxima de um gene alvo, e genes que são induzidos por SA foram selecionados para os demais estudos. Dentre estes genes, *At1g02450 (NIMIN-1)* apresentou a maior indução da expressão gênica quando tratado com SA de acordo com dados de microarranjo da Affymetrix (Thibaud-Nissen et al., 2006).

### **1.5. *NIMIN-1***

Os genes *NIMIN* (“NIM1 interacting”) foram primeiramente identificados pela sua interação com a proteína NPR1 (ou NIM1) usando a técnica de duplo-híbrido em levedura (Weigel et al., 2001). Os genes *NIMIN* são expressos em *Arabidopsis* em resposta a tratamentos indutores de SAR como SA e seu análogo INA. As proteínas codificadas por

esta família de genes possuem sinais de localização celulares que as direcionam para o núcleo da célula (Kinkema et al., 2000; Mou et al., 2003). Esses dados indicam que NIMIN-1, através da interação com NPR1, faz parte da via de sinalização que leva a expressão de genes relacionados a SAR. Plantas transgênicas superexpressando *NIMIN-1* apresentaram reduzida expressão dos genes de defesa (genes *PR*) induzidos por SA, comprometendo a SAR, além de reduzidas quantidades da proteína NPR1 (Weigel et al., 2005). Esses experimentos demonstraram que NIMIN-1 age como um regulador negativo de NPR1.

A indução da expressão do gene *NIMIN-1* em resposta a SA foi previamente demonstrada (Weigel et al., 2001; Glocova et al., 2005). Foi também observado um aumento do número de transcritos do gene *NIMIN-1* em resposta rápida (2 h) a SA em experimentos de microarranjo de DNA em *Arabidopsis* (Thibaud-Nissen et al., 2006). A expressão do gene repórter *GUS* sob o controle da região promotora de *NIMIN-1* (1.2 Kb anterior ao sítio de início da transcrição) em plantas transgênicas de tabaco demonstrou que essa região é sensível à indução por SA (Glocova et al., 2006).

Os elementos contidos na região promotora de *NIMIN-1* em *Arabidopsis* que regulam a ativação transcricional da expressão de *NIMIN-1* em resposta a SA continuam desconhecidos. A verificação funcional da região promotora do gene *NIMIN-1* é importante para sabermos se a ligação a fatores TGA é necessária para expressão de *NIMIN-1* induzida por SA. A confirmação dos elementos que regulem a expressão de genes induzidos por SA é um passo importante para desvendarmos novos genes alvo de fatores TGA que possam estar envolvidos em SAR, contribuindo para um melhor entendimento de como estas vias de transdução ou redes regulatórias operam em plantas. Para nossas análises iniciais da

expressão gênica das regiões promotoras selecionadas decidimos privilegiar a análise do gene *NIMIN-1* por mostrar a maior indução a SA de acordo com dados Affymetrix.

## 2. OBJETIVOS

Os principais objetivos desta primeira etapa do trabalho foram:

1. Avaliar o padrão de expressão de regiões promotoras dos cinco prováveis genes alvo do fator TGA em plantas transgênicas. Para tal propomos a geração de construções promotor::repórter através de clonagens Gateway e a análise de plantas transgênicas expressando essas construções.

2. A região promotora de *NIMIN-1* contém dois motivos TGACG que ficam separados por 70 pares de base. Um dos objetivos deste trabalho consistiu em desenvolver estudos funcionais para avaliar a mutação em um destes motivos na indução do gene *NIMIN-1*, situado a -360 pares de base do sítio de transcrição inicial de *NIMIN-1*. Com esta finalidade foi feita a quantificação da expressão do gene repórter *GUS* sob o domínio das regiões promotoras selvagem e mutante através de PCR em tempo real (qRT-PCR).

### 3. MATERIAIS E MÉTODOS

#### 3.1. Material Vegetal

Plantas de *Arabidopsis* do tipo selvagem (ecótipo Columbia-0) foram crescidas em meio de nutrientes Murashige e Skoog com 0.8% de Agar, a uma temperatura de 22°C e um fotoperíodo de 16 horas de luz para crescimento de longo período e oito horas de luz para crescimento de curto período. Plantas utilizadas para ensaios de expressão transiente foram crescidas diretamente de potes contendo solo em câmaras de crescimento sob condições de curto período por quatro a seis semanas. As mesmas plantas foram tratadas com 1 mM de ácido salicílico (Sigma), coletadas após duas horas e armazenadas a -80°C para posterior extração de RNA.

#### 3.2 Clonagem Gateway das regiões promotoras de cinco genes de *Arabidopsis*

Amplificação das regiões promotoras dos cinco genes alvo contendo o motivo de ligação de TGA2 (Tabela 1) com os sítios *attB*, bem como as reações LR e BP foram feitas de acordo com os protocolos presentes no manual de clonagem Gateway (Gateway Cloning Manual, Invitrogen, 2003). Os critérios usados para selecionar os genes alvo para uma validação funcional de sua região promotora foram: genes contendo o motivo de ligação do fator TGA2 próximo do sítio de início da transcrição e genes que apresentem indução a SA, dos quais *NIMIN-1* apresenta a maior indução, aumentando a chance de estarem envolvidos em SAR.

Tabela 1: Regiões promotoras de cinco genes candidatos clonados

Gene	Região clonada	Motivo	Localização
<i>At2g14610</i> (PR-1)	-914 a +7	TGACG	-677, -694
<i>At1g02450</i> (Nimin-1)	-805 a +7	TGACG	-360, -428
<i>At1g03850</i> (Glutaredoxina)	-1388 a +7	TGACGTCA	-1281
<i>At1g76600</i> (Proteína Expressa)	-990 a +7	TGACGTCA	-132
<i>At5g63790</i> (NAM)	-727 a +7	TGACGT	-353

Os promotores mutagenizados foram obtidos dos promotores nativos através de “site-directed mutagenesis” usando o kit “QuickChange Site Directed Mutagenesis kit” (Stratagene). Os primers usados para amplificação tanto dos promotores nativos quanto dos mutagenizados incluem os dois primeiros aminoácidos de cada gene, região 5’UTR, fragmento da região promotora e a sequência do sítio *attB* para a reação BP dos genes listados na tabela 2.

Tabela 2: Oligonucleotídeos usados para as clonagens Gateway. Nucleotídeos em azul: Região promotora; Nucleotídeos em vermelho: CDS do gene; nucleotídeos em preto: sítio de recombinação *attB*.

Gene	Descrição	Sequência
<i>Nimin-1</i>	AttB1-for	GGGGACAAGTTTGTACAAAAAAGCAGGCTGGAATCCAATTGTTCCACAC
<i>Nimin-1</i>	AttB2-rev	GGGGACCACTTTGTACAAGAAAGCTGGGTAGGATACATTTAGAGAAAGTGATTGAT
<i>PR-1</i>	AttB1-for	GGGGACAAGTTTGTACAAAAAAGCAGGCTGATATACGAAGGCGGTACAAAA
<i>PR-1</i>	AttB2-rev	GGGGACCACTTTGTACAAGAAAGCTGGGTAATTCATTTTCTAAGTTGATAATGGTT
<i>Glutaredoxina</i>	AttB1-for	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAGCAACATGGACCCTTTGT
<i>Glutaredoxina</i>	AttB2-rev	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTCATAATTGATGATAGAGAGACAAA
<i>Prot. Exp.</i>	AttB1-for	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTTTGCTACCTAAACCCGTGAAT
<i>Prot. Exp.</i>	AttB2-rev	GGGGACCACTTTGTACAAGAAAGCTGGGTGACCCATTTGATTTTGAAGTCTC
<i>NAM</i>	AttB1-for	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGTATTCTTCTCCAGTCAATGAAAA
<i>NAM</i>	AttB2-rev	GGGGACCACTTTGTACAAGAAAGCTGGGTAGTCCATGGAGCTTTTATAAACGAA

As regiões promotoras amplificadas com os sítios *attB* foram introduzidas no vetor de entrada pDONR207 (Invitrogen) através da reação Gateway BP e então transferidas para os vetores destino binários pYXT1 e pYXT2 (Xiao et al., 2005) via a reação Gateway LR. Especificamente, a amplificação por PCR foi feita com usando a Taq polimerase de alta fidelidade “Hot Start Phusion DNA polymerase” (New England Biolabs). Os produtos da reação BP foram então transformados em células competentes TOP10 (Invitrogen). Os insertos foram validados por seqüenciamento com os oligos direto 5’-TCGCGTTAACGCTAGCATGGATCTC-3’ e reverso 5’-GTAACATCAGAGATTTTGAGACAC-3’ do vetor pDONR207 (Figura 4). As construções resultantes (Tabela 3) contem as regiões promotoras, 5’UTR e os dois primeiros aminoácidos do gene nativo como uma fusão em fase para cada construção

repórter conectadas via o peptídeo *attB* codificado no sistema Gateway. Nesta mesma tabela podemos observar as construções pWTNIM::GFP e pWTNIM::GUS, contendo o promotor nativo do gene *NIMIN-1* fusionado a região codante de GFP e GUS respectivamente, enquanto as construções pMUTNIM::GUS contem o promotor mutagenizado de *NIMIN-1* fusionado ao gene repórter GUS por exemplo e assim sucessivamente para os outros quatro genes.

Tabela 3: Lista de construções promotoras::repórter inseridas em *Agrobacterium* geradas neste trabalho.

<b>gene</b>	<b>promotor</b>	<b>gene repórter</b>	<b>construção gerada</b>
<i>NIMIN-1</i>	selvagem	GFP	1- pWTNIM::GFP
	selvagem	GUS	2- pWTNIM::GUS
	mutante	GFP	3- pMUTNIM::GFP
	mutante	GUS	4- pMUTNIM::GUS
<i>PR-1</i>	selvagem	GFP	5- pWTPR1::GFP
	selvagem	GUS	6- pWTPR1::GUS
	mutante	GFP	7- pMUTPR1::GFP
	mutante	GUS	8- pMUTPR1::GUS
<i>Glutarredoxina</i>	selvagem	GFP	9- pWTGlut::GFP
	selvagem	GUS	10- pWTGlut::GUS
	mutante	GFP	11- pMUTGlut::GFP
	mutante	GUS	12- pMUTGlut::GUS
<i>NAM</i>	selvagem	GFP	13- pWTNAM::GFP
	selvagem	GUS	14- pWTNAM::GUS
	mutante	GFP	15- pMUTNAM::GFP
	mutante	GUS	16- pMUTNAM::GUS
<i>EP</i>	selvagem	GFP	17- pWTEP::GFP
	selvagem	GUS	18- pWTEP::GUS
	mutante	GFP	19- pMUTEp::GFP
	mutante	GUS	20- pMUTEp::GUS

O vetor pYXT1 foi construído a partir do vetor pBI121 e modificado para compatibilidade com o sistema Gateway (Invitrogen). Já o vetor Gateway-compatível pYXT2 foi montado a partir do vetor binário pET-15GAL4-VP16UASmGFP5ER (Bougourd et al., 2000). Um esquema das construções repórter das regiões promotoras gerados pode ser visualizado na Figura 5.

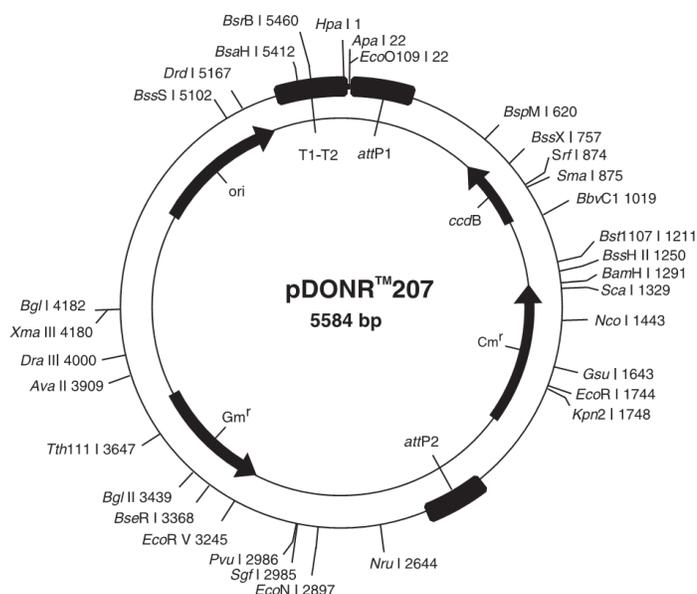


Figura 4: Mapa do vetor de entrada Gateway (“entry vector”) pDONR207 usado neste trabalho.

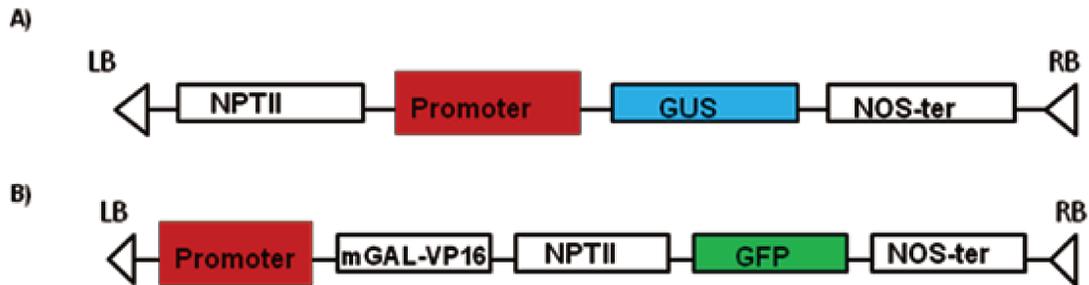


Figura 5: Esquema dos vetores destino Gateway (“destination vector”) usados neste trabalho para geração das construções promotor::repórter: A) pYXT1 e B) pYXT2. LB: borda esquerda; RB: Borda direita. NOS-ter; Terminador NOS; NPTII: gene de resistência a canamicina; GUS:  $\beta$ -glucuronidase; GFP: Proteína verde fluorescente.

### 3.3. Transformação de *Agrobacterium* por conjugação tri-parental

Os plasmídeos recombinantes pYXT1 e pYXT2 (Xiao et al., 2005) usados neste trabalho contendo as diferentes regiões promotoras foram transformados em células competentes *E. coli* TOP10 através do método de choque térmico depois da reação Gateway LR. Os plasmídeos recombinantes foram então transferidos para linhagem GV3101 de *Agrobacterium tumefaciens* por conjugação tri-parental. Linhagem GV3101 foi crescida overnight em meio LB líquido contendo 50 mg/mL rifampicina e 50 mg/mL gentamicina a uma temperatura de 28°C em um shaker a 250 rpm. Culturas de *E. coli* contendo o plasmídeo de interesse e *E. coli* helper strain (pRK 2013) também foram crescidos overnight em frascos separados com meio LB contendo 50 mg/mL de canamicina a 37°C a uma agitação de 250 rpm. No dia seguinte, 50  $\mu$ L de cada uma das três culturas foram plaqueadas em uma placa de Petri contendo meio LB Agar para a conjugação e incubadas durante dois dias a temperatura de 28 °C para produção de colônias. Para

confirmar a transformação foram feitas análises por PCR com primers específicos para os genes.

### **3.4. Transformação estável de Arabidopsis**

A transformação de Arabidopsis com as construções contendo as regiões promotoras dos cinco genes fusionados ao gene repórter foi feito de acordo com o método descrito por Clough e Bent (1998). Cepas de *A. tumefaciens* contendo as construções promotoras repórter foram crescidas em meio LB contendo o meio apropriado de seleção (rifampicina, gentamicina e canamicina a 50µg/mL cada um) por dois dias a uma temperatura de 28°C. A cultura foi então transferida para um tubo de plástico do tipo Falcon de 50 mL e centrifugado a 6000 rpm, 4°C por 5 minutos. O pellet foi ressuspensionado em uma solução de *Agrobacterium* (5% sacarose; 0,05% surfactante Silwet L-77). As inflorescências de plantas de Arabidopsis crescidas previamente em câmaras de crescimento (18 horas de luz) foram imersas na solução de *Agrobacterium* por aproximadamente 10 segundos. As plantas foram cobertas com um filme transparente e armazenadas nas câmaras de crescimento para geração de sementes. Sementes T0 foram coletadas e os transformantes selecionados em meio Murashige e Skoog contendo 50 mg/mL de canamicina e então transferidos para solo para obtenção de sementes T1.

### **3.5. Ensaio de expressão transiente por agroinfiltração**

Folhas de Arabidopsis tratadas e não tratadas com ácido salicílico foram Agroinfiltradas com uma seringa sem agulha contendo a solução de *Agrobacterium* carregando as construções pMUTNIM::GUS (promoter mutagenizado) e pWTNIM::GUS (promoter nativo) de acordo com o método descrito por Wroblewski et al. (2005) com

modificações. Depois de infiltração com as construções promotoras repórter, as plantas foram colocadas de volta na câmara de crescimento por um mínimo de dois dias para permitir a expressão das construções repórter. Folhas foram coradas com solução corante de GUS, contendo X-gluc, para inspeção visual. Culturas de *Agrobacterium* foram crescidas overnight em meio YEP líquido contendo o antibiótico para seleção a 29°C em um agitador. A cultura foi ajustada a um O.D.<sub>600</sub> = 0.3-0.6 antes da infiltração. Foram feitos ensaios histoquímicos para detecção de atividade de GUS a partir do terceiro dia pós infiltração (dpi). Amostras de tecido foliar foram incubados com solução corante de GUS (0,1 M X-gluc, 0,1 M fosfato de sódio (pH 7,0), 0,01 M EDTA e 0,1% (v/v) Triton X-100 ) durante a noite a 37°C. No dia seguinte as amostras foram lavadas cinco vezes com etanol 70% para remover a clorofila e preservar as amostras. Folhas também foram coletadas no terceiro dia pós-infiltração (dpi) e guardadas a -80°C para análises de quantificação de expressão gênica por qRT-PCR.

### **3.6. Extração de RNA e síntese de cDNA**

O RNA total foi isolado a partir de 50 mg de folhas de *Arabidopsis* maceradas em nitrogênio líquido usando o Kit de extração de RNA “RNeasy Plant Mini Kit” (Qiagen) com tratamento adicional de DNase (Qiagen) nas amostras de RNA para qRT-PCR. Concentração das amostras de RNA extraídas foram medidas em um espectrofotômetro NanoDrop ND-1000. A pureza do RNA verificado na absorbância A260/A280 foi acima de 1.9. A integridade do RNA foi avaliada em um gel de agarose a 1% (p/v) após extração do RNA. Cada amostra de cDNA foi sintetizada a partir de 2 µg de RNA usando a enzima SuperScript™ III (Invitrogen) de acordo com protocolo do fabricante.

### 3.7. PCR quantitativo em tempo real

Sequências de oligonucleotídeos específicos para os genes de *GUS*, *NPTII*, *NIMIN-1* e dos controles invariantes para os genes endógenos (*AT4G26410*, *AT5G15710*, *AT4G27960*, *AT5G46630*) usados para o PCR quantitativo em tempo real (qRT-PCR) estão listadas na Tabela 4. Todas as sequências de oligonucleotídeos usados neste trabalho foram comparados com o banco de dados de cDNA da TAIR para *A. thaliana* utilizando o algoritmo BlastN (Altschul et al., 1990) para confirmar a especificidade dos amplicons e evitar assim ampliações não específicas. Amplicons pequenos foram usados (100-150 pb) para aumentar a performance do experimento (Marino et al., 2003).

Tabela 4: Lista de oligonucleotídeos usados para PCR em tempo real (qRT-PCR)

Gene	Direto	Reverso
<i>GUS</i>	TAATGTTCTGCGACGCTCAC	TTCTCTGCCGTTTCCAAT
<i>NPTII</i>	CGTTGGCTACCCGTGATATT	CTCGTCAAGAAGGCGATAGAA
<i>NIMIN-1</i>	ATCTAACGGCGGAGAAAGGT	TGTGATCCGAAACAAACATCA
<i>At4g26410</i>	CCTGGAAGGGATGCTATCAA	GTCCGACATACCCATGATCC
<i>AT5G15710</i>	GCACTTCTGAGACTTTCGGC	ATGACTGAAGAGCACAACCG
<i>AT4G27960</i>	TAACCATCCATTTCCCTCCA	TGGAAATTGTGAGAGCAGGA
<i>AT5G46630</i>	GTGCCAATGTTACAGCATC	TGATCTCGTAAGATCCCCT

Reações de qRT-PCR foram feitas em um instrumento ABI PRISM® 7900HT (Applied Biosystems), usando SYBR® Green para monitorar síntese de DNA dupla-fita (dsDNA) de acordo com os seguintes parâmetros: 2 min. a 50°C, 10 min. a 95°C, 40 ciclos de 15 seg. a 95°C, 1 min. a 60°C. A curva de dissociação (Melting curve) gerada (95°C por 15 seg., 60°C durante 30 seg.) confirmou a presença de um único pico de produto de PCR

e não houve formação de dímeros de primers. Amostras com água substituindo o cDNA, denominadas de “NTC” (“no template control”), foram incluídas em cada corrida como controle para possíveis falsos positivos ou contaminação. O volume final da reação foi de 10 µL. Um “master mix” contendo 85 µl de 2x SYBR<sup>®</sup> Green (Applied Biosystems) e 45 µl de amostras de cDNA previamente diluídas foram preparados antes da distribuição na placa de qRT-PCR para reduzir erros de pipetagem. Uma multi-pipeta eletrônica automática foi usada para dispensar o master mix primeiro (7,5 µl por poço) e então 2,4 µl de oligos foram dispensados com o uso de uma pipeta de doze canais (Eppendorf). Os dados foram analisados usando o software SDS 2.0 (Applied Biosystems). O valor padrão do “threshold” utilizado foi de 0,1 para as três réplicas biológicas. A quantificação relativa dos genes analisados foi calculada usando o método  $\Delta\Delta C_T$  (Livak e Schmittgen, 2001; Czechowski et al., 2004) através da subtração dos valores  $C_T$  de genes alvo tratados com SA ( $T_2$ ) e não-tratados ( $T_1$ ), normalizados aos valores  $C_T$  dos controles invariantes ou genes “housekeeping”, estes também tratados com SA ( $C_2$ ) e não tratado ( $C_1$ ):  $\Delta\Delta C_t = ( T_{1(sa-)} - T_{2(sa+)} ) - ( C_{1(sa-)} - C_{2(sa+)} )$ . Este valor calculado é o logaritmo na base 2 dos valores de expressão relativos.

## 4. RESULTADOS E DISCUSSÃO

### 4.1. Padrão de expressão de quatro genes candidatos

Cada um dos quatro genes analisados apresentou um padrão de expressão por GFP específico, em diferentes tecidos e em diferentes etapas de desenvolvimento em *Arabidopsis*, sob o controle das regiões promotoras selvagem para cada um dos quatro genes analisados (Figuras 6-9). Plantas transgênicas expressando a construção pWTNAM::GFP (At5g63790) apresentaram expressão de GFP sob o controle da região promotora selvagem no tecido vascular foliar, hidátodos, nas sépalas e antera (Figura 6). Esse gene foi denominado *NAM* (“no apical meristem”) pelo fato de seu mutante não apresentar meristema apical (Aida et al., 1997). Posteriormente foi visto que esse gene codifica um fator de transcrição da família NAC, envolvido em processos de desenvolvimento do embrião, dos órgãos da flor e também está relacionado a defesa (Duval et al., 2002).

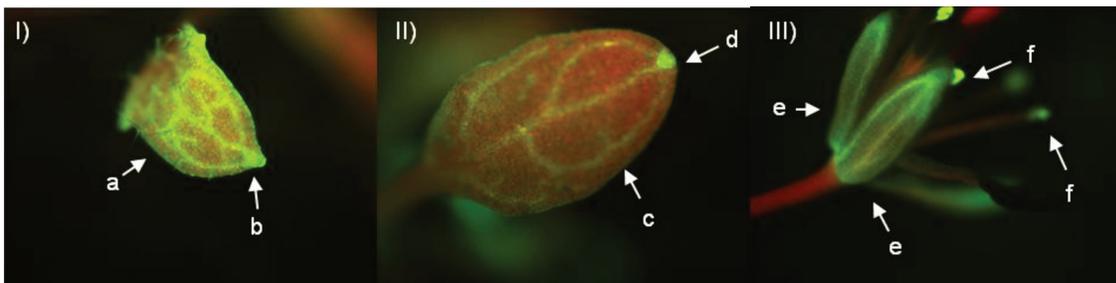


Figura 6: Padrão de expressão de GFP sob o controle da região promotora do gene At5g63790 (*NAM*). Plantas transgênicas de *Arabidopsis* com a construção pWTNAM::GFP foram produzidas. A expressão do gene repórter GFP foi avaliada em lupa sob luz ultravioleta. I e II) Expressão por GFP em: (a), (c) tecido vascular foliar; (b), (d) hidátodo. III) Expressão por GFP em: (e) sépala; (f) antera.

Plantas transgênicas expressando a construção pWTEP::GFP (At1g76600) apresentaram expressão por GFP na raiz e folhas de uma roseta jovem e antera de planta em estágio adulto (Figura 7). Este gene possui função ainda desconhecida.

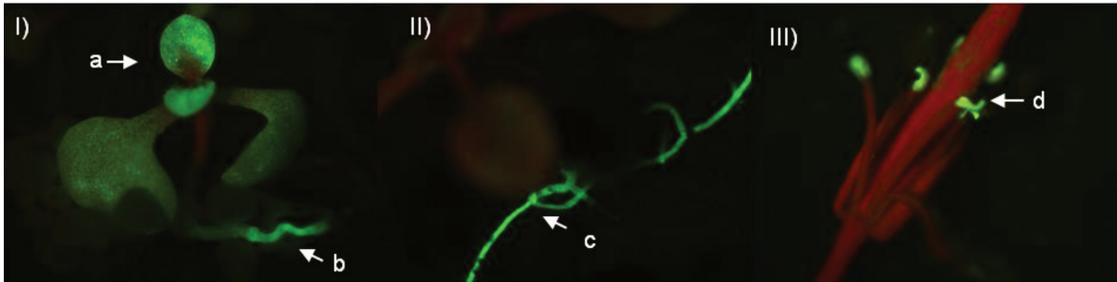


Figura 7: Padrão de expressão de GFP sob o controle da região promotora do gene At1g76600 (*Expressed protein* ou EP). Plantas transgênicas expressando a construção pWTEP::GFP foram produzidas e expressão por GFP observados em: I e II) (a) Folhas da roseta jovem; (b), (c) na raiz. III) (d) antera.

Expressão de GFP sob o controle da região promotora selvagem do gene *PR-1* (At2g14610) pode ser visualizada nos nódulos caulinares, filete e antera do estame de planta no estágio adulto. O gene *PR-1* está envolvido na defesa da planta e apresenta expressão induzida em resposta a SA e ataque por patógeno.

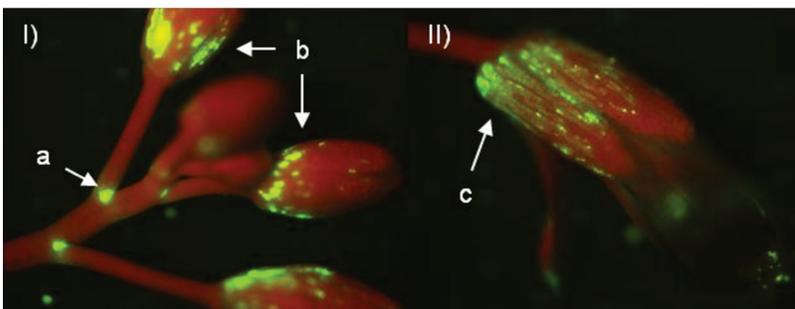


Figura 8: Padrão de expressão de GFP sob o controle da região promotora do gene At2g14610 (*PR-1*) em plantas transgênicas de *Arabidopsis* expressando a construção pWTPR1::GFP: I e II) Expressão por GFP em (a) nódulos; (b), (c) filete e antera do estame.

A expressão do gene *At1g03850*, que codifica uma proteína glutarredoxina envolvida em resposta a citoquininas apresentou padrão de expressão por GFP na raiz da plântula com cotilédone recém formado, sépala e pedículo (Figura 9).

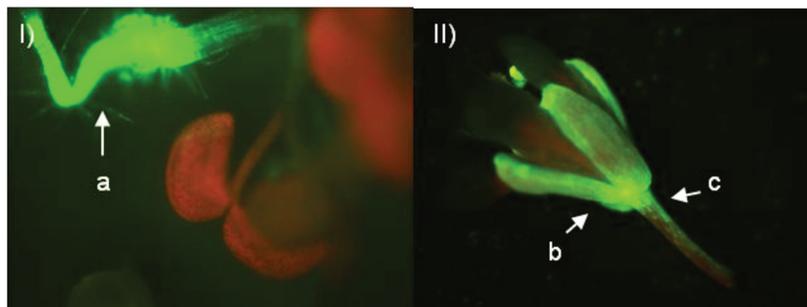


Figura 9: Padrão de expressão de GFP sob o domínio da região promotora do gene *At1g03850* (Glutarredoxina) em plantas transgênicas expressando a construção pWTGLU::GFP: I) Expressão por GFP na raiz (a) da plântula com cotilédone recém formado. II) Expressão por GFP na sépala (b) e pedículo (c).

#### 4.2 Análise funcional da região promotora de *NIMIN-1*

Uma análise funcional foi feita com o objetivo de verificar se a mutação gerada no sítio de ligação no fator TGA2 (TGACG, localizado -360 pb acima do sítio de início da transcrição do gene *NIMIN-1*) era essencial para expressão induzida por SA de *NIMIN-1*. Para tanto, dois pares de base foram alterados do motivo de ligação do fator TGA2 (TGACG) através de mutação sítio-dirigida. Esta versão mutagenizada da região promotora de *NIMIN-1* foi sequenciada de modo a confirmar as alterações feitas. O alinhamento das sequências nativa e mutagenizada da região promotora de *NIMIN-1* foi feito usando o programa de alinhamento ClustalW (Figura 10).

**pWTNIM: 5' ACACATCATCATCTGACGTAATTT 3'**  
**pMUTNIM: 5' ACACATCATCATCTGCTGTAATTT 3'**

Figura 10: Comparação das regiões selvagem e mutante do gene *NIMIN-1*. pWTNIMIN: região promotora selvagem de *NIMIN-1*; pMUTNIM: região promotora mutagenizada. A região de ligação do fator TGA2 está sublinhada. . Uma alteração de dois pares de base (em vermelho) foi gerada por mutação sítio-dirigida do sítio de ligação do motivo TGA2 localizado -360 pb acima do sítio de iniciação da transcrição do gene *NIMIN-1*.

As regiões promotoras mutagenizada e nativa do gene *NIMIN-1* que abrangem desde a posição +7 da CDS do gene até -805 pb à montante do sítio de início da transcrição de *NIMIN-1* foram isolados por PCR a partir dos vetores contendo os promotores nativo e mutagenizado. Através da utilização do sistema Gateway as regiões promotoras amplificadas para *NIMIN-1* foram clonadas no vetor de entrada pDONR207 (reação BP) e então inseridas nos vetores binários repórteres pYXT1 (GUS) e pYXT2 (GFP) através da reação Gateway LR. Os produtos da reação BP foram transformados em *E. coli* e a clonagem dos insertos foi validada através de sequenciamento. Clones positivos das reações LR foram confirmados através de PCR de colônia, com primers específicos das regiões promotoras, a partir de colônias transformadas de *E. coli* resultantes da reação Gateway LR. As construções foram então introduzidas em *Agrobacterium* através de conjugação tri-parental para posterior transformação de *Arabidopsis*. No total foram geradas quatro construções para análise da região promotora de *NIMIN-1*: As duas regiões promotoras nativas sendo pWTNIM::GFP e pWTNIM::GUS, e duas regiões promotora mutagenizadas pMUTNIM::GUS e pMUTNIM::GFP.

Foram geradas mais de cinquenta plantas transgênicas de *Arabidopsis* expressando a construção pWTNIM::GFP e plantas T1 foram observadas no microscópio para visualização do padrão de expressão do gene repórter sob controle da região promotora nativa de *NIMIN-1* (Figura 11). As linhagens segregantes expressando a construção pWTNIM::GFP apresentaram expressão de GFP no meristema apical, raiz, nervura principal da folha, nervuras secundárias da folha, hidátodo, tricoma e no pecíolo.

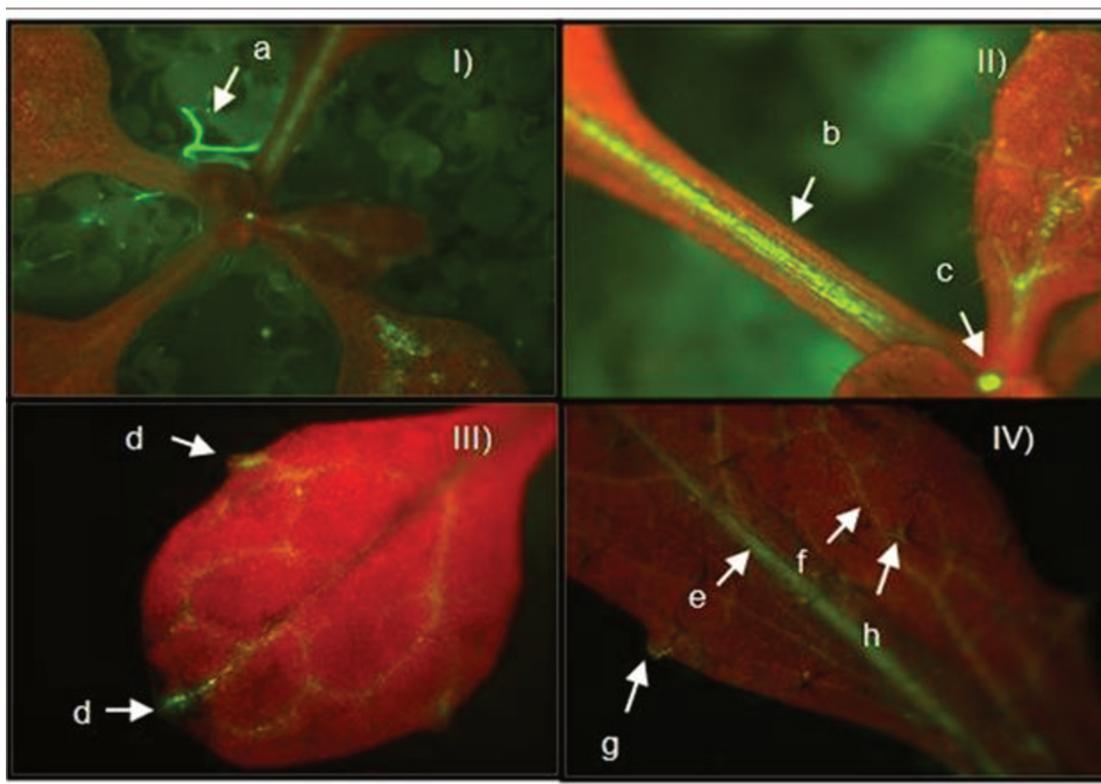


Figura 11: Padrão de expressão de linhagens de plantas transgênicas T1 em *Arabidopsis* contendo a construção pWTNIM::GFP. I) Expressão por GFP na raiz (a) de uma roseta de três semanas; II) Expressão por GFP no pecíolo (b) e meristema apical (c); III) Expressão por GFP nas nervuras primária e secundária e hidátodos (d) de uma folha de uma roseta de três semanas; IV) Expressão por GFP na veia média (e), veia secundária (f), hidátodo (g) e tricomas (h) de uma planta de seis semanas.

O padrão de expressão de GFP observado em plantas transgênicas expressando a construção pWTNIM::GFP está de acordo com os dados de expressão do gene endógeno

*NIMIN-1* do banco de dados de Arabidopsis eFP Browser tool (Winter et al., 2007). Este banco de dados possui dados de expressão do gene endógeno *NIMIN-1* a partir de experimentos de microarranjos de DNA dos chips ATH1 da Affymetrix.

Para verificar se o motivo de ligação do fator de transcrição TGA2 localizado na região promotora de *NIMIN-1* está envolvido na regulação da transcrição de *NIMIN-1* em resposta a SA foi feito uma análise funcional comparativa das regiões promotoras nativa e mutagenizada do gene *NIMIN-1*. Nós quantificamos o nível de expressão gênica por qRT-PCR dos promotores nativo e mutagenizado de *NIMIN-1* em relação ao gene endógeno de *NIMIN-1* usando para tanto as regiões promotoras descritas fusionadas ao gene repórter GUS.

O PCR quantitativo em tempo real foi o método escolhido para medir expressão de *NIMIN-1* induzida por SA uma vez que o objetivo era obter uma estimativa da resposta a nível transcricional em resposta rápida a SA (2 horas). De fato, apesar de revelar regiões de alta expressão no tecido da planta e confirmar que está ocorrendo expressão do gene repórter sob a região promotora sob seu controle, corar plantas com GUS não é um método quantitativo *per se* (Cho e Cosgrove., 2000). Portanto a expressão do gene repórter foi medida por qRT-PCR a partir de folhas agroinfiltradas no terceiro dia pós-infiltração (dpi), de maneira a permitir a expressão do gene repórter (Figuras 12 e 13).

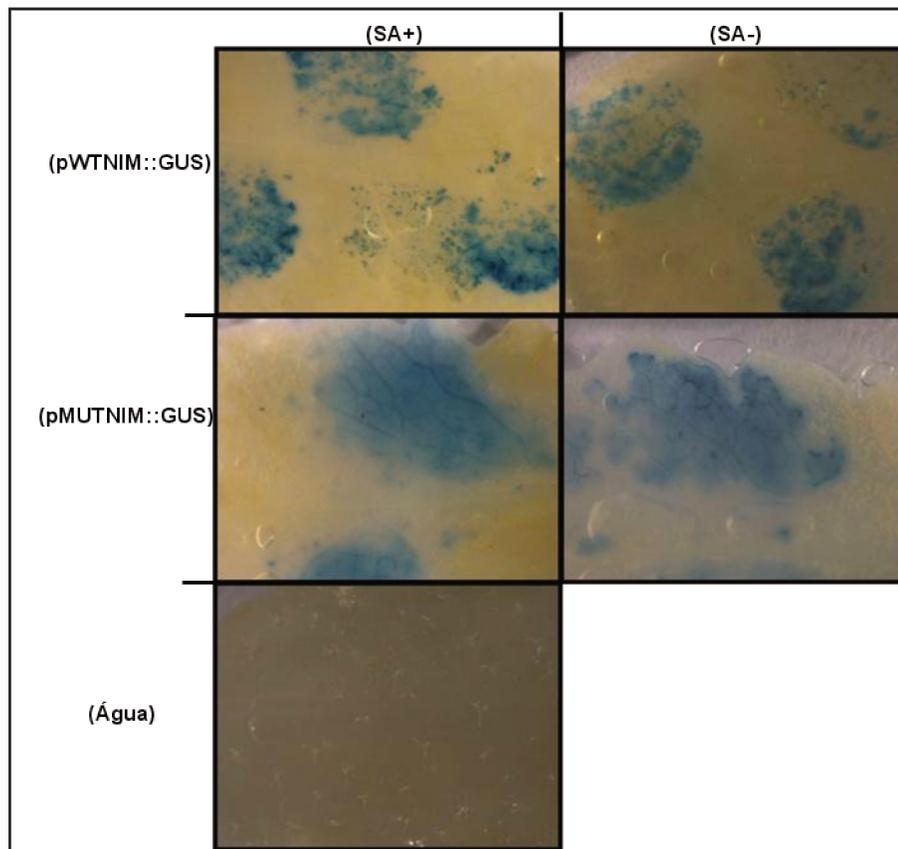


Figura 12: Ensaio histoquímico mostrando expressão de GUS em folhas de *Arabidopsis* agroinfiltradas com as construções promotoras selvagem (pWTNIM::GUS) e mutante (pMUTNIM::GUS) para o gene *NIMIN-1*. A expressão do gene repórter ocorreu nos sítios de infiltração.

De maneira a quantificar a resposta de ambas as regiões promotoras de *NIMIN-1* nativa e mutagenizada a indução por SA foram medidos os níveis dos transcritos das construções (pWTNIM::GUS e pMUTNIM::GUS) em folhas tratadas com 1 mM de SA por duas horas (SA+) e folhas não tratadas (SA-). Valores  $C_T$  para o gene endógeno *NIMIN-1* foram normalizados em relação aos valores  $C_T$  do gene controle (invariante) *At4g26410*, que foi o gene invariante mais constante de todos os quatro genes controle inclusos em cada corrida por qRT-PCR. O valor  $C_T$  médio para *At4g26410* foi de 20.44 ( $\pm$

0.32) em um “threshold” de 0.1 para todas as três réplicas biológicas. Valores  $C_T$  para as construções promotoras-repóter agroinfiltradas WTNIM::GUS e pMUTNIM::GUS foram normalizados em relação ao genes invariante *NPTII*. O valor  $C_T$  médio para *NPTII* foi de 26.46 ( $\pm 0.34$ ) para as três réplicas biológicas. A mudança nos níveis de expressão observados na resposta rápida (2h) a ácido salicílico foi calculada de acordo com o método  $\Delta\Delta C_T$  (Livak e Schmittgen, 2001; Czechowski et al., 2004).

Nós observamos que tanto a construção pWTNIM::GUS, contendo o promotor nativo quanto a cópia endógena do gene *NIMIN-1* apresentaram uma alta resposta a SA (Figura 13), de acordo com o que foi demonstrado em trabalhos anteriores (Weigel et al., 2001; Glocova et al., 2005). As análises por qRT-PCR mostraram que a mutação no motivo de ligação do fator TGA2 causou uma redução para apenas um terço da expressão de GUS sob o controle do promotor mutagenizado em relação ao promotor nativo, comprovando que o promotor mutante é menos sensível a indução por SA.

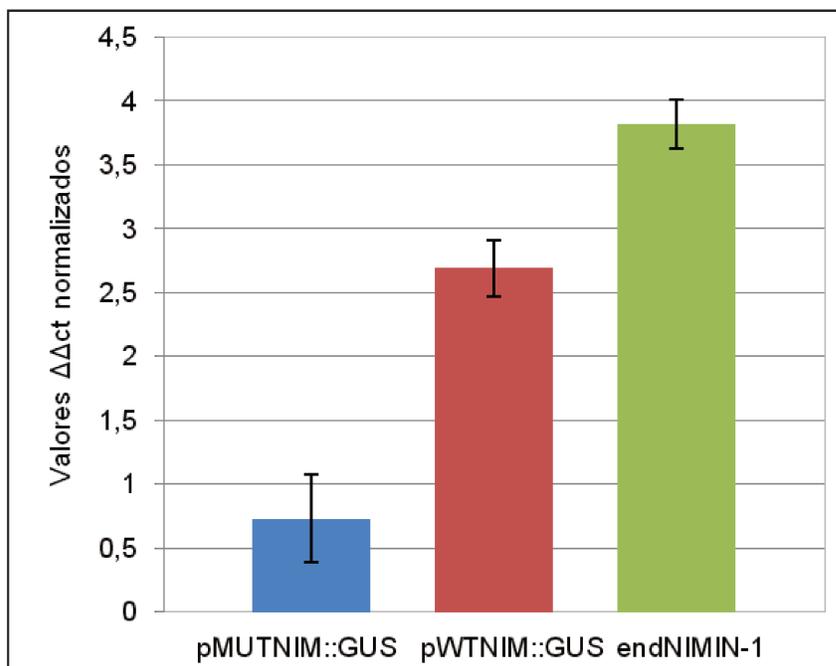


Figura 13: Efeito da mutação do sítio de ligação de TGA2 na região promotora do gene *NIMIN-1*. Barras representam mudança na expressão das regiões promotoras selvagem (pWTNIM::GUS), mutante (pMUTNIM::GUS) e do gene endógeno (endNIMIN-1) normalizados em relação aos controles internos (*NPTII* e *At5g15710*) em amostras tratadas com SA vs. não tratadas. Dados representam a média  $\pm$ DP para três réplicas biológicas independentes. GUS =  $\beta$ -glucuronidase.

A diferença observada na expressão do gene repórter *GUS* sob o controle das regiões promotoras nativa e mutagenizada comprova que a ligação do fator TGA2 ao motivo TGACG é relevante para a resposta de *NIMIN-1* a SA. É importante destacar que a fraca indução do promotor mutagenizado poderia indicar que a mutação nos dois pares de base apenas não seria suficiente para cancelar completamente a ligação do fator TGA2 ou talvez a presença do outro motivo TGACG na região promotora talvez tenha algum tipo de papel na mediação da indução de *NIMIN-1* por SA.

Em suma, nossos dados relacionados à análise funcional das regiões promotoras mutagenizada e selvagem de *NIMIN-1* comprovam que o motivo TGACG está diretamente envolvido na regulação da expressão gênica induzida por SA do gene *NIMIN-1*.

## 5. CONCLUSÕES

- O quatro genes que possuem o sítio de ligação do fator TGA2 apresentaram perfis de expressão distintos, cabendo desatacar: *At5g63790* (*NAM*): expressão por GFP no tecido vascular foliar, hidátodo, sépala e antera; *At1g76600* (*expressed protein*) apresentou expressão por GFP nos cotilédones da roseta jovem, raiz e antera; *At2g14610* (*PR-1*) apresentou expressão por GFP nos nódulos caulinares, no filete e na antera; *At1g03850* (glutarredoxina) apresentou expressão por GFP na raiz da plântula com cotilédones recém formado, sépala e pedículo.
- A expressão dirigida pela região promotora selvagem de *NIMIN-1* por GFP ocorre no meristema apical, raiz, pedículo, hidátodo, tecido vascular e tricomas de *Arabidopsis*.
- Uma análise funcional através de mutação no motivo de ligação do fator TGA2 mostrou uma redução de mais de três vezes na expressão de GUS sob o controle do promotor mutagenizado em relação ao promotor nativo, comprovando que o promotor mutante é menos sensível a indução por SA. Portanto, o motivo de ligação do fator TGA2 na região promotora de *NIMIN-1* está diretamente relacionado a modulação da expressão de *NIMIN-1* induzido por SA.

## 6. REFERÊNCIAS

- Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *Plant Cell*. 9: 841–857.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol*. 215: 403-410.
- Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP (1997). Signaling in plant-microbe interactions. *Science*. 276: 726-733.
- Bray EA, Bailey-Serres J, Weretilnyk E (2000). Responses to abiotic stresses. In: Gruissem W, Buchannan B, Jones R (Eds.). *Biochemistry and molecular biology of plants*. American society of plant biologists, Rockville, 2000. pp 1158-1249.
- Cao H, Li X, Dong XN (1998). Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. *Proc Natl Acad Sci USA*. 95: 6531-6536.
- Cho HT, Cosgrove DJ (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA*. 97: 9783-9788.
- Chuang CF, Running MP, Williams RW, Meyerowitz E (1999). The *Perianthia* gene encodes a bZIP protein involved in the determination of floral organ number in *Arabidopsis thaliana*. *Genes Dev*. 13: 334-344.
- Clough SJ, Bent AF (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J*. 16: 735-43.
- Czechowski T, Bari RP, Stitt M, Scheible WR, Udvardi MK (2004). Real-time RT-PCR profiling of over 1400 *Arabidopsis* transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J*. 38: 366-379.

Deppmann CD, Acharya A, Rishi V, Wobbes B, Smeekens S, Taparowsky EJ, Vinson C (2004). Dimerization specificity of all 67 B-ZIP motifs in *Arabidopsis thaliana*: a comparison to *Homo sapiens* B-ZIP motifs. *Nucleic Acids Res.* 32: 3435–3445.

Despres C, DeLong C, Glaze S, Liu E, Fobert PR (2000). The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. *Plant Cell.* 12: 279-290.

Durrant WE, Dong X (2004). Systemic acquired resistance. *Annu Rev Phyto.* 42: 185-209.

Duval M, Hsieh TF, Kim SY, Thomas TL (2002). Molecular characterization of AtNAM: A member of the Arabidopsis NAC domain superfamily. *Plant Mol Biol.* 50: 237–248.

Ellenberg TE, Brandl CJ, Struhl K, Harrison SC (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted helices: crystal structure of the protein–DNA complex. *Cell.* 71: 1223–1237.

Finkelstein RR, Lynch TJ (2000). The Arabidopsis abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *Plant Cell.* 12: 599-609.

Flowers TJ (2004). Improving crop salt tolerance. *J Exp Bot.* 55: 307-319.

Glazebrook J, Rogers EE, Ausubel FM (1996). Isolation of Arabidopsis mutants with enhanced disease susceptibility by direct screening. *Genetics.* 143: 973-982.

Glocova I, Thor K, Roth B, Babbick M, Pfitzner AJP, Pfitzner UM (2005). Salicylic acid (SA)-dependent gene activation can be uncoupled from cell death-mediated gene activation: the SA-inducible NIMIN-1 and NIMIN-2 promoters, unlike the PR-1a promoter, do not respond to cell death signals in tobacco. *Mol Plant Pathol.* 6: 299-314.

Hudson ME, Quail PH (2003). Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. *Plant Physiol* 133: 1605-1616.

Hurst HC. Leucine Zippers Transcription Factors. San Diego, Academic Press, 1996. p.72.

Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F (2002). bZIP transcription factors in Arabidopsis. *Trends Plant Sci.* 7: 106-111.

- Johnson C, Boden E, Arias J (2003). Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. *Plant Cell*. 15: 1846-1858.
- Katagiri F, Lam E, Chua NH (1989). Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. *Nature*. 340: 727-730.
- Kinkema M, Fan W, Dong X (2000). Nuclear localization of NPR1 is required for activation of PR gene expression. *Plant Cell*. 12: 2339-2350.
- König P, Richmond TJ (1993). The X-ray structure of the CGN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. *J Mol Biol*. 233: 139-154.
- Lam E, Benfey PN, Gilmartin PM, Fang RX, Chua NH (1989). Site-specific mutations alter in vitro factor binding and change promoter expression pattern in transgenic plants. *Proc Natl Acad Sci USA*. 86: 7890-7894.
- Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E (1998). Functional analysis of regulatory sequences controlling PR-1 gene expression in Arabidopsis. *Plant J*. 16: 223-233.
- Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CR, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne J, Volkert TL, Fraenkel E, Gifford DK, Young RA (2002). Transcriptional Regulatory Networks in *Saccharomyces cerevisiae*. *Science*. 298: 799-804.
- Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  method. *Methods*. 25: 402-408.
- Mahajan S, Tuteja N (2005). Cold, Salinity and drought stresses: An overview. *Arch Biochem Biophys*. 444: 139-158.
- Marino JH, Cook P, Miller KS (2003). Accurate and statistically verified quantification of relative mRNA abundances using SYBR green I and real-time RT-PCR. *J Immunol Methods*. 283: 291-306.

Mou Z, Fan W, Dong X (2003). Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell*. 113: 935-944.

Niggewed R, Thurow C, Kegler C, Gatz C (2000). Tobacco transcription factor TGA2.2 is the main component of as-1 binding factor ASF-1 and is involved in salicylic acid and auxin inducible expression of as-1-containing target promoters. *J Mol Chem*. 275: 19897-19905.

Osterlund MT, Wei N, Deng XW (2000). The roles of photoreceptor system and the COP-1 targeted destabilization of HY5 in the light control of Arabidopsis seedling development. *Plant Physiol*. 124: 1520-1524.

Miller KS (2003). Accurate and statistically verified quantification of relative mRNA abundances using SYBR Green I and real-time RT-PCR. *J Immunol Meth*. 283: 291-306.

Pandey GK (2008). Emergence of a novel calcium signaling pathway in plants: CBL-CIPK signaling network. *Physiol Mol Biol Plants*. 14: 51-68.

Riechmann JL, Heard J, Martin J, Reuber F, Jiang CZ, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandhari D, Sherman BK, Yu GL (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science*. 290: 2105-2110.

Rook F, Weisbeek P, Smekens S (1998). The light-regulated *Arabidopsis* bZIP transcription factor gene ATB2 encodes a protein with an unusually long leucine zipper domain. *Plant Mol Biol*. 37: 171-178.

Ross AF (1961). Systemic acquired resistance induced by localized virus infections in plants. *Virology*. 14: 340-358.

Ryals JA, Neuenschwander UH, Willits MG, Molina A, Steiner HY, Hunt MD (1996). Systemic acquired resistance. *Plant Cell*. 8: 1809-1819.

Schindler U, Menkens AE, Beckman H, Ecker JR, Cashmore AR (1992). Heterodimerization between light-regulated and ubiquitously expressed Arabidopsis GBF bZIP proteins. *EMBO J*. 11: 1261-1273.

Schlögl PS, Nogueira FT, Drummond RD, Felix JM, Rosa Jr VE, Leite A, Ulian EC, Menossi M (2008). Identification of new ABA- and MEJA-activated sugarcane bZIP genes by data mining in the SUCEST database. *Plant Cell Rep.* 27: 335-345

Staskawicz BJ, Ausubel FM, Baker BJ, Ellis JG, Jones JDG (1995). Molecular genetics of plant disease resistance. *Science.* 268: 661-667.

Thibaud-Nissen F, Wu HC, Richmond T, Redman JC, Johnson C, Green R, Arias J, Town CD (2006). Development of Arabidopsis whole-genome microarrays and their application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. *Plant J.* 47: 152-162.

Walsh JW, Feeling M (1997). The maize gene *liguleles 2* encodes a basic leucine zipper involved in the establishment of the leaf blade-sheath boundary. *Genes Dev.* 11: 208-218.

Ward ER, Uknes SJ, Williams SC, Dincher SS, Wiederhold DL, Alexander DC, Ahl-Goy P, Metraux JP, Ryals JA (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell.* 3: 1085-1094.

Weigel RR, Bauscher C, Pfitzner AJP, Pfitzner UM (2001). NIMIN-1, NIMIN-2 and NIMIN-3, members of a novel family of proteins from Arabidopsis that interact with NPR1/NIM1, a key regulator of systemic acquired resistance in plants. *Plant Mol. Biol.* 46: 143-160.

Weigel RR, Pfitzner UM, Gatz C (2005). Interaction of NIMIN1 with NPR1 Modulates PR Gene Expression in Arabidopsis. *Plant Cell.* 17: 1279-1291.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007). An “Electronic Fluorescent Pictograph” Browser for Exploring and Analyzing Large-scale Biological Data Sets. *PLoS ONE* 2: e718.

Wroblewski T, Tomczak A, Michelmore R (2005). Optimization of Agrobacterium mediated transient assays of gene expression in lettuce, tomato and Arabidopsis. *Plant Biotech J.* 3: 259-273.

Vinson C, Acharya A, Taparowsky EJ (2006) Deciphering B-ZIP transcription factor interactions in vitro and in vivo. *Biochim Biophys Acta*. 1759: 4–12.

Xiao YL, Smith SR, Ishmael N, Redman JC, Kumar N, Monaghan EL, Ayele M, Haas BJ, Wu HC, Town CD (2005). Analysis of the cDNAs of hypothetical genes on Arabidopsis chromosome 2 reveals numerous transcript variants. *Plant Physiol*. 139: 1323-1337.

Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, Pham P, Cheuk R, Karlin-Newmann G, Liu SX, Lam B, Sakano H, Wu T, Yu G, Miranda M, Quach HL, Tripp M, Chang CH, Lee JM, Toriumi M, Chan MM, Tang CC, Onodera CS, Deng JM, Akiyama K, Ansari Y, Arakawa T, Banh J, Banno F, Bowser L, Brooks S, Carninci P, Chao Q, Choy N, Enju A, Goldsmith AD, Gurjal M, Hansen NF, Hayashizaki Y, Johnson-Hopson C, Hsuan VW, Iida K, Karnes M, Khan S, Koesema E, Ishida J, Jiang PX, Jones T, Kawai J, Kamiya A, Meyers C, Nakajima M, Narusaka M, Seki M, Sakurai T, Satou M, Tamse R, Vaysberg M, Wallender EK, Wong C, Yamamura Y, Yuan S, Shinozaki K, Davis RW, Theologis A, Ecker JR. (2003). Empirical analysis of transcriptional activity in the Arabidopsis genome. *Science*. 302: 842-846.

Zhang Y, Fan W, Kinkeman M, Li X, Dong X (1999). Interaction of NPR1 with basic leucine zipper protein transcription factor that bind sequences required for salicylic acid induction of the PR-1 gene. *Proc Nat Acad Sci USA*. 96: 6523-6528.

Zhang YL, Tessaro MJ, Lassner M, Li X (2003). Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell*. 15: 2647-2653.

## **CAPÍTULO II - Caracterização da proteína ScCBL1 em cana-de-açúcar (*Saccharum spp.*).**

# 1. INTRODUÇÃO

## 1.1. O papel do cálcio na resposta a estresses bióticos e abióticos

O cálcio é um mensageiro secundário central tanto na transdução de sinais extracelulares, como luz ou hormônios, como na resposta a estresses bióticos e abióticos como seca, frio, corte, salinidade entre outros (Trewavas e Malho., 1998; Bush., 1995; Nagee et al., 2003; Sanders et al., 1999). Estes diferentes estímulos mencionados levam a mudanças na concentração de cálcio (assinaturas de cálcio) que é percebida por proteínas sensores de cálcio que mudam a sua conformação e interagem com outras proteínas sinalizadoras para transmissão do sinal (Iacovelli et al., 1999; Kim et al., 2000) (Figura 1).

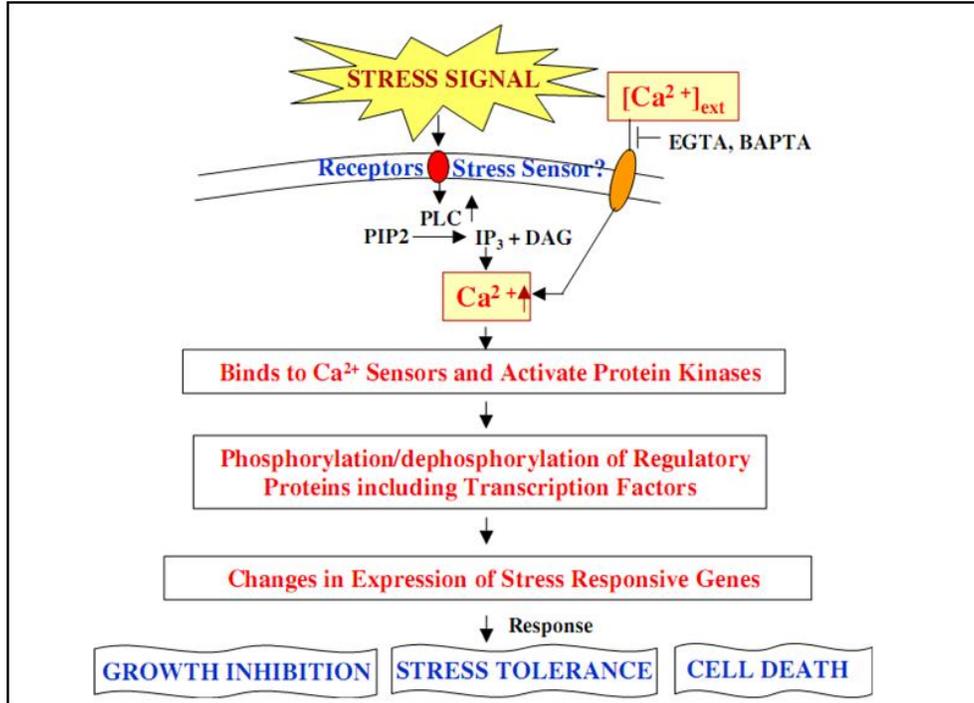


Figura 1: Esquema mostrando via de transdução ativada em resposta a estresse em plantas que tem o cálcio como mensageiro secundário. Adaptado de Mahajan et al. (2008).

## **1.2. Proteínas sensores de cálcio em plantas**

Sensores de cálcio em plantas pertencem a três classes: Sensores responsivos que possuem atividade enzimática própria (Harmon et al., 2000). Um exemplo de proteína pertencente a esta classe é a proteína quinase dependente de cálcio (CDPK) que possui tanto um domínio de ligação a cálcio do tipo calmodulina e um domínio de fosforilação Ser/Thr na mesma proteína.

Outra classe de proteínas sensores de cálcio são os sensores de interação que possuem apenas domínios de ligação de cálcio, não possuem atividade enzimática própria e, portanto dependem da interação com outras proteínas para transmissão do sinal (Luan et al., 2002). O exemplo melhor caracterizado deste tipo de sensor de cálcio é a calmodulina (CaM) e proteínas do tipo CaM.

Um terceiro tipo de proteínas sensores de cálcio denominadas de Calcineurinas do tipo B (CBL) foram posteriormente identificadas em *Arabidopsis* através de análises genéticas (Liu e Zhu., 1998) e por duplo-híbrido (Kudla et al., 1999). As proteínas CBL são estruturalmente similares a proteínas do tipo “calcineurim B-subunit ou CNB” e ao sensor neuronal de cálcio 1 (NCS-1) em animais (Kudla et al., 1999; Nagae et al., 2003).

## **1.3. Proteínas tipo Calcineurina B (CBL) em plantas**

Em *Arabidopsis* foram identificadas pelo menos dez membros destas CBL, denominadas AtCBL (Luan et al., 2002; Kudla et al., 1999) e que estão relacionadas a vias de sinalização em resposta a diversos estresses abióticos (Tabela 1 e Figura 2). As AtCBL são similares estruturalmente aos sensores de interação e até hoje não foi identificada atividade enzimática própria destas proteínas em *Arabidopsis* indicando que as CBL

funcionam através de interações com outras proteínas, como por exemplo as proteínas CIPK para transmissão do sinal.

Tabela 1. Expressão de genes CBL em resposta a estresse abiótico

CBL's	Resposta a estresse / hormônio / açúcares	Espécie
<i>CBL1</i>	Sal, seca e frio. Regulador negativo de sinalização por frio	<i>Arabidopsis</i>
<i>CBL2</i>	Luz	<i>Arabidopsis &amp; Oryza</i>
<i>CBL3</i>	Tardiamente induzido a frio, sal e SA	<i>Pisum sativum</i>
<i>CBL4/SOS3</i>	Estresse a sal	<i>Arabidopsis</i>
<i>CBL5</i>	ABA	<i>Arabidopsis</i>
<i>CBL9</i>	Regulador negativo de sinalização por ABA e promove germinação da semente	<i>Arabidopsis</i>

Foi demonstrado, por exemplo, a interação de AtCBL9 com CIPK3 na via de sinalização que regula negativamente resposta a ácido absísico (ABA) durante germinação de sementes (Pandey., 2008). ABA é um importante fito-hormônio para o crescimento vegetal e processos de desenvolvimento como germinação de sementes, maturação, desenvolvimento do embrião e resposta a estresse (Koornneef et al., 1998; Finkelstein et al., 2002; Pandey et al., 2008).

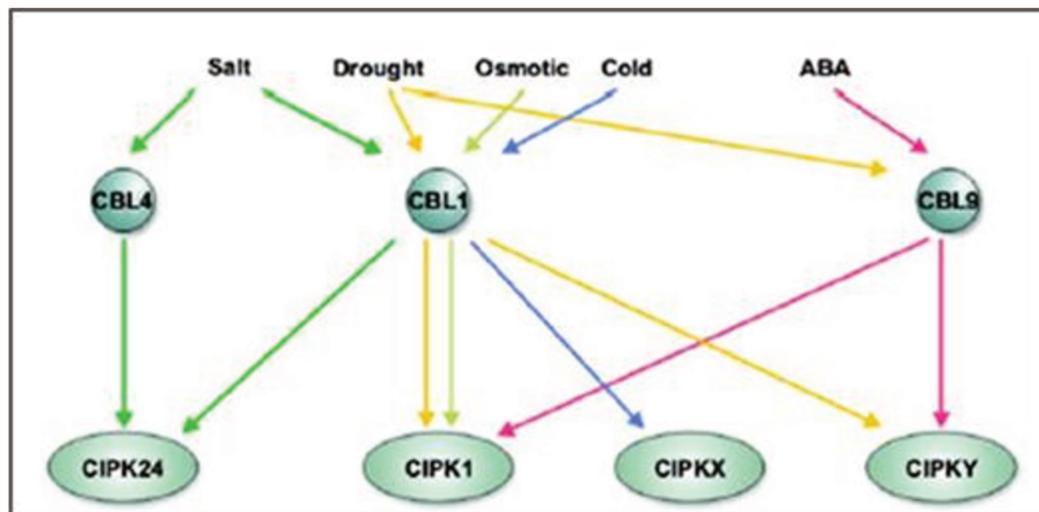


Figura 2: Sinalização de estresse abiótico através da rede CBL/CIPK (Batistic e Kudla., 2004).

O nível do transcrito de AtCBL1, cujo mRNA é preferencialmente expresso na raiz e no caule é altamente induzido por sinais de estresse específicos como seca, frio e ferimento (Kudla et al., 1999). *AtCBL4* ou “salt overly sensitive” (SOS3) foi relacionado a resistência a salinidade (Liu e Zhu., 1998). O complexo AtCBL4-AtCIPK24 (SOS2) está envolvido na ativação do transportador de sódio SOS1 através de fosforilação (Qiu et al., 2002). Outros experimentos mostraram o papel de AtCBL2 na via de transdução da luz através de interação com AtCIPK14 /AtSR1 (Nozawa et al., 2001; Chikano et al., 2001).

A estrutura de AtCBL2 sozinha e em complexo com AtCIPK14 (Figura 3) foi determinada por cristalografia de raios-x em *Arabidopsis* (Nagae et al., 2003; Akaboshi et al., 2008). A proteína AtCBL2 apresenta uma estrutura em alfa-hélice compacta com quatro motivos “EF-hand” de ligação de cálcio e a topologia geral apresentada pela proteína é similar a CNB e NCS-1 (Nagae et al., 2003). A interação entre AtCBL2 e AtCIPK14 ocorre através do motivo NAF/FISL de AtCBL2, ao mesmo tempo em que ocorre o deslocamento de  $\alpha$ -hélices (Akaboshi et al., 2008) (Figura 3).

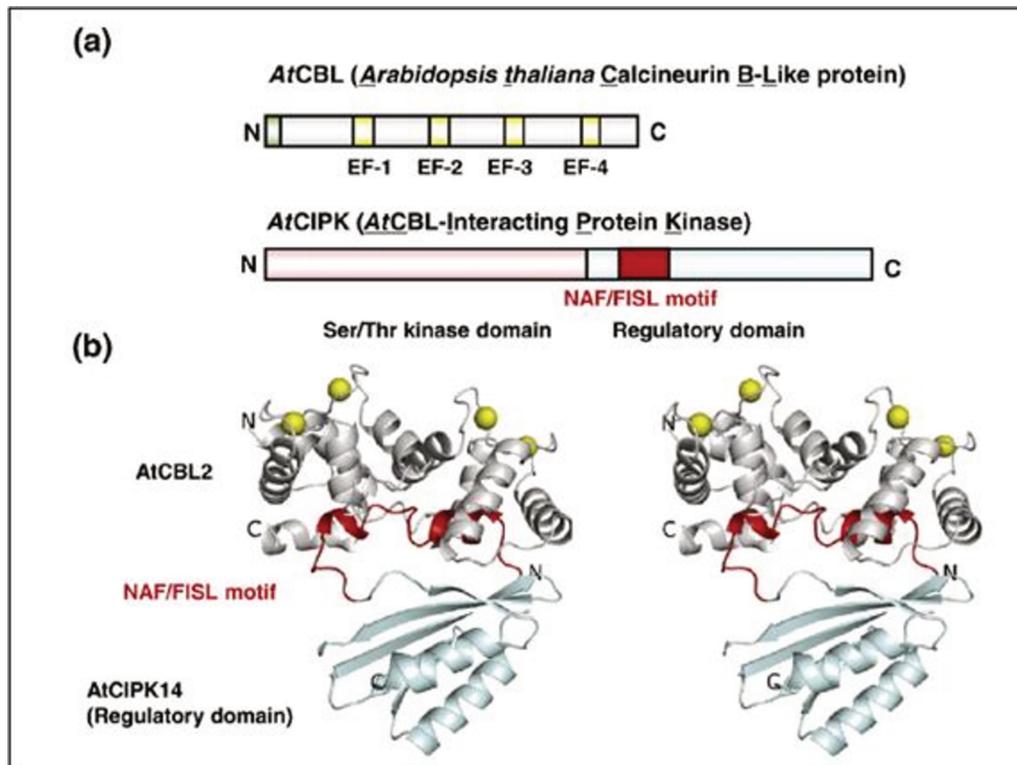


Figura 3: Diagrama mostrando principais domínios (a) e estrutura terciária (b) das proteínas AtCBL2 e AtCIPK14 (Akaboshi et al., 2008).

#### 1.4. Modificação pós-traducional das CBL's

Diversas proteínas CBL em *Arabidopsis* como AtCBL1, AtCBL9 e AtCBL4 (SOS3) foram associadas a membrana plasmática (Luan et al., 2002; Ishitani et al., 2000). Estas proteínas possuem uma importante característica que é modificação pós-traducional. Estas modificações covalentes consistem na adição de moléculas de lipídeo Meristato (miristoilação), composta por 14 carbonos e Palmitato (palmitolação), composta por 16 carbonos, na cadeia de polipeptídeo nascente, mais especificamente a resíduos de glicina e cisteína.

Estas modificações covalentes podem ter implicações em interações proteína-proteína (homo- e heterodimerização), como na ligação da proteína a membrana uma vez que palmitolação aumenta a hidrofobicidade da superfície da proteína modificada, aumentando assim a sua afinidade pela membrana. Na verdade, proteínas miristoiladas são apenas fracamente associadas à membrana, que para ocorrer de fato, necessita também de palmitolação (que ocorre no resíduo de cisteína adjacente a glicina miristoilada) que permitira uma ligação estável da proteína a membrana plasmática (Bijlmakers e Marsh, 2003).

## **1.5. Proteômica estrutural**

De maneira geral o ramo da proteômica estrutural dedica-se a resolver a estrutura das proteínas correspondente a todas as regiões codantes de DNA, mais especificamente a estrutura terciária, uma tarefa em larga medida auxiliada com os dados provenientes de diversos projetos genoma. Este esforço visa a montagem de uma “biblioteca” estrutural de proteínas de diversos organismos, complementando a informação contida em seu DNA. Quanto maior esta biblioteca de proteínas e sua diversidade filogenética maior será a probabilidade de derivarmos informação a respeito da função destas proteínas a partir de dados estruturais através de analogias com proteínas conhecidas e/ou homólogas (Chayen e Saridakis., 2008).

### **1.5.1. Dicroísmo circular (CD) de proteínas**

CD é uma técnica espectroscópica que mede diferença na absorção de luz polarizada à esquerda e à direita que aparecem devido a assimetrias estruturais de macromoléculas biológicas e é bastante utilizado em estudos da estrutura secundária de proteínas. Como as

pontes peptídicas que formam as proteínas são assimétricas, estas macromoléculas não apresentam um plano de simetria e exibem o fenômeno de dicroísmo circular. A estrutura secundária de proteínas pode ser determinada por CD na região “UV distante” (190-250 nm). Nestes comprimentos de onda o cromóforo são as pontes peptídicas sendo que o cromóforo amido da ponte peptídica domina o espectro de CD de proteínas no UV distante. Quando a radiação UV passa através de uma macromolécula que possui pontes peptídicas uma parcela da radiação é absorvida pelo composto. Esta absorção ocorre por uma subtração de energia do feixe de UV causado pela excitação dos elétrons de baixa energia para orbitais de energia mais elevada. A espectrometria de dicroísmo circular no UV distante provoca estas transições eletrônicas e esta energia absorvida pode ser quantificada em um gráfico de comprimento de onda em função da absorbância.

Os amidos possuem duas transições eletrônicas de baixa energia que são bem caracterizadas nas transições  $n \rightarrow \pi^*$  e  $\Pi_0 \rightarrow \Pi^*$  apresentando dicroísmo circular a 215-230 nm e 185-200 nm, respectivamente (Woody., 1995; Woody e Koslowski, 2002). Estes dois eventos dominam o fenômeno de dicroísmo circular no UV distante. A transição  $\Pi_0 \rightarrow \Pi^*$  é responsável pela banda positiva em aproximadamente 190 nm e uma negativa a 208 nm, característicos de um espectro de  $\alpha$ -hélice, além da banda positiva de 198 nm, característica de um espectro de folha  $\beta$  (Corrêa e Ramos, 2009). Podemos observar na figura 4 espectros característicos obtidos a partir de proteínas cujas estruturas secundárias são formadas predominantemente por  $\alpha$ -hélices, folhas  $\beta$  ou aleatória.

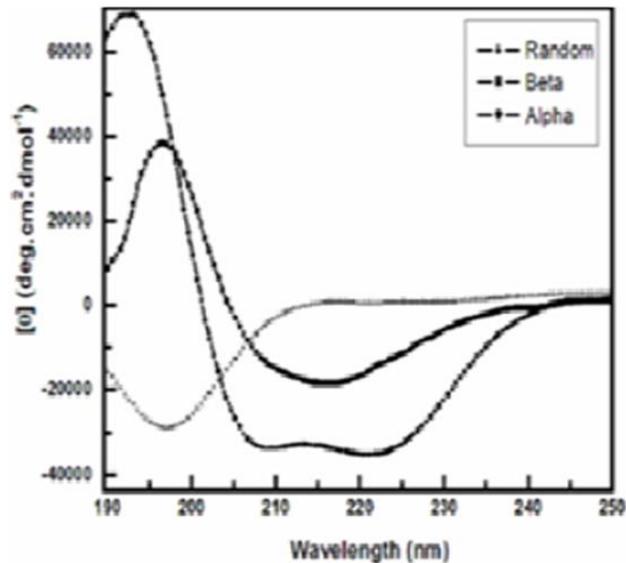


Figura 4: Espectros de CD medidos no UV distante característicos de proteínas predominantemente  $\alpha$ -hélice (“all alpha”), predominantemente  $\beta$  (“all beta”) e predominantemente aleatória (“random”). Observe que o espectro de uma proteína toda alfa (ver “Alpha” no gráfico) possui duas bandas negativas similares (a 222 e 208 nm) e uma banda positiva a ~ 190 nm (Corrêa e Ramos., 2009).

Em estudos de espectrometria por CD com proteínas a absorvância é geralmente expressa como a elipcidade residual média ( $\theta$ ):

$$[\theta] = \frac{\theta \times 100 \times M}{C \times l \times n} \quad (\text{Eq. 1})$$

Aonde  $\theta$  representa a elipcidade em graus,  $l$  é o caminho ótico (cm),  $C$  é a concentração em mg/mL,  $M$  é a massa molecular e  $n$  é o número de resíduos de uma proteína. A notação da elipcidade residual média  $[\theta]$  é dada em  $\text{deg.cm}^2.\text{dmol}^{-1}$ .

O dicroísmo circular pode ser utilizado também em estudos de estabilidade térmica para analisarmos, por exemplo, qual a temperatura que ocorre o desenovelamento de uma proteína, assim como interação de proteína com ligantes e/ou proteína-proteína (Johnson., 1985).

### **1.5.2. Espalhamento dinâmico de luz (“Dynamic Light Scattering”, DLS)**

O DLS é uma técnica utilizada em larga medida para avaliar a distribuição do tamanho de partículas em solução. Com ela é possível a medição do raio hidrodinâmico, polidispersão e efeitos de agregação de amostras de proteínas. Nesta técnica é medido o espalhamento de partículas (macromoléculas) em solução a partir de luz emitida por um diodo de laser. A partir das leis Brownianas para cálculo do movimento de macromoléculas em solução pode-se calcular a intensidade do espalhamento destas.

Foi determinado empiricamente que uma amostra de proteína considerada monodispersa – que apresenta um percentual de polidispersão menor do que 30% do tamanho total da partícula é mais provável de formar cristais do que uma amostra considerada polidispersa (mais de 30% de polidispersão) que contenha agregados não específicos (Zulauf e Dárcy., 1992; Ferre-D’amare e Burley., 1994). Uma típica amostra monodispersa geralmente apresenta um pico único de intensidade de espalhamento da proteína sendo medida, enquanto que amostras polidispersas, não recomendadas para cristalografia de proteínas, apresentam mais de um pico e necessitam de etapas de filtração/ purificação adicionais.

### **1.5.3. Cristalografia de Proteínas**

A determinação da estrutura de proteínas por cristalografia por raios-x envolve uma série de etapas desde a clonagem do gene de interesse, expressão, purificação, obtenção do cristal, coleta de dados de difração até a determinação da posição atômica (ver Figura 5).

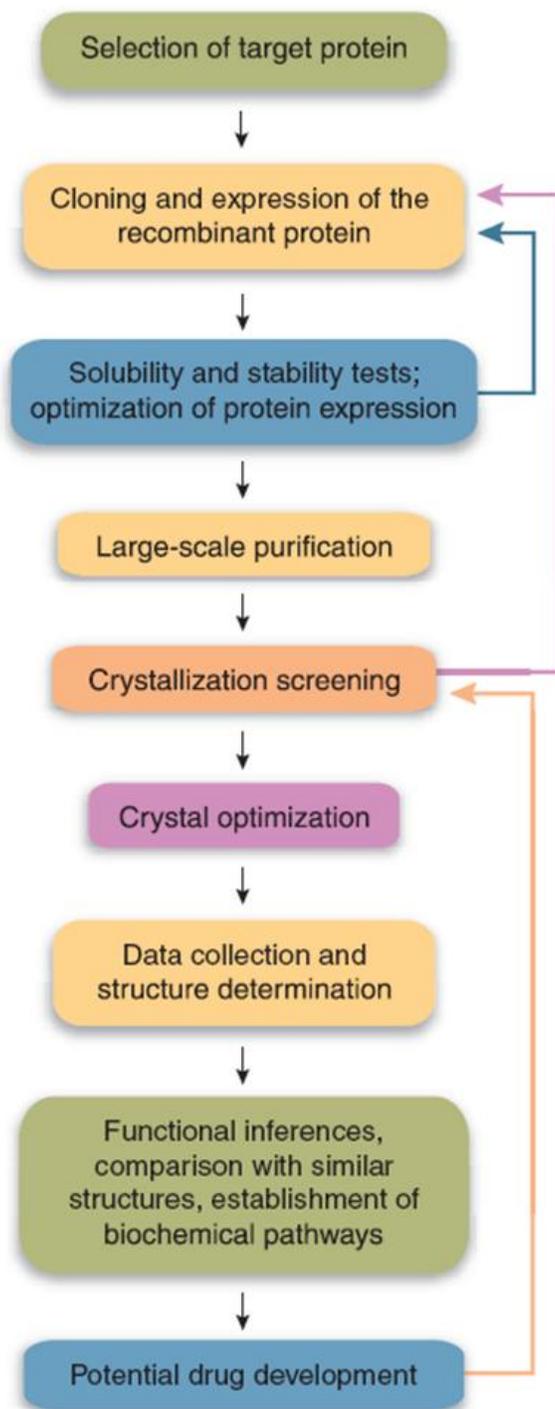


Figura 5: Pipeline mostrando etapas desde a escolha de um gene até obtenção de sua estrutura. A purificação da proteína e a obtenção do cristal são os maiores gargalos durante este processo (Chayen e Saridakis., 2008).

A cristalização de proteínas foi previamente definida como um fenômeno de transição de fases que pode ser melhor visualizado como um diagrama de fases (Figura 6). Esse diagrama ilustra em que estado físico (líquido, cristalino ou sólido amorfo – precipitado) a proteína é estável em diversos parâmetros de cristalização. Além da concentração da proteína, a temperatura e características do solvente (como pH, concentração do tampão e força iônica) são variáveis que estão diretamente ligadas ao processo de cristalização (Saridakis et al., 1994).

Técnicas de cristalização de proteínas baseiam-se no fato de que proteínas permanecem em solução somente até certa concentração e uma vez que este limite é alcançado a solução cessa de ser homogênea e um novo estado ou fase é alcançado (cristal). Portanto, através da mudança nas condições da solução tenta-se exceder o limite de solubilidade de uma proteína de modo a formar cristais (McPherson., 1999).

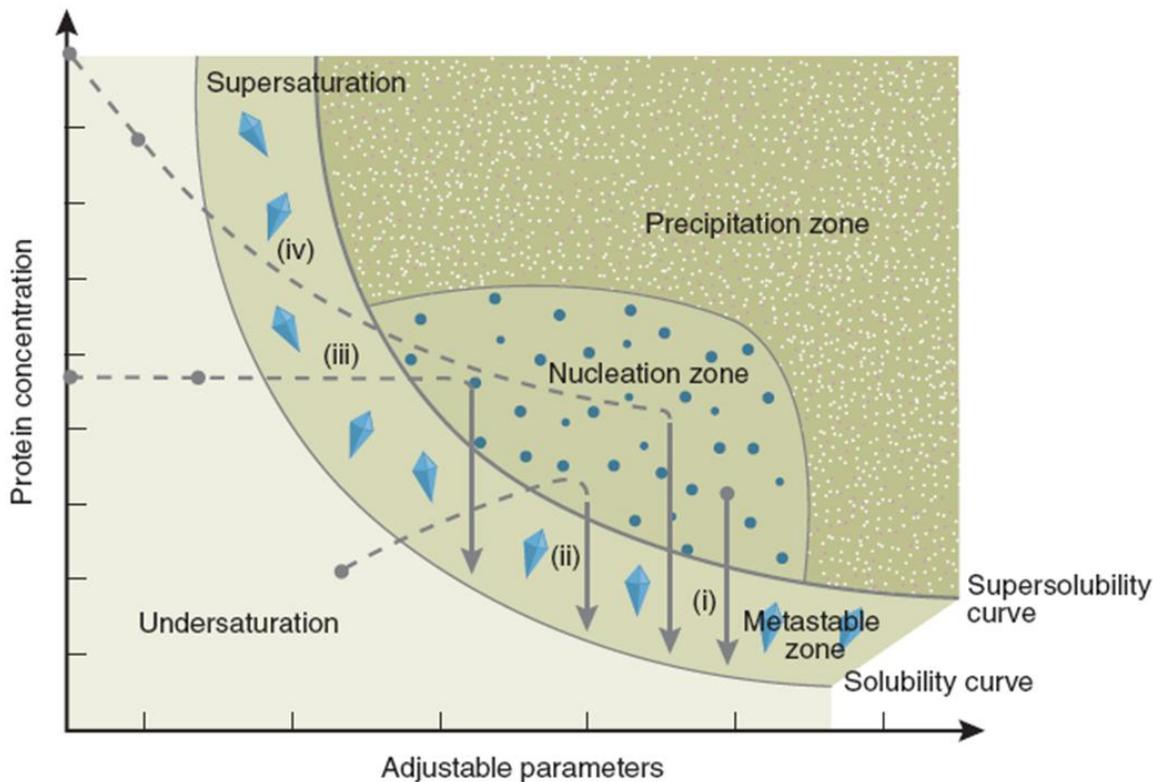


Figura 6: Diagrama de fase em cristalização de proteínas. Cristalização é uma função da concentração da proteína e parâmetros ajustáveis como temperatura pH, precipitante e concentração de aditivos. A solubilidade pode ser definida como concentração de proteína em solução que está em equilíbrio com o cristal. A curva de supersaturação separa as condições as quais ocorre nucleação ou mudança de fase daquelas onde a solução de cristalização permanece inalterada ou limpa, sem formação de estrutura (Chayen., 2004).

Em um típico diagrama de fases podemos distinguir quatro áreas: uma área onde a proteína irá precipitar, havendo a formação de estruturas desordenadas como agregados ou precipitados (*precipitation zone*); uma área onde nucleação espontânea ocorre (*nucleation zone*); uma área logo abaixo da zona de nucleação, onde cristais são estáveis e podem crescer (zona *metastable* –considerada melhor área para formação de cristais) e uma área de subsaturação (*undersaturated*) onde a proteína está dissolvida e nunca cristalizará pois a concentração está abaixo da solubilidade da proteína. De maneira geral pode-se afirmar que

quando a concentração da proteína excede sua solubilidade cristais são formados (Asherie., 2004).

Na prática, cristais raramente são formados enquanto a concentração da proteína não ultrapassar sua solubilidade por um fator de ao menos três (Chernov., 1997). A supersaturação é necessária para vencer a barreira de ativação de energia necessária para formação do cristal (Asherie., 2004). Essa barreira representa a energia livre necessária para criar os clusters microscópicos da proteína – conhecido como núcleo a partir do qual o cristal é formado (Kashchiev., 2000).

A difusão a vapor é a técnica mais utilizada em cristalização de proteínas. Nela, a proteína é uma gota sentada ou pendurada que se equilibra contra um reservatório contendo agentes de cristalização a uma concentração maior ou menor que a gota da proteína (Chayen e Saridakis., 2008). Tipicamente são usados volumes iguais 1:1 de proteína e solução precipitante (0.5-1  $\mu$ l) no *screening* cristalográfico.

#### **1.5.4. Espalhamento de raios-x a baixos ângulos (SAXS)**

SAXS é uma técnica de grande potencial que possibilita a obtenção de informações estruturais úteis a partir de macromoléculas em solução de maneira mais rápida e simples que cristalografia por raios-x. Em um típico experimento de SAXS é possível obtermos informação a respeito do formato, conformação da macromolécula em solução, massa molecular e estado oligomérico. A primeira informação essencial de um experimento de SAXS é sabermos o grau de monodispersão da amostra através das curvas de espalhamento e do gráfico de Guinier. As informações contidas nas curvas de espalhamento são bem menores que dados de cristalografia. Entretanto análises por SAXS podem ser usadas para

obtenção de estrutura de baixa resolução da molécula, a ser interpretada a partir de informações adicionais obtidas por técnicas de alta resolução como cristalografia por exemplo (Putnam et al., 2007).

De maneira sucinta, SAXS envolve a exposição da amostra a um feixe monocromático de raios-x e a medição dos ângulos espalhados onde a amplitude do vetor de espalhamento pode ser definida como  $q = 4\pi \sin\theta/\lambda$ , sendo  $2\theta$  o ângulo de espalhamento e  $\lambda$  o comprimento de onda do raio-x (1.488 Å em nossos experimentos; Figura 7). Podemos entender SAXS como sendo um método de contraste no qual o padrão de espalhamento final de uma proteína é obtido a partir da diferença ou subtração da densidade média de elétrons entre a amostra da proteína (soluto) e o tampão puro (solvente).

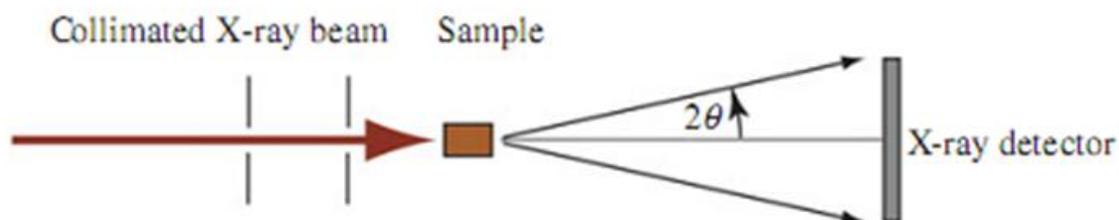


Figura 7: Esquema mostrando interação de raio-x e amostra. SAXS consiste na exposição da amostra a um feixe de raio-x e medição do espalhamento a um ângulo  $2\theta$  (Putnam et al., 2007).

Assumindo que a amostra é homogênea química e estruturalmente, monodispersa e desprezando interações intermoleculares (condição atingida a concentrações suficientemente baixas), o espalhamento é isotópico e a intensidade total será dada como a soma das intensidades dos espalhadores individuais.

A função de distribuição de distâncias (ou distribuição dos pares),  $P(r)$ , é calculada através da transformada de Fourier a partir da curva de espalhamento e seus resultados nos dão informações a respeito das distâncias entre elétrons nas partículas espalhadas nas amostras (Putnam et al., 2007).

A  $P(r)$  pode ser usada para análise do formato da proteína e pode ser definido como a probabilidade de encontrarmos um par de pontos (elementos) a uma dada distância  $r$  dentro do volume todo da partícula espalhadora (Guinier e Fournet., 1955; Glatter e Kratky., 1982). É possível obtermos informação acerca do formato de uma proteína a partir de uma simples inspeção visual de  $P(r)$  (Glatter e Kratky., 1982). Podemos observar na figura 8 a distribuição das distâncias  $P(r)$  de diferentes corpos geométricos. Podemos observar que proteínas prolatas (alongadas) possuem um máximo em distâncias curtas correspondendo ao raio da seção transversal e, à medida que as distâncias aumentam, a função decresce.

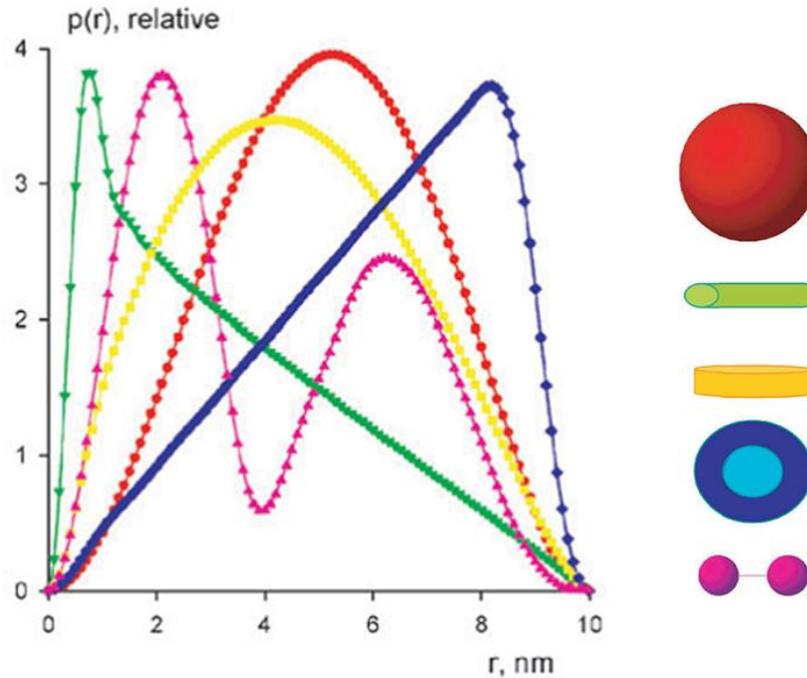


Figura 8: Gráfico de distribuição das distâncias  $P(r)$  de corpos geométricos Fonte: Svergun e Kock., 2003.

Outros dois parâmetros importantes que podem ser extraídos de análises por SAXS são a massa molecular (em  $kDa$ ) e o estado oligomérico de uma proteína. O parâmetro de tamanho, definido como raio de giro ( $R_G$ ), é dependente do formato do envelope protéico. No gráfico de Guinier ( $\log(I(q))$  contra  $q^2$ ) podemos obter uma linha reta com a qual podemos calcular o  $R_G$  e o  $I(0)$  no caso de amostras monodispersas. A faixa  $q$  sobre a qual a aproximação Guinier é válida ( $qR_G$ ), está em torno de  $qR_G < 1.3$  no caso de proteínas globulares e  $qR_G < 0.8$  no caso de proteínas alongadas (prolatas). Outro parâmetro importante que pode ser avaliado a partir de  $q$  é o  $I(0)$  que corresponde a intensidade do espalhamento de uma partícula e está diretamente relacionada ao peso molecular e concentração da proteína.  $I(0)$  pode ser calculado como a intensidade de espalhamento extrapolada no ângulo zero (Putnam et al., 2008). O SAXS também pode ser usado para

uma avaliação do enovelamento da proteína através do gráfico de Kratky. O gráfico de Kratky é calculado a partir da curva de espalhamento e no caso de proteínas enoveladas exibe um máximo bem definido a mais baixos ângulos. Proteínas desestruturadas não exibem este pico característico, havendo um platô a ângulos mais altos.

O SAXS permite a construção de um modelo *ab initio* de baixa resolução do formato do envelope protéico a partir dos perfis de espalhamento com o uso de alguns programas como DAMMIN, GASBOR e SASHA (Putnam et al., 2007).

Trabalhos anteriores feitos no Laboratório de Genômica Funcional (LGF) mostram que o gene ScCIPK8 está envolvido no metabolismo de carboidratos, mais especificamente que ScCIPK8 está associado ao teor de sacarose e é diferencialmente expresso em resposta ao fito-hormônio ABA (não publicado). Também foi feito um ensaio por duplo-híbrido usando a proteína ScCIPK8 como o híbrido-1 ou “isca” que identificou a interação com três proteínas sensores de cálcio em cana-de-açúcar (ScCBL1, ScCBL3 e ScCBL6). Destas três proteínas identificadas, ScCBL1 é a que apresenta uma interação mais forte com ScCIPK8.

O complexo CIPK-CBL foi relacionado em trabalhos anteriores à via de sinalização em resposta a diversos estresses abióticos, como seca e salinidade. O trabalho de caracterização dos componentes envolvidos nestas vias de sinalização é importante para o conhecimento sobre o funcionamento dessas redes regulatórias em plantas e em última análise para possíveis aplicações, como o uso destes genes como marcadores moleculares relacionados estresses abiótico bem como o desenvolvimento de plantas transgênicas resistentes a estresses.

## 2. OBJETIVO

O objetivo geral deste trabalho é a caracterização estrutural da proteína ScCBL1 de *Sacharum spp.* Mais especificamente objetivamos:

1. Clonagem, expressão e purificação de ScCBL1
2. Caracterização da estrutura secundária da proteína ScCBL1 por espectroscopia de dicroísmo circular (CD)
3. Caracterização inicial da estrutura terciária por espalhamento de raios-X a baixo ângulo (SAXS)
4. Preparação de ensaios de cristalografia para determinação da estrutura atômica de alta resolução.

### 3. MATERIAIS E MÉTODOS

#### 3.1. Clonagem e transformação de *Escherichia coli*

A região codificante completa, delimitada pelo ATG e pelo stop códon (CDS, de *coding sequence*) de cada um dos genes *ScCBL1*, *ScCBL3* e *ScCBL6* foi previamente clonada no vetor pGADT7 (Clontech). Os vetores foram digeridos com *NdeI* e *XhoI* (*ScCBL3*) e *NdeI*-*BamHI* (*ScCBL1* e *ScCBL6*) então ligados ao vetor pET28a digerido com as mesmas enzimas, de modo a incluir uma cauda de histidina na região N-terminal de cada proteína recombinante.

Os produtos da ligação foram então transformados em cepas de *E. coli* DH5 $\alpha$  por choque térmico, confirmados por digestão usando os respectivos sítios de restrição específicos e então transformados por choque térmico em BL21 (DE3) para expressão das proteínas (Sambrook et al., 1989).

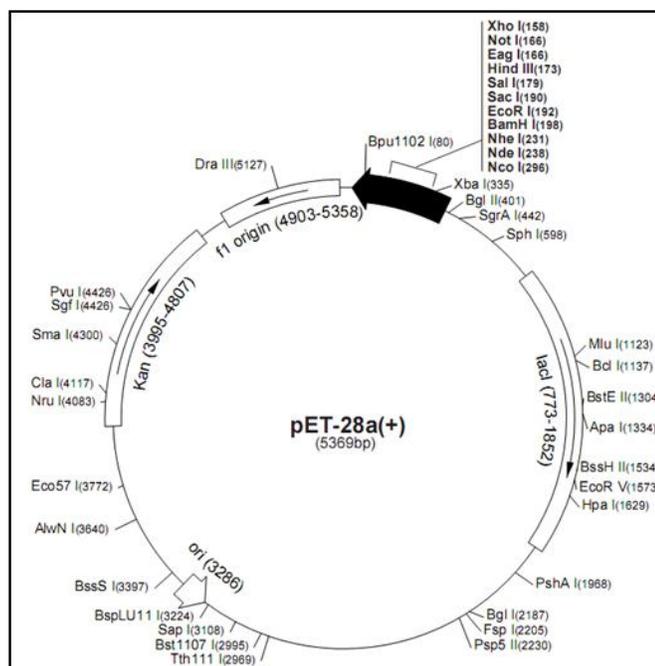


Figura 9: Mapa do plasmídeo pET28a(+). Fonte: Novagen.

### 3.2. Ensaio de expressão

Os ensaios de expressão das proteínas ScCBL1, 3 e 6 foram realizados em meio LB contendo 50 mg/ml de canamicina.. As células foram cultivadas a 37°C em um agitador a 240 rpm até atingirem uma OD<sub>600</sub> de aproximadamente 0,8 e o pellet não-induzido foi coletado. Neste ponto foi adicionado ao meio 1 mM de IPTG (isopropil-tio-β-D-galactopiranosídeo) e a temperatura de indução foi ajustada para 30°C. Após quatro horas de indução as células foram coletadas por centrifugação a 10.000 rpm, durante 10 minutos a 4°C e armazenados no freezer a -20°C. Análise das proteínas expressas foram feitas em gel de eletroforese (SDS-PAGE) a 12% (Laemmli., 1970) e confirmada por *western blot* usando o anticorpo contra His-tag.

### 3.3. Detecção da proteína recombinante

Para detecção das proteínas recombinantes foi corrido um gel de poliacrilamida (SDS-PAGE) das amostras das proteínas obtidas tanto a partir do extrato total, como do extrato solúvel, ressuspendidas em tampão 100 mM fosfato de sódio pH-7.4, 100 mM NaCl. As amostras foram aplicadas com um volume total de 25  $\mu$ L ao gel em uma proporção de 1:1 com tampão de amostra (Tris-HCL 0.5 M, pH-6.8; 2- $\beta$ Mercaptoetanol 10%; SDS 10%; glicerol 20%; azul de bromofenol 0,04%) e submetidos a eletroforese por aproximadamente uma hora a 150 V em tampão de corrida (24,8 mM tris base; 192 mM Glicina e 0,1% w/v SDS). Após a corrida, as proteínas presentes no gel foram transferidas para uma membrana de nitrocelulose Hybond™- C Extra (Amersham). Para tanto, o gel foi incubado por 15 min em tampão de transferência (Bjerrum e Schafer-Nielsen., 1986) (20% metanol; 39 mM Glicina; 48 mM Tris pH 9.2) e depois foi montado o aparato de transferência contendo o papel-filtro (“Extra Thick Blot Paper” – Bio-Rad), o gel contendo as proteínas e a membrana de nitrocelulose. A reação de transferência foi submetida a eletroforese durante o período de meia hora a 15 V em um aparelho de eletroforese horizontal Trans-Blot® SD (semi-Dry) (Bio-Rad). Após a transferência a membrana foi corada temporariamente usando o kit MemCode™ Reversible Protein Stain Kit (Thermo-scientific) de modo a confirmar a transferência das proteínas do gel para a membrana.

A membrana contendo agora as proteínas de interesse foi incubada durante a noite em 30 mL de solução bloqueadora (TBS 1X; Leite desnatado Molico 5%). No dia seguinte a membrana foi hibridada pelo período de 1 hora e meia com o anticorpo primário anti-his a uma concentração de 1: 4.000. A membrana foi lavada com TBS 1X duas vezes e com TBS 1X Tween 20 (0,05%). A membrana foi então incubada por 1 hora com o anticorpo

secundário (conjugado anti-his peroxidase de camundongo) a uma concentração de 1:20.000 e leite a 5% (leite desnatado em TBS 1X). A membrana foi lavada com TBS 1X duas vezes e com TBS 1X Tween 20 (0,05%). A revelação foi feita adicionando-se o substrato para reação quimioluminescente SuperSignal® West Pico (Thermo-scientific) na membrana por aproximadamente 5 min até o aparecimento das bandas. A membrana foi então exposta junto a um filme de revelação por cerca de 3 minutos e digitalizada.

### **3.4. Purificação por cromatografia de afinidade em resina de níquel (Ni-NTA)**

O pellet das culturas induzidas foi ressuspendido em tampão de lise (100 mM fosfato de sódio pH 7,4, 100 mM NaCl), adicionando-se 0,5 mg/ml de lisozima e do inibidor de protease PMSF sendo incubado por 30 min no gelo. Posteriormente a solução foi sonicada para quebra da parede celular a 160 K (três pulsos de 10 segundos) e centrifugada a 4°C por 10 minutos a 14.000 rpm. O pellet insolúvel (corpos de inclusão) foi separado para um tratamento posterior e o sobrenadante (fração solúvel) armazenado à 4°C.

O sobrenadante correspondente a fração solúvel foi novamente centrifugado à 4°C por 10 min a 14.000 rpm antes de ser aplicado a uma coluna de 1 mL contendo 1 mL de resina de níquel Ni-NTA (Qiagen). Após a ligação da proteína na coluna a resina foi lavada com tampão fosfato para remoção de proteínas que não se ligaram a resina (100 mM fosfato de sódio pH 7,4, 100 mM NaCl) e então eluídas da coluna com o tampão de lavagem contendo concentrações crescentes de imidazol (10 mM, 25 mM, 50 mM, 100 mM, 200 mM e 300 mM). Após a eluição da coluna a proteína de interesse fusionada foi visualizada em um gel SDS-PAGE 12% corado com *Comassie Brilliant*.

### **3.5. Espalhamento dinâmico de luz (DLS)**

Experimentos de espalhamento dinâmico de luz da proteína ScCBL1 em solução a concentrações de 3 mg/ml, 11 mg/ml e 15 mg/ml, com e sem DTT (3 mM) foram realizados em um volume final de 100 µL em um equipamento DynaPRO-MS/X (Protein Solutions) a uma temperatura de 18°C e os dados obtidos com o programa DYNAMICS, versão 6. Cada medição consistiu em cinquenta acumulações (leituras) independentes para cada amostra. O equipamento foi utilizado através de submissão da proposta para utilização do LEC, no LNBIO (antigo CEBIME/LNLS).

### **3.6. Espectroscopia de Dicroísmo Circular (CD)**

A coleta de dados para os experimentos de CD foram realizados em um espectropolarímetro Jasco J-720 com um controlador de temperatura tipo Peltier e um termostato localizado no IQ da UNICAMP. Os dados foram coletados sob uma temperatura constante de 4°C e o caminho óptico da cubeta usada foi de 0.1 cm. A espectroscopia por CD mede a absorbância de espectros de luz ultravioleta na região distante (190-260 nm) de proteínas determinando assim um espectro de CD característico de acordo com a estrutura secundária de cada proteína. Amostras da proteína ScCBL1 foram purificadas por cromatografia de afinidade a uma concentração de 4 µM (0,35 mg/mL) e dializadas durante a noite usando membranas de celulose para diálise (Sigma) contra um tampão fosfato de sódio pH 7.4, contendo 20 mM NaCl.

Os dados foram coletados usando os seguintes parâmetros: 100 nm por minuto de varredura, 1 nm de intervalo de dados. Para deconvolução do espectro da proteína ScCBL1 foi usado o programa Dychroweb (Whitmore e Wallace., 2004), usando o algoritmo CDSSTR (Compton e Johnson., 1986) com os dados de referência numero sete que

continham quarenta e oito proteínas de referência (Janes., 2008). O método PROFsec do programa PredictProtein (<http://www.predictprotein.org>) foi usado para uma predição da estrutura secundária de ScCBL1 a partir de sua sequência de aminoácidos (Rost et al., 2004).

### **3.7. Espalhamento de Raios-X a Baixos Ângulos (SAXS)**

Uma análise inicial da proteína ScCBL1 purificada previamente por cromatografia de afinidade foi realizada na linha SAXS2 do Laboratório Nacional de Luz Síncrotron (LNLS), equipada com um detector bidimensional MAR CCD 165, em colaboração com o Prof. Ricardo Aparicio (IQ/UNICAMP). As medidas foram feitas com feixe de raios-X monocromáticos no comprimento de onda  $\lambda = 1.488 \text{ \AA}$  e distância detector-amostra de 1979,68 mm, levando a um intervalo  $0,0153 \leq q \leq 0,3069$ , onde  $q$  é o módulo do vetor de espalhamento, definido como  $q = 4\pi \sin(\theta) / \lambda$ , e  $2\theta$  é o ângulo de espalhamento.

Foram coletados dados de espalhamento da proteína ScCBL1 concentrada a 3,5 e 11,0 mg/mL, em tampão fosfato de sódio pH 7.4, 100 mM NaCl, 200 mM Imidazol e 3 mM DTT. Diversas exposições consecutivas de 10 min. foram feitas para cada amostra, com o porta-amostras mantido a 298 K e o espalhamento do tampão sendo medido no mesmo tempo de exposição das amostras de proteína. Durante o experimento as amostras que aguardavam a coleta de dados foram mantidas em um isopor com gelo. Um padrão de calibração (albumina de soro bovino, BSA) foi utilizado para uma estimativa da massa molecular, de acordo com o método descrito em Mylonas e Svergun (2007). Uma segunda estimativa foi feita com base em um método novo baseado em uma única curva medida em escala relativa (Fischer et al., 2010).

As imagens foram integradas com o programa FIT2D (Hammersley et al., 1996), e o espalhamento foi normalizado pela intensidade do feixe e atenuação da amostra. O espalhamento do tampão foi então subtraído do espalhamento da amostra. As análises subsequentes e construção *ab initio* do envelope protéico a partir da curva de espalhamento foram realizadas com programas do pacote ATSAS (Konarev et al., 2006), entre eles GNOM/AUTOGNOM, DAMMIN e DAMAVER.

### **3.8. Ensaios de Cristalização de proteínas**

Ensaios de cristalização de amostras monodispersas da proteína ScCBL1 em três concentrações (3, 11 e 15 mg/mL) foram realizados através de proposta de pesquisa previamente submetida para testes de cristalização utilizando o RoboLab do LNLS. O equipamento usado para preparação das placas foi um robô Honeybee 961 (Genomic Solutions) e placas de 96 poços CrystalQuick™. Todas as amostras estavam diluídas em tampão fosfato de sódio pH 7,4, 100 mM NaCl, 200mM Imidazol. Foram testadas quinhentas e quarenta e quatro condições de cristalização para cada amostra de ScCBL1 totalizando aproximadamente 1.632 condições. Para tanto foram utilizados seis kits de cristalografia: Crystal screen I e II (Hampton Research), Wizard I e II (Emerald Biosystems), PACT (Nextal/Qiagen), JCSG (Nextal/Qiagen), SaltRx (Hampton Research) e Precipitant Synergy (Emerald Biosystems) através da técnica *sitting-drop*. As placas foram armazenadas à 18°C e observadas em um microscópio ótico após um mês de maneira a permitir formação dos cristais. Prováveis cristais de sal foram identificados pela adição de 0.5 µl do corante azul Izit (Hampton) que cora cristais macromoleculares, não havendo coloração de cristais inorgânicos ou de sal.

## 4. RESULTADOS E DISCUSSÃO

### 4.1. Expressão, detecção e purificação das proteínas ScCBL1, ScCBL3 e ScCBL6

A expressão das proteínas ScCBL1, ScCBL3 e ScCBL6 em *E. coli* BL21 (DE3) foi induzida com 1 mM IPTG. A indução da proteína ScCBL1 de aproximadamente 25 kDa foi monitorada por quatro horas a partir da fração total induzida (Figura 10). Dentre as três proteínas ScCBL analisadas neste trabalho a ScCBL1 foi a que apresentou a maior indução por IPTG..

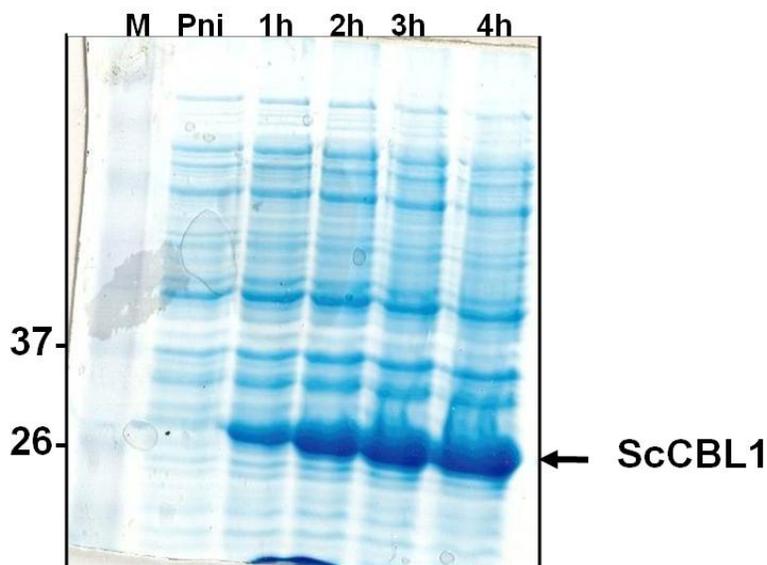


Figura 10: Expressão da proteína ScCBL1 em *E. coli*. A região codificante da proteína ScCBL1 foi clonada no vetor pET28a, que foi então transformado em *E. coli* BL21 (DE3). A indução da expressão da proteína ScCBL1 por IPTG foi feita por 1, 2, 3 e 4 h. Amostras da proteína foram separadas em SDS-PAGE a 12% e visualizadas com *Comassie Blue*. M : Marcador de peso molecular (kDa). Pni: Pellet não-induzido. A seta indica a posição da proteína ScCBL1.

A indução das proteínas ScCBL 3 e 6, de aproximadamente 25 kDa cada uma, foram monitoradas durante quatro horas a partir da fração total induzida (ver figura 11).

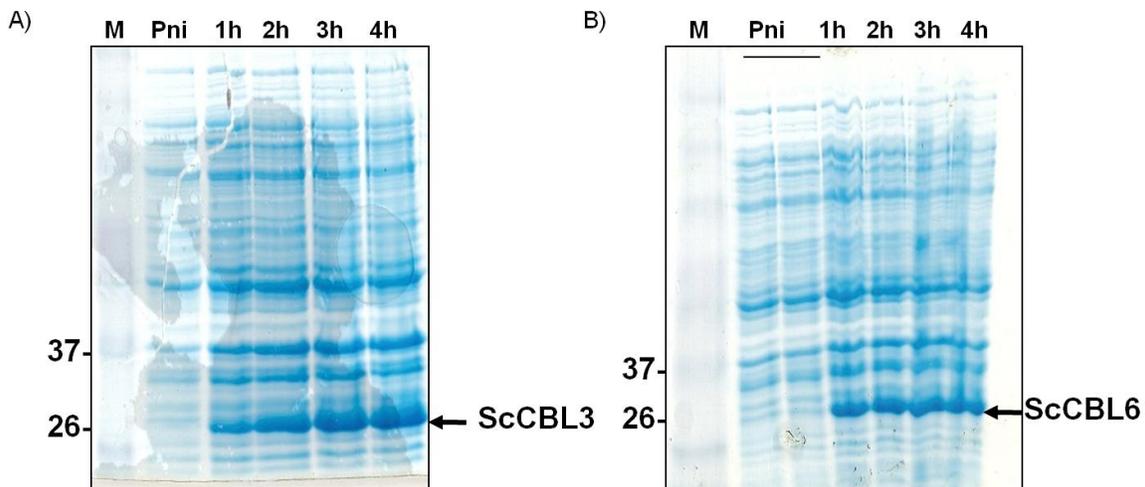


Figura 11: Expressão das proteínas ScCBL3 e ScCBL6 em *E. coli*. Análise de indução da expressão das proteínas por IPTG: A) ScCBL3 e B) ScCBL6 em SDS-PAGE a 12%. M: Marcador de peso molecular (kDa). Pni: Pellet não-induzido. Setas indicam a posição das proteínas ScCBL3 e ScCBL6.

Os ensaios do estado de solubilidade das proteínas ScCBL1, 3 e 6 clonadas em pET28a foram feitos com o objetivo de descobrirmos se as proteínas eram expressas no extrato solúvel (sobrenadante) ou no extrato insolúvel (pellet). Dentre as três proteínas, ScCBL1 é a única expressa na fração solúvel (Figura 12) enquanto as proteínas ScCBL 3 e 6 (Figura 13) são parcialmente solúveis.

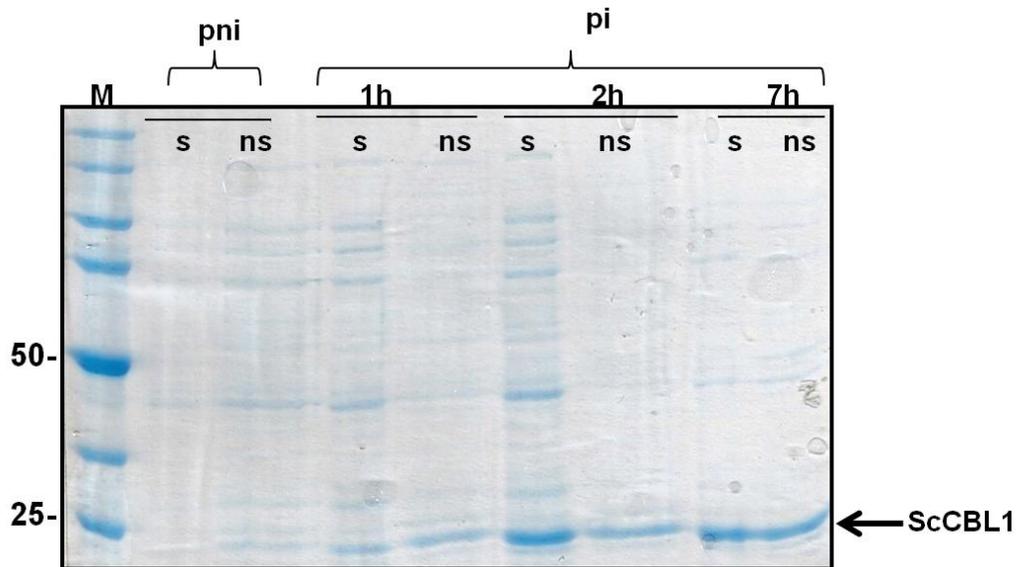


Figura 12: Ensaio do estado de solubilidade da proteína ScCBL1 em SDS-PAGE a 12% a partir do extrato induzido solúvel (s) e insolúvel (ns). M: marcador de peso molecular (kDa). Pni: pellet não induzido. Pi: pellet induzido

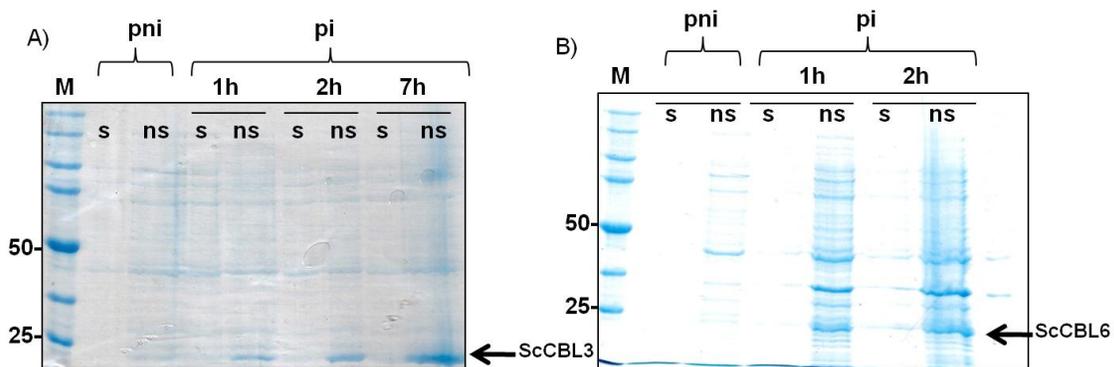


Figura 13: Ensaio do estado de solubilidade das proteínas ScCBL3 e ScCBL6 em SDS-PAGE a 12% a partir do extrato induzido solúvel (s) e insolúvel (ns). M: marcador de peso molecular (kDa). Pni: pellet não induzido. Pi: pellet induzido

A detecção da proteína ScCBL1 usando um anticorpo anti-his a partir de extrato solúvel foi feita de maneira a confirmarmos a especificidade da proteína expressa nos ensaios de indução. Podemos observar na figura 14 a banda da proteína ScCBL1 nas

frações induzidas a uma, duas e seis horas. Nenhuma banda foi detectada no pellet não induzido, conforme esperado.

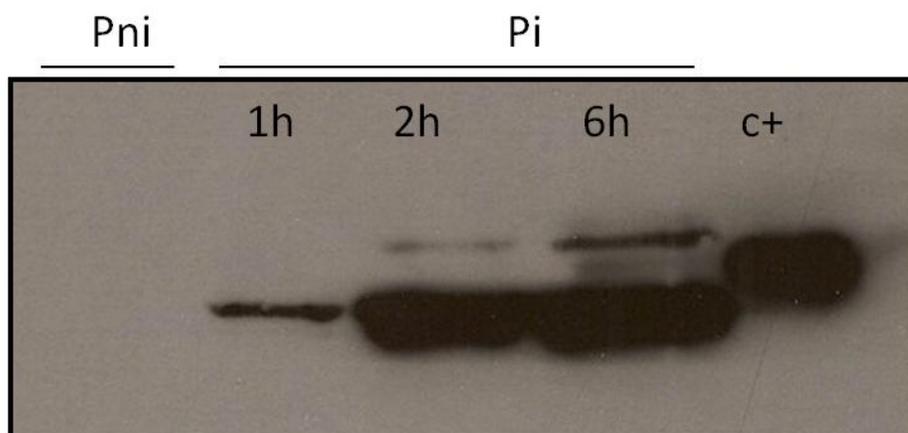


Figura 14: Detecção da proteína ScCBL1 por *western blot* com anticorpo anti-his. Detecção do sinal quimioluminescente gerado pela interação do conjugado HRP com anticorpo primário ligado a proteína ScCBL1 foi feito por revelação do filme sobreposto a membrana por raios-x . Pni: pellet não induzido; Pi: pellet induzido a 1, 2 e 6 horas. C+: controle positivo.

A detecção por *western blot* da proteína ScCBL3, tanto do extrato protéico total quanto do extrato solúvel, confirmam os testes de solubilidade, indicando que ScCBL3 é parcialmente solúvel havendo detecção da banda de ScCBL3 com anticorpo anti-His de uma a cinco horas no pellet induzido da fração total. Já a detecção de bandas ScCBL3 da fração solúvel ocorreu somente nos pellets induzidos a três e seis horas e com uma intensidade bem menor, indicando que a quantidade de proteína solúvel é menor (Figura 15). Testes para detecção da proteína ScCBL6 com anti-His confirmaram detecção da proteína a partir de extratos protéicos totais mas quase não houve detecção da proteína a partir do extrato solúvel (dados não mostrados).

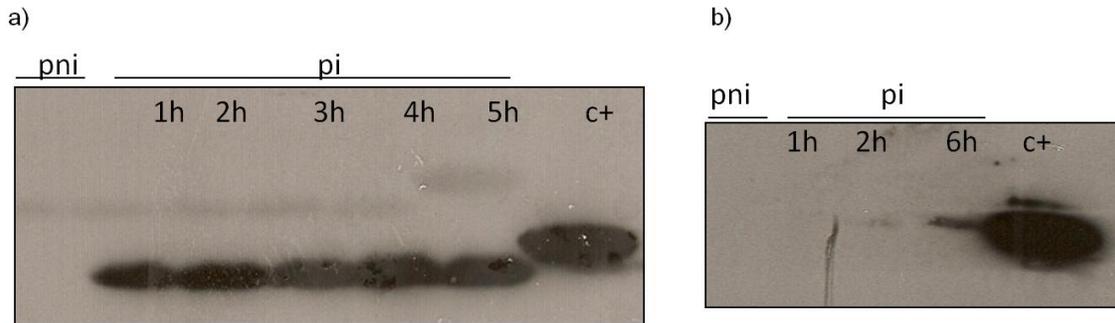


Figura 15: Detecção por *western blot* da proteína ScCBL3. a) fração total e b) fração solúvel. Detecção de bandas de ScCBL3 com anti-His da fração total ocorre com forte intensidade em todas as frações do pellet induzido (pi), já na fração solúvel ocorre em uma intensidade bem menor e somente a duas e seis horas do pellet induzido, confirmando que a proteína é apenas parcialmente expressa na fração solúvel.

Com o objetivo de obtermos a proteína ScCBL1 para uma caracterização estrutural, a fração solúvel de ScCBL1 foi purificada por cromatografia de afinidade usando uma resina de níquel Ni-NTA com um gradiente de concentração de 10 a 200 mM de imidazol (Figura 16). A proteína foi eluída da coluna de níquel a 200 mM de imidazol em um alto grau de pureza.

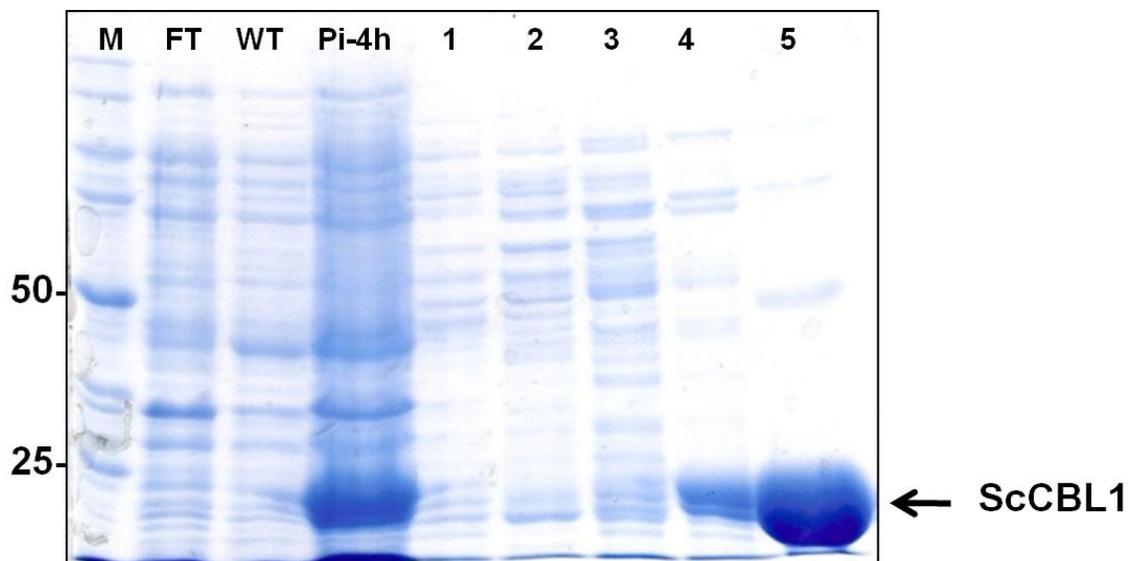


Figura 16: Purificação da proteína ScCBL1. A análise da purificação da proteína ScCBL1 foi feita em SDS-PAGE a 12%. Indução de 500 mL de cultura da cepa BL21 (DE3). Purificação foi feita por cromatografia de afinidade com resina Ni-NTA. M :marcador de peso molecular (kDa), FT: Flow-through, WT: Wash-through, Pi-4h: Pellet induzido 4 horas, 1: 10 mM imidazol, 2: 25 mM imidazol, 3: 50 mM imidazol, 4: 100 mM imidazol e 5: 200 mM imidazol.

As proteínas ScCBL3 e ScCBL6 também foram purificadas em uma resina Ni-NTA e eluídas da coluna a 200 mM de imidazol (Figuras 17 e 18, respectivamente). Essas proteínas não apresentaram um grau de pureza tão elevado quanto ScCBL1 sendo necessárias novas etapas de purificação por cromatografia de troca iônica ou gel filtração para obtenção de uma proteína mais pura. Como a proteína escolhida para uma caracterização estrutural mais detalhada neste trabalho foi a ScCBL1 estas etapas poderão ser realizadas futuramente de maneira mais detalhada.

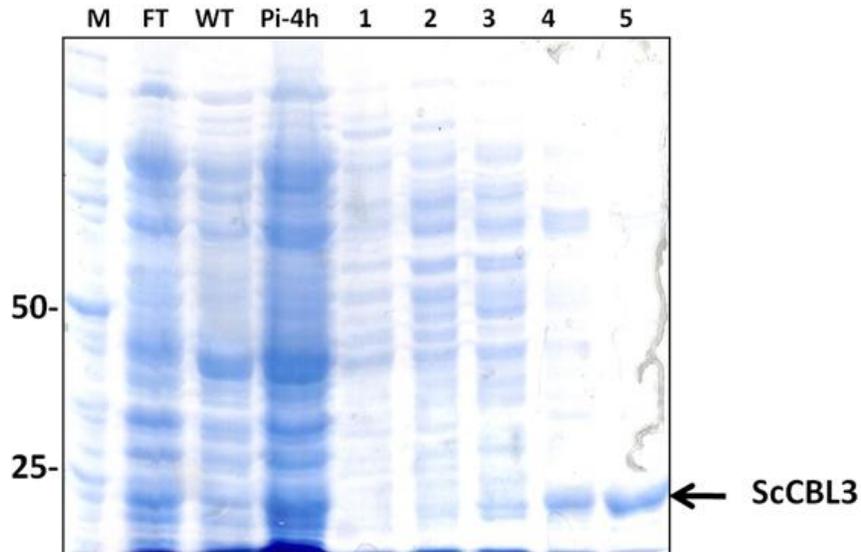


Figura 17: Purificação da proteína ScCBL3. A análise da purificação da proteína ScCBL3 foi realizada em SDS-PAGE a 12%. Indução de 500 mL de cultura da cepa BL21 (DE3). Purificação foi feita por cromatografia de afinidade com resina Ni-NTA. M :marcador de peso molecular (kDa), FT: Flow-through, WT: Wash-through, Pi-4h: Pellet induzido 4 horas, 1: 10 mM imidazol, 2: 25 mM imidazol, 3: 50 mM imidazol, 4: 100 mM imidazol e 5: 200 mM imidazol.

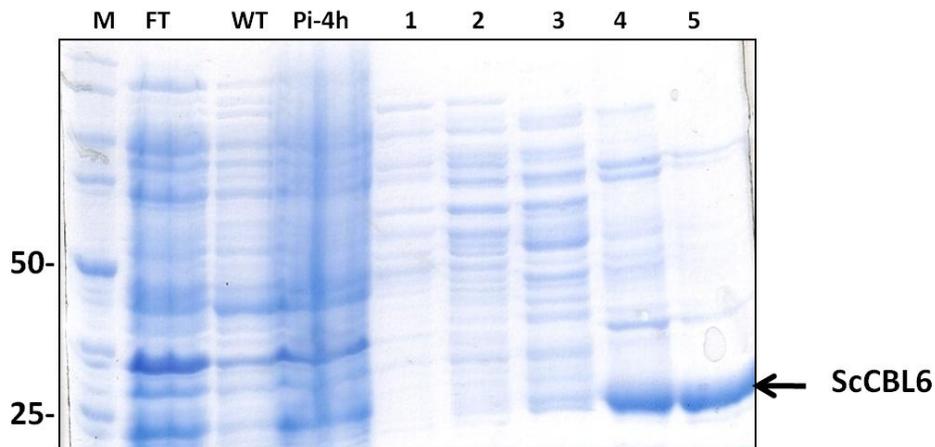


Figura 18: Purificação da proteína ScCBL6. A análise da purificação da proteína ScCBL6 foi realizada em SDS-PAGE a 12%. Indução de 500 mL de cultura da cepa BL21 (DE3). Purificação foi feita por cromatografia de afinidade com resina Ni-NTA. M :marcador de peso molecular (kDa), FT: Flow-through, WT: Wash-through, Pi-4h: Pellet induzido 4 horas, 1: 10 mM imidazol, 2: 25 mM imidazol, 3: 50 mM imidazol, 4: 100 mM imidazol e 5: 200 mM imidazol.

## 4.2. Caracterização da estrutura secundária de ScCBL1 por dicroísmo circular (CD)

O espectro gerado por CD da proteína ScCBL1 demonstrou que sua estrutura secundária é formada predominantemente por  $\alpha$ -hélices. As medidas de CD para ScCBL1 foram feitas no UV-distante em comprimento de onda de 190 a 260 nm. O espectro de ScCBL1 observado na figura 19 apresenta picos negativos a 208 e 220 nm e um pico positivo entre 190 e 195 nm, típicos de proteínas formadas predominantemente por  $\alpha$ -hélices. O valor da voltagem máxima medida para a amostra utilizada nos experimentos de CD da proteína ScCBL1 em HT [V] (*high tension voltage*) foi de 542,96 volts para o comprimento de onda de 190 nm, ficando portanto abaixo de 600 volts, a partir do qual o detector ficaria saturado.

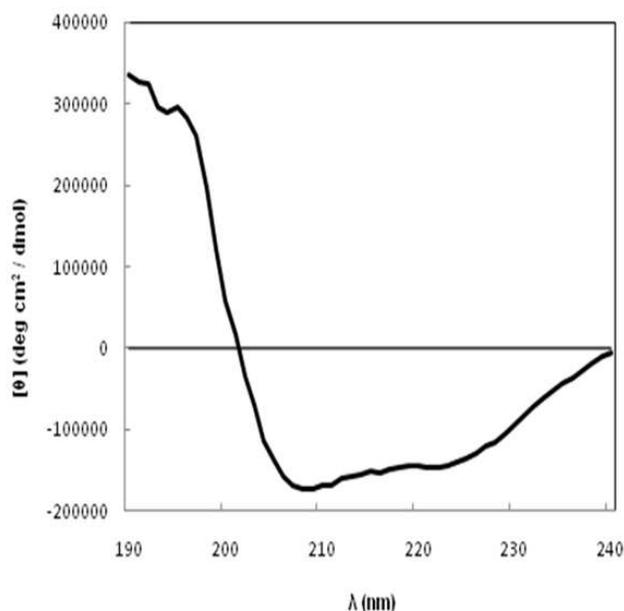


Figura 19: Análise da proteína ScCBL1 por dicroísmo circular. Espectro de CD da proteína ScCBL1 a 0.35 mg/mL (~14  $\mu$ mol/L) em 20 mM tampão fosfato de sódio pH 7.4 a 4°C obtido com o programa Dychroweb, usando o algoritmo CDSSTR.

A deconvolução do espectro de CD de ScCBL1 usando o algoritmo CDSSTR do programa DychroWeb indicou uma proteína composta por 70% de  $\alpha$ -hélices, 13% de folhas  $\beta$ , 5% de dobras  $\beta$  e 12% de estrutura randômica.

Estes dados estão próximos a predição da estrutura secundária a partir da sequência de aminoácidos de ScCBL1 feita com o programa PredictProtein Este programa estimou um conteúdo de 52% de  $\alpha$ -hélices, 2.3% de folhas  $\beta$  e 45.5% de estrutura randômica, caracterizando assim uma proteína formada predominantemente por  $\alpha$ -hélices.

#### **4.3. Caracterização por espalhamento dinâmico de luz (DLS)**

Ensaio de DLS de amostras da proteína ScCBL1 em solução foram feitos com o propósito de identificarmos amostras monodispersas para os experimentos de SAXS e ensaios de cristalização. A proteína ScCBL1 concentrada a 11 mg/ml com 3 mM do agente redutor DTT apresentou um pico predominante de espalhamento com uma intensidade de 88,8 %, representando 99,9 % da massa total em solução e um índice de polidispersão de 13%, caracterizando uma amostra monodispersa (Figura 20).

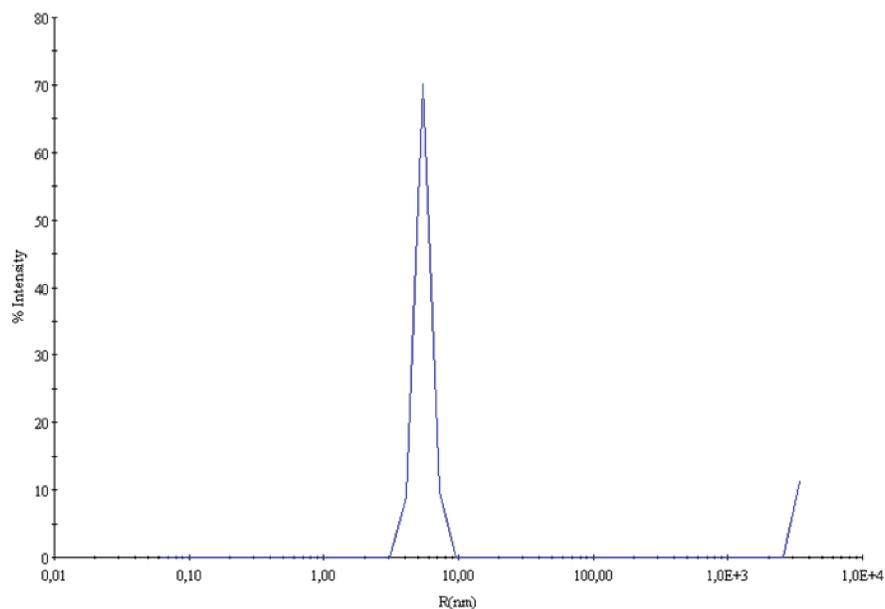


Figura 20: Análise do espalhamento de luz dinâmico da proteína ScCBL1. Gráfico característico de monodispersão da amostra da proteína ScCBL1 concentrada a 11 mg/ml.

A amostra da proteína ScCBL1 a 15 mg/mL apresentou um perfil semelhante com um pico predominante de espalhamento com uma intensidade de 89,6 %, representando 99,9 % da massa total em solução e um índice de polidispersão de 28,7 %, caracterizando uma amostra médio- dispersa (Figura 21).

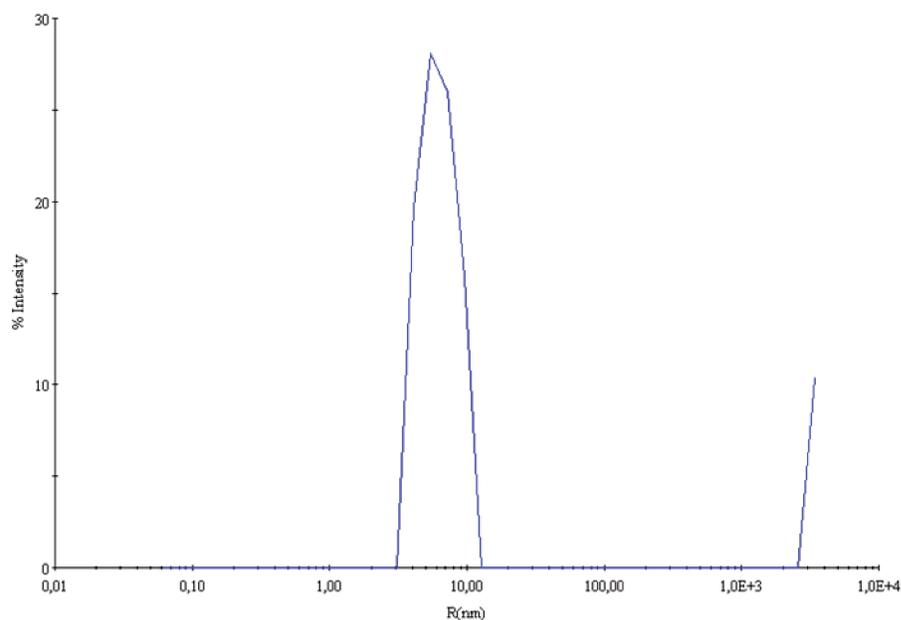


Figura 21: Análise do espalhamento de luz dinâmico da proteína ScCBL1. Gráfico característico de monodispersão da amostra da proteína ScCBL1 concentrada a 15 mg/ml.

Ambas as amostras concentradas a 11 e 15 mg/mL apresentaram um único pico predominante de espalhamento da proteína ScCBL1 característicos de amostras monodispersas e índices de polidispersão adequados (abaixo ou perto de 20%). Portanto, estas amostras foram encaminhadas para ensaios de cristalização.

#### 4.4. Ensaios de cristalização

Ensaios iniciais de cristalização foram realizados para a obtenção de cristais da proteína ScCBL1 para resolução da estrutura tridimensional da proteína via difração por raios-x. Em um *screening* inicial foram testadas três concentrações de amostras de ScCBL1

monodispersas (3, 11 e 15 mg/ml), seis kits e aproximadamente mil e seiscentas condições de cristalização da proteína ScCBL1 pela técnica de difusão de vapor através de gota sentada (“sitting drop”).

Houve formação de um monocristal simétrico da proteína ScCBL1 (Figura 22) a uma concentração de 15 mg/ml, de aproximadamente 0,05 mm (50 micrômetros) obtido com o kit Crystal Screen I e II (Hampton Research) na seguinte condição: 100 mM de tampão HEPES pH 7,5, 5% p/v MPD e 10% v/v PEG 6000.



Figura 22: Cristalização da proteína ScCBL1. Monocristal de aproximadamente 0,05 mm da proteína ScCBL1 a 15 mg/ml obtido com o kit Crystal screen I e II (Hampton Research) na seguinte condição: 100 mM de tampão HEPES pH 7.5, 5% p/v MPD e 10% v/v PEG 6000.

Foram observados também diversos sinais de cristalização promissores que serão refinados com a formação de micro-cristais e quasi-cristais (Figura 23). Algumas destas condições são promissoras para um refinamento via alteração de pH e concentração do precipitante. Outras amostras foram falso-positivos, como a presença de alguns cristais de sal que foram identificados pela adição do corante de proteínas Izit.

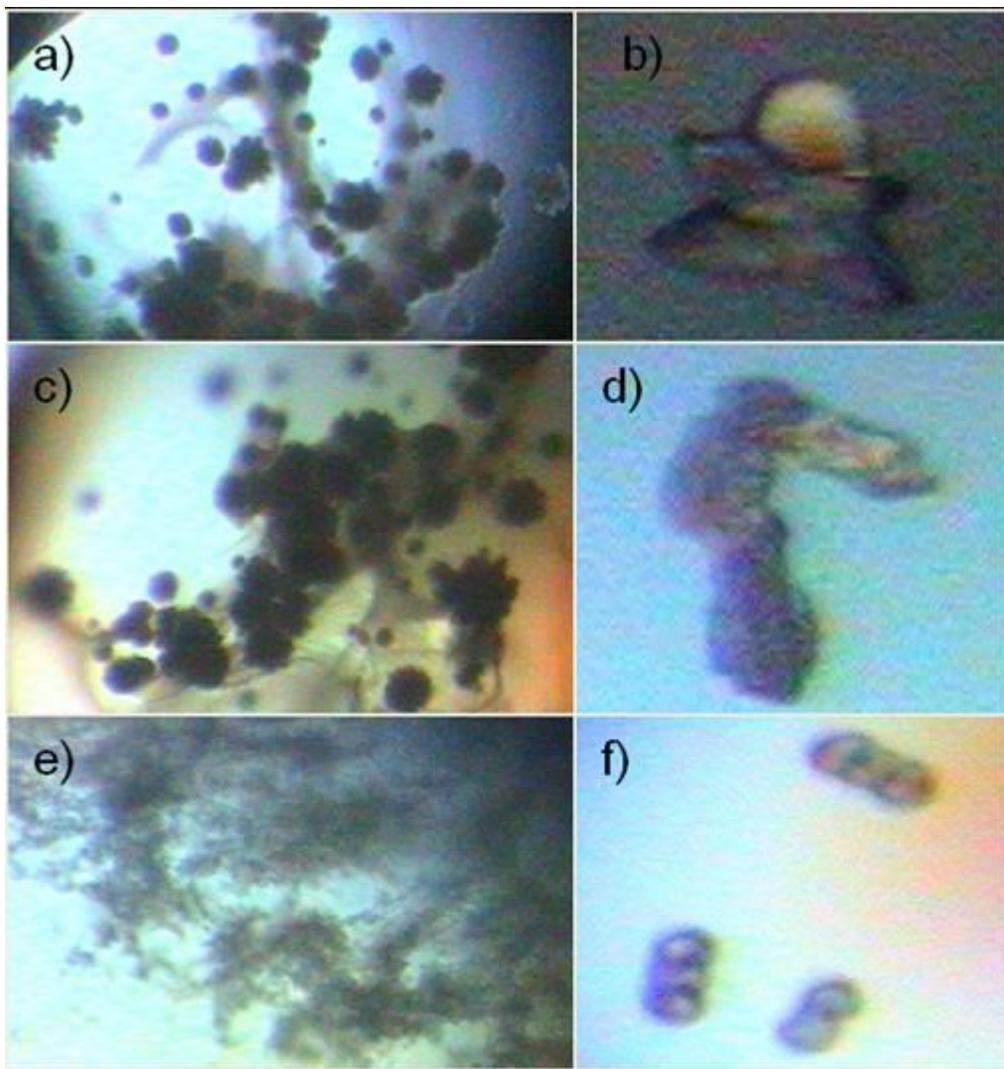


Figura 23: Sinais cristalográficos obtidos em um *screening* da proteína ScCBL1. A) e C) Microcristais; B) Quasi-cristal; D) Cristal amorfo; E) Microcristais / “Needle-cluster” F) Esferolitos (quasi-cristal).

#### 4.5. Caracterização estrutural por espalhamento de raios x a baixos ângulos (SAXS)

Análises iniciais de SAXS tiveram como objetivo a obtenção de informações estruturais da proteína ScCBL1 tais como: caracterização do envelope protéico de ScCBL1 (superfície 3-D que define o formato da proteína), o estado oligomérico das amostras tratadas com 3 mM de DTT e sem o agente redutor, estado de enovelamento e peso molecular das amostras em solução. As amostras de ScCBL1 utilizadas para SAXS foram

obtidas a uma alta concentração em uma etapa de purificação por afinidade usando uma coluna de níquel sem haver necessidade de concentrá-las (figura 24). As medidas de SAXS foram feitas com duas concentrações diferentes de ScCBL1 – 3,5 mg/ml e 11,0 mg /ml, com e sem adição de DTT, totalizando quatro amostras.

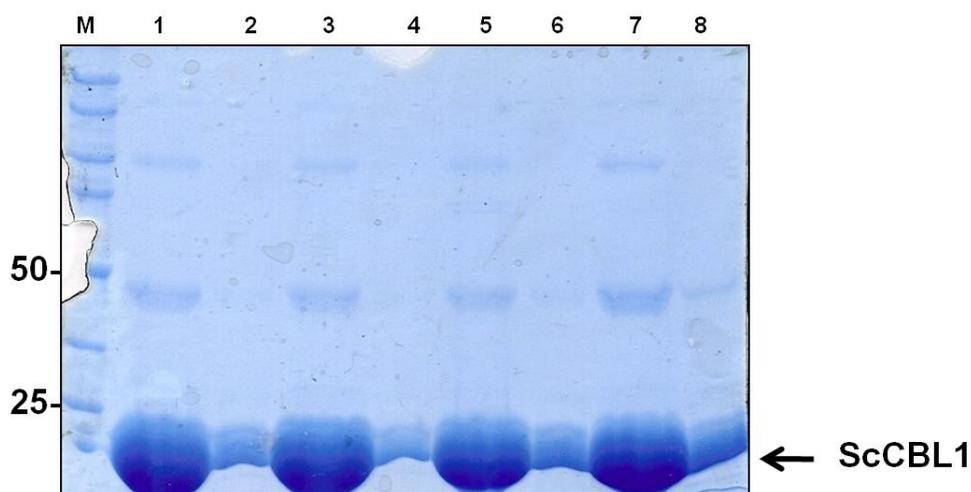


Figura 24: Proteína ScCBL1 usada nos experimentos de SAXS. Gel SDS-PAGE a 12% das amostras da proteína ScCBL1 usadas nas medidas de SAXS purificadas por uma rodada de afinidade em coluna Ni-NTA. Somente amostras eluídas a 200 mM de imidazol foram usadas (poços ímpares). 1, Amostra eluída a 200 mM de imidazol a 11 mg/ml; 2, amostra eluída a 300 mM imidazol a 1.2 mg/ml; 3, Eluída a 200 mM a 10.95 mg/ml; 4, 300 mM / 0.8 mg/ml; 5, 200 mM / 11mg/ml; 6, 300 mM/ 1.6 mg/ml; 7 200 mM/ 15 mg/ml; 8, 300 mM/ 2.6 mg/ml. M, Marcador de peso molecular (kDa).

Dados estruturais referentes ao estado de agregação da molécula em solução, peso molecular (massa) e estado oligomérico podem ser estimados por SAXS a partir de uma regressão linear na região de Guinier, bem como a partir da curva de espalhamento completa. As curvas de espalhamento geradas a partir de amostras da proteína ScCBL1 a diferentes concentrações, com e sem DTT, podem ser visualizadas na figura 25. Como a

concentração está relacionada à intensidade de espalhamento, pode-se observar que as amostras mais concentradas apresentaram uma intensidade de espalhamento maior do que as amostras menos concentradas. É interessante notar que as curvas de espalhamento de proteínas com a mesma concentração divergem no início da curva de espalhamento. Essa diferença pode ser explicada em termos de uma mudança no estado oligomérico das amostras tratadas com DTT. O DTT é um agente redutor de pontes dissulfeto formadas entre dois resíduos de cisteína, podendo mudar o estado oligomérico da proteína e portanto, a massa detectada.

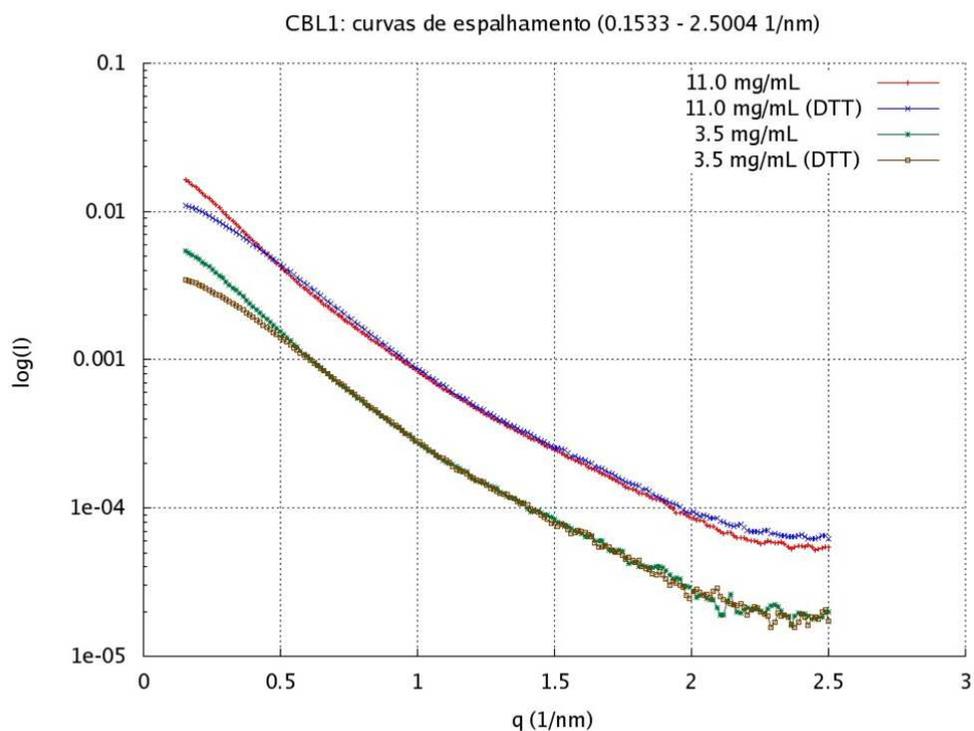


Figura 25: Curvas de espalhamento de intensidade obtidas por SAXS ( $I \times q$ ) a partir de amostras da proteína ScCBL1 medidas a diferentes concentrações (3,5 e 11 mg/mL), tratadas com 3 mM DTT e sem DTT (ver legenda). Pode-se notar que a intensidade do espalhamento aumenta em amostras mais concentradas.

O comportamento linear observado no gráfico de Guinier (Figura 26) indica que para todas as amostras, a proteína encontrava-se monodispersa em solução. A regressão linear na região de Guinier levou a estimativa de Raio de Giro ( $R_G$ ) e intensidade na origem ( $I(0)$ ) indicadas nas figuras. Podemos observar que houve uma diminuição na intensidade de espalhamento  $I(0)$  de amostras contendo 3 mM de DTT e amostras sem DTT (0,0194/0,0124 para a amostra de 11 mg/mL e 0,0063/0,0039 para a de 3,5 mg/mL), juntamente com uma significativa redução no raio de giro, indicando possivelmente uma dissociação da proteína e portanto alteração de seu peso molecular. Como o peso molecular do monômero de ScCBL1 é de 25 kDa, a partir destas últimas análises de SAXS e da avaliação de um padrão de calibração (BSA), foi possível estimar o peso molecular em aproximadamente 75 kDa para amostras com DTT (em ambas as concentrações de proteína) e aproximadamente 125 kDa para amostras sem DTT (em ambas as concentrações). Esses dados indicam uma alteração no estado oligomérico de ScCBL1 na presença do agente redutor DTT, independente da concentração da amostra. A partir dos valores  $I(0)$  obtidos é possível sugerir que um trímero é o maior espalhador em amostras com 3mM de DTT e um pentâmero em amostras sem DTT. As estimativas de massa foram corroboradas por um método recente (Fisher et al., 2010) usando dados experimentais de uma única curva de espalhamento medida por SAXS em uma escala relativa. Estas análises da massa molecular da proteína ScCBL1 representam estimativas iniciais coletadas na presença de imidazol cujo efeito no comportamento da proteína em solução ainda é desconhecido. Entretanto os dados confirmam que há uma mudança de massa da ScCBL1 em solução que é relacionada a uma alteração do estado oligomérico da proteína mediante a presença / ausência de DTT.

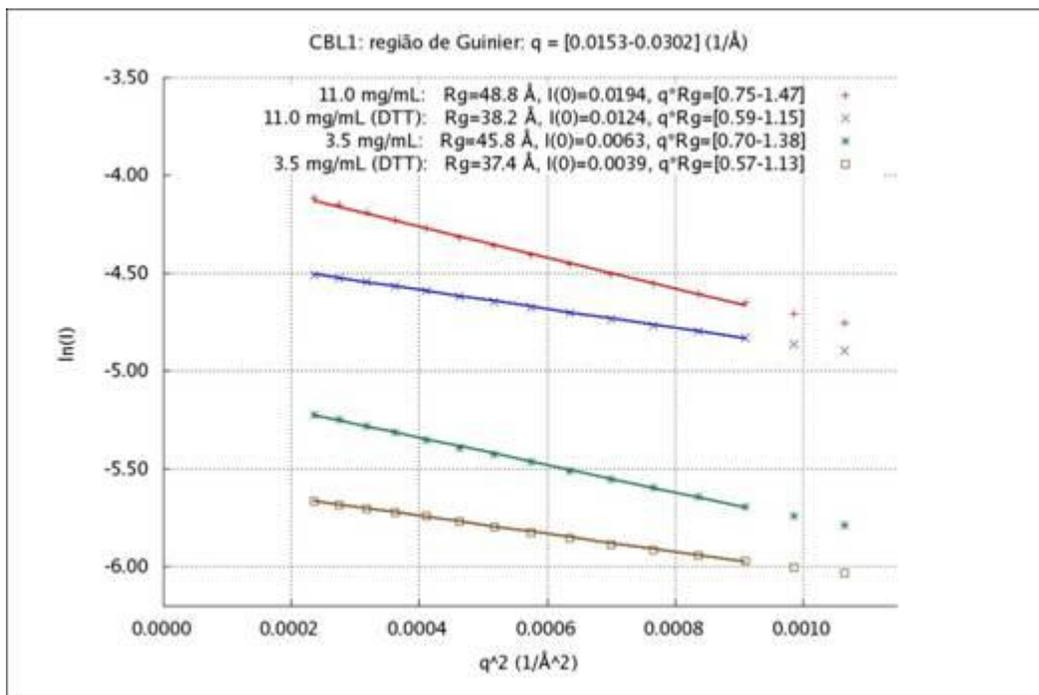


Figura 26: Gráfico de Guinier ( $\log(I(q))$  contra  $q^2$ ) para amostras da proteína ScCBL1 em diferentes concentrações, com e sem DTT (ver legenda) apresentaram um padrão linear que não indica agregação e é característico de amostras que apresentam comportamento monodisperso. As variáveis  $R_G$  e  $I(0)$  foram estimadas pela regressão linear.  $q$ : [ $q_{\min}$  –  $q_{\max}$ ] região utilizada para fazer a regressão linear (ver linha contínua).  $qR_G$ : Parâmetro que indica se estamos na região de validade na teoria de Guinier, o que foi atendido em todos os casos.

O gráfico de Kratky, que pode ser visualizado na figura 27, exibe um máximo e um decaimento característico de proteínas compactas, com enovelamento terciário estável em solução.

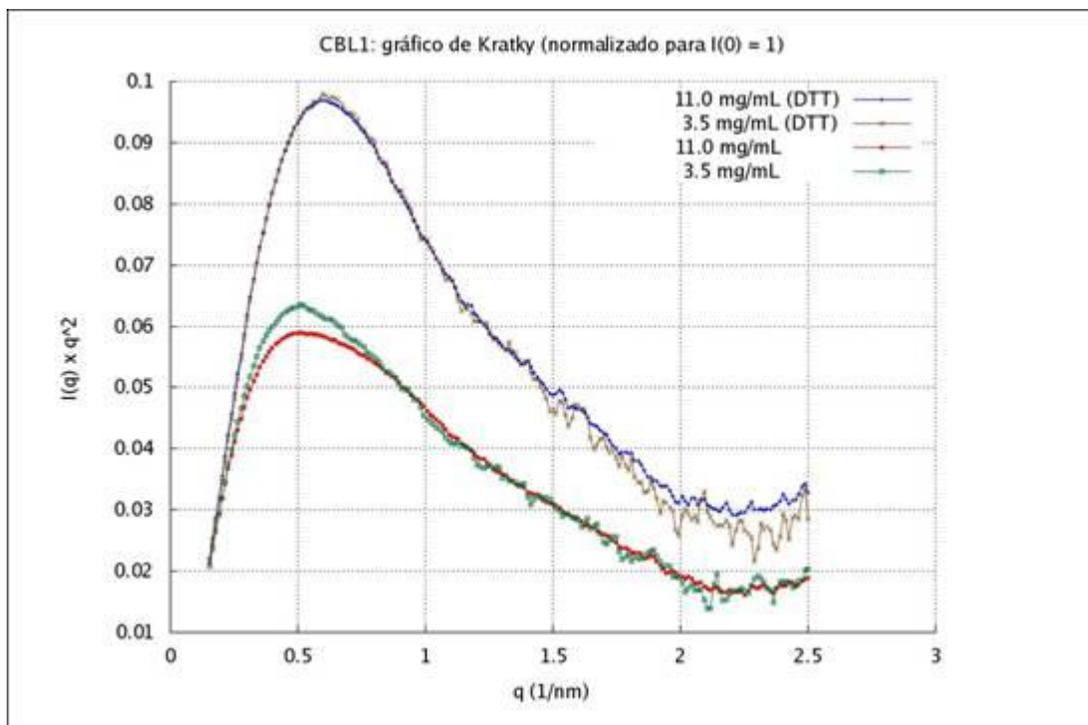


Figura 27: Gráfico de Kratky de amostras da proteína ScCBL1 a diferentes concentrações, tratadas com e sem DTT (ver legenda), representado neste gráfico como uma função de Intensidade ( $I(q) \times q^2$  (eixo x) contra  $q$  (eixo y). Os gráficos obtidos mostram um máximo para menores ângulos, característico de proteínas enoveladas.

A função de distribuição de distâncias,  $P(r)$ , calculada a partir da curva de espalhamento, indica que a proteína possui forma alongada (prolata) em solução como pode ser visualizado na figura 28. Proteínas prolatas possuem um máximo em distâncias menores correspondendo ao raio da seção transversal e, à medida que as distâncias aumentam, a função decresce. As estimativas de  $R_G$  resultantes a partir do cálculo da  $P(r)$  concordam bem com aquelas obtidas por regressão linear na região de Guinier (Figuras 23 e 24).

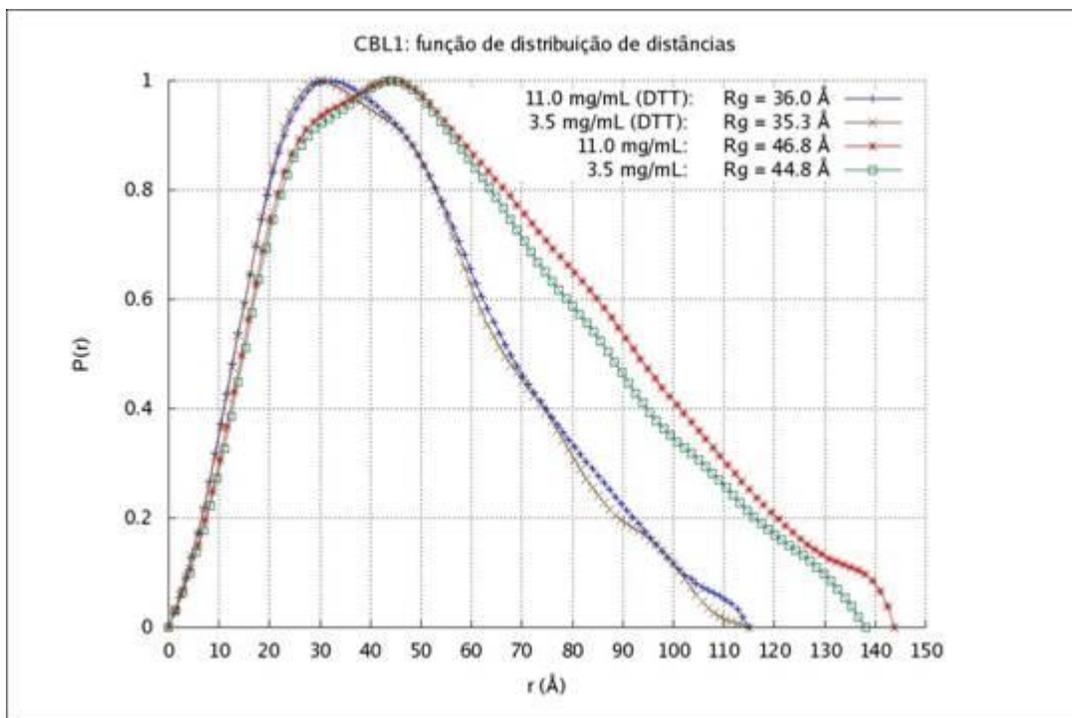


Figura 28: Gráfico da função de distribuição das distâncias  $P(r)$  de amostras da proteína ScCBL1 em diferentes concentrações, com e sem DTT (ver legenda), apresentam uma função com uma cauda longa a distâncias grandes  $r$ , característica de proteínas com um formato alongado ou prolato.  $r$  (Å): distância em angstroms.  $R_G$ : Raio de giro.

Os envelopes protéicos construídos a partir das curvas de espalhamento (sem imposição de simetria) podem ser visualizados nas figuras 29 a 31. Nestas figuras, para comparação de tamanho, foi incluída a estrutura de um monômero da homóloga AtCBL2 de *Arabidopsis thaliana* (Código PDB 1UHN; Nage et al., 2003). Como esperado, os modelos obtidos apresentam uma forma prolata com dimensão máxima e seção transversal compatíveis com as funções  $P(r)$  da figura 24.

Apesar de SAXS não ser a técnica mais adequada para uma estimativa precisa da massa da proteína em solução, novos experimentos de SAXS da ScCBL1, na ausência de

imidazol, que talvez possa interferir na análises de SAXS, poderão ser realizados futuramente de modo a entendermos melhor o comportamento da ScCBL1 em solução.

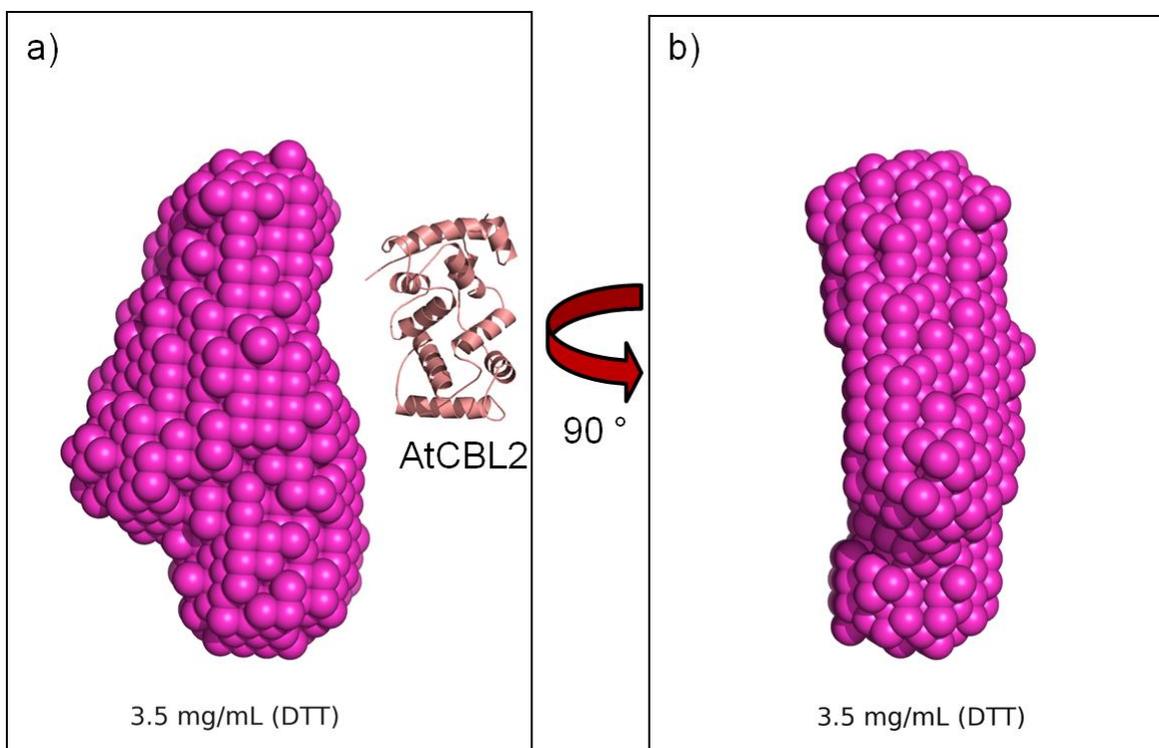


Figura 29: Envelope de baixa resolução apresentando um formato alongado da proteína ScCBL1 na presença de DTT. a) modelo do envelope protéico de ScCBL1 e estrutura resolvida do homólogo AtCBL2 para comparação de tamanho na mesma escala. b) rotação de 90° do modelo.

Análise do comportamento da ScCBL1 na presença de diferentes concentrações de cálcio por SAXS também seriam interessantes, uma vez que a proteína possui domínios de ligação de moléculas de cálcio em sua sequência de aminoácidos assim como suas homólogas em *Arabidopsis* (Nagae et al., 2003). Cálcio é um mensageiro secundário importante em diversas vias de sinalização em resposta a estresse em plantas, e são alterações na concentração de cálcio intracelulares que são detectadas por proteínas sensoras de cálcio como a ScCBL1 (Iacovelli et al., 1999; Kim et al., 2000; Mahajan et al., 2008). Portanto novas análises por SAXS da ScCBL1 serão importantes para indicar se a

presença de cálcio poderia causar alguma alteração conformacional na estrutura da ScCBL1 e como esta possível alteração estaria relacionada a uma regulação funcional da ScCBL1 relacionada a alguma via de sinalização abiótica.

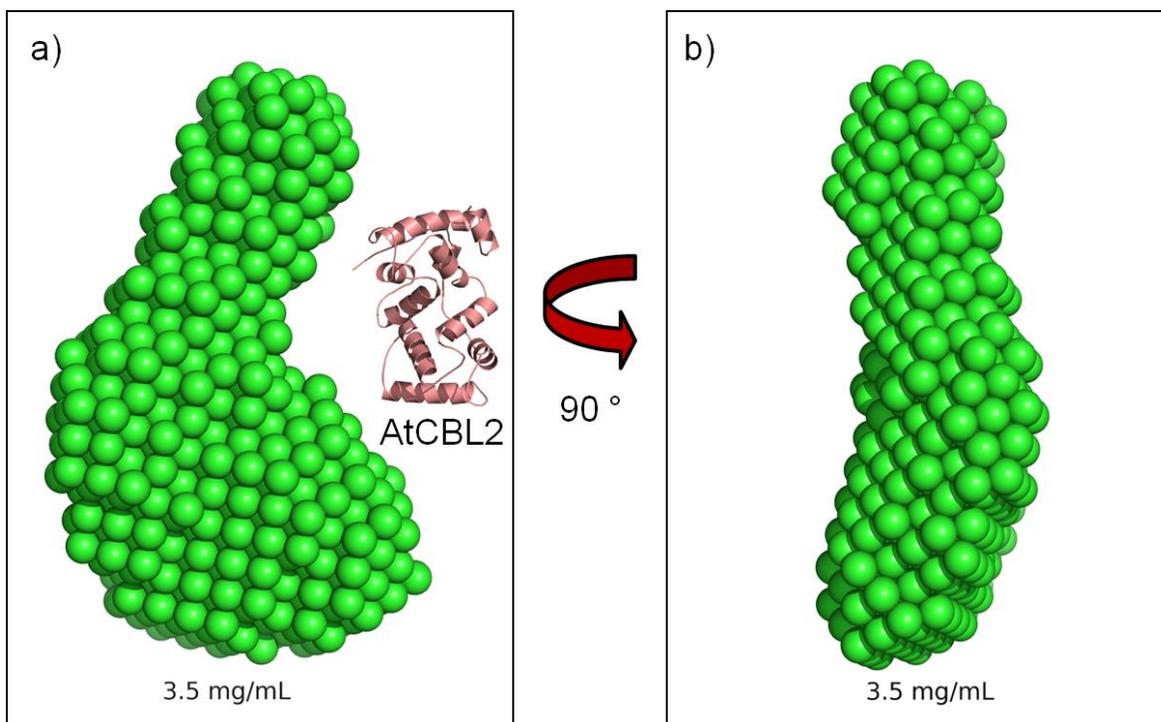


Figura 30: Envelope de baixa resolução apresentando um formato alongado da proteína ScCBL1 na ausência de DTT. a) modelo do envelope protéico de ScCBL1 e estrutura resolvida do homólogo AtCBL2 para comparação de tamanho na mesma escala. b) rotação de 90° do modelo.

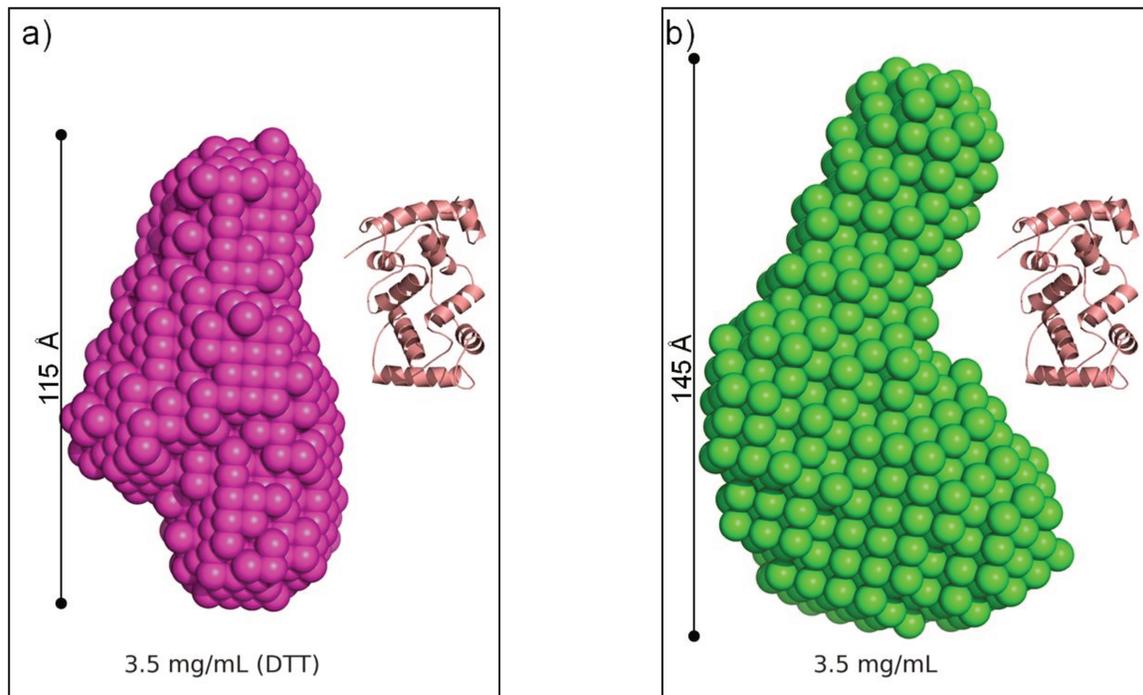


Figura 31: Comparação dos envelopes externos de baixa resolução gerados para amostras da ScCBL1 com DTT (a) cujas análises por SAXS indicam ser um trímero e sem DTT (b) que indica ser um pentâmero.

Dentre a família de proteínas CBL em plantas apenas uma, AtCBL2 de Arabidopsis, possui sua estrutura resolvida até hoje no banco de dados de proteínas PDB. Estudos demonstraram que o estado oligomérico desta proteína é de um monômero (Nagae et al., 2003; Akaboshi et al., 2008). As análises iniciais de SAXS para a proteína ScCBL1, indicando a presença de dois estados oligoméricos distintos em amostras com a mesma concentração – a de um trímero na presença de DTT e de um pentâmero sem DTT, são inesperados e merecem um estudo mais aprofundado. É interessante mencionar que uma percentagem significativa (35%) de proteínas celulares são oligoméricas, havendo uma suposição de que essas proteínas teriam características vantajosas do ponto de vista da evolução, como por exemplo novas oportunidades de controle funcional e regulações alostéricas (Goodsell e Olson., 2000; Mayssam e Imperiali., 2005). Diversas proteínas

ancestrais podem ter sido homo-oligoméricas ou hetero-oligoméricas para melhor desempenhar suas diversas funções, com o qual o estudo da natureza de oligomerização de proteínas poderia contribuir também para estudos da evolução das proteínas (Mayssam e Imperiali., 2005). Neste sentido, o estudo do estado oligomérico de ScCBL1 em cana de açúcar contribui para o entendimento da evolução de proteínas em plantas, adicionando informações interessantes do ponto de vista evolutivo e estrutural sobre a família de proteínas CBL de plantas.

Neste trabalho nós reportamos a clonagem, expressão, purificação, assim como uma análise estrutural inicial da proteína ScCBL1 de cana-de-açúcar por difração circular, SAXS, DLS e ensaios cristalográficos. A proteína ScCBL1 possui uma identidade de sequência de aminoácidos de 78% a AtCBL9, que está relacionada à regulação negativa em resposta a ABA, 77% de identidade a AtCBL1 que está envolvido em resposta a seca, frio e sal e 66% identidade a AtCBL2, cuja estrutura foi resolvida por cristalografia de raios-x. Em um alinhamento múltiplo da proteína ScCBL1 com proteínas AtCBL de *Arabidopsis* podemos visualizar as regiões ligantes de cálcio ou “EF Hands” e outras regiões conservadas entre estas sequências (Figura 32). Interessantemente, ScCBL1, assim como AtCBL1, AtCBL9 e AtCBL4 (SOS3), possui os motivos conservados de miristoilação/palmitolação em sua região N-terminal (ver anexo III), aumentando consideravelmente a probabilidade desta proteína estar associada a membrana plasmática (Mahajan et al., 2008; Bijlmakers e Marsh., 2003). ScCBL1 é a única das proteínas ScCBLs analisadas neste trabalho que possui um *hit* positivo para um sítio de miristoilação. Algumas vias de sinalização como tolerância a salinidade requerem esta modificação pós-traducional (Ishitani et al., 2000).

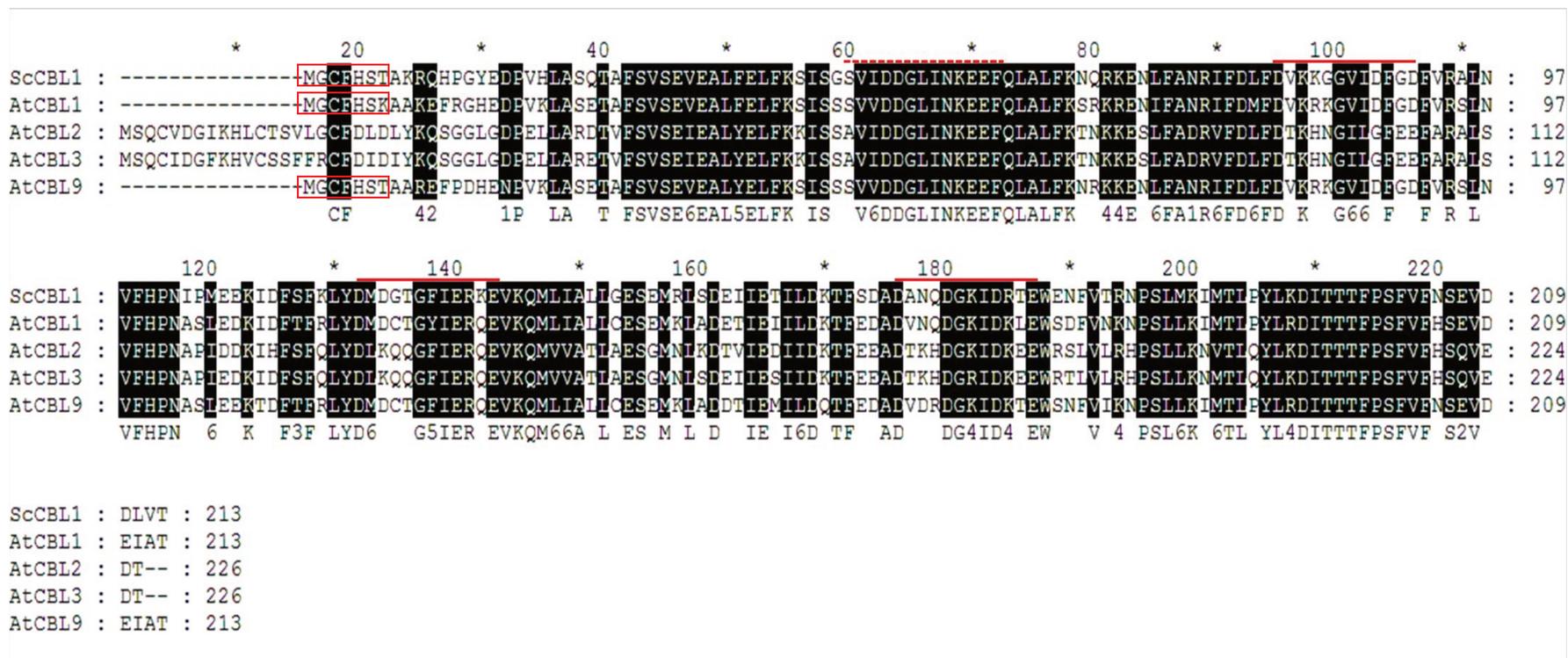


Figura 32: Alinhamento múltiplo da CDS da proteína ScCBL1 de cana-de-açúcar e proteínas AtCBL de Arabidopsis (AtCBL1, AtCBL2, AtCBL3 e AtCBL9). Alinhamento foi gerado com o programa ClustalX. Aminoácidos apresentando resíduos idênticos a sequência consenso do alinhamento estão marcados em preto. As sequências mostradas dentro das caixas representam motivos de miristoilação conservados (GXXXST/K) (Resh., 1999; Kudla et al., 1999). Números a direita indicam a posição dos aminoácidos. Linhas sólidas em vermelho acima do alinhamento representam a posição das regiões ligantes de cálcio ou “EF-Hands”. A linha pontilhada em vermelho representa uma variação da região de ligação de cálcio.

A caracterização estrutural da proteína sensor de cálcio ScCBL1 em cana de açúcar é importante para obtermos informação a respeito da função desta proteína na via de sinalização ativada pela planta em resposta a estresses abióticos; na qual o cálcio possui um papel chave (Iacovelli et al., 1999; Kim et al., 2000; Luan et al., 2002). De fato, sabe-se que alterações nas concentrações de cálcio promovem a interação de proteínas AtCBL com proteínas quinase (AtCIPKs) (Luan et al., 2002). O complexo CBL-CIPK é então direcionado para a membrana permitindo assim a fosforilação através de CIPK, de proteínas associadas à membrana (Luan et al., 2002). Este complexo CBL-CIPK foi relacionado a diversas vias de sinalização em resposta a estresses abióticos em Arabidopsis (Kudla et al., 1999; Pandey., 2008; Batistic e Kudla., 2004; Nagae et al., 2003; Akaboshi et al., 2008). A alta similaridade apresentada entre as proteínas AtCBL em Arabidopsis e a ScCBL1 em cana-de-açúcar caracterizada aqui sugere que esta proteína pode desempenhar um papel similar na resposta a estresses abióticos como seca, salinidade e ABA em cana-de-açúcar.

## 5. CONCLUSÕES

- A proteína ScCBL1 é solúvel, podendo ser obtida em um alto grau de pureza por uma etapa de cromatografia de afinidade, enquanto as proteínas ScCBL3 e ScCBL6 são parcialmente solúveis e necessitam de pelo menos uma segunda etapa de purificação.
- A proteína ScCBL1 possui uma estrutura secundária formada predominantemente por  $\alpha$ -hélices.
- Experimentos de SAXS realizados com amostras de ScCBL1 indicam que as amostras estavam homogêneas e monodispersas.
- Experimentos de SAXS da ScCBL1 indicam que em ambas as concentrações avaliadas, adição do agente redutor DTT causa uma mudança no estado oligomérico da proteína com redução da intensidade de espalhamento acompanhado de uma redução de seu raio de giro.
- Análises iniciais por SAXS do estado oligomérico e peso molecular de ScCBL1 indicam que a proteína em estado nativo é um pentâmero de aproximadamente 125 kDa, que pode ser convertido em um trímero de 75 kDa quando tratada com DTT.
- Análise do envelope protéico da proteína ScCBL1 por SAXS indicam que a proteína possui um formato prolato (alongado) e que ScCBL1 mantém sua estrutura em solução.

- A detecção de dois estados oligoméricos da proteína ScCBL1 abre a perspectiva de que este é ponto de regulação da sua atividade.
- Houve formação de um monocristal de aproximadamente 50 micrometros de diâmetro, demonstrando que a proteína forma cristais, abrindo perspectivas para estudos de cristalografia.

## II. Conclusão Geral e Perspectivas

Neste trabalho foi feita uma caracterização funcional do sitio de ligação do fator de transcrição TGA2 na região promotora do gene *NIMIN-1* bem como a caracterização da proteína ScCBL1. Tanto o gene *NIMIN-1* como a proteína ScCBL1 estão relacionadas a vias de sinalização em resposta a estresse biótico e abiótico, respectivamente. A caracterização destes componentes é um passo importante no estabelecimento de como estas redes de transcrição operam tanto na defesa de plantas como na resposta a estresses abióticos, como seca e salinidade por exemplo.

Foi demonstrado aqui que a mutação no motivo de ligação do fator TGA2 reduziu para apenas um terço a expressão de GUS induzida por SA. Este resultado é um importante passo na caracterização funcional da região promotora que regula a expressão de um gene relacionado a defesa em plantas. Experimentos adicionais através de *gel shift* e também usando mutante background triplo TGA2,5,6 seriam interessantes para aprofundar estas análises do impacto da mutação no motivo mutagenizado na ligação do fator TGA2. Os dados apresentados aqui relacionados à análise funcional das regiões promotoras mutagenizada e selvagem de *NIMIN-1* comprovaram que o motivo TGACG está

diretamente envolvido na regulação da expressão gênica induzida por SA do gene *NIMIN-1*, contribuindo assim para uma análise de seu significado biológico em SAR.

Neste trabalho foi feito também a caracterização de uma proteína de cana de açúcar que apresenta uma alta similaridade (>70%) com membros da família de proteínas sensores de cálcio em *Arabidopsis*, relacionadas a vias de sinalização em resposta a diversos estresses abióticos como seca (*AtCBL1*) e ABA (*AtCBL9*). A proteína *ScCBL1* apresenta regiões conservadas (*EF-hands*) em sua sequência, em comum com as *AtCBLs* que se ligam a moléculas de cálcio, um mensageiro secundário da via de sinalização de estresses abióticos em plantas. O gene *ScCBL1* foi clonado em um vetor de expressão e a proteína foi expressa e purificada a um alto teor de pureza por cromatografia de afinidade usando coluna de níquel para estudos estruturais. A estrutura secundária da *ScCBL1* foi determinada através de dicroísmo circular, sendo formada predominantemente por  $\alpha$ -hélices (*all alpha*) de maneira similar a um homólogo em *Arabidopsis* (*AtCBL2*) cuja estrutura foi resolvida por cristalografia. Experimentos de SAXS realizados com amostras da *ScCBL1* indicam que as amostras estavam homogêneas e monodispersas. Nas duas concentrações avaliadas, adição do agente redutor DTT causa uma mudança no estado oligomérico da proteína com redução da intensidade de espalhamento, acompanhado de uma redução de seu raio de giro, indicando uma possível dissociação do complexo protéico através da quebra das pontes dissulfeto. A detecção de dois estados oligoméricos da proteína *ScCBL1* abre a perspectiva de que este é ponto de regulação da sua atividade. Experimentos adicionais poderão ser realizados com amostras da *ScCBL1* na presença de cálcio para se observar uma possível alteração conformacional do envelope mediante presença de moléculas de cálcio ligas a *EF-hands*. Outras análises por SAXS indicaram que

a proteína ScCBL1 possui um formato alongado (prolato) e que a proteína mantém sua estrutura em solução. Análises cristalográficas da proteína ScCBL1 mostraram que esta forma cristais, abrindo perspectivas para estudos de cristalografia. Estas análises cristalográficas poderão ser utilizadas futuramente em conjunto com os resultados de SAXS para uma caracterização estrutural completa da proteína ScCBL1, permitindo assim um melhor entendimento de como sua estrutura esta relacionada a sua função na via de sinalização em resposta a estresses abióticos.

## 6. REFERÊNCIAS

Akaboshi M, Hashimoto H, Ishida H, Saijo S, Koizumi N, Sato M, Shimizu T (2008). The crystal structure of plant-specific calcium-binding protein AtCBL2 in complex with the regulatory domain of AtCIPK14. *J Mol Biol.* 377: 246-257.

Asherie A (2004). Protein crystallization and phase diagrams. *Methods.* 34: 266-272.

Batistic O, Kudla J (2004). Integration and channeling of calcium signaling through the CBL calcium sensor/CIPK protein kinase network. *Planta.* 219: 915-924.

Bijlmakers MJ, Marsh M (2003). The on-off story of protein palmitoylation. *Trends Cell Biol.* 13: 32-42.

Bjerrum OJ, Schafer-Nielsen C. Analytical Electrophoresis. Dunn MJ (Ed.). Berlin: Verlag Chemie, Weinheim, 1986. p.315.

Bush DS (1995). Calcium regulation in plant-cells and its role in signaling. *Plant Mol Biol.* 46: 95-122.

Chayen NE (2004). Turning protein crystallization from an art into a science. *Curr Opin Struct Biol.* 14: 577-583.

Chayen NE, Saridakis E (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nat Methods.* 5: 147-153.

Chernov AA (1997). Crystals built of biological macromolecules. *Phys Rep.* 288: 61-75.

Chikano H, Ogawa M, Ikeda Y, Koizumi N, Kusano T, Sano H (2001). Two novel genes encoding SNF-1 related protein kinases from *Arabidopsis thaliana*: Differential accumulation of AtSR1 and AtSR2 transcripts in response to cytokinins and sugars, and phosphorylation of sucrose synthase by AtSR2. *Mol Gen Genet.* 264: 674-681.

Compton LA, Johnson WC Jr (1986). Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem.* 155: 155-167.

Corrêa DHA, Ramos CHI (2009). The use of circular dichroism spectroscopy to study protein folding, form and function. *African J Biochem Res.* 3: 164-173.

Glatter O, Kratky O. Small Angle X-ray Scattering. London: Academic, 1982. p.525.

Goodsell DS, Olson AJ (2000). Structural Symmetry and Protein Function. *Annu Rev Biophys Biomol Struct.* 29: 105-153.

Guinier A, Fournet G. Small Angle Scattering of X-rays. New York: Wiley, 1955. p.268.

Ferre-D'amare AR, Burley SK (1994). Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies. *Structure.* 2: 357-359.

Finkelstein R, Gampala S, Rock C (2002). Abscisic acid signaling in seeds and seedlings. *Plant Cell.* 14: 15-45.

Fischer H, Oliveira Neto M, Napolitano HB, Polikarpov I, Craievich F (2010). Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J Appl Cryst.* 43: 101-109.

Hammersley AP, Svensson SO, Hanfland M, Fitch NA, Häusermann D (1996). Two-dimensional detector software: From real detector to idealized image or two-theta scan *High Press Res.* 14: 235-248.

Harmon AC, Gribskov M, Harper JF (2000). CDPKs - a kinase for every Ca<sup>2+</sup> signal? *Trends Plant Sci.* 5: 154-159.

Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL, Tsutakawa SE, Jenney Jr FE, Classen S, Frankel KA, Hopkins RC, Yang S, Scott JW, Dillard BD, Adams MWW, Tainer JA (2009). Robust, high-throughput solution structural analyses by small X-ray scattering (SAXS). *Nat Methods.* 6: 606-612.

Iacovelli L, Sallese M, Mariggio S, De Blasi A (1999). Regulation of G-protein-coupled receptor kinase subtypes by calcium sensor proteins. *Faseb J.* 13: 1-8.

Ishitani M, Liu JP, Halfter U, Kim CS, Shi WM, Zhu JK (2000). SOS3 function in plant salt tolerance requires N-myristoylation and calcium binding. *Plant Cell*. 12: 1667–1677.

Johnson WC Jr (1985). Circular dichroism and its empirical application to biopolymers. *Meth Biochem Anal*. 31: 62-163.

Jones HG, Jones MB. Introduction: some terminology and common mechanisms. In: Jones HG, Flowers TG, Jones MB (Eds.). *Plants Under Stress*. Cambridge, Cambridge University Press, 1989. pp 1-10.

Kashchiev D. *Nucleation: Basic Theory with Applications*. Oxford, Butterworth-Heinemann, 2000. p.544.

Kim KN, Cheong YH, Gupta R, Luan S (2000). Interaction specificity of Arabidopsis calcineurin B-like calcium sensors and their target kinases. *Plant Physiol*. 124: 1844-1853.

Konarev PV, Petoukhov MV, Volkov VV, Svergun DI (2006). ATSAS 21, a program package for small-angle scattering data analysis. *J Appl Cryst*. 39: 277-286.

Koornneef M, Léon-Kloosterziel KM, Schwartz SH, Zeevaart JAD (1998). The genetic and molecular dissection of abscisic acid biosynthesis and signal transduction in *Arabidopsis*. *Plant Physiol Biochem*. 36: 83–89.

Kudla J, Xu Q, Harter K, Gruissem W, Luan S (1999). Genes for calcineurin B-like proteins in Arabidopsis are differentially regulated by stress signals. *Proc Natl Acad Sci USA* 96: 4718–4723.

Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227: 680-685.

Liu J, Zhu JK (1998). A calcium sensor homolog required for plant salt tolerance. *Science*. 280: 1943-1945.

Liu J, Ishitani M, Halfter U, Kim CS, Zhu JK (2000). The *Arabidopsis thaliana* SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proc Natl Acad Sci USA*. 97: 3730–3734.

- Luan S, Kudla J, Rodríguez-Concepción M, Yalovsky S, Gruissem W (2002). Calmodulins and calcineurin B-like proteins: Calcium sensors for specific signal response coupling in plants. *Plant Cell*. 14: 389–400.
- Mahajan S, Pandey GK, Tuteja N (2008). Calcium- and salt-stress signaling in plants: shedding light on SOS pathway. *Arch Biochem Biophys*. 471: 146–158.
- Mayssam HA, Imperiali B (2005). Protein Oligomerization: How and Why. *Bioorg Med Chem*. 13: 5013-5020.
- McPherson A. *Crystalization of Biological Macromolecules*. Cold Spring Harbor, CSHL Press, 1999. p.586.
- Mylonas E, Svergun DI (2007). Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J Appl Cryst*. 40: 245-249.
- McCarty DR, Chory J (2000). Conservation and innovation in plant signaling pathways. *Cell*. 103: pp. 201–209.
- Nagae M, Nozawa A, Koizumi N, Sano H, Hashimoto H, Sato M, Shimizu T (2003). The crystal structure of the novel calcium-binding protein AtCBL2 from *Arabidopsis thaliana*. *J Biol Chem*. 278: 42240–42246.
- Nozawa N, Koizumi, Sano H (2001). An Arabidopsis SNF1-related protein kinase, AtSR1, interacts with a calcium-binding protein, AtCBL2, of which transcripts respond to light, *Plant Cell Physiol*. 42: 976–981.
- Pandey GK (2008). Emergence of a novel calcium signaling pathway in plants: CBL-CIPK signaling network. *Physiol Mol Biol Plants*. 14: 51-68.
- Putnam CD, Hammel M, Hura GL, Tainer JA (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Qua Rev Bio*. 40: 191-285.

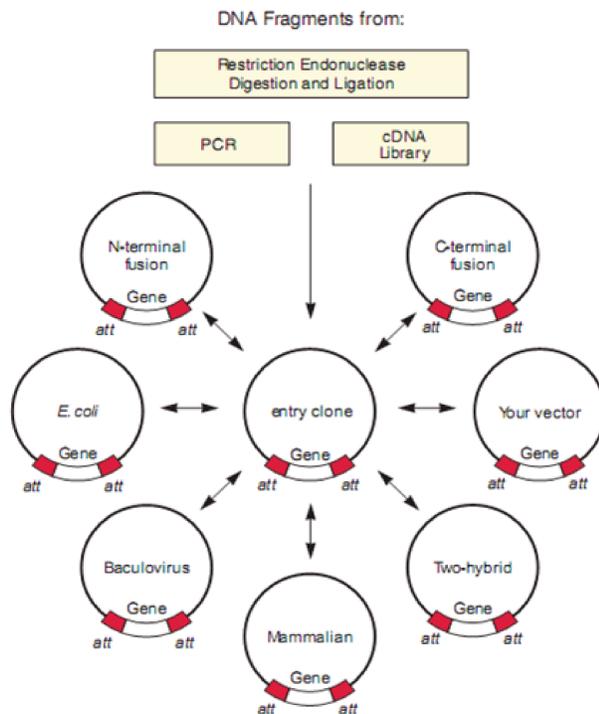
- Qiu QS, Guo Y, Dietrich MA, Schumaker KS, Zhu JK (2002). Regulation of SOS1, a plasma membrane Na<sup>+</sup>/H<sup>+</sup> exchanger in *Arabidopsis thaliana*, by SOS2 and SOS3. *Proc Natl Acad Sci USA*. 99: 8436–8441.
- Resh MD (1999). Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim Biophys Acta*. 1451: 1-16.
- Rost B, Yachdav G, Liu J (2004). The PredictProtein Server. *Nucleic Acids Res*. 32: 321-326.
- Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning: A Laboratory Manual*. 2<sup>nd</sup> Edition. New York, CSHL Press, 1989. p. 999.
- Sanders D, Brownlee C, Harper JF (1999). Communicating with calcium. *Plant Cell*. 11: 691-706.
- Saridakis EEG, Stewart PDS, Lloyd LF, Blow DM (1994). Phase diagram and dilution experiments in the crystallization of carboxypeptidase G<sub>2</sub>. *Acta Crystallogr*. 50: 293-297.
- Svergun D, Kock MHJ (2003). Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys*. 66: 1735-1782.
- Trewavas AJ, Malho R (1998). Ca<sup>2+</sup> signalling in plant cells: the big network! *Curr Opin Plant Biol*. 1: 428-433.
- Janes RW (2008). Reference Datasets Circular Dichroism and Synchrotron Radiation Circular Dichroism Spectroscopy of Proteins. In: *Modern Techniques in Circular Dichroism and Synchrotron Radiation Circular Dichroism Spectroscopy*. London, IOS Press, 2009. p. 244.
- Withmore L, Wallace BA (2004). DICHROWEB, an online server for protein secondary structure analysis from circular dichroism spectroscopic data. *Nuc Acd Res*. 32: 668-673.
- Woody RW (1995). Circular Dichroism. *Method Enzymol*. 246: 34-71.

Woody RW, Koslowski A (2002). Recent developments in the electronic spectroscopy of amides and alpha-helical polypeptides. *Biophys Chem.* 101-102: 535-551.

Zulauf M, D'arcy A (1992). Light scattering of proteins as a criterion for crystallization. *J Cryst Growth.* 122: 102-106.

### III. ANEXOS

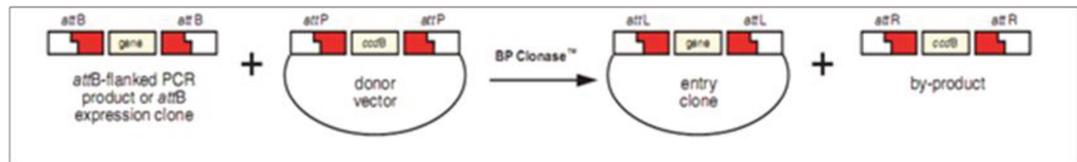
#### Anexo I. Descrição geral do sistema de clonagem Gateway



O sistema Gateway é um método de clonagem baseado nas propriedades de recombinação sitio-específicas (sequências de recombinação específicas ou sítios *att*) do bacteriófago lambda que permite a transferência de sequências de DNA a partir de um vetor de entrada para diversos tipos de vetores destino em apenas duas reações: BP e LR (Gateway Cloning Manual, Invitrogen, 2003).

## Anexo II. Descrição das reações BP e LR

A)



B)



A) Reação BP: produto de PCR flanqueado por sítios *attB* é incubado com um vetor doador (“donor vector”) contendo sítios de recombinação *attP* e reação é catalisada por enzima BP clonase resultando em um vetor de entrada (“entry clone”) contendo o gene de interesse flanqueado por sítios *attL*. B) Reação LR: Clone de entrada contendo o gene de interesse flanqueado por sítios *attL* é incubado com um vetor destino (“destination vector”), a reação é catalisada pela enzima LR clonase e resulta em um clone de expressão contendo o gene de interesse.

### Anexo III: Sequências dos genes *ScCBL1*, *ScCBL3* e *ScCBL6* e das proteínas deduzidas

A predição das CDS das proteínas *ScCBL1*, *ScCBL3* e *ScCBL6* foi feita com o programa ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>):

>*ScCBL1*

```
1 atgggggtgcttccattccacggcgaagcggcagcaccctggctac
  M G C F H S T A K R Q H P G Y
46 gaggaccccgtagcacctcgctcccagaccgccttcagcgtagc
  E D P V H L A S Q T A F S V S
91 gaggttgaggctctattcgagctgttcaagagcataagtggttcg
  E V E A L F E L F K S I S G S
136 gtgattgatgatggcctgatcaacaaggaagaattccagctagca
  V I D D G L I N K E E F Q L A
181 ttatttaagaatcagaggaaagaaaatcttttcgctaatcggata
  L F K N Q R K E N L F A N R I
226 ttgacctttttgatgtcaagaaaggggggtgtcattgattttggt
  F D L F D V K K G G V I D F G
271 gactttgtccgagctctaaatgtgtttcatccaaatatccaatg
  D F V R A L N V F H P N I P M
316 gaagaaaaaattgatttctcattcaagctatacgatattgatggc
  E E K I D F S F K L Y D M D G
361 acagggtttattgaacggaaggagggttaagcagattgtaattgct
  T G F I E R K E V K Q M L I A
406 ctctaggagaatcagaaattgaggctatctgatgagatcattgag
  L L G E S E M R L S D E I I E
451 acaatcctggataagacattttcggacgctgatgcaaattcaggat
  T I L D K T F S D A D A N Q D
496 gggaaaatagatagaacggagtgaggagaattttgttacaagaaat
  G K I D R T E W E N F V T R N
541 cttccttgatgaagataattgactctgccatacctcaaggacata
  P S L M K I M T L P Y L K D I
586 acaacaacattccccagctttgtgtttaactccgaagtcgatgat
  T T T F P S F V F N S E V D D
631 ctggtcagtgaa 642
  L V T *
```

>ScCBL3

```
1 atgttgacagtgcctggatggagtgaaagcacctacttggcgtgctc
M L Q C L D G V K H L L G V L
46 ctcaagtgctgcgacatcgacctcaagcagcccaaggggctagag
L K C C D I D L K Q P K G L E
91 gatccggagggtgctcgctagggagaccgtggttagcgtgagcgag
D P E V L A R E T V F S V S E
136 gttgaggcattgtatgagctgttcaagaagatcagcagtgccggtt
V E A L Y E L F K K I S S A V
181 attgatgatggactgatcaacaaggaggagtttcagttggcgttg
I D D G L I N K E E F Q L A L
226 ttcaagactagtaagaaggagagtttggttgctgatcgtgtat
F K T S K K E S L F A D R V F
271 gacttgtttgatacaaaacacaatggaattcctggatttgaagaa
D L F D T K H N G I L G F E E
316 tttgcccggggcactctcagtatttcatoccaatgctcccctagat
F A R A L S V F H P N A P L D
361 gagaagattgagttttcattccaactatatgacctcaagcaacaa
E K I E F S F Q L Y D L K Q Q
406 ggcttcattgaaagacaggaggtaaagcaaatgggttggtgctaca
G F I E R Q E V K Q M V V A T
451 cttgcagagtcaggaatgaatctttcagatgagggtcatagagagc
L A E S G M N L S D E V I E S
496 atcattgataagacttttgaggaggcagacacaaagcatgatggg
I I D K T F E E A D T K H D G
541 aaaattgacaaagaagaatggcgcaattagtagtgaggcatccc
K I D K E E W R N L V V R H P
586 tctttattgaagaacatgacactacaatatctcaaggacatcacc
S L L K N M T L Q Y L K D I T
631 acaacatttccgagctttgtcttccattctcagggtcgatgatacc
T T F P S F V F H S Q V D D T
676 tga 678
*
```

>ScCBL6

```
1 atggtggactttgttcgacggctcgcctctgcctgctcaagtgc
M V D F V R R L A S A L L K C
46 tgcgaccttgacataccaatcggccaaagggccttgaggacccc
C D L D I P N R P K G L E D P
91 gaacgcctggctagagagactgtttttaatgtgaatgaaatcgag
E R L A R E T V F N V N E I E
136 gctctgtatgagctgttcaagaagattagcagcgctgtggttgat
A L Y E L F K K I S S A V V D
181 gatggcttgattaacaaggaagaattccaattagccctgtttaag
D G L I N K E E F Q L A L F K
226 accaacagaaaggatagcatggtttgctgatcgggtatgttacttg
T N R K D S M F A D R V F D L
271 tttgacacgaagcacaatggaattcctgaatttgaggagtttgct
F D T K H N G I L E F E E F A
316 cgagccctttctgtgttccatccaagtgcaccaattgatgacaaa
R A L S V F H P S A P I D D K
361 attgattttgctttcaaattgtatgatctcaaacaacaaggtttc
I D F A F K L Y D L K Q Q G F
406 attgaaaagcaggagggtcaagcaaatggtggtcgcaacacttgct
I E K Q E V K Q M V V A T L A
451 gaatcaggaatgaatctttcagatgatattattgaaggcattatt
E S G M N L S D D I I E G I I
496 gataagacatttgaggaagcagatacaaagcacgatgggaaaatt
D K T F E E A D T K H D G K I
541 gacaaggaggaatggcgtagccttgtcttgaggcatccatctttg
D K E E W R S L V L R H P S L
586 ttaaaaaatatgaccctcccatatctccgggatatcaccacaaca
L K N M T L P Y L R D I T T T
631 tttccaagctttgtgttcaactctcaagtagaggatgcgtga 672
F P S F V F N S Q V E D A *
```

## Anexo IV: Modificações pós-traducionais da proteína ScCBL1

Programas de previsão de modificação pós-traducional indicam que a proteína ScCBL1 possui dois tipos de modificação na sua região N-terminal: Palmitolação e Miristoilação conforme predito respectivamente pelos programas:

### 1. CSS-PALM (<http://csspalm.biocuckoo.org/>)

Tabela mostrando predições positivas de sítios de palmitolação para as proteínas ScCBL1, 3 e 6 e AtCBL1: Somente ScCBL1 e 3 possuem sítios de palmitolação em sua região N-terminal (posições 3 e 4 respectivamente na sequência de aminoácidos).

Posição	Peptideo	Score	cutoff	tipo
<b>&gt;atCBL1</b>				
3	*MGCFHS	4,339	0,8	TypeIII: Others
<b>&gt;SsCBL1</b>				
3	*MGCFHS	3,817	0,8	TypeIII: Others
<b>&gt;SsCBL3</b>				
4	MLQCLDG	3,817	0,8	TypeIII: Others
18	LLKCCDI	4,141	0,7	TypeI: -CC-
19	LKCCDID	2,969	0,7	TypeI: -CC-
<b>&gt;ScCBL6</b>				
15	LLKCCDL	5,641	0,7	TypeI: -CC-
16	LKCCDLD	4,203	0,7	TypeI: -CC-

### 2. Myristoylator (<http://www.expasy.ch/tools/myristoylator/>)

O programa prediz que a proteína ScCBL1 é miristoilada apresentando uma pontuação de 0,99 em uma escala de 0 a 1. ScCBL1 é a única proteína analisada neste trabalho que possui um *hit* positivo para sítio de miristoilação.

**Anexo V. Artigo publicado: “Functional analysis of a TGA factor-binding site located in the promoter region controlling salicylic acid-induced *NIMIN-1* expression in Arabidopsis .”**



## Functional analysis of a TGA factor-binding site located in the promoter region controlling salicylic acid-induced *NIMIN-1* expression in *Arabidopsis*

J.P. Fonseca<sup>1</sup>, M. Menossi<sup>1</sup>, F. Thibaud-Nissen<sup>2</sup> and C.D. Town<sup>2</sup>

<sup>1</sup>Departamento de Genética, Evolução e Bioagentes,  
Universidade Estadual de Campinas, Campinas, SP, Brasil

<sup>2</sup>The Craig Venter Institute (JCVI), Rockville, MD, USA

Corresponding author: J.P. Fonseca  
E-mail: zepedrof@gmail.com

Genet. Mol. Res. 9 (1): 167-175 (2010)

Received October 1, 2009

Accepted October 29, 2009

Published February 2, 2010

**ABSTRACT.** TGA factors play a key role in plant defense by binding to the promoter region of defense genes, inducing expression. Salicylic acid (SA) induces the expression of the gene encoding *NIMIN-1*, which interacts with NPR1/NIM1, a key regulator of systemic acquired resistance. We investigated whether the TGA2-binding motif TGACG located upstream of the *NIMIN-1* gene is necessary for SA induction of *NIMIN-1* expression. A mutated version of the *NIMIN-1* promoter was created by site-directed mutagenesis. We generated T-DNA constructs in which native *NIMIN-1* and mutated promoters were fused to green fluorescent protein and  $\beta$ -glucuronidase reporters. We produced transgenic *Arabidopsis* plants and observed *NIMIN-1* promoter-driven green fluorescent protein expression in the roots, petiole and leaves. Constructs were agroinfiltrated into the leaves for transient quantitative assays of gene expression. Using quantitative real-time RT-PCR, we characterized the normal gene response to SA and compared it to the response of the mutant version of the *NIMIN-1* promoter. Both the native *NIMIN-1* construct and an endogenous copy of *NIMIN-1* were induced by SA. However, the mutated promoter construct was much less sensitive to SA than the native *NIMIN-1* promoter, indicating that

this TGA2-binding motif is directly involved in the modulation of SA-induced *NIMIN-1* expression in *Arabidopsis*.

**Key words:** NIMIN-1; TGA; Defense; Salicylic acid; *Arabidopsis*; Transient assays

## INTRODUCTION

Systemic acquired resistance (SAR) has been defined as a general plant defense response triggered by pathogen attack, which occurs in distal, non-infected parts of the plant (Glazebrook et al., 1996; Durrant and Dong, 2004). The signaling pathway involved in SAR begins with the synthesis and accumulation of the signal molecule salicylic acid (SA) in both infected and non-infected tissues of the plant in response to the pathogen. SA is required for the induction of pathogenesis-related (PR) genes and pathogen resistance during SAR (Zhang et al., 2003). SA accumulation activates an intermediate cytosolic protein NPR1 (NIM1), which in its monomeric form migrates to the nucleus (Mou et al., 2003), where it binds and activates TGA transcription factors. In turn, these factors bind to and regulate the expression of target PR genes (Johnson et al., 2003). Indeed, it has been shown that the treatment of plants with SA leads to induction of PR genes and establishment of SAR (Ward et al., 1991).

TGA factors are a conserved plant subfamily of bZIPs and were first described in tobacco by their ability to selectively bind to the *as-1* element of the CaMV 35S promoter and to promote transcription (Lam et al., 1989). The involvement of TGA factors in SA response is supported by the presence of the TGA-binding site in the PR-1 promoter (Lebel et al., 1998; Despres et al., 2000). *In vitro* experiments have shown that the motif TGACG is sufficient for TGA factor binding (Lam et al., 1989). In *Arabidopsis*, the TGA family comprises 10 members, and so far, 2 members (TGA2 and TGA3) were shown by chromatin immunoprecipitation (ChIP) to bind to the PR-1 gene promoter *in vivo* (Johnson et al., 2003).

In *Arabidopsis*, 51 putative binding sites for TGA2 were identified by hybridization of immunoprecipitated chromatin to ChIP-chip whole-genome arrays (Thibaud-Nissen et al., 2006). One of these sites lies upstream of *NIMIN-1*, an SA-induced gene from *Arabidopsis* (Weigel et al., 2001; Glocova et al., 2005; Thibaud-Nissen et al., 2006).

The *NIMIN* (NIM1 interacting) genes were first identified for their interaction with NPR1 (NIM1) in a yeast 2-hybrid screen (Weigel et al., 2001). *NIMIN-1* was shown to interact with NPR1 *in planta*, and transgenic plants overexpressing *NIMIN-1* show reduced SA-mediated PR gene induction and reduced amounts of NPR1, demonstrating that *NIMIN-1* acts like a negative regulator of NPR1 (Cao et al., 1998; Weigel et al., 2005).

Although *NIMIN-1* response to SA had been previously described (Weigel et al., 2001; Glocova et al., 2005), the element(s) regulating transcriptional activation of SA induced *NIMIN-1* expression at its promoter region remains unknown. *NIMIN-1* promoter contains two cis-acting TGACG motifs 70 bp apart. Here, we evaluated the role of mutation in one of them, located in the -360-bp region upstream of the *NIMIN-1* gene, in the regulation of SA-induced expression of *NIMIN-1*. In contrast with the high induction in response to SA observed for the native *NIMIN-1* promoter, SA-mediated induction of gene expression was much lower when 2 bp in the TGA binding site in the *NIMIN-1* promoter were altered by site-directed mutagenesis. These data indicate that this promoter element is involved in the regulation of SA-induced *NIMIN-1* expression.

## MATERIAL AND METHODS

### Plant material

*Arabidopsis* wild-type plants (Columbia-0) were grown on Murashige and Skoog medium and agar (0.8%), at 22°C under a 16-h daytime period for long-day growth and 8-h daytime for short-day growth. Plants used for transient assays were raised directly on soil in a growth chamber under short-day conditions for a period of 4 to 6 weeks. Plants used for transient assays were sprayed with 1 mM salicylic acid (Sigma), harvested after 2 h and stored at -80°C for RNA extraction.

### Gateway cloning of *NIMIN-1* promoter region

Primers used to polymerase chain reaction (PCR) amplify and Gateway clone the *NIMIN-1* (-804 to +7) promoter region into pDONR207 were as follows: *NIMIN-1* forward (*attB1*) 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGGAATCCAATTGTTCCA CAC-3'; *NIMIN-1* reverse (*attB2*) 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTAG GATACATTTAGAGAAAGTGATTGAT-3'.

Amplification was done using Hot Start Phusion DNA polymerase (New England Biolabs). The PCR product obtained (812 bp) was cloned into pDONR207 by the BP reaction (recombination and insertion of a PCR product flanked by *attB* sites into a Gateway entry clone catalyzed by the enzyme BP clonase). The product of the BP reaction was transformed into *Escherichia coli* TOP10 cells (Invitrogen). Inserts were validated by sequencing with pDONR207 forward 5'-TCGCGTTAACGCTAGCATGGATCTC-3' and reverse 5'-GTAACATCAGAGATTTTGAGACAC-3' primers. The mutagenized promoter was obtained from the cloned native promoter by site-directed mutagenesis using the Stratagene QuickChange Site Directed Mutagenesis kit. pDONR207 containing *NIMIN-1* promoter was then transferred by the LR reaction (recombination and insertion of an entry clone flanked by *attL* sites into a Gateway destination vector catalyzed by the enzyme LR clonase) to pYXT1 ( $\beta$ -glucuronidase, GUS) and pYXT2 (green fluorescent protein, GFP) binary reporter constructs. Both BP and LR reactions were stopped with proteinase K treatment. The resulting constructs contain the *NIMIN-1* promoter, 5'UTR and the first two amino acids of the native gene as an in-frame fusion to the each reporter construct connected via the Gateway *attB*-encoded peptide. pWTNIM::GFP and pWTNIM::GUS contain the wild-type *NIMIN-1* promoter fused to the coding regions from GFP and GUS, respectively, while pMUTNIM::GUS contains the mutagenized *NIMIN-1* promoter fused to the GUS gene.

### Plasmid transfer to *Agrobacterium* by triparental mating

The recombinant binary plasmids pYXT1 and pYXT2 (Xiao et al., 2005) used in this study containing the different promoters were transformed into *E. coli* TOP10 cells by chemical transformation after the Gateway LR reaction. Recombinant plasmids were then transferred to *Agrobacterium tumefaciens* strain GV3101 by triparental mating. *Agrobacterium* strain GV3101 was grown overnight in liquid LB medium containing 50 mg/mL rifampicin and 50 mg/mL gentamicin at 28°C on a shaker at 250 rpm. *E. coli* cultures containing the plasmid of interest and *E. coli* helper strain (pRK 2013) were also grown overnight in

separate flasks with LB medium containing 50 mg/mL kanamycin at 37°C at 250 rpm. On the next day, 50 µL of each of the three cultures were plated together on an LB agar plate for the mating and incubated for 2 days at 28°C to produce colonies. Further confirmation of agro-transformations was done by PCR analysis with gene-specific primers.

### Stable transformation of *Arabidopsis*

*Arabidopsis thaliana* transformation with the pWTNIM::GFP construct was done according to the floral dip method (Clough and Bent, 1998). *A. tumefaciens* carrying the promoter reporter construct was grown in LB medium containing 50 µg/mL each of rifampicin, gentamicin and kanamycin at 28°C for 2 days. The culture was transferred to a 50-mL plastic tube and centrifuged at 6000 rpm, 4°C for 5 min. We resuspended the pellet in 50 mL *Agrobacterium* suspension (5% sucrose solution with surfactant Silwet L-77 at a final concentration of 0.05%). We dipped the inflorescences into *Agrobacterium* suspension for 10 s and covered the plants with a transparent film. Seeds were harvested and transformants were screened on Murashige and Skoog medium containing 50 mg/mL kanamycin and then transferred to soil.

### Transient assay by agroinfiltration

We agroinfiltrated *Arabidopsis* leaves (SA-treated vs -non-treated) with a needleless syringe containing *Agrobacterium* suspension carrying constructs pMUTNIM::GUS (mutagenized promoter) and pWTNIM::GUS (native promoter) according to the method described by Wroblewski et al. (2005), with modifications. After infiltration with the promoter reporter constructs, the plants were placed back in the growth chamber for a minimum of 2 days to allow expression of the reporter constructs after which leaves were stained for visual inspection. Cultures of *Agrobacterium* were grown overnight in YEP liquid medium with antibiotic selection at 29°C on a shaker. The culture was adjusted to OD<sub>600</sub> = 0.3-0.6 prior to infiltration. We used only soil-grown *Arabidopsis* plants in the vegetative stage for agroinfiltration experiment since plants develop larger leaves than those from early flowering plants. Leaves were taken on the third day post-infiltration and stored at -80°C until gene expression analysis by quantitative real-time RT-PCR (qRT-PCR).

### RNA extraction and cDNA synthesis

Total RNA was isolated from 50 mg *Arabidopsis* leaves using RNeasy Plant Mini Kit (Qiagen) with additional RNase-free DNase (Qiagen) treatment of RNA samples for qRT-PCR. RNA concentration was measured using a NanoDrop ND-1000 spectrophotometer. The purity of RNA (A260/A280) was above 1.9. RNA integrity was checked on a 1% (w/v) agarose gel after RNA extractions. Each cDNA sample was synthesized from 2 µg RNA using SuperScript™ III First Strand Synthesis for RT-PCR (Invitrogen) according to the manufacturer protocol.

### Quantitative real-time RT-PCR

Sequences of specific primers for *GUS*, *NPTII*, *NIMIN-1*, and endogenous gene invariant controls (*AT4G26410*, *AT5G15710*, *AT4G27960*, *AT5G46630*) used for qRT-PCR are listed in Table 1. All primer sequences used in this study were blasted against TAIR's *A. thaliana* cDNA database to

**Table 1.** Primers used for quantitative real-time RT-PCR.

Gene	Forward	Reverse
<i>GUS</i>	TAATGTTCTGCGACGCTCAC	TTCTCTGCCGTTTCCAAT
<i>NPTII</i>	CGTTGGCTACCCGTGATATT	CTCGTCAAGAAGGCGATAGAA
<i>NIMIN-1</i>	ATCTAACGGCGGAGAAAGGT	TGTGATCCGAAACAACATCA
<i>AT4G26410</i>	CCTGGAAGGGATGCTATCAA	GTCCGACATACCCATGATCC
<i>AT5G15710</i>	GCACTTCTGAGACTTTCGGC	ATGACTGAAGAGCACAAACCG
<i>AT4G27960</i>	TAACCATCCATTCCCTCCA	TGGAAATTGTGAGAGCAGGA
<i>AT5G46630</i>	GTGCAATGTTACAGCATC	TGATCTCGTAAGATCCCGCT

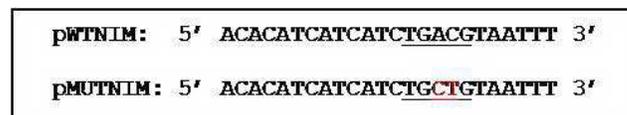
All primers are shown in the 5'-3' direction.

check for amplicon specificity and avoid nonspecific amplification. Short amplicons were used (100-150 bp) for increased assay performance (Marino et al., 2003).

qRT-PCR were performed in an ABI PRISM<sup>®</sup> 7900HT instrument (Applied Biosystems), using SYBR<sup>®</sup> Green to monitor dsDNA synthesis according to the following parameters: 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C, and 1 min at 60°C. Melting curves were generated to confirm the presence of a single PCR product specific peak and for detection of primer-dimer formation using incremental temperatures 95°C for 15 s plus 60°C for 30 s. No template control samples were included in each run as a control for possible contaminations. Final reaction volume was 10 µL. A master mix containing 85 µL 2X SYBR<sup>®</sup> Green (Applied Biosystems) and 45 µL of previously diluted cDNA was prepared prior to dispensing in individual wells to reduce pipetting errors. An electronic adjustable automatic multipipette was used to pipette the master mix (7.5 µL per well) and 2.4 µL of primer mix was dispensed per well with an Eppendorf 12-channel pipette. Data were analyzed using the SDS 2.0 software (Applied Biosystems). Relative quantitation values were calculated using the  $\Delta\Delta C_T$  method (Livak and Schmittgen, 2001; Czechowski et al., 2004) by measuring the difference in  $C_T$  values of target genes treated with SA ( $T_2$ ) and non-treated ( $T_1$ ), normalized to  $C_T$  values of housekeeping genes (internal control), also treated with SA ( $C_2$ ) and non-treated ( $C_1$ ):  $\Delta\Delta C_T = (T_{1(sa-)} - T_{2(sa+)}) - (C_{1(sa-)} - C_{2(sa+)})$ . This reports the logarithm to base 2 of the relative expression values.

## RESULTS AND DISCUSSION

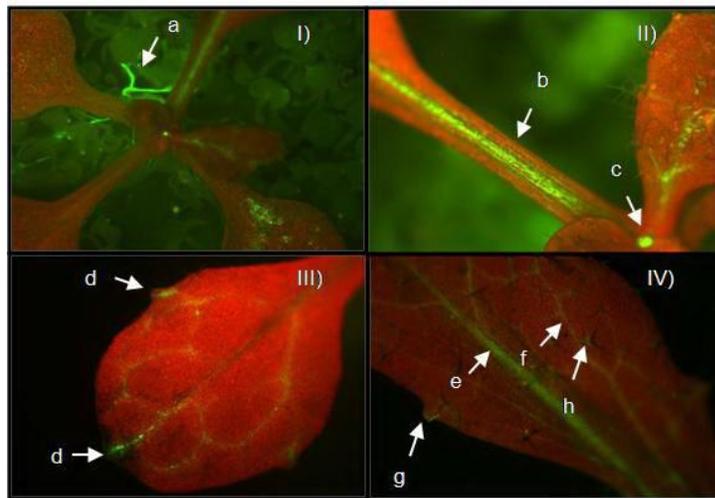
In order to functionally validate the TGA2-binding site TGACG located 360 bp upstream of the start site of the *NIMIN-1* coding region, 2 bp in the motif were altered by site-directed mutagenesis resulting in a “mutagenized” version of the *NIMIN-1* promoter. This mutagenized promoter version was sequenced and aligned to native promoter region using ClustalW (Figure 1).



**Figure 1.** The wild-type *NIMIN-1* promoter region with the TGA2-binding motif (top, pWTNIM) and mutagenized motif (bottom, pMUTNIM). A 2-bp alteration (in red) generated by site-directed mutagenesis of TGA2-binding motif (underlined) located -360 bp upstream of the translational start site of the *NIMIN-1* gene.

Native and mutagenized *NIMIN-1* promoter regions encompassing position +7 in the CDS until -805 bp upstream of the *NIMIN-1* start site were isolated using PCR from pre-existing Gateway entry clones containing native and mutagenized promoters. Using the Gateway system, the amplified promoter regions for *NIMIN-1* were cloned into the entry vector pDONR207 (BP reaction) and then into the reporter binary vectors pYXT1 (GUS) and pYXT2 (GFP) by the Gateway LR reaction. We transformed all BP and LR reactions into *E. coli* and validated cloning after BP reactions by sequencing the insert with pDONR207 flanking primers before proceeding to LR reaction. LR reactions were checked by colony PCR screening of transformed *E. coli* colonies using promoter-specific primers. In total, three constructs were generated for analysis of the *NIMIN-1* promoter: the two native versions of the promoter region being pWTNIM::GFP and pWTNIM::GUS, and one mutagenized: pMUTNIM::GUS.

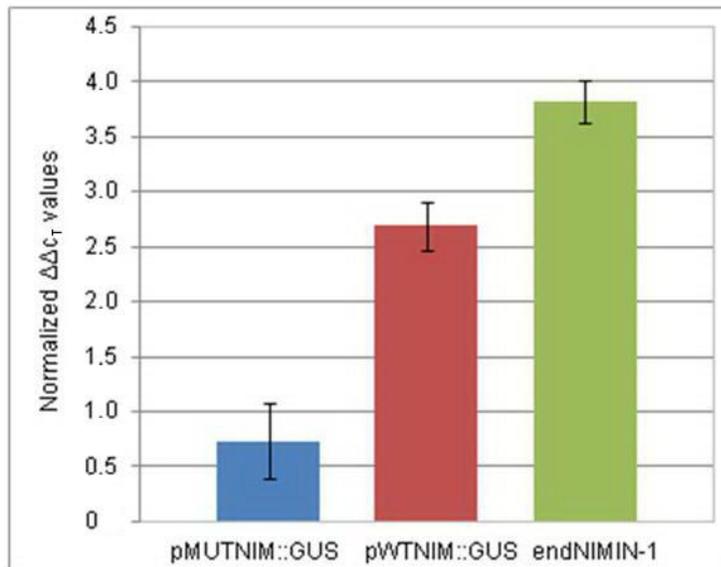
We generated more than 50 transgenic *Arabidopsis* plants expressing the pWTNIM::GFP promoter reporter construct, and T1 plants were observed under the microscope for the pattern of reporter gene expression (Figure 2). The independent segregating lines expressing pWTNIM::GFP showed GFP expression in the shoot apex, roots, midvein, secondary veins, hydathodes, and at the petiole. This expression pattern (e.g., GFP expression in roots and leaves of young *Arabidopsis* rosette) is in agreement with the data from the *Arabidopsis* eFP Browser tool (Winter et al., 2007), which depicts the endogenous *NIMIN-1* expression data from Affymetrix ATH1 GeneChip (data not shown).



**Figure 2.** Green fluorescent protein (GFP) expression patterns in T1 transgenic *Arabidopsis* lines containing the pWTNIM::GFP construct. I) GFP expression in roots (a) from a 3-week-old rosette; II) GFP expression at petiole (b) and shoot apex (c); III) GFP expression at primary and secondary veins and hydathodes (d) from a 3-week-old rosette leaf; IV) GFP expression in midvein (e), secondary vein (f), hydathodes (g), and trichomes (h) of a 6-week-old plant.

To determine if the TGA2-binding motif located in the promoter region of *NIMIN-1* is involved in the transcriptional regulation in response to SA of *NIMIN-1* expression, we performed comparative functional analysis of the native and mutagenized version of *NIMIN-1* promoter re-

gions. We compared levels of gene expression by qRT-PCR from native and mutant *NIMIN-1* promoters in relation to endogenous copy of *NIMIN-1* using promoter regions fused to the GUS reporter gene. qRT-PCR was the method of choice for measuring SA-induced *NIMIN-1* expression because we aimed to better estimate the response at the transcriptional level in a short period of time after SA exposure (2 h). Reporter gene expression was measured by qRT-PCR from agroinfiltrated leaves on the third day post-infiltration, to allow expression of the reporter gene (Figure 3).



**Figure 3.** Effect of mutation in the TGA2-binding site in the *NIMIN-1* promoter. Bars represent the change in expression of the wild-type (pWTNIM::GUS) and mutated (pMUTNIM::GUS) promoters and in the endogenous *NIMIN-1* gene (endNIMIN-1) relative to the internal controls *NPTII* (for promoter reporter constructs) and *At5g15710* (endogenous *NIMIN-1*) in SA-treated vs -untreated samples. Data are reported as means  $\pm$  SD of three independent biological replicates. GUS =  $\beta$ -glucuronidase.

We measured transcript levels from leaves sprayed with 1 mM SA at 2 h (SA+) and non-treated leaves at 0 h (SA-).  $C_T$  values for endogenous *NIMIN-1* were normalized to the  $C_T$  values of the internal control gene *At4g26410*, which was one of the most constant of the four housekeeping genes included in each qRT-PCR run. The average  $C_T$  value for *At4g26410* was  $20.44 \pm 0.32$  at a threshold of 0.1 for all three biological replicates.  $C_T$  values for agroinfiltrated promoter reporter constructs pWTNIM::GUS and pMUTNIM::GUS were normalized to the internal control *NPTII*. The average  $C_T$  value for *NPTII* was  $26.46 \pm 0.34$  for all three biological replicates. The changes in expression levels in response to SA was calculated using the  $\Delta\Delta C_T$  method (Livak and Schmittgen, 2001; Czechowski et al., 2004), as described in Methods.

We found that the pWTNIM::GUS construct containing the native promoter and the endogenous copy of *NIMIN-1* were highly responsive to SA (Figure 3), in agreement with previous study (Weigel et al., 2001; Glucova et al., 2005). qRT-PCR analysis also showed that

the SA-mediated induction of the native promoter was more than 3-fold higher compared to that observed in the mutagenized promoter.

This difference in reporter gene expression between mutant and native promoters indicates that TGA2 binding to the TGACG *cis*-element is relevant to the *NIMIN-1* response to SA. It is worth noting that the weak induction of the mutant promoter could suggest that the mutation in the 2 bp does not completely abolish the binding of the TGA2 factor or that the other TGACG motif present in the promoter region also plays a minor role in mediating SA induction of *NIMIN-1*.

In summary, our data on the mutagenized and wild-type *NIMIN-1* promoter indicate that the TGACG *cis*-element is involved in regulation of SA-induced expression.

## ACKNOWLEDGMENTS

We thank Dr. Yongli Xiao for kindly providing the vectors pYXT1 and pYXT2, Julia Redman for support with qRT-PCR and other members of the Plant Genomics group for their assistance. Research supported by JCVI and CAPES.

## REFERENCES

- Cao H, Li X and Dong X (1998). Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. *Proc. Natl. Acad. Sci. U. S. A.* 95: 6531-6536.
- Clough SJ and Bent AF (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735-743.
- Czechowski T, Bari RP, Stitt M, Scheible WR, et al. (2004). Real-time RT-PCR profiling of over 1400 *Arabidopsis* transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J.* 38: 366-379.
- Despres C, DeLong C, Glaze S, Liu E, et al. (2000). The *Arabidopsis* NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. *Plant Cell* 12: 279-290.
- Durrant WE and Dong X (2004). Systemic acquired resistance. *Annu. Rev. Phytopathol.* 42: 185-209.
- Glazebrook J, Rogers EE and Ausubel FM (1996). Isolation of *Arabidopsis* mutants with enhanced disease susceptibility by direct screening. *Genetics* 143: 973-982.
- Glocova I, Thor K, Roth B, Babbick M, et al. (2005). Salicylic acid (SA)-dependent gene activation can be uncoupled from cell death-mediated gene activation: the SA-inducible *NIMIN-1* and *NIMIN-2* promoters, unlike the *PR-1a* promoter, do not respond to cell death signals in tobacco. *Mol. Plant Pathol.* 6: 299-314.
- Johnson C, Boden E and Arias J (2003). Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in *Arabidopsis*. *Plant Cell* 15: 1846-1858.
- Lam E, Benfey PN, Gilmartin PM, Fang RX, et al. (1989). Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. *Proc. Natl. Acad. Sci. U. S. A.* 86: 7890-7894.
- Lebel E, Heifetz P, Thorne L, Uknes S, et al. (1998). Functional analysis of regulatory sequences controlling PR-1 gene expression in *Arabidopsis*. *Plant J.* 16: 223-233.
- Livak KJ and Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25: 402-408.
- Marino JH, Cook P and Miller KS (2003). Accurate and statistically verified quantification of relative mRNA abundances using SYBR Green I and real-time RT-PCR. *J. Immunol. Methods* 283: 291-306.
- Mou Z, Fan W and Dong X (2003). Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. *Cell* 113: 935-944.
- Thibaud-Nissen F, Wu H, Richmond T, Redman JC, et al. (2006). Development of *Arabidopsis* whole-genome microarrays and their application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. *Plant J.* 47: 152-162.
- Ward ER, Uknes SJ, Williams SC, Dincher SS, et al. (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell* 3: 1085-1094.
- Weigel RR, Bauscher C, Pfitzner AJ and Pfitzner UM (2001). NIMIN-1, NIMIN-2 and NIMIN-3, members of a novel family of proteins from *Arabidopsis* that interact with NPR1/NIM1, a key regulator of systemic acquired resistance

- in plants. *Plant Mol. Biol.* 46: 143-160.
- Weigel RR, Pfitzner UM and Gatz C (2005). Interaction of NIMIN1 with NPR1 modulates PR gene expression in *Arabidopsis*. *Plant Cell* 17: 1279-1291.
- Winter D, Vinegar B, Nahal H, Ammar R, et al. (2007). An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2: e718.
- Wroblewski T, Tomczak A and Michelmore R (2005). Optimization of *Agrobacterium* mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. *Plant Biotech. J.* 3: 259-273.
- Xiao YL, Smith SR, Ishmael N, Redman JC, et al. (2005). Analysis of the cDNAs of hypothetical genes on *Arabidopsis* chromosome 2 reveals numerous transcript variants. *Plant Physiol.* 139: 1323-1337.
- Zhang Y, Tessaro MJ, Lassner M and Li X (2003). Knockout analysis of *Arabidopsis* transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell* 15: 2647-2653.

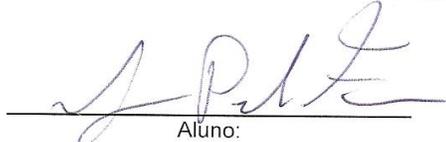
DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação/tese de Doutorado intitulada **“Vias de sinalização de estresses em plantas: análise da região promotora do gene *NIMIN-1* de *Arabidopsis thaliana* e da proteína ScCBL1 de cana-de-açúcar (*Saccharum* spp.)”**

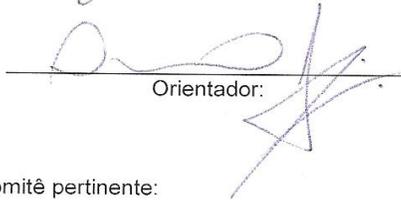
( ) não se enquadra no § 3º do Artigo 1º da Informação CCPG 01/08, referente a bioética e biossegurança.

( x ) tem autorização da(s) seguinte(s) Comissão(ões) de Bioética ou Biossegurança\*:  
Projeto CIBio/IB/UNICAMP, intitulado Genômica Funcional de Plantas, sob Protocolo(s) nº 2007/03.

\* Caso a Comissão seja externa à UNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.

  
Aluno:

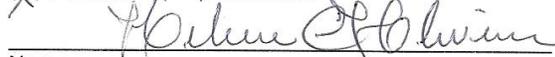
José Paulo Fonseca

  
Orientador:

Marcelo Menossi  
Teixeira

Para uso da Comissão ou Comitê pertinente:

(x) Deferido ( ) Indeferido



Nome:

Função:

Profa. Dra. HELENA COUTINHO F. DE OLIVEIRA  
Presidente  
Comissão Interna de Biossegurança  
CIBio/IB - UNICAMP