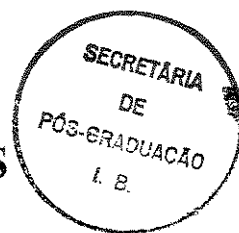


UNIVERSIDADE ESTADUAL DE CAMPINAS



CLAUDIA BANDEIRA KOBARG

UNICAMP

BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

Caracterização do gene BZ02H1 de *Arabidopsis thaliana* que codifica um fator de transcrição do tipo bZIP homólogo ao fator O2 do milho. Análise comparativa de sua especificidade de ligação ao DNA

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato(a)
Claudia Bandeira Kobarg
e aprovada pela Comissão Julgadora.

12/04/2001

Tese apresentada ao Instituto de Biologia para obtenção do título de Mestre em Biologia Funcional e Molecular na área de Bioquímica

Orientador: Prof. Dr. MICHAEL GEORGES ALBERT VINCENTZ

Campinas, 2001



UNIDADE	OC
N.º CHAMADA	T/ UNICAMP
	K792c
V.	Ex
TOMBO BC/	46038
PROC.	16-392/01
C	☐
D	☒
PREC.	R\$ 11,00
DATA	16-08-01
N.º CPD	

CM00158604-2

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO INSTITUTO DE BIOLOGIA – UNICAMP**

Kobarg, Claudia Bandeira

K792c

Caracterização do gene BZ02H1 de *Arabidopsis thaliana* que codifica um fator de transcrição do tipo bZIP homólogo ao fator 02 do milho: análise comparativa de sua especificidade de ligação ao DNA/Claudia Bandeira Kobarg. - - Campinas, SP. [s.n.], 2001. 87f: illus.

Orientador: Michel Vincentz

Dissertação (mestrado) – Universidade Estadual de Campinas. Instituto de Biologia.

1. Fator de transcrição. 2. Proteína. 3. DNA. I. Vincentz, Michel. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Campinas, 17 de abril de 2001.

BANCA EXAMINADORA:

Prof. Dr. MICHAEL GEORGES ALBERT VINCENTZ



Prof. Dr. ADILSON LEITE



Prof. Dr. CARLOS FRANCISCO SAMPAIO BONAFÉ



Prof. Dr. JOSÉ CAMILLO NOVELLO

AGRADECIMENTOS

Agradeço aos prezados professores:

Dr. Michel Georges Albert Vincentz, por disponibilizar seu projeto de pesquisa e espaço no laboratório para a realização de minha tese de mestrado.

Dr. José Camilo Novello por ter me deixado livre para escolher meu caminho.

Dr. Adilson Leite por sempre ter mantido as portas do laboratório de genética de plantas abertas e pela participação na banca de tese.

Dr. Carlos Francisco Sampaio Bonafé, pelo exemplo e por todas as agradáveis discussões sobre ciência, além da participação na banca de tese.

Agradeço também aos queridos colegas e amigos:

Rodrigo Siloto, pela paciência e boa vontade ao me ensinar a técnica de gel shift e por ter me ajudado, mesmo nas condições adversas, remando contra a maré, na escuridão; Dra. Alba Chiesse, pela proteína controle purificada O2 de *Coix* e por todas as frutíferas discussões; Dr. Marcelo José Surpili, pelo auxílio na purificação da proteína recombinante; Edmilson Rui, pelo apoio técnico com os géis de poliacrilamida que apresentaram problemas; Luciane Gauer, pelo apoio técnico e psicológico; Sandra Martins, pelo apoio técnico no laboratório; Patrícia Moura, pela boa vontade.

Ao meu querido Jörg, que além do exemplo, foi o maior contribuinte para a realização deste trabalho. Por todas as discussões, idéias, soluções de problemas, empréstimos de material e pelo apoio incondicional como marido, amigo e grande profissional acadêmico. *Vielen Dank*.

A todos meus amigos, colegas, pais e família, que de alguma forma, contribuíram para que este trabalho fosse finalizado.

À CAPES, pelo suporte financeiro.

Um dia, é preciso parar de sonhar e partir.

Amir Klink

Para Jörg.

Índice

Resumo	1
Abstract	2
Introdução	
Expressão gênica	3
Transcrição e sua regulação em eucariotos	3
Fatores de regulação do tipo bZIP	5
Fatores bZIP de plantas superiores	7
A família Multigênica de O2	8
Materiais e Métodos	
1-Material vegetal	11
2-Linhagens bacterianas	11
3-Isolamento do gene <i>BZO2H1</i>	11
4-Extração de RNA	12
5-Northern blot	14
6-Southern blot	14
7-Determinação da estrutura do gene	15
8-Análise das seqüências	15
9-Amplificação Rápida da Extremidade 5' do cDNA (5' RACE)	16
10-Clonagem e análise dos produtos de PCR	17
11-Transformação de <i>E. coli</i>	17
12-Mini-preparação de DNA	19
13-Construção do cDNA truncado BZH1 em vetor pET	19
14-Expressão das proteínas recombinantes em <i>E. coli</i> pelo sistema pET	20
15-Seleção de recombinantes	21
16-Crescimento de bactérias e indução de expressão	22
17-Purificação e identificação das proteínas recombinantes	22
a) Lise bacteriana	22
b) Purificação em coluna de cobalto	22
c) Determinação da concentração protéica	23
d) Géis de poliacrilamida (SDS-PAGE)	23
e) Western Blot	23
18-Preparação das sondas de DNA para EMSA	24
19-Ensaio de Retardamento da Mobilidade Eletroforética (EMSA)	24
Resultados	
I) Caracterização do gene <i>BZO2H1</i>	26
Isolamento e caracterização do cDNA <i>BZO2H1</i>	26
Localização do gene <i>BZO2H1</i> e seu padrão de expressão	26
O gene <i>BZO2H1</i> pertence a família multigênica O2	32
II) Comparação da especificidade de ligação da proteína <i>BZO2H1</i> e O2 de Coix	35
Expressão e purificação da proteína BZH1	35
Caracterização da especificidade de ligação da proteína recombinante BZH1	38
Otimização das condições de ligação da proteína BZH1 à sonda GP	40
A proteína BZH1 se liga especificamente a sonda GP	43

BZH1 reconhece os motivos GLM e Palíndrome	43
BZH1 e O2CX reconhecem G-box, mas BZH1 se liga a C-box com afinidade inferior que O2CX	45
Discussão	
A família multigênica O2	51
A especificidade de ligação ao DNA de BZH1	55
Perspectivas	59
Referências Bibliográficas	60

Índice de Figuras e Tabelas

Figura 1. Modelo de interação da estrutura “coiled coil” do zíper de leucinas e da interação do domínio bZIP de GCN4 com seu DNA alvo	6
Figura 2. Representação esquemática da técnica 5’RACE	18
Figura 3. Mapa de restrição do vetor pET28a	20
Figura 4. Sequência de nucleotídeos do cDNA BZO2H1 com seqüências de aminoácidos deduzida	27
Figura 5. Representação esquemática da localização cromossômica e estrutura do gene <i>BZO2H1</i>	29
Figura 6. Análise do DNA genômico de <i>Arabidopsis thaliana</i> por Southern blot	30
Figura 7. Análise da expressão do gene <i>BZO2H1</i> por Northern blot	31
Figura 8. Alinhamento das seqüências das proteínas bZIP da família Multigênica O2	33
Figura 9. Representação esquemática do cDNA inteiro de BZO2H1 e de sua respectiva proteína inteira e do domínio bZIP clonado	36
Figura 10. Proteína recombinante BZH1 está presente na fração solúvel do extrato bacteriano	37
Figura 11. O vetor pET28aBZH1 expressa dois polipeptídeos que são identificados devido à sua cauda hexamérica de histidina	39
Figura 12. Alinhamento das seqüências das proteínas recombinantes BZH1 e O2CX	41
Figura 13. EMSA da proteína recombinante BZH1 com a sonda GP	42
Figura 14. BZH1 se liga especificamente à sonda GP	44
Figura 15. EMSA das proteínas recombinantes O2CX e BZH1 com a sonda mutada GM	47
Figura 16. EMSA das proteínas recombinantes O2CX e BZH1 com a sonda mutada MP	48
Figura 17. Comparação da ligação das proteínas recombinantes BZH1 e O2CX à sonda G-box	49
Figura 18. Comparação da ligação das proteínas recombinantes BZH1 e O2CX a sonda C-box	50
Tabela I	13
Tabela II	25

Lista de Abreviaturas e Unidades utilizadas na tese

M = Molar
mM = milimolar
nmol = nanomol
pmol = picomol
l = litro
ml = mililitro
 μ l = microlitro
h = hora (s)
min = minuto(s)
s = segundo(s)
°C = grau Celsius
V = Volts
mV = milivolts
mA = miliamperes
OD600 = Absorbância em 600 nanômetros
g = grama(s)
pb = par(es) de base(s)
kb = kilobases
kDa = kiloDalton
cpm = cintilações por minuto
PCR = Reação de Polimerase em Cadeia
EMSA = Eletrophoresis Mobility Shift Assay
cDNA = DNA complementar
bZIP = Domínio básico e zíper de leucinas
o2 = Mutação Opaco2
O2= *Gene* Opaco2
O2 = Proteína Opaco2
mRNA = RNA mensageiro
BSA = Albumina bovina sérica
CaCl₂ = Cloreto de cálcio
CSPD = Disodium 3-(4-methoxy Spiro-1,2 dioxetano-3,2'-(5'-chloro)tricyclo [3.3.1.1]decan)-4yl- phenyl phosphate
DEPC = Dietilpirocarbonato
EDTA = Ácido Etileno Diaminotetracético
DTT = Ditiotreitól
IPTG = Isopropil tio- β -galactosídeo
KCl = Cloreto de potássio
MgCl₂ = Cloreto de magnésio
MOPS = 3(N-Morfolino)ácido propanosulfônico
NaCl = Cloreto de sódio
PMSF = Fenilmetilsulfonilfluorido
Rnase = Ribonuclease
SDS = Dodecil Sulfato de Sódio
SSC = Cloreto de sódio/ Citrato trissódico
Tris = Tris (hidroximetil)Aminometano
U = Unidade

RESUMO

O gene *Opaco2* (*O2*) é um fator de regulação da transcrição do tipo bZIP envolvido no controle da síntese de proteínas de reserva no endosperma do milho. Um cDNA de *Arabidopsis thaliana* que codifica uma proteína do tipo bZIP homóloga ao fator de transcrição *O2* de milho denominado BZO2H1 (Basic Leucine Zipper Opaque2 Homologous) foi isolado e caracterizado. A sequência do gene *BZO2H1* foi identificada nos bancos de dados do projeto genoma de *Arabidopsis thaliana*. O gene *BZO2H1* é composto por sete exons e seis íntrons e está presente em cópia única no genoma de *Arabidopsis thaliana*. A estrutura do gene *BZO2H1*, assim como a alta similaridade entre as sequências do domínio bZIP de BZO2H1 e *O2* confirmaram que BZO2H1 pertence à Família Multigênica de *O2*. Transcritos de RNA mensageiro de BZO2H1 foram detectados em todos os tecidos investigados e a incidência ou ausência de luz sobre as plantas jovens pareceu não afetar a expressão do gene. O domínio bZIP da proteína BZO2H1 foi produzido *em E. coli* para determinar sua especificidade de ligação ao DNA. O domínio bZIP de BZO2H1 reconhece sequências relacionadas ao sítio GCN4 como GLM (5'-TGACTA-3') e a sequência palíndrome (5'-GACATGTC-3'), além dos motivos G-box e, menos intensamente, C-box. A especificidade de ligação do domínio bZIP de BZO2H1 foi comparada com a da proteína *O2* de *Coix* e as diferenças encontradas foram discutidas.

ABSTRACT

The gene *Opaque2* (*O2*) encodes a bZIP transcription factor involved in storage protein synthesis control in maize endosperm. An *Arabidopsis thaliana* cDNA that encodes a bZIP protein homologous to the *O2* transcription factor, named BZO2H1 (Basic Leucine Zipper Opaque2 Homologous) has been isolated and characterized. The *BZO2H1* gene sequence was identified in the databases of the *Arabidopsis thaliana* genome project. The *BZO2H1* gene is composed of seven exons and six introns and is present as a single copy in the genome of *Arabidopsis thaliana*. The gene structure and the high sequence similarity of the bZIP domain shared by *BZO2H1* and *O2* defined *BZO2H1* as a member of the *O2* multigene family. BZO2H1 mRNA was detected in all tissues analyzed, but no differences of the expression levels were observed when comparing young plants exposed to light or darkness. The bZIP domain of BZO2H1 protein was expressed in *E. coli* to begin the characterization of its DNA binding specificity. The bZIP domain of BZO2H1 binds to the GCN4-related DNA binding site sequences GLM (5'-TGACTA-3') and "palindrome" (5'-GACATGTC-3'), as well as to the G-box and to a lesser extend to the C-box. The binding specificity of the BZO2H1 bZIP domain was compared to *O2* Coix bZIP domain and the differences observed were discussed.

INTRODUÇÃO

Expressão gênica

A expressão regulada dos genes é fundamental para a maioria dos processos biológicos como crescimento, diferenciação celular, desenvolvimento e resposta a estímulos ambientais. (Meshi & Iwabuchi, 1995). A regulação da expressão gênica pode ocorrer em diferentes passos da via DNA-RNA-proteína: pelo controle da transcrição, do processamento e transporte do RNA, da degradação do RNA mensageiro, da tradução da proteína e pelo controle da atividade da proteína (Darnell, 1982; Derman, 1981). Dentre os controles citados, o controle da transcrição é, provavelmente, um dos passos mais importantes para a regulação da expressão gênica. (Macfarlane, 2000).

Transcrição e sua regulação em eucariotos

O DNA de organismos eucarióticos se encontra empacotado com proteínas específicas, formando a cromatina no núcleo da célula. A unidade fundamental da cromatina é o nucleossomo. Cada partícula de nucleossomo é composta por 146 pb de DNA envolto num octâmero de moléculas de histona e outras proteínas, formando uma estrutura semelhante a um cilindro (Revisado por Lee & Young, 2000). Esse empacotamento do DNA dentro da cromatina representa uma barreira significativa para vários processos nucleares, entre eles, a transcrição. Os nucleossomos reprimem a transcrição por oclusão de sítios reconhecidos por proteínas ligantes de DNA, interferindo assim com a ligação do complexo de iniciação da transcrição (Kornberg, 1999).

A iniciação da transcrição ocorre em dois estágios. Primeiro, pelo alívio da repressão causada pelo nucleossomo e segundo, pela interação da RNA Polimerase e proteínas do complexo de iniciação da transcrição com o promotor do gene (Kornberg, 1999).

Presume-se que as proteínas designadas como Fatores de Ativação da transcrição sejam importantes no primeiro estágio, pois podem recrutar complexos modificadores de cromatina. Estes complexos possuem atividade catalítica, acetilando as histonas e causando assim uma modificação estrutural nas mesmas. Infere-se que essa modificação estrutural diminua a repressão sobre a transcrição exercida pelo nucleossomo (Kornberg, 1999).

Com o promotor exposto devido às modificações no nucleossomo, o segundo estágio do início da transcrição pode começar. Neste estágio, participam a RNA Polimerase II e os chamados fatores gerais de transcrição, designados como: TFIIB, TFIID, TFIIF e TFIIH. (Mitchel & Tjian, 1994).

O complexo TFIID é responsável pelo reconhecimento do promotor. A proteína TBP (TATA-Binding Protein) se liga a TATA-Box, enquanto que as outras subunidades do complexo TFIID, conhecidas como TAFs (TBP-Associated Factors), interagem com as seqüências adjacentes à TATA-box conferindo maior especificidade de ligação ao DNA (Buratowski, 1994; Kornberg, 1999). A curvatura do DNA da TATA-Box em volta da TBP cria um ambiente propício para a interação da TFIIB, que por sua vez, posiciona o promotor com a RNA Polimerase II (Ly *et al.*, 1994; Leuther *et al.*, 1996). O próximo passo é dado pelo complexo TFIIH que inclui as helicases ATP-dependentes que desenrolam o promotor em volta do sítio de iniciação para desencadear a iniciação da transcrição (Leuther *et al.*, 1996).

Os Ativadores, além de remodelar a estrutura da cromatina, recrutam o complexo de iniciação no promotor se ligando aos fatores TFIIB, TFIID e TFIIH. Essa ligação é mediada por um fator protéico adicional chamado Mediador. O fator Mediador oferece uma interface entre as proteínas ativadoras de genes e a RNA Polimerase II, porém sua influência sobre a ativação da transcrição ainda precisa ser determinada (Kornberg, 1999).

Durante todo o processo da iniciação de transcrição, fatores de transcrição não pertencentes à classe dos fatores gerais de transcrição, atuam de maneira tecido-específica, ativando ou reprimindo a transcrição de genes por reconhecimento de elementos em seus promotores e interagindo com o complexo de iniciação. Os fatores de transcrição são constituídos, basicamente, por uma região de ligação ao DNA, um sítio de dimerização, um domínio de regulação da transcrição e um sinal de localização nuclear (Washburn *et al.*, 1997; Hill *et al.*, 1996; Liu *et al.*, 1999; Tjian & Maniatis, 1989). Esses fatores podem ser classificados em função da estrutura ou seqüência primária do seu domínio de ligação ao DNA. Uma classificação completa, de acordo com a superclasse, classe, família e subfamília dos fatores de transcrição pode ser encontrada no banco de dados TRANSFAC (<http://transfac.gbf.de/TRANSFAC/>). Várias classes de fatores de transcrição já foram caracterizadas, dentro delas, os fatores de transcrição do tipo bZIP que é a classe de interesse nesse trabalho.

Fatores de regulação da transcrição do tipo bZIP

O domínio de ligação ao DNA dos fatores de regulação da transcrição do tipo “Basic Leucine Zipper” (bZIP) é formado por uma região básica, rica em aminoácidos básicos que faz contato com o DNA e um zíper de leucina, formado por uma repetição de resíduos hidrofóbicos a cada sete aminoácidos, promovendo a dimerização da proteína (revisado por Hurst, 1995).

A determinação da estrutura tridimensional do domínio bZIP do fator GCN4 de levedura complexado a sua sequência alvo é mostrada na figura 1. O domínio bZIP forma uma α -hélice e os dois zíperes de leucina formam uma estrutura “coiled coil” para dimerização. A região básica de cada subunidade se encaixa dentro do sulco maior da sequência alvo, onde cinco aminoácidos estabelecem ligações iônicas ou pontes de hidrogênio com as bases de um meio sítio alvo (König & Richmond, 1993; Ellenberg *et al.*, 1992; Alber, 1992; Hurst, 1995). É o conjunto dessas interações que determina a afinidade do domínio bZIP para uma sequência alvo específica. Esse modo de se ligar ao DNA também explica como os fatores bZIP reconhecem sequências do tipo pseudo-palíndrome ou palíndrome. Por exemplo, GCN4 apresenta uma alta afinidade para o pseudopalíndrome AP1 (5'ATGACTCAT3') e o palíndrome ATF/CREB (5'ATGACGTCAT3') (Suckow *et al.*, 1993).

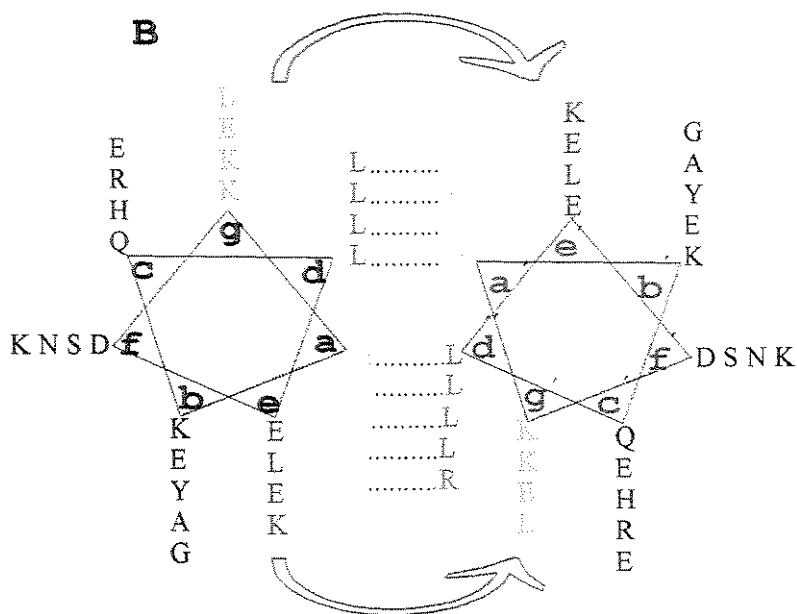
Como mostrado na figura 1, a dimerização e sua estabilidade resultam, principalmente, de interações de Van der Waals entre aminoácidos das posições a e d das duas α -hélices e interações eletrostáticas entre os aminoácidos das posições e e g. Portanto, o tipo de aminoácidos presentes nessas posições é determinante para o controle da dimerização. Por exemplo, cargas idênticas nas posições e e g irão desestabilizar ou impedir a dimerização da proteína (O'Shea *et al.*, 1991).

A capacidade dos fatores bZIP de formar homo ou heterodímeros pode ter consequência importante na sua função regulatória. Mais especificamente, a sua especificidade de ligação ao DNA pode ser alterada, mudando assim o espectro de genes regulados por esses fatores. (Lamb & McNight, 1991; Singh, 1998).

A

	a	b	c	d	e	f	g
				NH ₂	- D	P	A
Região básica	A	L	K	R	A	R	N
	T	E	A	A	R	R	S
	R	A	R	K	L	Q	R
Zíper de Leucinas	M	K	Q	L	E	D	K
	V	E	E	L	L	S	K
	N	Y	H	L	E	N	E
	V	A	R	L	K	K	L
	V	G	R	R	- COOH		

B



C

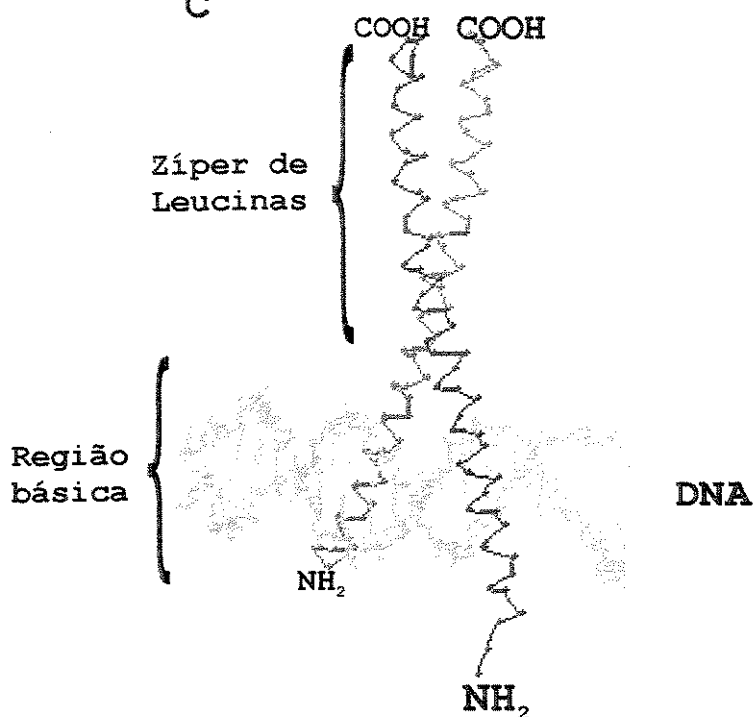


Figura 1. Modelo da interação da estrutura "coiled coil" do zíper de leucinas e da interação do domínio bZIP de GCN4 com seu DNA alvo. A) Seqüência peptídica do domínio bZIP de GCN4 (aminoácidos 226-281) apresentada como repetições de 7 aminoácidos em forma de tabela. As leucinas presentes na face interna da hélice estão destacadas em negrito. B) Diagrama das interações das alfa-hélices do zíper de leucinas do dímero GCN4. As duas hélices se encontram posicionadas de frente uma para a outra, de modo a permitir a interdigitação de suas leucinas formando o zíper. Os resíduos de um monômero das posições **a** e **d** se encontram de frente para os resíduos das posições **a'** e **d'** do outro monômero. As linhas descontínuas representam as interações hidrofóbicas entre os resíduos **a** e **d**, enquanto que as setas verdes representam as interações eletrostáticas entre os resíduos **g** e **e**. C) Modelo espacial da interação do domínio bZIP de GCN4 com seu DNA alvo. As duas alfa-hélices do dímero GCN4 formam uma estrutura "coiled coil" devido às interações hidrofóbicas mostradas. A região básica de cada monômero se encaixa dentro do sulco maior da seqüência alvo. (Ellenberger et al., 1992)

Como é o caso de outros fatores de regulação da transcrição, os fatores bZIP também apresentam motivos de localização nuclear, de ativação da transcrição e de interação com outras proteínas (Schwechheimer *et al.*, 1998; Sainz *et al.*, 1997). Outro aspecto importante é a presença de possíveis sítios específicos de fosforilação na sequência polipeptídica dos fatores bZIP que podem ser alvos de quinases. A fosforilação e defosforilação da proteína permitem regular vários aspectos funcionais como a translocação para o núcleo, a ligação ao DNA e a ativação das proteínas bZIP (Nicols *et al.*, 1992; Ciceri *et al.*, 1997).

Fatores bZIP de plantas superiores

Provavelmente, mais de 200 sequências de fatores do tipo bZIP de plantas superiores estão depositadas nos bancos de dados. Mais recentemente, foi possível determinar que o conjunto completo de fatores bZIP do genoma de *Arabidopsis* é formado por pelo menos 80 membros. Essas observações sugerem que a evolução das plantas se beneficiou amplamente da propriedade de regulação das bZIP.

Análises genéticas, moleculares e bioquímicas indicam que alguns fatores bZIP são reguladores importantes de processos fisiológicos e ontogenéticos de plantas. Entre esses processos, podemos citar a fotomorfogênese com a participação do fator HY5 (Osterlund *et al.*, 2000), o desenvolvimento de órgãos com os fatores PERIANTHIA e LIGULELESS2 (Walsh *et al.*, 1997; Chuang *et al.*, 1999), a elongação e a morfogênese celular com os fatores RSG e RF2a (Yin *et al.*, 1997; Fukazawa *et al.*, 2000), o controle do balanço de nitrogênio e carbono durante o desenvolvimento da semente por O2 (Ciceri *et al.*, 2000), o controle de desenvolvimento da semente pelos os fatores ROM 1 e ROM 2 (Chern *et al.*, 1996a; Chern *et al.*, 1996b) os mecanismos de defesa com os fatores TGA2.2 e OBFs (Niggeweg *et al.*, 2000; Zhang *et al.*, 1999), a via de sinalização da sucrose com ATB2 (Rook *et al.*, 1998), as vias de sinalização de hormônios com os fatores ABI5, ABF, TRAB e TGA (Choi *et al.*, 2000; Finkelstein & Linch, 2000; Uno *et al.*, 2000) e resposta à luz com os fatores CPRF2, CPRF4, GBF1, GBF2 e GBF3 (Schindler *et al.*, 1992; Wellmer *et al.*, 1999).

Todos esses fatores bZIP são prováveis ativadores de transcrição. Entretanto, podemos mencionar que os fatores ROM1 e ROM2 apresentam características predominantemente repressoras da transcrição, em função do estágio de desenvolvimento da semente (Chern *et al.*, 1996a e 1996b).

A caracterização da especificidade de ligação ao DNA das proteínas bZIP *in vitro* é uma das formas possíveis de se obter informações que possam ajudar a estabelecer seu significado funcional *in vivo* (Niu & Guiltinan, 1994). Sequências palindrômicas baseadas no motivo ACGT foram identificadas em promotores de genes de plantas regulados por fatores fisiológicos e ambientais (Foster *et al.*, 1994). A maior parte dos fatores bZIP de plantas já estudados, reconhece esse motivo e exibe especificidade e afinidade de ligação ao DNA variável, de acordo com as sequências adjacentes ao motivo ACGT. Isawa e colaboradores propuseram uma classificação para fatores bZIP de plantas, baseada nas características de ligação ao motivo ACGT e suas sequências adjacentes (Isawa *et al.*, 1993). Essa classificação propõe três grupos de proteínas bZIP segundo sua especificidade e afinidade de ligação: Grupo 1 exibe forte afinidade de ligação ao motivo G-box (5'-NNCACGTGNN-3'); Grupo 2, exibe afinidade comparável tanto para G-box, como para C-box (5'-NNGACGTCNN-3') e grupo 3, exibe forte afinidade de ligação a C-box (Isawa *et al.*, 1993). Alguns fatores bZIP, como O2 de milho, também se ligam a sequências pseudo-palíndromes relacionadas ao sítio AP1 (5'-ATGACTCAT-3') (Yunes *et al.*, 1994). Embora existam algumas informações sobre os elementos do domínio bZIP que regulam a especificidade de ligação ao DNA (Niu *et al.*, 1999), ainda é preciso conhecer melhor os mecanismos envolvidos para entender e definir o papel dos fatores bZIP *in vivo*. O reconhecimento de sequências alvos específicas para regular os genes adequados *in vivo*, corresponde a mais um nível de complexidade que envolve interações com outras proteínas de regulação.

A Família Multigênica de O2

Opaco2 (O2) é um gene que regula a expressão das proteínas de reserva no endosperma durante a formação da semente do milho (Schmidt, 1993). O2 codifica para um fator do tipo bZIP que é especificamente acumulado no endosperma, onde ele ativa a expressão dos genes das proteínas de reserva α e β -prolaminas (Schmidt *et al.*, 1992; Cord Neto *et al.*, 1995), o gene b-32 (Lohmer *et al.*, 1991) e o gene da piruvato ortofosfato diquinase citoplasmática (Gallusci *et al.*, 1996). O2 também está envolvido no controle da acumulação da lisina (Sodek & Wilson, 1970; Kemper *et al.*, 1999) e do metabolismo da treonina (Giroux *et al.*, 1994). Esses dados sugerem que O2 tem um papel importante na regulação coordenada da síntese das proteínas de reserva, do metabolismo do carbono e do nitrogênio durante o

desenvolvimento da semente. Essa hipótese é sustentada pela observação de que a atividade de ligação ao DNA de O2 é modulada por fosforilação/defosforilação segundo um ritmo circadiano relacionado a fluxos metabólicos (Ciceri *et al.*, 1999).

Prováveis fatores ortólogos a O2 de milho, que são fatores derivados de um evento de especiação (Graur & Li, 1999), foram isolados e caracterizados em outras plantas monocotiledôneas da subfamília *Panicoideae* como: O2 de sorgo (Pirovano *et al.*, 1994), O2 de *Coix* (Vettore *et al.*, 1998), SPA de trigo (Albani *et al.*, 1997) e BLZ2 de cevada (Onate *et al.*, 1999). Todos esses fatores são especificamente acumulados no endosperma e são capazes de ativar a transcrição *in vitro* a partir de seqüências de regulação de genes de proteínas de reserva. Essas proteínas são capazes de ligar pseudo-palíndromes relacionados a seqüências AP-1 ou tipo GCN4. Além disso, O2 de milho se liga a G-box e C-box com a mesma afinidade (Isawa *et al.*, 1993), enquanto O2 de *Coix* se liga de maneira cooperativa a dois sítios adjacentes, um palíndromo (5'-GACATGTC-3') e um pseudopalíndromo "GCN4-Like Motive" ou GLM (5'-TGACTA-3') que estão presentes no promotor do gene da proteína de reserva α -Coixina. É provável que essa cooperatividade contribua para estabelecer a seleção dos genes regulados por O2 *in vivo* (Yunes *et al.*, 1998). Outra maneira possível de estabelecer a especificidade de regulação *in vivo* é pela interação com outras proteínas de regulação, como foi mostrado no caso de O2 com PBF (Prolamin Box Factor). PBF é uma proteína "Zinc Finger" do tipo Dof que se liga a "Prolamin Box" presente no promotor de genes que codificam as Zeínas. Foi observado que PBF é capaz de se ligar a "Prolamin Box" da α -Zeína de 22 kDa de milho e simultaneamente interagir com a proteína O2. PBF interage especificamente com O2, mas não com outras bZIPs testadas (Vicente-Carbajosa *et al.*, 1997). A interação de PBF com O2 parece ser importante já que um sítio "Prolamin Box" funcional é necessário para a ativação da transcrição dependente de O2 em milho e para a ativação por outras bZIPs relacionadas a O2 em outros cereais (Albani *et al.*, 1997; Vicente-Carbajosa *et al.*, 1997, 1998).

O último exemplo comentado acima fornece evidência sobre a importância do controle combinatório na regulação transcricional da expressão gênica em plantas. Em animais, vários exemplos de controle combinatório já foram identificados com a formação de complexos "Enhanceosomes". "Enhanceosomes" é a designação de um conjunto de seqüências regulatórias presentes num promotor, posicionadas de maneira a permitir a interação das

proteínas que se ligam a essas seqüências (Singh, 1998). A associação de seqüências regulatórias e proteínas regulatórias que interagem entre si, podem possibilitar um controle mais fino da transcrição. Provavelmente, podemos aplicar o modelo do “Enhanceosome” para explicar a importância da interação O2/PBF e seus sítios de interação. Contudo, estudos ainda precisam ser feitos para esclarecer essa questão.

Sete outras proteínas do tipo bZIP de plantas monocotiledôneas e dicotiledôneas homólogas a O2 foram identificadas e parcialmente caracterizadas. Esse grupo de proteínas forma a família III ou Família Multigênica O2 (Vettore *et al.*, 1998). Entretanto, a existência de genes ortólogos a O2 em plantas dicotiledôneas ainda não foi estabelecido. Considerando a importância de O2 no desenvolvimento da semente, podemos supor que tal locus de regulação foi conservado no decorrer da evolução, estando, portanto, também presente em plantas dicotiledôneas.

Com o objetivo de buscar fatores bZIP ortólogos a O2 em plantas dicotiledôneas, foram identificados 4 cDNAs parciais que codificam proteínas do tipo bZIP na planta modelo *Arabidopsis thaliana* altamente similares à proteína O2.

O trabalho aqui desenvolvido, relata o isolamento e a caracterização de um desses cDNAs que designamos BZO2H1 (Basic Leucine Zipper Opaque2 Homologous). Também analisamos a estrutura do gene *BZO2H1* e seu padrão de expressão. Com o intuito de definir a especificidade de ligação do produto gênico de *BZO2H1* e compará-la com a de O2, cuja especificidade de ligação se encontra bem caracterizada, expressamos uma forma da proteína BZO2H1 como proteína recombinante em *E. coli* e a analisamos por Ensaio de Retardamento da Mobilidade Eletroforética (EMSA).

MATERIAIS E MÉTODOS

1-Material Vegetal

Arabidopsis thaliana (ecotipo Columbia) foi cultivada até a maturidade numa mistura de vermiculita: terra na proporção 1:3. O fotoperíodo foi de 16 h diárias de luz artificial e a temperatura foi de 22° C.

2-Linhagens bacterianas

As seguintes cepas de *Escherichia coli* foram utilizadas:

DH5α genótipo *supE44 ΔlacU169 (φ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1* para replicação de plasmídeos.

BL21(DE3)pLys S genótipo *hsdS gal (λclts857 ind1 Sam7 nin5 lac UV5-T7 gene1)* pLysS para alto nível de expressão de genes clonados em plasmídeos contendo o promotor T7.

3-Isolamento do gene *BZO2H1*

O gene *BZO2H1* foi isolado de uma biblioteca de cDNA de sementes em desenvolvimento de *Arabidopsis thaliana* cedida pelo Dr. G. Giraudat (Institut des Sciences Végétales, CNRS, Gif-sur-Yvette, Cedex, France). Os fragmentos de cDNA das sementes de *A. thaliana* foram clonados no sítio *Eco* RI do fago λZAP. O método utilizado foi a Amplificação Rápida das Extremidades 3' (3'RACE), que consiste de uma série de amplificações do cDNA a partir da extremidade 3' com oligonucleotídeos degenerados e sobrepostos, correspondentes a duas regiões conservadas das proteínas bZIP da família filogenética III descrita por Vettore *et al.*, em 1998. Na primeira etapa da reação de amplificação, aproximadamente 3.10^8 fagos λZAP recombinantes foram misturados com: 10 pmol do oligonucleotídeo sintético M13-reverso, 150 pmol dos oligonucleotídeos BC5 ou ZC3.1 (Tabela I), 1x Tampão de reação BRL, 1,5 mM MgCl₂, 3U Taq Polimerase (BRL), num volume final de 100 μl. As reações foram submetidas ao seguinte ciclo: uma etapa de desnaturação a 94°C por 4 min, seguido de 35 ciclos de uma etapa de desnaturação por 1 min a 94°C, uma etapa de anelamento dos oligonucleotídeos por 2 min a 42°C e uma etapa de alongamento por 2 min a 72°C.

Os produtos de PCR obtidos foram purificados com o kit QIAquick® PCR Purification (Quiagen). Um volume de 10 µl dos produtos purificados foi utilizado para a segunda fase de amplificação pelo acréscimo das seguintes concentrações finais de oligonucleotídeos e reagentes: 10 pmol M13 reverso, 75 pmol de BC6 ou BC7 ou ZC3.2 (Tabela I), 1x Tampão de reação BRL, 1,5 mM MgCl₂, 3U Taq Polimerase (BRL), num volume final de 50 µl. As reações foram submetidas a uma etapa de desnaturação a 94°C por 4 min, seguidas de 3 ciclos com temperatura de anelamento a 50°C e 33 ciclos com temperatura de anelamento a 56°C. As etapas de desnaturação e alongamento foram realizadas a 94°C por 1 min e a 72°C por 2 min respectivamente. Os produtos de PCR desta segunda fase que apresentaram tamanho superior a 200 pb foram purificados em gel de agarose, digeridos com *Eco* RI e clonados no vetor pBKS+ (Stratagene). Os clones que apresentaram tamanhos de insertos aparentemente diferentes entre si foram seqüenciados segundo a metodologia “Dye Terminator” (Big Dye Kit, Perkin Elmer). As seqüências foram analisadas com o programa BLAST no servidor do NCBI (www.ncbi.nlm.nih.gov). Os clones que apresentaram seqüências similares a de fatores bZIP conhecidos foram utilizados como sonda para um novo “screening” na biblioteca de cDNA. Dessa maneira, foi possível isolar o clone que apresentou o inserto de cDNA BZO2H1 de 1440 pb. Este clone de 1440pb foi utilizado como sonda nos ensaios de Southern and Northern blot.

4-Extração de RNA

Folha, flor, caule e raiz foram coletados em Nitrogênio líquido e congelados a -70°C. Para extração do RNA, cada órgão foi triturado em Nitrogênio líquido com auxílio de gral e pistilo até a formação de um pó. O tecido vegetal triturado foi homogeneizado com Trizol® (Life Technologies) numa proporção de 500 mg por 1 ml de Trizol® e incubado por 5 min a temperatura ambiente para permitir a dissociação total dos complexos nucleoprotéicos. (Nos seguintes passos, todos os volumes utilizados são referentes a 1 ml de Trizol® inicial). Um volume de 0,2 ml de clorofórmio foi adicionado à mistura e agitado vigorosamente num tubo, seguindo uma incubação de 3 min à temperatura ambiente e centrifugação a 12000 x g por 15 min a 4°C para separação de fases. O RNA da fase aquosa foi precipitado pelo acréscimo de 0,5 ml de isopropanol e incubação à temperatura ambiente por 15 min. A suspensão foi então centrifugada a 12000 x g por 10 min à 4°C, formando um precipitado de RNA gelatinoso no

fundo do tubo. O precipitado de RNA foi lavado com etanol 75%, seco à temperatura ambiente e dissolvido em água tratada com Dietilpirocarbonato (DEPC).

Para a extração de RNA da semente, 2g do material foi triturado e homogeneizado em 9 ml de uma solução alcalina de 1 M Tris-Cl pH 9,0 e SDS 1%. A extração ocorreu por agitação constante à temperatura ambiente por 10 min. O mesmo volume de fenol: clorofórmio: álcool isoamílico numa proporção 25: 24: 1 foi adicionado ao extrato e misturado por 5 min. A mistura foi então centrifugada a 8000 x g por 10 min a 4°C. Na fase aquosa, foi adicionado 2x o volume de etanol absoluto e 0,1x o volume de acetato de potássio 2,5 M pH 5,2 e o RNA foi precipitado a -20°C por 3 h. A suspensão foi então centrifugada a 10000 x g por 30 min a 4°C e o precipitado de RNA foi lavado com etanol 70% e seco à temperatura ambiente. O RNA foi diluído em solução 10 mM Tris-Cl (pH 8,0) e 1 mM EDTA (solução TE) e precipitado pelo acréscimo de igual volume de solução 4 M de cloreto de lítio e incubação a 4°C durante 12 h para eliminar o DNA que ficou solúvel. Em seguida, o RNA foi recuperado por centrifugação e lavado com etanol 70%, seco e suspenso em TE (Dean *et al*, 1985). O RNA foi quantificado por espectrofotometria a um $\lambda=260$ nm.

Primer	Seqüência (5'→3')	Posição# (pb)	Tm	Direção	Uso
O2L1	GTTTCTGGAGAACTAGTCG	460-479	56°C	Antisense	5'RACE
O2L2	CGAAGAGAAACAACAGTACG	415-434	58°C	Antisense	5'RACE
O2L3	GCACTCAGTCTCAAGCTTG	395-414	58°C	Antisense	5'RACE
AAP	GGCCACGCGTGCAGTACGGGIIIG GIIIGGGIIG	Extremidade Poly C 3'	67°C	Sense	5'RACE
T7	GTAATACGACTCACTATAGGGC	vetor	64°C	Sense	Seqüenciamento
SP6	ATTTAGGTGACACTATAGAATAC	vetor	60°C	Antisense	Seqüenciamento
T728REV	TAATACGACTCACTATAGGG	vetor	56°C	Sense	Seqüenciamento e PCR
T728DIR	GCTAGTTATTGCTCAGCGG	vetor	58°C	Antisense	Seqüenciamento e PCR
SES2	CGAGGTCACCTGTTTGC	754-777	52°C	Antisense	Seqüenciamento
ZIPS2N	GAATGAAACCATGGGTTCTTG	655-675	48°C	Sense	Construção
ZIPS2C	CACTTGATCTCTCGAGAAGCATCG	953-967	52°C	Antisense	Cosntrução
BC5*	TCHAAVMGDGARTCWGC		46°C	Sense	3'RACE e isolamento
BC6*	AAGGAATTCGARTCHGCHAGRAGGTC		50°C	Sense	3'RACE e isolamento
BC7*	AAGGAATTCGARTCWGCHAGRCGNTC		50°C	Sense	3'RACE e isolamento
ZC3.1*	GYNAAGGTRAAGATGG		42°C	Sense	3'RACE e isolamento
ZC3.2*	AAGGAATTCGTRAAGATGGSGARG		46°C	Sense	3'RACE e isolamento

Tabela I. Seqüência dos oligonucleotídeos sintéticos utilizados na amplificação da extremidade 5', seqüenciamento, PCR de colônia e construções do cDNA truncado. # significa posição em relação à seqüência da figura 4 e Tm representa a temperatura de anelamento utilizada. * significa oligonucleotídeos degenerados Biosynthesis®

5-Northern blot

A 20µg de RNA total dissolvidos em 11 µl de água foram adicionados 5 µl de tampão de corrida MOPS 10 x (0,4 M ácido 3-N-morfolino-propanosulfônico pH 7; 0,1 M acetato de sódio e 0,01 M EDTA), 9 µl de 12,3 M formaldeído, 25 µl de formamida deionizada e 10µl de tampão de aplicação (1 mM EDTA, 0,25% azul de bromofenol, 0,25% xilenocianol, 50% glicerol e 0,2µg/ml brometo de etídeo). A solução foi misturada com auxílio de vortex e incubada a 60°C por 20 min. Após breve centrifugação, as amostras de RNA foram aplicadas num gel de agarose em condições desnaturantes (1% agarose, tampão de corrida MOPS 1 x, 2,2 M Formaldeído) e submetidas a eletroforese a 80 V por aproximadamente 3 h (Ausubel, 1997). O RNA foi transferido para uma membrana de nylon Hybond® (Amersham) por capilaridade em tampão SSC 10 x (1,5 M NaCl; 0,15 M ácido cítrico dihidratado pH 7) e fixado à membrana através de UV "cross-linking". A membrana foi pré-hibridizada numa solução de hibridização [50% formamida deionizada, SSC 6x, solução de Denhardt 5x (0,1% Ficoll 400, 0,1% polivinilpirrolidona, 1 mg/ml albumina sérica bovina), 1% SDS e 500 µg/ml de DNA de esperma de salmão desnaturado] por, no mínimo, 3 horas a 42°C em forno de hibridização. Aproximadamente 20 ng da sonda BZO2H1, que consiste do fragmento de cDNA BZO2H1 de 1440 pb, foi marcada com [α -³²P]-dCTP por "random priming" utilizando-se o "kit Megaprime DNA labelling System" (Amersham) e seguindo-se o protocolo do fabricante. Após a marcação, a sonda foi purificada numa mini-coluna de Sephadex G-50, desnaturada a 95°C e adicionada à solução de hibridização. A membrana foi hibridizada por 16 horas a 42°C e na manhã seguinte, lavada 3 vezes, da seguinte maneira: em solução SSC 2x; SDS 0,5% por 15 min à temperatura ambiente; em solução SSC 0,1 x; SDS 0,5% por 15 min à temperatura ambiente; em solução SSC 0,1x; SDS 0,5% por 30 - 45 min a 65°C.

Após a lavagem, a membrana foi envolta numa folha de PVC e exposta a um filme Hyperfilm™MP (Amersham) a -70°C por, no mínimo, 72 h.

6-Southern blot

Aproximadamente 3 µg de DNA genômico de *Arabidopsis thaliana* Ecotipo Columbia foi digerido com as enzimas *Bam* HI, *Hind* III, *Eco* RI ou *Xba* I num volume total de 30 µl. O DNA digerido foi precipitado pelo acréscimo de 60 µl de etanol 96% e 3 µl de acetato de

potássio 3 M pH 5,2 e incubação a -20°C por 3 horas. O precipitado foi centrifugado e lavado com etanol 70% e dissolvido em TE. Os fragmentos de DNA foram separados por eletroforese num gel de agarose a 0,8% (Sambrook *et al*, 1989) e transferidos para uma membrana de nylon Hybond® (Amersham) por capilaridade em tampão SSC 10x. O DNA foi fixado à membrana por “cross-linking” e a membrana foi pré-hibridizada em SSC 6x, 0,5% SDS, solução de Denhardt 5x e 200 µg/ml DNA de espermatozóide de salmão desnaturado por 4 horas a 60°C . A sonda de cDNA BZO2H1 foi preparada de acordo com o método Northern blot já descrito anteriormente. A hibridização ocorreu a 60°C por 6 horas. A membrana foi lavada em solução SSC 2x e SDS 1% por 15 min a temperatura ambiente e em solução SSC 0,1x e SDS 1% por 40 min a 65°C . Foi então envolta numa folha de PVC e exposta a um filme Hyperfilm MP™ (Amersham) a -70°C por 48 h.

7-Determinação da estrutura do gene

A estrutura do gene foi determinada por comparação de seqüências de cDNA obtidas com as seqüências genômicas depositadas nos bancos de dados do projeto genoma de *Arabidopsis thaliana* (www.MIPS.biochem.mpg.de/proj/thal/).

8-Análise das seqüências

O alinhamento das seqüências de aminoácidos das proteínas foi feito com o programa CLUSTAL X (Thompson *et al*, 1997).

A determinação da porcentagem de identidade e similaridade das proteínas comparadas foi feita com auxílio do programa BOXSHADE que é baseado na matriz de Gonnet (Gonnet *et al.*, 1992).

A análise das seqüências de DNA foi feita com o programa DNASIS (Pharmacia).

A detecção dos motivos conservados entre os membros da família filogenética O2 foram feitos com auxílio do programa MEME (Bailey and Elkan, 1994; <http://meme.sdsc.edu/meme/website/>).

9-Amplificação Rápida da Extremidade 5' do cDNA (5'RACE)

Este método permitiu amplificar seqüências de um RNA mensageiro situadas entre um sítio interno definido e uma seqüência desconhecida na extremidade 5' com oligonucleotídeos sintéticos de DNA gene específicos (Figura 2).

Na primeira etapa, uma fita simples de cDNA foi sintetizada pela reação de transcrição reversa catalisada pela enzima Transcriptase Reversa. Nessa reação, 0,5 a 1,0 µg de RNA foi misturado com 2,5 pmol do oligonucleotídeo sintético O2L1 (tabela I) num volume final de 15, 5 µl. A mistura foi incubada por 10 min a 70 °C para desnaturar o RNA. Ao término da incubação, foram adicionados os seguintes reagentes: 20 mM Tris-HCL (pH 8.4), 50 mM KCl, 2,5 mM MgCl₂, 10 mM DTT, 400 µM dATP, 400µM dCTP, 400µM dGTP, 400µM dTTP e 5 Unidades de Transcriptase Reversa SuperscriptTM II num volume final de 25 µl. A mistura foi gentilmente homogeneizada e incubada a 42°C por 30 min. A reação foi então bloqueada por desnaturação da enzima a 70°C por 15 min. Após a síntese da fita simples de cDNA, o RNA inicial foi degradado pela adição de 0,5 µl de RNase H e 0,5 µl de RNase A seguindo incubação a 37°C por 30 min. O cDNA sintetizado foi purificado com auxílio do "kit Quiaquick Nucleotides Removal" (Quiagen), segundo instruções do fabricante.

A segunda etapa dessa técnica compreendeu a adição da cauda homopolimérica de deoxicitosina à extremidade 3' do cDNA pela enzima Deoxinucleotídeo terminal transferase (TdT). A mistura da reação foi composta por: 5 - 20 ng cDNA purificado, 10 mM Tris-HCl (pH 8.4), 25 mM KCl, 1,5 mM MgCl₂, 200 µM dCTP e 1 Unidade TdT em 25 µl de volume final. Após breve desnaturação do cDNA a 94°C, a reação foi incubada a 37°C por 10 min e a enzima inativada por mais 10 min a 65°C.

Na terceira etapa, o cDNA com a extremidade homopolimérica de deoxicitosina foi amplificado. Para isso foi utilizado um oligonucleotídeo sintético sense rico em deoxiguanina designado AAP (Abridged Anchor Primers) e o oligonucleotídeo gene-específico O2L2 (Tabela I) na direção antisense. A composição final da reação foi: 20 mM Tris-HCL (pH8.4), 50 mM KCl, 1,5 mM MgCl₂, 400 nM O2L2, 400 nM AAP, 200 µM de cada deoxinucleotídeo (dATP, dCTP, dGTP e d TTP), 5 µl de solução de cDNA com extremidade de deoxicitosina obtida na etapa anterior e 2 Unidades de *Taq* DNA polimerase num volume final de 50 µl. As reações foram submetidas ao seguinte ciclo: uma etapa de desnaturação a

94°C por 2 min, seguido de 35 ciclos de uma etapa de desnaturação por 1 min a 94°C , uma etapa de anelamento dos oligonucleotídeos por 2 min a 58°C e uma etapa de alongamento por 2 min a 72°C. A técnica de “hot start” foi utilizada em todas as reações de amplificação e consiste em equilibrar todos os reagentes, exceto a enzima Taq DNA polimerase, a 94°C antes de iniciar o primeiro ciclo.

Na quarta etapa, foi feita a segunda amplificação do cDNA com o terceiro oligonucleotídeo gene específico O2L3 de direção antisense e o oligonucleotídeo sense AAP. A composição final da reação foi: 20 mM Tris-HCL (pH8.4), 50 mM KCl, 1,5 mM MgCl₂, 200 nM O2L2, 200 nM AAP, 200 µM de cada deoxinucleotídeo (dATP, dCTP, dGTP e dTTP), 1 µl do produto de reação final de amplificação da etapa anterior e 2 U de Taq DNA polimerase. A técnica de “hot start” foi novamente utilizada e as reações foram submetidas ao seguinte ciclo: uma etapa de desnaturação a 94°C por 2 min, seguido de 35 ciclos de uma etapa de desnaturação por 1 min a 94°C, uma etapa de anelamento dos oligonucleotídeos por 2 min a 58°C e uma etapa de alongamento por 2 min a 72°C.

10-Clonagem e análise dos produtos de PCR

Os produtos das reações de amplificação foram separados em gel de agarose 1% e os fragmentos de DNA maiores de 400 pb, foram purificados com auxílio do “kit QUIquick Gel Extraction Kit” (Quiagen) segundo instruções do fabricante. Após a purificação, os fragmentos foram ligados no vetor pGem®-T-easy (Promega) e introduzidos na cepa DH5α de *E.coli* por transformação com PEG (Sambrook *et al.*, 1989).

11-Transformação de *E.coli*

Cepas competentes de *E. coli* foram transformadas pela diluição de 50 a 500 ng de DNA plasmidial em 20 µl de TE e adição de 80 µl de tampão de transformação (100 mM KCl, 30 mM CaCl₂, 50 mM MgCl₂ e 1,5% Polietilenoglicol 6000). Após breve incubação desta mistura no gelo, foram adicionados 100 µl de células competentes, seguindo incubação no gelo por 20 min e incubação à temperatura ambiente por mais 10 min. Para que as células pudessem se recuperar, 0,8 ml de meio de cultura LB líquido foram adicionados ao tubo que foi incubado a 37°C por 50 min. Após esse intervalo, as células bacterianas foram plaqueadas

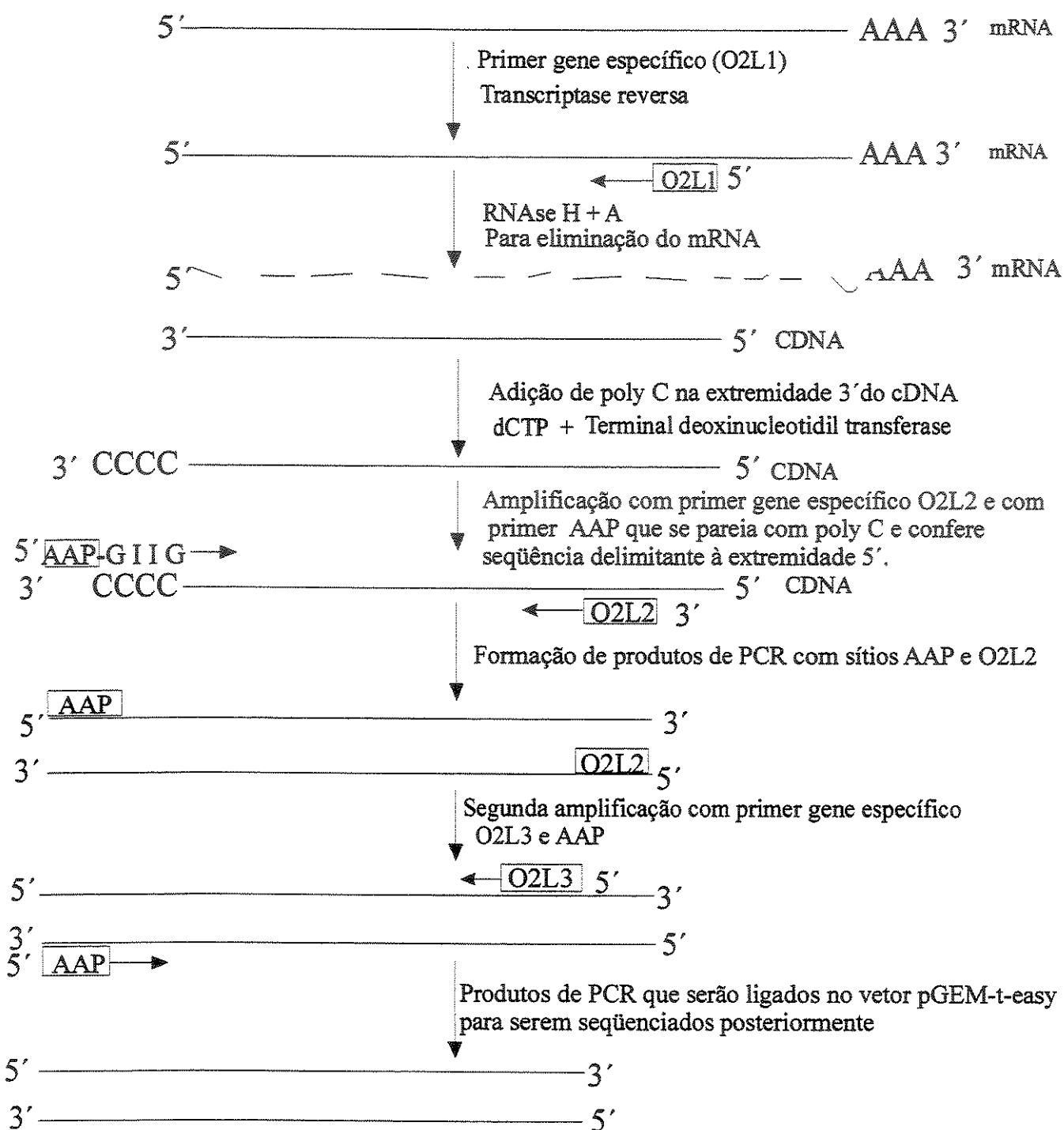


Figura 2. Representação esquemática da técnica 5' RACE

em meio de cultura LB sólido contendo o antibiótico de resistência do plasmídeo e incubadas por 12 h a 37°C para permitir o crescimento de colônias.

12-Mini-preparação de DNA

Após a transformação, as colônias positivas que cresceram no meio de seleção foram repicadas e incubadas em meio de cultura LB líquido contendo o antibiótico cuja resistência foi conferida pelo vetor. A mini-preparação de DNA seguiu a técnica de lise alcalina (Sambrook *et al.*, 1989).

13-Construção do cDNA truncado BZH1 em vetor pET para expressão de proteínas

A proteína truncada, que engloba somente o domínio básico e o zíper de leucina foi obtida pela amplificação do cDNA BZO2H1 com os oligonucleotídeos ZIPS2N e ZIPS2C (Tabela I). Para isso, aproximadamente 10 ng de DNA do vetor pBKS+ com inserto do cDNA inteiro de BZO2H1 foi misturado com 10 pmol de ZIPS2N, 10 pmol de ZIPS2C, 5 µl de tampão de PCR (Gibco), 1,5 mM MgCl₂, 0,2 mM de cada dNTP e 1 Unidade de *Taq* Polimerase recombinante (Gibco) em 50 µl de volume final. Após 2 min de incubação a 94°C para desnaturar o cDNA, a reação de amplificação ocorreu em 35 ciclos de 94°C por 1 min, 50°C por 1 min e 72°C por 2 min.

O oligonucleotídeo ZIPS2N conferiu sítio de restrição *Nco* I à extremidade 5' da sequência e ZIPS2C conferiu sítio de restrição *Xho* I à extremidade 3'. Após a amplificação, o produto de PCR de 314 pb foi purificado em gel de agarose e duplamente digerido com as enzimas *Nco* I e *Xho* I. Após a quantificação do fragmento de DNA, este foi clonado no vetor pET 28a (Figura 3) previamente digerido nos sítios *Nco* I e *Xho* I.

Outra vantagem deste sistema é o uso de cepas de expressão que apresentem resistência a determinado antibiótico, facilitando a seleção dos clones positivos. A cepa BL21DE3pLys S apresenta resistência ao Cloramfenicol pela presença do plasmídeo pLys S. O vetor pET28a foi escolhido por conferir uma cauda de histidina à proteína recombinante. Esta cauda de histina facilita a purificação da proteína por cromatografia de afinidade em coluna de cobalto ou níquel. A resistência conferida pelo vetor pET28a é ao antibiótico Kanamicina.

15-Seleção de recombinantes

Após a construção da fase codificante do cDNA no vetor pET28a, a cepa DH5 α de *E. coli* foi transformada com a mesma. Para testar as colônias transformadas que cresceram no meio, foi utilizado o método de PCR de colônia. Esse método utiliza oligonucleotídeos que amplificam as regiões adjacentes ao “polylinker”, onde o cDNA deve estar clonado. Os oligonucleotídeos sintéticos escolhidos foram T728dir e T728rev, cujas seqüências se encontram na Tabela I. Cada colônia foi tocada com uma ponteira estéril e ressuspensa em 20 μ l de água destilada num tubo de 0,5 ml. Cada tubo foi então aquecido a 95°C por 10 min e centrifugado a 12000 x g por 1 min. A 15 μ l do sobrenadante foram adicionados 35 μ l de mistura pré-PCR, a uma concentração final de: 10pmol T728dir, 10 pmol T728rev, 1,5 mM MgCl₂, 5 μ l de tampão de PCR (Gibco), 0,2 mM de cada dNTP e 1 U de Taq polimerase (Gibco). A reação de PCR começou com uma etapa de desnaturação a 94°C por 2 min seguida por 35 ciclos de 94 °C por 1 min, 58 °C por 1 min e 72 °C por 2 min. Os produtos de PCR foram visualizados por eletroforese em gel de agarose 1% confirmando a presença do vetor recombinante na bactéria hospedeira.

Após a seleção por PCR de colônia, foi necessário testar o DNA recombinante com enzimas de restrição, pois sendo PCR uma técnica muito sensível, pode haver contaminação de DNA não ligado nas placas, levando a resultados falsos positivos. Assim, os clones positivos foram inoculados em meio de cultura LB líquido (10 g/l de tripton, 5 g/l de extrato de levedura e 10 g/l de NaCl, pH 7,5) com 50 μ g/ml de Kanamicina e incubados a 37°C por 12 horas. O DNA plasmidial foi extraído e digerido com enzimas de restrição que liberassem fragmentos de tamanho conhecido, que foram analisados por eletroforese em gel de agarose.

16-Crescimento de bactérias e indução de expressão

Bactérias BL21(DE3)pLys S transformadas com as construções pET28a e pET28aBZH1 foram pré-inoculadas em 10 ml de meio de cultura 2YT acrescido de 50 µg/ml de Kanamicina e 34 µg/ml de cloramfenicol e incubadas a 37 °C em agitação constante durante a noite. O pré-inóculo foi diluído 1: 50 num volume de 800 ml de meio fresco e as bactérias foram crescidas até atingirem a fase logarítmica (OD 600~ 0,7).

A expressão foi induzida pelo acréscimo de IPTG a uma concentração final de 1 mM na cultura. Amostras foram retiradas a 2, 4, 5 e 6 horas de indução, para verificação do melhor tempo de indução que foi estabelecido em 4 horas.

17-Purificação e identificação das proteínas recombinantes

a) Lise bacteriana

Após a indução, as culturas foram divididas em duas alíquotas de 400 ml cada e centrifugadas a 4000 x g por 10 min a 4°C. O sobrenadante foi descartado e o precipitado de bactérias foi suspenso em 10 ml de tampão de lise N (50 mM Fosfato de sódio pH 7, 300 mM Cloreto de sódio, 1 mM PMSF, 5mM Benzamidina, 1mM Imidazol). As suspensões de bactérias foram então submetidas a 3 ciclos de congelamento em nitrogênio líquido e descongelamento a 37°C. Em seguida, o lisado bacteriano foi sonificado em gelo, por 3 pulsos consecutivos de 10 s cada e centrifugado a 4°C por 10 min a 10000 x g. O sobrenadante foi passado para um tubo novo e reservado para o próximo passo.

b) Purificação em coluna de Cobalto

As proteínas de fusão foram purificadas por cromatografia de afinidade em coluna de cobalto Talon® (Clontech). Um volume de 1 ml de resina Talon® foi aplicado numa seringa de 3 ml e equilibrado com 5 ml de tampão de lise N. O sobrenadante final da lise bacteriana foi passado pela coluna e a coluna foi lavada com 10 ml de tampão de lavagem (50 mM Fosfato de sódio pH 7, 300 mM Cloreto de sódio, 1 mM PMSF, 5mM Benzamidina, 5mM Imidazol). A proteína de fusão foi eluída da coluna com acréscimos de 1-2 ml de tampão de lavagem, contendo imidazol numa concentração crescente de 10 a 150 mM.

c) Determinação da concentração protéica

A concentração protéica dos eluatos foi medida pelo método de Bradford (Bradford, 1976) e por comparação visual das bandas no gel de poliacrilamida coradas com Azul de Coomassie com uma proteína purificada de concentração conhecida (BSA).

d) Gel de poliacrilamida (SDS-PAGE)

Amostras do lisado bacteriano e do eluato protéico foram submetidas à eletroforese em gel de poliacrilamida a 15%. A 20 µl de cada amostra foi adicionado 5 µl de tampão de amostra (0,28 M Tris-HCl pH 6,8, 30 % Glicerol, 2,5% SDS, 5% β-mercaptoetanol e 0,002% azul de bromofenol). As amostras foram incubadas a 95° C por 5 min, e centrifugadas por 3 min a 12000 x g antes de serem carregadas no gel. A eletroforese foi executada a temperatura ambiente, 100 V e a uma amperagem máxima de 100 mA em tampão de eletroforese (25 mM Tris Base, 192 mM Glicina e 1% SDS). Ao término da eletroforese, os géis foram corados com solução corante de proteínas (54% álcool etílico, 13,5% ácido Acético Glacial e 0,27% azul de Coomassie) por aproximadamente 2h, com leve agitação, e descorados em solução descorante (10% álcool etílico e 5% ácido acético glacial) até a visualização satisfatória das bandas protéicas.

e) Western Blot

As proteínas do eluato foram separadas por eletroforese em gel de poliacrilamida, e transferidas para uma membrana de Nylon (Hybond N® - Amersham) por transferência eletroforética em Mini Trans-blot® Electrophoretic Transfer Cell (Bio Rad). As condições de transferência foram: voltagem constante a 100 V, temperatura a 4°C em tampão de transferência (25 mM Tris pH 8,3; 192 mM glicina e 20% metanol (v/v)) por 1 h.

Após a transferência, a membrana foi cuidadosamente separada do gel e incubada, à temperatura ambiente, em solução bloqueadora TBST com leite (100 mM Tris pH 7,5; 0,9% NaCl; 0.01% Tween 20 e 5% leite em pó Molico®) por 1 hora. Em seguida, a membrana foi incubada com o anticorpo primário Anti-his(tag) policlonal de coelho (Santa Cruz) diluído 1:1000 em TBS com leite (sem Tween) durante toda a noite a 4 °C. Na manhã seguinte, a membrana foi lavada três vezes durante 5 min em TBST sem leite e incubada durante 1 hora com o anticorpo secundário Anti-IgG de coelho conjugado a fosfatase alcalina (Santa Cruz)

diluído 1:2000 em TBS com leite. A membrana foi novamente lavada 3 vezes com TBST e incubada numa solução 100 mM NaCl e 100 mM Tris-HCL pH 9,5 por 5 min. À solução anterior, foi adicionado CSPD® na proporção 1:2000 e a membrana foi novamente incubada por 5 min. A membrana foi então exposta a um filme T-Mat® S (Kodak) por 1 min e o filme foi revelado.

18-Preparação das sondas de DNA para EMSA

As sondas de DNA de dupla fita foram preparadas a partir de pares de oligonucleotídeos sintéticos (Gibco) (tabela II). Entre 0,5 a 1 nmol de oligonucleotídeos foram anelados em tampão de anelamento (100 mM NaCl, 10 mM Tris-HCL pH 7,5; 1 mM EDTA) por aquecimento a 94°C e resfriamento gradual à temperatura ambiente. Após o anelamento, a marcação radioativa da sonda foi feita por preenchimento das extremidades com o fragmento Klenow da DNA Polimerase I em presença dos nucleotídeos marcados radioativamente [α P³²]dATP e [α P³²]dCTP.

As sondas foram purificadas em colunas de Sephadex® G-25 (NAP™-5 Pharmacia Biotech) segundo instruções do fabricante e foram quantificadas em espectrofotômetro a 260 nm. As atividades específicas foram determinadas através de contagem de Cerenkoff em contador líquido de cintilação (Rackbeta 1217).

19-Ensaio de Retardamento da Mobilidade Eletroforética (EMSA)

O ensaio de retardamento de mobilidade eletroforética que também é conhecido por EMSA ou Gel Shift, é um método bastante utilizado na análise qualitativa e quantitativa da interação proteína-ácido nucléico. Este método é baseado na observação que a mobilidade eletroforética de um ácido nucléico em um gel de poliacrilamida pode ser alterada se tiver uma proteína ligada a ele (Carey, 1991).

A formação dos complexos DNA-Proteína foi analisada através da mistura de diferentes concentrações da proteína recombinante em tampão de ligação (10 mM Hepes pH 7,9, 50 mM KCl, 1mM EDTA, 10 mM DTT, 2 mg/ml albumina sérica bovina, 25 ng/ μ l DNA de timo de bezerro e 10% glicerol) à temperatura ambiente por 5 min. Cada reação foi adicionada de 200-1000 fmoles de sonda marcada radioativamente (5000 – 25000 cpm) num volume total de 50 μ l e incubada por 20 min a temperatura ambiente.

	Pares de oligonucleotídeos sintéticos de DNA dupla fita utilizados na preparação das sondas cujas extremidades são preenchidas com dinucleotídeos radioativos.
Sonda GP	GLM Palíndrome 5'-GCTCTAGATGACAAAATT GACTAGGAGACAT TCATCTCTAGCTTA-3' 3'-ACTGTTTTAACT GATCCTCTGTACAGTAGAGATCGAATCTTAAGG -5'
Sonda MM	5'-GCTCTAGATGACAAAATT TaAtTAGGAGACAT TCATCTCTAGCTTA-3' 3'-ACTGTTTTAA TaATCCTCTGTAAAGTAGAGATCGAATCTTAAGG -5'
Sonda GM	5'-GCTCTAGATGACAAAATT GACTAGGAGACAT TCATCTCTAGCTTA-3' 3'-ACTGTTTTAACT GATCCTCTGTAAAGTAGAGATCGAATCTTAAGG -5'
Sonda MP	5'-GCTCTAGATGACAAAATT TaAtTAGGAGACAT TCATCTCTAGCTTA-3' 3'-ACTGTTTTAA TaATCCTCTGTACAGTAGAGATCGAATCTTAAGG -5'
Sonda G-box	5'-AAGCTTAGCC CAGTGGCACTCGAG -3' 3'-TGAGCTC-5'
Sonda C-box	5'-AAGCTTAGT GACGTCACACTCGAG -3' 3'-TGAGCTC-5'

Tabela II. Sequências das duplas de oligonucleotídeos utilizados na preparação das sondas para os ensaios de EMSA. Os motivos GLM, Palíndrome, G-box e C-box estão destacados em negrito.

Após o término do tempo de incubação da proteína com o DNA para a formação do complexo, foram adicionadas às reações 5 µl de tampão de amostra (0,25 M Tris-Cl pH 7,5, 40 % glicerol , 0,2% azul de bromofenol, 0,2% Xileno cianol). As amostras foram então submetidas à eletroforese em gel não desnaturante de poliacrilamida 5% contendo 5% de glicerol e 0,25 x TBE (50 mM Trizma base 42 mM ácido bórico, 0,5 mM EDTA). O gel foi pré-corrido em tampão TBE 0,25 x por 30 min a ~20 V.cm⁻¹ e 4°C. A eletroforese foi realizada a 400 V e a 4°C por aproximadamente 90 min. O gel foi então seco à vácuo e exposto a um filme de Raio-X por 12 a 48 h à -70°C.

Nos ensaios de competição, frações molares crescentes do DNA competidor, em relação à sonda marcada, foram incubadas por 20 min, à temperatura ambiente, com uma concentração pré-definida da proteína. Terminado o tempo de incubação do DNA competidor com a proteína, uma concentração fixa da sonda marcada radioativamente foi adicionada a cada reação de ligação. As reações foram novamente incubadas por 20 min à temperatura ambiente e adicionadas de 5 µl de tampão de amostra e carregadas no gel de poliacrilamida como já descrito no parágrafo anterior.

RESULTADOS

I- Caracterização do gene *BZO2H1*

Isolamento e caracterização do cDNA *BZO2H1*

No contexto de um projeto visando identificar genes homólogos ao locus *O2* de milho, foi isolado um cDNA codificante para uma proteína bZIP que foi designado *BZO2H1* (Basic Leuцин Zipper Opaque-2 Homologous). A sequência desse cDNA é mostrada na figura 4. Sua fase de leitura prevista definiu uma proteína de 417 aminoácidos. Outro clone de cDNA de *BZO2H1* revelou a existência de um “splicing” diferencial que resulta da seleção de um sítio acceptor 3’ alternativo no íntron 2. Esse splicing determinou uma diferença de 17 pares de bases no RNA mensageiro ou 6 aminoácidos sobre o terceiro exon (Figura 5). Além disso, foram identificados dois sítios de poliadenilação situados a 162 e 190 pares de bases após o códon de terminação TGA respectivamente (Figura 5). O cDNA *BZO2H1* apresentou 100% de identidade com sequências do cromossomo IV de *Arabidopsis thaliana*.

Para determinar se os cDNAs isolados representavam a sequência completa dos RNAs mensageiros *BZO2H1*, a técnica 5’RACE foi aplicada sobre o RNA total da semente. Essa técnica permite, a princípio, definir a extremidade 5’ do RNA mensageiro. Os resultados obtidos pela técnica 5’RACE confirmaram a sequência do cDNA que já havia sido definida.

Localização do gene *BZO2H1* e seu padrão de expressão

Com o projeto de sequenciamento do genoma de *Arabidopsis thaliana* concluído (The Arabidopsis Genome Initiative, 2000), o gene *BZO2H1* foi identificado através da busca de sequências genômicas idênticas ao cDNA *BZO2H1*. Desta maneira, o gene *BZO2H1* foi localizado no BAC T1011 a aproximadamente 15 cM do mapa genético do cromossomo IV (Figura 5).

1 M N E I F S I T E F S E 12
 1 AGAAAGTGAAAAAGAGAGCGAGCGATGAACAGTATCTTCTCCATTGACGATTTCTCCGAT 60
 13 F F W E V E E I E L N E T S O V P V E A 61
 61 CCTTTCTGGGAAACTCCTCCGATTCTCTCAATCCCGACTCTTCTAAGCCTGTTACGGCG 12
 33 T E V S S S E E W T F I M F L E R I S 52
 121 GATGAAGTTAGCCAGAGTCAACCGGAATGGACTTTCGAGATGTTTCTCGAAGAGATTTCT 180
 53 F F A V F A V E L E N N V A I V E V S 72
 181 TCGTCGGCGGTGAGCTCTGAGCCACTTGGTAACAACAACAACGCGATCGTCGGTGTTCCT 240
 73 S A Q S L P A V A G Q V F E R T P S F F 92
 241 TCGGCGCAATCTCTTCTCTGTTTCCGGACAGAATGATTTTCGAGGATGATAGTCGATTT 300
 93 R E E S G V L P C A A S M V V V V I 112
 301 CGTGATCGCGATTCCGGGAAATTTGGATTGTGCTGCTCCCATGACGACGAAGACGGTGATT 360
 113 V E S E R Y F F V L N K I L T Y S C A T 132
 361 GTTGATTCCGATGATTATCGTCGTGTTCTTAAGAACAAGCTTGAGACTGAGTGCCTACT 420
 133 V V E L T V G S V V E E T S E P E E 152
 421 GTTGTTCCTCTTCGGGTGGGTCTGTGAAGCCTGAAGATTCGACTAGTTCTCCAGAAACT 480
 153 C L Q E V Q S E E L T Q G A I M T P G S 172
 481 CAACTTCAACCAGTTCATCCAGTCCTCTTACTCAAGGTTCTTTGATGACCCAGGAGAA 540
 173 L G V T S P L P A V V V G V S M N E 192
 541 CTTGGTGTACTTCTTCTTACCAGCTGAGGTGAAAAAACTGGTGTATCAATGAAGCAG 600
 193 V T S G S E F E Y S P E E T L P E I N E 212
 601 GTTACTAGTGGATCGTCGAGAGAATATTCTGATGACGAGGACCTTGATGAGAGAAATGAA 660
 213 T T G S L K E E D V K K S R R M L S N E 232
 661 ACCACCGGTTCTTGAAGCCAGAGGACGTTAAAAAATCTAGAAGGATGCTGTCAAATCGT 720
 DOMÍNIO BÁSICO
 233 E S A R R S R R R R R R R R R R R R R 252
 721 GAGTCAGCTAGGCGATCTAGAAGGAGAAAGCAGGAGCAACAAGTACCTCGAAACACAG 780
 ZÍPER DE LEUCINA
 253 V N D L K G E H S S L L K Q L S N M N H 272
 781 GTTAATGATCTAAAAGGTGAGCATTTCATCACTTCTTAAACAAGTGAACATGAATCAC 840
 273 K Y D E A A V G N R I L K A D I E T L R 292
 841 AAGTATGACGAGGCTGCTGTTGGCAATAGAATACTAAAGGCTGACATTGAGACATTAAGA 900
 293 A K V K M A E E T V K R V T G M N E M L 312
 901 GCTAAGGTGAAAATGGCGGAAGAAACCGTGAAGAGAGTAACAGGAATGAATCCGATGCTT 960
 313 L G S S G A N N N N K M P I T G N N F 332
 961 CTCGGAAGATCAAGTGGACATAACAACAACAACAAGTGAACATAACTGGTAACAACAGG 1020
 333 M C S S S I I D A Y Q S V S V L N E M S 352
 1021 ATGGATTCTTCTAGCATTATTCAGCTTATCAACCACACTCAAACCTAAACCATATGTCA 1080
 353 M Q N I G I F T I L E F A D G H D E A A 372
 1081 AACCAAAACATCGGGATCCCAACCATTTCTACCTCCAAGACTCGGAAACAATTTTCGCTGCT 1140
 373 F E S S S S E L Q P I A S G Q D F F V 392
 1141 CCTCCATCCCAAACAGCTCTCCCTTGAGAGAAATAGAAATGGGCAAAATCACCATGTT 1200
 393 T E S A N E Y G W Y I E E C P E S A W P 412
 1201 ACTCCAAGCGCCAACCCGTATGGCTGGAATACCGAACCTCAGAACGATTGAGCATGGCCG 1260
 413 K F C V D * 417
 1261 AAAAAATGCGTGGACTGATCAACAAGAAGCGGGTTTCGCACTATATTAATGTCTATGCA 1320
 1321 TCTGTAATTTGTAAGTGTATTAAAGTTACGAATCATGAGAAAACATCTTGTGAAAATACA 1380
 1381 GTCTCATGGCTTATATATATATAAGCTCTGTCTTATAACATTACAAGATTCTTATTTG 1440
 1441 AGAATCGTCTTTCTATTATAGCTAATAACAGAGTGTAAAAA 1477

└─AAAAAA (1440) Primeiro poly A

└─Segundo Poly A

Figura 4. Sequência de nucleotídeos do cDNA BZO2H1 com seqüências de aminoácidos deduzida. A seqüência sublinhada de azul claro representa o domínio básico e a sublinhada de amarelo, o zíper de leucina. As leucinas e os aminoácidos hidrofóbicos posicionados a cada sete resíduos do zíper de leucina estão destacados **em negrito**. Os diferentes grupos de aminoácidos estão representados pelas seguintes cores: de caráter ácido (verde escuro), básicos (amarelo), alifáticos (azul) , aromáticos (vermelho), com enxofre (rosa) ou hidroxilados (verde azulado). O asterisco representa o códon de terminação. Os dois sítios de poliadenilação estão demarcados na seqüência.

A análise do DNA genômico de *Arabidopsis thaliana* por Southern Blot com a sonda de cDNA BZO2H1 de 1440 pb, sugere que o gene *BZO2H1* esteja presente em cópia única no genoma de *Arabidopsis*. Além disso, o padrão de bandas obtidas nessa análise confere com o mapa de restrição deduzido a partir das seqüências do BAC T1011 (Figura 6).

A expressão do gene foi analisada por Northern blot. RNA total foi extraído de flores, raízes, caule e flores da planta adulta. RNA total de sementes foi extraído de plantas de quatro, cinco e seis semanas de idade. Além disso, RNA total da plântula com três semanas, crescidas sob condições de presença ou ausência de luz também foi extraído. *BZO2H1* é expresso em todos os tecidos analisados. Contudo, o menor nível de expressão verificado na aemente, parece estar relacionado com variações na concentração de RNA aplicado no gel. Comparando a expressão do gene em plantas expostas à luz e em plantas crescidas no escuro, não foi observada alteração na expressão do gene (Figura 7).

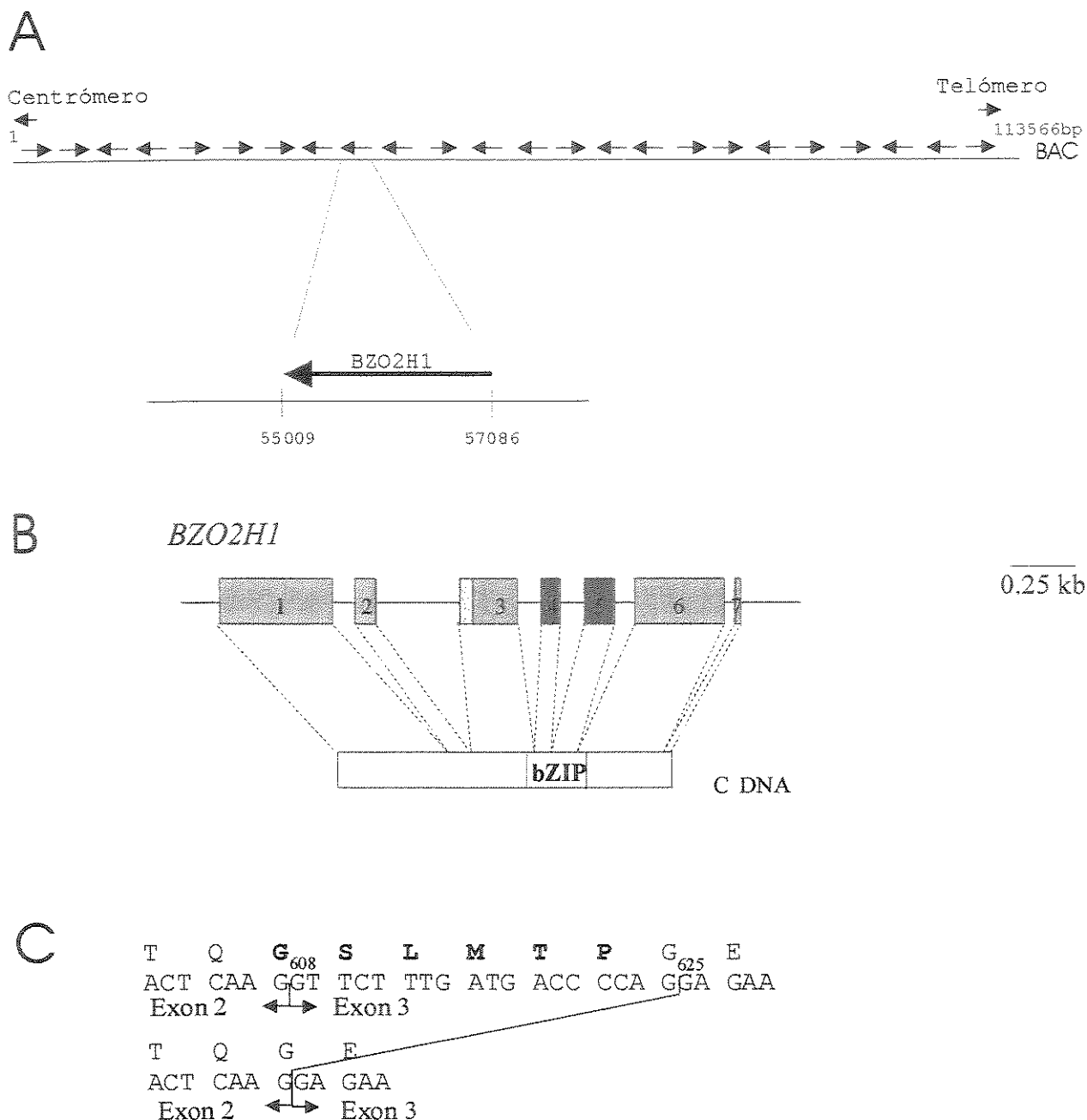


Figura 5. Representação esquemática da localização cromossômica e estrutura do gene *BZO2H1*

A) Localização do gene *BZO2H1* no cromossomo IV de *Arabidopsis thaliana*. A posição do gene no mapa genético é 15 cM no BAC T1011 e seu número de acesso pelo NCBI é AC002330. As setas representam genes presentes no intervalo de sequência e suas respectivas direções de transcrição.

B) Estrutura do gene *BZO2H1*. Os retângulos numerados representam os exons e as linhas, os íntrons. O retângulo amarelo representa a sequência envolvida no splicing diferencial (nucleotídeos 608 a 625).

C) Sequência de nucleotídeos que revelou existência de splicing diferencial entre os nucleotídeos 608 e 625, determinando a formação de duas isoformas da proteína *BZO2H1* que diferem pelos 6 aminoácidos destacados em negrito.

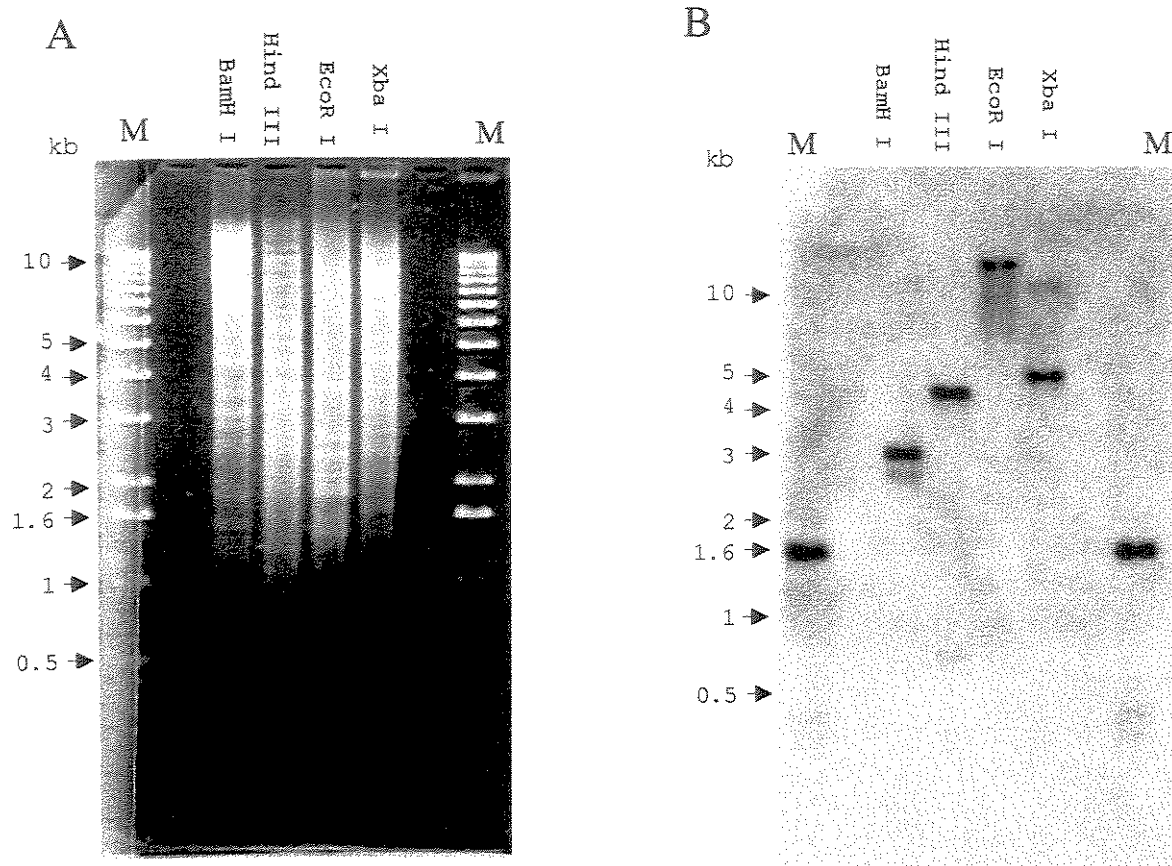


Figura 6. Análise do DNA genômico de *Arabidopsis thaliana* por Southern blot. A) DNA genômico foi digerido com enzimas *Bam* HI, *Hind* III, *Eco* RI e *Xba* I e separado num gel de agarose 0,8%. M = Marcador de Peso Molecular. B) Após a separação, o DNA foi transferido para uma membrana de nylon e hibridizado com sonda de cDNA BZO2H1. Os números ao lado da foto indicam os tamanhos dos fragmentos do padrão de DNA em kb.

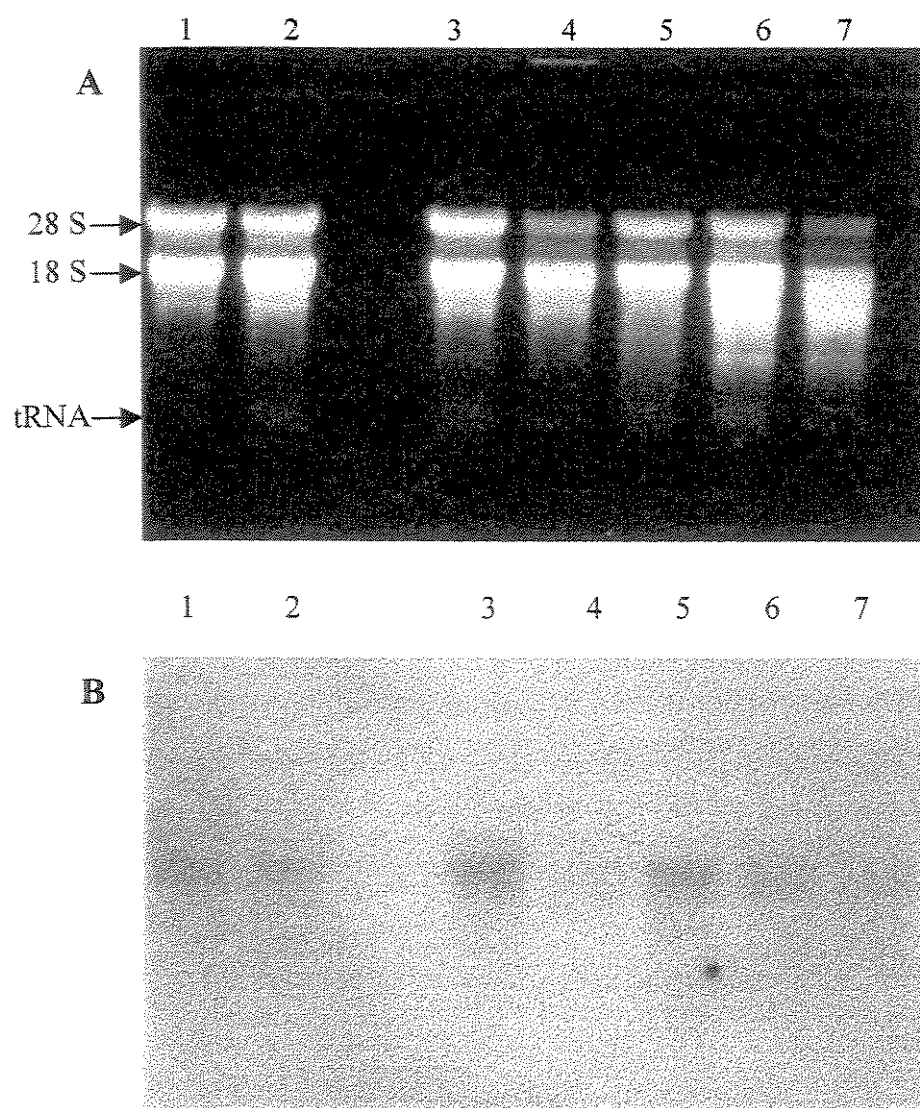


Figura 7. Análise da expressão do gene *BZO2H1* por Northern blot.
 A) Gel desnaturante de agarose 1% com 20 μ g de RNA total de *Arabidopsis thaliana* por poço. 1-Plântula crescida em ausência de luz, 2- Plântula crescida em presença de luz, 3- Flor, 4- Semente, 5- Raiz, 6- Caule, 7- Folha. Setas indicam RNA ribossômico de 28 e 18 S e RNA transportador.
 B) Autoradiografia da membrana correspondente ao gel hibridizada com sonda de cDNA *BZO2H1*

O gene *BZO2H1* pertence à Família Multigênica O2

A estrutura do gene foi deduzida por comparação do cDNA *BZO2H1* com o DNA genômico de *Arabidopsis thaliana*. O gene *BZO2H1* é composto por 7 exons e 6 introns (Figura 5). Pela análise da sequência dos exons, foi observado que 95% do domínio bZIP de *BZO2H1* é codificado pelos exons 4 e 5, cujo tamanho e posições se encontram também conservados entre os genes *BZL1* de cevada, *RTA1* de arroz, *O2* de Coix, *O2* de sorgo, *O2* de milho, *BZO2H2*, *BZO2H3* e *BZO2H4* de *Arabidopsis* (Figura 8). Essa característica é específica desse grupo de genes (Vettore *et al.*, 1998). Parece, portanto que todos esses genes são membros homólogos da Família Multigênica O2 (Vincentz *et al.*). Pelo alinhamento das proteínas codificadas pelos genes da família O2, foi observada uma alta similaridade de sequências do domínio bZIP entre seus membros (Figura 8). As proteínas *BZO2H1* e *O2* de milho compartilham 54,7% de aminoácidos idênticos e 72,1 % de aminoácidos similares nesse domínio, esses dados foram baseados na matriz similaridade de Gonnet (Gonnet *et al.*, 1992). Em razão dessa semelhança, foi previsto que *BZO2H1* poderia apresentar especificidade de ligação ao DNA parecida com a de *O2* como descrito na literatura (Isawa, *et al.*, 1993).

cdefgabcde f gabcde f gabcde f gabcde f gabcde f gabcde f gabcde f g

Intron 5 K M

AAGGTG

Exon 6

```

REB (Os)      403  ---LEHLOKRMCG-----GPASS-GSTS-----
BLZ1 (Hv)     373  ---LEHLOKRMCG-----GPASS-GSTS-----
OHP1 (Zm)     387  ---LEHLOKRMCG-----GPASS-GSTS-----
OHP2 (Zm)     392  ---LEHLOKRMCG-----GPASS-GSTS-----
O2 (Cij)      370  ---AMELIQKTMG-----AMPTS-SSA-LQESQLLGLGPDETINMDMY
O2 (Sb)       380  QPSMQLIQETMGA---MMPTSS-GST-LQESSELL-GPNETINMHMY
O2 (Zm)       397  ---AMELIQTMGS---MPPTSASGSTFPFPQDYELL-GPNGAIHMDMY
SPA (Ta)      387  ---TSAPTPQREST---TLDSTEIV-----NMVM-
BLZ2 (Hv)     386  ---TLVPTPQREST---TLDSNEIG-----NMVMQ
CPRF2 (Pc)    379  ---LEHLOKRIIRGG---VSSCEAQVSGKQ-----
BZO2H3 (At)   301  ---LEHLOKRIIRSV---GDQ-----
HBF1 (Gm)     356  ---LEHLOKRIIRGE---VSSCGTQGRGEQ-----
BZO2H4 (At)   383  ---QNPFAPDSNLY---ETLPHWNHKKH-----
BZO2H1 (At)   383  ---NHHVTPSANPYGWNTEPQNDASWPKKQVD-----
BZO2H2 (At)   261  ---LGNVSSEAVS---CVSDIWP-----
RITA1 (Os)    278  ---GGGEVGDAGW---WDNHSNGAMSK-----

```

Os cinco motivos conservados VFEW, EFWF, DYLLR, TSA e bZIP entre os membros da família estão indicados acima do alinhamento. A posição dos aminoácidos de cada repetição de 7 resíduos do zíper de leucina está indicada (a-g) e a posição (d) das leucinas e resíduos hidrofóbicos se encontra destacada em negrito. Aminoácidos idênticos e conservados em pelo menos 50% das proteínas estão demarcados pelo fundo cinza. Resíduos estritamente conservados entre os membros da família estão destacados pelo fundo preto. Proteínas cuja estrutura é conhecida se encontram sublinhadas e as posições dos íntrons 3, 4 e 5 estão demarcadas na figura. *At*: *Arabidopsis thaliana*; *Clj*: *Coix lacryma-jobi*; *Gm*: *Glycine max*; *Hv*: *Hordeum vulgare*; *Os*: *Oryza sativa*; *Pc*: *Petroselinum crispum*; *Sb*: *Sorghum bicolor*; *Ta*: *Triticum aestivum*; *Zm*: *Zea mays*. Os números de acesso são: SPA: T06767; BLZ2: CAA71795; REB: BAA36492; BLZ1: T04477; OHP1: JQ2147; OHP2: JQ2148; O2Zm: A34800; O2Jr: S42493; O2Sb: CAA50642; HBF1: T07154; CPRF2: Q99090; RITA1: AJ001267, BZO2H1: AAC78255; BZO2H2: F6A410; BZO2H3: AAF67360; BZO2H4: CAB77582.

II-Comparação da especificidade de ligação da proteína BZO2H1 e O2 de *Coix*

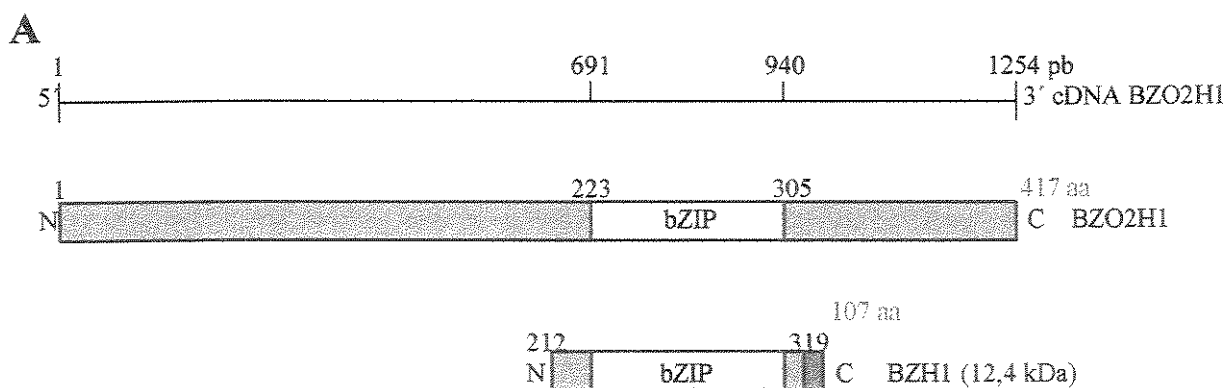
Para definir a especificidade de ligação ao DNA da proteína BZO2H1, construções gênicas permitindo sua expressão como proteína recombinante em *E. coli* foram realizadas.

Expressão e purificação da proteína BZH1

A fase codificante do cDNA BZO2H1 foi clonada no vetor de expressão pET28a e a expressão da proteína foi induzida em *E. coli*. Embora a proteína recombinante BZO2H1 tenha sido expressa na bactéria, o nível de expressão alcançado não foi suficiente para sua utilização em experimentos de Ensaio de Retardamento da Mobilidade Eletroforética (EMSA) (Resultado não mostrado).

Para contornar essa dificuldade, a construção realizada englobou apenas a região do domínio bZIP da fase codificante do cDNA BZO2H1 e foi ligada no vetor de expressão pET28a. Essa construção foi denominada de pET28aBZH1. Segundo indicações da literatura, o domínio bZIP apresenta a mesma especificidade de ligação ao DNA do que a proteína inteira correspondente (Suckow & Hollenberg, 1998). Baseada na construção pET28aBZH1, a proteína recombinante correspondente foi denominada de BZH1. Sua composição é de 107 aminoácidos e peso molecular teórico de 12,4 kDa (Figura 9). Devido ao grande número de aminoácidos polares e carregados (65,4%) e da fração relativamente pequena de aminoácidos apolares (34,6%), foi previsto que a proteína BZH1 seria expressa em forma solúvel na bactéria. Assim, a expressão da proteína recombinante foi induzida em bactérias transformadas com a construção pET28aBZH1 pelo acréscimo de IPTG. Após a lise bacteriana, foi verificada a expressão de BZH1 na fração solúvel (Figura 10). Na mesma figura, foi observado que as bactérias transformadas com o vetor pET28a induzidas com IPTG e as bactérias transformadas com a construção pET28aBZH1 não induzidas com IPTG, não expressaram a proteína recombinante BZH1.

A proteína recombinante BZH1 foi parcialmente purificada através de cromatografia de afinidade em colunas de cobalto imobilizado em sefarose (Talon® Clontech). Para isso, um hexâmero de histidina (6xHis-tag) foi adicionado à extremidade C-terminal da proteína recombinante BZH1 pelo vetor pET28a. Esse hexâmero de histidina exibe a propriedade de se ligar, reversivelmente, a metais de transição como o cobalto em condições de pH fisiológico.



B

MGSLKPEDVKKSRRLSNRESARRSRRRKQEQTSDLETQVNDLKGEHSS
LLKQLSNMNHKYDEAAVGNRILKADIETLRKVMAEETVKRVTGMNP
MLEHHHHHHH

Figura 9. Representação esquemática do cDNA inteiro de BZO2H1 e de sua respectiva proteína inteira e do domínio bZIP clonado

A) cDNA de BZO2H1 e respectivas proteínas inteira (BZO2H1) e truncada (BZH1). O retângulo amarelo representa o domínio bZIP e os retângulos verdes, as extremidades N- e C-terminal da proteína. O retângulo azul representa os aminoácidos conferidos pelo vetor de expressão. Os números indicam o tamanho da sequência codificante e da cadeia polipeptídica, assim como a posição relativa do domínio bZIP em destaque.

B) Sequência de BZH1. As metioninas das posições 1 e 15 estão destacadas em negrito, as leucinas e resíduos hidrofóbicos que fazem parte da repetição de 7 aminoácidos do zipper estão destacados em vermelho e os aminoácidos conferidos pelo vetor pET28a, em azul.

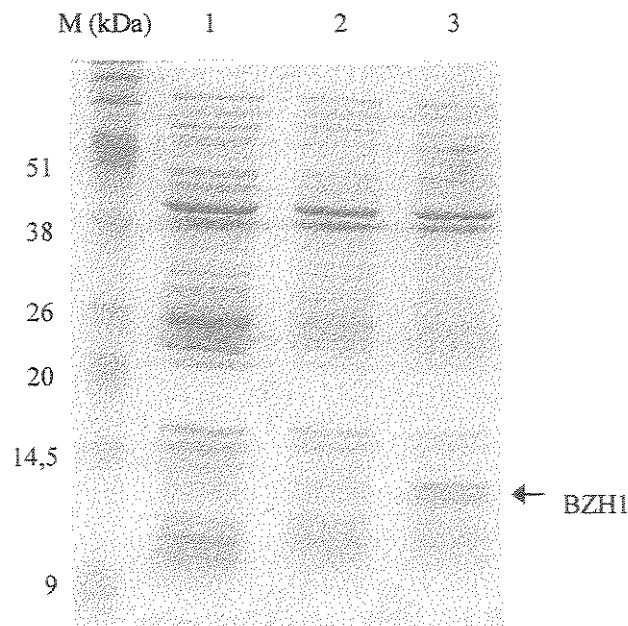


Figura 10. Proteína recombinante BZH1 está presente na fração solúvel do extrato bacteriano.

A expressão da proteína recombinante foi induzida pela adição de IPTG em culturas de bactérias transformadas com os vetores de expressão.

A foto mostra um gel de 15% de poliacrilamida carregado com 20 µl de eluatos das colunas de cobalto, corados com azul de Coomassie na seguinte ordem:

M: Marcador de peso molecular em kDa

- 1) Eluato de bactérias transformadas com vetor controle pET28a induzidas com IPTG
- 2) Eluato de bactérias transformadas com a construção pET28a BZH1 não induzidas
- 3) Eluato de bactérias transformadas com a construção pET28aBZH1 induzidas com IPTG

Essa ligação ocorre por compartilhamento de elétrons do nitrogênio do anel imidazólico da histidina com os orbitais deficientes em elétrons do cobalto. Dessa maneira, a proteína recombinante é ligada, reversivelmente, pela cauda de histidina na coluna e é eluída pelo acréscimo de imidazol. Assim, a proteína recombinante BZH1 foi eluída com concentrações de 50 e 100 mM de imidazol, em forma de uma dupla de bandas com peso molecular de aproximadamente 16 e 15 kDa (Figura 11). A presença desses dois polipeptídeos pode ser explicada pela presença de um segundo códon de iniciação na posição 15 da sequência codificante da proteína BZH1 (Figura 9). Provavelmente, este segundo AUG também seja reconhecido como sítio de iniciação da tradução no ribossomo, causando a formação de dois polipeptídeos (Figura 11). Pela análise da sequência, foi observado que o polipeptídeo menor preserva 65% do domínio de ligação ao DNA e 87% do domínio bZIP inteiro. Nas frações da eluição com 50 e 100mM de imidazol, também foi encontrada uma banda de proteína com peso molecular por volta de 27 kDa. Essa banda, provavelmente representa uma proteína contaminante da bactéria que foi co-purificada na coluna (Figura 11).

No experimento de Western Blot, a identidade dos dois polipeptídeos foi verificada pela detecção do hexâmero de histidina em suas extremidades C-terminais por um anticorpo Anti-6xHis. Os dois polipeptídeos de peso molecular 16 e 15 kDa foram, de fato, reconhecidos pelo anticorpo anti-6xHis confirmando sua identidade como duas formas da proteína recombinante BZH1 (Figura 11).

Caracterização da especificidade de ligação da proteína recombinante BZH1

Para começar a caracterizar alguns aspectos de ligação ao DNA das duas formas da proteína recombinante BZH1 encontradas, foi testado se BZH1 reconheceria os mesmos motivos de DNA reconhecidos por fatores bZIP da família Multigênica O2. Foi mostrado que O2 de *Coix* e de milho reconhecem uma sequência chamada de O2-box, presente no promotor do gene da α -Coixina (Yunes *et al.*, 1998). Essa sequência é formada por dois motivos de ligação ao DNA. Um deles é relacionado com o motivo de ligação do fator GCN4: GLM ou GCN4 Like Motiv (5'-TGACTA-3') e o outro é uma sequência palíndromica denominada Palíndrome (5'-GACATGTC-3') (Hill *et al.*, 1986). O estudo comparativo da ligação das proteínas recombinantes BZH1 e O2 de *Coix* (O2CX) foi iniciado com a sonda GP de 53 pb que contém a O2-box.

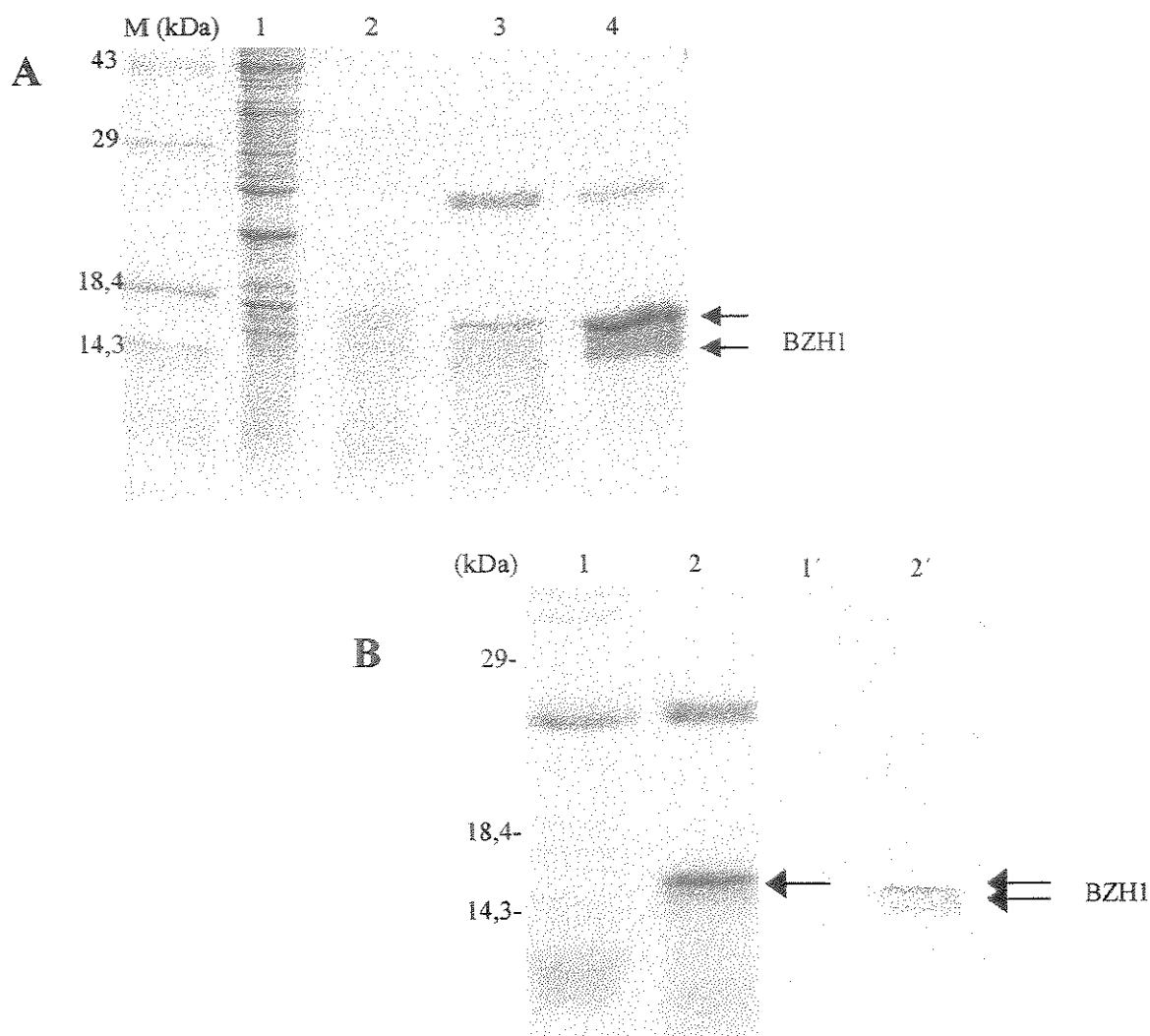


Figura 11. O vetor pET28aBZH1 expressa dois polipeptídeos

A) A fração solúvel das proteínas transformadas com pET28aBZH1 e induzidas com IPTG foi passada por uma coluna de cobalto. As proteínas ligadas a coluna foram eluídas com concentrações crescentes de imidazol em tampão N (Materiais e métodos). Os eluatos foram então submetidos a eletroforese em gel de poliacrilamida-SDS na seguinte ordem: M: Marcador de peso molecular em kDa; 1- Extrato bacteriano total; 2- Lavagem da coluna com 20mM de Imidazol; 3- Eluição da coluna com 50 mM de Imidazol; 4- Eluição da coluna com 100mM de Imidazol. B) Gel de poliacrilamida-SDS e respectivo filme de Western blot 1- Eluato de bactérias transformadas com o vetor pET28a, 2- Eluato da proteína recombinante BZH1, 1' e 2'- mesmas amostras transferidas para uma membrana de nylon que foi incubada com um anticorpo primário Anti-his(tag) e anticorpo secundário acoplado a fosfatase alcalina. As duas setas indicam a posição dos dois polipeptídeos BZH1.

Nos experimentos descritos em seguida, a proteína recombinante O2 de *Coix* (O2CX) foi utilizada como controle nos experimentos de Ensaio do Retardamento da Mobilidade Eletroforética (EMSA). A comparação das seqüências polipeptídicas de BZH1 e O2CX que foram utilizadas nos ensaios é mostrada na figura 12.

Otimização das condições de ligação da proteína BZH1 à sonda GP.

Nos experimentos preliminares visando estabelecer as condições adequadas de EMSA, foi constatado que a proteína recombinante BZH1 tem a capacidade de se ligar a sonda GP formando dois complexos Proteína/DNA (Não mostrado).

Uma curva de saturação da proteína recombinante BZH1 com a sonda GP foi estabelecida para, posteriormente, definir a especificidade de ligação da proteína recombinante BZH1 nos ensaios de competição. Com o aumento da concentração protéica, o aparecimento de duas bandas de retardamento indica a formação de dois complexos de ligação ao DNA. Um complexo inferior (I), menos intenso e outro superior (S), mais intenso (Figura 13). A interpretação deste resultado é que a formação do complexo I resulta da ligação de um homodímero da proteína BZH1 à sonda GP enquanto que a formação do complexo S resulta da ligação de um segundo homodímero da proteína BZH1 no motivo adjacente da mesma sonda. Este fenômeno de formação de dois complexos também já foi verificado com a proteína recombinante controle O2CX e a sonda GP (Yunes *et al.*, 1998). O aparecimento de dois complexos e a transição rápida do complexo I para o complexo S com o aumento da concentração da proteína, provavelmente, esteja relacionado com o fenômeno da cooperatividade de ligação da proteína recombinante aos sítios GLM e Palíndrome da sonda GP. Esse fenômeno indica que a ocupação de um sítio por um dímero da proteína recombinante BZH1 facilita a ocupação do sítio adjacente por um segundo dímero, como já foi descrito para a proteína O2CX (Yunes *et al.*, 1998). Nesse caso, a ligação de um dímero a um motivo da sonda parece aumentar a afinidade de ligação do segundo dímero pelo motivo adjacente. A diferença na intensidade dos complexos I e S observada, pode também estar relacionada com a diferença de afinidade das proteínas recombinantes BZH1 e O2CX pelos motivos GLM e Palíndrome da sonda GP.

BZH1 MGS LKPEDV **KKSRAMLSNAESARRSRRK**
O2CX MVRLATSSSSRDPSPSDEDMDGEVEILLGFKMPTEER**VEKKRKESNRESARRSRYRK**

QEQTSD**LE**TQVND**LK**GEHSS**LL**KQLSN**M**NHKYDE**AA**VGNRI**LK**ADIET**LR**AKVK**MA**EETV**KRV**
AAHL**KE**LEDQVE**QL**KAENSCLLRRLAAL**NQ**KYNE**AN**VDNR**VL**RADMET**LR**AKVK**MG**EDSL**KRV**

T-GMNPMLLEHHHHHH-----
MPAANKARKEAELAAATAEQ

Figura 12. Alinhamento das seqüências das proteínas recombinantes BZH1 (em vermelho) e O2CX (em preto). O domínio básico está destacado pelo fundo azul e o zíper de leucina pelo fundo amarelo. Os aminoácidos escritos em verde representam a região “hinge”. Os aminoácidos em negrito correspondem às leucinas e resíduos hidrofóbicos da repetição de 7 aminoácidos do zíper de leucina. Os aminoácidos sublinhados representam a seqüência conferida pelo vetor pET28a (para BZH1 em vermelho) e pelo vetor pET3b (para O2CX em preto).

A

Sonda GP

5'-GCTCTAGATGACAAAATTGACTAGGAGACATGTCATCTCTAGCTTAGAATTCC-3'

GLM

Palíndrome

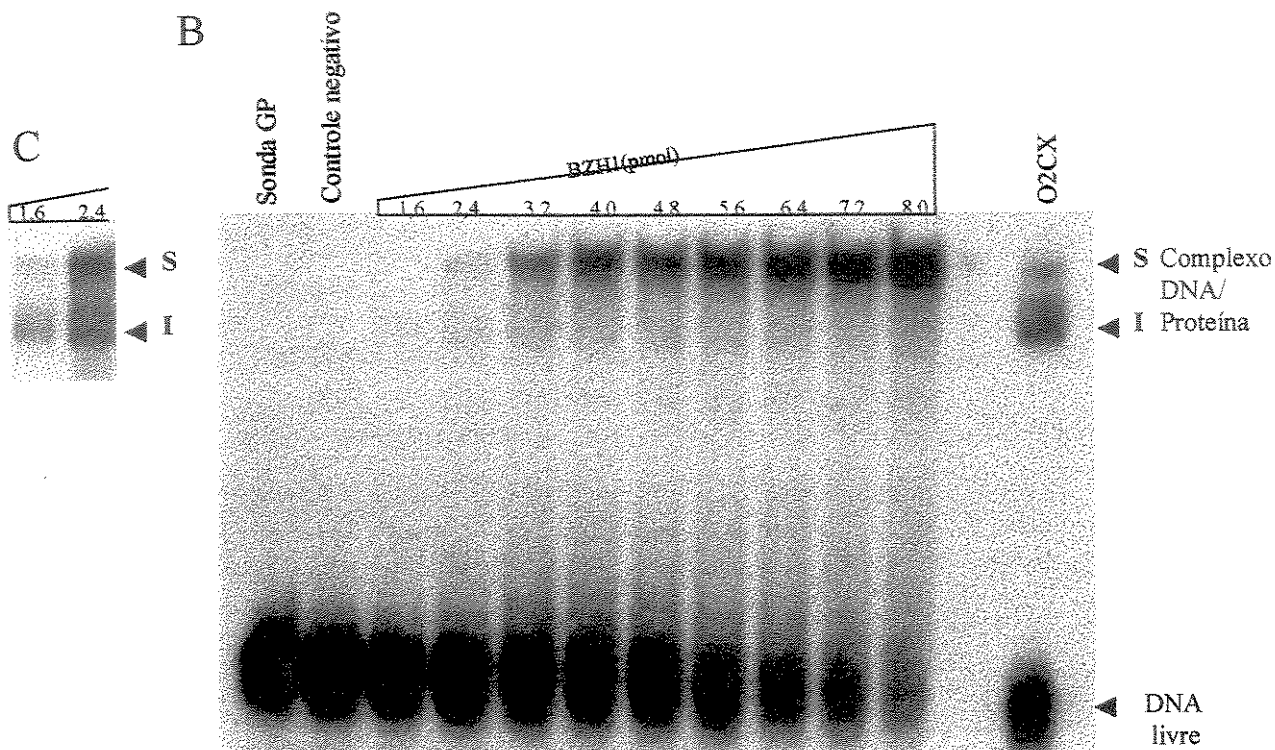


Figura 13. EMSA da proteína recombinante BZH1 com a sonda GP.

A) Sequência da sonda GP. As bases que compõem os motivos GLM e Palíndrome estão sublinhadas

B) EMSA da proteína recombinante BZH1 com o 300 fmol da sonda GP, a concentração da proteína foi gradativamente aumentada como indicada na figura. Os demais poços foram carregados como segue: Sonda GP: 300 fmol da sonda GP marcada radioativamente; Controle negativo: 100 ng de eluato de bactérias transformadas com o vetor pET28a; O2CX: 1,5 pmol da proteína recombinante O2CX.

C) Complexos S e I dos poços com 1,6 e 2,4 pmol de proteína BZH1 realçados pela exposição prolongada do gel ao filme.

A proteína BZH1 se liga especificamente a sonda GP.

No âmbito do método de EMSA, a especificidade é definida como a habilidade de uma determinada proteína de discriminar entre sua sequência alvo e outras sequências presentes no sistema. A especificidade também pode ser definida como a medida da razão da constante de afinidade da ligação ao DNA específica com a constante de afinidade da ligação inespecífica. Já a especificidade é a habilidade que uma proteína apresenta em discriminar entre seus sítios alvos de DNA e outros sítios que não sejam alvos de ligação numa mistura (Carey, 1991).

Para testar a especificidade de formação dos complexos entre uma proteína e uma sonda específica e inespecífica, ensaios de competição utilizando excessos molares de sondas competidoras foram realizados.

Ensaio de competição da proteína O2CX com a sonda GP específica e com a sonda MM inespecífica demonstraram que a ligação da proteína O2 de *Coix* à sonda GP é específica (Yunes *et al.*, 1998). A sonda MM apresenta duas substituições de bases no motivo GLM e uma substituição de base no motivo palíndrome da sonda GP, anulando a ligação da proteína O2CX à sonda MM.

O mesmo tipo de experimento foi realizado com a proteína recombinante BZH1. A sonda mutada MM inibiu eficientemente a interação da proteína recombinante BZH1 com a sonda GP, somente a partir de um excesso molar de 30 x (Figura 14), enquanto que a sonda GP não marcada, competiu eficientemente com a sonda GP marcada a partir de um excesso molar de 5 x. Esse resultado indica que a proteína recombinante BZH1 apresenta ligação específica à sonda GP. Assim o comportamento das proteínas BZH1 e O2CX em relação à sonda GP é similar, pois ambas se ligam à sonda GP de maneira específica.

BZH1 reconhece os motivos GLM e Palíndrome.

Na tentativa de confirmar qual dos motivos de DNA presentes na sonda GP a proteína recombinante BZH1 reconheceria, as sondas MP e GM foram utilizadas. Essas sondas apresentam mutações na sequência GP. A sonda MP possui duas substituições na sequência do motivo GLM e a sonda GM possui uma substituição na sequência do motivo designado Palíndrome (Tabela II).

Um ensaio de retardamento da mobilidade eletroforética da proteína recombinante BZH1 foi feito com a sonda GM. Para isso, uma quantidade fixa de sonda GM marcada radioativamente foi utilizada, enquanto que a quantidade da proteína BZH1 foi gradativamente aumentada (Figura 15). A formação de um complexo DNA/Proteína foi observada com 1,6 pmol da proteína recombinante BZH1. No entanto, quando a quantidade de proteína foi aumentada, ocorreu a formação de um segundo complexo superior instável em forma de rastro (Figura 15). No mesmo experimento, a proteína controle O2CX formou somente um complexo com a sonda GM em concentrações semelhantes a da proteína recombinante BZH1. Esse comportamento da proteína recombinante O2CX já era esperado, pois a substituição na sequência do motivo Palíndrome, assim como as duas substituições no motivo GLM, diminuem consideravelmente a afinidade de ligação da proteína O2CX a esses motivos, formando apenas um complexo da proteína O2CX com o motivo não mutado (Yunes *et al.*, 1998). Na figura 16, o mesmo experimento foi feito, porém com a sonda marcada MP. Resultados muito semelhantes aos do experimento anterior foram obtidos. A proteína BZH1 formou novamente um complexo inferior e um complexo instável superior à medida que sua concentração foi aumentada. A proteína recombinante controle O2CX, continuou apresentando um único complexo como esperado.

A formação de um complexo DNA/Proteína mais a formação de um complexo instável, em formato de rastro, da proteína recombinante BZH1 com as sondas mutadas GM e MP, sugere que as substituições feitas diminuem a afinidade de ligação da proteína recombinante BZH1 aos motivos Palíndrome e GLM. A comparação dos resultados obtidos com BZH1 e O2CX, também indica que as propriedades de ligação dessas duas proteínas com as sondas MP e GM são diferentes.

BZH1 e O2CX reconhecem G-box, mas BZH1 se liga a C-box com afinidade inferior que O2CX.

Dados da literatura demonstraram que a proteína O2 de milho se liga aos motivos G-box e C-box com afinidade de ligação comparável (Isawa *et al.*, 1993). Para testar se a proteína recombinante BZH1 também reconheceria os mesmos motivos, experimentos de Retardamento da Mobilidade Eletroforética com a proteína recombinante BZH1 utilizando sondas que continham os motivos G-box e C-box foram realizados. Na figura 17, foi

observado um aumento da intensidade da banda de complexo formada com o aumento gradativo da concentração das proteínas recombinantes BZH1 e O2CX a uma concentração fixa da sonda G-box marcada radioativamente. O mesmo experimento foi repetido com a sonda C-box e a formação do complexo entre a proteína recombinante BZH1 e a sonda C-box foi observado somente na concentração mais alta da proteína testada. Enquanto que a proteína recombinante controle O2CX formou uma banda de complexo com a sonda C-box em todas as concentrações testadas. Esse resultado demonstra uma diferença relativa da afinidade de ligação das proteínas recombinantes BZH1 e O2CX com a sonda C-box.

A

Sonda GM

5'-GCTCTAGATGACAAAATTGTACTAGGAGACATtTCATCCTCAGCTTAGAATTCC-3'

B

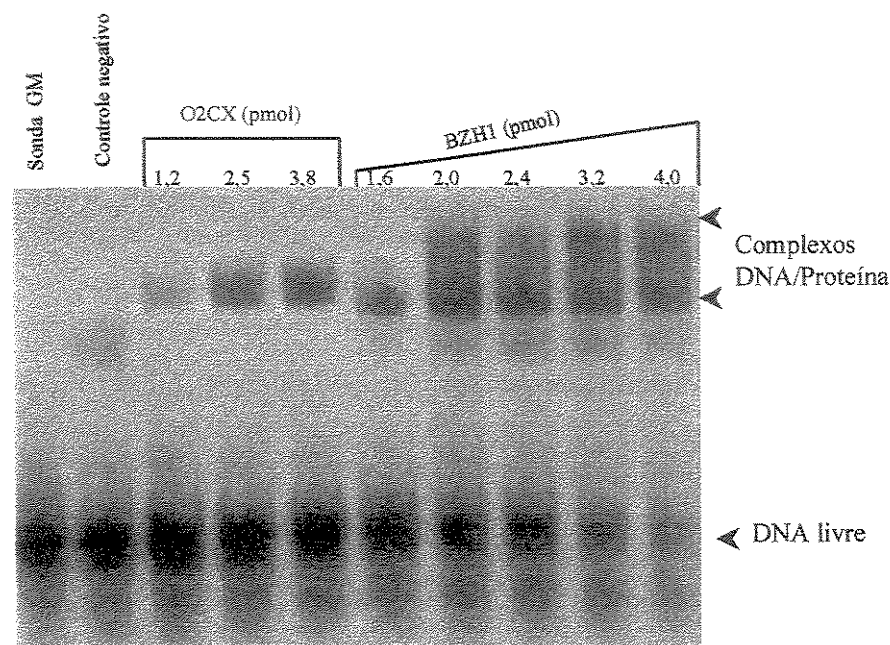


Figura 15. EMSA das proteínas recombinantes O2CX e BZH1 com a sonda mutada GM.

A) Sequência da sonda GM. Os motivos GLM e Palíndrome se encontram sublinhados e a substituição de uma baseno motivo Palíndrome está indicada em letra minúscula.

B) EMSA das proteínas recombinantes O2CX e BZH1 com 300 fmol da sonda mutada GM marcada com ^{32}P . No primeiro poço foi carregado apenas 300 fmol da sonda GM marcada; no poço controle negativo, foi carregado 100 ng de eluato de bactérias transformadas com o vetor pET28a incubado com 300 fmol da sonda GM; nos poços seguintes, foram carregadas as quantidades das proteínas recombinantes O2CX e BZH1 indicadas na figura, após terem sido incubadas com 300 fmol da sonda GM..

A

Sonda MP

5'-GCTCTAGATGACAAAATTaAtTAGGAGACATGTCATCCTCAGCTTAGAATTCC-3'

B

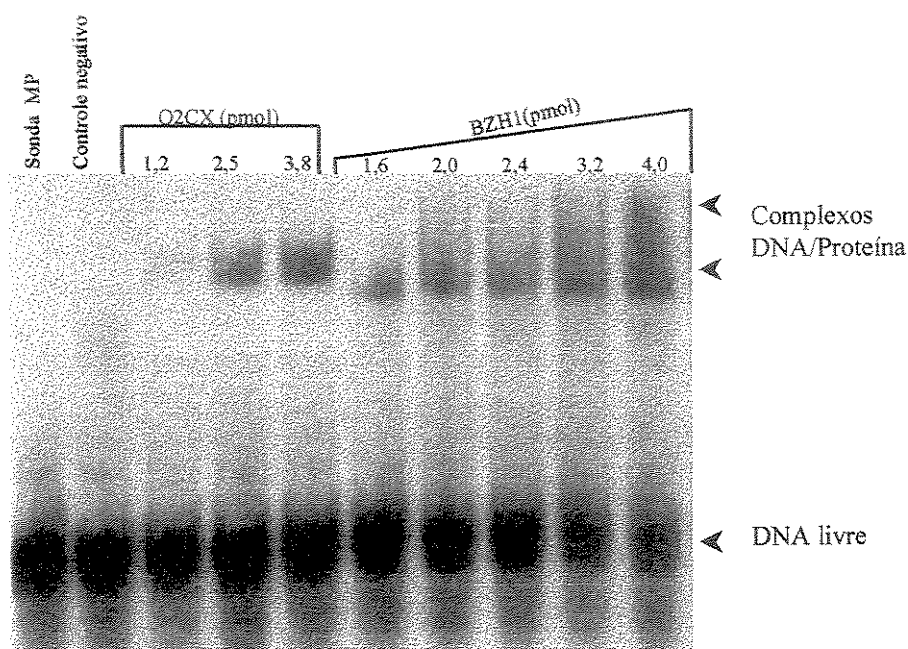


Figura 16. EMSA das proteínas recombinantes O2CX e BZH1 com a sonda mutada MP. A) Sequência do oligonucleotídeo MP. Os motivos GLM e Palíndrome se encontram sublinhados e a substituições de duas bases no motivo GLM se encontram indicadas em letra minúscula.

B) EMSA das proteínas recombinantes O2CX e BZH1 com 300 fmol da sonda mutada MP marcada com ^{32}P . No primeiro poço foi carregado apenas 300 fmol da sonda MP marcada; no poço controle negativo foi carregado 100 ng de eluato de bactérias transformadas com o vetor pET28a; as quantidades das proteínas recombinantes O2CX e BZH1 que foram carregadas nos poços seguintes estão indicadas na figura.

A

Sonda G-box

5'- AAGCTTAGCCACGTGGCACTCGAG-3'

B

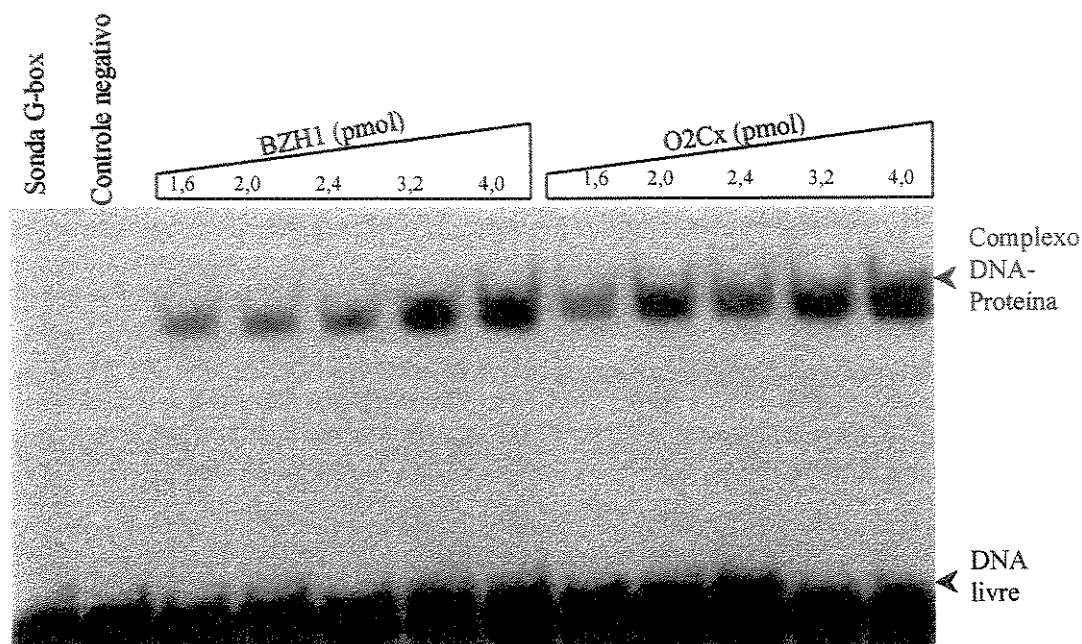


Figura 17. Comparação da ligação das proteínas recombinantes BZH1 e O2CX à sonda G-box.

A) Sequência da sonda G-box, o núcleo ACGT e as três bases adjacentes de cada lado, que formam uma sequência palíndrome do motivo G-box se encontram sublinhados.

B) EMSA das proteínas recombinantes BZH1 e O2CX com a sonda G-box.

O primeiro poço, foi carregado com apenas 800 fmol da sonda G-box marcada.

O segundo poço, foi carregado com 100 ng de eluato de bactérias transformadas com o vetor pET28a incubado com 800 fmol da sonda G-box. Os poços seguintes foram carregados com as quantidades das proteínas recombinantes BZH2 e O2CX indicadas na figura e incubadas com 800 fmol da sonda G-box marcada.

A

Sonda C-box

5'-AAGCTTAGTGACGTCACACTCGAG-3'

B

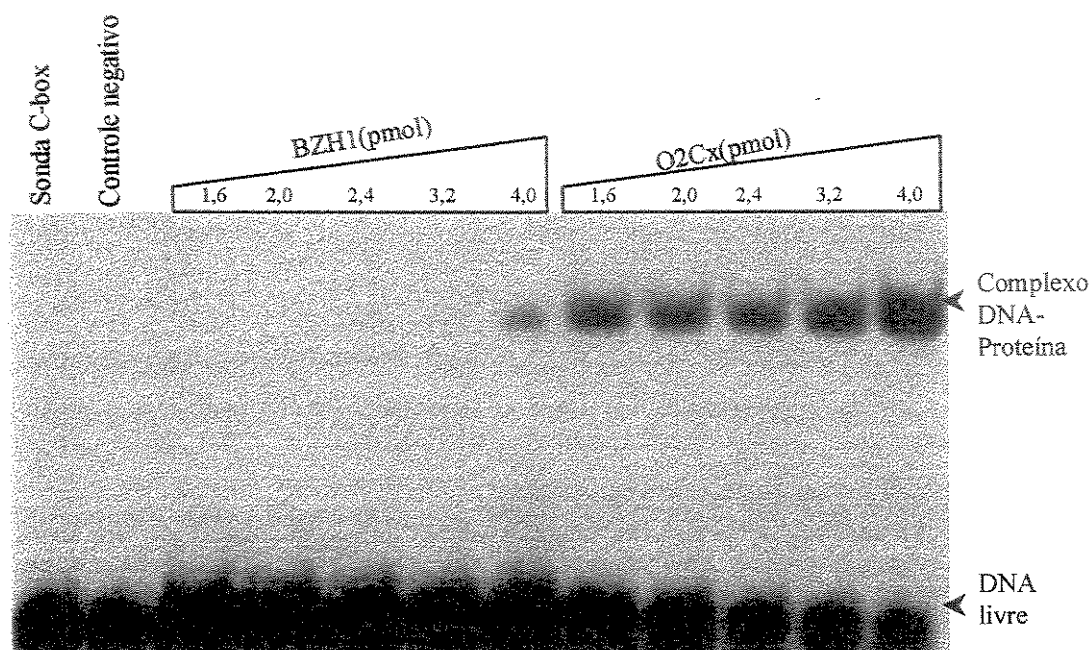


Figura 18. Comparação da ligação das proteínas recombinantes BZH1 e O2CX à sonda C-box. A) Sequência do oligonucleotídeo C-box, o núcleo ACGT e as três bases adjacentes de cada lado, que formam uma sequência palíndrome do motivo C-box se encontram sublinhados. B) EMSA das proteínas recombinantes BZH1 e O2CX com a sonda C-box. O primeiro poço, foi carregado com apenas 800 fmol da sonda C-box marcada. O segundo poço foi carregado com 100 ng de eluato de bactérias transformadas com o vetor pET28a incubado com 800 fmol da sonda C-box. Os poços seguintes foram carregados com as quantidades das proteínas recombinantes BZH2 e O2CX indicadas na figura e incubadas com 800 fmol da sonda C-box marcada.

DISCUSSÃO

A Família Multigênica O2

Um cDNA de *Arabidopsis thaliana* que codifica uma proteína do tipo bZIP homóloga ao fator de transcrição O2 de milho foi isolado e denominado BZO2H1. Nos próximos parágrafos deste trabalho, serão discutidos os resultados da caracterização da estrutura do gene, assim como a análise comparativa da especificidade de ligação da proteína recombinante BZH1.

Devido ao projeto genoma de *Arabidopsis thaliana* que recentemente completou o seqüenciamento deste organismo, foi possível definir a estrutura do gene *BZO2H1*, composto por sete exons e seis íntrons (Figura 5). Também foi observado que os exons quatro e cinco codificam 95% do domínio bZIP cuja posição e tamanho se encontram conservados e definem uma característica específica da Família Multigênica de Opaco2 (Figura 5). Um dos clones do mesmo cDNA de BZO2H1 apresentou o fenômeno do “splicing diferencial” do RNA mensageiro. Este fenômeno já foi documentado para vários mRNAs de genes de fatores de transcrição de plantas (Dietrich *et al.*, 1997; Grotewold *et al.*, 1991). Sua presença pode estar associada a uma função regulatória, por exemplo, o gene *myb7* de arroz, ajusta a razão entre os mRNAs de modo a atuar como uma chave de controle que afeta a expressão de outros genes. (Magaraggia *et al.*, 1997). A expressão do gene *BZO2H1* em vários tecidos da planta sugere que o fator BZO2H1 possa representar uma gama funcional diversificada, interagindo com diferentes promotores tecido-específicos.

Os resultados demonstraram que BZO2H1 pertence à Família Multigênica de O2 (Vincentz *et al.*). A Família Multigênica de *Opaco2* é composta pelos genes de plantas monocotiledôneas *O2*, *OHP1* e *OHP2* de milho, *SPA* de trigo, *RITA1* e *REB* de Arroz, *BZL1* e *BZL2* de cevada, *O2* de Coix e *O2* de Sorgo, enquanto que as dicotiledôneas estão representadas por *HBFI* de soja, *CPRF2* de salsa e *BZO2H1*, *BZO2H2*, *BZO2H3* e *BZO2H4* de *Arabidopsis thaliana*.

Todos os genes de plantas monocotiledôneas da Família Multigênica de *O2* são expressos no endosperma. A expressão desses genes no endosperma parece estar relacionada com o reconhecimento de elementos situados nos promotores de genes de proteínas de reserva da semente. BLZ1 e BLZ2 (*Barley bZIP 1* e *2*) podem se heterodimerizar e se ligam ao

elemento GLM, presente no promotor do gene *Hor-2* que codifica a proteína B1-hordeína, uma prolamina da cevada. Estudos feitos utilizando o sistema de dois híbridos em levedura, demonstraram que estas duas proteínas atuam como ativadores da transcrição (Vicente-Carbajosa *et al.*, 1998, Onâte *et al.*, 1999).

A proteína REB (Rice Endosperm bZIP) se liga a uma região do promotor de α -globulina que é homóloga a dos promotores dos genes de *Glutenina* e *D-hordeína* em cevada (Nakase *et al.*, 1996, Nakase *et al.*, 1997). RITA1 (Rice Transcription Activator 1) foi caracterizada por ser altamente expressa em endosperma de arroz durante o desenvolvimento da semente. Esta proteína apresenta ampla especificidade de ligação e alta afinidade para A-box, C-box e G-box (Isawa *et al.*, 1994). SPA (Storage Protein Activator) de trigo, assim como BLZ1 e O2, reconhece o motivo GLM presente no promotor dos genes de prolaminas (Albani *et al.*, 1997).

SPA é um forte ativador transcricional e apresenta alta similaridade de sequência com as proteínas OHP1 e OHP2 de milho (Auckerman & Schmidt, 1994, Albani *et al.*, 1997). OHP1 e OHP2 foram isolados de uma biblioteca de cDNA de endosperma de milho e reconhecem os elementos *cis* contidos no promotor das zeínas de 22kDa. OHP1 também se heterodimeriza com O2, porém a formação dos heterodímeros parece não afetar a afinidade de ligação ao DNA (Pysh *et al.*, 1993).

O2 de milho é um regulador da síntese das prolaminas durante o desenvolvimento da semente do milho. Análises genéticas e moleculares indicam que O2 parece ser importante para integrar o metabolismo do carbono e nitrogênio com a síntese das proteínas de reserva α - e β -zeínas (Varagona *et al.*, 1993, Schmidt *et al.*, 1993; Damerval & Le Guilloux, 1998; Cord Neto *et al.*, 1998). Fatores muito semelhantes a O2 de milho também foram isolados em *Coix* e sorgo. Além da sequência altamente conservada, O2 de *Coix* e O2 de sorgo, ainda reagem com os mesmos elementos presentes nos promotores das prolaminas de cada espécie (Vettore *et al.*, 1998; Yunes *et al.*, 1998; Pirovano *et al.*, 1994). Todas essas evidências aliadas ao fato que as três espécies de plantas pertencem a sub-família *Panicoideae* nos leva a crer que O2 de *Coix* e O2 de sorgo são ortólogos a O2 de milho.

Assim, pela análise do padrão de expressão dos fatores bZIP de plantas monocotiledôneas da Família Multigênica O2 e dos promotores de genes aos quais se ligam, a função deste grupo, assim como O2 de milho, parece estar relacionada com a regulação dos

genes expressos durante o desenvolvimento da semente. Porém, esta não deve ser a única função desempenhada por esse grupo de proteínas. A expressão de RITA-1, OHP-1 e BLZ1 em outros tecidos além do endosperma, sugere que estes fatores possam desempenhar funções diferentes da regulação dos genes expressos na semente, através de interações com promotores de outros genes dos tecidos onde são expressos.

Os fatores bZIP das plantas dicotiledôneas da Família Multigênica O2 ainda não foram suficientemente caracterizados para ter sua função conhecida. Estudos feitos com CPRF2 (Common Plant Regulatory Factor 2) de salsa demonstraram que essa proteína é translocada do citosol para o núcleo após incidência de luz no fitocromo, sugerindo seu envolvimento na transdução de sinais da luz envolvendo o fitocromo (Kircher *et al.*, 1999). Além disso, CPRF2 se liga a um elemento cis G-box presente no promotor do gene da Chalcona sintetase que é expresso em resposta à incidência de luz (Weisshaar *et al.*, 1991; Schulze-Lefert *et al.*, 1989; Fronmeyer *et al.*, 1992). CPRF2 é regulado pós-traducionalmente por fosforilação, em presença de luz, tanto *in vitro* como *in vivo* (Wellmer *et al.*, 1999). Porém, a fosforilação parece não afetar sua atividade de ligação ao DNA, mas pode estar envolvida com a regulação do transporte de CPRF2 para o núcleo da célula ou na modulação de transativação gênica devido a este fator (Wellmer *et al.*, 1999; Jans & Hübner, 1996; Karin, 1994; Faux & Scott, 1996).

HBF1 foi isolado de uma biblioteca de soja e caracterizado como mais membro da família Multigênica O2. HBF1 é fosforilado durante a indução do sistema de defesa da planta (Dröge-Laser *et al.*, 1997). Por ensaios de retardamento da mobilidade eletroforética, foi observado que sua fosforilação *in vitro* aumenta a ligação de HBF1 aos elementos cis presente no promotor do gene *chs15*, cuja proteína cataliza o primeiro passo da síntese de fitoalexinas derivadas de isoflavonóides (Lois *et al.*, 1989; Ohl *et al.*, 1990).

As proteínas de Arabidopsis BZO2H1, BZO2H2, BZO2H3 e BZO2H4 que estão sendo estudadas por nosso grupo, são expressas em todos os tecidos, com acúmulo de transcritos, preferencialmente em raízes, caule e flores. BZO2H1 é a primeira destas quatro proteínas a ser mais abragentemente caracterizada neste trabalho.

Pela análise da estrutura modular protéica dos fatores de transcrição da Família Multigênica O2 (Vincentz *et al.*), alguns motivos que podem definir nichos funcionais das proteínas foram identificados. O domínio bZIP, além de servir como ligante de DNA é o

elemento estrutural responsável pela dimerização, além disso apresenta um sinal de localização nuclear (NLS). Este NLS é um motivo duplo composto por uma região situada N-terminalmente ao motivo básico, o NLS-A, cuja sequência conservada é composta pelo motivo DYLLR e uma região bastante conservada na família multigênica O2, o NLS-B, que está situado dentro do motivo básico (Figura 8). Experimentos feitos com O2 de milho demonstraram que tanto NLS-A como NLS-B são importantes para a importação nuclear da proteína e que no domínio básico, as funções de ligação ao DNA e localização da proteína são independentes umas das outras (Varagona & Raikhel, 1994).

A aproximadamente 11 aminoácidos N-terminal do domínio básico, encontramos o motivo TSA (Figura 8). Esse motivo consiste de um agrupamento de serinas e treoninas seguidas por aminoácidos ácidos. Provavelmente, o motivo TSA contribua para a ativação transcricional da proteína bZIP. Em experimentos feitos usando o sistema de duplo-híbridos de levedura para testar diversas construções de BLZ1 de cevada, foi demonstrado que o motivo TSA contribui para a ativação de transcrição através de interações com outras proteínas (Vicente-Carbajosa *et al.*, 1998). Teoricamente uma outra contribuição desse domínio para a função da proteína poderia ser sua regulação através de fosforilação. Os resíduos de serina e treonina presentes no motivo poderiam servir como substrato de quinases e fosfatases. No entanto, ainda não foi relatado nenhum estudo que tratasse especificamente da fosforilação desse domínio e seus efeitos na proteína.

O domínio localizado N-terminalmente a TSA é caracterizado pela sequência de resíduos DYLLR e se encontra presente na família multigênica O2, com exceção de RITA1 e BZO2H2. O NLS-A de O2 se encontra dentro desse motivo. Estudos feitos com proteínas CPRF2 truncadas demonstraram que a deleção do motivo DYLLR impede a translocação regulada por luz de CPRF2 para o núcleo (Kircher *et al.*, 1999).

Outros motivos encontrados na porção N-terminal são os motivos VFEW e EFWF. O motivo VFEW está ausente nas proteínas HBF1, BZO2H2 e RITA1 e o motivo EFWF é pouco conservado entre essas proteínas. A combinação de resíduos hidrofóbicos com ácidos parece ter importância para a ativação da transcrição, como foi demonstrado para o fator VP1 do vírus *Herpes simplex* (Cress & Triezenber, 1991). Outra indicação da importância do motivo EWFEE é sua presença na parte N-terminal do domínio de ativação de transcrição de O2 de milho (Schmidt, 1997).

Com as informações descritas acima, mais os estudos filogenéticos feitos para caracterizar a Família Multigênica de O2 (Vincentz *et al.*), foi concluído que o gene *BZO2H1* de *Arabidopsis thaliana* provavelmente não seja ortólogo de *O2* de *Zea mays*, pois a presença do locus *O2* parece ser restrita a espécies monocotiledôneas. Assim, a função de *BZO2H1* ainda precisa ser definida por uma caracterização mais aprofundada. Essa caracterização inclui métodos de genética reversa como a superexpressão e o silenciamento do gene *in vivo*, além da determinação de seu padrão de expressão *in vivo* com auxílio de marcadores. Outra maneira de abordar a questão funcional é pela caracterização dos elementos *cis* reconhecidos por essa proteína *in vitro*. Com o conhecimento das seqüências de DNA que os fatores de transcrição se ligam, provavelmente, no futuro, será possível prever os genes com que esses fatores interagem. Contudo, isso só poderá ser realizado pelo acúmulo de informações sobre elementos regulatórios presentes nos promotores dos genes, tarefa que já começou a ser desempenhada por bancos de dados como TRANSFAC (www.transfac.gbf.de/TRANSFAC/), TRRD (www.bionet.nsc.ru/mgs/dbases/trrd4/), PlantCARE (www.sphinx.reg.ac.be:800/transferm/homepage.html) e PLACE (www.dna.affcr.go.jp/htdocs/PLACE).

A Especificidade de ligação ao DNA de BZH1

Considerando que o reconhecimento de determinados motivos de DNA por fatores de transcrição pode nos fornecer dados sobre o papel regulatório desempenhado por esses fatores, uma forma truncada da proteína *BZO2H1* foi expressa como proteína recombinante em *E. coli*. Essa proteína truncada englobou o domínio bZIP e foi denominada *BZH1* (Figura 9). A expressão e purificação da proteína recombinante *BZH1* foi necessária para possibilitar o estudo de alguns motivos de ligação ao DNA reconhecidos por ela.

Pelo alinhamento das proteínas da Família Multigênica *O2* foi observada a alta similaridade existente entre *BZO2H1* e *OPACO2* (Figura 8). Essa similaridade na estrutura primária sugeriu que essas proteínas fossem apresentar as mesmas características de ligação ao DNA. Assim, experimentos de EMSA foram realizados para definir alguns motivos de DNA que pudessem ser reconhecidos pela proteína *BZH1*.

Tendo em vista uma abordagem comparativa, em todos os experimentos, a proteína recombinante *O2* de *Coix*, que é ortóloga a *O2* de milho, foi utilizada como controle. Foi demonstrado que *O2* de *Coix* apresenta, além da alta conservação do domínio bZIP, a mesma

especificidade de ligação ao DNA de O2 de milho para alguns motivos que já foram bem caracterizados. Estudos feitos com uma região do promotor da α -Coixina denominada O2-box, revelaram a existência de sítios de DNA relacionados ao pseudopalíndrome AP-1 (5'-ATGACTCAT-3') que é reconhecido pelo fator GCN4 (Hill *et al.*, 1986). Os sítios que fazem parte da O2-box são o GCN4-Like Motiv ou GLM (5'TGACTA3') e a sequência palíndrome (5'GACATGTC3'). Estes dois motivos são reconhecidos tanto por O2 de milho, como O2 de Coix. Além disso, o motivo GLM também já foi encontrado nos promotores de outros genes de proteínas de reserva e é reconhecido pelos fatores BLZ1 e BLZ2 de cevada e SPA de trigo (Yunes *et al.*, 1994 e 1998; Vicente-Carbajosa *et al.*, 1998; Onãte *et al.*, 1999; Albani *et al.*, 1997). Em outro estudo de caracterização de proteínas bZIP por sua especificidade de ligação, foi demonstrado que o fator O2 de milho reconhece tanto o motivo G-box (5'-NNCACGTGNN-3') como C-box (5'-NNGACGTCNN-3') com afinidades de ligação comparáveis (Isawa *et al.*, 1993).

Baseados nestas informações, o primeiro experimento de EMSA realizado foi para verificar se a proteína BZH1 reconheceria a sonda GP que continha os motivos GLM e Palíndrome do O2-box. Os resultados demonstraram a formação de dois complexos, I (Inferior) e S (Superior) com a proteína recombinante BZH1 (Figura 13). À medida que a concentração da proteína recombinante BZH1 foi aumentada, foi observada uma transição do complexo I para o complexo S. A ligação de O2CX com a mesma sonda também resulta na formação de dois complexos DNA/Proteína (Yunes *et al.*, 1998). Entretanto, para uma mesma concentração protéica foi observado que a intensidade do complexo superior foi menor do que a do complexo inferior com a proteína O2CX em relação a BZH1, indicando uma diferença das características de ligação dessas duas proteínas com a sonda GP. A hipótese formulada para explicar tal resultado seria que a tendência de formação do complexo superior da proteína recombinante BZH1 com a sonda GP poderia ocorrer por um aumento da cooperatividade de ligação ao DNA. Isso significa que quando a proteína recombinante BZH1 se ligasse, sob forma dimerizada, a um dos motivos de DNA da sonda GP, promoveria a ligação de outro dímero no sítio adjacente. Seguindo esse raciocínio, também seria possível afirmar que a proteína BZH1 apresenta maior cooperatividade de ligação à sonda GP do que a proteína O2CX, pois o deslocamento do complexo inferior para o complexo superior ocorre a

uma concentração menor da proteína BZH1 do que com a proteína O2CX (Resultado não mostrado).

Os ensaios de competição da proteína recombinante BZH1 com a sonda GP indicaram que a proteína recombinante BZH1 se liga de maneira específica à sonda GP (Figura 15). Experimentos feitos com a proteína controle O2CX também demonstraram que a ligação da proteína recombinante O2CX à sonda GP era específica (Yunes *et al.*, 1998).

Na tentativa de confirmar a ligação da proteína BZH1 aos motivos GLM e Palíndrome presentes na sonda GP, foram realizados experimentos de EMSA utilizando sondas mutadas nesses motivos: a sonda MP, que apresenta duas substituições no sítio GLM e a sonda GM, que apresenta uma substituição no sítio Palíndrome. As substituições feitas em cada motivo suprimiram a ligação da proteína O2CX aos mesmos (Yunes *et al.*, 1998). Os dados da literatura foram confirmados nos experimentos das figuras 15 e 16, em que a proteína recombinante O2CX formou somente um complexo DNA/Proteína com o motivo não mutado de cada sonda. A proteína recombinante BZH1 também apresentou um complexo DNA/Proteína com as sondas GM e MP quando utilizada na concentração de 1,6 pmol (Figuras 15 e 16). Nos mesmos experimentos, foi observado que o aumento na concentração da proteína recombinante BZH1 provocou a formação de um segundo complexo DNA/Proteína instável, em forma de rastro. Provavelmente, a formação desse complexo instável ocorreu porque as mutações nos motivos GLM e Palíndrome não suprimiram a ligação da proteína recombinante BZH1 às sondas mutadas. As mutações podem ter apenas diminuído a afinidade da proteína recombinante BZH1 pelas sondas GM e MP, levando à formação do complexo instável observado.

Experimentos de EMSA utilizando as sondas G- e C-box que continham os respectivos motivos em suas seqüências foram realizados com o intuito de caracterizar a especificidade de ligação da proteína recombinante BZH1. Os resultados demonstraram que a proteína recombinante BZH1 reconhece o motivo G-box com afinidade relativa comparável à proteína recombinante O2CX, porém, o reconhecimento do motivo C-box por BZH1 ocorreu com afinidade menor do que com O2CX (Figuras 17 e 18).

Os experimentos envolvendo as sondas GM, MP, G-box e C-box demonstram que as proteínas recombinantes BZH1 e O2CX apresentam características de ligação diferentes.

Niu e colaboradores estudaram as regiões importantes para a especificidade de ligação ao DNA dos fatores bZIP EmBP-1 e TGA1a e concluíram que além da região básica, a região chamada “hinge” (= região dobradiça), situada entre os aminoácidos -5 e -1 em relação à primeira leucina do zíper, parece ter importância para a especificidade de ligação ao DNA. (Niu *et al.*, 1999; Isawa *et al.*, 1993). Com esses dados, as proteínas recombinantes BZH1 e O2CX foram alinhadas e analisadas (Figura 12). Foi observado que a região básica está bastante conservada entre as duas proteínas, no entanto, os cinco aminoácidos da região “hinge” são completamente diferentes. A região “hinge” da proteína recombinante BZH1 apresenta dois aminoácidos de caráter ácido (E, D) e três aminoácidos polares (Q, T, S), sendo que dois desses resíduos são hidroxilados (T, S). A região “hinge” da proteína O2CX apresenta dois aminoácidos de caráter básico (H, K) e dois de cadeia lateral apolar (A, L). É possível que a diferença na composição de aminoácidos dessa região esteja envolvida com a alta cooperatividade observada na ligação da proteína BZH1 e a sonda GP e pela menor afinidade de ligação da proteína BZH1 à sonda C-box observada em relação à proteína O2CX. Contudo, a possibilidade de que a diferença entre as regiões N- e C-terminal da proteína O2CX e BZH1, ou mesmo seu zíper de leucina, tenham algum efeito sobre a ligação ao DNA não poderia ser descartada (Figura 12). No entanto, somente experimentos de trocas de resíduos da região “hinge” e do zíper de leucina entre as proteínas O2CX e BZO2H1 poderiam definir quais são os resíduos responsáveis pelas diferenças de ligação.

PERSPECTIVAS

Para entender a função de fatores de transcrição do tipo bZIP, é possível analisar a questão por diferentes ângulos. Entre eles, o estudo da expressão do fator bZIP sob diferentes condições fisiológicas e ambientais, a caracterização da proteína bZIP e suas modificações sofridas durante a regulação e a identificação dos genes por ela ativados.

A caracterização de BZO2H1 teve por objetivo estudar um fator de transcrição, que por sua homologia com O2, é potencialmente importante e do qual não havia nenhuma informação disponível, além da sequência genômica. Neste estudo, foi enfocada principalmente a expressão e a especificidade de ligação de BZO2H1, assim como a comparação dos dados obtidos com os de outras proteínas da mesma família. Os próximos passos que serão levados adiante pelo grupo do Dr. Michel Vincentz, será a análise da expressão das proteínas BZO2H1-H4 sob diferentes condições de “stress” e cultivo de plantas em meios que apresentem variações na concentração de hormônios, como o ácido abscísico. Também serão produzidas plantas transgênicas que superexpressem ou que tenham a expressão destes fatores bZIP bloqueada para a análise das funções fisiológicas destas proteínas *in vivo*.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abel, T. & Maniatis, T. (1989). Action of leucine zippers. *Nature*, 341:24-25.
- Albani, D., Hammond-Kosack, M.C.U., Smith, C., Conlan, S., Colot, V., Holdsworth, M., and Bevan, M.W. (1997). The Wheat Transcriptional Activator SPA: A seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *Plant Cell*, 9:171-184.
- Alber, T. (1992). Structure of the Leucine Zipper. *Curr. Opin. Genet. Dev.*, 2 : 205-210.
- Aukerman, M.J., and Schmidt, R.J. (1994). Regulation of α -zein expression during maize endosperm development. In *Plant Promoters and Transcription Factors*, L. Nover, ed (Heidelberg, Germany: Springer-Verlag), pp. 2029-233.
- Ausubel, F. M. (1997). *Current Protocols en Molecular Biology V.1* John Wiley & Sons, Inc USA
- Bailey, T.L. & Elkan., C. (1994) The value of prior knowledge in discovering motifs with MEME. *Proc. Int. Conf. Intell.Syst.Mol.Biol.*, 3:21-29.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilising the principle of protein-dye binding. *Anal. Biochem.* , 72: 248-254.
- Brochetto-Braga, M.R., Leite, A., Arruda, P. (1992). Partial purification and characterization of lysine-ketoglutarate reductase in normal and Opaque2 maize endosperm. *Plant Physiol.*, 98: 1139-1147.
- Buratowski, S. (1994). The basics of basal transcription by RNA polymerase II. *Cell*, 77:1-3.
- Carey, J. (1991). Gel Retardation. *Methods in Enzimology*, 208: 103-117.
- Ciceri, P., Gianazza, E., Lazzari, B., Lippoli, G., Genga, A., Hoschek, G., Schmidt, R. J., and Viotti, A. (1997). Phosphorilation of Opaque2 changes diurnally and impacts its DNA binding activity. *Plant Cell*, 9: 97-108.
- Ciceri, P., Locatelli, F., Genga, A., Viotti, A., and Schmidt R. J. (1999). The activity of the mayze Opaque2 transcriptional activator is regulated diurnally. *Plant Phys.*, 121: 1321-1327.
- Chern, M-S., Bobb, A.J. and Bustos, M.M. (1996a). The regulator of MAT2 (ROM2) protein binds to early maturation promoters and represses PvALF-activated transcription. *Plant Cell*, 8: 305-321.
- Chern, M-S., Eiben, H.G. and Bustos, M.M. (1996b). The developmentally regulated bZIP factor ROM1 modulates transcription from lectin and storage protein genes in bean embryos. *Plant J.*, 10(1): 135-148.
- Choi, H., Hong, J., Ha, J., Kang, J. and Kim. S. Y. (2000). ABFs, a family of ABA-responsive element binding factors. *J. Biol. Chem.*, 275(3): 1723-1730.

- Chuang, C.F., Running, M.P., Williams, R.W. and Meyerowitz, E. (1999). The *Perianthia* gene encodes a bZIP protein involved in the determination of floral organ number in *Arabidopsis thaliana*. *Genes and Development*, 13:334-344.
- Cord Neto, G., Yunesm J.A., da Silva, M.J., Vettore, A.L., Arruda, P., and Leite, A. (1995). The involvement of Opaque2 in β -prolamin gene regulation in maize and *Coix* suggests a more general role for this transcriptional activator. *Plant Mol. Biol*, 27: 1015-1029.
- Cress A. & Trienzenberg, S.J. (1991). Nucleotide and deduced amino acid sequences of the gene encoding virion protein 16 of herpes simplex virus type 2. *Gene*, 103(2):235-238.
- Darnel, J.E. (1982). Variety in the level of gene control in eucaryotic cells. *Nature*, 297:365-371.
- Damerval, C. & Le Guilloux, M. (1998). Characterisation of novel proteins affected by the *O2* mutation and expressed during maize endosperm development. *Mol. Gen. Genet.*, 257-361.
- Dean C., Van der Elzen P., Tanaki S., Dunsmuir P & Bedbrook J. (1985). Differential Expression of eight genes of petunia ribulose biphosphate carboxylase small subunit multigene family. *Embo J*, 4: 3055-3061
- Derman, E. (1981). Transcription control in the production of liver specific mRNA's. *Cell*, 23:731-739.
- Dietrich, R.A., Richberg, M.H., Schmidt, R., Dean, C., and Dang, J.L. (1997). A novel zinc finger protein is encoded by the arabidopsis *LSD1* gene and functions as a negative regulator of plant cell death. *Cell*, 88: 685-694.
- Dröge-Laser W., Kaiser, A., Lindsay, P.W., Halkier, B.A., Loake, G.J., Doerner, P., Dixon, R.A., and Lamb, C. (1997). Rapid stimulation of a soybean protein-serine kinase that phosphorylates a novel bZIP DNA-binding protein, G/HBF-1, during the induction of early transcription-dependent defenses. *EMBO J.*, 16: 726-738.
- Ellenberger, T. E., Brandl, C. J., Struhl, K., and Harrison, S. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: cristal structure of the protein-DNA complex. *Cell*, 71: 1223-1237.
- Faux, M.C. & Scott, J. D. (1996). More on target with protein phosphorilation: conferring specificity by location. *Trends Biochem. Sci*, 21: 312-315.
- Finkelstein R.R. & Lynch, T.J. (2000). The Arabidopsis abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *Plant Cell*, 12(4): 599-609.
- Foster, R., Izawa, T., and Chua N.-H. (1994). Plant bZIP proteins gather at ACGT elements. *FASEB J.*, 8: 192-200.
- Fronhnemeyer, H., Ehmann, B., Kretsch, T., Rocholl, M., Harter, K., Nagatani, A., Furuya, M.,

- Batschauer, A., Hahlbrock, K., and Schäfer, E. (1992). Differential usage of photoreceptors for chalcon synthase gene expression during plant development. *Plant J.*, 2:899-906.
- Fukasawa, J., Sakai, T., Ishida, S., Yamaguchi, I., Kamiya, Y., and Takahashi, Y. (2000). REPRESSION OF SHOOT GROWTH, a bZIP transcriptional activator, regulates cell elongation by controlling the level of gibberellins. *Plant Cell*, 12: 901-915.
- Gallusci, P., Varrot, S., Matsuoka, M., Maddalini, M. and Thompson, R.D. (1996). Regulation of cytosolic pyruvate, orthophosphate dikinase expression in developing maize endosperm. *Plant Mol. Biol.*, 31:45-55.
- Giroux M. J., Boyer, C., Fleix, G., and Hannah, C. (1994). Coordinated transcriptional regulation of storage product genes in the maize endosperm. *Plant physiol.*, 106:713-722.
- Goeddel, D.G. (1991). *Gene Expression Technology - Methods in Enzymology*, v. 185 Academic press, San Diego USA
- Gonnet, G.H., Cohen, M.A., and Benner, S.A. (1992). Exhaustive matching of the entire protein sequence database. *Science* 256 (5062): 1443-1445
- Graur, D. & Li, W-H. (1999). *Fundamentals of molecular Evolution* 2nd ed. Chapter 6 pp.250-322. Sinauer Associates, Sunderland, Massachusetts.
- Grotewold, E., Athma, P., and Peterson, T. (1991). Alternatively spliced products of the maize P gene Encode proteins with homology to the DNA-binding domains of Myb-like transcription factors. *Proc. Natl. Acad. Sci. USA*, 88: 4587-4591.
- Harter, K., Kircher, S., Frohnmeyer, H., Krenz, M., Nagy, F., and Schäfer, E. (1994). Light-regulated modification and nuclear translocation of cytoplasmic G-box binding factors (GBFs) in parsley (*Petroselinum crispum* L.). *Plant Cell*, 6: 545-559.
- Hicks, G. R. & Raikhel, N.V. (1995). Nuclear Localization signal binding proteins in higher plant nuclei. *Proc. Natl. Acad. Sci. USA*, 92: 734-738.
- Hill, D.E., Hope, I.A., Macke, J.P. & Struhl, K. (1986). Saturation Mutagenesis of yeast his3 regulatory site: requirements of transcriptional induction and for binding by GCN4 activator protein. *Science*, 234: 451-457.
- Hurst, H.C. (1995). *Leucine Zippers Transcription Factors*. Academic Press, San Diego, USA.
- Isawa, T., Foster, R., Chua, N.H. (1993). Plant bZIP protein DNA binding specificity. *J.Mol.Biol.* 230: 1131-1144.
- Isawa, T., Foster, R., Nakajima, M., Shimamoto, K., and Chua, Nam-Hai. (1994). The Rice bZIP Transcriptional Activator RITA-1 Is Highly Expressed during Seed Development. *The Plant Cell*, 6:1277-1287.
- Jans, D.A. & Hübner, S. (1996). Regulation of protein transport to the nucleus: central role of

- phosphorilation. *Physiol. Rev.*, 76: 651-685.
- Karin, M. (1994). Signal transduction from the cell surface to the nucleus through the phosphorylation of transcription factors. *Curr. Opin. Biol.*, 6 : 415-424.
- Kemper, E.L., Cord Neto, G., Papes, F., Moraes, K.C.M., Leite, A. & Arruda, P. (1999). the role of Opaque2 in the control of lysine degrading activities in developing maize endosperm. *Plant Cell*, 11: 1981-1993.
- Kircher, S., Wellmer, f., Nick, P., Rügner, A., Schäfer, E., and Harter, K. (1999). Nuclear Import of the Parsley bZIP Transcription Factor CPRF2 is regulated by Phytochrome Photoreceptors *J Cell Biol.* 144(2): 201-211.
- Kornberg, R.D. (1999). Eukaryotic transcriptional control. *Trend Cell Biol.*, 12:46-49.
- Lamb, P., and McNight, S.L. (1991). Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization. *Trends Biochem. Sci.*, 16: 417-422.
- Landschulz, W.H., Johnson, P.F.,McKnight, S.L. (1988). The Leucine Zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science*, 240 (4860):1759-1764.
- Lee, T.I. & Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu Rev.Genet.*, 34:77-137.
- Leuther K. K., et al. (1996). Two dimensional crystallography of transcription factor IIB- and IIE- RNA polymerase complexes: implications for start site selection and initiation complex formation. *Cell*, 85: 773-779.
- Li, Y et al. (1994). RNA Polymerase II initiation factor interactions and transcription start site selection. *Science*, 263: 805-807.
- Liu, L., White, M. J., and MacRae, T.H. (1999). Transcription factors and their genes in higher plants, functional domains, evolution and regulation. *Eur. J. Biochem.*, 262:247-257.
- Lohmer, S. Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F., & Thompson, R. D. (1991). The maize regulatory locus Opaque2 encodes a DNA-binding protein which activates the transcription of the b-32 gene. *EMBO J.*, 10: 617-624.
- Lois, R., Dietrich, A., Hahlbrock, K , and Schulz, W. (1989). A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light-responsive cis-acting elements. *EMBO J.*, 8:1641-1648.
- Magaraggia, F. Solinas, G., Valle, G., Giovinazzo, G., and Coraggio, I. (1997). Maturation and translation mechanisms involved in the expression of a myb gene of rice. *Plant Mol. Biol.*, 35: 1003-1008.
- Macfarlane, W.M. (2000). Demystified...Transcription. *J. Clin. Pathol:Mol. Pathol*, 53-1-7.

- Nakayama, T., Okanami, M., Meshi, T and Iwabuchi, M. (1997). Dissection of the wheat transcription factor HBP-1a(17) reveals a modular structure for the activation domain. *Mol. Gen. Genet.*, 253:553-561.
- Meshi T. & Iwabuchi M. (1995). Plant Transcription Factors. *Plant Cell Physiol*, 36(8):1405-20.
- Meshi, T., Moda, I., Minami, M., Okanami, M., and Iwabuchi, M. (1998). Conserved Ser residues in the basic region of the bZIP-type transcription factor HBP1a(17): importance in DNA binding and possible targets for phosphorylation. *Plant. Mol. Biol.*, 36:125-136.
- Mitchell, P.J. & Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA-binding proteins. *Science*, 245:371-378.
- Nakase, M., Hotta, H., Adachi, T., Aoki, N., Nakamura, R., Masumura, T., Tanaka, K., e Matsuda, T. (1996). Cloning of the rice seed α -globulin-encoding gene: sequence similarity of the 5' flanking region to those of the genes encoding wheat high-molecular weight glutenin and barley D hordein. *Gene* 170: 223-226
- Nakase, M., Aoki, N., Matsuda, T., and Adachi, T. (1997). Characterization of a novel rice bZIP protein which binds to the α -globulin promoter. *Plant Molecular Biology*, 33:513-522.
- Niggeweg, R., Thurow C., Kegler, C., and 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. Biol. Chem.*, 275:19897-19905.
- Niu, X. & Guiltinan, M.J. (1994). DNA binding specificity of the wheat bZIP protein EmBP-1. *Nucl. Acid Res.*, 22 (23): 4969-4978.
- Niu, X., Renshaw-Gegg, L., Miller, L., and Guiltinan, M.J. (1999). Bipartite determinantes of DNA-binding specificity of plant basic leucine zipper proteins. *Plant Mol. Biol.*, 41: 1-13.
- Ohl, S., Hedrick, S. A., Chory, J. and Lamb, C. J. (1990). Functional properties of a phenylalanine ammonia-lyase promoter from *Arabidopsis*. *Plant Cell*, 2:837-848.
- Onãte, L., Vicente-Carbajosa, J., Lara, P., Díaz, I., and Carbonero, P. (1999). Barley BLZ2, a Seed-specific bZIP protein that interacts with BLZ1 in vivo and activates transcription from the GCN4-like motif of *B-hordein* Promoters in *Barley* endosperm. *J. Biol. Chem.*, 274 (14): 9175-9182.
- Osterlund, M.T., Wei, N., and Deng, X.W. (2000). The roles of photoreceptor system and the COP1-targeted destabilization of HY5 in light control of *Arabidopsis* seedling development. *Plant Physiol.*, 124: 1520-1524.
- O' Shea, E.K., Klemm, J.D., Kim, P.S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two stranded, parallel coiled coil. *Science*, 254, 539-544.
- Oyama, T. ET AL. (1997). The *Arabidopsis* *HY5* gene encodes a bZIP protein that regulates stimulus

- induced development of root and hypocotyl. *Gene Dev.*, 11:2983-2995.
- Ottoboni, L.M.M., Leite, A., Yunes, J.A. Targon, M.L.N.P., Souza Filho, G.A., and Arruda, P. (1993). Sequence analysis of 22 kD-like *Coixin* genes and their comparison with homologous zein and kaffirin genes reveals highly conserved protein structure and regulatory elements. *Plant Mol. Biol.*, 21:765-778.
- Pirovano, L., Lanzini, S., Hartings, H., Lazzaroni, M. Rossi, V., Joshi, R., Thompson, R.D. Salamini, F., and Motto, M. (1994). Structural and functional analysis of an Opaque-2 related gene from sorghum. *Plant Mol. Biol.*, 24:515-523.
- Pysh, L.D., Aukerman, M.J., and Schmidt, R. J. (1993). OHP1: a maize basic domain/leucin zipper protein that interacts with Opaque2. *Plant Cell*, 5: 227-236.
- Rook F., Weisbeck P., Smeeckens 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.
- Sainz, M.B., Goff. S.A., and Chandler, V.L. (1997). Extensive mutagenesis of a transcriptional activation domain identifies single hydrophobic and acidic amino acids for activation *in vivo*. *Mol. Cell. Biol.*, 17:115-122.
- Sambrook J., Fritsch E.F. and Maniatis T. (1989). *Molecular Cloning: a laboratory manual*. Cold Spring Laboratory Press, Cold Spring Harbor, NY.
- Schindler, U., Terzagui, W., Beckmann, H., Kadesch, t., and Cashmore, A. R. (1992). DNA binding site preferences and transcriptional activation properties of the Arabidopsis transcription factor GBF1. *EMBO J.*, 11 (4): 1275-1289.
- Singh, K.B. (1998). Transcriptional Regulation in plants: the importance of combinatorial control. *Plant Physiol.*, 118: 1111-1120.
- Schmidt, R.J. (1993). Opaque-2 and zein gene expression. In *control of plant gene expression*. D.P.S. Verma, Boca Raton: CRC Press. pp. 337-355.
- Schmidt, R.J.; Ketudat, M., Aukerman, M.J. and Hoschek, M.J. (1992). Opaque-2 is a transcriptional activator that recognizes a specific target site in 22KDa Zein genes. *Plant Cell*, 4: 689-700.
- Schulze-Lefert, P., Dangl, J.L., Becker André, M., Hahlbrook, K., and Schulz, W., (1989). Inducible *in vivo* footprints defence sequences necessary for UV light activation of the parsley chalcone synthase gene. *EMBO J.*, 8: 651-656.
- Schwechheimer, C., Smith, C., and Bevan, M.W. (1998). The activities of acidic and glutamine-rich transcriptional activation domains in plant cells: design of modular transcription factors for high-level expression. *Plant Mol. Biol.*, 36:195-204.
- Sodek, L. & Wilson, C.M. (1970). Incorporation of leucine-C14 and lysine -C14 into protein in the

- developing endosperm of normal and opaque2 corn. Arch. Biochem. Biophys. 140:29-38.
- Suckow, M., von Wilcken-Bergmann, B., and Mueller-Hill, B. (1993). Identification of three residues in the basic regions of the bZIP proteins GCN4, C/EBP and TAF-1 that are involved in specific DNA binding. EMBO J., 12(3):1193-1200.
- The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature, 408:796-815
- Thompson, J.D., Gibson, T.J., Plewiank, F., Jeanmougin, F. and Higgins, D.G. (1997). The CLUSTAL X windows interface: flexible strategies for multiple alignment aided by quality analysis tools. Nucl. Acids. Res., 25: 4876-4882.
- Uno Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000). Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. PNAS USA, 97:11632-11637.
- Varagona, M.J., Schmidt, R.J., and Raikhel, N.V. (1991). Monocot regulatory protein Opaque-2 is localized in the nucleus of maize endosperm and transformed tobacco plants. Plant Cell, 3: 105-113.
- Varagona, M.J. & Raikhel, N. V. (1994). The Basic domain in the bZIP regulatory protein Opaque2 serves two independent functions: DNA binding and nuclear localization. Plant J., 5(2):207-214.
- Vettore A.L., Yunes J.A., Neto G.C., da Silva M.J., Arruda P., Leite A. (1998). The molecular and functional characterization of an Opaque2 homologue gene from *Coix* and a new classification of plant bZIP proteins. Plant Mol Biol, 36:249-263.
- Vicente-Carbajosa, J., Moose, S.P., Parsons, R.L., and Schmidt, R. (1997). A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc. Natl. Acad. Sci. USA, 94: 7685-7690.
- Vicente-Carbajosa, J., Onate, L., Lara, P., Diaz, I., and Carbonero, P. (1998). Barley BLZ1: a bZIP transcriptional activator that interacts with endosperm-specific gene promoters. The Plant Journal, 13(5). 629-640.
- Vincentz, M., Bandeira-Kobarg, C., Gauer, L., and Leite, A. Identification of four Arabidopsis bZIP factors that share the same ancestral gene as the maize Opaque-2 regulatory protein. Phylogenetic characterization of the higher plants *Opaque-2* multigene family. EM PREPARAÇÃO.
- Walsh, J. Waters, C.A. and Feeling, M. (1997). The maize gene *liguleless2* encodes a basic leucine zipper protein involved in the establishment of the leaf blade-sheath boundary. Genes and Development, 11:208-218.

- Washburn, K. B. Davis, E. A. and Ackerman, S. (1997). Coactivators and TAFs of transcription activation in wheat. *Plant Mol. Biol.*, 35:1037-1043.
- Wellmer F., Kircher, S., Rügner, A., Frohnmeier, H., Schäfer, E., and Harter, K. (1999). Phosphorylation of the parsley bZIP transcription factor CPRF2 is regulated by light. *J. Biol. Chem.*, 274 (41): 29476-29482.
- Weishaar, B., Armstrong, G. A., Block, A., da Costa e Silva, O., and Hahlbrock, K. (1991). Light-inducible and constitutively expressed DNA-binding proteins recognizing a plant promoter element with functional relevance in light responsiveness. *EMBO J.*, 10:1077-1786.
- Yunes, J.A. Cord-Neto, G., Silva, M.J., Leite, A., Ottoboni, L.M.M. and Arruda, P. (1994). The transcriptional activator Opaque2 recognizes two different target sequences in the 22-kD-Like α -prolamin genes. *Plant cell*, 6:237-249.
- Yunes, J.A., Vettore, A. L., Silva, M.J., Leite, A., and Arruda, P. (1998). Cooperative DNA binding and discrimination by the Opaque2 bZIP factor. *Plant Cell*, 10:1941-1955.
- Yin Y., Zhu, Q., Da, S., and Beachy R. (1999). RF2a, a bZIP transcriptional activator of the phloem-specific rice tungro bacilliform virus promoter, function in vascular development. *EMBO J.*, 16: 5247-5259.
- Zhang Y., Fan, W., Kinkeman, M., Li, X., and 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. *PNAS USA*, 96: 6523-6528.