

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

CENTRO DE BIOTECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

**Ureases vegetais e suas chaperonas de ativação: um avanço  
na compreensão de suas propriedades estruturais e  
funcionais.**

Candidato: Msc. Rafael Real Guerra

Orientadora: Dra. Célia R. Carlini

Tese submetida ao Programa de Pós-graduação em Biologia Celular e Molecular da UFRGS como requisito parcial para obtenção do grau de Doutor em Ciências.

Porto Alegre, Março de 2011.

## **Membros da Banca Examinadora**

Profa. Dra. Leila Maria Beltramini

Depto. de Física e Informática, Instituto de Física de São Carlos – USP São Carlos

Profa. Dra. Ana Maria Oliveira Battastini

Depto. de Bioquímica, Instituto de Ciências Básicas da Saúde – UFRGS

Prof. Dr. Charley C. Staats

Depto. de Biologia Molecular e Biotecnologia, Instituto de Biociências – UFRGS

Dr. Deiber Olivera Severo (revisor e membro suplente)

Centro de Biotecnologia, Instituto de Biociências – UFRGS

Este trabalho foi desenvolvido em colaboração com o Dr. Dinesh Christendat da Universidade de Toronto, Canadá, Dr. Stefano Luciano Ciurli da Universidade de Bologna, Itália, e Dra. Leila Maria Beltramini do Instituto de Física de São Carlos, USP São Carlos.

## Agradecimentos

À minha família pelo apoio incondicional que me permitiu chegar até aqui.

À minha namorada Fernanda por não ter fugido durante o processo de confecção da tese.

À Dra. Leila Beltramini, pela disponibilidade e atenção com que sempre me recebe em seu laboratório.

Ao Dr. Stefano Ciurli, por me receber em seu laboratório e à Dra. Barbara Zambelli pelo apoio na realização de todos os experimentos durante minha estadia na Universidade de Bologna.

Ao Dr. Dinesh Christendat e aos membros de seu laboratório na Universidade de Toronto, pela disponibilidade em me receber durante o doutorado sanduíche e pelos valiosos ensinamentos em biologia molecular.

Ao Dr. Joseph Polacco, pelas valiosas ideias e discussões a respeito das proteínas acessórias e dos mistérios do universo.

À minha orientadora Célia Carlini, pelas oportunidades e por nos permitir escolher nossos próprios caminhos.

Às agências financiadoras CNPq, Capes, FINEP e FAPERGS.

# Índice

<b>Lista de abreviaturas</b>	2
<b>Resumo</b>	4
<b>Abstract</b>	6
<b>Introdução</b>	7
A urease de <i>Canavalia ensiformis</i> e seus aspectos estruturais	8
Mecanismo de ativação das ureases	12
UreE	15
UreG	15
UreF	16
UreD	16
Proteínas acessórias em plantas	17
<b>Objetivos</b>	18
<b>Capítulo I. Estudos cristalográficos da Urease de <i>Canavalia ensiformis</i> (JBU)</b>	
<b>Materiais e métodos</b>	
• Purificação de JBU	20
• Cristalização	20
• Coleta de Dados de Difração de Raios X	21
• Modificação química de resíduos de lisina	21

## **Estado da arte, resultados e discussão**

- Cristalização de JBU 22
  - Preparação da proteína 23
  - Co-cristalização na Presença de Ligantes 25
  - Proteólise *in situ* 27
  - Modificação Química de Resíduos de Lisina 29

## **Capítulo II. Análise da expressão dos genes das proteínas acessórias das ureases de soja (*Glycine max*)**

### **Materiais e métodos**

- PCR em tempo real 34
- Atividade ureásica 36

### **Resultados e discussão 36**

## **Capítulo III. Clonagem e expressão da proteína acessória UreD de soja (*Glycine max*)**

### **Materiais e métodos**

- Clonagem do gene *GmUreD20* 43
- Expressão de *GmUreD20* 44

### **Resultados e discussão 44**

- Produção de *GmUreD20* recombinante 44

## **Capítulo IV. Caracterização da proteína acessória UreF de soja (*Glycine max*)**

### **Materiais e métodos**

- Clonagem do gene *rGmUreF02* 47
- Expressão e purificação de *GmUreF02* 47
- Cromatografia de gel-filtração 48

• Análise por espectrometria de massas (MS/MS)	48
• Dicroísmo Circular	49
<b>Resultados e discussão</b>	49
• Produção de <i>GmUreF02</i> recombinante	49
• Análise estrutural de <i>GmUreF02</i>	52

## Capítulo V. Caracterização da proteína acessória UreG de soja (*Glycine max*)

### Materiais e métodos

• Purificação de <i>GmUreG</i> de sementes de soja	55
• Clonagem do gene <i>GmUreG</i>	55
• Expressão e purificação de <i>GmUreG</i> recombinante	56
• Clonagem e expressão de uma forma truncada de <i>GmUreG</i>	56
• Análise por espectrometria de massas	57
• Dicroísmo Circular	57
• Análise por Ressonância Magnética Nuclear (RMN)	57
• Medidas de Espalhamento de Luz	58
• Espectroscopia de Emissão Óptica com Plasma Acoplado Indutivo	58
• Titulação Calorimétrica Isotérmica	59
• Atividade GTPásica	59
• Modelagem molecular de <i>GmUreG</i>	60
<b>Resultados e discussão</b>	60
• Purificação de <i>GmUreG</i> das sementes de soja	60
• Expressão de <i>GmUreG</i> recombinante	62
• Análise estrutural de <i>GmUreG</i>	64
• Propriedades de ligação a metais	71
• Produção de uma forma truncada de <i>GmUreG</i>	76

• Atividade GTPásica	78
• Modelagem Molecular	82
<b>Conclusão geral e perspectivas</b>	85
<b>Referências bibliográficas</b>	86
<b>Anexos</b>	
I - Tabela 1 - Dados da quantificação de metais de <i>GmUreG</i> pela técnica de ICP-OES	96
II - Currículo	98
III - Artigo publicado	102
IV - Manuscrito	113



## Lista de abreviaturas

ABC – Complexo Dimetilamino-borano (*Dimethylamino-borano complex*)

BjHypB – Proteína HypB de *Bradyrhizobium japonicum*

BpUreG – Proteína UreG de *Bacillus pasteurii*

CD – Dicroísmo Circular (*Circular Dichroism*)

cDNA – Ácido Desoxirribonucleico Complementar (*Complementary Desoxyribonucleic Acid*)

CERM – Centro de Ressonância Magnética da Universidade de Firenze

DEAE – Dietilaminoetil

DNA – Ácido Desoxirribonucleico (*Dexosyribonucleic acid*)

dNTP – Desoxinucleotídeo Trifosfato (*Desoxinucleotide Triphosphate*)

EDTA – Ácido etilendiamino tetra-acético (*Ethylenediamine tetraacetic acid*)

ESI – Ionização por Eletrospray (*Eletrospray Ionization*)

EST – *Expressed Sequence Tag*

FPLC – Sistema para cromatografia líquida de proteínas (*Fast Protein Liquid Chromatography*)

GF – Gel-filtração

GmUreD02 – Proteína UreD de *Glycine max* do cromossomo 2

GmUreD20 – Proteína UreD de *Glycine max* do cromossomo 20

GmUreEe – Urease Embrião Específica de *Glycine max*

GmUreF02 – Proteína UreF de *Glycine max* do cromossomo 2

GmUreF14 – Proteína UreF de *Glycine max* do cromossomo 14

GmUreG – Proteína UreG de *Glycine max*

GmUreG $\Delta$ 56 – Proteína UreG de *Glycine max* truncada no seu 56º resíduo de amino ácido

GmUreG200 – Proteína UreG purificada a partir das sementes de *Glycine max* eluída da cromatografia de afinidade por metal com 200 mM de Imidazol

GmUreG300 - Proteína UreG purificada a partir das sementes de *Glycine max* eluída da cromatografia de afinidade por metal com 300 mM de Imidazol

GmUreUb – Urease Ubíqua de *Glycine max*

GTP – Guanosina Trifosfato

HEPES – Ácido 4-(2-hidroxietil)-1-piperazina etanossulfônico (*4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid*)

HpUreF – Proteína UreF de *Helicobacter pylori*

ICP-OES – Espectroscopia de Emissão Ótica de Plasma Indutivamente Acoplado (*Inductively coupled plasma optical emission spectroscopy*)

IPTG – Isopropil β-D-1-tiogalactopiranosídeo

ITC – Titulação Calorimétrica Isotérmica (*Isothermal Titration Calorimetry*)

JBU – Urease de *Canavalia ensiformis* (*Jack Bean Urease*)

KaUreG – Proteína UreG de *Klebsiella aerogenes*

LNLS – Laboratório Nacional de Luz Sincrotron

MALS – Espalhamento de Luz em Múltiplos Ângulos (*Multi-angle Light Scattering*)

MjHypB – Proteína HypB de *Methanococcus jannaschii*

mRNA – Ácido Ribonucleico Mensageiro (*Messenger Ribonucleic Acid*)

MS/MS – Espectrometria de Massas em Tandem (*Tandem Mass Spectrometry*)

MtUreG – Proteína UreG de *Mycobacterium tuberculosis*

nrmsd – Desvio quadrático médio normalizado (*Normalized Root Mean Square Deviation*)

PAGE – Eletroforese em gel de poliacrilamida (*Polyacrylamide Gel Electrophoresis*)

PCR – Reação em Cadeia da Polimerase (*Polymerase Chain Reaction*)

QELS – Espalhamento de Luz Quasi-Elastico (*Quasi-Elastic Light Scattering*)

Q-TOF – Analizadores Quadripolo – Tempo de Vôo (*Quadripole – Time Of Flight Analyzers*)

rGmUreG – Proteína UreG de *Glycine max* produzida de forma recombinante em *Escherichia coli*

R/HypB – Proteína HypB de *Rhizobium leguminosarum*

RMN – Ressonância Magnética Nuclear

RNA – Acido Ribonucleico (*Ribonucleic Acid*)

SAXS – Espalhamento de Raios X em Baixo Ângulo (*Small Angle X-ray Scattering*)

SDS – Dodecil Sulfato de Sódio (*Sodium Dodecyl Sulfate*)

UV – Luz Ultra Violeta

## Resumo

A ampla distribuição das ureases na natureza é um indício da grande importância desta enzima para os mais diferentes organismos, fato que levou diversos pesquisadores ao redor do globo a dedicarem-se exclusivamente à caracterização destas enzimas. Apesar de ser uma proteína estudada há quase um século, até o início deste trabalho apenas três ureases possuíam suas estruturas cristalográficas elucidadas e todas de origem bacteriana. A urease de *Canavalia ensiformis* (JBU) foi a primeira enzima desta família a ser descrita, em 1926 e ainda assim, apesar dos incansáveis esforços de diferentes grupos na caracterização estrutural e biológica desta proteína, sua estrutura cristalográfica só foi obtida no primeiro semestre de 2010. Tentativas anteriores de obtenção da estrutura cristalográfica de JBU falharam devido a baixa qualidade dos cristais formados por esta proteína através de técnicas convencionais de cristalização. Desta forma, na primeira parte deste trabalho relatamos o uso de técnicas alternativas de cristalização como o uso de ligantes ou aditivos, proteólise *in situ* e modificação química de resíduos de amino ácidos na obtenção de cristais de JBU com qualidade superior aos anteriormente obtidos, validando o uso destas metodologias na cristalização de proteínas recalcitrantes. Os cristais aqui obtidos, apesar de apresentarem qualidade superior aos previamente descritos na literatura, não foram otimizados ao ponto de obtenção da estrutura cristalográfica da enzima.

A atividade enzimática das ureases é dependente da presença de dois íons de níquel precisamente incorporados em seus sítios ativos. A biossíntese deste sítio ativo, bem como a incorporação de níquel na enzima consiste de um processo altamente regulado, cuja ocorrência depende da participação de diversas chaperonas (proteínas acessórias) agindo como reguladores pós traducionais. Apesar dos esforços já realizados, o mecanismo de ativação de ureases ainda permanece obscuro. Até hoje, os estudos concentraram-se em ureases de origem microbiana, sendo a informação disponível ainda extremamente limitada. Menos ainda é conhecido para o sistema de ativação de ureases de origem vegetal. Na segunda parte deste trabalho relatamos pioneiramente as primeiras tentativas de produção e caracterização de proteínas acessórias de origem vegetal. As proteínas UreD, UreF e UreG de soja (*Glycine max*) foram clonadas e expressas em sistema bacteriano. A proteína UreG de soja, purificada diretamente da planta e também produzida de forma recombinante em *E.coli*, foi caracterizada quanto a sua estrutura, capacidade de ligação a metais e atividade GTPásica, descrevendo algumas características nunca antes observadas para outras proteínas da mesma família. Também descrevemos o primeiro sucesso na obtenção da proteína acessória UreF tipo

selvagem na sua forma solúvel e a caracterização estrutural desta proteína foi realizada. As proteínas acessórias de soja também foram caracterizadas em relação ao seu perfil de expressão em diferentes tecidos, ao longo do desenvolvimento da planta, utilizando PCR em tempo real. Observamos uma grande variação dos níveis de expressão dos diferentes mRNAs nos tecidos avaliados. Uma possível relação entre estes níveis de expressão e a atividade ureásica de cada tecido foi investigada, visando uma melhor compreensão da dinâmica entre a expressão e atividade das proteínas responsáveis pela ativação das ureases vegetais.

## Abstract

The wide distribution of ureases in nature is an indication of the importance of this enzyme for the most different organisms, fact that lead researchers all over the world to dedicate their careers exclusively to the study of these enzymes. Despite of being studied for almost a century, until last year only three crystallographic structures of ureases had been solved, all from bacterial source. The urease from *Canavalia ensiformis* (JBU) was the first member of this family to be described, in 1926 and still, despite the restless efforts from different research groups on the structural and biological characterization of this protein, its structure was not solved until the beginning of 2010. Early trials to obtain the crystallographic structure of JBU have failed due to the poor quality crystals formed by this protein using conventional crystallization techniques. In this panorama, the first part of this work describes the use of alternative crystallization techniques, such as the use of ligands, *in situ* proteolysis and chemical modification of amino acid residues, for the obtention of JBU crystals with superior quality than the ones previously described in the literature, validating the use of such techniques on the crystallization of recalcitrant proteins. The crystals obtained here, despite of their superior characteristics, were not refined to the point of generating the crystallographic structure of JBU.

The enzymatic activity of ureases is dependent of the presence of nickel ions precisely incorporated in their active sites. The bioassembly of these active sites, including the nickel incorporation, consist of a tightly regulated process, whose occurrence depends on the participation of several chaperones (accessory proteins) that work as post-translational regulators. A great deal of effort has already been done on the study of the urease's activation process; however its details remain unclear. To date, all work done concentrates on the study of microbial ureases, but the available information is very limited and even less information is available for the plant ureases activation process. The second part this work presents the first trials on the production and characterization of plant urease accessory proteins. The proteins UreD, UreF and UreG from soybean (*Glycine max*) were cloned and expressed in bacterial system. The soybean UreG protein, purified directly from the plant seeds and also produced in recombinant fashion in *Escherichia coli*, was characterized regarding its structure, metal binding capacity and GTPase activity, describing a few characteristics never before observed for other proteins of the same family. It is also described the first success on the obtention of the wild type ureF accessory protein in its soluble form and a structural characterization was performed. All accessory proteins from soybean were also characterized for its expression profile in several tissues during its development utilizing real-time PCR and a correlation with the levels of urease activity were investigated. Altogether, these data improve the understanding of multiple factors involved on the urease activation process.

## Introdução

Ureases (uréia aminohidrolases EC 3.5.1.5) são metaloenzimas dependentes de  $\text{Ni}^{2+}$  responsáveis por catalizar a hidrólise da uréia à amônia e carbamato, o qual se decompõe espontaneamente para formar uma segunda molécula de amônia e dióxido de carbono [Carter et al., 2009; Krajewska et al., 2009]. São enzimas amplamente distribuídas na natureza, podendo ser encontradas em bactérias, fungos e plantas.

A existência de ureases em tecidos vegetais é bastante comum, sendo leguminosas, como a soja (*Glycine max*) e o feijão de porco (*Canavalia ensiformis*), particularmente ricas nessas proteínas [Polacco e Holland, 1993]. O papel exercido pelas ureases nos diferentes organismos onde são encontradas é amplo. Em bactérias, além da contribuição para o metabolismo de compostos nitrogenados, ureases também são consideradas importantes fatores de virulência que participam, por exemplo, na infecção gastroduodenal, formação de úlceras e até mesmo câncer de estômago por *Helicobacter pylori*, na formação de pedras nos rins e outras doenças do trato urinário causadas por muitas bactérias incluindo *Proteus mirabilis*, entre outras doenças, tanto de animais como de humanos [Mulrooney e Hausinger, 2003; Mobley et al., 1995; Carlini e Polacco, 2008]. Já em plantas, o papel das ureases ainda não está bem estabelecido. Propõe-se que ação das ureases seja o único meio de reaproveitamento de nitrogênio a partir da uréia proveniente de algumas rotas metabólicas (uréideos, purinas e arginina). Isso foi comprovado quando plantas de soja desprovidas de urease, quer induzidas geneticamente [Meyer-Bothling e Polacco, 1987; Polacco et al., 1989], pelo uso de inibidores, ou por remoção do  $\text{Ni}^{2+}$  do meio de crescimento [Polacco e Holland, 1993], apresentaram acúmulo de uréia nas folhas e raízes, assim como comprometimento do emprego da uréia como fonte de nitrogênio. Um fato bastante curioso é que, mesmo a uréia não sendo um metabólito majoritário em plantas, estas usualmente apresentam abundantes quantidades de urease em seus tecidos [Polacco e Holland, 1993]. A descoberta de uma segunda isoforma de urease em soja trouxe consigo uma série de ponderações a cerca da função desta enzima em plantas. Uma das isoformas, denominada embrião-específica, é sintetizada somente durante o desenvolvimento do embrião na semente em formação, e acumula-se na semente madura [Torisky e Polacco, 1990], enquanto a outra, denominada ubíqua, está presente em todos os tecidos da planta [Polacco e Winkler, 1984]. Estudos demonstraram que a urease ubíqua está envolvida na reciclagem da uréia, uma vez que mutantes negativos para esta isoforma apresentam anomalias típicas do acúmulo de uréia, como necrose de folhas e raízes e retardo na germinação, e acumulam uréia em folhas e

sementes [Polacco e Holland 1993]. Já mutantes negativos para urease embrião-específica não apresentaram qualquer fenótipo diferenciado em relação à planta normal, sendo necessário ressaltar que esta isoforma de urease de soja é cerca de 1.000 vezes mais abundante nas sementes do que a forma ubíqua. Neste contexto, surge a possibilidade de que ureases apresentem outra função em plantas, além da reciclagem de uréia. Duas isoformas de urease de *C. ensiformis* e a urease embrião-específica de soja são letais para algumas ordens de insetos [Carlini et al., 1997; Follmer et al., 2004a; Stanisçuaski et al., 2005], tornando a defesa uma das possíveis funções das ureases em plantas. Além da atividade inseticida, já foi descrita a atividade antifúngica da urease embrião-específica de soja, da urease de *C.ensiformis* (JBU) e de *Gossypium hirsutum*, além da urease da bactéria *H. pylori* [Becker-Ritt et al., 2007; Menegassi et al., 2008]. O efeito fungitóxico leva à lesões da parede e/ou membrana celular e, na maioria das vezes, à morte das células. Desde então, exaustivos estudos nessa área vem sendo realizados.

A atividade inseticida das ureases é parcialmente devida à liberação de peptídeos entomotóxicos por enzimas do trato digestório dos insetos, a partir da proteína intacta [Ferreira-DaSilva et al., 2000; Piovesan et al., 2008]. Um destes peptídeos, Jaburetox-2Ec, já foi caracterizado e expresso em sistema recombinante [Mulinari et al., 2007]. Hoje sabe-se que a urease intacta chega à hemolinfa dos insetos e também contribui para o efeito inseticida, alterando o transporte de fluidos nos túbulos de Malpighi e intestino anterior dos insetos [Staniscuaski et al., 2009; 2010].

## **A urease de *Canavalia ensiformis* e seus aspectos estruturais**

A urease de *Canavalia ensiformis* (JBU –Jack Bean Urease) é uma proteína de grande relevância histórica, tendo sido o personagem principal de importantes marcos na história da bioquímica moderna. A JBU foi a primeira enzima cristalizada da história, e através da análise de seus cristais foi possível evidenciar o caráter proteico das enzimas, derrubando a teoria vitalista sobre a origem dos compostos orgânicos; façanha realizada por James Sumner, publicada em 1926 e que lhe renderia o prêmio Nobel de química vinte anos depois. Em 1975, um outro marco foi alcançado com a descoberta de que  $Ni^{2+}$  era um componente essencial para a atividade de JBU, tendo sido esta a primeira demonstração de um papel biológico para  $Ni^{2+}$  [Dixon et al., 1975].

Apesar da primeira descrição da cristalização da JBU datar de cerca de 85 anos atrás e que durante esse tempo esta proteína foi alvo de exaustivas caracterizações do ponto de vista bioquímico, alguns aspectos relacionados a sua estrutura ainda são motivo de divergências e especulações como, por exemplo, sua forma oligomérica nativa em solução. Estudos de nosso grupo utilizaram espalhamento de luz dinâmico e estático para investigar a influência de alguns fatores, como a concentração de proteína, composição de tampões e agentes redutores sobre o comportamento de JBU, evidenciando que todos estes fatores podem alterar o estado de oligomerização desta proteína [Follmer et al., 2004b]. Estes dados são representativos da grande complexidade desta molécula, além de oferecer uma possível explicação para a grande discordância sobre o tema como observado entre diferentes artigos [Sumner e Somers, 1953; Gorin et al., 1967; Bailey e Boulter, 1969; Contaxis e Reithel, 1971; Dixon et al., 1980; Takishima et al., 1988].

A cadeia polipeptídica de JBU é composta por 840 amino ácidos, conforme dados de sequenciamento direto da proteína [Mamiya et al., 1985] e dedução a partir de seu cDNA [Riddles et al., 1991], com massa molecular de 90.770 Da. Sua forma mínima enzimaticamente ativa em solução é a de um trímero com 270 kDa e a conformação mais amplamente aceita na comunidade científica para sua forma nativa é de um hexâmero de 540 kDa [Zerner, 1991]. JBU apresenta em seu sítio ativo dois íons  $Ni^{2+}$ , sendo cada um deles coordenado por dois resíduos de histidinas e por uma lisina carboxilada compartilhada por ambos [Zerner, 1991; Jabri et al., 1995].

Até hoje, apenas quatro ureases tiveram sua estrutura cristalográfica elucidada, sendo três delas de origem bacteriana. São elas a urease de *Klebsiella aerogenes* [Jabri et al., 1995], *Bacillus pasteurii*, [Benini et al., 1998], *Helicobacter pylori* [Ha et al., 2001] e, recentemente o primeiro membro do reino vegetal, JBU [Balasubramanian e Ponnuraj, 2010]. A ausência de estruturas resolvidas por técnicas experimentais para outras ureases vegetais deve-se principalmente a dificuldade na obtenção destas proteínas com um alto grau de pureza a partir do organismo de origem. Essa dificuldade pode ser facilmente superada para ureases bacterianas, pois as mesmas apresentam seus genes organizados em clusters, o que torna simples sua super expressão em conjunto com suas proteínas acessórias, essenciais para a produção da enzima ativa (proteínas acessórias serão discutidas mais adiante no texto). Para ureases vegetais, a diferente organização de genes acessórios, a menor quantidade de genomas sequenciados e outros fatores, como o baixo grau de identidade entre proteínas acessórias de ureases de origem vegetal e bacteriana, tornam extremamente difícil a bem sucedida expressão heteróloga dessas ureases com 100% de ativação. Mesmo os poucos casos

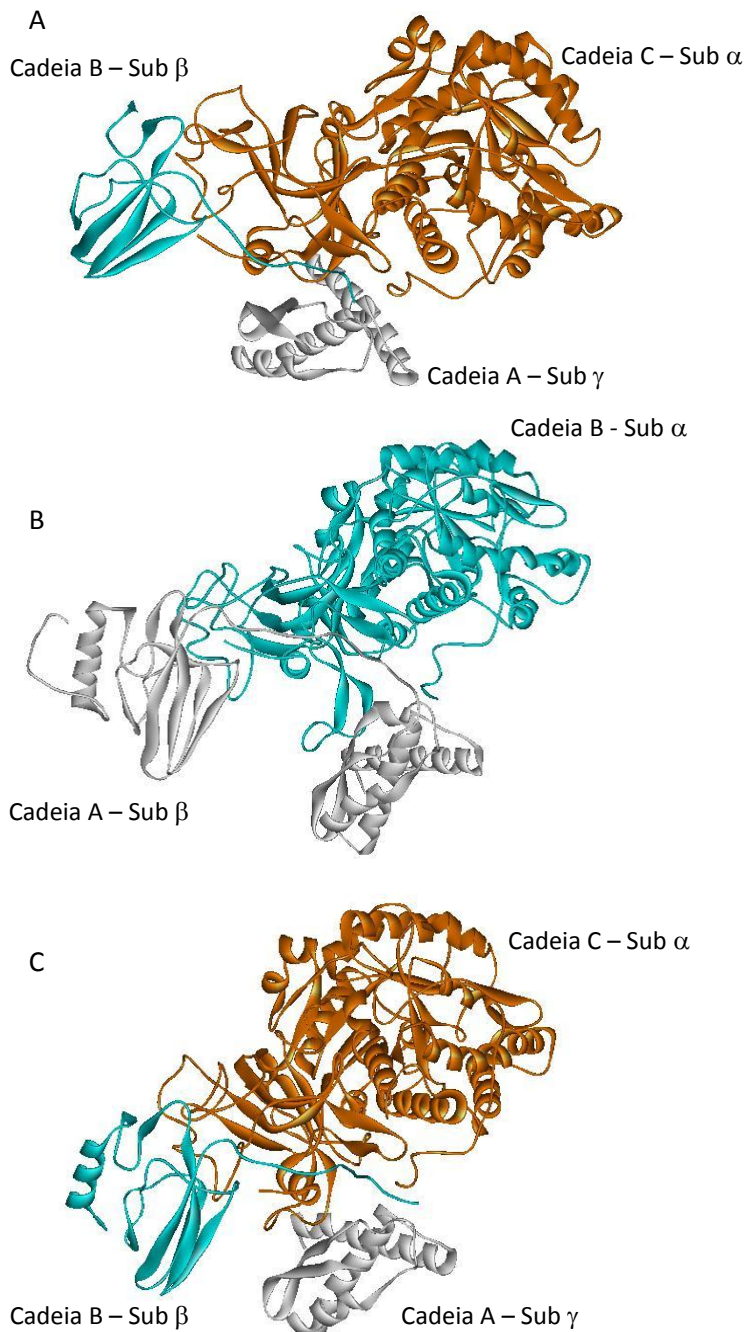


onde todos os genes envolvidos na ativação da enzima vegetal foram identificados, como em *Arabidopsis thaliana*, a expressão conjunta das proteínas acessórias com o gene estrutural da urease levou apenas a uma parcial ativação da enzima, em sistema bacteriano [Witte et al., 2005].

Comparando-se as estruturas cristalográficas de ureases microbianas (*K. aerogenes*, *H. pylori* e *B. pasteurii*), observa-se um alto grau de conservação tanto em termos de números de estados de estrutura secundária (quantidade de hélices  $\alpha$  e fitas  $\beta$ ), como na porcentagem de resíduos de aminoácidos envolvidos em cada estado (porcentagem de hélices  $\alpha$  e fitas  $\beta$ ). Ureases vegetais apresentem cerca de 50% de identidade de sequência com ureases microbianas, no entanto, uma importante diferença estrutural está no fato de que a unidade básica de ureases vegetais é composta por uma cadeia polipeptídica única agrupada em dímeros de trímeros ( $\text{urease}_3$ )<sub>2</sub> [Balasubramanian e Ponnuraj, 2010], enquanto unidades básicas de ureases microbianas são formadas por duas ou três cadeias (Figura 1), que podem se organizar em trímeros de trímeros ( $\alpha\beta\gamma$ )<sub>3</sub> [Jabri et al., 1995], hexâmeros de trímeros [( $\alpha\beta\gamma$ )<sub>3</sub>]<sub>2</sub> [Benini et al., 1998] ou até mesmo, para o caso da urease de *H. pylori*, dodecâmeros de dímeros [( $\alpha\beta$ )<sub>3</sub>]<sub>4</sub> [Ha et al., 2001]. As cadeias A (subunidade  $\gamma$ , com 100 aminoácidos) e B (subunidade  $\beta$ , com 106 aminoácidos) da urease de *K. aerogenes* alinham-se com a porção N terminal da urease de *C. ensiformis*, do resíduo 1 ao 232. Já a cadeia C (subunidade  $\alpha$ , com 567 aminoácidos) deste mesmo organismo, alinha-se com a porção C terminal da urease de *C. ensiformis*, do resíduo 274 ao 840, onde está localizado o sítio ativo de ambas. A urease de *H. pylori* apresenta apenas duas cadeias, A e B, sendo que a cadeia A (238 resíduos) corresponde à fusão das cadeias A e B de *K. aerogenes*, e a B (569 resíduos) alinha-se com a porção C-terminal das ureases de plantas. A urease de *B. pasteurii* apresenta 3 cadeias muito semelhantes a urease de *K. aerogenes*, alinhando-se da mesma forma que essa com a urease de *C. ensiformis*.

Diversas tentativas de obtenção da estrutura cristalográfica de JBU foram realizadas desde sua primeira cristalização em 1926, no entanto, os cristais obtidos eram incapazes de difratar raios X [Jabri et al., 1992; Jabri et al., 1995]. O primeiro sucesso nesta área foi obtido em 2002, quando Sheridan e colaboradores obtiveram cristais de JBU conjugados a um fragmento de anticorpo. Porém, estes cristais não eram estáveis o suficiente para obtenção de mais do que um único dado de difração (com resolução de cerca de 3.0 Å). Isso se deve ao fato de que os cristais estavam sendo expostos a raios X em condições não criogênicas uma vez que, segundo o mesmo trabalho, a presença de agentes crioprotetores eliminavam a difração

de raios X pelo cristal. Apenas recentemente (primeiro semestre de 2010) a estrutura cristalográfica de JBU pode ser obtida [Balasubramanian e Ponnuraj, 2010].



**Figura 1:** Estrutura de ureases microbianas determinadas por cristalografia de difração de raios X.

- A. Estrutura da urease de *K. aerogenes*, formada por três cadeias distintas: cadeia A (subunidade  $\gamma$ ), cadeia B (subunidade  $\beta$ ) e a cadeia C (subunidade  $\alpha$ ). Acesso PDB: 1FWJ.
- B. Estrutura da urease de *H. pylori*, formada por duas cadeias, cadeia A e cadeia B. Acesso PDB: 1E9Z.
- C. Estrutura da urease de *B. pasteurii*, formada por três cadeias distintas: cadeia A, cadeia B e cadeia C. Acesso PDB: 2UBP.

## Mecanismo de ativação das ureases

A biossíntese de metaloenzimas geralmente requer a participação de proteínas (chaperonas) específicas e indispensáveis para a correta construção de seus sítios ativos. Tais proteínas são geralmente chamadas de proteínas acessórias e sua função consiste da estabilização de uma determinada conformação da apoenzima, que permita a correta inserção do metal no sítio ativo, dissociando-se da enzima madura após a ativação [Kuchar e Hausinger, 2004].

No caso das enzimas dependentes de  $\text{Ni}^{2+}$ , como urease, monóxido de carbono desidrogenase, acetil coenzima A sintetase e hidrogenase, conjuntos de proteínas acessórias essenciais já foram identificadas [Kuchar e Hausinger, 2004; Mulrooney e Hausinger, 2003; Loke e Lindahl, 2003]. Embora cada uma destas enzimas necessite de proteínas acessórias específicas para sua ativação, é interessante observar que algumas propriedades são comuns entre elas. Para cada sistema parece existir o requerimento de proteínas acessória capazes de realizar as seguintes funções: hidrolizar nucleotídeo trifosfato, armazenar  $\text{Ni}^{2+}$ , entregar  $\text{Ni}^{2+}$  ao sítio ativo. Essas funções são exercidas por domínios específicos que podem estar contidos em uma única proteína, ou divididos em múltiplas chaperonas. A função de armazenamento de  $\text{Ni}^{2+}$  é normalmente exercida por regiões ricas em histidinas, sendo que estas regiões podem constituir uma proteína inteira, como o caso da Hpn - um pequeno acumulador de  $\text{Ni}^{2+}$  encontrado em *H. pylori* [Ge et al., 2006a;b; Seshadri et al., 2007] - ou estar localizado no N ou C terminal de chaperonas específicas, como no caso da proteína acessória da urease, UreE [Musiani et al., 2004], de algumas acessórias de hidrogenases, HypB [Olson e Maier, 2000; Rey et al., 1994] e SlyD [Wulfig et al., 1994], e de acessórias da monóxido de carbono desidrogenase, CooJ [Watt e Ludden, 1998]. A função de entrega do  $\text{Ni}^{2+}$  ao sítio ativo da enzima pode estar conjugada à proteína responsável pelo armazenamento deste íon, no entanto cada função é realizada através de sítios específicos. Exemplos de proteínas envolvidas na entrega de  $\text{Ni}^{2+}$  para diferentes enzimas são UreE [Remaut et al., 2001; Song et al., 2001; Stola et al., 2006], HypA [Mehta et al., 2003a; Blokesch et al., 2004; Atanassova e Zamble, 2005], HypB [Leach et al., 2005; Gasper et al., 2006], SlyD [Zhang et al., 2005] e CooJ [Watt e Ludden, 1998]. Finalmente, a hidrólise de nucleotídeo trifosfato é requerida para a bem sucedida formação do sítio ativo e as proteínas que realizam esta função normalmente apresentam um motivo P-loop homólogo, como é o caso da HypB [Maier et al., 1993; Mehta et al., 2003b] para hidrogenase, AcsF [Loke e Lindahl, 2003] para acetil coenzima A sintetase, CooC [Kerby et al., 1997] para monóxido de carbono desidrogenase e UreG [Moncrief e

Hausinger, 1997; Soriano e Hausinger, 1999] para urease. No entanto, apesar destes aspectos gerais serem mantidos entre as proteínas acessórias, a maior parte do processo de maturação de enzimas dependentes de  $\text{Ni}^{2+}$  permanece em completa obscuridade.

O processo de ativação das ureases é extremamente complexo, e está melhor caracterizado para bactérias do que para qualquer outro organismo, especialmente para a enterobactéria *Klebsiella aerogenes*. O conhecimento atual sobre o processo de ativação está resumido na Figura 2. O complexo de proteínas necessário para a ativação total da urease neste sistema envolve a participação de quatro proteínas acessórias, UreD, UreE, UreF e UreG, assim como uma molécula de  $\text{CO}_2$  e uma de GTP [Soriano et al., 2000].

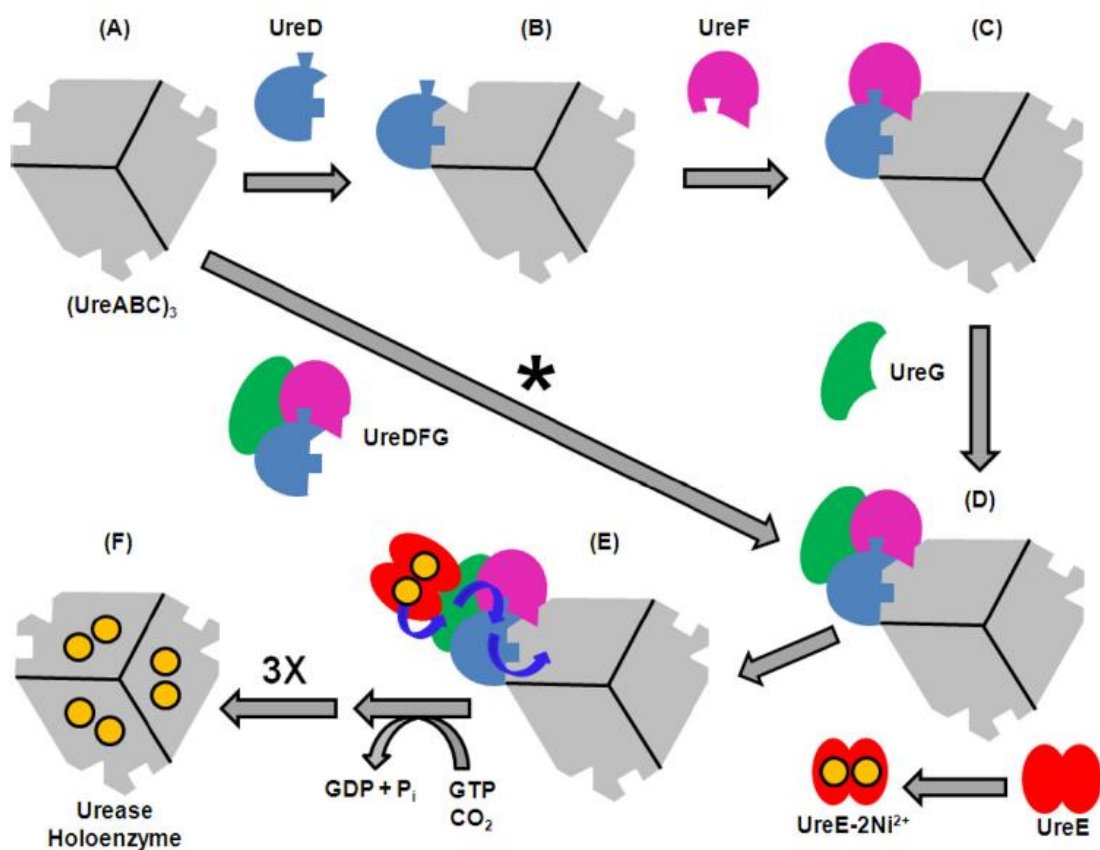


Figura 2: Modelos propostos para ativação da urease. A apoproteína, trimeros de trimeros (A) sequencialmente liga UreD, UreF e UreG para formar UreABC-UreD (B), UreABC-UreDF (C) e UreABC-UreDFG (D). Alternativamente, UreDFG pode existir como uma chaperona molecular pré-formada que liga-se diretamente à apoproteína, como indicado pelo asterisco. UreE entrega  $\text{Ni}^{2+}$  através de interação com UreG (E). *In vivo*, a ativação da urease necessita de  $\text{CO}_2$  e GTP, que é hidrolizado por UreG no complexo de ativação. É proposto que a entrega do  $\text{Ni}^{2+}$  ocorra de UreE para UreG para UreD e, finalmente, para a urease (F). Extraído de Carter e Hausinger, 2010.

Dois modelos são propostos para a participação destas proteínas acessórias no processo de ativação. O primeiro sugere um processo que ocorre passo a passo, onde uma a uma as proteínas acessórias se associam à apourease, formando a seguinte sequência de complexos: UreD-apourease [Park et al., 1994], UreDF-apourease [Moncrief e Hausinger, 1996], and UreDFG-apourease [Park e Hausinger, 1995]. O segundo modelo, no entanto, propõe que o complexo UreDFG seja formado previamente à sua ligação à apourease [Moncrief e Hausinger, 1997]. Em ambos os cenários, a proteína ureE, necessária para máxima ativação mas não considerada uma proteína acessória essencial, ligar-se-ia ao complexo UreDFG-apourease facilitando a ativação da enzima [Lee et al., 1993; Brayman e Hausinger, 1996; Soriano et al., 2000; Zambelli et al., 2005]. O papel específico exercido por cada uma das quatro proteínas acessórias ainda é alvo de estudos e a maior parte da informação disponível foi obtida através da caracterização *in vitro* de UreE e UreG. O conhecimento atual sugere que UreD se ligue a urease, e apenas promova algum tipo de mudança conformacional que a prepara para o próximo passo da ativação [Park et al., 1994; 1996]. No entanto, estudos envolvendo UreD de *K. aerogenes* conjugada à proteína ligadora de maltose mostram que UreD é capaz de ligar  $\text{Ni}^{2+}$ , e algum papel na entrega do  $\text{Ni}^{2+}$  ao sítio ativo da urease não pode ser descartado [Carter e Hausinger, 2010]. UreF liga-se ao complexo UreD-apourease e propõe-se ser ela a responsável por uma grande mudança conformacional na subunidade  $\beta$ , possibilitando o acesso direto ao sítio ativo da enzima e facilitando tanto a carboxilação da lisina do sítio ativo, como a inserção do  $\text{Ni}^{2+}$  no mesmo [Chang et al., 2004]. UreG liga-se ao complexo UreFD-apourease e, através da clivagem de GTP, fornece a energia necessária para o sistema [Soriano e Hausinger, 1999]. Duas hipóteses são propostas para explicar o papel da UreG no complexo de ativação. A primeira propõe que UreG poderia usar GTP e  $\text{CO}_2$  para sintetizar carboxifosfato, que seria um ótimo doador de  $\text{CO}_2$  para a lisina do sítio ativo [Soriano e Hausinger, 1999]. A outra hipótese sugere que a hidrólise de GTP estaria relacionada apenas à entrega do  $\text{Ni}^{2+}$  ao sítio ativo e que este papel poderia ser exercido pela própria UreG [Zambelli et al., 2005]. A proteína acessória UreE liga-se ao complexo UreGFD-apourease, sendo requerida para máxima ativação na maior parte dos sistemas bacterianos. No entanto, esta proteína não está presente em alguns microorganismos como o parasita *Mycobacterium tuberculosis* (conforme informação de genoma) e não é considerada essencial para a ativação da urease, uma vez que em sua ausência uma ativação parcial da enzima é observada [Lee et al., 1992]. Ainda assim, ambos papéis, de entrega e armazenamento de  $\text{Ni}^{2+}$  são atribuídos à esta proteína [Bellucci et al., 2009].

Os extensivos estudos descritos acima formam a base para o entendimento do mecanismo pelo qual as proteínas acessórias estão envolvidas na ativação da urease. O passo seguinte no entendimento da química envolvida na construção do sítio ativo e ativação da urease compreende o estudo do mecanismo de interação entre as proteínas acessórias e a enzima em nível molecular. Esse passo não pode ser dado sem que uma caracterização estrutural e bioquímica detalhada de cada uma das proteínas acessórias seja realizada.

## UreE

Estruturas cristalográficas para a proteína acessória UreE de *K. aerogenes* [Song et al., 2001], *H. pylori* [Shi et al., 2010] e *B. pasteurii* [Remaut et al., 2001] estão disponíveis, e a grande similaridade entre as estruturas destas proteínas, bem como das ureases destes organismos, são indicativos de um mecanismo conservado de ativação [Musiani et al., 2004]. A caracterização estrutural da UreE mostra que estas proteínas são homodímeros cuja interface, envolvendo o domínio C terminal da proteína, forma um sítio de ligação a metais [Bellucci et al., 2009; Song et al., 2001; Remaut et al., 2001; Shi et al., 2010]. A caracterização da capacidade de ligação a metais foi realizada para UreE de *K. aerogenes* [Colpas et al., 1999], *B. pasteurii* [Ciurli et al., 2002] e *H. pylori* [Bellucci et al., 2009; Shi et al., 2010], onde diferentes estequiometrias de ligação são observadas, devido às diferenças no conteúdo de histidinas no C terminal de cada proteína. No entanto, o sítio comum, e portanto considerado funcionalmente mais importante, corresponde aos dois resíduos de histidina localizados na interface entre os dois monômeros. A interação de UreE com seu mais provável parceiro, UreG, foi analisada para UreEs de *H. pylori* [Bellucci et al., 2009] e *K. aerogenes* [Boer et al., 2010] e em ambos os casos, a interação entre essas moléculas parece ser intermediada por íons  $Zn^{2+}$  e  $Ni^{2+}$ . Porém, inconsistências com relação a estequiometria de interação na formação do complexo ainda levantam dúvidas a respeito deste sistema.

## UreG

UreGs de diferentes microrganismos, como *M. tuberculosis* [Zambelli et al., 2007], *H. pylori* [Bellucci et al., 2009; Zambelli et al., 2009], *B. pasteurii* [Neyroz et al., 2006] e *K. aerogenes* [Boer et al., 2010; Zambelli et al., 2005] já foram estudadas. Estas proteínas apresentam um motivo P-loop, encontrado em GTPases, no entanto, sua atividade GTPásica é muito baixa ou não detectável [Boer et al., 2010; Moncrief e Hausinger, 1997; Zambelli et al., 2005; 2007; 2009]. Apesar de ser uma proteína solúvel e facilmente obtida através da super

expressão em *E. coli*, sua caracterização estrutural é bastante difícil e é alvo de muitas discordâncias na literatura, de forma que não há um consenso sobre o arranjo quaternário desta proteína em solução. Análises de RMN classificam UreG como membros da classe de proteínas intrinsecamente desordenadas, o que significa que esta proteína transita por diferentes conformações quando livre em solução, apenas adquirindo sua conformação ativa na presença de cofatores (proteicos ou não) [Zambelli et al., 2005]. Assim como UreE, UreG também é uma metalochaperona capaz de ligar diferentes metais, no entanto o papel fisiológico desta afinidade ainda é alvo de especulação [Boer et al., 2010; Zambelli et al., 2009].

## UreF

Diferentemente das proteínas UreE e UreG descritas acima, a quantidade de informação disponível para UreF é bastante reduzida. Isso se deve principalmente ao fato de UreF ser uma proteína expressa em quantidades muito pequenas e sem qualquer tipo de atividade mensurável caracterizada, que permita sua purificação a partir de fontes naturais. Adicionalmente, a maior parte destas proteínas mostram-se extremamente insolúveis quando produzidas em sistema bacteriano [Kim et al., 2006; 1999]. Baseado na modelagem molecular da estrutura de *BpUreF* [Salomone-Stagni et al., 2007] é proposto que esta proteína possa agir como uma ativadora de GTPases, possivelmente regulando a atividade de UreG (hipótese ainda não testada). Uma forma truncada de um membro desta família de proteínas, oriundo de uma cepa específica de *H. pylori* (cepa 26695), foi produzida em sua forma solúvel, dando origem a primeira estrutura cristalográfica de uma UreF [Lam et al., 2010]. No entanto, apesar da recente obtenção de sua estrutura, do ponto de vista bioquímico não há, até hoje, qualquer caracterização conclusiva a respeito da função desta proteína. Algumas tentativas de obtenção de UreF em sua forma solúvel foram feitas, nas quais UreF foi fundida a UreE [Kim et al., 2006] ou à proteína ligadora de maltose [Kim et al., 1999]. Apesar das proteínas híbridas serem capazes de promover uma ativação parcial da urease quando inseridas em seu cluster de genes, UreF é solúvel somente enquanto fusionada, fazendo sua caracterização bastante difícil. Ao contrário do observado para UreE, nenhum sítio de ligação a metal foi observado para UreF [Kim et al., 1999].

## UreD

Os estudos disponíveis sobre UreD são restritos a *K. aerogenes*. Essa proteína, assim como UreF, é insolúvel quando expressa isoladamente. Uma forma solúvel desta proteína, que

consiste de UreD conjugada à proteína ligadora de maltose [Carter e Hausinger 2010], permitiu que algumas caracterizações fossem feitas. Conforme este estudo, UreD também parece ser uma metalochaperona capaz de ligar tanto  $Ni^{2+}$  como  $Zn^{2+}$ . Sendo UreD a primeira proteína a se ligar diretamente à urease no processo de ativação, propõe-se que ela possa ter uma participação direta na entrega do  $Ni^{2+}$  ao sítio ativo da enzima. No entanto, é preciso ressaltar que quando o gene da proteína tipo selvagem foi substituído pelo gene da proteína fusionada, a ativação da urease foi apenas parcial, mostrando que esta proteína é apenas parcialmente funcional.

## Proteínas acessórias em plantas

Poucas proteínas acessórias de ureases vegetais foram identificadas até hoje. Em *Arabidopsis*, três proteínas necessárias para a ativação da urease foram encontradas: AtUreD, AtUreF e AtUreG [Witte et al., 2005]. UreG de batata (*Solanum tuberosum*) foi caracterizada em relação à localização gênica, sequência, número de cópias e expressão, em nível de RNA e proteína [Witte et al., 2001]. Em arroz (*Oryza sativa*), genes das proteínas acessórias UreD, UreF e UreG foram recentemente identificados e seus cDNAs correspondentes clonados [Cao et al., 2010].

Para soja, o gene Eu3 foi identificado como codificador de UreG [Freyermuth et al., 2000] e um segundo gene, Eu2, se mostrou essencial para a ativação da urease, mas a identidade e função da proteína codificada por esse gene ainda é desconhecida [Godraij et al., 2003]. De acordo com a informação genômica disponível no site Phytozome (<http://www.phytozome.net/>), a soja possui dois genes codificadores para UreD, localizados nos cromossomos 02 (ID: Glyma02g20690.1) e 20 (ID: Glyma20g17990.1), dois genes codificadores para UreF, localizados nos cromossomos 02 (ID: Glyma02g44440.1) e 14 (ID: Glyma14g04380.1), e um único gene codificador para UreG, localizado no cromossomo 08 (ID: Glyma08g08970.1). Até o momento, nenhum gene equivalente à UreE bacteriana foi identificado em plantas.

Vale ressaltar que, nos estudos acima, a caracterização do gene foi o principal enfoque e uma caracterização detalhada das proteínas em nível molecular ainda não foi realizada. Por esta razão, as informações a respeito do processo de ativação de ureases vegetais são bastante restritas, quando comparadas ao que se sabe para sistemas bacterianos.



## Objetivos

O conhecimento a respeito da estrutura de ureases vegetais e de suas proteínas acessórias ainda é bastante limitado. Por esta razão, o objetivo desta tese é ampliar este conhecimento, através de estudos estruturais e funcionais da urease de *C. ensiformis* (JBU) e das proteínas acessórias das ureases de soja.

### Objetivos específicos

- Obtenção de cristais de JBU
- Realizar estudos de cristalografia de raios X de JBU;
- Avaliar o perfil de expressão das proteínas acessórias de soja ao longo do desenvolvimento da planta, e associar esse perfil à atividade ureásica dos diferentes tecidos;
- Produzir formas recombinantes das proteínas acessórias das ureases de soja, UreD, UreF e UreG, para utilização na caracterização destas proteínas;
- Caracterizar estrutural e funcionalmente as proteína acessórias UreF e UreG de soja.

As sessões **Materiais e Métodos** e **Resultados e Discussão** estão divididas nos seguintes capítulos:

- I. Estudos cristalográficos da Urease de *Canavalia ensiformis*, JBU
- II. Análise da expressão dos genes das proteínas acessórias das ureases de soja
- III. Clonagem e expressão da proteína acessória UreD de soja
- IV. Caracterização da proteína acessória UreF de soja
- V. Caracterização da proteína acessória UreG de soja

## I. Estudos cristalográficos da Urease de *Canavalia ensiformis* (JBU)

### Materiais e métodos

#### Purificação de JBU

O método de purificação utilizado é o mesmo descrito por Blakeley et al. [1969], com as modificações introduzidas por Follmer et al. [2004a]. Resumidamente, após a moagem, a farinha das sementes de *C. ensiformis* foi delipidada com acetona e clorofórmio, e submetida à extração com tampão fosfato de sódio 20 mM, pH 7,5, EDTA 1 mM e 2-mercaptoetanol 2 mM por 1 h a 4 °C. O resíduo de farinha foi removido por centrifugação (30.000 x g, 20 min, 4 °C), e 28 % (v/v) de acetona (-20 °C) foi adicionada ao sobrenadante e a mistura permaneceu a 4 °C durante a noite. O material foi centrifugado, e o novo sobrenadante recolhido. A concentração de acetona foi elevada para 31,6 % (v/v) e o precipitado removido por centrifugação. O sobrenadante foi dialisado contra o tampão fosfato de sódio 20 mM, pH 7,5, EDTA 1 mM e 2-mercaptoetanol 5 mM. A amostra foi então submetida a uma cromatografia de troca iônica utilizando a resina Q-Sepharose (Amersham-Biotech). A eluição foi realizada em duas etapas, a primeira pela adição de 100 mM de NaCl ao tampão, e a segunda, contendo a fração enriquecida em urease, pela adição de 300 mM de NaCl. A segunda fração foi concentrada e aplicada em coluna de gel-filtração Superose 6 HR 10/30 (Amersham-Biotech). O pico contendo atividade ureásica foi dialisado contra fosfato de sódio 20 mM, pH 7,0, NaCl 0,5 M e submetido à cromatografia de afinidade por metal, em resina IDA-Sepharose (Amersham-Biotech) carregada com o íon  $\text{Co}^{2+}$ . A fração não retida na resina contém JBU. A fração purificada foi dialisada contra HEPES 5 mM, EDTA 1mM e 2-mercaptoetanol 5 mM e utilizadas nos testes de cristalização.

Alternativamente, urease comercial de jack bean tipo III (Sigma Aldrich) foi utilizada. O proteína foi solubilizada em tampão HEPES 5 mM, EDTA 1mM e 2-mercaptoetanol 5 mM e submetida à gel-filtração em coluna Superdex-200.

#### Cristalização

As cristalizações iniciais foram realizadas utilizando o método de varredura por matriz esparsa [Jancarik e Kim, 1991; Hampton Research], através da técnica de “Sitting Drop” de

difusão de vapores, em placas de 24 ou 96 poços. Todas as gotas consistiam de 2 µL de solução de proteína e 2 µL de solução do reservatório. As varreduras foram realizadas em duas temperaturas 25 °C e 4 °C utilizando os seguintes kits comerciais de cristalização: Crystal Screen Hampton Research 1 e 2 , Wizard 1 e 2 Jena Bioscience, Index 1 e 2 Hampton Research, JBScreen Bulks 1 à 10 Jena bioscience, SaltRx Hampton Research, Quick Screen, MembFac, Grid Screen Ammonium Sulfate, Crystal Screen Lite, Crystal Screen Cryo, Emerald Biostructures Cryo I e II, Preciptant Synergy. As condições que apresentavam microcristais ou alguma forma de precipitação organizada foram escolhidas e otimizadas (utilizando a técnica de “hanging drop”) para a produção de cristais com tamanho e qualidade para serem utilizados no difratometro. Detalhes sobre condições de cristalização e técnicas específicas serão encontrados ao longo do texto. Estas tentativas foram realizadas no LNLS (Laboratório Nacional de Luz Sincrotron, Campinas, SP) e no Department of Cell and Systems Biology, University of Toronto, Toronto, Canada.

#### **Coleta de Dados de Difração de Raios X**

Para a coleta de dados, diversas tentativas de congelamento dos cristais foram feitas usando diferentes crioprotetores. Coletas preliminares foram realizadas a 100 K sob fluxo de nitrogênio com a fonte de raios X Rigaku Micro-Max 007HF Microfocus, equipada com um detector Saturn 944+CCD (York University, Canada). Os cristais mais promissores foram utilizados para aquisição de dados na linha SBC19 no Advanced Photon Source (Argonne National Laboratory, Illinois, USA).

#### **Modificação química de resíduos de lisina**

JBU foi quimicamente modificada usando o protocolo publicado por Rayment [1997] e Walter et al. [2006]. Resumidamente, após a gel-filtração inicial, a proteína foi concentrada a 1 mg/ml e dialisada contra tampão HEPES 50 mM, NaCl 200 mM, pH7,5. Após a diálise, foram adicionados à amostra 20 µl de uma solução 1 M de ABC (dimethylamino-borano complex) e 40 µl de uma solução 1 M de formaldeído para cada mililitro de solução de proteína. A proteína foi mantida a 4 °C sob constante e suave agitação por 2 horas. Este procedimento foi repetido e a proteína mantida a 4 °C sob agitação por mais 2 horas. Finalmente, 20 µl de solução 1 M de ABC foram adicionados e a proteína mantida sob as mesmas condições até o dia seguinte. Com o objetivo de verificar a homogeneidade da preparação, bem como eliminar

o excesso de reagentes modificantes do meio, a proteína foi submetida a gel-filtração em coluna Superdex-200.

## **Estado da Arte, Resultados e Discussão**

### **Cristalização de JBU**

Estudos de genômica estrutural em larga escala mostram que, de todas as proteínas que avançam ao estágio de tentativa e cristalização, dois terços não irão cristalizar e, daquelas que cristalizam, metade formam cristais que não podem ser otimizados ao ponto da determinação estrutural. Estimativas são de sucesso em apenas 15% dos casos [<http://www.targetdb.pdb.org>; [http://thesgc.com/structures/target\\_progress.php](http://thesgc.com/structures/target_progress.php)].

Muitas tentativas de obtenção de cristais de boa qualidade para esta proteína foram realizadas, por diferentes grupos nas últimas décadas [Jabri et al., 1992; 1995; Sheridan et al., 2002]. Uma possível explicação para a dificuldade na obtenção de cristais de boa qualidade para JBU reside na presença de mais de uma isoforma de urease (canatoxina e JBU) nas preparações a partir das sementes de *C. ensiformis* [Follmer et al., 2001]. Follmer et al. [2004a], descreve um passo adicional ao protocolo padrão de purificação de JBU [Blakeley et al., 1969], capaz de eliminar a contaminação (cerca de 10%) pela canatoxina. Com esta possibilidade de se obter JBU sem a isoforma contaminante, o que possivelmente levaria a formação de cristais de melhor qualidade, nosso grupo realizou novas tentativas de cristalização de JBU, utilizando as técnicas tradicionais (método da matriz esparsa). Com esta abordagem, 2304 diferentes condições de cristalização foram testadas, variando-se pH, concentração salina, concentração proteica, temperatura e precipitantes, levando a formação de dezenas de cristais (Figura 3). Infelizmente, nossos esforços apenas nos conduziram a cristais incapazes de difratar raios X como os descritos anteriormente na literatura [Jabri et al., 1992; 1995], mostrando que a presença de mais de uma isoforma nas preparações de JBU não constitui a razão para a não obtenção de cristais de boa qualidade para esta proteína.

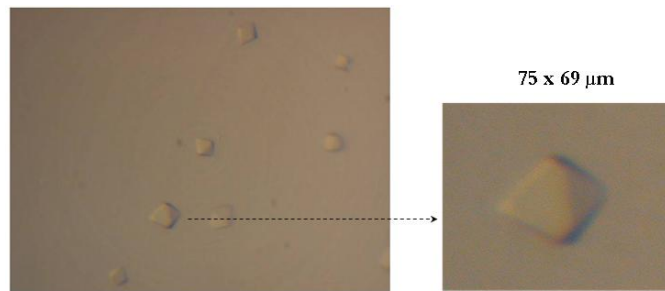


Figura 3: Exemplo de cristais obtidos para JBU, utilizando técnicas tradicionais de cristalografia (detalhes no texto).

Desde a última descrição de tentativas de obtenção da estrutura cristalográfica de JBU [Sheridan et al., 2002], muitos avanços ocorreram no mundo da cristalografia de macromoléculas, incluindo o desenvolvimento de técnicas alternativas específicas para obtenção de bons cristais de proteínas que se mostraram resistentes aos métodos convencionais de cristalização [Chruszcz et al., 2008]. Entre as possíveis abordagens alternativas estão: (i) adição de ligantes ou aditivos [McPherson e Cudney, 2006]; (ii) Proteólise *in situ*; (iii) uso da metilação redutiva de lisinas [Walter et al., 2006]. Tendo JBU se apresentado como uma proteína “difícil”, em termos de cristalização, estas metodologias alternativas foram utilizadas. Devido ao processo laborioso e demorado de purificação de JBU da planta, a proteína de origem comercial foi utilizada nos experimentos seguintes.

- **Preparação da proteína**

Trabalhos anteriores de nosso grupo com JBU de origem comercial mostraram que esta preparação da enzima não é completamente livre de contaminantes, principalmente concanavalina A. Por esta razão, e para garantia da homogeneidade da amostra, um passo adicional de purificação foi realizado. A preparação da proteína foi submetida a cromatografia de gel-filtração em tampão HEPES 5 mM, EDTA 1mM e 2-mercaptoetanol 5 mM, na presença de 300 mM de NaCl. Durante este procedimento, foi descoberto que a utilização de um tampão com alta força iônica era capaz de não apenas aumentar a pureza da proteína, como também separar dois estados oligoméricos da enzima, os estados hexamérico e trimérico. Estes dois estados já haviam sido descritos anteriormente na literatura [Sumner e Somers, 1953; Gorin et al., 1967; Bailey e Boulter, 1969; Contaxis e Reithel, 1971; Dixon et al., 1980; Takishima et al., 1988], no entanto o consenso era de que estas duas formas coexistiam em constante equilíbrio de associação/dissociação, o que nós viemos a descobrir ser uma

conclusão errônea. De acordo com nossos dados, esse processo de associação/dissociação é dependente da força iônica do meio, uma vez que a forma trimérica da proteína é apenas observada na presença de altas concentrações de NaCl (Figura 4). É interessante observar também que essa dissociação induzida por concentração salina parece separar a enzima em dois grupos com diferente grau de interação. Nossos experimentos mostram que a proporção entre hexâmeros e trímeros permanece inalterada apesar do aumento na força iônica do meio (Figura 4A), nos levando a conclusão de que estes dois estados oligoméricos formam dois grupos distintos da mesma proteína, com diferentes graus de interação entre si.

Quando os dois estados oligoméricos são coletados separadamente e re-cromatografados na presença e ausência de NaCl, dois comportamentos distintos foram observados. Em relação ao hexâmero separado na primeira cromatografia, a presença de sal não leva ao aparecimento de trímeros, indicando que não há um equilíbrio constante das formas oligoméricas, como proposto anteriormente. A retirada do sal do sistema não altera este perfil (Figura 4C). Para a forma trimérica, a presença de sal não causa alterações na oligomerização, ou seja, toda a fração continua como trímero. Entretanto, a retirada do sal leva à transição desta forma oligomérica para hexâmeros (Figura 4B). Observando este comportamento diferencial entre as formas oligoméricas de JBU e considerando que a homogeneidade oligomérica é essencial para a formação de cristais de boa qualidade, optou-se pela separação das duas formas nos ensaios de cristalização. É importante mencionar que as condições de cristalização envolvem diferentes forças iônicas, que poderiam levar a separação da proteína em diferentes formas oligoméricas na solução de cristalização caso as duas frações não fossem separadas, comprometendo a formação de cristais de boa qualidade.

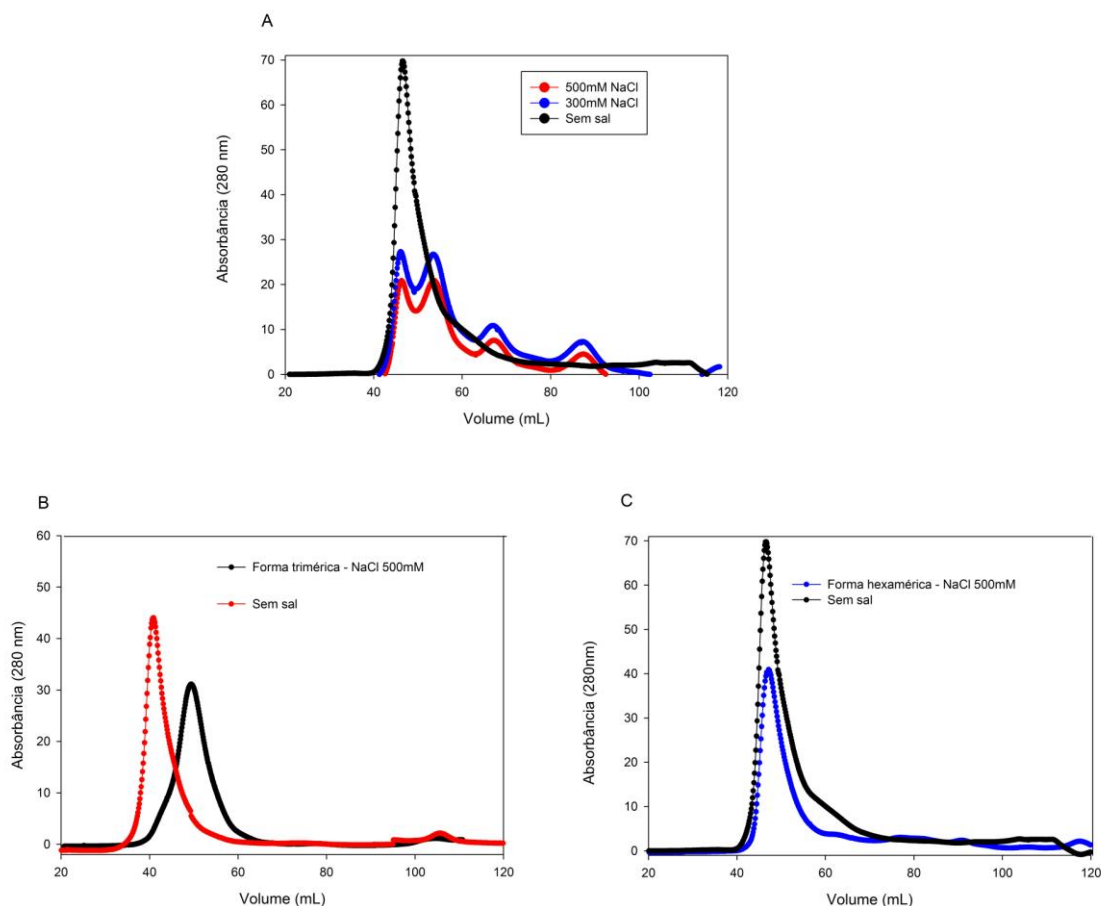


Figura 4: Perfil de gel-filtração de JBU. (A) efeito da concentração salina no estado de oligomerização da proteína. Efeito da remoção do sal em (B) forma trimérica e (C) forma hexamérica de JBU.

- **Co-cristalização na Presença de Ligantes**

Determinados compostos químicos ou pequenas moléculas podem ter um efeito dramático no sucesso da cristalização de uma proteína específica. A classe mais comum de aditivos (como são chamados) e a única para a qual existe alguma racionalização, é aquela das moléculas que, por razões fisiológicas, poderiam estar ligadas a proteína levando a mudanças em suas propriedades físico-químicas ou conformacionais [McPherson e Cudney, 2006]. Tais aditivos podem compreender coenzimas, grupos prostéticos, inibidores, produtos de reações enzimáticas, íons e outras moléculas efetoras. É tido como consenso que a forma ligada da proteína apresenta uma estrutura mais rígida e estável do que a forma não ligada, e geralmente levará a cristalização mais facilmente [McPherson e Cudney, 2006; Vedadi et al., 2006].



Muitos casos têm sido reportados onde pequenas, ou até mesmo grandes, moléculas são inesperadamente observadas fazendo interações intermacromoleculares, que ajudaram, guiaram ou asseguraram a boa formação do cristal. Essas moléculas, algumas vezes, apresentam uma relação fisiológica com a proteína de interesse, mas mais frequentemente, este não é o caso. Elas estão apenas formando interações que são benéficas ao cristal [McPherson e Cudney, 2006; Vedadi et al., 2006]. Diversos trabalhos relataram tentativas de cristalização feitas contra grandes bibliotecas de pequenas moléculas na busca da cristalização de proteínas difíceis [McPherson e Cudney, 2006; Vedadi et al., 2006; <http://www.ccp4.ac.uk/ccp4bb.php> (Collaborative Computational Project, 1994)].

Alguns estudos já foram realizados relatando os benefícios de determinados ligantes para a estabilidade de JBU. Baseado nesses estudos, decidimos investigar o efeito da presença destes aditivos na cristalização de JBU. A lista dos aditivos utilizados neste trabalho é apresentada abaixo.

- **Íons metálicos** –  $\text{Ni}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Cu}^{2+}$  [Zaborska et al., 2004; 2001]
- **Flavonoides** – Silibina e Morina [Illeova et al., 2003; Shabana et al., 2010]
- **Detergentes** – SDS (em concentrações submicelares) [Hirai et al., 1993]
- **Substratos** – Uréia e hidroxíuréia [Bennett e Wren, 1977]
- **Inibidores** – Ácido aceto hidroxâmico [Benini et al., 2000; Musiani et al., 2001]

A presença destes aditivos na solução de cristalização levou ao surgimento de diversos cristais de proteína (Figura 5), no entanto estes cristais apresentavam as mesmas características dos cristais previamente descritos para JBU, com tamanho e/ou qualidade insuficientes para utilização em estudos estruturais.

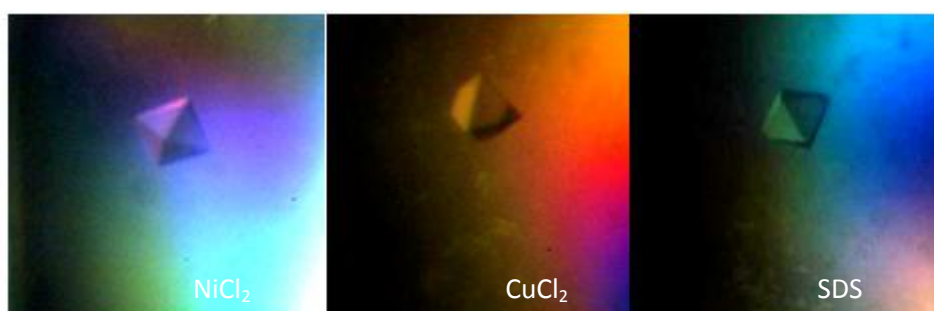


Figura 5: Exemplos de cristais obtidos pela cristalização de JBU na presença de aditivos. Os aditivos referentes a cada cristal estão indicados na figura.

- **Proteólise *in situ***

Por muitos anos, cientistas têm explorado o fato de que fragmentos ou domínios de proteínas geralmente cristalizam melhor, ou formam cristais que difratam com maior qualidade, comparados a proteína intacta. Também um fato sobre cristalografia de raios X é que a maior parte das estruturas depositadas em bancos de dados compreendem polipeptídios com menos de 30 kDa [Christendat et al., 2000; Edwards, 2000; Koth et al., 2003]. No início dos estudos com cristalografia de raios X, fragmentos protéicos eram preparados em larga escala a partir da proteína purificada e então cristalizados. Este método se mostrou eficiente em alguns casos [Huber et al., 1976; Journak et al., 1980], mas a ampliação do uso desta técnica é bastante limitada, devido à necessidade de purificação de grandes quantidades da proteína intacta e pela dificuldade adicional da purificação dos fragmentos derivados da proteólise, de forma homogênea [Dong et al., 2007]. Estas dificuldades foram superadas com o advento da produção de proteínas recombinantes e da espectrometria de massas aplicada à proteínas, possibilitando que a proteólise pudesse ser realizada em pequena escala, os fragmentos identificados por espectrometria de massa e posteriormente produzidos em larga escala de maneira recombinante. No entanto, devido ao complicado processo de ativação das ureases (principalmente de origem vegetal), este procedimento não se mostra adequado a estas proteínas.

A história nos mostra que fragmentos de proteínas também têm sido cristalizados “por sorte”. Na maior parte dos casos, a proteína purificada ou a solução de cristalização foram contaminados com quantidades mínimas de proteases e a proteólise ocorreu durante a cristalização. Existem diversos casos documentados da ocorrência deste fenômeno [Campbell et al., 2002; Sawaya et al., 1994] e, provavelmente, muitos mais não documentados. Devido a estas histórias de sucesso, recentemente a adição de pequenas quantidades de proteases à gota de cristalização tem sido explorada e se mostrou bem sucedida para diversas proteínas [Bai et al., 2007; Gaur et al., 2004; Johnson et al., 2006; Mandel et al., 2006; Taneja et al., 2006; Dong et al., 2007; Wernimont e Edwards, 2009].

O princípio por trás do método de proteólise *in situ* foi gerado a partir da simples observação de que proteínas grandes geralmente são compostas por múltiplos domínios. Em cristalografia, a heterogeneidade conformacional devido ao movimento destes domínios leva a formação de cristais de baixa qualidade [Koth, 2003]. O objetivo da proteólise *in situ* é a eliminação de determinadas regiões interdomínios com alta flexibilidade e, portanto, mais suscetíveis a proteólise do que domínios bem empacotados, permitindo que estes domínios

menores, e conseqüentemente mais propensos a formação de cristais de boa qualidade, sejam liberados em solução [Koth et al., 2003]. No entanto, é preciso garantir que a protease de escolha não seja capaz de hidrolisar completamente a proteína de interesse. Isso pode ser alcançado por testes prévios de hidrólise com diferentes proteases e análises por SDS-PAGE [Koth et al., 2003; Dong et al., 2007].

Observações anteriores do nosso grupo mostraram que JBU é surpreendentemente resistente a proteólise por tripsina, geralmente levando à formação de apenas dois fragmentos (Figura 6), de forma que esta foi nossa protease de escolha.

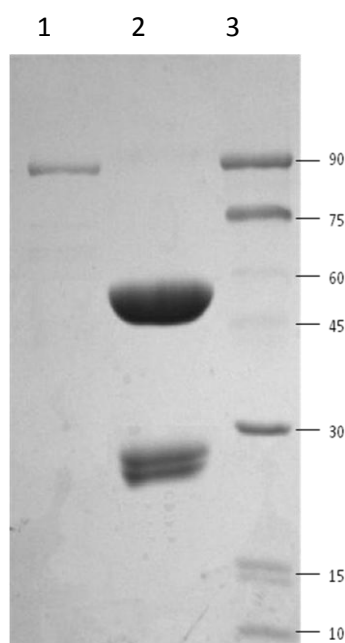


Figura 6: SDS-PAGE da hidrólise de JBU por tripsina (1:500). A hidrólise foi realizada a 37 °C, por 18 h. Canaletas 1: JBU intacta; 2: JBU tripsinizada; 3: padrão de massa molecular (indicado em kDa).

Tripsina foi adicionada à gota de cristalização em duas concentrações, 1:500 e 1:1000 (m/m) em relação à concentração de JBU. A concentração de 1:500 mostrou-se mais promissora, gerando cristais com características únicas nas seguintes condições:

- 0,1 M Bicina pH 9,0, 1,2 M  $(\text{NH}_4)_2\text{SO}_4$ , tripsina (1:500)
- 0,2 M  $(\text{NH}_4)_2\text{SO}_4$ , 0,1 M Tris-Base pH 8,5, 12% glicerol, tripsina (1:500)

No entanto, os cristais formados (Figura 7) utilizando esta abordagem necessitam de tempos muito longos (alguns meses) para sua formação e se mostraram extremamente instáveis. Esta característica impossibilitou que quaisquer dos cristais obtidos fossem testados em um difratômetro, uma vez que os mesmos dissolviam facilmente ao menor distúrbio da

solução de cristalização. Possivelmente, refinamentos das condições de cristalização poderiam levar à formação mais rápida e estável de cristais, e estas condições podem ser utilizadas em tentativas futuras.



0,1 M Tris pH 8,5  
0,2 M  $(\text{NH}_4)_2\text{SO}_4$   
12% glicerol  
Tripsina (1:500)



0,1 M Bicina pH 9,0  
1,2 M  $(\text{NH}_4)_2\text{SO}_4$   
Tripsina (1:500)

Figura 7: Cristais obtidos a partir da técnica de proteólise *in situ*. As condições de cristalização referentes aos cristais apresentados estão indicadas na figura.

- **Modificação Química de Resíduos de Lisina**

A modificação química de aminoácidos, ao longo da história da bioquímica, se mostrou uma ferramenta valiosa na caracterização de proteínas, no entanto até os últimos anos, não havia sido muito empregada para a obtenção de cristais de proteínas de maior qualidade. Isso se deve a pouca especificidade da maior parte das reações de modificação química, o que acabou gerando a crença de que este tipo de abordagem poderia introduzir ainda mais heterogeneidade às amostras, dificultando a cristalização. No entanto, uma classe de reações, a alquilação redutiva, pode ser empregada de maneira altamente específica e tem sido de grande valia na alteração de condições de cristalização [Rayment et al., 1997; 1993; Rypniewski et al., 1993].

Da mesma forma que as múltiplas conformações adotadas por proteínas multidomínios, ou movimentos de regiões flexíveis de proteínas, como alças e porções N e C terminais, tem efeitos negativos na formação de cristais altamente ordenados, pequenos movimentos e/ou contatos cristalinos, devidos às cadeias laterais flexíveis de amino ácidos,

podem ser igualmente disruptivos para a formação de cristais de alta qualidade. Diversos estudos já mostraram que os principais resíduos de amino ácidos envolvidos na má formação cristalina são resíduos de lisina e glutamina [Derewenda e Vekilov, 2006; Longenecker et al., 2001; McElroy et al., 1992; Mateja et al., 2002]. Neste contexto, o uso da alquilação redutiva aplicada à modificação química de resíduos da superfície, levando à redução da entropia de superfície das proteínas, apresenta-se como uma opção bastante vantajosa em relação a outros métodos, como a mutação sítio dirigida. Entre as vantagens deste método estão: (i) passos de clonagem, expressão e purificação são eliminados; (ii) a modificação da proteína madura elimina a possibilidade da indução de enovelamento incorreto da cadeia polipeptídica [Rypniewski et al., 1993]; (iii) apenas os resíduos expostos na superfície da proteína serão modificados, não aqueles que estão protegidos em seu interior e que provavelmente participam em interações inter- ou intra-proteína [Walter et al., 2006]. No entanto, é preciso ter cuidado no emprego desta técnica, para que resíduos importantes para a atividade da proteína alvo não sejam afetados e, portanto, testes preliminares são cruciais.

Conforme dados prévios [Real-Guerra, 2007], a modificação química de resíduos de lisina da superfície de JBU não causam qualquer modificação da atividade enzimática de JBU, provando que a proteína permanece totalmente funcional, e tornando válida a aplicação desta metodologia na obtenção de cristais de JBU. Através da metilação redutiva de resíduos de lisina (Figura 8) da superfície de JBU, foi possível a obtenção de cristais em diversas condições e com características distintas das previamente reportadas.

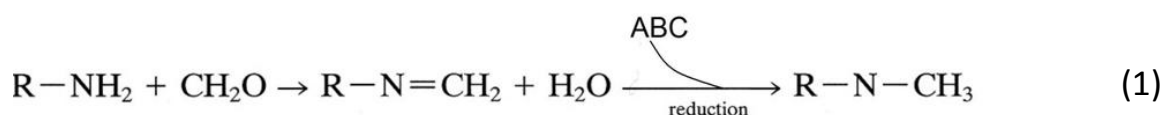


Figura 8: Reação de metilação redutiva de lisinas. (1) Indica a inicial formação de uma base de Schiff pela rápida reação inicial do grupamento amina com formaldeído, e a subsequente redução à amina secundária. (2) A sequência da reação forma, como produto final, um resíduo de lisina dimetilada.

Todos os cristais já obtidos para JBU apresentam características idênticas, tratando-se de bipirâmides tetragonais; no entanto, os cristais obtidos para a JBU modificada apresentam forma completamente distinta. Isso é um bom indicativo de alterações no padrão

de cristalização de JBU e, tendo em vista a pouca qualidade dos cristais anteriores, isso pode ser visto como um aspecto positivo. As condições consideradas mais promissoras foram utilizadas para refinamentos posteriores. As melhores condições estão apresentadas na Tabela 1 e os cristais obtidos estão apresentados na Figura 9.

Tabela 1: Refinamento das condições iniciais de cristalização de JBU modificada.

Condição Inicial		Condição de Refinamento		N°
Sal	Precipitante	Sal	Precipitante	
		Tiocianato de Potássio 0,40 M	PEG 3350 20%	1
		Tiocianato de Potássio 0,70 M	PEG 3350 20%	2
Tiocianato de Potássio 0,2 M	PEG 3350 20%	Tiocianato de Potássio 0,40 M	PEG 3350 19%	3
		Tiocianato de Potássio 0,60 M	PEG 3350 18%	4
		Tiocianato de Potássio 0,25 M	PEG 3350 18%	5
Iodeto de Potássio 0,2 M	PEG 3350 20%	Iodeto de Potássio 0,60 M	PEG 3350 24%	6
		Iodeto de Potássio 0,40 M	PEG 3350 15%	7

Os números indicam o cristal correspondente na Figura 9.

Os cristais obtidos foram utilizados para difração e se mostraram capazes de difratar raios X em uma fonte de laboratório com uma resolução de 6 Å, de maneira estável, utilizando 30 % de etilenoglicol ou glicerol como agente crioprotetor, o que até aquele momento jamais havia sido obtido para JBU. Quando os mesmos cristais foram irradiados em uma fonte de luz síncrotron (Advanced Photon Source - Argonne National Laboratory) a resolução obtida foi aumentada para 4,8 Å. Esses resultados mostram que, tanto a separação dos oligômeros como a modificação de resíduos de lisina da superfície de JBU, tiveram um grande efeito na obtenção de cristais de maior qualidade e, portanto, adequados para estudos de cristalografia de raios X. Mais refinamentos seriam necessários até a obtenção de cristais com resolução suficiente para a elucidação da estrutura da proteína. No entanto, no mês seguinte à obtenção destes resultados, uma comunicação de que a resolução da estrutura cristalográfica de JBU estava em andamento foi publicada por Balasubramanian e Ponnuraj [2009], seguida da publicação da estrutura em 2010.

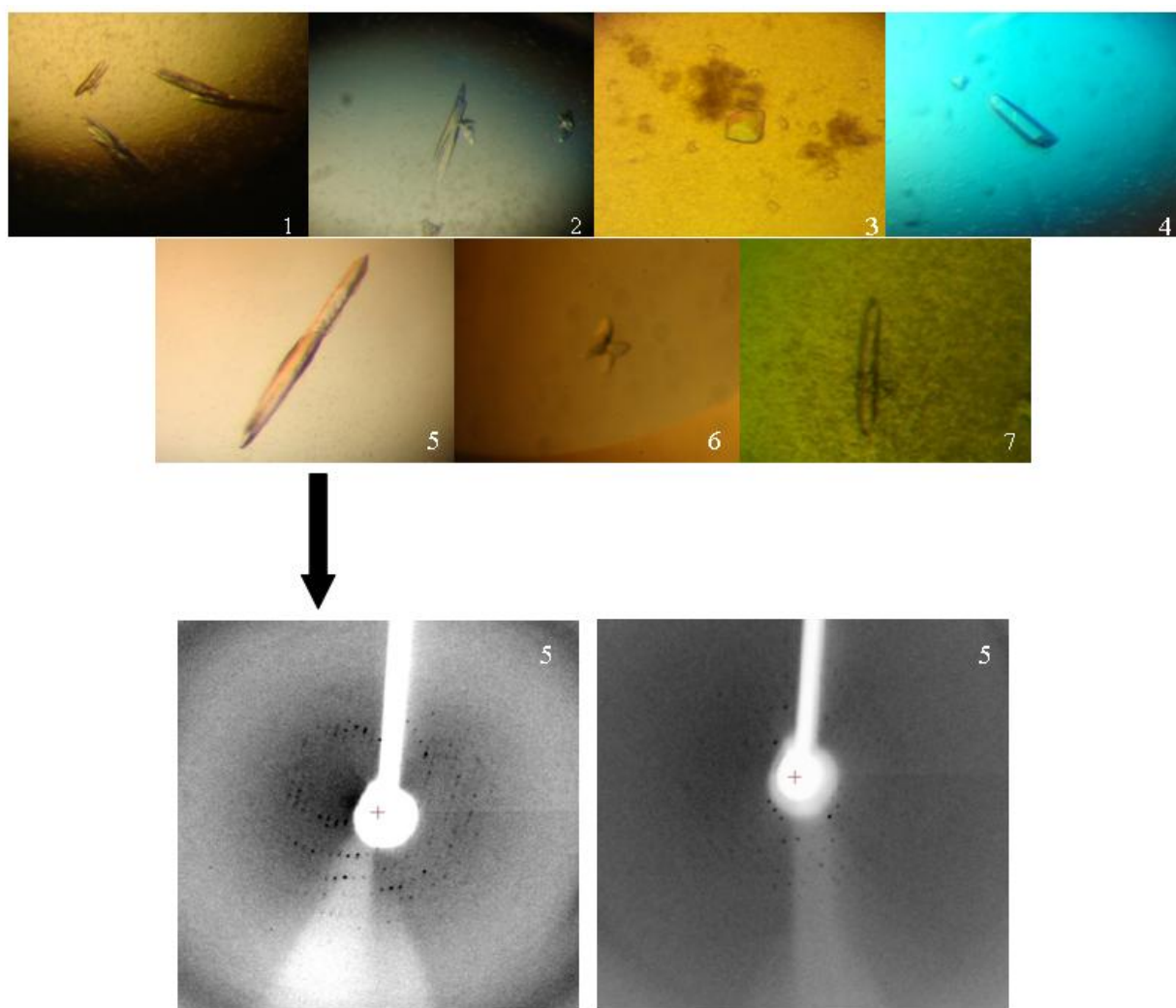


Figura 9: Cristais de JBU modificada quimicamente. Os números referem-se às condições descritas na Tabela 1. O cristal número 5 foi utilizado para coleta de dados em fonte de laboratório e de luz síncroton, gerando os padrões de difração mostrados (esquerda e direita, respectivamente).

Em sua abordagem, Balasubramanian e Ponnuraj realizaram modificações no protocolo de purificação descrito por Blakeley et al. [1969] pela adição de um passo extra de purificação, usando uma cromatografia de troca iônica em DEAE celulose. Com esta abordagem, obtiveram a proteína completamente pura e monodispersa na forma de um hexâmero. Esta observação vai de encontro ao previamente descrito por Sheridan et al. [2002] que observaram uma polidispersão da proteína em monômeros, trímeros e hexâmeros. Propõe-se que a obtenção de uma única forma oligomérica de JBU tenha sido a razão do sucesso na cristalização, no entanto é importante ressaltar que as condições de cristalização

utilizadas pelos autores são diferentes daquelas previamente reportadas na literatura, gerando cristais capazes de difratar raios X de maneira estável com resolução de até 2,05 Å.



## II. Análise da expressão dos genes das proteínas acessórias das ureases de soja (*Glycine max*)

### Materiais e métodos

#### PCR em tempo real

As alíquotas de cDNA dos diferentes tecidos utilizados nos experimentos de PCR em tempo real foram cedidas pela Dra. Maria Helena Zanettini, do Departamento de Genética da Universidade Federal do Rio Grande do Sul. Os tecidos avaliados estão descritos na Tabela 2. O cultivar de soja utilizado foi o MG/BR-46 Conquista.

Tabela 2: Descrição dos tecidos utilizados para análise do perfil de expressão

Tecido	Descrição
	Semente germinada em papel filtro por 24h (quebra de dormência)
Raíz jovem	Raíz de plântulas de 2 semanas
Caule jovem	Caule de plântulas de 2 semanas
Folha jovem	Folha de plântulas de 2 semanas
Raíz	Raíz de plantas adultas
Caule	Caule de plantas adultas
Folha	Folha de plantas adultas

Antes da realização da PCR, as alíquotas de cDNA foram diluídas 1:100 em água destilada autoclavada. A amplificação de um único produto por cada par de primers foi testada por PCR e analisada em gel de agarose. A reação de PCR em tempo real (25 µL) foi composta de 12,5 µL de cDNA molde diluído, tampão para PCR 1x (Invitrogen), 2,5 mM MgCl<sub>2</sub>, 25 µM dNTP (Fermentas), 2,5 µL de uma solução de SYBR-Green 1:1000 (Molecular Probes Inc.), 300 mU de Taq DNA Polimerase Platinum (Invitrogen) e 0,1 µM de cada primer. A reação de PCR em tempo real foi realizado no Termociclador StepOne (Applied Biosystem). As condições do PCR foram as seguintes: 5 min 94 °C; 40 ciclos de 15 s 94 °C, 10 s a 59 °C e 15 s a 72 °C; 1 min a 60 °C. Uma análise de curva de desnaturação foi realizada no final da PCR, na faixa de 55-99 °C, com incrementos de 0,1 °C a cada segundo. Todas as reações foram realizadas em quadruplicatas, e reações sem cDNA molde foram utilizadas como controles

negativos. As amplificações foram realizadas utilizando os primer gene-específicos mostrados na Tabela 3. Como controles internos para normalização da quantidade de mRNA presente em cada amostra, primers para os genes da proteína F-box e da enzima degradadora de insulina (Metaloprotease) foram utilizados. Estes genes já foram utilizados previamente como eficientes controles internos [Jian et al., 2008; Libault et al., 2008; Schimidt e Parrot, 2001]. A eficiência dos pares de primers foi verificada através da construção de uma curva padrão com diferentes quantidades do produto específico de cada par. A análise dos dados de expressão foi feita por quantificação comparativa, utilizando o método  $2^{-\Delta\Delta Ct}$  [Pffal, 2001; Livak e Schmittgen, 2001].

Tabela 3: Primers utilizados no PCR em tempo real.

Gene	Orientação	Sequência (5' - 3')	Eficiência
UreD02	Forward	ATGTCGTATGGTGCTGTTGTGACAC	1,90
	Reverso	TGAACCCATTTCCATTTCTCTGACTGC	
UreD20	Forward	GGTGCAGTATGTCCAGAATCGTGTG	1,89
	Reverso	CTTAGGACCAAAGGCACTGCAAGA	
UreF02	Forward	TGTTGCCCAATTTAGAACTTGGCAT	2,07
	Reverso	GGACCTCAGAAAACACAGCTGAGG	
UreF14	Forward	ATGAAGGACACAACAGACCTTGCTCA	1,88
	Reverso	TTAGACACTAAGTGGGACTGAACTGTTGC	
UreG	Forward	CTGATTTGGCAGTCATGCAGCG	1,89
	Reverso	GCTTATCCCTGTGGCTGCTTCC	
UreEs	Forward	ACCCAACACGCCCTTTAACACTCAATACTA	1,91
	Reverso	ATCGTTCCTTCTTATCCTTGAACATGC	
UreUb	Forward	CTGTGAAGAATGTGAGGAAGCTCACCA	1,91
	Reverso	AGTCTTGGCTGCGGTGCAG	
F-box	Forward	AGATAGGGAAATGTTGCAGGT	1,9
	Reverso	CTAATGGCAATTGCAGCTCTC	
Metaloprotease	Forward	ATGAATGACGGTTCATGTA	1,80
	Reverso	GGCATTAAAGGCAGCTCACTCT	

## Atividade ureásica

Os tecidos foram coletados nos intervalos de tempo escolhidos, congelados, liofilizados e macerados com gral e pistilo. A extração de proteínas foi feita em tampão fosfato de sódio 10 mM pH 7,5, NaCl 150 mM, por 2,5 h a 4 °C, sob agitação. O extrato foi então centrifugado por 30 min a 12.000 x *g*, e o sobrenadante dialisado contra o tampão de extração. A quantificação proteica foi feita por absorvância em 280 nm. A atividade ureásica dos tecidos estudados foi avaliada através do método colorimétrico do fenol-nitroprussiato-hipoclorito [Weatherburn, 1967], e uma curva padrão com sulfato de amônio foi utilizada para quantificação do produto formado.

## Resultados e Discussão

Poucos estudos foram realizados até hoje visando a avaliação da distribuição das proteínas acessórias e ureases em diferentes tecidos vegetais e ao longo do desenvolvimento da planta. Nos trabalhos existentes, a expressão de uma única proteína é avaliada. Neste trabalho, analisamos os níveis de mRNA de todas as proteínas acessórias de soja, assim como das ureases ubíqua e embrião-específica. Os dados preliminares destas avaliações estão apresentados nas Figuras 10 e 11.

Em uma primeira análise, os dados são mostrados relativos à expressão do gene de interesse na semente, visando avaliar a variação da expressão daquele gene ao longo do desenvolvimento da planta (Figura 10). Não observamos um padrão comum de acúmulo do mRNA para todos os genes ao longo do desenvolvimento da planta. No entanto, um padrão pode ser observado para *GmUreF14* e a urease ubíqua, que apresentaram a mesma cinética de expressão nos diferentes tecidos, sendo seu maior acúmulo observado na raiz adulta, seguida pela raiz jovem. Os genes *GmUreD20* e *GmUreF02* apresentaram um perfil bastante diferenciado dos outros genes, sendo expressos em altos níveis somente na semente e na raiz adulta.

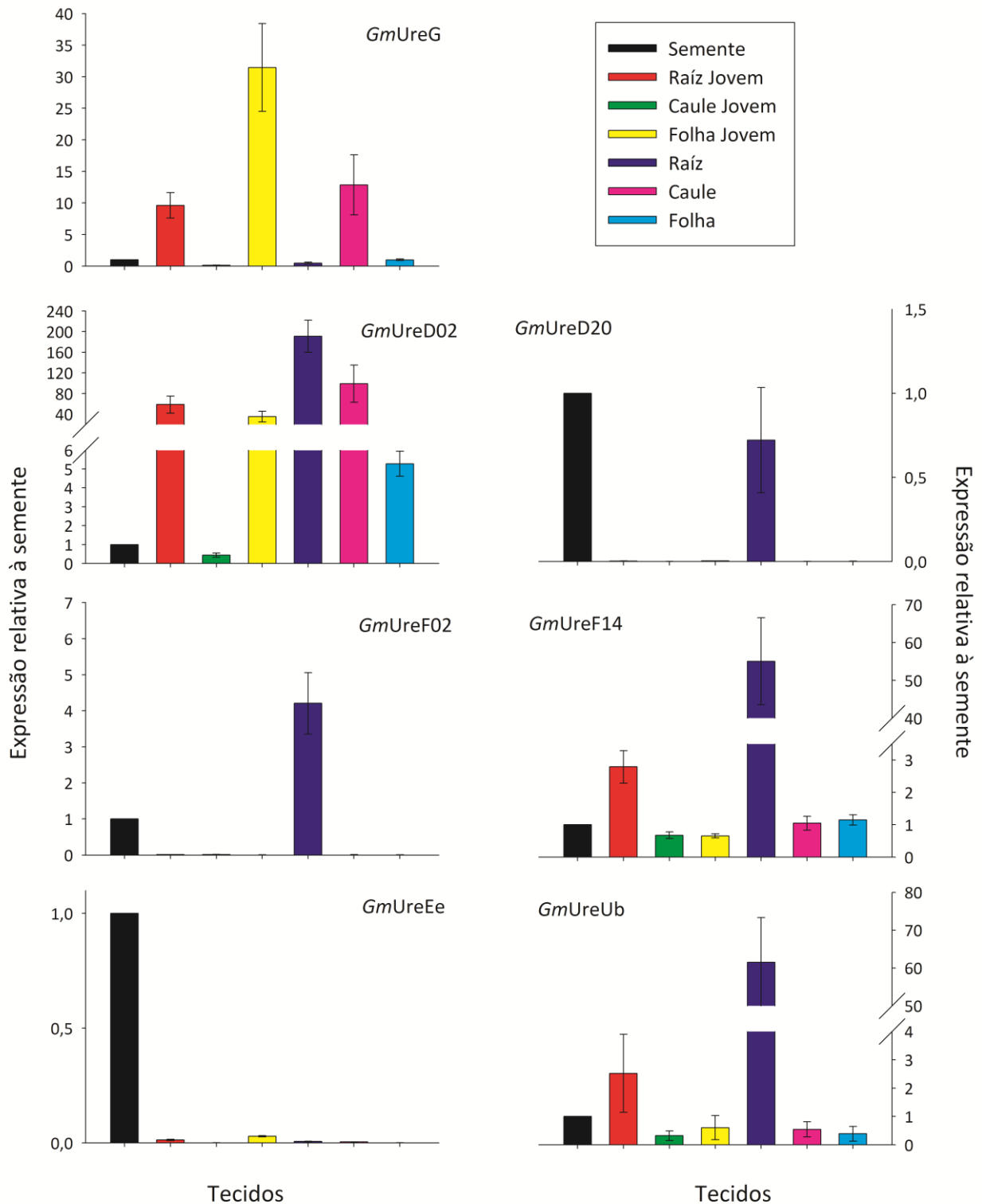


Figura 10: Análise da expressão dos genes das proteínas acessórias e ureases por PCR em tempo real ao longo do desenvolvimento da soja. A expressão nos diferentes tecidos é apresentada em relação aos níveis de expressão na semente. *GmUreEe* = urease embrião-específica, *GmUreUb* = Urease ubíqua. Resultados são apresentados como média  $\pm$  erro padrão.

A urease embrião-específica foi encontrada em níveis bastante baixos em outros tecidos além das sementes, como raiz e folha jovens (Figura 10). Em estudos realizados por Torisky e Polacco [1990], utilizando Western blots, foi descrita a presença da urease embrião-específica em raízes jovens e também demonstrado que esta urease é perdida ao longo do desenvolvimento da raiz, onde as plântulas de 10 e 20 dias possuem apenas traços desta proteína. Como observamos aqui, os níveis de expressão em tecidos outros além das sementes são bastante baixos, confirmando a hipótese levantada por Torisky e Polacco [1990] de que esta proteína seja sintetizada continuamente na raiz, mas em níveis tão baixos que as técnicas empregadas naquele trabalho não possuíam sensibilidade suficiente para detectá-los.

Na segunda análise dos dados realizada, os níveis de expressão são mostrados em relação a um dos genes de controle interno (F-Box) (Figura 11), o que possibilita uma comparação dos níveis de expressão dos genes nos diferentes tecidos.

Em batata, UreG - mRNA e proteína - foi detectada em todos os tecidos avaliados (folhas, caule, tubérculo, estolões e raiz) [Witte et al., 2001]. Um estudo proteômico com *C. ensiformis* [Demartini et al., 2011] detectou a presença de UreG nas sementes quiescentes e durante os primeiros 5 dias de germinação. Em soja, Freyermuth et al. [2000] não detectou a proteína Eu3 (UreG) em extratos brutos de tecidos não embriogênicos. Estes resultados contrastam com os níveis de mRNA de *GmUreG* observados em nosso trabalho, onde tecidos jovens apresentaram níveis de mRNA de *GmUreG* mais elevados do que a semente, sendo a folha jovem o tecido onde a expressão é maior.

Em relação as isoformas de UreD, *GmUreD20* é majoritariamente expressa em sementes e raiz adulta. Enquanto na raiz jovem a expressão de *GmUreD02* é apenas 3x maior do que a de *GmUreD20*, na raiz adulta ocorre um aumento significativo da expressão de *GmUreD20*, chegando a níveis 20x maiores do que *GmUreD02*. A expressão de ambas isoformas é similar em folhas jovens e adultas. No caule, uma aparente inversão da expressão parece ocorrer ao longo do desenvolvimento: no tecido jovem, observa-se uma expressão cerca de 10x maior de *GmUreD20* em relação à *GmUreD02*, enquanto no tecido adulto *GmUreD02* é 10x mais abundante que *GmUreD20*. É interessante observar que em todos os tecidos, com exceção da semente, a relação de expressão entre *GmUreUb* e *GmUreEe* é a mesma observada entre *GmUreD20* e *GmUreD02* (embora com razões diferentes): nos tecidos onde *GmUreUb* é mais abundante do que *GmUreEe*, *GmUreD20* é mais abundante que *GmUreD02*. Especula-se que ambas as proteínas *GmUreD* sejam funcionais [Polacco,

comunicação pessoal] e é possível que a presença de duas isoformas esteja relacionada a um controle diferenciado da ativação de cada isoforma de urease ao longo do desenvolvimento da planta. No entanto, estudos mais aprofundados, possivelmente envolvendo mutantes de soja, serão necessários para a confirmação desta hipótese.

Para *GmUreF02* e *GmUreF14*, observamos um padrão onde, em todos os tecidos, os níveis de *GmUreF02* são maiores do que os de *GmUreF14*, embora em raiz jovem e folha adulta os níveis sejam comparáveis. Esta relação corrobora a hipótese de que a forma predita como funcionalmente mais relevante é *GmUreF02*. Estes dados contrastam com os dados de expressão de ESTs (Microarrays) de soja, compilados no site Soybase (<http://soybase.org>), que indicam que os transcritos de *GmUreF14* são 5 vezes mais comuns que os de *GmUreF02* em tecidos como folhas, raízes, nódulos e vagens. No entanto, é importante salientar que, mesmo em níveis baixos, *GmUreF14* é encontrada em todos os tecidos avaliados. Assumindo que esta isoforma não seria funcional [Polacco, comunicação pessoal], a razão para a transcrição deste gene, em todos os tecidos da planta, é uma questão que ainda precisa ser respondida.

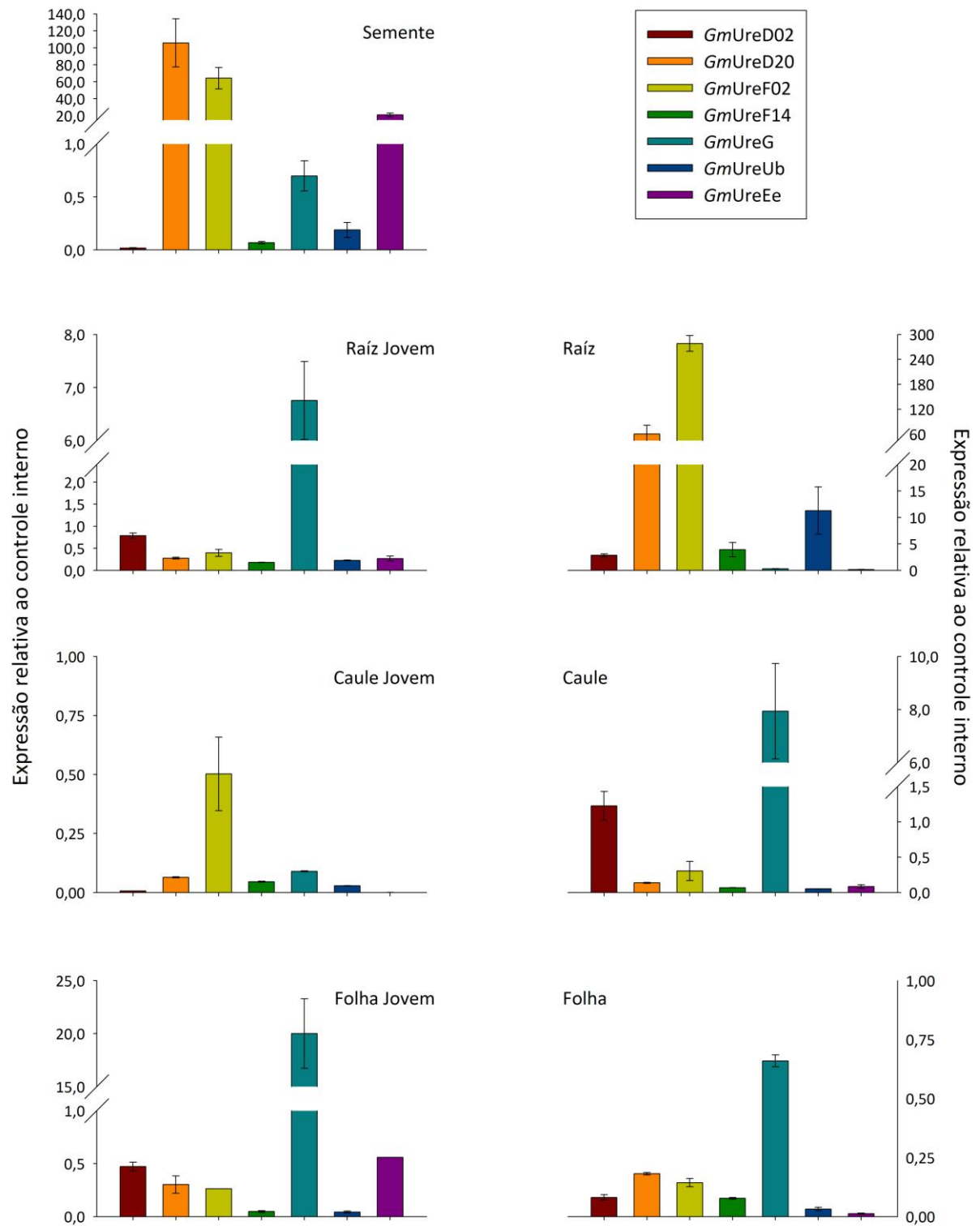


Figura 11: Análise da expressão dos genes das proteínas acessórias e ureases por PCR em tempo real ao longo do desenvolvimento da soja. A expressão nos diferentes tecidos é apresentada em relação aos níveis de expressão do gene F-Box (controle interno). *GmUreEe* = urease embrião-específica, *GmUreUb* = Urease ubíqua. Resultados são apresentados como média  $\pm$  erro padrão.

Podemos observar que a raiz das plantas adultas é o tecido onde a expressão de um maior número de genes está em seus níveis mais elevados, com a exceção de *GmUreG* e da urease embrião-específica. Em contraste, o caule jovem apresenta níveis baixos da expressão de todos os genes analisados. Buscando avaliar a relevância fisiológica dos níveis de expressão dos genes estudados, a atividade ureásica dos diferentes tecidos foi analisada (Figura 12).

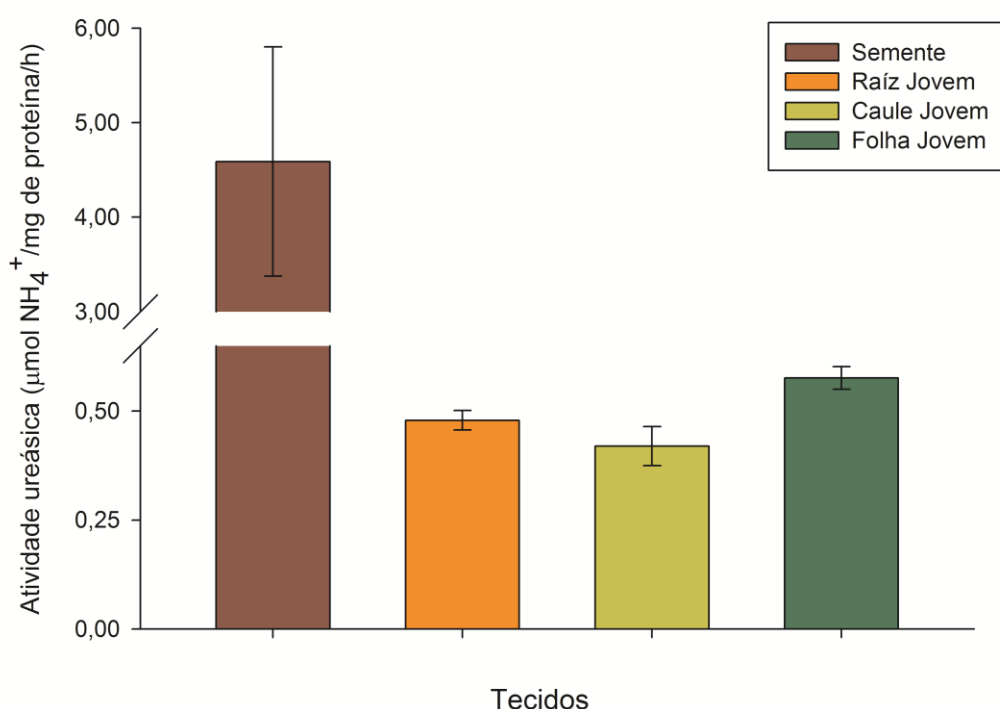


Figura 12: Atividade ureásica dos diferentes tecidos de soja. Resultados são apresentados como média  $\pm$  erro padrão.

Observamos que a semente apresenta uma atividade cerca de 10x maior do que os tecidos jovens. Comparando estes dados com os dados apresentados na Figura 11, uma possível relação entre os níveis de expressão da *GmUreD20* e *GmUreF02* e a atividade pode ser estabelecida. Na semente, estes genes são altamente expressos, enquanto nos tecidos jovens, a expressão é bastante reduzida. A atividade ureásica dos tecidos adultos ainda não foi avaliada, impedindo uma análise mais abrangente da relação dos níveis de expressão e a atividade ureásica.



Witte et al. [2001] avaliaram os níveis da atividade ureásica em diferentes tecidos de batata (*Solanum tuberosum*), mas não encontrou uma boa correlação com os níveis de expressão de UreG (a única proteína acessória para qual os níveis foram avaliados). Isto pode refletir a complexidade da dinâmica entre as proteínas acessórias e suas ureases, além de ressaltar que muito sobre o processo de ativação destas proteínas, principalmente em plantas, ainda permanece desconhecido.

É importante mencionar que os níveis de mRNA nem sempre equivalem aos níveis de proteína. Diferenças na tradução podem afetar esta relação, e a análise dos níveis protéicos é essencial para uma conclusão mais apurada sobre a relação entre os níveis das proteínas acessórias e a atividade ureásica observada nos diferentes tecidos. A utilização de anticorpos específicos para cada proteína, numa abordagem envolvendo Western blots ou ELISA, ou de ferramentas de proteômica, envolvendo análises por espectrometria de massas, poderão ajudar nesta área de estudo.

### III. Clonagem e expressão da proteína acessória UreD de soja (*Glycine max*)

#### Materiais e métodos

##### Clonagem do gene *GmUreD20*

O RNA total, extraído de brotos de soja (5 dias de germinação) utilizando o reagente de Trizol (Invitrogen), foi empregado para síntese de cDNA através do kit iScript (Bio-Rad) com primers Oligo(dT), de acordo com as recomendações do fabricante. Primers específicos para a amplificação do cDNA completo do gene *GmUreD02* (forward 5'-GGCGGCATATGGAAATGGGTTTCAGTGGTGGTT-3'; reverso 5'-CGCCAGATCTTTATCTGTATGGTGGCACCCC-3') foram desenhados com base na sequência disponível no site Phytozome (<http://www.phytozome.net/>). O cDNA completo de *GmUreD20* foi amplificado por PCR, nas seguintes condições: desnaturação inicial de 3 min a 94 °C; 35 ciclos de 30 s a 94 °C, 30 s a 57 °C e 60 s a 72 °C; extensão final de 10 min a 72 °C. A amplificação foi analisada por gel de agarose (1,2 %), revelado com brometo de etídio (10 % [v/v]). O produto de PCR foi purificado com o kit Illustra GFX (GE Lifesciences) de acordo com protocolo fornecido pelo fabricante. O produto purificado foi inserido no vetor pGEM-T Easy (Promega) utilizando-se a Ligase T4 (Promega), numa incubação por 18 h a 4 °C. Esta construção foi utilizada para transformação de *Escherichia coli* DH5 $\alpha$  por choque térmico. Bactérias transformadas foram plaqueadas em LB-ágar, contendo ampicilina (100  $\mu$ g/mL) e incubadas por 18 h a 37 °C. Colônias foram selecionadas para crescimento em meio LB líquido contendo ampicilina (100  $\mu$ g/mL) até o dia seguinte a 37 °C. Uma alíquota de 5  $\mu$ L da cultura de 18 h foram adicionados a 45  $\mu$ L de água destilada, fervidos por 5 min e centrifugados por 1 min a 12.000 x *g*. O sobrenadante foi utilizado como molde para PCR com primers gene-específicos, para seleção de transformantes positivos. A cultura dos transformantes positivos foi centrifugada por 10 min a 5.000 x *g* para recolhimento das células. Os plasmídeos contendo o inserto de interesse foram purificados através do kit Wizard Plus SV Minipreps (Promega), de acordo com as especificações do fabricante, e armazenados a -20 °C.

## Expressão de *GmUreD20*

O cDNA completo do gene *GmUreD20* foi retirado do vetor pGEM-T Easy por clivagens sequenciais com as enzimas de restrição NdeI e BglII. O inserto foi separado do vetor em gel de agarose (1,2 %) e purificado com o kit Illustra GFX. O inserto purificado foi ligado pela enzima ligase T4 ao vetor de expressão pET15b (Novagen), previamente clivado com as enzimas de restrição BamHI e NdeI. Esta construção foi utilizada para transformação de *Escherichia coli* BL21 - Codon Plus por choque térmico. Transformantes positivos foram selecionados por PCR com primers gene-específicos. A inserção correta do inserto no vetor foi confirmada por sequenciamento de DNA.

Um teste de expressão foi realizado, inoculando-se uma colônia positiva em 2 mL de meio LB contendo 100 µg/mL de ampicilina e 40 µg/mL de cloranfenicol. Quando a densidade óptica da cultura atingiu 0,6  $A_{600}$ , a expressão de *GmUreD20* foi induzida com IPTG (800 µM) a 37 °C, agitação de 200 rpm, por 2 h. Células foram coletadas por centrifugação (10 min a 5.000 x g), lisadas por fervura e o extrato celular total foi analisado por SDS-PAGE.

## Resultados e Discussão

Como visto anteriormente, UreD é uma das proteínas acessórias para a qual a quantidade de informação disponível é bem limitada. Até hoje nenhuma UreD foi obtida na forma solúvel isoladamente. Recentemente, um trabalho reportou a obtenção de uma proteína híbrida solúvel composta por UreD de *K. aerogenes* associada a proteína ligadora de maltose [Carter, 2010]. Este híbrido proteico permitiu que uma pequena caracterização fosse feita para ureD, constatando que esta chaperona apresenta capacidade de ligação a metais. Entretanto, estudos de complementação do cluster da urease mostraram que este híbrido proteico é apenas parcialmente funcional, e mais estudos são necessários para a compreensão do papel de UreD na ativação da urease.

## Produção de *GmUreD20* recombinante

Informações genômicas mostram a existência de genes para UreD em soja, um localizado no cromossomo 02 (*GmUreD02*) e outro localizado no cromossomo 20 (*GmUreD20*).

O alinhamento das proteínas preditas para estes genes mostra uma identidade de 81 % e similaridade de 83 % (Figura 13). A proteína codificada pelo gene *GmUreD20* foi escolhida aleatoriamente para os estudos de caracterização, pois segundo comunicação pessoal do Dr. Joseph Polacco (University of Missouri – USA), plantas mutantes para um ou outro gene de UreD ainda são capazes de produzir urease enzimaticamente ativa, indicando que ambas isoformas são funcionais.

```

GmUreD02      MSYGAVVTHQSGSCTSLIQSNQRVALPSVDGRGRIMRSGEAVREMEMGSV
GmUreD20      -----MEMGSV
                                     *****

GmUreD02      VVEKVGGRSSVTSCFSRYPLKFIIIPKKVGSSKTDVAVVYALNYGGGIVSG
GmUreD20      VVEKVGGRSSVTSCFSRYPLKFVIPPKKVGSSKTDVAVVYALNYGGGIVSG
*****:*****

GmUreD02      DNISCKFSVGDICTMVLTTQGSTKVKYKSVGSKCSQQILEARVGSNALLAI
GmUreD20      DNISCKFSVGDICTMVLTTQGSTKVKYKSVGSKCSQQILEARVGSNALLAI
*****

GmUreD02      IPDPVTCFSTARYCQKQVFCVLPDSNLVMVDWITSGRHESGEKWDFDLYR
GmUreD20      IPDPVTCFSTARYCQKQVFCVLPDSNLVMVDWITSGRHESGEKWDFDLYR
*****

GmUreD02      STNNIFLEDGQPLFLDTMLLDKEKIGCVQEHMHNYQVIAMIVLLGPKMQY
GmUreD20      SSNNIFLEDGQPLFLDTMLLDKEKIGCVQEHMHNYQVIAMIVLLGPKVQY
*:*****:

GmUreD02      IQNLVQDHSVKKVMSEQLQHPSAAWSHQDKADHFITKPSFIASCSAFGPK
GmUreD20      VQNRVQDHSVKEVMSEQLQHPSAAWSHQDKADHFITKPSFIASCSAFGPK
:* *****:

GmUreD02      VNYKFIQF
GmUreD20      -----

```

Figura 13: Alinhamento das proteínas preditas para os genes *GmUreD02* e *GmUreD20*. (\*) indicam resíduos idênticos, (:) indicam substituições conservadas e (.) indicam substituições semi-conservadas.

Neste trabalho, nós reportamos a clonagem e expressão de *GmUreD20*. A amplificação do cDNA completo de *GmUreD20* (768 pb) foi feita a partir do cDNA de brotos de soja (5 dias de germinação) (Figura 14A) e sua identidade confirmada por sequenciamento de DNA. O teste de expressão da proteína mostrou o aparecimento de uma banda de aproximadamente 29 kDa no extrato de células induzidas com IPTG (Figura 14B), equivalente ao peso molecular predito para *GmUreD20* (28.376,6 Da). A solubilidade da proteína, assim como a confirmação da sua sequência de amino ácidos, estão sendo avaliados.

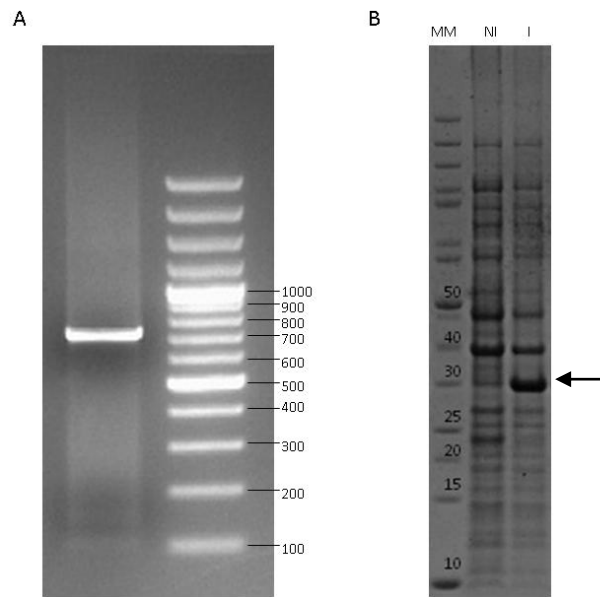


Figura 14: (A) Amplificação do cDNA de *GmUreD20* por PCR. Gel de agarose (1,2 %). Canaleta 1: *GmUreGD20*; Canaleta 2: Ladder, com tamanhos indicados em pares de bases. (B) SDS-PAGE (gradiente 4 – 20%) dos extratos de células não induzidas (NI) e induzidas (I) por IPTG. O padrão de massa molecular (MM) está indicado em kDa. A seta indica a banda da proteína *GmUreD20*.

## IV. Caracterização da proteína acessória UreF de soja (*Glycine max*)

### Materiais e métodos

#### Clonagem do gene r*GmUreF02*

RNA total, extraído de sementes de soja utilizando o reagente de Trizol (Invitrogen), foi empregado para síntese de cDNA através do kit iScript (Bio-Rad) com primers Oligo(dT), de acordo com as recomendações do fabricante. Primers específicos para a amplificação do cDNA completo do gene *GmUreF02* (forward 5'-GGCGGCATATGCAAGTAAATGAAGAACACAACAAACCTTG-3'; reverso 5'-CGCCGGATCCCTAAGATGAAAACAATCTTGAAAACAAGTAACC-3') foram desenhados com base na sequência disponível no site Phytozome (<http://www.phytozome.net/>). O cDNA completo de *GmUreF02* foi amplificado por PCR, nas seguintes condições: desnaturação inicial de 3 min a 94 °C; 35 ciclos de 30 s a 94 °C, 30 s a 57 °C e 60 s a 72 °C; extensão final de 10 min a 72 °C. A amplificação foi analisada por gel de agarose (1,2 %), revelado com brometo de etídio (10 % [v/v]). O produto de PCR foi purificado com o kit Illustra GFX (GE Lifesciences) de acordo com protocolo fornecido pelo fabricante. O produto purificado foi inserido no vetor pGEM-T Easy (Promega) utilizando-se a Ligase T4 (Promega), numa incubação por 18 h a 4 °C. Esta construção foi utilizada para transformação de *Escherichia coli* DH5 $\alpha$ , seguindo o protocolo descrito para *GmUreD20*.

#### Expressão e purificação de *GmUreF02*

O cDNA completo do gene *GmUreF02* foi retirado do vetor pGEM-T Easy, por clivagens sequenciais com as enzimas de restrição NdeI e BamHI. O inserto foi separado do vetor em gel de agarose (1,2 %) e purificado com o kit Illustra GFX. O inserto purificado foi ligado pela enzima ligase T4 ao vetor de expressão pET15b (Novagen), previamente clivado com as enzimas de restrição BamHI e NdeI. Esta construção foi utilizada para transformação de *Escherichia coli* BL21 - Codon Plus por choque térmico. Transformantes positivos foram selecionados como descrito para *GmUreD20*.

Um inóculo de 40 mL de cultura de células de 18 h contendo o inserto pET15b/*GmUreF02* foram adicionados a 2 L de meio LB contendo 100  $\mu$ g/mL de ampicilina e 40  $\mu$ g/mL de cloranfenicol. Quando a densidade óptica da cultura atingiu 0,6 A<sub>600</sub>, a expressão

de *GmUreF02* foi induzida com IPTG (800  $\mu\text{M}$ ) a 37 °C, agitação de 200 rpm, por 18 h, e ao final deste período as células foram coletadas por centrifugação (10 min a 5.000 x *g*). A purificação da proteína recombinante foi realizada em coluna de afinidade por  $\text{Ni}^{2+}$  (Chelating-Sepharose Fast Flow – GE Lifesciences). A concentração protéica foi medida por absorvância em 280 nm, e o valor do coeficiente de extinção molar ( $\epsilon_{280} = 26.470 \text{ M}^{-1}.\text{cm}^{-1}$ ) foi calculado a partir da sequência de aminoácidos, utilizando a ferramenta ProtParam ([au.expasy.org/tools/protparam.html](http://au.expasy.org/tools/protparam.html)). A pureza das amostras foi avaliada por SDS-PAGE.

### **Cromatografia de gel-filtração**

Frações coletadas da cromatografia de afinidade por níquel foram aplicadas em uma coluna Superdex-75 (GE LifeSciences) no sistema FPLC, previamente equilibrada com tampão Tris HCl 20 mM pH 7,5, NaCl 250 mM. A eluição das proteínas foi acompanhada por absorvância em 214 e 280 nm.

### **Análise por espectrometria de massas (MS/MS)**

Bandas de *GmUreF02* foram cortadas de géis corados com Coomassie Blue e tratadas com 180  $\mu\text{L}$  de bicarbonato de amônio 50mM e acetonitrila 50 % por 15 min (3 lavagens), seguido de uma lavagem com 180  $\mu\text{L}$  de acetonitrila. O gel foi então liofilizado e digerido por 24 h a 37 °C usando 12  $\mu\text{L}$  de tripsina suína modificada ( $10 \text{ mg.mL}^{-1}$ ) em  $\text{NH}_4\text{HCO}_3$  25 mM. Peptídeos foram extraídos em duas lavagens com 50  $\mu\text{L}$  de acetonitrila 50 % e TFA por 1 h, e então liofilizados e ressuspendido em 10  $\mu\text{L}$  de água Milli-Q. Análises de MS/MS foram feitas em Espectrômetro de Massas Q-TOF Micro™ (Micromass, Manchester, UK) acoplado a um sistema de cromatografia líquida (UPLC, Waters, Milford, US). Os dados foram adquiridos no modo dependente de dados, e peptídeos com múltiplas cargas (+2 and +3) foram automaticamente selecionados e dissociados. As condições LC e ESI foram fluxo de  $300 \text{ nL.min}^{-1}$ , voltagem capilar de 3.5 kV, temperatura de 100 °C e voltagem do cone de 50 V. Espectros de MS/MS processados usando o software ProteinLynx v. 2.0 (Waters, Milford, US) e os arquivos PKL gerados foram utilizados para buscas em bancos de dados do NCBI, utilizando o software MASCOT (Matrix Science). Todos experimentos de espectrometria de massas foram realizados na Unidade de Química de Proteínas e Espectrometria de Massas (Uniprote-MS), do Centro de Biotecnologia (UFRGS).

## Dicroísmo Circular

Espectros de dicroísmo circular (CD) na região de UV distante (190–260 nm) foram obtidos a 25 °C no equipamento Jasco J715 (Jasco Corporation, Japan) sob fluxo de N<sub>2</sub>, usando cubetas de quartzo de passo óptico de 1 mm, com *GmUreF02* 10 µM em fosfato de sódio 10 mM pH 7,5. Os espectros foram obtidos após acúmulo de 16 scans. A predição quantitativa da estrutura secundária foi feita por deconvolução dos espectros de CD utilizando o servidor DICHROWEB [Whitmore e Wallace, 2004; 2008]. Estas análises foram realizadas no laboratório do Grupo de Biofísica Molecular Sérgio Mascarenhas do Instituto de Física de São Carlos (USP-São Carlos)

## Resultados e Discussão

UreF e UreD são as proteínas acessórias para as quais existe a menor quantidade de informação disponível. Isso se deve ao fato de ambas as proteínas serem geralmente produzidas na forma de corpos de inclusão, quando super expressas [Carter e Hausinger, 2010; Moncrief e Hausinger, 1997] independente do seu organismo de origem. Até hoje, todas as tentativas de purificar estas proteínas na forma nativa, através de protocolos de desnaturação e renaturação, falharam, geralmente levando à formação de agregados gigantes (na ordem de mega daltons), que obviamente não correspondem a forma nativa desta proteína [Salomone-Stagni et al., 2007]. O contínuo fracasso na obtenção destas proteínas na forma solúvel levou pesquisadores a criarem proteínas híbridas, nas quais UreF era fusionada a alguma outra proteína, levando a obtenção de um híbrido solúvel. No entanto, estas proteínas híbridas mostraram-se extremamente limitadas para a caracterização da proteína alvo [Kim et al., 1999; 2006]. Recentemente, o primeiro sucesso na expressão de UreF na fração solúvel do extrato de *E. coli* foi obtido para uma forma truncada da UreF de *H. pylori* [Lam et al., 2010]. No entanto, ainda não foram realizados testes de complementação do cluster da urease com esta forma truncada e, portanto, sua funcionalidade ainda não foi provada.

## Produção de *GmUreF02* recombinante

Informações genômicas mostram a existência de dois genes para UreF em soja, um localizado no cromossomo 02 (*GmUreF02*) e outro localizado no cromossomo 14 (*GmUreF14*). O alinhamento das proteínas preditas para estes genes mostra uma identidade de 95 % e similaridade de 98 % (Figura 15). A proteína codificada pelo gene *GmUreF02* foi escolhida para



os estudos de caracterização pois, segundo uma comunicação pessoal do Dr. Joseph Polacco (University of Missouri – USA), esta isoforma da proteína é aquela cuja ausência leva à perda da capacidade de produção de urease ativa em plantas mutantes, e esta ausência não pode ser suplementada pela isoforma do cromossomo 14.

```

GmUreF02      MQVNEEHNKPCSDPFLQWSQWQLLDSLLPTGGFAHSFGLEAAVQCHLVSD
GmUreF14      MQVNEGHNRPCSDPFLQWSKWQLLDSLLPTGGFAHSFGLEATVQSHLVSN
                ***** *:*****:*****:*.****:

GmUreF02      SNDLKT FVIHILENTG SLLL PFVYSASMLPNLETWHKLDKILDATLTNEV
GmUreF14      SNDLKT FVIHILENTG SLLL PFVYSASMLPNLETWHKLDKILDATLTNEV
                *****

GmUreF02      GRKASISQGSALMRVASAVFSEVPSLKTMRDTSLGLGTVSFHHAPVFGLI
GmUreF14      GRKASISQGSALMRVASAVFSEVPSLKTMRDASLGLETVSFHHAPVFGLI
                *****:**** *****

GmUreF02      CGALGF DKTSSQRAYM FITMRDVISAATRLNLIGPLGAALLQH QVAPNAE
GmUreF14      CGALGF DRTSSQRAYM FITMRDVISAATRLNLIGPLGAALLQHRVAPIAE
                *****:*****:*** **

GmUreF02      VILEKWMNRVDEACQTMPLLDTVQGCHGYLFSRLFSS
GmUreF14      VILEKWMNRVVEEACQTMPLLDTVQGCHGYLFSRLFSS
                *****.*:*****

```

Figura 15: alinhamento das proteínas preditas para os genes *GmUreF02* e *GmUreF14*. (\*) indicam resíduos idênticos, (:) indicam substituições conservadas e (.) indicam substituições semi-conservadas.

O cDNA completo de *GmUreF02* (717 pb) foi amplificado a partir de RNA extraído de sementes de soja (Figura 16), utilizando-se primers específicos. A identidade do gene foi analisada através de sequenciamento, confirmando ser o cDNA de *GmUreF02*.

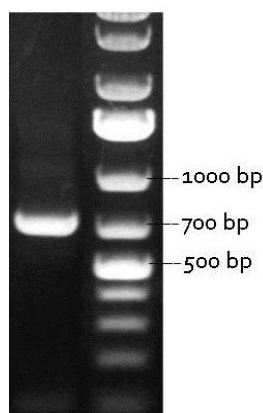


Figura 16: Amplificação do cDNA completo de *GmUreF02*. Gel de agarose (1,2%), corado com brometo de etídio. A segunda canaleta mostra padrões de pares de bases, com seus tamanhos indicados ao lado.

A expressão de *GmUreF02* após a indução por IPTG (0,8 mM) foi acompanhada em relação ao tempo pelo monitoramento do aparecimento de uma banda de aproximadamente 28 kDa em SDS-PAGE no extrato celular total (Figura 17A). Uma aparente estagnação e possível diminuição no conteúdo de *GmUreF02* após 2 horas de indução foi observada, de forma que este foi determinado como tempo ótimo de expressão.

Assim como todas as UreFs até hoje descritas, *GmUreF* também é superexpressa na forma de corpos de inclusão. Essa característica nos levou a usar um protocolo de desnaturação para solubilizar e permitir a purificação da proteína, que foi subsequentemente submetida a um processo de renaturação. A proteína foi purificada até homogeneidade em um único passo de cromatografia de afinidade por Ni<sup>2+</sup> (Figura 17B) e sua identidade confirmada através de sequenciamento por MS/MS (Tabela 4).

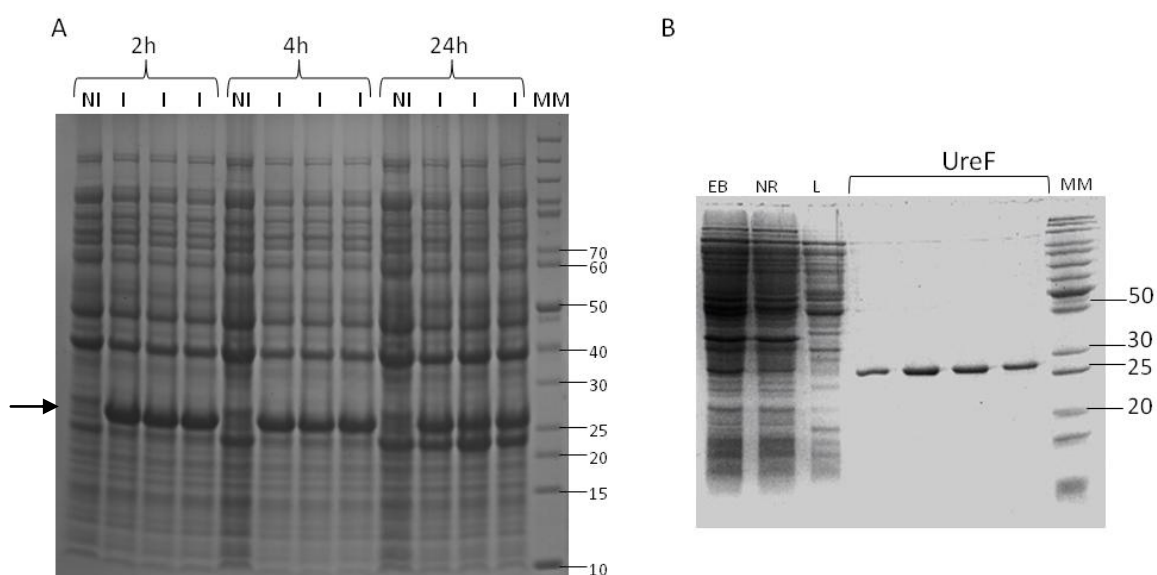


Figura 17: Expressão e purificação de *GmUreF02*. (A) Gel de gradiente (4 – 20 %) dos diferentes tempos de indução da expressão com IPTG. NI = não induzido, I = induzido. O tempo de indução está indicado no topo do gel. (B) SDS-PAGE (12 %) da purificação de *GmUreF02*. EB = extrato bruto, NR = não retido, L = lavagem. O padrão de massa molecular (MM) está indicado em kDa. A seta indica a banda da proteína *GmUreF02*.

Tabela 4: Sequenciamento de *GmUreF02* por ESI Q-TOF.

Sequência Peptídica	<i>GmUreF02</i>
<sup>91</sup> ILDATLTNEVGR <sup>102</sup>	+
<sup>103</sup> KASISQGSALMR <sup>114</sup>	+
<sup>104</sup> ASISQGSALMR <sup>114</sup>	+
<sup>127</sup> VASAVFSEVPSLK <sup>139</sup>	+
<sup>164</sup> AYMFITMR <sup>171</sup>	+
<sup>180</sup> LNLIGPLGAALLQHQPNAEVILEK <sup>205</sup>	+

+ indica peptídeo encontrado no sequenciamento.

Os peptídeos sequenciado correspondem a 30 % de cobertura total da proteína.

### Análise estrutural de *GmUreF02*

Apesar do sucesso na obtenção da proteína renaturada na forma solúvel, análises mais aprofundadas foram necessárias para confirmar o correto enovelamento da proteína. Como dito anteriormente, tentativas prévias de renaturação de UreF foram capazes de gerar um produto solúvel, no entanto correspondendo a formas agregadas da proteína [Salomone-Stagni et al., 2007]. Lam et al. [2010], reportam que a forma truncada de *HpUreF*, utilizada na cristalização da proteína, comporta-se como um dímero em solução. No entanto, na mesma publicação, a possibilidade de que a proteína nativa apresente este arranjo quaternário é posta em dúvida e existem três argumentos que suportam esta suspeita: (1) estudos do complexo ativador da urease envolvendo cross-link químico entre as proteínas acessórias em *K. aerogenes*, não identificaram qualquer cross-link envolvendo interações UreF-UreF [Chang et al., 2004]; (2) A confirmação da existência do complexo UreDF-apourease de *K. aerogenes* foi provada por uma combinação de análises densitométricas de géis de poli(acrilamida) nativos e desnaturantes, e a interpretação dos resultados confirmou uma estequiometria de (UreDF)<sub>0-3</sub>(apourease)<sub>3</sub>, inconsistente com a existência de um dímero de UreF [Moncrief e Hausinger, 1996]; (3) Análises de espalhamento de raios X em baixo ângulo (SAXS) do complexo ativador da urease de *K. aerogenes* geraram um modelo também incompatível com a existência do dímero UreF-UreF [Quiroz-Valenzuela et al., 2008].

Através de cromatografia de gel-filtração (GF) em coluna Superdex-75, observamos que *GmUreF02* recombinante se apresenta na forma monomérica (Figura 18).

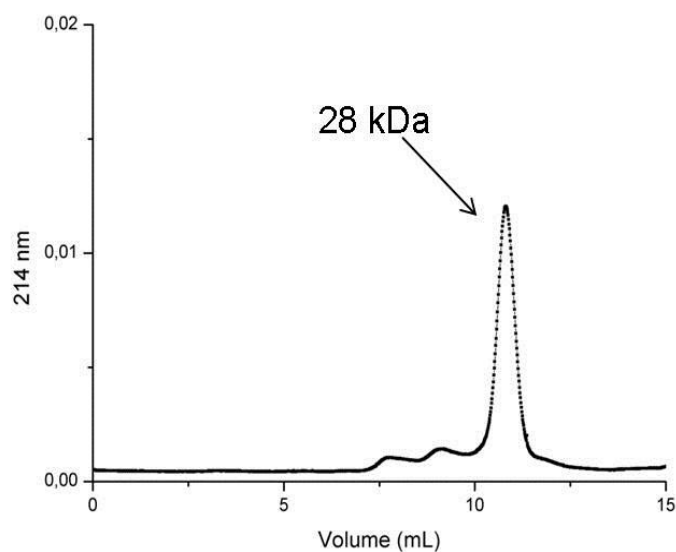


Figura 18: Gel-filtração em coluna Superdex-75 de *GmUreF02*, mostrando que esta proteína encontra-se em sua forma monomérica após a purificação em coluna de afinidade por níquel.

O simples fato de *GmUreF* não apresentar a formação de agregados após a renaturação não representa garantia de enovelamento correto da proteína e, por esta razão, análises estruturais adicionais foram realizadas por técnicas de espectroscopia de dicroísmo circular. O espectro de *GmUreF02* mostrou uma predominância de  $\alpha$ -hélices em seu conteúdo de estrutura secundária, com deflexões negativas em 208 e 222 nm, e um pico positivo em 194 nm (Figura 19B). O espectro foi quantitativamente analisado, e o melhor ajuste obtido estimou a composição de estrutura secundária em 68 %  $\alpha$ -hélices, 2 % folhas  $\beta$ , 13 % turns e 17 % random coil (nrmsd = 4 %; programa Contin; conjunto de referencia n° 4), em perfeita concordância com o conteúdo de estrutura secundária predito pelo algoritmo JPRED (Figura 19A). Ainda com o objetivo de confirmar o enovelamento de *GmUreF02* análises de desnaturação foram conduzidas e a perda de estrutura da proteína foi monitorada a 222 nm. Como evidenciado na Figura 19C, a proteína perde consideravelmente sua estrutura secundária com o acréscimo da concentração de desnaturante, reforçando a idéia de que o processo de renaturação da proteína realmente permite seu correto enovelamento. Desta forma, este trabalho constitui o primeiro relato da obtenção de uma proteína *UreF* tipo selvagem na forma solúvel completamente enovelada. Caracterizações bioquímicas desta proteína estão em progresso.

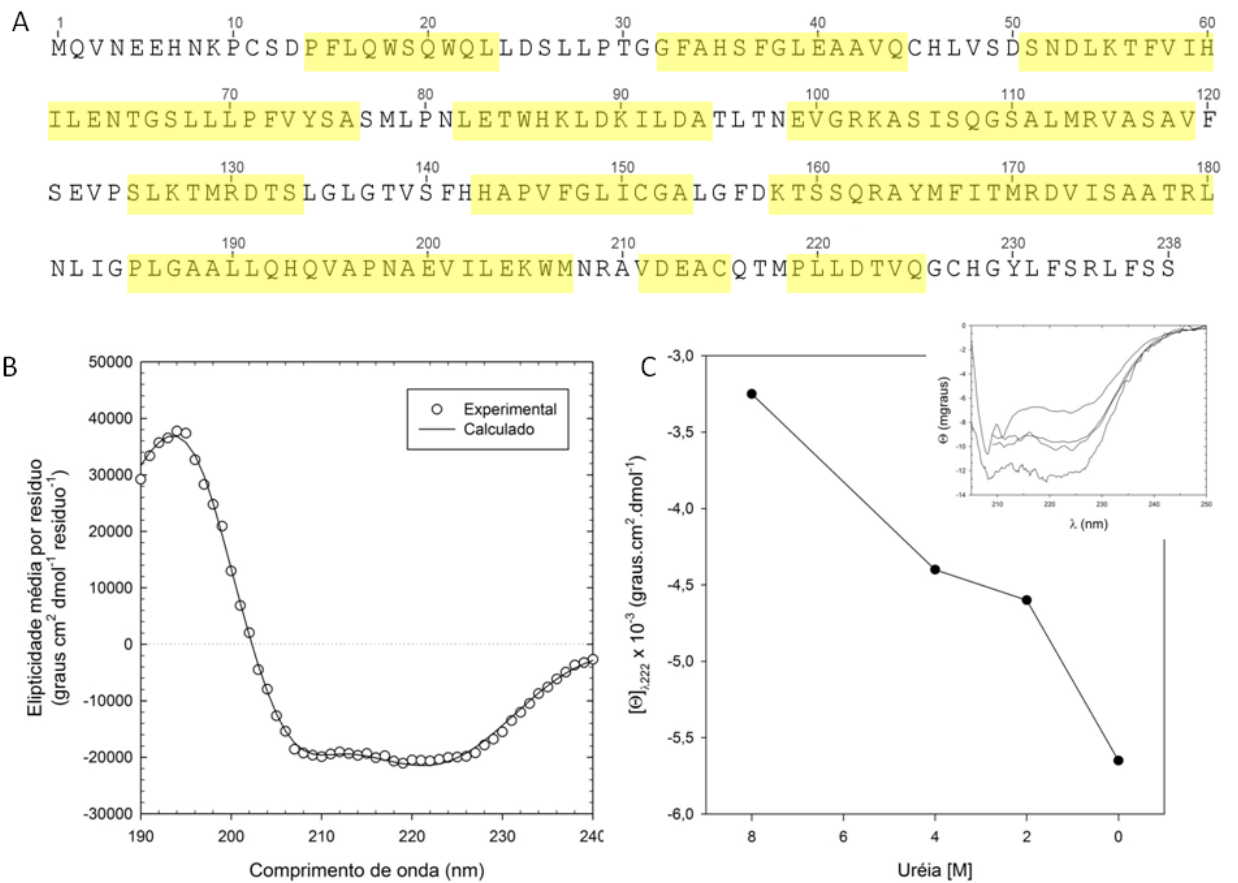


Figura 19: Análise da estrutura secundária de *GmUreF02*. (A) Predição do conteúdo de estrutura secundária pelo algoritmo JPRED, com 67 % de  $\alpha$ -hélices (em amarelo) e 33 % entre turns e random coil. (B) Análises de CD da proteína renaturada. (C) Análise da perda de estrutura secundária da proteína na presença de crescentes concentrações de uréia, demonstrada pela variação da elipticidade em 222 nm. Inset mostra o espectro original de onde C foi derivado.

## V. Caracterização da proteína acessória UreG de soja (*Glycine max*)

### Materiais e métodos

#### Purificação de *GmUreG* de sementes de soja

*GmUreG* foi purificada das sementes de soja de acordo com Freyermuth et al. [2000], com modificações. Sementes de soja (200 g) foram moídas em moinho analítico e as proteínas foram extraídas com 5 volumes de tampão A (Tris-HCl 50 mM pH 7,5, NaCl 500 mM) contendo imidazol 15 mM, a 4 °C sob agitação por 2 h. O extrato foi centrifugado por 30 min a 30.000 x g. O sobrenadante foi aplicado em uma coluna de resina Chelating-Sepharose Fast Flow (GE Lifesciences), previamente carregada com Ni<sup>+2</sup> e equilibrada com o mesmo tampão. Proteínas não-retidas foram eluídas com tampão A contendo imidazol 50 mM. Proteínas ligadas à coluna foram eluídas com tampão A contendo concentrações crescentes de imidazol (100, 200 e 300 mM). A concentração protéica foi medida por absorvância em 280 nm, e o valor do coeficiente de extinção molar ( $\epsilon_{280} = 19.940 \text{ M}^{-1} \cdot \text{cm}^{-1}$ ) foi calculado a partir da sequência de aminoácidos, utilizando a ferramenta ProtParam ([au.expasy.org/tools/protparam.html](http://au.expasy.org/tools/protparam.html)). A pureza das amostras foi avaliada por SDS-PAGE.

#### Clonagem do gene *GmUreG*

O RNA total, extraído de sementes de soja utilizando o reagente de Trizol (Invitrogen), foi empregado para síntese de cDNA através do kit iScript (Bio-Rad) com primers Oligo(dT), de acordo com as recomendações do fabricante. Primers específicos para a amplificação do cDNA completo do gene *GmUreG* (forward 5'- GCGGCCCATGGCTTCTGAAGGCGACCATCACC-3'; reverso 5'- CGCCGGATCCTTAATGACGCTTATCCCTGTGGCT-3') foram desenhados com base na sequência disponível no site Phytozome (<http://www.phytozome.net/>). O cDNA completo de *GmUreG* foi amplificado por PCR, nas seguintes condições: desnaturação inicial de 3 min a 94 °C; 35 ciclos de 30 s a 94 °C, 30 s a 57 °C e 60 s a 72 °C; extensão final de 10 min a 72 °C. A amplificação foi analisada em gel de agarose (1,2 %), revelado com brometo de etídio (10 % [v/v]). O produto de PCR foi purificado com o kit Illustra GFX (GE Lifesciences), de acordo com protocolo fornecido pelo fabricante. O produto purificado foi inserido no vetor pGEM-T Easy (Promega), utilizando-se a Ligase T4 (Promega) numa incubação até o dia seguinte a 4 °C. Esta

construção foi utilizada para transformação de *Escherichia coli* DH5 $\alpha$ , como descrito para *GmUreD20*.

### **Expressão e purificação de *GmUreG* recombinante**

O cDNA completo do gene *GmUreG* foi retirado do vetor pGEM-T Easy por clivagens sequenciais com as enzimas de restrição NcoI e BamHI. O inserto foi separado do vetor em gel de agarose (1,2 %) e purificado com o kit Illustra GFX. O inserto purificado foi ligado pela enzima ligase T4 ao vetor de expressão pET15b (Novagen) previamente clivado com as enzimas de restrição BamHI e NcoI (esta última empregada para remover a porção codificadora da cauda de histidina do vetor). Esta construção foi utilizada para transformação de *Escherichia coli* BL21 - Codon Plus(DE3)-RIL por choque térmico. Transformantes positivos foram selecionados por PCR com primers gene-específicos. A inserção correta do inserto no vetor foi confirmada por sequenciamento de DNA.

Um inóculo de 40 mL de cultura de células de 18 h contendo o inserto pET15b/*GmUreG* foram adicionados a 2 L de meio LB contendo 100  $\mu$ g/mL de ampicilina e 40  $\mu$ g/mL de cloranfenicol. Quando a densidade óptica da cultura atingiu 0,6  $A_{600}$ , a expressão de *GmUreG* foi induzida com IPTG (800  $\mu$ M) a 37 °C, agitação de 200 rpm, por 18 h. Células foram coletadas por centrifugação (10 min a 5.000 x *g*), lisadas com sonicação, centrifugadas novamente (10 min a 30.000 x *g*) e o sobrenadante utilizado para purificação da proteína recombinante. O sobrenadante foi aplicado em uma coluna de resina Chelating-Sepharose (GE Lifesciences), previamente carregada com Ni<sup>+2</sup> e equilibrada com tampão A contendo imidazol 15 mM. Proteínas não retidas foram eluídas com tampão A contendo imidazol 50 mM. Proteínas ligadas à coluna foram eluídas com tampão A contendo 300 mM de imidazol. A concentração proteica foi determinada como descrito anteriormente. A pureza das amostras foi avaliada por SDS-PAGE.

### **Clonagem e expressão de uma forma truncada de *GmUreG***

Primers específicos para a amplificação de *GmUreG* com deleção dos primeiros 56 resíduos de aminoácidos foram desenhados a partir da sequência do cDNA completo de *GmUreG* (forward 5'- GCGGCATATGCATTCTCACGAACCCATTTAC-3'). Esta forma truncada foi chamada *GmUreG* $\Delta$ 56. A clonagem deste cDNA foi realizada como descrito para o cDNA da

proteína completa, com uma única modificação na clivagem do vetor de expressão, utilizando as enzimas BamHI e NdeI, incorporando a cauda de histidina e o sítio de clivagem por trombina no N terminal da proteína truncada. A expressão e purificação foram realizadas com o mesmo protocolo estabelecido para a proteína completa.

### **Análise por espectrometria de massas**

Análises por ESI Q-TOF foram realizadas como descrito para *GmUreF02*.

### **Dicroísmo Circular**

Espectros de dicroísmo circular (CD) na região de UV distante (190–260 nm) foram obtidos no equipamento Jasco J715 (Jasco Corporation, Japan) a 25 °C, sob fluxo de N<sub>2</sub>, usando cubetas de quartzo de passo óptico de 1 mm, com *GmUreG* 10 µM em HEPES 20 mM pH 7,5, NaCl 150 mM, na presença ou ausência de Zn<sup>2+</sup> ou Ni<sup>2+</sup>. Todos espectros foram obtidos após acúmulo de 16 scans. A predição quantitativa da estrutura secundária foi feita por deconvolução dos espectros de CD utilizando o servidor DICHROWEB [Whitmore e Wallace, 2004; 2008]. Estas análises foram realizadas no laboratório do Grupo de Biofísica Molecular Sérgio Mascarenhas do Instituto de Física de São Carlos (USP-São Carlos) e no CERM (Centro Risonanze Magnetiche, Università degli Studi di Firenze, Sesto Fiorentino, Italia).

### **Análise por Ressonância Magnética Nuclear (RMN)**

*GmUreG* uniformemente marcada com <sup>15</sup>N foi produzida através crescimento das células em meio mínimo M9 (6 g.L<sup>-1</sup> Na<sub>2</sub>HPO<sub>4</sub>, 3 g.L<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub>, 0.5 g.L<sup>-1</sup> NaCl, 1 mM MgSO<sub>4</sub>, 4 g.L<sup>-1</sup> Glucose, pH 7,5) contendo <sup>15</sup>NH<sub>4</sub>Cl como única fonte de nitrogênio. Os espectros de RMN foram obtidos para uma amostra de *GmUreG* 0,7 mM enriquecida em <sup>15</sup>N, em tampão fosfato de sódio 20 mM, NaCl 50 mM, pH 7,0 e temperatura de 298 K, usando um espectrofotômetro Bruker Avance 900, operando a 899.07 MHz e equipado com probe criogênico TXI 5-mm, por gradiente de campo pulsado ao longo do eixo z. O espectro <sup>1</sup>H,<sup>15</sup>N-HSQC foi adquirido em modo sensível à fase usando seleção de gradiente Echo/Antiecho-TPPI (*time-proportional phase incrementation*) e pulsos “flip-back” para a supressão do sinal da água. Os espectros consistem de 8 scans, com janela espectral de 16/40 ppm nas dimensões próton/nitrogênio e com o carrier definido na frequência da água e 118 ppm, respectivamente. O delay de



relaxamento foi 1 s, mais 285 ms de tempo de aquisição. A matriz de 2048 x 256 pontos foi obtida em dimensões F2 e F1, respectivamente, e transformada em 2048 x 1024 pontos. Os dados foram processados utilizando o software iNMR (<http://www.inmr.net/>). Estes estudos foram realizados no CERM (Centro Risonanze Magnetiche, Università degli Studi di Firenze, Sesto Fiorentino, Italia)

### **Medidas de Espalhamento de Luz**

Estimativas da massa molecular absoluta de *GmUreG* foram feitas utilizando uma combinação de cromatografia de gel filtração, espalhamento de luz em múltiplos ângulos (MALS) e espalhamento de luz quase-elástico (QELS). *GmUreG* (100  $\mu$ M, 200  $\mu$ L) em Tris-HCl 20 mM, pH 7,5, NaCl 150 mM, foi aplicada em uma coluna Superdex-75 HR 10/30 (Amersham Biosciences), equilibrada com o mesmo tampão, e eluída em um fluxo de 0,5 ml.min<sup>-1</sup>. A coluna estava conectada a um fotômetro com laser multiangular (690.0 nm) DAWN EOS (Wyatt Technology). Dados de espalhamento de luz Quase-elástico (dinâmico) foram coletados em ângulo de 90° usando o equipamento Wyatt-QELS. A concentração da proteína eluída foi determinada utilizando um detector de índice de refração (Optilab DSP, Wyatt). Valores de 0,185 para o incremento do índice de refração ( $dn/dc$ ) e 1,330 para o índice refrativo do solvente foram utilizados. O peso molecular foi determinado a partir do gráfico de Zimm. Os dados foram analisados utilizando o software Astra 4.90.07 (Wyatt Technology), de acordo com as instruções do fabricante. Estas medidas foram realizadas no CERM (Centro Risonanze Magnetiche, Università degli Studi di Firenze, Sesto Fiorentino, Italia).

### **Espectroscopia de Emissão Óptica de Plasma Indutivamente Acoplado (ICP-OES - *Inductively coupled plasma optical emission spectroscopy*)**

A análise do conteúdo de metal foi realizada com um espectrômetro de emissão óptica Spectro Ciros CCD ICP (Spectro Analytical Instruments GmbH, Kleve, Germany) em combinação com um nebulizador de Lichte e uma bomba peristáltica para injeção da amostra, segundo o protocolo previamente descrito [Stola et al., 2006]. Estas análises foram realizadas no Dipartimento di Scienze e Tecnologie Agroambientali della Università di Bologna.

## Titulação Calorimétrica Isotérmica (ITC)

Experimentos de titulação foram realizados a 25 °C utilizando um microcalorímetro de alta sensibilidade VP-ITC (MicroCal LLC, Northampton, MA). As proteínas e os íons metálicos (estoques de 100 mM) foram diluídos em tampão Tris-HCl 20 mM, pH 7,5, NaCl 75 mM. Água deionizada foi utilizada como referência. Antes de cada experimento, a estabilidade da linha de base foi checada. Um intervalo de 300 s entre as injeções foi utilizado, para permitir que o sistema atingisse o equilíbrio térmico após cada adição. Para cada titulação, um experimento controle foi realizado, titulando-se o metal no tampão (sem proteína), nas mesmas condições. O calor de diluição foi negligenciável. A cubeta contendo a amostra (*rGmUreG* 15 µM), foi titulada com 60 injeções (5 µL cada) de uma solução de ZnSO<sub>4</sub> 300 µM ou NiSO<sub>4</sub> 600µM.

Os dados de calor integrados foram analisados com o software Origin (Microcal), e ajustados usando um algoritmo de minimização não linear por mínimos quadrados para curvas de titulação teóricas que envolvam diferentes modelos de ligação. O parâmetro qui-quadrado reduzido  $\chi_{v2}$  ( $\chi_{v2} = \chi_2/N$ , onde  $N$  são os graus de liberdade,  $N = N_{idp} - N_{par}$ ,  $N_{idp}$  = número de pontos e  $N_{par}$  = número de parâmetros estimados) foi utilizado para estabelecer o melhor ajuste entre os modelos testados. Valores de variação de entalpia da reação ( $\Delta H$ ), a constante de afinidade ( $K_d$ ) e o número de sítios ( $n$ ) foram os parâmetros do ajuste. A entropia da reação foi calculada usando as equações  $\Delta G = -RT \cdot \ln(K_d)$  ( $R = 1.9872 \text{ cal mol}^{-1} \text{ K}^{-1}$ ,  $T = 298 \text{ K}$ ) e  $\Delta G = \Delta H - T\Delta S$ . As constantes de dissociação e parâmetros termodinâmicos aqui mostrados não levam em consideração possíveis eventos de transferência de prótons relacionados à ligação do metal ou a presença, em solução, de complexos entre os íons metálicos e o tampão. Tais medidas foram realizadas no Dipartimento di Scienze e Tecnologie Agroambientali della Università di Bologna.

## Atividade GTPásica

GTP (2 mM) e *GmUreG* (5 µM) foram incubados em Tris-HCl 20 mM, pH 8,0, NaCl 75 mM, a 37 °C por intervalos de tempo de 1 a 6 h, na presença de MgCl<sub>2</sub> 5 mM. O fosfato liberado pela hidrólise do GTP foi medido pelo método colorimétrico de Fiske-Subbarow [1925]. O efeito de Zn<sup>2+</sup> e Ni<sup>2+</sup> na atividade GTPásica de *GmUreG* foi testado pela adição de 125 µM de cada metal à reação.

## Modelagem molecular de *GmUreG*

Um alinhamento inicial das sequências da HypB de *Methanococcus jannaschii* (*MjHypB*) e *GmUreG* foi produzido utilizando o software Modeller v9.8 [Marti-Renom et al., 2000] e este alinhamento foi ajustado manualmente para corresponder às estruturas primárias e secundárias das proteínas, determinadas usando o servidor JPRED [Cuff et al., 1998]. Os resíduos Cys e His envolvidos na ligação ao  $Zn^{2+}$  na estrutura de *MjHypB* foram alinhados aos resíduos correspondentes no motivo conservado Cys-Pro-His encontrado em todas as proteínas UreG [Zambelli et al., 2007]. Este alinhamento foi então utilizado para calcular 100 modelos estruturais do dímero de *GmUreG* (que inclui dois íons  $Mg^{2+}$ , os átomos de hidrogênio e a molécula  $GTP\gamma S$ ) utilizando o programa Modeller e a estrutura do dímero de *MjHypB* (PDB 2HF8) como molde. Os melhores modelos foram selecionados com base nos menores valores de DOPE no Modeller [Shen e Sali, 2006] e foram então submetidos a um passo de refinamento de otimização de loops, utilizando os parâmetros padrão do Modeller. Os resultados da análise no PROCHECK [Laskowski et al., 1993] para o modelo final foram completamente satisfatórios. A superfície molecular e os gráficos dos domínios de *GmUreG* foram calculados com o pacote UCSF Chimera [Pettersen et al., 2004]. O código eletrostático de cores foi gerado com o software DelPhi [Honig e Nicholls, 1995]. Todos os resíduos de histidina foram considerados neutros. A constante dielétrica interna da proteína foi estipulada como 4 em todos os cálculos, e a constante dielétrica do solvente como 80. A concentração de sal foi estabelecida como NaCl 0,15 M, que corresponde a força iônica fisiológica.

## Resultados e Discussão

### Purificação de *GmUreG* das sementes de soja

Apesar da considerável identidade de sequência (c.a. 60 %) observada entre UreG de plantas e de bactérias, os primeiros 25 % da sequência destas proteínas em plantas é ausente nas UreGs bacterianas. Essa região N terminal é extremamente rica em resíduos de histidina e, neste aspecto, se assemelha ao C terminal da proteína acessória UreE de bactérias. No caso da *GmUreG*, dos primeiros 58 resíduos, 23 são histidinas e 19 destas estão localizadas entre os resíduos 7 e 30 (Figura 20). Essa característica permite o uso de uma cromatografia de afinidade por níquel para sua purificação em único passo, a partir do extrato de sementes de soja.

```

1      10      20      30      40      50
MASEGDH HHHH QDHDH HDHDH HHHHH DGEGETNSWV GKDGKVYHSH
60      70      80      90      100
DGLAPHSHEP IYSPGYFTRR APPLLNRFNF ERAFTVGIGG PVGTGKTALM
110     120     130     140     150
LALCELLREN YSLAAVTNDI FTKEDGEFLV KHKALPEERI RAVETGGCPH
160     170     180     190     200
AAIREDISIN LGPLEELSNL FKADILLCES GGDNLAANFS RELADYIIYI
210     220     230     240     250
IDVSGGDKIP RKG GPGITQA DLLVINKTDL APAIGADLAV MQRDALRMRD
260     270     280     285
GGPFVFAQVK HKIGVEEIGN LVLQAWEEAT GNKRH

```

Figura 20: Sequência de *GmUreG*, conforme predição a partir da sequência gênica disponível no site Phytozome.

Devido a este alto conteúdo de histidinas, *GmUreG* não é eluída da resina até que a concentração de imidazol seja elevada para 200 mM. Como previamente descrito por Freyermuth et al. [2000], a proteína nativa é purificada em frações com diferentes graus de afinidade pela coluna de níquel. A fração majoritária de *GmUreG* é eluída com 200 mM de imidazol, no entanto uma fração menor só é eluída quando a concentração de imidazol é elevada para 300 mM (Figura 21). Estas frações foram denominadas *GmUreG200* e *GmUreG300*.

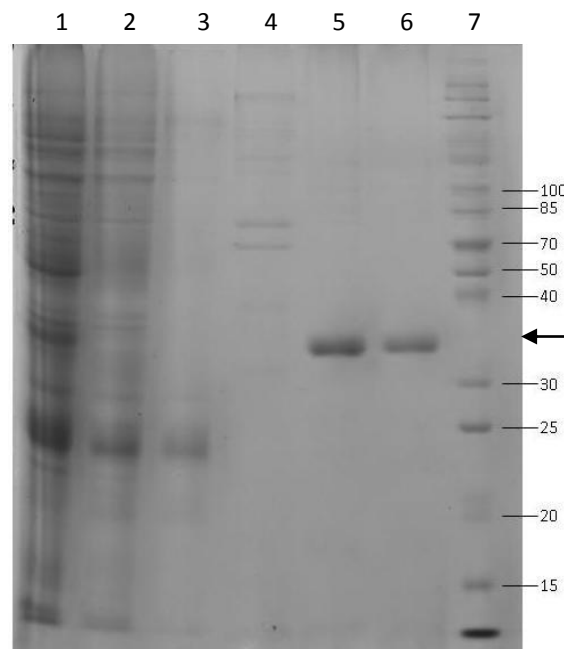


Figura 21: SDS-PAGE (12%) da purificação de *GmUreG* a partir do extrato de sementes de soja. Canaletas 1: extrato bruto; 2: não retido; 3: lavagem com 50 mM de imidazol; 4: eluição com 100 mM de imidazol; 5: eluição com 200 mM de imidazol (*GmUreG200*); 6: eluição com 300 mM de imidazol (*GmUreG300*); 7: padrão de peso molecular. A seta indica a banda das proteínas *GmUreG200* e *GmUreG300*.

Estudos anteriores [Freyermuth et al., 2000], pelo reconhecimento das duas frações por anticorpos policlonais contra *UreG* de *A. thaliana*, sugeriam que ambas as frações correspondessem a mesma proteína, no entanto nenhuma análise de sequência foi realizada para confirmar esta hipótese. Visando verificar a identidade destas duas frações, ferramentas de espectrometria de massas foram empregadas. Ambas as frações foram positivamente identificadas como *GmUreG* através do sequenciamento de diversos peptídeos internos por MS/MS (ESI Q-ToF) (Tabela 5).

Tabela 5: Sequenciamento de *GmUreG* por ESI-Q-TOF

Sequência Peptídica	<i>GmUreG200</i>	<i>GmUreG300</i>	<i>rGmUreG</i>
<sup>70</sup> RAPLLNR <sup>77</sup>	+	+	
<sup>83</sup> AFTVGIGGPVGTGK <sup>96</sup>	+	+	+
<sup>97</sup> TALMLALCELLR <sup>108</sup>	+		
<sup>109</sup> ENYSLAAVTNDIFTK <sup>123</sup>	+		+
<sup>124</sup> EDGEFLVK <sup>131</sup>	+		
<sup>133</sup> ALPEER <sup>138</sup>	+	+	
<sup>141</sup> AVETGGCPHAAIR <sup>153</sup>	+		
<sup>154</sup> EDISINLGPLEELSNLKF <sup>172</sup>	+	+	+
<sup>192</sup> ELADYIIYIIDVSGGDKIPR <sup>211</sup>	+		+
<sup>213</sup> KGGPGITQADLLVINK <sup>227</sup>	+	+	+
<sup>228</sup> TDLAPAIGADLAVMQR <sup>243</sup>	+	+	+
<sup>250</sup> DGGPFVFAQVK <sup>260</sup>	+		
<sup>263</sup> IGVEEIGNLVLQAWAATGNK <sup>283</sup>	+		+

+ indica peptídeos encontrados no sequenciamento da fração. A numeração indica a posição dos amino ácidos flanqueadores na sequência de aminoácidos

Os fragmentos sequenciados equivalem a 62, 27 e 36 % de cobertura de sequência para *GmUreG200*, *GmUreG300* e *rGmUreG*, respectivamente.

### Expressão de *GmUreG* recombinante

Devido às grandes quantidades de proteína necessárias para as caracterizações bioquímica e estrutural, tornou-se imprescindível a expressão de *GmUreG* em sistema bacteriano. O cDNA completo de *GmUreG* (858 pb) foi obtido a partir da amplificação por primers específicos, utilizando como molde o cDNA first-strand sintetizado do RNA total extraído de sementes de soja (Figura 22). O cDNA foi clonado no vetor pET15b modificado,

onde a porção codificadora da cauda 6-His foi previamente removida. Desta forma, a proteína recombinante não apresenta qualquer diferença de sequência em relação à proteína nativa.

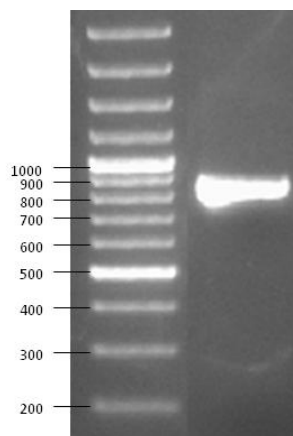


Figura 22: Amplificação por PCR do cDNA completo de *GmUreG*. Gel de agarose (1,2%). Canaleta 1: Ladder, com tamanho indicado em pares de bases. Canaleta 2: reação de PCR utilizando primers específicos para *GmUreG*.

A proteína recombinante tipo selvagem (*rGmUreG*) superexpressa em células de *E. coli* acumula-se na fração solúvel, como previamente descrito para *MtUreG* [Zambelli et al., 2007], *HpUreG* [Mehta et al., 2003a; Zambelli et al., 2008] e *KaUreG* [Moncrief e Hausinger, 1997; Boer et al., 2010], sendo diferente de *BpUreG*, que quando superexpressa é apenas encontrada na fração insolúvel na forma de corpos de inclusão [Zambelli et al., 2005]. *rGmUreG* foi purificada em dois passos, uma cromatografia de afinidade por níquel seguida de cromatografia de GF em Superdex-75 em sistema de FPLC (Figura 23). A proteína recombinante agrega parcialmente após sua eluição da coluna de níquel, e precisa ser imediatamente submetida à GF para remoção destes agregados. Este procedimento é necessário para evitar a precipitação da proteína que ocorre em apenas algumas horas após sua eluição, caso a remoção dos agregados não seja feita. A proteína resultante deste processo é estável e, portanto, adequada para as caracterizações subsequentes. A identidade de *rGmUreG* foi confirmada pela digestão trípica da banda de SDS-PAGE da proteína, seguida por sequenciamento por MS/MS (ESI Q-Tof). Os peptídios identificados estão apresentados na Tabela 5.

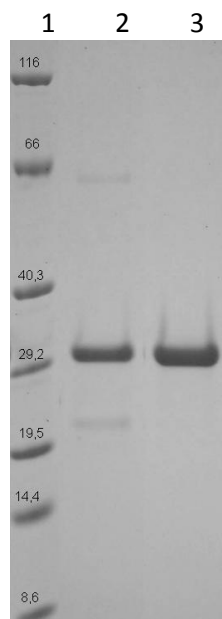


Figura 23: SDS-PAGE (gradiente 4-20 %) da purificação de *rGmUreG*. Canaletas 1: padrão de peso molecular; 2: fração eluída da cromatografia de afinidade por níquel; 3: gel-filtração da fração eluída da cromatografia de afinidade por níquel.

### **Análise estrutural de *GmUreG***

A predição de elementos de estrutura secundária de *GmUreG*, baseada no algoritmo JPRED (Figure 24), indica um conteúdo de 34% de  $\alpha$ -hélices e 15% de folhas  $\beta$ , com os 51% restantes sendo constituídos de turns, random coil, além dos primeiros 44 resíduos (composto pelo N terminal rico em histidinas, considerado pelo algoritmo como muito complexo para ser predito). Essa predição também sugere que *GmUreG* apresenta enovelamento similar ao de diversas proteínas *UreG*, como descrito anteriormente com base em estudos de modelagem molecular [Zambelli et al., 2007; 2005; 2009].

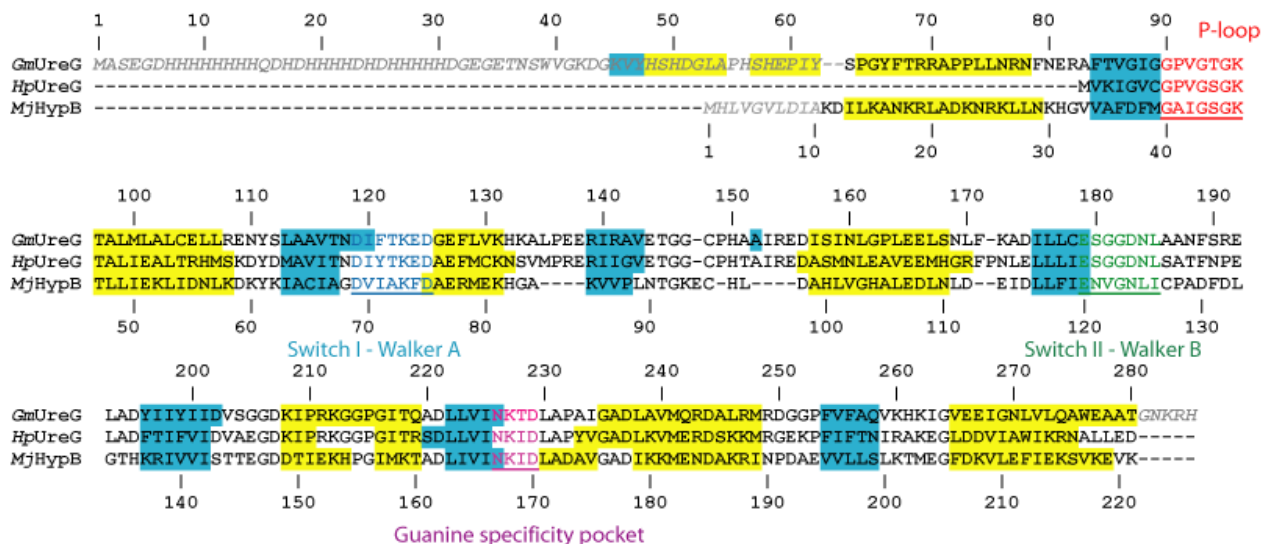


Figura 24: Alinhamento das sequências de *GmUreG*, *HpUreG* e *MjHypB*, com o conteúdo de estrutura secundária ( $\alpha$ -hélices em amarelo e folhas  $\beta$  em azul) predito pelo algoritmo JPRED para *GmUreG* e *HpUreG*, e a partir da estrutura (PDB 2HF8) para *MjHypB*; a localização das seqüências das porções funcionais das três proteínas está indicada.

O conteúdo de estruturas secundárias de *GmUreG200* e *GmUreG300* foi avaliado por espectroscopia de dicroísmo circular (CD) na faixa do UV distante. Surpreendentemente, as duas frações apresentam diferentes conteúdos de estrutura secundária (Figura 25A e B). Uma vez que as análises por espectrometria de massas confirmaram que ambas frações correspondem a mesma proteína, os espectros diferenciais de CD indicam que algum fator pode estar interferindo com o dobramento das proteínas *in vivo* e, talvez, o mesmo fator seja responsável por seu comportamento diferencial na resina de afinidade por níquel. O espectro da *GmUreG300* mostra a presença de ambos,  $\alpha$ -hélices e folhas  $\beta$ , com deflexões negativas em 208 nm e 221 nm, e um pico positivo em 190 nm, enquanto o espectro da *GmUreG200* mostrou uma deflexão única em 203 nm, típico de proteínas altamente desenoveladas.



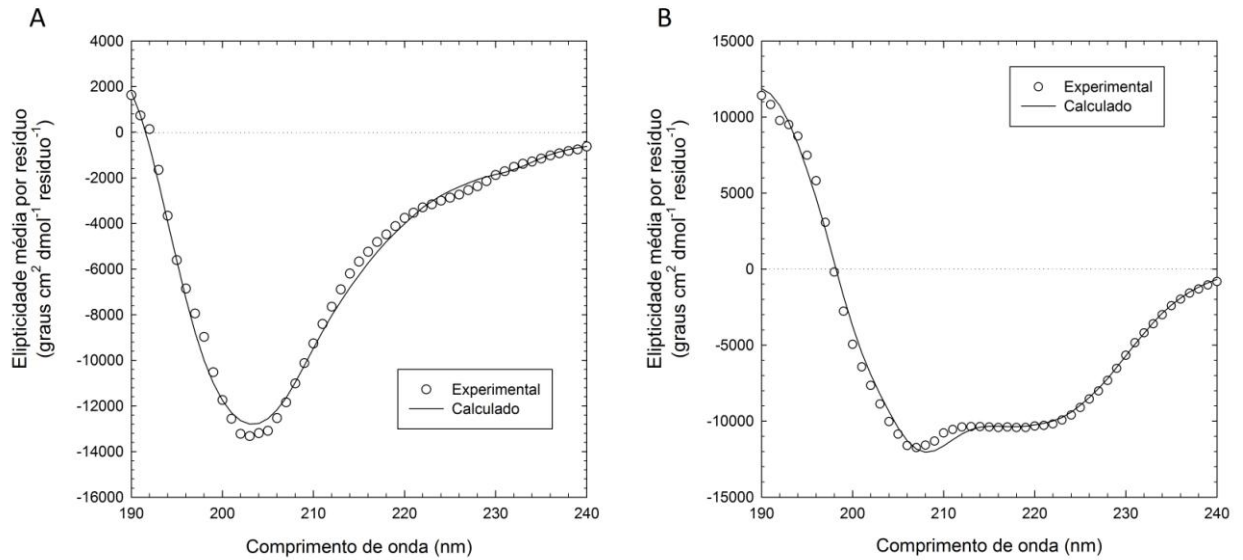


Figura 25: Análise de CD das proteínas nativas *GmUreG200* (A) e *GmUreG300* (B). Círculos abertos correspondem aos dados experimentais e as linhas sólidas correspondem ao melhor ajuste calculado.

Os espectros foram quantitativamente analisados e o melhor ajuste estimou a composição de estrutura secundária em 26 %  $\alpha$ -hélices, 18 % folhas  $\beta$ , 21 % turns e 35 % random coil (nrmsd = 6 %; programa Contin; conjunto de referência nº 4) para *GmUreG300* e em 16 %  $\alpha$ -hélices, 25 % folhas  $\beta$ , 15 % turns e 44 % random coil (nrmsd = 7 %; programa Contin; conjunto de referência SP175) para *GmUreG200*. Apesar da razoável quantidade de estrutura secundária estimada para *GmUreG200*, seu espectro é típico de uma proteína majoritariamente desenovelada. Problemas na quantificação acurada do conteúdo de estrutura secundária pelos programas atualmente disponíveis, para proteínas que apresentam grandes quantidades de estrutura desordenada, já foram observados [Whitmore e Wallace, 2008]. A qualidade na predição de estrutura secundária por tais programas depende, na maior parte, da utilização de bancos de dados de dicroísmo circular de proteínas com estrutura similar a da proteína em questão. No entanto, os bancos de dados disponíveis atualmente são, em sua maioria, compostos por espectros de proteínas que possuem estrutura cristalográfica resolvida e, portanto, apresentam um alto grau de estrutura secundária. Uma vez que proteínas desnaturadas não estão ainda adequadamente representadas nestes bancos de dados, a quantificação deste tipo de estrutura não pode ser considerada confiável. Desta forma, propomos que a quantidade de estrutura desordenada em *GmUre200* é consideravelmente maior do que o estimado. Neste cenário, é intrigante que *GmUreG200*, a forma majoritária de *GmUreG* encontrada no extrato da planta, apresente um maior conteúdo

de random coil do que *GmUreG300*. Isto pode sugerir um papel fisiológico importante deste comportamento. Ao mesmo tempo, *GmUreG300* apresenta um conteúdo de estrutura secundária mais próximo ao predito pelo algoritmo JPRED.

A *rGmUreG* também foi analisada por CD (Figura 26) e mostrou um comportamento bastante similar a *GmUreG300*. Os espectros foram quantitativamente analisados e o melhor ajuste estimou a composição de estrutura secundária em 29 %  $\alpha$ -hélices, 16 % folhas  $\beta$ , e 55 % entre turns e random coil (nrmsd = 6 %; programa K2d). Para verificar se a propriedade de ligação a metais das proteínas *UreG* pode afetar seu conteúdo de estrutura secundária, *rGmUreG* foi submetida a análise por CD na presença de  $Zn^{2+}$  e  $Ni^{2+}$ . Como mostra a Figura 26, a incubação da proteína com 10 equivalentes molares de  $Ni^{2+}$  teve pouca influência na estrutura geral da proteína, com apenas uma pequena mudança no conteúdo de estrutura secundária (36 %  $\alpha$ -hélices, 17 % folhas  $\beta$ , 22 % turns e 25 % random coil [rmsd = 2 %; programa CDSSTR; conjunto de referência n° 4]). Entretanto, quando a mesma quantidade de  $Zn^{2+}$  foi adicionada a solução de proteína, uma diminuição geral na intensidade do espectro, principalmente na região de 208 nm, foi observada. Este tipo de alteração no espectro é característico de uma diminuição substancial no conteúdo de  $\alpha$ -hélices e de um aumento de folhas  $\beta$ . Este comportamento foi confirmado pela quantificação do conteúdo de estrutura secundária para a proteína na presença de  $Zn^{2+}$ : 7 %  $\alpha$ -hélices, 25 % folhas  $\beta$  e 32 % turns e 36 % random coil (nrmsd= 3 %; programa CDSSTR; conjunto de referência n° 4). Um efeito similar foi descrito para *HpUreG*, onde o mesmo tipo de mudança na estrutura secundária devido a presença de  $Zn^{2+}$  foi evidenciado [Zambelli et al., 2009].

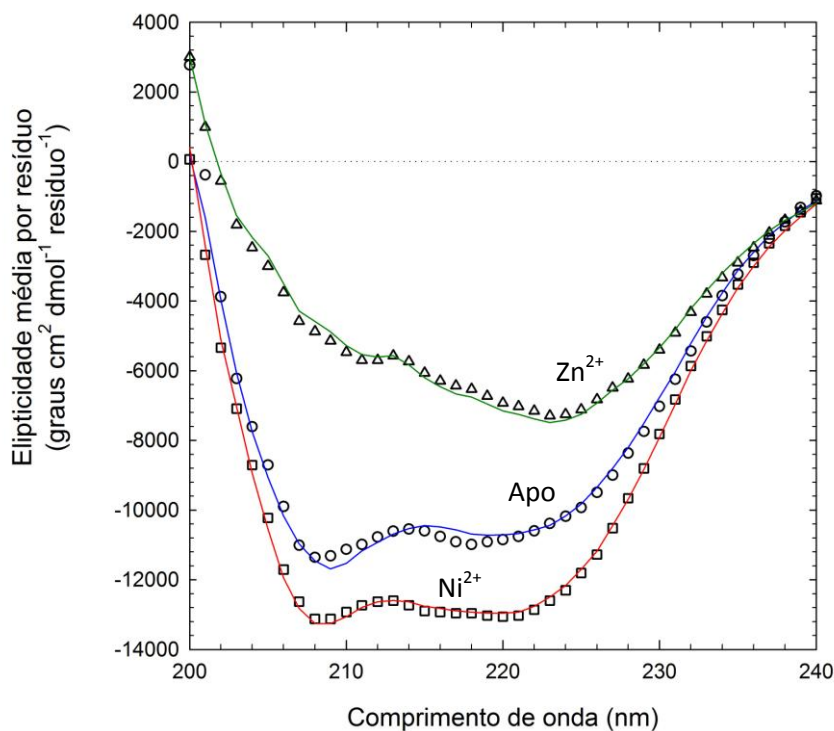


Figura 26: Análise de CD da proteína rGmUreG na ausência (círculos abertos) ou presença de 10 equivalentes de  $Zn^{2+}$  (triângulos abertos) ou  $Ni^{2+}$  (quadrados abertos). Os melhores ajustes obtidos são representados em linhas sólidas.

Apesar da estrutura secundária bem definida mostrada pelos dados de CD para rGmUreG, esta técnica não fornece nenhuma informação a respeito da estrutura terciária da proteína. Espectroscopia de ressonância nuclear magnética (RMN) foi realizada com  $^{15}N$ -rGmUreG (Figura 27). O espectro  $^1H, ^{15}N$ -HSQC de rGmUreG é caracterizado por ressonâncias pouco resolvidas, com poucas dispersões na dimensão  $^1H$  das amidas do esqueleto protéico, sendo que muitas delas caem na região correspondente a random coils (faixa de deslocamento químico de 7,6–8,5 ppm). Esta observação sugere que grandes porções do esqueleto da proteína estão passando por múltiplas mudanças de conformação, e não possuem uma estrutura terciária bem definida. Os sinais dos grupos  $NH_2$  das cadeias laterais de resíduos Asn e Gln presentes na proteína também são pouco resolvidos, gerando dois largos envelopes, centrados em 112,5 ppm ( $^{15}N$ ) e 7,4–6,7 ppm ( $^1H$ ). A falta de diferenciação entre estas ressonâncias sugere que estas cadeias laterais (Asn e Gln) não estão envolvidas em interações específicas, confirmando, portanto, a idéia de um comportamento fluxional em solução para GmUreG.

Os resultados de RMN confirmam dados publicados [Zambelli et al., 2005; 2009; 2007] relatando o caráter de alta flexibilidade dos membros da família de chaperonas UreG, mostrando que esta característica é também compartilhada por membros desta família de origem vegetal. A presença de grandes porções do esqueleto da proteína incorrendo em constantes mudanças conformacionais indica que *GmUreG* também pertence à classe de proteínas intrinsecamente desordenadas. Proteínas deste tipo transitam entre múltiplos estados de desordem e ordem, durante ou previamente, à sua função biológica [Dunker et al., 2002; Tompa, 2002; Uversky, 2002]. Esta característica pode estar relacionada ao papel fisiológico de UreGs *in vivo*, possivelmente agindo na regulação da atividade desta proteína, pois permitiria a minimização de hidrólise não desejada de GTP, a não ser que a proteína esteja corretamente inserida no complexo UreDFG-apourease, e pronta para realizar sua função na ativação da urease.

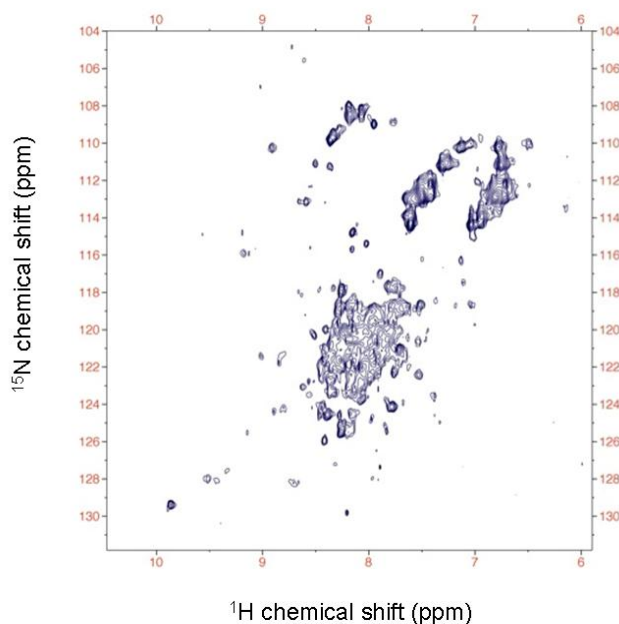


Figure 27: Espectro  $^1\text{H}$ - $^{15}\text{N}$ -HSQC de 700  $\mu\text{M}$  de apo-r*GmUreG*, adquirido em 899,07 MHz.

Para verificar o arranjo quaternário de *GmUreG* em solução, técnicas de GF e espalhamento de luz foram empregadas. O perfil de eluição de r*GmUreG* em GF é caracterizado pela presença de pico bastante amplo, com volume de retenção compatível com um intermediário entre monômero e dímero (c.a. 48 kDa) (Figura 28A). Este perfil não é alterado pela adição de agentes redutores como ditioneitol (dados não mostrados), mostrando que tal comportamento não envolve a participação de pontes dissulfeto. Estudos

anteriores relataram que *MtUreG* [Zambelli et al., 2007] e *BpUreG* [Neyroz et al., 2006] apresentam-se em solução como dímeros estabilizados por pontes dissulfeto, o que não parece ser o caso para *GmUreG*. Considerando a possibilidade de que o comportamento estrutural variável de *GmUreG* pudesse estar afetando seu perfil cromatográfico na GF, medidas de espalhamento de luz foram realizadas (MALS/QELS), confirmando uma massa média para *GmUreG* de 47,7 kDa, equivalente a um intermediário entre monômero e dímero (Figura 28B). Estes dados indicam que a proteína apresenta-se em solução em constante e rápido equilíbrio entre os estados monomérico e dimérico. Estudos prévios envolvendo a *HpUreG* [Zambelli et al., 2009] mostraram que esta proteína (geralmente monomérica em solução), é capaz de dimerizar-se na presença de íons de  $Zn^{2+}$ . Para investigar se *GmUreG* também compartilha esta característica, a proteína foi recromatografada por GF na presença de  $Zn^{2+}$ . O perfil cromatográfico de *GmUreG*, nestas condições, mostra um pico muito mais estreito e com volume de retenção compatível com a forma dimérica da proteína (64 kDa) (Figura 28A, linha pontilhada). Também se observa o aparecimento de um ombro com volume de retenção maior quando comparado ao volume de retenção da proteína livre de  $Zn^{2+}$ , que corresponde a uma porção residual da proteína em sua forma monomérica. Estes dados mostram que a dimerização de *GmUreG* não é causada por  $Zn^{2+}$ , mas sim estabilizada por sua presença, em contraste com o que foi previamente descrito para *HpUreG* [Zambelli et al., 2009]. As mudanças na estrutura secundária observadas nas análises de CD na presença de  $Zn^{2+}$  (Figura 26), indicam que o efeito estrutural da ligação do metal promove um rearranjo conformacional da proteína que provavelmente é o responsável pela estabilização da sua dimerização. Esta hipótese é corroborada pela observação de que, na presença de  $Zn^{2+}$ , o conteúdo de  $\alpha$ -hélices é reduzido de 26 para 7 %, e o de folhas  $\beta$  é aumentado de 17 para 25 %, o que representa uma composição similar à apresentada pelos apodímeros *BpUreG* (15 %  $\alpha$ -hélices e 29 % folhas  $\beta$ ) e *MtUreG* (8 %  $\alpha$ -hélices e 29 % folhas  $\beta$ ) [Zambelli et al., 2005; 2007]. O efeito de  $Ni^{2+}$  sobre *GmUreG* não pode ser avaliado, uma vez que, na concentração de proteína necessária para estes experimentos, *rGmUreG* precipita imediatamente após a adição do metal.

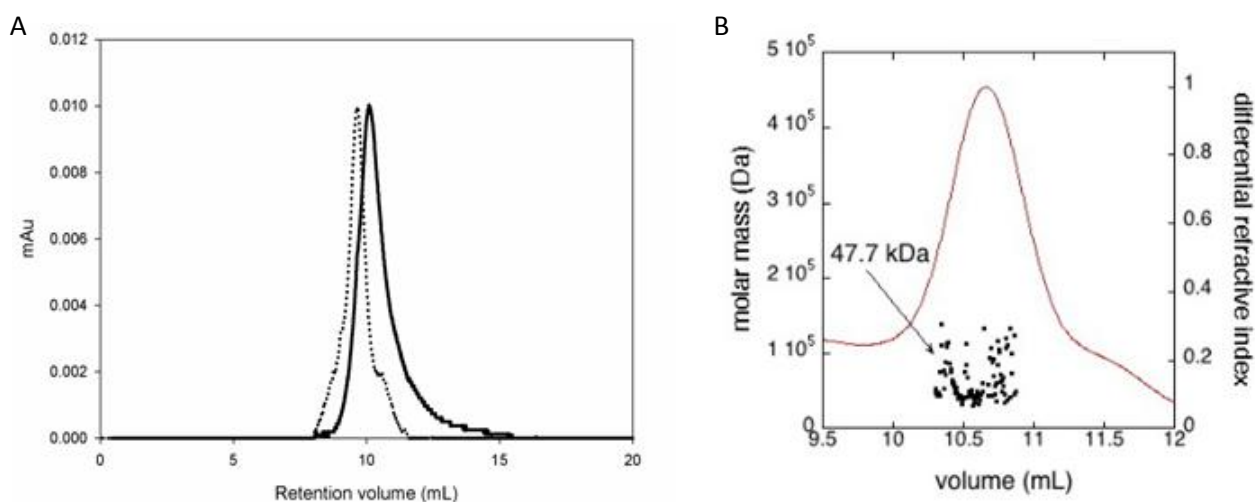


Figura 28: (A) cromatografia de gel-filtração em coluna Superdex-75 de rGmUreG na ausência (linha sólida) e presença de  $Zn^{2+}$  (linha pontilhada). (B) Gráfico da distribuição da massa molar de rGmUreG, derivado dos dados de MALS/QELS. A linha sólida indica o perfil da gel filtração em Superdex S-200, monitorado por detector de índice de refração, e os pontos são referentes à massa molecular média, medida a cada segundo. A massa molecular média resultante está indicada.

### Propriedades de ligação a metais

A capacidade de ligação a metais de UreGs bacteriana foi explorada em trabalhos anteriores mostrando que estas proteínas são capazes de ligar tanto  $Ni^{2+}$  quanto  $Zn^{2+}$  [Zambelli et al., 2005; 2007; 2008; Boer et al., 2010], característica essa compartilhada pela proteína acessória UreE [Lee et al., 1993, Brayman e Hausinger, 1996, Colpas et al., 1998; 1999; 2000]. Considerando que, em plantas, a proteína UreG parece acumular as funções de ambas proteínas bacterianas, sua capacidade de ligação à metal foi avaliada.

Lembrando que a característica de ligação a metal é o principal fator explorado para a purificação de GmUreG e rGmUreG, previamente aos ensaios de ligação o conteúdo metálico de cada proteína foi avaliado pela técnica de ICP-OES (*Inductively Coupled Plasma Optical Emission Spectroscopy*) (Anexo 1). As análises mostram que ambas frações purificadas diretamente da planta estão carregadas com  $Mg^{2+}$ , o que é um resultado esperado, considerando-se a propriedade de GTPase observada para UreGs bacterianas. GmUreG200 e rGmUreG mostraram apenas concentrações residuais de  $Ni^{2+}$  e  $Zn^{2+}$ , enquanto para GmUreG300, registrou-se a presença de  $Zn^{2+}$  (c.a. 0,3 equivalentes). A purificação da proteína ligada ao metal a partir de sua fonte natural suporta a hipótese de que  $Zn^{2+}$  seja um ligante fisiologicamente relevante. Inesperadamente, GmUreG200 parece carregar altas quantidades

de  $\text{Ca}^{2+}$  (cerca de 10 equivalentes), o que não foi observado para nenhuma outra amostra. A natureza das altas concentrações de  $\text{Ca}^{2+}$  encontradas na fração *GmUreG200* ainda não é conhecida, assim como o papel deste íon em relação às UreGs. Sabe-se que algumas GTPases, envolvidas nos mais variados processos, são afetadas por cálcio, em sua estrutura e/ou função. A proteína de espinafre PsbO (parte do complexo do Fotossistema II), também intrinsicamente desestruturada e com atividade GTPásica caracterizada [Lundin et al., 2007], sofre mudanças conformacionais induzidas por cálcio [Heredia e De Las Rivas, 2003] sendo proposto que sua interação com outras proteínas do complexo também sejam reguladas pela ligação com  $\text{Ca}^{2+}$  [Shutova et al., 2005]. A proteína FtsZ de *Escherichia coli*, uma GTPase chave no processo de divisão celular, também tem sua estrutura secundária e atividade enzimática afetadas por este metal [Jaiswal e Panda, 2009]. Entretanto, para qualquer inferência sobre a influência de íons  $\text{Ca}^{2+}$  na regulação de *GmUreG* sejam feitas, análises mais profundas serão necessárias. Do ponto de vista bioquímico, acreditamos que a presença de  $\text{Ca}^{2+}$  possa ser a razão da menor afinidade de *GmUreG200* na cromatografia de afinidade por  $\text{Ni}^{2+}$ , bem como de seu comportamento distinto em relação a *GmUreG300*, nas análises de CD.

Para melhor entender a capacidade de ligação a metais de *GmUreG*, a técnica de calorimetria de titulação isotérmica (ITC) foi empregada. Para  $\text{Zn}^{2+}$ , os dados de calor integrados (Figura 29B e C) mostram claramente a presença de três tipos distintos de eventos de ligação, terminados quando 1, 2,5 e 4 equivalentes de  $\text{Zn}^{2+}$  são adicionados à solução de proteína. Os ajustes obtidos indicam que, no primeiro evento, 0,5 equivalentes de  $\text{Zn}^{2+}$  ligam-se a cada monômero de *GmUreG*, com uma alta afinidade ( $K_{d1}(\text{Zn}) = 0,022 \pm 0,01 \mu\text{M}$ ). Esse processo é seguido por um segundo evento de ligação, onde um íon  $\text{Zn}^{2+}$  se liga a cada monômero, com menor afinidade ( $K_{d2}(\text{Zn}) = 1,31 \pm 0,05 \mu\text{M}$ ). Finalmente, 2,6 íons metálicos se ligam a cada monômero, com  $K_{d3} = 0,163 \pm 0,09 \mu\text{M}$ . No caso da titulação de  $\text{Ni}^{2+}$ , um único evento de ligação foi observado, com uma estequiometria de 2,15 equivalentes por monômero e  $K_d = 4,8 \pm 0,3 \mu\text{M}$  (Figura 29E). Todos os eventos de ligação observados são conduzidos por um fator entálpico grande e favorável, capaz de compensar os valores negativos de entropia (Tabela 6).

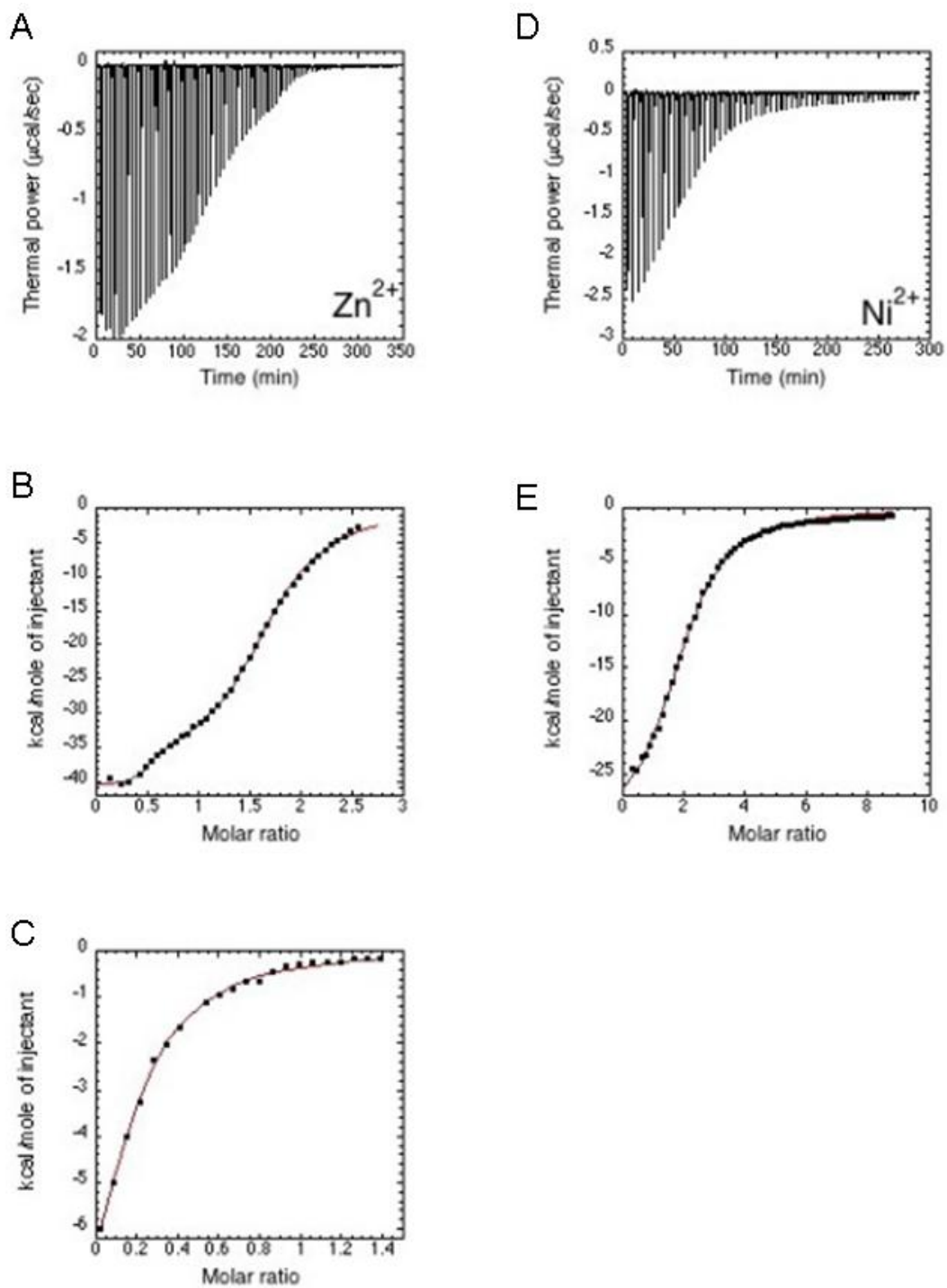


Figura 29: Gráficos representativos de dados típicos da titulação, mostrando o efeito térmico de 60 x 5  $\mu\text{L}$  injeções de  $\text{Zn}^{2+}$  (300  $\mu\text{M}$ ) (A) ou  $\text{Ni}^{2+}$  (600  $\mu\text{M}$ ) (D) na solução de *rGmUreG* (15  $\mu\text{M}$ ). Gráficos das curvas de titulação da interação entre *rGmUreG* e  $\text{Zn}^{2+}$  (B e C) ou  $\text{Ni}^{2+}$  (E), mostrando os melhores ajustes dos dados integrados (linhas sólidas). A curva de calor integrado de ligação do  $\text{Zn}^{2+}$  foi dividida nos gráficos B e C. Desta forma, o gráfico C representa a continuação do gráfico B.



Tabela 6: Parâmetros termodinâmicos da ligação de Zn<sup>2+</sup> e Ni<sup>2+</sup> à GmUreG

Metal	K <sub>gr</sub> (mM)	ΔH <sub>1</sub> (kcal.mol <sup>-1</sup> )	ΔS <sub>1</sub> (cal.mol <sup>-1</sup> .K <sup>-1</sup> )	K <sub>gr</sub> (μM)	ΔH <sub>2</sub> (kcal.mol <sup>-1</sup> )	ΔS <sub>2</sub> (cal.mol <sup>-1</sup> .K <sup>-1</sup> )	K <sub>gr</sub> (mM)	ΔH <sub>3</sub> (kcal.mol <sup>-1</sup> )	ΔS <sub>3</sub> (cal.mol <sup>-1</sup> .K <sup>-1</sup> )
Ni <sup>2+</sup>	4,8 ± 0,3	29,59 ± 0,40	-75	-	-	-	-	-	-
Zn <sup>2+</sup>	0,022 ± 0,01	40,59 ± 0,39	-101	1,31 ± 0,05	37,32 ± 0,44	-98,3	0,163 ± 0,09	12,14 ± 0,46	-9,66

Considerando que *GmUreG* apresenta um grande conteúdo de histidinas em seu N-terminal, é muito difícil determinar com exatidão quais eventos estão relacionados à ligação do metal ao domínio principal da proteína (domínio G) e quais estão relacionados à cauda de polihistidina; entretanto, fica claro que a proteína liga  $Zn^{2+}$  e  $Ni^{2+}$  de forma distinta, e algumas inferências podem ser feitas. O primeiro evento de ligação ao  $Zn^{2+}$  possui a maior afinidade já descrita para uma proteína UreG até o momento e, considerando esta característica, papéis funcionais e/ou regulatórios podem ser propostos. A ligação de 0,5 equivalentes por monômero reflete a dimerização estabilida por  $Zn^{2+}$ , uma vez que um único íon  $Zn^{2+}$  é compartilhado por dois monômeros, o que permite inferir também que este sítio de ligação está localizado na interface entre os dois monômeros. Baseando-se em publicações anteriores [Zambelli et al., 2009], acreditamos que este sítio de ligação corresponda ao motivo CPH (resíduos 148-150), altamente conservado em proteínas UreG e comprovadamente envolvido na ligação de  $Zn^{2+}$  [Zambelli et al., 2009; Boer et al., 2010]. É interessante observar que, para *GmUreG*, este sítio de ligação parece ser específico para  $Zn^{2+}$ , enquanto em UreGs bacterianas este sítio é capaz de ligar tanto  $Zn^{2+}$  quanto  $Ni^{2+}$ . É bastante provável que o terceiro evento, com a ligação de cerca de 2-3 íons de  $Zn^{2+}$  por monômero, reflita a interação dos íons com a cauda natural de polihistidina. O valor de  $K_d$  mais baixo observado para o terceiro evento de ligação em relação ao segundo pode ser o reflexo de uma superestimação, devido à divisão da curva de ligação (Figura 29B e C). Desta forma, esse dado deve ser simplesmente interpretado como uma indicação de que metais ainda ligam na proteína após o preenchimento dos dois primeiros sítios. Se o segundo evento de ligação de  $Zn^{2+}$  à proteína ocorre no domínio G ou na cauda de poli-histidinas, é uma questão para trabalhos futuros. O único evento observado para  $Ni^{2+}$  também é provavelmente relacionado ao N terminal rico em histidina. Estes dados parecem estar de acordo com o papel biológico de *GmUreG* no armazenamento de  $Ni^{2+}$ , suprimindo assim a ausência da proteína acessória UreE. Mais estudos são necessários para o entendimento completo das propriedades de ligação de metal da *GmUreG*; entretanto, se esta proteína liga  $Ni^{2+}$  somente através da sua porção N terminal, podemos inferir que *GmUreG* não esteja diretamente relacionada à entrega de  $Ni^{2+}$  ao metalocentro da urease. Em condições onde as concentrações citoplasmáticas de  $Ni^{2+}$  são limitantes, a maior afinidade por este metal deve pertencer à proteína com participação direta na sua entrega à urease, enquanto a proteína de armazenamento (*GmUreG*) seria preferencialmente carregada em sítios de baixa afinidade, com valores de  $K_d$  na faixa micromolar, como observado aqui. Esta característica também é encontrada em algumas proteínas HypB, que possuem uma extensão similar do N-terminal, rica em histidinas, como é o caso da HypB de *Rhizobium leguminosarum* (Figura 30), que também liga cerca de 2-3  $Ni^{2+}$  por monômero com afinidade semelhante na faixa de

micromolar [Rey et al., 1994]. Ainda assim, a interpretação dos dados termodinâmicos deve ser feita com cautela, uma vez que a construção do sítio ativo da urease é um processo dinâmico que envolve também a participação de outras proteínas acessórias, além da clivagem de GTP.

```

GmUreG  MASEGDHHHHHHHQQDHDHHHHDHDHHHHDGEGETNSWVGKDGKVVYHSFDGLAPHSHEPIYSPGYFTRR
BjHypB  MCTVCGCSDGKASIEHAHDHHHDHGHDDHGHGDGHHHHHGHQDHHHHDHAHGDAGLLDCGANPAGQK
RlHypB  MCTVCGCGTSAIEGTHHEVGDDGHHHHHDGHHDDHDDHHRGDHEHDDHHAEDGSVHYSKGIAGVIV
  
```

Figura 30: Comparação da sequência N terminal de *GmUreG*, *BjHypB* e *RlHypB*, mostrando a extensão rica em histidina.

### Produção de uma forma truncada de *GmUreG*

Algumas questões a respeito das propriedades de ligação a metal observadas para *GmUreG* ainda permanecem em aberto, como em relação ao segundo evento de ligação ao  $Zn^{2+}$ . Se este evento ocorre no domínio G ou não, e qual seria sua função para a proteína, são questões a serem respondidas em estudos futuros envolvendo diferentes formas truncadas desta proteína. O primeiro passo nesta direção foi dado com a produção de uma forma truncada de *GmUreG*. Esta forma não apresenta os primeiros 56 resíduos de aminoácidos de *GmUreG* (denominada portanto *GmUreG* $\Delta$ 56), compreendendo a porção N terminal rica em histidinas. O cDNA (690 pb) foi amplificado por PCR (Figura 31) e clonado no vetor pET15b que insere nesta proteína uma cauda 6-His removível por ação da trombina. Esta proteína é solúvel frente a superexpressão e não forma agregados após sua purificação, em contraste com o observado para a *GmUreG* tipo selvagem, portanto sendo, aparentemente, mais estável que a proteína completa.

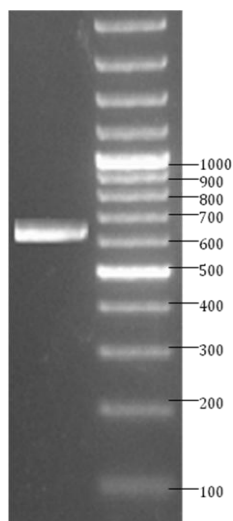


Figura 31: Amplificação de *GmUreG*Δ56 por PCR com primers específicos. Gel de agarose (1,2%). Canaleta 1: reação de PCR com primers específicos para *GmUreG*Δ56; Canaleta 2: Ladder, com tamanho indicado em pares de bases.

O perfil cromatográfico de *GmUreG*Δ56 em coluna Superdex-75 (Figura 32), indica que a maior parte da fração purificada apresenta-se na forma dimérica, e apenas uma pequena porção na forma monomérica (padrão não alterado mesmo após alguns dias de armazenamento). Este comportamento contrasta com aquele observado para a proteína completa, e a causa desta dimerização ainda não foi estudada. Esta diferença de comportamento observada entre a proteína tipo selvagem e a proteína truncada é motivo de preocupação, pois poderia refletir algum evento de dobramento incorreto da proteína, tornando-a inútil para estudos adicionais sobre o comportamento de *GmUreG* na ausência de seu N terminal rico em histidinas. Neste estágio do estudo, o processo de clivagem da cauda de histidina por trombina está sendo otimizado.

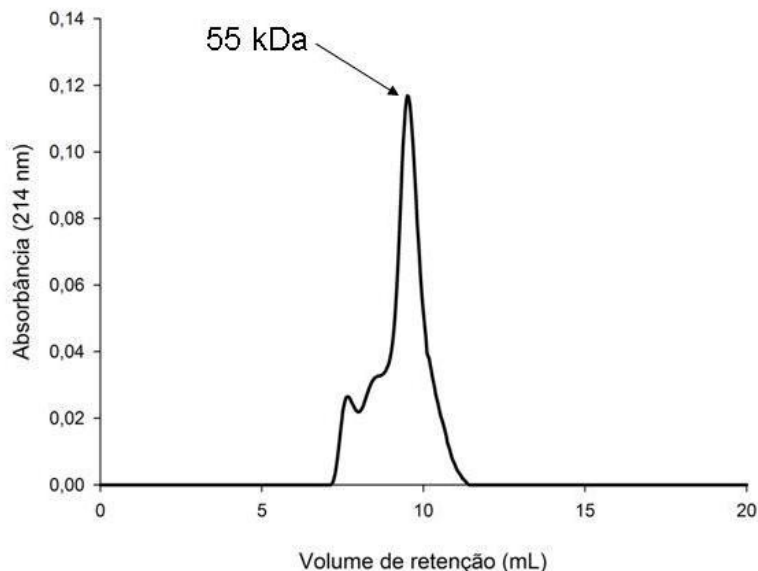


Figura 32: Perfil de gel-filtração de *GmUreG*Δ56, em coluna Superdex-75. A massa molecular estimada está indicada na figura.

### Atividade GTPásica

Proteínas UreG estão envolvidas na hidrólise de GTP, que ocorre concomitantemente à carboxilação de um resíduo de lisina no sítio ativo da urease [Park e Hausinger, 1995; Soriano e Hausinger, 1999]. O motivo P-loop de ligação a nucleotídeos é bastante conservado entre os membros da família de UreGs. A principal diferença entre as UreGs bacterianas e vegetais, com relação a este motivo, é a troca conservativa de uma serina por treonina, observada em *GmUreG* (Figura 24), nas UreGs de *A. thaliana* e *S. leguminosarum* [Witte et al., 2005; 2001], e também em outras proteínas G [Sprang, 1997]. A atividade GTPásica já foi demonstrada para *BpUreG* e *MtUreG* [Zambelli et al., 2005; 2007], enquanto *KaUreG* e *HpUreG* possuem atividade GTPásica abaixo dos níveis de detecção [Zambelli et al., 2009; Moncrief e Hausinger, 1997]. Neste trabalho, a atividade de *GmUreG* foi avaliada em ensaios utilizando uma concentração de GTP cerca de 3 ordens de magnitude maior do que a concentração da enzima, o que permite que valores de  $k_{cat}$  possam ser obtidos a partir de uma curva de atividade em relação ao tempo. *GmUreG*200, *GmUreG*300 e *rGmUreG* apresentam valores de  $k_{cat}$  de 0,017, 0,018 e 0,013  $\text{min}^{-1}$ , respectivamente (Figura 33A). Estes valores são comparáveis entre si e similares ao observado para *MtUreG* ( $k_{cat}$  de 0,01  $\text{min}^{-1}$  [Zambelli et al., 2007]), mas são menores do que o observado para *BpUreG* ( $k_{cat}$  de 0,04  $\text{min}^{-1}$  [Zambelli et al., 2005]). Esta

atividade pode ser considerada bastante baixa, quando comparadas a GTPases semelhantes, como as HypB de *B. japonicum* ( $k_{cat}$  de  $0,18 \text{ min}^{-1}$  [Fu et al., 1995]) e *E. coli* ( $k_{cat}$  de  $0,17 \text{ min}^{-1}$  [Maier et al., 1993]). As constantes mudanças conformacionais sofridas por *GmUreG* poderiam explicar sua baixa atividade, uma vez que na ausência dos cofatores apropriados, o dobramento correto da proteína não pode ser estabilizado, e somente uma pequena fração das moléculas estariam transitando por uma forma enzimaticamente favorável em um determinado ponto no tempo. Como discutido anteriormente, este comportamento é provavelmente benéfico, pois minimizaria a hidrólise desnecessária de GTP, até que ocorra a correta montagem do complexo de ativação da urease.

O efeito de  $\text{Zn}^{2+}$  e  $\text{Ni}^{2+}$  foi testado, em condições de excesso molar destes metais (25 vezes), em um ensaio de 6 h.  $\text{Ni}^{2+}$  não mostrou qualquer efeito significativo na atividade das proteínas nativas ou recombinante (Figura 33B, C e D).  $\text{Zn}^{2+}$ , entretanto, aumentou cerca de 89 e 115% a atividade de *GmUreG300* e *rGmUreG*, respectivamente (Figura 33C and D), enquanto o efeito sobre *GmUreG200* foi negligenciável (Figura 33B). As medidas de ICP mostraram que *GmUreG300* carrega  $\text{Zn}^{2+}$  após sua purificação a partir da planta, no entanto, em um quantidade baixa e que não parece ser suficiente para afetar a atividade GTPásica. Na presença de concentrações saturantes de  $\text{Zn}^{2+}$ , *GmUreG300* e *rGmUreG* atingem níveis similares de atividade. Estes resultados podem ser explicados em termos dos níveis de organização da proteína em presença de metais.  $\text{Zn}^{2+}$  parece estabilizar a dimerização de *GmUreG*, o que obviamente reduz a liberdade da proteína de transitar por diferentes conformações, tornando o evento de aquisição de seu dobramento correto mais favorável. Entretanto, a atividade na presença de  $\text{Zn}^{2+}$  ainda é bastante baixa, indicando que este metal, sozinho, não é capaz de estabilizar completamente o dobramento da proteína. As características moleculares descritas até agora para as proteínas *UreG* indicam que muitos fatores podem estar envolvidos na regulação de sua atividade GTPásica, como as transições estruturais, dimerização, ligação à metais e formação do complexo multiproteico [Zambelli et al., 2007]. A falta de efeito observada para *GmUreG200* é intrigante, e uma possível explicação envolve a presença de  $\text{Ca}^{2+}$ , que poderia, de alguma forma, estar interferindo na interação do  $\text{Zn}^{2+}$  com a proteína. É importante ressaltar que nenhuma mudança na atividade GTPásica da *BpUreG* foi observada quando  $\text{Zn}^{2+}$  foi adicionado à reação, porém, é preciso ter em mente que, naquele caso, a proteína já apresentava-se em sua forma dimérica (estabilizada por pontes dissulfeto) e apresentava uma maior (porém comparável) atividade do que *GmUreG* apresenta, mesmo após da adição de  $\text{Zn}^{2+}$  ( $\text{Zn}^{2+}$ -*rGmUreG*  $k_{cat} = 0,028 \text{ min}^{-1}$ ;  $\text{Zn}^{2+}$ -*GmUreG300*  $k_{cat} = 0,034 \text{ min}^{-1}$ ; *BpUreG*  $k_{cat} = 0,04 \text{ min}^{-1}$ ) [Zambelli et al., 2005]. Já na presença de  $\text{Ni}^{2+}$ , a

estrutura secundária da proteína é apenas levemente mais organizada, e considerando os dados de ITC, essa maior organização da proteína possivelmente ocorre no N terminal rico em histidinas. Além disso, apesar da impossibilidade de avaliação do efeito do  $\text{Ni}^{2+}$  sobre o arranjo quaternário de *GmUreG*, os dados de ITC sugerem que a proteína não compartilha nenhum íon  $\text{Ni}^{2+}$ , o que indica que a dimerização *GmUreG* não é estabilizada por este metal. Essa hipótese é suportada por dados publicados para *KaUreG* e *HpUreG*, onde nenhum efeito na estrutura quaternária destas proteínas é observado na presença de  $\text{Ni}^{2+}$  [Boer et al., 2010, Zambelli et al., 2009]. Os dados mostrados aqui suportam o papel de  $\text{Zn}^{2+}$  como um dos cofatores envolvidos na estabilização da conformação correta de *GmUreG* durante o processo de ativação da urease.

Apesar de *GmUreG* $\Delta$ 56 apresentar-se como um dímero em solução, em contraste com o observado para *GmUreG*, a proteína truncada parece ser funcional, apresentando atividade GTPásica nos mesmos níveis da proteína tipo selvagem (Figura 33E). Isso indica que, apesar do comportamento diferencial em solução, a proteína truncada é adequada para caracterizações adicionais de *GmUreG*.

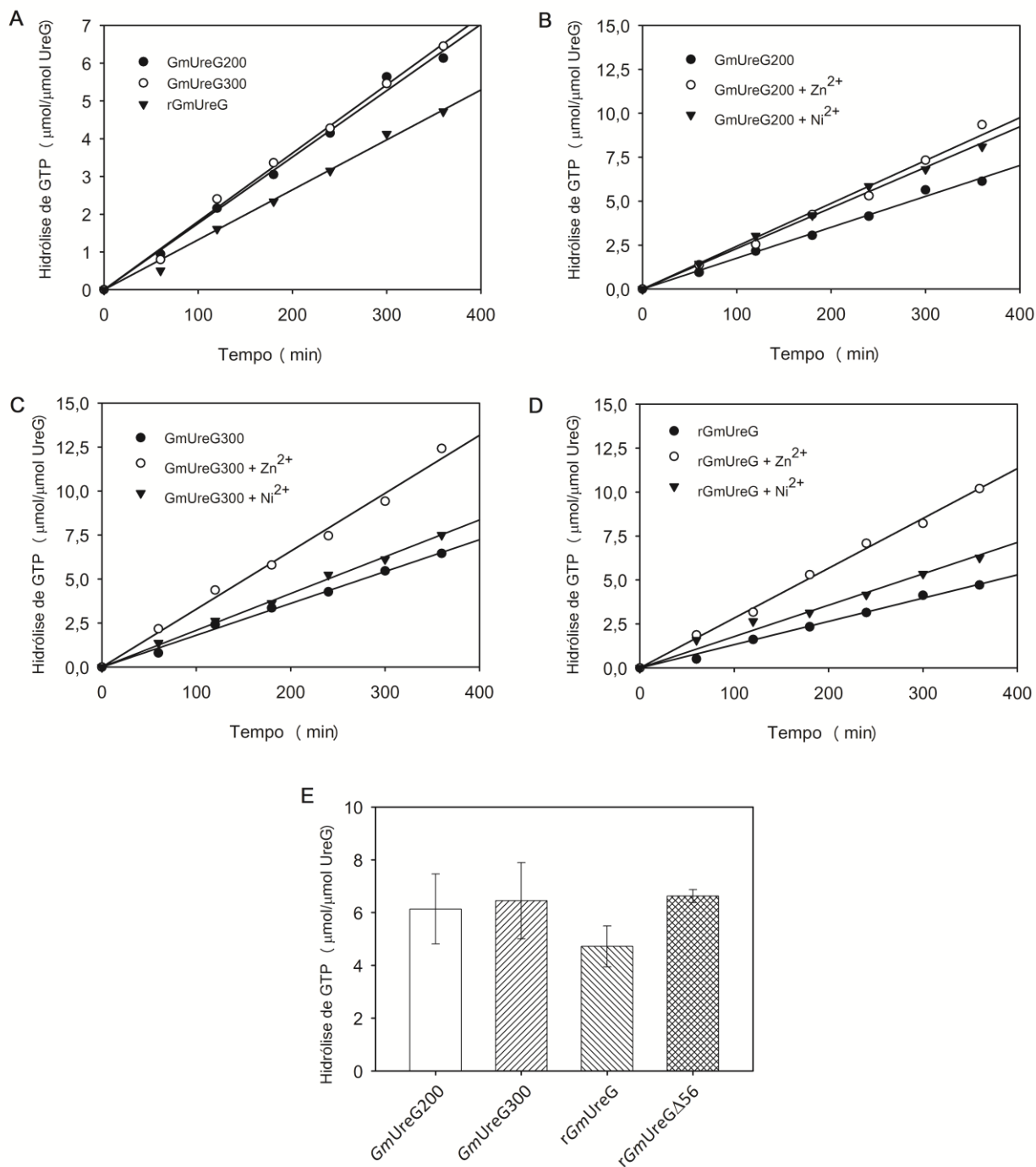


Figura 33: Atividade enzimática de *GmUreG*. (A) Curva de tempo da atividade GTPásica de *GmUreG200* (círculos fechados), *GmUreG300* (círculos abertos) e *rGmUreG* (triângulos). Curva de tempo da atividade GTPásica na presença de 125 μM Zn<sup>2+</sup> ou Ni<sup>2+</sup> para (B) *GmUreG200*, (C) *GmUreG300* e (D) *rGmUreG*. Atividade GTPásica das diferentes proteínas, em 6h (E).



## Modelagem Molecular

O alinhamento de *GmUreG* e *MjHypB*, resultante de uma análise de sequência e estrutura secundária das duas proteínas (Figura 24) mostrou uma boa correspondência das regiões funcionalmente ativas em *MjHypB*, como o domínio P-loop de ligação a GTP GAIGSGK na região N terminal (correspondendo a GPVGTGK em *GmUreG*), o motivo conservado NKTD, envolvido no reconhecimento específico de guaninas, os resíduos Thr, Asp e Glu que ligam GTP-complexado a  $Mg^{2+}$ , e os motivos DVIAKFD “switch-I” e ENVGNLI “switch-II” (correspondendo a DIFTKED e ESGGDNL em *GmUreG*), propostos como importantes para mudanças conformacionais funcionalmente essenciais que ocorrem durante a ligação ao GTP [Shimizu et al., 2000].

O modelo corresponde ao estado completamente enovelado de *GmUreG*, calculado a partir do alinhamento da Figura 24, constitui uma proteína globular formada por uma folha  $\beta$  composta por sete fitas paralelas, flanqueada por 10  $\alpha$ -hélices (Figura 34A). Como no caso do modelo de estrutura de *HpUreG* [Zambelli et al., 2009], a estrutura de *GmUreG* reproduz perfeitamente o molde *MjHypB*, não somente na arquitetura geral, mas também no arranjo de características funcionais essenciais, como as regiões do P-loop (também chamado “Walker-A”) e do “switch-II” (também chamado “Walker-B”). A hélice anfipática encontrada no N-terminal de *MjHypB*, que possui um padrão conservado de resíduos positivamente carregados e, portanto, proposta como importante nas interações proteína-proteína [Gasper et al., 2006] está presente na sequência de *GmUreG*, apesar de estar ausente em *UreGs* bacterianas [Zambelli et al., 2007; 2009]. A região rica em histidinas, que compreende os primeiros 62 resíduos da proteína, não foi modelada, devido a ausência de informações estruturais, assim como os últimos 5 resíduos do C terminal. Com base no modelo da estrutura de *GmUreG* (Figura 34A) e na acessibilidade diferencial ao solvente dos resíduos de amino ácidos quando na forma monomérica ou dimérica, é possível observar que a interface entre os dois monômeros parece ser composta por sete loops e três hélices. Estas regiões incluem os “switch-I” e “switch-II”, e o sítio de reconhecimento de guanina.

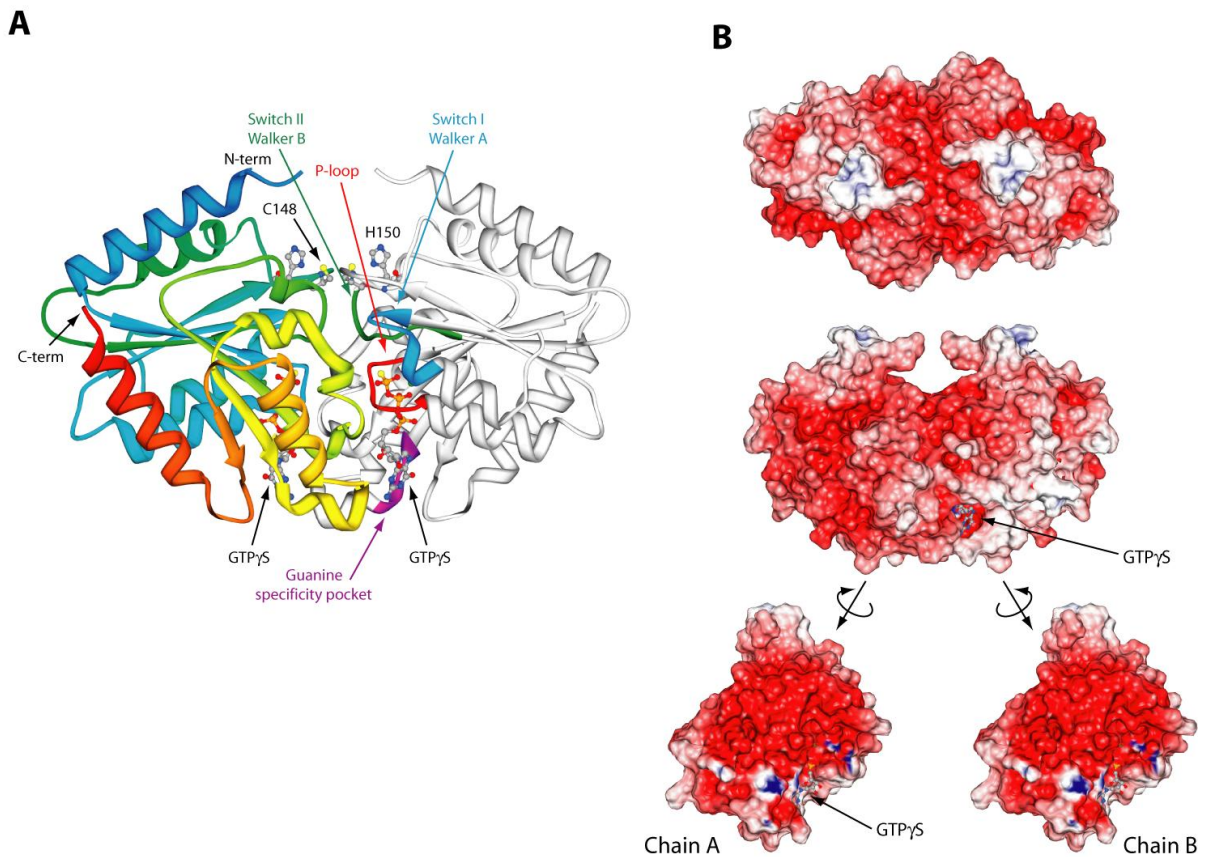


Figura 34: (A) Representação em fitas do modelo de estrutura dimérica de *GmUreG*, obtido utilizando o alinhamento de sequências mostrado na Figura 15. Na cadeia polipeptídica da esquerda, as fitas foram coloridas do azul, na porção N terminal, até o vermelho, no C terminal, enquanto na cadeia polipeptídica da direita as principais características funcionais da proteína estão coloridas em vermelho (P-loop), azul (switch-I), verde escuro (switch-II) e violeta (sítio de reconhecimento de guanina). A molécula de GTP $\gamma$ S e o íon Mg $^{2+}$  estão representados por bolas e traços coloridos, de acordo com o tipo de átomo (código CPK). (B): Superfícies coloridas com base no potencial eletrostático. A superfície exibida no painel do meio está na mesma orientação da molécula em (A), enquanto no painel superior ela foi girada 90° no eixo horizontal. O painel inferior mostra as duas metades do dímero, obtidas pela rotação de cada cadeia do monômero em 90°, mostrando a superfície de interação entre os monômeros e sua simetria C2.

A distribuição do potencial eletrostático na superfície acessível ao solvente mostra que *GmUreG* contém um extensa área de carga negativa, principalmente na zona central da interface monômero-monômero (Figura 34B, painel superior), enquanto uma área menos carregada é encontrada no lado oposto (Figura 34B, painel do meio) e na hélice N-terminal. Regiões de superfície carregadas positivamente correspondem aos resíduos Lys e Arg expostos ao solvente, com exceção da Lys227, encontrada na superfície interna do motivo conservado NKTD de reconhecimento de guanina (Figura 34B, painel inferior).

Este trabalho com *GmUreG* representa a primeira caracterização detalhada de uma proteína acessória da urease de planta. Estudos envolvendo formas truncadas e mutantes de *GmUreG* serão realizados para caracterização mais aprofundada dos diferentes aspectos desta proteína.

## Conclusão Geral e Perspectivas

Neste trabalho foi possível constatar a importância de técnicas alternativas de cristalização como a modificação química de amino ácidos, na obtenção de cristais de proteínas difíceis, como a urease de *Canavalia ensiformis*, que até poucos meses poderia ser considerada uma das proteínas mais resistente a estudos estruturais.

O processo de ativação das ureases, ainda tão pouco compreendido, principalmente para as enzimas de origem vegetal, foi estudado e uma caracterização bioquímica e estrutural de proteínas acessórias desta natureza foi realizada.

Pela primeira vez uma proteína UreG foi caracterizada quando purificada a partir de seu organismo de origem, fornecendo informações sobre seu possível comportamento *in vivo*. Apesar de manter um grande grau de semelhança em diversos aspectos com UreGs bacterianas, *GmUreG* apresenta algumas características distintas em sua propriedades. Comprovamos que esta proteína é uma GTPase, e também pertence a classe de proteínas intrinsecamente desordenadas. No entanto, sua capacidade de ligação a metais parece ser distinta do já observado para proteínas similares.

A obtenção de uma proteína UreF tipo selvagem em sua forma solúvel, fato até então não descrito na literatura, representa um grande passo em direção ao entendimento do papel desta proteína na ativação da urease, permitindo agora que sua caracterização seja realizada.

Este trabalho representa os primeiros passos no estabelecimento de nosso grupo como referência no estudo de proteínas acessórias de ureases de origem vegetal, e trás consigo uma série de possibilidades na investigação deste tema.

## Referências Bibliográficas

1. Atanassova, A., Zamble, D.B. *Escherichia coli* HypA is a zinc metalloprotein with a weak affinity for nickel. *Journal of Bacteriology*, 2005; 187, 4689–4697.
2. Bai, Y., Auperin, T.C., Tong, L. The use of *in situ* proteolysis in the crystallization of murine CstF-77. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 2007; 63, 135-138.
3. Bailey, C.J., Boulter, D. Subunit structure of jack-bean urease. *Biochemical Journal*, 1969; 113, 669-677.
4. Balasubramanian, A., Ponnuraj, K. Purification, crystallization and preliminary X-ray analysis of urease from jack bean (*Canavalia ensiformis*). *Acta Crystallographica Section F: Structural Biology and Crystallography Communication*, 2009; 65, 949-951.
5. Balasubramanian, A., Ponnuraj, K., Crystal structure of the first plant urease from jack bean: 83 years of journey from its first crystal to molecular structure. *Journal of Molecular Biology*, 2010; 400, 274-283.
6. Becker-Ritt, A.B., Martinelli, A.S., Mitidieri, S., Feder, V., Wassermann, G.E., Santi, L., Vaistein, M.H., Oliveira, J.A., Fiuza, L.M., Pasquali, G., Carlini, C.R. Antifungal activity of plant and bacterial ureases. *Toxicon*, 2007; 50, 971-983.
7. Bellucci, M., Zambelli, B., Musiani, F., Turano, P., Ciurli, S. Helicobacter pylori UreE, a urease accessory protein: specific Ni(2+)- and Zn(2+)-binding properties and interaction with its cognate UreG. *Biochemical Journal*, 2009; 422, 91-100.
8. Benini, S., Ciurli, S., Rypniewski, W.R., Wilson, K.S., Mangani, S. The complex of *Bacillus pasteurii* urease with  $\beta$ -mercaptoethanol from X-ray data at 1.65 Å resolution. *Journal of Biological Inorganic Chemistry*, 1998; 3, 268-273.
9. Benini, S., Rypniewski, W.R., Wilson, K.S., Miletti, S., Ciurli, S., Mangani, S. The complex of *Bacillus pasteurii* urease with acetohydroxamate anion from X-ray data at 1.55 angstrom resolution. *Journal of Biological Inorganic Chemistry*, 2000; 5, 110-118.
10. Bennett, J., Wren, E.A. Interaction of para-nitrophenyl carbamate with urease. *Biochimica et Biophysica Acta*, 1977; 482, 421-426.
11. Blakeley, R.L., Dixon, N.E., Zerner, B. Jack-bean urease (EC 3.5.1.5) - quaternary structure, contribution of light-scattering to the absorption-spectrum, and coordination of beta-mercaptoethanol to the active-site nickel ion(s). *Proceedings of the Australian Biochemical Society*, 1981; 14, 13-13.
12. Blakeley, R.L., Webb, E.C., Zerner, B. Jack Bean urase (EC 3.5.1.5). A new purification and reliable rate assay. *Biochemistry*, 1969. 8, 1984-1990.
13. Blokesch, M., Rohmoser, M., Rode, S., Bock, A. HybF, a zinc-containing protein involved in NiFe hydrogenase maturation. *Journal of Bacteriology*, 2004; 186, 2603–2611.
14. Boer, J., Quiroz-Valenzuela, S., Anderson, K.L., Hausinguer, R.P. Mutagenesis of *Klebsiella aerogenes* UreG to probe nickel binding and interactions with other urease-related proteins. *Biochemistry*, 2010; 49, 5859-5869.
15. Brayman, T.G., Hausinger, R.P. Purification characterization, and functional analysis of a truncated *Klebsiella aerogenes* UreE urease accessory protein lacking the histidine-rich carboxyl terminus. *Journal of Bacteriology*, 1996; 178, 5410-5416.
16. Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., Darst, S.A. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Molecular Cell*, 2002; 9, 527-539.

17. Cao, F.Q., Werner, A.K., Dahncke, K., Romeis, T., Liu, L.H., Witte, C.P. Identification and characterization of proteins involved in rice urea and arginine catabolism. *Plant Physiology*, 2010; 154, 98-108.
18. Carlini, C.R., Oliveira, A.E., Azambuja, P., Xavier-Filho, J., Wells, M.A. Biological effects of canatoxin in different insect models: evidence for a proteolytic activation of the toxin by insect cathepsin-like enzymes. *Journal of Economical Entomology*, 1997; 90, 340–348.
19. Carlini, C.R., Polacco, J.C. Toxic properties of ureases. *Crop Science*, 2008; 48, 1665-1672
20. Carter, E.L., Flugga, N., Boer, J.L., Mulrooney, S.M., Hausinger, R.P. Interplay of metal ions and urease. *Metallomics*, 2009; 1, 207-221.
21. Carter, E.L., Hausinger, R.P. Characterization of the *Klebsiella aerogenes* urease accessory protein UreD in fusion with the maltose binding protein. *Journal of Bacteriology*, 2010; 192, 2294-2304.
22. Chang, Z., Kuchar, J., Hausinger, R.P. Chemical Cross-linking and Mass Spectrometric Identification of Sites of Interaction for UreD, UreF, and Urease. *Journal of Biological Chemistry*, 2004; 279, 15305–15313.
23. Christendat, D., Yee, A., Dharamsi, A., Kluger, Y., Gerstein, M., Arrowsmith, C.H., Edwards, A.M. Structural proteomics: prospects for high throughput sample preparation. *Progress in Biophysics and Molecular Biology*, 2000; 73, 339-345.
24. Chruszcz, M., Wlodawer, A., Minor, W. Determination of protein structures - A series of fortunate events. *Biophysical Journal*, 2008; 95, 1-9.
25. Ciurli, S., Safarov, N., Miletti, S., Dikiy, A., Christensen, S.K., Kornetzky, K., Bryant, D.A., Vandenberghe, I., Devreese, B., Samyn, B., Remaut, H., Van Beeumen, J. Molecular characterization of *Bacillus pasteurii* UreE, a metal-binding chaperone for the assembly of the urease active site. *Journal of Biological Inorganic Chemistry*, 2002; 7, 623–631.
26. Colpas, G.J., Brayman, T.G., McCracken, J., Pressler, M.A., Babcock, G.T., Ming, L.J., Colangelo, C.M., Scott, R.A., Hausinger, R.P. Spectroscopic characterization of metal binding by *Klebsiella aerogenes* UreE urease accessory protein. *Journal of Biological Inorganic Chemistry*, 1998; 3, 150-160.
27. Colpas, G.J., Brayman, T.G., Ming, L.J., Hausinger, R.P. Identification of metal-binding residues in the *klebsiella aerogenes* urease nickel metallochaperone, UreE. *Biochemistry*, 1999; 38, 4078–4088.
28. Colpas, G.J., Hausinger, R.P. *In Vivo* and *in Vitro* Kinetics of Metal Transfer by the *Klebsiella aerogenes* Urease Nickel Metallochaperone, UreE. *Journal of Biological Chemistry* 2000; 275, 10731-10737.
29. Contaxis, C.C., Reithel, F.J. Studies on protein multimers. II. A study of the mechanism of urease dissociation in 1,2-propanediol: comparative studies with ethylene glycol and glycerol. *Journal of Biological Chemistry*, 1971; 246, 677-685.
30. Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., Barton, G.J. JPred: a consensus secondary structure prediction server. *Bioinformatics*, 1998; 14, 892-893.
31. Demartini, D.R., Carlini, C.R., Thelen, J.J. Global and targeted proteomics in developing jack bean (*Canavalia ensiformis*) seedlings: an investigation of urease isoforms mobilization in early stages of development. *Plant Molecular Biology*, 2011; 75, 53-65.
32. Derewenda, Z.S., Vekilov, P.G. Entropy and surface engineering in protein crystallization. *Acta Crystallographica D*, 2006; 62, 116-124.

33. Dixon, N.E., Gazzola, C., Blakeley, R.L., Zerner B Jack-bean urease (EC 3.5.1.5) - Metalloenzyme - simple biological role for nickel. *Journal of the American Chemical Society*, 1975; 97, 4131-4133.
34. Dixon, N.E., Riddles, P.W., Gazzola, C., Blakeley, R.L., Zerner, B. Jack bean urease (ec-3515) .4. the molecular-size and the mechanism of inhibition by hydroxamic acids - spectrophotometric titration of enzymes with reversible inhibitors. *Canadian Journal of Biochemistry*, 1980; 58, 1323-1334.
35. Dong, A.P., Xu, X.H., Edward, A.M. In situ proteolysis for protein crystallization and structure determination. *Nature Methods*, 2007; 4, 1019-1021.
36. Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry*, 2002; 41, 6573-6582.
37. Edwards, A.M., Arrowsmith, C.H., Christendat, D., Dharamsi, A., Friesen, J.D., Greenblatt, J.F., Vedadi, M. Protein production: feeding the crystallographers and NMR spectroscopists. *Nature Structural Biology*, 2000; 7, 970-972.
38. Ferreira-DaSilva, C.T., Gombarovits, M.E., Masuda, H., Oliveira, C.M., Carlini, C.R. Proteolytic activation of canatoxin, a plant toxic protein, by insect cathepsin-like enzymes. *Archives of Insect Biochemistry and Physiology*, 2000; 44, 162-171.
39. Fiske, C.H., SubbaRow, Y. The Colorimetric Determination of Phosphorus *Journal of Biological Chemistry*, 1925; 66, 375–400.
40. Follmer, C., Barcellos, G.B., Zingali, R.B., Machado, O.L., Alves, E.W., Barja-Fidalgo, C., Guimaraes, J.A., Carlini, C.R. Canatoxin, a toxic protein from jack beans (*Canavalia ensiformis*), is a variant form of urease (EC 3.5.1.5): biological effects of urease independent of its ureolytic activity. *Biochemical Journal*, 2001; 15, 217-224.
41. Follmer, C., Pereira, F.F.V., Silveira, N.P., Carlini, C.R. Jack bean Urease (EC 3.5.1.5) aggregation monitored by dynamic and static light scattering. *Biophysical Chemistry*, 2004b; 111, 79 - 87.
42. Follmer, C., Real-Guerra, R., Wasserman, G. E., Olivera-Severo, D. and Carlini, C. R. Jack bean, soybean and *Bacillus pasteurii* ureases: biological effects unrelated to ureolytic activity. *European Journal of Biochemistry*, 2004a; 271, 1357– 1363.
43. Freyermuth, S.K., Bacanamwo, M., Polacco, J.C. The soybean *Eu3* gene encodes an Ni-binding protein necessary for urease activity. *Plant Journal*, 2000; 21, 53-60.
44. Fu, C. L., Olson, J.W., Maier, R. J. HypB Protein of *Bradyrhizobium-japonicum* is a metal-binding GTPase capable of binding 18 divalent nickel ions per dimer. *Proceedings of the National Academy of Science USA*, 1995; 92, 2333-2337.
45. Gasper, R., Scrima, A., Wittinghofer, A. Structural insights into HypB, a GTP-binding protein that regulates metal binding. *Journal of Biological Chemistry*, 2006; 281, 27492–27502.
46. Gaur, R.K., Kupper, M.B., Fischer, R., Hoffmann, K.M., Preliminary X-ray analysis of a human V-H fragment at 1.8 angstrom resolution. *Acta Crystallographica D*, 2004; 60, 965-967.
47. Ge, R., Watt, R.M., Sun, X., Tanner, J.A., He, Q.Y., Huang, J.D., Sun, H. Expression and characterization of a histidine-rich protein. Hpn: potential for Ni<sup>2+</sup> storage in *Helicobacter pylori*. *Biochemical Journal*, 2006a; 393, 285-293.
48. Ge, R., Zhang, Y., Sun, X., Watt, R.M., He, Q.Y., Huang, J.D., Wilcox, D.E., Sun, H. Thermodynamic and kinetic aspects of metal binding to the histidine-rich protein, Hpn. *Journal of American Chemical Society*, 2006b; 128, 11330-11331.

49. Goldraij, A., Beamer, L.J., Polacco, J.C. Interallelic Complementation at the Ubiquitous Urease Coding Locus of Soybean. *Plant Physiology*, 2003; 132, 1801–1810.
50. Gorin, G., Mamiya, G., Chin, C.C. Urease and its interaction with sodium dodecyl sulfate. *Experientia*, 1967; 23, 443-445.
51. Grosseohme, N.E., Mulrooney, S.B., Hausinger, R.P., Wilcox, D.E. Thermodynamics of Ni<sup>2+</sup>, Cu<sup>2+</sup>, and Zn<sup>2+</sup> binding to the urease metallochaperone UreE. *Biochemistry*, 2007; 46, 10506–10516.
52. Ha, N.C., Oh, S.T., Sung, J.Y., Cha, K.A., Lee, M.H., Oh, B.H. Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. *Nature Structural Biology*, 2001; 8, 505-509.
53. Heredia, P., De Las Rivas, J. Calcium-dependent conformational change and thermal stability of the isolated PsbO protein detected by FTIR spectroscopy. *Biochemistry*, 2003; 42, 11831-11833.
54. Hirai, M., Kawahirai, R., Hirai, T., Ueki, T. Structural-change of jack bean urease induced by addition of surfactants studied with synchrotron-radiation small-angle x-ray-scattering. *European Journal of Biochemistry*, 1993; 215, 55-61.
55. Honig, B., Nicholls, A. Classical electrostatics in biology and chemistry. *Science*, 1995; 268, 1144-1149.
56. Huber, R., Deisenhofer, J., Colman, P.M., Matsushima, M., Palm, W. Crystallographic structure studies of an igg molecule and an fc fragment. *Nature*, 1976; 264, 415-420.
57. Illeova, V., Polakovic, M., Stefuca, V., Acai, P., Juma, M. Experimental modelling of thermal inactivation of urease *Journal of Biotechnology*, 2003; 105, 235-243.
58. Jabri, E., Carr, M.B., Hausinger, R.P., Karplus, P.A. The crystal structure of urease from *Klebsiella aerogenes*. *Science*, 1995; 268, 998-1004.
59. Jabri, E., Lee, M.H., Hausinger, R.P., Karplus, P.A. Preliminary Crystallographic Studies of urease from Jack Bean and from *Klebsiella aerogenes*. *Journal of Molecular Biology*, 1992; 227, 934-937.
60. Jaiswal, R., Panda, D. Differential assembly properties of *Escherichia coli* FtsZ and *Mycobacterium tuberculosis* FtsZ: an analysis using divalent calcium. *The Journal of Biochemistry*, 2009; 146, 733-742.
61. Jancarik, J., Kim, S.H. Sparse-matrix sampling - a screening method for crystallization of proteins. *Journal of Applied Crystallography*, 1991; 24, 409-411
62. Jian, B., Liu, B., Bi, Y., Hou, W., Wu, C., Han, T. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Molecular Biology*, 2008; 9, 1-14.
63. Johnson, S., Roversi, P., Espina, M., Deane, J.E., Birket, S., Picking, W.D., Blocker, A., Picking, W.L., Lea, S.M. Expression, limited proteolysis and preliminary crystallographic analysis of IpaD, a component of the Shigella flexneri type III secretion system. *Acta Crystallographica D*, 2006; 62, 865-868.
64. Journak, F., McPherson, A., Wang, A.H., Rich, A. Biochemical and structural studies of the tetragonal crystalline modification of the escherichia-coli elongation factor-tu. *Journal of Biological Chemistry*, 1980; 255, 6751-6757.



65. Kerby, R.L., Ludden, P.W., Roberts, G.P. *In vivo* nickel insertion into the carbon monoxide dehydrogenase of *Rhodospirillum rubrum*: molecular and physiological characterization of CooJ. *Journal of Bacteriology*, 1997; 179, 2259–2266.
66. Kim, J.K., Mulrooney, S.B., Hausinger, R.P. The UreEF fusion protein provides a soluble and functional form of the UreF urease accessory protein. *Journal of Bacteriology*, 2006; 188, 8413-8420.
67. Kim, K.Y., Yang, C.H., Lee, M.H. Expression of the recombinant *Klebsiella aerogenes* UreF protein as a MalE fusion. *Archives of Pharmacal Research*, 1999, 22, 274-278.
68. Koth, C.M., Orlicky, S.M., Larson, S.M., Edwards, A.M. Use of limited proteolysis to identify protein domains suitable for structural analysis. *Macromolecular Crystallography*, 2003; 368, 77-84.
69. Krajewska, B. Ureases I. Functional, catalytic and Kinetic properties: A review. *Journal of Molecular Catalysis B: Enzymes*, 2009; 59, 9-21.
70. Kuchar, J., Hausinger, R.P. Biosynthesis of metal sites. *Chemical Reviews*, 2004; 104, 509-525.
71. Lam, R., Romanov, V., Johns, K., Battaile, K.P., Wu-Brown, J., Guthrie, J.L., Hausinger, R.P., Pai, E.F., Chirgadze, N.Y. Crystal structure of a truncated urease accessory protein UreF from *Helicobacter pylori*. *Proteins*, 2010; 78, 2839-2848.
72. Laskowski, R.A., Macarthur, M.W., Moss, D.S., Thornton, J.M. Procheck - A Program To Check The Stereochemical Quality Of Protein Structures. *Journal of Applied Crystallography*, 1993; 26, 283-291.
73. Leach, M.R., Sandal, S., Sun, H., Zamble, D.B. Metal binding activity of the Escherichia coli hydrogenase maturation factor HypB. *Biochemistry*, 2005; 44, 12229–12238.
74. Lee, M.H., Mulrooney, S.B., Renner, M.J., Markowicz, Y., Hausinger, R.P. *Klebsiella aerogenes* urease gene-cluster - sequence of ured and demonstration that 4 accessory genes (ured, uree, uref, and ureg) are involved in nickel metallocenter biosynthesis. *Journal of Bacteriology*, 1992; 174, 4324-4330
75. Lee, M.H., Pankratz, H.S., Wang, S., Scott, R.A., Finnegan, M.G., Johnson, M.K. Ippolito, J.A., Christianson, D.W. and Hausinger, R.P. Purification and characterization of *Klebsiella aerogenes* ureE protein: A nickel-binding protein that function in urease metallocenter assembly. *Protein Science*, 1993; 2, 1042-1052.
76. Libault, M., Thibivilliers, S., Bilgin, D.D., Radwan, O., Benitez, M., Clough, S.J., Stacey, G. Identification of Four Soybean Reference Genes for Gene Expression Normalization. *The Plant Genome*, 2008; 1, 44–54.
77. Livak, K.J., Schmittgen, T.D. Analysis of relative gene expression data using realtime quantitative PCR and the  $2^{-\Delta\Delta C(T)}$  Method. *Methods*, 2001; 25, 402-408.
78. Loke, H.K., Lindahl, P.A. Identification and preliminary characterization of AcsF, a putative Ni-insertase used in the biosynthesis of acetyl-CoA synthase from *Clostridium thermoaceticum*. *Journal of Inorganic Biochemistry*, 2003; 93, 33-40.
79. Longenecker, K.L., Garrard, S.M., Sheffield, P.J., Derewenda, Z.S. Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI. *Acta Crystallographica D*, 2001; 57, 679-688.
80. Lundin, B., Thuswaldner, S., Shutova, T., Eshaghi, S., Samuelsson, G., Barber, J., Andersson, B., Spetea, C. Subsequent events to GTP binding by the plant PsbO protein:

structural changes, GTP hydrolysis and dissociation from the photosystem II complex. *Biochimica et Biophysica Acta*, 2007; 1767, 500-508.

81. Maier, T., Jacobi, A., Sauter, M., Bock, A. The product of the *hypB* gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein. *Journal of Bacteriology*, 1993; 175, 630–635.

82. Mamiya, G., Takishima, K., Masakuni, M., Kayumi, T., Ogawa, K., Sekita, T. Complete amino-acid sequence of jack bean urease. *Proceedings of the Japan Academy Series Biophysical and Biological Sciences*, 1985; 61, 395-398.

83. Mandel, C.R., Gebauer, D., Zhang, H., Tong, L., A serendipitous discovery that in situ proteolysis is essential for the crystallization of yeast CPSF-100 (Ydh1p). *Acta Crystallographica D*, 2006; 62, 1041-1045.

84. Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., Sali, A. Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, 2000; 29, 291-325.

85. Mateja, A., Devedjiev, Y., Krowarsch, D., Longenecker, K., Dauter, Z., Otlewski, J., Derewenda, Z.S. The impact of Glu → Ala and Glu → Asp mutations on the crystallization properties of RhoGDI: the structure of RhoGDI at 1.3 angstrom resolution. *Acta Crystallographica D*, 2002; 58, 1983-1991.

86. McElroy, H.E., Sisson, G.W., Schoettlin, W.E., Aust, R.M., Villafranca, J.E. Studies on engineering crystallizability by mutation of surface residues of human thymidylate synthase. *Journal of Crystal Growth*, 1992; 122, 265-272.

87. McPherson, A; Cudney, B Searching for silver bullets: An alternative strategy for crystallizing macromolecules. *Journal of Structural Biology*, 2006; 156, 387-406.

88. Mehta, N., Benoit, S., Maier, R.J. Roles of conserved nucleotide-binding domains in accessory proteins. HypB and UreG, in the maturation of nickel-enzymes required for efficient *Helicobacter pylori* colonization. *Microbial Pathogenesis*, 2003b; 35, 229–234.

89. Mehta, N., Olson, J.W., Maier, R.J. Characterization of *Helicobacter pylori* nickel metabolism accessory proteins needed for maturation of both urease and hydrogenase. *Journal of Bacteriology*, 2003a; 185, 726–734.

90. Menegassi, A., Wassermann, G.E., Olivera-Severo, D., Becker-Ritt, A.B., Martinelli, A.H.S., Feder, V., Carlini, C.R. Urease from cotton (*Gossypium hirsutum*) seeds: Isolation, physicochemical characterization, and antifungal properties of the protein. *Journal of Agriculture and Food Chemistry*, 2008; 56, 4399-4405.

91. Meyer-Bothling, L.E., Polacco, J.C. Mutational analysis of the embryo-specific urease locus of soybean. *Molecular and General Genetics*, 1987; 209, 439–444.

92. Mobley, H.L., Island, M.D., Hausinger, R.P. Molecular biology of microbial ureases. *Microbiology*, 1995; 10, 70-74.

93. Moncrief, M.B., Hausinger, R.P. Purification and activation properties of UreD-UreF-urease apoprotein complex. *Journal Bacteriology*, 1996; 178, 5417-5421.

94. Moncrief, M.C., Hausinger, R.P. Characterization of UreG, identification of a UreD-UreF-UreG complex, and evidence suggesting that a nucleotide-binding site in UreG is required for *in vivo* metallocenter assembly of *Klebsiella aerogenes* urease. *Journal of Bacteriology*, 1997; 179, 4081–4086.

95. Mulinari F., Stanisçuaski F., Bertholdo-Vargas L.R., Postal M., Oliveira-Neto O.B., Ridgen, D.J., Grossi-de-Sá M.F., Carlini C.R. Jaburetox-2Ec: an insecticidal peptide derived from an isoform of urease from the plant *Canavalia ensiformis*. *Peptides*, 2007; 28, 2042-2050.
96. Mulrooney, S.B., Hausinger, R.P. Nickel uptake and utilization by microorganisms. *FEMS Microbiology Reviews*, 2003; 27, 239-261.
97. Musiani, F., Arnofi, E., Casadio, R., Ciurli, S. Structure-based computational study of the catalytic and inhibition mechanisms of urease. *Journal of Biological Inorganic Chemistry*, 2001; 6, 300-314.
98. Musiani, F., Zambelli, B., Stola, M., Ciurli, S. Nickel trafficking: insights into the fold and function of UreE, a urease metallochaperone. *Journal of Inorganic Biochemistry*, 2004; 98, 803-813.
99. Neyroz, P., Zambelli, B., Ciurli, S. Intrinsically disordered structure of *Bacillus pasteurii* UreG as revealed by steady-state and time-resolved fluorescence spectroscopy. *Biochemistry*, 2006; 45, 8918-8930.
100. Olson, J.W., Maier, R.J. Dual roles of *Bradyrhizobium japonicum* nickelin protein in nickel storage and GTP-dependent Ni mobilization. *Journal of Bacteriology*, 2000; 182, 1702-1705.
101. Park, I.S., Carr, M.B., Hausinger, R.P. *In-vitro* activation of urease apoprotein and role of urea as a chaperone required for nickel metallocenter assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 1994; 91, 3233-3237.
102. Park, I.S., Hausinger, R.P. Evidence for the presence of urease apoprotein complex containing UreD, UreF, and UreG in cells that are competent for *in vivo* enzyme activation. *Journal of Bacteriology*, 1995; 177, 1947-1951.
103. Park, I.S., Hausinger, R.P. Metal ion interaction with urease and UreD-urease apoproteins. *Biochemistry*, 1996; 35, 5345-5352.
104. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E. UCSF Chimera - a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 2004; 25, 1605-1612.
105. Pffal, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 2001; 29, e45.
106. Piovesan, A., Staniscuaski, F., Marco-Salvadori, J., Real-Guerra, R., Defferrari, M.S., Carlini, C.R. Stage-specific gut proteinases of the cotton stainer bug *Dysdercus peruvianus*: role in the release of entomotoxic peptides from *Canavalia ensiformis* urease. *Insect Biochemistry and Molecular Biology*, 2008; 38, 1023-1032.
107. Polacco, J.C., Holland, M.A. Roles of urease in plant cells. *International Review of Cytology*, 1993; 145, 65-103.
108. Polacco, J.C., Judd, A.K., Dybing, J.K., Cianzio, S.R. A new mutant class of soybean lacks urease in leaves but not in leaf-derived callus or in roots. *Molecular General Genetics*, 1989; 217, 257-262.
109. Polacco, J.C., Winkler, R.G. Soybean Leaf Urease: A Seed Enzyme? *Plant Physiology*, 1984; 74, 800-803.
110. Quiroz-Valenzuela, S., Sukuru, S.C., Hausinger, R.P., Kuhn, L.A., Heller, W.T. The structure of urease activation complexes examined by flexibility analysis, mutagenesis, and small-angle X-ray scattering. *Archives of Biochemistry and Biophysics*, 2008; 480, 51-57.

111. Rayment, I. Reductive alkylation of lysine residues to alter crystallization properties of proteins *Macromolecular Crystallography*, 1997; 276, 171-179.
112. Rayment, I., Rypniewski, W.R., Schmidtbase, K., Smith, R., Tomchick, D.R., Benning, M.M., Winkelmann, D.A., Wesenberg, G., Holden, H.M. 3-dimensional structure of myosin subfragment-1 - a molecular motor. *Science*, 1993; 261, 50-58.
113. Real-Guerra, R. Aspectos biológicos e estruturais das ureases de *Canavalia ensiformis*. Dissertação de Mestrado do Programa de Pós-Graduação em Biologia Celular e Molecular da Universidade Federal do Rio Grande do Sul, 2007.
114. Remaut, H., Safarov, N., Ciurli, S. Van Beeumen, J.J. Structural basis for Ni transport and assembly of urease active site by the metallo-chaperone UreE from *Bacillus pasteurii*. *Journal of Biological Chemistry*, 2001; 276, 49365-49370.
115. Rey, L., Imperial, J., Palacios, J.M., Ruiz-Argueso, T. Purification of *Rhizobium leguminosarum* HypB, a nickel-binding protein required for hydrogenase synthesis. *Journal of Bacteriology*, 1994; 176, 6066-6073.
116. Riddles, P.W., Andrew, R.K., Blakeley, R.L. Zerner, B. Cloning and sequencing of jack bean urease-encoding cDNA. *Gene*, 1991; 108, 265-267.
117. Rypniewski, W.R., Holden, H.M., Rayment, I. Structural consequences of reductive methylation of lysine residues in hen egg-white lysozyme - an x-ray-analysis at 1.8-angstrom resolution. *Biochemistry*, 1993; 32, 9851-9858.
118. Salomone-Stagni, M., Zambelli, B., Musiani, F., Ciurli, S. A model-based proposal for the role of UreF as a GTPase-activating protein in the urease active site biosynthesis. *Proteins*, 2007; 68, 749-61.
119. Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H., Kraut, J. Crystal-structure of rat dna-polymerase-beta - evidence for a common polymerase mechanism. *Science*, 1994; 264, 1930-1935.
120. Schmidt, M., Parrott, W. Quantitative detection of transgenes in soybean [*Glycine max* (L.) Merrill] and peanut (*Arachis hypogaea* L.) by real-time polymerase chain reaction. *Plant Cell Reports*, 2001; 20, 422-428.
121. Seshadri, S., Benoit, S.L., Maier, R.J. Roles of his-rich hpn and hpn-like proteins in *Helicobacter pylori* nickel physiology. *Journal of bacteriology*, 2007; 189, 4120-4126.
122. Shabana, S., Kawai, A., Kai, K., Akiyama, K., Hayashi, H. Inhibitory Activity against Urease of Quercetin Glycosides Isolated from *Allium cepa* and *Psidium guajava*. *Bioscience Biotechnology and Biochemistry*, 2010; 74, 878-880.
123. Shen, M.Y., Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Science*, 2006; 15, 2507-2524.
124. Sheridan, L., Wilmot, C.M., Cromie, K.D., Van Der Logt, P., Phillips, S.E. Crystallization and preliminary X-ray structure determination of jack bean urease with a bound antibody fragment. *Acta Crystallographica Section D-Biological Crystallography*, 2002; 58, 374-376.
125. Shi, R., Munger, C., Asinas, A., Benoit, S.L., Miller, E., Matte, A., Maier, R.J., Cygler, M. Crystal structures of apo and metal-bound forms of the UreE protein from *Helicobacter pylori*: role of multiple metal binding sites. *Biochemistry*, 2010; 49, 7080-7088.
126. Shimizu, T., Ihara, K., Maesaki, R., Kuroda, S., Kaibuchi, K., Hakoshima, T. An open conformation of switch I revealed by the crystal structure of a Mg<sup>2+</sup>-free form of RhoA complexed with GDP. *Journal of Biological Chemistry*, 2000; 275, 18311-18317.

127. Shutova, T., Nikitina, J., Deikus, G., Anderson, B., Klimov, V., Samuelsson, G. Structural dynamics of the manganese-stabilizing protein – Effect of pH, calcium, and manganese. *Biochemistry*, 2005; 44, 15182-15192.
128. Song, H.K., Mulrooney, S.B., Huber, R., Hausinger, R.P. Crystal structure of *Klebsiella aerogenes* UreE, a Nickel-binding metallochaperone for urease activation. *Journal of Biological Chemistry*, 2001; 276, 49359-49364.
129. Soriano, A., Colpas, G. L. and Hausinger, R. P. UreE stimulation of GTP-dependent urease activation in the UreD-UreF-UreG-urease complex. *Biochemistry*, 2000; 39, 12435-12440.
130. Soriano, A., Hausinger, R.P. GTP-dependent activation of urease apoprotein in complex with the UreD, Ure F, and Ure G accessory proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 1999; 96, 11140–11144.
131. Sprang, S.P. G proteins mechanisms: Insights from structural analysis. *Annual Reviews in Biochemistry*, 1997; 66, 639-678.
132. Staniscuaski, F., Ferreira-DaSilva, C.T., Mulinari, F., Pires-Alves, M., Carlini, C.R. Insecticidal effects of canatoxin on the cotton stainer bug *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae). *Toxicon*, 2005; 45, 753–760.
133. Staniscuaski, F., Te Brugge, V., Carlini, C.R., Orchard, I. In vitro effect of *Canavalia ensiformis* urease and the derived peptide Jaburetox 2Ec on *Rhodnius prolixus* Malpighian tubules. *Journal of Insect Physiology*, 2009; 55, 255-263.
134. Staniscuaski, F., Te Brugge, V., Carlini, C.R., Orchard, I. Jack bean urease alters serotonin-induced effects on *Rhodnius prolixus* anterior midgut. *Journal of Insect Physiology*, 2010; 56, 1078 - 1086.
135. Stola, M., Musiani, F., Mangani, S., Turano, P., Safarov, N., Zambelli, B., Ciurli, S. The nickel site of *Bacillus pasteurii* UreE, a urease metallochaperone, as revealed by metal-binding studies and X-ray absorption spectroscopy. *Biochemistry*, 2006; 45, 6495-6509.
136. Sumner, J.B, Somers, G.F. *Chemistry and methods of enzymes*, 3<sup>rd</sup> ed., Academic Press, New York, 1953; 156-161.
137. Sumner, J.B. The isolation and crystallization of the enzyme urease. *Journal of Biological Chemistry*, 1926; 69, 435–441.
138. Takishima, K., Suga, T., Mamiya, G. The structure of jack bean urease - the complete amino-acid sequence, limited proteolysis and reactive cysteine residues. *European Journal of Biochemistry*, 1988; 175, 151-165.
139. Taneja, B., Patel, A., Slesarev, A., Mondragon, A. Structure of the N-terminal fragment of topoisomerase V reveals a new family of topoisomerases. *EMBO Journal*, 2006; 25, 398-408.
140. Tompa, P. Intrinsically unstructured proteins. *Trends in Biochemical Science*, 2002; 27, 527-533.
141. Torisky, R.S., Polacco, J.C. Soybean roots retain the seed urease isozyme synthesized during embryo development. *Plant Physiology*, 1990; 94, 681–689.
142. Uversky, V.N. Natively unfolded proteins: A point where biology waits for physics. *Protein Science*, 2002; 11, 739-756.
143. Vedadi, M., Niesen, F.H., Allali-Hassani, A., Fedorov, O.Y., Finerty, P.J., Wasney, G.A., Yeung, R., Arrowsmith, C., Ball, L.J., Berglund, H., Hui, R., Marsden, B.D., Nordlund, P., Sundstrom, M., Weigelt, J., Edwards, A.M. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences of the United States of America*, 2006; 103, 15835-15840.

144. Walter, T.S., Meier, C., Assenberg, R., Au, K.F., Ren, J.S., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I., Grimes, J.M. Lysine methylation as a routine rescue strategy for protein crystallization. *Structure*, 2006; 14, 1617-1622.
145. Watt, R.K., Ludden, P.W. The identification, purification, and characterization of CooJ. A nickel-binding protein that is co-regulated with Ni-containing CO dehydrogenase from *Rhodospirillum rubrum*. *Journal of Biological Chemistry*, 1998; 273, 10019-10025.
146. Weartherburn, M.W. Phenol-hypochlorite for determination of ammonia. *Analytical Chemistry*, 1967; 39, 971 – 974.
147. Wernimont, A., Edwards, A. *In situ* proteolysis to generate crystals for structure determination: an update. *PLoS One*, 2009; 4, e5094.
148. Whitmore, L, Wallace, B.A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, 2008; 89, 392-400.
149. Whitmore, L. Wallace, B.A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 2004; 32, W668-W673.
150. Witte, C.P., Isidore, E., Tiller, S.A., Davies, H.V., Taylor, M.A. Functional characterization of urease accessory protein G (ureG) from potato. *Plant Molecular Biology*, 2001; 45, 169-79.
151. Witte, C.P., Rosso, M.G., Romeis, T. Identification of three urease accessory proteins that are required for urease activation in Arabidopsis. *Plant Physiology*, 2005; 139, 1155-1162.
152. Wulfing, C., Lombardero, J., Pluckthun, A. An *Escherichia coli* protein consisting of a domain homologous to FK506-binding proteins (FKBP) and a new metal binding motif. *Journal of Biological Chemistry*, 1994; 269, 2895-2901.
153. Zaborska, W., Krajewska, B., Leszko, Olech, Z. Inhibition of urease by Ni<sup>2+</sup> ions. Analysis of reaction progress curves. *Journal of Molecular Catalysis B: Enzymatic*, 2001; 13, 103–108.
154. Zaborska, W., Krajewska, B., Olech, Z. Heavy metal ions inhibition of jack bean urease: potential for rapid contaminant probing. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 2004; 19, 65-69.
155. Zambelli, B., Musiani, F., Savini, M., Tucker, P., Ciurli, S. Biochemical studies on Mycobacterium tuberculosis UreG and comparative modeling reveal structural and functional conservation among the bacterial UreG family. *Biochemistry*, 2007; 46, 3171-3182.
156. Zambelli, B., Stola, M., Musiani, F., De Vriendt, K., Samyn, B., Devreese, B., Van Beeumen, J., Turano P., Dikiy A., Bryant D.A., Ciurli, S. UreG, a chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds Zn<sup>2+</sup>. *Journal of Biological Chemistry*, 2005; 280, 4684-4695.
157. Zambelli, B., Turano, P., Musiani, F., Neyroz, P., Ciurli, S. Zn<sup>2+</sup>-linked dimerization of UreG from *Helicobacter pylori*, a chaperone involved in nickel trafficking and urease activation. *Proteins*, 2009; 74, 222-239.
158. Zerner, B. Recent advances in the chemistry of an old enzyme, Urease. *Bioorganic chemistry*, 1991; 19, 116-131.
159. Zhang, J.W., Butland, G., Greenblatt, J.F., Emili, A., Zamble, D.B. A role for SlyD in the *Escherichia coli* hydrogenase biosynthetic pathway. *Journal of Biological Chemistry*, 2005; 280, 4360–4366.

## **Anexo I**

### **Tabela 1 – Análise por ICP-OES**

Tabela 1: Dados da quantificação de metais de GmUreG pela técnica de ICP-OES.

	<b>Ag 328.068</b>	<b>Al 167.078</b>	<b>Al 176.641</b>	<b>Ar 404.442</b>	<b>As 189.042</b>	<b>B 208.959</b>	<b>Ba 455.404</b>	<b>Be 313.042</b>	<b>Ca 183.801</b>
	ppm	ppm	ppm	cps [corr]	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	< 0.000	< 0.000	< -0.023	1782750	< 0.004	< -0.000	0,006	< -0.000	2,08
<b>GmUreG300</b>	n.d	n.d	< -0.921	n.d	n.d	0,007	n.d	n.d	< -5.95
<b>rGmUreG</b>	< 0.000	0,002	< -0.016	1743560	< 0.001	< -0.004	0,004	< -0.000	0,022
	<b>Cd 228.802</b>	<b>Co 228.616</b>	<b>Cr 267.716</b>	<b>Cu 324.754</b>	<b>Fe 241.331</b>	<b>Fe 259.941</b>	<b>Hg 184.950</b>	<b>K 766.491</b>	<b>Li 670.780</b>
	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	< -0.000	< 0.000	< -0.000	0,002	< -0.016	0,002	0,015	1,64	0
<b>GmUreG300</b>	< -0.006	< -0.009	< -0.010	< -0.000	n.d	< -0.686	n.d	1,6	n.d
<b>rGmUreG</b>	< -0.000	< 0.000	< -0.000	0,004	< -0.020	< 0.001	0,067	0,04	0
	<b>Mg 279.079</b>	<b>Mn 257.611</b>	<b>Mo 202.095</b>	<b>Na 330.237</b>	<b>Na 589.592</b>	<b>Ni 231.604</b>	<b>P 178.287</b>	<b>Pb 220.353</b>	<b>S 182.034</b>
	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	0,708	0,001	< 0.000	320	77,1	0,004	0,019	< 0.001	0,968
<b>GmUreG300</b>	0,083	< -0.009	< -0.010	464	45,1	< 0.001	< -0.247	< 0.002	1,47
<b>rGmUreG</b>	< 0.001	< -0.000	< 0.000	< 7.93	9,73	0,008	< 0.003	< -0.003	2,81
	<b>Sb 206.833</b>	<b>Se 196.090</b>	<b>Si 251.612</b>	<b>Sn 189.991</b>	<b>Sr 407.771</b>	<b>Ti 334.941</b>	<b>Tl 190.864</b>	<b>V 292.464</b>	<b>Zn 213.856</b>
	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	< -0.002	< 0.005	0,429	< 0.004	0,01	< -0.000	< -0.004	< -0.001	0,007
<b>GmUreG300</b>	n.d	n.d	n.d	n.d	n.d	n.d	n.d	n.d	0,096
<b>rGmUreG</b>	< -0.002	< 0.003	< 0.007	< -0.000	0	< -0.000	< -0.000	< -0.001	0,005

Dados são apresentados em parte por milhão (ppm.).

Metais com conteúdo diferenciado estão destacados.



**Anexo II**  
**Currículo**

## Rafael Real Guerra

Curriculum Vitae – Março/2011

---

### Dados Pessoais

**Nome** Rafael Real Guerra  
**Nome em citações** Real-Guerra, R.  
**Endereço profissional** Universidade Federal do Rio Grande do Sul, Instituto de Biociências,  
Departamento de Biofísica  
Avenida Bento Gonçalves 9500  
Agronomia - Porto Alegre  
91501-970, RS - Brasil  
Telefone: 51 3308-7606  
www.ufrgs.br/laprotox  
**Endereço eletrônico** rafael.guerra@ufrgs.br

---

### Formação Acadêmica/Titulação

**2007** Doutorado em Biologia Celular e Molecular - PPGBCM  
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil  
Estágio de doutorado sanduíche na Universidade de Toronto, supervisão do Dr. Dinesh Christendat  
“Caracterização funcional e estrutural de ureases vegetais e suas proteínas acessórias”  
Orientador: Célia R. Carlini  
Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

**2005 - 2007** Mestrado em Biologia Celular e Molecular - PPGBCM  
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil  
“Aspectos Biológicos e Estruturais das Ureases de *Canavalia ensiformis*”  
Orientador: Célia R. Carlini  
Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico

**2000 - 2004** Graduação em Química, Bacharelado.  
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil  
“Efeitos Físico-químicos e biológicos da Modificação Química de Aminoácidos em Ureases Vegetais”  
Orientador: Celia R. Carlini  
Bolsista da Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul

---

### Formação complementar

**2005** Curso de curta duração em Coleta de Dados na Linha MX1 (Cristalografia de Macromoléculas)

Associação Brasileira de Tecnologia de Luz Síncrotron, ABTLuS, Campinas, Brasil

**2005** Curso de curta duração em Treinamento em Espectrometria de Massas. Associação Brasileira de Tecnologia de Luz Síncrotron, ABTLuS, Campinas, Brasil

**2004** Extensão universitária em Bioinformática e Modelagem Molecular. Universidade Luterana do Brasil, ULBRA, Canoas, Brasil

---

## Produção em C, T& A

### Produção bibliográfica

#### Artigos completos publicados em periódicos

1. Piovesan, A., Staniscuaski, F., Marco-Salvadori, J., Real-Guerra, R., Defferrari, M., Carlini, C.R.

Stage-specific gut proteinases of the cotton stainer bug *Dysdercus peruvianus*. Role in the Release of Entomotoxic Peptides from *Canavalia ensiformis* Urease. *Insect Biochemistry and Molecular Biology* 38, 1023 - 1032, 2008.

2. Follmer, C., Real-Guerra, R., Wassermann, G.E., Olivera-Severo, D.O., Carlini, C.R.

Jackbean, soybean and *Bacillus pasteurii* ureases: biological effects unrelated to ureolytic activity. *European Journal of Biochemistry* 271, 1357 - 1363, 2004.

#### Trabalhos publicados em anais de eventos (resumo)

1. Real-Guerra, R., Staniscuaski, F., Christendat, D., Carlini, C.R.

Cloning, Expression and Preliminary Characterization of Soybean Urease Accessory Protein UreF In: XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2010, Foz do Iguaçu.

2. Staniscuaski, F., Real-Guerra, R., Paluzzi, J.P., Carlini, C.R., Orchard, I.

Aquaporins in Different Tissues of *Rhodnius prolixus*: Expression Analysis and Molecular Characterization. In: XXXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2010, Foz do Iguaçu.

3. Wolker, C., Real-Guerra, R., Carlini, C.R.

Caracterização do padrão de autofragmentação da urease de *Canavalia ensiformis*. In: IX Reunião Anual do Programa de Pós-graduação em Biologia Celular e Molecular do Centro de Biotecnologia - UFRGS, 2007, Porto Alegre.

4. Real-Guerra, R., Beltramini, L.M., Carlini, C.R.

Estudos da ligação de JBU e CNTX a glicoconjugados através de técnicas de ITC, SPR e

fluorescência. In: IX Reunião Anual do Programa de Pós-graduação em Biologia Celular e Molecular do Centro de Biotecnologia - UFRGS, 2007, Porto Alegre.

5. Real-Guerra, R., Stanisçuaski, F., Follmer, C., Carlini, C.R.

Chemical modification of Jackbean urease promoting differences in its biological activities. In: VIII Reunião Anual do Programa de Pós-graduação em Biologia Celular e Molecular do Centro de Biotecnologia - UFRGS, 2006, Porto Alegre.

6. Real-Guerra, R., Stanisçuaski, F., Follmer, C., Carlini, C.R.

Chemical modification of jackbean urease promoting differences in its biological activities. In: XXXV Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2006, Águas de Lindóia.

7. Real-Guerra, R., Carlini, C.R.

Estrutura de Ureases Vegetais. In: VII Reunião Anual do programa de Pós-graduação em Biologia Celular e Molecular do Centro de Biotecnologia da UFRGS, 2005, Porto Alegre.

8. Real-Guerra, R., Stanisçuaski, F., Carlini, C.R.

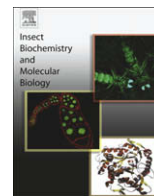
Stability and Oligomerization of Jaburetox-2Ec, an Insecticidal Recombinant Peptide derived from Canatoxin. In: XXXIV Reunião anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 2005, Águas de Lindóia.

9. Real-Guerra, R., Follmer, C., Carlini, C.R.

Aspectos físico-químicos e biológicos de Ureases vegetais. In: XV Salão de Iniciação Científica UFRGS, 2003, Porto Alegre.

### **Anexo III**

**Artigo publicado durante o período do doutorado**



## Stage-specific gut proteinases of the cotton stainer bug *Dysdercus peruvianus*: Role in the release of entomotoxic peptides from *Canavalia ensiformis* urease

Angela R. Piovesan<sup>a,1</sup>, Fernanda Stanisçuaski<sup>b,\*,1</sup>, Juliana Marco-Salvadori<sup>a</sup>, Rafael Real-Guerra<sup>a</sup>, Marina S. Defferrari<sup>a</sup>, Célia R. Carlini<sup>a,b</sup>

<sup>a</sup> Graduate Program in Cellular and Molecular Biology, Center of Biotechnology, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, Building 43421, CEP 91501-970 Porto Alegre, Rio Grande do Sul, Brazil

<sup>b</sup> Department of Biophysics and Center of Biotechnology, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, Building 43422, CEP 91501-970 Porto Alegre, Rio Grande do Sul, Brazil

### ARTICLE INFO

#### Article history:

Received 12 May 2008

Received in revised form 24 August 2008

Accepted 11 September 2008

#### Keywords:

*Dysdercus peruvianus*

Urease

Proteolytic enzyme

Insecticidal peptide

Enzyme inhibitor

### ABSTRACT

*Canavalia ensiformis* ureases are toxic to insects of different orders. The entomotoxicity of urease is due to a 10 kDa internal peptide released by proteinases in the insect digestive tract. We previously observed that, given orally, urease is toxic to nymphs of *Dysdercus peruvianus*, but does not affect adults. Here we characterized the major proteolytic activities of *D. peruvianus* midgut homogenates and investigated their *in vitro*-catalyzed release of the 10 kDa entomotoxic peptide from urease. Cysteine, aspartic and metalloproteinases are present in both homogenates. Variations in optimal pH and susceptibility to inhibitors indicated differences in the enzyme profiles in the two developmental stages. Only nymph homogenates released ~10 kDa fragment(s) from urease, recognized by antibodies against the entomotoxic peptide. Fluorogenic substrates containing urease partial sequences flanking the N-terminal or the C-terminal portion of the entomotoxic peptide were efficiently cleaved by homogenates from nymphs, but much more slowly by the adult homogenate. Different classes of enzymes in the homogenates cleaved both substrates suggesting that *in vivo* the release of the entomotoxic peptide results from the concerted action of at least two different proteinases. Our findings support the view that a differential processing of ingested urease by the insects explains at least in part the lack of toxicity in adults.

© 2008 Elsevier Ltd. All rights reserved.

### 1. Introduction

Ureases (urea amidohydrolase; EC 3.5.1.5), nickel-dependent enzymes that catalyze the hydrolysis of urea to ammonia and carbon dioxide, are synthesized by plants, fungi and bacteria (Mobley et al., 1995; Follmer, 2008). In plants, ureases are homotrimers or hexamers of a ~90 kDa subunit and participate in the use of urea as nitrogen source (Polacco and Holland, 1993; Sirko and Brodzik, 2000; Follmer, 2008). Canatoxin, a toxic protein isolated from *Canavalia ensiformis* seeds (Carlini and Guimarães, 1981) and more recently identified as a minor isoform of urease (Follmer et al., 2001), displays insecticidal properties (Carlini et al., 1997). Canatoxin and the major isoform of urease of *C. ensiformis* seeds (herein designated “urease”) are toxic to *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae), and this toxicity is independent of their ureolytic

activity (Follmer et al., 2004a; Stanisçuaski et al., 2005; Follmer et al., 2004b).

Insects with cathepsin-like digestive enzymes are susceptible to the toxic effects of urease, while insects with trypsin-like digestive enzymes are not (Carlini et al., 1997), possibly due to differences in the proteolytic processing of ureases in the insects. The entomotoxic activity is due to a 10 kDa internal peptide released from ureases by insect digestive enzymes. This peptide, called pepcanatox, was isolated and characterized (Ferreira-DaSilva et al., 2000). A recombinant peptide, Jaburetox-2Ec equivalent to pepcanatox, obtained by heterologous expression in *Escherichia coli* (Mulinari et al., 2004, 2007), is highly insecticidal and did not affect mice or neonate rats when administered by oral or intraperitoneal routes (Mulinari et al., 2007). Interestingly, Jaburetox-2Ec kills insects that are resistant to intact ureases, such as *Spodoptera frugiperda* (Mulinari et al., 2007) or adults of *Triatoma infestans* (Tomazzeto et al., 2007).

The cotton stainer bug *D. peruvianus* feeds on cotton seeds (*Gossypium hirsutum*), staining the cotton fibers, damaging the seeds, besides being a vector for phytopathogenic microorganisms, and thus can cause heavy losses in cotton production (Gallo, 1988).

\* Corresponding author. Tel.: +55 51 3308 7606; fax: +55 51 3308 7003.

E-mail address: [fernanda.staniscuaski@ufrgs.br](mailto:fernanda.staniscuaski@ufrgs.br) (F. Stanisçuaski).

URL: <http://www.ufrgs.br/laprotox>

<sup>1</sup> These authors have contributed equally to this work.

Nymphs are susceptible to canatoxin or urease, while adults are not. In a previous work, we showed that adults and nymphs have distinct repertoires of gut enzymes, and differences in the processing of ureases by stage-specific enzymes were proposed to explain the lack of susceptibility of adults (Stanisçuaski et al., 2005).

Digestive proteinases play two critical roles in insect physiology: breaking down proteins into amino acids essential for growth and development, and inactivating protein toxins ingested as a consequence of feeding (Terra et al., 1996). Endoproteinases belong to one of six groups based on the active site amino acid residue, or metal ion, involved in peptide bond hydrolysis (Barret et al., 1998): serine, cysteine, aspartic acid, glutamic acid, threonine, and metalloproteinases. Insects having midguts with pH in the neutral or alkaline range usually rely on serine proteinases. In contrast, insects having acidic or mildly acidic midguts generally utilize cysteine and aspartic proteinases (Terra et al., 1996; Terra and Ferreira, 1994; Silva et al., 1996; Wilhite et al., 2000).

In this study, we aimed to investigate further the processing of ureases by nymphs and adults of *D. peruvianus*. Because the release of the entomotoxic peptide could be due to the concerted or sequential action of more than one enzyme, we decided to study the enzymatic profile of whole midguts' homogenates. We characterized the enzymatic activities present in these homogenates and performed *in vitro* hydrolysis of urease with gut homogenates. We show that homogenates from nymphs and adults generate different fragments from urease and that a 10 kDa peptide recognized by anti-Jaburetox-2Ec antibodies is formed only by nymph homogenates.

## 2. Materials and methods

### 2.1. Reagents

Azocasein, *trans*-epoxysuccinyl-L-leucylamido (4-guanidino)-butane (E-64), pepstatin-A, phenylmethanesulphonylfluoride (PMSF), ethylenediamine tetraacetic acid (EDTA), methylcoumarin amide (MCA)-coupled substrates, anti-rabbit IgG conjugated with alkaline phosphatase, nitro-blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) were purchased from Sigma Chemicals Company (Saint Louis, USA). Hybond-P membrane (PVDF, 0.45  $\mu$ m) was from Amersham Biosciences (Little Chalfont, England). Molecular mass markers were from Invitrogen Life Technologies (California, USA). The synthetic peptide substrates coupled to Abz-EDDnp were a kind gift from Dr. Luis Juliano Neto and Dr. Maria Aparecida Juliano (Dept. Biophysics, Universidade Federal de São Paulo, São Paulo, Brazil).

### 2.2. Insects

A colony of *D. peruvianus*, established and operated in our laboratory for five years, is maintained as described (Stanisçuaski et al., 2005). The insects develop from eggs through five nymphal stages in about 20–25 days. The insects were fed with cotton seeds (*G. hirsutum*) and had free access to water.

### 2.3. Urease and Jaburetox-2Ec

*C. ensiformis* major urease was purified as previously described (Follmer et al., 2001, 2004b). Briefly, seed meal (100 g) was extracted with buffer A (20 mM sodium phosphate, pH 7.5, 1 mM EDTA, and 2 mM 2-mercaptoethanol) for 1 h at 4 °C. The meal was removed by centrifugation and 28% (v/v) ice-cold acetone (final concentration) was added to the supernatant. After removing the precipitate, the concentration of acetone in the supernatant was increased to 31.6% (v/v). The new precipitate was discarded, the supernatant was dialyzed against buffer B (20 mM sodium

phosphate, pH 7.5, 1 mM EDTA, and 5 mM 2-mercaptoethanol) and then mixed with 40 mL of Q-Sepharose in buffer B. The urease-enriched fraction was eluted by adding 300 mM NaCl in buffer B. This material was concentrated and gel-filtered on a Superose 6 HR 10/30 column in buffer B. The active peak was dialyzed against 20 mM sodium phosphate, pH 7.0, 0.5 M NaCl (buffer C) and submitted to immobilized metal affinity chromatography using 10 mL of Co<sup>2+</sup> loaded iminodiacetic acid-Sepharose (IDA-Co<sup>2+</sup>), equilibrated in buffer C. This step separates the isoforms of jack bean urease, the major urease being recovered in the non-retained fraction (Follmer et al., 2004b).

The recombinant urease-derived entomotoxic peptide Jaburetox-2Ec, produced and purified as described in Mulinari et al. (2007), was used to immunize rabbits and prepare anti-Jaburetox-2Ec polyclonal antibodies (Tomazzeto et al., 2007).

### 2.4. Bioassay of urease insecticidal activity

The bioassay was carried out essentially as described in Stanisçuaski et al. (2005). Third instars or adults of *D. peruvianus* were fed on artificial cotton seeds, consisting of gelatin capsules (size 2, Elli Lilly Co.) containing cotton seed flour and 0.01–0.04% (w/w) freeze-dried urease. Survival rates were recorded daily until day 16.

### 2.5. Gut homogenates

Homogenates of whole intestines of *D. peruvianus* were used for determination of proteolytic activities. For that, non-starved fourth instars or adult males were anaesthetized in ice and the whole intestines were removed and stored at –20 °C in 50% v/v glycerol in water, in a proportion of 20 intestines per milliliter. To prepare the homogenates, the material was thawed, homogenized manually with an ice-cold Potter pestle, centrifuged at 4 °C at 4000g (twice, 10 min each) and then at 12,000g for 5 min. The final supernatants were kept frozen at –20 °C until used. Homogenates of nymphs' guts and adults' guts are referred from now on as NH and AH, respectively.

### 2.6. Enzymatic assays

For determination of azocaseinolytic activity, the homogenates (2.5 mg protein/mL) were incubated with azocasein (final concentration 0.5%) in sodium acetate 0.05 M, pH 5.6, at 37 °C, in 300  $\mu$ L, with or without 1 mM PMSF. After 1 h, 250  $\mu$ L of 10% trichloroacetic acid (TCA) was added, followed by centrifugation (5 min, 18,000g). The supernatant (450  $\mu$ L) was mixed with 250  $\mu$ L of NaOH (2 M) and read at 420 nm. One unit of azocaseinolytic activity was defined as the amount of enzyme releasing 1.0 A<sub>420</sub> of acid-soluble azopeptides per hour at 37 °C.

Different fluorogenic synthetic substrates were employed to characterize the enzymatic activities in the homogenates. Aliquots of the homogenates were incubated in microplates at 37 °C, with 20  $\mu$ M of substrates Abz-AIAFFSRQ-EDDnp or N-Cbz-Phe-Arg-MCA, in 100  $\mu$ L final volume. Synthetic substrates corresponding to the N- and C-terminal regions flanking the entomotoxic peptide within urease sequence (Mulinari et al., 2007) were also tested, at 20  $\mu$ M final concentration. The substrates were Abz-NAIADGPVQ-EDDnp (equivalent to the N-terminal region) and Abz-KVIRDGMGQ-EDDnp (equivalent to the C-terminal region). The pH was varied using 50 mM citrate-phosphate buffer (pH 2.0–6.0) or Tris-HCl (pH 7.0–9.0). Reactions were monitored for 1 h in an F-Max fluorometer (Molecular Devices Inc.), using 320 nm excitation – 420 nm emission filters for EDDnp substrates and 370 nm excitation – 460 nm emission filters for the MCA substrate. Results are shown as relative fluorescence in milliunits (mRFU) per min per mg of homogenate.

To investigate the classes of the proteinases present in the homogenates, aliquots were pre-incubated 30 min at room temperature with either 10  $\mu$ M E-64, 10  $\mu$ M pepstatin-A, 1 mM EDTA or 1 mM PMSF, and then tested for proteolytic activity (Barret et al., 1998). Results are shown as percentage of the activity in the absence of the inhibitor.

### 2.7. In vitro hydrolysis of urease

Digestion of urease with *D. peruvianus* proteinases was performed as described in Ferreira-DaSilva et al. (2000) using a ratio of 0.5 mU azocaseinolytic activity of homogenates per microgram urease, in 5 mM ammonium formate, pH 5.6, at 37 °C, under continuous stirring. The reaction was stopped by freeze-drying the samples.

### 2.8. Western blot

The hydrolysis of urease was followed by Western blot (Towbin et al., 1979) after SDS-PAGE (Weber and Osborn, 1969) in 10–18% polyacrylamide gels. After being transferred to 0.2  $\mu$ m PVDF filters, the membranes were incubated for 2 h with anti-Jaburetox-2Ec antibodies (1:25,000) and then exposed for 2 h to anti-rabbit IgG coupled to alkaline phosphatase (1:30,000). The color reaction was developed with NBT and BCIP in buffer containing  $MgCl_2$ , pH 9.6.

### 2.9. Determination of cleavage site by LC–ESI-MS

The synthetic substrates (2 nmol) were incubated with the homogenates (100  $\mu$ g protein) at pH 4.0, 37 °C, overnight. The reaction was stopped by sample lyophilization. Digestion analysis was performed on a Q-TOF Micro instrument (Waters Corporation) with a nanoflow electrospray probe, capillary voltage at 3500 V and a cone voltage of 30 V, calibrated in the  $m/z$  region 50–2000 using phosphoric acid. A lock spray interface with phosphoric acid was employed, the lock mass data being averaged over 5 scans for correction. Chromatographic separations for LC–MS were performed in a Nano Acquity Ultra Performance LC (Waters Corporation), equipped with columns Symmetry C18 and BEH 130-C18 (Waters Corporation), and a linear gradient of 0–90% acetonitrile in 0.1% formic acid. The mass spectrometric data were collected in a full scan mode from  $m/z$  50 to 1400 in positive mode. Data were processed and analyzed using Mass Lynx software.

### 2.10. Statistical analysis

Results are expressed as mean  $\pm$  standard deviation. Significance of differences between means was determined using one-way ANOVA. Data were considered statistically different when  $p < 0.05$ .

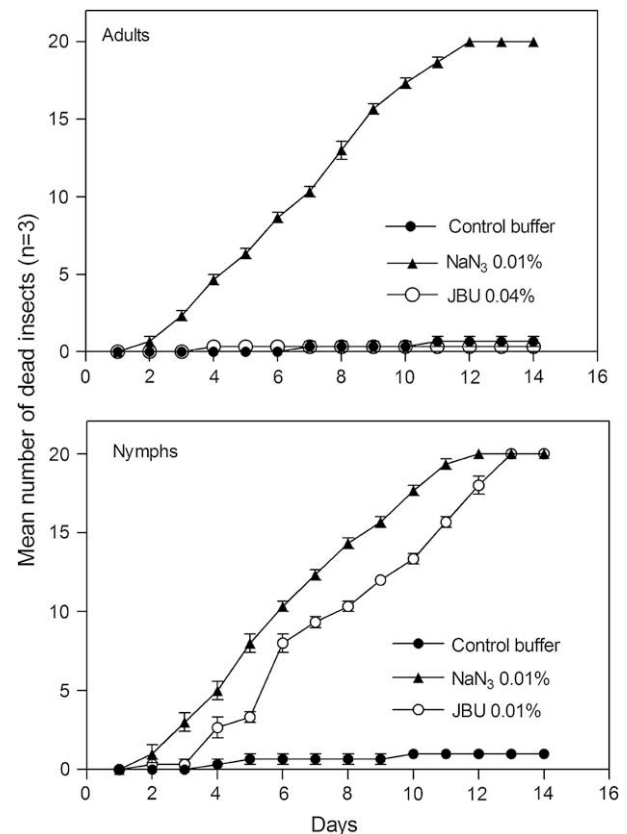
## 3. Results

### 3.1. Urease insecticidal effect

Fig. 1 shows the insecticidal effect of *C. ensiformis* urease fed to third instar *D. peruvianus* reaching 100% of mortality after 11 days. In contrast, adults survived 14 days ingesting a concentration of urease 4-fold higher. The adult insects, however, were as susceptible to sodium azide as the nymphs. These results are very similar to our previous observation for canatoxin, a minor isoform of urease present in the same seed (Stanišcuaski et al., 2005).

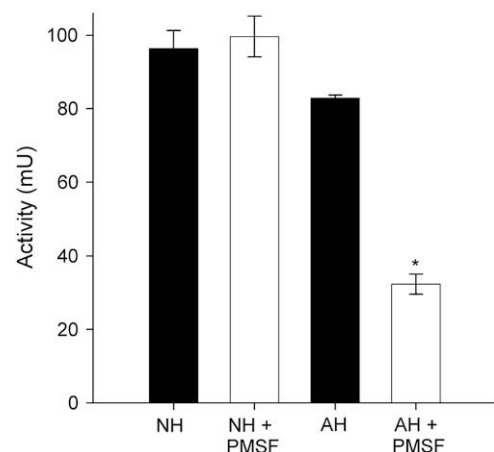
### 3.2. Enzymatic activity upon azocasein

Azocasein is a nonspecific protease substrate. Both homogenates have azocaseinolytic activity at pH 5.6. This activity is



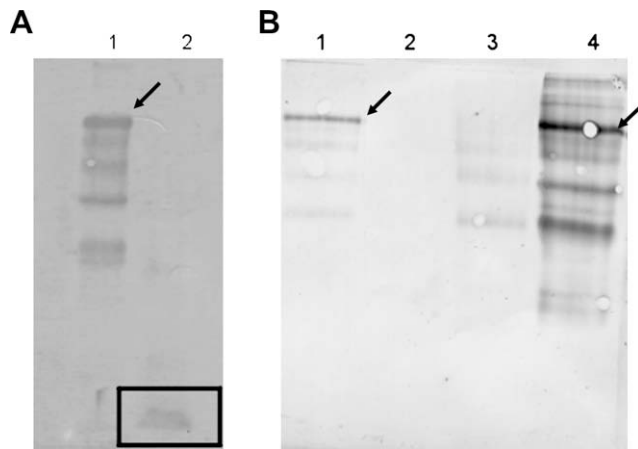
**Fig. 1.** Insecticidal effect of urease on *Dysdercus peruvianus*. Adults (upper panel) or nymphs (third instar) (lower panel) fed on artificial seeds containing urease, buffer alone (controls) or 0.01% w/v sodium azide. Surviving insects were counted daily up to 16 days. Results are expressed as mean and standard deviations of triplicated points ( $N = 20$  per point).

strongly inhibited by PMSF (a serine protease inhibitor) in adult's homogenate (AH) (~65% inhibition). In contrast, no inhibition by PMSF of nymph's homogenate (NH) activity was observed (Fig. 2).



**Fig. 2.** Hydrolysis of azocasein by NH and AH and effect of PMSF. The homogenates (2.5 mg protein/mL) were incubated with 0.5% w/v azocasein at 37 °C, pH 5.6. For the PMSF treatment, homogenates were pre-incubated with the inhibitor (1 mM) for 30 min before the addition of the substrate. The release of soluble azopeptides derived from azocasein was quantified colorimetrically. Activity is expressed as mU (described in Section 2). Values are means and standard deviation ( $N = 4$ ). \* Indicates a statistically significant difference between the control and the PMSF treatment ( $p < 0.05$ ).





**Fig. 3.** Western blot of urease *in vitro* hydrolysis developed with anti-jaburetox antibodies. (A) Urease hydrolysis by NH. Lane 1: urease incubated at 37 °C, pH 5.6 in the absence of NH; lane 2: urease incubated at 37 °C, pH 5.6 in the presence of NH. (B) Urease hydrolysis by AH. Lane 1: urease incubated at 37 °C, pH 5.6 in the absence of AH; lane 2: AH; lane 3: urease incubated at 37 °C, pH 5.6 in the presence of AH. Lane 4: urease incubated at 37 °C, pH 5.6 in the presence of AH and PMSF (1 mM). The arrows indicate intact urease. The square highlights the ~10 kDa fragment recognized by anti-jaburetox antibodies.

### 3.3. *In vitro* hydrolysis of urease

Because the release of the entomotoxic peptide could be due to the concerted or sequential action of more than one enzyme, we decided to analyze homogenates of whole midguts. The fragmentation patterns of urease produced by NH and AH were clearly distinct. After 24 h of incubation at 37 °C, pH 5.6, urease was extensively hydrolyzed by NH, and formed a ~10 kDa fragment, recognized by anti-jaburetox-2Ec antibodies (Fig. 3A). By size and antigenicity criteria this peptide is equivalent to the entomotoxic peptide characterized by Ferreira-DaSilva et al. (2000). Under the same reactions conditions, AH also hydrolyzed urease, but no 10 kDa fragment recognizable by anti-jaburetox-2Ec antibodies was formed (Fig. 3B, lane 3). When the hydrolysis was performed in the presence of PMSF (1 mM), the fragmentation of urease by AH was reduced (Fig. 3B, lane 4), but no inhibition was observed in hydrolysis produced by NH (data not shown).

### 3.4. Enzymatic activity upon synthetic substrates

Four synthetic substrates were used to characterize the proteinases present in the homogenates (data summarized in

Table 1). Abz-AIAFFSRQ-EDDnp was developed for cathepsin-D like aspartic proteinases (Sorgine et al., 2000). Both homogenates have similar activity upon this substrate, but each with a different pH optimum. For NH (Fig. 4, left panels), the optimal pH was 3.0, while AH (Fig. 4, right panels) showed maximal activity at pH 5.0. The activity of NH was completely inhibited by pepstatin-A between pH 2.0 and 6.0 (Fig. 4B). On the other hand, inhibition of AH activity by pepstatin-A was less significant, with a residual activity of 36% at pH 5.0 (Fig. 4F). Although both adult and nymph digestive tracts have aspartic proteinases, the differential susceptibility to inhibitors and the different pH optima suggest that different members of this proteinase family are present according to insect stage. EDTA was also effective in inhibiting AH activity upon this substrate, mainly at pH 6.0 (Fig. 4H), but no inhibition was observed for NH (Fig. 4D).

N-Cbz-Phe-Arg-MCA is a substrate suitable for the cysteine proteases cathepsins B and L (Sorgine et al., 2000; Renard et al., 2000). NH showed two distinct peaks of activity with this substrate (Fig. 5, left panels), one at pH 4.0 and a second at pH 6.0. AH activity was relatively equally distributed over pH 3.0–6.0 (Fig. 5, right panels). The activity of NH upon this substrate was about 2.5-fold higher than AH at pH 4.0 and 2-fold at pH 6.0 (Table 1). E-64 completely abolished the proteolytic activity of both homogenates in the pH 3.0–4.0 range, inhibiting to a lower extent the activity at pH 5.0–7.0 (Fig. 5C and F). Again, the data suggest the presence of different members of the cysteine protease family according to stage of *D. peruvianus*.

We assayed NH and AH protease action on synthetic substrates that correspond to the N- and C-terminal regions flanking the entomotoxic peptide within urease to gain further insights into gut processing of urease. NH has a higher activity upon the substrate Abz-NAIADGVPVQ-EDDnp (equivalent to the N-terminal region), with an optimal pH at 4.0. For AH, the pH optimum is 3.0 (Fig. 6A). While AH activity upon this substrate is inhibited by E-64 (Fig. 6B), the activity of NH is not (Fig. 6E). Pepstatin-A failed to inhibit either AH or NH activities (Fig. 6C and F), while EDTA was a strong inhibitor of both (Fig. 6D and G).

Both homogenates showed similar activities upon the C-terminal substrate, Abz-KVIRDGMGQ-EDDnp, with optimal pH at 4.0–4.5 (Fig. 7A). As observed for the N-terminal substrate, inhibition by E-64 was higher for AH than NH activity (Fig. 7E and B). At pH 4.0, E-64 inhibited about 50% of NH activity and almost 100% AH activity. Pepstatin-A had no effect (Fig. 7C and F) on either of the gut homogenates. Contrasting with the results with the N-terminal substrate, NH and AH activities were not significantly inhibited by EDTA (Fig. 7D and G).

**Table 1**

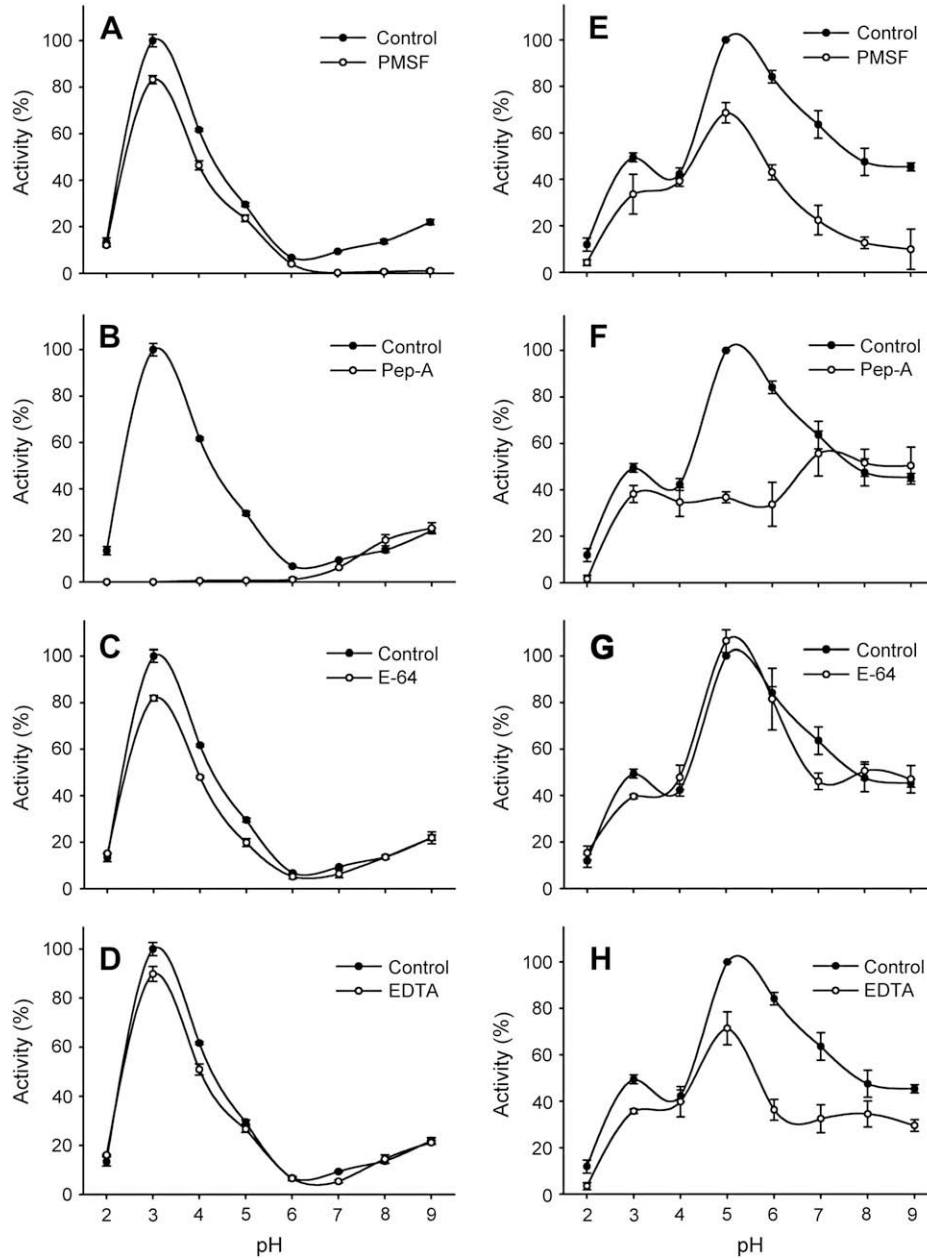
Major proteolytic activities of midgut homogenates of nymphs (third instars) or adult *Dysdercus peruvianus*.

Substrate	Nymph's midgut homogenates			Adult's midgut homogenates		
	Optimal pH	mU/min/mg or $A_{280}/3 \text{ h/mg}^a$	Inhibitor(s) <sup>b</sup>	Optimal pH	mU/min/mg or $A_{280}/3 \text{ h/mg}^a$	Inhibitor(s) <sup>b</sup>
Azocasein, pH 5.6	n.d.	95.0	Not inhibited by PMSF	n.d.	82.0	PMSF
Hemoglobin**	3.5	0.26	E-64, EDTA, pepstatin-A	5.0	0.20	Pepstatin-A
				8.5	0.11	PMSF
Abz-AIAFFSRQ-EDDnp (substrate designed for cathepsin-D)	3.0	820	Pepstatin-A	3.0	400	No inhibition
				5.0	800	PMSF, EDTA, Pepstatin-A
N-Cbz-Phe-Arg-MCA (substrate designed for cathepsin-B)	4.0	3700	E-64	3–5	1600	E-64
	6.0	2100	E-64			
Abz-NAIADGVPVQ-EDDnp (N-terminal flanking sequence)	4.5	22,000	EDTA	3.0	9000	E-64, EDTA
Abz-KVIRDGMGQ-EDDnp (C-terminal flanking sequence)	4.5	25,000	E-64	4.0	30,000	E-64

n.d. – Not determined.

<sup>a</sup> Hemoglobinolytic activity was measured as  $A_{280}/3 \text{ h/mg}$  (\*\* data taken from Stanisçuaski et al. (2005)).

<sup>b</sup> Inhibition greater than 40–50%. Inhibitors tested were 10  $\mu\text{M}$  pepstatin-A, 10  $\mu\text{M}$  E-64, 1 mM PMSF and 1 mM EDTA.



**Fig. 4.** Effect of inhibitors on Abz-AIAFFSRQ-EDDnp hydrolysis by NH (A–D) and AH (E–H). Final concentration of the substrate was 20  $\mu$ M. Activity was assayed at 37 °C for 1 h in the presence of 1 mM PMSF (A, E), 10  $\mu$ M pepstatin-A (B, F), 10  $\mu$ M E-64 (C, G) or 1 mM EDTA (D, H). Results are expressed as mean and standard deviation ( $N = 4$ ).

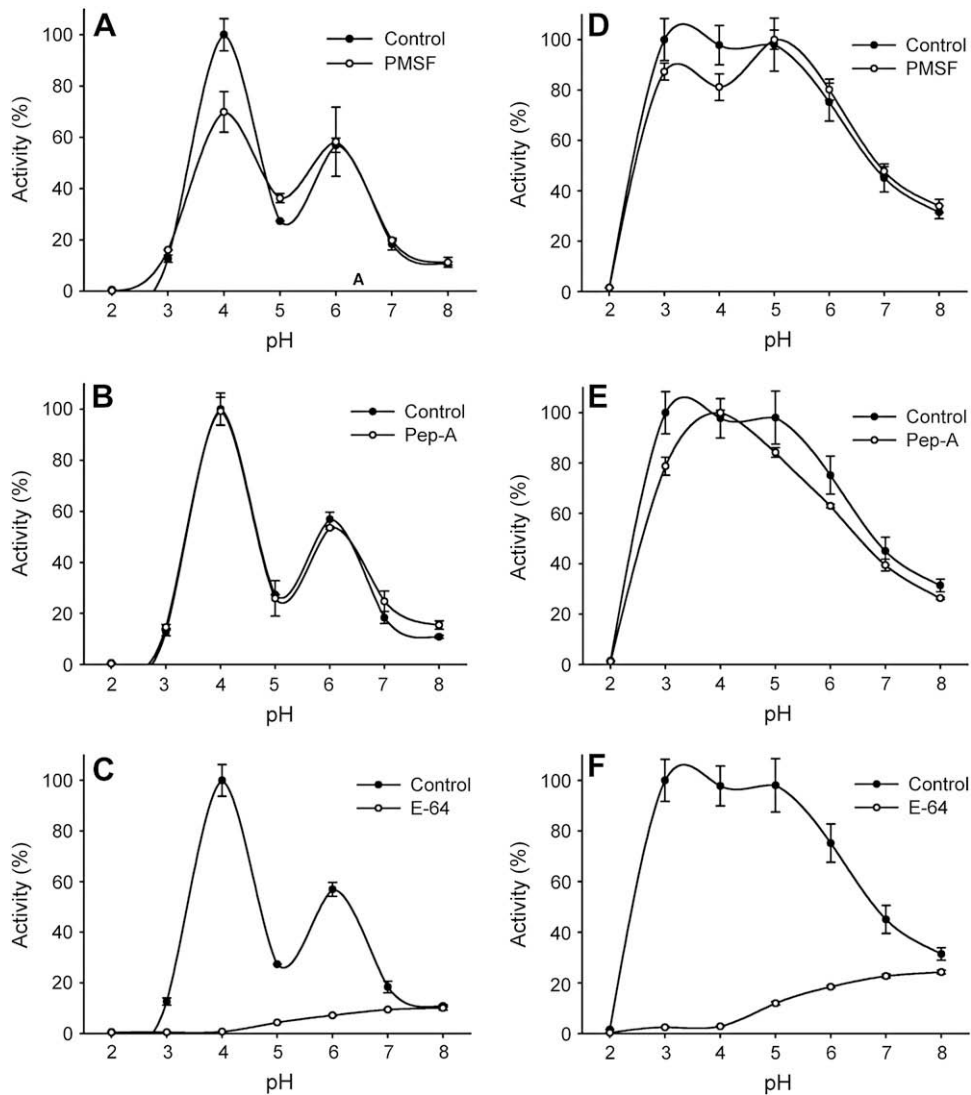
### 3.5. Determination of cleavage site by LC-ESI-MS

The cleavage site(s) of Abz-NAIADGPVQ-EDDnp and Abz-KVIRDGMGQ-EDDnp by the homogenates were determined by LC-ESI-MS (Table 2). For the hydrolysis of the N-terminal substrate Abz-NAIADGPVQ-EDDnp at pH 4.0 by either of the homogenates, the fragment DGPVQ-EDDnp ( $m/z$  723,302) was the most abundant. The fragment GPVQ-EDDnp was also present, but less abundant. These results indicate that both homogenates preferentially cleave between Ala and Asp residues in this substrate, but they are also effective in cleaving the bond between Asp and Gly residues. When the hydrolysis was performed at pH 5.6, which resembles the condition of the midgut of *D. peruvianus*, the preferred cleavage site in Abz-NAIADGPVQ-EDDnp was Asp–Gly for both homogenates (data not shown). In the case of the hydrolysis of the C-terminal substrate Abz-KVIRDGMGQ-EDDnp by the homogenates at pH 4.0,

the most abundant fragment was DGMGQ-EDDnp ( $m/z$  317,703), indicating that AH and NH preferentially cleave between Arg and Asp residues. For the hydrolysis with AH, a second fragment, Abz-KVIRD ( $m/z$  374,204), was observed with a lower intensity. This indicates that AH, but not NH, also cleaves the peptide bond between Asp and Gly residues in this substrate. At pH 5.6, most of Abz-KVIRDGMGQ-EDDnp was cleaved between the Ile and Arg residues by both homogenates (data not shown).

## 4. Discussion

The primary aim of this work was to investigate the major digestive proteolytic activities of nymphs and adults of *D. peruvianus* and to establish a possible relation between these activities and the stage-dependent resistance to toxic effects of ingested ureases. Because the release of the entomotoxic peptide could be



**Fig. 5.** Effect of inhibitors on N-Cbz-Phe-Arg-MCA hydrolysis by NH (A–C) and AH (D–F). Final concentration of the substrate was 20 μM. Activity was assayed at 37 °C for 1 h in the presence of 1 mM PMSF (A, D), 10 μM pepstatin-A (B, E) or 1 mM EDTA (C, F). Results are expressed as mean and standard deviation ( $N = 4$ ).

due to the concerted or sequential action of more than one enzyme, we decided to work with homogenates of whole midguts.

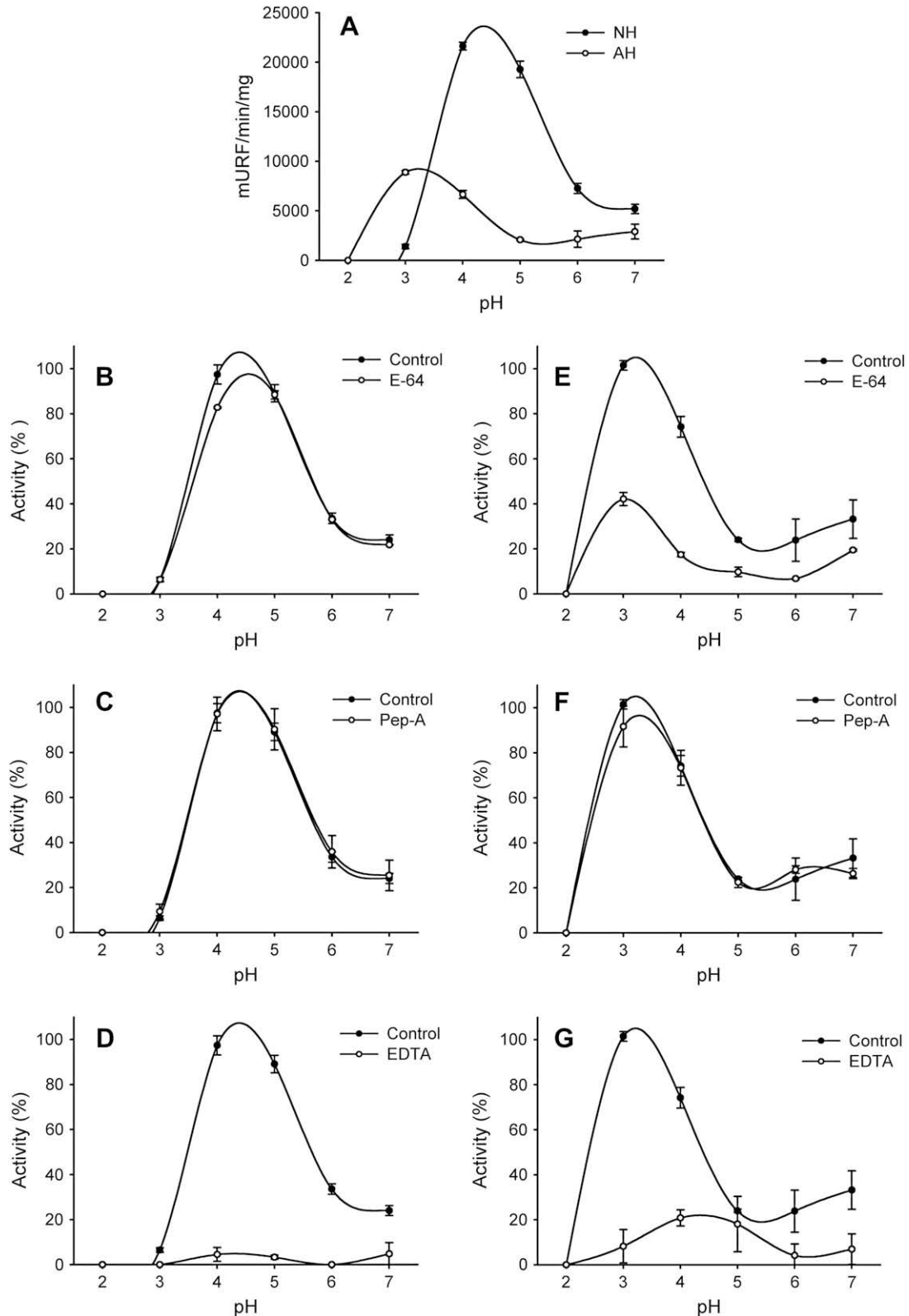
Ferreira-DaSilva et al. (2000) confirmed the proteolytic activation of canatoxin by enzymes from *Callosobruchus maculatus* larvae, and the release of a toxic peptide of 10 kDa. Here, we showed that *C. ensiformis* major urease is hydrolyzed by *D. peruvianus* NH and AH, but only NH is able to release a ~10 kDa peptide, recognized by anti-Jaburetox-2Ec antibodies. Although we cannot exclude at this point differences regarding midgut permeability, receptor(s), transport across the membranes of epithelial cells, or signal-transduction, this finding strongly suggests that adults of *D. peruvianus* are resistant to urease toxic effects due to differences in the proteolytic processing of the protein. Differences in the response to toxins during the life cycle of insects have been reported. Increased resistance to *Bacillus thuringiensis* Cry1C toxin in the late larval stages of *Spodoptera littoralis* as compared to young larvae was attributed to different enzymatic processing of the toxin (Keller et al., 1996).

Silva and Terra (1994) described digestive enzymes, absorptive sites and internal pH along the midgut, divided into three sections (V1–V3), of female adults *D. peruvianus*. They showed a luminal pH variation from 6.2 to 5.6 and that a soluble cysteine proteinase with a pH 5.5 optimum and produced in the V2 region was the most important proteolytic enzyme. However, this type of information is

not available for young forms of any hemipteran. It is reasonable to expect that more acidic microenvironments probably exist in which the enzymes we described here (particularly in nymphs) would approach their maximal activities.

Comparing the enzyme activities in homogenates of fed and starved *D. peruvianus*, Silva and Terra (1994) concluded that cotton meal contributes with an aminopeptidase activity found in the midgut contents. It is not known if symbionts present in hemipteran midgut contribute to the repertoire of enzymes found (Terra and Ferreira, 1994; Silva and Terra, 1994). Since in the present study we used non-starved insects, both stages fed on the same lot of cotton seeds, the differences we found in the repertoire of proteolytic enzymes are likely to be stage-specific. Thus, although there is no change in the feeding habits of young forms and adult *D. peruvianus*, our data point to important differences in the digestive physiology of the insect during its life cycle.

We showed that NH azocaseinolytic activity was not affected by the serine protease inhibitor PMSF, while AH activity was reduced almost 70% (Fig. 2), confirming similar observations in a previous study (Stanisçuaski et al., 2005). More interestingly, urease hydrolysis by AH is inhibited by PMSF, while NH is still able to fully degrade urease in the presence of this inhibitor (data not shown). The data indicate the presence of serine proteinases in adults, but

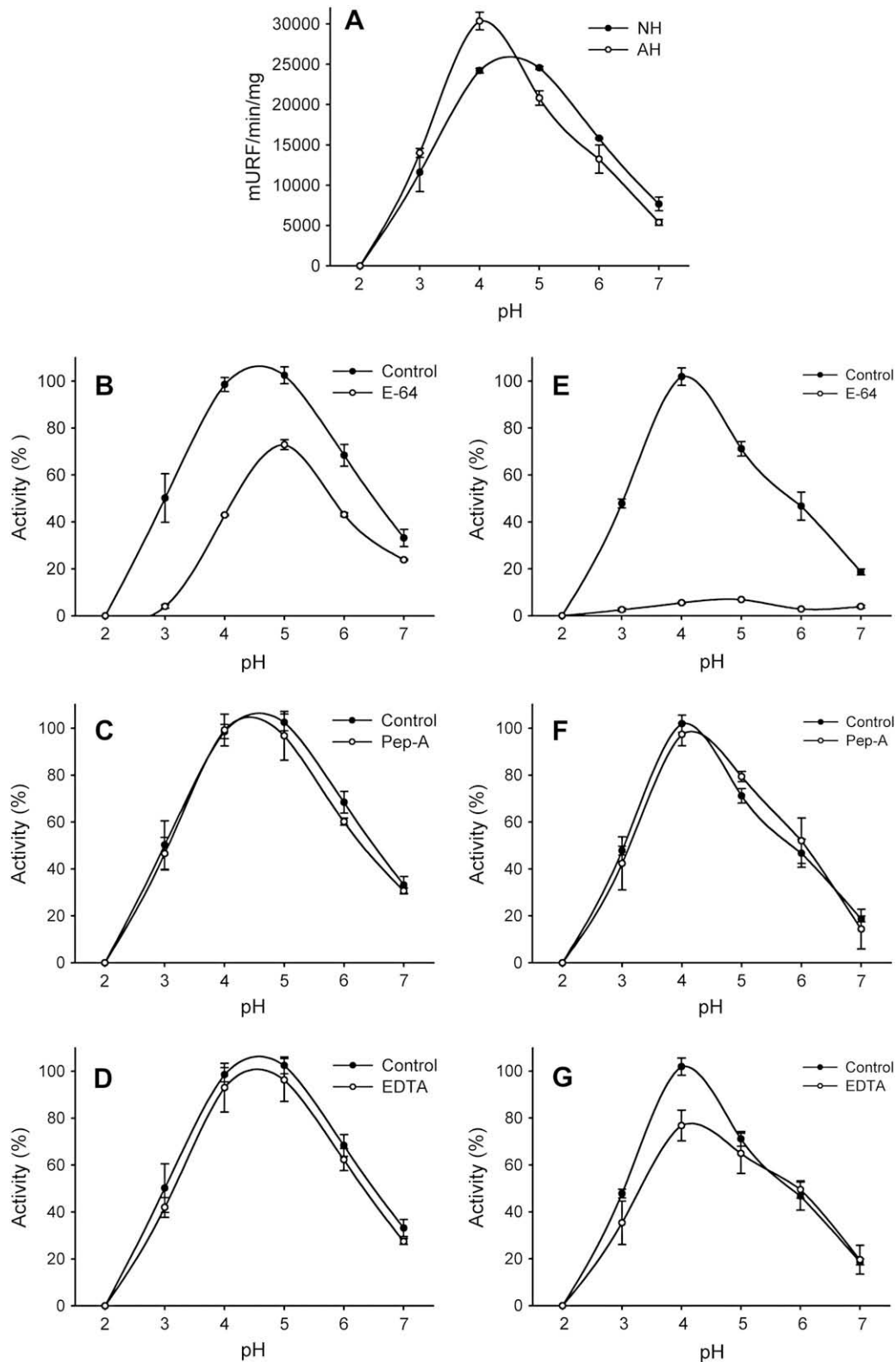


**Fig. 6.** Effect of pH and inhibitors on the hydrolysis of Abz-NAIADGPVQ-EDDnp (final concentration 20  $\mu$ M) by NH and AH. Activity was assayed at 37  $^{\circ}$ C for 1 h in the absence (A) or presence of 10  $\mu$ M E-64 (B, E), 10  $\mu$ M pepstatin-A (C, F) or 1 mM EDTA (D, G). Results are expressed as specific activity (A) or mean  $\pm$  s.d. (B–G) ( $N = 4$ ).

not in nymphs, of *D. peruvianus*. Hemipteran insects are said to rely on cathepsin-like digestive enzymes (Terra and Ferreira, 1994; Cristofolletti et al., 2003). To our knowledge, there are no published reports on the identification of serine trypsin-like digestive enzymes in any hemipteran insect. One alternative explanation

would be that PMSF is nonspecifically inhibiting cysteine proteases (Whitaker and Perez-Villasenor, 1968).

Midgut homogenates of both insect stages were able to cleave several substrates in acidic media up to pH 6.0. The activities upon N-Cbz-Phe-Arg-MCA and Abz-AIAFFSRQ-EDDnp indicate the



**Fig. 7.** Effect of pH and inhibitors on the hydrolysis of Abz-KVIRDGMGQ-EDDnp (final concentration 20  $\mu$ M) by NH and AH. Activity was assayed at 37 °C for 1 h in the absence (A) or presence of 10  $\mu$ M E-64 (B, E), 10  $\mu$ M pepstatin-A (C, F) or 1 mM EDTA (D, G). Results are expressed as specific activity (A) or mean  $\pm$  s.d. (B–G) ( $N = 4$ ).

presence of cysteine and aspartic proteinases in both homogenates. NH activity upon Abz-AIAFFSRQ-EDDnp was completely inhibited by pepstatin-A (Fig. 4B), while AH was still partially active in the presence of pepstatin-A (Fig. 4F). The data suggest the presence of a different isoform of aspartic proteinase in adults, less sensitive to

pepstatin-A, or alternatively, adults have another proteinase able to cleave this substrate, not susceptible to the inhibitor. Inhibition by E-64 of the cleavage of N-Cbz-Phe-Arg-MCA indicated the presence of cysteine proteinases in both AH and NH (Fig. 5), the activities differing in their pH optima and specific activities.

**Table 2**

Determination by LC–ESI–MS of cleavage sites of substrates by homogenates at pH 4.0.

Sequence	Mass (predicted)	Mass (observed)	
		Cleaved by AH	Cleaved by NH
<i>N-terminal</i>			
<b>Abz-NAIADGPVQ-EDDnp</b>	506,247	506,246	506,246
Abz-NAIADGPVQ-EDDnp	722,295	722,294 <sup>a</sup>	722,295 <sup>a</sup>
<b>Abz-NAIADGPVQ-EDDnp</b>	621,277	621,280	621,281
Abz-NAIADGPVQ-EDDnp	607,275	607,271	607,267
<i>C-terminal</i>			
<b>Abz-KVIRDGMGQ-EDDnp</b>	633,397	633,383 <sup>a</sup>	633,391 <sup>a</sup>
Abz-KVIRDGMGQ-EDDnp	714,235	714,238	n.o.
<b>Abz-KVIRDGMGQ-EDDnp</b>	748,417	748,409	n.o.
Abz-KVIRDGMGQ-EDDnp	599,215	n.o.	n.o.

The fragments observed are shown in bold. n.o. – not observed.

<sup>a</sup> The fragments that were more abundant in the spectrum.

Important differences were observed between NH and AH regarding their activity upon the substrates representing the regions flanking the entomotoxic peptide within the intact urease. We have previously shown that *in vitro* hydrolysis of canatoxin by cathepsins from *C. maculatus* larvae produced a family of entomotoxic peptides, the largest of which, pepcanatox, was the most toxic, the smaller ones probably representing further cleavage of the latter (Ferreira-DaSilva et al., 2000). The N-terminal sequence of pepcanatox was determined by Edman sequencing to be GPVQ- and based on its molecular mass its C-terminal was deduced to be –KVIRD (Mulinari et al., 2007). Here we observed that both homogenates cleaved the N-terminal substrate, and that this hydrolysis is not catalyzed by pepstatin-A-sensitive enzymes. Only AH was inhibited by E-64 when assayed with this substrate. Since EDTA was the only effective inhibitor of the hydrolysis of the N-terminal substrate catalyzed by NH, we suggest that a metalloprotease active at pH 4.0–5.0 is probably involved in the limited proteolysis of urease that releases the entomotoxic peptide in young forms of *D. peruvianus*. The LC–ESI–MS data (Table 2) indicated that both AH and NH at pH 4.0, in which all the activities studied are at or near to their maximal levels, cleave the substrate Abz-NAIADGPVQ-EDDnp preferentially between the residues Ala and Asp. Metalloproteases, as the peptidyl-Asp metallopeptidase, hydrolyzes bonds at the N-terminal side of aspartic acid (Drapeau, 1980; Hagmann, 2004), reinforcing our hypothesis that an enzyme of this class releases the N-terminal side of the entomotoxic peptide from within urease.

The cleavage of C-terminal substrate Abz-KVIRDGMGQ-EDDnp by AH and NH at pH 4.0 occurred between the residues Arg and Asp. The activity of both homogenates upon this substrate was inhibited significantly only by E-64, indicating the participation of cysteine proteinases. Cleavage of RD bonds is described for human cathepsin L (Dahl et al., 2001) and also for trypsin-1 (Nemoto et al., 1997). The data on azocasein hydrolysis suggested that adults have trypsin-like enzymes, while PMSF-sensitive enzymes were not found in nymphs. Diaz-Mendoza et al. (2005) reported that E-64 nonspecifically inhibited trypsin-1 enzymes from insects. The more pronounced effect of E-64 on AH activity upon the C-terminal substrate could thus represent the inhibition of cysteine proteinases plus nonspecific inhibition of trypsin-like enzymes. This may be an important difference in the limited proteolysis of urease by the two stages of *D. peruvianus*. Within the entomotoxic peptide sequence (Mulinari et al., 2007) there are 11 potential cleavage sites for trypsin. If adults in fact have a trypsin-like enzyme(s), it is possible that the entomotoxic peptide is released from urease, but then degraded into smaller, not toxic, fragments. Another fact that supports the hypothesis that the entomotoxic peptide is degraded is that Jaburetox-2Ec is not toxic to *D. peruvianus* adults when administered orally (Stanisçuaski et al., 2005).

A second cleavage site, between Asp and Gly, was observed for both homogenates with Abz-NAIADGPVQ-EDDnp and for AH with Abz-KVIRDGMGQ-EDDnp. When the hydrolysis was performed at pH 5.6, Asp–Gly was the only cleavage site identified in Abz-NAIADGPVQ-EDDnp (data not show). The hydrolytic activities upon these substrates at pH 5.0–6.0 are comparatively very low (Figs. 6A and 7A). Asp–Gly is not a consensus cleavage site for cathepsins either from the aspartic or cysteine protease families. The Peptidase Databank MEROPS (<http://merops.sanger.ac.uk>) provides literature showing that the bond between Asp–Gly can be cleaved by cysteine proteases such as caspases (Hawkins et al., 2000), metalloproteases as falcilysin (Eggleston et al., 1999), serine proteases as dipeptidase E (Miller, 2004) and also by aspartyl-peptidases (Wilk, 2004). Brinkworth et al. (2001) have shown that cathepsin-D from blood-feeding parasites has unusual cleavage sites, when compared to mammalian cathepsin-D. These differences in substrate cleavage preference reflected subtle but significant differences in the enzyme binding pockets. It is possible that, in the course of evolution, insects had some of their enzymes adapted to their feeding habits and hence diverged their substrates preferences from the expected.

Although we cannot exclude that a single enzyme in *D. peruvianus* nymphs is responsible for the limited proteolysis that releases the entomotoxic peptide from urease, our data indicate that more than one class of enzyme in their midguts is able to hydrolyze the synthetic peptides containing the potential cleavage points, particularly at the C-terminal portion of the peptide. Indeed, in a previous study (Carlini et al., 1997) we demonstrated that simultaneous administration of E-64 or pepstatin-A and canatoxin to third instars *Rhodnius prolixus* significantly reduced the lethality rate, suggesting that both cysteine and aspartic proteinases are involved in the proteolytic activation of the protein. Another important point concerning these results is that the whole native urease is actually not the only substrate to be considered as a partially cleaved protein could be a better substrate for the enzyme(s) releasing the internal entomotoxic peptide.

Table 1 summarizes our data showing the presence of multiple, stage-specific proteinases in the digestive system of *D. peruvianus*. Purification and physicochemical characterization of some of the enzymatic activities detected in this work will be dealt with elsewhere. Insect digestive proteolytic activity is dynamic and depends on the interactions among tissue type, pH, proteinase class and phase of life cycle (Wright et al., 2006). Such complexity could be directly involved in the processing and inactivation of toxic defense proteins produced by plants (Peumans and Van Damme, 1995). The difference in susceptibility to ureases of adults and nymphs of *D. peruvianus* is probably a complex event, with the contribution of multiple factors. Here we have demonstrated that stage-specific differences in the release of the entomotoxic peptide from the intact protein might contribute to the resistance of adult *D. peruvianus* to urease entomotoxic effects. The data also emphasize the need for adequate understanding of insect physiology when considering the potential biotechnological use of insecticidal proteins in protecting crops against insect pests.

## Acknowledgments

The authors are thankful to Antonio Polleti from Núcleo de Produção de Sementes de Aguai, SP, Brazil, for the cotton seeds supply, to Roseli Garcia, for maintaining the insect colony, and to Dr. Joseph C. Polacco, Univ. Missouri, Columbia, for critical reading of the manuscript. This project was supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and FAPERGS (Fundação de Amparo a Pesquisa do Rio Grande do Sul).

## References

- Barret, A.J., Rawlings, N.D., Woessner, J.F., 1998. Handbook of Proteolytic Enzymes. Academic Press, San Diego, 1666 pp.
- Brinkworth, R.I., Provic, P., Loukas, A., Brindley, P.J., 2001. Hemoglobin-degrading, aspartic proteases of blood-feeding parasites. *J. Biol. Chem.* 276, 38844–38851.
- Carlini, C.R., Guimarães, J.A., 1981. Isolation and characterization of a toxic protein from *Canavalia ensiformis* (jack bean) seeds, distinct from concanavalin A. *Toxicon* 19, 667–676.
- Carlini, C.R., Oliveira, A.E., Azambuja, P., Xavier-Filho, J., Wells, M.A., 1997. Biological effects of canatoxin in different insect models. Evidence for a proteolytic activation of the toxin by insect cathepsinlike enzymes. *J. Econ. Entomol.* 90, 340–348.
- Cristofaletti, P.T., Ribeiro, A.F., Deraison, C., Rahbé, Y., Terra, W.R., 2003. Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrtosiphon pisum*. *J. Insect Physiol.* 49, 11–24.
- Dahl, S.W., Halkier, T., Lauritzen, C., Dolenc, I., Pedersen, J., Turk, V., Turk, B., 2001. Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing. *Biochemistry* 40, 1671–1678.
- Díaz-Mendoza, M., Ortego, F., García de Lacoba, M., Magana, C., de la Poza, M., Farinos, G.P., Castanera, P., Hernandez-Crespo, P., 2005. Diversity of trypsins in the Mediterranean corn borer *Sesamia nonagrioides* (Lepidoptera: Noctuidae), revealed by nucleic acid sequences and enzyme purification. *Insect Biochem. Mol. Biol.* 35, 1005–1020.
- Drapeau, G.R., 1980. Substrate specificity of a proteolytic enzyme isolated from a mutant of *Pseudomonas fragi*. *J. Biol. Chem.* 255, 839–840.
- Eggleston, K.K., Duffin, K.L., Goldberg, D.E., 1999. Identification and characterization of falcylisin, a metalloproteinase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* 274, 32411–32417.
- Ferreira-DaSilva, C.T., Gombarovits, M.E., Masuda, H., Oliveira, C.M., Carlini, C.R., 2000. Proteolytic activation of canatoxin, a plant toxic protein, by insect cathepsin-like enzymes. *Arch. Insect Biochem. Physiol.* 44, 162–171.
- Follmer, C., 2008. Insights into the role and structure of plant ureases. *Phytochemistry* 69, 18–28.
- Follmer, C., Barcellos, G.B.S., Zingali, R.B., Machado, O.L.T., Alves, E.W., Barja-Fidalgo, C., Guimarães, J.A., Carlini, C.R., 2001. Canatoxin, a toxic protein from jack beans (*Canavalia ensiformis*), is a variant from of urease (EC 3.5.1.5): biological effects of urease independent of its ureolytic activity. *Biochem. J.* 360, 217–224.
- Follmer, C., Real-Guerra, R., Wassermann, G.E., Olivera-Severo, D., Carlini, C.R., 2004a. Jackbean, soybean and *Bacillus pasteurii* ureases: biological effects unrelated to ureolytic activity. *Eur. J. Biochem.* 271, 1357–1363.
- Follmer, C., Wassermann, G., Carlini, C.R., 2004b. Separation of Jack bean (*Canavalia ensiformis*) urease isoforms by immobilized metal affinity chromatography and characterization of insecticidal properties unrelated to ureolytic activity. *Plant Sci.* 167, 241–246.
- Gallo, D., 1988. Manual de Entomologia Agrícola, second ed. CERES, São Paulo, 649 pp.
- Hagmann, M.L., 2004. Peptidyl-Asp Metalloendopeptidase. In: Handbook of Proteolytic Enzymes, second ed. Elsevier, London, pp. 1037–1039.
- Hawkins, C.J., Yoo, S.J., Peterson, E.P., Wang, S.L., Vermooy, S.Y., Hay, B.A., 2000. The *Drosophila* caspase DRONC cleaves following glutamate or aspartate and is regulated by DIAP1, HID, and GRIM. *J. Biol. Chem.* 275, 27084–27093.
- Keller, M., Sneh, B., Strizhov, N., Prudovsky, E., Regev, A., Koncz, C., Schell, J., Zilberstein, A., 1996. Digestion of delta-endotoxin by gut proteases may explain reduced sensitivity of advanced instar larvae of *Spodoptera littoralis* to CryIC. *Insect Biochem. Mol. Biol.* 26, 365–373.
- MEROPS, <http://merops.sanger.ac.uk/>.
- Miller, C.G., 2004. Dipeptidase E. In: Handbook of Proteolytic Enzymes, second ed. Elsevier, London, pp. 2033–2035.
- Mobley, H.L., Island, M.D., Hausinger, R.P., 1995. Molecular biology of microbial ureases. *Microbiol. Rev.* 59, 451–480.
- Mulinari, F., Freitas-Silva, M.A., Grossi-De-Sa, M.F., Moraes, M.G., Kurtenbach, E., Carlini, C.R., 2004. Toxina Praguicida, Construção Gênica e Método de Controle de Pragas. Patent registered at National Institute for Intellectual Property (INPI). Brazil; 08/04/2004 [protocol BRPI0403435A].
- Mulinari, F., Stanisçuaski, F., Bertholdo-Vargas, L.R., Postal, M., Oliveira-Neto, O.B., Ridgen, D.J., Grossi-de-Sá, M.F., Carlini, C.R., 2007. Jaburetox-2Ec: an insecticidal peptide derived from an isoform of urease from the plant *Canavalia ensiformis*. *Peptides* 28, 2042–2050.
- Nemoto, T., Sato, N., Iwanari, H., Yamashita, H., Takagii, T., 1997. Domain structures and immunogenic regions of the 90-kDa heat-shock protein (HSP90). *J. Biol. Chem.* 272, 26179–26187.
- Peumans, W.J., Van Damme, E.J., 1995. Lectin as plant defense proteins. *Plant Physiol.* 109, 347–352.
- Polacco, J.C., Holland, M.A., 1993. Roles of urease in plant cells. *Int. Rev. Cytol.* 145, 65–103.
- Renard, G., Garcia, J.F., Cardoso, F.C., Richter, M.F., Sakanari, J.A., Ozaki, L.S., Termignoni, C., Masuda, A., 2000. Cloning and functional expression of a *Boophilus microplus* cathepsin L-like enzyme. *Insect Biochem. Mol. Biol.* 30, 1017–1026.
- Silva, C.P., Terra, W.R., 1994. Digestive and absorptive sites along the midgut of the cotton seed sucker bug *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae). *Insect Biochem. Mol. Biol.* 24, 493–505.
- Silva, C.P., Ribeiro, A.F., Terra, W.R., 1996. Enzyme markers and isolation of the microvillar and perimicrovillar membranes of *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae) midgut cells. *Insect Biochem. Mol. Biol.* 26, 1011–1018.
- Sirko, A., Brodzik, R., 2000. Plant ureases: role and regulation. *Acta Biochim. Pol.* 47, 1189–1195.
- Sorgine, M.H., Logullo, C., Zingali, R.B., Paiva-Silva, G.O., Juliano, L., Oliveira, P.L., 2000. A heme-binding aspartic proteinase from the eggs of the hard tick *Boophilus microplus*. *J. Biol. Chem.* 275, 28659–28665.
- Stanisçuaski, F., Ferreira-Dasilva, C.T., Mulinari, F., Pires-Alves, M., Carlini, C.R., 2005. Insecticidal effects of canatoxin on the cotton stainer bug *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae). *Toxicon* 45, 753–760.
- Terra, W.R., Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol.* 109, 1–62.
- Terra, W.R., Ferreira, C., Jordão, B.P., Dillon, R.J., 1996. Digestive Enzymes. In: Biology of the Insect Midgut. Chapman & Hall, London, pp. 153–194.
- Tomazzeto, G., Mulinari, F., Stanisçuaski, F., Settembrini, B.P., Carlini, C.R., Ayub, M.A.Z., 2007. Expression kinetics and plasmid stability of recombinant *E. coli* encoding urease-derived peptide with bioinsecticide activity. *Enzyme Microb. Technol.* 41, 821–827.
- Towbin, H., Staehelin, T., Gordon, J., 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76, 4350–4354.
- Weber, K., Osborn, M., 1969. The reliability of molecular weight determinations by dodecyl sulfate polyacrylamide gel electrophoresis. *J. Biol. Chem.* 244, 4406–4412.
- Whitaker, J.R., Perez-Villasenor, J., 1968. Chemical modification of papain. 1. Reaction with the chloromethyl ketones of phenylalanine and lysine and with phenylmethylsulphonyl fluoride. *Arch. Biochem. Biophys.* 124, 70–78.
- Wilhite, S.E., Elden, T.C., Brzin, J., Smigocki, A.C., 2000. Inhibition of cysteine and aspartyl proteinases in the alfalfa weevil midgut with biochemical and plant-derived proteinase inhibitors. *Insect Biochem. Mol. Biol.* 30, 1181–1188.
- Wilk, S., 2004. Beta-aspartyl peptidase. In: Handbook of Proteolytic Enzymes, second ed. Elsevier, London, pp. 2115–2116.
- Wright, M.K., Brandt, S.L., Coudron, T.A., Wagner, R.M., Habibi, J., Backus, E.A., Huesing, J.E., 2006. Characterization of digestive proteolytic activity in *Lygus hesperus* knight (Hemiptera: Miridae). *J. Insect Physiol.* 52, 717–728.

**Anexo IV**

**Manuscrito à ser submetido ao periódico Biochemical Journal**



**Biochemical and structural studies on native and recombinant *Glycine max* UreG: the first detailed characterization of a plant urease accessory protein.**

Rafael REAL-GUERRA\*<sup>†</sup>, Fernanda STANISCUASKI\*<sup>‡</sup>, Barbara ZAMBELLI<sup>§</sup>,  
Francesco MUSIANI<sup>§</sup>, Stefano CIURLI<sup>§,||</sup> and Célia R. CARLINI\*<sup>‡</sup>

\*Graduate program in Cellular and Molecular Biology, Center of Biotechnology, Institute of Biosciences, Universidade Federal do Rio Grande do Sul, Rio Grande do Sul, Brazil.

<sup>‡</sup>Center of Biotechnology & Department of Biophysics, Institute of Biosciences, Universidade Federal do Rio Grande do Sul, Rio Grande do Sul, Brazil.

<sup>§</sup>Laboratory of Bioinorganic Chemistry, Department of Agro-environmental Science and Technology, University of Bologna, Bologna, Italy.

<sup>||</sup>CERM (Centre for Magnetic Resonance), University of Florence, Sesto Fiorentino, Italy.

<sup>†</sup>Corresponding author: rafael.guerra@ufrgs.br

Address: Av. Bento Gonçalves, 9500 Building 43422 Room 204. Porto Alegre, RS, Brazil. 91501-970.

Phone: +55 51 3308-7600

Fax: +55 51 3308-7606

## Synopsis

Bacterial urease accessory proteins have been target of biochemical characterization *in vivo* and *in vitro*. In contrast, little is known about accessory proteins from plants. This is the first detailed biochemical characterization of a plant urease accessory protein, the soybean UreG. Native *Glycine max* UreG (*GmUreG*) was purified from seeds and a recombinant protein was produced in *Escherichia coli*. Native *GmUreG* separates into two fractions upon metal affinity chromatography, both “forms” showing identical sequences as determined by tandem mass spectrometry. The two forms of native *GmUreG* and the recombinant protein were compared regarding their secondary structure content by circular dichroism spectroscopy, metal binding affinity and the effects of  $\text{Ni}^{2+}$  and  $\text{Zn}^{2+}$  on their GTPase activity. Isothermal titration calorimetry showed that *GmUreG* binds  $\text{Zn}^{2+}$  and  $\text{Ni}^{2+}$ , with a very tight binding site for  $\text{Zn}^{2+}$  thus reinforcing the hypothesis that  $\text{Zn}^{2+}$  ions may play a role on the urease assembly process. Size exclusion chromatography and light scattering techniques showed that  $\text{Zn}^{2+}$  is able to stabilize the dimeric form of the protein. A homology model for the fully folded *GmUreG* was built. NMR measurements indicate that *GmUreG*, as its bacterial counterparts, belongs to the class of intrinsically unstructured proteins.

Short title: Biochemical characterization of *Glycine max* UreG, a urease accessory protein.

Key words: GTPase, metal binding, intrinsically unfolded, nucleotide binding, histidine-rich, chaperone

### Abbreviations

*GmUreG*: *Glycine max* UreG accessory protein

IPTG: Isopropyl- $\beta$ -D-thiogalactopyranoside

Q-TOF: Quadrupole Time-of-Flight tandem Mass Spectrometer

MALS: Multiple Angle Light Scattering

QELS: Quasi-Elastic Light Scattering

ITC: Isothermal Titration Calorimetry

ICP-OES: Inductively Coupled Plasma Emission Spectroscopy

*MjHypB*: *Methanococcus jannaschii* HypB

*MtUreG*: *Mycobacterium tuberculosis* UreG accessory protein

*BpUreG*: *Bacillus pasteurii* UreG accessory protein

*KaUreG*: *Klebsiella aerogenes* UreG accessory protein

*HpUreG*: *Helicobacter pylori* UreG accessory protein

$K_d$ : dissociation constant

$K_{cat}$ : catalytic constant

## INTRODUCTION

Despite the great amount of effort put on research in the field, the understanding of the machinery behind the maturation of nickel-containing enzymes has still a long way to go until the complete process can be understood. Considering the best characterized systems – hydrogenase, carbon monoxide dehydrogenase and urease –, all show a multistep, tightly regulated mechanism, involving the participation of a host of accessory proteins for the correct assembly of their active sites [1]. Interestingly, despite the specific particularities of each process, a few general aspects involving the participation of the accessory proteins seem to be very well conserved among them. In each system, there seems to be a requirement for a nucleotide triphosphate binding accessory protein as well as accessory proteins with a His-rich region, responsible for nickel storage and delivery into the buried cavity of these enzymes in a process that culminates with the activation of the target enzyme [1,2].

Ureases are one of the most studied nickel-dependent enzymes whose catalytic activity consists of hydrolyzing urea to form ammonia and carbamate, which then spontaneously decompose to yield carbonic acid and another molecule of ammonia [3]. These enzymes are widespread in nature, being present in most organisms such as bacteria, fungi and plants. In many microorganisms, this enzyme is related to the utilization of urea as a nitrogen source [1]. In other cases, the enzyme is a virulence factor that contributes to several human and others animals diseases, as the gastroduodenal infection by *Helicobacter pylori*, and the urinary stone formation by many bacteria, including *Proteus mirabilis* [1]. In higher organisms as plants, ureases are involved in the metabolism of urea arising from different metabolic pathways [4]. However, the presence of high amounts of specific isoforms of these enzymes where urea is not a major metabolite, rises the possibility that this protein may play other important role in these organisms [5]. In this panorama, several biological properties of ureases unrelated to the enzymatic activity have already been reported, including insecticidal [6], fungicidal [7] and exocytosis-inducing properties [8].

Biochemically, bacterial ureases are better characterized than the ones from any other organism. The same is true for their mechanisms of activation, especially in the case of the enterobacterium *Klebsiella aerogenes*. The complex of proteins required to reach maximum urease activation involves the participation of four accessory proteins conjugated to the apourease (UreD, UreE, UreF and UreG), and a GTP-dependent transfer of a CO<sub>2</sub> molecule necessary for lysine carbamylation [9]. Two models were proposed for the participation of the four accessory proteins in the

process. The first suggests a stepwise process where the accessory proteins altogether with apourease associate to form the following complexes: UreD-apourease [10], UreDF-apourease [11], and UreDFG-apourease [12]. The second model proposes that the UreDFG complex assembles itself and then binds to apourease [13]. In both propositions UreE, required for maximal activation but not essential to the process, is thought to bind the UreDFG-urease complex facilitating the activation of the enzyme [12,14,15].

Compared to bacteria, relatively little is known in plants about the process and the factors involved in nickel incorporation into apourease. The plant urease system has been best studied in soybean, where two structural genes codifying two different ureases have been identified [16,17] as well as two accessory proteins that are essential for both ureases activation, Eu2 and Eu3 [18], the later identified as the equivalent to the bacterial UreG [19]. According to genome information, soybean presents two distinct genes for UreD, two for UreF and only one for UreG [<http://www.phytozome.net/>]. However, the exact role played by each of these duplicated genes is still to be determined. Other plants, like *Arabidopsis thaliana* [20] and potato (*Solanum tuberosum*) [21], have also been studied in the search for urease accessory proteins. In plants, no UreE was ever identified and the two properties, Ni<sup>2+</sup> storage and nucleotide hydrolysis, are possibly displayed by UreG. In these organisms, UreG presents a polyhistidine stretch in its N-terminal that could fulfill UreE's function. This characteristic is also found in some HypB proteins, which are accessory proteins of the hydrogenase maturation system [22].

UreG has been described as an essential chaperone for urease activation, binding and hydrolyzing GTP, which provides the necessary energy for the process, and hypothetically it is also the main player in the carbamylation of the lysine in the urease active site [12]. UreG from a few microorganisms - *Mycobacterium tuberculosis*, *K. aerogenes*, *Bacillus pasteurii* and *Helicobacter pylori* - have been extensively studied, and several differences in what concern the biochemical and structural characteristics of these proteins, as well as its interaction with its partner UreE, were observed [23]. All the UreG proteins described so far bind GTP, and their GTPase activities have been measured and found to be very low or undetectable, suggesting the necessity of co-factors (protein or not) to exert their full activity [2,13,15,24]. The few structural studies of this protein family by NMR have shown they do not possess a rigid tertiary structure, and as such are assigned to the class of intrinsically unstructured proteins [2,15,24]. These proteins bind Ni<sup>+2</sup> and Zn<sup>2+</sup>, but the

stoichiometry of binding varies from one protein to the other, and the precise role exerted by each metal for the assembly of urease has not been assessed.

Despite of all work done so far, no member of the plant urease accessory proteins have been biochemically characterized. Here we presented the first characterization of the native and recombinant *Glycine max* UreG (*GmUreG*), including its GTPase, metal-binding and structural properties.

## **EXPERIMENTAL**

### **Purification of *GmUreG* from the seeds**

*GmUreG* was purified according to a protocol adapted from [19]. The full experimental details of protein purification are provided in the Supplementary Online Data. The purity of the sample was assessed by SDS-PAGE.

### **Cloning, expression and purification of recombinant full length *GmUreG***

Full length *GmUreG* cDNA was amplified by PCR using gene specific primers (forward 5'-GCGGCCCATGGCTTCTGAAGGCGACCATCACC-3'; reverse 5'-CGCCGGATCCTTAATGACGCTTATTCCCTGTGGCT-3'), designed according to the sequence available at the Phytozome website. Details on the cloning and expression of the recombinant *GmUreG* are given in the Supplementary Online Data. Cells were harvested by centrifugation (5.000 x *g* for 10 min), lysed by sonication and centrifuged (12.000 x *g* for 10 min) to separate cell debris from the supernatant. The cleared supernatant was applied into a Chelating-Sepharose resin, previously loaded with Ni<sup>+2</sup> and equilibrated with buffer A containing 15 mM imidazole. Non-retained proteins were washed from the column using buffer A containing 50 mM imidazole. Proteins bound to the column were eluted with the same buffer containing 300 mM imidazole. Eluted proteins were pooled and further purified by gel filtration on a Superdex-75 column equilibrated with 20 mM Tris-HCl pH 7.5, 75 mM NaCl. Protein concentration was measured as described above and purified samples were analyzed by SDS-PAGE.

## Mass spectrometry analysis

MS/MS analyses of trypsin-digested gel bands of *GmUreG* were performed in a Q-TOF Micro™ mass spectrometer (Micromass, Manchester, UK) coupled to a nanoflow capillary liquid chromatography system (UPLC, Waters, Milford, US). The experimental details of the data acquisition are provided in the Supplementary Online Data.

## Circular dichroism

Far-UV CD spectra (190–260 nm) were recorded at 25 °C in a Jasco J715 instrument (Jasco Corporation, Japan) flushed with N<sub>2</sub>, using quartz cells of 1 mm path length, with 10 μM *GmUreG* in 20 mM HEPES at pH 7.5, 150 mM NaCl, either in the absence or in the presence of Zn<sup>2+</sup> and Ni<sup>2+</sup>. All spectra were recorded after accumulation of 16 scans. Quantitative prediction of the secondary structure was performed by deconvolution of the CD spectra using the DICHROWEB [25,26].

## MNR analysis

Uniformly <sup>15</sup>N-labeled *GmUreG* was produced with the same expression and purification procedure described for the unlabelled proteins, using M9 minimal medium (6 g.L<sup>-1</sup> Na<sub>2</sub>HPO<sub>4</sub>, 3 g.L<sup>-1</sup> KH<sub>2</sub>PO<sub>4</sub>, 0.5 g.L<sup>-1</sup> NaCl, 0.12 g.L<sup>-1</sup> MgSO<sub>4</sub>, 4 g.L<sup>-1</sup> Glucose, at pH 7.5) containing <sup>15</sup>NH<sub>4</sub>Cl as sole nitrogen source. NMR spectra of *GmUreG* were recorded for a 0.7 mM <sup>15</sup>N-enriched sample in 20 mM sodium phosphate buffer, 50 mM NaCl, at pH 7.0 and 298 K using a Bruker Avance 900 spectrometer, operating at 899.07 MHz and equipped with a TXI 5-mm cryoprobe with pulsed-field gradients along the z-axis. <sup>1</sup>H, <sup>15</sup>N–HSQC spectra were acquired in phase sensitive mode using Echo/Antiecho-TPPI gradient selection and flip-back pulse for water signal suppression. The spectra consisted of 8 scans, spectral windows of 16/40 ppm in the proton/nitrogen dimensions and with the carrier set at the water frequency and 118 ppm, respectively. The relaxation delay was 1 s, plus 285 ms acquisition time. A matrix of 2048 x 256 points was acquired in F2 and F1 dimensions, respectively, and transformed into 2048 x 1024 points. The data were processed using iNMR (<http://www.inmr.net/>).

## **Inductively coupled plasma emission spectroscopy (ICP-OES)**

Protein metal content analyses were performed using a Spectro Ciros CCD ICP optical emission spectrometer (Spectro Analytical Instruments GmbH, Kleve, Germany) in combination with a Lichte nebulizer and a peristaltic pump for sample introduction, following the protocol previously described in [27].

## **Calorimetric measurements**

Titration experiments were performed at 25 °C using a high-sensitivity VP-ITC microcalorimeter (MicroCal LLC, Northampton, MA). The proteins and the metal ions (from 100 mM stock solutions) were diluted using the buffer (20 mM Tris-HCl, pH 7.5, 75 mM NaCl). The reference cell was filled with deionized water. Before starting each experiment, the baseline stability was checked. A delay of 300 sec between injections was applied in order to allow the system to reach thermal equilibrium after each addition. For each titration, a control experiment was carried out by titrating the metal solution into the buffer alone, under identical conditions. Heats of dilution were negligible. The sample cell contained *GmUreG* (15 μM), which was titrated with 60 injections (5 μL each) of a solution containing 300 μM ZnSO<sub>4</sub> or 600 μM NiSO<sub>4</sub>. The details of data analysis are given in the Supplementary Online Data. The dissociation constants and thermodynamic parameters provided in the article do not take into account possible events of proton transfer linked to metal binding, or the presence, in solution, of complexes between the metal ions and the buffer.

## **GTPase activity**

GTP (2 mM) and *GmUreG* (5 μM) were incubated in 20 mM Tris-HCl pH 8.0, 75 mM NaCl at 37 °C for up to 6 h in the presence of 5 mM MgCl<sub>2</sub>. The phosphate released from hydrolysis of GTP was measured using the Fiske-Subbarow method [28]. To test the effect of Zn<sup>2+</sup> and Ni<sup>2+</sup> in the GTPase activity, 125 μM of each metal was added to the reaction.

## **Molecular Modeling**

An initial alignment of *Methanococcus jannaschii* HypB (*MjHypB*) and *GmUreG* sequences was produced using Modeller v9.8, [29] and this alignment was manually adjusted in order to match up the primary and the secondary structure of the proteins,



as determined using the JPRED server [30]. The Cys and His residues involved in the binding of  $Zn^{2+}$  in the structure of *MjHypB* were aligned with the corresponding residues in the fully conserved Cys-Pro-His motif found in all UreG proteins [24]. This alignment was then used to calculate 100 structural models of the dimer of *GmUreG* (which included two  $Mg^{2+}$  ions, the hydrogen atoms, and the  $GTP\gamma S$  molecule) using the program Modeller and the *MjHypB* dimer structure (PDB code 2HF8) as template, imposing a structural identity for the two monomers. The best models were selected on the basis of the lowest value of the DOPE score in Modeller [31] and were subjected to a refining step of loop optimization using Modeller standard routines. Further details on the *GmUreG* model are provided in the Supplementary Online Data.

## RESULTS AND DISCUSSION

### Production of *GmUreG*

The long His-rich N-terminal of *GmUreG* (among the first 58 residues 23 are histidines, 19 of which are located between residues 7 – 30) (Figure 1A), allows the use of a nickel affinity column to purify this protein from the soybean extract in a single step. Due to its high histidine content, *GmUreG* is not significantly eluted from the resin until the imidazole concentration is raised to 200 mM. As previously described [19], the native protein displays fractions with different degrees of affinity to the Ni-column. A major fraction of *GmUreG* elutes with 200 mM imidazole, however a smaller fraction could only be eluted when the imidazole concentration is raised to 300 mM (Supplementary material, Figure S1A). These fractions were named *GmUreG*200 and *GmUreG*300. In order to verify the identity of these two fractions, we have used MS/MS tools. ESI Q-ToF tandem mass spectrometry positively identified both fractions as *GmUreG* by the sequencing of several internal fragments, covering 62 and 27 % of the *GmUreG* predicted sequence, for *GmUreG*200 and *GmUreG*300, respectively (Supplementary material, Table S1).

Due to the large amounts of protein required for the biochemical and structural characterization studies, expressing *GmUreG* in a heterologous system became necessary. Recombinant wild type *GmUreG* (*rGmUreG*) was extracted from *E. coli* cells, and most of the overproduced protein accumulated in the soluble fraction, as described for *MtUreG* [24], *HpUreG* [2] and *KaUreG* [13,23], and in contrast to *BpUreG* [15], which was isolated from inclusion bodies. *rGmureG* was purified in a two step protocol, consisting of a Ni-affinity chromatography followed by a size exclusion

chromatography (SEC) using high resolution FPLC Superdex-75 column (Supplementary material, Figure S1B). The protein partially aggregates after elution from the Ni-affinity column and should be promptly submitted to SEC to remove these aggregates to avoid precipitation. Removal of the aggregates stabilized the protein for the subsequent analysis. Identity of r*GmUreG* was confirmed by tryptic digestion of the protein band recovered from a SDS-PAGE, followed by ESI Q-ToF tandem mass spectrometry (Supplementary material, Table S1).

### ***GmUreG* structural analysis**

A prediction of the secondary structure elements of *GmUreG* (Figure 1A) based on JPRED algorithm indicates 34%  $\alpha$ -helix and 15%  $\beta$ -strands with the remaining 51% constituting turns and random coil, including the first 44 residues (composing most of the His-rich N-terminal of the protein). Also, this prediction suggests that *GmUreG* has a fold similar to several other *UreG* proteins, as previously reported on basis of modeling studies [2,15,24].

We assessed the secondary structure content of *GmUreG*200 and *GmUreG*300 by far-UV CD spectroscopy. Interestingly, despite the fact of containing the same protein (as confirmed by MS/MS analyses), the two fractions showed different degrees of secondary structure (Figure 2A), suggesting that some factor interferes with the proteins folding *in vivo* and may also be the responsible for their differential behavior on the Ni-affinity column. The *GmUreG*300 CD spectrum showed the presence of both  $\alpha$ -helices and  $\beta$ -strands, with negative deflections around 208 nm and 221 nm and a positive peak at 190 nm, while the *GmUreG*200 spectrum showed only one deflection at 203 nm, indicating a mostly unfolded protein. The data were quantitatively analyzed in the range of 190 – 240 nm, using all the different fitting programs available at the Dichroweb server [25] and all possible reference sets. The best fit was selected on the basis of the normalized root-mean-square deviation (nrmsd) between the experimental and calculated data. For *GmUreG*300 the estimated secondary structure composition was 26 %  $\alpha$ -helix, 18 %  $\beta$ -strands, 21 % turns and 35 % random coil (nrmsd = 0.058; Contin program and reference set 4), while the estimated content for *GmUreG*200 was 16 %  $\alpha$ -helix, 25 %  $\beta$ -strands, 15 % turns and 44 % random coil (nrmsd = 0.074; Contin program and reference set SP175). Despite the reasonable degree of secondary structure estimated for *GmUreG*200, its spectrum is typical of a mostly unfolded structure. Problems in the accurate prediction of secondary structure content for proteins with high amount of disordered structure have been previously noted [26]

since this kind of structure is not well represented on the available data bases. Based on that, we believe that the amount of unfolded polypeptide in *GmUreG200* is much higher than what is expressed in this estimative. Considering this scenario, it is very intriguing that *GmUreG200*, the most abundant *GmUreG* form found in the plant extract, presents a higher degree of random coil content than *GmUreG300*. This may suggest a relevant physiological role for this particular behavior. On the other hand, *GmUreG300* presents a secondary structure content closer to the one predicted by the JPRED algorithm.

*rGmUreG* was also analyzed by Far-UV CD spectroscopy (Figure 2B) and showed a behavior very similar to *GmUreG300*. The best fit estimated a secondary structure composition of 29 %  $\alpha$ -helix, 16 %  $\beta$ -strands, 55 % turns and random coil (nrmsd = 0.058; K2d method). In order to verify if the metal binding property of *UreG* proteins could affect their secondary structure content, *rGmUreG* was submitted to CD analysis in the presence of  $Zn^{2+}$  and  $Ni^{2+}$ . As shown in Figure 2B, the incubation of the protein with ten molar equivalents of  $Ni^{2+}$  had very little influence on the overall structure of the protein, causing only a small change in the secondary structure content when quantitatively analyzed: 36 %  $\alpha$ -helix, 17 %  $\beta$ -strands, 22 % turns and 25 % random coil (nrmsd = 0.02; CDSSTR program and reference set 4). However, when the same amount of  $Zn^{2+}$  was added to the protein solution, a decrease in the overall intensity of the spectrum, mostly in the region of 208 nm, is observed. Such spectral changes correspond to a substantial reduction in the  $\alpha$ -helix content and an increase in the  $\beta$ -strands. The estimated content for the  $Zn^{2+}$  bound protein was 7 %  $\alpha$ -helix, 25 %  $\beta$ -strands, 32 % turns and 36 % random coil (nrmsd = 0.027; CDSSTR program and reference 4). This same effect was reported for *HpUreG*, for which  $Zn^{2+}$  binding caused a similar change in the secondary structure of the protein [2].

Despite the well defined secondary structure of *rGmUreG* according to the CD spectroscopy results, this technique provides no information about its tertiary structure. NMR spectroscopy was performed on the  $^{15}N$ -*rGmUreG* (Figure 3). The  $^1H$ ,  $^{15}N$ -HSQC spectra of *rGmUreG* is characterized by poorly resolved resonances with little dispersion in the  $^1H$  dimension of the backbone amides, many of them falling in the random coil region (7.6–8.5 ppm chemical shift range). This observation suggests that large portions of the protein backbone undergo multiple conformations and lack a well defined tertiary structure. The signals of the side chain  $NH_2$  groups of the Asn and Gln residues present in the protein are also poorly resolved and give rise to two broad envelopes, centered at 112.5 ppm ( $^{15}N$ ) and 7.4–6.7 ppm ( $^1H$ ). Lack of differentiation in

these resonances suggests lack of specific interactions for the Asn and Gln side chains and therefore confirms the fluxional behavior of *GmUreG* in solution.

The NMR results presented here seem to confirm previously published data [2,15,24] regarding the high flexibility of the UreG chaperone family, showing that this feature is also shared by the plant members of the family. The presence of large portions of the protein backbone subjected to conformational changes indicates that *GmUreG* also belongs to the class of intrinsically unstructured proteins, that undergo disorder to order transitions (among random coil, pre-molten globule, molten globule or fully folded states) during or prior to their biological function [32-34]. This characteristic could be related to the functional role of UreG *in vivo*, as it would allow minimization of unwanted hydrolysis of GTP unless the protein is ready to perform its role together with its partner chaperones in a complex formed by UreDFG-apouase.

In order to assess the quaternary arrangement of *GmUreG* in solution, both SEC and light scattering were employed. The elution profile of r*GmUreG* in SEC shows a broad peak with a retention volume compatible with the protein behaving as an intermediate between monomer and dimer (Figure 4), which was further confirmed by light scattering data (MALS/QELS) (Supplementary material, Figure S2), suggestive of a fast equilibrium between the two states. This behavior is not altered by dithiothreitol (data not shown), showing it is not a dimer stabilized by disulfide bonds, in contrast to what is known for *MtUreG* [24] and *BpUreG* [35]. The SEC profile in the presence of  $Zn^{2+}$  shows a narrower peak with a reduced retention time, with an estimated molecular mass of 64 kDa, corresponding to the protein dimer (Figure 4, dotted line). The appearance of a small shoulder with increased retention volume suggests the presence of trace amounts of the monomeric protein. The data show that *GmUreG* dimerization is not caused by  $Zn^{2+}$ , but rather stabilized in the presence of the metal, contrasting to what was previously described for *HpUreG* [2]. The variation in the secondary structure composition observed in the presence of  $Zn^{2+}$  (Figure 2B) indicates that the structural effect of metal binding produces a conformational rearrangement likely responsible for the stabilization of the protein dimeric form. This is supported by the  $Zn^{2+}$ -induced change in the secondary structure of *GmUreG* when the  $\alpha$ -helix content is reduced from 26 to 7 % and the  $\beta$ -strand is increased from 17 to 25 %, apparently yielding a conformational state similar to the apodimers *BpUreG* and *MtUreG* [15,24] or the  $Zn^{2+}$ -induced dimer of *HpUreG* [2]. The effect of  $Ni^{2+}$  ions on oligomerization state of *GmUreG* could not be evaluated since at the protein concentration needed for the experiment, it readily precipitates upon addition of this metal.

## Metal binding properties

Previous reports have shown that bacterial UreG are capable of binding both  $Zn^{2+}$  and  $Ni^{2+}$  [2,15,23,24], a feature also observed for UreE [14,36]. Since *GmUreG* seems to accumulate both proteins roles, its metal binding capacity was assessed.

Prior to the metal binding assays, the metal content of each sample was assessed by ICP-OES. Both samples purified from the plant seem to be loaded with  $Mg^{2+}$ , which is related to the GTPase property observed for many UreG proteins. On the other hand, *GmUreG200* also carries high amounts of  $Ca^{2+}$  (ca. 10 equivalents), which was not observed for any other sample. We believe that the  $Ca^{2+}$  presence may be the reason for the lower affinity of this fraction to the  $Ni^{2+}$ -loaded column. *GmUreG200* and *rGmUreG* carry only residual concentrations of other metals, including  $Ni^{2+}$  and  $Zn^{2+}$ , with *GmUreG300* containing ca. 0.3 equivalents of  $Zn^{2+}$  (Supplementary material, Table S2). The purification of this  $Zn^{2+}$ -loaded protein fraction from the plant supports the hypothesis that  $Zn^{2+}$  may be a physiologically relevant ligand. To better understand the metal binding properties of *GmUreG*, ITC measures were carried out. In the case of  $Zn^{2+}$ , the integrated heat data (Figure 5B and C) clearly shows three different types of binding events, completed when 1, 2.5 and 4 equivalents of  $Zn^{2+}$  were added to the protein. The calculated fits indicated that initially 0.5 equivalents of  $Zn^{2+}$  bind each *GmUreG* monomer very tightly with a  $K_{d1}(Zn)$  of  $0.0217 \pm 0.010 \mu M$ . This process is followed by a second binding event of one  $Zn^{2+}$  ion per protein monomer with a lower affinity ( $K_{d2}(Zn) = 1.31 \pm 0.05 \mu M$ ). Finally, ca. 2.6  $Zn^{2+}$  ions bind each monomer with  $K_{d3} = 0.163 \pm 0.09 \mu M$ . In the case of  $Ni^{2+}$  titration, a single binding event was observed, with a stoichiometry of 2.15 equivalents per protein monomer and a  $K_d = 4.8 \pm 0.3 \mu M$  (Figure 5E). These processes are driven by a large favorable enthalpic factor that compensates the negative entropic value (Table 1). Since *GmUreG* presents such a large content of histidines in its N-terminal, it is not possible to say with certainty which events are related to metal binding to the protein core (G-domain) or to the polyhistidine stretch; however, it is clear that the protein binds  $Zn^{2+}$  and  $Ni^{2+}$  differently and some inferences can be made. The first binding event shown for  $Zn^{2+}$  has the highest affinity described so far for an UreG protein and implicates that  $Zn^{2+}$  could possibly play a functional and/or structural role. The half-integer stoichiometry of  $Zn^{2+}$  bound per monomer reflects the  $Zn^{2+}$ -stabilized dimerization, since one  $Zn^{2+}$  ion is shared by two monomers, indicating that this  $Zn^{2+}$  binding site is located at the interface between the two monomers. Based on previous publications [2] we infer that this site corresponds to the CPH motif (residues 148-150), well conserved among UreG proteins and shown to bind  $Zn^{2+}$  [2,23]. Interestingly, in

the case of *GmUreG* this binding site seems to be specific for  $Zn^{2+}$ , contrasting to bacterial UreGs, which bind both  $Zn^{2+}$  and  $Ni^{2+}$ . It is very likely that the third binding event, with a stoichiometry of 2.6  $Zn^{2+}$  ions bound per monomer, reflects the interaction with the His-stretch. The lower  $K_d$  observed for this event, when compared to the second one, could be an overestimation due to the quality of the data fit. Nevertheless, these data indicate that additional metal ions bind to the protein after filling the first two metal sites. The single binding event observed for  $Ni^{2+}$  is also proposed to be related to the His-rich portion. These data seem to agree with *GmUreG* also acting as a  $Ni^{2+}$  storage protein. Although further studies are necessary to fully understand *GmUreG* metal binding properties, if *GmUreG* binds  $Ni^{2+}$  only in its N-terminal portion, this is probably not directly involved in the deliver of  $Ni^{2+}$  into urease metallocenter, a role that should be played by another accessory protein. Under limited concentrations of  $Ni^{2+}$  in the cytoplasm, the highest affinity should be that of the protein that actually delivers the  $Ni^{2+}$  to urease while the storage protein (*GmUreG*) would be loaded on weaker sites, with  $K_d$  values in the micromolar range. This feature is also observed for some HypB proteins that have similar N-terminal His-extension as *Rhizobium leguminosarum* HypB (Figure 1B), which binds a maximum of ca. 4  $Ni^{2+}$  per monomer with affinity in the micromolar range [37]. Whether the second binding event of  $Zn^{2+}$  occurs in the G-domain or not, and what could be its function, is a matter for further studies. The role of the high amounts of  $Ca^{2+}$  co-purified with *GmUreG200* also demands additional work to be elucidated. Some GTPases have their structure and/or activity affected by calcium. The spinach PsbO protein, part of the Photosystem II (PSII) complex and also an intrinsically unstructured protein with GTPase activity [38], undergoes structural changes [39] and is proposed to have its interaction with partner proteins regulated by  $Ca^{2+}$  binding [40]. *Escherichia coli* FtsZ, a key GTPase for bacterial cell division, also has its secondary structure and enzymatic activity affected by  $Ca^{2+}$  binding [41].

### **GTPase activity**

UreG is involved in the hydrolysis of GTP concomitantly with the carbamylation of a lysine residue in the urease active site [12]. The P-loop nucleotide binding motif is conserved among all members of this family. The main difference between plant and bacterial UreGs in this motif seems to be a conservative serine to threonine change, observed in *GmUreG* (Figure 1A) as well as in *A. thaliana* and *S. leguminosarum* UreGs [20,21], and also in other small G-proteins [42]. GTPase activity has been demonstrated for *BpUreG* and *MtUreG* [15,24] while *KaUreG* and *HpUreG* have shown

negligible GTPase activity [2,13]. *GmUreG* activity was assessed in an assay in which the GTP concentration was 3 orders of magnitude higher than the enzyme concentration, allowing the use of time course data to calculate the  $k_{cat}$  values. *GmUreG200*, *GmUreG300* and *rGmUreG* presented  $k_{cat}$  values of 0.017, 0.018 and 0.013  $\text{min}^{-1}$ , respectively (figure 6A). These values are comparable to those found for *MtUreG* ( $k_{cat}$  0.01  $\text{min}^{-1}$ ; [24]) and *BpUreG* ( $k_{cat}$  0.04  $\text{min}^{-1}$ ; [15]). This level of activity is about 10 times lower than other similar GTPases such as HypB from *B. japonicum* ( $k_{cat}$  0.18  $\text{min}^{-1}$ ) [43] and *E. coli* ( $k_{cat}$  0.17  $\text{min}^{-1}$ ) [44]. The conformational changes carried out by this protein might explain its lower activity, considering that in the absence of its co-factors the enzymatically active folding can not be stabilized, and only a small fraction of the molecules would be in an enzymatic favorable form at a given time. As discussed previously, this is probably beneficial since it minimizes unnecessary GTP hydrolysis before the urease activating complex is correctly assembled.

The effect of  $\text{Zn}^{2+}$  and  $\text{Ni}^{2+}$  on the GTPase activity was tested in conditions of molar excess of metals (25 molar equivalents), during the course of 6 h.  $\text{Ni}^{2+}$  did not show a significant effect on the activity of either the native or recombinant proteins (Figure 6B, C, D).  $\text{Zn}^{2+}$ , in the other hand, increased 89 and 115% the activity of *GmUreG300* and *rGmUreG*, respectively (Figures 6B, C and D), while the effect on *GmUreG200* was negligible. As demonstrated by the ICP measurements, *GmUreG300* purified from the plant carries  $\text{Zn}^{2+}$ , however, in amounts not enough to ensure full GTPase activity. When saturating amounts of  $\text{Zn}^{2+}$  are added to the reaction, both *GmUreG* and *rGmUreG* reach similar levels of activity. These results could be explained in terms of the level of organization of the protein in the presence of the metals. Despite the fact that  $\text{Zn}^{2+}$  reduces the amount of secondary structure in the protein, it seems to stabilize the dimeric form of *GmUreG*, thus reducing its flexibility, making it easier for the protein to acquire its active folding. However, the still very low activity displayed by this protein indicates that  $\text{Zn}^{2+}$  alone is not enough to stabilize completely the folding of *GmUreG*. The molecular properties so far described for UreG proteins imply that several factors may be involved in the regulation of their GTPase activity, such as structural transitions, dimerization, metal binding, and multiprotein complex formation [24]. The lack of effect observed for  $\text{Zn}^{2+}$  on GTPase activity of *GmUreG200* is intriguing and a possible explanation could be that the presence of  $\text{Ca}^{2+}$  somehow interferes with  $\text{Zn}^{2+}$  interaction. It is worth to mention that no change in GTPase activity was observed for *BpUreG* when  $\text{Zn}^{2+}$  was added to the reaction mixture, but the protein was already in its dimeric form and presented slightly higher activity than *GmUreG*, even after  $\text{Zn}^{2+}$  addition ( $\text{Zn}^{2+}$ -*rGmUreG*  $k_{cat}$  = 0.028  $\text{min}^{-1}$ ;  $\text{Zn}^{2+}$ -

*GmUreG*300  $k_{cat} = 0.034 \text{ min}^{-1}$ ; *BpUreG*  $k_{cat} = 0.04 \text{ min}^{-1}$ ) [15]. In the case of  $\text{Ni}^{2+}$ , the secondary structure of the protein is a little more organized in the presence of this metal, and considering the ITC data, this probably reflects some organization of the His-rich N-terminal. Although it was not possible to evaluate the  $\text{Ni}^{2+}$  effect on the quaternary arrangement of *GmUreG*, the ITC data suggests that the protein does not share any  $\text{Ni}^{2+}$  ions, leading us to think that the metal-stabilized dimerization of *GmUreG* does not happen in presence of  $\text{Ni}^{2+}$ . This assumption is supported by previously published data showing that there are no alterations in the quaternary structures of *KaUreG* and *HpUreG* in the presence of  $\text{Ni}^{2+}$  [2,23]. Altogether, these data support a role for  $\text{Zn}^{2+}$  as one of the co-factors involved in the stabilization of the correct folding of *GmUreG* during urease activation.

### ***GmUreG* structural modeling**

The alignment of *GmUreG* and *MjHypB*, resulting from an analysis of the sequences and the secondary structure of the two proteins (Figure 1A), showed a good match of regions known to be functionally active in *MjHypB*, such as the GTP-binding P-loop GAIGSGK at the N-terminal portion of the protein (corresponding to GPVGTGK in *GmUreG*), the conserved NKTD motif involved in the guanine-specific recognition of the aromatic portion of GTP, the residues Thr, Asp, and Glu that bind GTP complexed to  $\text{Mg}^{2+}$ , and the switch-I (DVIAKFD) and switch-II (ENVGNLI) motifs (DIFTKED and ESGGDNL in *GmUreG*, respectively), thought to be important for the essential conformational changes occurring upon GTP-binding.

The folded state model of *GmUreG*, calculated using this alignment, shows a globular protein consisting of a seven-stranded parallel  $\beta$ -sheet flanked by ten  $\alpha$ -helices (Figure 7A). As in the *HpUreG* model [2], the structure largely reproduces the template *MjHypB*, not only in the overall architecture, but also in the arrangement of key functional features, such as the P-loop (or Walker-A) and the switch-II (or Walker-B) regions. The amphipathic helix found at the N-terminus of *MjHypB*, which features a conserved pattern of positively charged residues and was therefore proposed to be important for protein-protein interactions [45] is present in the sequence of *GmUreG*, while is absent in bacterial *UreGs* [2,24]. The His-rich region of *GmUreG*, comprising the first N-terminal 62 residues, was not modeled due to the absence of available structural information, as well as the last 5 residues at the C-terminal. Based on the model structure of *GmUreG* (Figure 7A), and on the differences in residue-by-residue solvent accessibility between the monomer and the dimer, the interface between the



two monomers of *GmUreG* appears to be composed by seven loops and three helices. These regions include the P-loop, the switch-I and switch-II, and the guanine-specific pocket.

The distribution of the electrostatic potentials on the solvent-accessible surface shows that *GmUreG* contains an extended area of negative charge, located mainly around the central zone of the monomer-monomer interface (Figure 7B, top panel), while a less charged area is found on the opposite side (Figure 7B, middle panel) and on the N-terminal helix. Positively charged regions at the protein surface correspond to solvent exposed Lys and Arg residues, with the exception of Lys227, found in the inner surface of the conserved NKTD motif constituting the guanine binding pocket (see Figure 7B, bottom panel).

## **Conclusions**

We presented here the first detailed characterization of a plant urease accessory protein, *GmUreG*. We have demonstrated its GTPase activity, its capacity of binding  $Zn^{2+}$ , and have provided more evidences that support the previous proposed hypothesis that  $Zn^{2+}$  has a structural and functional role in UreG proteins. The data showed here also indicates that plant UreGs share several properties with the bacterial members of the family, such as their intrinsically unfolded structure. All the characteristics exploited here are essential features of UreG proteins, and can possibly affect the role played by these proteins in urease activation.

## **ACKNOWLEDGMENTS**

The authors are thankful to the Brazilian agencies CNPQ (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior) e FAPERGS (Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul) for grants and fellowships. We acknowledge Dr. Joseph Polacco, Dept. Biochemistry, University of Missouri at Columbia, USA, for insightful discussions on topics of this work.

## REFERENCES

- 1 Molrooney, S. B. and Hausinger, R. P. (2003) Nickel uptake and utilization by microorganisms. *FEMS Microbiol. Rev.* 27, 239-261
- 2 Zambelli, B., Turano, P., Musiani, F., Neyroz, P. and Ciurli, S. (2008) Zn<sup>2+</sup>-linked dimerization of ureG from *Helicobacter pylori*, a chaperone involved in nickel trafficking and urease activation. *Proteins* 74, 222-239
- 3 Krajewska, B. (2009) Ureases I. Functional, catalytic and Kinetic properties: A review. *J. Mol. Catal. B: Enzym.* 59, 9-21
- 4 Polacco, J. C. and Holland, M. A. (1993) Roles of urease in plant cells. *Int. Rev. Cytol.* 145, 65–103
- 5 Carlini, C. R. and Polacco, J. C. (2008) Toxic properties of urease. *Crop Science* 48, 1665-1672
- 6 Follmer, C., Real-Guerra, R., Wasserman, G. E., Olivera-Severo, D. and Carlini, C. R. (2004). Jack bean, soybean and *Bacillus pasteurii* ureases: biological effects unrelated to ureolytic activity. *Eur. J. Biochem.* 271, 1357– 1363
- 7 Becker-Ritt, A. B., Martinelli, A. S., Mitidieri, S., Feder, V., Wassermann, G. E., Santi, L., Vaistein, M. H., Oliveira, J. A., Fiuza, L. M., Pasquali, G. and Carlini, C. R. (2007) Antifungal activity of plant and bacterial ureases. *Toxicon* 50, 971-983
- 8 Wassermann, G. E., Olivera-Severo, D., Uberti, A. F. and Carlini, C. R. (2010) *Helicobacter pylori* urease activates blood platelets through a lipoyxygenase-mediated pathway. *J. Cell. Mol. Med.* 14, 2025 – 2034.
- 9 Soriano, A., Colpas, G. L. and Hausinger, R. P. (2000) UreE stimulation of GTP-dependent urease activation in the UreD-UreF-UreG-urease complex. *Biochemistry* 39, 12435-12440
- 10 Park, I. S., Carr, M. B. and Hausinger, R. P. *in-vitro* activation of urease apoprotein and role of urea as a chaperone required for nickel metallocenter assembly. *Proc. Nacional. Acad. Sci. EUA* 91, 3233-3237, 1994.
- 11 Moncrief, M. B. and Hausinger, R. P. (1996) Purification and activation properties of ureD-ureF-urease apoprotein complex. *J. Bacteriol.* 178, 5417-5421
- 12 Park, I. S. and Hausinger, R. P. (1995) Evidence for the presence of urease apoprotein complex containing UreD, UreF, and ureG in cells that are competent for in vivo enzyme activation. *J. Bacteriol.* 177, 1947-1951
- 13 Moncrief, M. B. and Hausinger, R. P. (1997) Characterization of UreG, identification of a UreD-ureF-UreG complex and evidence suggesting that a nucleotide-binding site in UreG is required for in vivo metallocenter assembly of *Klebsiella aerogenes* Urease. *J. Bacteriol.* 179, 4081-4086
- 14 Lee, M. H., Pankratz, H. S., Wang, S., Scott, R. A., Finnegan, M. G., Johnson, M. K. Ippolito, J. A., Christianson, D. W. and Hausinger, R. P. (1993) Purification and characterization of *Klebsiella aerogenes* ureE protein: A nickel-binding protein that function in urease metallocenter assembly. *Protein Sci.* 2, 1042-1052
- 15 Zambelli, B., Stola, M., Musiani, F., De Vriendt, K., Samyn, B., Devreese, B., Van Beeumen, J., Turano P., Dikiy A., Bryant D. A., and Ciurli, S. (2005) UreG, a

chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds  $Zn^{2+}$ . J. Biol. Chem. 280, 4684-4695

16 Meyer-Bothling, L. E. and Polacco J. C. (1987) Mutational analysis of the embryo-specific urease locus of soybean. Mol. Gen. Genet. 209, 439-444

17 Torisky, R. S., Griffin, J. D., Yenofsky, R. L. and Polacco, J. C. (1994) A single gene (Eu4) encodes the tissue ubiquitous urease of soybean. Mol. Gen. Genet. 242, 404-414

18 Meyer-Bothling, L. E., Polacco, J. C. and Cianzio, S. R. (1987) Pleiotropic soybean mutants defective in both urease isoenzymes. Mol. Gen. Genet. 209, 432-438

19 Freyermuth, S. K., Bacanamwo, M. and Polacco, J. C. (2000) The soybean Eu3 gene encodes a Ni-binding Protein necessary for urease activity. The Plant J. 21, 53-60

20 Witte, C. P., Rosso, M. G. and Romeis, T. (2005) Identification of three urease accessory proteins that are required for urease activation in *Arabidopsis*. Plant Physiol. 139, 1155-1162

21 Witte, C. P., Isidore, E., Tiller, S. A., Davies, H. V. and Taylor, M. A. (2001) Functional characterization of urease accessory protein G (ureG) from potato. Plant Mol. Biol. 45, 169-179

22 Casalot, L. and Rousset, M. (2001) Maturation of the [NiFe] hydrogenases. Trends Microbiol. 9, 228-237

23 Boer, J., Quiroz-valenzuela, S., Anderson, K. L. and Hausinguer, R. P. (2010) Mutagenesis of *Klebsiella aerogenes* UreG to probe nickel binding and interactions with other urease-related proteins. Biochemistry 49, 5859-5869

24 Zambelli, B., Musiani, F., Savini, M., Tucker, P. and Ciurli, S. (2007) Biochemical studies on *Mycobacterium tuberculosis* UreG and comparative modeling reveal structural and functional conservation among the bacterial UreG family. Biochemistry 46, 3171-3182

25 Whitmore, L. Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Research 32, W668-W673

26 Whitmore, L. Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases Biopolymers 89, 392-400

27 Stola, M., Musiani, F., Mangani, S., Turano, P., Safarov, N., Zambelli, B. and Ciurli, S. (2006) The Nickel Site of *Bacillus pasteurii* UreE, a Urease Metallo-Chaperone, As Revealed by Metal-Binding Studies and X-ray Absorption Spectroscopy. Biochemistry 45, 6495-6509

28 Fiske, C. H., and SubbaRow, Y. (1925) The Colorimetric Determination of Phosphorus J. Biol. Chem. 66, 375-400

29 Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. 29, 291-325

- 30 Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. and Barton, G. J. (1998) Jpred: A consensus secondary structure prediction server. *Bioinformatics* 14, 892-893
- 31 Shen, M.Y. and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507-2524.
- 32 Dunker, A. K., Brown, C. J., Lawson, J.D., Iakoucheva, L. M. and Obradovic, Z. (2002) Intrinsic disorder and protein function. *Biochemistry* 41,6573-6582
- 33 Tompa, P. (2002) Intrinsically unstructured proteins. *Trends Biochem. Sci.* 27, 527-533
- 34 Uversky, V. N. (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* 11, 739-756
- 35 Neyroz, P., Zambelli, B. and Ciurli, S. (2006) Intrinsically disordered structure of *Bacillus pasteurii* UreG as revealed by steady-state and time-resolved fluorescence spectroscopy. *Biochemistry* 45, 8918-8930
- 36 Colpas, G. J. and Hausinger, R. P. (2000) *In Vivo* and *in Vitro* Kinetics of Metal Transfer by the *Klebsiella aerogenes* Urease Nickel Metallochaperone, UreE. *J. Biol. Chem.* 275, 10731-10737
- 37 Rey, L., Imperial, J., Palacios, J. M. and Ruizargueso, T. (1994) Purification of rhizobium-leguminosarum HypB, a nickel-binding protein required for hydrogenase synthesis. *J. Bacteriol.* 176, 6066-6073
- 38 Lundin, B., Thuswaldner, S., Shutova, T., Eshaghi, S., Samuelsson, G., Barber, J., Andersson, B. and Spetea, C. (2007) Subsequent events to GTP binding by the plant PsbO protein: Structural changes, GTP hydrolysis and dissociation from the photosystem II complex. *Biochim Biophys Acta.* 1767, 500-508
- 39 Heredia, P. and De Las Rivas, J. (2003) Calcium-dependent conformational change and thermal stability of the isolated PsbO protein detected by FTIR spectroscopy. *Biochemistry* 42, 11831-11833
- 40 Shutova, T., Nikitina, J., Deikus, G., Anderson, B., Klimov, V. and Samuelsson, G. (2005). Structural dynamics of the manganese-stabilizing protein – Effect of pH, calcium, and manganese. *Biochemistry* 44, 15182-15192
- 41 Jaiswal, R. and Panda, D. (2009) Differential assembly properties of *Escherichia coli* FtsZ and *Mycobacterium tuberculosis* FtsZ: an analysis using divalent calcium. *J. Biochem.* 146, 733-742
- 42 Sprang, S. P. (1997). G proteins mechanisms: Insights from structural analysis. *Annu. Rev. Biochem.* 66, 639-678
- 43 Fu, C. L., Olson, J. W. and Maier, R. J. (1995) HypB Protein of *Bradyrhizobium japonicum* is a metal-binding GTPase capable of binding 18 divalent nickel ions per dimer. *Proc. Nat. Acad. Sci. USA* 92, 2333-2337
- 44 Maier, T., Jacob, A., Sauter, M. and Bock, A. (1993) The product of the HypB gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein. *J. Bacteriol.* 175, 630-635

45 Gasper, R., Scrima, A. and Wittinghofer, A. (2006) Structural insights into HypB, a GTP-binding protein that regulates metal binding. *J. Biol. Chem.* 281, 27492-27502

**Table 1. Thermodynamic parameters of Zn<sup>2+</sup> and Ni<sup>2+</sup> binding to GmUreG.**

Metal ion	K <sub>g1</sub> (μM)	ΔH <sub>1</sub> (kcal.mol <sup>-1</sup> )	ΔS <sub>1</sub> (cal.mol <sup>-1</sup> .K <sup>-1</sup> )	K <sub>g2</sub> (μM)	ΔH <sub>2</sub> (kcal.mol <sup>-1</sup> )	ΔS <sub>2</sub> (cal.mol <sup>-1</sup> .K <sup>-1</sup> )	K <sub>g3</sub> (μM)	ΔH <sub>3</sub> (kcal.mol <sup>-1</sup> )	ΔS <sub>3</sub> (cal.mol <sup>-1</sup> .K <sup>-1</sup> )
Ni <sup>2+</sup>	4.8 ± 0.3	29.59 ± 0.40	-75	-	-	-	-	-	-
Zn <sup>2+</sup>	0.022 ± 0.01	40.59 ± 0.39	-101	1.31 ± 0.05	37.32 ± 0.44	-98.3	0.163 ± 0.09	12.14 ± 0.46	-9.66

The data is presented in terms of cal, where 1 cal ≈ 4.189 J.

## FIGURE LEGENDS

Figure 1: **Sequence analysis of GmUreG.** Alignment of the sequences of *GmUreG*, *HpUreG* and *MjHypB*, with the secondary structure assignments ( $\alpha$ -helix, yellow;  $\beta$ -strand, cyan) derived from JPRED prediction for *GmUreG* and *HpUreG*, and from the PDB structure (code 2HF8) for *MjHypB*; functional portions of the three proteins are indicated (A). Comparison of the N-terminal sequences of *GmUreG*, *BjHypB* and *RjHypB*, showing the His-stretch (B).

Figure 2: **Far-UV Circular Dichroism Spectrum of GmUreG.** The experimental data are shown as dots and the solid lines represent the best fit calculated using the Dichroweb server. Mean residue ellipticity units are expressed in degrees.cm<sup>2</sup>.dmol<sup>-1</sup>.residue<sup>-1</sup>. (A) *GmUreG*200 (open circles) and *GmUreG*300 (filled circles) analysis. (B) apo-r*GmUreG* (open circles), Zn<sup>2+</sup>-r*GmUreG* (open triangles) and Ni<sup>2+</sup>-r*GmUreG* (open squares).

Figure 3: **NMR spectroscopic properties of GmUreG.** <sup>1</sup>H-<sup>15</sup>N-HSQC spectrum of apo-r*GmUreG* (700  $\mu$ M), acquired at 899.07 MHz.

Figure 4: **Size-exclusion chromatography of GmUreG.** r*GmUreG* size exclusion chromatography was performed on a Superdex-75 column, in the presence (dotted line) or absence (solid line) of Zn<sup>2+</sup>. Molecular markers are indicated in kDa.

Figure 5: **Metal binding properties of GmUreG.** Representative plots of raw titration data showing the thermal effect of 60 x 5  $\mu$ L injections of Zn<sup>2+</sup> (300  $\mu$ M) (A) or Ni<sup>2+</sup> (600  $\mu$ M) (D) on to a r*GmUreG* solution (15  $\mu$ M). Titration curves plots for the interaction between r*GmUreG* and Zn<sup>2+</sup> (B and C) or Ni<sup>2+</sup> (E), together with the best fits of the integrated data, represented as solid lines, obtained by a non-linear least squares procedure.

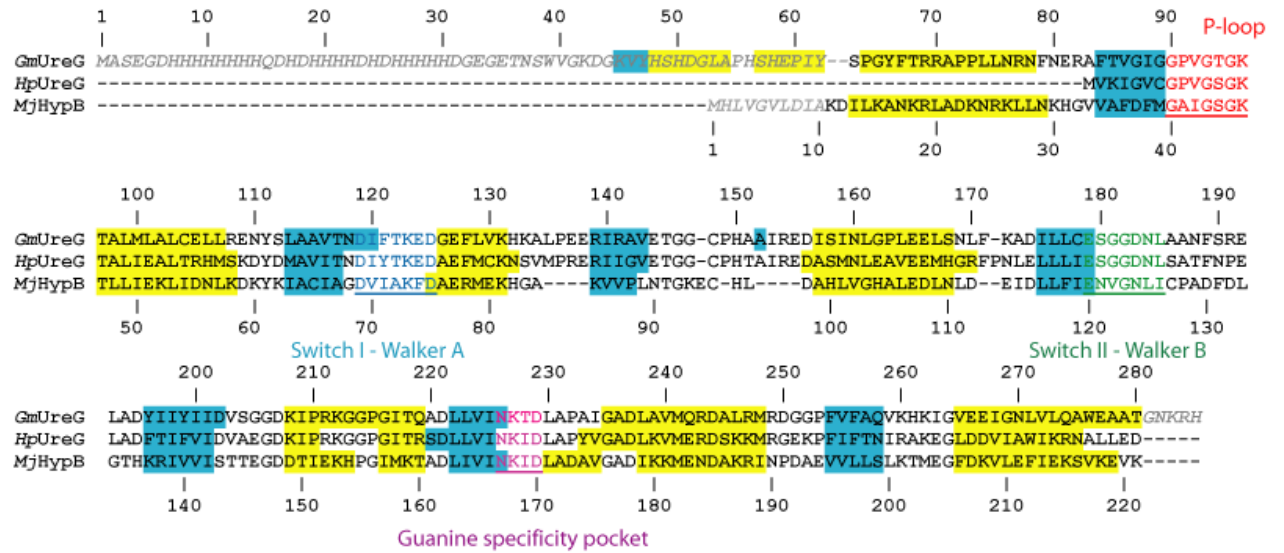
Figure 6: **GTPase activity of GmUreG.** (A) Time course of GTP hydrolysis by GmUreG200 (closed circles), GmUreG300 (open circles) and rGmUreG (triangles). Time course of GTPase activity in the presence of 125  $\mu\text{M}$   $\text{Zn}^{2+}$  or  $\text{Ni}^{2+}$  for (B) GmUreG200, (C) GmUreG300, and (D) rGmUreG.

Figure 7: **Structural modeling of GmUreG.** (A) Ribbon scheme of dimeric GmUreG model structure, derived from the sequence alignments reported in Figure 1A. In the monomer at the left, the ribbons are colored from blue in the proximity of the N-terminal to red at the C-terminus, whereas in the monomer at the right, the main functional features of the protein are color coded as in the alignment shown in Figure 1A (P-loop, red; switch-I, blue; switch-II, dark green; guanine specificity pocket, violet). The GTP $\gamma$ S molecule and the  $\text{Mg}^{2+}$  ion are shown as ball and stick and colored according to atom type (CPK code). (B): Solvent-excluded surfaces of the structural model of dimeric GmUreG colored on the basis of the electrostatic potential. The surface shown in the middle panel is in the same orientation as in A, whereas the surface in the top panel is rotated by 90° around the horizontal axis. Bottom panel displays the two halves of the dimeric GmUreG model structure obtained by rotating each monomer chain by 90°, as indicated by the curved arrows, in order to display the interacting surface of each monomer and the C2 symmetry relating the two monomers.



Figure 1

A



B

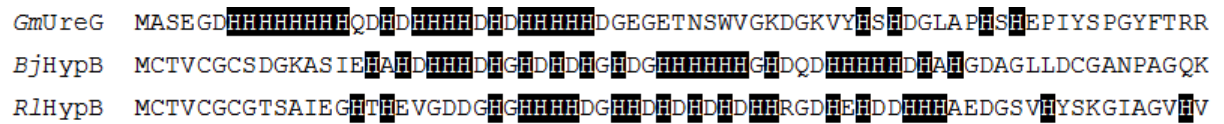


Figure 2

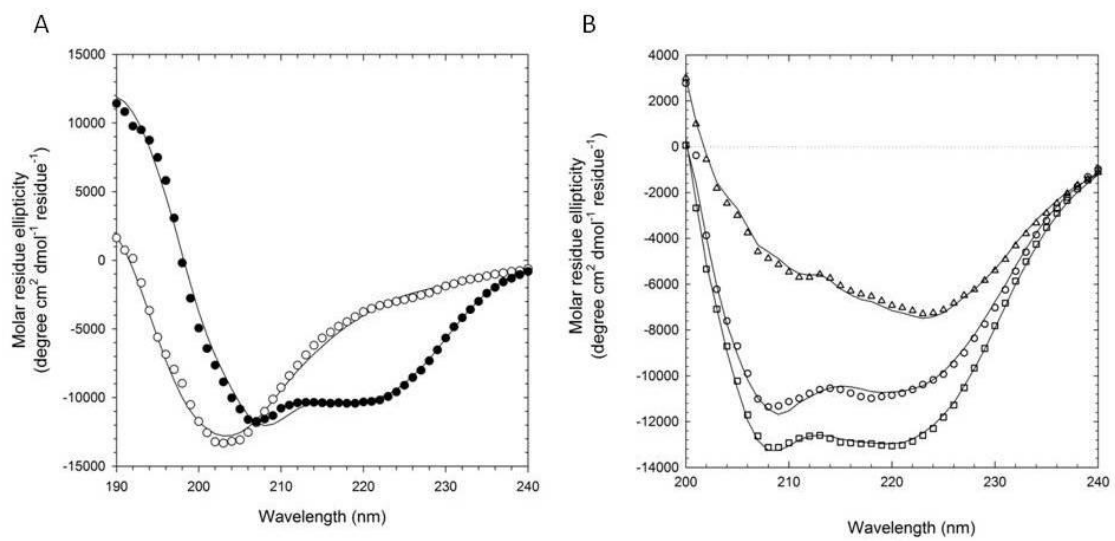


Figure 3

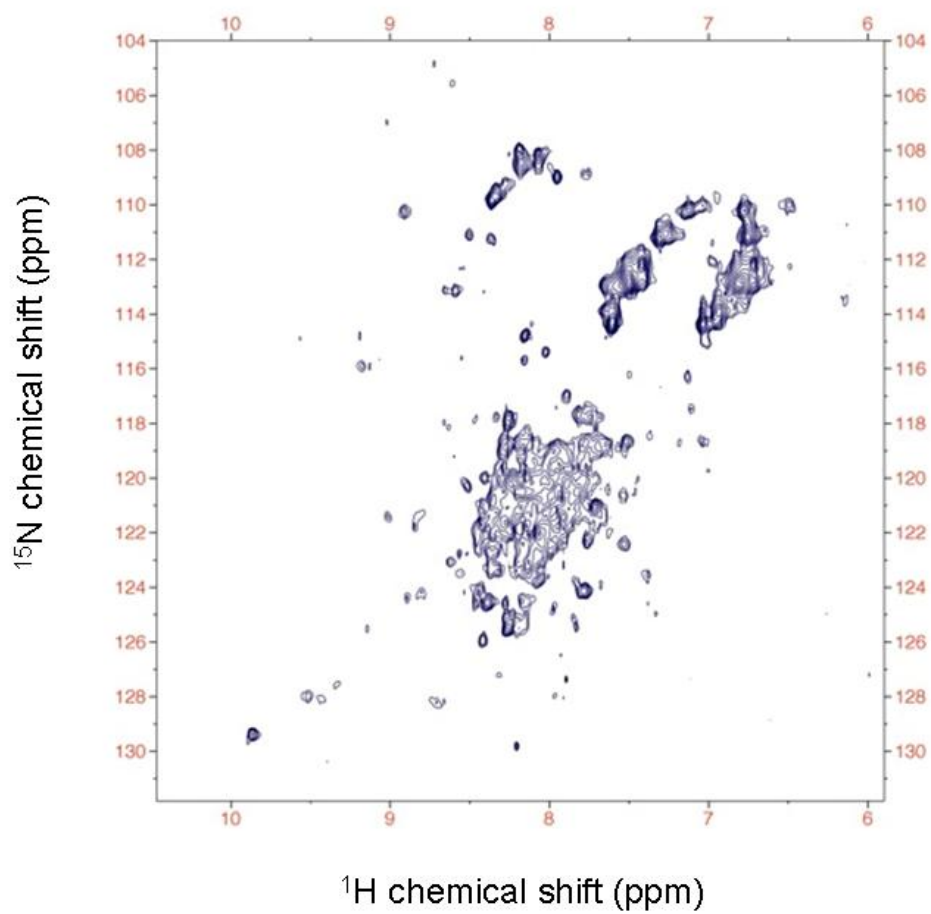


Figure 4

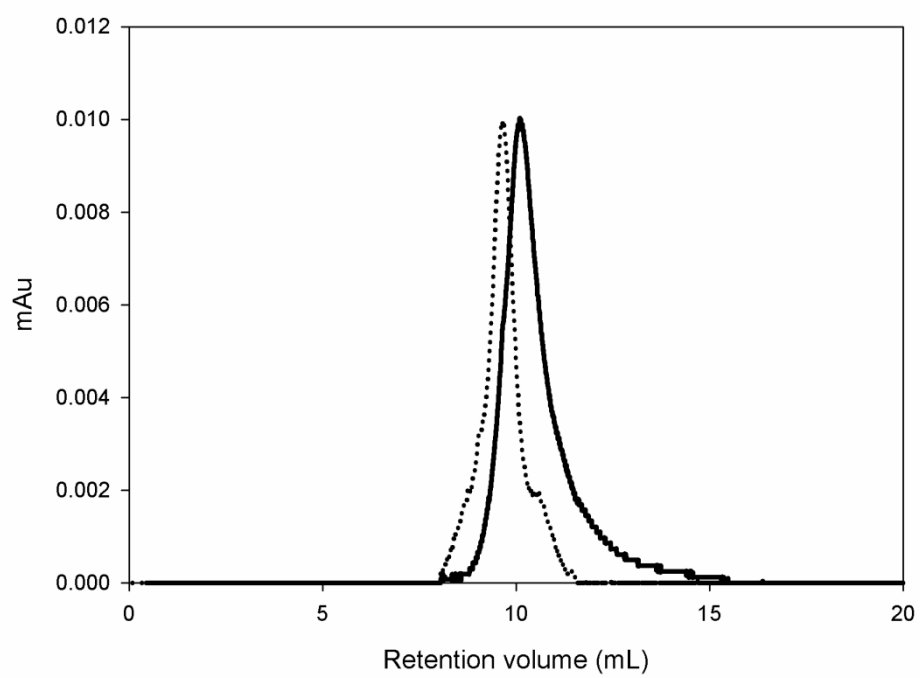


Figure 5

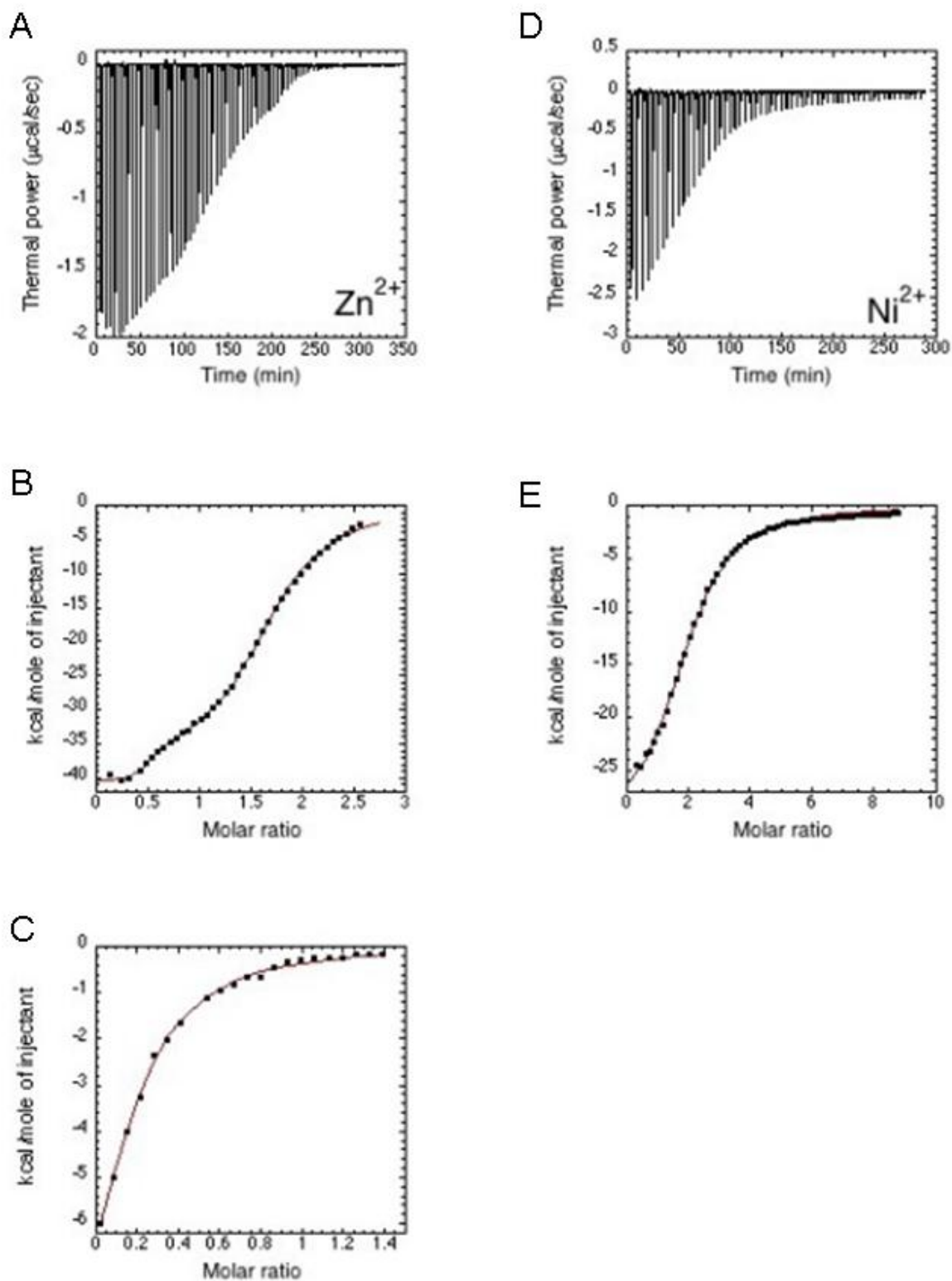


Figure 6

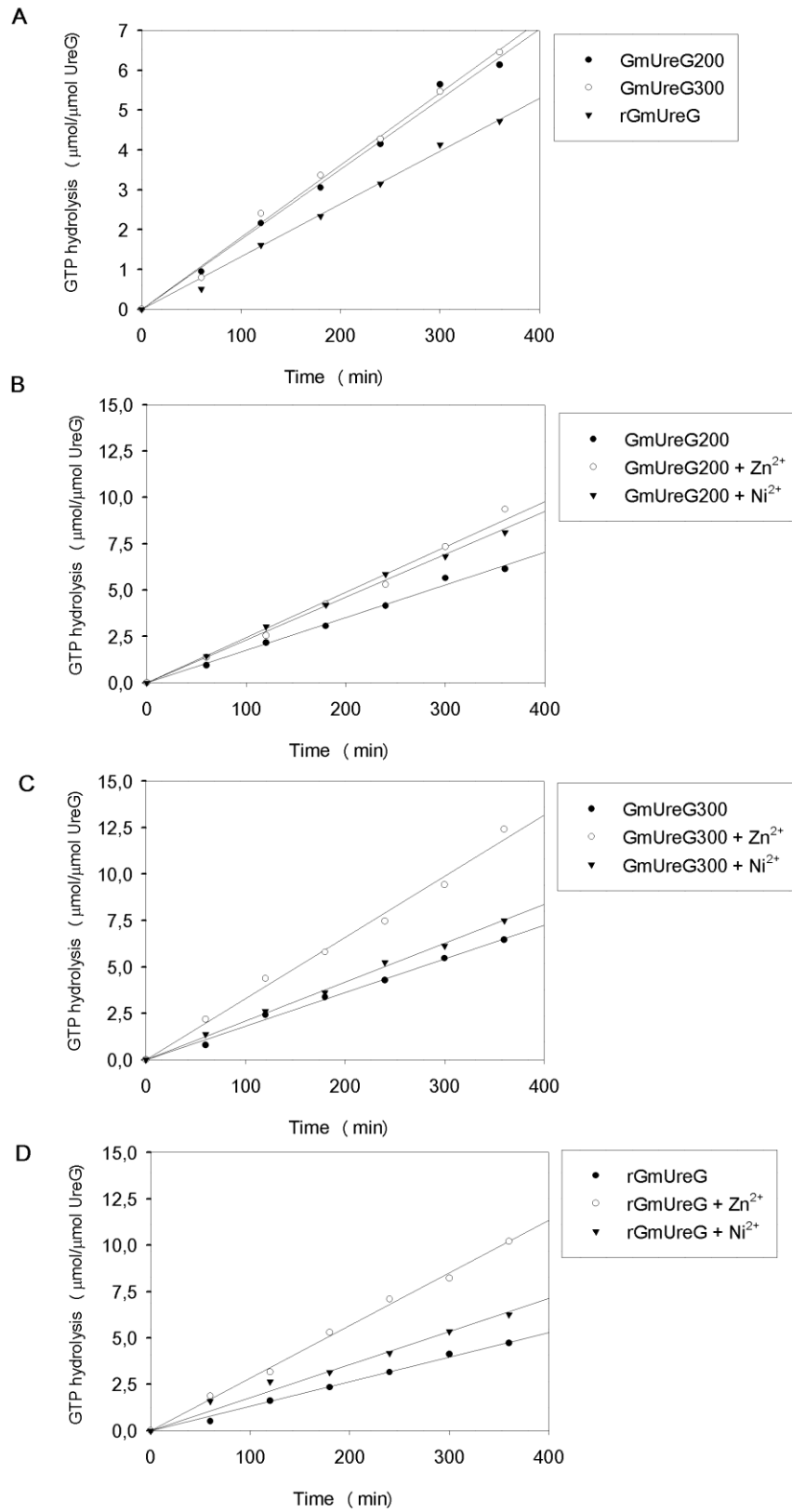
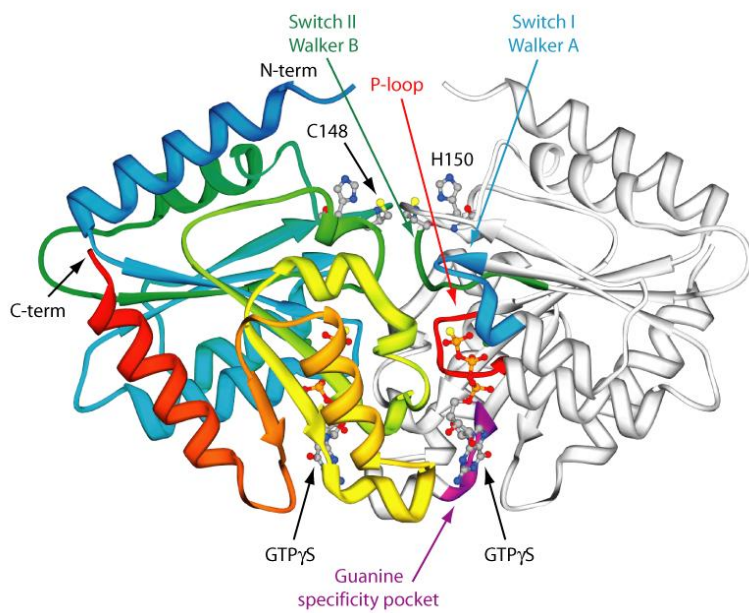
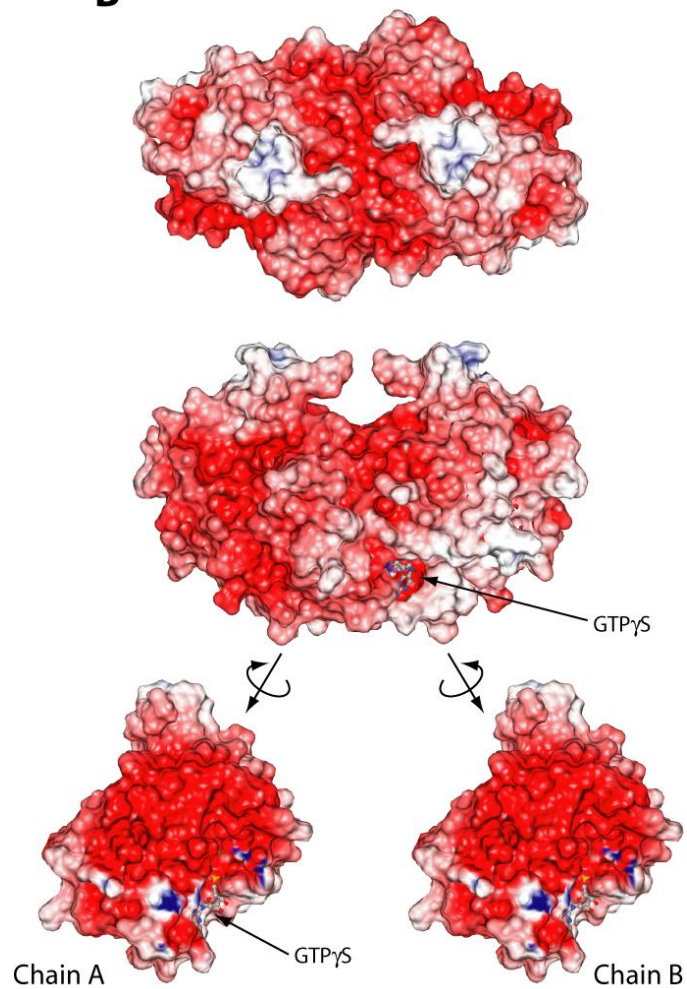


Figure 7

**A**



**B**



## Supplementary online material

### **Biochemical and structural studies on native and recombinant *Glycine max* UreG: the first detailed characterization of a plant urease accessory protein.**

Rafael REAL-GUERRA\*<sup>†</sup>, Fernanda STANISCUASKI\*<sup>‡</sup>, Barbara ZAMBELLI<sup>§</sup>,  
Francesco MUSIANI<sup>§</sup>, Stefano CIURLI<sup>§,||</sup> and Célia R. CARLINI\*<sup>‡</sup>

#### Material and Methods

##### **Purification of *GmUreG* from the seeds**

*GmUreG* was purified according to a protocol adapted from [1]. Soybean seeds (200 g) were grounded in a mill and proteins were extracted with 5 volumes of buffer A (50 mM Tris-HCl pH 7.5, 500 mM NaCl) containing 15 mM imidazole, under stirring at 4 °C for 2 h. The extract was centrifuged for 30 min at 14,000 x g. The supernatant was applied to a Chelating-Sepharose resin, previously loaded with Ni<sup>+2</sup> and equilibrated with the same buffer. Non-retained proteins were washed from the column with buffer A containing 50 mM imidazole. Proteins bound to the column were eluted stepwise with buffer A containing increasing concentrations of imidazole (100, 200 e 300 mM). Protein concentration was measured by absorbance at 280 nm, and the value for the extinction coefficient ( $\epsilon_{280} = 19,940 \text{ M}^{-1}\cdot\text{cm}^{-1}$ ) was calculated from the amino acid sequence using the ProtParam tool ([au.expasy.org/tools/protparam.html](http://au.expasy.org/tools/protparam.html)). The purity of the sample was assessed by SDS-PAGE.

##### **Cloning, expression and purification of recombinant full length *GmUreG***

Full length *GmUreG* cDNA was amplified by PCR using gene specific primers, designed according to the sequence available at the Phytozome website. The PCR product was purified using the Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences), cloned into the pGEM T-Easy vector (Promega) and transformed into *Escherichia coli* (DH5a) by heat-shock. Colonies were screened by



PCR with the gene specific primers and the plasmid containing the desired insert was purified from these cells using the Wizard® Plus SV Miniprep Kit (Promega). The insert was cleaved from the pGEM T-Easy cloning vector with the restriction enzymes NcoI and BamHI (Promega). The cleaved product was ligated into the expression vector pET15b (Novagen) previously cleaved with NcoI and BamHI (in order to remove the plasmid cleavable His-tag codifying portion) and dephosphorylated using Shrimp Alkaline Phosphatase (Invitrogen). This construction was sequenced to confirm the correct ligation of the insert and then transformed into *Escherichia coli* BL21-CodonPlus (DE3)-RIL. Transformants were selected by PCR using the gene specific primers. 10 mL of overnight culture were inoculated into 400 mL of LB medium, containing ampicillin ( $100 \mu\text{g}\cdot\text{mL}^{-1}$ ) and chloramphenicol ( $40 \mu\text{g}\cdot\text{mL}^{-1}$ ). The culture was grown until the optical density (600 nm) reached 0.6. Expression of *GmUreG* was induced by adding IPTG (0.8 mM) to the culture and expression was carried overnight.

### **Light Scattering Measurements**

Absolute estimates of molecular mass of *GmUreG* were determined using a combination of size exclusion chromatography, multiple angle light scattering (MALS), and quasi-elastic light scattering (QELS). *GmUreG* (100  $\mu\text{M}$ , 200  $\mu\text{L}$ ) in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, was loaded onto a Superdex-75 HR 10/30 (Amersham Biosciences), equilibrated in the same buffer, and eluted at a flow rate of  $0.5 \text{ mL}\cdot\text{min}^{-1}$ . The column outlet was connected to a multiple angle laser light (690.0 nm) scattering DAWN EOS photometer (Wyatt Technology). Quasi-elastic (dynamic) light scattering data were collected at a  $90^\circ$  angle by using a Wyatt-QELS device. The concentration of the eluted protein was determined using a refractive index detector (Optilab DSP, Wyatt). Values of 0.185 for the refractive index increment ( $dn/dc$ ) and 1.330 for the solvent refractive index were used. Molecular weights were determined from a Zimm plot. Data were analyzed using the Astra 4.90.07 software (Wyatt Technology), following the manufacturer's indications.

### **Mass spectrometry analysis**

*GmUreG* bands were excised from Coomassie-stained gels and treated in three washing steps with 180  $\mu\text{L}$  of 50 mM ammonium bicarbonate in 50% acetonitrile for 15 min, followed by one washing step with 180  $\mu\text{L}$  of acetonitrile. Gel slices were dried and digested for 18–24 h at  $37^\circ\text{C}$  using 12  $\mu\text{L}$  of  $10 \text{ mg}\cdot\text{mL}^{-1}$  modified porcine trypsin

in 25 mM  $\text{NH}_4\text{HCO}_3$ . Peptides were extracted in two washing steps with 50  $\mu\text{L}$  of 50% acetonitrile 5% TFA for 1 h, and then dried and resuspended in 10  $\mu\text{L}$  of Milli-Q water. MS/MS analyses were performed in a Q-TOF Micro™ mass spectrometer (Micromass, Manchester, UK) coupled to a nanoflow capillary liquid chromatography system (UPLC, Waters, Milford, US). Data were acquired in the data-dependent mode, and multiple charged peptide ions (+ 2 and + 3) were automatically mass selected and dissociated in MS/MS experiments. Typical LC and ESI conditions were: flow of 300  $\text{nL}\cdot\text{min}^{-1}$ , nanoflow capillary voltage of 3.5 kV, block temperature of 100 °C, cone voltage of 50 V. The MS/MS spectra were processed using ProteinLynx v. 2.0 software (Waters, Milford, US) and the generated PKL files were used to perform database searches using the MASCOT (Matrix Science) against the NCBI database. All mass spectrometry data were acquired at the Unidade de Química de Proteínas e Espectrometria de Massas (Uniprote-MS), Centro de Biotecnologia, UFRGS.

### Calorimetric experiments data analysis

The integrated heat data were analyzed using the Origin software package (Microcal), and fitted using a non-linear least-squares minimization algorithm to theoretical titration curves that involved different binding models. The reduced chi-square parameter  $cv2$  ( $cv2 = c2/N$ , where  $N$  is the degrees of freedom,  $N = N_{idp} - N_{par}$ ,  $N_{idp}$  = number of points, and  $N_{par}$  = number of parameters floating in the fit) was used to establish the best fit among the tested models. Values for the enthalpy change of reaction ( $\Delta H$ ), the binding affinity constant ( $K_b$ ) and the number of sites ( $n$ ) were the parameters of the fit. The reaction entropy was calculated using the equations  $\Delta G = -RT \cdot \ln(K_b)$  ( $R = 1.9872 \text{ cal mol}^{-1} \text{ K}^{-1}$ ,  $T = 298 \text{ K}$ ) and  $\Delta G = \Delta H - T\Delta S$ .

### Molecular Modeling

An initial alignment of *Methanococcus jannaschii* HypB (*MjHypB*) and *GmUreG* sequences was produced using Modeller v9.8, [2] and this alignment was manually adjusted in order to match the primary and secondary structure of the proteins, as determined using the JPRED server [3]. The Cys and His residues involved in the binding of  $\text{Zn}^{2+}$  in the structure of *MjHypB* were aligned with the corresponding residues in the fully conserved Cys-Pro-His motif found in all UreG proteins [4]. This alignment was then used to calculate 100 structural models of the dimer of *GmUreG* (which included two  $\text{Mg}^{2+}$  ions, the hydrogen atoms, and the  $\text{GTP}\gamma\text{S}$  molecule) using the

program Modeller and the *Mj*HypB dimer structure (PDB code 2HF8) as template, imposing a structural identity for the two monomers. The best models were selected on the basis of the lowest value of the DOPE score in Modeller [5] and were subjected to a refining step of loop optimization using Modeller standard routines. The results of the PROCHECK analysis [6] for the final model were fully satisfactory. The model of *GmUreG* structure is available upon request to the authors. The molecular (solvent-excluded) surfaces and graphics of the *GmUreG* domains were calculated using the UCSF Chimera package [7]. The electrostatic potential color-code was generated using the DelPhi software [8]. All histidine residues were considered neutral. The protein internal dielectric constant was set to 4 in all calculations, and the solvent dielectric constant was 80. The salt concentration was set to 0.15 M NaCl, which corresponds to the physiological ionic strength.

## References

- 1 Freyermuth, S. K., Bacanamwo, M. and Polacco, J. C. (2000) The soybean *Eu3* gene encodes an Ni-binding Protein necessary for urease activity. *The Plant J.* 21, 53-60
- 2 Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29, 291–325
- 3 Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. and Barton, G. J. (1998) Jpred: A consensus secondary structure prediction server. *Bioinformatics* 14, 892-893
- 4 Zambelli, B., Musiani, F., Savini, M., Tucker, P. and Ciurli, S. (2007) Biochemical studies on *Mycobacterium tuberculosis* *UreG* and comparative modeling reveal structural and functional conservation among the bacterial *UreG* family. *Biochemistry* 46, 3171-3182
- 5 Shen, M.Y. and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507-2524.
- 6 Laskowski, R. A., Macarthur, M. W., Moss, D. S. and Thornton, J. M. (1993). Procheck - A Program To Check The Stereochemical Quality Of Protein Structures. *J. Appl. Crystal.* 26, 283-291

7 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF Chimera- a visualization system for exploratory research and analysis. *J Comput Chem.* 25, 1605-1612

8 Honig, B. and Nicholls, A. (1995) Classical electrostatics in biology and chemistry. *Science* 268, 1144-1149

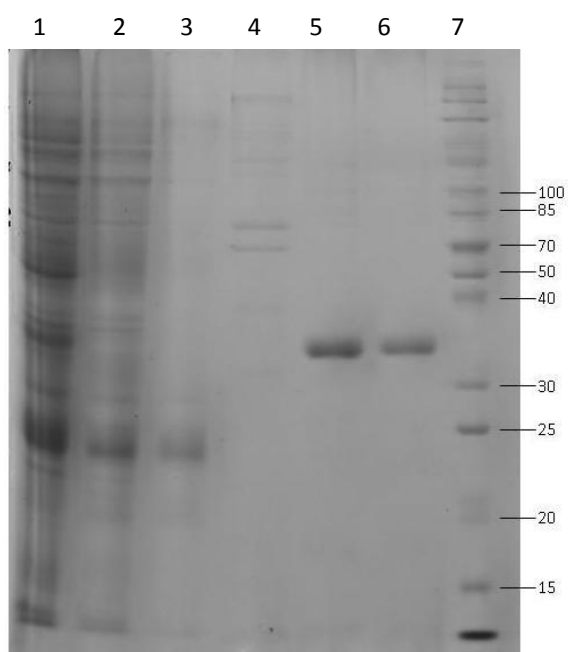
## Figure Legends

Figure S1: **Electrophoresis analysis.** Proteins were analyzed in a 12% SDS-PAGE. (A) Purification of *GmUreG* from soybean seeds. Lanes 1: Crude Extract; 2: non-retained proteins; 3: 50 mM imidazole washes; 4: elution with 100 mM imidazole; 5: elution with 200 mM de imidazole (*GmUreG*200); 6: elution with 300 mM imidazole (*GmUreG*300); 7: molecular weight standard in kDa. (B) Purification of r*GmUreG*. Lanes 1: molecular weight standard; 2: protein eluted from the Ni<sup>2+</sup> affinity chromatography; 3: protein after gel-filtration.

Figure S2: **MALS/QELS of *GmUreG*.** Plot of the molar mass distribution for r*GmUreG*. The solid lines indicate the Superdex S-200 size-exclusion elution profile monitored by the refractive index detector, and the dots are the weight-averaged molecular masses for each slice, measured every second. The average molecular mass is indicated.

Figure S1

**A**



**B**

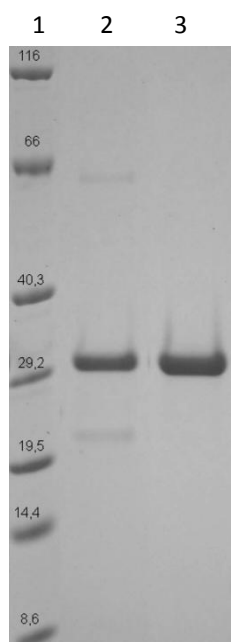


Figure S2

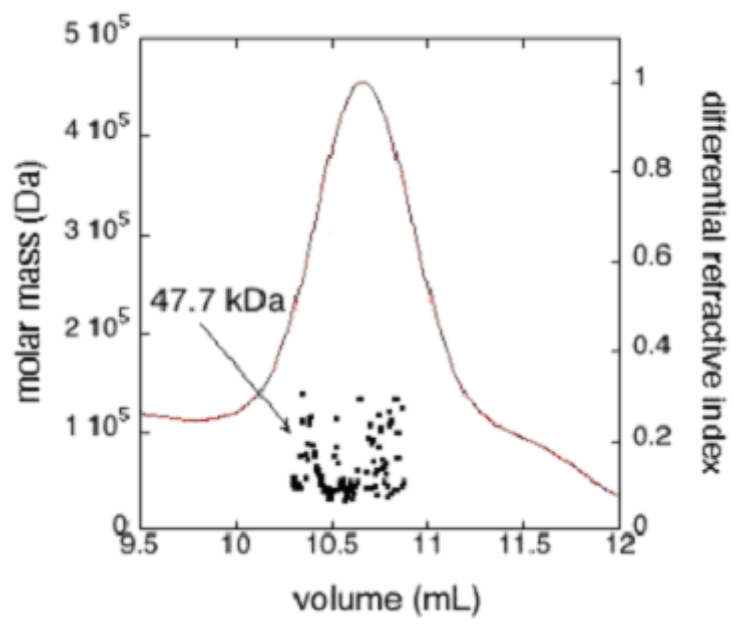


Table S1

MS analysis of *GmUreG*.

Peptide Sequence	GmUreG200	GmUreG300	rGmUreG
RAPLLNR	+	+	
AFTVGIGGPVGTGK	+	+	+
TALMLALCELLR	+		
ENYSLAAVTNDIFTK	+		+
EDGEFLVK	+		
ALPEER	+	+	
AVETGGCPHAAIR	+		
EDISINLGPLEELSNLFK	+	+	+
ELADYIIYIIDVSGGDKIPR	+		+
GGPGITQADLLVINK	+	+	+
TDLAPAIGADLAVMQR	+	+	+
DGGPFVFAQVK	+		
IGVEEIGNLVLQAWAATGNK	+		+

+ indicates that the peptide was found in the MS analysis of the protein.

Table S2 - ICP analysis of GmUreG

	<b>Ag 328.068</b>	<b>Al 167.078</b>	<b>Al 176.641</b>	<b>Ar 404.442</b>	<b>As 189.042</b>	<b>B 208.959</b>	<b>Ba 455.404</b>	<b>Be 313.042</b>	<b>Ca 183.801</b>
	ppm	ppm	ppm	cps [corr]	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	< 0.000	< 0.000	< -0.023	1782750	< 0.004	< -0.000	0,006	< -0.000	2,08
<b>GmUreG300</b>	n.d	n.d	< -0.921	n.d	n.d	0,007	n.d	n.d	< -5.95
<b>rGmUreG</b>	< 0.000	0,002	< -0.016	1743560	< 0.001	< -0.004	0,004	< -0.000	0,022
	<b>Cd 228.802</b>	<b>Co 228.616</b>	<b>Cr 267.716</b>	<b>Cu 324.754</b>	<b>Fe 241.331</b>	<b>Fe 259.941</b>	<b>Hg 184.950</b>	<b>K 766.491</b>	<b>Li 670.780</b>
	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	< -0.000	< 0.000	< -0.000	0,002	< -0.016	0,002	0,015	1,64	0
<b>GmUreG300</b>	< -0.006	< -0.009	< -0.010	< -0.000	n.d	< -0.686	n.d	1,6	n.d
<b>rGmUreG</b>	< -0.000	< 0.000	< -0.000	0,004	< -0.020	< 0.001	0,067	0,04	0
	<b>Mg 279.079</b>	<b>Mn 257.611</b>	<b>Mo 202.095</b>	<b>Na 330.237</b>	<b>Na 589.592</b>	<b>Ni 231.604</b>	<b>P 178.287</b>	<b>Pb 220.353</b>	<b>S 182.034</b>
	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	0,708	0,001	< 0.000	320	77,1	0,004	0,019	< 0.001	0,968
<b>GmUreG300</b>	0,083	< -0.009	< -0.010	464	45,1	< 0.001	< -0.247	< 0.002	1,47
<b>rGmUreG</b>	< 0.001	< -0.000	< 0.000	< 7.93	9,73	0,008	< 0.003	< -0.003	2,81
	<b>Sb 206.833</b>	<b>Se 196.090</b>	<b>Si 251.612</b>	<b>Sn 189.991</b>	<b>Sr 407.771</b>	<b>Ti 334.941</b>	<b>Tl 190.864</b>	<b>V 292.464</b>	<b>Zn 213.856</b>
	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
<b>GmUreG200</b>	< -0.002	< 0.005	0,429	< 0.004	0,01	< -0.000	< -0.004	< -0.001	0,007
<b>GmUreG300</b>	n.d	n.d	n.d	n.d	n.d	n.d	n.d	n.d	0,096
<b>rGmUreG</b>	< -0.002	< 0.003	< 0.007	< -0.000	0	< -0.000	< -0.000	< -0.001	0,005

Data is presented as part per million (p.p.m.). Metals with differential content are highlighted.