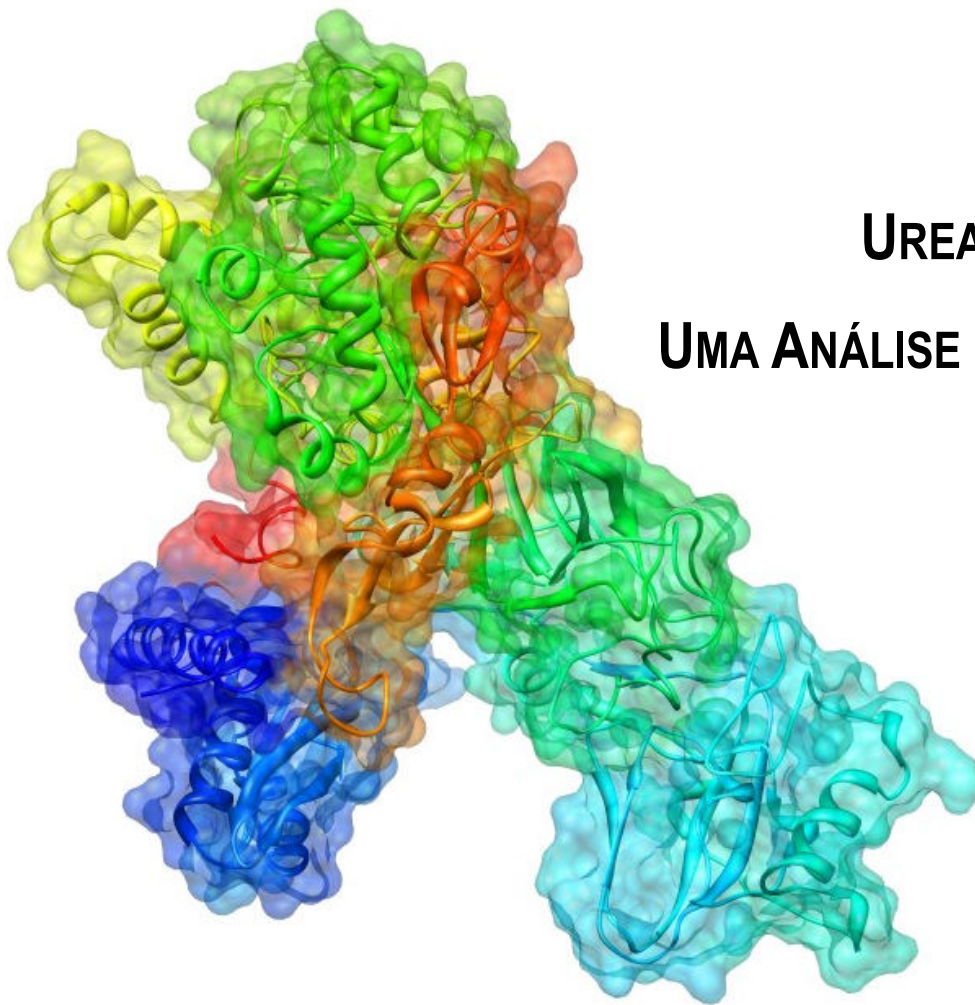


RODRIGO LIGABUE BRAUN



**UREASES:
UMA ANÁLISE COMPARATIVA**

DRA. CÉLIA REGINA RIBEIRO DA SILVA CARLINI
Orientadora

DRA. EVELYN KOECHE SCHROEDER
Co-orientadora

PORTO ALEGRE

2008

RODRIGO LIGABUE BRAUN

**UREASES:
UMA ANÁLISE COMPARATIVA**

Monografia apresentada como requisito parcial para obtenção do grau de Bacharel em Ciências Biológicas – Ênfase Molecular, Celular e Funcional. Universidade Federal do Rio Grande do Sul Instituto de Biociências; Departamento de Biofísica e Centro de Biotecnologia do Estado do Rio Grande do Sul.

DRA. CÉLIA REGINA RIBEIRO DA SILVA CARLINI
Orientadora

DRA. EVELYN KOECHE SCHROEDER
Co-orientadora

PORTO ALEGRE

2008

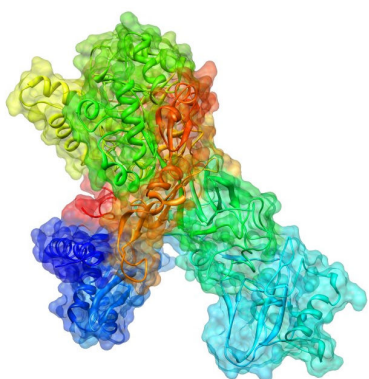


Figura da capa: estrutura tridimensional da isoforma alternativa (JBURE-IIB) da urease de *Canavalia ensiformis*, colorida de acordo com a sucessão da estrutura secundária (N-terminal: azul, C-terminal: vermelho)

Agradecimentos

Agradeço especialmente às minhas orientadoras, por terem me acolhido e direcionado, sendo sempre sinceras, construtivas e divertidas.

Agradeço, também, aos colegas do Laprotox, por sempre me incentivarem (direta ou indiretamente), e aos membros do Grupo de Bioinformática Estrutural (CBiot, UFRGS), por estarem sempre dispostos a compartilhar seus conhecimentos.

Por fim, agradeço a BDGIJLLLLMPVW, por permanecerem comigo ao longo de todo esse processo.

“The effort to understand the universe is one of the very few things that lifts human life a little above the level of farce, and gives it some of the grace of tragedy”

Steven Weinberg

Nota: com a finalidade de tornar o texto mais fluido e informativo ao leitor, preferiu-se adotar um formato diferenciado para a presente monografia. Os resultados obtidos e sua discussão foram combinados e constiuem uma seção única do trabalho. Esta seção também aborda alguns aspectos metodológicos específicos.

RESUMO

Ureases (EC 3.5.1.5) são enzimas níquel-dependentes que catalisam a hidrólise da uréia em amônia e dióxido de carbono. São amplamente distribuídas em plantas, fungos e bactérias. Em plantas e fungos, as ureases são homopolímeros de subunidades contendo cerca de 840 resíduos de aminoácidos, formando hexâmeros ou trímeros. Em bactérias, as ureases possuem duas ou três subunidades menores, que se alinham na cadeia polipeptídica única das ureases vegetais, com as quais têm 50-60% de identidade de seqüência. Para ter-se uma visão geral das relações que podem ser estabelecidas entre as ureases, buscando padrões, similaridades e diferenças que possam auxiliar na análise e predição dos resultados experimentais, foi realizada uma ampla análise das seqüências de nucleotídeos e aminoácidos de ureases. Após busca, extração, interpretação e edição destas, árvores filogenéticas foram construídas. Por meio de alinhamentos de seqüência múltiplos, a grande similaridade entre as seqüências oriundas de plantas e fungos foi confirmada, e permitiu a análise das regiões de conservação próximas ao sítio ativo, além da verificação da conservação de todos os resíduos de aminoácidos considerados essenciais para atividade catalítica da enzima. Diferenças foram encontradas nas regiões correspondentes ao peptídeo entomotóxico originalmente descrito para a urease de *Canavalia ensiformis* L. (feijão-de-porco). Como as estruturas de ureases de plantas e fungos não foram ainda determinadas, um modelo de urease de *C. ensiformis* foi construído por modelagem molecular comparativa, utilizando como molde uma estrutura de urease bacteriana. O modelo foi avaliado e utilizado para determinar a localização espacial do peptídeo entomotóxico e dos resíduos essenciais para a catálise.

Palavras-chave:

ureases – filogenia – modelagem molecular comparativa

ABSTRACT

Ureases (EC 3.5.1.5) are nickel-dependent enzymes which catalyze hydrolysis of urea to ammonia and carbon dioxide. They are found in plants, fungi and bacteria. In plants and fungi, ureases are homopolymers of subunits with ~840 amino acids, forming hexamers or trimers. Bacterial ureases have two or three smaller subunits which align within the single chain of plants and fungi ureases with 50-60% sequence identity. To set an overview of relationships that can be established among ureases, in search for patterns, similarities and differences that may help analysis and prediction of experimental results, a comprehensive analysis of amino acid and nucleotide sequences of ureases was carried out. After search, extraction, interpretation and edition of these sequences, they were used to build phylogenetic trees. By means of multiple sequence alignments, the high level of similarity among plant and fungi sequences was confirmed, allowing analysis of conservation of regions near the active site and confirming the conservation of all residues taken as essential for catalysis. Differences were found in regions related to the entomotoxic peptide originally described for *Canavalia ensiformis* L. (jack bean). Since structures of plants and fungi ureases are still not solved, a model for a *C. ensiformis* urease was built via comparative molecular modeling using a bacterial urease structure as template. The model was evaluated and used to spatially locate the entomotoxic peptide and those residues essential for catalysis.

Keywords:

ureases – phylogeny – comparative molecular modeling

LISTA DE FIGURAS E TABELAS

Figura 1: Esquema da reação catalisada pelas ureases.	13
Figura 2: Comparação esquemática das subunidades estruturais das ureases de organismos selecionados	15
Tabela 1: Seqüências de aminoácidos de ureases utilizadas neste trabalho	23
Figura 3: Relação evolutiva entre 24 seqüências de aminoácidos de ureases.....	27
Figura 4: Relação evolutiva entre 22 seqüências de nucleotídeos de ureases.....	28
Figura 5: Alinhamento múltiplo de seqüências de aminoácidos de ureases oriundas de plantas.....	31
Figura 6: Alinhamento múltiplo de seqüências de aminoácidos de ureases oriundas de fungos e de <i>C. ensiformis</i> , para comparação.....	34
Figura 7: alinhamento múltiplo de seqüências de aminoácidos da subunidade C de ureases oriundas de bactérias e da região correspondente na JBURE-IIB e na isoforma majoritária de urease de <i>C. ensiformis</i>	40
Figura 8: Passos na modelagem comparativa por homologia	43
Figura 9: Mapa de Ramachandran para as estruturas do molde e do modelo.....	46
Figura 10: Resultados do Verify3D transpostos às estruturas tridimensionais do molde e do modelo.....	47
Figura 11: Estrutura tridimensional do modelo de JBURE-IIB, colorido de acordo com a estrutura secundária.....	48
Figura 12: Sobreposição do modelo da JBUREIIB com seu molde 2UBP	48
Figura 13: Estrutura tridimensional do modelo de JBURE-IIB.	48
Figura 14: Representação espacial dos resíduos do sítio catalítico conservados em todas as seqüências de ureases analisadas, dos átomos de níquel e do <i>cluster</i> de água	50
Figura 15: Sobreposição dos resíduos do sítio catalítico conservados em todas as seqüências de ureases analisadas	50
Figura 16: Distribuição das diferentes classes de aminoácidos na JBURE-IIB e na urease de <i>Bacillus pasteurii</i>	51
Figura 17: Predição da estrutura secundária do peptídeo entomotóxico derivado da JBURE-IIB.	52
Figura 18: Determinação da estrutura secundária da JBURE-IIB a partir do modelo tridimensional construído neste trabalho.....	53
Figura 19: Predição da estrutura secundária da JBURE-IIB a partir de sua seqüência primária.....	54

LISTA DE ABREVIATURAS E SIGLAS

2UBP	estrutura tridimensional da urease de <i>Bacillus pasteurii</i> em suas conformação nativa
cDNA	DNA complementar
CNTX	canatoxina
Jaburetox	<i>jack bean urease toxin</i>
Jaburetox-2Ec	<i>jack bean urease toxin 2</i> [expressada em] <i>E. coli</i>
JBURE-I	<i>jack bean urease I</i> (urease de <i>Canavalia ensiformis</i> , isoforma predominante)
JBURE-II	<i>jack bean urease II</i> (urease de <i>C. ensiformis</i> , isoforma alternativa truncada)
JBURE-IIB	<i>jack bean urease IIB</i> (urease de <i>C. ensiformis</i> , isoforma alternativa)
KCX	lisina carbamilada
NCBI	National Center for Biotechnology Information
NJ	<i>neighbor-joining</i>
OTU	unidade taxonômica operacional
PDB	Protein Data Bank
Pepcanatox	peptídeo derivado da canatoxina
Pré-mRNA	precursor do mRNA maduro
RMN	ressonância magnética nuclear
UPGMA	<i>unweighted pair group method with arithmetic means</i>
UreD-K	proteínas acessórias das ureases

ÍNDICE

1. Introdução.....	13
1.1 Ureases.....	13
1.2 Ureases de <i>Canavalia ensiformis</i>	16
1.3 Justificativa.....	17
2. Objetivos.....	18
3. Metodologia.....	19
3.1 Seqüência primária e inferência filogenética.....	19
3.2 Modelagem molecular comparativa.....	21
4. Resultados e Discussão.....	23
4.1 Inferência filogenética.....	23
4.2 Comparação de seqüências primárias.....	30
4.3 Determinação de estruturas terciárias de proteínas.....	42
4.4 Modelagem molecular comparativa.....	42
4.5 Estrutura tridimensional das ureases.....	47
5. Conclusões e perspectivas.....	56
6. Referências bibliográficas.....	57
APÊNDICES.....	63
APÊNDICE A: arquivos de entrada utilizados no programa Modeller 9v4.....	64
APÊNDICE B: listagem das seqüências aminoácidos de ureases bacterianas obtidas no presente trabalho.....	65
APÊNDICE C: Tabela comparativa de dados cinéticos para ureases.....	70

1. Introdução

1.1 Ureases

Ureases (uréia amidohidrolases, EC 3.5.1.5) são enzimas níquel-dependentes que catalisam a hidrólise da uréia em amônia e dióxido de carbono (Zerner, 1991), a uma taxa 10^{14} vezes mais rápida que a reação não catalisada (Andrews *et al.*, 1984; Mobley *et al.*, 1995). O mecanismo catalítico continua sendo de interesse, pois não está definido se a reação envolve um intermediário do tipo ácido carbâmico, um intermediário cianato ou se ela procede por decomposição direta (Benini *et al.*, 1999; Todd & Hausinger, 2000; Estiu *et al.*, 2006). Sabe-se que a reação não catalisada ocorre por um mecanismo de eliminação, tendo sido identificado um carbamato como intermediário (Barrios & Lippard, 2001; Estiu *et al.*, 2006) (Figura 1).

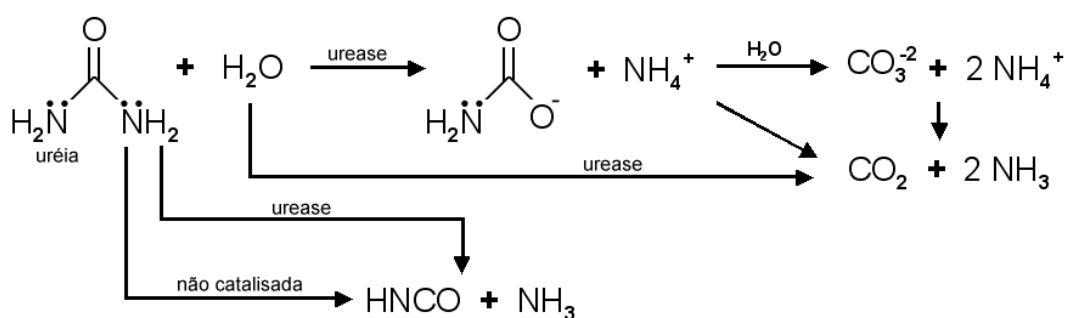


Figura 1: Esquema da reação catalisada pelas ureases (adaptado de Estiu *et al.*, 2006).

Tanto a uréia quanto a urease representam marcos da investigação científica (Mobley *et al.*, 1995). A uréia foi a primeira molécula a ser sintetizada, em 1828 por Friedrich Wöhler (Andrews *et al.*, 1984), e a urease de *Canavalia ensiformis* (feijão-de-porco ou *jack bean*) foi a primeira enzima a ser cristalizada (Sumner, 1926), além de ter sido a primeira enzima com presença de níquel demonstrada (Dixon, 1975).

O papel principal da urease é permitir que os organismos utilizem a uréia, gerada interna ou externamente, como fonte de nitrogênio (Jabri *et al.*, 1995). Além deste papel, todavia, as ureases apresentam grande versatilidade de aplicações, podendo atuar no transporte de nitrogênio e defesa em plantas e na neutralização do ambiente gástrico de mamíferos para colonização bacteriana (Ha *et al.*, 2001).

Além das plantas, algumas bactérias, fungos e possivelmente invertebrados sintetizam ureases. As ureases animais são pouco caracterizadas, tendo sido

identificadas como coadjuvantes no processo de formação de inclusões de carbonato de cálcio no molusco *Aplysia californica* (Pedrozo, 1996a,b). É possível, nesse caso, que a urease seja oriunda de alguma bactéria simbiote, e não produzida pelo próprio molusco. Além disso, foi demonstrado que o bicho-da-seda (*Bombyx mori*) não sintetiza urease, utilizando em seu metabolismo a enzima absorvida da dieta (constituída exclusivamente por folhas da amoreira *Morus alba*), indicando que a urease passa intacta do intestino à hemolinfa, estando associada à degradação de uréia após o estágio ingestivo do inseto (Hirayama, 2000a).

As ureases têm papel importante na patogênese de várias espécies bacterianas, incluindo *Proteus mirabilis*, *Staphylococcus saprophiticus*, *Yersinia enterocolitica* e *Ureaplasma urealiticum* (Mobley *et al.*, 1995). O exemplo mais freqüente na literatura é o da urease de *Helicobacter pylori* (Sirko & Brodzik, 2000), por seu papel essencial na patogênese desse microorganismo, além de sua alta prevalência como patógeno humano (Eaton *et al.*, 1991). Os melhores dados estruturais, entretanto, estão disponíveis para *Klebsiella aerogenes* (Jabri *et al.*, 1995).

Atividade ureásica foi encontrada em várias espécies de fungos, mas a descrição genética correspondente está disponível para apenas algumas delas (Sirko & Brodzik, 2000). Exemplos incluem o patógeno humano *Coccidioides immitis* (Yu *et al.*, 1997) e *Schizosaccharomyces pombe* (Tange & Niwa, 1997).

Em relação a ureases de plantas, os dados genéticos mais abrangentes estão disponíveis para soja (*Glycine max*) (Polacco & Holland, 1993; 1994). Genes separados, que codificam duas isoformas da urease, uma ubíqua e uma embrião-específica, assim como genes não ligados, codificadores de proteínas acessórias, foram identificados em soja (Meyer-Bothling & Polacco, 1987). A urease embrião específica é uma proteína abundante em sementes de várias espécies de plantas, incluindo soja, feijão-de-porco (Polacco & Holland, 1994) e *Arabidopsis* (Zonia *et al.*, 1995), enquanto a forma ubíqua é encontrada em baixas quantidades nos tecidos vegetativos da maioria das plantas (Sirko & Brodzik, 2000). Bioquimicamente, a urease vegetal melhor caracterizada é a de *C. ensiformis* (Hirai *et al.*, 1993; Karmali & Domingos, 1993; Riddles *et al.*, 1991; Takishima *et al.*, 1988). Recentemente, a urease das folhas da amoreira *M. alba* também foi purificada e caracterizada (Hiraiama, 2000b), bem como a urease das sementes de algodão, *Gossypium hirsutum* (Menegassi *et al.*, 2008).

A alta similaridade das seqüências primárias (50-60%) indica que todas as ureases devem ser variantes da mesma enzima e que, por isso, devem possuir estruturas terciárias e mecanismos catalíticos similares (Mobley *et al.*, 1995). Estruturalmente, as ureases bacterianas são multímeros formados por complexos de duas ou três subunidades, sendo exemplos, respectivamente, *Helicobacter pylori* e *K. aerogenes*. Em plantas e fungos, a enzima é homo-oligomérica, sendo formada por subunidades idênticas de aproximadamente 90kDa. (Mobley *et al.*, 1995; Tange & Niwa, 1997) (Figura 2). A urease de *C. ensiformis* existe como homotrímero, capaz de agregar-se em homohexâmeros. As ureases bacterianas possuem estruturas similares às da urease de *C. ensiformis*, sendo trímeros ou hexâmeros de complexos de subunidades, podendo também existir na forma de agregados (Sirko & Brodzik, 2000). O sítio catalítico encontra-se na subunidade C da enzima de *K. aerogenes* e nas regiões colineares (Figura 2) das ureases dos outros organismos¹.

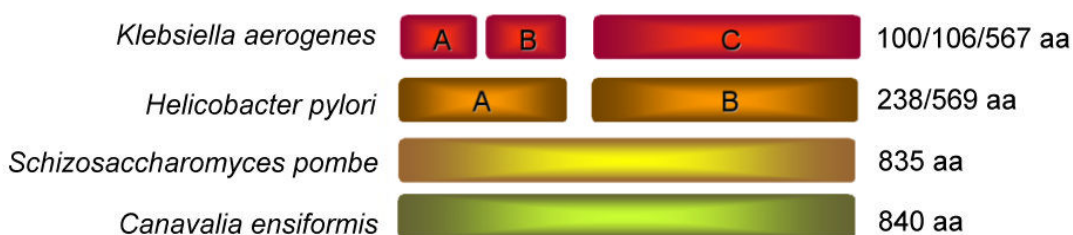


Figura 2: Comparação esquemática das subunidades estruturais das ureases de organismos selecionados. O número de aminoácidos é apresentado à direita (adaptado de Sirko & Brodzik, 2000)

Cada subunidade catalítica contém no sítio ativo dois íons de níquel que, no caso da urease cristalizada de *K. aerogenes*, estão separados por 3.5Å (Jabri *et al.*, 1995). Nesta enzima, um átomo de níquel está ligado a dois resíduos de histidina (His246 e His272), enquanto o segundo átomo de níquel está ligado a três resíduos: duas histidinas (His134 e His136) e um ácido aspártico (Asp360). Além destes, um ligante carbamato, derivado da Lys217 interage com os dois íons de níquel (Mobley *et al.*, 1995). O processo de incorporação de níquel requer a participação de muitas proteínas acessórias (ou ativadoras), que parecem agir como chaperonas urease-específicas (Mobley *et al.*, 1995). Tais proteínas são necessárias para a montagem

¹ A nomenclatura das subunidades das ureases é encontrada de duas maneiras na literatura: utilizando caracteres romanos e gregos. As subunidades A, B e C são equivalentes às subunidades γ , β e α , respectivamente. No presente trabalho, os caracteres gregos serão reservados para a discriminação dos domínios das ureases monoméricas.

de uma urease ativa, tanto em bactérias quanto em plantas (Polacco & Holland, 1994; Mobley *et al.*, 1995).

Até o momento, apenas ureases bacterianas tiveram suas estruturas tridimensionais definidas por cristalografia de raios X. As estruturas das enzimas de *K. aerogenes* (código PDB 1FWJ), *Bacillus pasteurii* (código PDB 2UBP) e *H. pylori* (código PDB 1E9Z) foram determinadas e analisadas (Jabri *et al.*, 1995; Benini *et al.*, 1999; Ha *et al.*, 2001). Mesmo sendo a primeira enzima cristalizada (Sumner, 1926), apenas a cristalização e a análise preliminar por raios X da urease de *C. ensiformis* foram descritas (Jabri *et al.*, 1992; Sheridan *et al.*, 2002). Dados preliminares sobre cristalização da urease de *Cajanus cajan* (feijão-guandu) foram publicados recentemente (Balasubramanian & Ponnuraj, 2008).

1.2 Ureases de *Canavalia ensiformis*

Assim como a soja, a *C. ensiformis* também apresenta isoformas de urease (Pires-Alves *et al.*, 2003). Inicialmente, determinaram-se a seqüência primária da mais abundante urease de feijão-de-porco, com 840 aminoácidos (Mamiya *et al.*, 1985), e a seqüência nucleotídica de seu cDNA (Riddles *et al.*, 1991). Convencionou-se chamar esta urease de JBURE-I (*jack bean urease I*).

Posteriormente, mostrou-se que a canatoxina (CNTX), uma proteína neurotóxica (dímero de subunidades de 95 kDa), é uma isoforma da urease de *C. ensiformis* (Carlini *et al.*, 1997; Follmer *et al.*, 2001). A canatoxina apresenta atividade inseticida, sugerindo que as ureases podem estar envolvidas na defesa das plantas (Ferreira-DaSilva *et al.*, 2000; Carlini & Grossi-de-Sá, 2002). As principais isoformas de urease de feijão-de-porco e a urease embrião-específica de soja (mas não a urease bacteriana de *B. pasteurii*) também apresentaram propriedades inseticidas em testes com *Dysdercus peruvianus* (praga da cultura de algodão). O efeito entomotóxico das ureases é independente da atividade ureolítica (Follmer *et al.*, 2004). Sua toxicidade se baseia em um peptídeo interno (pepcanatox), com 10kDa, liberado pela hidrólise da canatoxina por catepsinas do trato digestório de insetos suscetíveis (Carlini *et al.*, 1997; Ferreira-DaSilva *et al.*, 2000).

Uma possível terceira isoforma de urease de *C. ensiformis*, a JBURE-II, foi proposta, baseada em cDNA clonado. A JBURE-II apresentaria uma cadeia

polipeptídica mais curta (78kDa) que todas as ureases vegetais ou fúngicas conhecidas, um ponto isoelétrico teórico mais básico e seqüência nucleotídica diferindo daquela disponível para canatoxina (Pires-Alves *et al.*, 2003). Em estudo posterior, Mulinari (2008) identificou uma forma alternativa da JBURE-II clonada originalmente, com diferenças de seqüência nas regiões 3' e 5'. Há indícios de possível ocorrência de *splicing* alternativo do pré-mRNA de JBURE-II, especialmente na região 5'. A ocorrência extensiva de *splicing* alternativo no pré-mRNA do gene de urease de batata-inglesa (*Solanum tuberosum*) (Witte *et al.*, 2005) apóia esta teoria. Ainda assim, a hipótese de ter ocorrido um seqüenciamento incompleto da JBURE-II, resultando em uma proteína truncada no estudo inicial, não foi descartada. A forma alternativa de JBURE-II analisada por Mulinari (2008) foi, então, chamada de JBURE-IIB.

Um peptídeo recombinante, equivalente ao pepcanatox e com propriedades inseticidas muito similares, foi obtido por clonagem e expressão heteróloga de um fragmento equivalente de cDNA de JBURE-IIB (Mulinari *et al.*, 2007). Esse peptídeo recombinante foi chamado de Jaburetox-2Ec (*Jack bean urease toxin 2* [expressada em] *E. coli*).

1.3 Justificativa

Apesar da grande quantidade de informações a respeito de ureases em particular, pouco se sabe a respeito das relações que podem ser estabelecidas entre elas. Nesse contexto, uma ampla análise das seqüências de nucleotídeos e aminoácidos de ureases, buscando padrões, similaridades e diferenças, pode auxiliar na análise e predição dos resultados experimentais, tanto daqueles obtidos no Laboratório de Proteínas Tóxicas (UFRGS), quanto daqueles já descritos na literatura. Além disso, a modelagem de uma urease de interesse (no caso, uma urease de *C. ensiformis*) pode contornar o problema da não-disponibilidade de estruturas tridimensionais de ureases de plantas e fungos. A predição da estrutura tridimensional pode, ainda, esclarecer certos aspectos referentes a essa urease, como a conformação de seu peptídeo entomotóxico e das regiões da proteína que encontram-se ausentes nas ureases bacterianas.

2. Objetivos

O presente trabalho teve como objetivo principal o estudo comparativo das ureases, a partir de dados já disponíveis a seu respeito, e o estudo computacional de uma das ureases de *C. ensiformis*, buscando evidenciar similaridades e diferenças entre as ureases já estudadas, bem como guiar futuros estudos comparativos em relação a essa enzima.

Como objetivo secundário, este trabalho procura fazer uma breve revisão das técnicas necessárias para o desenvolvimento deste projeto e para a interpretação dos resultados dele obtidos.

3. Metodologia

3.1 Seqüência primária e inferência filogenética

Inicialmente, foi realizada uma busca por seqüências de aminoácidos referentes à palavra-chave “urease” na base de dados de proteínas do National Center for Biotechnology Information - NCBI (Wheeler *et al.*, 2008) (<http://www.ncbi.nlm.nih.gov>). As seqüências foram baixadas em blocos de duzentas, no dia 15 de novembro de 2007. As 9046 seqüências obtidas foram agrupadas em arquivos de texto, contendo mil seqüências cada. Estes arquivos serviram de base para as etapas subseqüentes do trabalho e foram verificados, de maneira automatizada ou não, diversas vezes, com diferentes intenções.

Inicialmente, buscaram-se seqüências de proteínas acessórias das ureases (UreD-K) presentes entre as seqüências obtidas. Todas elas foram removidas da listagem. A seguir, cada bloco foi inspecionado visualmente, buscando seqüências de tamanho incomum. Cada seqüência selecionada foi comparada às demais, via Blast (Altschul *et al.*, 1997), tendo destinos diferentes conforme a situação. Seqüências curtas sem equivalentes oriundos da mesma espécie foram excluídas, por serem resultado de possíveis seqüenciamentos incompletos. Seqüências curtas com seqüências equivalentes maiores, oriundas da mesma espécie, foram agrupadas junto a essas.

Todas as seqüências não-únicas, oriundas de um mesmo organismo, foram agrupadas, comparadas por alinhamentos múltiplos via ClustalW (Larkin *et al.*, 2007) (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html) e, a partir desta comparação, uma das seqüências foi eleita como seqüência-tipo (ou seja, aquela que representava de maneira mais satisfatória a urease em questão). Este ciclo, de agrupamento, comparação e seleção da seqüência-tipo, foi repetido para a maioria das seqüências. No caso de bactérias, cada subunidade foi tratada independentemente, com as seqüências das duas ou três subunidades oriundas de cada bactéria sendo agrupadas ao final do processo. Casos especiais, como o das diferentes ureases de batata-inglesa (*S. tuberosum*), foram analisados posteriormente.

Após o refinamento de todas as seqüências, com cada organismo sendo associado a apenas uma seqüência de aminoácidos (para cada uma de suas subunidades de urease), a nomenclatura científica dos organismos foi substituída por um código, geralmente formado por seis caracteres. A seguir, uma busca pelas

seqüências de desoxirribonucleotídeos correspondentes às seqüências-tipo de plantas e fungos e de algumas bactérias selecionadas foi realizada na base de dados de nucleotídeos do NCBI. Para melhor visualização, foi estabelecida uma listagem, contendo o organismo-fonte, o código do organismo, a referência interna do NCBI para a seqüência-tipo, indicação sobre a disponibilidade da seqüência de DNA correspondente à seqüência de aminoácidos selecionada e o táxon ao qual o organismo pertence.

Após estarem todas catalogadas, as seqüências de aminoácidos de plantas, fungos e de algumas bactérias selecionadas (*Helicobacter pylori*, *Helicobacter hepaticus*, *Klebsiella aerogenes*, *Klebsiella pneumoniae*, *Bacillus pasteurii*, *Bacillus subtilis*) foram alinhadas pelo ClustalW. O alinhamento foi inspecionado visualmente, por meio do programa BioEdit 8 (Hall, 1999).

A versão final do alinhamento foi utilizada para a construção de uma árvore filogenética. Tal construção foi feita por meio do programa MEGA4 (Tamura *et al.*, 2007), pelo método de *neighbor-joining* (ou ligação com vizinhos) (Saitou & Nei, 1987), com modelo de correção de Poisson (Zuckerandl & Pauling, 1965), e com a filogenia sendo testada por um *bootstrap* (Felsenstein, 1985) de 10.000 replicatas. Todas as posições contendo *gaps* foram removidas (opção de deleção completa).

Da mesma forma, as seqüências de nucleotídeos de ureases de plantas, fungos e das cadeias C de bactérias selecionadas foram alinhadas e editadas, sendo utilizadas para a construção de uma árvore filogenética, por meio do programa MEGA4. A história evolutiva foi inferida pelo método de *neighbor-joining* (Saitou & Nei, 1987), com a distância evolutiva sendo estimada pelo método de Jukes-Cantor (Jukes & Cantor, 1969). As seqüências foram tratadas como codificantes e todas as posições contendo *gaps* foram removidas (opção de deleção completa).

Alinhamentos adicionais das seqüências de aminoácidos foram produzidos para possibilitar a identificação e a sinalização de posições conservadas específicas, além de permitir que diferenças de interesse fossem evidenciadas.

A escolha das seqüências de bactérias a serem consideradas foi arbitrária, tomando as três ureases com estrutura tridimensional definida (de *Bacillus pasteurii*, *Klebsiella aerogenes* e *Helicobacter pylori*) e três ureases de espécies dos mesmos gêneros (de *Bacillus subtilis*, *Klebsiella pneumoniae* e *Helicobacter hepaticus*), com

o intuito de utilizá-las também como um controle interno das árvores filogenéticas (pois, em teoria, elas deveriam agrupar-se de maneira muito próxima).

3.2 Modelagem molecular comparativa

A isoforma JBURE-IIB da urease de *Canavalia ensiformis* foi escolhida como alvo para modelagem comparativa. Sua seqüência primária foi obtida de Mulinari (2008). Em uma abordagem inicial, verificou-se qual seria a sugestão de moldes fornecida pelo Swiss-Model (Guex & Peitsch, 1997) (<http://swissmodel.expasy.org>). Posteriormente, uma busca por seqüências similares à JBURE-IIB que tivessem estrutura tridimensional definida foi feita por BlastP (Altschul *et al.*, 1997), via NCBI (Wheeler *et al.*, 2008). Como os resultados do BlastP confirmaram moldes que haviam sido pré-selecionados (as únicas estruturas tridimensionais de ureases disponíveis, de *K. aerogenes*, *H. pylori* e *B. pasteurii*) como sendo os mais adequados, a estrutura do molde foi escolhida principalmente por sua resolução. A estrutura tridimensional do molde (2UBP) foi obtida do Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB, ou simplesmente PDB) (Berman *et al.*, 2000) (<http://www.rcsb.org/pdb>). Os três moldes selecionados originalmente apresentavam grau de identidade de seqüência com o alvo superior a 30% (~55%), valor adequado para uma modelagem bem sucedida (Eswar *et al.*, 2000).

O alinhamento das seqüências molde e alvo foi realizado pelo ClustalW (Larkin *et al.*, 2007). A modelagem molecular foi realizada por meio do programa Modeller 9v4 (Eswar *et al.*, 2000; Martí-Renom *et al.*, 2000), por meio de sua rotina “model-loop-define” (Apêndice A). Esta rotina é indicada para casos onde alças (ou outras regiões curtas) do modelo não apresentam equivalentes no molde, requerendo atenção especial. Este é exatamente o caso das duas regiões de ligação de domínios presentes nas ureases de plantas e fungos, que não ocorrem em bactérias. Foram construídos dez modelos, cada um deles com dez sub-modelos, representando possibilidades alternativas para as regiões sem molde.

Em um primeiro momento, os modelos foram inspecionados visualmente, sendo descartados aqueles que apresentavam padrões estruturais não realísticos, causados por modelagem inadequada de regiões sem molde maiores que quinze resíduos de aminoácidos (Šali, 2008). Na etapa seguinte, os modelos restantes

foram avaliados pelo mapa de Ramachandran (Ramachandran *et al.*, 1963)², gerado pelo Procheck (Laskowski *et al.*, 1993), e pelos resultados do Verify3D (Lüthy *et al.*, 1992), transpostos às estruturas tridimensionais por meio do GeneSilico Metaserver (Kurowski & Bujnicki, 2003) (<http://genesilico.pl>). Foi eleito como melhor modelo aquele que apresentou melhor avaliação pelo mapa de Ramachandran e organização espacial mais verossímil pelo Verify3D. A determinação da estrutura secundária a partir da estrutura terciária modelada foi realizada pelo PDBsum (Laskowski *et al.*, 2005) (<http://www.ebi.ac.uk/pdbsum>), enquanto a predição da estrutura secundária a partir da seqüência primária foi realizada pelo PSIPRED (McGuffin *et al.*, 2000) (<http://bioinf.cs.ucl.ac.uk/psipred>).

A visualização e a manipulação das estruturas tridimensionais das ureases foram realizadas nos programas Swiss-PDB Viewer versão 3.7 (Guex & Peitsch, 1997) e UCSF Chimera versão 1 (*build* 2540) (Pettersen *et al.*, 2004).

² Também chamado de Gráfico de Ramachandran, Ramachandran Plot, Mapa de Sasisekharan-Ramakrishnan-Ramachandran, ou simplesmente Ramachandran; em homenagem a seu propositor, Gopalsamudram Narayana Iyer Ramachandran.

4. Resultados e Discussão

4.1 Inferência filogenética

Após a exaustiva análise das 9046 seqüências de aminoácidos obtidas originalmente, foram selecionadas 237 seqüências de ureases de bactérias e 16 seqüências de ureases de plantas e fungos. Dentre as seqüências do primeiro grupo, foram escolhidas seis seqüências para as análises posteriores. A Tabela 1 apresenta a listagem detalhada das seqüências obtidas, enquanto o Apêndice B apresenta a listagem completa das 237 seqüências de ureases de bactérias.

Tabela 1: Seqüências de aminoácidos de ureases utilizadas neste trabalho.

Organismo	Código	GI da seqüência-tipo	Seqüência de DNA disponível	Vegetal, Fungo ou Bactéria
<i>Agaricus bisporus</i>	AGABIS	108859293	sim	F
<i>Ajellomyces capsulatus</i>	AJECAP	154281373	sim	F
<i>Arabidopsis thaliana</i>	ARATHA	15220459	sim	V
<i>Aspergillus clavatus</i>	ASPCLA	121702879	sim	F
<i>Aspergillus fumigatus</i>	ASPFUM	70990710	sim	F
<i>Aspergillus terreus</i>	ASPTER	114193850	não	F
<i>Bacillus pasteurii</i>	BACPAS	498708	sim	B
<i>Bacillus subtilis</i>	BACSUB	50812173	sim	B
<i>Canavalia ensiformis</i>	CANENS_1	167228	sim	V
<i>Canavalia ensiformis</i> (JBURE-IIB)	CANENS_JBURE-IIB	22831365	sim	V
<i>Coccidioides posadasii</i>	COCPPOS	60280325	sim	F
<i>Cryptococcus neoformans</i> var. <i>grubii</i>	CRYNGR	3688063	sim	F
<i>Cryptococcus neoformans</i> var. <i>neoformans</i>	CRYNNE	58270418	sim	F
<i>Glycine max</i> (embrião-específica)	GLYMAX_EMBR	32170829	sim	V
<i>Glycine max</i> (ubíqua)	GLYMAX_UBI	32170831	sim	V
<i>Helicobacter hepaticus</i>	HELHEP	32265499	sim	B
<i>Helicobacter pylori</i>	HELPHYL	209406223	sim	B
<i>Klebsiella aerogenes</i>	KLEAER	149335	sim	B
<i>Klebsiella pneumoniae</i>	KLEPNE	206575712	sim	B
<i>Neosartorya fischeri</i>	NEOFIS	119477773	sim	F
<i>Neurospora crassa</i>	NEUCRA	28926785	não	F
<i>Oryza sativa</i> (indica cultivar-group)	ORYSAT	17402589	sim	V
<i>Ostreococcus lucimarinus</i>	OSTLUC	144576703	não	P
<i>Schizosaccharomyces pombe</i>	SCHPOM	1944148	sim	F
<i>Solanum tuberosum</i>	SOLTUB	14599437	sim	V

Todas as seqüências foram utilizadas para construção de árvores filogenéticas, buscando evidenciar relações de similaridade presentes entre elas. É importante ressaltar que as árvores construídas representam relações de similaridade entre as seqüências, e não entre os organismos das quais elas provêm.

Em casos de reconstituição da filogenia de organismos, os genes de rRNA são os mais úteis: estão presentes em todos os organismos, e podem apresentar-se nas formas nuclear e mitocondrial. Genes de rRNA mitocondrial têm resolução na faixa de dez a cem milhões de anos, enquanto genes de rRNA nucleares têm resolução na faixa de centenas de milhões de anos (Ridley, 2004).

A reconstrução filogenética consiste em estimar as relações de ancestralidade para um determinado número de Unidades Taxonômicas Operacionais (OTUs), sejam populações, famílias, gêneros ou espécies, podendo basear-se em caracteres morfológicos, funcionais ou moleculares. Filogenias moleculares, com intenção de definir ancestralidade entre organismos baseando-se em seqüências de aminoácidos ou nucleotídeos, são consideradas casos especiais dentro do estudo filogenético tradicional (Miyaki *et al.*, 2001).

Uma árvore filogenética representa a história evolutiva das OTUs presentes nela, graficamente apresentada como pontos (ou nós) ligados por linhas (ou ramos). Os nós terminais representam as OTUs estudadas, enquanto os nós internos representam o ancestral comum mais recente dessas OTUs (Miyaki *et al.*, 2001). Nas árvores obtidas no presente trabalho, os nós internos representam a seqüência comum mais recente antes de ocorrer divergência entre as seqüências apresentadas nos nós terminais; ou seja, as seqüências individuais são tratadas como OTUs, sendo elas mesmas o alvo das análises, e não os organismos que as contém.

Existem três classes de técnicas para inferência de filogenias a partir de dados moleculares³. Em ordem crescente de complexidade, temos: distância molecular, parcimônia, e máxima verossimilhança (Nei & Kumar, 2000).

- distância molecular: é um método geométrico. Utiliza matrizes de diferenças e similaridades entre as seqüências, sendo consideradas mais distantes evolutivamente aquelas seqüências que apresentarem o maior número de diferenças entre si. Este é considerado um método rápido e menos preciso, pois pressupõe um relógio molecular: as taxas evolutivas são consideradas constantes entre todas as seqüências analisadas (Ridley, 2004). Essa classe inclui os métodos

³ Uma quarta classe, proposta em 1996, baseia-se em redes neurais. Sua utilidade prática ainda não foi evidenciada (Nei & Kumar, 2000).

de junção com vizinhos (*neighbor-joining*, NJ), UPGMA (*unweighted pair group method with arithmetic means*)⁴ e evolução mínima.

- parcimônia: é um método probabilístico. Corresponde ao princípio de que a filogenia que exige o menor número de mudanças evolutivas para obter a diferenciação das OTUs analisadas é a melhor estimativa da filogenia real. De maneira simplificada, geram-se todas as árvores possíveis e, entre elas, selecionam-se aquelas que apresentam a menor quantidade de mudanças evolutivas. O método está de acordo com a observação de que uma mudança evolutiva é um evento raro (Myaki *et al.*, 2001).
- máxima verossimilhança: é um método probabilístico. Consiste em, usando um modelo de evolução de seqüências pré-estabelecido, calcular a probabilidade de que os dados analisados tenham sido originados conforme o modelo evolutivo adotado (Pereira *et al.*, 2001).

No presente trabalho, optou-se por utilizar o método de *neighbor-joining* para a construção da árvore filogenética das seqüências de aminoácidos, sendo essa opção baseada principalmente no estudo comparativo de Hasegawa & Fujiwara (1993). Tal estudo indica o *neighbor-joining* como método confiável para obtenção da filogenia, mesmo quando as taxas de variação evolutiva não são totalmente uniformes entre as linhagens (o que seria pré-requisito para aplicação do método). Para a construção da árvore filogenética das seqüências de nucleotídeos, optou-se também por utilizar o método de *neighbor-joining*. Essa escolha baseou-se no estudo comparativo de Saitou & Imanishi (1989) que indica o *neighbor-joining* como o método de escolha quando aspectos práticos são considerados, mas sugere que abordagens mais complexas sejam escolhidas quando possível (isto é, quando a capacidade computacional não for uma restrição e quando o operador possuir conhecimento avançado das teorias subjacentes ao método).

O uso dos métodos de distância, parcimônia e de máxima verossimilhança, nessa ordem, exigem quantidades de dados e poder computacional cada vez maiores. Essa ordem reflete a seqüência histórica do uso dos métodos de pesquisa filogenética molecular (Ridley, 2004). Atualmente existe muito debate sobre as eficiências relativas dos diferentes métodos de construção de árvores, especialmente *neighbor-joining*, parcimônia e máxima verossimilhança. Está claro

⁴ Raramente definido por extenso, o nome do método pode ser traduzido por “método não-pesado de agrupamento par a par usando médias aritméticas”.

que nenhum método é superior aos outros em todas as condições (Nei & Kumar, 2000). Em estudos com dados reais (e não simulações baseadas em reconstrução de uma filogenia já conhecida), com uso de um número substancial de seqüências pouco divergentes, os três métodos produzem topologias similares ou idênticas. Sugere-se, portanto, que árvores que apresentem valores adequados de *bootstrap*, construídas por meio de qualquer um dos métodos, fornecem essencialmente a mesma conclusão sobre as relações filogenéticas dos organismos, genes ou proteínas (Nei & Kumar, 2000).

A árvore construída a partir das seqüências de aminoácidos das ureases de diferentes organismos é apresentada na Figura 3. Observa-se na árvore um claro agrupamento das seqüências de plantas, fungos e bactérias entre si; com a alga unicelular *Ostreococcus lucimarinus* agrupando-se junto às plantas. Além disso, a árvore indica que as seqüências de aminoácidos de plantas e bactérias são mais semelhantes entre si do que semelhantes às seqüências de fungos. Tal similaridade pode indicar que algumas (ou muitas) das propriedades não-catalíticas dessas enzimas sejam comuns a ureases de plantas e bactérias, mas encontrem-se ausentes ou diferenciadas em ureases fúngicas. Estes resultados poderão guiar futuros estudos focados em ureases, além de auxiliar a compreensão de resultados obtidos experimentalmente. A adição das seqüências de bactérias no cálculo das árvores filogenéticas forçou a comparação a restringir-se à região equivalente em todas as seqüências. Optou-se pela região correspondente à subunidade C de bactérias (domínio α em plantas e fungos), por ser esta a maior subunidade e, principalmente, por conter o sítio catalítico das ureases. Por precaução, árvores sem a inclusão das seqüências de bactérias foram calculadas, e verificou-se que a inclusão das seqüências de ureases não alterou o padrão geral de distribuição dos ramos.

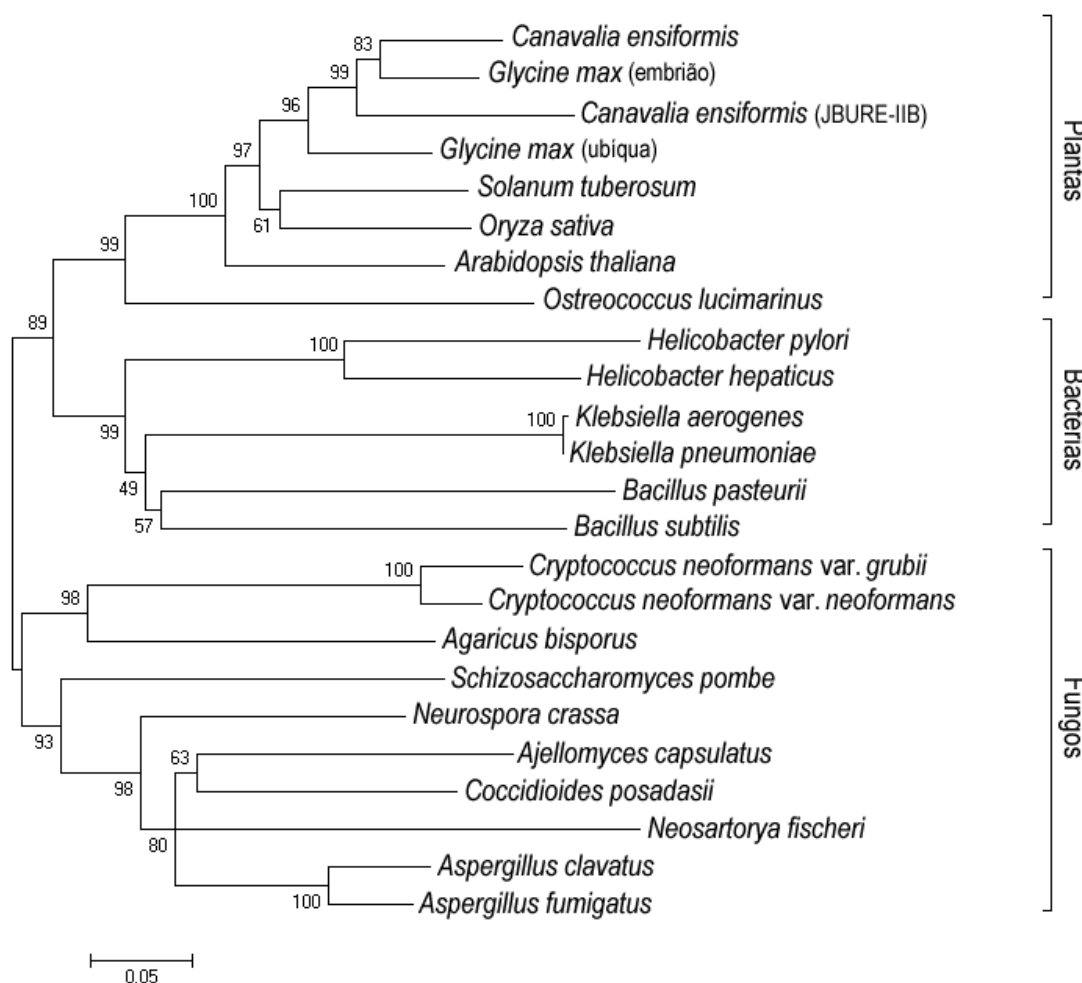


Figura 3: Relação evolutiva entre 24 seqüências de aminoácidos de ureases. A barra de escala representa número de substituições de aminoácidos por sítio. Comparação realizada utilizando-se as seqüências completas de ureases de plantas e fungos e as seqüências das cadeias C de bactérias (correspondentes ao domínio α de plantas e fungos). A árvore apresentada é o consenso, obtido após 10000 replicatas de *bootstrap*, e está representando a história evolutiva das seqüências analisadas. A percentagem de árvores em que as seqüências associadas se agruparam é apresentada próxima às ramificações. A árvore está representada em escala, com o comprimento dos ramos nas mesmas unidades das distâncias evolutivas usadas para inferir a árvore filogenética (número de substituições de aminoácidos por sítio). Os grupos taxonômicos dos organismos cujas ureases foram comparadas encontram-se à direita da árvore.

A árvore construída a partir das seqüências de nucleotídeos das ureases de diferentes organismos é apresentada na [Figura 4](#). Esta árvore também apresenta um agrupamento das seqüências de plantas, fungos e bactérias entre si, mas com algumas peculiaridades. Diferentemente da árvore de aminoácidos, esta indica que as seqüências de nucleotídeos de plantas e fungos são mais semelhantes entre si do que semelhantes às seqüências de bactérias. Além disso, há o agrupamento de três seqüências de ureases fúngicas junto às ureases de plantas. A seqüência de nucleotídeos correspondente à urease da alga *O. lucimarinus* não encontrava-se disponível e por este motivo não foi incluída na construção desta árvore. A

comparação das seqüências de nucleotídeos não é tão relevante para estudos estruturais quanto a comparação das seqüências de aminoácidos, mas é interessante para aplicações que envolvam o próprio material genético. Os dados de similaridade obtidos podem ser úteis, por exemplo, ao prever-se o comportamento de determinado fragmento de DNA ao ser inserido em um plasmídeo em estudos de clonagem de genes de urease.

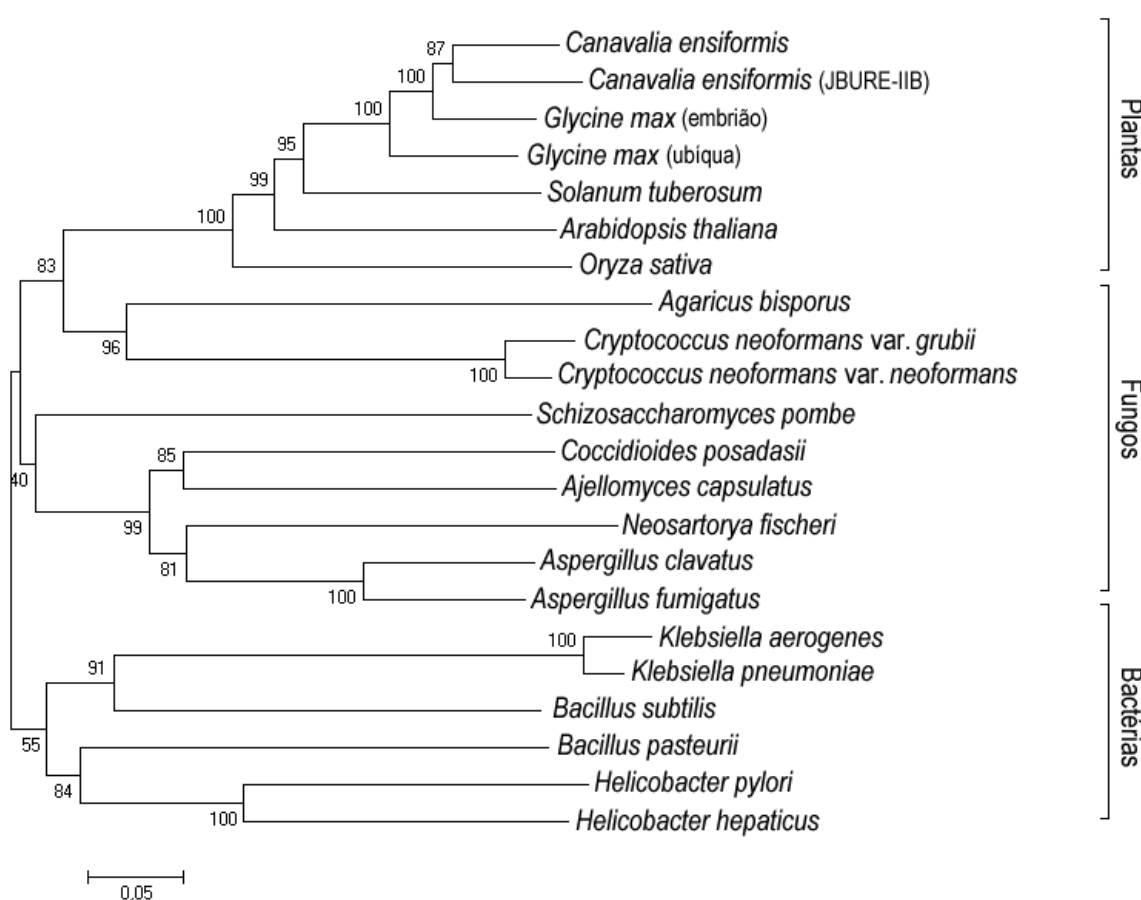


Figura 4: Relação evolutiva entre 22 seqüências de nucleotídeos de ureases. A barra de escala representa número de substituições de aminoácidos por sítio. Comparação realizada utilizando-se as seqüências completas de ureases de plantas e fungos e as seqüências das cadeias C de bactérias (correspondentes ao domínio α de plantas e fungos). A árvore apresentada é o consenso, obtido após 10000 replicatas de *bootstrap*, e está representando a história evolutiva das seqüências analisadas. A percentagem de árvores em que as seqüências associadas se agruparam é apresentada próxima às ramificações. A árvore está representada em escala, com o comprimento dos ramos nas mesmas unidades das distâncias evolutivas usadas para inferir a árvore filogenética (número de substituições de bases por sítio).

Foi observada uma diferença entre as árvores construídas a partir de aminoácidos e nucleotídeos. Entretanto, as duas árvores não deveriam, necessariamente, ser idênticas. As classes de dados utilizadas são diferentes, e teoricamente uma seqüência de DNA apresenta maior probabilidade de diferenciar-

se do ancestral do que uma seqüência de aminoácidos. Isso se deve especialmente à degeneração do código genético (Ridley, 2004). A árvore construída a partir de seqüências de nucleotídeos tem, portanto, um caráter complementar ao restante das análises realizadas.

Adicionalmente, uma árvore considerando apenas as seqüências de aminoácidos correspondentes ao peptídeo entomotóxico derivado de JBURE-IIB foi construída. Surpreendentemente, esta árvore refletia a topologia da árvore construída para as seqüências de nucleotídeos, agrupando os fungos *A. bisporus*, *C. neoformans v. neoformans* e *C. neoformans v. grubii* junto às seqüências oriundas de plantas. Tal agrupamento poderia ser uma indicação de que as ureases destes fungos apresentam propriedades tóxicas semelhantes às das ureases vegetais. Devido ao baixo suporte estatístico desta filogenia (~30% em algumas ramificações), tal hipótese permanece apenas em nível especulativo.

Como na imensa maioria dos casos de análise filogenética não se conhece a verdadeira topologia da árvore filogenética em análise, a confiabilidade na topologia obtida é geralmente testada ao examinar-se o grau estatístico de confiança no padrão de ramificação da árvore obtida (Nei & Kumar, 2000). A confiabilidade nas árvores construídas neste trabalho foi testada pelo método de *bootstrap*.

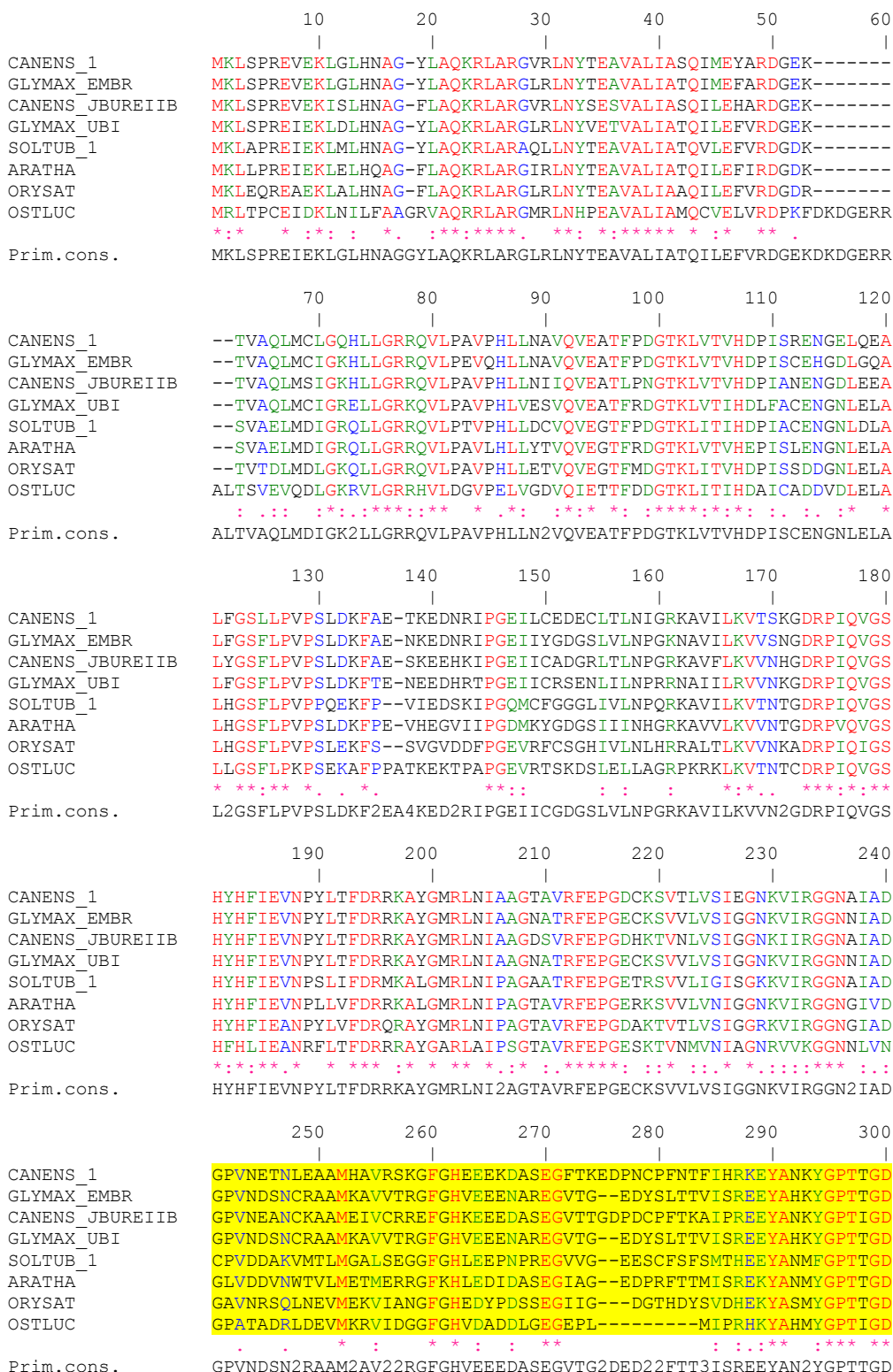
O *bootstrap* (Felsenstein, 1985) consiste em uma reamostragem com reposição pseudoaleatória dos dados. Em cada reamostragem, o número de dados amostrados mantém-se constante, e cada posição tem a mesma chance de ser amostrada; conseqüentemente, algumas posições serão contadas mais de uma vez, enquanto algumas não serão amostradas. A cada reamostragem, uma árvore réplica é construída (Russo *et al.*, 2000). Nas árvores apresentadas nas Figuras 3 e 4, a percentagem com que cada agrupamento ocorreu nas dez mil réplicas aparece como índices sobre os pontos de ramificação. A confiança do *bootstrap*, entretanto, depende da consistência do método usado para construir-se a árvore. Se o método não for adequado para uma dada faixa de parâmetros, os valores de *bootstrap* podem ser altos sem que a topologia esteja correta. No caso oposto, com parâmetros adequados ao método, o *bootstrap* é uma subestimativa da probabilidade daquele ramo ser verdadeiro (Russo *et al.*, 2001).

4.2 Comparação de seqüências primárias

Por meio de alinhamentos para cada um dos grupos de seqüências de aminoácidos de ureases (vegetais, fúngicas e bacterianas) foi confirmada a conservação dos seis resíduos considerados essenciais para a catálise em ureases (conforme Mobley *et al.*, 1995) em todas as seqüências analisadas. Estes resíduos encontram-se destacados em verde nas Figuras 5, 6 e 7, com indicações sobre sua localização na seqüência primária da JBURE-IIB, e correspondem ao resíduos His134, His136, Lys217, His246, His272 e Asp360 da urease de *K. aerogenes*. A conservação destes resíduos em todas as seqüências analisadas é um forte indicativo de que eles exerçam a mesma função em todas as ureases.

A região correspondente ao peptídeo entomotóxico de *C. ensiformis* foi localizada e analisada nas seqüências de plantas e fungos (a região é inexistente em ureases bacterianas). Verificou-se a baixa identidade deste trecho das seqüências (28,26% para seqüências de plantas e 19,19% para seqüências de fungos, comparadas junto às de *C. ensiformis*) e seu alto grau de diferença (51,09% para plantas e 62,63% para fungos, comparadas junto às de *C. ensiformis*). A região do peptídeo encontra-se destacada em amarelo nas Figuras 5 e 6, e os dados referentes a seus alinhamentos, realizados independentemente dos trechos adjacentes, encontram-se junto aos dados dos alinhamentos totais. O alto grau de diferença desse segmento entre as ureases, especialmente ao comparar-se ureases de plantas e fungos, pode indicar que nem todos os peptídeos devam, necessariamente, apresentar propriedades inseticidas.

Figura 5: Alinhamento múltiplo de seqüências de aminoácidos de ureases oriundas de plantas. Códigos dos organismos de acordo com a Tabela 1.



```

          310      320      330      340      350      360
CANENS_1      KIRLGDTNLLAEIEKDYALYGDCEVFGGGKVI RDGMGQSCGHPPAISLDTVITNAVIIDY
GLYMAX_EMBR   KIRLGDTDLFAEIEKDFALYGDCEVFGGGKVI RDGMGQSCGDPPAISLDTVITNAVIIDY
CANENS_JBUREIIB KIRLGDTDLIAEIEKDFALYGDCEVFGGGKVI RDGMGQSSGHPPPAMSLDTVITSAVIIDY
GLYMAX_UBI    KIRLGDTDLFAEIEKDFAVYGDCEVFGGGKVI RDGMGQSSGHPPEGLDTVITNAVIIDY
SOLTUB_1     RIRLGDTDLFAEIEKDFGIFGDCEVFGGGKVI RDGMGQACGYPPADCLDTVITNAVIIDY
ARATHA       KIRLGDTNLYARIEKDYTVYGDCEVFGGGKVI RDGMGQGEQAEALS LDTVITNSVIIDY
ORYSAT       KIRLGDTDLFAEIEKDYAIYGDCEVFGGGKVI RDGMGQSAGYPASDCLDTVVITNAVIIDY
OSTLUC       RVRLGDTNLYITPERDLTMKGEESKFGGGKTL RDGMSQQAGVGDADSLDTIITNALIVDH
:****:*      *:*      :*:*      *****:*:*:*      *****:*:*:*
Prim.cons.    KIRLGDTDLFAEIEKDFALYGDCEVFGGGKVI RDGMGQSCGHPPADSLDTVITNAVIIDY

```

```

          370      380      390      400      410      420
CANENS_1      TGIKADIGIKDGLIASIGKAGNPDIMNGV--FSNMIIGANTEVIAGEGLIVTAGAIDC His407
GLYMAX_EMBR   SGIKADIGIKDGLIVSIGKAGNPDIMDDV--FFNMIIGANTEVIAGEGLIVTAGAIDC
CANENS_JBUREIIB TGIKADIGIKDGLIASIGKAGNPDIMNGV--FPNMIIGVNTTEVICGEGLIVTAGGIDC
GLYMAX_UBI    TGIKADIGIKDGLIISTGKAGNPDIMNDV--FPNMIIGANTEVIAGEGLIVTAGAIDC
SOLTUB_1     TGIKADIGIKDGHIVSLCKAGNPDIMD----SDAIIIGVNTTEVIAGEGMIVTAGAIDC
ARATHA       SGIYKADIGIKNGHIVGIGKAGNPDIMHGV--QNNMLIGNKTEVIAGEGMIVTAGAIDC
ORYSAT       TGIYKADIGINGGLIVAIIGKAGNPDVMDMGVNEEMIVGVNTTEVIAAEGMIVTAGGIDC
OSTLUC       SGIYKADIGIKDGHIVGIGHGNPDVADV---PGMVVGVNTEAIAAEGCIVTAGALDIT
:*:*:*:*:*   *:*:*   :*:*   :*:*   :*:*   :*:*   :*:*   :*:*   :*:*   :*:*
Prim.cons.    TGIKADIGIKDGLIVSIGKAGNPDIMDGVGVFPNMIIGVNTTEVIAGEGLIVTAGAIDC

```

```

          430      440      450      460      470      480
CANENS_1      VYICPQLVYEAISSGITTLVGGGTGPAAGTRATTCTPSPQMRLMLQSTDDLPLNFGFT His409
GLYMAX_EMBR   VYICPQLVDEAISSGITTLVGGGTGPTAGTRATTCTPAPSQMKMLQSTDDLPLNFGFT
CANENS_JBUREIIB VYICPQSLDEAISSGITTVVGGGTGPTDGSRATTCTPAPTQMKMLQSTDDIPLNFGFT
GLYMAX_UBI    VYICPQLVYDAVTSGITTLVGGGTGPAADGTRATTCTPAPNQMKMLQSTDDMPLNFGFT
SOLTUB_1     VYFICPQLAYEAISSGITTMVGGGTGPAHGTTRATTCTPGHVHMLMLQSTDEIPLNFGFT
ARATHA       VYFICPQLVYEAVSSGITTMVGGGTGPAYGTRATTCTPSPFDMKMLQSTDSLPLNFGFT
ORYSAT       VYFICPQLAEEAIASGITTLVGGGTGPAHGTCATTCTPSPSHMKMLQSTDELPIINMGFT
OSTLUC       VYFICPQLCTEAVASGITTLVGGGTGPAASGTCATTCTPAAHMQFMLETTDALPLNFAFT
:*:*:*:*   :*:*:*:*   :*:*:*:*   :*:*   :*:*   :*:*   :*:*   :*:*   :*:*
Prim.cons.    VH2ICPQLVYEAISSGITTLVGGGTGPA3GTRATTCTPSP2QMKMLQSTDDLPLNFGFT

```

```

          490      500      510      520      530      540
CANENS_1      GKGSSSKPDELHEI IKAGAMGLLHEDWGSTPAAIDNCLTIAEHHDIQINI TDTLNEAG Lys490 His519
GLYMAX_EMBR   GKGSSSKPDELHDI IKAGAMGLLHEDWGSTPAAIDSCLTVADQYDIQINI TDTLNEAG
CANENS_JBUREIIB GKGSGSHPDELHEI IKAGAMGLLHEDWGSTPAAIDNCLVAEQHDIQINI TDTVNESG
GLYMAX_UBI    GKGNSAKPDELHEI IRAGAMGLLHEDWGSTPAAIDSCLTVADQYDIQINI TDTLNEAG
SOLTUB_1     GKGNSSKADGLHEI IKAGAMGLLHEDWGSTPAAIDMCLTVADQYDIQINI TDTLNEAG
ARATHA       GKGNTAKPLELRHIVEAGAMGLLHEDWGSTPAAIDNCLVAEEYDIQINI TDTLNEAG
ORYSAT       GKGNTTKPDGLAEI IKAGAMGLLHEDWGSTPAAIDNCLVAEAFDIQINI TDTLNEAG
OSTLUC       GKGNTASPEGLHEI IKAGAVGMLHEDWGSTPAAIDNCLTIAEEYDVAVTI TDTLNEAG
**:*   :   *   :*:*:*:*   :*:*:*   :*:*   :*:*   :*:*   :*:*
Prim.cons.    GKGNSSKPDELHEI IKAGAMGLLHEDWGSTPAAIDNCLTVAEQYDIQINIHTDTLNEAG

```

```

          550      560      570      580      590      600
CANENS_1      FVEHSIAAFKGRTIHTY SEGAGGGHAPDIKVCGIKNVLPSSSTNPTREPLTNSNTIDEHL D His545
GLYMAX_EMBR   FVEHSIAAFKGRTIHTY SEGAGGGHAPDIKVCGMKNVLPSSSTNPTREPLTNTIDEHL
CANENS_JBUREIIB FVEHTIAAFNGRTIHTY SEGAGGGHAPDIKVCSMKNVLPSSSTNPTREPLTNSNTIDEHL
GLYMAX_UBI    FVEHTIAAFKGRTIHTY SEGAGGGHAPDIKVCGEKNVLPSSSTNPTREPTHTNTIDEHL
SOLTUB_1     FVEHTIAAFKGRTIHTY SEGAGGGHAPDIKVCVKNVLPSSSTNPTREPTLNTIDEHL
ARATHA       FVEHTINAFRGRTIHTY SEGAGGGHAPDIRVCGVKNVLPSSSTNPTREPTKNTIDEHL
ORYSAT       CVEHTIAAFKDRTIHTY SEGAGGGHAPDIKVCVKNVLPSSSTNPTREPTLNTIDEHL
OSTLUC       CVEKSI EAFKGRTIHTY SEGAGGGHAPDIKVCGEKNVLPSSSTNPTREPTKNTIDEHL
:*:*:*   *:*   :*:*:*   :*:*:*   :*:*   :*:*   :*:*   :*:*
Prim.cons.    FVEHTIAAFKGRTIHTYSEGAGGGHAPDIKVCVKNVLPSSSTNPTREPT2LNTIDEHL

```

```

          610      620      630      640      650      660
CANENS_1      MLMVCHHLDR EIPEDLAF AHSRIRKKTIAAEDVLNDIGAI S I S S S QAMGRVGEVISRT Asp633
GLYMAX_EMBR   MLMVCHHLNREIPEDLAFACSRIREGTAAEDILHDI GAI S I S S S QAMGRVGEVISRT
CANENS_JBUREIIB MLMVCHKLNREIPEDLAFASSRVREQTAAEDILHDI GGI S I S S S AQAVGRIGEVISCT
GLYMAX_UBI    MLMVCHHLNKNIPEDVAF AFSRIRAETIAAEDILHDKGAI S I S S S QAMGRIGEVISRT
SOLTUB_1     MLMVCHHLCKNSREDVAF AFSRIRAETIAAEDILHDMGAI S I S S S QAMGRIGEVICRT
ARATHA       MLMVCHHLDKNIPEDVAF AFSRIRAETIAAEDILHDMGAI S I S S S QAMGRIGEVISRT
ORYSAT       MLMVCHHLDRNIPEDVAF AFSRIRAETIAAEDILHDMGAI S I S S S QAMGRIGEVITRT
OSTLUC       MLMVCHHLDP EIPEDVAF AFSRIRAETIAAEDVLHDMGALS I M A S S QAMGRVGEVIRRT
*****:*   :*:*:*   :*:*:*   :*:*   :*:*   :*:*   :*:*   :*:*
Prim.cons.    MLMVCHHLDR2IPEDVAF AFSRIRAETIAAEDILHDMGAI S I S S S QAMGRIGEVISRT

```



```

          670      680      690      700      710      720
          |        |        |        |        |        |
CANENS_1  WQTADKMK AQTGPLKCD-----SSDNDNFRIRRYIAKYTINPAIANGFSQYVGSVEVGKL
GLYMAX_EMBR WQTANKMKVQRGPLQPG-----ESDNDNFRIRRYIAKYTINPAIANGFSQYVGSVEVGKL
CANENS_JBUREIIB WQTADKMK AERGPLQPD-----GSDNDNFRIRRYIAKYTINPAIANGISQYVGSVEVGKL
GLYMAX_UBI  WQTADKMK SQRGPLQP-----GEDNDNFRIRRYIAKYTINPAIANGLSQYVGSVEAGKL
SOLTUB_1   WQTAHKMK LFRGPLDID-----GSDNDNFRIRRYIAKYTINPAIANGISQYVGSVEVGKL
ARATHA     WQTADKMK AQRGAIDPN-----MADDNFRIRRYIAKYTINPAIANGFADLIGSVEVKKL
ORYSAT     WQTANKMK RQRGRPLPISSPDAAEDNDNFRIRRYIAKYTINPAIANGFSDFVGSVEVGKL
OSTLUC     WQTAHSNK EQRGFLEEDAN---SGADNVRVKRYIAKYTINPAIAHGMSHKVGSLEVGKF
          * * * * * : * * * * * * * * * * * * * * * * * * * * * *
Prim.cons. WQTADKMK AQRGPLQPD22PDAGSDNDNFRIRRYIAKYTINPAIANGFSQYVGSVEVGKL

```

```

          730      740      750      760      770      780
          |        |        |        |        |        |
CANENS_1  ADLVMWKPSFFGTPKPEMVIKGGMVAWADIGDPNASIPTPEPVKMRPMYGTLGKAGGALS I
GLYMAX_EMBR ADLVMWKPSFFGAKPEMVIKGGVVAWADMGDPNASIPTPEPVKMRPMYGTLGKAGGALS I
CANENS_JBUREIIB ADLVIWKPSFFGAKPDIKGGSIWADMGDPNGSIPTPEPVLMRPMYGTLGKAGSALS I
GLYMAX_UBI  ADLVVWKPSFFGAKPEMVIKGGEVAYANMGDPNASIPTPEPVIMRPMYGTLGKAGSALS I
SOLTUB_1   ADLVVWKPSFFGAKPEMVIKGGVIAWSNMGDPNASIPTPEPVIMRPMYGTLGKAGSALS I
ARATHA     ADLVIWQPAFFGAKPEMIKGGNIWANMGDANASIPTPEPVISRPMYGTLGKAGSENSV
ORYSAT     ADLVIWKPSFFGAKPEMVIKGGAIACANMGDPNASIPTPEPVIMRPMYGTLGKAGSALS I
OSTLUC     ADVVIWKPAFFGAKPEIVVKGQIAWAQMGDPNASIPTPEPVIMRPMYGTLGKAGSALS I
          * * * * * : * * * * * * * * * * * * * * * * * * * * * *
Prim.cons. ADLVIWKPSFFGAKPEMVIKGGVIAWANMGDPNASIPTPEPVIMRPMYGTLGKAGSALS I

```

```

          790      800      810      820      830      840
          |        |        |        |        |        |
CANENS_1  AFVSKAALDQRVNVLVYGLNKRVEAVSNVRKLTKLDMLNDALPEITVDPEPTYTVKADGKL
GLYMAX_EMBR AFVSKAALDQRVHVALYGLNKRVEAVGNVRKLTKLDMLNDALPEITVDPEPTYTVTADGEV
CANENS_JBUREIIB AFVSKAALDLGVKVLVYGLNKRVEAVSNVRKLTKLDLKLNSLPEITVCPETFTVTVDDGQA
GLYMAX_UBI  AFVSKAALDEGVKASYGLNKRVEAVKVRKLTKRDMLNDLTLPEITVDPEPTYTVTADGEV
SOLTUB_1   AFVSKAALDAGIKDSYRLNKRVEAVTNVRNISKLDMLNDALPEIKVDPEPTYTVTADGTA
ARATHA     AFVSKAALRKGKVELYGLKRRVAVSNVRQLTKLDMLNDALPEITVDPEPTYTVTADGEV
ORYSAT     AFVSKAAKEAGVAVQYKLGKRVEAVGRVRLTKLNMKLNLDALPKIDVDPEPTYTVTADGEV
OSTLUC     AFVSAAAAADVGAEYGLNKRVEAVVKCRGLSKDDMLNACPRIEVDPEPTYTVRADGVV
          * * * * * : * * * * * * * * * * * * * * * * * * * * * *
Prim.cons. AFVSKAALDAGVK2LYGLNKRVEAVSNVRKLTKLDMLNDALPEITVDPEPTYTVTADGEV

```

```

          850
          |
CANENS_1  LCVSEATTVPLSRNYFLF
GLYMAX_EMBR LTSFATTFVPLSRNYFLF
CANENS_JBUREIIB LSSEAVTTLPLSQNYFIF
GLYMAX_UBI  LTCCTAAKTVPPLSRNYFLF
SOLTUB_1   LTCPPATTVPLSRNYFLF
ARATHA     LTCAPADSVPLSRNYFLF
ORYSAT     LRCQPPTVPPLSRNYFLF
OSTLUC     LKSQPAQELPLARRYFIV
          * * * * * : * * * * * * * * * * * * * * * * * * * * * *
Prim.cons. LTCQPATTVPLSRNYFLF

```

Dados do alinhamento das seqüências completas:

Comprimento do alinhamento: 858
 Idênticos (*): 376 ou 43.82 %
 Fortemente similares (:): 166 ou 19.35 %
 Fracamente similares (.): 67 ou 7.81 %
 Diferentes: 249 is 29.02 %
 Seqüência 0001: CANENS_1 (840 resíduos).
 Seqüência 0002: GLYMAX_EMBR (838 resíduos).
 Seqüência 0003: CANENS_JBUREIIB (840 resíduos).
 Seqüência 0004 : GLYMAX_UBI (837 resíduos).
 Seqüência 0005 : SOLTUB_1 (834 resíduos).
 Seqüência 0006 : ARATHA (838 resíduos).
 Seqüência 0007 : ORYSAT (843 resíduos).
 Seqüência 0008 : OSTLUC (840 resíduos).

Dados do alinhamento da região do

peptídeo, destacada em amarelo:
 Comprimento do alinhamento : 92
 Idênticos (*) : 26 ou 28.26 %
 Fortemente similares (:): 14 ou 15.22 %
 Fracamente similares (.): 5 ou 5.43 %
 Diferentes : 47 ou 51.09 %

Figura 6: Alinhamento múltiplo de seqüências de aminoácidos de ureases oriundas de fungos e de *C. ensiformis*, para comparação. Códigos dos organismos de acordo com a Tabela 1.

```

          10      20      30      40      50      60
          |      |      |      |      |      |
ASPCLA -----
ASPFUM -----
ASPTER MPPITKPFFAFFFAFFAILVLA TLMLQLQA AVAEVSKAYINHSNTVLRGSSDLEDGPLANIAN
NEUCRA -----
AJECAP -----
COCPOS -----
NEOFIS -----
SCHPOM -----
CRYNGR -----
CRYNNE -----
AGABIS -----
CANENS_1 -----
CANENS_JBUREIIB -----

Prim.cons. MPPITKPFFAFFFAFFAILVLA TLMLQLQA AVAEVSKAYINHSNTVLRGSSDLEDGPLANIAN

          70      80      90      100     110     120
          |      |      |      |      |      |
ASPCLA -----
ASPFUM -----
ASPTER LLQENGLFGFPFGGARATSPGAKSEVGEVGGKKRKRAPHDPNAPKRALT PFFFLYMQHN
NEUCRA -----
AJECAP -----
COCPOS -----
NEOFIS -----
SCHPOM -----
CRYNGR -----
CRYNNE -----
AGABIS -----
CANENS_1 -----
CANENS_JBUREIIB -----

Prim.cons. LLQENGLFGFPFGGARATSPGAKSEVGEVGGKKRKRAPHDPNAPKRALT PFFFLYMQHN

          130     140     150     160     170     180
          |      |      |      |      |      |
ASPCLA -----
ASPFUM -----
ASPTER RAKISAE LGS GARPKEVSDEGTRRWQTMPQAERDVWKQMYAENLALYKERMRLYKAGLPI
NEUCRA -----
AJECAP -----
COCPOS -----
NEOFIS -----
SCHPOM -----
CRYNGR -----
CRYNNE -----
AGABIS -----
CANENS_1 -----
CANENS_JBUREIIB -----

Prim.cons. RAKISAE LGS GARPKEVSDEGTRRWQTMPQAERDVWKQMYAENLALYKERMRLYKAGLPI

          190     200     210     220     230     240
          |      |      |      |      |      |
ASPCLA -----
ASPFUM -----
ASPTER TDDARAANQLQQEAEATNAEESDETGDDDLAAEPPREPTPPRSGKRRRSEVKSTPKQAE T
NEUCRA -----
AJECAP -----
COCPOS -----
NEOFIS -----
SCHPOM -----
CRYNGR -----
CRYNNE -----
AGABIS -----
CANENS_1 -----
CANENS_JBUREIIB -----

Prim.cons. TDDARAANQLQQEAEATNAEESDETGDDDLAAEPPREPTPPRSGKRRRSEVKSTPKQAE T

```

	250	260	270	280	290	300
ASPCLA	-----					
ASPFUM	-----					
ASPTER	PKIPAPEAKTPAEKKASPEKKASPEKRRGPKAAAAAQAQDEASTPAAAGRKS LGGGEAPK					
NEUCRA	-----					
AJECAP	-----MKPRLETSQEFLEGRNKQHFLAFTGDPLQWR					
COCPOS	-----					
NEOFIS	-----					
SCHPOM	-----					
CRYNGR	-----					
CRYNNE	-----					
AGABIS	-----					
CANENS_1	-----					
CANENS_JBUREIIB	-----					

Prim.cons. PKIPAPEAKTPAEKKASPEKKASPEKRRR22222222Q2222222222222222G222222

	310	320	330	340	350	360
ASPCLA	-----					
ASPFUM	-----MHLIPKEVR					
ASPTER	ETTKRAKKRRLYYGVMKSGVISPPAANQRPNQPYQLISAVWPTNPPSFLFSLIPVAGPL					
NEUCRA	-----					
AJECAP	PEDREKAKGLLEESLLNQLRPVTLASNRIDDRQEC SRRPARMIFRNVPVPDQSAFQRG					
COCPOS	-----					
NEOFIS	-----					
SCHPOM	-----					
CRYNGR	-----					
CRYNNE	-----					
AGABIS	-----					
CANENS_1	-----					
CANENS_JBUREIIB	-----					

Prim.cons. 2222222K22L2223333333333

	370	380	390	400	410	420
ASPCLA	-----					
ASPFUM	DRDLLPTTYIHIIHISKVVSTNHFAQAARHLAIGISRPASSCARCAAQSCRSGMDFP					
ASPTER	AVVCGVRRLLLLLPLSLSLFRFRLLRFFFS S S P L L L E P Q P C R V R G D I W F L A T F L P Q					
NEUCRA	-----					
AJECAP	PVDCSVPRAFTCNDLGT AQSDAKPCKATRYPKLPPPHFACSI PAAF GH W S S L A T S L L N P D					
COCPOS	-----					
NEOFIS	-----					
SCHPOM	-----					
CRYNGR	-----					
CRYNNE	-----					
AGABIS	-----					
CANENS_1	-----					
CANENS_JBUREIIB	-----					

Prim.cons. 3VDC3V3R333333L333S333F333A33333L333333333333AR3333S333S333P3

	430	440	450	460	470	480
ASPCLA	-MHLVPKELDKLVISQLGFLAQRRLARGVRLNHAEAVALI SSNIHELIRDGHYSVADLMS					
ASPFUM	LFSLRDPLRRLIEANSRPGSHLVESARGIFPTSELEAGYHGAYSQQLIRDGHYSVADLMS					
ASPTER	DRTLTLRLDKLAISQLGFLAQRRLARGVRLNHAEAVALI SSNLQELIRDGHYSVADLMS					
NEUCRA	-MHLIPKELDKLAISQLGFLAQRRLARGIKLNHSEATALIANNLQELIRDGNHTVADLMA					
AJECAP	QMQLVPRELDKLVTSQLGFLAQRRLARGVRLNHAEATALIASNIQELIRDGNHTVADLMA					
COCPOS	-MQLVPREIDKLTISNLGFLAQRRLARGVRLNHAEATALIAS-NLQELIRDGNNSVADLMT					
NEOFIS	-MHLIPRELDKLTIAHFGNLAQRRLSRGIRLNHAEAAALISSLHHELIRDGHYAVADLMA					
SCHPOM	---MQPRELHKLT LHQLGSLAQRRLCRGVKLNKLEATS LIASQIQEYV RDGNHSVADLMS					
CRYNGR	-MHLLPRETDKLI LTTLTGTLAQRRLARGLILNRAETIALI SSQIQEFV RDGRH SVAELMD					
CRYNNE	-MHLLPRETDKLI VTTLTGTLAQRRLARGLILNRAETIALI SSQIQEFIRDGRH SVAELMD					
AGABIS	-MRLLPREEAKVLLHQVGFI AQRLARGVRLNHAEAVALI SSVLERIRDGRH SVAELMQ					
CANENS_1	-MKLSPREVEKLGHNAGYLAQRLARGVRLN YTEAVALI ASQIMEYARDGEKTVAQLMC					
CANENS_JBUREIIB	-MKLSPREVEKISLHNAGFLAQRRLARGVRLN YSESVALI ASQILEHARDGEKTVAQLMS					
	:	:	.*.*:	.	.	: **.*:***.*
Prim.cons.	3MHL2PRELDKLT2SQLGFLAQRRLARGVRLNHAEAVALI ASQIQELIRDGHHSVADLMS					

```

          490      500      510      520      530      540
ASPCLA      IGKTMLRRRHVLPSVVHSLVELQVEGTFLTGTYLVTVVHHPISSDDGDLEKALYGSFLPVP
ASPFUM      IGKTMLRRRHVLPSVVHTLVELQVEGTFFTGTYLVTVVHHPISSDDGDLEKALYGSFLPIP
ASPTER      IGKTMLRRRHVLPSVPFTLIELQVEGTFFTGTYLVTVVHHPISSDDGDLEKALYGSFLPVP
NEUCRA      LGATMLRRRHVLPEVTSTLHEIQVEGTFPSGTYLVTVVHNPISDDGDLARALYGSFLPIP
AJECAP      IGKMLLRRRHVLPSVTSTVHEVFVEGTFPSGTYLVTIHEPISTEDGDLEKALYGSFLPVP
COCPOS      IGKEMLRRRHVLPSVVATLKQVQVEGTFFTGTNLITVVNPVCSDDGDLEKALYGSFLPVP
NEOFIS      LGRSILLRRRHVLPSVWFTLRQLQIEGTFFTGTYLVTVQHPIANEEGDLAKALYGSGLPVP
SCHPOM      LGKDMLGRRRHVQPNVVHLLHEIMIEATFPDGTYLITIHDPICTTDGNLEHALYGSFLPTP
CRYNGR      LGKKMLGRRRHVRKGVPESIHTIQVEGTFPDGVFLVTVVDDPISDDGDLNNAFYGSFLPIP
CRYNNE      LGKKMLGRRRHVRKGVPESIHTIQVEGTFPDGVFLVTVVDDPISDDGDLNNAFYGSFLPIP
AGABIS      HGKKILLRRRHVLDPVPALLHEIQVEGTFLDGVFLVTVVHQPICTEDGDLEAALYGSFLPIP
CANENS_1    LGQHLLGRRQVLPVPHLLNAVQVEATFPDGTKLVTVVHDPISRENGELQEALFGSLLPVP
CANENS_JBUREIIB  IGKHLLRRRQVLPVPHLLNIQVEATLPNGTKLVTVVHDPIANENGLEALYGSFLPVP
          * : ** : * * * : * : * : * : * : * : * : * : * : * : * : * : * : *
Prim.cons.  2GKTMLRRRHVLPSVPHTLHEIQVEGTFFTGTYLVTVVHDPISDDGDLEKALYGSFLPVP

```

```

          550      560      570      580      590      600
ASPCLA      PAD--AFPDPNPEDYRPEKMPGAVLPLKE-----ERITLSEGRKRIRLKVTSKGDRPIQV
ASPFUM      PAD--TFPDPNPDDYLPEKMPGAVLPVKN-----ERITLNDGRKRIRLKVMSKGDRPIQV
ASPTER      PSD--AFPDPKPEDYELEKMPGAVIPVKN-----ERITLNAGRKRIKLKVMSKGDRPIQV
NEUCRA      PNDSDVFLPPAAAYEASKQPGAVVCVKE-----AKVKLNEGRKRIRLRVTSKGDRPIQV
AJECAP      SKD--IFPDDPEDYHPLKMPGAVIPVKQ-----SRIMLNEGRKRISLKVTSKGDRPIQV
COCPOS      PKE--TFPDDPDDYQPEKMPGAVIPLKTS----KKIELNAGRNRIMLKVTSRGDRPIQV
NEOFIS      PSD--PPFSSDPADYDEEKAPGAVLLLKGEDGKVPVITLNEGRTRRCLEVTNCGSRPIQV
SCHPOM      SQE--LFPLEEEKLYAPENSPGFVEVLEG-----EIELLNPLRPTPELVRNMGDRPIQV
CRYNGR      SAD--VFPAAPEP---ADTLLGALICRKE-----TVKINASRRRFRLEVKNAGDRPVQV
CRYNNE      SAD--VFPAAPEP---ADTLLGALICRKE-----PIKINASRRRFRLEVKNAGDRPIQV
AGABIS      PQD--DFPVAPDSDYLPEKTAGAIIPKQE-----DIVLNQGRERIRLRITNTGDRPIQV
CANENS_1    SLD-----KFAETKEDNRIPGEILCEDE-----CLTLNIGRKAVILKVTSKGDRPIQV
CANENS_JBUREIIB  SLD-----KFAESKEEHKIPGEIICADG-----RLTLNPGRKAVFLKVVNHGDRPIQV
          . : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *
Prim.cons.  PADDSVFPDPPPEDYEPEKMPGAVIP2KE2DGKPERITLN2GRKRIRLKVTSKGDRPIQV

```

```

          610      620      630      640      650      660
ASPCLA      GSHYHFIETNPQLHFDRLRAYGYRLDIPAGTSVRFEPGDKTVTLVEIAGHRIKGGNCL
ASPFUM      GSHYHFIETNPQLHFDRLRAYGYRLDIPAGTSVRFEPGDKTVTLVEIAGNRIKGGNSI
ASPTER      GSHYHFIETNPQLHFDRLRAYGYRLDIPAGTSVRFEPGDKTVTLVEIAGNRVIKGGNFI
NEUCRA      GSHYHFIETNPQLEFDRARAYGYRLDIPAGTSVRFEPGDKTVTLVEIGGNKVIRGGNNL
AJECAP      GSHYHFIEVNPHLDFDRLRSYGYRLDIPAGTAIRFEPGDSKTVTLVEIGGNKVIHGGNFL
COCPOS      GSHYHFIEVNPQLDFDRAKAYGYRLDIPAGTSIRFEPGATKAIPLVEIGGNKRIRGGNHI
NEOFIS      GSHYHFIETNPKLLFNRIHSYGYRLDKPAGASVRFEPGEKKVVNLVKIAGHKIIQGGNWI
SCHPOM      GSHYHFIETNEKLCFDRSKAYGKRLDIPSGTAIRFEPGVMKIVNLIPIGGAKLIQGGNSL
CRYNGR      GSHYHFLETNPALIFDRLLSYGYHLDIPAGTAVRFEPGEKKTVTMVFFGGKIFHGGSGL
CRYNNE      GSHYHFLETNPALIFDRLLSYGYHLDIPAGTAVRFEPGEKKTVTMVFFGGKIFHGGSGL
AGABIS      GSHYHFIETNRALSFDRLSKYGYRLDIAAGTAVRFEPGDKAVTLVLSIGNKVISGGNSL
CANENS_1    GSHYHFIEVNPYLTFDRRKAYGMRLNIAAGTAVRFEPGDCKSVTLVLSIGNKVIRGGNAI
CANENS_JBUREIIB  GSHYHFIEVNPYLTFDRRKAYGMRLNIAAGDSVRFEPGDHKTVNLVLSIGNKIIRGGNAI
          ***** : * * : * : * : * : * : * : * : * : * : * : * : * : * : * :
Prim.cons.  GSHYHFIETNPQLHFDRL2AYGYRLDIPAGTSVRFEPGDKTVTLVEIGGNKIIRGGNSL

```

```

          670      680      690      700      710      720
ASPCLA      ASGKVEVSRAEEIMQRLQVDGFAHIPEPLS-----ANAAPIAPFTMDREAYARMFGPTT
ASPFUM      ASGKVDISRAEELQRLQVEGFAHVPEPAPT-----ADSALIAPFTMDREAYARMFGPTT
ASPTER      ASGKVDLARAEEIMQRLQVEGFAHTPDPTA-----DAALVEPFTMDREIYARMFGPTT
NEUCRA      ASGKLDLTRAEEIIAKLQKAGFAHAPEEPVGD---AAAFISEPFTMDRRAYAVMFGPTV
AJECAP      ASGKVDLTRADEILERLQKAGFAHTPEPAS-----GITHIDPHSMDRESYMRFGATTT
COCPOS      AVGQVDFRRVDEIIMRLQKAGFAYTPEPKQ-----DAHLIEPFSMDREAYARMFGPTT
NEOFIS      APGKVDQGKIREIEERLRSKEFCHDGRPVR-----LPAPDPCVMEFEKYCEMFGPTE
SCHPOM      SKGVFDDSRTREIVDNLMKQGFHMHPESPLN----MPLQSARPFVVPRKLYAVMFGPTT
CRYNGR      GNGSFDENLRETKVKEMVEKVGFGHKEQEK-----IEEGPVTEMNREVYASMFGPTT
CRYNNE      ASGSFDENLRETKVKEMVEKGGFGHKDQEK-----VEEGPTEMNREVYASMFGPTT
AGABIS      ASGHIGTFRSEVLLEDILRRDFAHVSEPGA-----LEVMEDTKIGRETYISMYGPTT
CANENS_1    ADGPVNETNLEAAMHAVRSKGFGEHEEKDASEGFTKEDPNCPFNTFIHRKEYANKYGPTT
CANENS_JBUREIIB  ADGPVNEANCKAAMEIVCRREFGHEEEDASEGVTTGDPDCPFTKAIPREYANKYGPTT
          . * . : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *
Prim.cons.  ASGKVDE2RAEEIMERLQKKGFAHTPEP4ASEG2T2ADA2LIEPFTMDREAYARMFGPTT

```

	730	740	750	760	770	780
ASPCLA	GDLVRLGLTNLWIRVEKDYT---VYGDECAFGGGKTI	REGMGQASERSEK-----ECLD				
ASPFUM	GDLVRLGLTNLWVIRVEKDYT---VYGDECAFGGGKTI	REGMGQSSERSAT-----ECLDT				
ASPTER	GDLVRLGLTNLWVIRVEKDYT---MYGDECSFGGGKSL	RDGMAQASGRSSK-----DCLDT				
NEUCRA	GDMVRLGATDLWIKVEKDYT---VYGDECKFGGGKTL	REGMGQATGRSDE-----ETLDM				
AJECAP	GDLIRLGSTDLWIKVEKDLT---SFGDECTFGGGKTL	REGMGQASGRSD-----DVLDT				
COCPOS	GDVVKLGTTDLWIKVEKDLT---VYGDECSFGGGKTI	IRDGMGQATGRHSV-----DVLDT				
NEOFIS	GDLVRLGLTNLWVEVERDCRKSKEYGDECNFGGGKTI	IRDGMGQTSNGSSGGGKYTCADT				
SCHPOM	NDKIRLGDTNLIVRVEKDFD---EYGNESVFGGGKVI	RDGTGQSSSKSMD-----ECLDT				
CRYNCR	GDKIKLADMDLWIEVEKDYT---VYGDECKFGGGKVI	RDGGQASGRHDH-----EVLDT				
CRYNNE	GDKIKLADMDLWIEVEKDYT---VYGECKFGGGKVI	RDGGQASGRHEH-----EVLDT				
AGABIS	GDRVRLGDTLWIEVEHDET---VYGDEVKFGGGKVI	REGMGQATNRSSN-----ETLDT				
CANENS_1	GDKIRLGDTNLIAIEKDYA---LYGDECVFGGGKVI	RDGMGQSCGHPPAIS-----LDT				
CANENS_JBUREIIB	GDKIRLGDTDLIAIEKDFD---LYGDESVFGGGKVI	RDGMGQSSGHPPAMS-----LDT				
	. * : : * : * : * : * : * : * : * : * : * : *	: * : * : * : * : * : * : * : * : * : *	: * : * : * : * : * : * : * : * : * : *	: * : * : * : * : * : * : * : * : * : *	: * : * : * : * : * : * : * : * : * : *	: * : * : * : * : * : * : * : * : * : *
Prim.cons.	GD2VRLGDT2LWIEVEKDYTKSKVYGDECKFGGGK2IRDGMGQASGRSS43SKGYECLDT					

	790	800	810	820	830	840
ASPCLA	VITNALIIDWSGI	FKADIGIKDGI	IIV	IGKAGNPDV	MDGVHPDM	IIGTST
ASPFUM	VITNALIIDWSGI	YKADIGIKNGL	IS	ATGKAGNPD	MDGVHPDM	IVGSS
ASPTER	VVTNALIIDWTGI	YKADIGIRNG	VIV	IGKAGNPD	MDGVHPDM	VVGS
NEUCRA	VVTNALVVDWTGI	YKADIGKNG	VIV	IGKAGNPD	MDGVTPGM	IVGSC
AJECAP	VITNALIVDWTGI	YVADIGIKD	NI	VIGKAGNPD	MDGVYPM	FVGS
COCPOS	VLVNALIVDWTGI	YKADIGLKD	GL	ICGKAGNPD	MDGVTPNM	IVGSS
NEOFIS	VITNALIIDHSGI	FKADIGVKDK	IS	IGRAGNPD	TMND	IDEG
SCHPOM	VITNAVLI	DHTGIYKADIG	KNG	VIV	IGKAGNPD	MDNI
CRYNCR	VITNALIVDWTGI	YKADIGVKN	GI	VIGKAGNPD	MDGVTDGM	IVGSS
CRYNNE	VITNALIVDWN	GIYKADIGVKN	GI	VIGKAGNPD	MDGVTDGM	IVGSS
AGABIS	VITNALIVDWSGI	YKADIGVKN	GF	ICGKAGNPD	VMSNI	HP
CANENS_1	VITNAVLI	DYTGIIKADIG	KDGL	IAS	IGKAGNPD	IMNG
CANENS_JBUREIIB	VITSAVLI	DYTGIIKADIG	KDGL	IAS	IGKAGNPD	IMNG
	* : : * : * : * : * : * : * : * : * : * : *	* : * : * : * : * : * : * : * : * : *	* : * : * : * : * : * : * : * : * : *	* : * : * : * : * : * : * : * : * : *	* : * : * : * : * : * : * : * : * : *	* : * : * : * : * : * : * : * : * : *
Prim.cons.	VITNALIIDWTGIYKADIGIK2GLIVIGKAGNPD2MDGV2PNMIVGSSDVIAGENKIV					

	850	860	870	880	890	900
	His407	His409				
ASPCLA	TAGGFD	IVHFICPQ	QVE	EEALASG	VTTFLGGGTG	PPTG
ASPFUM	TAGGFD	IVHFICPQ	QVDE	EEALASG	ITFLGGGTG	PSTG
ASPTER	TAGGFD	IVHFICPQ	QVDE	EEALASG	ITSMGGGTG	PSTG
NEUCRA	TAGGID	IVHFICPQ	QA	EEALAG	ITFLGGGTG	PSAG
AJECAP	TAGGID	IVHLITP	DQV	EEALASG	VTTMLGGGTG	PSTG
COCPOS	TAGGID	IVHFICPQ	QVDE	EEALASG	VTTLLGGGTG	PTEGS
NEOFIS	TAGGID	IVHFLDP	QVME	EGLCNGL	TTLGGGTG	PSTG
SCHPOM	TYGMD	SVHFICPQ	QIE	EEALASG	ITMYGGGTG	PSTG
CRYNCR	TAGRLD	IVVYIS	PQLM	TEALASG	ITTVGGGTG	PADGS
CRYNNE	TAGALD	IVVYICP	QLM	TEALASG	ITTVGGGTG	PADGS
AGABIS	TAGGID	IVHFICP	QLV	DEALASG	LTTLLGGGTG	PSAG
CANENS_1	TAGAI	DCVYVYICP	QLV	EAISSG	ITTLVGGGTG	PAAG
CANENS_JBUREIIB	TAGGID	IVVYICP	QSL	DEAIS	SGITTVGGGTG	P
	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *
Prim.cons.	TAGGIDTHVHFICPQQVDEALASGITTLGGGTGPTGTNATTCTPGPTYMRQMIQATD2					

	910	920	930	940	950	960
			Lys490			His519
ASPCLA	LPIN	VGITG	KGNDC	GGQSIEEQ	ILAGA	AGL
ASPFUM	LPIN	VGITG	KGNDS	GGKSIEEQ	IRAGA	AGL
ASPTER	LPVNI	GITG	KGNDC	EGLSIEEQ	IHAGA	AGL
NEUCRA	LP	LNIGIT	GKGNDS	SPSALRE	QILAG	ACGL
AJECAP	LP	INFGIT	GKGNDC	QPSLR	QIRAGA	AGL
COCPOS	LP	INVGLT	GKGNDS	GLP	SLRDCR	AGA
NEOFIS	LP	VNI	GITG	KGNNS	CKEGL	EEQ
SCHPOM	YPM	NI	GLTG	KGNDS	GSSE	LEI
CRYNCR	I	PLN	FGFT	GKGS	SDSG	TN
CRYNNE	V	PLN	FGFT	GKGNDS	SGT	NAL
AGABIS	L	PM	NFGFT	GKGN	DAGP	TA
CANENS_1	L	PLN	FGFT	GKGS	SSK	DEL
CANENS_JBUREIIB	I	PLN	FGFT	GKGS	SGSH	DEL
	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *	* : * : * : * : * : *
Prim.cons.	LPLNFGITGKGNDSGP4SL2EQIRAGAAGLKLHEDWGSTPAAIDTCLSVCEYDVQC3IH					


```

                1210      1220      1230      1240      1250      1260
                |        |        |        |        |        |
ASPCLA      VPSTSV---MWVSQASIDAGIVQSYGIKKRIEAVKNCRNIGKKDMKFNDLMPRHVDPES
ASPFUM      VPSTSI---MWVSQASIDDGIVQSYGLKKRIEAVRNCRNIGKKDMKFNDVMPKMRVDPES
ASPTER      -----
NEUCRA      VAKSKV---LFVSQASLDGSAIEGYGLKSRVEAVKGCRTVGGKDMKFNERMPRMKVDPER
AJECAP      LPRTSI---MFVSQASKSLGIIDSYGLKRRVEVVKNCRNISKKDMKFNDVMPKMRVDPES
COCPOS      NPKASI---MFVSQASIKLGIIDSYHLKKRIEAVKNCRNISKKDMKFNDIMPKMRVDPES
NEOFIS      VPRTSI---MFVSQASKNK-VEEYGLKKTIEAVKHCRTIDTNQIQDPKPKPEFLSKKH
SCHPOM      DPERSI---AFVSKASITSGVIESYGLHKKRVEAVKSTRNIGKKDMVYNSYMPKMTVDPEA
CRYNGR      LRPLHS-IQLFVSQASLDKDLVKRYRLRKRRAEAVKNCRSIGKKDMKWNDTMPKMTVDPET
CRYNNE      PEAAALNSIVVVSQASLDKDLVKRFNIKRAEAVKNCRAIGKKDMKWNDSMPKMTVDPET
AGABIS      PGSAAALNSVAFVSQVSITSRVIESYGLSKKIEAVRGCARDIGKKDMKWNDTTPAMKVDPEP
CANENS_1    GKAGGALSIAFVSKAALDQRVNVLYGLNKRVEAVSNVRKLTCLKLMDLNDALPEITVDPEP
CANENS_JBUREIIB  GKAGSALSIAFVSKAALDLGVKVLYGLNKRVEAVSNVRKLTCLKLMDLNDALPEITVCPET

```

Prim.cons. VP2TSI2SIFVSQAS2D3GVIESYGLKKR2EAVKNCRNIGKKDMKFND3MPKMTVDPEP

```

                1270      1280      1290      1300      1310      1320
                |        |        |        |        |        |
ASPCLA      YRVEADGVLCEAEPAVELPLTQDYFVY-----
ASPFUM      YVVEADGVLCEAEPAELPLTQDYFVY-----
ASPTER      -----
NEUCRA      YTVEADGEVCTAEPSSSELPLTQSYFVY-----
AJECAP      YDRTVQHITYQAKGENGWSREDVLLVLEKQKSRKAGER-----
COCPOS      YVVEADGEECTAEPVSELPLTQDYFVY-----
NEOFIS      MKYNSEMPEMEVDPEPTEFSEPTVKYATWLLQRRCLRKTISYSESGTEVC SHDMGANFMEL
SCHPOM      YTVTADGKVMCEPEVDKLPKLSQSYFIF-----
CRYNGR      YDVRADGVLCDVPPADKLPTRRYFVY-----
CRYNNE      YDVHADGVLCDVPPADKLPKRYFVY-----
AGABIS      YEVRADGVLMDVKPVERVALATPYNLF-----
CANENS_1    YTVKADGKLLCVSEATTVPLSRNYFLF-----
CANENS_JBUREIIB  FTVTVDGQALSSEAVTTPLSQNYFIF-----

```

Prim.cons. YTVEADGVLCD2EPA22LPLTQDYFVY22Q2222222ISYSESGTEVC SHDMGANFMEL

```

                1330
                |
ASPCLA      -----
ASPFUM      -----
ASPTER      -----
NEUCRA      -----
AJECAP      -----
COCPOS      -----
NEOFIS      FGPAFLSVLPFGEA
SCHPOM      -----
CRYNGR      -----
CRYNNE      -----
AGABIS      -----
CANENS_1    -----
CANENS_JBUREIIB  -----

```

Prim.cons. FGPAFLSVLPFGEA

Dados do alinhamento das seqüências completas:

Comprimento do alinhamento: 1334
 Idênticos (*): 246 ou 18.44 %
 Fortemente similares (:): 135 ou 10.12 %
 Fracamente similares (.): 72 ou 5.40 %
 Diferentes: 881 ou 66.04 %
 Seqüência 0001 : ASPCLA (837 resíduos).
 Seqüência 0002 : ASPFUM (908 resíduos).
 Seqüência 0003 : ASPTER (1171 resíduos).
 Seqüência 0004 : NEUCRA (842 resíduos).
 Seqüência 0005 : AJECAP (999 resíduos).
 Seqüência 0006 : COCPOS (836 resíduos).
 Seqüência 0007 : NEOFIS (899 resíduos).
 Seqüência 0008 : SCHPOM (835 resíduos).
 Seqüência 0009 : CRYNGR (832 resíduos).
 Seqüência 0010 : CRYNNE (833 resíduos).
 Seqüência 0011 : AGABIS (838 resíduos).
 Seqüência 0012 : CANENS_1 (840 resíduos).
 Seqüência 0013 : CANENS_JBUREIIB (840 resíduos).

Dados do alinhamento da região do peptídeo, destacada em amarelo:

Comprimento do alinhamento: 99
 Idênticos (*): 19 ou 19.19 %
 Fortemente similares (:): 11 ou 11.11 %
 Fracamente similares (.): 7 ou 7.07 %
 Diferentes : 62 ou 62.63 %

Figura 7: alinhamento múltiplo de seqüências de aminoácidos da subunidade C de ureases oriundas de bactérias e da região correspondente na JBURE-IIB e na isoforma majoritária de urease de *C. ensiformis*. Códigos dos organismos de acordo com a Tabela 1.

	10	20	30	40	50	60
HELPHYL	MKKISRKEYASMYGPTTGDKVRLGDTDLIAEVEHDYTIYGEELKFGGGKTLREGMSQS--					
HELHEP	MKISRQYASMYGPTTGDKVRLGDTNLF AEIEKDYTTYGEEIKFGGGKTIRDGMAQS--					
BACPAS	-MKINRQQYAESYGPTTVGDEVRLADTDLWIEVEKDYTTYGDEVNFGGGKVLREGMGNGT					
BACSUB	-MKMSREEYAELEFGPTTGDKIRLGDSDLWIEVEKDFTVYGEEMI FGGGKTIRDGMGQNGR					
KLEAER	MSNISRQAYADMF GPTTVGDKVRLADTELWIEVEDDLTTYGEEVKFGGGKVIRDMGMGQG--					
KLEPNE	MSNISRQAYADMF GPTTVGDKVRLADTELWIEVEDDLTTYGEEVKFGGGKVIRDMGMGQG--					
CANENS_1	NTFIHRKEYANKYGPTTGDKIRLGDSDLWIEVEKDYALYGDECVFGGGKVIRDMGMGQS-C					
CANENS_JBUREIIB	TKAIPREYANKYGPTTGDKIRLGDSDLIAEIEKDFALYGDESVFGGGKVIRDMGMGQS-S					
	: * : * * : * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * *					
Prim.cons.	M3KISR2EYA4MYGPTTGDKVRLGDTDLW2EVEKDYT2YGEEVKFGGGKVIRDMGMGQSG4					
	70	80	90	100	110	120
HELPHYL	NNPSKEELDLIITNALIVDYTG IYKADIGIKDGKIAGIGKGGNKDMQDGVKNL SVGPAT					
HELHEP	ASTYTNELDAVITNAMIIDYTG IYKADIGIKGGKIVGIGKAGNPDTQDSVNEAMVVGAAAT					
BACPAS	YTRTENVLDLLITNALILDYTG IYKADIGVKGDIYVIGKGGNPDIMDGVTPNMIVGTAT					
BACSUB	ITGKDGALDLVITNVVLLDYTG IYKADVGVKDGRIYVGVKSGNPDIMDGVDPHVMVIGAGT					
KLEAER	QMLAADCVDLVLTNALIVDHWGIVKADIGVKGDRIFAIGKAGNPDIQPNVT--IPIGAAT					
KLEPNE	QMLAADCVDLVLTNALIVDHWGIVKADIGVKGDRIFAIGKAGNPDIQPNVT--IPIGAAT					
CANENS_1	GHPPAISLDTVITNAVIDYTG IYKADIGIKDGLIASIGKAGNPDIMNGVFSNMIIGANT					
CANENS_JBUREIIB	GHPPAMSLDTVITSAVIDYTG IYKADIGIKDGLIASIGKAGNPDIMNGVFPNMIIGVNT					
	: * : * * : * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * *					
Prim.cons.	23P2A23LDLVITNALI2DYTG I2KADIG2KDGRI2GIGKAGNPDI2DGVTPNMIIGAAT					
	130	140	150	160	170	180
		His407	His409			
HELPHYL	EALAGEGLIVTAGGIDT I I F I S P Q Q I P T A F A S G V T T M I G G G T G P A D G T N A T T I T P G R R N					
HELHEP	EVIAGEGQIITAGGIDT I I F I S P T Q I P T A L Y S G V T T M I G G G T G P A A G T N A T T C T P G K W N					
BACPAS	EVIAAEGKIVTAGGIDT I I F I N P D Q V D V A L A N G I T T L F G G G T G P A E G S K A T T V T P G P W N					
BACSUB	EVI S G E K I L T A G G V D T I I F I C P Q Q M E V A L S S G V T T L G G G T G P A T G S K A T T C T S G A W Y					
KLEAER	EVIAAEGKIVTAGGIDT I I F W I C P Q Q A E E A L V S G V T T M V G G G T G P A A G T H A T T C T P G P W Y					
KLEPNE	EVIAAEGKIVTAGGIDT I I F W I C P Q Q A E E A L V S G V T T M V G G G T G P A A G T H A T T C T P G P W Y					
CANENS_1	EVIAGEGLIVTAGAIDC I V Y I C P Q L V Y E A I S S G I T T L V G G G T G P A A G T R A T T C T P S P T Q					
CANENS_JBUREIIB	EVI C G E L I V T A G G I D C I V Y I C P Q S L D E A I S S G I T T V G G G T G P T D G S R A T T C T P A P T Q					
	: * : * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * *					
Prim.cons.	EVIAGEGKIVTAGGIDTHIHFICPQQ3EEALSSGVTTMVGGGTGPAAGT4ATTCTPGPW2					
	190	200	210	220	230	240
				Lys490		
HELPHYL	LKFMRLRAAEYSMNFGFLAKGNVSN D A S L A D Q I E A G A I G F I H E D W G T T P S A I N H A L D V A					
HELHEP	MHQMLRAAESYAMNLF G F G K G N S S N E E G L E E Q I K A G A L G L V H E D W G S T P A A I N H A L N V A					
BACPAS	IEKMLKSTEGLEP I N V G I L G K G H S S I A P I M E Q I D A G A G L I H E D W G A T P A S I D R S L T V A					
BACSUB	MARMLEAAEEFP I N V F L G K G N A S D K A P L I E Q V E A G A I G L L H E D W G T T P S A I K T C M E V V					
KLEAER	ISRM L Q A A D S L P V N I G L L G K G N V S Q P D A L R E Q V A A G V I G L I H E D W G A T P A A I D C A L T V A					
KLEPNE	ISRM L Q A A D S L P V N I G L L G K G N V S Q P D A L R E Q V A A G V I G L I H E D W G A T P A A I D C A L T V A					
CANENS_1	MRLMLQSTDDLPLNFGFTGKGS S K P D E L H E I I K A G A M G L L H E D W G S T P A A I D N C L T I A					
CANENS_JBUREIIB	MKLMLQSTDDIPLNFGFTGKGS S H P D E L H E I I K A G A M G L L H E D W G C T P A A I D N C L A V A					
	: * * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * *					
Prim.cons.	M2RMLQAA2SLP4NFGFLGKGNVS2PD3L2EQIKAGAIGLKI HEDWGATPAAID3ALTVA					
	250	260	270	280	290	300
		His519		His545		
HELPHYL	DKYDVQVAIHTDTLNEAGCVEDTMAA I A G R T M H T F F T E G A G G G H A P D I I K V A G E H N I L P A					
HELHEP	QKYDVQVAIHTDTLNEAGCVEDT M K A I D G R T I H T F F T E G A G G G H A P D I I K A G E F N I L P A					
BACPAS	DEADVQVAIHTDTLNEAGFLEDTLRA I N G R V I H S F F V E G A G G G H A P D I M A M A G H P N V L P S					
BACSUB	DEADIQVAIHTDTINEAGFLENTLDA I G D R V I H T Y F I E G A G G G H A P D I M K L A S Y A N I L P S					
KLEAER	DEM DI Q V A I H S D T L N E S G F V E D T L A A I G G R T I H T F F T E G A G G G H A P D I I T A C A H P N I L P S					
KLEPNE	DEM DI Q V A I H S D T L N E S G F V E D T L A A I G G R T I H T F F T E G A G G G H A P D I I T A C A H P N I L P S					
CANENS_1	EHHDIQNIHTDTLNEAGFVEHSIAA F K R T I H T Y S E G A G G G H A P D I I K V C G I K N V L P S					
CANENS_JBUREIIB	EQHDIQVNIHTDTVNESGFVEHTIAA F N G R T I H T Y S E G A G G G H A P D I I K V C S M K N V L P S					
	: * * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * * : * * * * * * * *					
Prim.cons.	DE4DIQVAIHTDTLNEAGFVEDTLAAIGGRTIHTFHTEGAGGGHAPDI I K22GHPNILPS					


```

          310      320      330      340      350      360
HELPYL  STNPTIPFTVNTAEAHMDMLMVCHHLDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSI
HELHEP  STNPTIPFTKNTADEHLDMVMVCHHLDKIKEDVAFADSRIRPETIAAEDTLHDMGIFSI
BACPAS  STNPTTRPFTVNTIDEHLDMVMVCHHLDKQIPEDVAFADSRIRPETIAAEDTLHDLGIISM
BACSUB  STTPTIPYTVNTMDEHLDMMMVCHHLDKVPEDVAFSHSRIRRAATIAAEDILHDIGAIISM
KLEAER  STNPTLPYTLNTIDEHLDMVMVCHHLDPAEDVAFESRIRRETIAAEDVLHDLGAFSL
KLEPNE  STNPTLPYTLNTIDEHLDMVMVCHHLDPAEDVAFESRIRRETIAAEDVLHDLGAFSL
CANENS_1 STNPTTRPLTSNTIDEHLDMVMVCHHLDREIPEDLAFASRIRKKTIAAEDVLNDIGAIISI
CANENS_JBUREIIB STNPTTRPLTSNTVDEHLDMVMVCHHLDREIPEDLAFASSRVREQTIAAEDILHDIGGISI
Prim.cons. STNPT2P2TVNTIDEHLDMVMVCHHLD33IPEDVAFADSRIRPETIAAED2LHD2GA2SI

```

```

          370      380      390      400      410      420
      Asp633
HELPYL  TSSISQAMGRVGEVITRTWQTADKNKKEFGRLKEEKGDNDNFRICKRYLSKYTINPAIAHG
HELHEP  TSSISQAMGRVGEVITRTWQTADCKNEFGALKEECGENDNFRICKRYSKYTINPAIAHG
BACPAS  MSTIALAMGRAGEMVLRWQTADKMKKQRGPLAEEKNGSDNFRICKRYSKYTINPAIAQG
BACSUB  TSSISQAMGRVGEVITRTWQVADKMKKQRGALAGE-NGNDNVRKRYIAKYTINPAITHG
KLEAER  TSSISQAMGRVGEVILRTWQVAHRMKVQRGALAEETGDNDFRVKRYIAKYTINPALTHG
KLEPNE  TSSISQAMGRVGEVILRTWQVAHRMKVQRGALAEETGDNDFRVKRYIAKYTINPALTHG
CANENS_1 ISSISQAMGRVGEVISRWQTADKMKKAQTGPLKCDSSDNDNFRIRRYIAKYTINPAIANG
CANENS_JBUREIIB ISSIAQAVGRIGEVIISCTWQTADKMKKAERGPLQPDGSDNDFRICKRYIAKYTINPAIVNG
Prim.cons. TSSDSQAMGRVGEVILRTWQTADKMKKQRGALAE2GDNDNFRICKRYIAKYTINPAIAHG

```

```

          430      440      450      460      470      480
HELPYL  ISEYVGSVEVGKVADLVWSPAFFGVKPNMIIKGGFIALSQMGDANASIPTPQPVVYREM
HELHEP  ISEYVGSVEVGKFADLVWKPSPMFGIKPEMILKNGMIVAAGKSDNASIPTPEPVVYAPM
BACPAS  IAHEVGSIEEGKFADLVWPEKFFGVKADRVIKGGIIAYAQIGDPSASIPTPQPVMGRRM
BACSUB  LSHDEVGSVEGKLADLVWDPVFFGVKPELVKGGMIIARAQMGDPNASIPTPEPVFMRQM
KLEAER  IAHEVGSIEVGKLADLVWSPAFFGVKPATVIKGGMIIAIPMGDINASIPTPQPVHYRPM
KLEPNE  IAHEVGSIEVGKLADLVWSPAFFGVKPATVIKGGMIIAIPMGDINASIPTPQPVHYRPM
CANENS_1 FSQYVGSVEVGKLADLVWKPSPFFGKPEMVIKGGMVAWADIGDPNASIPTPEPVKMRPM
CANENS_JBUREIIB ISQYVGSVEVGKLADLVWKPSPFFGAKPDIVIKGGSIAWADMGPNGSIPTPEPVLMRPM
Prim.cons. IISH2VGSVEVGKLADLVW2P2FFGVKPEMVIKGGMIIA2AQMGPNASIPTP2PVHYRPM

```

```

          490      500      510      520      530      540
HELPYL  FAHHGKAKYDANITFVSAAYDKGIKEELGLERQVLPVKNCRNITKKDMQFNDTAAHIEV
HELHEP  FGSYGKAKYNCAITFVSKIAAYDCHIKEELGLERILLPVKNCRNITKKDMKFNDVITPIEV
BACPAS  YGTVGDLIHDNTITFMSKSSIQQGVPAKLGKRRIGTVKNCRNIGKKDMKNDVTTDDIDI
BACSUB  YASYGKANRSTSIITFMSQASIERGVAESLGLERKRIKSPVKNIRKLSKLDMKLNSALPKIEI
KLEAER  FGALGSAARHHCRLTFLSQAAAANGVAERLNLRSAIAVVKGCRTVQKADMVHNSLQPNITV
KLEPNE  FGALGSAARHHCRLTFLSQAAAANGVAERLNLRSAIAVVKGCRTVQKADMVHNSLQPNITV
CANENS_1 YGTLGKAGGALSIAFVSKAALDQRVNVVLYGLNKRVEAVSNVRKLTKLDLKLNNSLPEITV
CANENS_JBUREIIB YGTLGKAGSALSIAFVSKAALDLGVKVLYGLNKRVEAVSNVRKLTKLDLKLNNSLPEITV
Prim.cons. 2GTLGKA3H3CSITFVSA4A2D2GV2E3LGLER2RI3PVKNCR22TK2DMKLN3LP2ITV

```

```

          550      560      570
HELPYL  NSETYHVFVDGKEVTSKPAKVSLSLAQLFSIF
HELHEP  NPETYEVVRVNNKITSKPEKVSLSGLQLYCLF
BACPAS  NPETYEVKVDGEVLTCEPVKELPMAQRYFLF
BACSUB  DPKTYQVVFADGEEELSCQPVVDYVPLGQRYFLF
KLEAER  DAQTYEVRVDGELITSEPADVLPMAQRYFLF
KLEPNE  DAQTYEVRVDGELITSEPADVLPMAQRYFLF
CANENS_1 DPESYTVKADGKLLCVSEATTVPKSRNYFLF
CANENS_JBUREIIB CPETFVTVVDGQALSSEAVTTLPKSNYFIF
Prim.cons. DPETYEVVRVDGELLTSEP2D32PLAQRYFLF

```

Dados do alinhamento:

```

Comprimento do alinhamento: 571
Idênticos (*) : 216 ou 38.00 %
Fortemente similares (.) : 123 ou 21.54 %
Fracamente similares (.) : 69 ou 12.08 %
Diferentes : 162 ou 28.37 %
Seqüência 0001 : HELPYL (569 resíduos).
Seqüência 0002 : HELHEP (569 resíduos).
Seqüência 0003 : BACPAS (570 resíduos).
Seqüência 0004 : BACSUB (569 resíduos).
Seqüência 0005 : KLEAER (567 resíduos).
Seqüência 0006 : KLEPNE (567 resíduos).
Seqüência 0007 : CANENS_1 (570 resíduos).
Seqüência 0008 : CANENS_JBUREIIB (570 resíduos).

```

4.3 Determinação de estruturas terciárias de proteínas

As abordagens para determinação e predição de estruturas tridimensionais de proteínas podem ser classificadas de acordo com a técnica predominantemente utilizada para calcular o modelo (Šali & Blundell, 1993). Os métodos experimentais incluem a cristalografia de raios X e a ressonância magnética nuclear (RMN).

Na cristalografia de raios X, a interação dos raios X com elétrons das proteínas organizadas em cristais é utilizada para obtenção de um mapa de densidade eletrônica da proteína, que pode ser traduzido em termos de um modelo atômico. A cristalização de proteínas, entretanto, é um processo lento, complexo e, algumas vezes, impossível de realizar-se. As dificuldades abrangem tanto o processo de produção e purificação da proteína de interesse quanto a indução da cristalização e o transporte dos cristais (Branden & Tooze, 1999).

No RMN, as propriedades magnéticas dos spins dos núcleos atômicos das proteínas são utilizadas para obtenção de uma lista de restrições espaciais entre os átomos da molécula, da qual podem ser extraídos dados que levam à definição da estrutura tridimensional da proteína. Esta técnica é aplicada a proteínas de até ~35kDa em solução, obtendo estruturas múltiplas para uma mesma proteína (refletindo sua flexibilidade em meio biológico) (Branden & Tooze, 1999).

Na cristalografia de raios X, apenas uma estrutura média é obtida, e esta corresponde à proteína empacotada em uma matriz cristalina, com potencial diferenciação de sua estrutura tridimensional em relação àquela obtida em solução (Branden & Tooze, 1999).

As abordagens teóricas para determinação e predição de estruturas tridimensionais de proteínas podem ser divididas em métodos físicos e empíricos. Os métodos físicos de predição baseiam-se em interações entre átomos e incluem a dinâmica molecular e a minimização de energia, enquanto os métodos empíricos dependem de estruturas de proteínas que já tenham sido determinadas experimentalmente (Šali & Blundell, 1993). Esses últimos incluem as modelagens combinatória e comparativa.

4.4 Modelagem molecular comparativa

A modelagem molecular comparativa utiliza estruturas protéicas determinadas experimentalmente (moldes) para prever a conformação de outras proteínas com seqüências similares de aminoácidos (alvos). Isso é possível por que uma pequena

mudança na seqüência de aminoácidos geralmente resulta em uma pequena mudança na estrutura (Chothia & Lesk, 1986) e estima-se que um terço das seqüências conhecidas sejam relacionadas a pelo menos uma estrutura conhecida (Martí-Renom *et al.*, 2000). A modelagem comparativa não é tão acurada quanto a cristalografia de raios X ou RMN (Šali & Blundell, 1993), mas é o melhor método disponível para predição de estruturas de proteínas (Jones, 2000). Além disso, mesmo modelos de baixa resolução podem ser úteis para a biologia, visto que alguns aspectos da função podem, algumas vezes, serem preditos a partir de aspectos estruturais grosseiros do modelo (Martí-Renom *et al.*, 2000).

Todos os métodos de modelagem comparativa consistem em quatro passos seqüenciais (Figura 8): seleção do molde, alinhamento molde-alvo, construção do modelo e avaliação do modelo. Se o modelo não for satisfatório, os três primeiros passos podem ser repetidos até que um modelo satisfatório seja obtido (Martí-Renom *et al.*, 2000).

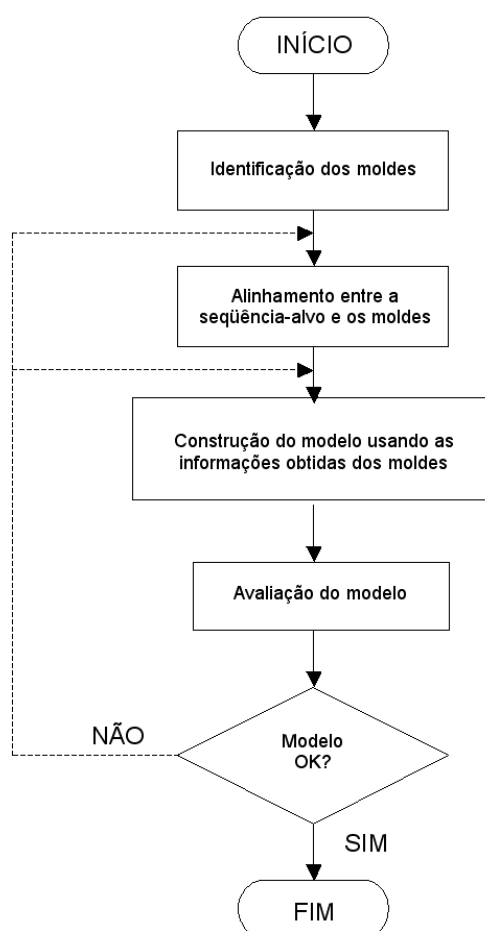


Figura 8: Passos na modelagem comparativa por homologia (retirado de Fernandes, 2002; originalmente adaptado de Martí-Renom *et al.*, 2000).

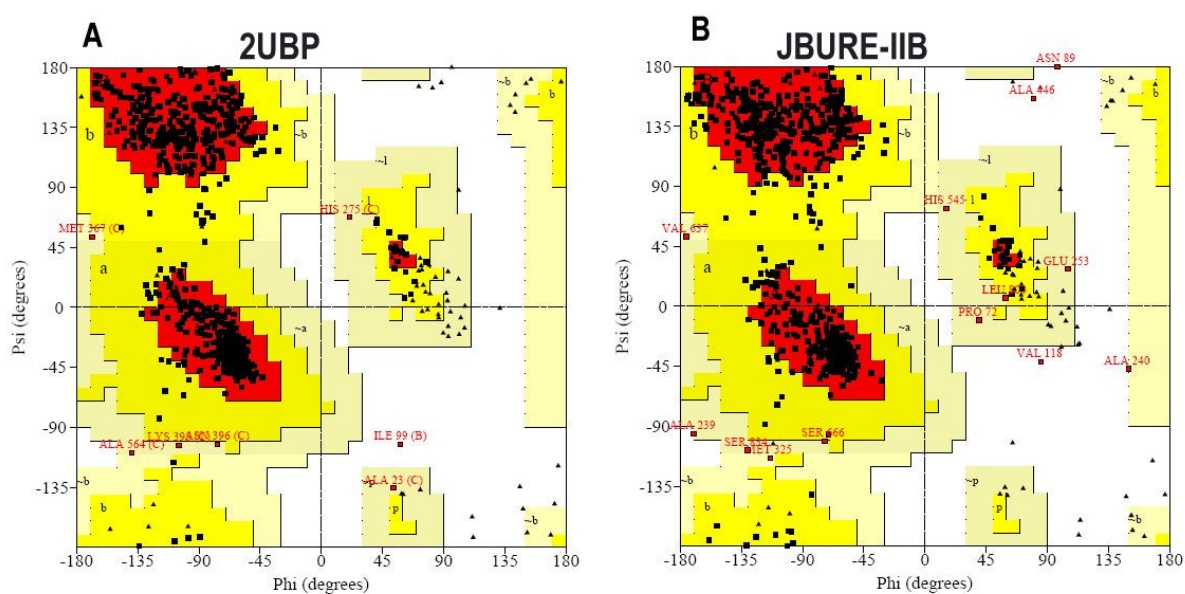
Uma vez alinhados molde e alvo, uma variedade de métodos pode ser utilizada para construir o modelo tridimensional da proteína alvo. O método original e ainda amplamente utilizado é a modelagem por montagem de corpo rígido. Outra família de métodos, a modelagem por similaridade de segmentos, se baseia nas posições aproximadas de átomos conservados nos moldes. O terceiro grupo de métodos, a modelagem por satisfação de restrições espaciais, usa distâncias geométricas ou técnicas de otimização para satisfazer restrições espaciais extraídas do alinhamento (Sánchez & Šali, 2000). Tais restrições derivadas da homologia são geralmente complementadas por restrições estereoquímicas dos comprimentos de ligação, ângulos de ligação, ângulos diedro e contatos entre átomos não ligados, obtidas do campo de força de mecânica molecular. O modelo é então derivado ao minimizarem-se as violações de todas as restrições obtidas. No caso do programa Modeller, isso é feito por um método de otimização de espaço real: o modelo é construído usando restrições de distância e de ângulos diedro para a seqüência alvo, obtidos de seu alinhamento com as estruturas-molde. Então, as restrições espaciais e os termos do campo de força CHARMM22, que impõem a estereoquímica adequada, são combinados em uma função objetiva. Finalmente, o modelo é gerado ao otimizar-se a função objetiva no espaço cartesiano.

Quanto menor a similaridade entre alvo e molde, maior a propensão a erro da modelagem (Sánchez & Šali, 2000). Erros em modelagem comparativa podem ser explicados baseando-se nos fatos de que o modelo assemelha-se aos moldes tanto quanto possível, e que o procedimento de modelagem não consegue superar erros de alinhamento. Os erros típicos em modelagem comparativa incluem: erros de empacotamento de cadeias laterais, distorções e alterações em regiões alinhadas corretamente, erros em regiões sem molde, erros devidos a alinhamentos ruins e utilização de moldes incorretos (Sánchez & Šali, 2000).

Após sua construção, é importante verificar os possíveis erros do modelo. As avaliações podem ser de dois tipos: internas, que avaliam a auto-consistência do modelo, se ele satisfaz ou não as restrições usadas para calculá-lo; ou externas, baseadas em informações que não foram utilizadas para calcular o modelo (Sánchez & Šali, 2000). Entre as variedades de avaliação externa, tem-se o reconhecimento de regiões não confiáveis do modelo. Essa avaliação pode ser feita pelo cálculo de perfis de energia do modelo ou por avaliação estereoquímica (Martí-Renom *et al.*, 2000; Sánchez & Šali, 2000).

A avaliação estereoquímica pode considerar vários aspectos do modelo, como comprimento de ligações, ângulos de ligação, planaridade da ligação peptídica e de anéis de cadeias laterais, quiralidade, ângulos de torção das cadeias laterais e principal e contatos entre átomos não ligados. Para a avaliação estereoquímica do modelo apresentado no presente trabalho, foi utilizado o mapa de Ramachandran gerado pelo programa Procheck (Laskowski *et al.*, 1993).

Em geral, espera-se que um modelo de qualidade tenha 90% de seus resíduos nas regiões mais favoráveis (Laskowski *et al.*, 1993). O modelo construído no presente trabalho apresenta 87,3% dos resíduos nas regiões favoráveis e 11,0% nas regiões adicionalmente favoráveis, enquanto o molde utilizado (2UBP) apresenta 89,6% dos resíduos em regiões favoráveis e 9,4% nas regiões adicionalmente favoráveis. Considerando o tamanho maior da seqüência do modelo (842 aminoácidos) em relação à seqüência do molde (790 aminoácidos) e a natureza das regiões envolvidas nesse aumento de seqüência (trechos ligadores semelhantes a alças, aparentemente mais flexíveis), considerou-se o valor obtido no mapa de Ramachandran adequado para validação deste modelo. Os mapas de Ramachandran do modelo apresentado neste trabalho e do molde utilizado são apresentados na [Figura 9](#).



	A		B	
Resíduos em regiões mais favoráveis	594	89,5%	619	87,3%
Resíduos em regiões adicionalmente favoráveis	62	9,4%	78	11,0%
Resíduos em regiões generosamente favoráveis	6	0,9%	7	1,1%
Resíduos em regiões não permitidas	1	0,2%	4	0,6%
Número total de resíduos que não prolina e glicina:	663	100,0%	709	100,0%
Número de resíduos de glicina (mostrados como triângulos)	84		84	
Número de resíduos de prolina	37		45	
Número total de resíduos (incluindo terminais)	790		840	

Figura 9: Mapa de Ramachandran para as estruturas do (A) molde e do (B) modelo. Vermelho: regiões mais favoráveis, amarelo: regiões adicionalmente favoráveis, amarelo claro: regiões generosamente favoráveis, branco: regiões não permitidas.

Métodos de avaliação de modelos baseados em perfis tridimensionais e potenciais estatísticos de força média consideram, de forma implícita, muitos dos aspectos estereoquímicos também testados nos métodos de avaliação estereoquímica (Martí-Renom *et al.*, 2000). Os programas avaliam o ambiente de cada resíduo do modelo, comparando com o ambiente esperado, conforme se apresentam em estruturas de alta resolução obtidas por cristalografia de raios X. O Verify 3D (Lüthy *et al.*, 1992) foi o programa escolhido para avaliar o modelo por perfis. Este programa faz uma análise teórica da estrutura tridimensional de cada aminoácido (perfil 3D) dentro da estrutura da proteína, sendo esta análise feita por comparação da estrutura do aminoácido junto a um banco de conformações possíveis para este aminoácido (perfil 1D). Para a comparação é computado um escore, indicando a viabilidade da conformação encontrada (escore 3D-1D). Quanto maior este escore, melhor será a estrutura. Os resultados do Verify 3D são geralmente apresentados na forma de um gráfico, com os aminoácidos no eixo x e seus escores no eixo y. Neste trabalho, preferiu-se representar os escores visualmente, nas próprias estruturas tridimensionais do molde e do modelo, por meio de ferramenta específica do GeneSilico Metaserver (Kurowski & Bujnicki, 2003). Esta opção é justificada pela dificuldade em comparar os gráficos obtidos para o molde e o modelo (devido às estruturas apresentarem tamanhos e organização diferentes), além de facilitar a visualização das regiões que necessitam de maior cuidado. Nos resultados obtidos pelo Verify 3D para o modelo construído neste trabalho em comparação com seu molde (Figura 10), nota-se que as áreas com maior desvio em relação ao esperado (representado pela cor vermelha) são, principalmente, os

segmentos de ligação entre os domínios (ausentes no molde). Por tratarem-se de regiões sem molde, maiores que vinte aminoácidos, o que é considerado o limite máximo para modelagem de alças sem molde em proteínas⁵ (Fiser *et al.*, 2000), é bastante possível que a organização espacial dessas áreas do modelo exijam refinamentos e análises mais detalhadas.

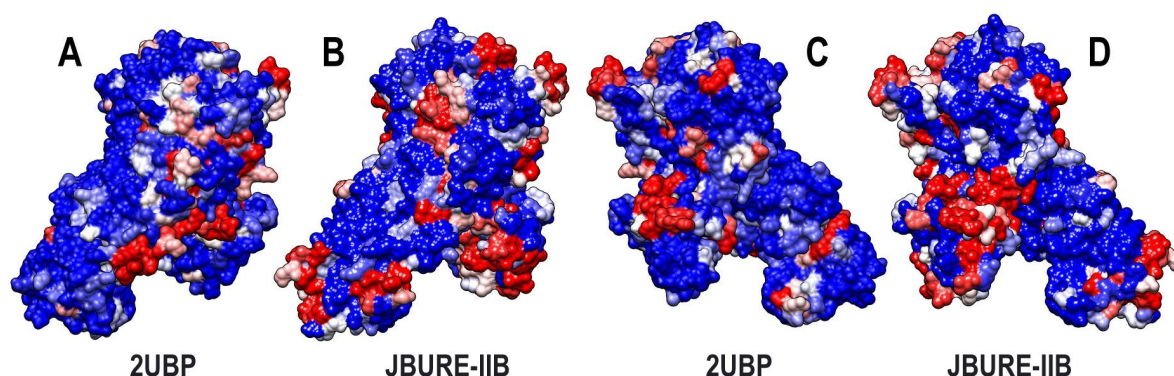


Figura 10: Resultados do Verify3D transpostos às estruturas tridimensionais do molde e do modelo. A cor azul representa áreas com conformação adequada (conforme o esperado), a cor vermelha representa áreas em não conformidade com o esperado, a cor branca representa valores intermediários entre as duas situações. (A) e (C) visualização de duas faces complementares (rotadas de 180° sobre o eixo y) da estrutura tridimensional do molde 2UBP. (B) e (D) visualização de duas faces complementares (rotadas de 180° sobre o eixo y) da estrutura tridimensional do modelo construído para a JBURE-IIB. Estruturas representadas por suas superfícies moleculares.

4.5 Estrutura tridimensional das ureases

As ureases apresentam um *fold* típico das amidohidrolases, caracterizado pela presença, no C-terminal, de um barril $(\alpha\beta)_8$ (barril TIM), onde se localiza o sítio ativo com os dois átomos de níquel, seguido por um conjunto de folhas β (Jabri *et al.*, 1995; Nagano *et al.*, 2002), ambos formando o domínio α . O domínio β apresenta predominantemente estruturas β , enquanto o domínio γ é formado por estruturas $\alpha\beta$ (Jabri *et al.*, 1995) (Figura 11).

O modelo da urease de JBURE-IIB construído neste trabalho apresenta uma estrutura muito similar àquela das ureases com estrutura tridimensional já definida (todas oriundas de bactérias), com exceção de duas regiões de ligação, entre os domínios α e β , e entre β e γ (Figuras 12 e 13). Estas regiões, por não estarem envolvidas na atividade catalítica nem na associação de subunidades, devem

⁵ São consideradas confiáveis as modelagens de alças menores que vinte aminoácidos. Em condições de simulação, alças com 4 a 8 aminoácidos foram modeladas da forma correta com uma precisão de ~90%. Esse valor caiu para 30% na modelagem de alças com 14 aminoácidos (Fiser *et al.*, 2000).

apresentar taxa de divergência mais elevada que outras regiões da proteína (Mulinari *et al.*, 2007), e estão associadas a outras propriedades das ureases, como a toxicidade. A determinação da estrutura tridimensional das regiões ligadoras da proteína é de grande importância para a compreensão das diferentes propriedades das ureases em sistemas biológicos, especialmente aquelas não encontradas em ureases bacterianas.

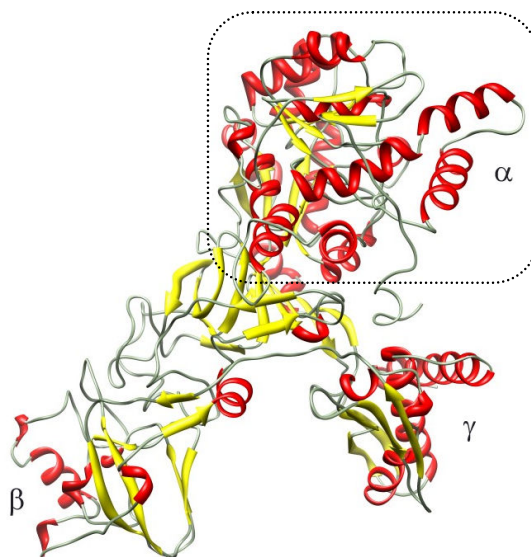


Figura 11: Estrutura tridimensional do modelo de JBURE-IIB, colorido de acordo com a estrutura secundária (α -hélices em vermelho, β -folhas em amarelo, *loops* em cinza). A região do barril TIM está destacada pelo retângulo. α , β , γ : domínios estruturais.

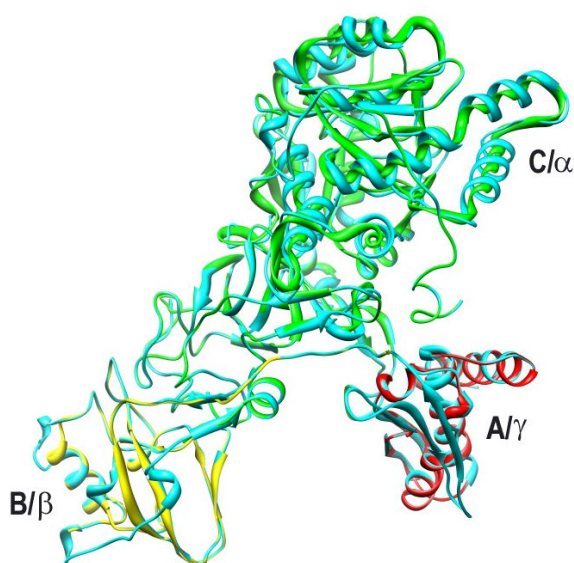


Figura 12: Sobreposição do modelo da JBUREIIB (ciano; α , β , γ : domínios estruturais) com seu molde 2UBP (vermelho: subunidade A, amarelo: subunidade B, verde: subunidade C).

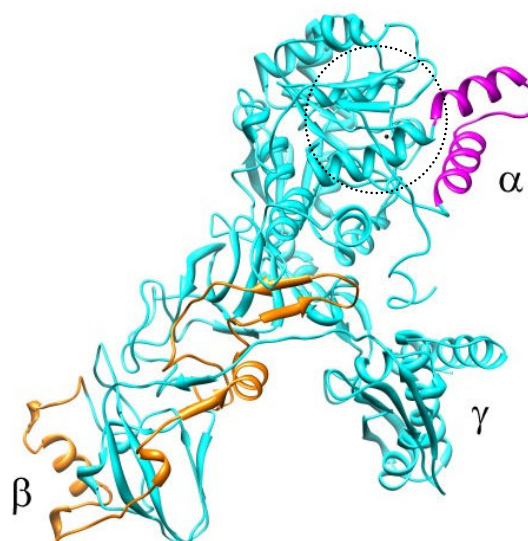


Figura 13: Estrutura tridimensional do modelo de JBURE-IIB. A região do peptídeo está destacada em laranja e a aba flexível está destacada em magenta. A região do sítio ativo está circulado. α , β , γ : domínios estruturais.

Uma estrutura do tipo hélice-volta-hélice forma uma aba (*flap*) sobre o sítio ativo, sendo associada à alta especificidade das ureases pelo seu substrato, a uréia (Jabri *et al.*, 1995; Benini *et al.*, 1999; Benini *et al.*, 2001). Esta aba, refletindo o molde utilizado, encontra-se presente e em uma conformação aberta (Benini *et al.*, 2001) no modelo construído (Figura 13). Por mecanismos ainda não totalmente elucidados, esta aba encontraria-se aberta enquanto a enzima não apresentasse substrato no sítio ativo, fechando-se de forma simultânea à aproximação e entrada da uréia (Benini *et al.*, 1999; Benini *et al.*, 2001). A movimentação desta aba seria facilitada pelas regiões sem estrutura secundária definida (alças), que são conformacionalmente mais flexíveis.

O modelo de JBURE-IIB foi utilizado, também, para localizar espacialmente os aminoácidos essenciais para a catálise, determinados como presentes em todas as seqüências analisadas neste trabalho (Figuras 14 e 15). Há, ainda, debate na literatura quanto ao critério de essencialidade para tais aminoácidos. Optou-se, neste trabalho, pela utilização de todos aqueles aminoácidos que, ao serem mutados, tornam a enzima inativa. Verificou-se que estes aminoácidos encontram-se muito próximos espacialmente, mesmo estando afastados na seqüência primária. A semelhança da organização espacial do sítio ativo do molde e do modelo permitiu o *docking* geométrico dos átomos de níquel no modelo. O *docking* foi realizado por sobreposição das estruturas 2UBP e do modelo de JBURE-IIB, com a transferência das coordenadas cartesianas correspondentes aos átomos de níquel. A identificação da estrutura 3D da apoenzima de urease de *K. aerogenes* (código PDB 1KRA) (Jabri & Karplus, 1996) indicou que o sítio ativo é pré-organizado de maneira totalmente favorável à ligação dos átomos de níquel, ocorrendo baixíssimo nível de reorganização conformacional após sua inserção na enzima.

Além dos átomos de níquel, quatro moléculas de água, organizadas de maneira específica no sítio ativo, foram adicionadas ao modelo seguindo a mesma metodologia utilizada para o níquel (Figura 14). Estas moléculas de água formam um cluster altamente organizado (Benini *et al.*, 1999; Benini *et al.*, 2001): três delas ocupam um espaço equivalente ao da uréia, enquanto uma quarta é considerada essencial para a catálise. A organização e real participação do cluster de água no sítio ativo ainda é objeto de debate na literatura.

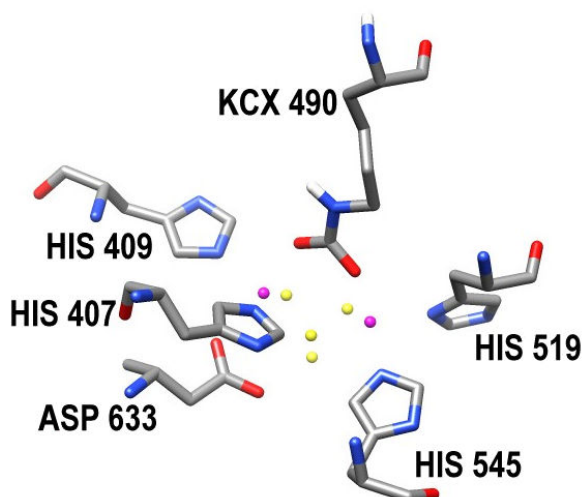


Figura 14: Representação espacial dos resíduos do sítio catalítico conservados em todas as seqüências de ureases analisadas, dos átomos de níquel (esferas em magenta) e do *cluster* de água (esferas em amarelo). A numeração dos resíduos refere-se à seqüência da JBURE-IIB (KCX: lisina carbamilada). Aminoácidos coloridos de acordo com o tipo de átomo (CPK).

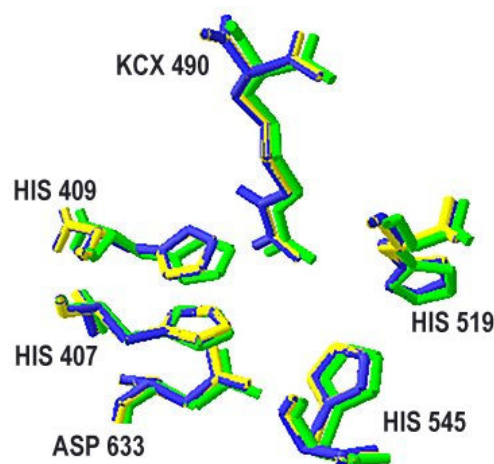


Figura 15: Sobreposição dos resíduos do sítio catalítico conservados em todas as seqüências de ureases analisadas. Em amarelo, urease de *Bacillus pasteurii* (2UBP); em azul, urease de *C. ensiformis* (JBURE-IIB); em verde, urease de *Klebsiella aerogenes* (1FWJ). A numeração dos resíduos refere-se à seqüência da JBURE-IIB (KCX: lisina carbamilada).

A associação das ureases em homo ou heteropolímeros lhes confere uma propriedade incomum: sua relação superfície/volume é uma das mais baixas conhecidas (Jabri *et al.*, 1996). Na urease de *K. aerogenes*, por exemplo, 59% de seus átomos estão ocultos, e sua superfície corresponde a 81% da superfície esperada para uma proteína de seu tamanho (Jabri *et al.*, 1996). A JBURE-IIB apresenta uma grande quantidade de aminoácidos hidrofóbicos (Figura 16), provavelmente associados à interação das subunidades (que podem formar homotrímeros e homohexâmeros). O caso da urease de *H. pylori*, que agrupa-se em dodecâmeros (Ha *et al.*, 2001), é um exemplo da especialização das ureases também no que se refere à sua estrutura quaternária.

Sondi & Salopek-Sondi (2005) verificaram diferenças significativas na formação de polimorfos de carbonato de cálcio quando na presença das ureases de *B. pasteurii* e de *C. ensiformis*. Estas diferenças foram associadas à abundância de aminoácidos carregados negativamente na superfície da urease de *B. pasteurii*, que estariam em menor quantidade na urease de *C. ensiformis*. O modelo da JBURE-IIB construído no presente trabalho está de acordo com essa observação, apresentando uma quantidade menor destes aminoácidos em sua superfície quando comparado à urease de *B. pasteurii* (Figura 16). O efeito exato destes aminoácidos sobre a formação dos polimorfos de carbonato de cálcio é ainda debatido.

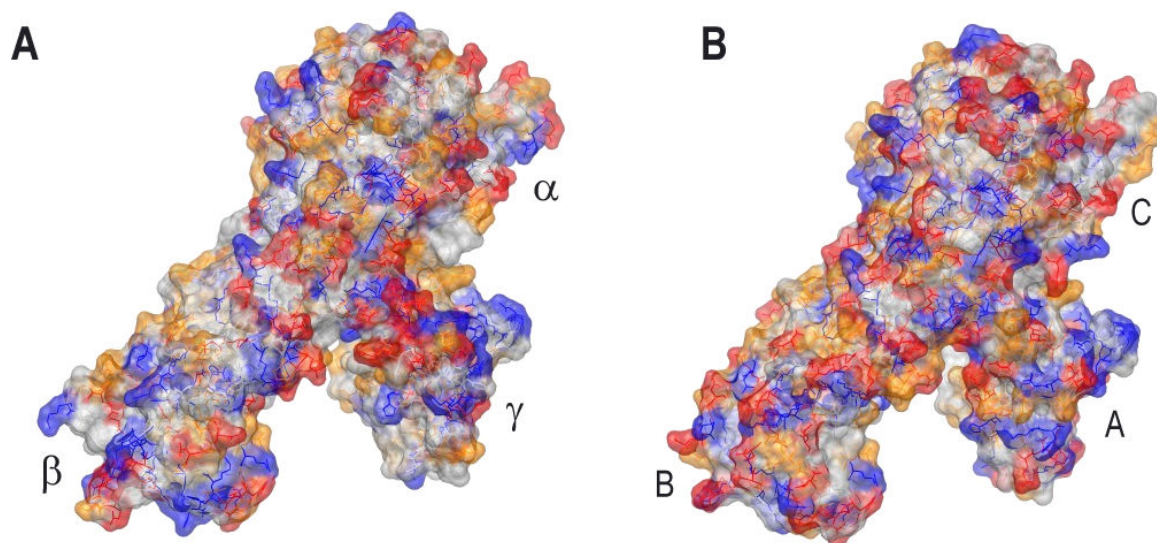


Figura 16: Distribuição das diferentes classes de aminoácidos na (A) JBURE-IIB e na (B) urease de *Bacillus pasteurii*. Aminoácidos apolares em cinza, aminoácidos polares em laranja; aminoácidos carregados negativamente em vermelho, aminoácidos carregados positivamente em azul. As estruturas estão representadas por suas superfícies moleculares em transparência; α , β , γ : domínios estruturais; A, B, C: subunidades.

A região correspondente ao peptídeo entomotóxico foi identificada espacialmente (Figura 13). A intenção ao determinar a estrutura tridimensional do peptídeo tóxico reside em obter indicações de sua ação *in vivo*, baseando-se em casos conhecidos. Sabe-se que pepcanatox não é um inibidor das enzimas digestivas dos insetos (Ferreira-DaSilva *et al.*, 2000). Porém, sua ingestão por insetos suscetíveis causa paralisia momentânea e movimentos descoordenados de antenas precedentes à morte, sugerindo que a ação do peptídeo clivado seja neurotóxica ou de indução de morte celular por alteração da permeabilidade de membranas (Mulinari *et al.*, 2007).

Ao realizar um estudo por modelagem *ab initio* de vários fragmentos do peptídeo tóxico, Mulinari *et al.* (2007) identificaram um motivo β -hairpin (grampo- β) recorrente. Em peptídeos tóxicos conhecidos, este motivo é associado à inibição de canais iônicos (como ocorre na neurotoxina charibdotoxina) e à formação de poros em membranas (Mulinari *et al.*, 2007), efeitos sugeridos para os peptídeos entomotóxicos derivados das ureases de *C. ensiformis*.

A predição da estrutura secundária pelo PSIPRED (McGuffin *et al.*, 2000) mostrou diferenças ao analisar a seqüência completa da proteína e apenas o peptídeo, de maneira isolada (Figura 17). Ao ser considerado isoladamente, o peptídeo apresentaria duas folhas- β adicionais (Lys43-Ile45, Ala78-Leu79) em relação à estrutura secundária predita para a região do peptídeo na proteína. A

definição da estrutura secundária a partir da estrutura modelada (Figura 18), realizada pelo PDBsum (Laskowski *et al.*, 2005), identificou na região do peptídeo vários β -turns (voltas- β), um γ -turn (volta- γ) e um β -hairpin. É interessante notar que a presença de β -turns, motivo associado à inserção em membranas pela toxina aerolisina, foi considerada inexistente neste peptídeo por Mulinari *et al.* (2007).

A conformação do peptídeo quando ainda inserido na urease serve apenas como indicativo de sua possível conformação ativa, pois é esperado que, ao ser liberado da cadeia protéica, ele adote uma conformação diferenciada. O estudo da conformação do peptídeo livre em solução poderia ser realizado via simulação por dinâmica molecular.

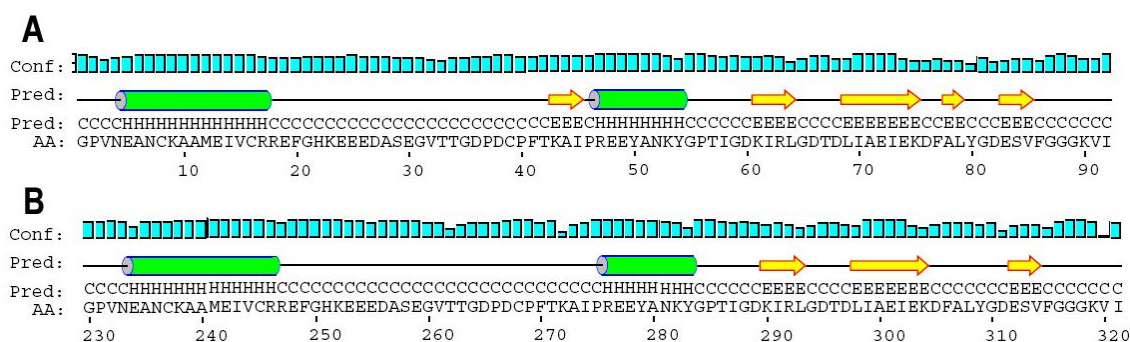
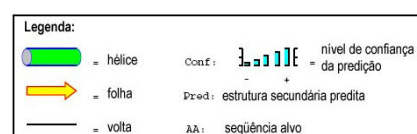


Figura 17: Predição da estrutura secundária do peptídeo entomotóxico derivado da JBURE-IIB. **(A)** predição da estrutura secundária considerando apenas o peptídeo, isolado do restante da proteína. **(B)** predição da estrutura secundária da região do peptídeo na JBURE-IIB. Adaptado de originais obtidos do PSIPRED (McGuffin *et al.*, 2000).



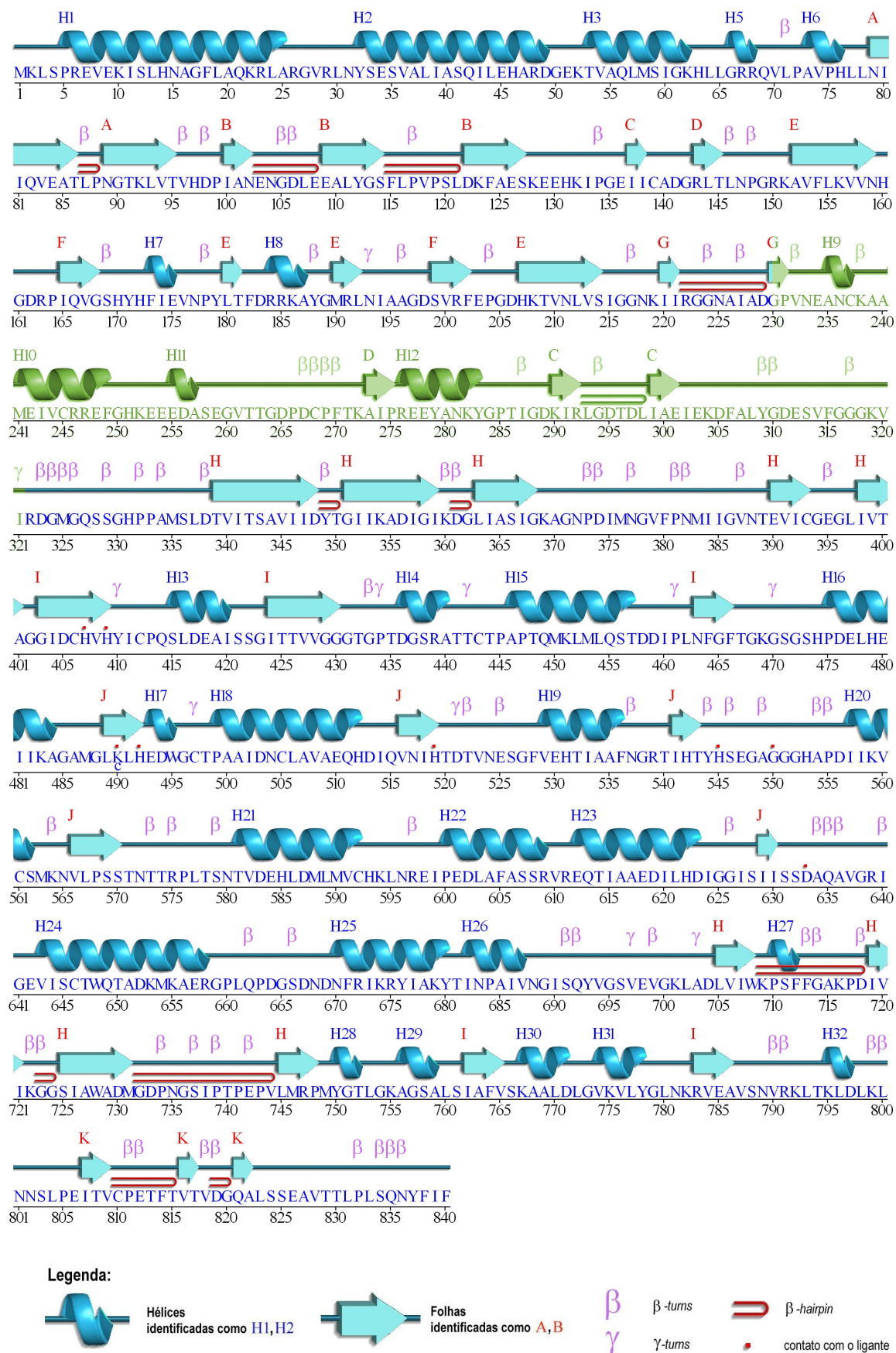


Figura 18: Determinação da estrutura secundária da JBURE-IIB a partir do modelo tridimensional construído neste trabalho. A região correspondente ao peptídeo entomotóxico encontra-se destacada em verde. Adaptado do original obtido do PDBsum (Laskowski *et al.*, 2005).

Em termos gerais, a predição da estrutura secundária a partir da seqüência primária (Figura 19) e a definição da estrutura secundária a partir da estrutura terciária (Figura 18) estão de acordo. Tomando a definição da estrutura secundária a partir da terciária como referência, as exceções de correspondência nos resultados encontram-se na ausência da α -hélice e maior quantidade de folhas- β no trecho Ile81-His160 e na ausência da folha- β entre as α -hélices do trecho Leu577-Ser608. Como a estruturação de todas as outras regiões parece muito similar, estas diferenças de estrutura secundária em relação ao esperado a partir da seqüência primária podem ser devidas ao processo de enovelamento, que propicia contatos diferenciados entre os resíduos de aminoácidos. Entretanto, a definição dos critérios para classificação de estruturas secundárias é arbitrária, apresentando variações entre os diferentes programas utilizados. Desta forma, pequenas variações entre seus resultados podem ser apenas reflexo de critérios implícitos de classificação.

5. Conclusões e perspectivas

A análise filogenética realizada no presente trabalho indicou um alinhamento preferencial das seqüências de plantas entre si e das seqüências de fungos entre si, com separação clara entre elas. Tal separação foi mais evidente na comparação das seqüências de aminoácidos de ureases. A comparação das diferentes seqüências de ureases evidenciou, também, a conservação de resíduos considerados essenciais para a atividade ureásica. A presença destes resíduos no sítio catalítico foi confirmada espacialmente na estrutura modelada para JBURE-IIB.

A utilização da estrutura 2UBP (urease de *Bacillus pasteurii*) como molde para a predição da estrutura tridimensional da JBURE-IIB mostrou-se adequada, pois levou à construção de um modelo tridimensional de boa qualidade para esta enzima. A JBURE-IIB de *Canavalia ensiformis* apresenta uma estrutura tridimensional semelhante à das ureases de bactérias, com exceção das regiões de ligação entre seus domínios. Devido à inexistência de um molde apropriado para estas regiões da proteína, especial atenção deverá ser dada ao seu refinamento. Para tanto, pretendemos utilizar técnicas de Dinâmica Molecular.

Apesar de bastante complexa, conforme apresentada no Apêndice C, espera-se que a associação dos dados teóricos obtidos neste trabalho com os dados experimentais já disponíveis ou que venham a ser obtidos no futuro, permita uma melhor compreensão das ureases em termos estruturais e cinéticos.

Como perspectivas para a continuidade do trabalho, tem-se a construção de um modelo de urease fúngica de interesse do nosso laboratório (do fungo *Cryptococcus gattii*), minimização de energia dos sistemas modelados e o refinamento das regiões de ligação entre domínios dos modelos de ureases. Estão planejados, também, estudos de dinâmica molecular dos sistemas modelados, com ênfase na conformação das alças entre domínios. A continuidade do trabalho prevê, ainda, estudos especificamente voltados ao peptídeo entomotóxico, em especial à sua conformação quando isolado da enzima, em solução. Como um objetivo final do estudo teórico das ureases esperamos obter resultados que ajudem no mapeamento das regiões protéicas envolvidas nas diferentes propriedades biológicas dessas enzimas.

6. Referências bibliográficas

- Altschul S. F., Madden T. L., Schäffer A. A., Zhang J., Zhang Z., Miller W., Lipman D. J., 1997.** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402.
- Andrews R. K., Blakeley R. L., Zerner B., 1984.** Urea and urease. *Adv. Inorg. Biochem.* **6**, 245–283.
- Balasubramanian A., Ponnuraj K., 2008.** Purification, crystallization and preliminary X-ray analysis of urease from pigeon pea (*Cajanus cajan*). *Acta Cryst.* **F64**, 662–664.
- Barrios A. M., Lippard, S. J., 2001.** Phthalazine-based dinucleating ligands afford dinuclear centers often encountered in metalloenzyme active sites. *Inorg. Chem.* **40**, 1060-1064.
- Benini S., Rypniewski W. R., Wilson K. S., Miletti S., Ciurli S., Mangani S., 1999.** A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels. *Structure.* **7**, 205-216.
- Benini S., Rypniewski W. R., Wilson K. S., Ciurli S., Mangani S., 2001.** Structure-based rationalization of urease inhibition by phosphate: novel insights into the enzyme mechanism. *J. Biol. Inorg. Chem.* **6**, 778-90.
- Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E., 2000.** The Protein Data Bank. *Nucleic Acids Research*, **28**, 235-242.
- Branden C., Tooze, J.** *Introduction to protein structure*. 2.ed. New York: Garland, 1999. p. 373-392.
- 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 cathepsin-like enzymes. *J. Econ. Entomol.* **90**, 340-348.
- Carlini C.R., Grossi-de-Sá M.F., 2002.** Plant toxic proteins with insecticidal properties. A review on their potentialities as bioinsecticides. *Toxicon.* **40**, 1515-1539.
- Chothia C., Lesk A. M., 1986.** The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5**, 823-826.
- Dixon N. E., Gazzola C., Blakeley R. L., Zerner B., 1975.** Jack bean urease (EC 3.5.1.5). A metalloenzyme. A simple biological role for nickel? *J. Am. Chem. Soc.* **97**, 4131–4133.
- Eaton K.A., Brooks C.L., Morgan D.R., Krakowka S., 1991.** Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infect. Immun.* **59**, 2470–2475.
- Estiu G., Suárez D., Merz Jr K.M., 2006.** Quantum Mechanical and Molecular Dynamics Simulations of Ureases and Zn β -Lactamases. *J. Comput. Chem.* **27**, 1240–1262.

- Eswar N., Marti-Renom M. A., Webb B., Madhusudhan M. S., Eramian D., Shen M., Pieper U., Šali A., 2000.** Comparative protein structure modeling with MODELLER. *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc., **Suppl. 15**, 5.6.1-5.6.30.
- Felsenstein J., 1985.** Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*. **39**, 783-791.
- Fernandes C. L.** *Predição da estrutura tridimensional da EPSP sintase de Mycobacterium tuberculosis utilizando modelagem molecular comparativa por homologia*. Porto Alegre: UFRGS, 2002. Monografia (Graduação em Farmácia), Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, 2002.
- 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.
- Fiser A., Do R. K. G., Sali A., 2000.** Modeling of loops in protein structures. *Protein Sci.* **9**, 1753-1773.
- Follmer C, Barcellos GB, Zingali RB, Machado OL, Alves EW, Barja-Fidalgo C, Guimarães J.A., Carlini C.R., 2001.** 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. *Biochem. J.* **360**, 217-224.
- Follmer C., Real-Guerra R., Wasserman G.E., Olivera-Severo D., Carlini, C.R., 2004.** Jackbean, soybean and *Bacillus pasteurii* ureases: biological effects unrelated to ureolytic activity. *Eur. J. Biochem.* **271**, 1357-1363.
- Guex N., Peitsch M. C., 1997.** SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis.* **18**, 2714-2723.
- Ha N.-C., Oh S.-T., Sung J. Y., Cha K. A., Lee M. H., Oh B.-H., 2001.** Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. *Nature Struct. Biol.* **8**, 505-509.
- Hall T. A., 1999.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98.
- Hasegawa M, Fujiwara M., 1993.** Relative efficiencies of the maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. *Mol. Phylogenet. Evol.* **2**, 1-5.
- Hirai M., Kawai-Hirai R., Hirai T., Ueki T., 1993.** Structural change of jack bean urease induced by addition of surfactants studied with synchrotron-radiation small-angle X-ray scattering. *Eur. J. Biochem.* **215**, 55-61.
- Hirayama C., Sugimura M., Saito H., Nakamura M., 2000a.** Host plant urease in the hemolymph of the silkworm, *Bombyx mori*. *J. Insect Physiology.* **46**, 1415-1421.
- Hirayama C., Sugimura M., Saito H., Nakamura M., 2000b.** Purification and properties of urease from the leaf of mulberry, *Morus alba*. *Phytochemistry.* **53**, 325-330.

- Jabri E., Lee M. H., Hausinger R. P., Karplus P. A., 1992.** Preliminary crystallographic studies of urease from jack bean and from *Klebsiella aerogenes*. *J. Mol. Biol.* **227**, 934–937.
- Jabri E., Carr M.B., Hausinger R.P., Karplus P.A., 1995.** The crystal structure of urease from *Klebsiella aerogenes*. *Science*. **268**, 998–1004.
- Jabri E., Karplus P. A., 1996.** Structures of the *Klebsiella aerogenes* urease apoenzyme and two active-site mutants. *Biochemistry*. **35**, 10616-10626.
- Jones D. T., 2000.** A practical guide to protein structure prediction. *Methods Mol Biol.* **143**, 131-54.
- Jukes T. H., Cantor C. R.** Evolution of protein molecules. In: **Munro H. N.** (ed.). *Mammalian Protein Metabolism*. New York: Academic Press, 1969.
- Karmali A., Domingos A., 1993.** Monoclonal antibodies against urease from *Canavalia ensiformis*. *Biochimie*. **75**, 1001–1006.
- Kurowski M. A., Bujnicki J. M., 2003.** GeneSilico protein structure prediction meta-server. *Nucleic Acids Res.* **31**, 3305-3307
- Larkin M. A., Blackshields G., Brown N. P., Chenna R., McGettigan P. A., McWilliam H., Valentin F., Wallace I. M., Wilm A., Lopez R., Thompson J. D., Gibson T. J., Higgins D. G., 2007.** ClustalW and ClustalX version 2. *Bioinformatics*, **23**(21), 2947-2948.
- Laskowski R. A., McArthur M. W., Moss D. S., Thornton J. M., 1993.** PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**, 283-291.
- Laskowski R. A., Chistyakov V. V., Thornton J. M., 2005.** PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res.*, **33**, D266-D268.
- Lüthy R., Bowie J., Eisenberg D., 1992.** Assessment of protein models with three-dimensional profiles. *Nature*, **356**, 83-85.
- Mamiya G., Takishima K., Masakuni M., Kayumi T., Ogawa K., Sekita, T., 1985.** Complete amino acid sequence of jack bean urease. *Proc. Japan Acad.* **61**, 395-398.
- Martí-Renom M. A., Stuart A. C., Fiser A., Sánchez R., Melo F.; Šali A., 2000.** Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291-325.
- McGuffin L. J., Bryson K., Jones D. T., 2000.** The PSIPRED protein structure prediction server. *Bioinformatics*. **16**, 404-405.
- Menegassi A., Wassermann G. E., Olivera-Severo D., Becker-Ritt A. B., Martinelli A. H., Feder V., Carlini C. R., 2008.** Urease from cotton (*Gossypium hirsutum*) seeds: isolation, physicochemical characterization, and antifungal properties of the protein. *J. Agric. Food Chem.* **56**, 4399-4405.
- Meyer-Bothling L.E., Polacco J.C., 1987.** Mutational analysis of the embryo-specific urease locus of soybean. *Mol. Gen. Genet.* **209**, 439–444.

- Miyaki C. Y., Russo C. A. M., Pereira S. L.** Reconstrução filogenética. Introdução ao método da máxima parcimônia. In: **Matioli, S. R.** (ed.). *Biologia Molecular e Evolução*. 1. ed. Ribeirão Preto: Holos, 2001.
- Mobley H.L., Island M.D., Hausinger R.P., 1995**, Molecular biology of microbial ureases. *Microbiol. Rev.* **59**, 451–480.
- Mulinari F., Stanisçuaski F., Bertholdo-Vargas L.R., Postal M., Oliveira-Neto O.B., Rigden 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.
- Mulinari F.** *Ureases de Canavalia ensiformis e peptídeo inseticida derivado*. Porto Alegre: UFRGS, 2008. Tese (Doutorado em Biologia Celular e Molecular), Centro de Biotecnologia do Estado do Rio Grande do Sul, Universidade Federal do Rio Grande do Sul, 2008.
- Nagano N., Orengo C. A., Thornton J. M., 2002**. One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J Mol Biol.* **321**, 741-765.
- Nei M., Kumar S.** *Molecular evolution and phylogenetics*. 1.ed. New York: Oxford, 2000. p. 73-164.
- Pedrozo H.A., Schwartz Z., Dean D.D., Wiederhold M.L., Boyan B.D., 1996a**. Regulation of statocyst mineralization in *Aplysia californica* in vitro. *Connect. Tissue Res.* **35**, 317–323.
- Pedrozo H.A., Schwartz Z., Luther M., Dean D.D., Boyan B.D., Wiederhold M.L., 1996b**. A mechanism of adaptation to hypergravity in the statocyst of *Aplysia californica*. *Hear Res.* **102**, 51–62.
- Pereira S. L., Miyaki C. Y., Russo C. A. M.** Reconstrução filogenética: métodos probabilísticos. In: **Matioli, S. R.** (ed.). *Biologia Molecular e Evolução*. 1. ed. Ribeirão Preto: Holos, 2001.
- Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., Greenblatt D. M., Meng E. C., Ferrin T. E., 2004**. UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612.
- Pires-Alves M., Grossi-de-Sá M.F., Barcellos G.B., Carlini, C.R., Moraes, M.G., 2003**. Characterization and expression of a novel member (JBURE-II) of the urease gene family from jackbean [*Canavalia ensiformis* (L.) DC]. *Plant Cell Physiol.* **44**: 139-145.
- Polacco J.C., Holland, M.A.** Roles of urease in plant cells. In **Jeon K. W. & Jarvik J.** (eds.). *International Review of Cytology*, vol. 145. San Diego: Academic Press, 1993.
- Polacco J.C., Holland M.A.** Genetic control of plant ureases. In Setlow J.K. (ed.). *Genetic Engineering*, vol. 16. New York: Plenum Press, 1994.
- Ramachandran G.N., Ramakrishnan C., Sasisekharan V., 1963**. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, **7**, 95-99.
- Riddles P.W., Whan V., Blakeley R.L., Zerner B., 1991**. Cloning and sequencing of a jack bean urease-encoding cDNA. *Gene*. **108**, 265–267.
- Ridley M.** *Evolução*. 3.ed. Porto Alegre: Artmed, 2004. p. 445-495.

- Russo C. A. M., Miyaki C. Y., Pereira S. L.** Reconstrução filogenética: métodos geométricos. In: **Matioli, S. R.** (ed.). *Biologia Molecular e Evolução*. 1. ed. Ribeirão Preto: Holos, 2001.
- Saitou N., Imanishi T., 1989.** Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. *Mol. Biol. Evol.* **6**, 514-525;
- Saitou N., Nei M., 1987.** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425.
- Šali A., 2008.** Modeller: a program for protein structure modeling - release 9v4 (manual). *Salilab*. Disponível em «<http://salilab.org/modeller/9v4/manual>» (acesso em 28out08)
- Šali A., Blundell T. L., 1993.** Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815.
- Sánchez R, Šali A., 2000.** Comparative protein structure modeling. Introduction and practical examples with Modeller. *Methods Mol. Biol.* **143**, 97-129.
- Sheridan L., Wilmot C. M., Cromie K. D., van der Logt P., Phillips S. E. V., 2002.** Crystallization and preliminary X-ray structure determination of jack bean urease with a bound antibody fragment. *Acta Cryst.* **D58**, 374–376.
- Sirko A., Brodzik R., 2000.** Plant ureases: roles and regulation. *Acta Bioch. Pol.*, **4**, 1189-1195.
- Sondi I., Salopek-Sondi B., 2005.** Influence of the primary structure of enzymes on the formation of CaCO₃ polymorphs: a comparison of plant (*Canavalia ensiformis*) and bacterial (*Bacillus pasteurii*) ureases. *Langmuir.* **21**, 8876-8882.
- Sumner J.B., 1926.** The isolation and cristallization of the enzyme urease. *J. Biol. Chem.* **69**, 435–441.
- Takishima K., Suga T., Mamiya G., 1988.** The structure of jack bean urease. The complete amino acid sequence, limited proteolysis and reactive cysteine residues. *Eur. J. Biochem.* **175**, 151–165.
- Tamura K., Dudley J., Nei M., Kumar S., 2007.** MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, **24**, 1596-1599.
- Tange Y., Niwa O., 1997.** Identification of the *ure1+* gene encoding urease in fission yeast. *Curr. Genet.* **32**, 244–246.
- Todd M. J., Hausinger R. P., 2000.** Fluoride inhibition of *Klebsiella aerogenes* urease: mechanistic implications of a pseudo-uncompetitive, slow-binding inhibitor. *Biochemistry*, **39**, 5389.
- Wheeler D. L., Barrett T., Benson D. A., Bryant S. H., Canese K., Chetvernin V., Church D. M., Dicuccio M., Edgar R., Federhen S., Feolo M., Geer L. Y., Helmsberg W., Kapustin Y., Khovayko O., Landsman D., Lipman D. J., Madden T. L., Maglott D. R., Miller V., Ostell J., Pruitt K. D., Schuler G. D., Shumway M., Sequeira E., Sherry S. T., Sirotkin K., Souvorov A., Starchenko G., Tatusov R. L., Tatusova T. A., Wagner L., Yaschenko E.,**

- 2008.** Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **36** (D), 13-21.
- Witte C.P., Tiller S., Isidore E., Davies H.V., Taylor M.A., 2005.** Analysis of two alleles of the urease gene from potato: polymorphisms, expression, and extensive alternative splicing of the corresponding mRNA. *J. Exp. Bot.* **56**: 91-99.
- Yu J.J., Smithson S.L., Thomas P.W., Kirkland T.N., Cole G.T., 1997.** Isolation and characterization of the urease gene (URE) from the pathogenic fungus *Coccidioides immitis*. *Gene.* **198**, 387–391.
- Zerner B., 1991.** Recent advances in the chemistry on an old enzyme, urease. *Bioorg. Chem.*, **19**, 116-131.
- Zonia L.E., Stebbins N.E., Polacco J.C., 1995.** Essential role of urease in germination of nitrogen- limited *Arabidopsis thaliana* seeds. *Plant Physiol.* **107**, 1097–1103.
- Zuckerlandl E., Pauling L.** Evolutionary divergence and convergence in proteins. In: **Bryson V., Vogel H. J.** (eds.). *Evolving Genes and Proteins*, New York: Academic Press, 1965.

APÊNDICES

APÊNDICE A: arquivos de entrada utilizados no programa Modeller 9v4.**A1:** arquivo *model-loop-define.py*

```

from modeller import *
from modeller.automodel import *

log.verbose()
env = environ()

env.io.atom_files_directory = ['.']

# Create a new class based on 'loopmodel' so that we can redefine
# select_loop_atoms
class MyLoop(loopmodel):
    # This routine picks the residues to be refined by loop modeling
    def select_loop_atoms(self):
        # Two residue ranges (both will be refined simultaneously)
        return selection(self.residue_range('102:', '132:'),
                        self.residue_range('234:', '264:'))

a = MyLoop(env,
           alnfile = 'alinha.ali',      # alignment filename
           knowns   = '2UBP',          # codes of the templates
           sequence = 'jbureiib',      # code of the target
           loop_assess_methods=assess.DOPE) # assess each loop with DOPE
a.starting_model= 1                    # index of the first model
a.ending_model  = 10                   # index of the last model

a.loop.starting_model = 1              # First loop model
a.loop.ending_model  = 10              # Last loop model

a.make()                                # do modeling and loop refinement

```

A2: arquivo *alinha.ali*

```

>P1;2UBP
structureX:2UBP:1 :A:570 :C: : : :
MHLNPAEKEKQLQIFLASELLRRKARGLKLNYPEAVAIITSFIMEGARDG
-KTVAMLMEEGKHLVLRDDVMEGVPEMIDDIQAEATFPDGTKLVTVHNPI
SN-----YIVPGEYRVAEAGEIEINAGR
EKTTRIRVSNVTGDRPIQVGSYHIFVEVNKELLFDRAGIGRRLNIPSGTAA
RFEPGEEMEVELTELGGNREVFGISDLTNGSVDNK---ELILQRAKELG
YKG-----VEMKINRQQYAESYGPTVGDVRLADTDLWI
EVEKDYTTYGDEVNFGGGKVLREGMGNGTYTRTENVLDDLLLTNALILDY
TGIYKADIGVKDGYIVGIGKGGNPDIMDGVTPNMIIVGTATEVIAAEGKIV
TAGGIDTHVHFINPDQVDVALANGITTLFGGGTGPAGESKATTVTPGPWN
IEKMLKSTEGPLPINVIGILGKGGSSIAPIIMEQIDAGAAGLKIHEWDGATP
ASIDRSLTVADEADVQVAIHSDTLNEAGFLEDTLRAINGRVIHSFHVEGA
GGGHAPDIMAMAGHPNVLPSTNPTRPFTVNTIDEHLDMLMVCHHLKQNI
PEDVAFADSRIRPETIAAEDILHDLGIIISMSTDALAMGRAGEMVLRWQ
TADKMKKQRGPLAEEKNGSDNFRKRYVSKY TINPAIAQGI AHEVGSIEE
GKFADLVLWEPKFFGVKADRVIKGGIIAYAQIGDPSASIPFPQVPMGRRM
YGTVGDLIHDTNITFMSKSSIQQGVPAKLGKRRIGTVKNCRNIGKKDMK
MNDVTTDIDINPETYEYVVKVDGEVLTCEPVKELPMAQRYFLE
*

```

```

>P1;jbureiib
sequence:jbureiib:1 :A:840 :A: : : :
MKLSPREVEKISLHNAGFLAQKRLARGVRLNYSSEVALIASQILEHARDG
EKTVAQLMSIGKHLGRRQVLPVPHLLNIIQVEATLPNGTKLVTVHDPI
ANENGDL EEAALYGSFLPVP SLDKFAESKEEHKIPGEIICADGRLTLNPGR
KAVFLKVVNHGDRPIQVGSYHIFIEVNPYTLFDRRKAYGMRLNIAAGDSV
RFEPGDHKTVNLVSI GGNKIIRGGNAIADGPVNEANCKAAMEIVCRREFG
HKEEEDASEGVTTGDPDCPFTKAI PREEYANKYGPTIGDKIRLGD TDLIA
EIEKDFALYGDSESVFGGGKVI RDGMGQSSGHP-PAMSLDVTITS AVIIDY
TGIKADIGIKDGLIASIGKAGNPDIMNGVFPNMIIGVNTEVICGEGGLIV
TAGGIDCHVHYICPQSLDEALSSGITTVVGGGTGPTDGSRATTCTPAPTQ
MKLMLQSTDDIPLNFGFTGKGGSGSHPDELHEIIKAGAMGLKLHEDWGCTP
AAIDNCLAVAEQHDIQVNIHTDTVNESGFVEHTIAAFNGRTIHTYHSEGA
GGGHAPDIKVCMSKKNVLPSTNTRPLTSNTVDEHLDMLMVCHKL NREI
PEDLAFASSRVREQTIAAEDILHDI GGISIISSDAQAVGRIGEVISCTWQ
TADKMAERGLQPDGSDNDNFRIKRYIAKY TINPAIVNGISQYVGSVEV
GKLADLVIWKPSFFGAKPDIVIKGGSIAWADMGPNGSIPTEPEPVMRPM
YGTGLGKAGSALSIAFVSKAALDLGVKVLVYGLNKRVEAVSNVRKLT KLDLK
LNNSLPEITVCPETFTVTVDGQALSSEAVTTLPLSQNYFIF
*

```


APÊNDICE B: listagem das seqüências aminoácidos de ureases bacterianas obtidas no presente trabalho. Os números representam a quantidade de seqüências substancialmente diferentes encontradas para cada uma das cadeias.

Organismo	Cadeia C	Cadeia B	Cadeia A	Código
<i>Acaryochloris marina</i>	1	1	1	ACAMAR
<i>Acidovorax avenae</i> subsp. <i>citrulli</i>	1	1	1	ACIAVE
<i>Acinetobacter calcoaceticus</i>	1	1	1	ACICAL
<i>Acinetobacter baumannii</i>	1	1	1	ACIBAU
<i>Actinobacillus pleuropneumoniae</i>	1	1	1	ACTPLE
<i>Actinomyces naeslundii</i>	1	1	1	ACTNAE
<i>Agrobacterium tumefaciens</i>	1	1	1	AGRTUM
<i>Alcanivorax borkumensis</i>	1	1	1	ALCBOR
<i>Alkalilimnicola ehrlichei</i>	1	1	1	ALKAHR
<i>Alteromonas macleodii</i> 'Deep ecotype'	1	1	1	ALTMAC
<i>Anabaena variabilis</i>	1	1	1	ANAVAR
<i>Anaeromyxobacter</i> sp. Fw109-5	1	0	1	ANAESP
<i>Arcobacter butzleri</i>	1	0	0	ARCBUT
<i>Arthrobacter aurescens</i>	1	1	1	ARTAUR
<i>Arthrobacter</i> sp. FB24	1	1	1	ARTHSP
<i>Arthrospira maxima</i>	0	1	0	ARTMAX
<i>Aurantimonas</i> sp. SI85-9A1	1	1	1	AURASP
<i>Aureococcus anophagefferens</i>	1	0	0	AURANO
<i>Azoarcus</i> sp. BH72	1	1	1	AZOASP
<i>Azorhizobium caulinodans</i>	1	1	1	AZOCAU
<i>Azotobacter vinelandii</i>	1	1	1	AZOVIN
<i>Bacillus amyloliquefaciens</i>	1	1	1	BACAMY
<i>Bacillus cereus</i>	1	1	1	BACCER
<i>Bacillus coagulans</i>	1	0	0	BACCOA
<i>Bacillus halodurans</i>	1	1	1	BACHAL
<i>Bacillus pasteurii</i>	1	1	1	BACPAS
<i>Bacillus</i> sp. B14905	1	1	1	BACISP
<i>Bacillus subtilis</i>	1	1	1	BACSUB
<i>Bacillus subtilis</i> subsp. <i>subtilis</i>	1	1	1	BACSSU
<i>Blochmannia floridanus</i> [candidatus]	1	1	1	BLOFLO
<i>Blochmannia pennsylvanicus</i> [candidatus]	1	1	1	BLOPEN
<i>Bordetella bronchiseptica</i>	1	1	1	BORBRO
<i>Bordetella parapertussis</i>	1	1	1	BORPAR
<i>Bordetella pertussis</i>	1	1	1	BORPER
<i>Bradyrhizobium japonicum</i>	1	1	1	BRAJAP
<i>Brucella abortus</i> (<i>Brucella melitensis</i> biovar <i>Abortus</i>)	2	2	2	BRUABO
<i>Brucella melitensis</i>	2	2	2	BRUMEL
<i>Brucella ovis</i>	2	2	2	BRUOVI
<i>Brucella suis</i> (<i>Brucella melitensis</i> biovar <i>Suis</i>)	2	2	2	BRUSUI
<i>Burkholderia ambifaria</i>	1	1	1	BURAMB
<i>Burkholderia cenocepacia</i>	1	1	1	BURCEN
<i>Burkholderia cepacia</i>	1	1	1	BURCEP
<i>Burkholderia dolosa</i>	1	1	1	BURDOL
<i>Burkholderia mallei</i>	1	1	1	BURMAL
<i>Burkholderia multivorans</i>	1	1	1	BURMUL
<i>Burkholderia phymatum</i>	1	1	1	BURPHY

<i>Burkholderia phytofirmans</i>	2	1	1	BURPHT
<i>Burkholderia pseudomallei</i>	1	1	1	BURPSE
<i>Burkholderia thailandensis</i>	1	1	1	BURTHA
<i>Burkholderia vietnamiensis</i>	1	1	1	BURVIE
<i>Burkholderia xenovorans</i>	1	1	1	BURXEN
<i>Caldicellulosiruptor saccharolyticus</i>	1	1	1	CALSAC
<i>Cenarchaeum symbiosum</i>	1	1	1	CENSYM
<i>Chlamydomonas</i> sp. CCMP 222	1	0	0	CHLASP
<i>Chromohalobacter salexigens</i>	1	1	1	CHRSAL
<i>Clostridium perfringens</i>	1	1	1	CLOPER
<i>Clostridium phytofermentans</i>	1	1	1	CLOPHY
<i>Clostridium thermocellum</i>	1	1	1	CLOTHE
<i>Comamonas testosteroni</i>	1	1	1	COMTES
<i>Congregibacter litoralis</i>	1	1	1	CONLIT
<i>Corynebacterium efficiens</i>	1	1	1	COREFF
<i>Corynebacterium glutamicum</i>	1	1	1	CORGLU
<i>Cytophaga hutchinsonii</i>	1	0	0	CYTHUT
<i>Dechloromonas aromatica</i>	1	1	1	DECARO
<i>Deinococcus radiodurans</i>	1	0	0	DEIRAD
<i>Delftia acidovorans</i>	1	1	1	DELACI
<i>Dinoroseobacter shibae</i>	1	1	1	DINSHI
<i>Edwardsiella ictaluri</i>	1	1	1	EDWICT
<i>Enterobacter</i> sp. 638	1	1	1	ENTESP
<i>Escherichia coli</i> O157:H7	1	1	1	ESCCOL
<i>Flavobacterium johnsoniae</i>	1	1	1	FLAJOH
<i>Flexispira rappini</i> [nota: helicobacteriaceae]	1	0	0	FLERAP
<i>Frankia alni</i>	1	1	1	FRAALN
<i>Frankia</i> sp. EAN1pec	1	1	1	FRANSP
<i>Fulvimarina pelagi</i>	1	1	1	FULPEL
<i>Geobacillus kaustophilus</i>	1	1	1	GEOKAU
<i>Granulibacter bethesdensis</i>	1	1	1	GRABET
<i>Haemophilus influenzae</i>	1	1	1	HAEINF
<i>Hahella chejuensis</i>	1	1	1	HAHCHE
<i>Haloarcula aidinensis</i>	1	1	0	HALAID
<i>Haloarcula argentinensis</i>	1	1	0	HALARG
<i>Haloarcula californiae</i>	1	1	0	HALCAL
<i>Haloarcula hispanica</i>	1	1	0	HALHIS
<i>Haloarcula japonica</i>	1	1	0	HALJAP
<i>Haloarcula marismortui</i>	1	1	1	HALMAR
<i>Haloarcula quadrata</i>	1	1	0	HALQUA
<i>Haloarcula sinaiensis</i>	1	1	0	HALSIN
<i>Haloarcula vallismortis</i>	1	1	0	HALVAL
<i>Haloquadratum walsbyi</i>	1	1	1	HALWAL
<i>Helicobacter acinonychis</i>	1	1	0	HELACI
<i>Helicobacter baculiformis</i>	1	1	0	HELBAC
<i>Helicobacter bilis</i>	1	0	0	HELBIL
<i>Helicobacter bizzozeronii</i>	1	1	0	HELBIZ
<i>Helicobacter cynogastricus</i>	1	1	0	HELCCYN
<i>Helicobacter felis</i>	2	2	0	HELFEL
<i>Helicobacter hepaticus</i>	1	1	0	HELHEP
<i>Helicobacter mustelae</i>	1	1	0	HELMUS
<i>Helicobacter pylori</i>	1	1	0	HELPYL

<i>Helicobacter salomonis</i>	1	1	0	HELSAL
<i>Helicobacter</i> sp. TD1	1	1	0	HELISP
<i>Helicobacter suis</i> [candidatus]	1	0	0	HELSUI
<i>Herpetosiphon aurantiacus</i>	1	1	1	HERAUR
<i>Jannaschia</i> sp. CCS1	1	1	1	JANNSP
<i>Kineococcus radiotolerans</i>	1	0	0	KINRAD
<i>Klebsiella aerogenes</i>	1	1	1	KLEAER
<i>Klebsiella pneumoniae</i>	1	1	1	KLEPNE
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>	1	1	1	KLEPPN
<i>Lactobacillus fermentum</i>	1	1	1	LACFER
<i>Lactobacillus reuteri</i>	1	1	1	LACREU
<i>Loktanella vestfoldensis</i>	1	1	1	LOKVES
<i>Lyngbya</i> sp. PCC 8106	1	1	1	LYNGSP
<i>Magnetococcus</i> sp. MC-1	1	1	1	MAGNSP
<i>Magnetospirillum gryphiswaldense</i>	1	1	1	MAGGRY
<i>Marinobacter algicola</i>	1	1	1	MARALG
<i>Marinobacter aquaeolei</i>	1	1	1	MARAQU
<i>Marinobacter</i> sp. ELB17	1	1	1	MARISP
<i>Marinomonas</i> sp. MED121	1	0	1	MARDSP
<i>Marinomonas</i> sp. MWYL1	1	1	1	MARWSP
<i>Mesorhizobium loti</i>	1	1	1	MESLOT
<i>Mesorhizobium</i> sp. BNC1	1	1	1	MESOSP
<i>Metallosphaera sedula</i>	1	0	1	METSED
<i>Methylibium petroleiphilum</i>	1	1	1	METPET
<i>Methylobacillus flagellatus</i>	1	1	1	METFLA
<i>Methylobacterium chloromethanicum</i>	3	1	3	METCHL
<i>Methylobacterium extorquens</i>	3	1	3	METEXT
<i>Methylobacterium populi</i>	3	1	3	METPOP
<i>Methylobacterium</i> sp. 4-46	1	0	1	METHSP
<i>Minibacterium massiliensis</i>	1	1	1	MINMAS
<i>Moritella</i> sp. PE36	1	1	1	MORISP
<i>Mycobacterium bovis</i>	1	1	1	MYCBOV
<i>Mycobacterium gilvum</i>	2	2	1	MYCGIL
<i>Mycobacterium smegmatis</i>	2	1	1	MYCSME
<i>Mycobacterium tuberculosis</i>	1	1	1	MYCTUB
<i>Mycobacterium ulcerans</i>	1	1	1	MYCULC
<i>Mycobacterium vanbaalenii</i>	1	1	1	MYCVAN
<i>Natronomonas pharaonis</i>	1	1	1	NATPHA
<i>Neptuniibacter caesariensis</i>	1	1	1	NEPCAE
<i>Nitrosococcus oceani</i>	1	1	1	NITOCCE
<i>Nitrosomonas cryotolerans</i>	1	0	0	NITCRY
<i>Nitrosospira briensis</i>	1	0	0	NITBRI
<i>Nitrosospira multiformis</i>	1	1	1	NITMUL
<i>Nitrosospira tenuis</i>	1	0	0	NITTEN
<i>Nocardia farcinica</i>	1	1	1	NOCFAR
<i>Nodularia spumigena</i>	1	1	1	NODSPU
<i>Nostoc muscorum</i>	1	1	1	NOSMUS
<i>Nostoc punctiforme</i>	2	1	1	NOSPUN
<i>Oceanicola batsensis</i>	2	1	1	OCEBAT
<i>Oceanicola granulosus</i>	1	1	1	OCEGRA
<i>Oceanobacter</i> sp. RED65	1	1	1	OCEASP
<i>Ochrobactrum anthropi</i>	2	2	2	OCHANT

<i>Paracoccus denitrificans</i>	1	1	1	PARDEN
<i>Photobacterium damsela</i> subsp. <i>damsela</i>	1	0	0	PHODDA
<i>Photobacterium profundum</i>	1	1	1	PHOPRO
<i>Photorhabdus luminescens</i> subsp. <i>laumondii</i>	1	1	1	PHOLLA
<i>Polaromonas naphthalenivorans</i>	1	1	1	POLNAP
<i>Polaromonas</i> sp. JS666	1	1	1	POLASP
<i>Polynucleobacter</i> sp. QLW-P1DMWA-1	1	1	1	POLYSP
<i>Prochlorococcus marinus</i>	1	1	1	PROMAR
<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i>	1	1	1	PROMPA
<i>Proteus mirabilis</i>	1	1	1	PROMIR
<i>Proteus vulgaris</i>	1	1	1	PROVUL
<i>Pseudoalteromonas haloplanktis</i>	1	1	1	PSEHAL
<i>Pseudomonas aeruginosa</i>	1	1	1	PSEAER
<i>Pseudomonas entomophila</i>	1	1	1	PSEENT
<i>Pseudomonas fluorescens</i>	1	1	1	PSEFLU
<i>Pseudomonas mendocina</i>	1	1	1	PSEMEN
<i>Pseudomonas putida</i>	1	1	1	PSEPUT
<i>Pseudomonas stutzeri</i>	1	1	1	PSESTU
<i>Pseudomonas syringae</i> pv. <i>phaseolicola</i>	1	1	1	PSESYR
<i>Psychrobacter cryohalolentis</i>	2	2	2	PSYCRY
<i>Psychromonas ingrahamii</i>	1	1	1	PSYING
<i>Ralstonia eutropha</i>	1	1	1	RALEUT
<i>Ralstonia metallidurans</i>	1	1	1	RALMET
<i>Ralstonia pickettii</i>	1	1	1	RALPIC
<i>Ralstonia solanacearum</i>	1	1	1	RALSOL
<i>Reinekea</i> sp. MED297	1	1	1	REINSP
<i>Rhizobium etli</i>	1	1	1	RHIETL
<i>Rhizobium leguminosarum</i>	1	1	1	RHILEG
<i>Rhodobacter capsulatus</i>	1	1	1	RHOCAP
<i>Rhodobacter sphaeroides</i>	1	1	1	RHOSPH
<i>Rhodococcus</i> sp. RHA1	1	1	1	RHODSP
<i>Rhodoferax ferrireducens</i>	1	1	1	RHOFER
<i>Rhodomonas salina</i>	1	0	0	RHOSAL
<i>Rhodopseudomonas palustris</i>	1	1	1	RHOPAL
<i>Roseobacter denitrificans</i>	1	1	1	ROSDEN
<i>Roseobacter</i> sp. AzwK-3b	1	1	1	ROSESP
<i>Saccharophagus degradans</i>	1	1	1	SACDEG
<i>Saccharopolyspora erythraea</i>	1	1	1	SACERY
<i>Saccharopolyspora spinosa</i>	1	1	1	SACSPI
<i>Sagittula stellata</i>	1	1	1	SAGSTE
<i>Silicibacter pomeroyi</i>	1	1	1	SILPOM
<i>Silicibacter</i> sp. TM1040	1	1	1	SILISP
<i>Sinorhizobium medicae</i>	1	1	1	SINMED
<i>Sinorhizobium meliloti</i>	1	1	1	SINMEL
<i>Sporosarcina pasteurii</i>	1	1	1	SPOPAS
<i>Staphylococcus aureus</i>	1	1	1	STAAUR
<i>Staphylococcus aureus</i> subsp. <i>aureus</i>	1	1	1	STAAAU
<i>Staphylococcus epidermidis</i>	1	1	1	STAEPI
<i>Staphylococcus leei</i>	1	1	0	STALEE
<i>Staphylococcus saprophyticus</i> subsp. <i>saprophyticus</i>	1	1	1	STASSA
<i>Staphylococcus xylosus</i>	1	1	1	STAXYL
<i>Stappia aggregata</i>	1	1	1	STAAGG

<i>Stigmatella aurantiaca</i>	1	0	0	STIAUR
<i>Streptococcus alactolyticus</i>	1	0	0	STRALA
<i>Streptococcus hyointestinalis</i>	1	0	0	STRHYO
<i>Streptococcus salivarius</i>	1	1	1	STRSAL
<i>Streptococcus thermophilus</i>	1	1	1	STRTHE
<i>Streptococcus vestibularis</i>	1	1	1	STRVES
<i>Streptomyces avermitilis</i>	2	1	1	STRAVE
<i>Streptomyces coelicolor</i>	2	1	1	STRCOE
<i>Streptomyces kanamyceticus</i>	1	0	0	STRKAN
<i>Sulfitobacter</i> sp. EE-36	1	1	1	SULESP
<i>Sulfitobacter</i> sp. NAS-14.1	1	1	1	SULNSP
<i>Sulfolobus tokodaii</i>	1	0	0	SULTOK
<i>Synechococcus elongatus</i>	1	1	1	SYNELO
<i>Synechocystis</i> sp. PCC 6803	1	1	1	SYNESP
<i>Thermosynechococcus elongatus</i>	1	1	1	THEELO
<i>Trichodesmium erythraeum</i>	1	1	1	TRIERY
<i>Ureaplasma parvum</i>	1	1	1	UREPAR
<i>Ureaplasma urealyticum</i>	1	1	1	UREURE
<i>Verminephrobacter eiseniae</i>	1	1	1	VEREIS
<i>Vibrio fischeri</i>	1	1	1	VIBFIS
<i>Vibrio parahaemolyticus</i>	1	1	1	VIBPAR
<i>Vibrio shilonii</i>	1	1	1	VIBSHI
<i>Vibrio splendidus</i>	1	1	1	VIBSPL
<i>Xanthobacter autotrophicus</i>	1	0	1	XANAUT
<i>Yersinia aldovae</i>	1	1	1	YERALD
<i>Yersinia bercovieri</i>	1	1	1	YERBER
<i>Yersinia enterocolitica</i>	1	1	1	YERENT
<i>Yersinia enterocolitica</i> subsp. <i>enterocolitica</i>	1	1	1	YEREEN
<i>Yersinia frederiksenii</i>	1	1	1	YERFRE
<i>Yersinia intermedia</i>	1	1	1	YERINT
<i>Yersinia kristensenii</i>	1	1	1	YERKRI
<i>Yersinia mollaretii</i>	1	1	1	YERMOL
<i>Yersinia pestis</i>	1	1	1	YERPES
<i>Yersinia pseudotuberculosis</i>	1	1	1	YERPSE
<i>Yersinia rohdei</i>	1	1	1	YERROH

APÊNDICE C: Tabela comparativa de dados cinéticos para ureases.

A tabela apresenta dados compilados a partir de aproximadamente 200 artigos referentes a aspectos bioquímicos das ureases. Por terem sua comparação dificultada (devido a condições experimentais diferentes em cada artigo, por exemplo), estes dados não foram considerados na elaboração da versão final do presente trabalho. A tabela tem caráter preliminar em relação à distribuição do conteúdo e sua formatação. As referências citadas na tabela encontram-se listadas abaixo da mesma.

Organismo	K_m	K_{cat} e pH	K_i
<i>Bacillus pasteurii</i>	Km 40-130 mM [8, 13] Km 55.3 mM [21,22]		Quinonas: (non-competitive, K _i em mM) [15] p-Benzoquinone 80 p-Floranil 0.1 p-Chloranil 0.2 o-Chloranil 4.8 p-Bromanil 0.15 p-Iodanil 1.0 Di-Br,di-I-benzoquinone 0.2 Chloranilic acid >1000 2-CH ₃ ,-6-OCH ₃ -benzoquinone 200 2,3-Cl,-5,6-CN-benzoquinone 600 o-Naphthoquinone 70 p-Naphthoquinone 10 2,3-Di-Cl-naphthoquinone 13
<i>Canavalia ensiformis</i> : JBU	Km (pH 6.5) 1.5±0. mM [16] Km (pH 7.5) 2.9±0.6 mM [16] Km (pH 8.5) 7.1±1.7 mM [16] Km 2-3.5 mM [16, 17]	Kcat [Molecular mass of 270 kDa was taken as the minimal molecular form expressing enzymic activity for urease]: (pH 6.5) 9700 s ⁻¹ [16] (pH 7.5) 61650 s ⁻¹ [16] (pH 8.5) 120150 s ⁻¹ [16] pH ótimo: 7 [19,20]	K _i (Boric acid , B(OH) ₃ , acts as a competitive inhibitor): 0.12 mM [2,3] Phosphate is a competitive inhibitor in the pH range 5.8-7.5, with K _i values increasing from 0,53mM at pH 5.8 to 123mM at pH 7.5. [6,5] Quinonas: (non-competitive, K _i em mM) [15] p-Benzoquinone 10 p-Floranil 0.1 p-Chloranil 0.2 o-Chloranil 3,5 p-Bromanil 0.3 p-Iodanil 0.1 Di-Br,di-I-benzoquinone 0.1 Chloranilic acid 100 2-CH ₃ ,-6-OCH ₃ -benzoquinone 120 2,3-Cl,-5,6-CN-benzoquinone 70 o-Naphthoquinone 5 p-Naphthoquinone 7 2,3-Di-Cl-naphthoquinone 11 ** ****
<i>Canavalia ensiformis</i> : Canatoxina	Km (pH 6.5) 1.5±0.2 mM [16] Km (pH 7.5) 1.8±0.1 mM [16] Km (pH 8.5) 7.7±1.5 mM [16] Km 2-7 mM [16]	Kcat [Molecular mass of 184 kDa was taken as the minimal molecular form expressing enzymic activity for canatoxin]: (pH 6.5) 9200 s ⁻¹ [16] (pH 7.5) 18093 s ⁻¹ [16] (pH 8.5) 43240 s ⁻¹ [16]	
<i>Glycine max</i> : embryo-specific	Km 0.2-0.6 mM [17]		
<i>Helicobacter</i>	Km 0,17mM [8, 9]		*

<i>pylori</i>	Km 0,2-0,3 mM [18] Km 2,4 ±0.1 mM [1]	Kcat (The kcat is based on Mr=83000 for the urease catalytic unit.): 2970 ±100 s ⁻¹ [1] pH ótimo: 8 [1]	*** Phosphate: within the pH range 5.0-7.0 the inhibition is pH dependent, with values of Ki increasing from circa 0.1 mM at pH 5 to ca. 50 mM at pH 7. The inhibition is weak and not purely competitive at pH>7.0, while the enzyme is labile at pH<5.0. [6,5] Ki (Boric acid , B(OH) ₃ , acts as a competitive inhibitor): 0.33 mM [2,5] Ki (aceto hydroxamic acid : slow, tight-binding competitive inhibitor): 2.6 μM. [8, 14] Ki (phenylphosphorodiamidate : slowly binding competitive inhibitor): 94 pM [8, 14] *****
<i>Morus alba</i>	Km 0.16mM [19,20, 23]	pH ótimo: 9 [23]	
<i>Proteus mirabilis</i>	Km 13 mM [8, 11]		Ki (Boric acid , B(OH) ₃ , acts as a competitive inhibitor): 0.1mM [2,4]
<i>Selenomonas ruminantium</i>	Km 2,2mM [8, 10]		
<i>Sporosarcina ureae</i>	Km 60 mM [8, 12]		

Notas gerais: microbial ureases possess an optimum pH of near neutrality and are often irreversibly denatured by exposure to pH values below 5. [8]

* IC₅₀ de vários inibidores dipeptídicos com ácido hidroxâmico na referência [24].

** Dados sobre inibição com compostos experimentais contendo organobismuto na referência [25].

*** Dados sobre inibição com fenilfosforodiamidatos na referência [26].

**** Dados sobre inibição com cetonas-a-b-insaturadas na referência [27].

***** Dados sobre inibição com fluoretos na referência [28].

Atividades específicas de algumas: [8] pg459.

Referências:

- 1) Pearson, A.M. et al., 2000. Kinetic and Structural Characterization of Urease Active Site Variants. *Biochemistry*, 39: 8575-8584.
- 2) Benini, S et al., 2004. Molecular Details of Urease Inhibition by Boric Acid: Insights into the Catalytic Mechanism. *J. Am. Chem. Soc.*, 126: 3714-3715
- 3) Krajewska, B. et al., 1999. *Polish J. Chem.* 73: 359-366.
- 4) Breitenbach, J. M.; Hausinger, R. H. 1988. *Biochem. J.*, 250: 917-920.
- 5) Todd, M. J.; Hausinger, R. P. 1989. *J. Biol. Chem.*, 264: 15835-15842.
- 6) Benini, S et al., 2001. Structure-based rationalization of urease inhibition by phosphate: novel insights into the enzyme mechanism. *J. Biol. Inorg. Chem.*, 6: 778-790
- 7) Krajewska, B.; Zaborska, W., 1999. *J. Mol. Catal. B.* 6: 75-81.
- 8) Mobley, H.L.T.; Island, M.D.; Hausinger, R.P., 1995. Molecular Biology of Microbial Ureases. *Microbiol. Rev.*, 59: 451-480.
- 9) Hu, L.-T., and H. L. T. Mobley. 1990. Purification and N-terminal analysis of urease from *Helicobacter pylori*. *Infect. Immun.* 58:992-998.
- 10) Hausinger, R. P. 1986. Purification of a nickel-containing urease from the rumen anaerobe *Selenomonas ruminantium*. *J. Biol. Chem.* 261:7866-7870.
- 11) Breitenbach, J. M., and R. P. Hausinger. 1988. *Proteus mirabilis* urease: partial purification and inhibition by boric acid and boronic acids. *Biochem. J.* 250:917-920.
- 12) McCoy, D. D., A. Cetin, and R. P. Hausinger. 1992. Characterization of urease from *Sporosarcina ureae*. *Arch. Microbiol.* 157:411-416.
- 13) Larson, A. D., and R. E. Kallio. 1954. Purification and properties of bacterial urease. *J. Bacteriol.* 68:67-73.

- 14) Todd, M. J., and R. P. Hausinger. 1989. Competitive inhibitors of *Klebsiella aerogenes* urease. Mechanisms of interaction with the nickel active site. *J. Biol. Chem.* 264:15835–15842.
- 15) Ashiralieva, A.; Kleiner, D., 2003. Polyhalogenated benzo- and naphthoquinones are potent inhibitors of plant and bacterial ureases. *FEBS Letters.* 555:367-370.
- 16) Follmer, C. et al., 2001. 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. *Biochem. J.* 360:217-224
- 17) Follmer, C. et al., 2004. Jackbean, soybean and *Bacillus pasteurii* ureases: Biological effects unrelated to ureolytic activity. *Eur. J. Biochem.* 271:1357–1363
- 18) Mobley, H.L.T.; Hu, L.-T.; Foxall, P.A., 1991. *Helicobacter pylori* urease: properties and role in pathogenesis. *Scand. J. Gastroenterol. Suppl.* 187:39-46
- 19) Hirayama, C. et al., 2000. Host plant urease in the hemolymph of the silkworm, *Bombyx mori*. *Journal of Insect Physiology.* 46:1415–1421
- 20) Fishbein, W.N., 1969. Urease catalysis. III. Stoichiometry, kinetics, and inhibitory properties of a third substrate dihydroxyurea. *Journal of Biological Chemistry.* 244:1188–1193.
- 21) Ha, N.-C. et al., 2001. Supramolecular assembly and acid resistance of *Helicobacter pylori* urease. *Nature Structural Biology.* 8:505-509.
- 22) Morsdorf, G. & Kaltwasser, H.. 1989. *Arch. Microbiol.* 152:125–131.
- 23) Hirayama, C. et al., 2000. Purification and properties of urease from the leaf of mulberry, *Morus alba*. *Phytochemistry.* 53: 325-330.
- *24) Mishra, H.; Parrill, A.L.; Williamson, J.S., 2002. Three-Dimensional Quantitative Structure-Activity Relationship and Comparative Molecular Field Analysis of Dipeptide Hydroxamic Acid *Helicobacter pylori* Urease Inhibitors. *Antimicrobial Agents and Chemotherapy*, 8:2613–2618
- **25) Murafuji, T. et al., 2006. Inhibition of jack bean urease by organobismuth compounds. *Bioorganic & Medicinal Chemistry Letters.* 16:1510–1513
- ***26) Faraci, W.S. et al., 1995. Inhibition of *Helicobacter pylori* Urease by Phenyl Phosphorodiamidates: Mechanism of Action. *Bioorganic & Medicinal Chemistry.* 5:605-610
- ****27) Tanaka, T.; Kawase, M.; Tani, S., 2003. Urease inhibitory activity of simple α,β -unsaturated ketones. *Life sciences.* 73: 2985–2990
- *****28) Todd, M.J.; Hausinger, R.P., 2000. Fluoride Inhibition of *Klebsiella aerogenes* Urease: Mechanistic Implications of a Pseudo-uncompetitive, Slow-Binding Inhibitor. *Biochemistry.* 39:5389-5396