

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
DISSERTAÇÃO DE MESTRADO**

SHEYLA CARLA BARBOSA DA SILVA LIMA

**Isolamento e caracterização *in silico* de ciclotídeos em
milho (*Zea mays*) e centeio (*Secale cereale*)**

**Recife
2015**

SHEYLA CARLA BARBOSA DA SILVA LIMA

Isolamento e caracterização *in silico* de ciclotídeos em milho (*Zea mays*) e centeio (*Secale cereale*)

Dissertação apresentada ao programa de Pós Graduação em Ciências Biológicas da Universidade Federal de Pernambuco, como requisito final exigido para a obtenção do título de Mestre em Ciências Biológicas, área de concentração: Biotecnologia.

Orientadora: Prof^a Dr^a Valesca Pandolfi

Coorientadora: Prof^a Dr^a Ana Maria Benko Iseppon

**Recife
2015**

Catálogo na fonte
Elaine Barroso
CRB 1728

Lima, Sheyla Carla Barbosa da Silva

Isolamento e caracterização *in silico* de ciclotídeos em milho (*Zea mays*) e centeio (*Secale cereale*) Sheyla Carla Barbosa da Silva Lima—
Recife: O Autor, 2015.

149 folhas: il., fig., tab.

Orientadora: Valesca Pandolfi

Coorientadora: Ana Maria Benko Iseppon

Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Ciências Biológicas. Biotecnologia, 2015.

Inclui bibliografia e anexos

1. Bioinformática 2. Peptídeos 3. Gramínea I. Pandolfi, Valesca (orientadora) II. Iseppon, Ana Maria Benko (coorientador) III. Título

660.6

CDD (22.ed.)

UFPE/CCB-2015-203

**Isolamento e caracterização *in silico* de ciclotídeos
em milho (*Zea mays*) e centeio (*Secale cereale*)**

Dissertação apresentada ao programa de Pós Graduação em Ciências Biológicas da Universidade Federal de Pernambuco, como requisito final exigido para a obtenção do título de Mestre em Ciências Biológicas, área de concentração: Biotecnologia.

Data de Aprovação: 24/02/2015

COMISSÃO EXAMINADORA

Profa. Dr^a Valesca Pandolfi

(Orientadora - Universidade Federal de Pernambuco)

Profa. Dr^a. Ana Maria Benko-Iseppon

(Membro interno - Universidade Federal de Pernambuco)

Dr. José Ribamar Costa Ferreira Neto

(Membro externo - Universidade Federal de Pernambuco)

**Recife
2015**

À minha família, por todo amor, dedicação e apoio.

Dedico.

AGRADECIMENTOS

Primeiramente, agradeço a Deus pelo dom da vida e por me proporcionar mais uma conquista, apesar de tantos obstáculos Ele está ao meu lado guiando-me e dando-me força.

Aos meus pais, Lucilia e Robson que estiveram comigo desde a primeira vez da escrita do meu nome e não mediram esforços para que hoje estivesse onde estou. Mostraram o caminho difícil, mas recompensador dos estudos.

Ao meu irmão que amo tanto pelos momentos únicos que me proporcionou e por todo apoio.

Ao meu esposo Cassiano Neto pelos anos de dedicação, amor, apoio e em especial compreensão nos momentos mais difíceis desta jornada. Obrigada meu amor, tudo que tenho conquistado também devo a você.

À minha orientadora Dr^a Valesca Pandolfi, pela oportunidade de ser orientanda, por toda ajuda e por acreditar na minha capacidade.

À minha coorientadora e chefe do Laboratório de Genética e Biotecnologia Vegetal (LGBV) Dr^a Ana Maria Benko-Iseppon, pela oportunidade de integrar a equipe deste laboratório, pela confiança e ajuda.

Ao meu coorientador não documentado M.Sc. João Pacífico por todo conhecimento passado e construído de forma paciente e didática. Por me ajudar a desvendar os “mistérios” da bioinformática, por toda ajuda no desenvolvimento deste trabalho, além dos momentos de risadas e pela carona de quase todos os dias.

Aos amigos que encontrei neste caminho: Stephani Soares e Cecília Bernardo pelas conversas e companhia nos primeiros momentos dentro do LGBV; à Karla Kamila por ser tão solícita em transmitir conhecimentos e pelas risadas que tivemos; aos meninos da “bioinfo” Marx, Pedro e Rômulo por toda ajuda; à Rebeca Rivas e Ana Rafaela pela participação especial.

A Bruna, Flávia, Hévila e Silvany pelos momentos incríveis que passamos juntas no laboratório e fora dele, pelas risadas e brincadeiras que tornaram meus dias mais leves e os congressos mais animados. Pelas diversas conversas a respeito do futuro e tantos outros assuntos. Obrigada pela amizade!

A todos que fizeram e fazem parte do LGBV: Polyana, Mitaly, Lidiane, Anne, Manu, Santelmo, Diego, Neto, Wilson, Ivamberto, Jaiza, pelos momentos de crescimento profissional e momentos de intensa alegria.

À Vanessa pela ajuda na resolução de problemas e pelas palavras mágicas “vai dar tempo” nos momentos de desespero.

À Universidade Federal de Pernambuco e Pós-Graduação em Ciências Biológicas, por todo suporte neste período e à Adenilda que tenta nos ajudar do jeito dela (risos).

Aos familiares pelo apoio e incentivo, primos, primas, tias, tios, minhas madrinhas Ione e Lucineide. Ao meu sogro Cassiano Filho pelas palavras de incentivo e minha sogra Girleide que direta ou indiretamente contribuíram para essa conquista. Às minhas amigas desde o ensino médio Anne, Izabel, Vera, Fernanda e Andreza que mesmo distantes, se preocuparam, me apoiaram e sempre torcem por mim, obrigada pela amizade.

Aos professores que compartilharam o conhecimento e suas experiências.

Ao suporte financeiro concedido pelas agências de fomento à pesquisa, CAPES (Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e FACEPE (Fundação de Amparo à Ciência e Tecnologia de Pernambuco).

Enfim, a todos que de alguma maneira direta ou indiretamente contribuíram para a realização deste trabalho.

Muito obrigada!

*“Ó minha alma se alegra somente no Senhor,
porque d’Ele vem a minha esperança. ”*

Salmos 62-5

RESUMO

Ciclotídeos são uma classe de peptídeos antimicrobianos (AMPs - do inglês *Antimicrobial peptide*) cíclicos de plantas, compostos de, aproximadamente, 30 resíduos de aminoácidos, sendo seis cisteínas conservadas e conectadas por três pontes de dissulfeto. Sua expressão é constitutiva, tendo sua principal função na defesa vegetal contra patógenos, que podem causar perdas significativas em culturas importantes para a agricultura, como no caso da família Poaceae que apresenta destacada importância econômica no Brasil e no mundo. Nesse estudo foi conduzida uma busca por genes relacionados a ciclotídeos vegetais, disponíveis em bancos de dados de acesso restrito e público, com vistas ao isolamento e caracterização *in silico* desses peptídeos. Através da busca nos genomas de *Hevea brasiliensis*, *Manihot esculenta*, *Ricinus communis*, *Sorghum bicolor* e *Zea mays*; bem como no transcriptoma de *Vigna unguiculata* foi verificado que apenas o genoma de *Zea mays* apresentou dois possíveis genes codificadores de ciclotídeos. Assim, *primers* foram desenhados para o isolamento destes genes em milho. Além da espécie *Z. mays*, as espécies *Triticum aestivum* (trigo) e *Secale cereale* (centeio), foram utilizadas para a tentativa de isolamento a partir dos pares de primers desenhados. Foram obtidos 19 fragmentos (amplicons), sendo quatro deles (zm315, zm316, zm317, sc359) com o domínio ciclotídeo, os três primeiros de milho e o último de centeio. Essas quatro sequências foram, então, submetidas a uma caracterização *in silico*, para predição da estrutura secundária, terciária e função predita. Verificou-se que esses peptídeos apresentam as seis cisteínas conservadas, três pontes dissulfeto e o padrão de aminoácidos entre as cisteínas, similar aos encontrados em ciclotídeos. Ainda foi possível a predição de algumas propriedades físico-químicas e modelagem por homologia para as quatro proteínas, o que mostrou a qualidade e confiabilidade dos modelos. Sugere-se que dois dos ciclotídeos isolados (zm315, zm316) pertençam a uma nova classe de peptídeos lineares, mas com características de ciclotídeos.

Palavras-Chave: Bioinformática, Modelagem por homologia, AMPs, Poaceae.

ABSTRACT

Cyclotides are a class of cyclic antimicrobial peptides (AMPs) present on plants, composed by approximately 30 amino acid residues, including six conserved cysteines connected by three disulphide bridges. Its expression is constitutive, with main function on plant defense against pathogens, that may cause significant losses in important cultivars, as in the case of Poaceae, a family that presents economic importance for the agriculture in Brazil and worldwide. This study performed a search for genes related to plant cyclotides, available in restricted and public access databases, aimed at their *in silico* isolation and characterization. Searching for these peptides in *Hevea brasiliensis*, *Manihot esculenta*, *Ricinus communis*, *Sorghum bicolor*, *Vigna unguiculata* and *Zea mays* genomes, we obtained two possible genes encoding Cyclotides in *Z. mays*. Thus, primers were designed for the isolation of these genes in maize as well in wheat (*Triticum aestivum*) and rye (*Secale cereale*) species. We obtained 19 amplicons and four of them (zm315, zm316, zm317, sc359) presented cyclotide domain. These four sequences were then subjected to *in silico* characterization, for predicting their secondary and tertiary structures, as well their function. It was found that these peptides present six conserved cysteines, three disulphide bridges and the amino acid pattern between the cysteines similar to those found in cyclotides. It was also possible to predict some physical chemical properties and also building a 3D protein by homology modeling for the four peptides, presenting high quality and reliability. Our analysis indicates that two isolated cyclotides (zm315, zm316) appear to belong to a new class of linear peptides, but with cyclotide features.

Keywords: bioinformatics, homology modeling, AMPs, Poaceae.

LISTA DE FIGURAS

Página

REVISÃO BIBLIOGRÁFICA

- Figura 1.** Alguns representantes de ciclotídeos alinhados e separados por subfamílias, mostrando resíduos conservados, em amarelo as cisteínas numeradas I-VI, *loops* 1-6, presença da prolina (vermelho) na subfamília Möbius, principal diferença entre as duas famílias. O resíduo de glicina em verde está associado ao nó de cistina cíclico ou estabilização da estrutura e a região em azul é a sequência identificada no precursor como flaqueadora do sítio de processamento e ciclização (PICCHI et al., 2009)..... 27
- Figura 2.** Comparação de arranjos genéticos entre diferentes precursores de ciclotídeo em diferentes espécies e famílias as quais pertencem. Em destaque: região com sequência sinal (ER, em preto); N-terminal pró-domínio (NTPD, em cinza); a região NTR (em roxo); peptídeo maduro (em verde); o domínio A1a (em rosa) e a região C-terminal (CTR), indicado como “Tail (em azul) (NGUYEN, et al., 2011)..... 29

CAPÍTULO I

- Fig. (1).** Schematic illustration of three different cyclotide conformations. (A) The arabic numbers above the vertical alignment regard the primary structure of: (1) Circulin A, (2) Cyclopsychotride A, (3) Kalata B1 and (4) Varv peptide. The conserved cysteine residues are represented by Roman numerals (I to IV), forming three disulfide bonds (vertical black lines). The inter-cysteine loops are numbered from 1 to 6. Other amino acids residues are in grayscale, whereas those chemically similar are showed in same gray intensity; (B) Topology of the cyclic cysteine knot (CCK) motif and summary of the conserved and variable residues of all known vegetables cyclotides: the disulfide bonds are represented by black lines connecting the cysteine residues (Subscript I-IV); (C) Secondary structure of cyclotides, showing beta sheet and disulfide bonds connecting cysteine residues (I-IV Subscript) and the inter-cysteine loops (L1-6)..... 73
- Fig. (2).** Alignment of cyclotide sequences available in GenBank (NCBI), showing conserved cysteines residues (gray), represented by Roman numerals (I to VI),

forming three disulfide bonds (black lines at the bottom of the Figure). Protein loops are numbered from 1 to 6 at the top of the alignment. Black columns at the bottom indicate conservation. Black arrow indicates residue possibly introduced by mutation that lead to cyclization.....

74

Fig. (3). Alignment of trypsin inhibitors (TI) of the squash family, based on publically available sequences (GenBank, NCBI) showing conserved cysteine residues (gray), represented by Roman numerals (I to VI), forming three disulfide bonds (black lines at the bottom of the figure). Legend to species abbreviations: CPTI: *Cucurbita pepo* trypsin inhibitor; MCTI: *Momordica charantia* trypsin inhibitor; EETI: *Ecballium elaterium* trypsin inhibitors; CMTI: *Cucurbita maxima* trypsin inhibitors; MCoTI: *Momordica cochinchinensis* trypsin inhibitors.....

75

Fig. (4). Alignment of SFTI sequences obtained from GenBank (NCBI), showing conserved residues, cysteines (numbered I-II) and the disulfide bond (gray line at the bottom of the figure). Sequences isolated from *Helianthus annuus* (sunflower).....

76

Fig (5). Schematic representation of chimeric genetic arrangements of cyclotide precursors in different plant families. (A) Albumin 1 domain from *Pisum sativum*, Fabaceae (pA1a and pA1b); (B) cliotide from *Clitoria ternatea*, Fabaceae (cT7); (C) cyclotide from *Oldenlandia affinis*, Rubiaceae (Oak1), including representation of introns (black triangle) and (D) cyclotides from *Viola odorata*, Violaceae (Voc1). ER, endoplasmic reticulum signal; Pro-region (NTPD, N-terminal precursor domain; NTR, N-terminal repeat; cyclotide domain and Tail, the C-terminal pro-peptide tail. Note that A1a domain is present only in albumin and cliotides.....

77

Fig (6). Schematic representation of cyclotide post transcriptional processing. Hypothetical steps regarding cyclotide intracellular processing (I to IV) and proposed mechanism for AEP mediation in the cyclization reaction expanded in IV. Legend for abbreviations: AEP = Asparaginyl Endopeptidase; NTPP = N-terminal propeptide; NTR = N-Terminal Repeat; PDI = Protein Disulfide Isomerase.....

78

Fig. (7). General outline of the main mechanism of interaction between antimicrobials peptides and pathogen cell membranes. (A) “Barrel” model; (B) “Carpet” model; (C) “Toroidal pore” model; (D) “Toroidal disorderly pore” model;

(E) “Membrane thinning” model; (F) “Anion carrier” model; (G) “Non-lytic depolarization” model; (H) Electroporation model.....	79
--	----

GRAPHICAL ABSTRACT	83
---------------------------------	----

CAPÍTULO II

Figura 1. Predição estrutural dos fragmentos isolados de milho (A) Zm315, (B) Zm316, (C) Zm317 e de centeio (D) Sc359, a partir de dados gerados no <i>Philius</i> e <i>Smart</i> . As regiões ao longo das sequências estão delimitadas por letras coloridas, as quais representam: a região do peptídeo sinal (verde); região pro-peptídeo N-terminal (NTPP) (azul) e o domínio ciclotídeo (vermelho) Observação: as sequências foram omitidas por motivo de propriedade intelectual.....	97
--	----

Figura 2. Alinhamento das sequências similares encontradas no GenBank, através da ferramenta BLASTp. As cores em destaque mostram os resíduos com maior conservação. Observação: as sequências foram omitidas por motivo de propriedade intelectual.....	98
---	----

Figura 3. Alinhamento das sequências isoladas e identificadas como possíveis ciclotídeos com duas sequências caracterizadas como ciclotídeos disponíveis no GenBank/NCBI. (Destaque: seis cisteínas em rosa) Observação: as sequências foram omitidas por motivo de propriedade intelectual.....	99
---	----

Figura 4. Porcentagem de aminoácidos presentes em cada uma das sequências isoladas (Zm315, Zm316, Zm317 e Sc359).....	99
--	----

Figura 5. Sequências isoladas neste estudo e sequências da família Poaceae disponíveis nos bancos de dados do GenBank/ NCBI e Phytozome. Observação: as sequências foram omitidas por motivo de propriedade intelectual.....	100
---	-----

Figura 6. Estrutura secundária predita através da ferramenta Alphapred, para as sequências zm315, zm316, zm317 e sc359 isoladas (sequências superiores) e a estrutura encontrada (sequências inferiores). Folhas- β (seta azul; E) e α -hélice (estrutura helicoidal azul, H). Observação: as sequências foram omitidas por motivo	
--	--

de	propriedade	intelectual.	
.....			101
Figura 7.	Estrutura predita das pontes dissulfeto (parte inferior da sequência) e as seis cisteínas indicadas por números romanos (parte superior da sequência). Observação: as sequências foram omitidas por motivo de propriedade intelectual.....		102
Figura 8.	Estrutura 3D modelada por homologia das quatro sequências isoladas. Estrutura em seta caracteriza a folha beta e uma pequena estrutura helicoidal representa a alfa hélice. Observação: os detalhes foram omitidos por motivo de propriedade intelectual.....		106
Figura 9.	Gráficos representativos do Z-score dos modelos gerados, comparativamente aos modelos de cristalografia de raios-x (cinza claro) e ressonância magnética (cinza escuro) depositados no PDB. Os pontos marcados em vermelho referem-se aos modelos gerados dos prováveis ciclotídeos.....		107

LISTA DE TABELAS

	Página
REVISÃO BIBLIOGRÁFICA	
Tabela 1. Características da cultivar BR 1 (MINISTÉRIO DA AGRICULTURA, 2015).....	35
 CAPÍTULO I	
Table 1. Selected examples of antimicrobial peptides and putative modes of membrane interaction.	80
Table 2. Plant cyclic peptides, including their sources and known activities.....	81
 CAPÍTULO II	
Tabela 1. Sequências caracterizadas nos genomas de <i>Vigna unguiculata</i> (Vu); <i>Zea mays</i> (Zm), <i>Sorghum bicolor</i> (Sb), <i>Hevea brasiliensis</i> (Hb), <i>Ricinus communis</i> (Rc) e de <i>Manihot esculenta</i> (Me), utilizando os modelos construídos para os diferentes grupos taxonômicos de ciclótídeos.....	93
Tabela 2. Domínios encontrados nas espécies candidatas: <i>Vigna unguiculata</i> (Vu); <i>Sorghum bicolor</i> (Sb), <i>Zea mays</i> (Zm), <i>Hevea brasiliensis</i> (Hb), <i>Manihot esculenta</i> (Me) e <i>Ricinus communis</i> (Rc), obtidos após a busca por homólogos via HMMER, utilizando a ferramenta de busca de domínios Batch.....	94
Tabela 3. Descrição das sequências isoladas de ciclótídeos com domínio (<i>cyclotide Family</i>) conservado com similaridade no banco do NCBI.....	96

Tabela 4. Predição das propriedades físico-químicas das sequências isoladas: Zm315 e Zm316 (possíveis lineares), Zm317 e Sc359 (circulares), e representantes de ciclotídeos já caracterizados Kalata B1 e Circulin B. Cada coluna está representada por valores preditos no peptídeo completo (PepC) e por valores preditos no peptídeo maduro (PepM).....	104
Tabela 5. Valores obtidos para identidade, qualidade e energia para os modelos estruturais de ciclotídeos em milho e centeio obtidos via SWISS-MODEL e Procheck.....	106

LISTA DE ABREVIATURAS

AMPs	<i>Antimicrobial peptides</i> Peptídeos Antimicrobianos
APD	<i>Antimicrobial Peptide Database</i> Banco de dados Peptídeos Antimicrobianos
BLAST	<i>Basic Local Alignment Search Tool</i> Ferramenta de Busca por Alinhamento Local
CCK	<i>Cyclic Cystine Knot</i> Nó Cistina Cíclico
ctc	<i>Clitoria ternatea</i> ciclotídeo
CTR	<i>C-terminal repeat</i> Repetição C-terminal
EMBL	<i>European Molecular Biology Laboratory</i> Laboratório Europeu de Biologia Molecular
GOLD	<i>Genome OnLine Database</i> Banco de dados Online Genoma
HR	<i>Hipersensitive Response</i> Reação Hipersensibilidade
KEGG	<i>Kyoto Encyclopedia of genes and Genomes</i> Enciclopédia de genes e genomas de Kyoto
NCBI	<i>National Center for Biotechnology Binformation</i> Centro Nacional para Informação Biotecnológica
NTPD	N-terminal precursor domain Pro-domínio N-terminal
NTR	N-terminal repeat Repetição N-terminal
SAR	Systemic Acquired Resistance Resistência Sistêmica adquirida
VMT	Vírus do Mosaico do Trigo
VNAC	VNAC Vírus do Nanismo Amarelo da Cevada

SUMÁRIO

	Página
I. INTRODUÇÃO.....	18
II. OBJETIVOS.....	20
2.1. Objetivo geral.....	20
2.2. Objetivos específicos	20
III. REVISÃO BIBLIOGRÁFICA.....	21
3.1. Mecanismos de defesa das plantas.....	21
3.2. Peptídeos Antimicrobianos de plantas.....	22
3.3. Ciclotídeos.....	25
3.3.1. Descoberta e caracterização.....	25
3.3.2. Características estruturais e sequência conservada.....	26
3.3.3. Biossíntese.....	28
3.3.4. Atividade biológica	29
3.3.5. Evolução e distribuição	30
3.4. Espécies foco do estrudo.....	31
3.4.1. Considerações gerais sobre a cultura de <i>Zea mays</i>	31
3.4.2. Considerações gerais sobre a cultura de <i>Secale cereale</i>	33
3.5. Bancos de dados e ferramentas de bioinformática.....	35
IV. REFERÊNCIAS BIBLIOGRÁFICAS.....	40
CAPÍTULO I:	
<i>Plants Defense-related Cyclic Peptides: Diversity, Structure and Applications.....</i>	50
CAPÍTULO II:	
<i>Isolamento e caracterização in silico de ciclotídeos em membros da família Poaceae.....</i>	84

V. CONCLUSÕES GERAIS.....	122
VI. ANEXOS.....	123
Anexo 1. Instruções para autores: <i>Current Protein & Peptide Science</i>	123
Anexo 2. Instruções para autores: <i>Genetics and Molecular Biology</i>	136
Anexo 3. Tabela: Sequências de ciclotídeos obtidas no GenBank/ NCBI e utilizadas para geração de motivos através da ferramenta HMMER.....	141
Anexo 4. Tabela: BLASTp realizado do banco de dados Phytozome, para análise de similaridade.....	149

I. INTRODUÇÃO

Ao longo de sua evolução, os vegetais estão sujeitos ao ataque de uma variedade de microrganismos patogênicos (bactérias, fungos e vírus) presentes no ambiente. Como forma de defesa, as plantas têm desenvolvido numerosos e eficientes mecanismos físicos e químicos contra tais fatores de estresse. Respostas químicas incluem moléculas produzidas constitutivamente (como a lignina, por exemplo) ou induzidas após percepção dos patógenos (LOON et al., 2006), estes são elementos vitais nos processos de defesa (MENEZES et al., 2002; REDDY et al., 2004). Nesta categoria de moléculas destacam-se os peptídeos antimicrobianos-AMPS, representados pelas defensinas, as tioninas, as proteínas transferidoras de lipídeos, as proteínas tipo-heveína, proteínas tipo-knotina, ciclotídeos, entre outros (GARCÍA-OLMEDO et al., 2001).

No que se refere a ciclotídeos, esses são uma classe incomum de peptídeos circulares ricos em cisteínas, expressas naturalmente, apresentando atividade antimicrobiana, inseticida, anti-helmíntico, antiviral, antifúngica (BASSE 2005; DALY et al., 2009) e também como inibidores de proteases (AVRUTINA et al., 2005). Devido a tais propriedades, estes peptídeos têm sido reconhecidos nos vegetais como importantes moléculas tanto no desenvolvimento de produtos terapêuticos, como no desenvolvimento de novas estratégias ou para o aprimoramento de técnicas já existente no melhoramento de culturas importantes, tanto no âmbito econômico como no social (KEYMANESH et al., 2009).

No âmbito econômico, entre espécies de grande importância econômica, destacam-se membros da família Poaceae, dentre as quais se destacam o arroz (*Oryza sativa*), trigo (*Triticum aestivum*), milho (*Zea mays*), aveia (*Avena sativa*), cevada (*Hordeum vulgare*), centeio (*Secale cereale*) e cana-de-açúcar (*Saccharum officinarum*).

Devido ao seu potencial de produção, composição química e valor nutricional, a cultura do milho (*Zea mays*), é um dos mais importantes cereais cultivados no Brasil. Este cereal é cultivado em todas as regiões do Brasil, chegando a uma produção de mais de 78 milhões de toneladas de grãos (CONAB, 2015).

O centeio (*Secale cereale*), espécie originária da Ásia, pode ser cultivado sob condições ambientais bem diferentes quando comparado aos outros cereais de inverno. Seu cultivo pode ser observado desde o Círculo Polar Ártico até o extremo sul da

América do Sul, em locais próximos ao nível do mar ou à 4.300 m de altitude (BAIER, 1988). Este cereal também pode ser encontrado em regiões de clima seco ou frio e em solos arenosos e pouco férteis (DE MORI et al., 2013).

Com os avanços das “ômicas” e, conseqüentemente, a maior disponibilidade de dados gerados a partir de projetos de sequenciamento, a bioinformática tem se tornado cada vez mais essencial na manipulação e análise rápida deste grande número de informações. Diante desta quantidade de dados e informações geradas, por envolver ferramentas que agregam conceitos de biologia, química, estatística e ciências da computação, a bioinformática se mostra como uma importante e indispensável área para organização e manipulação de análise desse grande número de dados biológicos gerados (EDWARDS & BATLEY 2004; DROIT et al., 2005).

O objetivo deste trabalho foi a identificação, isolamento e caracterização de sequências codificadoras de ciclotídeos, em duas espécies pertencentes à família Poaceae, *Z. mays* e *S. cereale*.

II. OBJETIVOS

2.1. Geral

- ✓ Este estudo teve como principal objetivo identificar, isolar e caracterizar estruturalmente genes relacionados a ciclotídeos em plantas, em especial, membros da família Poaceae.

2.2. Específicos

- ✓ Identificar os principais ciclotídeos em bancos de dados de nucleotídeos e de aminoácidos de diversas espécies de interesse econômico, utilizando os dados obtidos para amplificação e isolamento de tais peptídeos nas espécies-alvo;
- ✓ Caracterizar e descrever regiões conservadas, como domínios e motivos, nas sequências identificadas;
- ✓ Isolar regiões codificantes, sequenciar e analisar comparativamente as sequências de nucleotídeos e a estrutura do gene obtido com os disponíveis em bancos de dados;
- ✓ Estudar e caracterizar a estrutura tridimensional dos prováveis ciclotídeos, procurando diferenças e similaridades com indivíduos de outras espécies;

III. REVISÃO BIBLIOGRÁFICA

3.1. Mecanismos de defesa das plantas

Diferentes de outros organismos, as plantas não processam moléculas de imunoglobulinas ou células imunes circulantes, não realizam processos fagocíticos e nem mostram imunidade adaptativa (SPOEL & DONG, 2012). Para compensar a ausência destes, os vegetais desenvolveram uma variedade de mecanismos de defesa inata, desencadeando respostas adequadas, geralmente através da secreção de uma mistura de compostos químicos primários e/ou secundários (PRITHIVIRAJ et al., 2006; BENKO-ISEPPON et al., 2010). O primeiro obstáculo à penetração do patógeno são as barreiras físicas das plantas, que incluem a produção de camadas foliares epicuticularizadas e tecidos epidérmicos suberizados, cutinizados e lignificados (CANO-DELGADO et al., 2003).

Diante de fitopatógenos, as plantas, de um modo geral, reconhecem os invasores e ativam rapidamente os mecanismos que reprimem a ação do invasor. Esta percepção é essencial para o sucesso da defesa vegetal, assim, as células vegetais são capazes de reconhecer uma variedade de “assinaturas” moleculares destes microrganismos (LEÓN & MONTESANO, 2013).

A resposta de defesa pode ocorrer de forma constitutiva, a qual já está presente nas plantas, ou induzida, ocorrendo após o ataque de fitopatógenos e envolvem a ativação de genes e indução de uma rede de transdução sinais, relacionadas com estresse abiótico (HEATH, 2000). A ativação destes mecanismos de defesa ocorre por meio de uma cascata de sinais que se inicia no reconhecimento do agente agressor, culminando na ativação das barreiras físicas e químicas envolvidas no processo.

Dentre os mecanismos de defesa utilizados pelas plantas estão a resposta hipersensível (HR), a resistência sistêmica adquirida (SAR), a indução de proteínas relacionadas à patogênese (Proteínas PR) e compostos sinalizadores, tais como ácido salicílico e peróxido de hidrogênio. A resposta hipersensível ou reação de hipersensibilidade em plantas é considerada como um dos principais eventos da resposta de defesa contra o ataque de patógenos, sendo caracterizado por uma resposta rápida e localizada, ou seja, no sítio de infecção do patógeno. Tal mecanismo de defesa faz com

que ocorra o colapso do tecido vegetal (ao redor do sítio de infecção), em decorrência da liberação de compostos tóxicos, os quais também atuam, em alguns casos, diretamente sobre o patógeno, ocasionando sua morte (AGRIOS, 2004).

Fungos, bactérias e vírus também podem ativar a SAR em plantas, a qual depende do agente envolvido. Neste caso, a infecção inicial precisa resultar na formação de lesões necróticas, decorrentes da HR (acúmulo de peróxido de hidrogênio) ou como sintoma da doença (HAMMOND-KOSACK; JONES, 2000). O mecanismo da SAR envolve uma cascata de eventos e sinais, os quais iniciam no momento da interação planta/patógeno, levando a alterações no metabolismo celular, que ativam sinais moleculares dirigidos para outras partes da planta, promovendo a redução da severidade da doença (AGRIOS, 2004).

Atualmente, sabe-se que uma grande variedade de organismos é capaz de produzir peptídeos, os quais atuam na primeira linha de sua defesa. Alguns desses peptídeos são caracterizados pela presença de cisteínas (4, 6, 8 e 12), conectadas em pares, formando pontes dissulfeto, conferindo alta estabilidade à molécula (BROEKAERT et al., 1997). Tais peptídeos podem ser sintetizados constitutivamente pelo organismo, ou ser induzidos somente após uma lesão ou infecção, apresentando amplo espectro de atividade biológica. Milhares destas moléculas já foram isoladas e tiveram suas atividades biológicas comprovadas. Destas, a grande maioria apresenta atividade antimicrobiana de grande importância no sistema de defesa do organismo (JENSSEN et al., 2006; WANG et al., 2008). Tal resposta reflete a estrutura genética e proteica dessas moléculas a exemplos das proteínas de tamanho pequeno, de carga geral positiva (catiônicos em pH fisiológico), comumente ricas em um tipo de aminoácido. Tais características permitem a sua classificação em várias subclasses de proteínas responsáveis pela ação antimicrobiana, tolerância a solventes ácidos e orgânicos, estabilidade térmica e estereogeometria anfipática, além de ampla atividade biológica (YEAMAN & YOUNT, 2003; BROGDEN et al., 2005).

3.2. Peptídeos antimicrobianos de plantas.

Os peptídeos antimicrobianos de plantas (AMPs – *antimicrobial peptides*) e de animais geralmente são compostos por 12 a 120 resíduos de aminoácidos,

compreendendo moléculas catiônicas ($\text{pH} \leq 7$, devido ao excesso de resíduos de lisina, arginina e histidina), anfipáticos e relativamente pequenos (muitas vezes menores que 50 kDa) (HANCOCK & SCOTT, 2000; BULET et al., 2004; GANZ, 2004; BOIX & NOGUES, 2007).

Nas plantas, os principais grupos de AMPs são representados por ciclotídeos, defensinas, proteínas de transferência de lipídios (LTPs), esnaquinas, tioninas, proteínas do tipo heveína e knotina (Da ROCHA PITTA et al., 2010). Tais peptídeos tendem a se apresentar sem estrutura definida em solução aquosa, adotando uma forma anfipática ao interagir com membranas, atributo crucial tanto para atividade quanto para citotoxicidade. Apesar da estrutura tridimensional e propriedades físicas serem bastante semelhantes entre os diversos AMPs, em geral, há pouca homologia entre sequências primárias, principalmente entre peptídeos de classes distintas (HADLEY & HANCOCK, 2010). Membros da mesma família apresentam um padrão dobrável que é globalmente comparável entre as famílias, adotando uma estrutura tridimensional que envolve a formação de elementos estruturais secundários, tais como folhas beta pregueadas e alfa-hélices que são estabilizadas por pontes dissulfeto intramoleculares, impondo uma estrutura mais rígida, permitindo maior estabilidade da molécula (GIUDICI et al., 2006; BENKO-ISEPPON et al., 2010).

Apesar dos peptídeos antimicrobianos apresentarem uma grande diversidade estrutural, é possível agrupá-los em quatro classes distintas de estruturas secundárias (JENSSEN et al., 2006). As estruturas mais comumente observadas são as α -hélices (formadas apenas quando o peptídeo entra em contato com a membrana), e as folhas- β (estabilizadas por 2 a 4 ligações dissulfeto, as quais, ocasionalmente, apresentam pequenas porções de estrutura α -hélice. As demais estruturas (menos comuns) são estruturas curvadas, formadas por ligações dissulfeto simples ou devido à presença de resíduos de prolinas na molécula (RESENDE et al., 2008).

Os AMPs atuam na membrana plasmática dos microrganismos alvos por meio da interação eletrostática com as cadeias negativas dos fosfolípidos presentes nas membranas dos microrganismos patogênicos (PAPO & SHAI, 2003), contudo, o dano à membrana seria apenas a primeira ação desses peptídeos (JELINEK & KOLUSHEVA, 2005; LOBO et al., 2007). Eles possuem diversos mecanismos de ação, descritos com base na possível interação dessas moléculas com as membranas e a formação de poros

quando alcançam uma concentração limiar. No “modelo de barril”, os AMPs se inserem de forma paralela na membrana formando um poro. Já no “modelo de tapete”, esses peptídeos inserem porções na membrana e induzem a formação de poros, enquanto que no “modelo de poro toroidal” também ocorre a formação de poros com os peptídeos inseridos na membrana de forma paralela; porém, nesse caso, eles são intercalados com lipídeos (NGUYEN et al., 2011).

Atualmente, novos modelos foram propostos para tentar entender seus mecanismos de ação. Entre eles está o modelo de “poro toroidal desordenado”, que ocorre com uma menor concentração de peptídeos, comparada à quantidade presumivelmente ao que ocorre no modelo poro toroidal, onde os peptídeos estão dispostos de forma desordenada, formando um poro na membrana celular. Outro modelo é o de afinamento da membrana, em que a atração entre os fosfolipídios carregados gera regiões nas membranas mais finas e mais frágeis. Outro modelo que tenta explicar o mecanismo de ação dessas moléculas é o “modelo de atração de ânions”, o qual sugere que os AMPs são capazes de atrair pequenos ânions, fazendo-os atravessar a membrana de forma a causar perda do potencial eletrostático celular. Ainda em outro modelo, a perda do potencial eletrostático pode ocorrer sem a atração de íons. Já, no “modelo da eletroporação”, os peptídeos afetam o potencial de membrana de forma a diminuir sua permeabilidade a várias moléculas (NGUYEN et al., 2011).

Dentre os peptídeos antimicrobianos existentes, uma classe interessante são aqueles com uma estrutura cíclica, os quais apresentam uma grande variedade de atividade biológica, destacando-se como importantes agentes antivirais. O primeiro peptídeo cíclico de plantas estruturalmente caracterizado foi o kalata B1, constituído por 29 aminoácidos, isolado de *Oldenlandia affinis*, uma planta tropical africana muito usada como chá na medicina popular para facilitar e acelerar o parto (CRAIK, 2001). Apresenta vasta gama de atividade biológica, incluindo citotóxicas, sendo um potente agente anticancerígeno, anti-HIV, inseticida e antimicrobiano, o que confere a estas proteínas um potencial bastante promissor em pesquisas farmacêuticas e/ou agroquímicas (HALLOCK et al., 2000; TRABI et al., 2004; CHEN et al., 2005; JENNINGS et al., 2005).

3.3 Ciclotídeos

3.3.1 Descoberta e caracterização

Relatos com ciclotídeos tiveram seu início a partir de duas observações realizadas na África, com a planta *Oldenlandia affinis* (Rubiaceae), a qual era utilizada pela medicina popular, através de uma decocção, para favorecer as contrações uterinas, facilitando o trabalho de parto. Em 1965, o pesquisador Finn Sandberg, descreveu esta decocção como sendo de “wetegere”, que quer dizer *O. affinis*, na língua Gbaya, na República Centro-Africana (SANDBERG, 1965). Em seguida esta decocção foi associada a outra decocção, chamada *kalata-kalata* (na língua Tsjiluba), esta associação foi realizada pelo médico norueguês Lorents Gran, sendo detectada também como uma decocção aquosa de *O. affinis* (GRAN, 1970), sendo atrelada a uma série de complicações obstétricas. Mais tarde, a sequência parcial deste peptídeo foi desvendada, sendo chamada de Kalata B1 (GRAN, 1973; SLETTEN & GRAN, 1973). Ao longo da década de 90, vários novos peptídeos foram descritos e caracterizados (SCHOEPKE et al., 1993; GUSTAFSON et al., 1994; WITHERUP et al., 1994; CLAESON et al., 1998; GORANSSON et al., 1999) como “ciclotídeos” (ciclo-peptídeos) (CRAIK et al., 1999).

Ciclotídeos variam de 28 a 37 aminoácidos, com estrutura cíclica (do tipo “cabeça-cauda”), com seis cisteínas conservadas e ligadas numa topologia de “nó”, formando três pontes dissulfeto (CRAIK et al., 2002). Estas ligações de dissulfeto são estabelecidas da seguinte forma: CysI-CysIV e CysII-CysV, enquanto que a terceira ligação dissulfeto (CysIII-CysVI) penetra nesse anel (CRAIK, 2010). O motivo estrutural compartilhado entre o ciclotídeos é o “nó cistina cíclico” (*Cyclic Cystine Knot* - CCK) (CRAIK et al., 1999). Alguns variantes acíclicos dos ciclotídeos encontrados nas famílias Fabaceae e Poaceae apresentam um arranjo semelhante das ligações entre cisteínas, formando nó (BASSE, 2005; NGUYEN et al., 2011; 2013)

Com base na estrutura primária da cadeia polipeptídica, os ciclotídeos são divididos, de acordo com alguns autores, em duas subfamílias: Möbius e Bracelet (CRAIK et al., 1999). Para outros, uma terceira subfamília é conhecida como “inibidores de tripsina”, sendo MCoTI-I *Momordica cochinchinensis* inibidor de tripsina-I) e MCoTI-II isolados de *M. cochinchinensis* (HERNANDEZ, 2000; TSOI et

al., 2005; CHAN, et al., 2009). Apesar da similaridade de sequência com os ciclotídeos ser baixa, e a extensão dos *loops* serem diferentes, as proteínas cíclicas de *Momordica cochinchinensis* foram classificadas como “ciclotídeos”, com base no motivo de CCK conservado (FELIZMENIO-QUIMIO, et al. 2001).

3.3.2 Característica estrutural e sequência conservada

Os ciclotídeos apresentam seis *loops*, os quais são regiões topologicamente fundamentais para as suas atividades biológicas, onde as cadeias laterais dos aminoácidos são expostas, modulando tais atividades (PICCHI, et al., 2009).

A subfamília Möbius (nomenclatura recebida em homenagem ao matemático August Ferdinand Möbius) se caracteriza por uma ligação peptídica *cis*-prolina (no *loop* 5), conferindo uma torção na sua estrutura terciária. Já, na subfamília Bracelet, esta torção não ocorre, uma vez que não há prolina na posição correspondente em sua sequência primária, por isso são consideradas similares a um “bracelete” (pulseira) (CRAIK et al., 1999).

Entre os resíduos que compõem a sequência dos ciclotídeos, o ácido glutâmico (E), no *loop* 1, destaca-se entre os mais conservados, embora outros resíduos de conservação como uma glicina (G) ou alanina (A) e uma treonina (T) ou serina (S) também podem estar presentes na primeira e terceira posição, respectivamente (figura 1). A treonina ou a serina na 3ª posição da conservação citada pode ainda variar de subfamília, o resíduo de treonina ocorre na subfamília Möbius e a serina na Bracelet. No *loop* 2, quatro aminoácidos apesar de conservados, apresentam diferença na posição dos resíduos a depender da subfamília, sendo na subfamília Möbius apresentando dois resíduos GG centrais, seguido por, na maioria das vezes, um resíduo T. Por outro lado, a subfamília Bracelet, este mesmo *loop* apresenta os três primeiros resíduos hidrofóbicos, cujo quarto resíduo representado por uma prolina (P). Para os *loops* 3 e 5 os números de aminoácidos variam, ou seja, na Möbius são quatro aminoácidos, sendo que na Bracelet são sete. Ainda no *loop* 3 o resíduo de G que flanqueia a C da quarta posição, mantém esta conservação na maioria dos ciclotídeos, porque desempenha uma função estrutural, este resíduo exibe um ângulo ϕ importante para o estabelecimento da ligação entre CysI-CysIV.

O *loop* 4 mostra apenas um resíduo em todas as sequências de ciclotídeos analisadas. Entre os resíduos mais frequentes estão: serina (S), treonina (T) e lisina (K), a depender da subfamília. No *loop* 5 (a maior parte hidrofóbica) pode existir uma valina (V) para as duas subfamílias, na subfamília Möbius antes da valina ocorre a sequência SWP (serina triptofano e prolina) e na Bracelet são dois resíduos carregados positivamente.

Por fim, no *loop* 6, a variação de aminoácidos ocorre a depender da subfamília, embora uma sequência RNG(I/L)P mostra-se mais conservada. Esta sequência pode ser identificada como flanqueadora do sítio do sítio de processamento e ciclização.

	I	II	III	IV	V	VI	
MÖBIUS	loop 1	loop 2	loop 3	loop 4	loop 5	loop 6	
Kalata B1	C G E T C	V G G T -	C - - - N T P	G C T C	- - S W -	P V C T	R N G - L P V -
Kalata B2	C G E T C	F G G T -	C - - - N T P	G C S C	- - T W -	P I C T	R D G - L P V -
Kalata B3	C G E T C	F G G T -	C - - - N T P	G C T C	- D P W -	P I C T	R D G - L P T -
Kalata S	C G E T C	V G G T -	C - - - N T P	G C S C	- - S W -	P V C T	R N G - L P V -
varv peptídeo A	C G E T C	V G G T -	C - - - N T P	G C S C	- - S W -	P V C T	R N G - L P V -
varv peptídeo B	C G E T C	F G G T -	C - - - N T P	G C S C	- D P W -	P M C S	R N G - L P V -
varv peptídeo C	C G E T C	V G G T -	C - - - N T P	G C S C	- - S W -	P V C T	R N G - V P I -
varv peptídeo D	C G E T C	V G G S -	C - - - N T P	G C S C	- - S W -	P V C T	R N G - L P I -
varv peptídeo E	C G E T C	V G G T -	C - - - N T P	G C S C	- - S W -	P V C T	R N G - L P I -
violapeptídeo 1	C G E T C	V G G T -	C - - - N T P	G C S C	- - S R -	P V C T	X N G - L P V -
vodo M	C G E S C	F T G K -	C - - - Y T V	Q C S C	- - S W -	P V C T	R N G - A P I -
vodo N	C G E T C	T L G K -	C - - - Y T A	G C S C	- - S W -	P V C Y	R N G - L P V -
BRACELET	loop 1	loop 2	loop 3	loop 4	loop 5	loop 6	
circulina A	C G E S C	V W I P -	C I S A A -	L G C S C	- - - K N K V	C Y	R N G - I P - -
circulina B	C G E S C	V F I P -	C I S T L -	L G C S C	- - - K N K V	C Y	R N G V I P - -
circulina C	C G E S C	V F I P -	C I T S V -	A G C S C	- - - K S K V	C Y	R N G - I P - -
circulina D	C G E S C	V W I P -	C V T S I -	F N C K C	- - - E N K V	C Y	H D K - I P - -
circulina E	C G E S C	V W I P -	C L T S V -	F N C K C	- - - E N K V	C Y	H D K - I P - -
circulina F	C G E S C	V W I P -	C I S A A -	I G C S C	- - - K N K V	C Y	Y R A - I P - -
cicloviolina A	C G E S C	V F I P -	C I S A A -	I G C S C	- - - K N K V	C Y	R N G V I P - -
cicloviolina B	C G E S C	Y V L P -	C F T V - -	G C T C	- - - T S S Q	C F	K N G T A - - -
cicloviolina C	C G E S C	V F I P -	C L T T V -	A G C S C	- - - K N K V	C Y	R N G - I P - -
cicloviolina D	C G E S C	V F I P -	C I S A A -	I G C S C	- - - K N K V	C Y	R N G - F P - -
cicloviolacina O1	C A E S C	V Y I P -	C T V T A L L	G C S C	- - - S N R V	C Y	- N G - I P - -
cicloviolacina O2	C G E S C	V W I P -	C I S S A -	I G C S C	- - - K S K V	C Y	R N G - I P - -
cicloviolacina O3	C G E S C	V W I P -	C L T S A -	I G C S C	- - - K S K V	C Y	R N G - I P - -
cicloviolacina O4	C G E S C	V W I P -	C I S S A -	I G C S C	- - - K N K V	C Y	R N G - I P - -
vhr1	C A E S C	V W I P -	C T V T A L L	G C S C	- - - S N K V	C Y	- N G - I P - -
kalata B5	C G E S C	V Y I P -	C I S G V -	I G C S C	- - - T D K V	C Y	L N G - T P - -
hupa A	C A E S C	V Y I P -	C T I T A L L	G C S C	- - - K N K V	C Y	- N G - I P - -
ciclopsicotrídeo A	C G E S C	V F I P -	C T V T A L L	G C S C	- - - K S K V	C Y	K N S - I P - -
pallcoureína	C G E T C	R V I P V	C T Y S A A L	G C T C	D D R S D G L	C K	R N G D - P T F

Figura 1. Alguns representantes de ciclotídeos alinhados e separados por subfamílias, mostrando resíduos conservados em amarelo as cisteínas numeradas I-VI loops 1-6,

presença da prolina (em vermelho) na subfamília Möbius principal diferença entre as duas famílias. O resíduo de glicina em verde está associado ao nó de cistina cíclico ou estabilização da estrutura e a região em azul é a sequência identificada no precursor como flaqueadora do sítio de processamento e ciclização (PICCHI, et al., 2009).

3.3.3 Biossíntese

A primeira evidência de que os ciclotídeos seriam sintetizados nos ribossomos foi sugerida por JENNINGS et al., (2001) através da identificação de genes codificadores dos ciclotídeos Kalata B1, B2, B6 e B7 de *O. affinis*.

Quanto à arquitetura dos diferentes precursores de ciclotídeos e suas famílias, verifica-se a presença de: 1) uma sequência sinal (**ER**) de, aproximadamente, 20 aminoácidos a qual direciona o peptídeo para o retículo endoplasmático; 2) um N-terminal pro-domínio (**NTPD**), de tamanho variável, com nenhuma função atualmente relacionada a maturação destes peptídeos cíclicos; 3) uma região correspondente a repetição N-terminal (**NTR**); 4) regiões correspondentes ao domínio maduro do ciclotídeo; e 5) uma região C-terminal propeptídeo *tail* (**Tail**). Com exceção dos peptídeos encontrados em *Clitoria ternatea* (Fabaceae) - denominados de “cliotides” pelos autores, os quais são compostos pelo peptídeo sinal imediatamente seguido pelo domínio ciclotídeo que é flanqueado na extremidade C-terminal por ligante, e logo após uma cadeia de albumina e a extremidade C-terminal.

Estudos desenvolvidos por NGUYEN et al., (2011), relatam novos ciclotídeos em *Clitoria ternatea* (Fabaceae), denominados “cliotídeos” (*ctc*: *C. ternatea* *cliotide*) cuja representação estrutural do precursor deste peptídeo é mostrada na figura 2 (*ctc3* e *ctc7*).

Ainda na figura 2 é possível verificar uma comparação entre os ciclotídeos de diferentes famílias. Na família Fabaceae uma típica albumina (*pa1*) de *Pisum sativum*, exhibe os domínios A1b e A1a, o domínio A1b de *C. ternatea* (*ctc3* e *ctc7*) foi “substituído” pelo domínio ciclotídeo. As sementes de ervilhas codificam um peptídeo de 37 resíduos, contendo três ligações dissulfeto, estas, são algumas características que apresentam. A homologia de genes de *cliotides* e genes *ciclotídeos* (Rubiaceae e Violaceae) indica que os genes codificantes de ciclotídeos em Fabaceae são quiméricos,

metade de um ciclótideo (Rubiaceae e Violaceae) e outra metade albumina-1. Ainda neste estudo foi discutido que a ausência da região NTR no precursor de cliotídeos, sugere que esta região não seja essencial para sua biossíntese. Os genes *cliotides* contém uma região híbrida contendo domínios que não são relacionados e não foi possível detectar se a região A1b foi totalmente substituída.

Segundo MYLNE (2011), as cisteínas conservadas, os resíduos de asparagina (N) ou ácido aspártico (D) na região C-terminal, bem como a alanina ou glicina na região N-terminal são igualmente importantes, uma vez que parecem estar fortemente envolvidas no processo de síntese dos ciclótídeos.

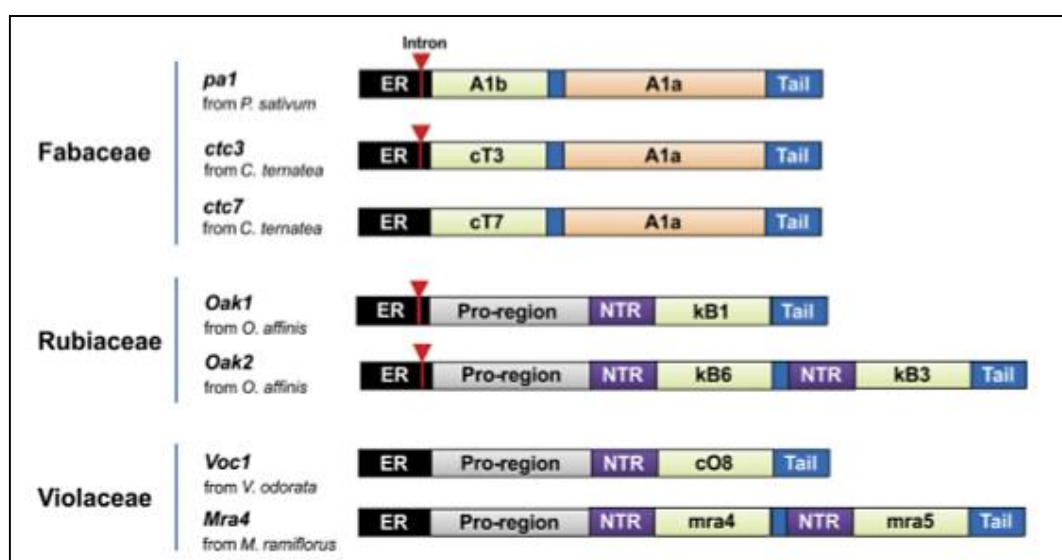


Figura 2. Comparação de arranjos genéticos entre diferentes precursores de ciclótideo em diferentes espécies e famílias as quais pertencem. Em destaque: região com sequência sinal (ER, em preto); N-terminal pró-domínio (NTPD, em cinza); a região NTR (em roxo); peptídeo maduro (em verde); o domínio A1a (em rosa) e a região C-terminal (CTR), indicado como “Tail (em azul) (NGUYEN, et al., 2011).

3.3.4 . Atividade biológica

O primeiro caso de atividade dos ciclótídeos foi associada a efeitos uterotônicos, da espécie *O. affinis* (GRAN, 1970).

Entretanto, nas últimas décadas, o número de atividades associadas aos ciclótídeos só vem aumentando. Exemplos incluem: o “*clopsichotride A*” isolado da planta *Psychotria*

longipes (Rubiaceae), com atividade de inibição-ligação da neurotensina (WITHERUP et al., 1994); as “circunlinas A-F” de *Chassalia praviflora* (Rubiaceae), com propriedades anti-HIV (GUSTAFSON et al., 1994; 2000), “cicloviolinas A-D”, a partir de *Leonia cymosa* (Violaceae) (HALLOCK et al., 2000) e “palicoureina” a partir de *Palicourea condensata* (Rubiaceae) (BOKESCH et al., 2001); ciclotídeos com atividade inseticida BARBETA et al., 2008; JENNINGS et al., 2001); atividade moluscida ((PLAN et al., 2008) bem como atividade anti-helmíntico (COLGRAVE et al., 2009) e antimicrobiana (TAM et al., 1999).

3.3.5 . Evolução e distribuição

Algumas questões relacionadas à evolução dos ciclotídeos e das proteínas circulares remetem a algumas evidências em termos dos três mecanismos comumente conhecidos, os quais podem elucidar a distribuição observada de ciclotídeos no reino vegetal: (1) vários ganhos independentes da funcionalidade cíclica, (2) a transferência lateral de genes e (3), descendência de um ancestral comum com perdas. A explicação mais parcimoniosa para a distribuição observada é a evolução convergente, a partir de precursores lineares *cyclotide-like* com, pelo menos, quatro origens independentes de ciclotídeos dentro da família Rubiaceae (GRUBER et al., 2008).

Dentre as famílias vegetais onde os ciclotídeos são mais encontrados, destacam-se Rubiaceae e Violaceae onde, segundo GRUBER, et al. (2008), não há relações filogenéticas diretas, o que sugere que ciclotídeos devem ter se originado de forma independente, mais de uma vez, e em famílias de parentesco distante. As diferenças na estrutura dos genes precursores dos diferentes ciclotídeos e presença de íntrons em alguns genes descartam a possibilidade de transferência lateral. Essa diferença estrutural nos ciclotídeos dessas duas famílias não favorece a existência de um ancestral comum entre estas duas famílias de plantas, fato que argumenta contra a evolução divergente. As informações combinadas e o conhecimento da estrutura genética até o momento permitem inferir que ciclotídeos provavelmente tenham surgido da evolução convergente, com a aquisição da capacidade de ciclização (GRUBER et al., 2008).

A descoberta de seqüências lineares, semelhantes a ciclotídeos em monocotiledôneas (Poaceae) (BASSE, 2005; MULVENNA et al., 2006), indica a

possibilidade (ou evidência) de uma mutação na posição onde estaria os resíduos N ou D (C-terminal) por um códon de parada, sugerindo assim, uma explicação para a biossíntese destes peptídeos lineares.

3.4. Espécies foco do estudo

3.4.1. Considerações gerais sobre a cultura de *Zea mays*.

O milho é uma cultura antiga, com centro de origem no México, sugere-se que o milho moderno seja resultado da domesticação de vários genomas de “teosinte” (uma gramínea selvagem), cuja domesticação possa ter ocorrido, provavelmente, há cerca de 9.000 anos atrás (WANG et al., 1999; MATSUOKA et al., 2002; KANE & RIESEBERG, 2005). De acordo com a classificação botânica, o milho é uma monocotiledônea, pertencente à família Poaceae, subfamília Panicoideae, tribo Maydeae, genero *Zea* e espécie *Zea mays*. (SILOTO, 2002). É uma espécie herbácea, anual, com caule do tipo colmo, constituído de nós e entrenós.

O genoma do milho está em torno de 2.500 Mb, podendo ser comparado ao tamanho do genoma humano, mas sua complexidade é maior devido à grande parte do genoma do milho ser composto por sequências repetitivas, sendo de 65 a 70% de retrotransposons, elementos transponíveis mediados por RNA (BENNETZEN, 1996). O milho apresenta 10 cromossomos que são estruturalmente diferentes e passam por mudanças dinâmicas na composição da cromatina (SANMIGUEL et al., 1998)

Considerado um dos cereais mais importantes cultivados no mundo, o milho apresenta-se como um cereal de alta qualidade nutricional estando para o Brasil, entre as graníferas mais importantes na produção agrícola, sendo destinada tanto para consumo humano como animal (SANTOS, 2002). Além disso, em países industrializados, derivados do milho podem ser encontrados em muitos produtos como: cosméticos, roupas, papel, pasta dental, tintas, loções, entre outros. Mais recentemente, milho vem sendo utilizado, também, na confecção de polímeros biodegradáveis, matérias-primas para as indústrias de embalagens e automobilística (MIRANDA PEIXOTO, 2002).

O milho também tem relações com aspectos sociais e culturais, servindo como parte da subsistência de muitas famílias rurais (KIRINO, 2003). Segundo o 8º levantamento do USDA (*United States Department of Agriculture*), haverá recorde na produção mundial de milho 2014/15, com cerca de 991,6 milhões de toneladas, superando a produção de 2013/14 que foi de 989,3 milhões de ton.

O Brasil aparece em 3º lugar de produção mundial, com 75 milhões de ton, perdendo para EUA e China (353,7 e 218,5 milhões de ton, respectivamente). Na produção 2013/14 o Brasil obteve uma maior produtividade 79,3 ton, em comparação a produção de 2013/14, essa queda na produção 2014/15 foi influenciada pela previsão de redução na área cultivada pelo cereal. No *ranking* da exportação, o Brasil ocupa o 2º lugar, com 19,5 milhões de ton, e em 1º o EUA, com 44,5 milhões de ton. Em relação ao consumo, o Brasil aparece na 4ª posição com 56,5 milhões de ton, o EUA lidera devido à previsão de consumo recorde 302,8 milhões de ton (USDA, 2014). Tal volume é projetado devido ao aumento esperado pela demanda entre as indústrias alimentícias e de etanol.

No Brasil a produção de milho é caracterizada por duas épocas para o plantio, primeira safra (ou safra de verão) e segunda safra (ou safrinha). O plantio de verão é realizado em todos os estados, durante o período chuvoso, que varia de agosto, na região Sul, até os meses de outubro/novembro, no Sudeste e Centro-Oeste. Porém, na região Nordeste, esse período ocorre no início do ano. Já a safrinha, a Conab (Companhia Nacional de Abastecimento) classifica como a plantada em Rondônia, Tocantins e em determinadas regiões da Bahia (CONAB, 2015). A safrinha refere-se ao milho de sequeiro, plantado extemporaneamente, geralmente de janeiro a março ou abril, quase sempre depois da soja precoce, predominantemente na região Centro-Oeste e nos estados do Paraná, São Paulo e Minas Gerais. (EMBRAPA, 2012)

Entretanto, a cultura sofre com doenças e pragas, o que pode comprometer a produção e qualidade. Relatos de perdas na produtividade devido ao ataque de patógenos têm sido frequentes. Algumas modificações ocorridas no sistema de produção, que resultaram no aumento da produtividade dessa cultura, foram, também, responsáveis pelo aumento da incidência e da severidade das doenças. As doenças que se destacam no plantio dessa cultura no Brasil são mancha branca, a cercosporiose, a ferrugem polissora, a ferrugem tropical, os enfezamentos vermelho e pálido, as

podridões de colmo e os grãos ardidos (EMBRAPA, 2012). As pragas que acometem a cultura do milho apresentam-se em duas fases, as de fase vegetativa e de fase reprodutiva, que variam de acordo com vários fatores como condições edafoclimáticas, sistemas de cultivo e fatores bióticos localizados.

A escolha da semente deve ser correta, com o intuito de alcançar alta produtividade, evitar doenças e pragas, adequação ao sistema de produção em uso e as condições edafoclimáticas, por isso na safra 2013/14, a Embrapa disponibilizou 467 cultivares, sendo 253 transgênicas e 214 cultivares melhoradas. Observa-se o maior número de transgênicas, sendo a primeira vez que isto ocorreu (EMBRAPA, 2012).

3.4.2. Considerações gerais sobre a cultura de *Secale cereale*

Embora ainda não se saiba a origem principal do centeio, acredita-se tenha partido do sudoeste da Ásia, a mesma região de origem do trigo, aveia e cevada (BUSHUK, 2001). Segundo Baier (1994), o centeio apresenta dois centros de origem, um na Anatólia e no Cáucaso, a leste da Turquia e ao norte do Irã, e outro, ao norte do Afeganistão e a oeste do Irã. Observa-se que nestas áreas ocorre grande diversidade genética, ocorrendo uma continua variação entre populações cultivadas e silvestres.

Inicialmente considerado um invasor do cultivo de trigo e cevada, o centeio foi disseminado pelo centro e norte da Europa, se expandindo para outras partes do mundo. Atualmente a cultura do centeio ocorre em todo o mundo, mas a maior produção está no hemisfério norte, ente os Montes Urais e o Mar Nórdico.

No Brasil o centeio foi introduzido por alemães e poloneses, no século XIX (BAIER, 1994), sendo o cultivo iniciado por cultivares estrangeiras. Em 1986 foi lançada a primeira cultivar brasileira – cultivar BR 1 pela Embrapa Trigo.

O centeio é classificado na família Poaceae (Gramineae), subfamília Pooideae, tribo Triticeae, gênero *Secale*, espécie *cereale* (AGEITEC, 2015). É utilizado diretamente ou em pré-misturas de pães, biscoitos, cereais matinais e produtos dietéticos. A forte vinculação do centeio com tradições culturais alimentares está presente na Finlândia, Polônia, Rússia, Alemanha, entre outros com o pão de centeio que é um pão escuro e ácido.

Ainda, na indústria farmacêutica, a cravagem do centeio ou esporão, estruturas arqueadas (esclerotos) oriundas da infecção do centeio pelo fungo *Claviceps purpurea*, dá origem a importantes alcaloides [(ergotamina, ergometrina, metilergonovina, metilsergida), os quais possuem propriedades no tratamento de hemorragias pós-parto, cefaleias de origem vascular, doença de Parkinson, patologias associadas à hiperprolactinemia, etc.) os quais podem causar intoxicação (ergotismo)]. Existem registros de uma epidemia na França nos anos de 94-95, com 20 mil pessoas mortas (TAVEIRA; CRUZ, 2008)

O centeio ocupa o oitavo lugar na lista mundial em termos de área colhida e de produção no mundo, com produção de 14,11 milhões de ton em 2012/2013. No período de 2008-2012 a União Europeia correspondeu a 58,8 % do consumo mundial de centeio (DE MORI et al., 2013). No período entre 2010-2012, o Brasil produziu 3,6 milhões de ton, sendo a média de área/ano de 2,3 mil hectares, correspondendo a 1,7 mil hectares (71,9%) no Rio Grande do Sul e 646 hectares (27,7%) no Paraná. O mercado do centeio é limitado e inconstante, sendo os moinhos os principais compradores do produto para a produção de farinhas para a panificação (EMBRAPA, 2013). O cultivo desta cultura no sul do Brasil é acometido por doenças como a ferrugem do colmo, a cravagem, a helmintosporiose, nanismo amarelo, entre outras.

As cultivares registradas no Ministério da Agricultura são a cultivar BR 1, BRS Progresso, BRS Serrano (da Embrapa), IPR 89 (do IAPAR, Instituto Agrônômico do Paraná) e Temprano (da Atlântica Sementes S.A.) (MINISTÉRIO DA AGRICULTURA, 2015). Algumas características da cultivar BR 1 são apresentadas na tabela 1.

Tabela 1. Características da cultivar BR 1 (MINISTÉRIO DA AGRICULTURA, 2015).

CARACTERÍSTICAS	
Cultivar	BR 1
Cruzamento	Seleção de plantas em populações coloniais
Ciclo	Precoce
Área de adaptação	RS
Potencial produtivo	> 2.200 kg/ha
REAÇÃO A DOENÇAS	
Ferrugem da folha	Resistente
Ferrugem do colmo	Suscetível
Giberela	Moderadamente resistente
Oídio	Resistente
Septoriose	Resistente
Mancha marrom	Resistente
Bacteriose	Resistente
Carvão	Resistente
Brusone	Sem informação
VMT	Moderadamente resistente
VNAC	Moderadamente resistente

3.5 Bancos de dados e ferramentas de bioinformática

A bioinformática, apresentou maior importância na área genômica na segunda metade da década de 80, com a geração de uma quantidade expressiva de sequências, devido aos primeiros projetos de sequenciamento em larga escala, aumentando a demanda por recursos computacionais, cada vez mais eficientes, no armazenamento e suas ferramentas para interpretação dos dados biológicos (ROUZÉ, et al., 1999; PROSDOCINI et al., 2002). De forma sucinta, a bioinformática é definida como sendo uma área multidisciplinar que engloba diversas áreas como, biologia, matemática, física, tecnologia da informação entre outras disciplinas (SABBATINI, 1999).

Informações sobre 7.262 genomas completos e 10.809 incompletos até 2014, podem ser acessados no banco do GOLD (*Genome OnLine Database*, disponível em <https://gold.jgi-psf.org/>). Tais projetos envolvem representantes da área ambiental, médica, industrial e organismos modelos importantes para desenvolvimento de pesquisa científica. Essa maior disponibilidade de dados gerados a partir de sequenciamento tem tornado a bioinformática cada vez mais essencial na manipulação e análise rápida deste grande volume de informação.

A concretização da bioinformática e da biologia computacional nos últimos anos tem favorecido aos pesquisadores acesso a um grande número de dados, propiciando abordagens globais e inovadoras no que diz respeito ao estudo da organização, estrutura e evolução de genomas (BINNECK, 2004; ABBY e DAUBIN, 2007) da predição de função de proteínas, interações e estrutura tridimensional (GINALSKI, 2006; LEE et al., 2007; SKRABANEK et al., 2008).

Focando no armazenamento desses grandes volumes de dados, além de permitir o acesso rápido e de forma eficiente, vários bancos de dados foram sendo criados, tanto de acesso público como privado, além de redes de acesso que admitissem a interação desses grupos (PROSDOCINI et al., 2002). Embora ainda existam muitos bancos de acesso restrito, é notória a tendência que dados gerados sejam disponibilizados e sejam acessíveis a toda comunidade científica, onde alguns bancos já disponibilizam seus dados como o *GenBank*, do NCBI (*National Center for Biotechnology Information*) dos EUA, que se destaca como um dos bancos mais consultados, uma vez que inclui aproximadamente 181.336.445 milhões de sequências (NCBI, 2015). O *GenBank* está integrado com duas organizações, o EMBL (*European Molecular Biology Laboratory*; Laboratório de Biologia Molecular Europeu) e o DDBJ (*DNA DataBank of Japan*; Banco de Dados de DNA do Japão), sendo que as três trocam dados diariamente. (NCBI, 2015).

Outro banco de dados criado para facilitar os estudos de genômica comparativa de plantas é o Phytozome (<http://phytozome.jgi.doe.gov/pz/portal.html>), o qual abriga mais de 41 genomas de plantas sequenciados e anotados, agrupados em 20 famílias de genes e, frequentemente anotados com auxílio de outros bancos de proteínas, tais como: o PFAM e o KEGG (*Kyoto Encyclopedia of genes and Genomes*) (PHYTOZOME, 2015). Informações acerca de estruturas 3D de proteínas e de ácidos nucleicos

apresentam-se, também, disponíveis em bancos como o PDB (*Protein Data Bank*, <http://www.rcsb.org/pdb/>) o qual tem sido fundamental no entendimento dos aspectos moleculares (PDB, 2015).

Com o campo dos peptídeos antimicrobianos crescendo rapidamente por causa da demanda por novos agentes antimicrobianos e, para melhor entender a função destes peptídeos, novos bancos de dados foram desenvolvidos para reunir informações sobre tais peptídeos. Informações sobre peptídeos em geral podem ser encontradas no APD - *The Antimicrobial Peptide Database* (<http://aps.unmc.edu/AP/about.php>), no qual estão disponibilizados 2.513 peptídeos antimicrobianos de diversos organismos (bactérias, fungos, plantas e animais); o CAMP - *Collection of Anti-Microbial Peptides* (<http://www.camp.bicnirrh.res.in/dbStat.php>), composto por 6.756 sequências de AMPs (das quais, 2.602 validadas experimentalmente, 1.716 patentes e 682 estruturas de vários organismos). Ainda, dentre os outros bancos de AMPs, destacam-se: aqueles isolados de bactérias (BAGEL <http://bagel2.molgenrug.nl/>; BACTIBASE <http://bactibase.pfba-lab-tun.org/main.php>), Peptídeos inibidores de HIV (HIPdb <http://crdd.osdd.net/servers/hipdb/>), Defensinas base de conhecimento (<http://defensins.bii.a-star.edu.sg/>), de plantas (PhytAMP <http://phytamp.pfba-lab-tun.org/main.php>), além do banco com acesso a informações sobre abordagens recombinantes de AMPs (RAPD - *Recombinantly-produced Antimicrobial Peptides Database*, <http://faculty.ist.unomaha.edu/chen/rapd/>).

Dentre os AMPs conhecidos, os peptídeos cíclicos vêm se destacando devido ao seu interesse pela indústria farmacêutica e agrícola, sua estabilidade térmica e enzimática, sua conformação (cíclica), trazendo novas perspectivas de uso. Interessantemente, os ciclotídeos tem se apresentado um excelente candidato a inseticidas, como confirmado em estudos no combate insetos causadores de doenças e pragas em plantas cultivadas (PINTO et al, 2012).

Um banco de dados direcionado a tais peptídeos cíclicos é o Cybase (*The Database of Cyclic Proteins*), desenvolvido pelo grupo do Instituto de Biociências Moleculares da Universidade de Queensland, Austrália e liderado pelo professor David Craik. Este grupo se dedica à descoberta e caracterização de proteínas circulares, em especial da família que nomearam de ciclotídeos. O interesse envolve a determinação da estrutura por Ressonância Magnética Nuclear de uma gama de proteínas ricas em

ligações dissulfeto, que apresentam aplicações no planejamento de fármacos, incluindo organismos como cobras, aranhas, rãs e plantas. Este banco possui acesso a mais de 800 proteínas de 103 espécies, dentre estas 532 são ciclotídeos de 55 espécies, além de acesso a sequências nucleotídicas e estruturas.

Além do armazenamento e disponibilização de um grande número de informações, um banco de dados também disponibiliza várias ferramentas destinadas ao processamento e análise desta grande quantidade de dados. Estas ferramentas permitem, por exemplo, a análise de predição *in silico* de sequências desconhecidas e novos genes, por meio de análises de alinhamentos comparativos (KENT et al., 2002).

O alinhamento local (BLAST – *Basic Local Alignment Search Tool*), ferramenta mais popular e acessível (ALTSCHUL et al., 1990), permite comparar e alinhar uma sequência de DNA ou proteína “*query*” com todas as sequências, retornando assim, aquelas sequências (DNA ou proteína) com maior homologia encontradas no banco de dados pesquisado.

O BLAST é dividido em várias funcionalidades, o que depende do tipo de busca e retorno que é desejado. Por exemplo, o BLASTn compara sequências de nucleotídeos contra o banco de dados de nucleotídeos; o BLASTp, compara sequências de aminoácidos com o banco de proteínas; o BLASTx, traduz a sequência de DNA para proteína e a compara com banco de dados proteico; o tBLASTn, compara sequências proteicas com o banco de dados de nucleotídeos traduzido; e o tBLASTx traduz tanto a sequência nucleotídica quanto o banco de dados de nucleotídeos e os compara.

As sequências nucleotídicas também podem ser traduzidas, sendo possível analisar sua sequência proteica. Um exemplo de ferramenta que traduz as sequências nucleotídicas e encontra todos os quadros abertos de leitura (NCBI, 2015) é o ORF-Finder (*Open Reading Frame Finder*).

Para uma análise comparativa de sequências, seja DNA ou proteína, a ferramenta de alinhamentos múltiplos mostra a homologia posicional entre as bases nucleotídicas ou aminoácidos presentes nas sequências, que podem ser de espécies diferentes, o CLUSTAL (THOMPSON et al., 1997) é um dos softwares mais utilizados para efetuar estes alinhamentos múltiplos.

Outras ferramentas e software podem ser acessados no ExPASy (*Bioinformatics Resource Portal*) que foi disponibilizado pelo SIB (*Swiss Institute of Bioinformatics*) e

podem ser acessados para análises em diversas áreas das ciências da vida como proteômica, genômica, biologia de sistemas, filogenia, transcriptômica e genética de populações, entre outras. Entre as ferramentas disponíveis está o *Swiss Model* - utilizada para modelagem por homologia de proteínas. A modelagem por homologia é um método no qual se obtém a estrutura tridimensional de uma sequência proteica, a partir da similaridade desta com uma ou mais estruturas proteicas já conhecidas (ROST et al., 1996; KOLINSKI et al., 1999).

Os recursos e bancos de dados descritos são uma pequena parte do que se tem disponível na *internet* de forma gratuita, além dos recursos que podem ser baixados e utilizados localmente, a depender de que análise se busca. Tais ferramentas e abordagens têm auxiliado o estudo em diversas áreas da comunidade científica.

IV. REFERÊNCIAS BIBLIOGRÁFICAS

- ABBY, S.; DAUBIN, V. Comparative genomics and the evolution of prokaryotes. **Trends Microbiology**, v. 15, p. 135-141, 2007.
- AGRIOS, G. N. Plant Pathology. 5. Ed. San Diego, Califórnia: Elsevier Academic Press, 2004. 922p.
- APD - *The Antimicrobial Peptide Database*. Disponível em: <<http://aps.unmc.edu/AP/about.php>>. Acesso em: 18 fev. de 2015.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.
- AVRUTINA, O.; SCHMOLDT H. U.; GABRIJELCIC-GEIGER, D.; et al. Trypsin inhibition by macrocyclic and open-chain variants of the squash inhibitor MCoTI-II. **Biological chemistry**, v. 386, p. 1301–1306, 2005.
- BACTIBASE – Disponível em: <<http://bactibase.pfba-lab-tun.org/main.php>>, Acesso em: 18 fev. de 2015.
- BAGEL - Disponível em: <<http://bagel2.molgenrug.nl/>>. Acesso em: 18 fev. de 2015
- BAIER, A. C. **Centeio**. Passo Fundo: EMBRAPA-CNPT, 1994, 29p. (EMBRAPA-CNPT. Documentos, 15).
- BAIER, A. C. Centeio. In: BAIER, A. C; FLOSS, E. L.; AUDE, M. I. S. **As lavouras de inverno 1: aveia, centeio, triticale, colza, alpiste**. Rio de Janeiro: Globo, 1988. p. 107-130.
- BARBETA, B. L. et al. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. **Proceedings of the National Academy of Sciences**, v. 105, n. 4, p. 1221–1225, 2008.
- BASSE, C. W. Dissecting defense-related and developmental transcriptional responses of maize during *Ustilago maydis* infection and subsequent tumor formation. **Plant physiology**, v. 138, n. 3, p. 1774–1784, 2005.
- BELARMINO, L. C. et al. EST-Database Search of Plant Defensins – An Example Using Sugarcane, a Large and Complex Genome. **Current protein & peptide science**, v. 11, p. 1–7, 2010.
- BENKO-ISEPPON, A. M. et al. Prospecção de Genes de Resistência à Seca e à Salinidade em Plantas Nativas e Cultivadas (Exploration of Genes for Resistance to

- Drought and Salinity in Native and Cultivated Plants). **Revista Brasileira de Geografia Física**, v. 4, n. 6, p. 1112–1134, 2012.
- BENKO-ISEPPON, A.M. et al. Overview on plant antimicrobial peptides. **Current Protein & Peptide Science**, v. 11, n. 3, p. 181-188, 2010.
- BINNECK, E. As ômicas: integrando a bioinformação - O papel da bioinformática em expansão. **Biotecnologia Ciência**, v. 32, p. 28-37, 2004.
- BENNETZEN, Jeffrey L. The contributions of retroelements to plant genome organization, function and evolution. **Trends in microbiology**, v. 4, n. 9, p. 347-353, 1996.
- BOIX, E.; NOGUES, M. V. Mammalian antimicrobial proteins and peptides: overview on the RNase a superfamily members involved in innate host defence. **Molecular bioSystems**, v. 3, n. 5, p. 317-35, 2007.
- BOKESCH, H. R. et al. A Novel Anti-HIV Macrocyclic Peptide from *Palicourea condensata*. **Journal of natural products**, v. 64, n. 2, p. 249–250, 2001.
- BROEKAERT, W. F. et al. Antimicrobial peptides from plants. **Critical Reviews in Plant Sciences**, v. 16, n. 3, p. 297-323, 1997.
- BROGDEN, K. A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? **Nature Reviews Microbiology**, v. 3, n. 3, p. 238-250, 2005.
- BULET, P.; STÖCKLIN, R.; MENIN, L. Anti-microbial peptides: from invertebrates to vertebrates. **Immunological Reviews**, v. 198, p. 169–84, 2004.
- BUSHUK, W. Rye Production and Uses Worldwide1. 2001.
- CAMP - *Collection of Anti-Microbial Peptides*. Disponível em: <<http://www.camp.bicnirrh.res.in/dbStat.php>> Acesso em: 18 fev. de 2015.
- CANO-DELGADO, A.; PENFIELD, S.; SMITH, C.; CATLEY, M.; BEVAN, M. Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*. **The Plant Journal**, v. 34, p. 351-362, 2003.
- CHAN, L.Y. et al. Isolation and characterization of peptides from *Momordica cochinchinensis* seeds. **Journal of Natural Products**, v. 72, p. 1453–8, 2009.
- CHEN, B. et al. Isolation and Characterization of Novel Cyclotides from *Viola hederacea*. Solution structure and anti-HIV activity of vh-1, a leaf-specific expressed cyclotide. **Journal Biological Chemistry**, v. 280, n. 23, p. 22395-22405, 2005.

- CLAESON, P.; GOÈRANSSON, U.; JOHANSSON, S.; LUIJENDIJK, T.; BOHLIN, L. Fractionation Protocol for the Isolation of Polypeptides from Plant Biomass. **Journal of Natural Products**, v. 61, p. 77-81, 1998.
- COLGRAVE, M. L. et al. Anthelmintic activity of cyclotides: in vitro studies with canine and human hookworms. **Acta tropica**, v. 109, n. 2, p. 163–166, 2009.
- COMPANHIA NACIONAL DE ABASTECIMENTO. Acompanhamento da safra brasileira de grãos, v. 2 – safra 2014/2015, n. 6 – Sexto Levantamento, mar. 2015. Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/15_03_11_14_07_48_boletim_graos_marco_2015.pdf>.
- COSTA, R. V.; CASELA, C. R.; COTA, L.V. Doenças. In: CRUZ, J. C. (Ed.). Cultivo do milho. 6. ed. Sete Lagoas: Embrapa Milho e Sorgo, 2010. (Embrapa Milho e Sorgo. Sistema de produção, 1).
- CRAIK, D.J. Plant cyclotides: circular, knotted peptide toxins. **Toxicon**, v. 39, n. 12, p. 1809-1813, 2001.
- CRAIK, D. J.; DALY, N. L.; BOND, T.; WAINE, C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. **Journal of molecular biology**. v. 294, p. 1327–1336, 1999.
- CRAIK, D.J. et al. Discovery and applications of the plant cyclotides. **Toxicon**, v. 56, p. 1092–1102, 2010.
- CRAIK, D.J. et al. Discovery and structures of the cyclotides: novel from plants, **Letters in Peptide Science**, v. 8, p. 119-128, 2002.
- CYBASE - *The Database of Cyclic Proteins*. Disponível em: <<http://www.cybase.org.au/>>. Acesso em: 18 fev. de 2015.
- DA ROCHA PITTA, M.G.; DA ROCHA PITTA, M.G.; GALDINO, S.L. Development of novel therapeutic drugs in humans from plant antimicrobial peptides. **Current Protein & Peptide Science**, v. 11, n. 3, p. 236-247, 2010.
- DALY, N.L.; ROSENGREN, K.J.; CRAIK, D.J. Discovery, structure and biological activities of cyclotides. **Advanced Drug Delivery Reviews**, v. 61: p. 918-930, 2009.
- DEFENSINS KNOWLEDGEBASE. Disponível em: <<http://defensins.bii.a-star.edu.sg/>>. Acesso em: 18 fev. de 2015.

- DE MORI, C.; NASCIMENTO, A.J.; MIRANDA, M.Z Aspectos econômicos e conjunturais da cultura do centeio no mundo e no Brasil. Passo Fundo: EMBRAPA Trigo, 2013, 27p (EMBRAPA Trigo. Documentos, 142)
- DROIT, A.; POIRIER, G. G.; HUNTER, J. M. Experimental and bioinformatic approaches for interrogating protein-protein interactions to determine protein function. **Journal of Molecular Endocrinology**, v. 34, n. 2, p. 263–80, abr. 2005.
- EDWARDS, D.; BATLEY, J. Plant bioinformatics: from genome to phenome. **Trends in Biotechnology**, v. 22, n. 5, p. 232–7, 2004.
- FELIZMENIO, Q.M.E.; DALY, N.L.; CRAIK, D.J. Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis* **Journal Biological Chemistry**, v. 276, p. 22875-22882, 2001.
- GANZ, T. Defensins: antimicrobial peptides of vertebrates. **Comptes Rendus Biologies**, v. 327, p. 539- 549, 2004.
- GARCÍA-OLMEDO, F., PALENZUELA, R. P., MOLINA A. Antibiotic activities, hydrogen peroxide and peroxynitrite in plant defense. **FEBS Letters**, v. 498, p. 219–222, 2001.
- GINALSKI, K. Comparative modeling for protein structure prediction. **Current Opinion in Structural Biology**, v. 16, n. 2, p. 172-177, 2006.
- GIUDICI, M. et al. Antifungal effects and mechanism of action of viscotoxin A3. **FEBS Journal**, v. 273, p. 72-83, 2006.
- GODOY, C. V.; AMORIM, L.; BERGAMIN FILHO, A. Alterações na fotossíntese e na transpiração de folhas de milho infectadas por *Phaeosphaeria maydis*. **Fitopatologia Brasileira**, v. 26, n. 2, p. 209–213, 2001.
- GOLD Genomes on line database. Disponível em: <<https://gold.jgi-psf.org/>>. Acesso em: 18 de fev 2015.
- GORASSON, U.; LUIJENDIJK, T.; JOHANSSON, S.; BOHLIN, L.; CLAESON, P. Seven novel macrocyclic polypeptides from *Viola arvensis*. **Journal of Natural Products**, v. 62, p. 283–6, 1999,
- GRAN, L. On the effect of a polypeptide isolated from "Kalata-kalata" (*Oldenlandia affinis* DC) on the oestrogen dominated uterus. **Acta Pharmacol Toxicol**, v. 33, p. 400-408, 1973.

- GRAN, L. Oxitocic principles found in *Oldenlandia affinis*. **Meddelelser Norsk Farmasøytisk Selskap**, v. 32, p.173–180, 1970.
- GRUBER, C. W. et al. Distribution and evolution of circular miniproteins in flowering plants. **The Plant Cell Online**, v. 20, n. 9, p. 2471–2483, 2008.
- GUSTAFSON, K. R. et al. Circulins A and B: Novel HIV-inhibitory macrocyclic peptides from the tropical tree *Chassalia parvifolia*. **Journal of the American Chemical Society**, v. 116, p. 9337-9338, 1994.
- HADLEY, E. B.; HANCOCK, R. E. Strategies for the discovery and advancement of novel cationic antimicrobial peptides. **Current Topics in Medicinal Chemistry**, v. 10, n. 18, p. 1872-81, 2010.
- HALLOCK, Y.F. et al. Cycloviolins A-D, anti-HIV macrocyclic peptides from *Leonia cymosa*. **Journal of Organic Chemistry**, v. 65, p.124-8, 2000.
- HANCOCK, R. E. W.; SCOTT, M. G. The role of antimicrobial peptides in animal defenses. **Proceedings of the National Academy of Sciences**, v. 97, n. 16, p. 8856-8861, 2000.
- HEATH, M.C. Hypersensitive response-related death. **Plant Molecular Biology**, v. 44, p. 321-334, 2000.
- HERNANDEZ, J.F. et al. Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. **Biochemistry**, v. 39, p. 5722–30, 2000.
- HIPdb. Disponível em: <<http://crdd.osdd.net/servers/hipdb/>>. Acesso em: 18 fev. de 2015.
- JELINEK, R.; KOLUSHEVA, S. Membrane interactions of host-defense peptides studied in model systems. **Current Protein & Peptide Science**, v. 6, p. 103–114, 2005.
- JENNINGS, C. et al. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. **Proceedings of the National Academy of Sciences**, v. 98, n. 19, p. 10614-10619, 2001.
- JENNINGS, C. V. et al. Isolation, solution structure, and insecticidal activity of kalata B2, a circular protein with a twist: do Möbius strips exist in nature? **Biochemistry**, v. 44, n. 3, p. 851-60, 2005.
- JENSSEN, H.; HAMILL, P. E.; HANCOCK, R. E. W. Peptide antimicrobial agents. **Clinical microbiology reviews**, v. 19, n. 3, p. 491-511, 2006.

- KANE, N. C.; RIESEBERG, L. H. Maize genetics: the treasure of the Sierra Madre. **Current Biology**, v. 15, n. 4, p. R137–R139, 2005.
- KENT, W.J.; SUGNET, C.W.; FUREY, T.S.; ROSKIN, K.M. et al. The human genome browser at UCSC. **Genome Research**, v. 12, p. 996-1006, 2002.
- KEYMANESH, K.; SOLTANI, S.; SARDARI, S. Application of antimicrobial peptides in agriculture and food industry. **World Journal of Microbiology and Biotechnology**, v.25, p. 933–944, 2009.
- KOLINSKI, A.; ROTKIEWICZ, P.; ILKOWSKI, B.; SKOLNICK, J. A method for the improvement of threading-based protein models. **Proteins**, v. 37, n. 4, p. 592-610, 1999.
- LEE, D.; REDFERN, O.; ORENCO, C. Predicting protein function from sequence and structure. **Nature Reviews Molecular Cell Biology**, v. 8, p. 995-1005, 2007.
- LOBO, D. et al. Antifungal Pisum sativum defensin one interacts with *Neurospora crassa* cyclin F related to cell cycle. **Biochemistry**, v. 46, p. 987–996, 2007.
- MATSUOKA, Y. et al. A single domestication for maize shown by multilocus microsatellite genotyping. **Proceedings of the National Academy of Sciences**, v. 99, n. 9, p. 6080–6084, 2002.
- MENEZES, H.; JARED, C.; Immunity in plants and animals: common ends through different means using similar tools. **Comparative Biochemistry and Physiology - Part C: Toxicology & Pharmacology**, v. 132, p.1-7, 2002.
- MINISTÉRIO DA AGRICULTURA. Cultivares registradas. Disponível em: <http://extranet.agricultura.gov.br/php/snpc/cultivarweb/cultivares_registradas.php>. Acesso em:
- MIRANDA-PEIXOTO, C. O milho: o rei dos cereais – da sua descoberta há 8000 anos até as plantas transgênicas. Seed News, 2002. Disponível em <http://www.seednews.inf.br/portugues/seed62/milho62.shtml>. Acesso em: 10 de dezembro de 2014.
- MULVENNA, J. P. et al. Discovery of cyclotide-like protein sequences in graminaceous crop plants: ancestral precursors of circular proteins? **The Plant Cell Online**, v. 18, n. 9, p. 2134–2144, 2006.
- MYLNE, J. S. et al. Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. **Nature Chemical Biology**, v. 7, n. 5, p. 257–259, 2011.

- NCBI National Center for Biotechnology Information. Disponível em: <<http://www.ncbi.nlm.nih.gov/>>. Acesso em: 18 de fev 2015.
- NGUYEN, G. K. T. et al. Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the Fabaceae family. **Journal of Biological Chemistry**, v. 286, n. 27, p. 24275-24287, 2011.
- NGUYEN, L.T.; HANEY, E.F.; VOGEL, H.J. The expanding scope of antimicrobial peptide structures and their modes of action. **Trends in Biotechnology**, v. 29, n. 9, p. 464-72, 2011.
- NGUYEN, Giang Kien Truc et al. Discovery of linear cyclotides in monocot plant *Panicum laxum* of Poaceae family provides new insights into evolution and distribution of cyclotides in plants. **Journal of Biological Chemistry**, v. 288, n. 5, p. 3370-3380, 2013.
- OELKE, E. A. et al. **Rye**. Alternative field crops manual. Disponível em: <<https://www.hort.purdue.edu/newcrop/afcm/rye.html>>. Acesso em: 3 de jan de 2015.
- PAPO, N.; SHAI, Y. Can we predict biological activity of antimicrobial peptides from their interactions with model membranes? **Peptides**, v. 24 p. 1693-1703, 2003.
- PhytAMP. Disponível em: <<http://phytamp.pfba-lab-tun.org/main.php>>. Acesso em: 18 fev. de 2015.
- PICCHI, D.G. et al. Peptídeos cíclicos de biomassa vegetal: características, diversidade, biossíntese e atividades biológicas. **Química Nova**, v. 32, n. 5, p. 1262-1277, 2009.
- PINTO, M.F. et al. Identification and structural characterization of novel cyclotide with activity against an insect pest of sugar cane. **Journal of Biological Chemistry**, v. 287, p. 134-47, 2012.
- PLAN, M. R. R. et al. Backbone cyclised peptides from plants show molluscicidal activity against the rice pest *Pomacea canaliculata* (golden apple snail). **Journal of agricultural and food chemistry**, v. 56, n. 13, p. 5237–5241, 2008.

- PONCE DE LEÓN, I.; MONTESANO, M. Activation of Defense Mechanisms against Pathogens in Mosses and Flowering Plants. **International journal of molecular sciences**, v. 14, n. 2, p. 3178–3200, 2013.
- PORTO, W. F. et al. In silico identification of novel hevein-like peptide precursors. **Peptides**, v. 38, n. 1, p. 127–36, 2012.
- PRITHIVIRAJ, B.; PASCHKE, M.W.; VIVANO, J.M. Root Communication: The Role of Root Exudates. **Encyclopedia of Plant and Crop Science**, v. 1, p. 1-4, 2006.
- PROSDOCINI, F. et al. Bioinformática: manual do usuário. **Biotecnologia, Ciência & Desenvolvimento**, v. 29, p. 12-25, 2002.
- RAPD - *Recombinantly-produced Antimicrobial Peptides Database*. Disponível em: <<http://faculty.ist.unomaha.edu/chen/rapd/>>. Acesso em: 18 fev. de 2015.
- REDDY, K. V. R.; YEDERY, R. D.; ARANHA, C.; Antimicrobial peptide: Premises and promises. **International Journal of Antimicrobial Agents**, v. 24, p. 536, 2004.
- REIS, E. M. **Diagnose, patometria e controle de doenças de cereais de inverno**. [S.l.]: Universidade de Passo Fundo, Faculdade de Agronomia e Medicina Veterinária, 2001.
- RESENDE, J. M. et al. "Solution NMR structures of the antimicrobial peptides phylloseptin-1,-2, and-3 and biological activity: the role of charges and hydrogen bonding interactions in stabilizing helix conformations". **Peptides**, v. 29, n. 10, p. 1633-1644, 2008.
- ROST, B.; FARISELLI, P.; CASADIO, R. Topology prediction for helical transmembrane proteins at 86% accuracy. **Protein Science**, v. 5, n. 8, p. 1704-1718, 1996.
- ROUZÉ, P.; PAVY, N.; ROMBAUTS, S. Genome annotation: wich tools do we have for it? **Current Opinion in Plant Biology**, v. 2, p. 90-95, 1999.
- SABBATINI, R. M. E. A bioinformática chegou. **Revista Informática Médica**, v. 2, p. 4-6, 1999.
- SKRABANEK, L.; SAINI, H.K.; BADER, G.D.; ENRIGHT, A.J. Computational prediction of protein-protein interactions. **Molecular biotechnology**, v. 38, n. 1, p. 1-17, 2008.

- SANDBERG, F., Étude sur les plantes médicinales et toxiques d'Afrique équatoriale. Cahiers de la Mabokey. 12 Rue de Buffon, Paris, Tome III, Fascicule 1, p. 27. 1965.
- SANMIGUEL, Phillip et al. The paleontology of intergene retrotransposons of *maize*. **Nature genetics**, v. 20, n. 1, p. 43-45, 1998.
- SANTOS, A. O.; PRADO, H. DO. Análise de interações solo-planta-clima em zonas diferenciadas de áreas de cultivo de milho. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, n. 1, p. 101–106, 2002.
- SILOTO, R. C. **Danos e biologia de Spodoptera frugiperda (JE Smith, 1797) (Lepidoptera: Noctuidae) em genótipos de milho**. [S.l.]: Universidade de São Paulo, 2002.
- SLETTEN, K.; GRAN, L. Some molecular properties of kalatapeptide B-1. **Meddelelser Norsk Farmasøytisk Selskap**, v. 35, p. 69-82, 1973.
- SPOEL, S.H.; DONG, X. How do plants achieve immunity? Defence without specialized immune cells. **Nature Reviews Immunology**, v. 12, n. 2, p. 89-100, 2012.
- TAM, J. P. et al. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. **Proceedings of the National Academy of Sciences**, v. 96, n. 16, p. 8913–8918, 1999.
- TAVEIRA, M.; CRUZ, S. **Alcalóides da cravagem do centeio**. 2008. Disponível em: <http://www.ff.up.pt/toxicologia/monografias/ano0708/g25_centeio/cut.html>. Acesso em: 20 dez. 2012. In: DE MORI, C.; NASCIMENTO, A.J.; MIRANDA, M.Z Aspectos econômicos e conjunturais da cultura do centeio no mundo e no Brasil. Passo Fundo: EMBRAPA Trigo, 2013, 27p (EMBRAPA Trigo. Documentos, 142)
- THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, J. et al. The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Research**, v. 25, p. 4876-4882, 1997.
- TRABI, M. et al. Variations in Cyclotide Expression in Viola Species. **Journal Natural Products**, v. 67, p. 806-810, 2004.

- TSOI, A.Y.K.; NG, T.B.; FONG, W.P. Antioxidative effect of a chymotrypsin inhibitor from *Momordica cochinchinensis* (Cucurbitaceae) seeds in a primary rat hepatocyte culture. **Journal of Peptide Science**, v. 11, p. 665–8, 2005.
- VAN LOON, L. C.; REP, M.; PIETERSE, C. M. J. Significance of inducible defense-related proteins in infected plants. **Annual review of phytopathology**, v. 44, p. 135-62, 2006.
- WANG, G.; LI, X.; WANG, Z. "APD2: the updated antimicrobial peptide database and its application in peptide design". **Nucleic Acids Research**, v. 37, n. 4, p. 933-937, 2008.
- WANG, R.-L. et al. The limits of selection during maize domestication. **Nature**, v. 398, n. 6724, p. 236–239. 1999.
- WITHERUP, K. M. et al. Cyclopsychotride A, a biologically active, 31- residue cyclic peptide isolated from *Psychotria longipes*. **Journal Natural Products**, v. 57, p. 1619- 1625, 1994.
- YEAMAN, M. R.; YOUNT, N. Y. "Mechanisms of antimicrobial peptide action and resistance". **Pharmacological Reviews**, v. 55, n. 1, p. 27-55, 2003.
- AGEITEC - Agência Embrapa de informação tecnológica. 2015. <http://www.agencia.cnptia.embrapa.br/gestor/centeio/Abertura.html>. Acesso em: 15 de dezembro de 2014.
- EMBRAPA. 2012. http://www.cnpms.embrapa.br/publicacoes/milho_8_ed/index.htm Acesso em: 10 de dezembro de 2014.
- KIRINO, P. Farmers and food security. IFA – FAO. Agriculture Conference: Rome – Italy, 2003. Disponível: <http://www.fao.org>. Acesso em: 10 de dezembro de 2014.
- PHYTOZOME, 2015 <http://www.phytozome.net/>. Acesso em 3 de janeiro de 2015.
- SISTEMA DE PRODUÇÃO EMBRAPA, 2015, <https://www.spo.cnptia.embrapa.br/home>. Acesso em: 15 de dezembro de 2014.
- USDA (United States Department of Agriculture) disponível em: http://az545403.vo.msecnd.net/uploads/2014/12/boletim_safra-mundial-milho_dezembro2014.pdf. Acesso em: 10 de dezembro de 2014.

CAPÍTULO I

Artigo aceito na Revista *Current Protein & Peptide Science*

Plants Defense-related Cyclic Peptides: Diversity, Structure and Applications

Running Title: Defense related cyclic peptides from plants

Sheyla Carla Barbosa da Silva Lima¹, Ana Maria Benko-Iseppon¹, João Pacifico Bezerra Neto¹, Lidiane Lindinalva Barbosa Amorim², José Ribamar Costa Ferreira Neto¹, Sergio Crovella^{1,3} and Valesca Pandolfi^{1,*}.

¹*Federal University of Pernambuco, Center of Biological Sciences, Department of Genetics, Av. Prof. Moraes Rêgo, 1235, CEP 50670-420, Recife, PE, Brazil;*

²*Federal Institute of Education, Science and Technology of Piauí, Campus Oeiras, Rua Projetada, CEP 64500-000, Oeiras, PI, Brazil.*

³*Genetic Service, IRCCS Burlo Garofolo and Department of Developmental and Reproductive Sciences, University of Trieste -Via dell'Istria, 65/1- 34137 Trieste, Italy.*

**Corresponding author:* E-mail: valescapandolfi@gmail.com

ABSTRACT

Plant growth is prone to several unfavorable factors that may compromise or impair development and survival, including abiotic or biotic stressors. Aiming at defending themselves, plants have developed several strategies to survive and adapt to such adversities. Cyclotides are a family of plant-derived proteins that exhibit a diverse range of biological activities including antimicrobial and insecticidal activities that actively participate in plant defense process. Three main categories of peptides have been described: (i) Cyclotides (ii) Sunflower Trypsin Inhibitor (SFTI) and (iii) peptides MCoTI-I and II, from *Momordica cochinchinensis*. They comprise proteins of approximately 30 amino acids, containing a head-to-tail cyclized backbone, with three disulfide bonds configured in a cystine knot topology, therefore bearing greater peptide stability. Given their features and multifunctionality, cyclotides stand out as promising sources for the discovery of new antimicrobial agents. The present review infers on cyclotide occurrence, abundance and action in plants, also their diversity and evolution. Considerations regarding their use in the context of biomedical and agronomical sciences uses are also carried out.

Key Words: cyclotides; MCoTI-I/II; SFTI; disulfide bonds; protease mediated defense.

INTRODUCTION

Along their evolutionary processes plants have naturally developed effective and widely diversified mechanisms (including the production of a series of primary and secondary compounds) to defend themselves against the attack from a wide variety of pathogens, as fungi, bacteria and viruses [1, 2]. Usually, the plant protection is accomplished after pathogen recognition, followed by activation of defense mechanisms that help overcoming the deleterious effects caused by the invading microorganism. Even if the plant cells are able to recognize a variety of molecular signatures of microorganisms, the speed and efficiency of this activation are essential for the success of plant defense [3]. Regardless of whether the defense mechanism is constitutive or induced (by pathogenic microorganisms), defense responses in plants involve gene activation and induction of a signal transduction network, as well as responses related to abiotic stress [4], leading to the establishment of physical and chemical barriers.

Among the main mechanisms used in the course of plant defense some should be highlighted: Hypersensitive Response (HR), Systemic Acquired Resistance (SAR), induction of Pathogenesis-Related proteins (PR-proteins) and synthesis of signaling compounds such as salicylic acid (SA) and hydrogen peroxide (H_2O_2). HR is characterized by a rapid and localized response that occurs at the site of infection by the pathogen. This process leads to a "collapse" of the tissue due to toxic compounds released around infection site. Such compounds may also act directly on the pathogen, causing its death [5,6]. SAR mechanism involves a cascade of signaling events (associated to the plant-pathogen interaction) promoting changes in cell metabolism, resulting in reduced disease severity [6,7].

A wide variety of organisms are also able to synthesize antimicrobial peptides (AMPs) that work in the first line of defense. These peptides belong to the PR-families PR-12, PR-13 and PR-14 [2] and are characterized by the presence of cysteine (Cys) residues (4, 6 and 8) connected by disulfide bridges, providing structural stability to the molecule [8]. Expressed constitutively or induced by biotic stimuli, AMPs are associated with a wide spectrum of biological roles, with emphasis on their antimicrobial activity [9,10] in most living organisms, including humans [11].

Reports indicated effectiveness of these molecules as inhibitors of digestive enzymes such as serine protease [12] and α -amylases [13]. Another representatives are known to inhibit protein synthesis [14,15], besides demonstrating efficacy as inhibitors of viruses and cancer cell growth [16-18], with the advantage of generally presenting a low toxicity to mammalian and plant cells [19,20]. AMP properties reflect their genetic and protein structure (i.e. small proteins, 12-50 aa; ~50 kDa; encoded by single genes; presenting positive charge, cationic at physiological pH, due to lysine, arginine and histidine excess), commonly rich in a type of amino acid and assigned in several subclasses that share some features responsible for antimicrobial activity, tolerance to acids and organic solvents, thermal stability, and broad biological activity) [21-25]. Even though AMPs present quite similar dimensional structure and physical properties, they generally exhibit low sequence homology, especially when comparing distinct peptide classes [26]. However, members of a given family show a relatively conserved folding pattern. They adopt a three-dimensional structure that involves the formation of secondary structural elements (beta-sheets and alpha helices) stabilized by intramolecular disulfide bridges, forming a rigid structure and giving, therefore, greater stability [2, 27].

Despite the great structural diversity displayed by AMPs, they may be grouped into distinct classes considering their secondary structures [6]. The prevalent structures regard α -helices, formed only when the peptide is in contact with the cell membrane, being the β -sheets stabilized by 2-4 disulfide bonds that occasionally present a small portion of α -helical structure. On the other hand, the less frequent structures are bent, formed by simple disulfide bonds or due to the presence of proline residues [28] or still extended, characterized by the predominance of one or two amino acid residues in the primary sequence. Major plant AMP groups are represented by cyclotides, defensins, lipid transfer proteins (LTPs), snakins, thionins, hevein- and knottin-type of protein [2,20]. Some of these AMPs play an important role in biocontrol of plant diseases, as verified in transgenic plants expressing partial or total resistance [29].

A fascinating plant AMP subgroup includes the cyclic peptides (CPs). They regard proteins ranging from 12 to 80 amino acids, with a “head-to-tail” cyclic backbone, stabilized by one or more disulfide bonds [30]. Such physical and chemical properties have shown advantages over open-chain (linear) peptides [31]. The structural conformation is created due to the lack of free N- and C- termini, creating a continuous, rigid and stable structure, resistant to thermal degradation and cleavage by proteolytic enzymes [32-34]. Moreover, the cyclization of peptides and proteins reduces their flexibility, giving to such compounds a higher binding affinity to receptors [35].

This review will focus on the current knowledge of main cyclic peptides described in plants with a known role in plant-pathogen defense: Cyclotides, SFTI (Sunflower Trypsin Inhibitor) and the MCoTI-I/II (*M. cochinchinensis* Trypsin Inhibitors I and II, also referred to as cyclic knottins).

1. PLANT CPs WITH ANTIMICROBIAL PROPRIETIES

1.1. Structural aspects and classification

Over the years, hundreds of linear peptides, including defensins, thionins, glycine-rich proteins, snakins, 2S albumin-type and hevein-like proteins, have been discovered and characterized [36-39]. According to Colan and Anderson [40], the three major classes of plant CPs include (i) small CPs such as Segetalina A (*Vaccaria segetalis*) [41], (ii) a class of trypsin inhibitors named SFTI (Sunflower Trypsin Inhibitor) and (iii) cyclotides. In turn, Thorstholm and Craik [30] grouped the CPs into only two classes: SFTI (Sunflower Trypsin Inhibitor) and cyclotides. The present review will focus on features of CPs considering three groups: (i) cyclotides, (ii) SFTI (considered a homologue of the Bowman-Birk family) and (iii) MCoTI-I / II (both treated as members of the Squash TI family).

1.2. Cyclotides:

Cyclotides were originally discovered in extracts of the medicinal plant *Oldenlandia affinis* (Rubiaceae), known as “kalata-kalata” by the indigenous healers in Africa [42] where the plant was used as uterotonic agent (causing uterine contractions) to accelerate child birth [43]. In the early 1970’s studies uncovered the partial sequence of this peptide, identified as “Kalata B1” [44,45]. Since then, for several

years many related peptides were identified [46-50]; part of them, characterized as “macrocyclic peptides”, today known as cyclotides or cyclo-peptides [51].

The cyclotides are one of the largest families of circular proteins. They present 28 to 37 structurally cyclic amino acids (head-tail type), containing six conserved cysteine residues (numbered from I to VI) and linked in a "node" topology by three disulfide bonds, forming six loops [39,52]. This configuration occurs due to disulfide bonds between cysteine bridges at established positions, i.e. CysI-CysIV and CysII-CysV, whereas the third bridge (CysIII-CysVI) penetrates the ring, promoting the cyclization of the molecule [53]. The shared structural motif between cyclotides regards Cyclic Cystine Knot (CCK), whereas it is evident that the six loops represent topologically key regions for its biological activity (Fig. 1) [51].

Cyclotides are divided into two subfamilies: Möbius and Bracelet. The subfamily Möbius (assigned in honor to the mathematician August Ferdinand Möbius) is characterized by a cis-proline peptide binding that occurs in the fifth loop, giving a twist in its tertiary structure. In turn, the members of the Bracelet subfamily do not possess this ligation (due to lack of proline in the corresponding position in their primary sequence), and therefore have no twist, giving them the aspect of a "strap", hence the name Bracelet. However, some conserved residues are common to both subfamilies. The most conserved residue is glutamic acid (E) in the loop 1. Loops 1 and 4 show higher residue conservation, exhibiting a consensus (G/A)E(T/S) in the first loop, whereas for loop 4 a single residue is found for all cyclotide sequences, corresponding to a threonine (T), a Serine (S) or a Lysine (K) (Fig. 2). Most conserved regions regard loops 1, 4 and 6, based on the sequences of Kalata B1 and cycloviolacin O1 that represent the families Möbius and Bracelet, respectively. For the loop 6, its conservation is probably related to the ligation or cleavage processing. For example, N and C-terminal connection points occur after the Asn/Asp conserved residue that may play a role in cleavage of the C-terminal tail of the precursor protein, or in the ligation of the extremities or, still, in both [54].

1.3 *Momordica cochinchinensis* trypsin inhibitors I and II (MCoTI-I/II)

Other class of knottin (which involves a cystine-knot arrangement of their three conserved disulfide bonds) regards the squash trypsin inhibitor, isolated from squash (*Cucurbita maxima*) seeds [55]. The unique members of this squash family that present a cyclic cystine knot (CCK) are the MCoTI-I and MCoTI-II (*M. cochinchinensis* Trypsin Inhibitors I and II) [56-59]. These peptides were isolated from seeds of both mentioned Curcubitaceae members, which are commonly used in the Chinese traditional medicine against pain and inflammatory processes [60]. Proteins belonging to this family contain 28 to 32 aa residues with six conserved cysteine residues and molecular weight of approximately 3.0 kDa [61]. Although referred to as a member of the cyclotide subfamily due to the presence of a CCK motif [57-63] MCoTI exhibited low sequence homology when compared to other cyclotides and, therefore, was referred by other authors as member of squash trypsin inhibitor group [56,61,64], whose reference will also be adopted here. The TI squash family is characterized by containing an inhibitor cystine knot (ICK) [65,66] composed by small triple-stranded chain of antiparallel β -sheets and a half-turn of a 3_{10} helix, two β turns

and the inhibitory loop [61], formed by three disulfide bonds that participate in the stabilization of the molecule. This structural motif comprises a ring formed by disulfide bonds I-IV, II-V and by the third disulfide bond III-VI in the same way as reported by Craik [53] for cyclotides.

MCoTI-I and II present similar sequences to other members of linear squash TI family with the exception of an additional short linker, consisting predominantly of Ser and Gly residues which complete the structure of these CPs (Fig. 3). The presence of a conserved Gly after the last Cys residue is common to all representatives of the squash TI family (Fig. 3) and appears to be the processing point of the C-terminus of the molecule. The similarity to CysIV with CysV in Cyclotides and squash TI seems to allow closer positioning of the N and C-termini before cyclization [67].

1.4. Sunflower Trypsin Inhibitor (SFTI-1)

Another small bioactive plant CP was discovered and characterized in the mid-1990s, belonging to the proteinase inhibitor family BBI (Bowman-Birk Inhibitor) [68,69]. A peptide of this category, the Sunflower Trypsin Inhibitor (SFTI-1) was firstly isolated from sunflower (*Helianthus annuus*, Asteraceae family) seeds [70,71], comprising 14-residues with a single disulfide bridge (Fig. 4) that displays the strongest trypsin inhibitory activity among all Bowman-Birk inhibitors [70,71]. It has been considered that the disulfide bonds and cyclic backbone of SFTI-1 contribute to the activity and stability of this peptide. When the disulfide bond is removed, the trypsin inhibitory activity is considerably reduced [72]. The same occurred when the cyclic structure was opened (between Asp14 and Gly 1) [73] (Fig. 4).

1.5. Biosynthesis and cyclization

Plant CPs are encoded and cleaved from a precursor protein prior to cyclization. Its genetic origin, as ribosomal synthesized protein, was firstly reported by Jennings et al. [74], with the identification of genes encoding Kalata (B1, B2, B6 and B7) cyclotides from *Oldenlandia affinis* (Rubiaceae), whose precursor protein usually exhibit a basic and highly conserved organization. As illustrated in Fig. 5A-D, it includes: (1) a signal addressed to the endoplasmic reticulum (ER), of around 25 aa residues; (2) in the pro-region a N-terminal precursor domain (NTPD) of variable length and with no function currently assigned in regard to maturation of the CP; (3) a small linker region; (4) a mature cyclotide domain; (5) a N-terminal repeat (NTR) region and (6) a C-terminal pro-peptide tail [75,76]. An exception is the cliotide from *Clitoria ternatea* (Fabaceae), which exhibit an ER signal peptide immediately followed by the cliotide domain (cT7), flanked at the C-terminus by a peptide linker and an albumin a-chain (Ala) (Fig. 5B) [76]. Also a typical albumin 1 from *Pisum sativum* (pa1) (a plant of the Fabaceae family) presents domains A1a and A1b (Fig. 5A). The last domain (A1b) encodes a peptide of 37 residues, formed by three disulfide bonds, making them homologous to the cyclotides.

Homology of cliotide genes for both A1 genes (Fabaceae) and other cyclotide genes from Rubiaceae and Violaceae (Fig. 5C and Fig. 5D, respectively) reveal its chimeric nature. The absence of NTR region in the structure of cliotides suggests that NTR is essential for their biosynthesis [76]. Although a wide range of studies on the biosynthesis of these peptides have been carried out [33,40,74,

77-81] a lot of aspects remain to be discovered. A common feature is that they all appear to be derived from a precursor protein and are post-translationally processed to produce a cyclic backbone. In general, the mechanism involves a wide variety of amino acids and proteases (from the cysteine or serine class) which catalyze the transpeptidation reactions. In a recent work Craik and Malik [82] clearly described some of the main aspects that underlie the biosynthesis and cyclization of cyclotides. Processing occurs from the transcription of DNA into RNA and translation to prepropeptides.

Steps regarding prepropeptides cyclization (Fig. 6) include: (Fig 6-I) protein is assigned to the ER when the signal sequence is removed and the folding cyclotide precursor is generated possibly aided by protein disulfide isomerase (PDI); (Fig. 6-II) shortly after the propeptide is directed to the vacuole where the cyclotide domain is excised and cyclization is presumably promoted by NTR (N-terminal repeat) or NTPP (N-terminal propeptide); (Fig. 6-III) cleavage occurs at the N-terminus aided by a still unknown enzyme that exposes the amino acid Gly (or Ala); (Fig. 6-IV) transpeptidation occurs aiming to release the CTR (C-terminal repeat) or CTPP (C-terminal propeptide) and form the CP. This cyclizing is promoted by asparaginyl endopeptidase (AEP) which is specific for asparagine (N) or aspartic acid (D) in the S1 site, followed by a short side chain amino acid (glycine, serine or alanine) and followed by a leucine (L) and a proline (P). The result is shown in Fig. 6-IV in regard to Kalata B1, presenting in S1', S2' and S3' the following amino acids: glycine (G), leucine (L) and proline (P). After recognition of the sites and their respective amino acids, CTR is cleaved with the formation of an acyl-enzyme intermediate through the AEP attack in the thiol active site of the carbonyl group of the peptide bond. The cleaved CTR leaves the AEP active site, being replaced by a sterical N-terminal region of cyclotide domain whereas the primary amine is a glycine (G) that acts as a nucleophile and attacks the intermediate acyl-enzyme. Finally the enzyme is regenerated and the already cyclized product leaves the active site. The importance of the cyclization model should be highlighted for both the N-terminal tripeptide motif as well as for the C-terminal region flanking the cyclotide domain.

1.6. Mechanisms of action

A range of studies have confirmed a clear correlation between membrane binding and biological activity of AMPs. In general, this interaction may result in membrane structure alterations, either by forming a discrete pore or by disrupting the bilayer structure, promoting the reorientation of the peptide in the membrane, which may cross the membrane and diffuse into the cytoplasm to reach intracellular targets [83-88].

Regarding the association of AMPs to membranes, several mechanisms of action have been proposed; some of them are represented in Fig. 7, based on membrane interaction and also pore formation. The most known model regards the "barrel model" (Fig. 7A), that correspond to the insertion of peptides parallel to the membrane, forming a pore. Another one is represented by the "carpet model" (Fig. 7B), where insertion portions occurs in the membrane surface, leading to pore formation; or, still, in the "toroidal-pore model" (Fig. 7C) where peptide helices penetrate the membrane and induce peptide monolayers to bend through the pore, but in this case interspersed with lipids. In addition to these examples, other models have been proposed, such as "disordered-toroidal-pore model" (Fig. 7D), which

occurs with a lower concentration of peptides (compared to the toroidal pore model), arranged in a disorderly manner, forming a pore in the cell membrane; the "membrane thinning model" (Fig. 7E), where the attraction between the charged phospholipids generate thinner regions and more fragile membranes; the "anions attraction model" (Fig. 7F) that suggest that AMPs may be able to attract small anions making them cross the membrane so as to cause loss of cell electrostatic potential; while in the "non-lytic depolarization model" (Fig. 7G) the loss of the electrostatic potential can occur without the attraction of ions. Finally, in the "electroporation model" (Fig. 7H) peptides affect membrane potential in order to reduce their permeability to various molecules [89].

Although most AMPs interact directly with cell membrane lipids forming transmembrane pores, currently, an increasing amount of studies bring evidence that membrane permeabilization alone appears insufficient to cause cell death and therefore other complementary processes may be involved. More recently it was proposed that AMP driven microbial death can be caused by other mechanisms in addition to membrane disruption, followed by cell lysis. Indeed, evidences indicated that some AMPs can interact with intracellular targets inducing cell damages, by interfering on bacterial or fungal biosynthesis and binding abilities, also making important lipid precursors of peptidoglycan synthesis unavailable. Suppression of protein synthesis has been also proposed by inhibiting nucleic acid synthesis as well as transcription and, therefore, affecting the translation process [90,91].

The relationship between the cyclotide structures and their actions with the membrane is increasingly evident [92-98], as summarized in Table 1. According to Wang et al. [92] cyclotides (as monomers or oligomers) bind to the membrane surface and insert deeply into lipid bilayers forming a "barrel-stave" or "toroidal" pore, resulting in membrane disruption. The formation of a structural and defined pore has also been suggested by Pinto et al. [93]. In the mentioned study, a bracelet cyclotide from *Palicourea rigida* (Rubiaceae), identified as "parigidin-br1", exhibited a potent *in vivo* and *in vitro* insecticidal activity, reinforcing the role of cyclotide as potential target for bioinsecticide development. Interactions, via membrane thinning and pore-formation have been evidenced in members of the bracelet family (cycloviolacin O2) whereas in Möbius (i.e. kalata B1/B2) pore-formation occurs [94].

Simulations conducted with biophysical characterization and molecular-dynamics confirmed this relationship, suggesting that cationic charges and amphipathicity, as well as variety in the secondary structure, rigidity, and size of the AMPs are essential features for the complexity and diversity of the pore structure and activity of peptide on the integrity of target membrane [99,100]. Membrane peculiarities between the bodies (i.e. composition of different phospholipids, presence or absence of sterol, presence of fillers) also confer specificities in regard to interactions with peptides [101]. Several studies reported on selectivity of cyclotides for certain subtypes of biological membranes. Regarding Kalata B1, for example, there is considerable affinity for 'disordered' membranes, i.e. more rigid membranes rich in cholesterol and sphingomyelin, being able to break the membranes by formation of multimeric pores [101-103].

Another group of the cyclic peptides of microbial origin (Gram-positive bacteria, phylum Firmicutes) are the bacteriocins. These proteins are synthesized in the ribosomes and act by forming pores in the target cell membrane (i.e. Carnocyclin A). Unlike plant cyclic peptides, bacteriocins present higher molecular weight (approximately 5.6 to 7.2 kDa) [67,104]. In addition, the antimicrobial peptide RTD-1

(rhesus theta defensin-1, isolated from leukocytes of rhesus monkeys) is also a cyclic peptide. Similarly to SFTI, RDT-1 is a small peptide, composed by 18 amino acid residues, whose six Cys form three disulfide bonds (like cyclotides and MCoTI), but characterized by a Cyclic Cystine Ladder (CCL) motif and presenting more flexibility than plant cyclic peptides here described [67].

1.7. Evolution and occurrence among plants

According to Gruber et al. [105] the most parsimonious interpretation for the observed cyclotide distribution within the plant kingdom is convergent evolution, starting from linear precursors of the cyclotide-like type with at least four independent sources within Rubiaceae. Among the plant families with more cyclotide representatives Rubiaceae, and Violaceae stand out, whereas the lack of direct phylogenetic relationships between these groups may indicate that these peptides evolved independently in far related families.

According to Gruber et al. [105] differences in structure of precursor cyclotide genes, including the presence of introns in some representatives remove the possibility of lateral transfer, as shown, for example, in fig. 5C where a representative of Rubiaceae family (Oak1) presents one intron (black triangle) while a member of the Violaceae family (Voc1) presents no intron. Such differences between families of plants evidence the lack of a common ancestor, and therefore turn the possibility of divergent evolution to be unlikely. On the other hand, the discovery of linear sequences in monocots (Poaceae) with similarity to cyclotides [35,106] have strengthened the hypothesis that an Asn/Asp residue would have been introduced by mutation (Fig. 2, arrow) in a fundamental region for the cyclization of the linear ancestor near the C-terminus ancestor. It has also been proposed that the cyclization mechanism successfully evolved in molecules related to SFTI (Sunflower Trypsin Inhibitor) of the Cucurbitaceae family [56].

Additionally, it has been pointed out that the processing of Squash TI family seems to be variable [107]. As shown in Fig. 3, members of the same species, often exhibit almost complete sequence identity with the exception of the addition of an N-terminal segment, which exhibit many glutamine residues (Glu). If the existence of two TI peptides is a result of two distinct genes, or the result of post-translational modifications is yet unknown, although in the peptides of *Cucurbita maxima* (CMTI) and *Momordica charantia* (MCTI) enough replacements can be observed beyond the N-terminus, suggesting individual genes. Interestingly in MCoTI-III, the presence of a precursor sequence in cyclic versions was considered as a consequence of mutations in the N-terminal region of the precursor protein in the short version of TI resulting in cyclization [67]. It has been also proposed that the cyclization process in *M. cochinchinensis* has been acquired to confer resistance to proteolytic activity and/or to increase the peptide stability [57].

However, it should be emphasized that the linear squash inhibitors present a very stable structure, as in the case of EETI-I that bears a melting temperature of 140°C [108] and, therefore, the conferred advantage would be increased resistance to exopeptidases. Since there is a linear homolog of *M. cochinchinensis* peptide and if other Cucurbitaceae species do not contain cyclic trypsin inhibitors, so it is

very likely that cyclization is a relatively recent event, possibly caused by mutations in a precursor sequence [67].

1.8. Databanks and prevalence of CPs in plants

Cyclic proteins are widely distributed among living organisms, from prokaryotes to eukaryotes (including fungi, animals and plants) [40]. Cybase - the Database of Cyclic Proteins (<http://www.cybase.org.au/index.php>; Wang et al. [67] - is the most comprehensive databank that includes more than 800 entries from 103 different species.

Using the CyPred (<http://biomine.ece.ualberta.ca/CyPred/.method>) method, Kedarisetti et al. [109] predicted cyclic proteins in 640 complete proteomes of the three domains of life (Archaea, Bacteria and Eukarya). Depending on the group, 89-98 % of proteomes presented at least one predicted cyclic protein (or 45-56 % of proteomes, if considering high confidence; score > 0.9). However, only a small fraction of proteomes had larger counts of CPs, i.e., only between 7 % (for archaea proteomes) and 16 % (for eukaryotic proteomes) comprised over 10 cyclic proteins. There were no proteomes in archaea with more than 10 CPs predicted with high confidence, while 2 % and 11 % of proteomes in bacteria and eukaryote, respectively, exhibited at least 10 CPs that were identified with high confidence. When only the Eukaria domain was analyzed, a variation in proteome size was evident between species. Considering this domain, animals showed high proportions of CPs with high confidence levels when compared to fungi. In turn, plants presented a wide range of situations, including species that have no putative CP up to some with high number of candidate molecules as in Violaceae and Rubiaceae, that exhibit a large number of cyclic proteins, reaffirming previous observations of Gruber et al. [105]

Among prevalent plant CPs, cyclotides stand out, with 532 entries, regarding 55 different taxa (Cybase: <http://www.cybase.org.au/index.php>), although it is estimated that this number may well achieve 50,000 representatives [105]. In turn, Sunflower Trypsin Inhibitor (SFTI-1) was the less represented, with only three peptides described for two species (Cybase: <http://www.cybase.org.au/index.php>), being considered the most potent BBI inhibitor homolog actually known [73]. Regarding MCoTI-I and II that belong to squash family, they have been grouped within the cyclotides, even though some structural evidences consider this classification as contradictory [110].

2. PERSPECTIVES AND POTENTIAL APPLICATIONS

Based on vast biological properties, such as bactericidal, insecticidal, antiviral, antifungal, nematocidal, molluscicidal, antitumoral and hemolytic, evaluated in different plant species within the families Rubiaceae, Violaceae, Cucurbitaceae, Asteraceae and Fabaceae (summarized in Table 2), CPs represent promising potential targets for agricultural development and pharmaceutical applications [103, 111-114].

Natural and linear peptides are considered promising resources for pharmaceuticals purposes, presenting often perceptible effects and high specificity to physiological targets. Despite of that, some of them have been reported to present some disadvantages with respect to low bioavailability and *in vivo* stability. In addition, peptides have been regarded as expensive to produce compared to drugs with small

molecules. In turn, CPs present advantages over their acyclic counterparts. For example, they show resistance to exopeptidases. The cyclization also allows the blocking of conformation of another peptide, reducing entropic losses, thus resulting in a more efficient binding interaction, besides being small, what reduces production costs [33,114].

Among naturally occurring CPs, there are potent trypsin inhibitors such as the Bowman-Birk inhibitor SFTI-1 and squash family of cystine-knotted peptides [115]. A member of squash family isolated from the bitter gourd *M. cochinchinensis* (MCoTI) and the squirting cucumber *Ecballium elaterium* (EETI) act as molecular scaffolds that are important to design drugs with better stability for uses as, for example, agents for non-invasive molecular imaging of tumors in living subjects [116,117]. MCoTI are interesting from a pharmaceutical perspective because of their ability to penetrate cells and interact with intracellular targets [118]. Studies carried out by Greenwood et al. [119] showed that MCoTI-II is capable to enter human macrophages and are non-hemolytic and non-toxic to human cells. In addition, SFTI-1 presented potential applications in the treatment of prostate cancer because it selectively inhibits human KLK4 (kallikrein-related peptidase 4) [120].

Considering CP toxicity with the RTD-1 peptide indicated low cytotoxic and hemolytic activity in human red cells with concentrations up to 100 µg/mL [121]. Many approaches have tested CP activity against tumor line cells, as in the case of the study by Park et al. [122], indicating that most cyclotides present high cytotoxic activity against GTB-line U937 lymphoma cells, while Huang et al. [123] using a modified MCoTI-II cyclic peptide identified characteristics required as a model for graft desired bioactivities and identified no cytotoxicity against HeLa cells at concentrations up to 64 µM.

It has been proposed that the natural function of cyclotides is associated to protection of the host plants from pathogens or pests, especially insects [103]. Some studies have been carried out *in vivo* or *in vitro*, to determine the effects of cyclotides on insect survival and development. The cyclotides kalata B1 and B2, isolated from *Oldenlandia affinis* (Rubiaceae) presented effects on the development of Lepidoptera (*Helicoverpa armigera*), causing larval growth delay and consequently reducing insect population [74,124]. The insecticidal activity of cyclotides appears to occur by depletion of cell membranes, as judged from recent microscopic examination of the guts of *H. armigera* larvae after ingestion of kalata B1 in artificial diets [125]. Similarly, the bracelet cyclotide, cycloviolacin O2 caused potent membrane disruption and showed correlation between membrane interaction and biological activity [94]. In a subsequent study, a cyclotide from *Palicourea rigida* (Rubiaceae), named “parigidin-br1”, showed potent insecticidal activity against neonate larvae of *Diatraea saccharalis*, causing 60% mortality. The effect of parigidin-br1 was observed on insect cell lines from *Spodoptera frugiperda* (SF-9) and was supported by *in vivo* trials of insecticidal activity [93]. Since cyclotides exhibit potential insecticidal activity, testing against a range of insect species, including agricultural pests, are needed as well as the generation of transgenic plants expressing cyclotides [126].

The antimicrobial activity of cyclotides has been reported by different groups. For kalata B1 conflicting results have been reported. For example, Tam et al. [127] reported that kalata B1 was active against *Staphylococcus aureus* and inactive against *Escherichia coli*, whereas Gran et al. [128] showed that this peptide had the reverse effect over strains of both mentioned bacteria. Similarly, the kalata

peptides KB1, KB2, KB5-9, cycloviolacin O2 and tricyclon A (tcA) presented no activity against *E. coli* and *S. aureus*, except for cycloviolacin O2 that was active against *E. coli*. Interestingly, cyclotides CT1 and CT4 from *Clitoria ternatea* (Fabaceae) showed antimicrobial activity against strains of *E. coli*, *Klebsiella pneumonia* and *Pseudomonas aeruginosa* [76].

Some authors suggested that interaction between cyclotides and microbial cell membrane is salt dependent, and that the lack of antimicrobial activity in some trials may be related to the physiological relevant salt conditions, suggesting electrostatic interactions [101,127]. No activity was observed against any microorganisms with 100 μ M of KB1 in physiologically relevant saline conditions (150 mM NaCl). The lack of antimicrobial activity under physiological salt conditions was consistent with previous studies. In further studies with Kalata B1 when the level of salt (100 mM NaCl) was lower, an action against Gram-positive bacteria was observed [101]. Further studies are necessary to investigate the mode of action of cyclotides given the growing occurrence of antibiotic-resistant microorganisms in human medicine. Moreover, there is a critical need to understand the effect of cyclotides in bacterial plant pathogens [129]. Cyclotides with cytotoxic and antitumoral activities have also been reported, especially from *Viola* species and in *C. ternatea*. Three cyclotides named vitri A, varv A and varv E were isolated from *Viola tricolor* and they all showed cytotoxicity and applicability in cancer therapy [130]. Another CP, the Cycloviolacin O2 isolated from *V. odorata* has been a particular focus of these studies due to its cytotoxic activity against a variety of human tumor cell lines, tested with different primary cultures of human tumor cells *in vitro* relative to normal cells, including myeloma, leukemia, lymphoma and renal adenocarcinoma [131,132]. Additionally, five cyclotides isolated from *C. ternatea* (named CT2, CT4, CT7, CT10 and CT12) showed significant cytotoxic activity against human lung cancer cells [133].

A raw extract of *V. tricolor* was fractionated, guided by the *fluorometric microculture* cytotoxicity assay (FMCA) using cancer cells [GTB U-937 (lymphoma) RPMI-8226/s (myeloma)]. After fractioning, three cyclotides (vitri A, Varv A and E) presented cytotoxicity after 72 h treatment depending on the dose. Another test on the effect of seven *Clitoria ternatea* cyclotides in human lung cancer cells of the lineage A549 (and its sub-line A549/paclitaxel) pointed out that five among them showed significant cytotoxicity [130,133].

Another area of research that is of particular interest regarding the therapeutic use of CPs is the inhibitory activity of some cyclotides against human immunodeficiency virus (HIV). Wang et al. [134] identified five new and three known cyclotides from *Viola yedoensis* using RP-HPLC (Reversed Phase-High Performance Liquid Chromatography), and observed the existence of correlation between hydrophobicity and an anti-HIV action of new described cyclotides. They proposed that this trend tracks with their ability to disrupt membranes, which was inferred from hemolytic assays on human erythrocytes. Recently, the bracelet cyclotide cycloviolacin Y5 from the same plant species was demonstrated to be the most potent anti-HIV cyclotide tested so far [134]. Moreover, cyclotides named cycloviolin and palicourein [isolated from *Leonia cymosa* (Violaceae) and *Palicourea consensata* (Rubiaceae), respectively] also displayed activity against HIV [17,135].

Circular retrocyclin peptides, synthesized from RTD-1, have antimicrobial and anti-HIV activity, being smaller than cyclotides (only 18 amino acid residues, with six cysteine and three disulfide bonds).

Due to their significant stability, cyclotides can be synthesized with a number of substituted amino acids in the sequence, allowing them to be modified or optimized for their inherent activity or, still, to be used as a scaffold for biologically active epitopes. Similar synthetic approaches have been used to create retrocyclins congeners, increasing their antiviral or antimicrobial activity or to select other desirable attributes such as the development of therapeutic or preventive agents [136], thereby allowing cytotoxic effects are controlled. As previously reported, some cyclotides have toxic effects [74,124,125,137-141], although none of these activities are directed to mammals, being therefore not a concern in the development of human therapeutical cyclotides, although the existence of any degree of toxicity should be taken into account. For example, in accordance to Henriques and Craik [142] the toxic effects of Kalata B1 are not an impediment to the future pharmaceutical development synthetically modified cyclotides using bioengineering.

CONCLUDING REMARKS

The importance of CPs for plant survival is evident and their use has been considered for the production of transgenic plants, more resistant to microorganism attack or as natural insecticides, potentially benefiting farmers and considerably reducing losses and costs. In addition, the high biodiversity, small size, cyclic structure, stability and vast bioactivity revealed by cyclotides have triggered a great interest in the pharmaceutical area, as precious source in the production of new drugs.

Although many answers still need to be answered, as those addressed to the evolution of these peptides in plants, surveys are increasingly committed to unravel the evolutionary paths taken by these peptides, especially considering the large scale generation of genomic, transcriptomic and proteomic data, and their use in biotechnological approaches .

CONFLICT OF INTEREST:

The authors declare that they have no conflict of interest.

ACKNOWLEDGEMENTS

The authors thank Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support and fellowships.

ABBREVIATIONS

A1b = Albumin 1 chain b

aa = amino acids

AEP = Asparaginyl Endopeptidase

AMPs = Antimicrobial Peptides

BBI = Bowman-Birk Inhibitor

CCK = Cyclic Cystine Knot
 CCL = Cyclic cystine ladder
 CMTI = *Cucurbita maxima* Trypsin Inhibitor
 CPs = Cyclic Peptides
 cT = Clotide
 CTPP = C-terminal propeptide
 CTR = C-terminal repeat
 EETI = *Ecballium elaterium* Trypsin Inhibitor
 ER = Endoplasmic Reticulum
 FMCA = Fluorometric Microculture Cytotoxicity Assay
 HR = Hypersensitive Response
 ICK = Inhibitor Cystine Knot
 kDa = kilodalton
 KLK4 = kallikrein-related peptidase 4
 LTPs = Lipid Transfer Proteins
 MCoTI-I = *Momordica cochinchinensis* Trypsin Inhibitor I
 MCoTI-II = *Momordica cochinchinensis* Trypsin Inhibitor II
 MCTI = *Momordica charantia* Trypsin Inhibitor
 NTPD = N-terminal precursor domain
 NTPP = N-terminal propeptide
 NTR = N-terminal repeat
 pa1 = Albumin 1 from *Pisum sativum*
 PDI = Protein Disulfide Isomerase
 PR = Pathogenesis Related
 PR-proteins = Pathogenesis-Related proteins
 RTD-1 = *Rhesus Theta* Defensin-1
 SAR = Systemic Acquired Resistance
 SF-9 = *Spodoptera frugiperda*-9
 SFTI = Sunflower Trypsin Inhibitor
 TI = Trypsin Inhibitor

REFERENCES

- [1] Prithiviraj, B.; Paschke, M.W.; Vivano, J.M. *Root Communication: The Role of Root Exudates*. Encyclopedia of Plant and Crop Science, **2006**, *1*, 1-4.
- [2] Benko-Iseppon, A.M.; Galdino, S.L.; Calsa, T. JR; kido, E.A.; Tossi, A.; Belarmino, L.C.; Crovella, S. Overview on plant antimicrobial peptides. *Curr Protein Pept Sci.*, **2010**, *11*, 181-188.
- [3] Ponce De León, I.; Montesano, M. Activation of Defense Mechanisms against Pathogens in Mosses and Flowering Plants. *Int J Mol Sci.*, **2013**, *14*(2), 3178–3200.
- [4] Heath, M.C. Hypersensitive response-related death. *Plant Mol Biol.*, **2000**, *44*, 321-334.

- [5] Agrios, G.N. *Plant Pathology*, 5th ed.; Elsevier Academic Press: San Diego, California, **2004**.
- [6] Kachroo, A.; Robin, G.P. Systemic signaling during plant defense. *Curr. Opin. Plant Biol.*, **2013**, *16*, 527-533.
- [7] Dempsey, D.A.; Klessig, D.F. SOS – too many signals for systemic acquired resistance? *Trends Plant Sci.*, **2012**, *17*(9), 538-545.
- [8] Broekaert, W. F.; Cammue, B. P. A.; De Bolle, M. F. C.; Thevissen, K.; De Samblanx, G. W.; Osborn, R. W.; Nielson, K. Antimicrobial peptides from plants. *CRC Crit Rev Plant Sci.*, **1997**, *16*(3), 297-323.
- [9] Jenssen, H.; Hamill, P. E.; Hancock, R. E. W. Peptide antimicrobial agents. *Clin. Microbiol. Rev.*, **2006**, *19*(3), 491-511.
- [10] Wang, G.; Li, X.; Wang, Z. APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.*, **2009**, *37*(4), 933-937.
- [11] Brogden, K.A.; Ackermann, M.; Mccray, P.B.; Tack, B.F. Antimicrobial peptides in animals and their role in host defences. *Int. J. Antimicrob. Agents.*, **2003**, *22*, 465-478.
- [12] Melo, F.R.; Rigden, D.J.; Franco, O.L.; Mello, L.V.; Ary, M.B.; De Sá, M.F.G.; Bloch, C.Jr. Inhibition of trypsin by cowpea thionin: characterization, molecular modeling, and docking. *Proteins*, **2002**, *48*, 311-319.
- [13] Liu, K.; Jiang, H.; Moore, S.L.; Watkins, C.B.; Jahn, M.M. Isolation and characterization of a lipid transfer protein expressed in ripening fruit of *Capsicum chinense*. *Planta*, **2006**, *223*, 672-683.
- [14] Méndez, E.; Rocher, A.; Calero, M.; Girbés, T.; Citores, L.; Soriano, F. Primary structure of x-hordothionin, a member of a novel family of thionins from barley endosperm, and its inhibition of protein synthesis in eukaryotic and prokaryotic cell-free systems. *Eur. J. Biochem.*, **1996**, *239*, 67-73.
- [15] Chen, J.J.; Chen, G.H.; Hsu, H.C.; Li, S.S.; Chen, C.S. Cloning and functional expression of a mungbean defensin VrD1 in *Pichia pastoris*. *J. Agric. Food Chem.*, **2004**, *52*, 2256-2261.
- [16] Ngai, P.H.; Ng, T.B. Phaseococcin, an antifungal protein with antiproliferative and anti-HIV-1 reverse transcriptase activities from small scarlet runner beans. *Biochem. Cell Biol.*, **2005**, *83*, 212-220.
- [17] Ireland, D.C.; Colgrave, M.L.; Craik, D.J. A novel suite of cyclotides from *Viola odorata*: sequence variation and the implications for structure, function and stability. *Biochem. J.*, **2006**, *400*, 1-12.
- [18] Wang, S.; Rao, P.; Ye, X. Isolation and biochemical characterization of a novel leguminous defense peptide with antifungal and antiproliferative potency. *Appl. Microbiol. Biotechnol.*, **2009**, *82*(1), 79-86.
- [19] Carvalho, A.O.; Gomes, V.M. Plant defensins - Prospects for the biological functions and biotechnological properties. *Peptides*, **2009**, *30*, 1007-1020.
- [20] Da Rocha Pitta, M.G.; Galdino, S.L. Development of novel therapeutic drugs in humans from plant antimicrobial peptides. *Curr. Protein Pept. Sci.*, **2010**, *11*(3), 236-247.
- [21] Hancock, R. E. W.; Scott, M. G. The role of antimicrobial peptides in animal defenses. *Proc. Natl. Acad. Sci. U.S.A.*, **2000**, *97*(16), 8856-8861.
- [22] Yeaman, M. R.; Yount, N. Y. Mechanisms of antimicrobial peptide action and resistance. *Pharmacol. Rev.*, **2003**, *55*(1), 27-55.

- [23] Ganz, T. Defensins: antimicrobial peptides of vertebrates. *C. R. Biol.*, **2004**, 327, 539- 549.
- [24] Brogden, K. A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.*, **2005**, 3(3), 238-250.
- [25] Boix, E.; Nogues, M. V. Mammalian antimicrobial proteins and peptides: overview on the RNase A superfamily members involved in innate host defence. *Mol Biosyst.*, **2007**, 3, 317-335.
- [26] Hadley, E. B.; Hancock, R. E. Strategies for the discovery and advancement of novel cationic antimicrobial peptides. *Curr Top Med Chem.*, **2010**, 10(18), 1872-1881.
- [27] Giudici, M.; Poveda, J.A.; Molina, M.L.; De La Canal, L.; Gonzalez-Ros, J.M.; Pfuller, K.; Pfuller, U.; Villalain, J. Antifungal effects and mechanism of action of viscotoxin A3. *FEBS J.*, **2006**, 273, 72-83.
- [28] Resende, J. M.; Moraes, C. M.; Prates, M. V.; Cesar, A.; Almeida, F. C. L.; Mundim, N.; Valente, A. P.; Bemquerer, M. M.; Piló-Veloso, D.; Bechinger, B. Solution NMR structures of the antimicrobial peptides phylloseptin-1,-2, and-3 and biological activity: the role of charges and hydrogen bonding interactions in stabilizing helix conformations. *Peptides*, **2008**, 29, 1633-1644.
- [29] Montesinos, E. Antimicrobial Peptides and Plant Disease Control. *FEMS Microbiol. Lett.*, **2007**, 270, 1–11.
- [30] Thorstholm, L.; Craik, D. J. Discovery and applications of naturally occurring cyclic peptides. *Drug Discov Today Technol.*, **2012**, 9, e1–e70.
- [31] Bogdanowich-Knipp, S. J.; Jois, D. S. S.; Siahaan, T. J. The effect of conformation on the solution stability of linear vs. cyclic RGD peptides. *J. Pept. Sci.*, **1999**, 53, 523-529.
- [32] Craik, D.J. Plant cyclotides: circular, knotted peptide toxins. *Toxicon*, **2001**, 39, 1809-1813.
- [33] Trabi, M.; Craik, D. J.; Circular proteins - no end in sight. *Trends Biochem. Sci.*, **2002**, 27, 132-138.
- [34] Tugyi, R.; Mezo, G.; Fellingner, E.; Andreu, D.; Hudecs, F. The effect of cyclization on the enzymatic degradation of herpes simplex virus glycoprotein D derived epitope peptide. *J. Pept. Sci.*, **2005**, 11, 642-649.
- [35] Mulvena, J. R.; Mylne, J. S.; Bharathi, R.; Burtonb, R.A.; Shirley, N.J.; Fincher, G.B.; Anderson, M.A.; Craik, D.J. Discovery of Cyclotide-Like Protein Sequences in Gramineous Crop Plants: Ancestral Precursors of Circular Proteins? *Plant Cell.*, **2006**, 18, 2134-2144.
- [36] Selitrennikoff, C.P. Antifungal Proteins. *Appl. Environ. Microbiol.*, **2001**, 67(7), 2883–2894.
- [37] Pelegri, P.B.; Franco, O.L. Plant γ -thionins: novel insights on the mechanism of action of a multi-functional class of defense proteins. *Int. J. Biochem. Cell Biol.*, **2005**, 37(11), 2239–2253.
- [38] Witkowska, D.; Bartys, A.; Gamian, A. Defensins and cathelicidins as natural peptide antibiotics. *Postepy Hig Med Dosw.*, **2008**, 62, 694–707.
- [39] Daly, N.L.; Rosenfren, K.J.; Craik, D.J. Discovery, structure and biological activities of cyclotides. *Adv. Drug Deliv. Rev.*, **2009**, 61(11), 918-930.
- [40] Conlan, B.; Anderson, M. Circular micro-proteins and mechanisms of cyclization. *Curr. Pharm. Des.*, **2011**, 17, 4318-28.

- [41] Condie, J.A.; Nowak, G.; Reed, D.W.; Balsevich, J.J.; Reaney, M.J.T.; Arnison, P.G.; Covello, P.S. The biosynthesis of Caryophyllaceae-like cyclic peptides in *Saponaria vaccaria* L. from DNA-encoded precursors. *Plant J.*, **2011**, 67(4), 682-690.
- [42] Gran, L. Oxitocic principles found in *Oldenlandia affinis*. *Medd Nor Farm Selsk.*, **1970**, 32, 173–180.
- [43] Sandberg, F. Etude sur les plantes médicinales et toxiques de l'Afrique équatoriale. *Cahiers de la Maboké*, **1965**, 3, 5 – 49.
- [44] Gran, L. On the effect of a polypeptide isolated from "Kalata-kalata" (*Oldenlandia affinis* DC) on the oestrogen dominated uterus. *Acta Pharmacol Toxicol*, **1973**, 33, 400-408.
- [45] Sletten, K.; Gran, L. Some molecular properties of kalatapeptide B-1. *Medd Nor Farm Selsk*, **1973**, 35, 69-82.
- [46] Schöpke, T.; Hasan Agha, M. I.; Kraft, R.; Otto, A.; Hiller, K. Hämolytisch aktive komponenten aus *Viola tricolor* L. und *Viola arvensis* Murray. *Sci Pharm.*, **1993**, 61, 145-153.
- [47] Gustafson, K. R.; Sowder, R. C. II.; Henderson, L. E.; Parsons, I. C.; Kashman, Y.; Cardellina, J. H. II.; McMahon, J. B.; Buckheit, R. W.; Pannell, L. K.; Boyd, M. R.; Circulins A and B: Novel HIV-inhibitory macrocyclic peptides from the tropical tree *Chassalia parvifolia*. *J. Am. Chem. Soc.*, **1994**, 116, 9337-9338.
- [48] Witherup, K. M.; Bogusky, M. J.; Anderson, P. S.; Ramjit, H.; Ransom, R. W.; Wood, T.; Sardana, M. Cyclopsychoptide A, a biologically active, 31- residue cyclic peptide isolated from *Psychotria longipes*. *J. Nat. Prod.*, **1994**, 57, 1619- 1625.
- [49] Claeson, P.; Goëransson, U.; Johansson, S.; Luijendijk, T.; Bohlin, L. Fractionation Protocol for the Isolation of Polypeptides from Plant Biomass. *J. Nat. Prod.*, **1998**, 61, 77-81.
- [50] Göransson, U.; Luijendijk, T.; Johansson, S.; Bohlin, L.; Claeson, P. Seven novel macrocyclic polypeptides from *Viola arvensis*. *J. Nat. Prod.*, **1999**, 62, 283–286.
- [51] Craik, D. J.; Daly, N. L.; Bond, T.; Waive, C. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.*, **1999**, 294, 1327–1336.
- [52] Craik, D.J.; Anderson, M.A.; Barry, D.G.; Clark, R.J.; Daly, N.L.; Innings, C.V.; Mulvenna, . Discovery and structures of the cyclotides: novel from plants. *Lett. Pept. Sci.*, **2002**, 8, 119-128.
- [53] Craik, D.J. Discovery and applications of the plant cyclotides. *Toxicon*, **2010**, 56(7), 1092-1102.
- [54] Rosengren, K.J.; Daly, N.L.; Plan, M.R.; Waive, C.; Craik, D.J. Twists, knots, and rings in proteins. Structural definition of the cyclotide framework. *J. Biol. Chem.*, **2003**, 278(10), 8606-8616.
- [55] Polanowski, A.; Wilusz, T.; Nienartowicz, B.; Cieslar, E.; Slominska, A.; Nowak, K. Isolation and partial amino acid sequence of the trypsin inhibitor from the seeds of *Cucurbita maxima*. *Acta Biochim. Pol.*, **1980**, 27, 371-382.
- [56] Hernandez, J.F.; Gagnon, J.; Chiche, L.; Nguyen, T.M.; Andrieu, J.P.; Heitz, A.; Hong, T.T.; Pham, T.C.; Nguyen, D.L. Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. *Biochemistry*, **2000**, 39, 5722–5730.
- [57] Felizmenio-Quimio, M. E.; Daly, N. L.; Craik, D. J. Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. *J Biol Chem.*, **2001**, 276(25), 22875–22882.

- [58] Tsoi, A.Y.K.; Ng, T.B.; Fong, W.P. Antioxidative effect of a chymotrypsin inhibitor from *Momordica cochinchinensis* (Cucurbitaceae) seeds in a primary rat hepatocyte culture. *J. Pept. Sci.*, **2005**, *11*, 665–668.
- [59] Chan, L.Y.; Wang, C.K.; Major, J.M.; Greenwood, K.P.; Lewis, R.J.; Craik, D.J.; Daly, N.L. Isolation and characterization of peptides from *Momordica cochinchinensis* seeds. *J. Nat. Prod.*, **2009**, *72*, 1453–1458.
- [60] Chen, J.; Chen, S.K.; *Spermatophyta*. In: Wu ZY, editor. Flora Yunnanica. 1st ed. Beijing: Science Press; **1995**; pp. 329–330.
- [61] Heitz, A.; Hernandez, J.F.; Gagnon, J.; Hong, T.T.; Trân, T.; Pham, C.; Nguyen, T.M.; Nguyen, D.L.; Chiche, L. Solution Structure of the Squash Trypsin Inhibitor MCoTI-II. A New Family for Cyclic Knottins. *Biochemistry*, **2001**, *40*, 7973–7983.
- [62] Chiche, L.; Heitz, A.; Gelly, J. C.; Gracy, J.; Chau, P. T. T.; Ha, P. T.; Hernandez, J. F.; Nguyen, D.L. Squash inhibitors: From structural motifs to macrocyclic knottins. *Curr. Protein Pept. Sci.*, **2004**, *5*, 341 – 349.
- [63] Thongyoo, P.; Roque-Rosell, N.; Leatherbarrow, R.J.; Tate, E.W. Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org. Biomol. Chem.*, **2008**, *6*, 1462–1470.
- [64] Cemazar, M.; Daly, N. L.; Haggblad, S.; Lo, K. P.; Yulyaningsih, E.; Craik, D. J. Knots in rings – the circular knotted protein *Momordica cochinchinensis* trypsin inhibitor-II folds via a stable two-disulfide intermediate. *J Biol Chem.*, **2006**, *281*, 8224 – 8232.
- [65] Craik, D.J.; Daly, N.L.; Waive, C. The cystine knot motif in toxins and implications for drug design. *Toxicon*, **2001**, *39*, 43–60.
- [66] Gracy, J.; Le-Nguyen, D.; Gelly, J.C.; Kaas, Q.; Heitz, A.; Chiche, L. KNOTTIN: the knottin or inhibitor cystine knot scaffold in 2007. *Nucleic Acids Res.*, **2008**, *36*, D314–D319.
- [67] Wang, C.K.L.; Kaas, Q.; Chiche, L.; Craik, D.J. CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. *Nucleic Acids Res.*, **2008**, *36*, 206–210.
- [68] Bowman, D. E. Differentiation of soybean antitryptic factors. *Proc. Soc. Exp. Biol. Med.*, **1946**, *63*, 547–550.
- [69] Birk, Y.; Gertler, A.; Khalef, S. A pure trypsin inhibitor from soya beans. *Biochem. J.*, **1963**, *87*(2), 281–284.
- [70] Luckett, S.; Garcia, R.S.; Barker, J.J.; Konarev, A.V.; Shewry, P.R.; Clarke, A.R.; Brady, R.L. High resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J. Mol. Biol.*, **1999**, *290*, 525–533.
- [71] Korsinczky, M.L.; Schirra, H.J.; Rosengren, K.J.; West, J.; Condie, B.A.; Otvos, L.; Anderson, M.A.; Craik, D.J. Solution structures by 1H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant *J. Mol. Biol.*, **2001**, *311*(3), 579–91.
- [72] Zablotna, E.; Kazmierczak, K.; Jaskiewicz, A.; Stawikowski, M.; Kupryszewski, G.; Rolka, K. Chemical synthesis and kinetic study of the smallest naturally occurring trypsin inhibitor SFTI-1 isolated from sunflower seeds and its analogues. *Biochem. Biophys. Res. Commun.*, **2002**, *292*, 855–859.

- [73] Korsinczky, M.L.J.; Clark, R.J.; Craik, D.J. Disulfide bond mutagenesis and the structure and function of the head-to-tail macrocyclic trypsin inhibitor SFTI- 1. *Biochemistry*, **2005**, *44*, 1145-1153.
- [74] Jennings, C.; West, J.; Waite, C.; Craik, C.J.; Anderson, M. Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proc. Natl. Acad. Sci. U.S.A.*, **2001**, *98*, 10614-10619.
- [75] Dutton, J. L.; Renda, R. F.; Waite, C.; Clark, R. J.; Daly, N. L.; Jennings, C. V.; Anderson, M. A.; Craik, D. J.; Conserved Structural and Sequence Elements Implicated in the Processing of Gene-encoded Circular Proteins. *J Biol Chem.*, **2004**, *279*, 46858.
- [76] Nguyen, G.K.T.; Zhang, S.; Nguyen, N.T.K.; Nguyen, P.Q.T.; Chiu, M.S.; Hardjojo, A.; Tam, J.P. Discovery and Characterization of Novel Cyclotides Originated from Chimeric Precursors Consisting of Albumin-1 Chain a and Cyclotide Domains in the Fabaceae Family. *J Biol Chem.*, **2011**, *286*(27), 24275–24287.
- [77] Saska, I.; Gillon, A.D.; Hatsugai, N.; Dietzgen, R.G.; Hara-Nishimura, I.; Anderson, M.A.; Craik, D.J. An asparaginyl endopeptidase mediates in vivo protein backbone cyclization. *J Biol Chem.*, **2007**, *282*, 29721-29728.
- [78] Gillon, A.D.; Saska, I.; Jennings, C.V.; Guarino, R.F.; Craik, D.J.; Anderson, M.A.; Biosynthesis of circular proteins in plants. *Plant J.*, **2008**, *55*, 505-515.
- [79] Qin, Q.; McCallum, E.J.; Kaas, Q.; Suda, J.; Saska, I.; Craik, D.J.; Mylne, J.S. Identification of candidates for cyclotide biosynthesis and cyclisation by expressed sequence tag analysis of *Oldenlandia affinis*. *BMC Genomics*, **2010**, *11*, 1-11.
- [80] Conlan, B.F.; Gillon, A.D.; Craik, D.J.; Anderson, M.A. Circular proteins and mechanisms of cyclization. *Biopolymers*, **2010**, *94*, 573-583
- [81] Mylne, J.S.; Colgrave, M.L.; Daly, N.L.; Chanson, A.H.; Elliott, A.G.; McCallum, E.J.; Jones, A.; Craik, D.J. Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nat. Chem. Biol.*, **2011**, *7*, 257-259
- [82] Craik, D. J.; Malik, U. Cyclotide biosynthesis. *Curr Opin Chem Biol.*, **2013**, *17*, 546–54.
- [83] Papo, N.; Shai, Y. Can we predict biological activity of antimicrobial peptides from their interactions with model membranes? *Peptides*, **2003**, *24*, 1693-1703.
- [84] Jelinek, R.; Kolesheva, S. Membrane interactions of host-defense peptides studied in model systems. *Curr. Protein Pept. Sci.*, **2005**, *6*, 103–114.
- [85] Lobo, D.; Pereira, I.B.; Fragel-Madeira, L.; Medeiros, L.N.; Cabral, L.M.; Faria, J.; Campos, R.C.; Linden, R.; Kurtenbach, E. Antifungal *Pisum sativum* defensin one interacts with *Neurospora crassa* cyclin F related to cell cycle. *Biochemistry*, **2007**, *46*(4), 987–996.
- [86] Huang, Y.; Huang, J.; Chen, Y. Alpha-Helical Cationic Antimicrobial Peptides: Relationships of Structure and Function. *Protein Cell*, **2010**, *1*, 143–152.
- [87] Cirac, A. D.; Moiset, G.; Mika, J. T.; Koçer, A.; Salvador, P.; Poolman, B.; Marrink, S.; Sengupta, D. The molecular basis for antimicrobial activity of pore-forming cyclic peptides. *Biophys. J.*, **2011**, *100*(10), 2422–31.
- [88] Fjell, C.D.; Hiss, J. A.; Hancock, R.E.W.; Schneider, G. Designing Antimicrobial Peptides: Form Follows Function. *Nat Rev Drug Discov.*, **2012**, *11*, 37–51.

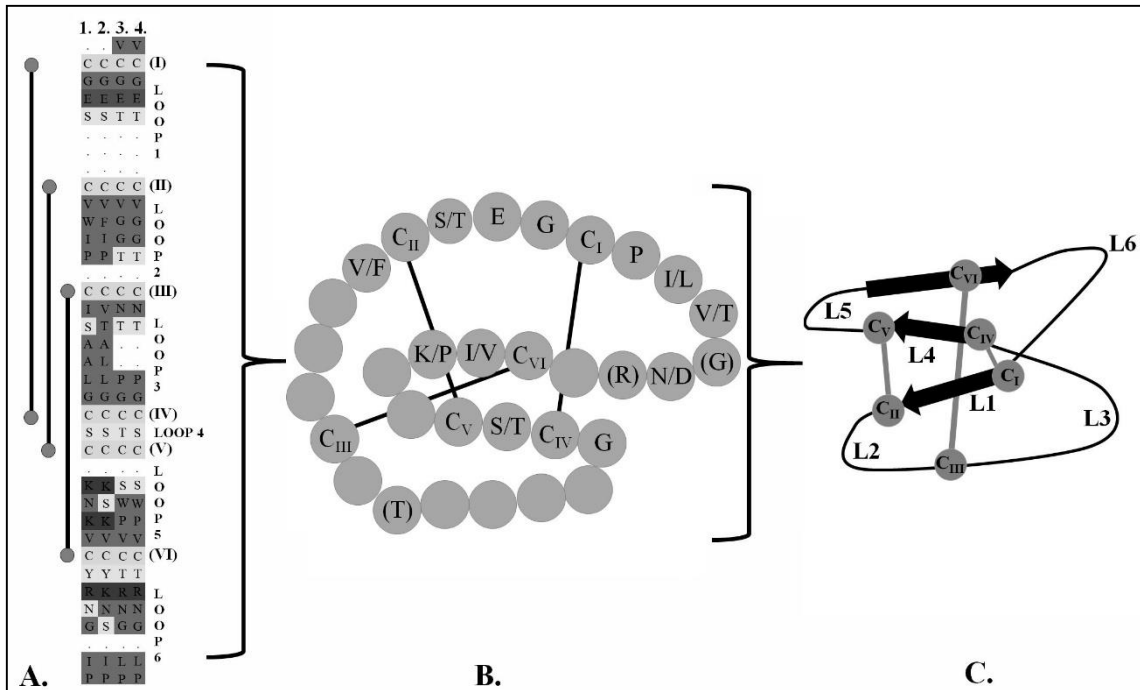
- [89] Nguyen, L.T.; Haney, E.F.; Vogel, H.J. The expanding scope of antimicrobial peptide structures and their modes of action. *Trends Biotechnol.*, **2011**, 29(9), 464-72.
- [90] Teixeira, V.; Feio, M. J.; Bastos, M. Role of lipids in the interaction of antimicrobial peptides with membranes. *Prog Lipid Res.*, **2012**, 51(2), 149–177.
- [91] Guilhelmelli, F.; Vilela, N.; Albuquerque, P.; Derengowski, L.; Silva-Pereira, I.; Kyaw, C. Antimicrobial development challenges: the various mechanisms of action of antimicrobial peptides and of bacterial resistance. *Front Microbiol.*, **2013**, 4, 1-12.
- [92] Göransson, U.; Burman, R.; Gunasekera, S.; Strömstedt, A.A.; Rosengren, K.J. Circular Proteins from Plants and Fungi. *J Biol Chem.*, **2012**, 287(32), 27001–27006.
- [93] Pinto, M.F.; Fensterseifer, I.C.; Migliolo, L.; Sousa, D.A.; De Capdville, G.; Arboleda-Valencia, J.W.; Colgravel, M.L.; Craik, D.J.; Magalhães, B.S.; Dias, S.C.; Franco, O.L. Identification and structural characterization of novel cyclotide with activity against an insect pest of sugar cane. *J Biol Chem.*, **2012**, 287, 134-47.
- [94] Burman, R.; Stromstedt, A. A.; Malmsten, M.; Göransson, U. Cyclotide-membrane interactions: defining factors of membrane binding, depletion and disruption. *Biochim. Biophys. Acta, Biomembranes* **2011**, 1808, 2665–2673.
- [95] Svängård, E.; Burman, R.; Gunasekera, S.; Lövborg, H.; Gullbo, J.; Göransson, U. Mechanism of action of cytotoxic cyclotides: cyclviolacin O2 disrupts lipid membranes. *J. Nat. Prod.*, **2007**, 70, 643–647
- [96] Shenkarev, Z.O.; Nadezhdin, K.D.; Lyukmanova, E.N.; Sobol, V.A.; Skjeldal, L.; Arseniev, A.S. Divalent cation coordination and mode of membrane interaction in cyclotides: NMR spatial structure of ternary complex Kalata B7/Mn²⁺/DPC micelle. *J. Inorg. Biochem.*, **2008**, 102, 1246–1256.
- [97] Huang, Y.H.; Colgrave, M.L.; Daly, N.L.; Keleshian, A.; Martinac, B.; Craik, D.J. The biological activity of the prototypic cyclotide Kalata B1 is modulated by the formation of multimeric pores. *J Biol Chem.*, **2009**, 284, 20699–20707.
- [98] Wang, C.K.; Colgrave, M.L.; Ireland, D.C.; Kaas, Q.; Craik, D.J. Despite a conserved cysteine knot motif, different cyclotides have different membrane binding modes. *Biophys. J.*, **2009**, 97, 1471–1481.
- [99] Moiset, G.; Cirac, A. D.; Stuart, M. C. A.; Marrink, S. J.; Sengupta, D.; Poolman, B. Dual action of BPC194: a membrane active peptide killing bacterial cells. *PLoS one*, **2013**, 8(4), e61541.
- [100] Chen, L.; Li, X.; Gao, L.; Fang, W. Theoretical Insight into the Relationship Between the Structures of Antimicrobial Peptides and Their Actions on Bacterial Membranes. *J Phys Chem B.*, **2014**.
- [101] Henriques, S.T.; Huang, Y.H.; Rosengren, K.J.; Franquelim, H.G.; Carvalho, F.A.; Johnson, A.; Souza, S.; Tachedjian, G.; Castanha, M.A.R.B.; Daly, N.L.; Craik, D.J. Decoding the membrane activity of the cyclotide kalata B1: The importance of phosphatidylethanolamine phospholipids and lipid organization hemolytic and anti-HIV activities. *J Biol Chem.*, **2011**, 286, 24231–24241.
- [102] Simonsen, S.M.; Sando, L.; Rosengren, K.J.; Wang, C.K.; Colgrave, M.L.; Daly, N.L.; Craik, D.J. Alanine scanning mutagenesis of the prototypic cyclotide reveals a cluster of residues essential for bioactivity. *J Biol Chem.*, **2008**, 283, 9805–9813.
- [103] Craik, D. J. Host-defense activities of cyclotides. *Toxins*, **2012**, 4(2), 139–156.

- [104] Gabrielsen, C.; Brede, D.A.; Nes, I.F.; Diep, D.B. Circular bacteriocins: biosynthesis and mode of action. *Appl. Environ. Microbiol.*, **2014**, *80*(22), 6854-6862.
- [105] Gruber, C.W.; Elliott, A.G.; Ireland, D.C.; Delpreteb, P.G.; Desseinc, S.; Göransson, U.; Trabia, M.; Wanga, C.K.; Kinghorna, A.B.; Robbrecht, E.; Craik, D.J. Distribution and Evolution of Circular Miniproteins in Flowering Plants. *Plant Cell.*, **2008**, *20*, 2471-2483.
- [106] Basse, C. W. Dissecting Defense-Related and Developmental Transcriptional Responses of Maize during *Ustilago maydis* Infection and Subsequent Tumor Formation I. *Society*, **2005**, *138*, 1774-1784.
- [107] Wieczorek, M.; Otlewski, J.; Cook, J.; Parks, K.; Leluk, J.; Wilimowska-Pelc, A.; Polanowski, A.; Wilusz, T.; Laskowski, M. The squash family of serine proteinase inhibitors. Amino acid sequences and association equilibrium constants of inhibitors from squash, summer squash, zucchini, and cucumber seeds. *Biochem. Biophys. Res. Commun.*, **1985**, *126*, 646-652.
- [108] Heitz, A.; Le-Nguyen, D.; Chiche, L. Min-21 and min-23, the smallest peptides that fold like a cystine-stabilized β -sheet motif: design, solution structure, and thermal stability. *Biochemistry*, **1999**, *38*, 10615-10625.
- [109] Kedarisetia, P.; Miziantya, M.J.; Kaasb, Q.; Craik, D.J.; Kurgan, L. Prediction and characterization of cyclic proteins from sequences in three domains of life. *BBA-Proteins Proteom*, **2014**, *1844*(1), part B, 181-190.
- [110] Picchi, D. G.; Altei, W. F.; Saito, M. S.; da Silva Bolzani, V.; Cilli, E. M. Peptídeos cíclicos de biomassa vegetal: características, diversidade, biossíntese e atividades biológicas. *Quim Nova*, **2009**, *32*(5), 1262-1277.
- [111] Pomilio, A.B.; Battista, M.E.; Vitale, A.A. Naturally-Occurring Cyclopeptides: Structures and Bioactivity. *Curr Org Chem.*, **2006**, *10*, 2075-2121.
- [112] Brogden, N.K.; Brogden, K.A. Will new generations of modified antimicrobial peptides improve their potential as pharmaceuticals. *Int. J. Antimicrob. Agents.*, **2011**, *38*, 217-25.
- [113] Pinto, M. F. S.; Almeida, R. G.; Porto, W. F.; Fensterseifer, I. C. M.; Lima, L. A.; Dias, S. C.; Franco, O. L. Cyclotides: From Gene Structure to Promiscuous Multifunctionality. *J Evid Based Complementary Altern Med.*, **2011**, *17*(1), 40-53.
- [114] Cemazar, M.; Kwon, S.; Mahatmanto, T.; Ravipati, A. S.; Craik, D. J. Discovery and applications of disulfide-rich cyclic peptides. *Curr Top Med Chem.*, **2012**, *12*(14), 1534-1545.
- [115] Quimbar, P.; Malik, U.; Sommerhoff, C.P.; Kaas, Q.; Chan, L.Y.; Huang, Y.H.; Grundhuber, M.; Dunse, K.; Craik, D.J.; Anderson, M.A.; Daly, N.L. High-affinity Cyclic peptide matriptase inhibitors. *J Biol Chem.*, **2013**, *288*, 13885-13896.
- [116] Daly, N.L.; Craik, D.J. Bioactive cystine knot proteins. *Curr Opin Chem Biol.*, **2011**, *15*, 362-368.
- [117] Moore, S.J.; Leung, C.L.; Cochran, J.R. Knottins: Disulfide-bonded therapeutic and diagnostic peptides. *Drug Discov Today Technol.*, **2011**, *9*, 3-11.
- [118] Contreras, J.; Elnagar, A.Y.; Hamm-Alvarez, S.F.; Camarero, J.A. Cellular Uptake of cyclotide MCoTI-I follows multiple endocytic pathways. *J Control Release.*, **2011**, *155*, 134-143.
- [119] Greenwood, K.P.; Daly, N.L.; Brown, D.L.; Stow, J.L.; Craik, D.J. The Cyclic cystine knot miniprotein MCoTI-II is internalized into cells by macropino cytosis. *Int. J. Biochem. Cell Biol.*, **2007**, *39*, 2252-2264.

- [120] Stefansson, K.; Brattsand, M.; Roosterman, D.; Kempkes, C.; Bocheva, G.; Steinhoff, M.; Egelrud, T. Activation of proteinase-activated receptor-2 by human kallikrein-related peptidases. *J. Invest. Dermatol.*, **2008**, *128*, 18–25.
- [121] Lehrer, R.I.; Cole, A.M.; Selsted, M.E. θ - Defensins: Cyclic Peptides with Endless Potential. *J Biol Chem.*, **2012**, *287*(32), 27014–27019.
- [122] Park, S.; Strömstedt, A.A.; Göransson, U. Cyclotide Structure–Activity Relationships: Qualitative and Quantitative Approaches Linking Cytotoxic and Anthelmintic Activity to the Clustering of Physicochemical Forces. *PLoS ONE*, **2014**, *9*(3) e91430.
- [123] Huang, Y.; Chaousis, S.; Cheneval, O.; Craik, D.J.; Troeira_henriques, S. Optimization of the cyclotide framework to improve cell penetration properties. *Front Pharmacol.*, **2015**, *6*:17.
- [124] Jennings, C. V.; Rosengren, K. J.; Daly, N. L.; Plan, M.; Stevens, J.; Scanlon, M. J.; Waive, C.; Norman, D. G.; Anderson, M. A.; Craik, D. J. Isolation, solution structure, and insecticidal activity of kalata B2, a circular protein with a twist: do Mobius strips exist in nature? *Biochemistry*, **2005**, *44*(3), 851–60.
- [125] Barbeta, B. L.; Marshall, A. T.; Gillon, A. D.; Craik, D. J.; Anderson, M. A. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. *Proc. Natl. Acad. Sci. U.S.A.*, **2008**, *105*, 1221–1225.
- [126] Gruber C.W.; Cemazar M.; Anderson M.A.; Craik D.J. Insecticidal plant cyclotides and related cystine knot toxins. *Toxicon*, **2007**, *49*, 561–575.
- [127] Tam, J.P.; Lu, Y.A.; Yang, J.L.; Chiu, K.W. An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides. *Proc. Natl. Acad. Sci. U.S.A.*, **1999**, *96*, 8913–8918.
- [128] Gran, L.; Sletten, K.; Skjeldal, L. Cyclic Peptides from *Oldenlandia affinis* DC. Molecular and Biological Properties. *Chem. Biodivers.*, **2008**, *5*, 2014–2022.
- [129] Kaas, Q.; Westermann, J-C.; Henriques, S.T.; Craik, D.J. Antimicrobial peptides in plants. In Wang G. ed., *Antimicrobial Peptides: Discovery, Design and Novel Therapeutic Strategies*. G. Wang (ed.). Chapter 3, pp. 4071, **2010**. CABI: Oxfordshire, UK.
- [130] Svängård, E.; Göransson, U.; Hocaoglu, Z.; Gullbo, J.; Larsson, R.; Claeson, P.; Bohlin, L. Cytotoxic Cyclotides from *Viola tricolor*. *J. Nat. Prod.*, **2004**, *67*, 144–147.
- [131] Lindholm, P.; Göransson, U.; Johansson, S.; Claeson, P.; Gullbo, J.; Larsson, R.; Bohlin, L.; Backlund, A. Cyclotides: A novel type of cytotoxic agents. *Mol. Cancer Ther.*, **2002**, *1*, 365–369.
- [132] Burman, R.; Svedlund, E.; Felth, J.; Hassan, S.; Herrmann, A.; Clark, R.J.; Craik, D.J.; Bohlin, L.; Claeson, P.; Göransson, U.; Gullbo, J. Evaluation of toxicity and antitumor activity of cycloviolacin O2 in mice. *Biopolymers*. **2010**, *94*, 626–634.
- [133] Sen, Z.; Zhan, X. K.; Jing, J.; Yi, Z.; Wanqi, Z. Chemosensitizing activities of cyclotides from *Clitoria ternatea* in paclitaxel-resistant lung cancer cells. *Oncol Lett.*, **2013**, *5*, 641–644.
- [134] Wang, C. K. L.; Colgrave, M. L.; Gustafson, K. R.; Ireland, D. C.; Goransson, U.; Craik, D. J. Anti-HIV cyclotides from the Chinese medicinal herb *Viola yedoensis*. *J. Nat. Prod.*, **2007**, *71*, 47–52.
- [135] Hallock, Y.F.; Sowder II, R.C.; Pannell, L.K.; Hughes, C.B.; Johnson, D.G.; Gulakowski, R.; Cardellina II, J.H.; Boyd, M.R. Cycloviolins A-D, anti-HIV macrocyclic peptides from *Leonia cymosa*. *J. Org. Chem.*, **2000**, *65*, 124–128.

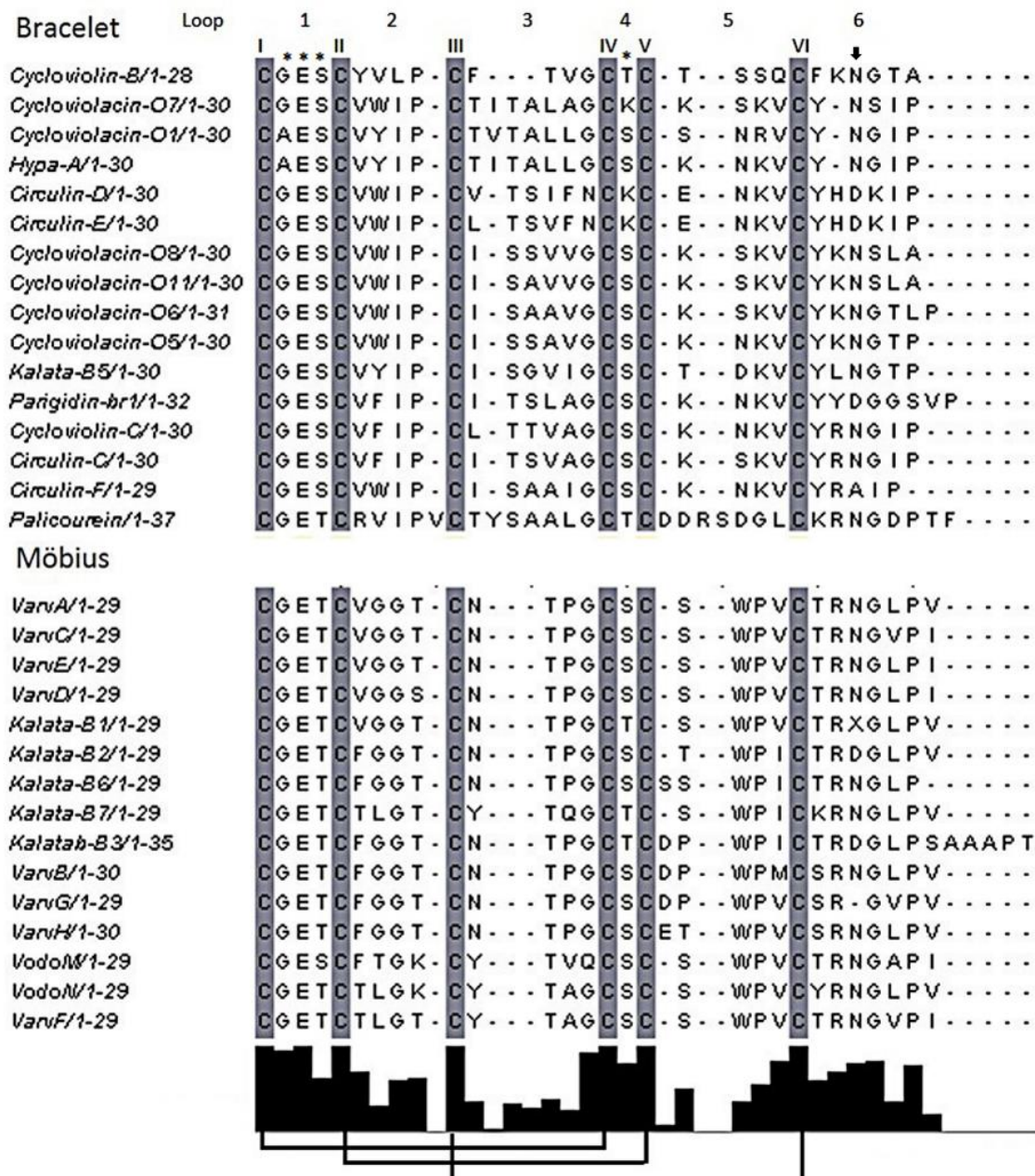
- [136] Owen, S.M.; Rudolph, D.L.; Wang, W.; Cole, A.M.; Waring, A.J.; Lal, R.B.; Lehrer, R.I. RC-101, a retrocyclin-1 analogue with enhanced activity against primary HIV type 1 isolates. *AIDS Res. Hum. Retroviruses*, **2004**, *20*, 1157–1165.
- [137] Plan, M. R.; Saska, I.; Cagauan, A.G.; Craik, D.J. Backbone cyclised peptides from plants show molluscicidal activity against the rice pest *Pomacea canaliculata* (golden apple snail). *J. Agric. Food Chem.*, **2008**, *56*, 5237–5241.
- [138] Göransson, U.; Sjögren, M.; Svangård, E.; Claeson, P.; Bohlin, L. Reversible antifouling effect of the cyclotide cycloviolacin O2 against barnacles. *J. Nat. Prod.*, **2004**, *67*, 1287–90.
- [139] Colgrave, M.L.; Kotze, A.C.; Huang, Y.H.; O'Grady, J.; Simonsen, S.M.; Craik, D.J. Cyclotides: natural, circular plant peptides that possess significant activity against gastrointestinal nematode parasites of sheep. *Biochemistry*, **2008**, *47*, 5581–5589.
- [140] Colgrave, M.L.; Kotze, A.C.; Ireland, D.C.; Wang, C.K.; Craik, D.J. The anthelmintic activity of the cyclotides: natural variants with enhanced activity. *Chembiochem*, **2008**, *9*, 1939–1945.
- [141] Colgrave, M.L.; Kotze, A.; Kopp, S.; McCarthy, J.S.; Coleman, G.T.; Craik, D.J. Anthelmintic activity of cyclotides: *In vitro* studies with canine and human hookworms. *Acta Trop.*, **2009**, *109*, 163–166.
- [142] Henriques, S.T.; Craik, D.J. Cyclotides as templates in drug design. *Drug Discovery Today*, **2010**, *15*, 57–64.
- [143] Nawae, W.; Hannongbua, S.; Ruengjitchachawalya, M. Defining the membrane disruption mechanism of kalata B1 via coarse-grained molecular dynamics simulations. *Sci Rep.*, **2014**, Article number: 3933.
- [144] Vriens, K.; Cammue, B.P.A.; Thevissen, K. Antifungal Plant Defensins: Mechanisms of Action and Production. *Molecules*, **2014**, *19*, 12280–12303.
- [145] Tang, M.; Hong, M. Structure and mechanism of b-hairpin antimicrobial peptides in lipid bilayers from solid-state NMR spectroscopy. *Mol Biosyst.*, **2009**, *5*, 317–322.
- [146] Rokitskaya, T.I.; Kolodkin, N.I.; Kotova, E.A.; Yuri, N. Antonenko Indolicidin action on membrane permeability: Carrier mechanism versus pore formation. *Biochim. Biophys. Acta*, **2011**, *1808*, 91–97.
- [147] Gifforda, J.L.; Huntera, H.N.; Vogel, H.J. Lactoferricin: a lactoferrin-derived peptide with antimicrobial, antiviral, antitumor and immunological properties. *Cell. Mol. Life Sci.*, **2005**, *62*, 2588–2598.
- [148] Chan, D.I.; Prenner, E.J.; Vogel, H.J. Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. *Biochim. Biophys. Acta*, **2006**, *1758*, 1184–1202.
- [149] Raghuraman, H.; Chattopadhyay, A. Melittin: a Membrane-active Peptide with Diverse Functions. *Biosci. Rep.*, **2007**, *27*, 189–223.
- [150] Sengupta, D.; Leontiadou, H.; Mark, A.E.; Marrink, S.J. Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochim. Biophys. Acta*, **2008**, *1778*, 2308–2317.
- [151] Kredics, L.; Szekeres, A.; Czifra, D.; Vágvölgyi, C.; Leitgeb, B. Recent Results in Alamethicin Research. *Chem. Biodivers.*, **2013**, *10*, 744–771.
- [152] Zhang, J.; Liao, B.; Craik, D.J.; Li, J.T.; Hu, M.; Shu, W.S. Identification of two suites of cyclotide precursor genes from metallophyte *Viola baoshanensis*: cDNA sequence variation, alternative RNA splicing and potential cyclotide diversity. *Gene*, **2009**, *431*, 23–32.

- [153] Gustafson, K.R.; Walton, L.K.; Sowder, R.C.; Johnson, D.G.; Pannell, L.K.; Cardellina II J.H.; Boyd, M.R. New circulin macrocyclic polypeptides from *Chassalia parvifolia* *J. Nat. Prod.*, **2000**, *63*(2), 176–178.
- [154] Ireland, D.C.; Wang, C.K.; Wilson, J.A.; Gustafson, K.R.; Craik, D.J. Cyclotides as natural anti-HIV agents. *Biopolymers*, **2008**, *90*, 51–60.
- [155] Favel, A.; Mattras, H.; Coletti-Previero, M.A.; Zwiling, R.; Robinson, E.A.; Castro, B. Protease inhibitors from *Ecballium elaterium* seeds. *Int. J. Pept. Protein Res.*, **1989**, *33*, 202–208.
- [156] Broussalis, A.M.; Göransson, U.; Coussio, J.D.; Ferraro, G.; Martino, V.; Claeson, P. First cyclotide from *Hybanthus* (Violaceae). *Phytochemistry*, **2001**, *58*, 47–51.
- [157] Pränting, M.; Loov, C.; Burman, R.; Göransson, U.; Andersson, D.I. The cyclotide cycloviolacin O2 from *Viola odorata* has potent bactericidal activity against Gram-negative bacteria. *J. Antimicrob. Chemother.*, **2010**, *65*, 1964–1971.
- [158] Ovesen, R.G.; Brandt, K.K.; Göransson, U.; Nielsen, J.; Hansen, H.C.; Cedergreen, N. Biomedicine in the environment: cyclotides constitute potent natural toxins in plants and soil bacteria. *Environ. Toxicol. Chem.*, **2011**, *30*, 1190–1196.
- [159] Plan, M.R.; Rosengren, K.J.; Sando, L.; Daly, N.L.; Craik, D.J. Structural and Biochemical Characteristics of the Cyclotide Kalata B5 from *Oldenlandia affinis*. *Biopolymers*, **2010**, *94*, 647–658.
- [160] Daly, N.L.; Clark, R.J.; Plan, M.R.; Craik, D.J. Kalata B8, a novel antiviral circular protein, exhibits conformational flexibility in the cystine knot motif. *Biochem. J.*, **2006**, *393*, 619–626.
- [161] Bokesch, H.R.; Pannell, L.K.; Cochran, P.K.; Sowder, R.C.; McKee, T.C.; Boyd, M.R. A novel anti-HIV macrocyclic peptide from *Palicourea condensata*. *J. Nat. Prod.*, **2001**, *64*, 249–250.
- [161] Yeshak, M.Y.; Burman, R.; Asres, K.; Göransson, U. Cyclotides from an extreme habitat: characterization of cyclic peptides from *Viola abyssinica* of the Ethiopian highlands. *J. Nat. Prod.*, **2011**, *74*, 727–731.
- [162] Tang, J.; Wang, C.K.; Xulin, P.; He, Y.; Guangzhi, Z.; Wenyan, X.; Hea, W.; Daly, N.L.; Craik, D.J.; Tan, N. Isolation and characterization of cytotoxic cyclotides from *Viola tricolor*. *Peptides*, **2010**, *31*, 1434–40.
- [163] Chen, B.; Colgrave, M.L.; Daly, N.L.; Rosengren, K.J.; Gustafson, K.R.; Craik, D.J. Isolation and Characterization of Novel Cyclotides from *Viola hederacea*. Solution structure and anti-HIV activity of vh-1, a leaf-specific expressed cyclotide. *J Biol Chem.*, **2005**, *280*(23), 22395–22405.
- [164] Herrmann, A.; Burman, R.; Mylne, J.S.; Karlsson, G.; Gullbo, J.; Craik, D.J.; Clark, R.J.; Göransson, U. The alpine violet, *Viola biflora*, is a rich source of cyclotides with potent cytotoxicity. *Phytochemistry*, **2008**, *69*, 939–952.
- [165] Svängård, E.; Göransson, U.; Smith, D.; Verma, C.; Backlund, A.; Bohlin, L.; Claeson, P. Primary and 3-D modeled structures of two cyclotides from *Viola adorata*. *Phytochemistry*, **2003**, *64*, 135–142.
- [166] Tang, J.; Wang, C.; Pan, X.; Yan, H.; Zeng, G.; Xu, W.; He, W.; Daly, N.; Craik, D.; Tan, N. Isolation and characterization of bioactive cyclotides from *Viola labridorica*. *Helv. Chim. Acta.* **2010**, *93*, 2287 - 2295.



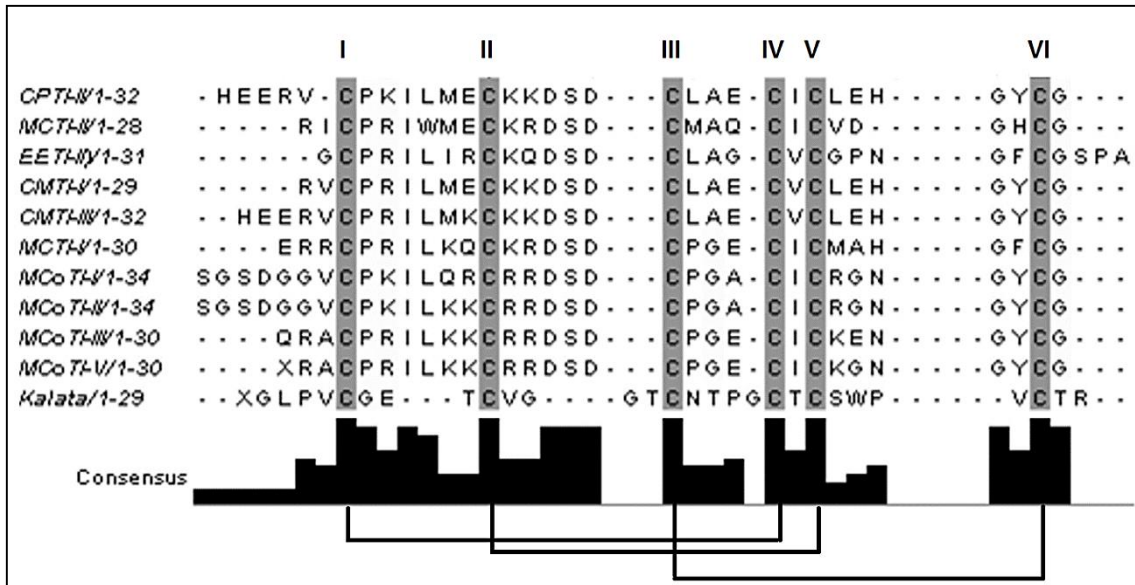
1
2
3
4
5
6
7
8
9
10
11
12
13

Fig. (1). Schematic illustration of three different cyclotide conformations. **(A)** The arabical numbers above the vertical alignment regard the primary structure of: (1) Circulin A, (2) Cyclopsychotride A, (3) Kalata B1 and (4) Varv peptide. The conserved cysteine residues are represented by Roman numerals (I to IV), forming three disulfide bonds (vertical black lines). The inter-cysteine loops are numbered from 1 to 6. Other amino acids residues are in grayscale, whereas those chemically similar are showed in same gray intensity; **(B)** Topology of the cyclic cystine knot (CCK) motif and summary of the conserved and variable residues of all known vegetables cyclotides: the disulfide bonds are represented by black lines connecting the cysteine residues (Subscript I-IV); **(C)** Secondary structure of cyclotides, showing beta sheet and disulfide bonds connecting cysteine residues (I-IV Subscript) and the inter-cysteine loops (L1-6).



14

15 **Fig. (2).** Alignment of cyclotide sequences available in GenBank (NCBI), showing conserved cysteines
 16 residues (gray), represented by Roman numerals (I to VI), forming three disulfide bonds (black lines at
 17 the bottom of the Figure). Protein loops are numbered from 1 to 6 at the top of the alignment. Black
 18 columns at the bottom indicate conservation. Black arrow indicates residue possibly introduced by
 19 mutation that lead to cyclization.



20

21 **Fig. (3).** Alignment of trypsin inhibitors (TI) of the squash family, based on publically available sequences
 22 (GenBank, NCBI) showing conserved cysteine residues (gray), represented by Roman numerals (I to VI),
 23 forming three disulfide bonds (black lines at the bottom of the figure). Legend to species abbreviations:
 24 *CPTI*: *Cucurbita pepo* trypsin inhibitor; *MCTI*: *Mormodica charantia* trypsin inhibitor; *EETI*: *Ecballium*
 25 *elaterium* trypsin inhibitors; *CMTI*: *Cucurbita maxima* trypsin inhibitors; *MCoTI*: *Momordica*
 26 *cochinchinensis* trypsin inhibitors.

27

28

29

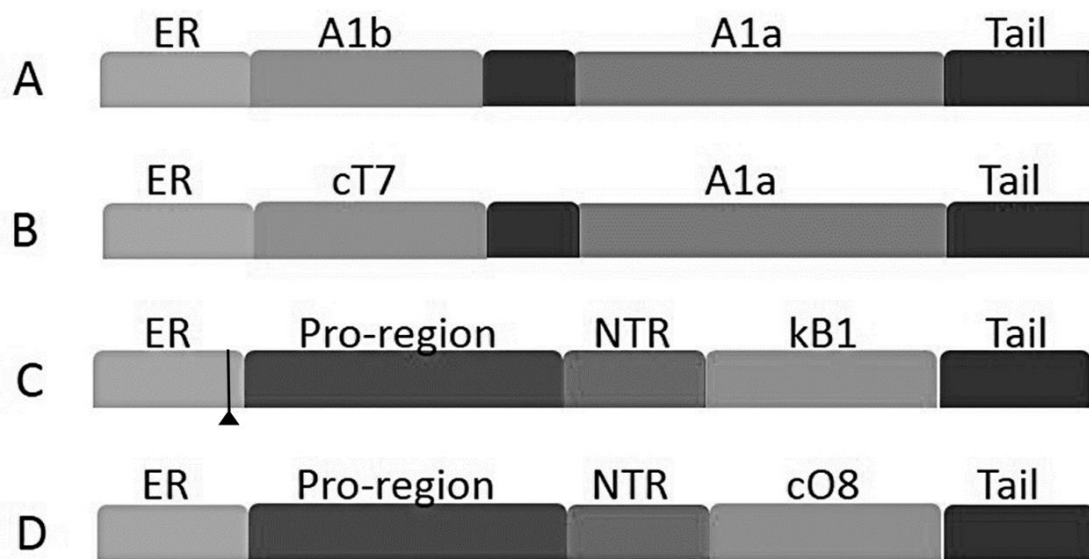


30

31 **Fig. (4)** Alignment of SFTI sequences obtained from GenBank (NCBI), showing conserved residues,
 32 cysteines (numbered I-II) and the disulfide bond (gray line at the bottom of the figure). Sequences
 33 isolated from *Helianthus annuus* (sunflower).

34

35

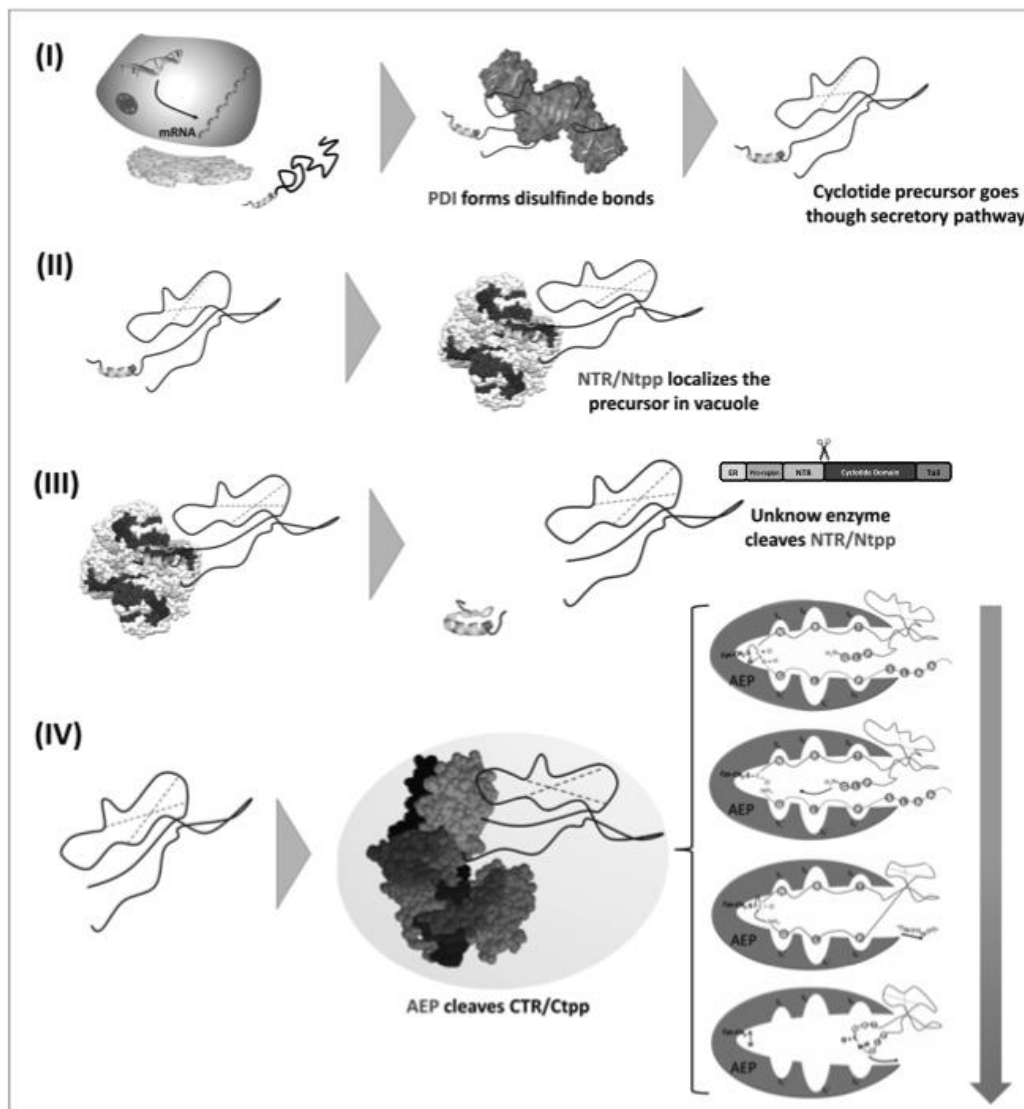


36

37 **Fig. (5).** Schematic representation of chimeric genetic arrangements of cyclotide precursors in different
 38 plant families. (A) Albumin 1 domain from *Pisum sativum*, Fabaceae (pA1a and pA1b); (B) cliotide from
 39 *Clitoria ternatea*, Fabaceae (cT7); (C) cyclotide from *Oldenlandia affinis*, Rubiaceae (Oak1), including
 40 representation of introns (black triangle) and (D) cyclotides from *Viola odorata*, Violaceae (Voc1). ER,
 41 endoplasmic reticulum signal; Pro-region (NTPD, N-terminal precursor domain; NTR, N-terminal repeat;
 42 cyclotide domain and Tail, the C-terminal pro-peptide tail. Note that A1a domain is present only in
 43 albumin and cliotides.

44

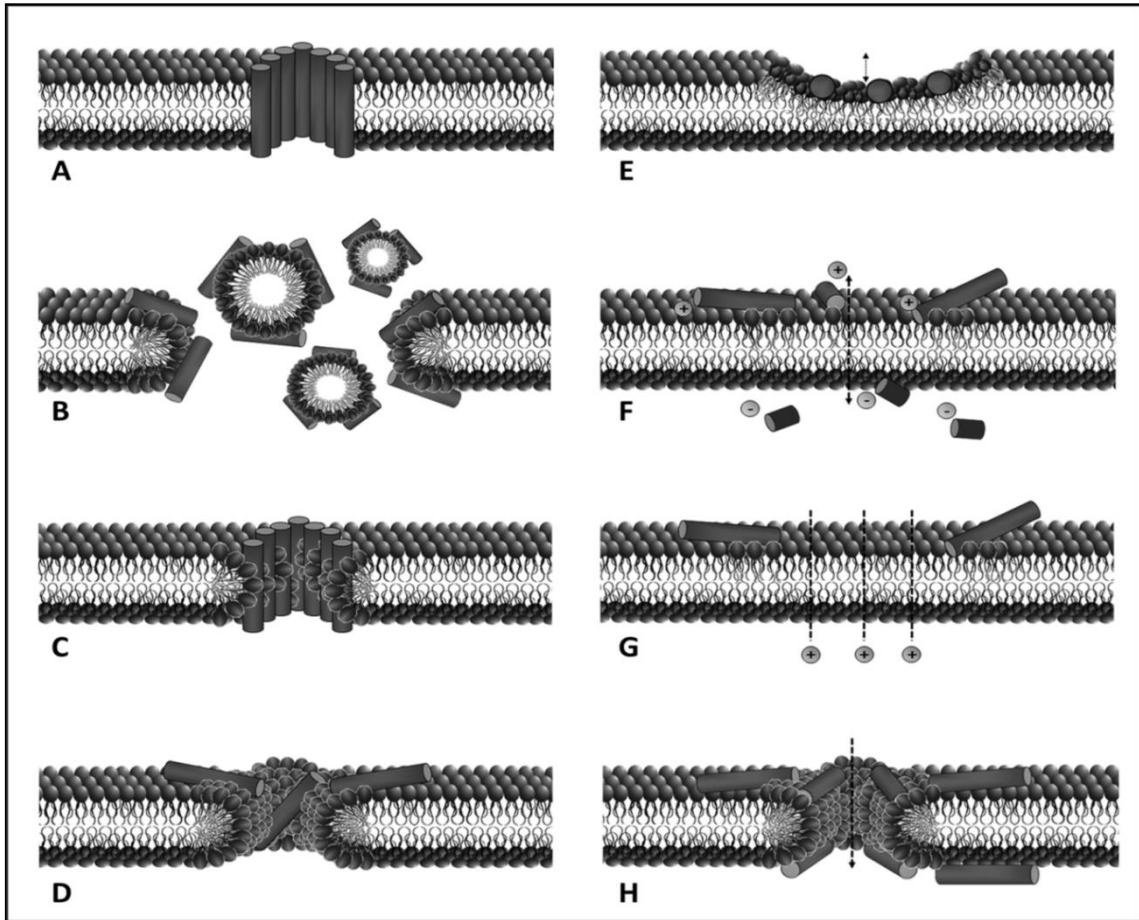
45



46
 47
 48
 49
 50
 51
 52
 53

Fig. (6). Schematic representation of cyclotide post transcriptional processing. Hypothetical steps regarding cyclotide intracellular processing (I to IV) and proposed mechanism for AEP mediation in the cyclization reaction expanded in IV. Legend for abbreviations: AEP = Asparaginyl Endopeptidase; NTPP = N-terminal propeptide; NTR = N-Terminal Repeat; PDI = Protein Disulfide Isomerase.

54



55

56

57 **Fig. (7).** General outline of the main mechanism of interaction between antimicrobials peptides and
 58 pathogen cell membranes. (A) “Barrel” model; (B) “Carpet” model; (C) “Toroidal pore” model; (D)
 59 “Toroidal disorderly pore” model; (E) “Membrane thinning” model; (F) “Anion carrier” model; (G)
 60 “Non-lytic depolarization” model; (H) Electroporation model.

61

62

63 Table 1. Selected examples of antimicrobial peptides and putative modes of membrane
64 interaction.

65

Groups	AMPs*	Mode of action	Refs
Plants	Cyclotides	Barrel-stave, toroidal pore, membrane thinning	[92,143]
	Defensins	Barrel-stave	[144]
Mammal	Protegrin I	Toroidal pore	[145]
	Indolicidin	Anion carrier	[146]
	Bovine Lactoferricin	Non-lytic membrane depolarization	[147]
	NK-lysin	Electroporation	[148]
Insect	Melittin	Disordered toroidal pore	[149,150]
	Cecropin	Detergent micellization	[24]
Fungi	Alamethicin	Barrel-stave pore	[24,151]

*The peptides listed in this table are not limited to their modes of action and related herein may employ multiple mechanisms simultaneously.

66

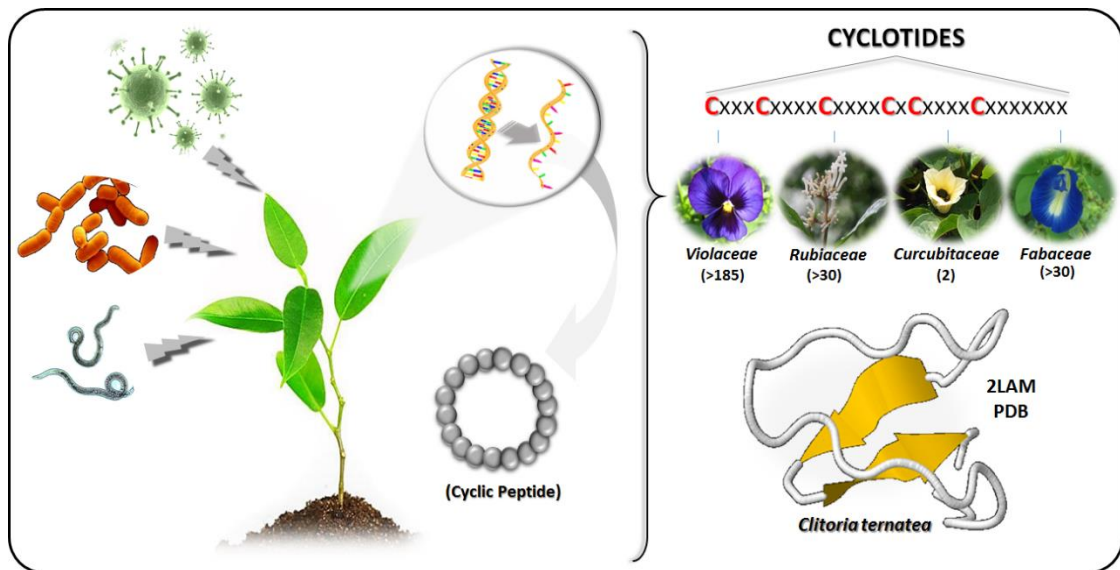
67

Table 2. Plant cyclic peptides, including their sources and known activities.

Peptide	Source species	Family	Activity	References
Cycloviolacin O12	<i>Viola abyssinica</i> <i>Viola arvensis</i> , <i>Viola baoshanensis</i> , <i>Viola tricolor</i> , <i>Viola yedoensis</i> ,	Violaceae	Antitumor, hemolytic	[152]
Circulin A/B	<i>Chassalia parvifolia</i>	Rubiaceae	Antibacterial, anti-HIV, hemolytic	[47]
Circulin C/D/E/F			Anti-HIV	[153]
CT1/CT4	<i>Clitoria ternatea</i>	Fabaceae	Antibacterial	[76]
CT2/CT4/CT7/CT12			Cytotoxic, antitumor	[133]
Cyclopsychotride A	<i>Oldenlandia affinis</i>	Rubiaceae	Antimicrobial	[127]
Cycloviolacin (Y1/Y4/Y5)	<i>Viola yedoensis</i>	Violaceae	Anti-HIV, hemolytic	[134]
Cycloviolacin H3	<i>Viola hederaceae</i>		Nematocidal	[140]
Cycloviolacin H4	<i>Viola hederaceae</i>		Hemolytic	[17]
Cycloviolacin O1	<i>Viola odorata</i>		Anthelmintic, molluscicidal	[137,140]
Cycloviolacin O13-O14	<i>Viola odorata</i>		Anti-HIV, hemolytic, nematicidal	[140,154]
Cycloviolacin O15-O16	<i>Viola odorata</i>		Nematocidal	[140]
Cycloviolacin O2	<i>Viola biflora</i> <i>Viola odorata</i>		Antibacterial, molluscicidal, cytotoxic, antitumor, hemolytic, Anthelmintic	[17,138-140;165;167]
Cycloviolacin O24	<i>Viola odorata</i>		Anti-HIV, hemolytic	[154,132]
Cycloviolacin O3-O8	<i>Viola odorata</i>		Nematocidal	[141]
Cycloviolacin A-D	<i>Leonia cymosa</i>		Anti-HIV	[135]
EETI	<i>Ecballium elaterium</i>	Curcubitaceae	Trypsin inhibitor	[155]
Hypa A	<i>Hybanthus parviflorus</i>	Violaceae	Insecticidal	[156]
Kalata B1	<i>Oldenlandia affinis</i> <i>Viola odorata</i>	Rubiaceae Violaceae	Antibacterial, anti-HIV, hemolytic, insecticidal, Anthelmintic, molluscicidal	[74,101,127,133-137]

Kalata B2	<i>Oldenlandia affinis</i>	Rubiaceae	Antibacterial, insecticidal, Anthelmintic, molluscicidal	[103,124,157-158]
Kalata B5			Hemolytic	[159]
Kalata B6/B7			Anthelmintic	[139]
Kalata B8			Anti-HIV, antitumor	[160]
MCoTI I/II	<i>Momordica cochinchinensis</i>	Curcubitaceae	Trypsin inhibitor	[56,63]
Palicourein	<i>Palicourea condensate</i>	Rubiaceae	Anti-HIV	[161]
Parigidin-br-1	<i>Palicourea rigida</i>		Insecticidal	[93]
SFTI-1	<i>Helianthus annuus</i>	Asteraceae	Trypsin inhibitor	[70]
Vaby A/D	<i>Viola abyssinica</i>	Violaceae	Cytotoxic	[162]
Varv A/F	<i>Viola arvensis</i>		Antitumor	[131]
	<i>Viola odorata</i>			
Varvpeptide D/E/H	<i>Viola tricolor</i>		Cytotoxic	[163]
Vhl-1	<i>Viola hederaceae</i>		Anti-HIV	[164]
Vibi D/E/G/H	<i>Viola biflora</i>		Cytotoxic	[165]
Vila A/B	<i>Viola labridorica</i>		Cytotoxic	[166]
Vitri B/C/D/E/F	<i>Viola tricolor</i>		Cytotoxic	[163]

Graphical abstract



CAPÍTULO II

(Artigo a ser submetido na Revista *Genetics and Molecular Biology*)

Isolamento e caracterização *in silico* de ciclotídeos em membros da família Poaceae

Sheyla Carla Barbosa da Silva Lima¹, João Pacifico Bezerra-Neto¹, Ana Maria Benko-Iseppon¹ and Valesca Pandolfi^{1*}

¹ Universidade Federal de Pernambuco, Centro de Ciências Biológicas, Departamento de Genética, Av. Prof. Moraes Rêgo 1235, CEP 50670-420, Recife, PE, Brasil;

Autor correspondente:

Valesca Pandolfi, Universidade Federal de Pernambuco, Centro de Ciências Biológicas, Av. Prof. Moraes Rego 1235, CEP 50.670-420, Recife, PE, Brasil. Tel. +55 81-2126-7816; Fax: 55-82-2126-8522; Email: valescapandolfi@gmail.com

Resumo

Ciclotídeos são pequenas proteínas, presentes em plantas, que apresentem seis cisteínas conservadas, as quais formam três pontes dissulfeto. O motivo estrutural encontrado nesta molécula é o nó cistina cíclico (CCK - *Cyclic Cystine Knot*), o que o torna estável e confere aos ciclotídeos resistência à degradação térmica e enzimática. Os ciclotídeos possuem múltiplas atividades biológicas, como bactericida, inseticida, antiviral e antifúngica, gerando um interesse cada vez maior por esta família de peptídeos antimicrobianos. Neste estudo quatro sequências gênicas de duas espécies da família Poaceae foram isoladas, sendo três de milho, *Zea mays* (zm315, zm316 e zm317) e uma de centeio, *Secale cereale* (sc359). As referidas sequências foram caracterizadas *in silico* como prováveis ciclotídeos, uma vez que apresentam propriedades físico-químicas similares aos peptídeos já descritos, com estrutura homóloga à uma estrutura de ciclotídeo de uma espécie pertencente a família Rubiaceae, disponível no PDB databank. Os modelos tridimensionais gerados por homologia apresentaram similaridade entre 65.39 % e 73.08 % com o template. As análises indicam ainda, que dois destes possíveis ciclotídeos parecem ser lineares (zm315 e zm316), pois não apresentam na região C-terminal os aminoácidos N/D que estão relacionados com a ciclização das moléculas. O presente estudo envolve a descrição da possível estrutura de membros deste grupo de peptídeos antimicrobianos em bancos de dados para a família Poaceae.

Palavras-Chave: AMPs, milho, centeio, modelagem por homologia

Introdução

Estudos de caracterização dos ciclotídeos iniciaram-se na década de noventa (Craik *et al.* 1999). Tais peptídeos variam de tamanho entre 28 e 37 aminoácidos, apresentando uma estrutura cíclica, do tipo “cabeça-cauda” e contendo seis cisteínas conservadas, ligadas numa topologia de nó e formando três pontes dissulfeto (Craik *et al.*, 2002). Os resíduos de cisteínas são numerados por algarismos romanos do I ao VI (Daly *et al.*, 2009). Quanto às ligações dissulfeto são estabelecidas da seguinte forma: CysI-CysIV, CysII-CysV e CysIII-CysVI, essa última penetrando no anel I (Craik, 2010). O motivo estrutural compartilhado entre os ciclotídeos é o nó cistina cíclico, do inglês *Cyclic Cystine Knot* (CCK). Com base na sua estrutura primária da cadeia polipeptídica, os ciclotídeos são divididos em duas subfamílias: Möbius e a Bracelet (Craik *et al.*, 1999).

Devido à sua vasta atividade biológica, os ciclotídeos apresentam-se como moléculas de grande potencial tanto para a indústria farmacêutica, no desenvolvimento de novas drogas, como em aplicações biotecnológicas no setor agrícola. Tais atividades incluem ação bactericida, inseticida, antiviral, antifúngica, nematicida, moluscicida, aplicações antitumorais e hemolíticas (Pomilo *et al.*, 2006; Brodgen and Brodgen 2011; Pinto *et al.*, 2011; Craik, 2012; Cemazar *et al.*, 2012). Estas atividades foram descritas com base em propriedades destes peptídeos encontrados em diferentes espécies de plantas das famílias Rubiaceae, Violaceae, Curcubitaceae e Fabaceae. Entretanto, pouco se sabe sobre a diversidade e estrutura desses peptídeos em outras famílias vegetais como, por exemplo em membros da família Poaceae.

A família Poaceae (Gramineae) é considerada uma das famílias de angiospermas mais amplamente distribuídas, representando aproximados 40 % da cobertura vegetal no mundo (Gibson, 2009) conhecida como a quarta maior família, com cerca de 700 a 800

gêneros e 10.000 a 11.000 espécies (Peterson, 2005; GPWG, 2011; Souza and Lorenzi 2012). Estão entre os grupos de plantas de grande importância econômica mundial, pois se destacam como importantes alimentos para humanidade, com destaque para o arroz (*Oryza sativa* L.), o trigo (*Triticum aestivum* L.), o milho (*Zea mays* L.), além da cana-de-açúcar (*Saccharum officinarum*), sendo estes últimos grandes fornecedores de matéria prima para fabricação de biocombustível (Schmidt and Longhi-Wagner 2009; Salla *et al.*, 2010)

O milho é um dos cereais mais importantes cultivados no mundo, se destacando como um cereal de alta qualidade nutricional, com grande importância para a produção agrícola do Brasil, entre as graníferas, sendo destinado para consumo humano e animal (Santos and Prado, 2002). Em países industrializados, o milho pode ser encontrado em muitos produtos como: cosméticos, roupas, papel, pasta dental, tintas, loções, entre outros. Além disso, o milho vem sendo utilizado na confecção de polímeros biodegradáveis que são matérias-primas para as indústrias de embalagens e automobilística (Miranda-Peixoto, 2002).

Segundo o oitavo levantamento do USDA (*United States Department of Agriculture*) haverá recorde na produção mundial de milho 2014/15 com cerca de 991,6 milhões de ton, superando a produção de 2013/14 que foi de 989,3 milhões de toneladas (ton.). O Brasil aparece em terceiro lugar com 75 milhões de ton., perdendo para EUA e China, 353,7 e 218,5 milhões de ton., respectivamente.

Não menos importante está o centeio, o qual ocupa o oitavo lugar na lista mundial em termos de área colhida e de produção no mundo (14,11 milhões de ton, safra 2012/13) (Mori *et al.*, 2013).

Entretanto, como toda grande cultura, tanto milho como o centeio apresentam limitações de cultivo em decorrência de doenças e pragas, cujo ataque vem provocando

o comprometimento da produção, qualidade e produtividade. Dessa forma, a identificação e caracterização de peptídeos antimicrobianos, em especial ciclotídeos, nessas culturas, fornece base para o melhoramento de ambas, frente a estresses bióticos, uma vez que são responsáveis por perdas significativas em ambas. Assim, neste estudo foram isoladas e caracterizadas *in silico* sequências de possíveis ciclotídeos cíclicos e uma “variedade” de ciclotídeos lineares nessas duas espécies da família Poaceae.

Materiais e métodos

Busca e análise de sequências de ciclotídeos

Inicialmente foi realizada a busca por sequências homólogas a ciclotídeos em bancos de dados para diferentes espécies vegetais, visando a identificação de possíveis genes codificantes para ciclotídeos nas mesmas. Esta busca foi realizada por meio do método probabilístico fundamentado no Modelo Oculto de Markov (HMM – *Hidden Markov Model*; Eddy, 1996) com o pacote de ferramentas HMMER 3.1. Para o desenho dos modelos, via HMMER, foram utilizadas todas as sequências proteicas de ciclotídeos disponíveis no GenBank/NCBI, sendo estas alocadas em um banco de dados local e curadas manualmente para seu agrupamento, seguindo a classificação taxonômica dos organismos de origem. Ao total foram montados sete grupos: Solanales, Malphigiales, Gentianiales, Fabales, Poales, sintéticos (sintetizados experimentalmente) e um grupo misto (composto por todas as sequências obtidas). Foi realizado, então, o alinhamento das sequências constituintes de cada um desses grupos, sendo o resultado de cada um dos alinhamentos o *input* do programa HMMER, no desenho dos diferentes motivos proteicos a serem utilizados como *query* na busca por possíveis homólogos nos bancos

locais que foram compostos pelo genoma das espécies, com exceção de *Vigna unguiculata* que foi utilizado o transcriptoma.

Para a realização da busca por ciclotídeos, foram montados bancos locais para diferentes organismos vegetais, visando maximizar as chances de sucesso na obtenção de homólogos a estes peptídeos em diferentes espécies vegetais. A construção do banco local foi realizada a partir do *download* dos genomas completos para cinco espécies vegetais: *Hevea brasiliensis*, *Manihot esculenta*, *Ricinus communis* (JCVI - J.Craig Venter Institute); *Sorghum bicolor* e *Zea mays* (Phytozome), além do genoma expresso (RNAseq e EST) de *Vigna unguiculata*, disponível no banco NordEST. A busca por sequências ortólogas foi realizada via HMMER, confrontando os motivos construídos para os grupos descritos acima, contra cada um dos genomas traduzidos das seis espécies selecionadas, nos seis quadros (*frames*) de leitura, utilizando um ponto de corte “*cut-off*” $\leq e^{-05}$.

As sequências homólogas a ciclotídeos (obtidas na prospecção via HMMER) foram curadas manualmente, para a retirada de redundâncias, e então traduzidas no quadro de leitura indicado pelo HMMER por meio do programa ORF-*finder* (*Open Reading Frame Finder*) Em seguida, as sequências anotadas foram submetidas a ferramenta *online* Batch (Marchler-Bauer, 2011) para análise da presença e integridade do domínio conservado.

Amplificação das sequências homólogas a ciclotídeos

Desenho de *primers*

A partir das sequências dos genes encontrados e caracterizados via pacote de ferramentas HMMER e Batch em milho (*Zea mays*), foi conduzido o desenho de pares de *primers* através da ferramenta Primer3 (Untergasser *et al.*, 2012), a serem aplicados

no isolamento de sequências homólogas a ciclotídeos em membros da família Poaceae. Os *primers* foram desenhados levando-se em consideração os seguintes parâmetros: tamanho do oligonucleotídeo (mínimo: 18 bases; ótimo: 20 bases; máximo: 22 bases); temperatura de Melting (mínimo: 57° C; ótimo: 60° C; máximo 63° C); e conteúdo em GC (mínimo: 40%; ótimo: 50%; máximo: 60%). Com base nesses parâmetros, foram obtidos dois pares de *primers*: ZmCicloI (F:... e R: ...) e ZmCicloII (F: ... e R: ...). [Observação: as sequências foram omitidas por motivo de propriedade intelectual].

Material vegetal e extração de DNA

Folhas jovens de *Zea mays*, cultivar São José - BR 5026 (mantidas no Instituto Agrônômico de Pernambuco – IPA) foram coletadas e armazenadas em ultrafreezer (-80°C). A extração do DNA genômico foi realizada conforme o protocolo descrito por Weising *et al.* (2005), com modificações no que se refere à quantidade de material vegetal (aproximadamente 2,5 g). Em seguida, foi realizada precipitação seletiva de polissacarídeos conforme descrito por Michaels *et al.*, (1994). Procedeu-se, então, à quantificação e à análise de qualidade do DNA em gel de agarose 1 % (contendo blue green), por comparação com DNA lambda (λ -DNA) em diferentes concentrações (20, 50 e 100 ng/ μ L). Os DNAs de centeio (cv. BR1) e trigo (IAC-5), já estavam disponíveis no banco de DNAs no Laboratório de Genética e Biotecnologia Vegetal (LGBV) tendo sido extraídos com a mesma metodologia citada acima.

Reação em Cadeia da polimerase – PCR

A partir dos dois primers obtidos (ZmCicloI e ZmCicloII) foram realizadas as reações de PCR nas seguintes condições: 2 μ L de DNA (20 ng de cada espécie alvo), 1,5 μ L de MgCl₂ (25 mM), 0,25 μ L de dNTPs (10 mM), 1 μ L de cada *primer* (10 mM), 0,2 U Taq polimerase (5U / μ L), 2,5 μ L de solução tampão 10x (500 mM KCl, 100 mM

Tris-HCl, pH 8.3 e 1,4 mM MgCl₂) em volume final de 25 µL. Inicialmente as reações de PCR foram realizadas em gradiente de temperatura, para definir a melhor temperatura e, portanto, aumentando a especificidade dos *primers*. Após eletroforese em gel de agarose a 2,5 %, as bandas (amplicons) foram isoladas do gel, purificadas com o kit “GFX PCR DNA and Gel Band Purification Kit” (GE Healthcare) e sequenciadas utilizando o kit BigDye Terminator v.3.1 Cycle Sequence Kit (Applied Biosystems), em um sequenciador automático ABI 3130XL (Applied Biosystems, Foster City, CA-USA).

Análises de bioinformática

As sequências obtidas foram avaliadas no programa Bioedit (Hall, 1999) e submetidas ao ORF *Finder*, para identificação do quadro de leitura aberta e tradução da sequência para aminoácidos. Após essa etapa, as sequências de nucleotídeos e proteínas foram confrontadas com sequências disponíveis nos bancos de dados do GenBank/NCBI e Phytozome, utilizando os algoritmos tBLASTx e BLASTp para análise de similaridade das sequências com possíveis homólogos depositados nos bancos de dados.

Os peptídeos hipotéticos (obtidos após tradução) foram submetidos ao programa de predição do sítio de clivagem do peptídeo sinal ‘Philius Transmembrane Prediction Server’ (Reynolds *et al.*, 2008) e ao SignalP Server (Petersen *et al.*, 2011). A formação das pontes dissulfetos foi predita através do programa DiANNA (Ferrè and Clote, 2005) e o alinhamento múltiplo de proteínas foi feito no programa Clustal Omega (Sievers *et al.*, 2011). Para identificação do domínio, as sequências foram submetidas à ferramenta *online* “Smart” (Letunic *et al.*, 2015). Adicionalmente, as sequências foram submetidas ao Blast2GO a fim de predizer a possível função biológica das proteínas.

Análises comparativas das propriedades físico-químicas de ambas as proteínas - completas e apenas com o domínio, foram realizadas a partir do programa ProtParam (Gasteiger *et al.*, 2005), sendo sua estrutura secundária obtida via ferramenta Alphapred (Kaur and Raghava, 2004).

Em seguida, procedeu-se a uma análise de modelagem comparativa das sequências proteicas obtidas, visando à determinação de sua estrutura terciária, contribuindo para o entendimento de sua estrutura. Para tal, a estrutura tridimensional de um ciclotídeo de uma espécie pertencente a família Rubiaceae (depositada no PDB - *Protein Data Bank* e obtida por Espectrometria de Ressonância Magnética - NMR) foi utilizada como *template* para a construção do modelo. Foi utilizada a ferramenta *online Protein Structure & Model Assessment* da plataforma SWISS-MODEL Workspace (Arnold *et al.*, 2006), onde ainda foi avaliada a qualidade do modelo, juntamente com os dados para a avaliação via Procheck (LAskowski *et al.*, 1993), que determina a qualidade estereoquímica do modelo, através da análise do gráfico de Ramachandran e QMEAN (Benkert *et al.*, 2009), comparando-se o modelo gerado com estruturas depositadas de mesmo tamanho. As estruturas modeladas foram visualizadas no Jmol 13.0.

Resultados

Busca e análise de sequências candidatas a ciclotídeos

Na busca por sequências homólogas a ciclotídeos vegetais, via ferramenta HMMER, foram selecionadas 324 sequências de ciclotídeos, obtidas no GenBank/NCBI, as quais foram utilizadas para geração de motivos (anexo 3). Após a geração dos consensos e, realizada a comparação com os genomas de seis espécies vegetais, foram obtidas: 12 sequências candidatas para *H. brasiliensis*, 48 para *M. esculenta*, 30 para *R. communis*, 36 em *S. bicolor*, 34 em *V. unguiculata* e 64 em *Z. mays*, obtendo-se 224

sequências no total (Tabela 1), seguindo-se então a caracterização por meio da análise do domínio conservado (Tabela 2).

Tabela 1. Sequências caracterizadas nos genomas de *Vigna unguiculata* (Vu); *Zea mays* (Zm), *Sorghum bicolor* (Sb), *Hevea brasiliensis* (Hb), *Ricinus communis* (Rc) e de *Manihot esculenta* (Me), utilizando os modelos construídos para os diferentes grupos taxonômicos de ciclotídeos.

Grupos taxonômicos	Número de sequências por espécie					
	Vu	Zm	Sb	Hb	Rc	Ms
Malpighiales	29	61	32	12	28	42
Gentianales	0	2	0	0	0	0
Solanales	0	2	0	0	0	0
Fabales	1	2	0	0	0	0
Monocotiledôneas	0	3	0	0	0	0
Sintéticos	0	2	0	0	0	0
Grupo Misto	36	62	36	12	30	48
Total*	34	64	36	12	30	48

*Total de sequências caracterizadas por genoma.

A análise da presença e integridade dos domínios conservados (via ferramenta Batch CD-search) resultou na obtenção de 21 domínios, sendo dois deles com domínio característico de ciclotídeos ambos encontrados no genoma de *Z. mays* (Tabela 2). O domínio de maior ocorrência, no entanto foi o *Thioredoxin_like superfamily*, que é um grupo grande de proteínas que apresentam o domínio TRX (Thioredoxin) com um motivo redox CXXC. Tais proteínas funcionam como oxidoreduases dissulfeto, alterando o estado redox de proteínas alvo. Nesta classe ainda se encaixam as proteínas dissulfeto isomerase que estão relacionadas ao dobramento dos ciclotídeos. Com base nesse resultado, partiu-se para a caracterização dos dois genes no banco Phytozome. A

predição do gene realizada via fgenesh (HMM-based gene structure prediction) permitiu observar que um dos genes apresentava 849 pb (sendo a região codificante para ciclotídeos de 252 pb), enquanto que o outro gene apresentava 528 pb (sendo 255 pb a região codificante para ciclotídeos).

Tabela 2. Domínios encontrados nas espécies candidatas: *Vigna unguiculata* (Vu); *Sorghum bicolor* (Sb), *Zea mays* (Zm), *Hevea brasiliensis* (Hb), *Manihot esculenta* (Me) e *Ricinus communis* (Rc), obtidos após a busca por homólogos via HMMER, utilizando a ferramenta de busca de domínios Batch.

Domínios	Espécies					
	Vu	Sb	Zm	Hb	Me	Rc
<i>AANH_like superfamily</i>	4	1	2	-	1	1
<i>APH_ChoK_like superfamily</i>	1	-	-	-	-	-
<i>Albumin_I superfamily</i>	1	-	-	-	-	-
<i>COPIIcoated_ERV superfamily</i>	1	-	-	-	-	-
<i>Cyclotide superfamily</i>	-	-	2	-	-	-
<i>DUF2847 superfamily</i>	-	-	-	-	-	1
<i>DUF594 superfamily</i>	-	-	1	-	-	-
<i>ERGIC_N superfamily</i>	1	-	-	-	-	-
<i>ERp29c superfamily</i>	1	2	5	-	2	2
<i>Iso_dh superfamily</i>	-	-	1	-	-	-
<i>NAD_binding_8 superfamily</i>	1	-	2	-	2	1
<i>Periplasmic_Binding_Protein-Type_1 superfamily</i>	-	-	-	-	-	1
<i>PKc_like superfamily</i>	1	-	-	-	1	-
<i>PRK13728 superfamily</i>	-	-	2	-	-	2
<i>Pyr_redox superfamily</i>	1	-	2	-	2	1
<i>SANT superfamily</i>	-	-	2	-	-	-
<i>SBP superfamily</i>	-	-	1	-	-	-
<i>Thioredoxin_5 superfamily</i>	-	-	3	-	-	-
<i>Thioredoxin_like superfamily</i>	53	55	97	12	68	45
<i>TPR superfamily</i>	1	-	-	-	2	1
<i>Wax2_C superfamily</i>	-	-	1	-	-	-

Amplificação das sequências homólogas e caracterização *in silico*

Dois pares de *primers* foram desenhados a partir das duas sequências encontradas e caracterizadas em milho. As reações de PCR resultaram na amplificação de 19 fragmentos (amplicons), com tamanho variando entre 200 e 450 pb.

A análise destas sequências revelou que apenas quatro fragmentos gênicos, sendo três (zm315, zm316 e zm317) de *Z. mays* e um (sc359) de *Secale cereale* (centeio), apresentaram não somente similaridade a ciclotídeos, como também apresentaram o domínio ciclotídeo (via *Conserved Domain Search Service - CD Search*), permitindo a continuidade da análise via *ORF Finder* e demais ferramentas de predição. Os fragmentos gênicos apresentaram pouca variação, sendo 366, 245, 323 e 410 pares de bases (pb) para zm315, zm316, zm317 e sc359, respectivamente, apresentando maior similaridade com os seguintes acessos de *Z. mays* (disponíveis no GenBank), com variação entre 79 e 86 % como observado na Tabela 3. Quando estas sequências foram analisadas no banco de dados Phytozome, as quatro sequências apresentaram similaridade com sequências de *Z. mays*, de *Setaria italica* (gramínea) e de *Sorghum bicolor*, além da similaridade com todas já citadas também houve similaridade com uma sequência de *Panicum virgatum* (anexo 4). Nos dois bancos confrontados a família ciclotídeos aparece como domínio correspondente as sequências isoladas, com exceção da sequência de *Panicum virgatum*, o que indica a similaridades destas sequências com as de ciclotídeos já disponíveis em bancos de dados.

Tabela 3. Descrição das sequências isoladas de ciclotídeos com domínio (*cyclotide Family*) conservado com similaridade no banco do NCBI.

Nome	Tamanho			Maior Similaridade (%)	
	(nt)	(RC)	(aa)	(DC)	NCBI*
zm315	366	43 ao 294	83	30	-
zm316	253	1 ao 201	65	31	-
zm217	323	28 ao 285	85	31	-
sc359	410	51 ao 338	95	39	-

Nt = nucleotídeo; RC = região codificante; aa = aminoácidos; DC = domínio conservado.

*Observação: as sequências foram omitidas por motivo de propriedade intelectual.

Análise de similaridade, função e estrutura secundária.

Os fragmentos zm315, zm316, zm317 e sc359 após a tradução via ORF Finder, apresentaram uma região codificante de 252, 201, 258 e 288 pb e peptídeos de 83, 65, 85 e 95 aminoácidos, respectivamente. Com exceção de zm316, todos apresentaram no início da sequência de nucleotídeos uma região que não codificou aminoácidos, de comprimento variável entre 28 pb (zm317) e 73 pb (zm315 e sc359). Após a predição do peptídeo sinal (SignalP e Phyllius), verificou-se que ambos os peptídeos apresentaram a posição do sítio de clivagem entre os aminoácidos 29 e 30, para zm315, e entre os resíduos 27 e 28, para zm317, já as sequências zm316 e sc359 não apresentaram o peptídeo sinal.

Também através do programa Phyllius foi verificado que as proteínas codificadas pelos fragmentos gênicos zm315, zm317 e sc359 não são citoplasmáticas, ao passo que zm316 não obteve resultado. A região do domínio conservado foi identificada por meio da ferramenta *online Smart*, a qual revelou que o domínio de zm315 tem início a partir do resíduo de aminoácido 54 até a posição 83 (Figura 1A). Para zm316 o domínio se estende entre os resíduos 35 a 65 (Figura 1B), enquanto que para zm317 e sc359, este é compreendido entre as posições 55 a 85 (Figura 1C e D, respectivamente). A estrutura predita para todos os candidatos descritos encontra-se ilustrada na Figura 1.

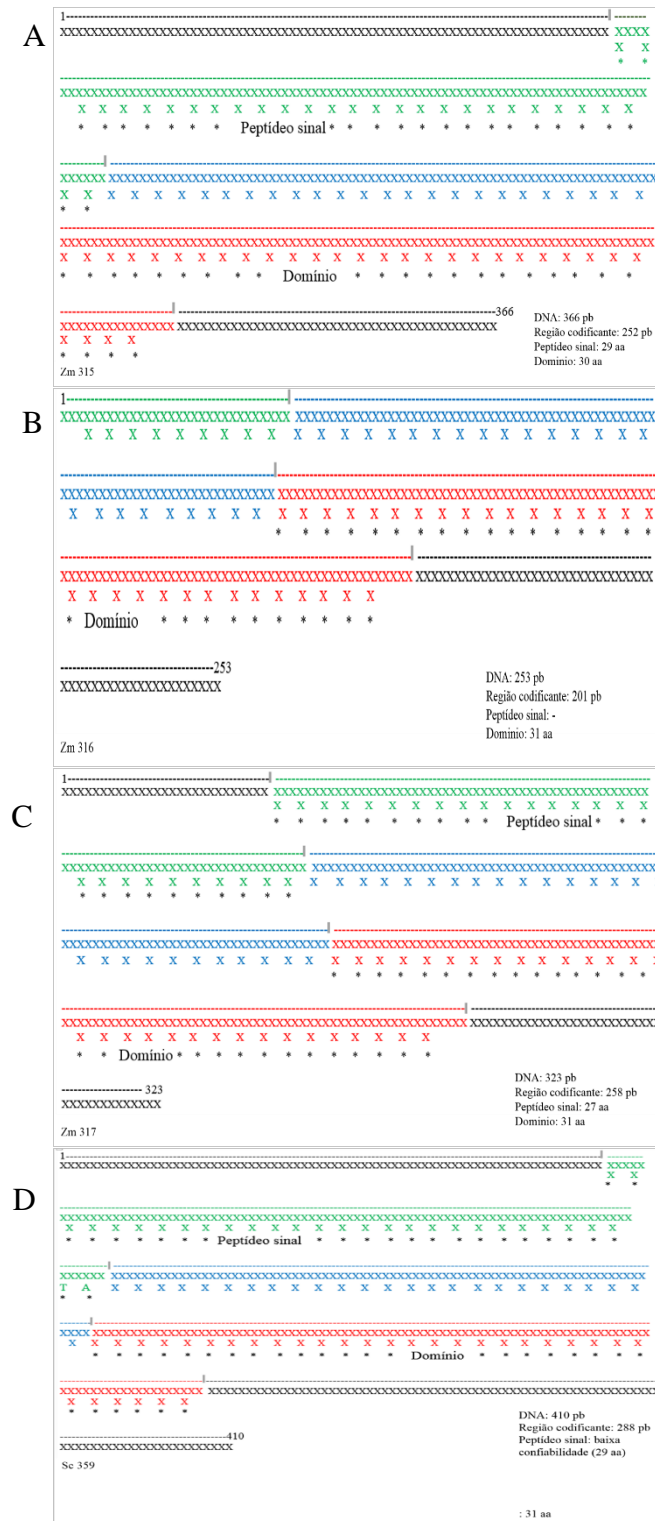


Figura 1. Predição estrutural dos fragmentos isolados de milho (A) Zm315, (B) Zm316, (C) Zm317 e de centeio (D) Sc359, a partir de dados gerados no *Philius* e *Smart*. As regiões ao longo das sequências estão delimitadas por letras coloridas, as quais representam: a região do peptídeo sinal (verde); região pro-peptídeo N-terminal (NTPP) (azul) e o domínio ciclotídeo (vermelho). Observação: as sequências foram omitidas por motivo de propriedade intelectual.

Os domínios dos peptídeos hipotéticos (alinhados contra as sequências do GenBank/ NCBI) apresentaram similaridade entre 100 – 61 %. A sequência zm315 mostrou 100 % de similaridade com a proteína hipotética de *Z. mays*, e 70 % com uma proteína precursora de ciclotídeo, isolada de uma espécie arbórea, da família Violaceae: *Melicytus ramiflorus*.

A Sequência zm316 apresentou maior similaridade (71 %) com um ciclotídeo, isolado de *Gloeospermum blakeanum*, também uma de Violaceae. O peptídeo predito ainda apresentou 75 % de similaridade com um precursor de ciclotídeo, identificado na espécie *Chassalia chartacea*.

Já a sequência zm317 com 77 % de similaridade com a proteína hipotética de *Z. mays*, também mostrou 61 % de ciclotídeo isolado de *Palicourea rígida* (Rubiaceae). Por fim, o peptídeo gerado para sc359 apresentou, respectivamente, 65 % e 72 % de similaridade com dois precursores de ciclotídeo: um de *Oldenlandia affinis* e o outro de *Clitoria ternatea* (Figura 2).

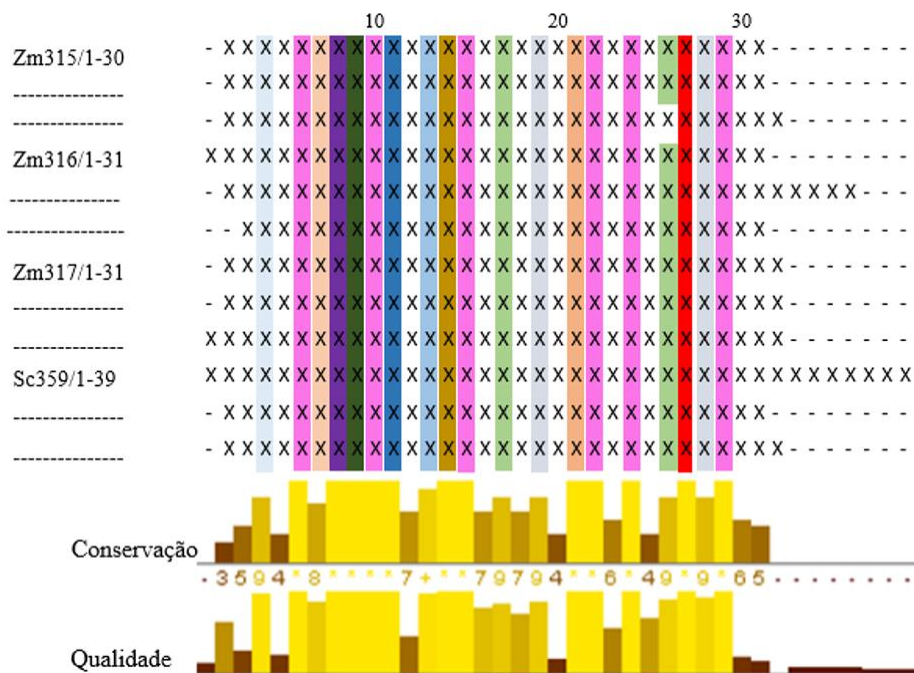


Figura 2. Alinhamento das sequências similares encontradas no GenBank, através da ferramenta BLASTp. As cores em destaque mostram os resíduos com maior conservação. Observação: as sequências foram omitidas por motivo de propriedade intelectual.

Os detalhes da composição de aminoácido de cada sequência são mostrados na Figura 3 e Figura 4. Verifica-se um maior número de resíduos de leucina em zm315, alanina em zm316, leucina em zm317 e em sc359, assim como valina está em maior presença (Figura 4).

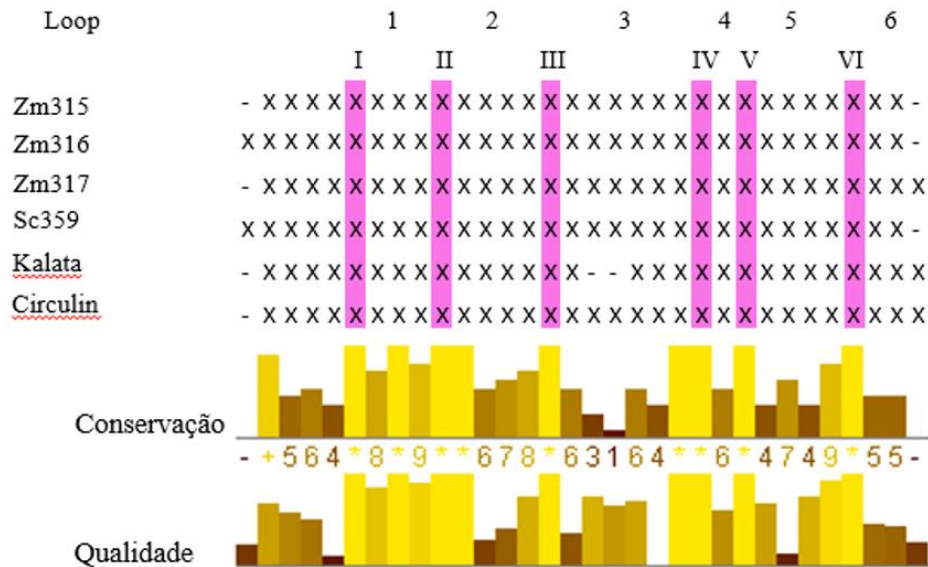


Figura 3. Alinhamento das sequências isoladas e identificadas como possíveis ciclotídeos com duas sequências caracterizadas como ciclotídeos disponíveis no GenBank/NCBI (Destaque: seis cisteínas em rosa) Observação: as sequências foram omitidas por motivo de propriedade intelectual.

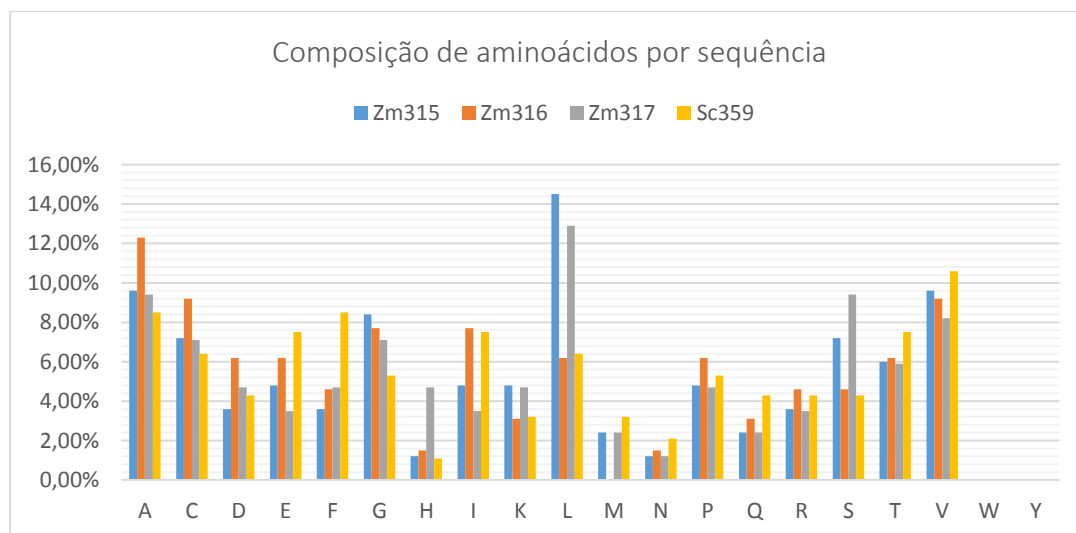


Figura 4. Porcentagem de aminoácidos presentes em cada uma das sequências isoladas (Zm315, Zm316, Zm317 e Sc359)

Quando alinhadas com sequências de supostos ciclotídeos da família Poaceae, disponíveis nos bancos de dados do GenBank/NCBI e Phytozome (Figura 5), verifica-se a notável conservação de diversos aminoácidos, em especial as seis cisteínas conservadas (posições 6, 10, 15, 22, 24 e 29).

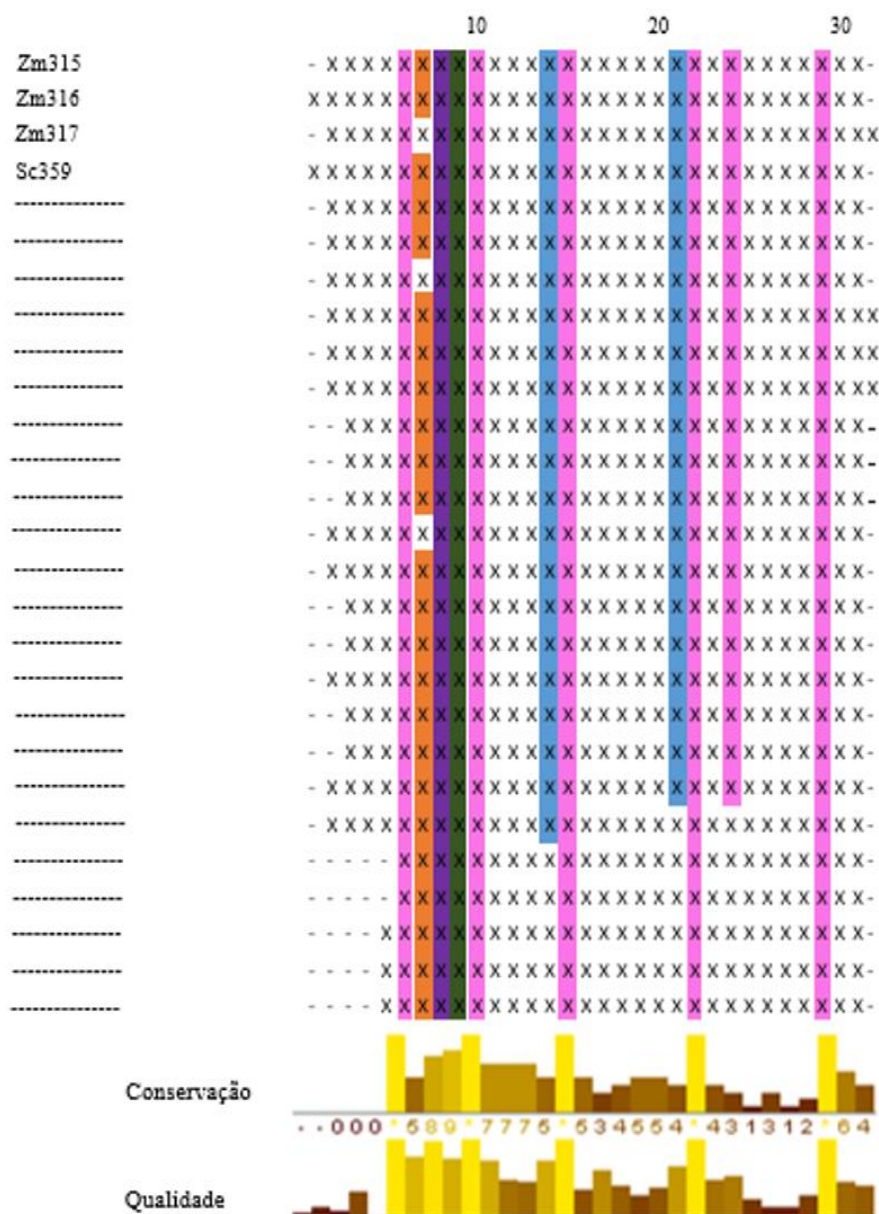


Figura 5. Sequências isoladas neste estudo e sequências da família Poaceae disponíveis nos bancos de dados do GenBank/ NCBI e Phytozome. Observação: as sequências foram omitidas por motivo de propriedade intelectual.

Em relação à função dos ciclotídeos preditos, foi realizada uma análise via Blast2GO, onde a caracterização funcional é feita *in silico*. Todos os possíveis ciclotídeos foram similares a “Cyclotide vitri-a-like” com função em resposta de defesa.

Predição da estrutura secundária

A conformação secundária (predita pelo programa Alphapred) das quatro sequências foram assim estruturadas: zm315, com três folhas- β nos resíduos (Cys₉, Val₁₀, Ile₁₁) (Arg₂₂) e (Leu₂₇) e uma α -hélice (resíduo Ile₁₂ ao Leu₁₉). Já a zm316 apresentou quatro folhas- β (Ile₃ ao Cys₆) (Cys₁₀ ao Phe₁₂) (Arg₂₃) (Val₂₈) e uma α -hélice (Ile₁₃ ao Leu₂₀). Igualmente, a zm317 apresentou 4 folhas- β (Ile₃ ao Cys₅) (Val₁₀, Phe₁₁) (Arg₂₂) (Lys₂₆ ao Cys₂₈) e uma α -hélice (Leu₁₃ ao Phe₂₀). Por fim sc359, com três folhas- β (Ile₃ ao Cys₆) (Cys₁₀ ao Cys₁₅) (Val₂₈) e 2 α -hélices (Thr₁₈) e (Gln₂₀) (Figura 6).

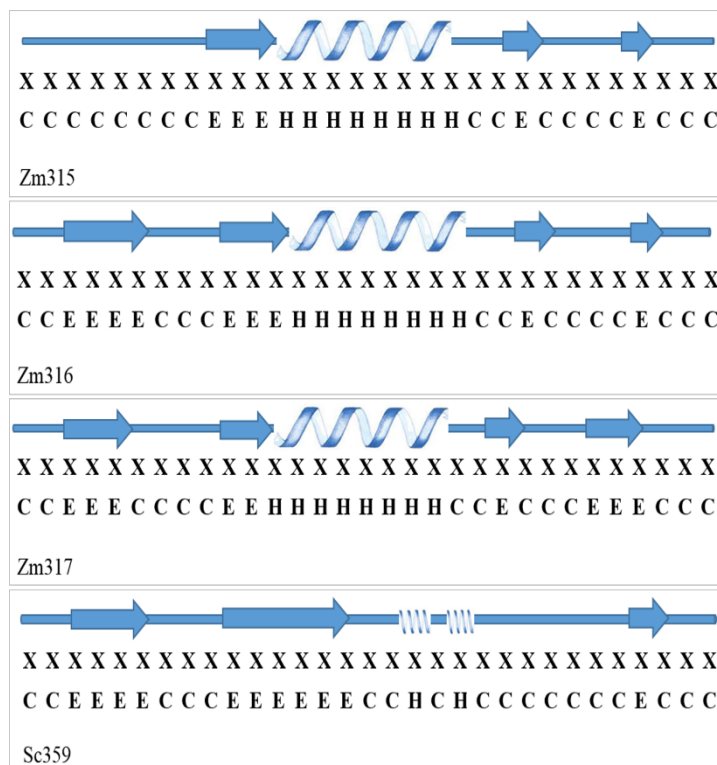


Figura 6. Estrutura secundária predita através da ferramenta Alphapred, para as sequências zm315, zm316, zm317 e sc359 isoladas (sequências superiores) e a estrutura encontrada (sequências inferiores).

Folhas- β (seta azul; E) e α -hélice (estrutura helicoidal azul, H). Observação: as sequências foram omitidas por motivo de propriedade intelectual.

A análise realizada com a ferramenta *online* DiANNA, estas sequências apresentaram as três pontes de dissulfeto, características de AMPs, sendo formadas pelas seguintes cisteínas de cada sequência: (Cys I – IV), (Cys II – V) e (Cys III – VI), para zm315 e zm316; assim como (Cys I –IV), (Cys II –III) e (Cys V- VI) para zm317 e sc359 (Figura 7).

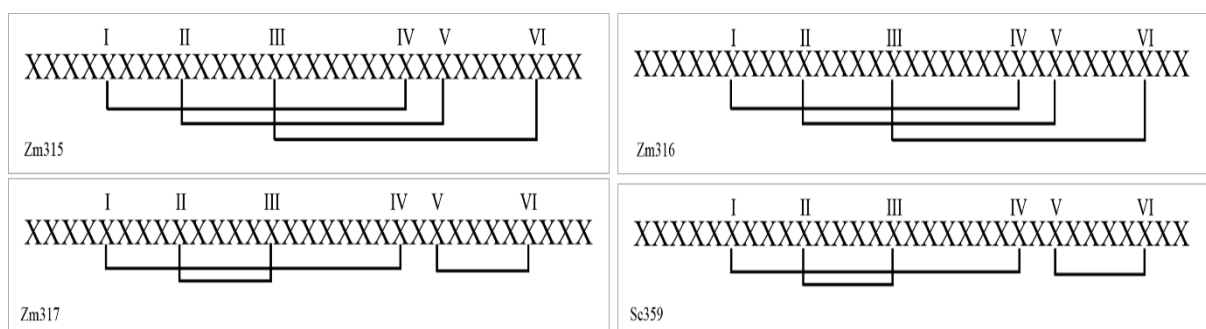


Figura 7. Estrutura predita das pontes dissulfeto (parte inferior da sequência) e as seis cisteínas indicadas por números romanos (parte superior da sequência). Observação: as sequências foram omitidas por motivo de propriedade intelectual.

Propriedades físico-químicas das proteínas

Os dados fornecidos a partir de preditores, mostram a diferença dos resultados quando comparadas as sequências completas ou apenas o domínio (ou seja, o peptídeo maduro sem o peptídeo sinal) (Tabela 4). O peso molecular dos peptídeos completos variou de 6809.8 a 10317.0 Da, enquanto que para os peptídeos maduros a variação foi de 3099.7 a 3260.8.0 Da. O ponto isoelétrico (pI) variou de 4.75 a 6.87 e de 6.10 a 7.79 nos para os peptídeos completos e para os peptídeos maduros (sem o peptídeo sinal), respectivamente. O índice de hidrofobicidade (GRAVY- do inglês *Grand average of hydropathicity*) variou dependendo de ser o peptídeo completo ou somente o peptídeo maduro. Para o peptídeo completo verificou-se uma variação de 0.371 a 0.651, enquanto

que no peptídeo maduro o índice aumentou para 0.632 e 0.935. O índice alifático mostrou variação de 93.08 a 112.77 no peptídeo completo e 84.52 a 110.33 no peptídeo maduro. Já o índice de estabilidade apresentou variação de 25.15 a 36.36 com peptídeo sinal e variou de 8.57 a 43.69 no peptídeo maduro. E por fim a meia vida estimada foi de 30, 20 e 10 horas com e sem o peptídeo sinal (Tabela 4).

Tabela 4. Predição das propriedades físico-químicas das sequências isoladas: Zm315 e Zm316 (possíveis lineares), Zm317 e Sc359 (circulares), e representantes de ciclotídeos já caracterizados Kalata B1 e Circulin B. Cada coluna está representada por valores preditos no peptídeo completo (PepC) e por valores preditos no peptídeo maduro (PepM).

Acesso	Aa		PM (Da)		pI		Res (-)		Res (+)		GRAVY		IA		IE		MVE*	
	PepC	PepM	PepC	PepM	PepC	PepM	PepC	PepM	PepC	PepM	PepC	PepM	PepC	PepM	PepC	PepM	PepC	PepM
zm315	83	30	8679.3	3099.7	6.52	7.77	7	2	7	3	0.651	0.853	112.77	110.33	29.72	25.57	30h >20h >10h	30h >20h >10h
zm316	65	31	6808.8	3229.9	4.75	7.77	8	2	5	3	0.371	0.935	93.08	103.55	25.15	12.12	1.1h, 3min, >10h	30h >20h >10h
zm317	85	31	9015.5	3334.9	6.87	7.79	7	2	7	3	0.394	0.690	97.53	84.52	36.36	43.69	30h >20h >10h	30h >20h >10h
sc359	94	31	10317.0	3260.8	4.75	6.10	11	2	7	2	0.461	0.632	93.30	87.74	26.85	13.05	30h >20h >10h	30h >20h >10h
Circulin B		31		3307.9		8.33		1		3		0.642		90.97		28.76		30h, >20, >10
Kalata B1		29		2913.5		5.96		1		1		0.272		43.45		41.99		-

Aa: número de resíduos de aminoácidos; PM: Peso Molecular; pI: Ponto Isoelétrico; Res. (-/+) Resíduos negativos e positivos; GRAVY: *Grand average of hydropathicity index*; IA: Índice Alifático; IE: Índice de Estabilidade; MVE: Meia Vida Estimada para testes em: reticulócitos de mamíferos *in vitro*; levedura *in vivo*; *Escherichia coli in vivo*.

Modelagem comparativa - Predição de estrutura terciária

Para a predição da estrutura 3D das proteínas foi realizada uma modelagem comparativa por homologia, método que consiste basicamente em utilizar uma estrutura proteica já conhecida como modelo (*template*) para modelar a proteína ainda desconhecida, a partir da similaridade entre os aminoácidos e estrutura secundária. Nessa análise as sequências foram submetidas à ferramenta *online* Swiss Model, resultando em maior similaridade com o modelo (*template*) de ciclotídeo, isolado de uma espécie da família Rubiaceae disponível no banco de dados de proteínas - PDB (Protein Data Bank). As identidades das sequências observadas, após o alinhamento com o modelo, foram de 73 % (zm315, zm316 e em sc359) e 65 % em zm317, indicando que estes são modelos satisfatórios. Os modelos gerados podem ser observados na figura 8, que mostra as folhas beta (seta) e uma pequena alfa hélice (verde) nos modelos zm316, zm317 e sc359. Os valores da qualidade QMEAN Z-score foram -2.02, -1,9, -1,48 e -1,61 respectivamente, onde uma estrutura baseada em no método RMN (Ressonância Magnética Nuclear), deve ser maior que -5 e preferencialmente maior que -2 (neste caso usou-se esse parâmetro porque a estrutura utilizada como modelo deriva desse método) (Benker et al., 2011) (Tabela 5). Ainda que os valores de z-score sejam negativos indicando um bom modelo, a plotagem do modelo no gráfico de scores de todas as estruturas proteicas depositadas no PDB precisa ser considerada, as estruturas são representadas nas cores cinza claro para método de cristalografia de raios-x e cinza escuro para ressonância magnética nuclear (figura 9). Com base nesta análise, os ciclotídeos modelados situam-se dentro da área mencionada, obtendo padrões de qualidade desejada, muitas delas próximas ao modelo estrutural utilizado na modelagem comparativa (Figura 9). A distribuição dos aminoácidos (via gráfico de Ramachandran, programa Procheck) apresentou valores de qualidade estereoquímica da estrutura de 81 % para zm315, zm316 e sc359, e de 77 % para zm317, nos resíduos de aminoácidos na região mais favorável (Tabela 5).

Tabela 5. Valores obtidos para identidade, qualidade e energia para os modelos estruturais de ciclotídeos em milho e centeio obtidos via SWISS-MODEL e Procheck.

Nome	(aa)	Identidade*	Z-score	QMEAN	Ramachandran**			
					RMF	RPA	RGP	RNP
zm315	30	73.08%	-2.02	0.17	81%	19%	0%	0%
zm316	31	73.08%	-1.9	0.2	81%	19%	0%	0%
zm217	31	65.39%	-1.48	0.29	77.3%	18.2%	4.5%	0%
sc359	31	73.08%	-1.61	0.26	81%	19%	0%	0%

(*) Identidade com modelo (Observação: modelo omitido por motivo de de propriedade intelectual); (**) RMF: Região Mais Favorável; RPA: Região Permitida Adicional; RGP: Região Generosamente Permitida; RNP: Região Não Permitida.

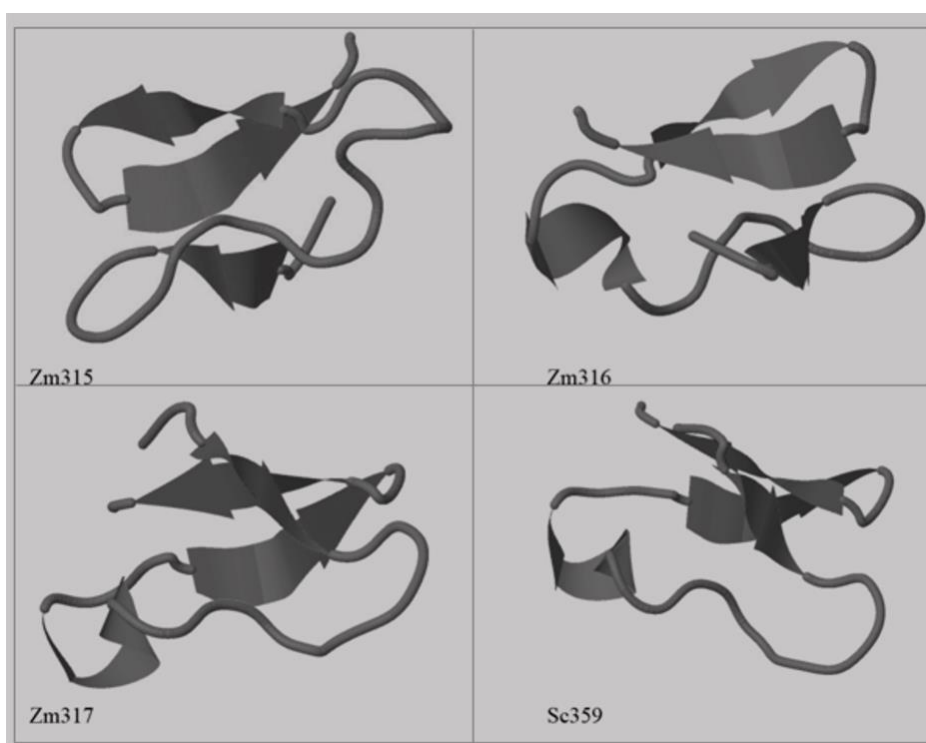


Figura 8. Estrutura 3D modelada por homologia das quatro sequências isoladas. Estrutura em seta caracteriza a folha beta e uma pequena estrutura helicoidal representa a alfa hélice. Observação: os detalhes foram omitidos por motivo de propriedade intelectual.

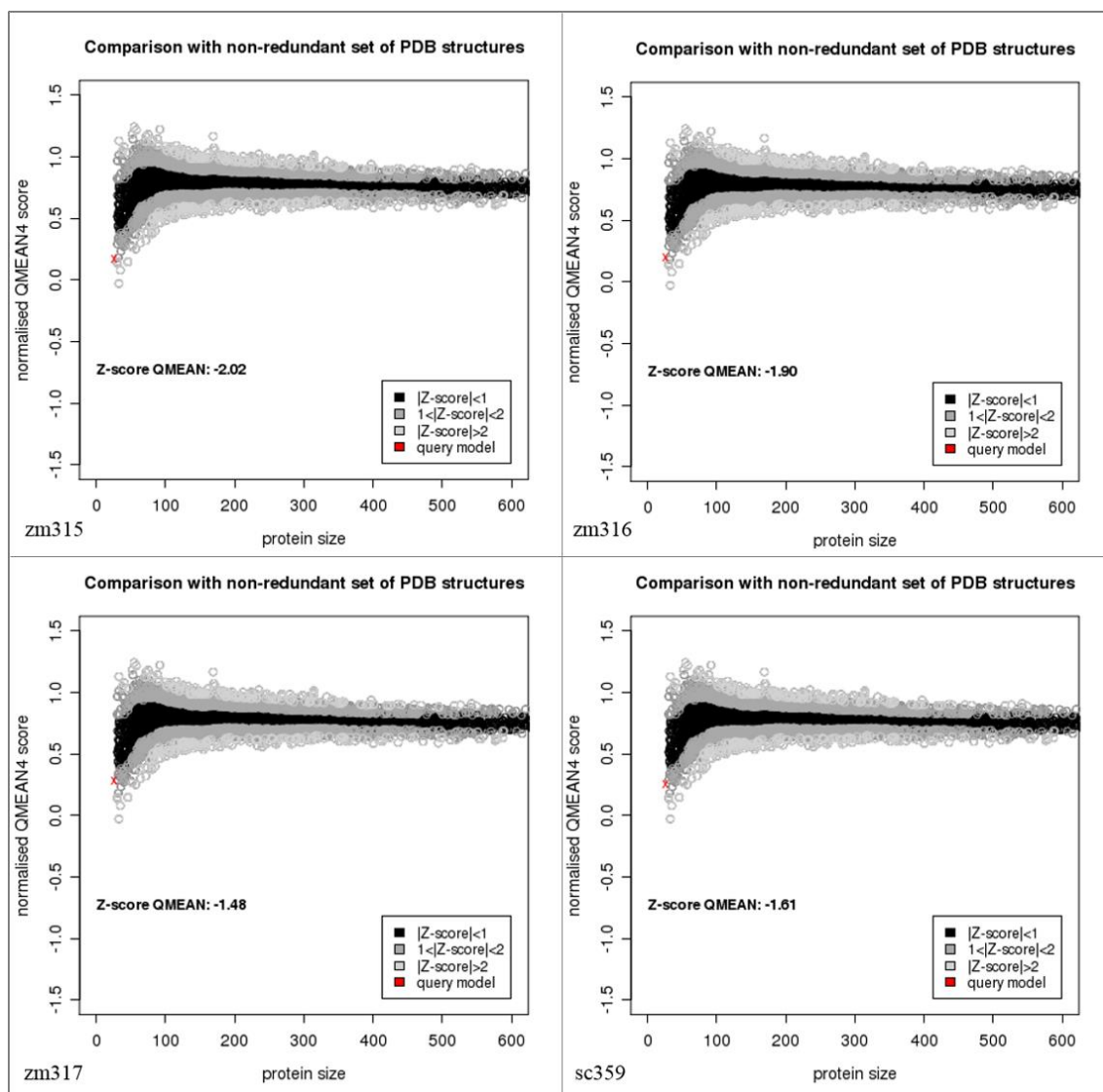


Figura 9. Gráficos representativos do Z-score dos modelos gerados, comparativamente aos modelos de cristalografia de raios-x (cinza claro) e ressonância magnética (cinza escuro) depositados no PDB. Os pontos marcados em vermelho referem-se aos modelos gerados dos prováveis ciclotídeos.

Discussão

Identificação de ciclotídeos em bancos de dados e análise das sequências

Ciclotídeos fazem parte do sistema de defesa das plantas, apresentando uma vasta gama de atividades biológicas contra patógenos e insetos (Craik et al 2004). No sentido de conhecer melhor esses peptídeos, realizamos uma busca, identificação e caracterização estrutural dessas importantes moléculas em diferentes espécies vegetais, com ênfase em membros da família Poaceae.

Por meio de pesquisa e análises de sequências em bancos de dados (GenBank/NCBI) foram obtidos 224 candidatos a possíveis genes codificantes para ciclotídeos. A ferramenta utilizada (HMMER) usa modelos probabilísticos baseados no Modelo Oculto de Markov (EDDY, 1996), permitindo maior eficiência na identificação de homólogos pela geração de modelos a partir de alinhamentos múltiplos de proteínas (EDDY, 2009). As 224 sequências candidatas apresentaram uma maior similaridade com os modelos gerados para cada grupo de ciclotídeos. Foram eles: Solanales, Malphigiales, Gentiniales, Fabales, Poales, sintéticos, ou seja, sintetizados experimentalmente e um grupo misto (composto por todas as sequências obtidas). O critério para esta seleção de candidatos foi o ponto de corte “*cut-off*” $\leq e^{-05}$.

Duas sequências identificadas no genoma de *Z. mays* (disponível no Phytozome,) apresentaram o domínio característico de ciclotídeos (*cyclotide Family*) sendo, portanto utilizadas para desenho de *primers*, gerando a amplificação de quatro fragmentos genômicos de 366, 245, 323 e 410 pb, com pró-peptídeos de 83, 65, 85 e 95 aminoácidos, respectivamente para as sequências zm315, zm316, zm317 (de *Z. mays*) e sc359 (de *Secale cereale*). Os genes que foram encontrados na análise com o HMMER apresentam 849 e 528 pb. Quando submetidos ao FGENESH para predição estrutural destes genes observou-se que ambos apresentam apenas um éxon, para o primeiro gene o tamanho foi de 252 pb que se inicia no nucleotídeo 69 e termina no 320, codificando um peptídeo de 83 aa. Já o segundo gene apresenta 255 pb, iniciando no nucleotídeo 83 e terminando no 337, codificando um peptídeo de 85 aa, tamanhos similares aos encontrados para os fragmentos zm315, zm316, zm317 e sc539, que apresentou 252, 201, 258 e 288 pb como possível éxon, já que não foi possível a predição pelo FGENESH, quando submetidas, o resultado não apresentou confiabilidade. Jennings *et al.*, (2001) obtiveram resultados semelhantes quando isolaram ciclotídeos de *Oldenlandia affinis* a partir de cDNA com variação de 677 pb a 993 pb, que codificaram sequências com 111 a 210 aminoácidos, enquanto que Mulvena *et al.*, (2006)

relata a descoberta de ciclotídeos na família Poaceae, onde genes codificaram proteínas de aproximadamente 70 aminoácidos, o que corrobora com os dados obtidos neste estudo.

Sugere-se que a alta frequência de aparecimento do domínio Thioredoxin_like superfamily neste estudo é devido à sua relação com a sequência do peptídeo sinal, a qual é direcionado para o RE, já que segundo Lu and Holmgren (2014) proteínas com este domínio, assim como a proteína dissulfeto isomerase (PDI), estão envolvidas no dobramento dessas proteínas. Além disso, envolve a formação de ligações dissulfeto, como já foi relatado sua possível participação em estudo no dobramento de ciclotídeos (Gruber *et al.*, 2007). Tais proteínas desempenham suas funções fisiológicas em compartimentos celulares oxidativos, como o Retículo Endoplasmático (RE) ou o periplasma em bactérias (Lu and Holmgren 2014)

Quando confrontados com ciclotídeos de banco de dados, os baixos índices de similaridade encontrados para algumas das sequências (zm316, zm317 e sc359) podem se justificar devido ao baixo número de ciclotídeos encontrados nesta família (25) no banco de proteínas do NCBI e no banco do Phytozome, já que para a família Violaceae estão disponíveis no banco de dados de proteínas do NCBI até o momento 188 entradas para ciclotídeos.

De acordo com a estrutura primária e conservação dos resíduos de aminoácidos os ciclotídeos se dividem em duas subfamílias, Möbius e a Bracelet (Craik *et al.*, 1999). A subfamília Möbius se caracteriza por uma ligação péptica cis-prolina que ocorre no loop 5, conferindo uma torção na sua estrutura terciária, esta torção que ocorre na fita de Möbius (que foi estudada pelo matemático August Ferdinand Möbius por isso o nome desta subfamília). Na terceira posição logo após a primeira cisteína encontra-se uma treonina e antecedendo a terceira cisteína também se encontra uma treonina, sendo características da família Möbius. Na subfamília Bracelet não ocorre a prolina no loop 5 na posição correspondente em sua sequência primária, e o resíduo encontrado na terceira posição diferente do que foi encontrado

na Möbius é uma serina e na posição que antecede a terceira cisteína é uma prolina. Tais características conservadas tornam as sequências isoladas neste estudo, possíveis membros da subfamília Bracelet (pulseira) (Craig *et al.*, 1999).

Caracterização estrutural

Segundo Nguyen *et al* (2011), a região “sinal”, que corresponde ao peptídeo sinal, é responsável pelo direcionamento da proteína até o lúmen do retículo endoplasmático, para processo de dobramento e subsequentemente introdução na via secretora, [o peptídeo sinal foi encontrado nas análises para cada uma das sequências analisadas, possuindo aproximadamente 29 a 27 aminoácidos], como observado em outros ciclotídeos isolados que tiveram de 26 a 29 aminoácidos (Mulvena *et al.*, 2006; Nguyen *et al.*, 2011). A partir da ferramenta Philius foi possível detectar sítio de clivagem do peptídeo sinal e para onde a proteína é endereçada. Com exceção de zm316 (que não apresentou peptídeo sinal) e de sc359 (o qual mostrou baixa confiabilidade do peptídeo sinal) as proteínas codificadas pelos possíveis genes zm315 e zm316 são secretadas, podendo assim desempenhar o papel de defesa, combatendo patógenos.

A predição estrutural dessas sequências apresentou as regiões típicas de ciclotídeos caracterizados em estudos anteriores, ou seja, uma sequência sinal, um pro-peptídeo N-terminal (NTPP), um domínio ciclotídeos maduro e um pró-peptídeo C-terminal (CTPP) (Dutton *et al.*, 2004; Nguyen, 2011), com exceção do CTPP (como observado na Figura 1), o qual também não foi identificado em algumas sequências de ciclotídeos de membros da família Poaceae (Mulvena *et al.*, 2006; Nguyen *et al.*, 2011).

Embora ainda não se saiba ao certo a função da região pró-peptídeo C-terminal (NTPP), sabe-se que tal região varia tanto em tipo como em número de aminoácidos que a compõem, mesmo entre espécies da mesma família. Por exemplo, essa região pode chegar a

65 resíduos, para Kalata B1 (Jennings *et al.*, 2001), de 15 a 20 resíduos em chassatides (Nguyen *et al.*, 2012), ou até mesmo, não apresentar esta região, como em cliotides (Nguyen *et al.*, 2011b). Para os ciclotídeos isolados neste estudo a região do NTPP variou de 24 a 27 resíduos de aminoácidos, números similares aos encontrados em panitides (ciclotídeos lineares encontrados na espécie *Panicum laxum* da família Poaceae) (Nguyen *et al.*, 2011).

A sequência do peptídeo maduro foi confirmada com ajuda da ferramenta online Smart, sendo composta por 30 a 31 aminoácidos e, portanto, de acordo com o encontrado para outros ciclotídeos, cujo tamanho varia de 28 - 37 resíduos de aminoácidos (Craik *et al.*, 1999). O peptídeo maduro ou domínio é a região que confere à proteína sua função. Quando esta proteína se dobra, este processo é muito importante, pois modifica cadeias lineares de polipeptídeos em estruturas tridimensionais, as quais permitem que as proteínas exerçam todas as suas atividades vitais (Ellis and Hartl 1996)

As cisteínas conservadas, os resíduos de Asn/Asp (na região proto C-terminal) e Ala/Gly (na região proto N-terminal) são igualmente importantes, uma vez que parecem estar envolvidas no processo de síntese destas moléculas (Mylne 2010). Com base nisso, pode-se afirmar que os candidatos isolados zm315 e zm316 apresentam-se como ciclotídeos lineares, uma vez que não apresentam nenhum dos resíduos na região proto C-terminal (Mylne *et al.*, 2010). Ciclotídeos lineares também foram encontrados em *Panicum laxum*, espécie também da família Poaceae (Nguyen *et al.*, 2011). Por outro lado, a presença dos resíduos Asp e Asn, nas sequências zm317 e sc359, respectivamente, na região C-terminal, permite que estes peptídeos sofram ciclização da proteína resultante. De acordo com Gruber *et al.* (2008) tal mudança pode ter ocorrido em decorrência de uma mutação, a qual promoveu a introdução de um resíduo de Asn/ Asp na molécula, permitindo ciclização perto do C-terminal ancestral, permitindo, assim, uma possível a ciclização de proteínas lineares.

O motivo estrutural compartilhado entre os ciclotídeos já descritos é o nó cistina cíclico (do inglês *Cyclic Cystine Knot* - CCK), onde as seis cisteínas conservadas nos ciclotídeos formam ligações estabelecidas da seguinte forma (CysI-CysIV e CysII-CysV), enquanto que a terceira ligação dissulfeto (CysIII-CysVI) penetra nesse anel (CRAIK, 2010), o que confere uma maior estabilidade à estrutura. (Craik *et al.*, 1999). As sequências zm315 e zm316 apresentaram as 3 pontes de dissulfeto típico dos ciclotídeos, uma vez estas são formadas pelas cisteínas nas posições (Cys I-IV), (Cys II-V) e (Cys III-VI) (Craik, 2010). Por outro lado, os outros possíveis ciclotídeos (zm317 e sc359) apenas a primeira ligação entre (CysI-CysIV) se manteve.

Predição de propriedades físico-químicas

A partir do preditor utilizado para propriedades dos possíveis ciclotídeos, foi possível uma análise comparativa entre as propriedades físico-químicas do pró-peptídeo e do domínio. De um modo geral a massa molecular dos ciclotídeos encontra-se por volta dos 3700 Da (Poth *et al.*, 2011), em cicloviolins de A-D as massas variaram de 2887.1 a 3213.4 Da, em ciclotídeos isolados de *Clitoria ternatea* variou de 3071.34 a 3267.49 Da, para o parigidin-br1 isolado de *Policourea rigida* apresentou 3178.1 Da, e em representantes da família Poaceae observou-se variação entre 2500 a 4000 Da (Hallock *et al.*, 2000; Poth *et al.*, 2011; Nguyen *et al.*, 2011; Pinto *et al.*, 2012). A massa molecular encontrada nos possíveis ciclotídeos diminuiu quando analisado apenas seu domínio. O ponto isoelétrico predito variou de 6.10 a 7.79 para análise sem o peptídeo sinal, onde valores similares foram vistos em cicloviolins de A-D que apresentavam variação entre 5.90 a 7.93 (Hallock *et al.*, 2000). Segundo Daly *et al.*, (2009), os ciclotídeos não são altamente carregados positivamente, o que justifica os valores equilibrados encontrados para os dados deste estudo, apenas um resíduo a mais como positivo foi observado.

O índice de GRAVY se apresentou em valores positivos o que revela que os possíveis ciclotídeos cíclicos e lineares são hidrofóbicos. Os ciclotídeos apresentam um trecho de resíduos hidrofóbicos o que, aliado com as diversas atividades a eles relacionadas, indica um modo de ação que envolve atos contra a membrana biológica. Assim como relatado em estudo realizado com larvas de *Spodoptera frugiperda* a ação de parrigidin-br1 como inseticida e sua ação ocorre através da membrana celular, que também apresentou regiões hidrofóbicas (Simonsen *et al.*, 2008; Pinto *et al.*, 2102).

O índice de instabilidade proporciona a estimativa da proteína em um tubo de ensaio, sendo que uma proteína com valor de instabilidade menor que 40 é prevista como estável e quando o valor é acima de 40 é instável (Guruprasad *et al.*, 1990). Com exceção da sc359, todas as outras apresentaram este índice menor que 40. Em relação ao índice alifático, definido como o volume relativo de uma proteína com resíduos de cadeia lateral alifática (A, V, L, I), o mesmo é considerado como fator para aumento da estabilidade térmica de proteínas globulares, onde índices elevados, são considerados como ponto positivo para proteínas, podendo ser estáveis em um amplo campo de temperaturas, atendendo a esse requisito zm315 e zm316 apresentam índices mais elevados em comparação a zm317 e sc359.

A meia vida estimada mostra a estabilidade dos possíveis ciclotídeos, nos ambientes propostos (mamíferos, levedura e bactérias). Apenas zm316 apresentou valores diferentes quando testado, sendo que esta sequência apresenta apenas um fragmento do suposto peptídeo sinal.

Predição de estrutura terciária

Para que seja possível a compreensão do mecanismo e da função de uma proteína, comumente é requerido o conhecimento de sua estrutura tridimensional (Blundell *et al.*, 1978; Weber, 1990). Em ciclovioiacin foi encontrada uma estrutura tridimensional compacta com três folhas beta e um pequeno segmento helicoidal (Craik *et al.*, 1999), o que parece estar nos

modelos gerados neste estudo. O ciclotídeo utilizado como *template* para a modelagem por homologia, apresenta duas folhas beta, formando um β -hairpin, além de apresentar uma volta 3_{10} helicoidal encontrada no loop 3. Por sua vez as estruturas modeladas zm316, zm317 e sc359 apresentaram essa estrutura helicoidal no loop 3. Em parigidin-br1 também foi encontrada a volta 3_{10} helicoidal no loop 3, o que se assemelha também a membros da subfamília Bracelet (Pinto *et al.*, 2012).

Os ciclotídeos isolados neste trabalho apresentaram uma estrutura similar à encontrada em outros ciclotídeos já bem estabelecidos em termos de sua estrutura tridimensional. Como o gráfico de Ramachandran é útil para definição dos resíduos que se encontram nas regiões energeticamente mais favoráveis e desfavoráveis, orientando a avaliação da qualidade de modelos teóricos ou experimentais de proteínas. Assim como visto em Circulina B, 80% dos resíduos na região mais favorável, 17.6% na região permitida adicional, 2.4% na região generosamente permitida e 0% na região não permitida, ainda outro ciclotídeo isolado por Pinto *et al.*, 2012, o paridigin-br1, obteve valores de 84%, 12%, 4% e 0% respectivamente. Os modelos aqui apresentados obtiveram mais de 70% dos resíduos na região mais favorável. Estes dados apontam para o fato de que os modelos propostos apresentam boa qualidade e características confiáveis comuns aos ciclotídeos, podendo ser considerados membros deste grupo de peptídeos antimicrobianos

Agradecimentos

A todos que contribuíram para que este trabalho se realizasse e à FACEPE pelo suporte financeiro.

Referências

- Arnold K, Bordoli L, Kopp J and Schwede T (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22:195-201.
- Benkert P, Schwede T and Tosatto SCE (2009) QMEANclust: Estimation of protein model quality by combining a composite scoring function with structural density information. *BMC Structural Biology*, 20:9-35.
- Brogden NK and Brogden KA (2011) Will new generations of modified antimicrobial peptides improve their potential as pharmaceuticals *Int. J. Antimicrob. Agents*, 38:217-25.
- Cemazar M, Kwon S, Mahatmanto T, Ravipati AS and Craik DJ (2012) Discovery and applications of disulfide-rich cyclic peptides. *Curr Top Med Chem*, 12:1534–1545.
- Craik DJ (2010) Discovery and applications of the plant cyclotides. *Toxicon*, 56:1092-1102.
- Craik DJ (2012) Host-defense activities of cyclotides. *Toxins*, 4:139–156.
- Craik DJ, Anderson MA, Barry DG, Clark RJ, Daly NL, Innings CV and Mulvenna J (2002) Discovery and structures of the cyclotides: novel from plants. *Lett. Pept. Sci.*, 8:119-128.
- Craik DJ, Daly NL, Bond T and Waine C (1999) Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.*, 294:1327–1336
- Craik DJ, Daly NL, Bond T and Waine C (1999) Plant cyclotides: A unique Family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *J. Mol. Biol.* 294:1327–1336.
- Daly NL, Rosenfren KJ and Craik DJ (2009) Discovery, structure and biological activities of cyclotides. *Adv. Drug Deliv. Rev.*, 61:918-930.

- Dutton JL, Renda RF, Waine C, Clark RJ, Daly NL, Jennings CV, Anderson MA and Craik DJ (2004) Conserved structural and sequence elements implicated in the processing of gene-encoded circular proteins. *J. Biol. Chem.* 279:46858–46867.
- Eddy SR (1996) Hidden Markov models. *Curr Opin Struct Biol*, 6:361–365.
- Eddy, SR (2009) A new generation of homology search tools based on probabilistic inference. *Genome informatics. International Conference on Genome Informatics*, v. 23, n. 1, p. 205-11, 2009.
- Ellis RJ, Hartl FU (1996) Protein folding in the cell: competing models for chaperonin function. *FASEB J*, 10:20-6.
- Ferrè F. and Clote P (2005) DiANNA: a web server for disulfide connectivity prediction. *Nucleic Acids Res*, 33, Web Server issue.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD and Bairoch A (2005) Protein Identification and Analysis Tools on the ExPASy Server; In: John M. Walker - *The Proteomics Protocols Handbook*, Humana Press. 571-607.
- Gibson DJ (2009) *Grasses and grassland ecology*. New York: Oxford Univ. Press, 305 pp.
- GPWG (Grass Phylogeny Working Group) (2011) New grass phylogeny resolves deep evolutionary relationships and discovers C4 origins. *New Phytologist* 193:304-312.
- Gruber CW et al (2007) A novel plant protein-disulfide isomerase involved in the oxidative folding of cystine knot defense proteins. *J Biol Chem*, 282:20435-20446.
- Gruber CW et al (2008) Distribution and Evolution of Circular Miniproteins in Flowering Plants. *The plant cell*. 20:2471–2483.
- Guruprasad K, Reddy BV, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng Des Sel* 4:155–161.

- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser*, 41:95-98.
- Hallock YF, Sowder RC, Pannell LK, Hughes CB, Johnson DG, Gulakowski R, Cardellina JH and Boyd MR (2000) Cycloviolins AD, Anti-HIV Macrocyclic Peptides from *Leonia cymosa*. *J. Org. Chem.*, 65:124-128.
- Jennings C, West J, Waine C, Craik D and Anderson M (2001) Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from *Oldenlandia affinis*. *Proc. Natl. Acad. Sci. U.S.A.* 98:10614–10619.
- Kaur H and Raghava GPS (2004) Prediction of α -turns in proteins using PSI-BLAST profiles and secondary structure information. *Proteins: Struct., Funct., Bioinf.* 55:83–90.
- Laskowski RA, Macarthur MW, Moss DS and Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr*, 26:283-291.
- Letunic I, Doerks T and Bork P (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res*, 43:D257-D260.
- Lu J and Holmgren A (2014). The thioredoxin superfamily in oxidative protein folding. *Antioxidants & redox signaling*, 21:457-470.
- Marchler-Bauer A et al. (2011), "CDD: a Conserved Domain Database for the functional annotation of proteins.", *Nucleic Acids Res.* 39:225-9.
- Michaels SD, John MC and Amasino RM (1994) Removal of polysaccharides from plant DNA by ethanol precipitation. *Biotechniques* 17:274-276.
- Miranda-Peixoto C (2002) O milho: o rei dos cereais – da sua descoberta há 8000 anos até as plantas transgênicas. *Seed News.*, Disponível em <http://www.seednews.inf.br/portugues/seed62/milho62.shtml>.

- Mori C, Nascimento A and Miranda MZ (2013) Aspectos econômicos e conjunturais da cultura do centeio no mundo e no Brasil. ISSN 1518-6512.
- Mulvenna JP, Mylne JS, Bharathi R, Burton RA, Shirley NJ, Fincher GB, Anderson MA and Craik DJ (2006) Discovery of cyclotide-like protein sequences in graminaceous crop plants: ancestral precursors of circular proteins?. *Plant Cell*, 18:2134-2144.
- Mylne JS, Wang CK, van der Weerden NL and Craik DJ (2010) Cyclotides are a component of the innate defense of *Oldenlandia affinis*. *Peptide Science*, 94:635–646.
- Nguyen GK, Lim WH, Nguyen PQ, Tam JP (2012) Novel cyclotides and uncyclotides with highly shortened precursors from *Chassalia chartacea* and effects of methionine oxidation on bioactivities. *J Biol Chem*. 287:17598-607.
- Nguyen GK, Zhang S, Wang W, Wong CT, Nguyen NT and Tam JP (2011) Discovery of a linear cyclotide from the bracelet subfamily and its disulfide mapping by top-down mass spectrometry. *J. Biol. Chem*. 286:44833–44844.
- Nguyen, G.K., Zhang, S., Nguyen, N. T., Nguyen, P. Q., Chiu, M. S., Hardjojo, A., and Tam, J. P. (2011b) Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the Fabaceae family. *J. Biol. Chem*. 286, 24275–24287.
- Petersen TN, Brunak S, von Heijne G and Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods*, 8:785-786.
- Peterson PM (2005) Grasses: family Poaceae. pp. 104-108. In: KRUPNICK, G.A.
- Pinto MF, Fensterseifer IC, Migliolo L, Sousa DA, de Capdville G, Arboleda-Valencia JW, Colgrave ML, Craik DJ, Magalhães BS, Dias SC and Franco OL. (2012) Identification and structural characterization of novel cyclotide with activity against an insect pest of sugar cane. *J Biol Chem*. 287:134-47.

- Pinto MFS, Almeida RG, Porto WF, Fensterseifer ICM, Lima LA, Dias SC and Franco OL (2011) Cyclotides: From Gene Structure to Promiscuous Multifunctionality. *J Evid Based Complementary Altern Med*, 17:40–53.
- Pomilio AB, Battista ME and Vitale AA (2006) Naturally-Occurring Cyclopeptides: Structures and Bioactivity. *Curr Org Chem*, 10:2075-2121.
- Poth AG, Colgrave ML, Philip R, Kerenga B, Daly NL, Anderson MA and Craik DJ (2011) Discovery of cyclotides in the Fabaceae plant family provides new insights into the cyclization, evolution, and distribution of circular proteins. *ACS Chem. Biol.*, 6:345-355.
- Reynolds SM, Käll L, Riffle ME, Bilmes JÁ and Noble WS (2008) Transmembrane Topology and Signal Peptide Prediction Using Dynamic Bayesian Networks. *PLoS Computational Biology*, 4, e1000213.
- Salla DA, Furlaneto FPB, Cabello C and Kanthack RAD (2010) Estudo energético da produção de biocombustível a partir do milho. *Ciência Rural*, 40:2017-2022.
- Santos AO and Prado H (2002) Análise de interações solo-planta-clima em zonas diferenciadas de área de cultivo de milho, *Rev. bras. eng. agríc. Ambiente*. 6:101–106.
- Santos Filho OA and Alencastro RBD (2003). Modelagem de proteínas por homologia. *Química Nova*, 26:253-259.
- Schmidt R and Longhi-Wagner HM (2009) A tribo Bambuseae (Poaceae-Bambusoideae) no Rio Grande do Sul, Brasil. *R. bras. Bioci.*,7:71-128.
- Sievers F, Wilm A, Dineen D, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
- Simonsen SM, Sando L, Rosengren KJ, Wang CK, Colgrave ML, Daly NL and Craik DJ (2008) Alanine scanning mutagenesis of the prototypic cyclotide reveals a cluster of residues essential for bioactivity, *J. Biol. Chem.* 283:9805–9813.

Souza VC and Lorenzi H (2012) Botânica sistemática: guia ilustrado para a identificação das famílias de fanerógamas nativas e exóticas no Brasil, baseado em APG III. Instituto Plantarum. Nova Odessa, São Paulo.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M and Rozen SG (2012) Primer3—new capabilities and interfaces. *Nucleic Acids Res*, 40, e115-e115.

USDA (United States Department of Agriculture) disponível em: http://az545403.vo.msecnd.net/uploads/2014/12/boletim_safra-mundial-milho_dezembro2014.pdf

Weising K, Nybom H, Wolff K and Kahl G (2005) DNA fingerprinting in plants: principles, methods, and applications. Second edition. CRC Press, Boca Raton, US.

Recursos da internet

Alphapred, <http://www.imtech.res.in/raghava/alphapred/index.html> (22 de dezembro de 2014)

Banco de dados de proteínas do NCBI, <http://www.ncbi.nlm.nih.gov/protein> (12 de novembro de 2014)

Batch, <http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi> (11 de setembro de 2014)

CD Search, <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> (19 de novembro de 2014)

fgenesh (HMM-based gene structure prediction) <http://linux1.softberry.com/berry.phtml> (28 de dezembro de 2013)

GenBank, <http://blast.ncbi.nlm.nih.gov/Blast.cgi>, (12 de novembro de 2014)

HMMER3.1, <http://hmmer.janelia.org/> (26 de novembro de 2013)

JCVI - J. Craig Venter Institute, <http://www.jcvi.org/cms/home/> (28 de novembro de 2013)

Jmol 13.0, <http://jmol.sourceforge.net/> (24 de dezembro de 2014)

NordEST Project, <http://bioinfo03.ibi.unicamp.br/vigna/> (28 de novembro de 2013)

ORF-finder, <http://www.ncbi.nlm.nih.gov/projects/gorf/> (11 de setembro de 2014)

Philius Transmembrane Prediction Server,

<http://www.yeastrc.org/philius/pages/philius/runPhilius.jsp> (17 de novembro de 2014)

Phytozome; <http://www.phytozome.net/> (28 de novembro de 2013)

Primer3, <http://frodo.wi.mit.edu/primer3/> (30 de junho de 2014)

SWISS-MODEL Workspace, <http://swissmodel.expasy.org/> (22 de dezembro de 2014)

V. CONCLUSÕES GERAIS

- As análises *in silico* realizadas neste estudo através do HMMER revelam que nos bancos de dados de nucleotídeos e de proteínas das espécies de *Vigna unguiculata*, *Sorghum bicolor*, *Hevea brasiliensis*, *Manihot esculenta* e *Ricinus communis* não existe, até o presente estudo, nenhuma sequência similar a ciclotídeos;
- Neste estudo foi possível o isolamento e a caracterização de quatro fragmentos gênicos de sequências homólogas a ciclotídeos, sendo três de milho e uma de centeio;
- As quatro sequências caracterizadas neste estudo revelam semelhanças em suas estruturas primária, secundária e terciária, bem como semelhanças funcionais, além de resíduos conservados em posições relevantes para caracterização de suas subfamílias e atividades antimicrobianas semelhantes às de ciclotídeos;
- Os peptídeos zm315 e zm316 parecem estar dentre os ciclotídeos lineares já que em suas sequências não foram encontrados na região C-terminal nenhum dos resíduos Asn/Asp, sendo estes encontrados em zm317 e sc359;
- Todas as sequências identificadas apresentam características que indicam que as mesmas pertencem à subfamília Bracelet dos ciclotídeos;
- Os resultados obtidos neste trabalho podem contribuir para o conhecimento a respeito dos ciclotídeos e suas características, nas espécies de importância agrícola, deixando claro que mais estudos devem ser realizados para o melhor aproveitamento destas moléculas, para fins biotecnológicos.

VI. ANEXOS

Anexo 1 : Instruções para autores: *Current Protein & Peptide Science*

ONLINE MANUSCRIPT SUBMISSION:

An online submission and tracking service *via* Internet facilitates a speedy and cost-effective submission of manuscripts. The full manuscript has to be submitted online via Bentham's Content Management System (CMS) at bsp-cms.eurekaselect.com / [View Submission Instructions](#).

Manuscripts must be submitted by one of the authors of the manuscript, and should not be submitted by anyone on their behalf. The principal/corresponding author will be required to submit a Copyright Letter along with the manuscript, on behalf of all the co-authors (if any). The author(s) will confirm that the manuscript (or any part of it) has not been published previously or is not under consideration for publication elsewhere. Furthermore, any illustration, structure or table that has been published elsewhere must be reported, and copyright permission for reproduction must be obtained.

For all online submissions, please provide soft copies of all the materials (main text in MS Word or Tex/LaTeX), figures / illustrations in TIFF, PDF or JPEG, and chemical structures drawn in ChemDraw (CDX) / ISISDraw (TGF) as separate files, while a PDF version of the entire manuscript must also be included, embedded with all the figures / illustrations / tables / chemical structures etc. It is advisable that the document files related to a manuscript submission should always have the name of the corresponding author as part of the file name, i.e., "Cilli MS text.doc", "Cilli MS Figure 1", etc.

It is imperative that before submission, authors should carefully proofread the files for special characters, mathematical symbols, Greek letters, equations, tables, references and images, to ensure that they appear in proper format.

References, figures, tables, chemical structures *etc.* should be referred to in the text at the appropriate place where they have been first discussed. Figure legends/captions should also be provided.

A successful electronic submission of a manuscript will be followed by a system-generated acknowledgement to the principal/corresponding author. Any queries therein should be addressed to ilyas@benthamscience.org

Editorial Policies:

The editorial policies of Bentham Science Publishers on publication ethics, peer-review, plagiarism, copyrights/licenses, errata/corrections, and article retraction/ withdrawal can be viewed at [Editorial Policy](#)

MANUSCRIPTS PUBLISHED:

The Journal accepts mini- and full-length review articles written in English. Single topic/thematic issues may also be considered for publication.

Single Topic Issues:

These special issues are peer-reviewed and may contain invited or uninvited review/mini-review articles. Single Topic Special Editor will offer a short Perspective and coordinate the solicitation of manuscripts between 5-8 (for a mini-thematic issue) to 8-12 (for full-length thematic issue) from leading scientists. Authors interested in editing a single topic issue in an emerging topic of protein and peptide science may submit their proposal to the Editor-in-Chief at cpaps@benthamscience.org for consideration. Visit www.benthamscience.com/cpaps/Special-Issues.htm for viewing forthcoming issues.

Conference Proceedings:

For proposals to publish conference proceedings in this journal, please contact us at email: proceedings@benthamscience.org

MANUSCRIPT LENGTH:

Review Articles:

The total number of words for a published comprehensive review article is from 8000 to 40000 words, and for mini-review articles from 3000 to 6000 words.

There is no restriction on the number of figures, tables or additional files e.g. video clips, animation and datasets, that can be included with each article online. Authors should include all relevant supporting data with each article (Refer to Supplementary Material section).

MANUSCRIPT PREPARATION:

The manuscript should be written in English in a clear, direct and active style. All pages must be numbered sequentially, facilitating in the reviewing and editing of the manuscript.

MICROSOFT WORD TEMPLATE:

It is advisable that authors prepare their manuscript using the template available on the Web, which will assist in preparation of the manuscript according to Journal's Format. [Download the Template.](#)

Our contracted service provider [Eureka Science](#) can, if needed, provide professional assistance to authors for the improvement of English language and figures in manuscripts.

MANUSCRIPT SECTIONS FOR PAPERS:

Manuscripts may be divided into the following sections:

- Copyright Letter
- Title
- Title page
- Abstract
- Graphical Abstract
- Keywords
- Text Organization
- List of Abbreviations (if any)
- Conflict of Interest
- Acknowledgements
- References
- Appendices
- Figures/Illustrations (if any)
- Chemical Structures (if any)
- Tables (if any)
- Supportive/Supplementary Material (if any)

Copyright Letter:

It is mandatory that a signed copyright letter also be submitted along with the manuscript by the author to whom correspondence is to be addressed, delineating the scope of the submitted article declaring the

potential competing interests, acknowledging contributions from authors and funding agencies, and certifying that the paper is prepared according to the 'Instructions for Authors'. All inconsistencies in the text and in the reference section, and any typographical errors must be carefully checked and corrected before the submission of the manuscript. The article should not contain any such material or information that may be unlawful, defamatory, fabricated, plagiarized, or which would, if published, in any way whatsoever, violate the terms and conditions as laid down in the copyright agreement. The authors acknowledge that the publishers have the legal right to take appropriate action against the authors for any such violation of the terms and conditions as laid down in the copyright agreement. Copyright letter can be downloaded from the journal's Web site. [Download the Copyright letter](#)

Title:

The title of the article should be precise and brief and must not be more than 120 characters. Authors should avoid the use of non-standard abbreviations. The title must be written in title case except for articles, conjunctions and prepositions.

Authors should also provide a short 'running title'. Title, running title, byline, correspondent footnote and keywords should be written as presented in original manuscripts.

Title Page:

Title page should include paper title, author(s) full name and affiliation, corresponding author(s) names complete affiliation/address, along with phone, fax and email.

Abstract:

The abstract should not exceed 250 words for review papers summarizing the essential features of the article.

Graphical Abstract:

A graphic must be included with each manuscript for use in the Table of Contents (TOC). This must be submitted separately as an electronic file (preferred file types are EPS, PDF, TIFF, Microsoft Word, PowerPoint and CDX etc.). A graphical abstract, not exceeding 30 words along with the illustration, helps to summarize the contents of the manuscript in a concise pictorial form. It is meant as an aid for the rapid viewing of the journals' contents and to help capture the readers' attention. The graphical abstract may feature a key structure, reaction, equation, etc. that the manuscript elucidates upon. It will be listed along with the manuscript title, authors' names and affiliations in the contents page, typeset within an area of 5 cm by 17 cm, but it will not appear in the article PDF file or in print.

Graphical Abstracts should be submitted as a separate file (must clearly mention graphical abstract within the file) online *via* Bentham's Content Management System by selecting the option "supplementary material".

Keywords:

Provide 6 to 8 keywords in alphabetical order.

Text Organization:

The main text should begin on a separate page and should be divided into separate sections. The manuscript should be divided into title page, abstract and the main text. The text may be subdivided further according to the areas to be discussed, which should be followed by the Acknowledgements and Reference sections. The review article should mention any previous important recent and old reviews in the field and contain a comprehensive discussion starting with the general background of the field. It should then go on to discuss the salient features of recent developments. The authors should avoid presenting material which has already been

published in a previous review. The authors are advised to present and discuss their observations in brief. The manuscript style must be uniform throughout the text and 10 pt Times New Roman fonts should be used. The full term for an abbreviation should precede its first appearance in the text unless it is a standard unit of measurement. The reference numbers should be given in square brackets in the text. Italics should be used for Binomial names of organisms (Genus and Species), for emphasis and for unfamiliar words or phrases. Non-assimilated words from Latin or other languages should also be italicized e.g. *per se*, *et al.* etc.

Standard Protocol on Approvals, Registrations, Patient Consents & Animal Protection:

All clinical investigations must be conducted according to the Declaration of Helsinki principles. For all manuscripts reporting data from studies involving human participants, formal review and approval by an appropriate institutional review board or ethics committee is required. For research involving animals, the authors should indicate whether the procedures followed were in accordance with the standards set forth in the eighth edition of *Guide for the Care and Use of Laboratory Animals* (grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals_prepub.pdf); published by the National Academy of Sciences, The National Academies Press, Washington, D.C.).

A specific declaration of such approval must be made in the copyright letter and in a stand-alone paragraph at the end of the Methods section especially in the case of human studies where inclusion of a statement regarding obtaining the written informed consent from each subject or subject's guardian is a must. The original should be retained by the guarantor or corresponding author. Editors may request to provide the original forms by fax or email.

Greek Symbols and Special Characters:

Greek symbols and special characters often undergo formatting changes and get corrupted or lost during preparation of manuscript for publication. To ensure that all special characters used are embedded in the text, these special characters should be inserted as a symbol but should not be a result of any format styling (*Symbol* font face) otherwise they will be lost during conversion to PDF/XML.

Authors are encouraged to consult reporting guidelines. These guidelines provide a set of recommendations comprising a list of items relevant to their specific research design. Chemical equations, chemical names, mathematical usage, unit of measurements, chemical and physical quantity & units must conform to SI and Chemical Abstracts or IUPAC.

All kinds of measurements should be reported only in International System of Units (SI).

List of Abbreviations:

If abbreviations are used in the text either they should be defined in the text where first used, or a list of abbreviations can be provided.

Conflict of Interest:

Financial contributions to the work being reported must be clearly acknowledged, as should any potential conflict of interest under the heading 'Conflict of Interest'. Authors must list the source(s) of funding for the study, for each author, and for the manuscript preparation.

Acknowledgements:

All individuals listed as authors must have contributed substantially to the design, performance, analysis, or reporting of the work and are required to indicate their specific contribution. Anyone

(individual/company/institution) who has substantially contributed to the study for important intellectual content, or who was involved in the article's drafting the manuscript or revising must also be acknowledged.

Guest or honorary authorship based solely on position (e.g. research supervisor, departmental head) is discouraged.

The specific requirements for authorship have been defined by the International Committee of Medical Journal Editors (ICMJE; www.icmje.org). Examples of authors' contributions are: 'designed research/study', 'performed research/study', 'contributed important reagents', 'collected data', 'analyzed data', 'wrote paper' etc. This information must be included in the submitted manuscript as a separate paragraph under the heading 'Acknowledgements'. The corresponding author is responsible for obtaining permission from all co-authors for the submission of any version of the manuscript and for any changes in the authorship.

References:

References must be listed in the ACS Style only. All references should be numbered sequentially [in square brackets] in the text and listed in the same numerical order in the reference section. The reference numbers must be finalized and the bibliography must be fully formatted before submission.

See below few examples of references listed in the ACS Style:

Journal Reference:

- [1] Bard, M.; Woods, R.A.; Bartón, D.H.; Corrie, J.E.; Widdowson, D.A. Sterol mutants of *Saccharomyces cerevisiae*: chromatographic analyses. *Lipids*, **1977**, *12*(8), 645-654.
- [2] Zhang, W.; Brombosz, S.M.; Mendoza, J.L.; Moore, J.S. A high-yield, one-step synthesis of o-phenylene ethynylene cyclic trimer *via* precipitation-driven alkyne metathesis. *J. Org. Chem.*, **2005**, *70*, 10198-10201.

Book Reference:

- [3] Crabtree, R.H. *The Organometallic Chemistry of the Transition Metals*, 3rd ed.; Wiley & Sons: New York, **2001**.

Book Chapter Reference:

- [4] Wheeler, D.M.S.; Wheeler, M.M. In: *Studies in Natural Products Chemistry*; Atta-ur-Rahman, Ed.; Elsevier Science B. V: Amsterdam, **1994**; Vol. *14*, pp. 3-46.

Conference Proceedings:

- [5] Jakeman, D.L.; Withers, S.G.E. In: *Carbohydrate Bioengineering: Interdisciplinary Approaches*, Proceedings of the 4th Carbohydrate Bioengineering Meeting, Stockholm, Sweden, June 10-13, 2001; Teeri, T.T.; Svensson, B.; Gilbert, H.J.; Feizi, T., Eds.; Royal Society of Chemistry: Cambridge, UK, **2002**; pp. 3-8.

URL(WebPage):

- [6] National Library of Medicine. Specialized Information Services: Toxicology and Environmental Health. sis.nlm.nih.gov/Tox/ToxMain.html (Accessed May 23, **2004**).

Patent:

- [7] Hoch, J.A.; Huang, S. Screening methods for the identification of novel antibiotics. U.S. Patent 6,043,045, March 28, **2000**.

Thesis:

- [8] Mackel, H. *Capturing the Spectra of Silicon Solar Cells*. PhD Thesis, The Australian National University: Canberra, December **2004**.

E-citations:

- [9] Citations for articles/material published exclusively online or in open access (free-to-view) , must contain the exact Web addresses (URLs) at the end of the reference(s), except those posted on an author's Web site unless editorially essential, e.g. 'Reference: Available from: URL'.

Some important points to remember:

- All references must be complete and accurate.
- All authors must be cited and there should be no use of the phrase *et al*.
- Online citations should include the date of access.
- Journal abbreviations should follow the *Index Medicus/MEDLINE*.
- Special care should be taken of the punctuation convention as described in the above-mentioned examples.
- Superscript in the in-text citations and reference section should be avoided.
- Abstracts, unpublished data and personal communications (which can only be included if prior permission has been obtained) should not be given in the reference section but they may be mentioned in the text and details provided as footnotes.
- The authors are encouraged to use a recent version of EndNote (version 5 and above) or Reference Manager (version 10) when formatting their reference list, as this allows references to be automatically extracted.

Appendices:

In case there is a need to present lengthy, but essential methodological details, appendices must be used, which can be a part of the article. An appendix must not exceed three pages (Times New Roman, 12 point fonts, 900 max. words per page). The information should be provided in a condensed form, ruling out the need of full sentences. A single appendix should be titled APPENDIX, while more than one can be titled APPENDIX A, APPENDIX B, and so on.

Figures/Illustrations:

All authors must strictly follow the guidelines below for preparing illustrations for publication in **Current Protein & Peptide Science**. If the figures are found to be sub-standard, then the manuscripts will be rejected and the authors offered the option of figure improvement professionally by [Eureka Science](#). The costs for such improvement will be charged to the authors.

The authors should provide the illustrations as separate files, as well as embedded in the text file, numbered consecutively in the order of their appearance. Each figure should include a single illustration. Each figure should be closely cropped to minimize the amount of white space surrounding the illustration.

If a figure consists of separate parts, it is important that a single composite illustration file be submitted, containing all parts of the figure.

Photographs should be provided with a scale bar if appropriate, as well as high-resolution component files.

Scaling/Resolution:

For Line Art image type, which is generally an image based on lines and text and does not contain tonal or shaded areas, the preferred file format is TIFF or EPS, with colour mode being Monochrome 1-bit or RGB, in a resolution of 900-1200 dpi.

For Halftone image type, which is generally a continuous tone photograph and contains no text, the preferred file format is TIFF, with colour mode being or RGB or Grayscale, in a resolution of 300 dpi.

For Combination image type, which is generally an image containing halftone in addition to text or line art elements, the preferred file format is TIFF, with colour mode being or RGB or Grayscale, in a resolution of 500-900 dpi.

Formats:

For illustrations, the following file formats are acceptable:

- **Illustrator**
- **EPS** (preferred format for diagrams)
- **PDF** (also especially suitable for diagrams)
- **PNG** (preferred format for photos or images)
- **Microsoft Word** (version 5 and above; figures must be a single page)
- **PowerPoint** (figures must be a single page)
- **TIFF**
- **JPEG** (conversion should be done using the original file)
- **BMP**
- **CDX** (ChemDraw)
- **TGF** (ISISDraw)

Bentham Science does not process figures submitted in GIF format.

If the large size of TIFF or EPS figures acts as an obstacle to online submission, authors may find that conversion to JPEG format before submission results in significantly reduced file size and upload time, while retaining acceptable quality. JPEG is a 'lossy' format, however in order to maintain acceptable image quality, it is recommended that JPEG files are saved at High or Maximum quality.

Files should not be compressed with tools such as Zipit or Stuffit prior to submission as these tools will in any case produce negligible file-size savings for JPEGs and TIFFs, which are already compressed.

Please do not:

1. Supply embedded graphics in your word processor (spreadsheet, presentation) document;
2. Supply files that are optimized for screen use (like GIF, BMP, PICT, WPG); the resolution is too low;
3. Supply files that are too low in resolution;
4. Submit graphics that are disproportionately large for the content.

Image Conversion Tools:

There are many software packages, many of them freeware or shareware, capable of converting to and from different graphics formats, including PNG.

Good general tools for image conversion include GraphicConverter on the Macintosh, PaintShop Pro, for Windows, and ImageMagick, which is available on Macintosh, Windows and UNIX platforms.

Note that bitmap images (e.g. screenshots) should not be converted to EPS, since this will result in a much larger file size than the equivalent JPEG, TIFF, PNG or BMP, with no increase in quality. EPS should only be used for images produced by vector-drawing applications such as Adobe Illustrator or CorelDraw. Most vector-drawing applications can be saved in, or exported as, EPS format. In case the images have been originally prepared in an Office application, such as Word or PowerPoint, then the original Office files should be directly uploaded to the site, instead of being converted to JPEG or another format that may be of low quality.

Color Figures/Illustrations:

- The cost for each individual page of color figures/plates/illustrations is **US\$ 950**.
- Color figures should be supplied in CMYK not RGB colors.

Chemical Structures:

Chemical structures must be prepared in ChemDraw/CDX and provided as separate file.

Structure Drawing Preferences:

[As according to the ACS style sheet]

Drawing Settings:

Chain angle	120°
Bond spacing	18% of width
Fixed length	14.4 pt (0.500cm, 0.2in)
Bold width	2.0 pt (0.071cm, 0.0278in)
Line width	0.6 pt (0.021cm, 0.0084in)
Margin width	1.6 pt (0.096cm)
Hash spacing	2.5 pt (0.088cm, 0.0347in)

Text settings:

Font	Times New Roman / Helvetica
Size	8 pt

Under the Preference Choose:

Units	Points
Tolerances	3 pixels

Under Page Setup Use:

Paper	US letter
-------	-----------

Scale 100%

Tables:

- Data Tables should be submitted in Microsoft Word table format.
- Each table should include a title/caption being explanatory in itself with respect to the details discussed in the table. Detailed legends may then follow.
- Table number in bold font *i.e.* Table **1**, should follow a title. The title should be in small case with the first letter in caps. A full stop should be placed at the end of the title.
- Tables should be embedded in the text exactly according to their appropriate placement in the submitted manuscript.
- Columns and rows of data should be made visibly distinct by ensuring that the borders of each cell are displayed as black lines.
- Tables should be numbered in Arabic numerals sequentially in order of their citation in the body of the text.
- If a reference is cited in both the table and text, please insert a lettered footnote in the table to refer to the numbered reference in the text.
- Tabular data provided as additional files can be submitted as an Excel spreadsheet.

Supportive/Supplementary Material:

We do encourage to append supportive material, for example a PowerPoint file containing a talk about the study, a PowerPoint file containing additional screenshots, a Word, RTF, or PDF document showing the original instrument(s) used, a video, or the original data (SAS/SPSS files, Excel files, Access Db files etc.) provided it is inevitable or endorsed by the journal's Editor.

Supportive/Supplementary material intended for publication must be numbered and referred to in the manuscript but should not be a part of the submitted paper. In-text citations as well as a section with the heading "Supportive/Supplementary Material" before the "References" section should be provided. Here, list all Supportive/Supplementary Material and include a brief caption line for each file describing its contents.

Any additional files will be linked to the final published article in the form supplied by the author, but will not be displayed within the paper. They will be made available in exactly the same form as originally provided only on our Web site. Please also make sure that each additional file is a single table, figure or movie (please do not upload linked worksheets or PDF files larger than one sheet). Supportive/ Supplementary material must be provided in a single zipped file not larger than 4 MB.

Authors must clearly indicate if these files are not for publication but meant for the reviewers'/editors' perusal only.

PERMISSION FOR REPRODUCTION:

Bentham Science has collaborated with the Copyright Clearance Center to meet our customer's licensing, besides rights & permission needs.

The Copyright Clearance Center's RightsLink[®] service makes it faster and easier to secure permission from Bentham Science's journal titles. Simply visit [Journals by Title](#) and locate the desired content. Then go to the article's abstract and click on "Rights and Permissions" to open the RightsLink's page. If you are unable to locate the content you wish to use or you are unable to secure the rights you are seeking, please e-mail us at permissions@benthamscience.org

Published/reproduced material should not be included unless written permission has been obtained from the copyright holder, which should be forwarded to the Editorial Office in case of acceptance of the article for publication.

AUTHORS AND INSTITUTIONAL AFFILIATIONS:

The author will be required to provide their full names, the institutional affiliations and the location, with an asterisk in front of the name of the principal/corresponding author. The corresponding author(s) should be designated and their complete address, business telephone and fax numbers and e-mail address must be stated to receive correspondence and galley proofs.

PAGE CHARGES:

No page charges will be levied to authors for the publication of their review articles.

LANGUAGE AND EDITING:

Manuscripts submitted containing many English typographical errors will not be published. Manuscripts which are accepted for publication on condition that the written English submitted is corrected, will be sent a quote by Eureka Science, a professional language editing company. Authors from non-English language countries who have poor English language written skills, are advised to contact the language editing company prior to submitting their manuscript to the journal. Please contact Eureka Science for a language editing quote at e-mail: info@eureka-science.com stating the total number of words of the article to be edited.

언어 및 편집:

영문 오타가 많은 원고는 출판되지 않을 것입니다. 영문 오타를 없애겠다는 조건으로 받은 원고는 영어 편집 전문회사인 유럽 공동 기술개발 기구로부터 가격 견적서가 보내 질 것입니다. 영어 작문에 어려움이 있는 비영어권 국가의 저자들은 원고를 학술지에 제출하기 전에 영어 편집회사와 접촉할 것을 권합니다. 영어 편집 견적서를 받기 위해서 교정될 원고의 단어수를 적은 메일을 유럽 공동 기술개발 기구 메일인 info@eureka-science.com 로 보내시기 바랍니다.

语言和编辑:

含有很多英文印刷错误的提交稿将不予发表。接受发表的稿件其英文写作应是正确的；专业的语言编辑公司（尤里卡科学），可对稿件的英文润色提供报价。建议非英语国家、且英文写作欠佳的作者在投稿前先与语言编辑公司联系。请与尤里卡科学联系 info@eureka-science.com。

EDITION ET LANGUE:

Les manuscrits soumis avec plusieurs erreurs typographiques en Anglais ne seront pas publiés en l'état. Les manuscrits sont acceptés pour publication à la condition que l'anglais utilisé soit corrigé après la soumission et seront envoyés pour examen à Eureka Science, une société d'édition de langue professionnelle. Les auteurs en provenance de pays où la langue est différente de l'anglais et qui ont de médiocres compétences en anglais écrit, sont priés de contacter la société d'édition de langue avant de soumettre leur manuscrit à la revue. Merci de contacter Eureka Science à info@eureka-science.com pour un devis en indiquant le nombre total de mot de l'article à éditer.

PROOF CORRECTIONS:

Authors will receive page proofs of their accepted paper before publications. To avoid delays in publication, proofs should be checked immediately for typographical errors and returned **within 48 hours**. Major changes are not acceptable at the proof stage. If unable to send corrections **within 48 hours** due to some reason, the author(s) must at least send an acknowledgement on receiving the galley proofs or the article will be published

exactly as received and the publishers will not be responsible for any error occurring in the published manuscript in this regard.

The corresponding author will be solely responsible for ensuring that the revised version of the manuscript incorporating all the submitted corrections receives the approval of all the co-authors of the manuscript.

REPRINTS:

Printed reprints and e-prints may be ordered from the Publisher prior to publication of the article. First named authors may also order a personal print and online subscription of the journal at 50% off the normal subscription rate by contacting the subscription department at e-mail: subscriptions@benthamscience.org

OPEN ACCESS PLUS:

Accepted articles can be published online for free open access for all to view. Open access publishing provides the maximum dissemination of the article to the largest audience. Authors must pay for this service. All authors will be asked to indicate whether or not they wish to pay to have their paper made freely available on publication. If authors do not select the "Open Access Plus" option, then their article will be published with standard subscription-based access at no charge, subject always to the applicable Standard Terms & Conditions and Copyright Letter signed by the author.

Bentham Science offers authors the choice of "Open Access Plus" publication of their articles at a fee of **US\$ 2,900** per published article.

All editors, board members and those authors who have contributed more than two articles in Bentham Science publications are entitled to a **40% discount** on "Open Access Plus" fees.

For more information please contact us at e-mail: openaccess@benthamscience.org.

REVIEWING AND PROMPTNESS OF PUBLICATION:

All manuscripts submitted for publication will be immediately subjected to peer-reviewing, usually in consultation with the members of the Editorial Advisory Board and a number of external referees. Authors may, however, provide in their Copyright Letter the contact details (including e-mail addresses) of four potential peer reviewers for their paper. Any peer reviewers suggested should not have recently published with any of the authors of the submitted manuscript and should not be members of the same research institution.

All peer-reviewing will be conducted *via* the Internet to facilitate rapid reviewing of the submitted manuscripts. Every possible effort will be made to assess the manuscripts quickly with the decision being conveyed to the authors in due course.

COPYRIGHT:

Authors who publish in Bentham Science print & online journals will transfer copyright to their work to **Bentham Science Publishers**. Submission of a manuscript to the respective journals implies that all authors have read and agreed to the content of the Copyright Letter or the Terms and Conditions. It is a condition of publication that manuscripts submitted to this journal have not been published and will not be simultaneously submitted or published elsewhere. Plagiarism is strictly forbidden, and by submitting the article for publication the authors agree that the publishers have the legal right to take appropriate action against the authors, if plagiarism or fabricated information is discovered. By submitting a manuscript the authors agree that the copyright of their article is transferred to the publishers if and when the article is accepted for publication. Once submitted to the journal, the author will not withdraw their manuscript at any stage prior to publication.

SELF-ARCHIVING

By signing the Copyright Letter the authors retain the rights of self-archiving. Following are the important features of self-archiving policy of Bentham Science journals:

1. Authors can deposit the first draft of a submitted article on their personal websites, their institution's repositories or any non-commercial repository for personal use, internal institutional use or for permitted scholarly posting.
2. Authors may deposit the ACCEPTED VERSION of the peer-reviewed article on their personal websites, their institution's repository or any non-commercial repository such as PMC, arXiv after **12 MONTHS of publication** on the journal website. In addition, an acknowledgement must be given to the original source of publication and a link should be inserted to the published article on the journal's/publisher's website.
3. If the research is funded by NIH, Wellcome Trust or any other Open Access Mandate, authors are allowed the archiving of published version of manuscripts in an institutional repository after the mandatory embargo period. Authors should first contact the Editorial Office of the journal for information about depositing a copy of the manuscript to a repository. Consistent with the copyright agreement, Bentham Science does not allow archiving of FINAL PUBLISHED VERSION of manuscripts.
4. The link to the original source of publication should be provided by inserting the DOI number of the article in the following sentence: "The published manuscript is available at EurekaSelect via <http://www.eurekaselect.com/>[insert DOI]."
5. There is no embargo on the archiving of articles published under the OPEN ACCESS PLUS category. Authors are allowed deposition of such articles on institutional, non-commercial repositories and personal websites immediately after publication on the journal website.

PLAGIARISM PREVENTION:

Bentham Science Publishers uses the iThenticate software to detect instances of overlapping and similar text in submitted manuscripts. iThenticate software checks content against a database of periodicals, the Internet, and a comprehensive article database. It generates a similarity report, highlighting the percentage overlap between the uploaded article and the published material. Any instance of content overlap is further scrutinized for suspected plagiarism according to the publisher's Editorial Policies. Bentham Science allows an overall similarity of 20% for a manuscript to be considered for publication. The similarity percentage is further checked keeping the following important points in view:

Low Text Similarity:

The text of every submitted manuscript is checked using the Content Tracking mode in iThenticate. The Content Tracking mode ensures that manuscripts with an overall low percentage similarity (but which may have a higher similarity from a single source) are not overlooked. The acceptable limit for similarity of text from a single source is 5%. If the similarity level is above 5%, the manuscript is returned to the author for paraphrasing the text and citing the original source of the copied material.

It is important to mention that the text taken from different sources with an overall low similarity percentage will be considered as a plagiarized content if the majority of the article is a combination of copied material.

High Text Similarity:

There may be some manuscripts with an overall low similarity percentage, but a higher percentage from a single source. A manuscript may have less than 20% overall similarity but there may be 15 % similar text taken from a single article. The similarity index in such cases is higher than the approved limit for a single source. Authors are advised to thoroughly rephrase the similar text and properly cite the original source to avoid plagiarism and copyright violation.

Types of Plagiarism:

We all know that scholarly manuscripts are written after thorough review of previously published articles. It is therefore not easy to draw a clear boundary between legitimate representation and plagiarism. However, the following important features can assist in identifying different kinds of plagiarized content. These are:

- Reproduction of others words, sentences, ideas or findings as one's own without proper acknowledgement.
- Text recycling, also known as self-plagiarism. It is an author's use of a previous publication in another paper without proper citation and acknowledgement of the original source.
- Paraphrasing poorly: Copying complete paragraphs and modifying a few words without changing the structure of original sentences or changing the sentence structure but not the words.
- Verbatim copying of text without putting quotation marks and not acknowledging the work of the original author.
- Properly citing a work but poorly paraphrasing the original text is considered as unintentional plagiarism. Similarly, manuscripts with language somewhere between paraphrasing and quoting are not acceptable. Authors should either paraphrase properly or quote and in both cases, cite the original source.
- Higher similarity in the abstract, introduction, materials and methods, and discussion and conclusion sections indicates that the manuscript may contain plagiarized text. Authors can easily explain these parts of the manuscript in many ways. However, technical terms and sometimes standard procedures cannot be rephrased; therefore Editors must review these sections carefully before making a decision.

Plagiarism in Published Manuscripts:

Published manuscripts which are found to contain plagiarized text are retracted from the journal website after careful investigation and approval by the Editor-in-Chief of the journal. A 'Retraction Note' as well as a link to the original article is published on the electronic version of the plagiarized manuscript and an addendum with retraction notification in the journal concerned.

E-PUB AHEAD OF SCHEDULE:

Bentham Science Publishers are pleased to offer electronic publication of accepted papers prior to scheduled publication. These peer-reviewed papers can be cited using the date of access and the unique DOI number. Any final changes in manuscripts will be made at the time of print publication and will be reflected in the final electronic version of the issue. Articles ahead of schedule may be ordered by pay-per-view at the relevant links by each article stated *via* the [E-Pub Ahead of Schedule](#).

Disclaimer:

Articles appearing in E-Pub Ahead-of-Schedule sections have been peer-reviewed and accepted for publication in this journal and posted online before scheduled publication. Articles appearing here may contain statements, opinions, and information that have errors in facts, figures, or interpretation. Accordingly, Bentham Science Publishers, the editors and authors and their respective employees are not responsible or liable for the use of any such inaccurate or misleading data, opinion or information contained in articles in the E-Pub Ahead-of- Schedule.

Member of Cope



Anexo 2: Instruções para autores: *Genetics and Molecular Biology*

SUBMISSION OF PAPERS

There is a publication charge for manuscripts once they are accepted. For price information, exemptions and waiver policies, please consult the journal homepage <http://www.gmb.org.br>.

1. Manuscripts must be submitted through our online submission platform hosted at: <http://mc.manuscriptcentral.com/gmb>

The cover letter should be addressed to:
Carlos C. F. Menck, Editor-in-Chief, Genetics and Molecular Biology

2. For submission the following instructions must be observed:

- The manuscript must be submitted by the Corresponding Author. This is the person who will also check the page proofs, and arranges for any payment that may incur during the editorial process.
- Entering the following metadata is required: (i) the manuscript title, (ii) a short running title (max. 35 characters), (iii) the Abstract, and (iv) up to five keywords. All these items must be exactly the same as those figuring in the first two pages of the manuscript file. Furthermore, a cover letter addressed to the Editor is required. It must be edited by the corresponding author inserted within the reserved data field.
- Statements are required informing that the data have not been published and are not under consideration elsewhere, and that all authors have approved the submission of the manuscript. Furthermore, possible conflicts of interest (e.g. due to funding, consultancies) must also be disclosed.
- The names of all co-authors, including institutional affiliations and e-mail addresses must be entered, as contact information for the Editorial Office.
- In the referee suggestions field, up to five reviewer names can be entered by the author(s); valid e-mail contact addresses for these are required, in case they are selected by the editor. These suggestions can be made separately as preferred and not-preferred reviewer(s).
- Files must be uploaded separately and identified according to file types. The main text file must include references and, if applicable, figure legends, which must be typed on a separate page following the References and Internet Resources sections. Each table, figure and element containing supplementary material must be saved and uploaded in a separate file. Formats for text and tables are Word or RTF in Windows platform. Figures should be in TIFF or JPEG formats (see detailed instructions in 3.1.h).
- Manuscripts including photos or any other identifiable data of human subjects must be accompanied by a copy of the signed consent by the individual or his/her guardian.

Failure to adhere to these guidelines can delay the handling of your contribution and manuscripts may be returned before being reviewed. Special attention should be given to the structuring of the manuscript and correct language usage. These are important factors in the smooth running of the editorial and peer-review process, and can result in faster publication.

[Scope and policy](#)
[Submission of papers](#)
[Gene/Protein Nomenclature Guidelines](#)

3. Categories of Contribution

3.1. Research Articles

Manuscripts must be written in English in double-spaced, 12-point type throughout; marked with consecutive line and page numbers, beginning with the cover page. The following elements must start on a new page and be ordered as they are listed below:

a) **The title page** must contain: a concise and informative title; the authors' names (first name at full length); the authors' institutional affiliation, including department, institution, city, state or province, and country; different affiliations indicated with superscript Arabic numbers; a short running title of up to 35 characters (including spaces); up to five key words; the corresponding author's name, full postal, and email address.

b) **The Abstract** must be a single paragraph that does not exceed 200 words and summarizes the main results and conclusions of the study. It should not contain references.

c) **The text** must be as succinct as possible. *Text citations*: articles should be referred to by authors' surnames and date of publication; citations with two authors must include both names; in citations with three or more authors, name the first author and use *et al.* List two or more references in the same citation in chronological order, separated by semi-colons. When two or more works in a citation were published in the same year, list them alphabetically by the first author surname. For two or more works by the same author(s) in a citation, list them chronologically, with the years separated by commas. (Example: Freire-Maia *et al.*, 1966a, 1966b, 2000). Only articles that are published or in press should be cited. In the case of personal communications or unpublished results, all contributors must be listed by initials and last name (*et al.* should not be used). *Numbers*: In the text, numbers nine or less must be written out except as part of a date, a fraction or decimal, a percentage, or a unit of measurement. Use Arabic numerals for numbers larger than nine. *Binomial Names*: Latin names of genera, species and infraspecific taxa must be printed in italics; names of orders and families should appear in the Title and also when first mentioned in the text. URLs for programs, data or other sources should be listed in the Internet Resources Section, immediately following the References Section, not in the text. The text includes the following elements:
Introduction - Description of the background that led to the study.
Material (or Subjects) and Methods - Details relevant to the conduct of the study. Statistical methods should be explained at the end of this section.
Results - Undue repetition in text and tables should be avoided. Statistical analyses should be presented as complete as possible, i.e. not only *P*-values should be shown, but also all other test variables required for full appreciation of the results by the reviewers and readers. Comments on relevance of results are appropriate but broader discussion should be part of the Discussion section.
Discussion - The findings of the study should be placed in context of relevant published data. Ideas presented in other publications should not be discussed solely to make an exhaustive presentation. Some manuscripts may require different formats appropriate to their content.

d) **The Acknowledgments** must be a single paragraph that immediately follows the discussion and includes references to grant support.

e) **The References Section**: References must be ordered alphabetically by the first author surname; references with the same first author should be ordered as follows: first, as single author in chronological order; next, with only one more co-author in alphabetical order by the second author; and finally followed by references with more than two co-authors, in

chronological order, independent of the second author surnames. In references with more than 10 authors only the first ten should be listed, followed by *et al.* Use standard abbreviations for journal titles as suggested by NCBI (<http://www.ncbi.nlm.nih.gov/journals/>).

Only articles that are published or in press should be included in this section. Works submitted for publication but not yet accepted, personal communications and unpublished data must be cited within the text. "Personal communication" refers to information obtained from individuals other than the authors of the manuscript being submitted; "unpublished data" refers to data produced by at least one of the authors of the manuscript under consideration. Works of restricted circulation (e.g., theses not available in public databases, congress abstracts not published in regular journals or public databases) should not be listed in this section.

Sample	journal	article	citation:
Breuer ME and Pavan C (1955)		Behaviour of polytene chromosomes of <i>Rhynchosciara angelae</i> at different stages of larval development.	Chromosoma 7:371-386.
Yonenaga-Yassuda Y, Rodrigues MT and Pellegrino KCM (2005)		Chromosomal banding patterns in the eyelid-less microteiid lizard radiation: The X1X1X2X2:X1X2Y sex chromosome system in <i>Calyptommatus</i> and the karyotypes of <i>Psilophthalmus</i> and <i>Tretioscincus</i> (Squamata, Gymnophthalmidae).	Genet Mol Biol 28:700-709.

Sample	book	citation:
Dobzhansky T (1951)	Genetics and Origin of Species. 3rd edition. Columbia University Press, New York, 364 pp.	

Sample	chapter-in-book	citation:
Crawford DC and Howard-Peebles PN (2005)	Fragile X: From cytogenetics to molecular genetics. In: Gersen SL and Keagle MB (eds) The Principles of Clinical Cytogenetics. 2nd edition. Humana Press, New Jersey, pp 495-513.	

Sample	electronic	article	citation:
Gotzek D, Ross KG (2009)		Current status of a model System: The gene <i>Gp-9</i> and its association with social organization in fire ants.	PLoS One 4:e7713.

f) **Internet Resources Section:** this section should contain a list of URLs referring to data presented in the text, as well as software programs and other Internet resources used during data processing. Date of consultation must be stated.

Sample	Internet	resource	citation:
Online Mendelian Inheritance in Man (OMIM),	http://www.ncbi.nlm.nih.gov/OMIM		(September 4, 2009)
LEM Software,	http://dir.niehs.nih.gov/dirbb/weinbergfiles/hybrid_design.htm		(September 4, 2009)

g) **Tables:** must be in Word format prepared with the table tool (do not use space bar or tabulator). A concise title should be provided above the table. Tables must be numbered consecutively in Arabic numerals. Each column must have a title in the box head. Footnotes typed directly below the table should be indicated in lowercase superscript letters. Tables that are to appear in the printed version must be saved in Word format and not as figures, so that they can later be fitted during typesetting. Each table must be saved and uploaded as a separate file.

h) **Figures** must be numbered consecutively using Arabic numerals. Images should be in TIFF or JPEG format. Figures in Word, PowerPoint or Excel format cannot be published. Only sequence data can be presented in Word format. Journal quality reproduction will require grayscale resolution yielding 300 dpi, color figures should be at 600 dpi. These resolutions refer to the output size of the file, that is the size in which it will appear printed in the journal; if it is anticipated that images will be enlarged or reduced, the resolutions should be adjusted accordingly. Figures composed of several elements should be sent as a single panel, obeying the print size definitions of the journal (single or two columns width). Scanned figures should not be submitted. Color illustrations are accepted. Each figure/panel must be saved and uploaded as a separate file. When uploading, identify each illustration by the first author name and the number of the respective figure. Figure legends must be included at the end of the main text file and should be typed on a new page.

i) **Nomenclature:** Taxonomic names should be in accordance with current international standards. For rules concerning gene names and gene symbols, please see separate Instruction form.

j) **Sequences** may appear in text or in figure. DNA, RNA and protein sequences equal to or greater than 50 units must be entered into public databases and accession numbers must be provided upon acceptance of the article. Failure to do so will inadvertently delay publication.

k) **Data access:** reference should be made to availability of detailed data and materials used for reported studies.

l) **Ethical issues:** Reports of experiments on live vertebrates must include a statement in the text that the institutional review board approved the work and the protocol number must be provided. For experiments involving human subjects, authors must also include a statement that informed consent was obtained from all subjects. If photos or any other identifiable data are included, a copy of the signed consent must be uploaded during manuscript submission.

m) **Supplementary Material:** Data that the authors consider of importance for completeness of a study, but which are too extensive to be included in the published version, can be submitted as Supplementary Material. At publication, this material will be made available together with the electronic version. In case a manuscript contains such material, it should be appropriately identified within the text file. Supplementary material in tables should be identified as Table S1, Table S2, etc., in case of figures they should be named accordingly, Figure S1, Figure S2. In addition, a list of this material should be presented at the end of the manuscript text file, containing the following statement:

Supplementary material - the following online material is available for this article:

- *Table S1 – < short title >*
- *Figure S1 – < short title >*

This material is available as part of the online article from <http://www.scielo.br/gmb>

3.2 Short Communications

Short Communications present brief observations that do not warrant full-length articles.

They should not be considered preliminary communications;

- should be 15 or fewer typed pages in double spaced 12-point type, including literature cited;
- should include an Abstract;
- but no further subdivision, with introduction, material and methods, results and discussion; all in a single section and without headers.
- up to four items (tables and/or figures) may be submitted;

Note: The title page, abstract and reference section format is that of a full-length Research Article. For Supplementary Material see instructions in item 3.1.m.

3.3 Letters to the Editor

Relate or respond to recent published items in the journal. Discussions of political, social and ethical issues of interest to geneticists are also welcome in this form.

3.4 Review Articles

Review Articles are welcome. The Editor must be contacted prior to submission. Please, provide an Abstract and a list of your recent publications in the area.

3.5 Book Reviews

Publishers are invited to submit books on Genetics, Evolution and related disciplines, for review in the journal. Aspiring reviewers may propose writing a review.

3.6 History, Story and Memories

These are accounts on historical aspects of Genetics relating to Brazil.

4. Articles accepted for publication

Once an article is accepted, the Editorial Office will send it to copy editor for language and technical corrections. If major corrections were proposed, the manuscript with the highlighted corrections will be returned to the corresponding author for approval. The final version approved by the authors must be free of any text/correction markings when returned to the Editorial Office. After typesetting, page proofs will be sent to the corresponding author. Changes made to page proofs, apart from typesetting errors, will be charged to the authors. Notes added in proof require Editorial approval. Together with the proofs, a form of consent to publish and transfer of copyright is sent to the corresponding other. The latter will have sign this form, also on behalf of any co-authors, and send it by fax to the Editorial Office.

5. Reprints

Reprints are free of charge and will be provided as a pdf-file.

Anexo 3

(Tabela) Sequências de ciclotídeos obtidas no GenBank/ NCBI e utilizadas para geração de motivos através da ferramenta HMMER.

Acesso	Espécie	Ordem
ACG69861.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69860.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69859.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69858.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69857.1	<i>Viola baoshanensis</i>	Malpighiales
2KNN_A	<i>Viola odorata</i>	Malpighiales
2KNM_A	<i>Viola odorata</i>	Malpighiales
AAU04395.1	<i>Viola odorata</i>	Malpighiales
ABQ96648.1	<i>Viola baoshanensis</i>	Malpighiales
P83937.2	<i>Viola hederacea</i>	Malpighiales
P84522.1	<i>Viola hederacea</i>	Malpighiales
1ZA8_A	<i>Viola hederacea</i>	Malpighiales
ACG69855.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69854.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69853.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69852.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69851.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69850.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69849.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69848.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69847.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69846.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69845.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69844.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69843.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69842.1	<i>Viola baoshanensis</i>	Malpighiales
ACG69841.1	<i>Viola baoshanensis</i>	Malpighiales
P85231.1	<i>Viola hederacea</i>	Malpighiales
ADB03554.1	<i>Gloeospermum pauciflorum</i>	Malpighiales
ADB03553.1	<i>Gloeospermum pauciflorum</i>	Malpighiales
ADB03552.1	<i>Gloeospermum pauciflorum</i>	Malpighiales
ADB03551.1	<i>Gloeospermum pauciflorum</i>	Malpighiales
ADB03550.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ADB03549.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ADB03548.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ADB03547.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ADB03546.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ADB03545.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ABI48959.1	<i>Viola baoshanensis</i>	Malpighiales
ABI48958.1	<i>Viola baoshanensis</i>	Malpighiales

ABI48957.1	<i>Viola baoshanensis</i>	Malpighiales
ABI48956.1	<i>Viola baoshanensis</i>	Malpighiales
ABO21629.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21628.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21627.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21626.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21620.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21619.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21616.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABQ96647.1	<i>Viola baoshanensis</i>	Malpighiales
AGI97142.1	<i>Viola odorata</i>	Malpighiales
ABB04013.1	<i>Hybanthus epacroides</i>	Malpighiales
ABB04010.1	<i>Hybanthus enneaspermus</i>	Malpighiales
ABB04009.1	<i>Hybanthus debilissimus</i>	Malpighiales
ABB04003.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB04002.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB04001.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB04000.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB03999.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB03998.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB03997.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB03996.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ABB03995.1	<i>Hybanthus floribundus</i> subsp. <i>Floribundus</i>	Malpighiales
ACG69856.1	<i>Viola baoshanensis</i>	Malpighiales
AGI97141.1	<i>Viola odorata</i>	Malpighiales
AGI97140.1	<i>Viola odorata</i>	Malpighiales
AGI97139.1	<i>Viola odorata</i>	Malpighiales
1YP8_A	<i>Viola tricolor</i>	Malpighiales
1VB8_A	<i>Viola hederacea</i>	Malpighiales
ABW08095.1	<i>Viola biflora</i>	Malpighiales
ABW08094.1	<i>Viola biflora</i>	Malpighiales
ABW08093.1	<i>Viola biflora</i>	Malpighiales
ABW08092.1	<i>Viola biflora</i>	Malpighiales
ABW08091.1	<i>Viola biflora</i>	Malpighiales
ABW08090.1	<i>Viola biflora</i>	Malpighiales
AAU04394.1	<i>Viola odorata</i>	Malpighiales
AAU04393.1	<i>Viola odorata</i>	Malpighiales
AAU04392.1	<i>Viola odorata</i>	Malpighiales
ABI48960.1	<i>Viola baoshanensis</i>	Malpighiales
ACI29323.1	<i>Viola odorata</i>	Malpighiales
ACI29317.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21630.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21625.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21624.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21623.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21622.1	<i>Melicytus ramiflorus</i>	Malpighiales

ABO21621.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21618.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21617.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21615.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABO21614.1	<i>Melicytus ramiflorus</i>	Malpighiales
ABB51617.1	<i>Hybanthus floribundus subsp. Floribundus</i>	Malpighiales
ABB51616.1	<i>Hybanthus floribundus subsp. Floribundus</i>	Malpighiales
ABB04012.1	<i>Hybanthus epacroides</i>	Malpighiales
ABB04008.1	<i>Hybanthus calycinus</i>	Malpighiales
ABB04007.1	<i>Hybanthus vernonii subsp. vernonii</i>	Malpighiales
ABB04006.1	<i>Hybanthus stellarioides</i>	Malpighiales
ABB04005.1	<i>Hybanthus monopetalus</i>	Malpighiales
ABB04004.1	<i>Hybanthus monopetalus</i>	Malpighiales
ABB03994.1	<i>Hybanthus floribundus subsp. Floribundus</i>	Malpighiales
ABB04011.1	<i>Hybanthus enneaspermus</i>	Malpighiales
ABB51618.1	<i>Hybanthus floribundus subsp. Floribundus</i>	Malpighiales
P58445.1	<i>Hybanthus parviflorus</i>	Malpighiales
B1NRR3.1	<i>Viola biflora</i>	Malpighiales
B1NRR1.1	<i>Viola biflora</i>	Malpighiales
B1NRR0.1	<i>Viola biflora</i>	Malpighiales
B1NRQ9.1	<i>Viola biflora</i>	Malpighiales
P85246.1	<i>Viola biflora</i>	Malpighiales
P85245.1	<i>Viola biflora</i>	Malpighiales
P85244.1	<i>Viola biflora</i>	Malpighiales
B1NRQ8.1	<i>Viola biflora</i>	Malpighiales
P84649.1	<i>Hybanthus floribundus</i>	Malpighiales
P84648.1	<i>Hybanthus floribundus</i>	Malpighiales
P84647.1	<i>Hybanthus floribundus</i>	Malpighiales
P84636.1	<i>Viola cotyledon</i>	Malpighiales
P84635.1	<i>Viola cotyledon</i>	Malpighiales
ACI29316.1	<i>Viola odorata</i>	Malpighiales
P58440.2	<i>Viola odorata</i>	Malpighiales
P58443.2	<i>Viola odorata</i>	Malpighiales
Q5USN8.1	<i>Viola odorata</i>	Malpighiales
2KCG_A	<i>Viola odorata</i>	Malpighiales
P85525.1	<i>Viola biflora</i>	Malpighiales
Q2HY54.1	<i>Viola odorata</i>	Malpighiales
P85242.1	<i>Viola biflora</i>	Malpighiales
P85241.1	<i>Viola biflora</i>	Malpighiales
P85240.1	<i>Viola biflora</i>	Malpighiales
P85239.1	<i>Viola biflora</i>	Malpighiales
P85233.1	<i>Viola hederacea</i>	Malpighiales
2FQA_A	<i>Viola odorata</i>	Malpighiales
ABC94585.1	<i>Viola odorata</i>	Malpighiales
P83840.1	<i>Viola arvensis</i>	Malpighiales
P85526.1	<i>Viola biflora</i>	Malpighiales

B1NRR2.1	<i>Viola biflora</i>	Malpighiales
P58436.2	<i>Viola odorata</i>	Malpighiales
P85188.1	<i>Viola odorata</i>	Malpighiales
P85183.1	<i>Viola odorata</i>	Malpighiales
P85180.1	<i>Viola odorata</i>	Malpighiales
P84640.1	<i>Leonia cymosa</i>	Malpighiales
P84639.1	<i>Leonia cymosa</i>	Malpighiales
P84637.1	<i>Leonia cymosa</i>	Malpighiales
P58433.1	<i>Viola hederacea</i>	Malpighiales
P58442.1	<i>Viola odorata</i>	Malpighiales
P58441.1	<i>Viola odorata</i>	Malpighiales
P58439.1	<i>Viola odorata</i>	Malpighiales
P58438.1	<i>Viola odorata</i>	Malpighiales
P58437.1	<i>Viola odorata</i>	Malpighiales
P58435.1	<i>Viola odorata</i>	Malpighiales
P58434.1	<i>Viola odorata</i>	Malpighiales
CAJ56225.1	<i>Viola sp. WO2005118628</i>	Malpighiales
P85234.1	<i>Viola hederacea</i>	Malpighiales
P85232.1	<i>Viola hederacea</i>	Malpighiales
P82230.2	<i>Viola odorata</i>	Malpighiales
2GJO_A	<i>Viola odorata</i>	Malpighiales
1NBJ_A	<i>Viola odorata</i>	Malpighiales
ACZ95475.1	<i>Gloeospermum pauciflorum</i>	Malpighiales
ACZ95474.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ACZ95473.1	<i>Gloeospermum blakeanum</i>	Malpighiales
ACZ95472.1	<i>Viola biflora</i>	Malpighiales
ACZ95471.1	<i>Viola biflora</i>	Malpighiales
ACZ95470.1	<i>Viola biflora</i>	Malpighiales
ACZ95469.1	<i>Viola biflora</i>	Malpighiales
P0C589.1	<i>Viola arvensis</i>	Malpighiales
P83839.1	<i>Viola odorata</i>	Malpighiales
P83838.1	<i>Viola odorata</i>	Malpighiales
P58453.1	<i>Viola arvensis</i>	Malpighiales
P58452.1	<i>Viola arvensis</i>	Malpighiales
P58451.1	<i>Viola arvensis</i>	Malpighiales
P58449.1	<i>Viola arvensis</i>	Malpighiales
P58448.1	<i>Viola arvensis</i>	Malpighiales
P58447.1	<i>Viola arvensis</i>	Malpighiales
P58446.1	<i>Viola arvensis</i>	Malpighiales
P85187.1	<i>Viola odorata</i>	Malpighiales
P85186.1	<i>Viola odorata</i>	Malpighiales
P85185.1	<i>Viola odorata</i>	Malpighiales
P85184.1	<i>Viola odorata</i>	Malpighiales
P85182.1	<i>Viola odorata</i>	Malpighiales
P85181.1	<i>Viola odorata</i>	Malpighiales
P85179.1	<i>Viola odorata</i>	Malpighiales

P85178.1	<i>Viola odorata</i>	Malpighiales
P85177.1	<i>Viola odorata</i>	Malpighiales
P84638.1	<i>Leonia cymosa</i>	Malpighiales
P83836.1	<i>Viola odorata</i>	Malpighiales
P83835.1	<i>Viola arvensis</i>	Malpighiales
1DF6_A	<i>Viola odorata</i>	Malpighiales
CAN08378.1	<i>Viola sp. WO2006041917</i>	Malpighiales
CAJ56194.1	<i>Viola sp. WO2005118628</i>	Malpighiales
CAJ56226.1	<i>Viola sp. WO2005118628</i>	Malpighiales
Q5USN7.1	<i>Viola odorata</i>	Malpighiales
AGQ04614.1	<i>Carapichea ipecacuanha</i>	Gentianales
AGQ04613.1	<i>Carapichea ipecacuanha</i>	Gentianales
1R1F_A	<i>Palicourea condensata</i>	Gentianales
AFH57365.1	<i>Chassalia chartacea</i>	Gentianales
AFH57364.1	<i>Chassalia chartacea</i>	Gentianales
AFH57363.1	<i>Chassalia chartacea</i>	Gentianales
AFH57362.1	<i>Chassalia chartacea</i>	Gentianales
AFH57361.1	<i>Chassalia chartacea</i>	Gentianales
AFH57360.1	<i>Chassalia chartacea</i>	Gentianales
AFH57359.1	<i>Chassalia chartacea</i>	Gentianales
AFH57358.1	<i>Chassalia chartacea</i>	Gentianales
AFH57357.1	<i>Chassalia chartacea</i>	Gentianales
AFH57356.1	<i>Chassalia chartacea</i>	Gentianales
AFH57355.1	<i>Chassalia chartacea</i>	Gentianales
AFH57354.1	<i>Chassalia chartacea</i>	Gentianales
AFH57352.1	<i>Chassalia chartacea</i>	Gentianales
AFH57353.1	<i>Chassalia chartacea</i>	Gentianales
B3EWF1.1	<i>Palicourea rigida</i>	Gentianales
P84644.1	<i>Chassalia parvilfora</i>	Gentianales
P84643.1	<i>Chassalia parvilfora</i>	Gentianales
P84642.1	<i>Chassalia parvilfora</i>	Gentianales
P84641.1	<i>Chassalia parvilfora</i>	Gentianales
P56879.2	<i>Chassalia parvilfora</i>	Gentianales
P56872.2	<i>Chassalia parvilfora</i>	Gentianales
P56871.2	<i>Chassalia parvilfora</i>	Gentianales
AEX01723.1	<i>Hedyotis biflora</i>	Gentianales
AEX01722.1	<i>Hedyotis biflora</i>	Gentianales
AEX01721.1	<i>Hedyotis biflora</i>	Gentianales
AEX01720.1	<i>Hedyotis biflora</i>	Gentianales
P84645.1	<i>Palicourea condensata</i>	Gentianales
1BH4_A	<i>Chassalia parvilfora</i>	Gentianales
2ERI_A	<i>Chassalia parvilfora</i>	Gentianales
AAB34024.1	<i>Psychotria longipes</i>	Gentianales
2KUX_A	<i>Oldenlandia affinis</i>	Gentianales
ADJ17473.1	<i>Oldenlandia affinis</i>	Gentianales
ADJ17472.1	<i>Oldenlandia affinis</i>	Gentianales

ADJ17471.1	<i>Oldenlandia affinis</i>	Gentianales
ADJ17470.1	<i>Oldenlandia affinis</i>	Gentianales
1ZNU_A	<i>Oldenlandia affinis</i>	Gentianales
1PT4_A	<i>Oldenlandia affinis</i>	Gentianales
ADJ17469.1	<i>Oldenlandia affinis</i>	Gentianales
AAL05480.1	<i>Oldenlandia affinis</i>	Gentianales
P56254.3	<i>Oldenlandia affinis</i>	Gentianales
P58456.2	<i>Oldenlandia affinis</i>	Gentianales
P85135.1	<i>Oldenlandia affinis</i>	Gentianales
P85134.1	<i>Oldenlandia affinis</i>	Gentianales
P58457.1	<i>Oldenlandia affinis</i>	Gentianales
2KCH_A	<i>Oldenlandia affinis</i>	Gentianales
P58455.1	<i>Oldenlandia affinis</i>	Gentianales
P58454.1	<i>Oldenlandia affinis</i>	Gentianales
P85127.1	<i>Oldenlandia affinis</i>	Gentianales
P85175.1	<i>Oldenlandia affinis</i>	Gentianales
P85133.1	<i>Oldenlandia affinis</i>	Gentianales
P85132.1	<i>Oldenlandia affinis</i>	Gentianales
P85131.1	<i>Oldenlandia affinis</i>	Gentianales
P85130.1	<i>Oldenlandia affinis</i>	Gentianales
P85129.1	<i>Oldenlandia affinis</i>	Gentianales
P85128.1	<i>Oldenlandia affinis</i>	Gentianales
P83938.2	<i>Oldenlandia affinis</i>	Gentianales
1NB1_A	<i>Oldenlandia affinis</i>	Gentianales
AAL05479.1	<i>Oldenlandia affinis</i>	Gentianales
AAL05477.1	<i>Oldenlandia affinis</i>	Gentianales
ACI29320.1	<i>Oldenlandia affinis</i>	Gentianales
ACI29319.1	<i>Oldenlandia affinis</i>	Gentianales
P58458.1	<i>Oldenlandia affinis</i>	Gentianales
2JWM_A	<i>Oldenlandia affinis</i>	Gentianales
2B38_A	<i>Oldenlandia affinis</i>	Gentianales
AFM52764.1	<i>Petunia x hybrida</i>	Solanales
AFM52763.1	<i>Petunia x hybrida</i>	Solanales
AFM52762.1	<i>Petunia x hybrida</i>	Solanales
B3EWH5.1	<i>Petunia x hybrida</i>	Solanales
B3EWH6.1	<i>Petunia x hybrida</i>	Solanales
B3EWH7.1	<i>Petunia x hybrida</i>	Solanales
2LAM_A	<i>Clitoria ternatea</i>	Fabales
AEK26407.1	<i>Clitoria ternatea</i>	Fabales
AEK26403.1	<i>Clitoria ternatea</i>	Fabales
AEK26412.1	<i>Clitoria ternatea</i>	Fabales
AEK26411.1	<i>Clitoria ternatea</i>	Fabales
AEK26410.1	<i>Clitoria ternatea</i>	Fabales
AEK26409.1	<i>Clitoria ternatea</i>	Fabales
AEK26408.1	<i>Clitoria ternatea</i>	Fabales
AEK26406.1	<i>Clitoria ternatea</i>	Fabales

AEK26405.1	<i>Clitoria ternatea</i>	Fabales
AEK26404.1	<i>Clitoria ternatea</i>	Fabales
AEK26402.1	<i>Clitoria ternatea</i>	Fabales
AEB92229.1	<i>Clitoria ternatea</i>	Fabales
P86842.2	<i>Clitoria ternatea</i>	Fabales
P86903.1	<i>Clitoria ternatea</i>	Fabales
P86904.1	<i>Clitoria ternatea</i>	Fabales
P86902.1	<i>Clitoria ternatea</i>	Fabales
P86901.1	<i>Clitoria ternatea</i>	Fabales
P86899.1	<i>Clitoria ternatea</i>	Fabales
P86852.1	<i>Clitoria ternatea</i>	Fabales
P86851.1	<i>Clitoria ternatea</i>	Fabales
P86850.1	<i>Clitoria ternatea</i>	Fabales
P86849.1	<i>Clitoria ternatea</i>	Fabales
P86848.1	<i>Clitoria ternatea</i>	Fabales
P86847.1	<i>Clitoria ternatea</i>	Fabales
P86845.1	<i>Clitoria ternatea</i>	Fabales
P86844.1	<i>Clitoria ternatea</i>	Fabales
P86843.1	<i>Clitoria ternatea</i>	Fabales
P86841.1	<i>Clitoria ternatea</i>	Fabales
P86900.1	<i>Clitoria ternatea</i>	Fabales
P86846.1	<i>Clitoria ternatea</i>	Fabales
XP_004962819.1	<i>Setaria italica</i>	Poales
XP_004962817.1	<i>Setaria italica</i>	Poales
AAW23948.1	<i>Zea mays</i>	Poales
AFZ39122.1	<i>Steinchisma laxum</i>	Poales
AFZ39121.1	<i>Steinchisma laxum</i>	Poales
AFZ39118.1	<i>Steinchisma laxum</i>	Poales
DAA55293.1	<i>Zea mays</i>	Poales
DAA55290.1	<i>Zea mays</i>	Poales
XP_004962820.1	<i>Setaria italica</i>	Poales
XP_004962816.1	<i>Setaria italica</i>	Poales
NP_001144691.1	<i>Zea mays</i>	Poales
NP_001105812.1	<i>Zea mays</i>	Poales
ACG47570.1	<i>Zea mays</i>	Poales
ACG45070.1	<i>Zea mays</i>	Poales
ACG43881.1	<i>Zea mays</i>	Poales
ACG42356.1	<i>Zea mays</i>	Poales
ACG26826.1	<i>Zea mays</i>	Poales
CAL31949.1	<i>Zea mays</i>	Poales
AFZ39120.1	<i>Steinchisma laxum</i>	Poales
AFZ39119.1	<i>Steinchisma laxum</i>	Poales
3E4H_A	Construção sintética	-
2K7G_A	Construção sintética	-
2F2J_A	Construção sintética	-
2F2I_A	Construção sintética	-

1WN8_A	Construção sintética	-
1WN4_A	Construção sintética	-
2LUR_A	Construção sintética	-
1N1U_A	Construção sintética	-
1ORX_A	Construção sintética	-
CAN08379.1	Construção sintética	-
1IB9_A	<i>Momordica cochinchinensis</i>	Cucurbitales
2M86_B	<i>Homo sapiens</i>	Primata

Anexo 4

(Tabela) BLASTp realizado do banco de dados Phytozome, para análise de similaridade.

Nome	Organismo	Nome do locus*	Score	E-value	Identity*	Positive	Frame	Nome transcrito*	Localização*	Pfam	GO
zm315	<i>Zea mays</i>	-	131.7	6.8E-39	-	100%	0	-	-	3784	6952
	<i>Zea mays</i>	-	110.2	1.7E-30	-	90.5%	0	-	-	3784	6952
	<i>Sorghum bicolor</i>	-	67.0	9.7E-14	-	74.1%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	60.1	3.2E-11	-	68.4%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	59.3	5.3E-11	-	71.7%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	58.2	1.6E-10	-	70.2%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	51.6	3.9E-8	-	63.8%	0	-	-	3784	6952
zm316	<i>Zea mays</i>	-	104.4	1.5E-28	-	88.7%	0	-	-	3784	6952
	<i>Zea mays</i>	-	104.4	1.6E-28	-	88.7%	0	-	-	3784	6952
	<i>Sorghum bicolor</i>	-	73.2	2.3E-16	-	69.7%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	59.3	2.9E-11	-	61.9%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	60.5	1.1E-11	-	68.5%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	60.1	1.5E-11	-	66.7%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	52.8	9.5E-9	-	63.5%	0	-	-	3784	6952
zm317	<i>Zea mays</i>	-	98.6	6.1E-26	-	86.9%	0	-	-	3784	6952
	<i>Zea mays</i>	-	107.8	1.4E-29	-	90.2%	0	-	-	3784	6952
	<i>Sorghum bicolor</i>	-	67.0	9.5E-14	-	70.7%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	52.4	2.00E-08	-	60.7%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	52.8	1.7E-8	-	66%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	54.3	3.8E-9	-	63.8%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	45.8	5.1E-6	-	59.3%	0	-	-	3784	6952
sc359	<i>Zea mays</i>	-	105.5	1.5E-28	-	84.7%	0	-	-	3784	6952
	<i>Zea mays</i>	-	107.5	3.1E-29	-	83.5%	0	-	-	3784	6952
	<i>Sorghum bicolor</i>	-	84.0	5.2E-20	-	65.5%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	61.2	1.8E-11	-	61.2%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	60.5	2.9E-11	-	58.2%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	62.4	6.3E-12	-	66.2%	0	-	-	3784	6952
	<i>Setaria italica</i>	-	55.5	2.00E-09	-	64.6%	0	-	-	3784	6952
	<i>Panicum virgatum</i>	-	42.7	1.00E-04	-	46.4%	0	-	-	Sem anotação	Sem anotação

*Dados omitidos por motivo de propriedade intelectual.