

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP  
CÂMPUS DE JABOTICABAL**

**ANÁLISE E EXPRESSÃO DE GENES DE PKS II E DE  
CARBOIDRASES PROVENIENTES DE BIBLIOTECAS  
METAGENÔMICAS**

**Elisângela Soares Gomes**

Biólogo

**2015**

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP  
CÂMPUS DE JABOTICABAL**

**ANÁLISE E EXPRESSÃO DE GENES DE PKS II E DE  
CARBOIDRASES PROVENIENTES DE BIBLIOTECAS  
METAGENÔMICAS**

**Elisângela Soares Gomes**

Orientadora: Profa. Dra. Eliana Gertrudes de Macedo Lemos

Tese de Doutorado apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Microbiologia Agropecuária.

**2015**

G633a Gomes, Elisângela Soares  
Análise e Expressão de Genes de PKS II e de Carboidrases  
Provenientes de Bibliotecas Metagenômicas / Elisângela Soares  
Gomes. -- Jaboticabal, 2015  
ix, 107 p. ; 28 cm

Tese (doutorado) - Universidade Estadual Paulista, Faculdade de  
Ciências Agrárias e Veterinárias, 2015

Orientadora: Eliana Gertrudes de Macedo Lemos

Banca examinadora: André Ricardo de Lima Damásio, Janete  
Aparecida Desidério, Igor Polikarpov, Tiago Santana Balbuena.

Bibliografia

1.  $\beta$ -glicosidase. 2. Celulase. 3. Metagenômica. I. Título. II.  
Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 576.8:575.111

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação –  
Serviço Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

**CERTIFICADO DE APROVAÇÃO**

**TÍTULO:** ANÁLISE E EXPRESSÃO DE GENES DE PKSII E DE CARBOIDRASES  
PROVENIENTES DE BIBLIOTECAS METAGENÔMICAS

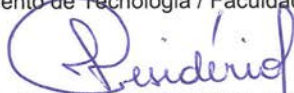
**AUTORA:** ELISÂNGELA SOARES GOMES

**ORIENTADORA:** Profa. Dra. ELIANA GERTRUDES DE MACEDO LEMOS

Aprovada como parte das exigências para obtenção do Título de DOUTOR EM MICROBIOLOGIA  
AGROPECUÁRIA, pela Comissão Examinadora:



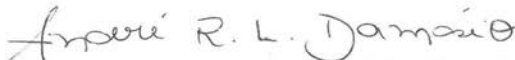
Profa. Dra. ELIANA GERTRUDES DE MACEDO LEMOS  
Departamento de Tecnologia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal



Profa. Dra. JANETE APPARECIDA DESIDERIO  
Departamento de Biologia Aplicada - Agropecuária / Faculdade de Ciências Agrárias e  
Veterinárias de Jaboticabal



Prof. Dr. TIAGO SANTANA BALBUENA  
Departamento de Tecnologia / Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal



Prof. Dr. ANDRE RICARDO DE LIMA DAMASIO  
Laboratório Nacional de Ciência e Tecnologia do Bioetanol / Campinas/SP



Prof. Dr. IGOR POLIKARPOV  
Universidade de São Paulo / São Carlos/SP

Data da realização: 12 de fevereiro de 2015.

## DADOS CURRICULARES DO AUTOR

**ELISÂNGELA SOARES GOMES** – Nasceu em Londrina (PR, Brasil) no dia 05 de janeiro de 1985. Ingressou na “Faculdade de Ciências Agrárias e Veterinárias” – FCAV-UNESP/Jaboticabal em março de 2005, no curso de Ciências Biológicas (modalidade Bacharel), o qual foi concluindo em dezembro de 2008. No ano de 2009 ingressou na modalidade de Licenciatura do curso de Ciências Biológicas, com término em dezembro de 2009. Durante os cinco anos de graduação realizou atividades de estágio no Laboratório de Bioquímica de Micro-organismos e Plantas (LBMP), sob orientação da Profa. Dra. Eliana Gertrudes de Macedo Lemos. No período de 2005-2006 realizou estágio voltado para o cultivo *in vitro* e conservação de *Orchidaceae*, acompanhando o trabalho do então graduando Renato Fernandes Galdiano Júnior, participando de trabalhos envolvendo o cultivo *in vitro* de sementes de *Cattleya walkeriana*, cultivo por gêmula floral de *Phalaenopsis*, cultivo por calos de raízes de *Cattleya walkeriana* e documentação por Microscopia Eletrônica de Varredura de estágios iniciais de germinação de *Cattleya walkeriana*. No período de julho de 2006 a julho de 2007 acompanhou a então doutoranda Viviane Schuch em seu trabalho de prospecção de genes envolvidos na síntese de antibióticos do sistema PKS I (Modular), envolvendo a construção de bibliotecas metagenômicas e busca por sequências conservadas das mega-sínteses, tendo oportunidade de aprendizagem de técnicas de biologia molecular e uso de ferramentas de bioinformática. No período de agosto de 2007 a julho de 2009 desenvolveu uma atividade de prospecção de clones metagenômicos abrigando agrupamentos gênicos relacionados à síntese de antibióticos policetílicos aromáticos (PKSII) com Bolsa de auxílio FAPESP (processo 07/53708-0). Em março de 2010 ingressou no curso de Mestrado em Microbiologia Agropecuária (FCAV-UNESP) realizando, como atividade de pesquisa, a predição funcional *in silico* de genes envolvidos na síntese de antibióticos policetílicos aromáticos, com Bolsa de Auxílio do Programa oferecida pelo CNPq (processo 130454/2010-9), concluído em fevereiro de 2011, mesmo ano em que ingressou no curso de Doutorado em Microbiologia Agropecuária (FCAV-UNESP), desenvolvendo o presente trabalho de pesquisa, com suporte da Fapesp (Processo 2011/10981-3).

## EPÍGRAFE

*“Águas profundas são as palavras da boca do homem,  
e ribeiro transbordante é a fonte da sabedoria”.*

*Provérbios, 18:4.*

## DEDICATÓRIA:

À minha família querida, Luíza de Castro S. Clemente (mãe), José Cicero Gomes e José Clemente (pais); Jaqueline S. Clemente, Cristiano S. Clemente, Jéssica Fernanda Gomes (irmãos)

Ao Meu grande amor, Paulo Sergio de Souza Pepe

À Deus, por toda a ajuda e força.

## **AGRADECIMENTOS:**

À minha orientadora, Profa. Dra. **Eliana G. de Macedo Lemos**, pelo constante incentivo, pela sabedoria das palavras diante de cada impasse, pela paciência e pela confiança em mim depositada;

À minha família: **Luiza** (mãe), **José Cícero** (pai), **José Clemente** (pai de coração), **Cristiano, Jaqueline e Jéssica** (Irmãos), pela estrutura e estabilidade emocional.. já viram uma casa sem alicerce?

Ao meu amor **Paulo Sergio de Souza Pepe** pela paciência, ajuda e amor incondicional;

Aos amigos com os quais eu tive a honra de trabalhar direta e indiretamente: **João Carlos Campanharo, Lúcia M. Carareto Alves, Jackson M. de Souza, Simone C. Pick, Ester Wekert, Silvana Pompeia, Viviane Schuch, Viviane Vieira, Renata Andrade; Tereza Cristina L. Castellane; Eliamar N. Pedrinho; Renato F. Galdiano Jr., Gisele Regina Rodrigues; Wellington Moreira, Karla Cristina Stropa; Fernanda Paganelli, Mauricio R. Demitrove, Douglas Paixão, Mariana Rangel, Thais Maester, Erica Mendes Lopes, Elwi M. Sierra, Aliandra, André, Andressa, Wellington Moreira, Wellington Omeri, Milena, Rosmeriana, Micheli Persona, Maria Luíza,** enfim: **a todos os meus amigos do LBMP**, pela ajuda de sempre, pelo ambiente amistoso, propiciando o crescimento mútuo pessoal, afetivo e profissional, e pela amizade de sempre, sem a qual não existe a satisfação pessoal;

À **Fapesp, CNPq**, e ao **Programa de Microbiologia Agropecuária**, pelo suporte a essa pesquisa.



## SUMÁRIO

	<b>Página</b>
SUMÁRIO .....	i
RESUMO .....	iv
ABSTRACT .....	v
LISTA DE ABREVIATURAS .....	vi
LISTA DE TABELAS .....	vii
LISTA DE FIGURAS .....	viii
1. INTRODUÇÃO .	1
1.1. Objetivos do trabalho .....	4
.....1.1.1. Geral .....	4
1.1.2 Específicos .....	4
1.2. Fluxograma das Atividades desenvolvidas .....	4
2. REVISÃO DE LITERATURA. ....	6
2.1. Acesso ao potencial de micro-organismos incultiváveis .....	6
2.2. Potencial industrial microbiano .....	10
2.3. Enzimas envolvidas na degradação de biomassa lignocelulósica e a produção de etanol de segunda geração .....	13
2.4. Vias de Degradação de biomassa lignocelulósica .....	18
2.5. Multicopper oxidases .....	21
2.6. Motivação para a busca de novos antibióticos .....	22
2.7. Policetídeos .....	25
2.8. Tipos de PKS e Enzimologia dos Grupos .....	27
2.9. As Policetídeo Sintases II .....	31
2.10. Biossíntese combinatória como estratégia para a produção de novos fármacos .....	32
3.11. O uso da metagenômica para busca de novas vias de antibióticos .....	35
3. MATERIAL E MÉTODOS .....	37
3.1. Origem da amostra metagenômica e análise in silico .....	37
3.2. Linhagens e vetores utilizados .....	39
3.3. Meios de cultura .....	40
3.4. Técnicas gerais .....	40
3.4.1. Eletroforese de DNA .....	40

3.4.2. Eletroforese para amostras proteicas .....	41
3.5. Clonagem dos genes em vetores de expressão .....	42
3.5.1. Síntese de oligonucleotídeos iniciadores .....	42
3.5.2. Reação de amplificação por PCR .....	43
3.5.3. Preparo do vetor pET28a EcoRI/HindIII .....	44
3.5.4. Preparo dos insertos e reação de ligação .....	45
3.5.5. Preparo das células competentes .....	46
3.5.6. Transformação por choque térmico .....	47
3.6. Ensaios de Expressão .....	47
3.6.1. Protocolo de extração empregando tampão Tris .....	48
3.6.2. Protocolo de extração empregando tampão Fosfato .....	49
3.6.3. Protocolo para extração empregando tampão Hepes .....	49
... 3.7. Métodos de lise celular utilizados .....	50
3.7.1. Extração empregando ultrassom (lise por cavitação) .....	50
3.7.2. Extração por congelamento e descongelamento .....	50
3.8. Purificação através cromatografia de afinidade por metal imobilizado (Ni <sup>2+</sup> ) .....	51
3.9. Determinação da concentração de proteínas solúveis .....	51
3.10. Ensaio de Atividade enzimática para a β-glicosidase .....	52
3.10.1. Ensaio de atividade enzimática empregando substratos cromogênicos .....	52
3.10.2. Determinação de unidade enzimática (U) .....	53
3.10.3. Temperatura e pH ótimos .....	53
3.10.4. Determinação de Parâmetros cinéticos .....	53
3.10.5. Efeito de sais e outros reagentes sobre a atividade enzimática .....	54
3.10.6. Determinação de unidade enzimática (U) para β-glicosidase .....	55
3.10.7. Zimograma para β-glicosidase (E.C. - 3.2.1.21) .....	55
3.11. Preparo de materiais para os Ensaios de Expressão para o “cluster” de PKS II .....	55
4. RESULTADOS E DISCUSSÃO .....	58
4.1. Predição in silico das enzimas metagenômicas envolvidas na degradação de material lignocelulósico .....	58
4.1.1. β-Glicosidase Metagenômica .....	59
4.1.2. Análise in silico da celulase .....	64
4.1.3. Análise In Silico da Lacase .....	72

4.2. Amplificação das ORFs metagenômica relacionadas ao metabolismo de biomassa lignocelulósica .....	78
4.3. Expressão Enzimática e Purificação das carboidrases .....	81
4.3.1- Celulase .....	81
4.3.2 Multicopper oxidase (Lacase) .....	83
4.3.3 $\beta$ -glicosidase .....	85
A. Detecção de Atividade $\beta$ -glicolítica e estimativa do tamanho molecular da enzima .....	85
B. Ensaio para diferentes substratos .....	86
C. Influência do pH e da temperatura .....	87
D. Efeitos de Sais e outros reagentes sobre a atividade enzimática .....	90
E. Determinação dos parâmetros cinéticos para a $\beta$ -glicosidase .....	95
4.4. Preparo de materiais para a expressão do cluster de PKSII .....	96
5. CONCLUSÕES E PERSPECTIVAS .....	99
6. REFERÊNCIAS .....	101

## ANÁLISE E EXPRESSÃO DE GENES DE PKS II E DE CARBOIDRASES PROVENIENTES DE BIBLIOTECAS METAGENÔMICAS

**RESUMO** - O clone metagenômico B5pl37 (inserto de 23 Kb) foi obtido por meio de uma prospecção por PCR para o conjunto gênico de “Policetídeos Sintases tipo II” (PKSII) em uma biblioteca cosmidial de amostra de solo de bosque de eucaliptos (S.E.). As PKSII são responsáveis pela produção de inúmeras moléculas bioativas, dentre os quais se destacam os antibióticos policetídeos. Parte da sequência do inserto (cobertura de 27-43%) apresentou similaridade de 77-78% com bactérias da família *Streptomycetaceae* (ferramenta BlastN, Genbank). Além da PKS II, a anotação gênica permitiu identificar genes relacionados à degradação de compostos lignocelulósicos importantes para a produção de combustíveis de segunda geração (2G): uma  $\beta$ -glicosidase, uma celulase e uma “Multicopper oxidase” (MCO). Considerando a importância industrial destes genes, este trabalho teve como objetivo a caracterização *in silico* e *in vitro/vivo* dos mesmos. As análises *in silico* auxiliaram a identificar a arquitetura do sítio ativo (modelo de estrutura proteica obtida por homologia) e famílias enzimáticas (com base em agrupamentos fenéticos com sequências de enzimas de atividade conhecida). Quanto às análises *in vitro/vivo*, foram realizados experimentos de expressão em vetor pET28a em *Escherichia coli* para as enzimas celulase (linhagens de *E.coli* BL21; C41 e Artics) e  $\beta$ -glicosidase (*E.coli* BL21), enquanto o gene MCO foi submetido à síntese *in vitro* para um estudo posterior. Foi feita a caracterização cinética da  $\beta$ -glicosidase (melhor rendimento na expressão e purificação), que obteve os parâmetros catalíticos:  $V_{max}$  de 0,79  $\mu$ M/min;  $k_m$  de 0,46 mM. Os respectivos valores de temperatura e pH ótimos para a enzima foram 37°C e pH7-7,5; e a enzima foi capaz de manter uma atividade acima de 80% para uma faixa de pH 6,5-8,0 e temperatura 35-40°C. Com relação à PKSII, foram iniciados os processos preparatórios para a expressão do inserto no hospedeiro homólogo especializado para expressão de antibióticos *Streptomyces coelicolor* M1152 (cortesia de Juan P. Gomez-Ezcribano, John Innes Centre, Norwich, Reino Unido). Foi realizada a troca de vetor cosmidial para um vetor conjugativo/integrativo (pSET152/*E.coli* DH10B); seleção dos clones positivos; eletroporação da preparação em bactéria não metilante de DNA (*E.coli* ET12567 pUZ8002), além do cultivo e preparo de suspensão de esporos de *S. coelicolor* M1152 para conjugação Inter gênica e integração sítio específica.

**Palavras-chave:**  $\beta$ -glicosidase, celulase, Metagenômica, “multicopper oxidase”, Policetídeos.

## ANALYSIS AND GENE EXPRESSION OF PKS II AND CARBOHYDRASES FROM METAGENOMIC LIBRARIES

**ABSTRACT** – The Clone metagenomic B5pl37 (insert of 23kb length) was obtained by a PCR prospect for gene set "Polyketide Synthases type II" (PKSII) in a library cosmial eucalyptus grove soil sample (SE). The PKSII are responsible for producing many bioactive molecules, among which the antibiotics were noticeable. Part of the sequence of the insert (27-43% coverage) showed similarity of 77-78% with bacteria *Streptomycetaceae* family (BLASTN tool, Genbank). In PKS II gene annotation was possible to identify genes related to the degradation of lignocellulosic important compounds for the production of second generation biofuel (2G): A  $\beta$ -glucosidase, a cellulose and a Multicopper oxidase (MCO). Considering the industrial importance of these genes, this study aimed to characterize *in silico* and *in vitro / in vivo* thereof. The *in silico* analyzes helped identify the architecture of the active site (protein structure obtained by homology) and enzyme families (based on phenetic groups with sequences of known enzyme activity). The *in vitro / vivo* experiments were performed in pET28a expression vector in *Escherichia coli* for cellulase (*E. coli* strains BL21; Artics and C41) and  $\beta$ -glucosidase (*E. coli* BL21), while the gene MCO in vitro synthesis for a subsequent study it was submitted. We had done the kinetic characterization of  $\beta$ -glucosidase (best performance in the expression and purification), which obtained the catalytic parameters:  $V_{max}$  0.79  $\mu$ M / min;  $k_m$  from 0.46 mM. The respective temperature and the pH optimum for the enzyme was 37 ° C and pH7-7,5; and the enzyme was able to maintain an activity above 80% at a pH of 6.5-8.0 and temperature range 35-40 ° C. With respect to PKSII, the preparatory processes have been started to insert the expression in the same host specialized for expression of *Streptomyces coelicolor* antibiotic M1152 (courtesy of Juan P. Gomez-Escribano, John Innes Centre, Norwich, UK). Cosmidial exchange vector was carried out for conjugative / integrative vector (pSET152 / *E.coli* DH10B); selection of positive clones; Preparation of electroporation in non-DNA methylating bacterium (*E. coli* ET12567 pUZ8002) plus spore suspension culture and preparation for *S. coelicolor* M1152 intergenyc site-specific integration and conjugation.

**Key Words:**  $\beta$ -glicosidase, cellulose, Metagenomics, multicopper oxidase, Poliketides.

## LISTA DE ABREVIATURAS

**BGL:**  $\beta$ -glicosidase;

**CMC:** carboxi-metil-celulose;

**EDTA:** ácido etileno - diamino - tetracético;

**Kan50:** antibiótico Sulfato de Kanamicina na concentração final de 50  $\mu\text{g/mL}$ .

**kDa:** quilodalton;

**PAGE:** eletroforese (não-desnaturante) em gel de poliacrilamida;

**pNP:** *para*-nitrofenol;

**pNP- $\beta$ -Cb:** pNP- $\beta$ -D-celobiosídeo;

**pNP- $\beta$ -Fuc:** pNP- $\beta$ -D-fucopiranosídeo;

**pNP- $\beta$ -G:** pNP- $\beta$ -D-glicopiranosídeo;

**q.s.p:** quantidade suficiente para volume pretendido;

**SDS-PAGE:** eletroforese desnaturante em gel de poliacrilamida;

**SDS:** dodecil sulfato de sódio;

**TCA:** ácido tricloroacético;

**Tris:** hidroximetilaminometano;

**UV:** ultravioleta.

**LISTA DE TABELAS**

Tab.1 - Análise Metagenômica de enzimas celulolíticas .....	9
Tab.2 - Dados sobre algumas glicosil-hidrolases-GHs .....	17
Tab.3 - Bancos de dados e ferramentas úteis para análise in silico de genes .....	38
Tab.4 - Linhagens bacterianas .....	39
Tab.5 - Vetores .....	40
Tab.6 - Oligonucleotídeos para as carbohidrases .....	43
Tab.7 - Preparo do vetor de clonagem pET28a .....	45
Tab.8 - Reação de Restrição enzimática e ligação .....	46
Tab.9 - Testes de Otimização da expressão .....	48
Tab.10 - Resultado para alinhamento entre a CBM metagenômica e CBMs tipo 2.....	70

## LISTA DE FIGURAS

	Página
Fig.1 - Mapa da distribuição de ORFs e Domínios conservados obtidos para amostra metagenômica .....	3
Fig.2 - Fluxograma do desenvolvimento deste trabalho de pesquisa .....	5
Fig.3 - Avaliação de enzimas metagenômicas seguindo critérios de um perfil biocatalítico ideal para o interesse industrial .....	11
Fig.4 - Projeções para o Mercado Global de enzimas .....	12
Fig.5 - Faturamento em bilhões de dólares para os antibacterianos mais utilizados ...	13
Fig.6 - Distribuição das famílias de GHs em Clãs com base na arquitetura do seu sítio ativo .....	16
Fig.7 - Atuação da enzima endo-1,4- $\beta$ -glucanases (EC 3.2.1.4) sobre o metabolismo da celulose .....	19
Fig.8 - Atuação da enzima exo-1,4- $\beta$ -glucanases (EC 3.2.1.91) sobre o metabolismo da celulose .....	19
Fig.9 - Atuação da enzima 1,4- $\beta$ -glucosidases (EC 3.2.1.21) sobre o metabolismo de derivados de degradação de celulose .....	20
Fig.10 - Padrão de mobilidade do conector do módulo catalítico para a EngD e seu CBM .....	21
Fig.11 - Distribuição por família química e por estágio de pesquisas para novas .....	24
Fig.12 - Médias obtidas para o tempo de retenção de novas moléculas de acordo com a área patológica, obtidos no período de 2003 a 2007 .....	25
Fig.13 - Estrutura de policetídeos representativos .....	26
Fig.14 - Esquema de reações e dinâmica catalítica das PKSs I, II e III na síntese da cadeia policetídicica característica de cada grupo .....	29
Fig.15 - Padrão de mobilidade do conector do módulo catalítico para a EngD e seu CBM Alguns exemplos de moléculas de novos fármacos obtidos através de estratégias metagenômica .....	36
Fig.16 - Esquema ilustrativo apresentando características do vetor pET28a .....	45



Fig.17 - Cosmidio pWEB-TNC, que abriga o inserto metagenômico para o clone B5p137 .....	56
Fig.18 - Vetor integrativo conjugativo para transferência intergêneros entre <i>E. coli</i> e <i>Streptomyces sp</i> .....	57
Fig.19 - Árvore fenética para família de glicosil hidrolase GH1.....	60
Fig.20 - Alinhamento entre molde (Bgl3, estrutura 1GNX) e modelo de estrutura para a ORF metagenômica .....	62
Fig.21 - Modelo da estrutura da $\beta$ -glicosidase metagenômica .....	63
Fig.22 - Modelo do sitio ativo estrutura da $\beta$ -glicosidase metagenômica .....	64
Fig.23 - Dendrograma fenético de comparação entre a sequência metagenômica da ORF11 e as sequências de carboidrases obtidas junto ao Uniprot .....	66
Fig.24 - Alinhamento para a celulase metagenômica e outras de estruturas proteicas conhecidas .....	68
Fig.25 - Comparação entre os respectivos módulos estruturais catalíticos e CBMs para Endoglucanase EngD e os obtidos para a ORF11 metagenômica .....	69
Fig.26 - Alinhamento entre CBMs tipo 2 e o modelo de estrutura proposto para a CBM metagenômica .....	71
Fig.27 - Domínios catalíticos obtidos para a ORF11 metagenômica (base de dados <i>Prodom</i> ) .....	72
Fig.28 - . Exemplo da estrutura da ligação organometálica presente na multicopper oxidase CotA .....	73
Fig.29 - Motifs conservados entre as “multicopper oxidases” do tipo Slac relativos aos ligantes de cobre .....	74
Fig.30 - Modelo de estrutura para a MCO .....	75
Fig. 31 - Alinhamento entre as estruturas do molde e modelo para a lacase metagenômica .....	76
Fig.32 - Dendrograma fenético de comparação entre a sequência metagenômica para MCO e sequências presentes no banco de dados “The laccase and multi copper oxidase engineering database) .....	77
Fig.33 - .Confirmação de pareamento correto entre a sequencia do “primer” e a região de interesse para a $\beta$ -glucosidase metagenômica .....	78

Fig.34 - Confirmação de pareamento correto entre a sequencia do “primer”2 e a região de interesse para a Celulase metagenômica .....	78
Fig.35 - Confirmação de pareamento correto entre a sequencia do par de “primer”3 e a região de interesse para a MCO metagenômica .....	78
Fig.36 - Amplificação por PCR do gene para $\beta$ -glucosidase (1516pb) .....	79
Fig.37 - . Amplificação por PCR obtida para a celulase metagenômica .....	79
Fig.38 - . Região “upstream” e “downstream” ao gene da lacase .....	80
Fig.39 - Teste de diferentes tempos de indução, de 0 á 6 horas (T0-T6) .....	81
Fig.40 - Expressão da Celulase 37 °C .....	82
Fig.41 - Expressão da Celulase 30 °C .....	82
Fig.42 - Detecção de peptídeo sinal na sequência proteica para a enzima metagenômica celulase .....	83
Fig.43 - Expressão para Multicopper oxidase .....	78
Fig.44 - Expressão e extração para Multicopper oxidase .....	79
Fig.45 - Perfil eletroforético da enzima $\beta$ -glicosidase metagenômica (Met-Bg) .....	86
Fig.46 - Ensaio enzimático para $\beta$ -glicosidase frente a diferentes substratos .....	87
Fig.47 - . Influência do pH e da temperatura sobre a atividade enzimática da $\beta$ -glicosidase Met-Bg .....	89
Fig.48 - Efeitos de sais sobre a atividade enzimática .....	91
Fig.49 - Presença de grupamento tiol no sítio ativo da $\beta$ -glicosidase metagenômica ...	92
Fig.50 - Efeitos do EDTA (0,1 mM) sobre sais inibidores (0,5 mM) da atividade da $\beta$ -glicosidase Met-Bg .....	93
Fig.51 - Atividade da enzima $\beta$ -glicosidase metagenômica na presença de alcoóis de cadeia curta etanol (A) e Metanol (B) .....	94
Fig.52 - Efeitos da concentração de substrato sobre a velocidade inicial da $\beta$ -glicosidase metagenômica .....	95
Fig.53 - Cosmidio pWEB-TNC, que abriga o inserto metagenômico para o clone B5pl37 .....	98

Fig.54 - Restrição com enzima EcoRI de dois clones da Biblioteca metagenômica S.E. O clone B5 é o alvo deste trabalho .....	98
Fig.55 - Morfologia da colônia do hospedeiro de expressão <i>Streptomyces coelicolor</i> M1152 .....	99

## 1. INTRODUÇÃO

Não restam dúvidas que a Metagenômica, uma abordagem de acesso à microbiota focada no DNA independente de cultivo microbiano, permite potencializar a busca por novas enzimas de interesse industrial, sobretudo por ampliar o acesso à riqueza metabólica presente em micro-organismos (FU et al., 2013; HESS et al., 2011; XIA et al., 2013).

Esta metodologia tem mostrado seu valor na busca por vias complexas de metabólitos bioativos como as dos antibióticos, sobretudo as que são expressas por grandes agrupamentos gênicos. Este é o caso das vias de síntese de policetídeos conhecidas como PKSs, podendo ser do tipo I, II ou III (DENG e BAI, 2006; CHEN *et al.*, 2012; GOMES *et al.*, 2013).

Em relação aos compostos bioativos, a classe dos policetídeos constitui-se de produtos do metabolismo secundário de micro-organismos (algumas espécies de bactérias, fungos, microalgas como os dinoflagelados) e mesmo algumas plantas. Os policetídeos compõem a estrutura química de muitos fármacos: antibacterianos, antifúngicos, antineoplásicos, imunossuppressores, agentes redutores de colesterol, analgésicos, neurotoxinas, antiparasitas e determinantes de virulência (RIX et al., 2002).

Tal versatilidade de funções dos policetídeos é refletida em um faturamento anual superior a U\$ 20 bilhões, chegando a mais de um terço do total das novas drogas lançadas no mercado no período de 2005 a 2007 (WEISSMAN, 2009). Embora os números sejam significativos, entretanto, eles não significam uma abundância sobressalente na disponibilidade de novas moléculas, refreada pelo processo lento de obtenção de formulações farmacológicas finais, contrastando com plasticidade de adaptação dos micro-organismos patogênicos (HAMAD, 2010).

Tendo em conta a demanda constante de novas moléculas, várias são as estratégias adotadas pelos pesquisadores na tentativa de sobrepor o “gap” entre a descoberta de novos compostos e todos os procedimentos até que a mesma se encontre disponível no mercado (HAMAD, 2010; KNELLER, 2010; LING et al., 2015).

A própria complexidade das vias de PKS gera uma fonte bastante interessante para a engenharia genética, uma vez que é possível “criar” novas moléculas através de biossíntese recombinatória por meio da substituição de genes em diferentes funções, desde a escolha dos substratos, tamanho da cadeia policetídica e modificações na estrutura principal, tais como glicosilações, metilações, reduções, ciclizações e aromatisações. A busca por novas vias policetídicas empregando estratégias metagenômicas amplia ainda mais este campo de possibilidades para as pesquisas em torno de novos antibióticos (GOMES *et al.*, 2013; HAMAD, 2010; WEISSMAN, 2009).

Além dos antibióticos, dentre as várias aplicações para enzimas no contexto biotecnológico, pode-se destacar a crescente demanda gerada pelo desenvolvimento de biocombustíveis de segunda geração (2G). Esta abordagem surgiu como alternativa de ampliação da produção desatrelada da expansão das fronteiras agrícolas e do comprometimento de culturas alimentícias, gerando energia a partir de resíduos industriais, geralmente biomassa lignocelulósica (AMORIM *et al.* 2011; DASHTBAN *et al.* 2010; HOWARTH e BRINGEZU 2009; NIGAM e SINGH 2011).

Entretanto, a implementação efetiva da tecnologia 2G como atividade comercial competitiva ainda avança a passos moderados devido ao alto custo de processos envolvendo o uso de enzimas (HUFFER *et al.*, 2012; KLEIN-MARCUSCHAMER *et al.*, 2012; PERALTA-YAHYA *et al.*, 2012). Durante estes processos, a degradação da matéria celulósica é feita por reações multienzimáticas mediada por inúmeras glicosil hidrolases (GHs, como endo-glicanases, exo-glicanases e as  $\beta$ -glicosidases). Uma vez que a biomassa presente em resíduos industriais (ex. bagaço, folhas e a palha de cana-de-açúcar) é complexa, contendo diferentes polímeros além de celulose (lignina, hemicelulose, ceras, etc.), outras enzimas participam do processo em caráter auxiliar, como é o caso das enzimas “Multicopper oxidases” (ou MCO), que facilitam o acesso às fibras de celulose através da degradação da lignina (LEMOS e STRADIOTTO, 2012; LI *et al.*, 2012; PERCIVAL ZHANG, HIMMEL e MIELENZ, 2006).

Neste contexto, o clone metagenômico objeto de estudo deste trabalho (B5pl37) foi obtido por meio de uma prospecção por PCR para agrupamentos

biossintéticos de PKS tipo II, uma via de policetídeos aromáticos (GOMES, 2008). A prospecção foi realizada a partir de uma biblioteca construída previamente utilizando DNA ambiental de solo sob arboreto de eucaliptos, a Biblioteca S.E. (SCHUCH, 2007).

Este inserto metagenômico foi sequenciado pelo método de Sanger, estratégia de “shotgun” e submetido à anotação manual de fases abertas de leitura (ORFs, Figura 1). Os resultados indicaram um agrupamento gênico bastante promissor do ponto de vista biotecnológico, apresentando tanto genes para policetídeos (ORF4 e ORFs de 14-22), quanto genes relacionados ao metabolismo de compostos lignocelulósicos: uma  $\beta$ -glicosidase (ORF10), uma celulase/quitinase (ORF11) e uma “Multicopper oxidase” ou MCO (ORF13).

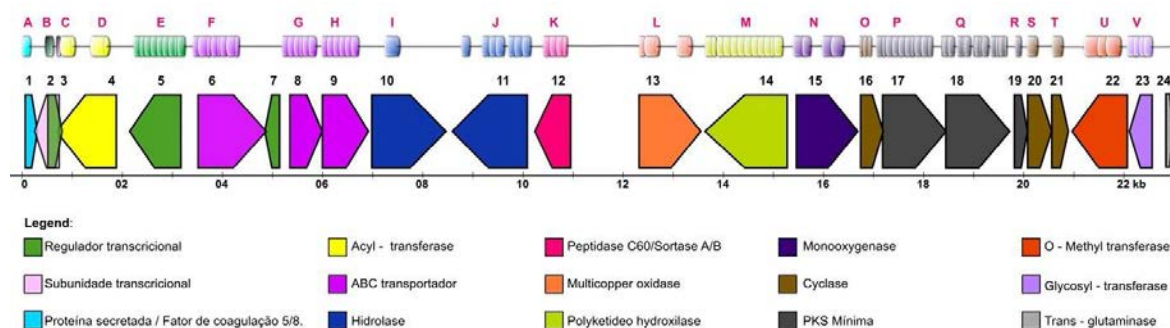


Figura 1. Mapa da distribuição de ORFs (1-22: “ORFfinder”/ “Blastp”) e Domínios conservados (A-V: Base de domínios “PRODOM”) obtidos para amostra metagenômica de biblioteca de solo sob bosque de eucaliptos. A sequência foi obtida por sequenciamento tipo Sanger, estratégia de Biblioteca “shotgun” (GOMES, 2011).

Considerando que a pesquisa dos genes relacionados ao metabolismo de biomassa lignocelulósica se constitui em um importante foco da biotecnologia, optou-se por desenvolver um trabalho de caracterização *in silico* e *in vitro* para as três ORFs relacionadas à atividade de carboidrase (ORFs 10, 11 e 13).

É inegável o apelo intrínseco que também permeia a temática dos antibióticos sendo compostos extremamente importantes para a saúde humana e animal. Entretanto, devido ao caráter mais complexo da manipulação de “constructos” de expressão de “clusters” gênicos de antibióticos quando comparado à expressão de

genes individuais (KIESER et al., 2000), a parte experimental deste trabalho ainda está em curso, em fase de preparo de hospedeiro e vetores de transferência.

A seguir, encontram-se descritos os objetivos (item 1.1), as atividades desenvolvidas e as perspectivas deste trabalho (item 1.2).

## **1.1. Objetivos:**

**1.1.1. Geral:** Estudar genes de interesse industrial (via de PKSII e carboidrases) provenientes de biblioteca metagenômica de solo de bosque de eucaliptos (Biblioteca Metagenômica SE), empregando métodos “*in silico*” e/ou “*in vitro*”

### **1.1.2. Específicos:**

- Análise dos genes obtidos para as vias de PKSII e de carboidrases metagenômicas empregando métodos *in silico*;
- Clonagem e expressão heteróloga de genes individuais de carboidrases em *Escherichia coli*;
- Ensaio bioquímico e de cinética enzimática para enzimas super-expressas;
- Preparo do inserto contendo a via de PKSII para a expressão em *S. coelicolor* M1152, através de subclonagem em vetor de integração e bactéria não metilante do DNA;
- Preparo de vetores, células competentes e suspensão de esporos para a posterior clonagem dos genes de antibiótico.

## **1.2. Fluxograma das Atividades desenvolvidas**

A sequência de eventos que antecedem este trabalho, as etapas concluídas do mesmo, bem como suas perspectivas, podem ser observadas na figura 2. O bosque de eucaliptos que originou a biblioteca S.E. foi plantado no campus da Unesp Jaboticabal em 1969, e foi mantido sem tratamentos culturais até a ocasião da coleta das amostras. Um estudo de diversidade microbiana com base na sequência de *16SrDNA* (SILVEIRA et al., 2006) indicou a presença expressiva de *Actinobacteria* neste ambiente, e esta foi a motivação para a construção da

Biblioteca S.E., uma vez que estes micro-organismos são conhecidos pela produção de diversas moléculas de policetídeos.

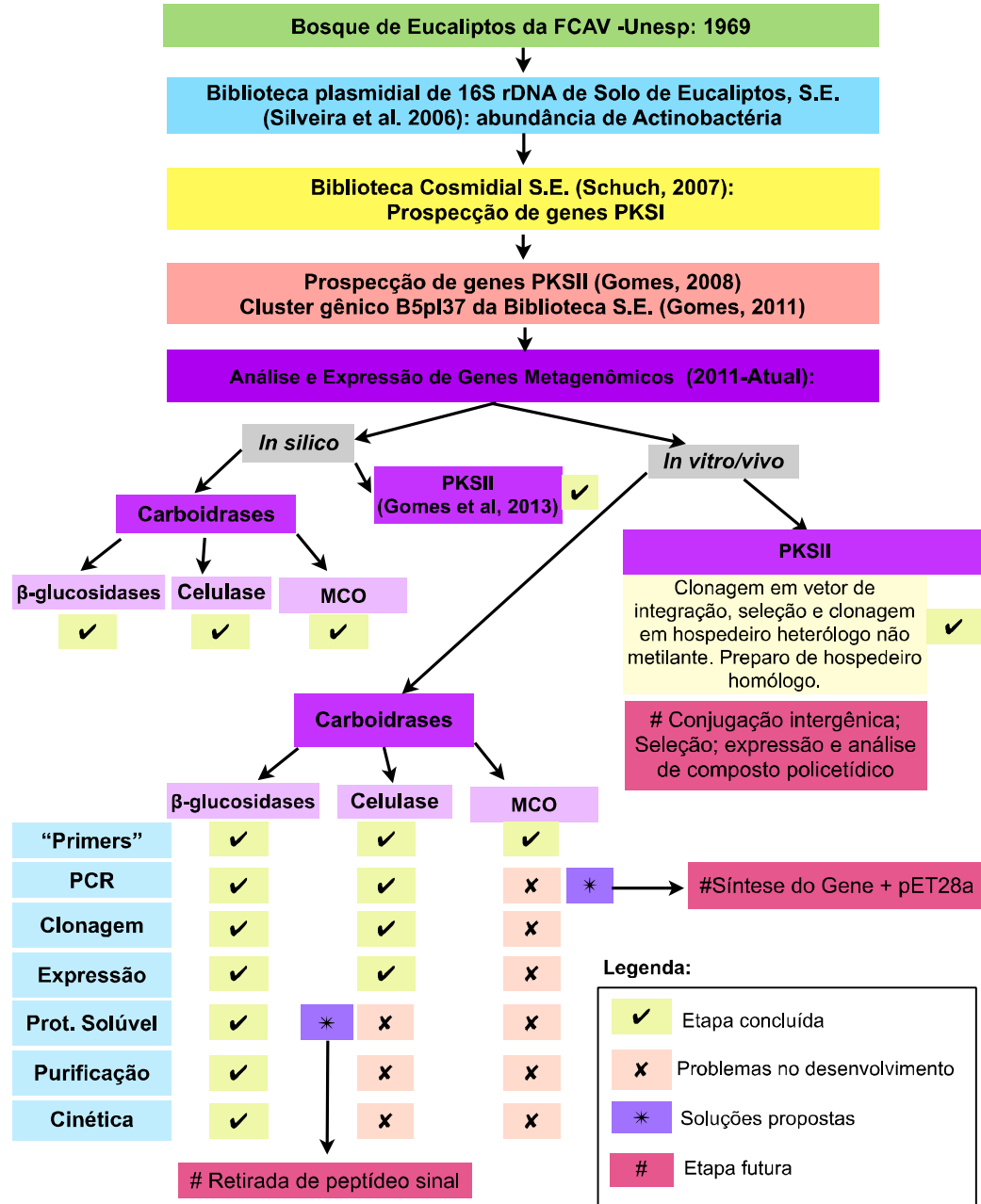


Figura 2. Fluxograma do desenvolvimento deste trabalho de pesquisa. O esquema ilustra os passos desde a origem da amostra até as etapas planejadas para seu desenvolvimento futuro. MCO: "Multicopper oxidase".



## 2. REVISÃO DE LITERATURA

### 2.1. Acesso ao potencial de micro-organismos incultiváveis

A grande maioria dos micro-organismos possuem necessidades complexas demais para serem reproduzidas em ambientes artificiais, apresentando sinergismos/sintrofismos dependentes de todo um conjunto de fatores físico-químicos produzidos pelo metabolismo de outros micro-organismos, carecendo de suplementação nutricional de aminoácidos, fontes de carbono e outros fatores de crescimento; solubilidade de sais, pH e temperatura ideais, além da disponibilidade de formas acessíveis de nitrogênio e fósforo, redução de resíduos metabólicos nocivos, criação de ambientes aeróbicos ou anaeróbicos, dentre outras condições específicas do ambiente (MOREIRA e SIQUEIRA, 2006).

Técnicas de microscopia e sequenciamento demonstraram que a diversidade genética de micro-organismos na natureza é muito maior que a refletida nas coleções e bancos de linhagens obtidas por isolamento e cultivo artificial, estimando que apenas uma pequena parcela destes, em torno de 1%, é cultivável pelas técnicas padrões em meios de cultura em placa de Petri (AMANN, LUDWING e SCHLEIFER, 1995; COMMITTEE ON METAGENOMICS, 2007; STEWART 2012; LING et al 2015).

Buscando um acesso a este tesouro desconhecido, algumas abordagens foram desenvolvidas dentre as quais podemos destacar a cultura de consórcio enriquecido; o cultivo *in situ*; e os estudos em metagenômica, ou seja, o estudo dos genomas diretamente extraído dos ambientes naturais sem cultivo prévio (STEWART, 2012).

A técnica de consórcio enriquecido utiliza uma pressão seletiva, como por exemplo, fontes alternativas de carbono ou nitrogênio para permitir o crescimento de uma parcela da comunidade microbiana. Esta pressão seletiva visa recriar uma condição do ambiente na qual estes microrganismos foram coletados, e pode representar uma via metabólica de interesse. A técnica auxilia o entendimento de parte da fisiologia desta comunidade através de estudos bioquímicos, ainda que não

seja possível se obter cultura pura de todos os organismos presentes (PAIXÃO et al. 2010; STEWART, 2012; ROMANO et al., 2013).

O cultivo *in situ*, por sua vez, é uma abordagem que visa cultivar o organismo em seu *habitat* natural, de forma que o ambiente possa fornecer os fatores de crescimento necessários ainda desconhecidos. Uma destas metodologias envolve o uso um dispositivo denominado "Ichip" (NICHOLS et al. 2010), que se constitui em uma lâmina com microcâmaras nas quais uma diluição de amostra ambiental é colocada de forma a distribuir uma única célula por câmara. Em seguida, estas câmaras são revestidas com uma membrana semipermeável, e o dispositivo é devolvido ao ambiente. Posteriormente o Ichip é recuperado, e as colônias puras obtidas podem ser utilizadas para visualização em microscópio, sequenciamento, estoques e ensaios funcionais (degradação de um composto específico ou detecção de halos de inibição para antibióticos em potencial). Esta metodologia tem permitido a recuperação de até 50% dos micro-organismos presentes em uma determinada amostra. Aliado ao sequenciamento destes organismos, o método permite entender melhor sobre os fatores limitantes do cultivo, favorecendo o desenvolvimento de novas técnicas de cultivo artificial (NICHOLS et al. 2010; STEWART, 2012; LING et al 2015).

A metagenômica, por sua vez é uma metodologia que se fundamenta em análises genéticas e funcionais da microbiota de forma independente de cultivo, interagindo diretamente sobre o "pool" de moléculas de ácidos nucléicos do ambiente (e mesmo um consórcio já estabelecido), e se constitui em uma abordagem extremamente útil para a biotecnologia, uma vez que permite o acesso à coleção de genomas de amostras ambientais praticamente em sua totalidade (HANDELSMAN *et al.*, 1998; WOOLEY et al. 2010; STEWART, 2012).

Estudos em metagenômica podem envolver a extração de DNA e construção de bibliotecas através da clonagem deste material em cosmídeos, fosmídeos ou mesmo em vetores que comportem grandes insertos, tais como BACs e YACs (Cromossomos Artificiais Bacterianos e de Levedura, respectivamente). Os genes de interesse ecológico e/ou biotecnológico são então acessados nestas bibliotecas através de prospecção por técnicas de detecção funcional, como as realizadas em placas de meio de cultura com substrato específico, ou buscando características

morfológicas distintas das colônias. Outros métodos de busca se fundamentam em técnicas de genética molecular, como PCR, DGGE ou hibridização de sondas (COMMITTEE ON METAGENOMICS, 2007; GOMES et al., 2013; CASTRO et al., 2014).

Recentemente, o sequenciamento de nova geração e o avanço das metodologias de análise de dados proveniente destes novos métodos para montagem de metagenomas tem permitido explorar esta riqueza metabólica sem a necessidade de construção de bibliotecas. O sequenciamento nos permite ter um acesso virtual aos genes presentes na microbiota, os quais podem ser avaliados *in silico* quanto ao seu papel biológico ou seu interesse de aplicação biotecnológica. Posteriormente, o gene escolhido pode ser amplificado da amostra original, ou mesmo ser submetido à síntese *in vitro* e podem ser expressos em hospedeiros homólogos ou heterólogos (WOOLEY et al. 2010; STEWART, 2012; SIM & KIM, 2015).

Um exemplo comparativo entre os rendimentos em número de genes obtidos por estratégias de prospecção funcional de bibliotecas e estratégia de sequenciamento de nova geração de metagenomas pode ser observado abaixo (Tab.1). Os trabalhos exemplificados tiveram como objetivo a busca de genes relacionados à degradação de celulose.

Como pode ser observado na tabela 1, o sequenciamento de nova geração ampliou muito a nossa capacidade de encontrar genes. Por outro lado, esta imensa quantidade de dados desafia os pesquisadores a encontrar métodos cada vez mais eficientes de triagem *in silico*, sejam estes o reconhecimento de domínios conservados, sítios catalíticos ou estrutura proteica diferenciada das disponíveis até então.

**Tabela 1 – Análise Metagenômica de enzimas celulolíticas: da prospecção ao sequenciamento em larga escala**

<b>Amostra</b>	<b>Abordagem</b>	<b>Enzima alvo</b>	<b>Referência</b>
<b>Prospecção Funcional de bibliotecas metagenômica</b>			
Biblioteca metagenômica fosmidial de trato digestivo de <i>Macropus eugenii</i>	Prospecção funcional e por hibridação	Celulases (17); endo-hemicelulases (13); desramificadoras (1); degradação de oligossacarídeos (20).	POPE et al., 2010
Biblioteca cosmidial de cécum de coelho.	Prospecção funcional	Celulases (4); degradação de oligossacarídeos (7).	FENG et al., 2007
Biblioteca cosmidial de rúmen de búfalo.	Prospecção funcional	Celulases (14); endo-hemicelulases (2).	DUAN et al., 2009
Biblioteca fosmidial de rúmen de búfalo.	Prospecção funcional	Celulases (6); endo-hemicelulases (3); degradação de oligossacarídeos (2); outras GHs (1).	GUYEN et al., 2012
<b>Sequenciamento de nova geração para amostras ambientais</b>			
Amostra ambiental de rúmen de Yac (consórcio)	Pirosequenciamento	Celulases (2.114); endo-hemicelulases (4.929); desramificadoras (1.627); degradação de oligossacarídeos (10.857); outras GHs (18.036);	DAI et al., 2012
Trato digestivo <i>Macropus eugenii</i>	Pirosequenciamento	Celulases (10); endo-hemicelulases (28); desramificadoras (42); degradação de oligossacarídeos (184); outras GHs (293).	POPE et al., 2010
Rúmen bovino	Illumina (total de 268 Gb)	Celulases (2.365); endo-hemicelulases (2.360); desramificadoras (1.381); degradação de oligossacarídeos (6.591); outras GHs (15.058).	HESS et al., 2011
Rúmen bovino comparativo (três animais)	Pirosequenciamento	Celulases (13); endo-hemicelulases (39); desramificadoras (99); degradação de oligossacarídeos (549); outras GHs (257).	BRULC et al., 2009
Lodo celulolítico termofílico	Illumina (Consórcio)	Celulases (21); endo-hemicelulases (30); degradação de oligossacarídeos (83); outras GHs (100).	XIA et al., 2013

A despeito da facilidade do sequenciamento direto das amostras, nem sempre esta é a melhor abordagem. Quando o alvo das pesquisas corresponde a grandes “clusters” gênicos (como é o caso de vias de PKS) as bibliotecas metagenômicas se caracterizam como a melhor opção, uma vez que permitem um acesso físico (os fragmentos de DNA são mantidos em bibliotecas de clones em estoques) destas vias, enquanto o sequenciamento por si só nos fornece um acesso virtual a este potencial (obtemos apenas a sequência do DNA).

Sendo assim, a escolha da técnica (ou conjunto destas) deve ser feita de

acordo com as características do estudo que se pretende fazer, e a utilização da metagenômica como abordagem de busca por genes de interesse amplia em grande escala a perspectiva de sucesso destas estratégias, uma vez que abrange uma diversidade não cerceada pela possibilidade ou não de cultivo laboratorial. Visto que a ampla diversidade microbiana do solo constitui-se de um enorme “pool” genético e biológico que pode ser explorado para a descoberta de novos genes, vias metabólicas inteiras e seus produtos, a metagenômica oferece um caminho para acessar a diversidade microbiana de um ambiente em sua totalidade; (CASTRO et al. 2014; GOMES et al, 2013; BALTZ, 2006, COWAN, 2005).

## **2.2. Potencial industrial microbiano**

A microbiota representa um imenso reservatório de enzimas e metabólitos bioativos para utilização industrial, com grande potencial de expansão (LING et al., 2015; CASTRO et. al., 2014, BCC Research, 2013; HAMAD 2010 STUART, 2013; WOOLEY et al., 2010).

Entretanto, os processos industriais carecem de uma constante busca por um equilíbrio entre o custo/benefício de seus meios de produção, sem o qual é impossível viabilizar determinada atividade comercial. Como não poderia ser diferente, a utilização de enzimas pelo setor segue as mesmas normas de conduta para aplicação, necessitando que a busca por enzimas mais eficientes, de baixo custo de produção, e com capacidade de resistência a uma série de fatores inerentes do processo industrial seja a principal meta das pesquisas voltadas para o desenvolvimento de formulações enzimáticas para a indústria (Fig. 3).

Essa busca por um equilíbrio se torna indispensável para o cenário atual de crise para o setor canavieiro, que engloba restrições orçamentárias tanto para a produção de etanol de primeira quanto de segunda geração (Nova Cana, acesso em 06 de abril de 2015; disponível em [http://www.novacana.com/n/etanol/2-geracelulosecao-/euforia-etanol-celulosico-acaba-desafios-larga-escala-310315/?kmi=trinkabio2005@yahoo.com.br&utm\\_source=Etanol&utm\\_campaign=ed8f9e8a0c-](http://www.novacana.com/n/etanol/2-geracelulosecao-/euforia-etanol-celulosico-acaba-desafios-larga-escala-310315/?kmi=trinkabio2005@yahoo.com.br&utm_source=Etanol&utm_campaign=ed8f9e8a0c-)

Euforia\_E2G\_acaba\_Mar\_31\_2015&utm\_medium=email&utm\_term=0\_9fda3940f1-ed8f9e8a0c-71133353).

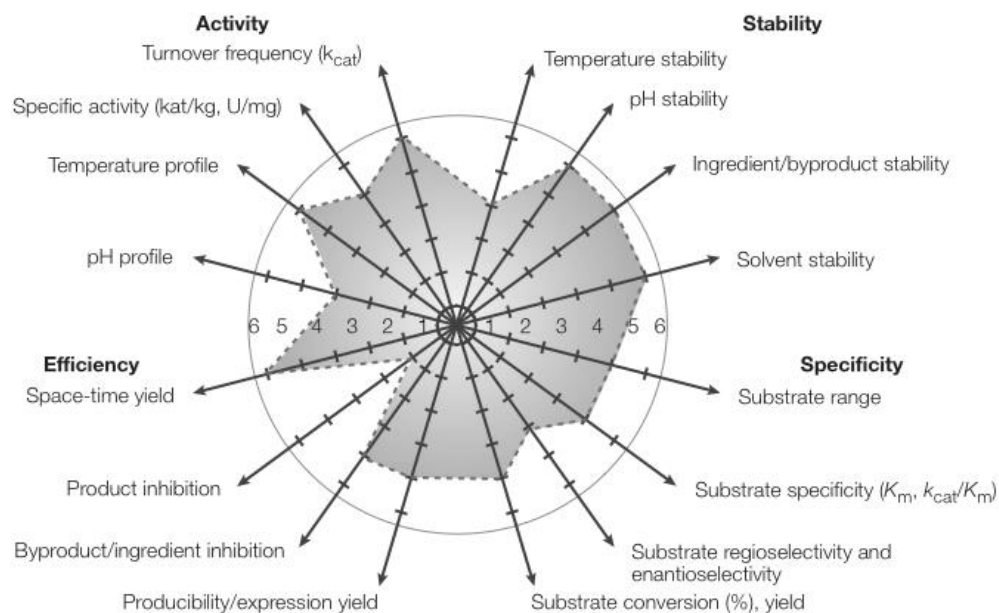


Figura 3. Avaliação de enzimas metagenômicas seguindo critérios de um perfil biocatalítico ideal para o interesse industrial (fonte: LORENZ e ECK, 2005).

Conforme pode ser visualizado abaixo na Figura 4, as melhores projeções de crescimento no mercado de enzimas são para a classe das celulases: as perspectivas são de duplicação de divisas até o ano de 2017, alcançando um rendimento de 400 milhões de dólares no mercado enzimático mundial. O principal uso das celulases está relacionado com expansão de novas tecnologias para a produção de combustíveis de segunda geração a partir de resíduos industriais (Moraes, 2011; Nigam & Singh, 2011; Howarth & Bringezu, 2009 ;Vasudevan & Briggs, 2008).

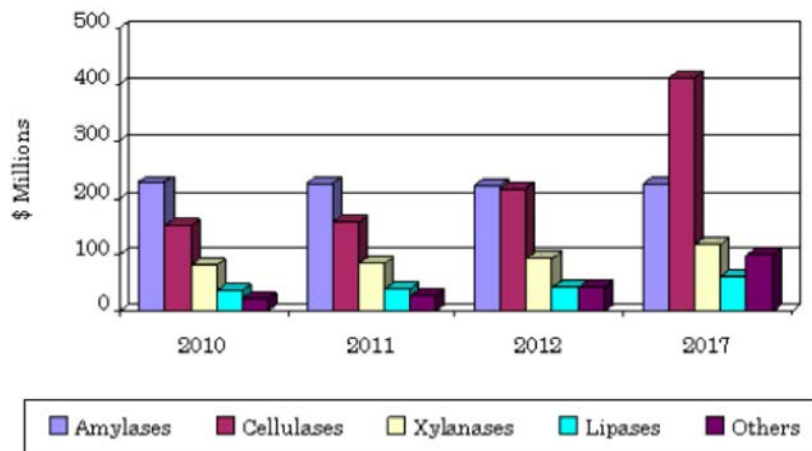


Figura 4. Projeções para o Mercado Global de enzimas. Fonte BCC Research (2013).

Dentre os bioativos de origem microbiana, por sua vez, os fármacos (sobretudo antibióticos e os antitumorais) movimentam bilhões de dólares na economia mundial, haja vista nossa constante necessidade por novas moléculas para o controle de doenças em humanos e animais (LING et. al., 2015; CASTRO et al. 2014; GOMES et al., 2013; HAMAD 2010). A figura 5 ilustra os rendimentos obtidos para cada um dos principais princípios ativos dos fármacos mundialmente comercializados referentes ao ano de 2009, com destaque para os policetídeos da classe das cefalosporinas, responsáveis por 11,9 bilhões de dólares.

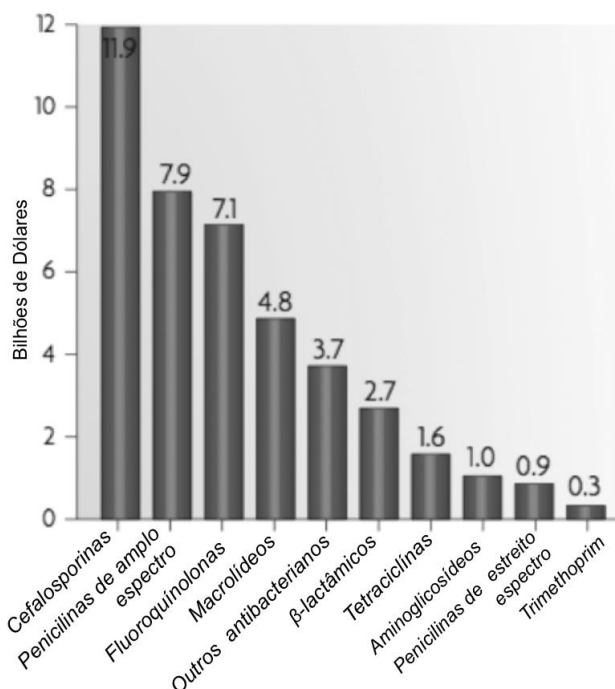


Figura 5. Faturamento em bilhões de dólares para os antibacterianos mais utilizados (Fonte: HAMAD, 2010).

Estes dados nos permitem visualizar o quanto os produtos de origem microbiana exercem um papel forte na economia global, sendo indispensáveis para a vida moderna.

### 2.3. Enzimas envolvidas na degradação de biomassa lignocelulósica e a produção de etanol de segunda geração

Tendo em vista a demanda especializada exercida pela indústria, cada vez mais pesquisas são empreendidas para facilitar a busca de novas enzimas. Como exemplo, podemos citar a busca por novas fontes de energia que tem atraído o olhar da comunidade científica em todos os âmbitos da pesquisa, impulsionada por uma demanda cada vez mais crescente (VASUDEVAN e BRIGGS, 2008).

Se considerarmos os dados estatísticos referentes ao ano de 2013 divulgados pela “BP Statistical Review of World Energy” (<http://www.bp.com/content/dam/bp/pdf/Energy-economics/statistical-review->



2014/BP-statistical-review-of-world-energy-2014-full-report.pdf edição 63 de junho de 2014) o nível de consumo diário atual de energia atinge 91 milhões de barris de óleo diesel (41 milhões de toneladas), com os maiores consumidores sendo os Estados Unidos (20%) e China (12%). Para o mesmo período, o consumo de energia renovável (incluindo energia eólica, geotermal, de biomassa e outros resíduos) atingiu ao equivalente de 279 milhões de toneladas de óleo diesel.

Ainda que seja claro o papel altamente poluidor dos combustíveis fósseis, com o passar do tempo ficou evidente que os impactos sócio-econômico-ambientais também não isentavam totalmente os biocombustíveis, que exerciam sua parcela de culpa sobre o aumento dos custos de alimentos pela diminuição da área cultivada ou desvios de culturas alimentícias para sua produção (HOWARTH & BRINGEZU, 2009). Neste cenário, surgiu uma nova visão de ampliação da produção de combustíveis não ligadas à expansão da fronteira agrícola: a produção de energia limpa fundamentada em biocombustíveis a partir de biomassa lignocelulósica nos combustíveis de Segunda Geração (“second generation biofuel” ou 2G).

O Brasil é o maior produtor mundial de etanol, com uma vasta experiência na escala de produção, o que faz com que assuma papel de referência para o desenvolvimento do setor (NATURE, 2009). A produção anual de etanol no país já atinge uma capacidade de 28 bilhões de litros (IBGE/LSPA, 2013). Dentre os subprodutos da produção do etanol no país, destaca-se a produção de bagaço, que gira em torno de 160 milhões de toneladas. Essa biomassa lignocelulósica (atualmente empregue na geração de energia térmica para a própria usina) possui um grande potencial no aumento da produção de combustíveis, de forma que seu uso pode incrementar a produção de etanol brasileiro em 40% (NOVAENZYME, 2012).

Entretanto, os desafios do custo de produção do etanol 2G são ainda uma importante barreira para esse avanço, seja qual for a fonte de matéria-prima inicial. Neste sentido, o etanol 2G vem captando esforços para maximizar sua consolidação, sobretudo na diminuição do custo e aumento da eficiência da digestão enzimática utilizada no processo, através da produção mais rentável de enzimas tais como endo-/exo- celulasas e  $\beta$ -glicosidasas (carboidrases).

Como a degradação da matéria lignocelulósica é feita por reações multienzimáticas, vários são os pontos passíveis de interferência pela engenharia genética em busca de aumentar a produção, com destaque para as Glicosídeo-Hidrolases (GHs). As GHs englobam 132 famílias, dentre outras enzimas não classificadas segundo a base de dados “The Carbohydrate-Active Enzymes database” (CAZY: <http://www.cazy.org/>). Essa base de dados permite um acesso online e atualizado de classificação das carboidrases em famílias fundamentada em similaridade de sequências, bem como fornece links a informações de atuação e estrutura 3D proteica para enzimas relacionadas à montagem, modificação ou quebra de oligo ou polissacarídeos. A nomenclatura adotada pela base Cazy é uma referência para a comunidade científica, permitindo uma maior padronização das informações geradas e sua comparação com demais dados da literatura. Nesta base de dados estão reunidas informações sobre as seguintes classes de enzimas:

- Glicosil-hidrolases (GHs): hidrólise e/ou rearranjo de ligações glicosídicas;
- Glicosil-Transferases (GTs): formação de ligações glicosídicas;
- Polissacarídeo-Liases (PLs): clivagem não hidrolítica de ligações glicosídicas;
- Carboidrato Esterase (CEs): hidrólise de ésteres de carboidrato.
- Atividade auxiliar (AAs): enzimas tipo redox que agem em conjunto com as carboidrases.
- Modulos ligantes de Carboidrato (Carbohydrate binding module CBMs): adhesion to carbohydrates.

Este sistema de classificação foi proposto por Henrissat e colaboradores (1996), e leva em conta a similaridade entre as sequências e os padrões de dobramento entre diversas enzimas, nos chamados HCA (Análise de Clusters Hidrofóbicos). Assim, as 132 famílias de GHs foram estabelecidas através da análise da restauração da homologia primária nas regiões adjacentes ao sítio ativo das sequências por filogenia (análise de similaridade entre sequências) e análise na distribuição das estruturas secundárias  $\alpha$ -hélices e folhas- $\beta$  no enovelamento do bolsão catalítico. Com base na arquitetura do enovelamento do sítio ativo e

similaridades nos mecanismos de atuação, algumas GHs são ainda reagrupadas nos Clãs de GH-A até GH-N, de acordo com a figura 6.

► GH Clans of Related Families

<b>GH-A</b>	$(\beta/\alpha)_8$	1	2	5	10	17	26	30	35	39	42	50	51	53	59	72	79	86	113	128
<b>GH-B</b>	$\beta$ -jelly roll	7	16																	
<b>GH-C</b>	$\beta$ -jelly roll	11	12																	
<b>GH-D</b>	$(\beta/\alpha)_8$	27	31	36																
<b>GH-E</b>	6-fold $\beta$ -propeller	33	34	83	93															
<b>GH-F</b>	5-fold $\beta$ -propeller	43	62																	
<b>GH-G</b>	$(\alpha/\alpha)_6$	37	63																	
<b>GH-H</b>	$(\beta/\alpha)_8$	13	70	77																
<b>GH-I</b>	$\alpha+\beta$	24	46	80																
<b>GH-J</b>	5-fold $\beta$ -propeller	32	68																	
<b>GH-K</b>	$(\beta/\alpha)_8$	18	20	85																
<b>GH-L</b>	$(\alpha/\alpha)_6$	15	65	125																
<b>GH-M</b>	$(\alpha/\alpha)_6$	8	48																	
<b>GH-N</b>	$\beta$ -helix	28	49																	

Figura 6. Distribuição das famílias de GHs em Clãs com base na arquitetura do seu sítio ativo.

No caso específico da degradação de materiais lignocelulósicos, as famílias mais importantes são representadas pelos seguintes grupos funcionais. (DAI *et al.*, 2012).

- Celulases (GHs 5-7, 9, 44, 45 e 48);
- Endo-hemi-celulases (GHs 8, 10-12, 26, 28 e 53);
- Enzimas desramificadoras (GHs 51, 54, 62, 67 e 78);
- enzimas degradadoras de oligossacarídeos (GHs 1-3, 29, 35, 38, 39, 42, 43, 52).

Entretanto, quando se leva em conta as informações sobre sequências de glicosil-hidrolases da base Cazy, pode-se perceber que existe uma enorme quantidade de dados de sequência para enzimas de potencial interesse industrial ainda não caracterizadas funcionalmente, como pode ser exemplificado nos dados abaixo (Tab.2).

Tabela 2- Dados sobre algumas glicosil-hidrolases-GHs (Fonte: CAZY, <http://www.cazy.org/>)

<b>GH</b>	<b>Sequências disponíveis</b>	<b>Caracterizadas experimentalmente</b>	<b>Estruturas 3D proteicas</b>
GH1	5487	279	41
GH2	3882	134	14
GH5	4514	398	43
GH7	4990	73	10
GH10	1678	263	28
GH51	943	64	6

Desta forma, o sucesso das pesquisas envolvendo carboidrases não deve se restringir apenas à busca de novas sequências gênicas, e sim abranger, sobretudo a busca por métodos efetivos de seleção de melhor enzima potencial. Quando a conjuntura de fatores envolvem uma imensa quantidade a de dados contrapondo métodos morosos de identificação experimental de propriedades catalíticas, a utilização de métodos *in silico* de análises se interpõem para equilibrar melhor as chances de êxito na obtenção de um produto de aplicação biotecnológica.

Outro desafio para a produção de etanol 2G reside na dificuldade em gerar uma formulação comercial efetiva para aplicação industrial. Isso porque a eficiência desta dependerá não apenas da determinação das condições ideais para a cinética enzimática individual, mas também em um ajuste de condições que favorecerão todas as enzimas que serão utilizadas em uma mesma etapa do processo. Essa adaptação de condições é possível em função da plasticidade de determinadas enzimas em trabalhar em amplas faixas de pH, temperatura, concentração de substrato e salinidade, sendo que esta versatilidade é uma condição desejada na prospecção de novas enzimas (KLEIN-MARCUSCHAMER *et al.*, 2012; BRULC *et al.* 2009; LORENZ e ECK, 2005).

Além de formulações enzimáticas, existem pesquisas em biocombustíveis focadas em aperfeiçoar um micro-organismo para alcançar o perfil ideal como “biofábrica” de biocombustíveis. Dentre outras características, o micro-organismo ideal deverá possuir alta capacidade de utilização de substrato e processamento, vias rápidas e constitutivas de transporte de açúcares, capacidade de tolerância a inibidores e produtos, e alto fluxo metabólico, além de gerar um único produto no processo de fermentação (ALPER e STEPHANOPOULOS, 2009).

## 2.4. Vias de Degradação de biomassa lignocelulósica

A celulose se constitui na fonte renovável de energia mais abundante produzida na biosfera, com uma produção anual de 100 bilhões de toneladas de matéria seca (Zang et al., 2006). A degradação da celulose por celulasas e celulosomas representa o maior fluxo dos reservatórios de carbono fixado até carbono atmosférico, o que lhe confere extrema importância nos processos agrícolas e de tratamento de resíduos, bem como no potencial de utilização para produção de fontes alternativas aos combustíveis fósseis (Zang et al., 2006).

Bioquimicamente, a celulose se constitui homopolissacarídeo de resíduos de  $\beta$ -D-glicopiranosídeos ligados por ligações glicosídicas  $\beta$ -1 $\rightarrow$ 4, nos quais a glicose é a menor unidade repetitiva, em um polímero linear que pode conter até 10 mil resíduos de glicose. Se as regiões cristalinas da celulose são impermeáveis à água e resistentes a degradação química e biológica. Na maioria dos casos, as enzimas que atuam diretamente sobre a degradação da celulose possuem três tipos básicos de ação, sendo estes: endo-1,4- $\beta$ -glucanase; exo-1,4- $\beta$ -glucanases (ou celobiohidrolase) e 1,4- $\beta$ -glucosidases. Outras enzimas, como a lacase podem atuar por meio da degradação de lignina e outros componentes facilitando o acesso ao material celulósico (NASCIMENTO et al. 2010).

De forma geral, a degradação de biomassa lignocelulósica envolve a interação entre endo-celulasas, exo-celulasas e  $\beta$ -glicosidases (LYND et al. 2002).

As endo-1,4- $\beta$ -glucanases (EC 3.2.1.4) atuam randomicamente nas ligações internas da cadeia de celulose (Fig.7).

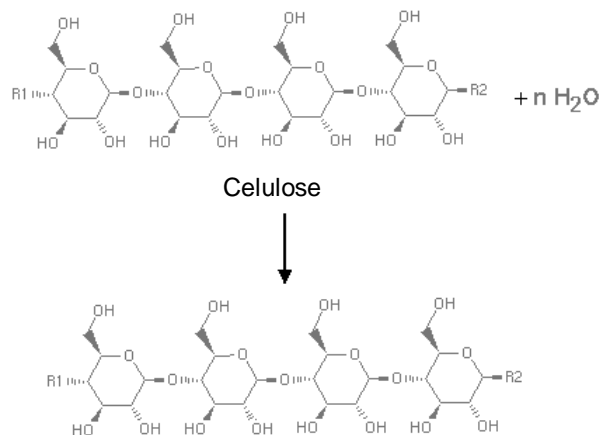


Figura 7. Atuação da enzima endo-1,4- $\beta$ -glucanases (EC 3.2.1.4) sobre o metabolismo da celulose (Fonte: BioCyc).

Já as exo-1,4- $\beta$ -glucanases ou celobiohidrolases (EC 3.2.1.91) atuam liberando unidades de celobiose de extremidades redutoras e não redutoras da cadeia, não sendo capazes de quebras no interior da molécula (Fig.8).

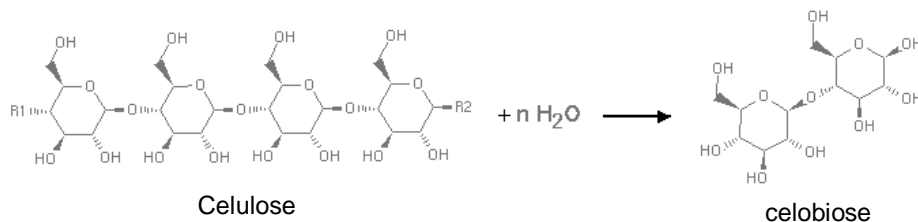


Figura 8. Atuação da enzima exo-1,4- $\beta$ -glucanases (EC 3.2.1.91) sobre o metabolismo da celulose. (Fonte: BioCyc).

Por sua vez, as 1,4- $\beta$ -glucosidases (EC 3.2.1.21) atuam sobre as moléculas de celobiose e celobio-oligossacarídeos para liberação de unidades simples de glicose, atuando, sobretudo no final da via das celulases e de outros açúcares (Fig.9).

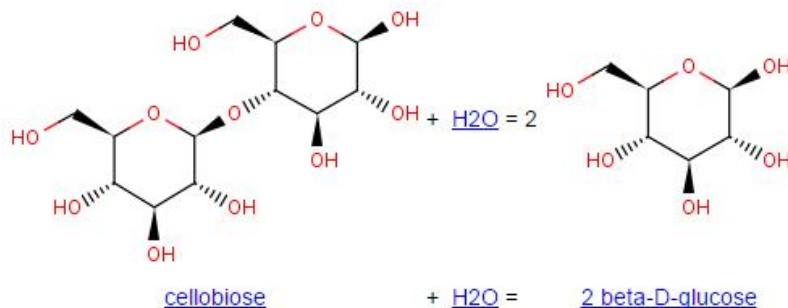


Figura 9. Atuação da enzima 1,4- $\beta$ -glucosidases (EC 3.2.1.21) sobre o metabolismo de derivados de degradação de celulose (celobiose e celbio-oligossacarídeos) (Fonte: BioCyc).

Além de seu papel no setor bioenergético, as enzimas  $\beta$ -glucosidase possuem aplicação para a indústria de alimentos, através da melhoria das características organolépticas de sucos e vinhos, através da quebra de glicoconjugados flavorizantes (MATEUS, 2009).

Tanto a endo- quanto a exo-1,4- $\beta$ -glucanase são inibidas por altas taxas de celobiose, o que torna a ação das enzimas 1,4- $\beta$ -glucosidases um fator limitante importante para a atividade celulolítica. Por outro lado, a maioria das  $\beta$ -glucosidases é altamente inibida pela glicose (produto), de forma que enzimas tolerantes a elevadas concentrações são desejáveis, favorecendo uma completa sacarificação da celulose em monômeros de glicose.

O entendimento das características estruturais e dinâmicas que as celulasas empregam para se ligar às cadeias de celulose e hidrolisá-la a açúcares fermentescíveis é, sem dúvida, um dos maiores objetivos das pesquisas que buscam a engenharia de celulasas para conversão de biomassa em larga escala. Neste sentido, Bianchetti e colaboradores (2012) realizaram uma importante contribuição através da elucidação de estruturas cristalográficas para a endoglucanase EngD (*Clostridium cellulovorans*, resolução de 1 Å) nos estados ligado e desligado do substrato (celotriose). Essa endoglucanase apresenta um padrão modular, com um domínio catalítico N-terminal, e um domínio C-terminal “carbohydrate-binding module” (CBM), ao qual é ligado através de um conector

flexível que garante mobilidade à proteína em solução (comprovado por SAXS), conforme pode ser visualizado na figura 10.

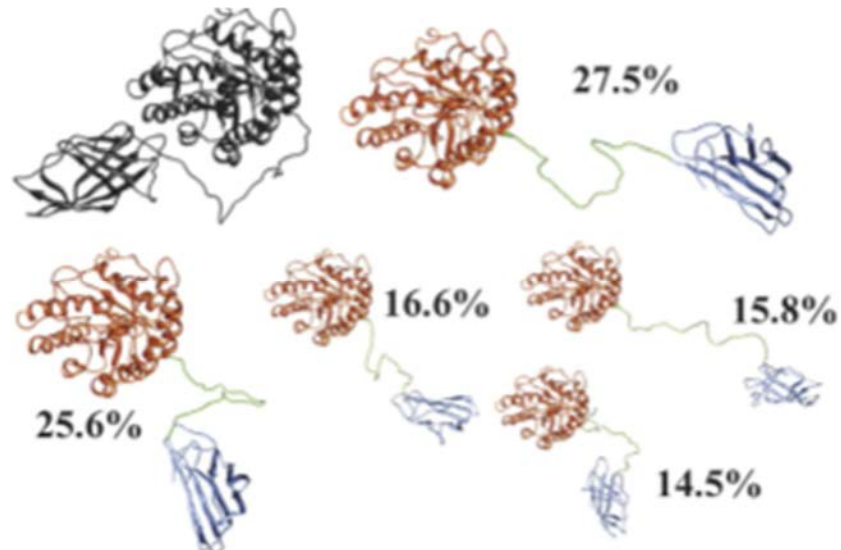


Figura 10. Padrão de mobilidade do conector do módulo catalítico para a EngD e seu CBM (BIANCHETTI, et al. 2012).

Tal mobilidade aparentemente exerce papel chave na forma em que a enzima interage com o substrato, de forma que junto ao módulo CBM, se comporte de forma análoga à um “braço” que permite “aproximar” a enzima ao substrato (BIANCHETTI et al., 2012).

Além das enzimas do complexo das celulasas, outras enzimas possuem papel importante na degradação desta biomassa, como são os casos de algumas enzimas da classe das Multicopper oxidases (MCO) que atuam sobre a lignina.

## 2.5. Multicopper oxidases

As multicopper oxidases são enzimas que oxidam seus substratos aceitando elétrons em um centro cúprico mononuclear e os transferindo para um centro cúprico trinuclear. O dioxigênio se liga a este centro trinuclear, doando quatro elétrons e reduzindo-se a duas moléculas de água (BENTO et al., 2005).

Existem três centros cúpricos espectroscopicamente diferentes: tipo 1 (ou



azul); tipo 2 (ou normal) e tipo 3 (ou binuclear acoplado) (MESSERSCHMIDT e HUBER, 1990; OUZOUNIS e SANDER, 1991). As multicopper oxidases consistem em 2, 3 ou 6 destes domínios homólogos, os quais também compartilham homologia com as cupredoxinas azurin e plastocianina. Estes domínios possuem um dobramento similar à cupredoxina, um motivo sanduiche  $\beta$ , consistindo em sete cadeias em duas folhas  $\beta$ , arranjadas em um barril  $\beta$  tipo “chave grega” (ROBERTS et al., 2002). As principais classes das “multicopper oxidases” incluem:

Ceruplasmim (E.C. 1.16.3.1; ferroxidase), enzima de seis domínios encontrada no soro dos mamíferos e aves que oxidam diferentes substratos orgânicos e inorgânicos, e possui homologia interna de sequência que parece ter surgido de uma triplicação do domínio de ligação do cobre similar ao das lacases e da ascorbato oxidase.

Lacase (E.C. 1.10.3.2; urishiol oxidase) é uma enzima de três domínios encontrada em fungos e plantas, os quais oxidam diferentes fenóis e diaminas. A enzima CueO é uma lacase encontrada em *E. coli*, envolvida com a resistência a cobre (ROBERTS et al., 2002).

Ascorbato oxidase E.C. 1.10.3.3, uma enzimas de três domínios presente em plantas superiores.

Nitrito redutase E.C. 1.7.2.1, uma enzima de dois domínios, sendo um com o centro cúprico tipo 1 e outro tipo 2 (NAKAMURA et al., 2003; SUZUKI et al., 2000).

A diversidade de aplicação das enzimas MCO tem gerado inúmeros trabalhos voltados tanto ao estudo de desenvolvimento de drogas para controle de micro-organismos resistentes a defensivos agrícolas à base de cobre, até enzimas envolvidas com a degradação de material lignocelulósico.

## **2.6. Motivação para a busca de novos antibióticos**

Desde a descoberta da estreptomicina na década de 50, produzida por um actinomicete isolado do solo, microbiologistas e farmacêuticos do mundo inteiro passaram a investir grandes esforços no cultivo de diversos micro-organismos para pesquisar novas linhagens capazes de sintetizar antibióticos. A busca por novos

compostos levou as companhias farmacêuticas a isolarem dezenas de milhares de linhagens produtoras, e ainda hoje micro-organismos cultiváveis de solo representam a principal fonte de antibióticos e outros compostos ativos. Entretanto, os métodos tradicionais para pesquisa de novas drogas que envolvem o cultivo de micro-organismos isolados do solo não são mais tão promissores, devido principalmente à alta taxa de redescoberta de antibióticos já conhecidos, que chega a 99,9% (ZAEHNER e FIELDLER, 1995). Durante um grande período de tempo houve uma refreada na descoberta de novas moléculas, chegando ao limite da metodologia clássica baseada somente em isolamento, cultivo e bioensaios. Esse “Gap” de inovação se estendeu desde o início da década de 1970, estagnando na descoberta das quinolonas e estreptograminas e foi superado no ano 2000, com a descoberta de uma nova classe: oxazolidinona. Entre 2000 e 2009, foi ainda descoberta mais uma classe, a dos lipopeptídeos, além de novas estruturas para a classe das mutilinas (FISCHBACH e WALSH, 2009).

Consequentemente, nos últimos 30 anos, apenas duas classes de antibióticos inéditas foram lançadas no mercado: a oxazolidinona linezolidina e o lipopeptídeo cíclico daptomicina. Como agravante, a telavancina (um glicopeptídeo), a retapamulina (uma pleuromutilina), e a tigeciclina (última geração de tetraciclina) foram os únicos princípios ativos de antibióticos admitidos como moléculas totalmente novas pela “US Food and Drug Administration” (FDA) lançados no mercado nos últimos cinco anos (HAMAD, 2010). Em adição a estas, são poucos os antibióticos em circulação ainda considerados seguros quanto à retenção da capacidade ativa contra micropatógenos multidrogas resistentes (HAMAD, 2010).

A descoberta de uma nova molécula publicada este ano, a teixobactina, trouxe um novo alento para as pesquisas: este antibiótico foi eficiente contra uma série de bactérias gram-positivas consideradas como patógenos recalcitrantes ou multidroga resistentes, tais como *S. aureus* (MSSA e MRSA); *Enterococcus faecalis*, *E. faecium*. Mesmo quando submetidos a concentrações abaixo da mínima concentração inibitória (MIC) sucessivas passagens de *S. aureus* não geraram mutantes resistentes. Esta molécula foi descoberta na bactéria *Eleftheria terrae* através de estratégia de cultivo *in situ* (LING et al., 2015).

O cenário de crise do setor transparece na morosidade do caminho percorrido

entre a prospecção e a chegada ao mercado consumidor. A Figura 11 ilustra o perfil da distribuição de moléculas em fase de pré-implementação em 2010, e demonstra que mais do que o próprio processo de descoberta de novos bioativos, o grande gargalo do processo é representado pelo período pré-clínico (HAMAD, 2010).

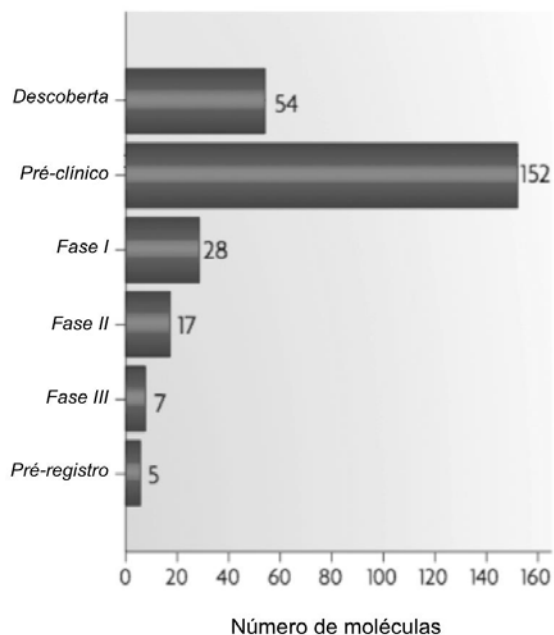


Figura 11. Distribuição por família química e por estágio de pesquisas para novas drogas (HAMAD 2010), representando o “Pipeline” da produção de novos fármacos.

Outro fator que contribui muito para o tempo de retenção de cada droga em cada estágio é a natureza de sua aplicação e a urgência na demanda de certos grupos clínicos (Fig. 12, KAITIN, 2010).

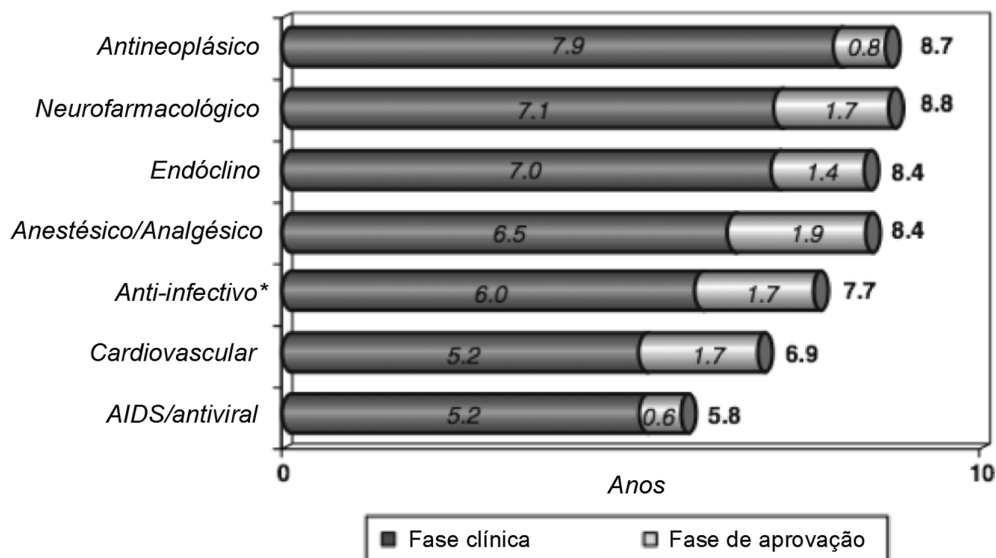


Figura 12. Médias obtidas para o tempo de retenção de novas moléculas de acordo com a área patológica, obtidos no período de 2003 a 2007. \* A categoria Anti-infeccioso exclui agentes antivirais para a AIDS (detentores de categoria própria) (KAITIN, 2010).

Como pode ser observado na Figura 12, os medicamentos para o combate ao HIV são os mais rapidamente implementados, devido em parte à rapidez em que as drogas utilizadas atingem o estágio de ineficácia, frente a novos sorotipos prevalentes da doença. Tanto essa classe de medicamentos quanto os antineoplásicos, apresentam também o menor tempo necessário a regulamentação (Fase de aprovação), refletindo uma clara redução do processo burocrático envolvendo a liberação destes fármacos, garantido um intercâmbio mais dinâmico entre o setor acadêmico/técnico/científico e a comunidade.

## 2.7. Policetídeos

Os compostos policetídeos assumem lugar de destaque entre as drogas de origem natural conhecidas. Mais de um terço dos produtos naturais e seus derivados aprovados como drogas entre 2005 e 2007 foram moléculas deste tipo químico (BUTLER et al. 2008), sendo que os rendimentos anuais de produtos farmacêuticos

frequentemente ficam acima de US\$ 20 bilhões. Dentre os medicamentos de maior destaque, podem-se citar a estatina, avaliada como o cardiovascular mais efetivo; e o enedieno, considerado o agente anti-carcinogênico mais poderoso já descoberto, devido a sua alta reatividade (WEISSMAN, 2009).

A riqueza das estruturas de policetídeos deriva das modificações pós-síntese das cadeias policetílicas, responsáveis na maioria das vezes pela própria reatividade da molécula. A principal forma de modificação é por oxidação, e em seguida vêm as reações de transferência de grupos químicos, tais como as glicosilações e acilações. A diversidade de moléculas também abrange o tamanho diferenciado das cadeias de policetídeos (Fig. 13), indo das menores moléculas com seis a oito carbonos, como o ácido triacético lactona (TAL) e o ácido 6-metil salicílico (6-MSA), até as gigantescas moléculas da maitoxina, que possui um peso molecular de 3.422 Daltons, e uma cadeia que incorpora 164 átomos de carbono (WEISSMAN, 2009).

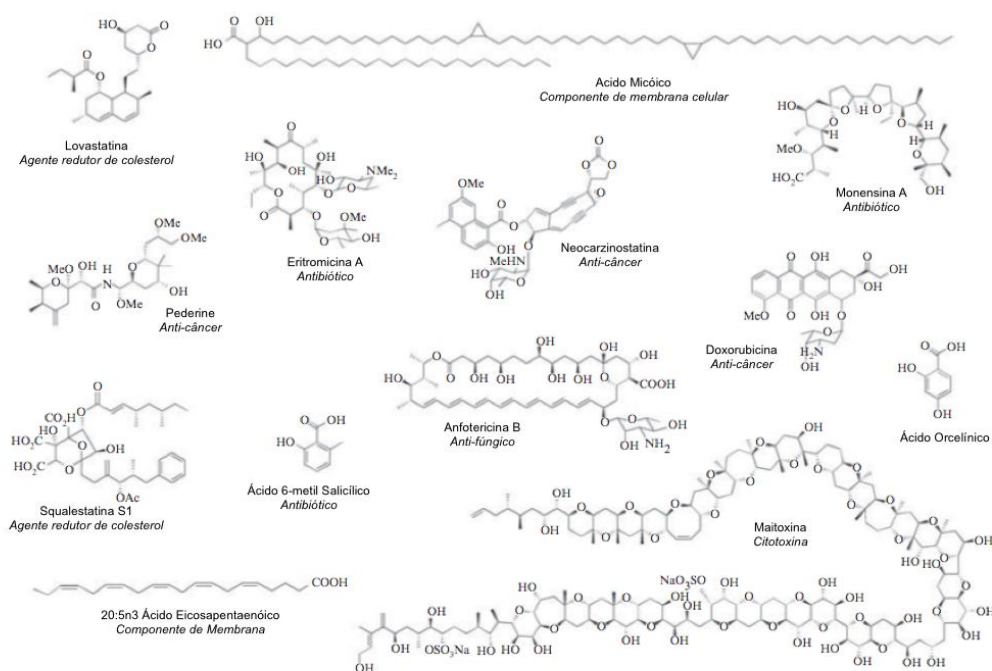


Figura 13. Estrutura de policetídeos representativos, incluindo bioativos proeminentes. (Fonte: WEISSMAN, 2009).

Os policetídeos são originalmente produtos do metabolismo secundário de algumas espécies de bactérias, fungos, plantas, e mesmo microalgas como dinoflagelados, e compõem a estrutura química de muitos fármacos: antibacterianos, antifúngicos, anticarcinomas, imunossupressores, agentes redutores de colesterol, analgésicos, neurotoxinas, antiparasitas e determinantes de virulência (RIX et al., 2002). Eles são sintetizados por complexos enzimáticos variados, denominados Policetideo Sintases (PKSs).

## 2.8. Tipos de PKS e Enzimologia dos Grupos

O complexo de enzimas que estão relacionadas à produção de policetídeos é dividido em três grandes famílias homologas às “Fatty Acid Synthases” (FAS): as “Polyketide Synthases” (PKSs) do tipo I, II e III. Tal divisão não é rígida ou definitiva, podendo haver subfamílias e “clusters” híbridos (SHEN, 2003; WEISSMAN, 2009; JENKE-KODAMA & DITTMANN, 2009). Dentre os actinomicetos encontram-se linhagens que abrigam simultaneamente em seu genoma genes que codificam para a PKS I e a PKS II (BUSTI *et al.*, 2006).

Como características distintivas, tem-se que as PKSs do tipo I (PKS Modular/Interativas) são enzimas grandes, compostas por multi-domínios, que comportam uma série de sítios para as etapas de síntese de policetídeos. Exemplos de produtos naturais derivados dessas enzimas multi-catalíticas são as classes dos macrolídeos (eritromicina) ou polienos (nistatina). As PKSs do tipo II (PKS aromática) são compostas por enzimas separadas mono ou bifuncionais, as quais agem de maneira interativa durante a síntese para formar uma estrutura policetídica, que então é convertida enzimaticamente na forma cíclica para formar compostos aromáticos policíclicos, como a tetraciclina ou a doxorubicina (HUTCHINSON & FUJJI, 1995).

Já as PKSs tipo III, também chamadas de Chalconas Sintases, são proteínas relativamente pequenas, que estão principalmente envolvidas na produção de compostos importantes para as plantas, tais como flavonóides, estilbenos (composto fenólico), mas também são encontradas em bactérias e fungos. Ao contrário das outras PKSs, os intermediários policetídeos da PKS III não são ligados

enzimaticamente, mas por tioésteres de coenzima - A livres (BRACHMANN *et al.*, 2007).

Dentre as PKSs, a via mais estudada para biossíntese combinatória são as PKS tipo I, o que possibilitou o desenvolvimento de banco de dados específicos para essa família, e um sistema eficiente de comparação de domínios enzimáticos capaz de predição da maioria dos substratos por meio de softwares baseados em homologia (YADAV, GOKHALE & MOHANTY, 2003). Tal nível de conhecimento ainda não foi atingido para o grupo de PKS II, e em virtude da sua importância ecológica e biotecnológica é imprescindível que metodologias similares de predição sejam desenvolvidas para auxiliar o direcionamento de pesquisas em proteômica e engenharia molecular de forma mais rápida.

A Figura 14 apresenta esquemas de mecanismo de reação para os três tipos de complexos enzimáticos, ilustrando a dinâmica geral da forma de atuação entre os grupos catalíticos.

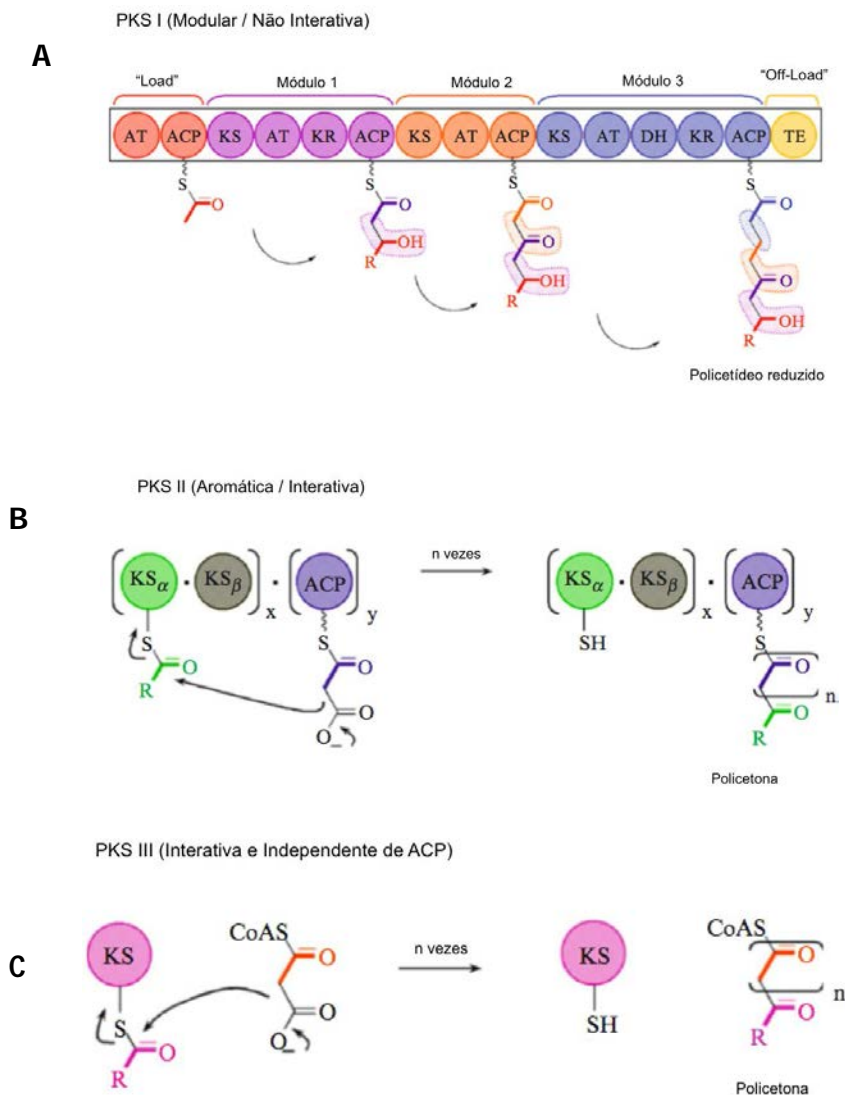


Figura 14. Esquema de reações e dinâmica catalítica das PKSs I, II e III na síntese da cadeia policetílica característica de cada grupo. Legenda: AT, Acil transferase; ACP, Proteína carreadora de acila; KS, ceto-sintase; KR, ceto-redutase; DH, desidratase; TE, Tioesterase; CoA, coenzima A(WEISSMAN, 2009).

Como pode ser observado na Figura 14, a ação das PKSs I Modulares (Fig. 14.A) é do tipo não interativo. A reação se inicia no módulo de ativação, o “loading module”, pela ligação tioéster do precursor inicial ao ACP por ação do primeiro sítio AT. Existe uma transferência do substrato ativado para o ACP do próximo módulo, em uma reação conhecida como tipo “ping-pong” entre sítios ativos, e as enzimas deste módulo irão atuar para extensão de um passo da reação para adição de um novo monômero à cadeia cetônica, e para modificação desse grupo por ações óxido-



reduzidas e de transferases. Essa nova molécula será, por conseguinte, o substrato do próximo módulo. O passo final da reação é a ação da Tio-esterase (TE) que irá liberar a molécula de policetídeo da ACP do módulo de terminação: “Off Module” (WEISSMAN, 2009). Vale lembrar que esse não é o único tipo de reação de grupos de PKS I: Existem as PKS I *cis* AT e *trans* AT, e as PKS I interativas (um ou dois módulos revezam entre si os passos de reação, ou o mesmo gene dá origem a várias cópias do polipeptídeo, que irá atuar de forma interativa nas reações).

Por conseguinte, tanto a PKS II (Fig. 14.B) quanto a PKS III (Fig. 14.C) são interativas. Contudo, a ação das PKS II na síntese da cadeia cetônica inicial envolve a atuação interativa de três subunidades enzimáticas: KS $\alpha$ , KS $\beta$  e ACP. Primeiro ocorre uma ligação tioéster do precursor inicial (maioria Acetil-CoA) ao ACP, reação catalisada pela KS $\alpha$ , mas que pode também ocorrer de forma espontânea *in vitro*, depois a primeira unidade da cadeia é construída por uma reação de condensação tipo “Claisen” (denominação da reação entre dois ésteres, ou um éster e um grupo cetona, na presença de uma base forte ou um sítio catalítico, produzindo um  $\beta$ -cetoéster ou  $\beta$ -dicetona, respectivamente), de um novo substrato (Malonil-CoA e Metilmalonil-CoA), catalisada pelas subunidades KS $\alpha$  e KS $\beta$ , sobre o início da cadeia ainda ancorada na ACP. A reação tipo “Claisen” se repete muitas vezes, com a cadeia nascente sempre ancorada à ACP, e com atuação constante das mesmas enzimas, o que lhe confere o título de interativa. Durante a síntese, unidades enzimáticas adicionais atuam sobre essa cadeia, produzindo grupos cíclicos e aromáticos, ou reduzindo os grupos cetônicos. Após a liberação da molécula da ACP, muitas outras enzimas atuam sobre a cadeia em modificações pós sintéticas variadas, como a ação de oxidases, transferases e hidroxilases (WEISSMAN, 2009).

A PKS III, por sua vez, é surpreendentemente constituída apenas por uma unidade KS, e o precursor inicial é derivado por ligação direta da acetil-CoA livre ao sítio catalítico da KS, que realiza sozinha todos os passos da ligação tipo “Claisen”. O potencial redox do sítio catalítico e a instabilidade de alguns grupos químicos da cadeia em formação irão influir em ciclizações espontâneas na cadeia (WEISSMAN, 2009; BRACHMANN *et al.*, 2007).

## 2.9. As Policetídeo Sintases II

Um grupo de três enzimas, comumente chamado de Policetídeo Sintase Mínima ou PKS mínima, é compartilhado por todas as vias da Família PKS do tipo II: a Cetoacil-Sintase (KS $\alpha$ ), um Fator Elongador de Cadeia (KS $\beta$ ) e a ACP, uma Proteína Carreadora de Acila (WAWRIK et al., 2005). As moléculas policetílicas produzidas por PKSs do tipo II são normalmente glicosiladas, e as alterações no padrão destas glicosilações geram uma diversidade adicional a este grupo (HOPWOOD, 1997). Outros genes conferem a diversidade na estrutura da cadeia policetílica e podem estar presentes em alguns organismos e ausentes em outros (HUTCHINSON & FUJJI, 1995).

Os policetídeos aromáticos constituem uma importante família de bioativos de emprego farmacológico, muitos dos quais com ação antibiótica e anticâncer, e incluem as tetraciclínas, antraciclínas, ácidos aureólicos, tetracenomicinas, anguciclínas, benzo-iso-cromoquinonas, e os policetídeos mais conhecidos são os produzidos pelo metabolismo dos *Actinomycetes* (ZHANG e TANG, 2009).

A despeito da prospecção por novos *clusters* de PKS II objetivar frequentemente novos antibióticos, essa é apenas uma descrição parcial ou imprecisa de sua função, uma vez que podem estar envolvidos na síntese de pigmentos de esporo, na diferenciação fisiológica em condições ambientais adversas (JENKE-KODAMA e DITTMANN, 2009) e mesmo na degradação de compostos lignínico-celulósicos (HSIAO e KIRBY, 2008).

Por outro lado, mesmo trabalhos empregando expressão de genes relacionados à produção de pigmentos demonstram o quanto a biossíntese combinatória destes “clusters” pode contribuir com a farmacologia. Como exemplo, SHEN (1999) descreveu um interessante trabalho em que organismos recombinantes foram produzidos a partir da introdução de uma *pks mínima* derivada de PKS II de esporo de pigmento, resultando em uma variedade de novos esqueletos carbônicos muito superiores aos encontrados em ensaios que empregaram a *pks mínima* de produtores naturais de antibióticos. Um composto policetílico componente de pigmento pode ainda apresentar atividade farmacológica “*in natura*”. Alguns podem ser antitumorais ou antibióticos fracos, como é o caso de

alguns pigmentos antraquinônicos (BRACHMANN et al. 2007). Os pigmentos do tipo nafitoquinônicos (policetídeos de origem fúngica) são conhecidos por sua atividade antibacteriana, antifúngica e inseticida (GRAZIANI et al. 2004).

Segundo HSIAO e KIRBY (2008), o gênero *Streptomyces* ocupa um importante papel no solo, atuando na reciclagem dos materiais de natureza lignocelulósica, e algumas enzimas envolvidas no processo parecem derivar de vias de policetídeos de pigmento. A determinação de quais e quantas enzimas, precisamente, estariam envolvidas na rota metabólica de degradação de celulose e lignina (incluindo genes não relacionados a PKS) permanece obscura, mas é possível evidenciar a participação de: catalases/peroxidases; endoglucanases; celulases; proteínas ligantes de celulase; mono-oxigenases; dioxigenases; hemeoxigenases; tirosinases (monofenol/mono-oxigenases) e seus cofatores; celobiose-hidrolases; oxigenase; alcanos-mono-oxigenases; e reguladores sensores de peróxido de hidrogênio, além de enzimas de metabolismo de celobiose. Algumas ainda podem também se relacionar novamente às vias de PKS de forma indireta, através do fornecimento de unidades precursoras para modificação da cadeia de policetídeos (HSIAO & KIRBY, 2008).

## **2.10. Biossíntese combinatória como estratégia para a produção de novos fármacos**

A formação em “*cluster*” dos genes para as PKS II favorece a prospecção por PCR de novas rotas biosintéticas em bibliotecas metagenômicas. Isso se justifica pelo fato de que localizar a sequência conservada de uma única região gênica em um dado clone metagenômico pode implicar em encontrar não apenas a enzima para qual o oligonucleotídeo iniciador foi desenhado, mas todas as outras que estão localizadas nas adjacências suportadas pelo tamanho do inserto do clone. Dessa forma, existem grandes chances de recuperação de uma via biosintética na íntegra, incluindo promotores, reguladores, genes que conferem resistência ao hospedeiro heterólogo para o antibiótico em questão, enzimas de modificação pós-síntese da estrutura policetídica, dentre outros (REF de agrupados fisicamente no cromossomo).

Além da abordagem clássica de prospecção de fármacos que se fundamenta na descoberta de moléculas naturais, outra estratégia na busca por novos fármacos seria o uso da engenharia molecular. Essa engloba o conjunto de manipulações gênicas inerentes às práticas de expressão heteróloga e de biossíntese combinatória, embora não se restrinja à apenas essas abordagens.

No caso da expressão gênica heteróloga, um exemplo de emprego da estratégia seria a introdução de uma maquinaria gênica responsável por determinada molécula em um hospedeiro mais eficiente do que o original para a produção do composto em escala industrial (MARTIN & DEMAIN, 1980). Outra variação do emprego dessa técnica ganha destaque a partir da expansão do uso de abordagens metagenômicas, nas quais muitas vezes o material gênico promissor encontrado por homologia pertence a um organismo desconhecido, sem possibilidades de seu isolamento e cultivo laboratorial pelas técnicas disponíveis. Nesse caso, a expressão heteróloga se configura como a via mais acessível para tentativa de expressão e obtenção dos compostos enzimáticos ou químicos de interesse.

Um fator importante para as atividades de prospecção por novas moléculas, utilizando expressão heteróloga, é o fato de atualmente o pesquisador dispor do acesso às previsões parciais da estrutura tridimensional de enzimas, e mesmo de estruturas químicas de compostos que possam ser produzidos por maquinarias bioquímicas similares (BALTZ, 2006). Essa previsão pode ser obtida por representações *in silico* criadas com o auxílio de ferramentas de bioinformática e da engenharia molecular, a partir da sequência nucleotídica estudada, e das informações reunidas pelo pesquisador em relação à família enzimática de interesse.

A partir do “esboço” molecular do candidato a fármaco são fornecidas as pistas sobre a relevância deste metabólito, através de uma comparação dos prováveis grupos químicos com estruturas similares conhecidas. Indo um pouco além, algumas vezes é possível uma formulação de hipóteses sobre a interação de tal composto com células vivas (bioatividade) por meio da comparação de suas características com bancos de moléculas com funções já elucidadas. Uma vez que se tem uma ideia de quão promissor um determinado “cluster” seja, a etapa seguinte

seria a introdução deste em hospedeiros diferentes para testes de expressão. No caso específico de genes para o gênero *Streptomyces*, existe a disponibilidade para a pesquisa de hospedeiros heterólogos aptos a receber o DNA exógeno, com vantagens de serem selecionados e modificados para não silenciar a expressão de genes heterólogos. Como exemplos, têm-se linhagens de *Streptomyces lividans*, *Streptomyces albus* e *Streptomyces coelicolor*, além de linhagens mutantes de *Escherichia coli* desenvolvidas especialmente para expressão de genes do grupo de *Streptomyces*, através da modificação do “códon usage” para atender a expressão de genes exógenos, aumentando a quantidade de RNA transportador para códons GC, frequentes em *Streptomyces* (MARTIN & DEMAINE, 1980; WANG et al., 2000; BALTZ, 2006).

Por sua vez, a biossíntese combinatória voltada para a busca de novos fármacos também é um campo de abordagem para pesquisadores que trabalham em Metagenômica.

A produção de um dado composto policetídeo específico é resultado da interação enzimática peculiar a um grupo de organismos que compartilhem de uma mesma rota bioquímica para tal metabólito secundário. Para certos tipos de antibióticos, existem organismos produtores do mesmo composto entre linhagens pertencentes a espécies distintas. Existem duas situações que exemplificam o compartilhamento do mesmo metabólito por um dado conjunto de organismos. Um dos exemplos é o que ocorre dentre espécies filogeneticamente próximas entre si, que porventura possuam uma mesma via biosintética herdada de um grupo ancestral comum. A outra possibilidade se sucede dentre indivíduos de relativo distanciamento filogenético, mas que, no entanto, tenham incorporado “clusters” enzimáticos do outro organismo por transferência gênica horizontal. Esta troca de genes entre organismos não detentores de relações íntimas de parentesco se constitui em um fenômeno de relatada ocorrência dentro do gênero *Streptomyces* (METSÄ-KETELÄ et al., 1999).

No caso de antibióticos da PKSs, o emprego da biossíntese combinatória se beneficia do fato de que, embora cada rota bioquímica normalmente resulte em um tipo específico de moléculas, irão existir dentro de uma mesma família de PKS

enzimas comuns a todas as rotas, derivadas de genes homólogos e detentores de regiões conservadas, ainda que não sejam idênticas entre si.

Em virtude desta semelhança enzimática existirão etapas bioquímicas similares nos mecanismos básicos envolvidos na formação da cadeia policetídic. Esta conservação de certas propriedades dentre as enzimas homólogas permite certa compatibilidade entre grupos enzimáticos de “clusters” diferentes. No caso de antibióticos como os da Família PKS II, em que a produção da molécula final é resultado da interação de várias enzimas pertencentes a um “cluster”, a biossíntese combinatória permite criar moléculas, através da recombinação de módulos enzimáticos de organismos produtores já conhecidos, bem como combinações de módulos conhecidos com genes novos obtidos pela prospecção em metagenômica.

Assim, a engenharia molecular ou biossíntese combinatória engloba um conjunto de manipulações da genética molecular e outras manipulações biosintéticas, que basicamente podem ser entendidos como estratégias para acelerar os processos evolutivos em muitas ordens de magnitude visando “competir” com a evolução natural dos antibióticos. Como exemplo deve-se considerar a imensa quantidade de complexos peptídeos, policetídeos e outros arcabouços químicos conhecidos, amplamente diversificados pela presença e disposição de estruturas tais como: glicosilações, metilações, halogenizações, grupos acilas e hidroxilações, e que se desenvolveram ao longo de bilhões de anos em actinomicetos e fungos (BALTZ, 2006).

### **3.11. O uso da metagenômica para busca de novas vias de antibióticos**

Uma característica interessante do arranjo em cluster para os genes relacionados à produção de antibióticos é a melhor coordenação do processo de regulação gênica. Do ponto de vista da engenharia genética de novas moléculas, este arranjo permite que a recombinação de vias seja possível, com substituição de subunidades em pontos estratégicos para o desenho racional de bioativos.

Em alguns casos, a arquitetura dos agrupamentos gênicos permite deduzir, até certo ponto, qual o tipo de molécula resultante da ação daquela via biosintética. Tal predição pode muitas vezes identificar potenciais substratos que serão

incorporados na cadeia policetídica, potenciais glicosilações, acilações, reduções, hidroxilações, aminações, aromatização de estruturas cíclicas de policetídeos, etc.

Este agrupamento físico no cromossomo dos genes relacionados à produção de policetídeos é uma grande vantagem no ponto de vista da possibilidade de acesso de vias inteiras de antibióticos por estratégias de metagenômica. Aliás, a metagenômica tem assumido um papel cada vez mais relevante como estratégia de acesso ao potencial microbiano de bioativos, aliada à engenharia racional de bioativos (FERNANDEZ-MORENO et al., 1991; WANG et al., 2000; BALTZ, 2006; IQBAL et al., 2012). A título de ilustração, alguns exemplos de novas moléculas de antibióticos obtidas através da metagenômica são apresentados na figura 15.

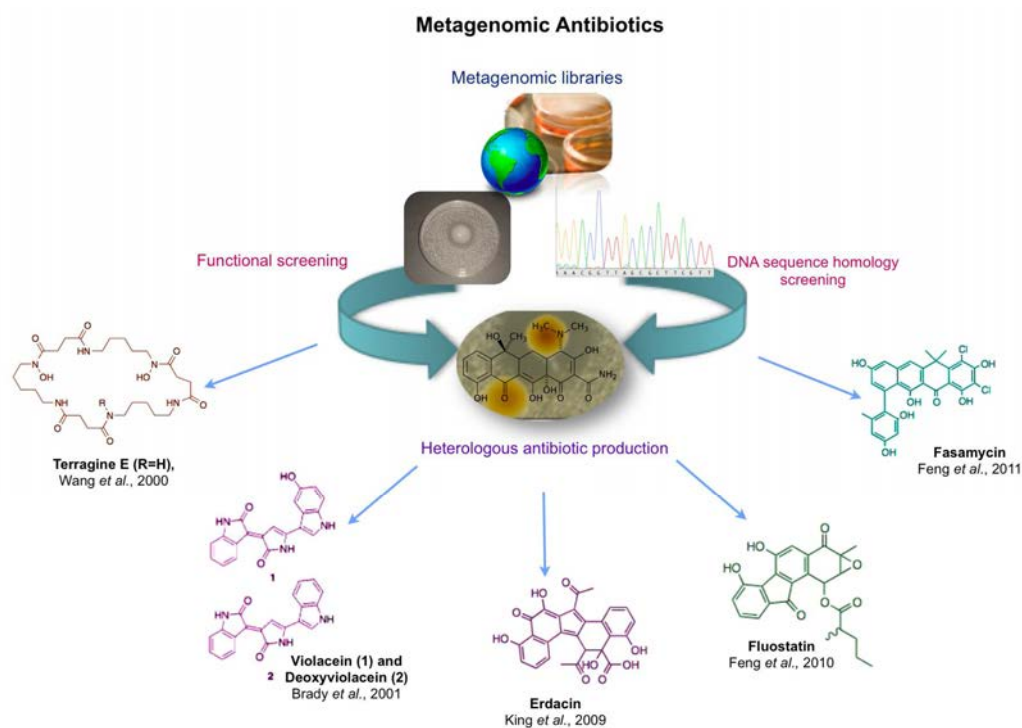


Figura 15. Alguns exemplos de moléculas de novos fármacos obtidos através de estratégias metagenômicas (Fonte: GOMES et al., 2013/disponibilizada em março de 2014)

### 3. MATERIAL E MÉTODOS

Este trabalho foi realizado no Laboratório de Bioquímica de Micro-organismos e Plantas (LBMP), da Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal (FCAV) da UNESP. Os protocolos experimentais, quando não especificados pelo fornecedor dos *Kits* ou reagentes, foram delineados conforme metodologia descrita por SAMBROOK & RUSSELL (2001); KIESER e colaboradores (2000). Tais técnicas não oferecem riscos para o meio ambiente e a saúde humana, encaixando-se nos padrões de biossegurança (NB-1).

#### 3.1. Origem da amostra metagenômica e análise *in silico*

O clone metagenômico B5pl37 estudado no presente trabalho pertence à Biblioteca Metagenômica de Solo de Eucalipto (S.E). A biblioteca S.E. (Schuch, 2007) foi construída empregando amostras de solo coletado sob serapilheira de bosque de eucaliptos da “Faculdade de Ciências Agrárias e Veterinárias” FCAV-UNESP, Campus de Jaboticabal, sob as coordenadas 21°14'S, 48°17' W.

Este clone originalmente foi selecionado com base em uma prospecção por PCR de genes relacionados à produção de antibióticos, pertencentes à via de Policetídeos Sintases do tipo II (PKS II). Após o sequenciamento do clone cosmidial (método de Sanger, estratégia de sub-biblioteca “shotgun”), as fases abertas de leitura (“open reading frames”: ORFs) foram identificadas através do programa ORFfinder (NCBI) e submetidas à triagem manual com base nos maiores resultados de similaridade com sequências conhecidas. Além dos genes relacionados à PKSII, a anotação gênica resultou em sequências similares a genes de enzimas relacionadas à degradação de biomassa lignocelulósica: uma  $\beta$ -glicosidase, uma celulase e uma lacase (GOMES, 2011; GOMES et al., 2013). Os genes relacionados ao metabolismo de biomassa foram escolhidos como alvos do presente estudo.

As sequências peptídicas deduzidas para os genes de interesse foram submetidas ao programa “Swiss-Model” [12] para a geração de modelos de estrutura tridimensional proteica, através da ferramenta “Swiss-Model Automatic Modelling



Mode”. O programa “Protparam” [13] foi utilizado para predição de massa molecular, e coeficiente de extinção molar teórico para as enzimas deduzidas.

As sequências de aminoácidos correspondentes a cada ORF foram submetidas às análises de agrupamentos fenéticos com base em métodos de distância entre agrupamentos de sequências. O alinhamento múltiplo global foi realizado utilizando-se o método de otimização por vizinhos próximos (Neighbor joining), com análise de confiança por reamostragem de mil réplicas (bootstrap 1000). Para tanto, foram utilizados os programas EbioEX (versão 1.5.1(25), LAGERCRANTZ, 2008) e ClustalX (LARKIN et al., 2007).

As sequências padrões comparadas às ORFs metagenômicas para  $\beta$ -glicosidase e celulase foram obtidas junto à base de dados “Carbohydrate-Active enZYmes Database” – CAZy ([www.cazy.org](http://www.cazy.org)). As sequências padrões para as lacases foram obtidas na base de dados “The laccase and multi copper oxidase engineering database” – LCCED (<http://www.lcced.uni-stuttgart.de/cgi-bin/LccED1.2/index.pl>).

Além dos programas e bases de dados descritos, outros foram utilizados com o objetivo de obter mais informações sobre os genes estudados e sobre as vias metabólicas nas quais os mesmos se inserem, conforme pode ser visualizado na tabela 03.

**Tabela 03:** Bancos de dados e ferramentas úteis para análise *in silico* de genes

Ferramenta	Recurso utilizado
“Brenda”	Sistema de informação sobre enzimas (substratos, inibidores, cinética, classificação, artigos, etc.)
“Kegg”	“Kyoto Encyclopedia of Genes and Genomes”: sequencias gênicas e vias metabólicas.
“LCCED”	“The laccase and multi copper oxidase engineering database”: famílias de lacases
“BioCyc”	Informações sobre vias metabólicas permitindo comparação entre indivíduos, auxiliando na definição do melhor hospedeiro homólogo/heterólogo de expressão.
“CAZY”	Informações sobre como classificar famílias de carboidrases; acesso à sequências e vias metabólicas.
“Swiss model”	Identificação de similaridade entre porções de uma sequência de aminoácidos de interesse e os bancos de dados de proteínas disponíveis.
“IMG”	Banco de dados de anotação automática de sequências, mapeamento e comparação entre vias similares.
Uniprot	Base integrada de dados obtidos para proteínas identificadas experimentalmente, com interface para análise de similaridade entre sequências através do alinhamento local (BlastUniprot) e parâmetros físico-químicos dos resíduos de aminoácidos.
Ebioex	Pacote de Ferramentas para análise de sequências através de alinhamento múltiplo global, par à par ou local.

<i>ClustalW</i>	<i>Pacote de Ferramentas para análise de sequências através de alinhamento múltiplo global, par à par ou local.</i>
<i>Bioedit</i>	<i>Pacote de Ferramentas para análise de sequências através de alinhamento múltiplo global, par à par ou local.</i>
<i>MacPymol</i>	<i>Ferramenta para visualização e edição de estrutura tridimensional de proteínas.</i>
<i>Prisma</i>	<i>Ferramenta de análise e construção de gráficos para interpretação de parâmetros cinéticos enzimáticos.</i>
<i>Sigmaplot</i>	<i>Ferramenta com recursos avançados para a análise gráfica de dados.</i>
<i>Prodom</i>	<i>Base de dados para identificação de domínios enzimáticos.</i>
<i>Pfam</i>	<i>Base de dados para identificação de famílias de domínios enzimáticos.</i>
<i>Figtree</i>	<i>Visualização e edição de cladogramas e fenogramas.</i>
<i>PDB (Protein Data Base)</i>	<i>Base de dados para estruturas protéicas resolvidas por métodos experimentais.</i>
<i>NCBI/Blast</i>	<i>Banco Mundial de dados sobre sequências (Americano).</i>
<i>Protparam</i>	<i>Dedução de parâmetros físico-químicos de proteínas através da estrutura primária protéica (sequências de resíduos de aminoácidos).</i>
<i>Phyre2</i>	<i>Identificação de moldes e desenho de modelos de estrutura tridimensional de proteínas in silico, com base em homologia e <i>ab initio</i>.</i>

### 3.2. Linhagens e vetores utilizados

As linhagens bacterianas e os vetores utilizados neste trabalho encontram-se relacionados nas tabelas 4 e 5. O material biológico foi manipulado obedecendo a boas práticas de contenção e manipulação laboratorial, sendo que todos os procedimentos foram realizados em câmara de fluxo laminar e em condições assépticas, precedido e seguido de desinfecção (etanol 70% e U.V.). Os materiais descartados em cada procedimento experimental foram devidamente submetidos à esterilização por vapor quente durante 20 minutos a 121°C em autoclave. O armazenamento das linhagens comerciais e modificadas foi realizado em estoques de glicerol 20% submetidos à ultra freezer -80°C.

Tab 4. Linhagens bacterianas

<b>Linhagem</b>	<b>Fonte ou Referência</b>
<i>Escherichia coli</i> BL21 (DE3)	Novagem
<i>Escherichia coli</i> C41(DE3)	Miroux et al., 1996
<i>Escherichia coli</i> Artic Express (DE3)	Agilent
<i>Escherichia coli</i> DH5 F <sup>-</sup> [ <i>supE44</i> , $\Delta$ <i>lacU169</i> ( $\phi$ 80 <i>lacZ</i> $\Delta$ M15), <i>hsdR17</i> , <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>relA1</i> ].	SAMBROOK & RUSSELL, 2001.
<i>E.coli</i> ET12567 pUZ8002	John Innes Centre
<i>S. coelicolor</i> M1152	John Innes Centre

Tab. 5: Vetores

Vetor	Fonte ou Referência
Vetor de clonagem pUC19DNA/ <i>Sma</i> I	Fermentas
Vetor de expressão PET28a	Novagen
Vetor de integração pSET152	John Innes Centre

### 3.3. Meios de cultura:

a. **Luria Bertani (LB):** Triptona (10 g/L); Extrato de Levedura (5 g/L); NaCl (10 g/L); pH 7,0; agar (9 g/L) para meio sólido. Suplementado com ampicilina (LB+Amp 70-100 µg/mL) ou kanamicina (LB+Kan 50 µg/mL), de acordo com o vetor em uso (SAMBROOK & RUSSELL, 2001);

b. **Meio SOC:** Triptona (20 g/L); Extrato de Levedura (5g/L); NaCl (0,5 g/L); 10 mL/L de KCl 250 mM; pH 7,0; 10 mL/L de MgCl<sub>2</sub> 1M (adicionado após esterilização do meio); 20 mM de glicose.

c. **Meio CMC:** Peptona (5 g/L); Extrato de levedura (3g/L); Carboxi-metil-celulose: CMC (0,2% m/v); ágar (13 g/L); pH 7,0. Meio sólida: ágar (13 g/L).

d. **Meio MS:** Manitol (20g/L); farinha de soja: 20g/L; agar 20g/L, pH 7,2.

### 3.4. Técnicas gerais

#### 3.4.1. Eletroforese de DNA:

Amostras de DNA foram analisadas por eletroforese em gel de agarose (0,8-1%) dissolvida em tampão TBE (Tris 89 mM, Ácido Bórico 89 mM e EDTA 2,5 mM; pH 8,0). A eletroforese foi realizada no mesmo tampão, em voltagem pré-fixada entre 70-90 volts (de acordo com o tamanho do gel), empregando tampão de carregamento da amostra composto por glicerol (50%); Azul de bromofenol (0,02%) em tampão HCl 0,1 M pH 7,0 (q.s.p). O DNA foi rotineiramente visualizado através do uso de brometo de etídio (0,5 µg/mL) como agente intercalante de DNA, seguido

de exposição à luz ultravioleta em aparelho foto documentador Molecular Imager® Gel Doc™ XR+ (Bio-Rad, número de catálogo 170-8195).

As eletroforeses preparativas seguiram o mesmo procedimento, entretanto não foi utilizado brometo de etídeo para a confecção do gel de agarose. Após a corrida, apenas uma parte do gel contendo uma alíquota de DNA e o padrão de tamanho molecular foi corada com brometo de etídeo. A banda de interesse foi visualizada em luz ultravioleta, e foi feita uma marcação da região com auxílio de um palito de madeira. Posteriormente o restante do gel foi colocado sobre uma película plástica ao lado do gel marcado, e a região correspondente foi cortada do gel não corado. A região de interesse foi eluída do gel com auxílio do kit “Wizard SV Gel and PCR Clean-Up System” (Promega) seguindo especificações do fabricante.

### **3.4.2. Eletroforese para amostras proteicas**

As amostras de proteínas foram analisadas pela técnica de eletroforese em gel de poliacrilamida conforme descrito a seguir.

#### **a. Eletroforese desnaturante em gel de poliacrilamida**

O padrão proteico das amostras extraídas e purificadas foi analisado através de eletroforese desnaturante em gel de poliacrilamida (Laemmli, 1970). As amostras foram previamente incubadas a 100°C por cinco minutos em tampão de amostra (glicerol, 20%; SDS, 4%; de  $\beta$ -mercaptoetanol, 5%; Azul de bromofenol [0,02%]; Tris-HCl, 62mM pH 6,8), a seguir, aplicada em gel de poliacrilamida 10% contendo SDS. A separação de proteínas foi realizada através da aplicação de campo elétrico (125 V e 25 mA). As amostras foram coradas pelo método de Azul de Comassie (0,2% Azul de Comassie, 40% Metanol, 10% Ácido Acético).

#### **b. Eletroforese não desnaturante em gel de poliacrilamida (PAGE) e detecção de atividade catalítica**

A proteína  $\beta$ -glicosidase foi submetida a esta eletroforese, sendo que alíquotas

da amostra proteica purificada foram submetidas à eletroforese (120V constante; 1,5 horas) em condição não desnaturante (ausência de SDS ou  $\beta$ -mercaptoetanol), conduzida em banho de gelo. Após a eletroforese, parte do gel contendo uma alíquota da enzima foi submergido em tampão fosfato de sódio (20mM, pH 7) acrescido de 5mM de pNP- $\beta$ G e incubado por 20 min a 37°C para revelar a ação catalítica (Zimograma), enquanto outra parte contendo albumina sérica bovina (BSA) e uma alíquota da amostra foi corada pelo método de Azul de Comassie. O perfil eletroforético de migração obtido foi então comparado ao perfil em condição desnaturante.

### **3.5. Clonagem dos genes em vetores de expressão**

#### **3.5.1. Síntese de oligonucleotídeos iniciadores**

Os oligonucleotídeos iniciadores foram desenhados pelo método manual sobre a sequência nucleotídica tendo por base a anotação gênica obtida através do programa ORFfinder (NCBI). Estes iniciadores foram sintetizados na região imediatamente externa ao gene, incluindo sítios de reconhecimento para enzimas de restrição EcoRI (oligonucleotídeos iniciadores sense) e HindIII (oligonucleotídeos iniciadores antisense) para possibilitar a clonagem direcionada em vetor de expressão. Tais sítios enzimáticos foram escolhidos por não apresentarem sítios de restrição dentro do inserto metagenômico, conforme análise no programa “SMS Restriction map” ([http://www.bioinformatics.org/sms2/rest\\_map.html](http://www.bioinformatics.org/sms2/rest_map.html)), e por fazerem parte do sítio de múltipla clonagem presente no vetor utilizado, o pET28a (Novagen, Fig. 16).

Os pares de oligonucleotídeos obtidos foram checados através do programa “PCR Primer Stats” ([http://www.bioinformatics.org/sms2/pcr\\_primer\\_stats.html](http://www.bioinformatics.org/sms2/pcr_primer_stats.html)), para avaliar as características de afinidade de pareamento entre as moléculas e possibilidades de formação de estruturas secundárias estáveis indesejadas. A ferramenta também permite a determinação do conteúdo G/C e predição da temperatura de pareamento. Os oligonucleotídeos iniciadores estão descritos na tabela 6.

Tabela 6: Oligonucleotídeos para as carboidrases

Gene/pb	Oligonucleotídeo	sequências 5' -> 3' *	Tm-5	%G/C
β-glucosidase: ORF8; 1500 pb (F.L. positiva)	GLB1gene1 EcoRIF	TGG <b>GAATTC</b> AGTCCCCGCATGACAGC	64,3	57,7
	GLB1gene1 HindIII <sup>R</sup> *	CACA <u>AAGCTT</u> CGTTCCTTCAGGACGTC	62,7	53,8
Celulase/quitinase: ORF9; 1500 pb (F.L. negativa)	GLB1gene2 EcoRIR*	CGAGGCG <b>GAATTC</b> AGGCAGCACATGGCAAC	66,7	58,6
	GLB1gene2 HindIIIF	CCGGA <u>AAGCTT</u> TGCGGGCGGTTCAGCCGG	71,8	74,1
Lacase: ORF11; 1750 pb (F.L. positiva)	GLB1gene3 EcoRIF	AGCC <b>GAATTC</b> GGCGCAGCGATGGCCAG	68,7	66,7
	GLB1gene3 HindIII <sup>R</sup> *	GTGCGTA <u>AAGCTT</u> GCGCGGTGGTCAGCC	68,7	66,7

\* Negrito: sítio de restrição da EcoRI; sublinhado: sítio de restrição da HindIII. F.L.: Fase de Leitura.

Para confirmação do reconhecimento de sítio de pareamento sobre os genes avaliados, foi feito o alinhamento da sequência alvo (ou sua sequência reversa/complementar) e os oligonucleotídeos iniciadores através do programa EbioX. O alto conteúdo GC foi tolerado em virtude de escassez de regiões AT no inserto metagenômico (homólogo à *Streptomyces sp.*), mas a ausência de “hairpins” e dímeros foi buscada sempre que possível. Contudo, a natureza palindrômica dos sítios enzimáticos inseridos nos iniciadores contribuem para a formação de auto pareamentos, de forma que foram escolhidos iniciadores que envolvessem o menor número de bases nucleotídicas em tais estruturas. Os demais oligonucleotídeos iniciadores empregues neste trabalho foram os pares universais T7/SP6; M13F/M13R e o oligo do pET28aF (Novagen).

### 3.5.2. Reação de amplificação por PCR

As condições de amplificação por PCR foram padronizadas levando-se em conta testes de tipo de tampão de reação (padrão ou acrescido de KCl), concentrações de cloreto de magnésio, DNA molde. Foram realizados testes com diferentes temperaturas, tanto empregando reação de PCR convencional, cujos ciclos contem a mesma temperatura de pareamento, quanto PCR “touchdown”,

cujos ciclos iniciais são com aumento gradual de temperatura de pareamento. Os testes foram realizados de forma sequencial, visando melhorar progressivamente a amplificação. Quando as adaptações realizadas não surtiam efeito, novos oligonucleotídeos eram sintetizados. Reações otimizadas foram utilizadas para amplificação dos amplicons com uma DNA Polimerase de alta especificidade, a “PFU DNA Polymerase” (Thermo scientific).

### **3.5.3. Preparo do vetor pET28a EcoRI/HindIII**

Células BL21DE3 contendo vetor de clonagem pET28a (Fig.16) sem inserto foram cultivados em placas de Petri com meio LB acrescido de Kan 50 µg/mL (LB+Kan50). As culturas foram incubadas em B.O.D. 37 °C por 20 h. Após o cultivo, uma colônia isolada foi utilizada para o preparo de um pré-inóculo (18h em agitador de bancada, a 200 rpm e 37 °C). Este material foi utilizado para inóculo a 1% em 100 ml de meio LB+ Kan50, mantido nas mesmas condições descritas por 20h. Após este período, o material foi centrifugado (20 min, 3220 xg, 22 °C), as células foram lavadas com salina 0,85% e divididas em frações de 2ml de cultivo para a posterior extração empregando o kit da Promega “Wizard Plus SV Minipreps DNA Purification System”.

As amostras foram quantificadas em aparelho Nanodrop e visualizadas através de eletroforese em gel de agarose 0,8%. As preparações de DNA plasmidial foram então submetidas a uma reação de restrição seguida de desfosforilação conforme descrito na tabela abaixo (Tab. 7), e posteriormente purificadas com auxílio de uma eletroforese preparativa.

Tabela 07: Preparo do vetor de clonagem pET28a

Reação	Reagentes	Volume	TR/T*
A. Restrição "Fast Digest" para o vetor	Tampão Fermentas (10X)	6,5 µl	37°C/30 min, seguido de desnaturação à 80°C /10 min.
	Fermentas Fast Digest EcoRI	8,0 µl	
	Fermentas Fast Digest HindIII	8,0 µl	
	DNA plasmidial 165 ng/µl	39,0 µl	
	H <sub>2</sub> O milli-Q autoclavada e deionizada	3,5 µl	
B. Desfosforilação	"Shrimp Alkaline phosphatase" (SAP) – Sigma 1 U/ µl	5,0 µl	37 °C/75 min, seguido de desnaturação à 65 °C /15 min.
	Tampão SAP	10,0 µl	
	H <sub>2</sub> O milli-Q autoclavada e deionizada	55,0 µl	
	Reação A (após ensaio de restrição)	30,0 µl	

\*TR/T: Temperatura de Reação/Tempo de incubação.

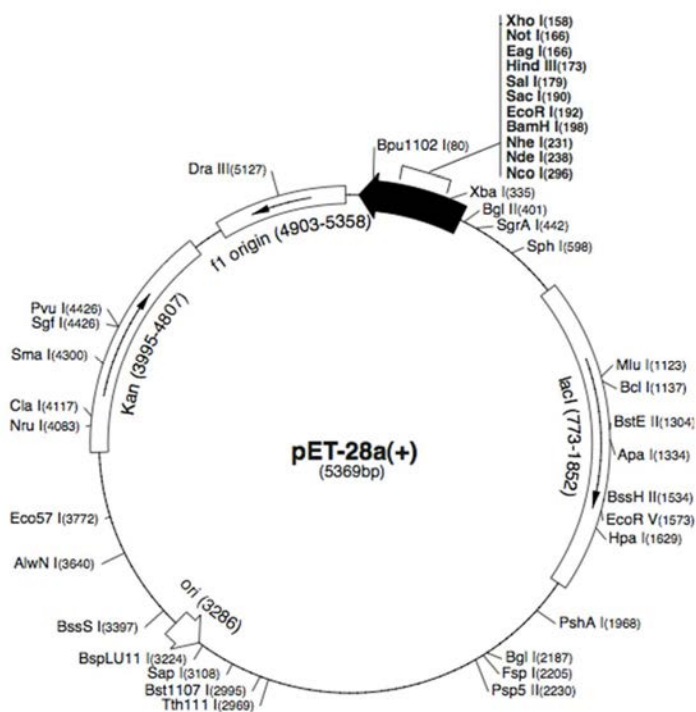


Figura 16. Esquema ilustrativo apresentando características do vetor pET28a (Novagen).

### 3.5.4. Preparo dos insertos e reação de ligação

O produto de amplificação por PCR foi submetido a uma reação de restrição empregando as mesmas enzimas do vetor, *EcoRI* e *HindIII*. Em seguida os fragmentos foram purificados em colunas (Wizard® SV Gel and PCR Clean-Up



System; Promega) e submetidos à reação de ligação com o vetor pET28a. A tabela 8 trás as condições que foram utilizadas.

Tabela 8: Reação de Restrição enzimática e ligação

Reação	Reagentes	Volume	TR/T*
<b>Restrição “Fast Digest”</b>	Tampão 10X concentrado Fermentas	03,0 µl	37 °C/30 min, seguido de desnaturação à 80 °C/10 min.
	Fermentas Fast Digest <i>EcoRI</i>	01,5 µl	
	Fermentas Fast Digest <i>HindIII</i>	01,5 µl	
	Amplicon de 1500 pb (180 ng/µl)	6,0 µl	
	H <sub>2</sub> O milli-Q autoclavada e deionizada	18,0 µl	
<b>Reação de ligação pET28a 3:1</b>	Enzima T4 DNA Ligase – Biolabs 600 U/ µl	01,5 µl	16 °C/16 horas.
	Tampão da T4 DNA ligase (Biolabs)	02,0 µl	
	H <sub>2</sub> O milli-Q autoclavada e deionizada	5,2 µl	
	vetor pET28aEcoRI/HindIII desfosforilado purificado <sup>a</sup> (55 ng/µl)	9,3 µl	
	Amplicon restrito EcoRI/HindIII e purificado <sup>a</sup> (25 ng/µl)	02,0 µl	

\*TR/T: Temperatura de Reação/Tempo de incubação; a: Purificação em eletroforese preparativa.

### 3.5.5. Preparo das células competentes

Para o preparo das células utilizadas como hospedeiras de expressão, foram utilizados estoques de células BL21 DE3; C41 ou Artics. Todas as etapas de manipulação foram realizadas em condições assépticas em câmara de fluxo laminar.

Inicialmente, obteve-se colônias isoladas através de estriagem de células do estoque em meio LB puro (sem antibiótico) e cultivado em B.O.D. 37°C por 20 h. Em seguida, foi feito o pré-inóculo através de repasse de uma colônia bem desenvolvida em 10 ml de meio LB puro seguido de cultivo por 18h em agitador de bancada à 180 rpm a 37 °C.

Um mililitro deste pré-inóculo foi utilizado para inocular 50 mL de meio LB puro, e submetido ao cultivo em agitador orbital à 180 rpm e 37°C, sendo monitorado periodicamente a cada 10-20 min até atingir a D.O.<sub>600</sub> de 0,3-0,4. Alíquotas de 10 mL do material foram então repassado à tubos falcons de 50 ml e mantidos no gelo por 10 minutos. Em seguida, este material foi centrifugado por 20 min, 3220x.g, 4°C. O sobrenadante foi descartado e o precipitado foi delicadamente ressuspenso em

50% do volume de cultivo de  $\text{CaCl}_2$  50 mM, e incubado no gelo por 20 min, seguido de centrifugação por 15 min nas mesmas condições já descritas. O sobrenadante foi descartado, e o precipitado ressuspendido gentilmente em 1/15 do volume de cultivo em  $\text{CaCl}_2$  50 mM, e 15% deste volume de glicerol em banho de gelo. A suspensão celular foi então alíquotada em tubos de centrifuga (150  $\mu\text{l}$ ), congeladas em nitrogênio líquido e mantidas em freezer  $-80^\circ\text{C}$  até o uso. Amostras deste material foram estriadas em placas de meio LB com antibiótico (meios LB Amp/cloram/kan, como controles negativos) e em LB puro (controle positivo) para validação.

### **3.5.6. Transformação por choque térmico**

Para a clonagem do vetor+inserto metagenômico nos hospedeiros heterólogos procedeu-se da seguinte forma em condições assépticas em câmaras de fluxo laminar: 20  $\mu\text{l}$  de vetor pET28a ligado à carboidrase foram acrescentados em uma das alíquotas de células competentes. A reação foi suavemente agitada por pipetagem e mantida em gelo por 5 min. Em seguida a reação foi incubada por 90 segundos à  $42^\circ\text{C}$  e novamente resfriada em gelo por 5 min. A esta preparação foram acrescentados 370  $\mu\text{l}$  de meio SOC, e a suspensão celular foi cultivada em agitador orbital à  $37^\circ\text{C}$ , 240 rpm por uma hora, sendo que 60  $\mu\text{l}$  de células desenvolvidas foram plaqueados por placa de Petri com meio LB+Kan50 e mantidas em B.O.D. a  $37^\circ\text{C}$  por 22 horas. Cinco colônias bem desenvolvidas foram coletadas, cultivadas em agitador orbital à  $37^\circ\text{C}$ , 240 rpm por 22 horas. Em seguida, alíquotas foram estocadas em glicerol 20% à  $-80^\circ\text{C}$ , e o restante foi submetido a extração plasmidial pelo protocolo da lise alcalina (Sambrook e Russel, 2003), para confirmação da presença do inserto por PCR.

### **3.6. Ensaio de Expressão**

Para os ensaios de expressão, partiu-se de colônias isoladas dos transformantes para preparo de pré-inóculo (cultivo por 16 horas) à 2% para um volume de 1000 ml de cultivo final, ambos mantidos em agitador orbital à  $37^\circ\text{C}$ , 240 rpm. Após adição do pré-inóculo, a suspensão celular foi cultivada até atingir D.O.<sub>600</sub>

0,5-0,6. Após este período, as células foram cultivadas sob indução com IPTG de acordo com o respectivo tratamento descrito na tabela 9. Para o primeiro ensaio de tempo de expressão (E0), as amostras não foram previamente extraídas, e sim rompidas no procedimento inicial de preparo das amostras antes da eletroforese. Para as demais condições foram feitas extrações com tampão Tris, Fosfato ou Hepes, conforme descrito a seguir.

**Tabela 09:** Testes de Otimização da expressão

Tratamento	IPTG	Período de indução (horas)	Temperatura de indução °C	Tampão empregue na extração
E0	0,1 mM	0,2,3,4,5 e 6	37	lise térmica
1	0,1 mM	2	37	Tris/Fosfato/Hepes
2	0,1 mM	2	30	Tris
3	0,1 mM	4	29	Fosfato
4	0,1 mM	4	28	Fosfato

### 3.6.1. Protocolo de extração empregando tampão Tris

Após o cultivo, as células foram centrifugadas à 10000xg, 10 °C por 30 minutos; o sobrenadante foi descartado e o precipitado (“pellet”) foi ressuspensionado em 10 mL de Tampão A1 (10 mM de Tris-Cl [7,4]; 100 mM de NaCl; 1 mM de EDTA [pH 7,35]; 0,5 mM de PMSF; 1 mg/L de lisozima) com auxílio de uma pipetadora manual de 5mL com ponteira descartável estéril. Todo o procedimento de extração foi realizado em banho de gelo. As amostras foram inicialmente adicionadas de 1mg/L de lisozima, seguido de incubação por 30 min. Alíquotas de 5mL foram submetidas á homogeneização e lise por sonicação, empregando-se 4 pulsos de 20 segundos intercalados de 20 segundos de repouso entre os pulsos (300W). Após a sonicação, as amostras foram centrifugadas por 20 min 10000xg, 4°C, o sobrenadante foi decantado em um tubo estéril e o “pellet” foi ré extraído com 5ml do tampão A2 (Tampão A1 acrescido de 2% de Triton X-100) seguindo o mesmo passo de extração anterior, o qual foi repetido até o atingir o quarto extrato.

### **3.6.2. Protocolo de extração empregando tampão Fosfato**

De forma similar ao descrito para extração acima, o cultivo celular foi centrifugado à 10000xg, 10 °C, por 30 minutos. Em seguida, o sobrenadante foi descartado e o precipitado (“pellet”) foi ressuscitado em 10 mL de Tampão B1 (Tampão fosfato/Tampão de lise: NaH<sub>2</sub>PO<sub>4</sub> (50mM); NaCl (300 mM); imidazol (10 mM); pH 8) com auxílio de uma pipetadora manual de 5mL com ponteira descartável estéril. Todos os procedimentos de extração foram realizados em banho de gelo. As amostras foram inicialmente adicionadas de 1mg/L de lisozima, seguido de incubação por 30 min.

Entretanto, para este procedimento alíquotas de 5 mL foram submetidas à homogeneização e lise por sonicação, empregando-se 6 pulsos de 10 segundos intercalados de 10 segundos de repouso entre os pulsos (300W). Após a sonicação, as amostras foram centrifugadas por 20 min 10000xg, 4°C,, o sobrenadante foi decantado em um tubo estéril e o “pellet” foi re-extraído com 5ml do tampão B2 (Tampão B1 acrescido de 0,25% de Triton X-100), seguindo o mesmo passo de extração anterior, o qual foi repetido até o atingir o quarto extrato. Alíquotas dos extratos e do “pellet” foram retiradas para serem analisadas por corrida eletroforética em gel SDS-Page.

### **3.6.3. Protocolo para extração empregando tampão Hepes**

As células transformantes contendo a proteína ligada a uma cauda de hexa-histidina (6xHis-Tag) foram recuperadas através de centrifugação (10.000xg, 4°C por 30 minutos). O sobrenadante foi descartado e o precipitado foi ressuscitado em 10 mL de tampão HEPES (Sigma) 50mM pH 7, contendo 20mM de imidazol (Sigma) e lisozima (1mg/L de cultivo inicial), e mantido sob incubação a 4°C por 30 min. Em seguida, a amostra foi submetida à lise por congelamento a -20°C por 10 min seguido de descongelamento a 37°C por 5 min. Posteriormente, o extrato celular foi centrifugado por 20 minutos 10.000 xg, 4°C, o sobrenadante foi decantado em um tubo estéril, enquanto o “pellet” foi ressuscitado em 5 mL do mesmo tampão.

### **3.7. Métodos de lise celular utilizados**

O cultivo celular foi centrifugado a 10.000xg, 10°C por 30 minutos. Em seguida, o sobrenadante foi descartado e o precipitado (“pellet”) foi ressuspensionado em 10 mL de tampão de lise com auxílio de uma pipetadora manual de 5mL com ponteira descartável estéril. Todos os procedimentos de extração foram realizados em banho de gelo. As amostras foram inicialmente adicionadas de 1mg/L de lisozima, seguido de incubação por 30 min. Após o tratamento enzimático com lisozima as amostras foram submetidas a um dos métodos de lise descrito abaixo.

#### **3.7.1. Extração empregando ultrassom (lise por cavitação)**

Alíquotas de 5mL foram submetidas à homogeneização e lise por sonicação (Sonicador *Sonifier 250A*), empregando-se 6 pulsos de 10 segundos intercalados de 10 segundos de repouso entre os pulsos (300W). Após a sonicação, as amostras foram centrifugadas por 20 minutos 10.000xg, 4°C, o sobrenadante foi decantado em um tubo estéril (primeiro extrato) e armazenado a 15°C. O “pellet” foi re-extraído com 5mL do mesmo tampão acrescido de detergente (Triton X-100), seguindo o mesmo passo de extração anterior, o qual foi repetido até atingir o quarto extrato. Alíquotas dos extratos e do “pellet” foram retiradas para serem analisadas por corrida eletroforética em gel de poliacrilamida com dodecil sulfato de sódio (SDS-PAGE).

#### **3.7.2. Extração por congelamento e descongelamento**

A lise celular foi realizada incubando-se a amostra em 10 mL de tampão Hepes pH 7 seguido procedimento de lise celular por congelamento (10 minutos à -20°C) e descongelamento (5 minutos à 37°C). Em seguida, o extrato celular foi centrifugado por 20 minutos 10.000xg, 4 °C, o sobrenadante foi decantado em um tubo estéril, enquanto o “pellet” foi ressuspensionado em 5 mL do mesmo tampão. As amostras foram mantidas à 15 °C, ou congeladas à -20 °C na presença de 10% de

glicerol, ou imediatamente congeladas com nitrogênio líquido e armazenadas à -80 °C.

### **3.8. Purificação através cromatografia de afinidade por metal imobilizado (Ni<sup>2+</sup>)**

O extrato referente à extração de cada enzima recombinante foi purificado através da técnica de cromatografia de afinidade utilizando a resina de níquel Ni-NTA (Qiagen), que possibilita a purificação por afinidade em proteínas recombinantes contendo uma cauda de hexa-histidina (6xHis-Tag). Assim, para 35 mL de extrato, correspondente a 1 L de indução, foi utilizado 1 mL da resina Ni-NTA (2mL de uma solução estoque etanoica). A resina foi centrifugada a 13.000 xg, por 1 minuto, a 4°C, o sobrenadante descartado e a resina ressuspensa com água destilada/deionizada/filtrada tipo Milli-Q. Este procedimento foi repetido 3 vezes para retirar o álcool, sendo que, na última vez, a resina foi ressuspensa no tampão Hepes 50 mM (pH 7,0) para equilibrar a resina, acrescido de 10mM de imidazol para evitar a formação de complexos inespecífica de proteínas aos sítios ligantes. A mistura foi levada para geladeira (15°C) para incubação por 1 hora sob agitação orbital.

Após o período, o material incubado foi colocado em uma coluna "Poly-Prep Chromatography Columns 9cm" (Bio-Rad) e a fração não complexada foi coletada ("flow-through"). À coluna foi adicionado o tampão de eluição, o mesmo usado para equilibrar a resina, com gradiente na concentração final do imidazol de 20 mM até 1M. As frações foram coletadas e estocadas a 4°C para análise em gel de poliacrilamida (SDS-PAGE). Após a purificação a enzima foi submetida à dessalinização através de emprego de filtros concentradores de proteína da marca *Millipore* consistindo em sistemas de falcons de 50mL acoplados com filtros Millipore com capacidade de retenção de moléculas com tamanho molecular superior a 10.000 MW. Sucessivas diluições seguidas de concentração por filtração segundo recomendações do fabricante permitiram retirar o imidazol das amostras.

### **3.9. Determinação da concentração de proteínas solúveis**

A concentração de proteína total foi determinada através do método de

Bradford (1976), empregando uma curva-padrão preparada com albumina sérica bovina em diferentes concentrações. A proteína pura, por sua vez, foi estimada por espectrofotometria, através da leitura da absorbância em comprimento de onda de 280nm em aparelho espectrofotômetro modelo Nanodrop1000 (Thermo Scientific) empregando os valores estimados no software Protparam para o coeficiente de extinção molar e a massa molecular da proteína metagenômica, inserindo estes dados no aparelho antes da leitura da proteína purificada

### **3.10. Ensaio de Atividade enzimática para a $\beta$ -glicosidase**

#### **3.10.1. Ensaio de atividade enzimática empregando substratos cromogênicos**

Para a determinação da atividade enzimática sobre substratos cromogênicos através da medida da absorbância após a reação para estimar a liberação do íon *p*-nitro-fenolato (pNP;  $\epsilon_{405}=0,0187\mu\text{M}^{-1}\text{cm}^{-1}$ , pH 10), procedeu-se conforme adaptações do protocolo descrito por Daroit (2008). A reação inicial (100  $\mu\text{L}$ ) consistiu de 82 $\mu\text{L}$  de tampão fosfato de sódio 20mM, pH 7; 1 mM do substrato (concentração final); e 18  $\mu\text{L}$  da enzima diluída a 0,02 mg/mL (total de 0,00036 mg). A adição da enzima ao meio reacional foi temporizada (adições por coluna com intervalos de 1 min) de forma a permitir que todas as amostras apresentassem o mesmo tempo de reação. As reações foram incubadas em aparelho termociclador a 37°C ou nas respectivas temperaturas especificadas na descrição dos experimentos.

Após a incubação sob as respectivas condições avaliadas, alíquotas de 30 $\mu\text{L}$  foram adicionadas em 150 $\mu\text{L}$  de tampão carbonato de sódio (500 mM; pH 10), em placas para cultura de tecidos (96 poços; fundo chato: JetBiofil) e homogeneizadas por pipetagem utilizando um pipetador multicanal semiautomático (Eppendorf). Logo em seguida, foi realizada a leitura no comprimento de onda 405 nm do *p*-nitrofenolato liberado, em espectrofotômetro SpectraMax<sup>®</sup> M2<sup>e</sup> (Molecular Devices). Todos os ensaios foram feitos em triplicata, e na presença dos seguintes controles: reações sem enzima e reação sem o respectivo composto avaliado. A atividade foi calculada empregando a derivação da fórmula de cálculo da concentração de partículas por absorbância de Lambert-Beer, e a equação da atividade de Michaelis-Menten (reação em equilíbrio dinâmico), levando-se em consideração a quantidade

de enzima equivalente a 30µL de reação (0,000108 mg, equivalente a 5,4 µL de enzima a 0,02 mg/mL), em um volume final de 180 µL, com um passo óptico vertical de 0,5454545 cm, conforme a fórmula de cálculo a seguir:

$$\text{Atividade} = \left( \frac{\text{Abs}_{405} \div \text{tempo [min]} \times \text{volume aferido [mL]}}{\text{volume da enzima [mL]} \times \text{passo óptico [cm]}} \right) \div (\epsilon \text{ pNP } [\mu\text{M}^{-1} \text{ cm}^{-1}]) \times$$

### 3.10.2. Determinação de unidade enzimática (U)

O substrato padrão utilizado para a determinação dos parâmetros cinéticos foi o de pNP-β-D-glicopiranosídeo (pNP-βG). Uma unidade (U) de β-glicosidase foi considerada como a quantidade de enzima necessária para a hidrólise de 1 µmol de pNP-βG por minuto, sob as respectivas condições ótimas.

### 3.10.3. Temperatura e pH ótimos

Para a determinação da temperatura e pH ótimos da enzima parcialmente purificada, a atividade β-glicosídica foram realizados experimentos em triplicata nas condições de temperatura variando de 4 a 50°C e o pH de 3,5 a 10,5. A atividade relativa foi expressa em percentual de atividade correspondente à condição ótima de reação para cada experimento (controle, 100%), e foi calculada conforme descrito no item anterior.

### 3.10.4. Determinação de Parâmetros cinéticos

A atividade da β-glicosidase purificada foi avaliada em triplicata utilizando diferentes concentrações de pNP-β-D-glicopiranosídeo (0,1-10mM, totalizando 23 concentrações) em condições ótimas de temperatura e pH, de 0-60 min com intervalos de 10 min. O produto da reação foi estimado através do cálculo de produto liberado empregando a fórmula a seguir:



$$\text{Produto} = (\text{Abs}_{405} \div (\epsilon \text{ pNP } [\mu\text{M}^{-1}\text{cm}^{-1}]) \times \text{passo óptico})$$

A velocidade inicial ( $V_0$ ) de formação do produto *p*-nitro-fenolato foi calculada através da fórmula de regressão linear obtida para o gráfico de formação do produto em relação ao tempo, onde  $V_0$  é igual ao ângulo de inclinação da reta obtida para o intervalo linear correspondente a cada concentração de substrato. Os parâmetros cinéticos  $K_m$  (constante de Michaelis-Menten) e  $V_{\max}$  (velocidade máxima de reação) para o substrato *p*-Nitrofenil- $\beta$ -D-glicopiranosídeo foram calculados a partir do uso do programa "Graph Pad Prism".

### 3.10.5. Efeito de sais e outros reagentes sobre a atividade enzimática

O efeito de sais (e respectivos íons) sobre a atividade  $\beta$ -glicosídica foi determinado através da pré-incubação de 18  $\mu\text{L}$  da enzima convenientemente diluída adicionada de 82  $\mu\text{L}$  de tampão fosfato de sódio (20 mM; pH 7,0) em microplacas de PCR em termociclador a 37 °C por 10 minutos, contendo os respectivos sais de forma a atingir as concentrações finais de 0-5mM. Após a pré-incubação, 20 $\mu\text{L}$  do substrato pNP- $\beta$ G (10mM) foi adicionado a cada coluna com a pré-incubação finalizada, e homogeneizada através de pipetagens sucessivas empregando a pipetadora multicanal. As amostras permaneceram no termociclador por mais 30 minutos a 37°C. Após a incubação, alíquotas de 30 $\mu\text{L}$  foram adicionadas em 150 $\mu\text{L}$  de tampão carbonato de sódio (500 mM; pH 10) em microplacas de fundo chato de polipropileno e foram homogeneizadas por pipetagem. Logo em seguida, foi realizada a leitura em espectrofotômetro (405 nm) do *p*-nitrofenol liberado. Em todos estes testes, a atividade enzimática foi expressa como percentual da atividade apresentada sem as respectivas adições (controle).

De forma similar, foi avaliada a influência de outros reagentes, como o  $\beta$ -Mercaptoetanol (1-5 mM); o Dodecil Sulfato de Sódio (SDS, 0,1-2%); etanol (0,5mM-1 M) e metanol (0,5-500 mM).

### **3.10.6. Determinação de unidade enzimática (U) para $\beta$ -glicosidase**

O substrato padrão utilizado para a determinação dos parâmetros cinéticos foi o de 4-Nitrofenil- $\beta$ -D-glicopiranosídeo (também conhecido como p-Nitrofenil- $\beta$ -D-glicopiranosídeo, ou pNP- $\beta$ -D-glicopiranosídeo, ou ainda pNP $\beta$ G). Uma unidade (U) de  $\beta$ -glicosidase foi considerada como a quantidade de enzima necessária para a hidrólise de 1  $\mu$ mol de pNP $\beta$ G por minuto, sob as respectivas condições ótimas.

### **3.10.7. Zimograma para $\beta$ -glicosidase (E.C. - 3.2.1.21)**

Para estimar o tamanho molecular e o perfil da migração da banda com atividade catalítica da proteína  $\beta$ -glicosidase, alíquotas da amostra proteica purificada foram submetidas a uma eletroforese em condição de proteína nativa (não dissociativa). A eletroforese em condição não desnaturante foi conduzida em banho de gelo e na ausência de SDS e  $\beta$ -Mercapta-etanol (120V constante; 1,5 horas). Para evidenciar a possível existência de polimorfismos de bandas, após a eletroforese, parte do gel contendo uma alíquota da enzima foi submergida em tampão fosfato de sódio (20mM, pH 7) acrescido de 5mM de 4-Nitrofenil- $\beta$ -D-glicopiranosídeo e incubado por 20 min a 37°C para revelar a ação catalítica (Zimograma), enquanto outra parte contendo BSA como padrão de tamanho molecular e uma alíquota da amostra foi corada pelo método de Azul de Comassie. O perfil eletroforético de migração foi então comparado com um perfil em condição desnaturante, para se estimar o tamanho molecular da enzima na condição nativa.

## **3.11. Preparo de materiais para os Ensaio de Expressão para o “cluster” de PKS II**

Conforme já explanado na introdução, a expressão dos genes para antibióticos presentes no inserto metagenômico será realizada posteriormente em hospedeiro homólogo *Streptomyces coelicolor* M1152, especialmente modificada por engenharia molecular para expressão de vias de antibióticos, uma vez que as vias endógenas de produção de policetídeos foram retiradas para otimizar a produção de

compostos exógenos e facilitar sua detecção (Gomez-Escribano & Bibb, 2011). Neste trabalho, foram realizados experimentos preparatórios visando a inserção destes genes no hospedeiro homólogo.

Inicialmente, foi realizada a troca de vetor cosmidial para um vetor conjugativo/integrativo (pSET152/*E.coli* DH10B), através de digestão do clone cosmidial B5pl37 com enzima *EcoRI* (esta enzima não corta o inserto e o libera do vetor cosmidial inicial, vide fig. 17), purificação do inserto seguido de ligação com vetor pSET152 (Fig. 18) também digerido com *EcoRI*.

Figure 2. pWEB-TNC Cosmid Vector.

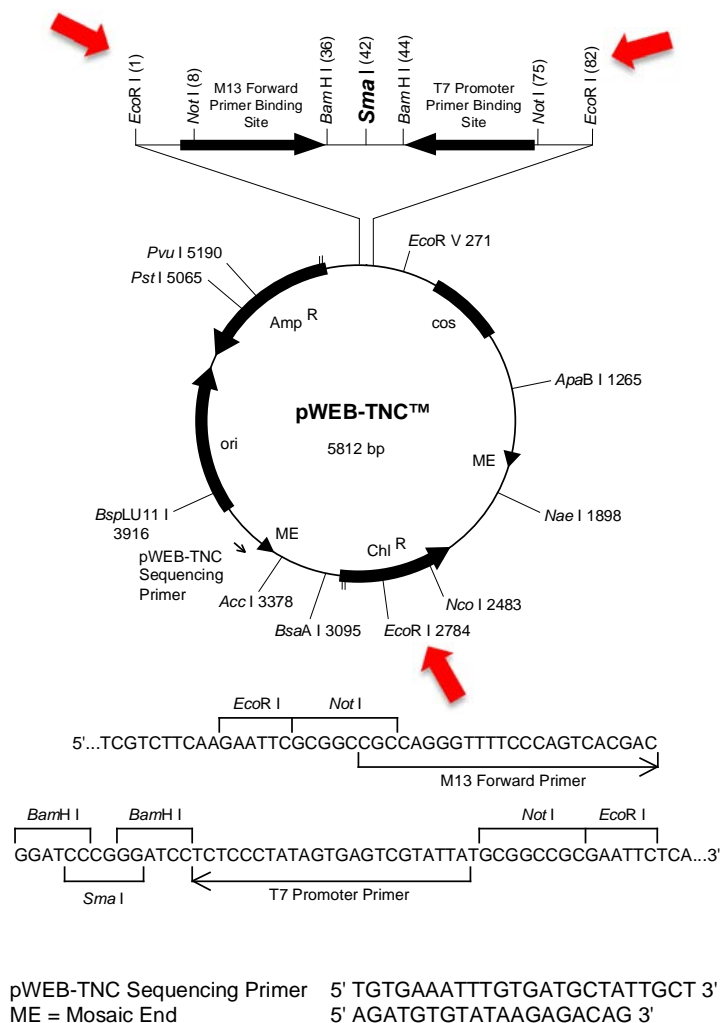


Figura 17. Cosmidio pWEB-TNC, que abriga o inserto metagenômico para o clone B5pl37.

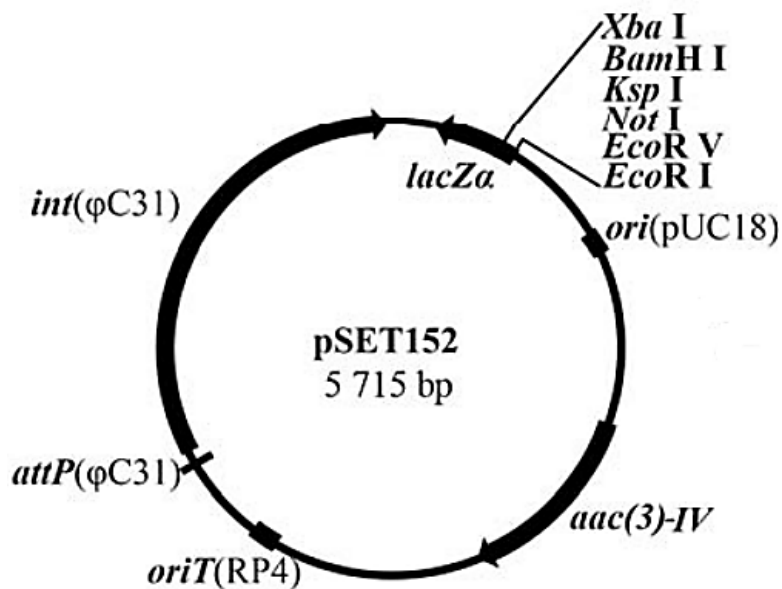


Figura 18. Vetor integrativo conjugativo para transferência intergêneros entre *E. coli* e *Streptomyces sp.*

Posteriormente foi realizada a seleção dos clones positivos empregando o sistema de seleção *lacZ* presente no vetor, sendo selecionadas colônias brancas isoladas e bem desenvolvidas após 18 horas de cultivo em B.O.D. 37°C. O DNA foi então extraído, a inserção do fragmento de interesse foi verificado por restrição enzimática com a enzima *EcoRI*. O inserto clonado no vetor pSET152 foi inserido por eletroporação em células eletrocompetentes de bactéria não metilante de DNA *E.coli* ET12567 pUZ8002.

A suspensão de esporos para o hospedeiro homólogo final *Streptomyces coelicolor* M1152 foi feita em glicerol 20% estocada em freezer -80°C. Inicialmente, as células foram reativadas em meio MS (B.O.D. 30°C.), e posteriormente uma alçada foi inoculada em 5 ml de meio MS líquido, e mantida sob agitação por 24 horas em agitador orbital (200 rpm, 30°C). Alíquotas de 100 µl desta suspensão foram inoculadas em placas de petri contendo meio MS sólido, espalhadas com auxílio de um soabe estéril e mantida em B.O.D 30°C por três semanas até que os esporos estavam bem visíveis. Estes foram então ressuspensos em glicerol e estocados.

## 4. RESULTADOS E DISCUSSÃO

### 4.1. Predição *in silico* das enzimas metagenômicas envolvidas na degradação de material lignocelulósico

A busca por novas enzimas mais eficientes e de baixo custo de obtenção é motivada pela dinâmica dos processos industriais, extremamente dependente de práticas competitivas e autossustentáveis.

À medida que novas enzimas são caracterizadas, estas informações são armazenadas em banco de dados, e o estudo das sequências dos genes expressos para estas enzimas permitem que relações de similaridades possam servir como ferramenta de busca e triagem de enzimas com potencial de ação parecido em grandes projetos de sequenciamento.

Sendo assim, a busca por genes *in silico*, através de estudos de predição funcional e de estrutura tridimensional proteica, é de grande importância para triagens de enzimas de interesse industrial, porque é uma abordagem rápida e pouco dispendiosa, que permite ter uma noção de como uma enzima pode atuar antes de ir para o experimento de bancada. Muitas vezes, a busca por domínios conservados relacionados a uma característica especial, como a resistência a altas temperaturas ou um maior ajuste ao substrato desejado, permite criar um “pipeline” para a busca deste perfil ideal (Lorez e ECK, 2005).

Este procedimento adquire valor especial com o surgimento de inúmeras abordagens de sequenciamento em larga escala que geraram um imenso banco de dados de sequências gênicas (LI et al., 2013; MACLEAN, WOOSI et al. 2010; JONES and STUDHOLME, 2009; MASON et al., 2012; SIMON and DANIEL, 2011; SMITH, 2011).

Utilizando como exemplo os dados para glicosil hidrolases (GHs), a base de dados CAZy (LAMBERD et al., 2013) já reúne 194596 entradas classificadas por similaridade de sequência como módulos catalíticos distribuídos em 132 famílias e 2314 sequências não classificadas. Dentre todas estas GHs, apenas 3,5% (6891)

foram caracterizadas experimentalmente *in vitro*, e apenas 0,5% (938) possuem estrutura resolvida experimentalmente.

A julgar pela pequena parcela de enzimas caracterizadas experimentalmente, é de se esperar que somente a pesquisa clássica não seja capaz de atender a demanda por novas enzimas. Assim, a predição *in silico* é fundamental para selecionar o alvo mais promissor para caracterização cinética, através de estratégias rápidas, passíveis de automatização e validação (MILLER and EISENBERG, 2008; SLEATOR and WALSH, 2010).

Sendo assim, optou-se por iniciar o estudo dos genes relacionados ao metabolismo de degradação de compostos lignocelulósicos ( $\beta$ -glicosidase (ORF8); uma celulase/quitinase (ORF9); e uma lacase (ORF11)) primariamente através desta abordagem, tendo por base a análise de similaridade com sequência disponíveis em banco de dados tais como o Cazy e o NCBI/Uniprot. A construção de modelos tridimensionais para as proteínas também foi adotada para reconhecimento de *motifs* conservados e geometria do sítio catalítico.

#### 4.1.1. $\beta$ -Glicosidase Metagenômica

Dentre as fases abertas de leitura (“open reading frame”: ORFs) obtidas no clone B5pl37, encontramos uma sequência de aminoácidos deduzida com até 77% de similaridade (cobertura de 98%) com  $\beta$ -glicosidases (EC 3.2.1.21) da família GH1 do gênero *Streptomyces*.

A partir de um alinhamento global (empregando a ferramenta ClustalX do software Ebioex) entre a ORF8 metagenômica e sequências de GH1 disponíveis na base de dados Cazy, foi feito um dendograma fenético pelo método de otimização de distância entre as sequências Neighbor Joining, com análise de confiança de distribuição de ramos por re-amostragem bootstrap de 5000 réplicas. A árvore fenética obtida pode ser visualizada na figura 19. A sequência metagenômica se agrupou no clado correspondente as sequências de GH1 presentes em Actinobactéria, incluindo a sequência da  $\beta$ -Glicosidase Bgl3, que é uma GH1 presente em *Streptomyces* para a qual foi elucidado o mecanismo catalítico de atuação.

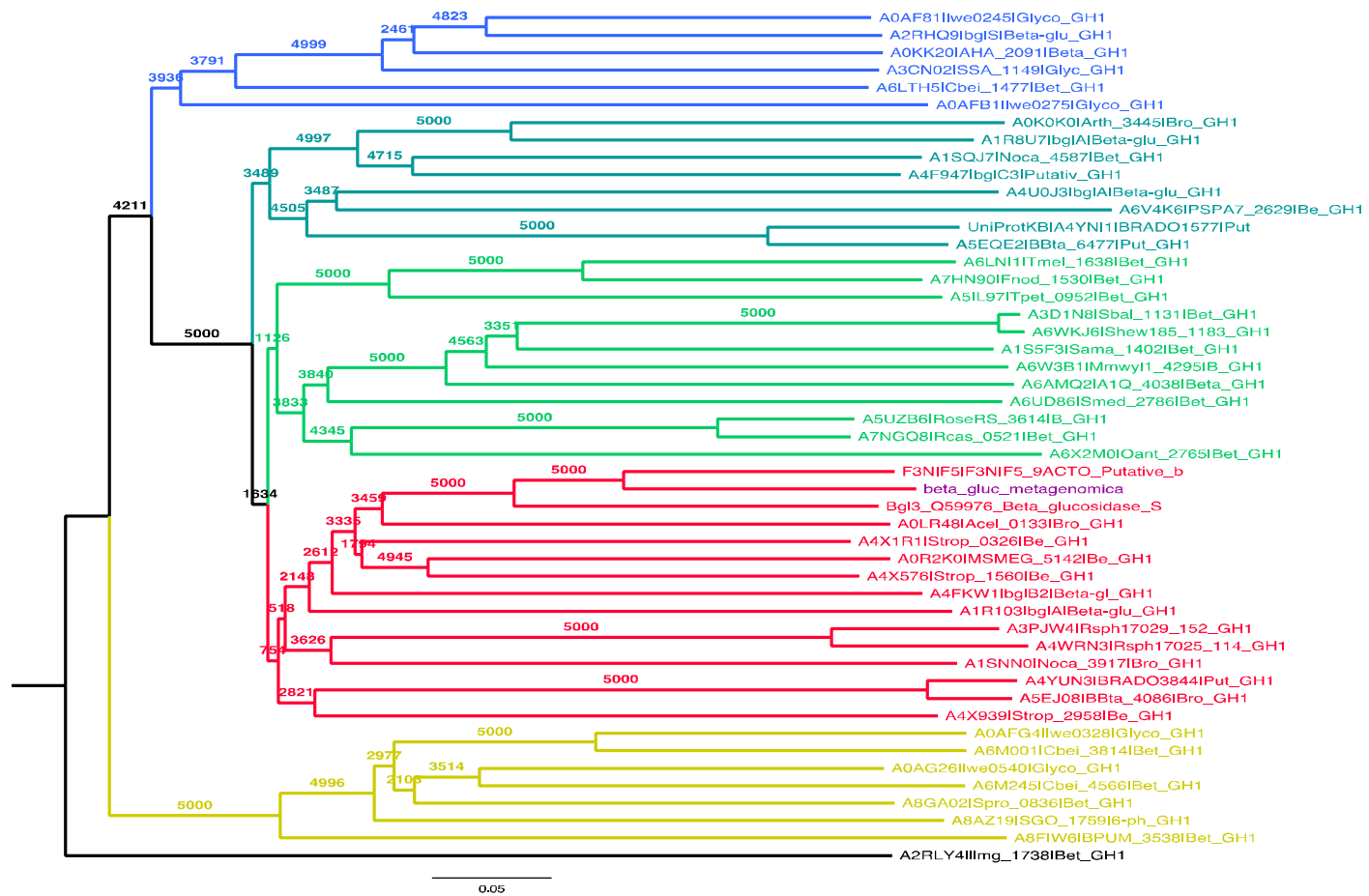


Figura 19. Árvore fenética para família de glicosil hidrolase GH1. Método: Alinhamento ClustalW, Matrix de substituição Blosum, penalidade para gap de abertura: 10; penalidade para gap de extensão: 5; Algoritmo de clusterização: neighbor Joining; índice de confiança para clusterização dos ramos por reamostragem bootstrap com 5 mil réplicas.

A análise de modelos preditivos de estrutura tridimensional *in silico* permite avaliar qual a arquitetura dos domínios conservados presentes na sequência, incluindo a geometria de seu sítio ativo. Pode-se visualizar também o balanço de cargas presentes na nova sequência, e se existem características únicas nesta que não equivalem a(s) proteína(s) com maior proximidade de homologia (Prosdociimi, 2001).

Sendo assim, a construção do modelo tridimensional para a ORF metagenômica foi realizada utilizando-se a interface de análise de estruturas proteicas do programa Swiss-model, através da ferramenta “Automated Modeling”. O programa detectou maior similaridade com a estrutura 1GNX correspondente à  $\beta$ -glicosidase Bgl3 de *Streptomyces sp.*, uma estrutura homodimérica (VALLMITJANA *et al.*, 2001). Uma vez que houve uma boa qualidade no alinhamento entre molde e modelo (identidade de 69% de acordo com o programa Phyre (Kelley; Sternberg, 2009)), a estrutura tridimensional predita para a proteína metagenômica também retornou um modelo com duas cadeias (homodiméricas) para a estrutura predita para a  $\beta$ -glicosidase metagenômica. Entretanto, esta informação não é passível de ser obtida apenas com base em dados de similaridade de sequência, sendo necessária a sua confirmação por técnicas *in vitro*, como por exemplo, uma cromatografia por exclusão molecular.

Abaixo é possível visualizar o alinhamento entre as sequências de aminoácidos de 1GNX (GUASCH *et al.*, 1999) e da  $\beta$ -glicosidase metagenômica (Fig.20; apenas uma parte do alinhamento é ilustrada). Indicados pelas setas estão os dois resíduos de glutamato (E) que fazem parte do sítio catalítico conforme determinado experimentalmente para a enzima Bgl3 (VALLMITJANA *et al.*, 2001). O primeiro resíduo de glutamato (Bgl3 é o Glu178) está relacionado com a ação geral de resíduo ácido/básico para interação com a ligação glicosídica beta (1->4); enquanto o segundo resíduo de glutamato (Bgl3 é o Glu388) está envolvido na ação de nucleofílico catalítico nos mecanismos de retenção da Bgl3 (GUASCH *et al.*, 1999). Estes dois resíduos são extremamente conservados entre  $\beta$ -glicosidasas pertencentes à família GH1, ao passo em que  $\beta$ -glicosidase da família GH3 (com base em algumas poucas estruturas resolvidas) utilizam um aspartato no ataque



nucleofílico e um glutamato como doador de prótons (CANTAREL et al., 2009; COSTA, 2013).

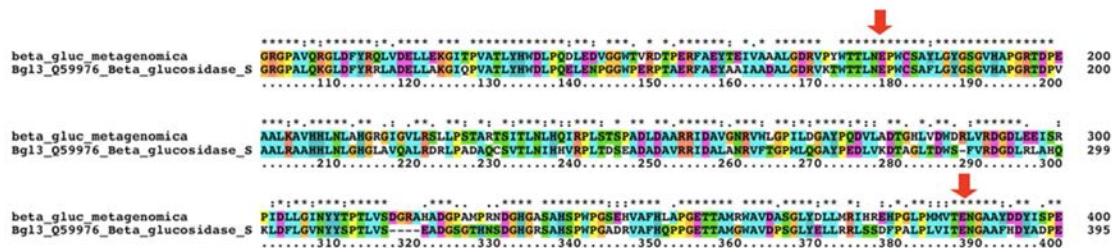


Figura 20. Alinhamento entre molde (Bgl3, estrutura 1GNX) e modelo de estrutura para a ORF metagenômica. Setas: resíduos envolvidos na hidrólise da ligação glicosídica (glutamatos, E).

Conforme pode ser observado na Figura 21, o modelo predito de estrutura apresentou cadeias homodiméricas com bolsão catalítico organizado em um enovelamento tipo TIM/Barril, com estruturas secundárias organizadas no arranjo em estrutura supersecundária (Motivo) ( $\alpha/\beta$ ) $_8$ . Esta estrutura é compartilhada pelas glicosil-hidrolases dos clãs *GH-A*, *GH-D*, *GH-H* e *GH-K*. A enzima  $\beta$ -1,4-glucosidase (E.C. 3.2.1.21) é uma atividade que está presente no clã *GH-A* (o maior e mais diverso destes) englobando as famílias 1, 5, 30. Estes dados corroboram a evidência de homologia entre a sequência metagenômica e as *GH1* (Clã *GH-A*). No modelo é possível visualizar os resíduos do sítio catalítico no interior do bolsão catalítico (o “pocket”, Fig.21, 2C).

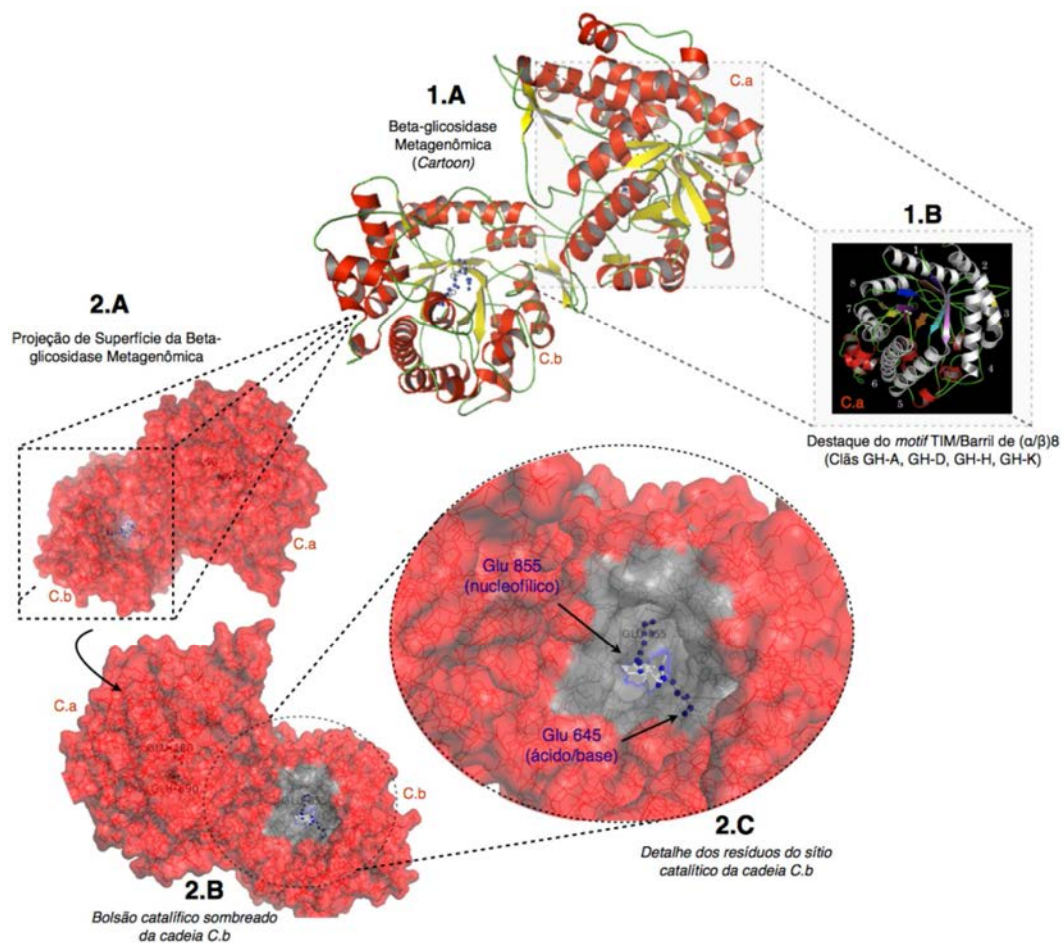


Figura 21. Modelo da estrutura da  $\beta$ -glicosidase metagenômica. Destaques para o sítio ativo e superfície (2.A-C) e “motif” de estrutura super-secundária, 1.A-B.

Além dos resíduos principais diretamente envolvidos na catálise, as enzimas  $\beta$ -glicosidases compartilham resíduos que atuam sobre o direcionamento do substrato para a correta interação do sítio catalítico com a porção glicona do substrato, os quais foram localizados na enzima metagenômica, fazendo parte do bolsão catalítico (Fig. 22). A localização de todos os resíduos fundamentais para determinada catálise é um bom indício de que a enzima é funcional, uma vez que mutações podem levar a perda de determinadas funções em muitos micro-organismos, através de deleções ou mudanças no quadro de leitura.

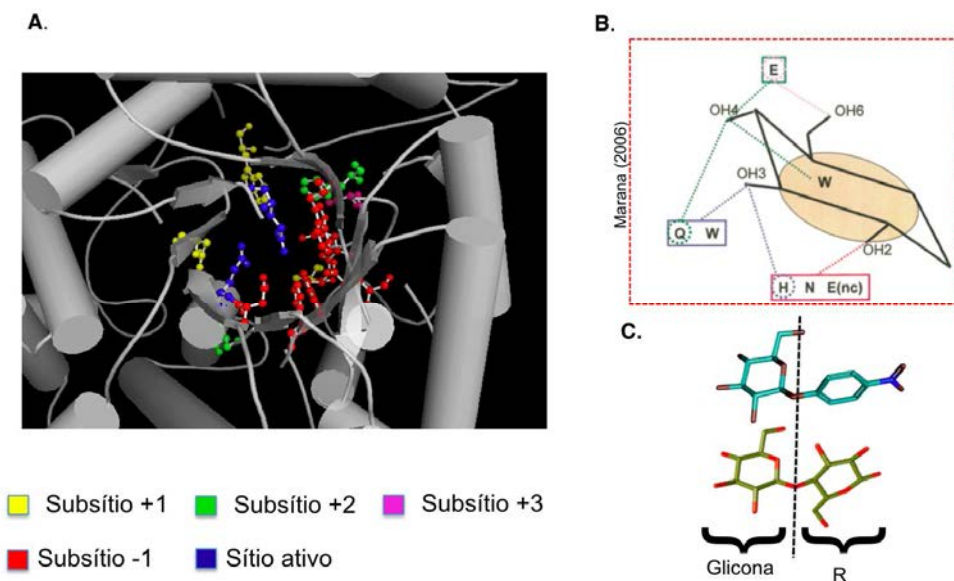


Figura 22. Modelo do sítio ativo estrutura da  $\beta$ -glicosidase metagenômica. Destaques para o sítio ativo (A). B: interação dos resíduos dos subsítio -1 para com a porção glicona do substrato (Marana, 2006). C: Exemplos de substratos glicosídeos porções gliconas no substrato p- nitrofenolato.

#### 4.1.2. Análise *in silico* da celulase

As celulasas propriamente ditas compõem um grupo de carboidrases com diversos mecanismos de ação, estruturas e relações de homologia. As principais classes de celulasas são as endo- e exo- celulasas, com atuação randômica amorfa ou nas extremidades não redutoras, respectivamente (LEMOS e STRADIOTTO, 2012; LI et al., 2012; PERCIVAL ZHANG, HIMMEL e MIELENZ, 2006).

A despeito da identificação conclusiva da classe enzimática só possa ser obtida através de abordagens experimentais frente a diferentes substratos para desvendar o perfil catalítico da enzima, as análises *in silico* são de caráter fundamental para traçar o delineamento experimental de forma mais eficiente. Muitas vezes nos deparamos com perfis catalíticos muito diferentes dos já descritos, de forma que apenas estudos detalhados de estrutura associadas a ligantes, e mutações pontuais em candidatos aos resíduos catalíticos podem esclarecer quais os mecanismos envolvidos na catálise. Nestas situações, as análises *in silico* se mostram bastantes úteis para identificar divergências entre os *motifs* catalíticos, e escolher quais resíduos substitutos poderiam interferir menos possível no restante da molécula a ponto de serem bons candidatos para identificar o sítio catalítico.

Em vista do exposto, a sequência obtida para a ORF11 do inserto metagenômico foi submetida às análises de similaridade de sequência em busca de resultados de homologia para domínios conservados que permitissem elaborar uma hipótese funcional.

A submissão da sequência metagenômica da ORF11 ao blast do Uniprot retornou similaridade com carboidrases (“sugar hydrolases”) de micro-organismos do gênero *Streptomyces*, com um valor máximo de similaridade de 72%. Dentre as sequências que apresentaram maior similaridade à sequência metagenômica, foram selecionadas 23 sequências para uma análise de alinhamento global (ClustalW, programa Bioedit, matrix de substituição blosum, com respectivas penalidades 10 e 5, para os Gaps de abertura e extensão).

Uma vez comprovada a existência de regiões de conservação no alinhamento, este foi submetido ao programa ClustalX para o cálculo da matrix de distância entre as sequências empregando o critério de otimização “Neighbor Joining”, e a construção do dendrograma fenético (análise de confiança de agrupamento dos ramos “bootstrap”, com 1000 re-amostragens).

A análise agrupou as sequências analisadas em quatro clados, sendo que a celulase metagenômica apresentou maior proximidade com a sequência F8JT81, uma carboidrase isolada de *Streptomyces cattleya* (Fig.23, clado em rosa). O clado representado em verde é o mais distante dos demais, e apresentou a maioria das sequências identificadas como ligantes de celulose do tipo II, sendo que a sequência mais dissimilar dentro do clado foi uma quitinase bi-funcional.

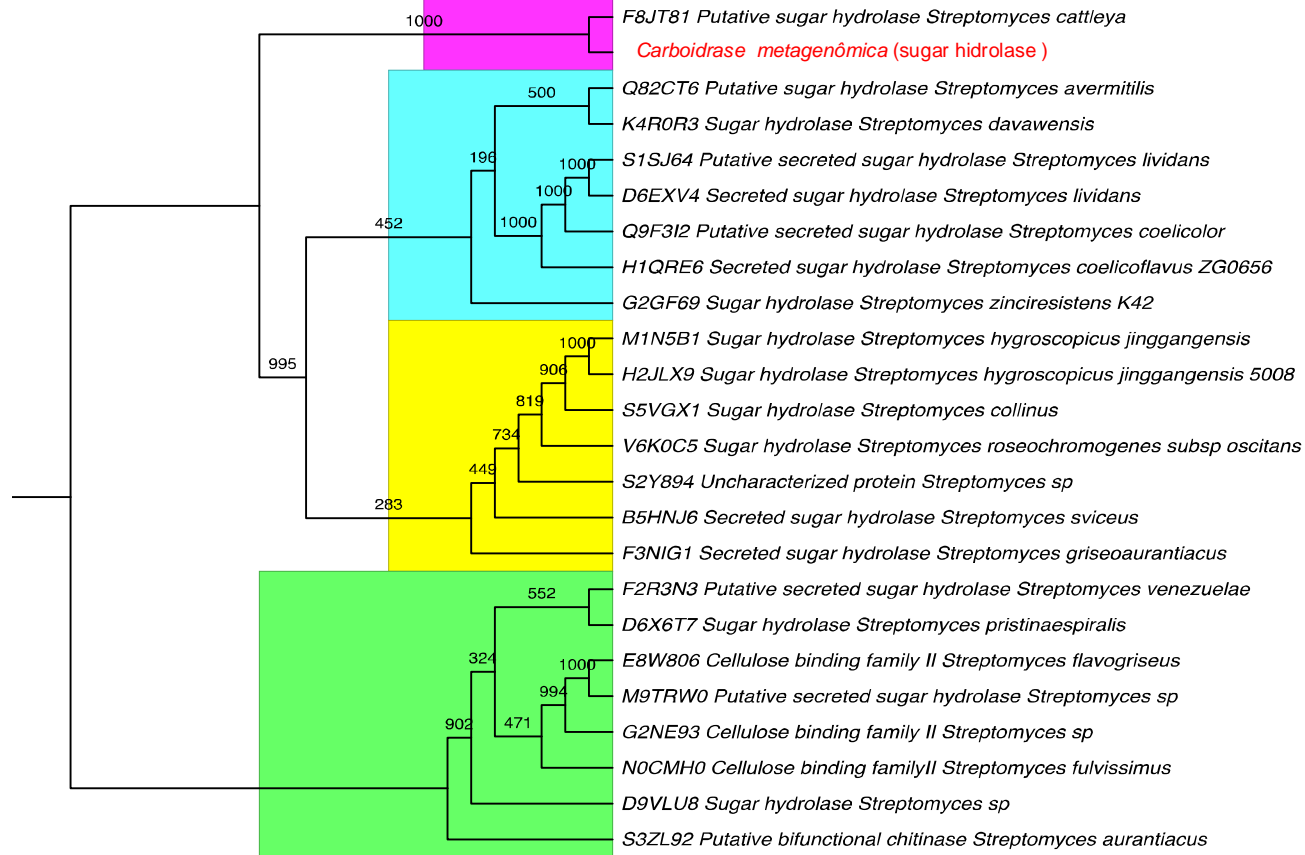


Figura 23. Dendrograma fenético de comparação entre a sequência metagenômica da ORF11 e as sequências de carboidrases obtidas junto ao Uniprot.

De forma similar à descrita para a  $\beta$ -glicosidase metagenômica, foi feita a predição *in silico* da estrutura proteica da ORF correspondente à celulase, sob o intuito de tentar descobrir quais as características estruturais poderiam fornecer eventuais pistas para a classificação da enzima.

Através dos dados obtidos a partir do modelo tridimensional proteico para a ORF11 metagenômica, e da análise de alinhamento global, foram delimitadas as três regiões que compõem a celulase ligada ao CBM (CBM, conector e módulo catalítico) nas sequências analisadas. A análise iniciou-se de um alinhamento com 25 sequências, incluindo as sequências metagenômica e Endoglucanase-EngD. Estas duas últimas serviram para delimitar as porções da sequência candidatas a cada uma das referidas regiões, tendo por base as informações dos resíduos que participam do dobramento estrutural equivalente ao modelo obtido para a estrutura metagenômica e o observado no EngD.

Após o alinhamento e identificação dos pontos conservados, cinco das sequências com maior similaridade no alinhamento foram mantidas junto à sequência metagenômica. Em seguida, os gaps foram excluídos e a sequência foi realinhada. Conforme pode ser visualizado na figura 24, a sequência para os módulos catalíticos da celulase e do CBM apresentaram um elevado nível de conservação, enquanto que a porção equivalente ao conector apresentou baixa conservação. É conhecido que regiões proteicas envolvidas diretamente na catálise tendem a ser mais conservadas, uma vez que alterações nos bolsões catalíticos frequentemente causarem perda de função enzimática. Aparentemente as regiões correspondentes aos conectores apresentam maior plasticidade evolutiva, potencialmente decorrente de sua função estrutural não depender de um arranjo tão específico e ordenado de aminoácidos quanto o presente em sítios catalíticos.

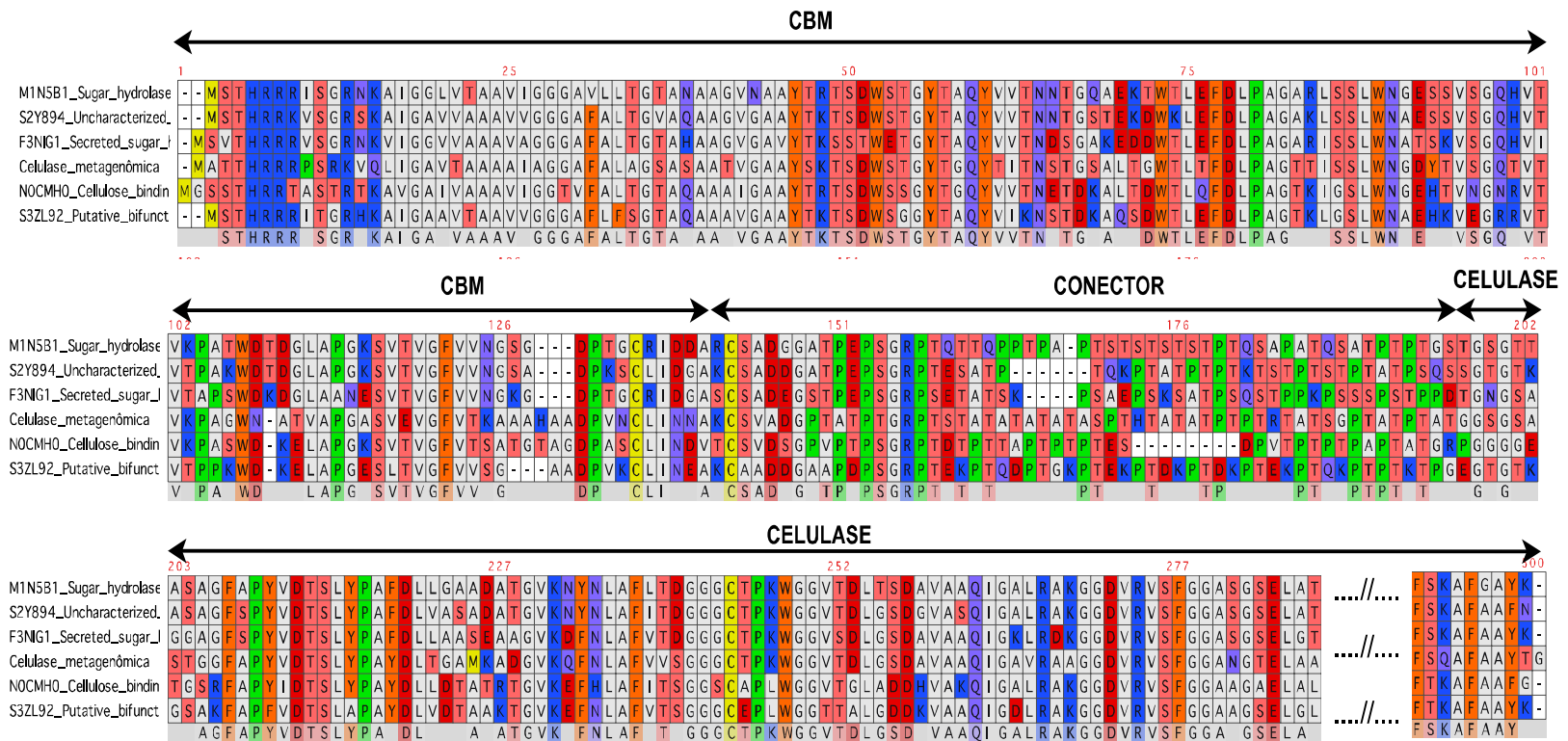


Figura 24. Alinhamento para a celulase metagenômica e outras de estruturas proteicas conhecidas. Evidência de conservação respectiva entre os módulos CBM e catalítico Celulase.



A figura 25 ilustra uma comparação entre a estrutura do módulo catalítico da enzima EngD e seu CBM, e os modelos das respectivas estruturas propostas para a enzima metagenômica. A estrutura para o *motif* catalítico propriamente dito para a celulase (Fig. 25.B) se mostrou bastante similar à estrutura obtida para a  $\beta$ -glicosidase (Fig21-2C), com o mesmo padrão Tim/Barril (arranjo em estrutura super-secundária no Motivo  $(\alpha/\beta)_8$ ). Este motivo é compartilhado por diversas carboidrases (glicosil-hidrolases dos clãs GH-A, GH-D, GH-H e GH-K), e não equivale a uma correspondência de sequências de aminoácidos, mas apenas de compartilhamento de estrutura super-secundária para estas famílias de GHs.

Com relação ao módulo CBM, as estruturas apresentaram bastante similaridade. Uma diferença importante entre elas é a substituição do terceiro resíduo conservado, que é uma tirosina no CBM EngD (tríade W393,W430 e Y448), enquanto no módulo CBM metagenômico é encontrado um triptofano ocupando o lugar deste resíduo (tríade catalítica W51, W88, W106).

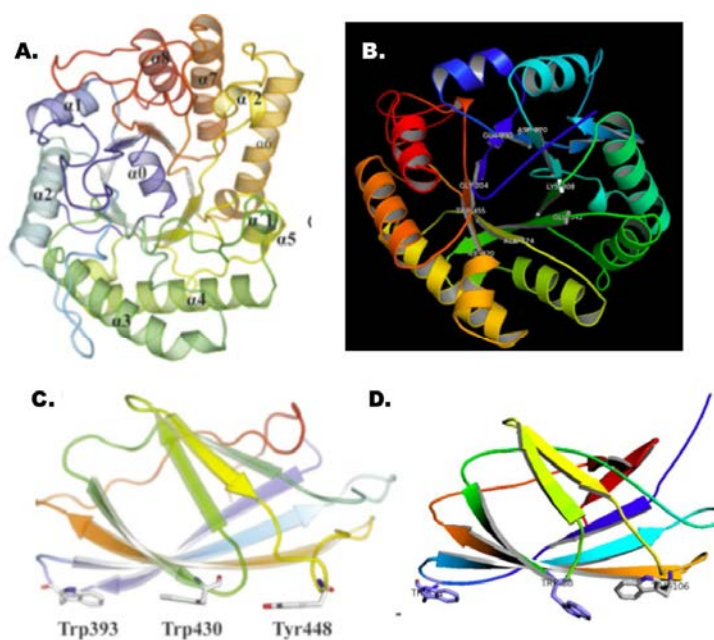







Figura 25. Comparação entre os respectivos módulos estruturais catalíticos (figuras dispostas acima) e CBMs (dispostas abaixo) para Endoglucanase EngD (à esquerda, Bianchetti, et al. 2012) e os obtidos para a ORF11 metagenômica (direita).



A análise de domínios indicou quatro estruturas de CBM2 compatíveis com o domínio metagenômico. Buscando identificar qual destas estruturas era mais similar a CBM metagenômica, foi realizado alinhamento entre elas no programa Pymol. O resultado do desvio entre os ângulos dos aminoácidos alinhados (RMS) e os três resíduos conservados para as CBM (propostos por BIANCHETTI et al., 2012) ou seus respectivos substitutos no alinhamento foram apresentados na tabela 10 e na figura 26.

Tabela 10. Resultado para alinhamento entre a CBM metagenômica e CBMs tipo 2

Estrutura	Cadeia	RMS do alinhamento individual (Å)	Triade	Cor
Metagenômica	A	-	W51, W88, W106	
1exh	A	0,263 (86 para 86)	W17, W54, W74	
3ndy	E	1,490 (98 para 98 )	W393, W430, Y448	
1hej	C	1,745 (50 para 50)	W570, W602, S620	
2cwr	A	3,094 (71 para 71)	W274,W308, W326	
<b>RMS para Todas</b>			<b>3.094 (71 para 71)</b>	

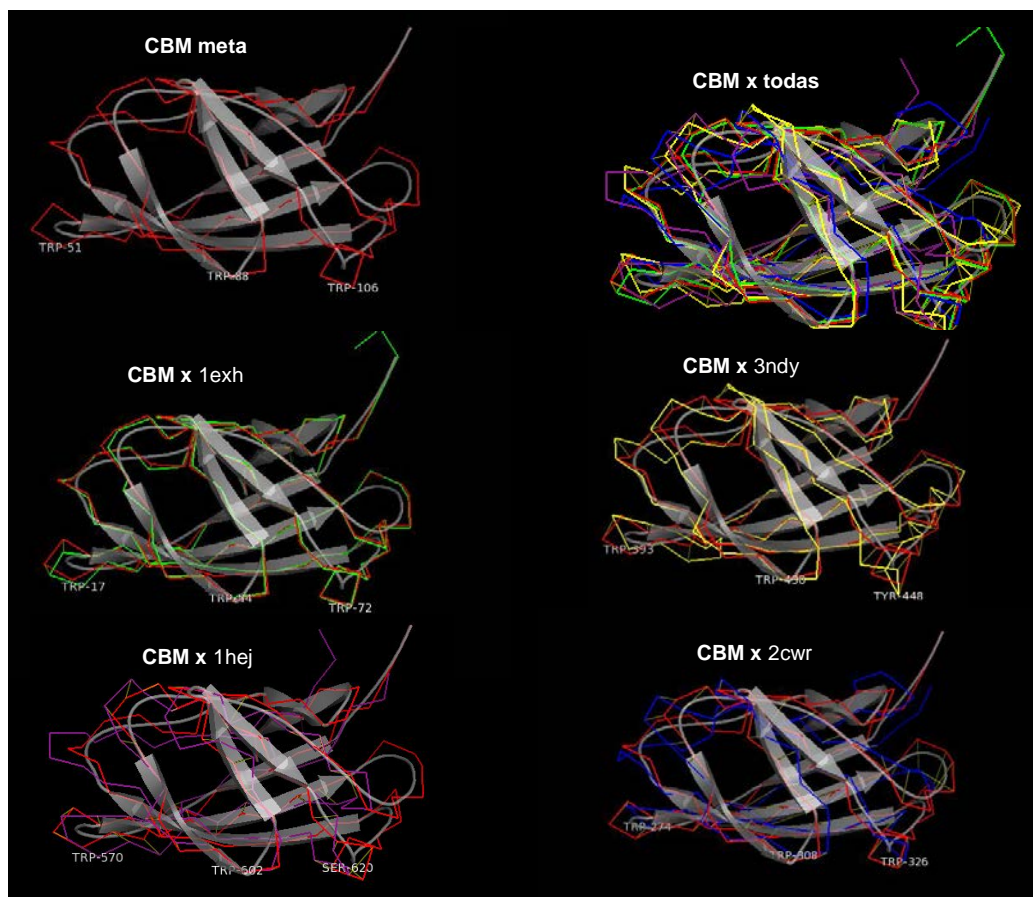


Figura 26. Alinhamento entre CBMs tipo 2 e o modelo de estrutura proposto para a CBM metagenômica.

Como pode ser visto na tabela 10, a estrutura que apresentou um dobramento mais similar à da CBM metagenômica foi a **1exh**, com RMS de 0,263 Å. Dentre as CBMs, houve a conservação dos dois primeiros resíduos dos três propostos por BIANCHETTI e colaboradores (2012). O terceiro resíduo, a tirosina (Y), foi substituída por resíduos de triptofano nas estruturas **metagenômica**, **1exh** e **2cwr**, e por uma serina na estrutura **1hej**.

A despeito dos resultados obtidos para similaridade com celulases, a sequência apresentou domínios conservados presentes e quitinases, na família GH18. Os domínios conservados foram analisados através do uso da base de dados Prodom ([www.prodom.prabi.fr/prodom/current/html/home.php](http://www.prodom.prabi.fr/prodom/current/html/home.php)). O gráfico obtido para os domínios de maior pontuação identificados na sequência de aminoácidos da ORF11 é apresentado na figura 27.

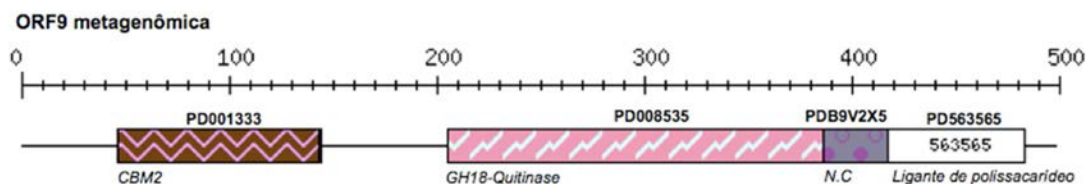


Figura 27. Domínios catalíticos obtidos para a ORF11 metagenômica (base de dados *Prodom*).

Considerando estes dados, a possibilidade da enzima metagenômica ser uma quitinase não deve ser descartada, de forma que estudos empregando substratos celulose e quitina serão importantes para investigar a função desta enzima.

#### 4.1.3. Análise *In Silico* da Lacase

As lacases (EC 1.10.3.2) pertencem à superfamília das “Multi-copper oxidases” (MCOs), que são um grupo de enzimas que oxidam substratos em um centro mononuclear cúprico (T1). Os elétrons obtidos são então transferidos internamente para um centro cúprico tri-nuclear (T2/T3), onde uma molécula de oxigênio é reduzida por quatro elétrons, liberando duas moléculas de água. Os íons de cobre são classificados em T1, T2 ou T3 de acordo com suas propriedades espectroscópica. Desta forma, pelo menos quatro átomos de cobre estão presentes em MCOs e são compartilhados entre os sítios T1 e T2/T3 da enzima (QUINTANAR et al., 2007; REISS et al., 2013).

Como pode ser visualizado na figura 28, o átomo de cobre T1 é ligado à enzima por dois resíduos de histidina e uma cisteína, formando uma ligação organometálica. Além disso, as cadeias laterais de uma metionina, uma leucina e uma isoleucina (os dois últimos resíduos não estão representados) se situam bastante próximos ao sítio cúprico T1 (ENGUITA et al., 2003). Os átomos de cobre T2 do sítio cúprico tri-nuclear, por sua vez, é coordenado por dois ligantes de histidina, enquanto os dois átomos de cobre tipo T3 (que também compõem o sítio T2/T3) têm no total seis histidinas ligantes (REISS et al., 2013).

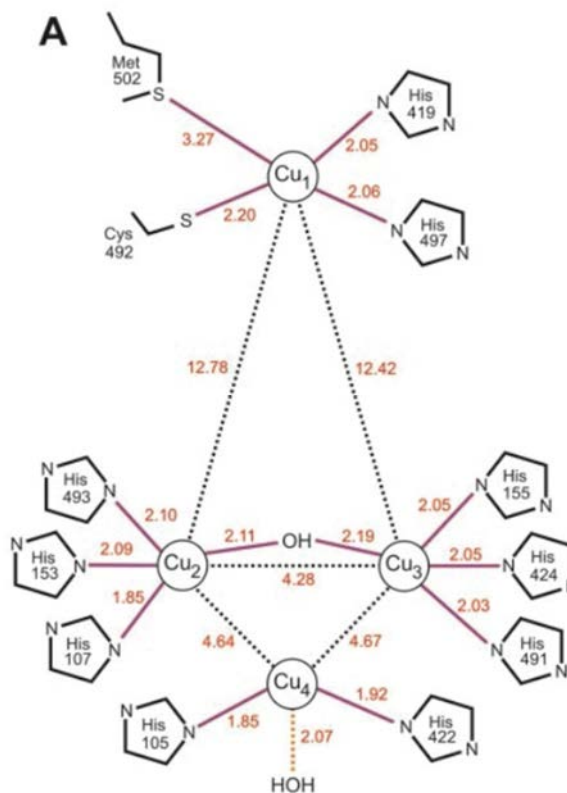


Figura 28. Exemplo da estrutura da ligação organometálica presente na multicopper oxidase CotA. O esquema representa os dois centros cúpricos, incluindo as distâncias inter-atômicas mais relevantes. Fonte: ENGUITA et al., 2003.

Embora a homologia de sequências entre as diferentes famílias de “Multicopper Oxidases” (MCOs) seja baixa, o alinhamento entre as sequências de aminoácidos revela que algumas estruturas e *motifs* relacionados a ligação de cobre são altamente conservadas (REISS et al., 2013; GIARDINA et al., 2010). Dentre estas ilhas de conservação, podemos destacar quatro *motifs* subsequentes e estritamente conservados dentro da sequência primária de aminoácidos: HXHG, HXH, HXXHXH e HCHXXXHXXXM/L/F. Com relação a estrutura tridimensional, por sua vez, as MCOs compreendem 2, 3 ou 6 domínios homólogos, cada um compartilhando um dobramento similar à cupredoxina, que possui estrutura supersecundária Barril  $\beta$  Chave Grega, com oito fitas distribuídas em um padrão de folhas beta, contendo cada uma quatro fitas.

A sequência obtida para a lacase metagenômica é bastante interessante pelo fato de ser muito diferente das outras depositadas no banco de dados, com um valor de identidade de apenas 39% com uma “Multicopper” oxidase tipo 3 (UP U5QPS1) *Gloeobacter kilaueensis* JS1. Na figura 29 é possível ver o alinhamento entre a sequência metagenômica e algumas sequências de “Small laccases”, as SLACs (Ebiox, algoritmo Muscle). O alinhamento demonstra baixa conservação entre a sequência metagenômica e as demais, entretanto é possível detectar os quatro *motifs* extremamente conservados das “multicopper oxidases” relacionados à ligação de cobre.

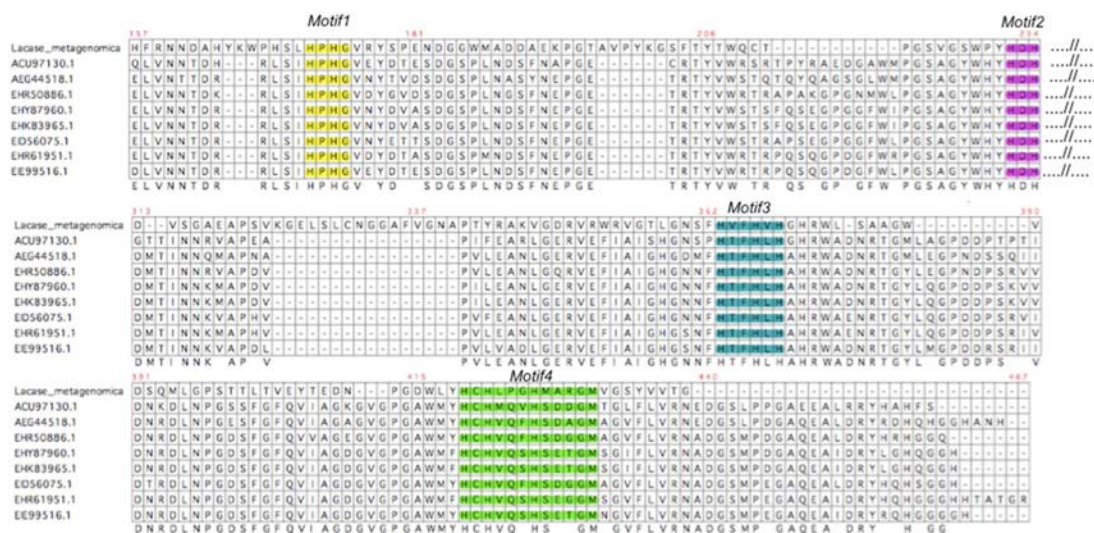


Figura 29. Motifs conservados entre as “multicopper oxidases” do tipo Slac relativos aos ligantes de cobre.

A submissão da sequência à base de dados “The Laccase and Multicopper Oxidase Engineering Database”- LccED (<http://www.lcced.uni-stuttgart.de/cgi-bin/LccED1.2/index.pl?page=fam>) retornou uma maior semelhança com sequências pertencentes a família das SLAC (“Small laccases”), que correspondem a subfamília K.

Tendo em vista que se trata de uma sequência muito diferente das outras depositadas, houve uma dificuldade maior em se obter um modelo de simulação de estrutura tridimensional proteica para a lacase metagenômica. O melhor modelo obtido é apresentado abaixo (Fig. 30).



Figura 30. Modelo de estrutura para a MCO. Modelo de estrutura tridimensional gerado pelo software *Swiss-Model* para a Lacase metagenômica, obtido através da similaridade com o molde 2j5W.1.A, com 30% de identidade com a ORF metagenômica.

As regiões possivelmente mais problemáticas no modelo estrutural são evidenciadas no gráfico da figura 30, correspondentes aos valores de similaridade ao molde abaixo de 0,4. Contudo, o modelo apresentou qualidade aceitável para inferências de homologia e permitiu verificar que houve uma boa correspondência entre o arcabouço das duas proteínas, uma vez que o alinhamento entre as estruturas do molde (3tbc, da família SLAC, com dois domínios para multi-copper oxidase) e modelo (Fig. 31) retornou uma boa pontuação de RMS: 0.670 (157 para 157 átomos), considerando que o valor máximo aceitável para um bom alinhamento é de até 1.5 Å de desvio.

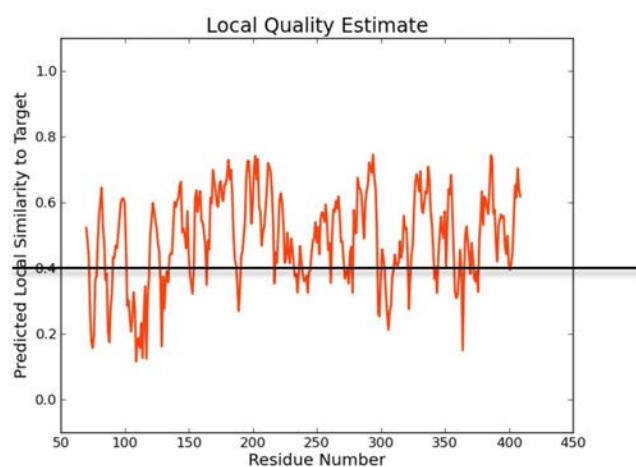


Figura 36. Estimativa de qualidade local para o modelo obtido para a lacase metagenômica.





Figura 31. Alinhamento entre as estruturas do molde (3tbc, da família SLAC, com dois domínios para multi-copper oxidase) e modelo para a lacase metagenômica. (Modelo gerado por homologia no Swiss-model).

Buscando identificar qual a família de lacase mais próxima à sequência metagenômica, esta foi alinhada à representantes de lacases encontradas em bactérias disponíveis na base de dados “The laccase and multi copper oxidase engineering database” (LCCED). O alinhamento obtido no programa EbiOx, utilizando a ferramenta clustalW foi utilizado para a construção de árvore pelo método de distância “Neihgboor joining”. O resultado pode ser observado na figura 32.

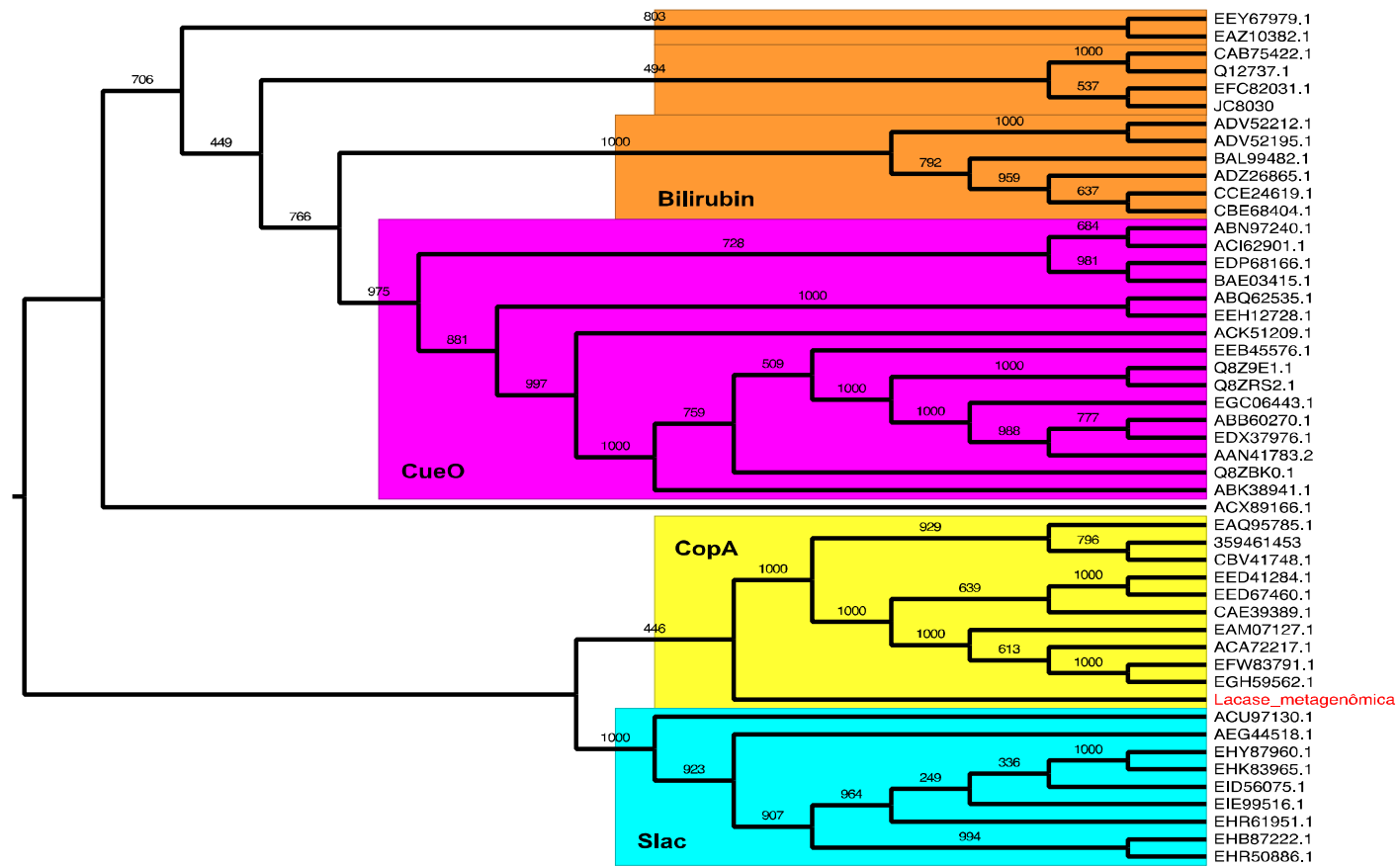


Figura 32. Dendrograma fenético de comparação entre a sequência metagenômica para MCO e sequências presentes no banco de dados "The laccase and multi copper oxidase engineering database" (LCCED).



## 4.2. Amplificação das ORFs metagenômica relacionadas ao metabolismo de biomassa lignocelulósica

Os pares de oligonucleotídeos desenhados para as sequências das ORFs metagenômicas foram inicialmente alinhados a sequência do *contig* B5p137, para confirmar pareamento correto com os genes de interesse. Para a representação do alinhamento, as sequências tomadas na fita de DNA de orientação 3'→5' foram representadas na forma direta, embora tenham sido sintetizadas na fita reversa e complementar (Fig. 33-35).

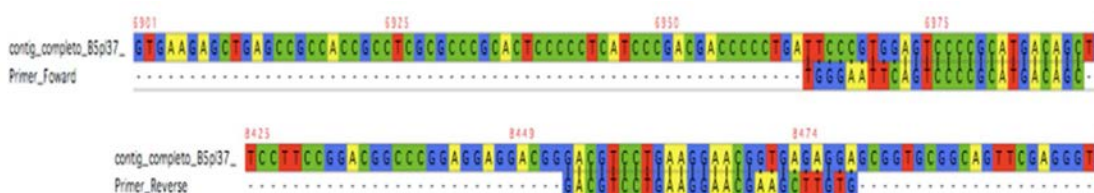


Figura 33. Confirmação de pareamento correto entre a sequência do “primer” e a região de interesse para a β-glucosidase metagenômica. O gene possui fase de leitura positiva, ou seja orientação na fita 5'→3'.

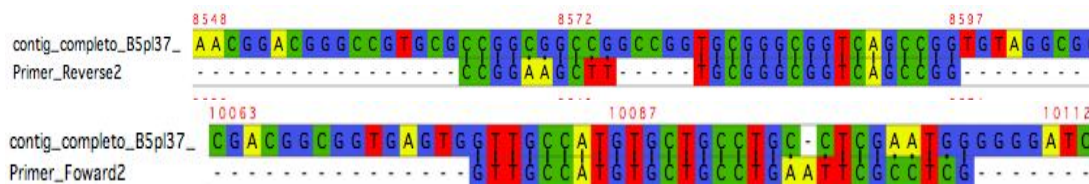


Figura 34. Confirmação de pareamento correto entre a sequência do “primer”2 e a região de interesse para a Celulase metagenômica. O gene possui fase de leitura negativa, ou seja orientação na fita 3'→5'.

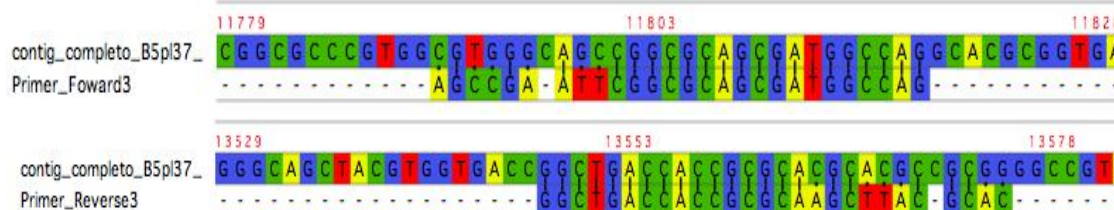


Figura 35. Confirmação de pareamento correto entre a sequência do par de “primer”3 e a região de interesse para a MCO metagenômica. O gene possui fase de leitura positiva, ou seja orientação na fita 5'→3'.

Adotando práticas laboratoriais rotineiras de otimização de reação de amplificação por PCR (temperatura, concentração de cloreto, oligonucleotídeos,

DNA, entre outras) foi possível obter a amplificação para os genes da  $\beta$ -glicosidase e celulase, conforme pode ser observado nas figuras 36 e 37, respectivamente.

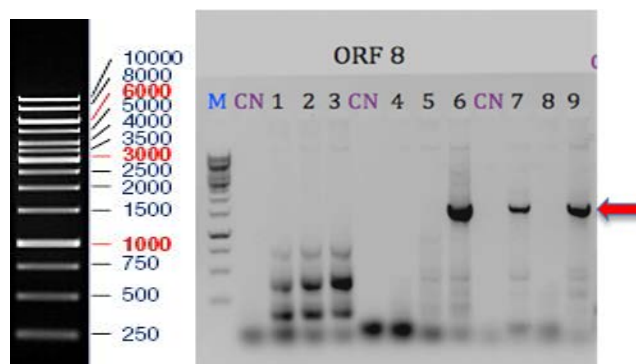


Figura 36. Amplificação por PCR do gene para  $\beta$ -glicosidase (1516pb). M.: Marcador de tamanho molecular 1kb plus, Fermentas; CN: controle negativo. Amostras de DNA do clone B5pl37 variando na concentração de 50, 100 e 200 ng para cada um dos respectivos genes em três condições de reação: 1-3: Reação com tampão comum; 4-6: Reação com tampão KCl, ambas reações com 1,6 mM de cloreto de magnésio; 7-9: Reação de KCl com 1,12 mM de cloreto de magnésio.

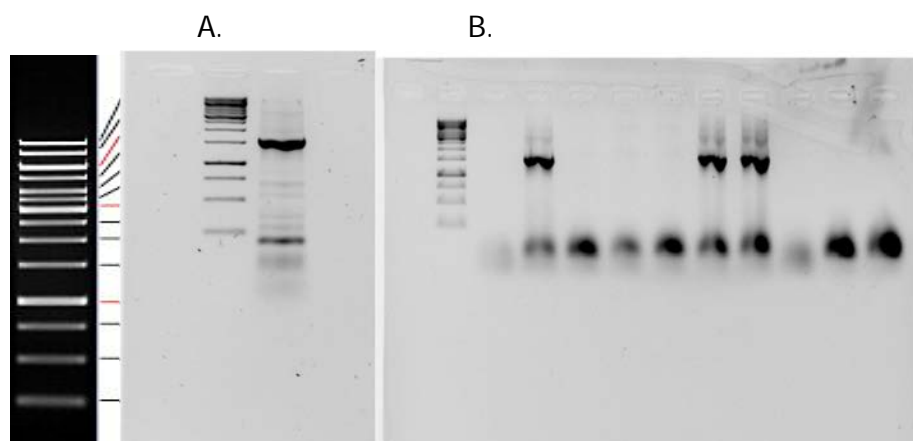


Figura 37. Amplificação por PCR obtida para a celulase metagenômica. M.: Marcador de tamanho molecular 1kb plus, Fermentas; A - amplificação obtida para o gene de interesse a partir do clone cosmidial metagenômico B5pl37. B - PCR de confirmação de ligação do inserto para os clones em vetor pET28a.

Não foi possível obter uma amplificação positiva para a “*Multicopper oxidase*”. A região onde esta ORF se encontra é uma região que não favorece a síntese de oligonucleotídeos de expressão, tendo em vista que as bases C/G são

predominantes em todo incerto metagenômico (72,6%), mas exibem uma predominância ainda mais acentuada nesta região (Figura 38). A predominância destas bases diminui o sinal filogenético da sequência, propiciando vários blocos idênticos de nucleotídeos. Estas regiões dificultam a síntese de oligos uma vez que a repetição de bases complementares favorece a formação de grampos e dímeros, o alto conteúdo G/C aumentam a temperatura de pareamento, e “primers” com base nos blocos repetitivos amplificam muitas regiões inespecíficas. A incorporação de sítios de restrição já confere um caráter “inespecífico” ao oligonucleotídeo, de forma que algumas regiões inespecíficas podem ainda ser mais facilmente ligadas ao “primer” do que a própria região de interesse.

```

GTCGCCCCGTCAGCAGGCGTCAACTCGGTCCGGTGGCACGCCTGACGACG
GGACGTCAGCCGGTCCGGGGCGCTGCGCGGGCACGGACCCCGAGGTCAC
CCGCCCGGCGGGACCGGCCGGTGGGGCGGGCCGGGTAATGACTGCAGTGG
TGGGCGGGAGGCGACCGCGGCCGGGCCTGGACGGGTCCGGGCGGGCGCCG
TCCGCCCCCGCGGGCGGGGCCCGCCGTCCCGCACTCGCACCCCCACCCCGA
GAGGACGTCATG . . . ACCGGCTGACCACCGCGCACGCACGCCGCGGGGCC
GTACGCACGGTGCGTACGGCCCCGCGGCGCGTGCGGGCCAGGTGTCCCGC
GGGCGTCAGCCGCGGCGGAACACCC

```

Figura 38. Região “upstream” e “downstream” ao gene da lacase. Em Vermelho: bases G/C, verde conteúdo A/T; cinza: gene.

Após várias tentativas de solucionar o problema (síntese de novos oligonucleotídeos, adição de DMSO, Nested, entre outras) optou-se por realizar a síntese do gene através de métodos *in vitro* junto à empresa “Genone”.

### 4.3. Expressão Enzimática e Purificação das carboidrases

#### 4.3.1- Celulase

Conforme pode ser visualizado na figura 39 foi possível obter a expressão para a enzima celulase em vetor pET28a.

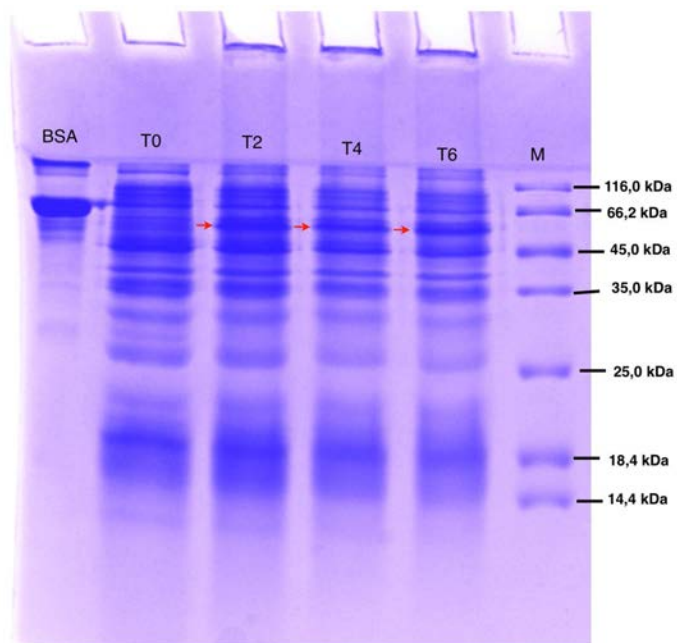


Figura 39. Teste de diferentes tempos de indução, de 0 á 6 horas (T0-T6). Cultivo de indução (0,1 mM de IPTG) à 37 °C. As amostras foram lisadas por aquecimento. Seta: Banda de interesse, com aproximadamente 55 kDa.

Entretanto, não foi possível superexpressar esta enzima na fração solúvel para as três linhagens analisadas (BL21, C41 e Artics), mesmo variando as condições de temperatura, concentração de indutor, tipo de extração, conforme exemplos abaixo (Fig. 40-41)

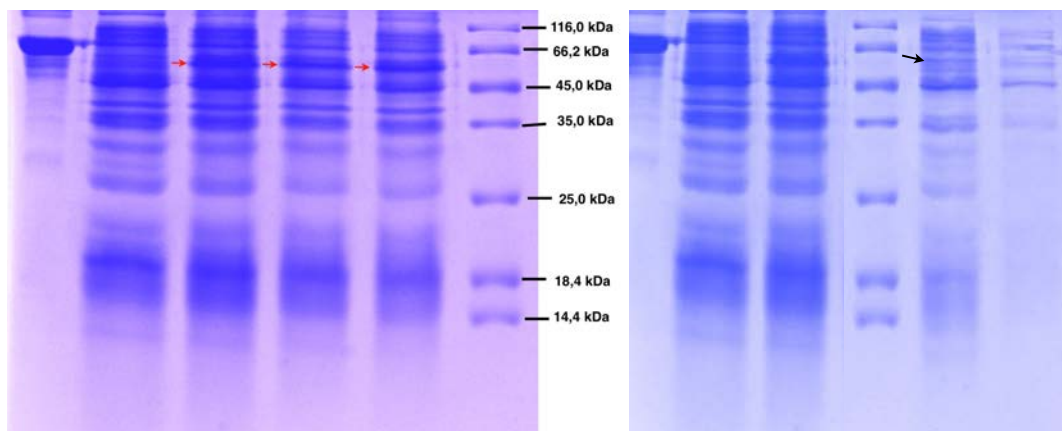


Figura 40. Expressão da Celulase 37 °C. **A:** Teste de diferentes tempos de indução, de 0 à 6 horas (T0-T6). **B.** Cultivo de indução (0,1 mM de IPTG) à 37 °C; **C.** Seta: Banda de interesse, com aproximadamente 54 kda.

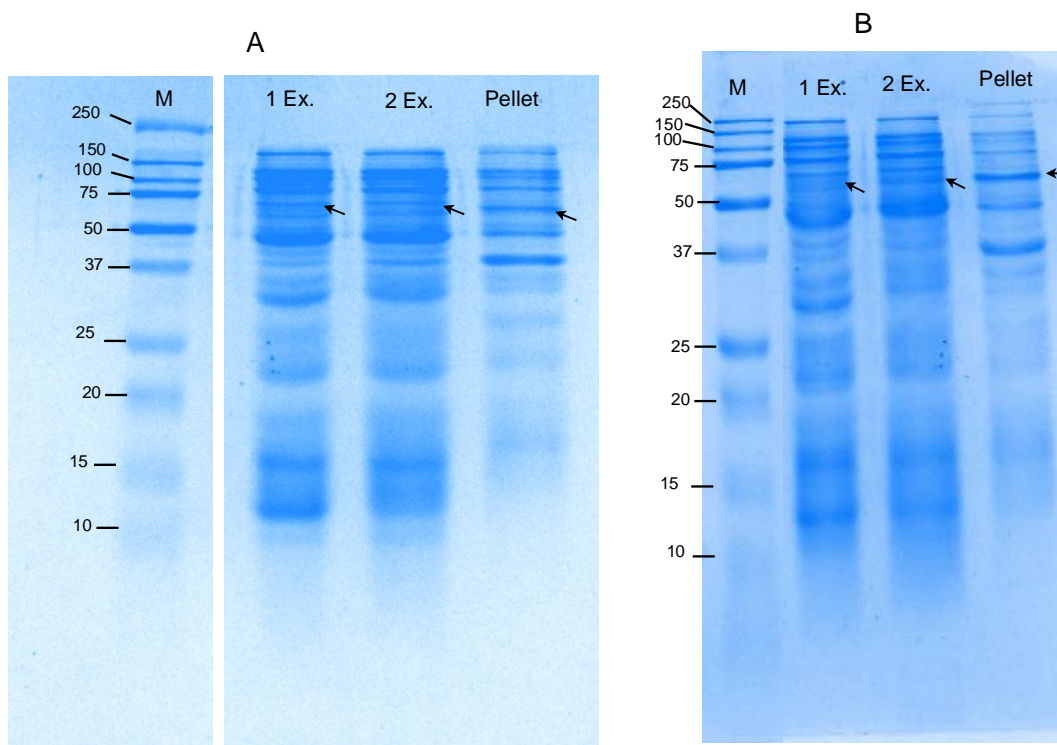


Figura 47. Expressão da Celulase 30 °C. **A:** Cultivo de indução (0,1 mM de IPTG) à 30 °C; **B:** Cultivo de indução (0,2 mM de IPTG) à 30°C. Seta: Banda de interesse, com aproximadamente 54 kda.

Uma análise da sequência de celulase em suas extremidades permitiu verificar que a mesma apresenta peptídeo sinal na região N-terminal (Fig. 42), o que



passou despercebido nas análises anteriores. Como estes possuem caráter fortemente hidrofóbico, além de um sítio de clivagem que pode ser reconhecido pelo hospedeiro a presença do peptídeo sinal pode dificultar a solubilidade da enzima de interesse, bem como pode levar a perda da cauda de histidina necessária para o passo posterior de purificação. Tendo em vista estes problemas, novos oligonucleotídeos foram sintetizados de forma a excluir esta região para posterior prosseguimento do trabalho.

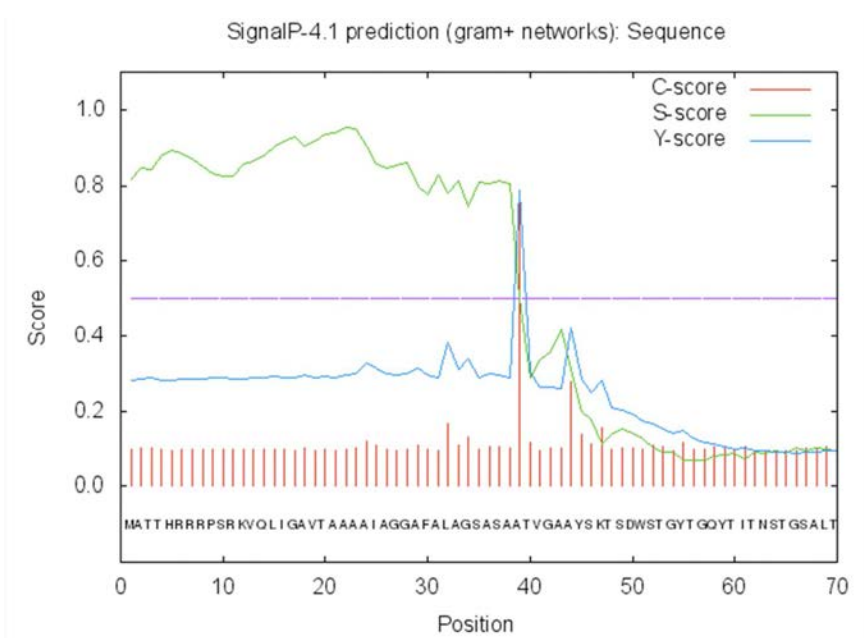


Figura 42. Detecção de peptídeo sinal na sequência proteica para a enzima metagenômica celulase. Programa signalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>).

#### 4.3.2 Multicopper oxidase (Lacase)

A enzima Multicopper oxidase foi transformada em células competentes de *E. coli* BL21 e expressa em vetor pET28A, à 37°C por seis horas e a sua extração com fosfato de sódio (pH 7,5) conforme pode ser visualizado abaixo, figuras 43 e 44.

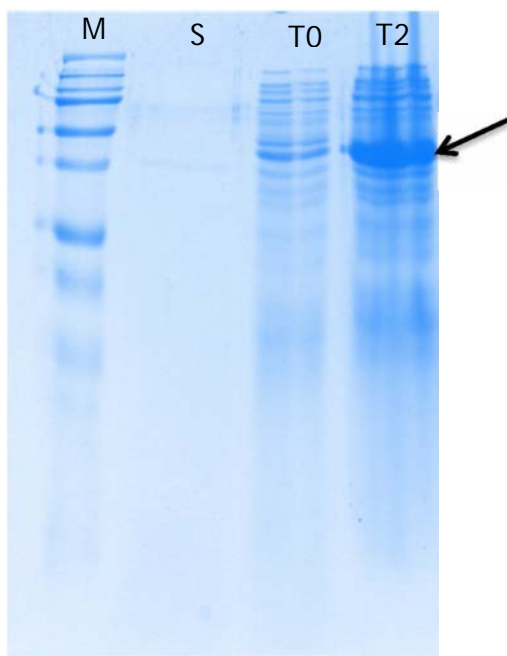


Figura 43. Expressão para Multicopper oxidase. M: Marcador de proteínas Bio-RAd (Cat. #161-0374); S: Sobrenadante; T0: Amostra sem indução, T2: duas horas após indução. seta: banda de interesse.

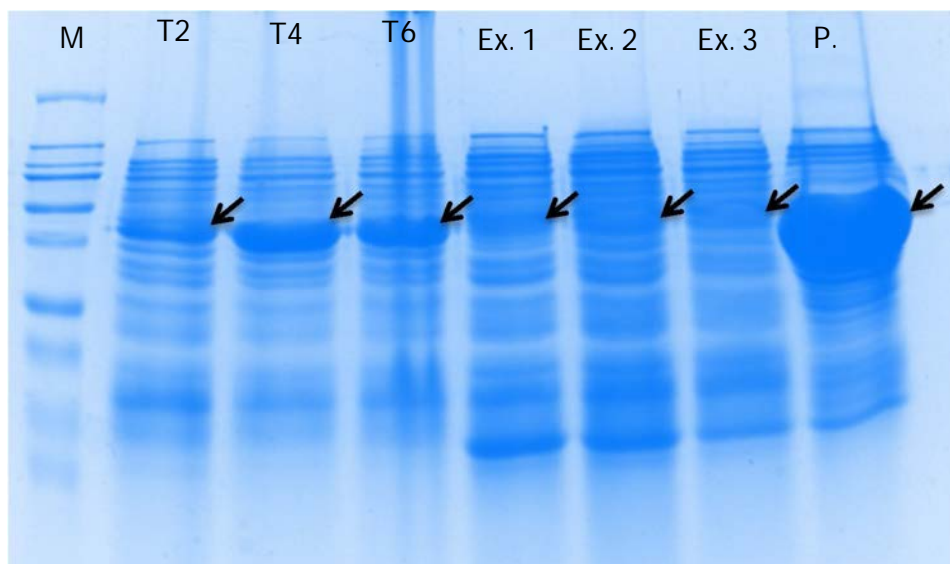


Figura 44. Expressão e extração para Multicopper oxidase. M: Marcador de proteínas Bio-RAd (Cat. #161-0374); T2-6: amostras retiradas após 2-6 horas de indução, respectivamente. Ex.1-3: extratos solúveis, P. *Pellet* de células.

Como pode ser observado nas figuras, houve a expressão da enzima de interesse. A melhoria da solubilidade da proteína poderá ser feita através de ensaios de expressão à temperaturas mais baixas.

### **4.3.3 $\beta$ -glicosidase**

#### **A. Detecção de Atividade $\beta$ -glicolítica e estimativa do tamanho molecular da enzima**

A enzima metagenômica Met-Bg foi clonada em vetor pet28A em *E. coli* BL21DE3 para expressão heteróloga de uma enzima de fusão com uma cauda hexa-histidina. Uma submissão da sequência de aminoácidos de Met-Bg ao programa Protparam permitiu estimar a massa molecular correspondente a um polipeptídeo de 58,33 kDa, com um coeficiente de extinção molar de 109780 M<sup>-1</sup> cm<sup>-1</sup> (Abs 0,1%; 280 nm).

Após a eletroforese em condição de proteína nativa (PAGE), a detecção da atividade  $\beta$ -glicolítica foi realizada para uma canaleta através de um Zimograma. A outra parte do gel contendo a enzima e o padrão BSA foi corada pelo método de Azul de Comassie. O padrão obtido foi então comparado ao encontrado para uma eletroforese SDS-PAGE, com desnaturação térmica da enzima na presença de  $\beta$ -mercaptoetanol, conforme pode ser visualizado na figura 45. Conforme pode ser visualizado no zimograma, a atividade enzimática está concentrada apenas na região correspondente a uma única banda (Fig. 45.C), o que não indica polimorfismos de estrutura quaternária.





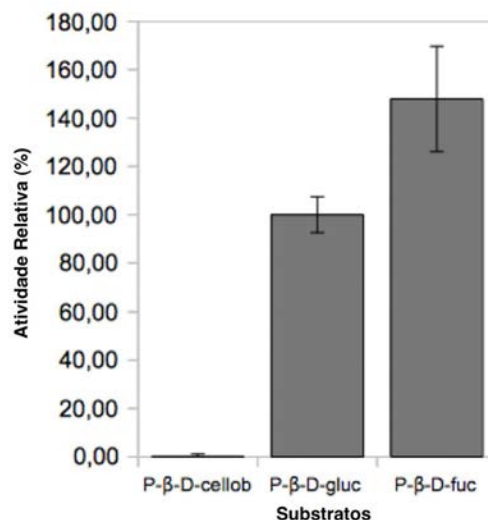


Figura 46. Ensaio enzimático para  $\beta$ -glicosidase frente a diferentes substratos. pNP- $\beta$ -D-glicopiranosídeo (pNP- $\beta$ G, padrão); pNP- $\beta$ -D-fucopiranosídeo (pNP- $\beta$ Fuc) e pNP- $\beta$ -D-celobiosídeo (pNP- $\beta$ Cb). Meio reacional: 1mM de substrato; Tampão fosfato de sódio pH7,5; temperatura de 37°C (10 min de incubação).

A atividade enzimática foi superior em 47% para o substrato  $p$ - $\beta$ -D-fuc em relação ao obtido para o padrão  $p$ - $\beta$ -D-gluc. Segundo Marana (2006), as enzimas  $\beta$ -glicosidase da família GH1 normalmente tem preferência por estes substratos, além do substrato  $\beta$ -galactosídeo, e esta preferência está diretamente associada a capacidade de melhor interação entre o glutamato reativo do subsítio -1 e a hidroxila OH2 destes glicosídios. Apesar de aparentemente compartilhar um mesmo sítio enzimático, uma atividade enzimática superior para o substrato  $p$ - $\beta$ -D-fuc, quando comparado ao substrato padrão  $p$ - $\beta$ -D-gluc também foi observada para a enzima  $\beta$ -glicosidase Bgl3. Os autores concluíram através de experimentos de cinética enzimática que a atividade diferencial se dava pelo fato de que a enzima sofria inibição alostérica pelo produto da degradação do substrato  $p$ - $\beta$ -D-gluc, o que não acontecia para o substrato  $p$ - $\beta$ -D-fuc.

Nenhuma atividade foi observada para o substrato  $p$ -Nitro-fenil- $\beta$ -D-celobiosídeo para enzima  $\beta$ -glicosidase nas condições avaliadas.

### C. Influência do pH e da temperatura

Dentre os fatores que influenciam a atividade enzimática, o pH assume papel importante nos diferentes níveis de protonação da cadeia peptídica e grupos

prostéticos. Essa influência assume um caráter mandatório na conformação do sítio catalítico e na sua capacidade de interação com substratos, cofatores, inibidores, ativadores, e mesmo com o produto em estágios intermediários de catálise ou inibição alostérica. Os substratos também podem conter grupos ionizáveis para os quais a enzima é estereo-específica. A temperatura, por sua vez, exerce um papel duplo na cinética da reação enzimática: a elevação gradual da temperatura leva à aumentos graduais de atividade até que temperaturas ligeiramente mais altas do que a de atividade ótima levam ao rompimento das pontes de hidrogênio, desencadeando uma cascata de alterações estruturais, levando a enzima a uma nova conformação, ou a um estado sem estrutura definida (Mariotto, 2006).

A figura 47 expressa os valores de atividade relativa (%) obtidos para a enzima Met-Bg nos experimentos do efeito do pH (Fig.47.A) e da temperatura (Fig.47.B). Conforme pode ser visualizado no gráfico (Fig. 47.A), o pH ótimo da enzima ficou entre pH 7,0-7,5. A influência do tipo de tampão utilizado foi nítida, uma vez que a atividade variou bastante para um mesmo pH, sendo que o tampão fosfato de sódio foi o melhor tampão para a faixa de pH 6,5-8,0.

De acordo com dados explorados na base de dados sobre enzimas “Brenda”, através da ferramenta “Functional Enzyme Parameters”, para  $\beta$ -glicosidase (E.C. 3.2.1.21) bacterianas, a maior parte das enzimas depositada apresenta pH ótimo entre 6-7 (26 enzimas), seguido da faixa pH5-6 (10 enzimas) e pH7-8 (8 enzimas). Apenas uma enzima apresentou pH ótimo muito alcalino (pH8-9), enquanto que as faixas mais ácidas de pH 3-4 e pH4-5 contam com três enzimas cada.

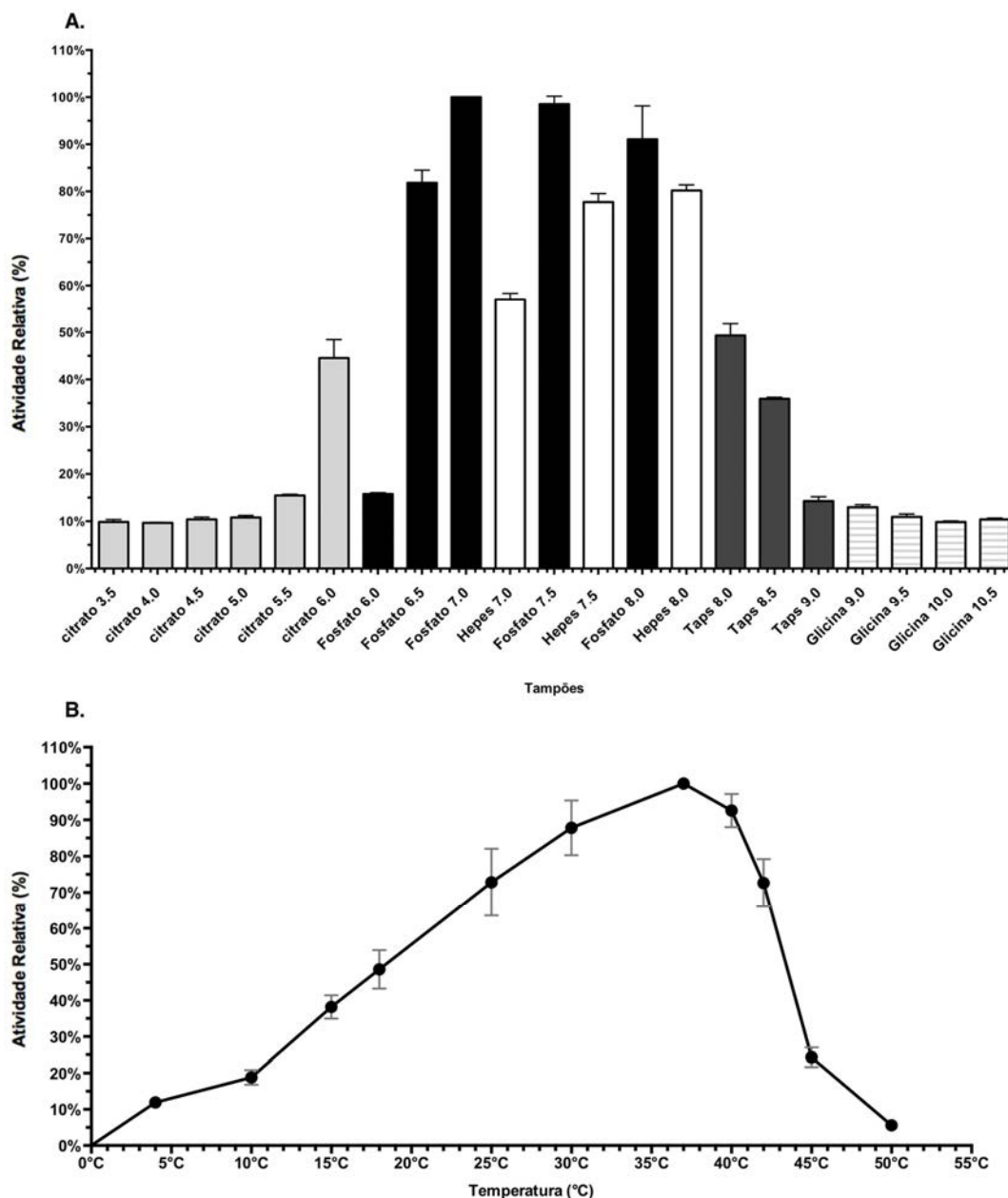


Figura 47. Influência do pH e da temperatura sobre a atividade enzimática da  $\beta$ -glicosidase Met-Bg. A enzima purificada foi avaliada para degradação do substrato pNP- $\beta$ -D-glicopiranosídeo. **A**, ensaio do pH: as reações foram realizadas à 37 °C, e os reagentes assinalados como Citrato e fosfato são sais de sódio. **B**, ensaio da temperatura: As reações foram realizadas empregando tampão fosfato de sódio pH 7.

Considerando estes valores de pHs do Brenda, podemos dizer que a enzima  $\beta$ -glicosidase apresentou pH ótimo em uma faixa comum. Entretanto, quando analisamos o “pH range”, ou seja, a amplitude da faixa de pH que uma mesma

enzima pode atuar, a maioria das enzimas bacterianas depositadas no Brenda atuam em uma faixa estreita de pH de 6-7. Conforme pode ser observado no gráfico (Fig.47.A), a enzima Met-Bg apresentou atividade acima de 80% para uma amplitude de pHs acima da média (pH 6,5-8,0).

Já em relação ao efeito da temperatura (Fig. 47.B), os dados do gráfico permitem assumir que a temperatura ótima da enzima é de 37 °C. Considerando um percentual mínimo de corte para uma atividade de 80%, a atividade enzimática foi expressiva para as temperaturas de 30-40 °C, a qual também corresponde a faixa encontra para a maior parte das enzimas  $\beta$ -glicosidase bacterianas depositadas na base de dados Brenda.

#### **D. Efeitos de Sais e outros reagentes sobre a atividade enzimática**

A atividade enzimática pode ser afetada por diversos íons, que podem atuar como cofatores ou inibidores, se ligando à enzima e alterando a forma ao sítio ativo, ou mesmo através de precipitados que afetem a disponibilidade e afinidade de ligantes à enzima. O balaço geral destes interferentes à ação enzimática (sobretudo íons metálicos) pode ser medido através da comparação da atividade relativa na presença do interferente em relação à reação sem aditivos (controle).

Buscando avaliar o efeito de alguns sais e seus respectivos cátions sobre a atividade da  $\beta$ -glicosidase metagenômica, foram avaliadas diferentes concentrações (0-5mM) de doze sais comumente conhecidos por sua interferência sobre a atividade enzimática (tratamentos avaliados em triplicata). O panorama geral obtido para ativadores/inibidores sobre a atividade enzimática pode ser visualizado no gráfico abaixo (Fig. 48).

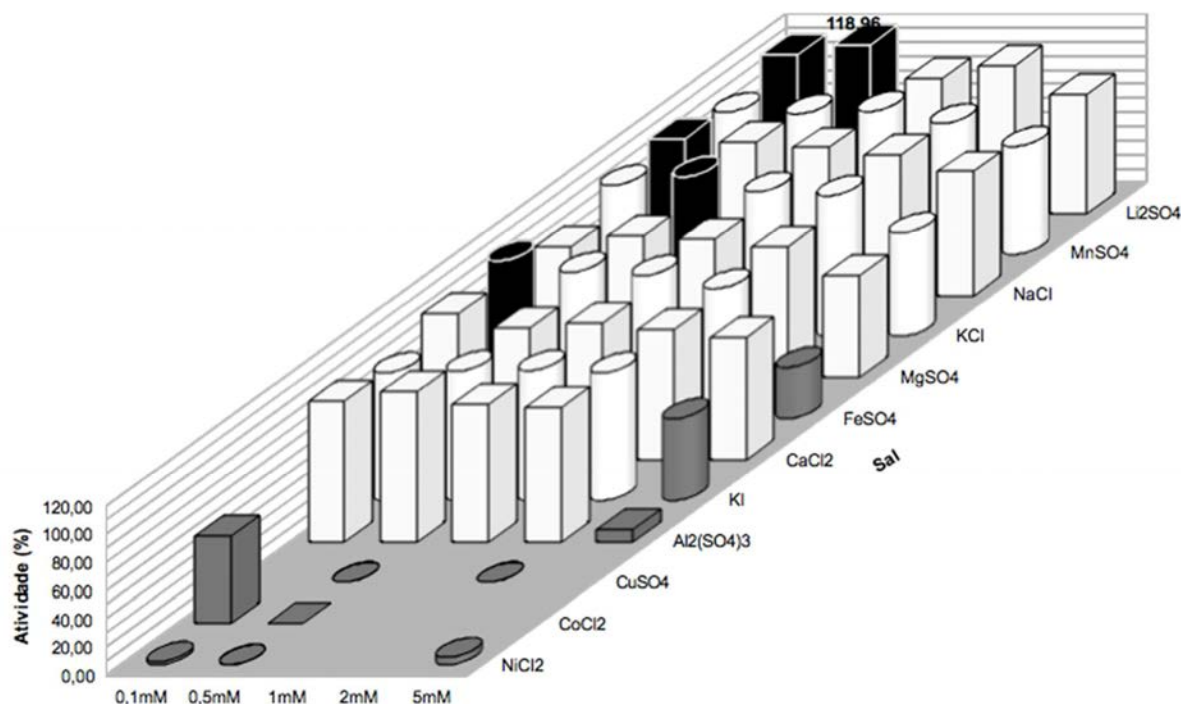


Figura 48. Efeitos de sais sobre a atividade enzimática. Os valores foram calculados com base no tratamento controle (sem aditivos). Branco: valores de atividade relativa entre 70-110%; preto: efeito ativador superior à 10%; cinza: efeito inibidor acima de 30%. O uso de alternância entre os estilos “caixa” e “cilindro” tem como único propósito diferenciar tratamentos adjacentes. Tampão fosfato de sódio pH7 (20mM), substrato p-Nitro-fenil- $\beta$ -D-glicopiranosideo (1mM).

A atividade da enzima foi fortemente inibida pelos sais de cobalto, níquel e cobre. Em um menor nível, a inibição também ocorreu para iodeto de potássio, e os sais de magnésio e alumínio, sobretudo em altas concentrações. O sulfato de ferro apresentou um efeito duplo, agindo como um bom ativador e potente inibidor, nas respectivas concentrações mínima e máxima.

Segundo RIOU et al. (1998) a modificação da atividade enzimática de  $\beta$ -glicosidasas por cátions como  $\text{Ag}^{2+}$ ,  $\text{Cu}^{2+}$ , e  $\text{Fe}^{3+}$ , pode indicar a presença de grupos tióis nos sítios catalíticos, pois estes grupos são alvos de metais pesados. Conforme pode ser observado na Figura 49, existe mesmo um grupo tiol (S, seta) no centro catalítico da enzima estudada, presente na cisteína em azul claro.

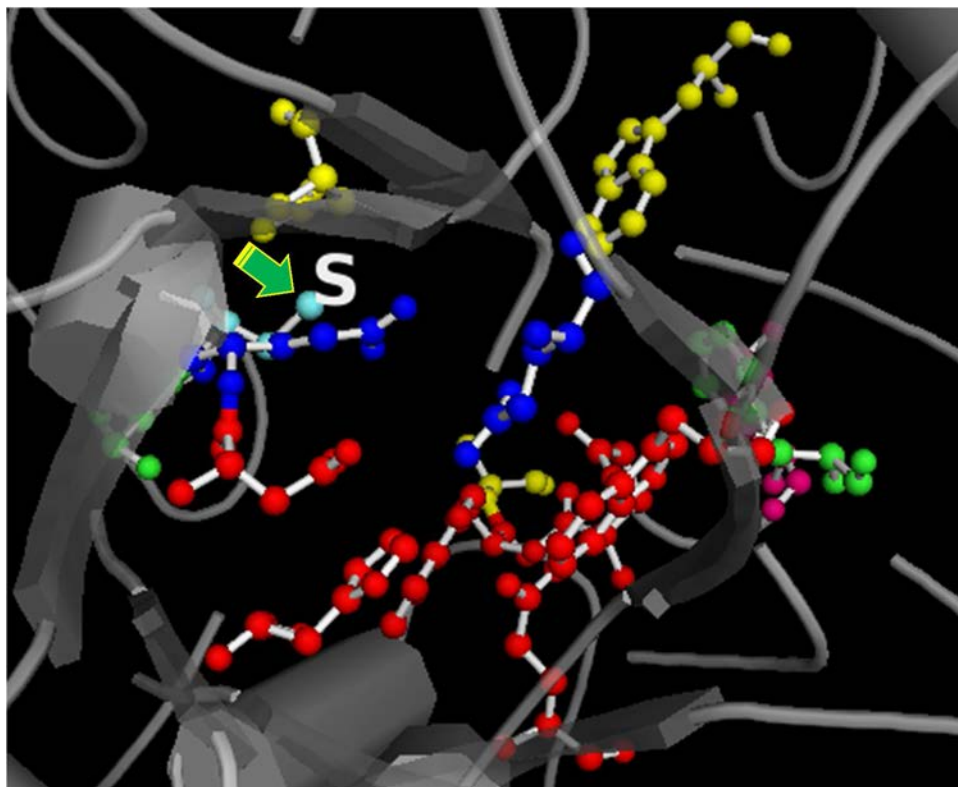


Figura 49. Presença de grupamento tiol (seta, aminoácido cisteína em azul claro) no sítio ativo da  $\beta$ -glicosidase metagenômica (modelo predito por homologia). As cores dos demais aminoácidos indicam a interação dos mesmos com o substrato segundo Marana (2006), sendo: Azul, centro catalítico; vermelho, subsítio -1; amarelo, subsítio +1; verde claro, subsítio +2 e lilás: subsítio +3.

O sulfato de lítio foi o sal que apresentou maior efeito ativador (aumento de quase 19%, para 0,5 mM), seguido do cloreto de potássio (aumento de aproximadamente 13,5%).

O uso de EDTA em baixa concentração (0,1 mM) foi praticamente capaz de restaurar a atividade catalítica nos sais de cobre e cobalto (Fig.50). Isso indica que o efeito inibitório observado nos ensaios se remete a interferência de cátions ( $\text{Cu}^{2+}$  e  $\text{Co}^{2+}$ ) e não aos ânions ( $\text{SO}_4^{2-}$  ou  $\text{Cl}^{-}$ ).

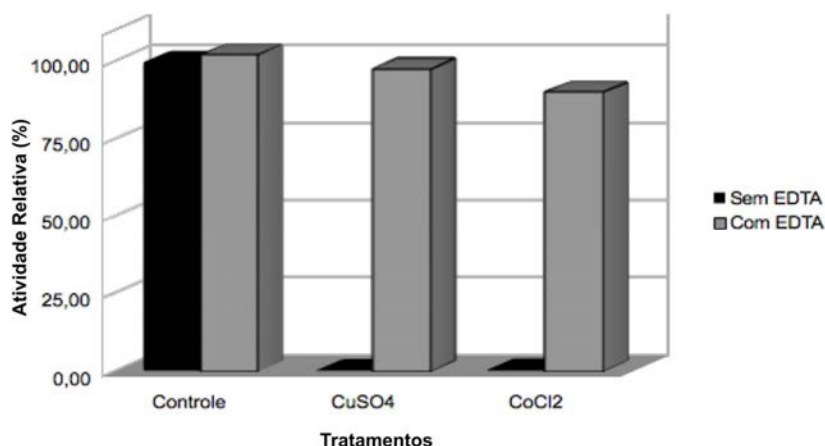


Figura 50. Efeitos do EDTA (0,1 mM) sobre sais inibidores (0,5 mM) da atividade da  $\beta$ -glicosidase Met-Bg. Tampão fosfato de sódio pH7 (20mM), substrato p-Nitro-fenil- $\beta$ -D-glicopiranosideo (1mM).

Em relação aos controles com e sem EDTA (Fig.50), a ausência de variação indica que a enzima aparentemente não é dependente de um cofator catiônico, que possivelmente estivesse sendo suprido da reação através do tampão (fosfato de sódio), ou elementos traços nos reagentes utilizados. A capacidade de atuar na presença do EDTA pode ser promissora para explorar o uso da enzima em condições industriais em que a tolerância a íons inibitórios sejam fatores determinante para o sua utilização, bastando incluir o EDTA na formulação do tampão enzimático.

Uma das características importantes para a aplicação biotecnologia de uma enzima é a sua capacidade de tolerar a presença de diversos contaminantes inerentes do processo industrial. Dentre estes compostos, a influência de alcoóis é bastante relevante, uma vez que as  $\beta$ -glicosidase podem ser expostas a concentrações substanciais durante diversas aplicações comerciais.

A influência dos alcoóis etanol (0,5mM a 1M) e metanol (0,5mM-400M) sobre a atividade  $\beta$ -glicolítica foi avaliada (Fig. 51).



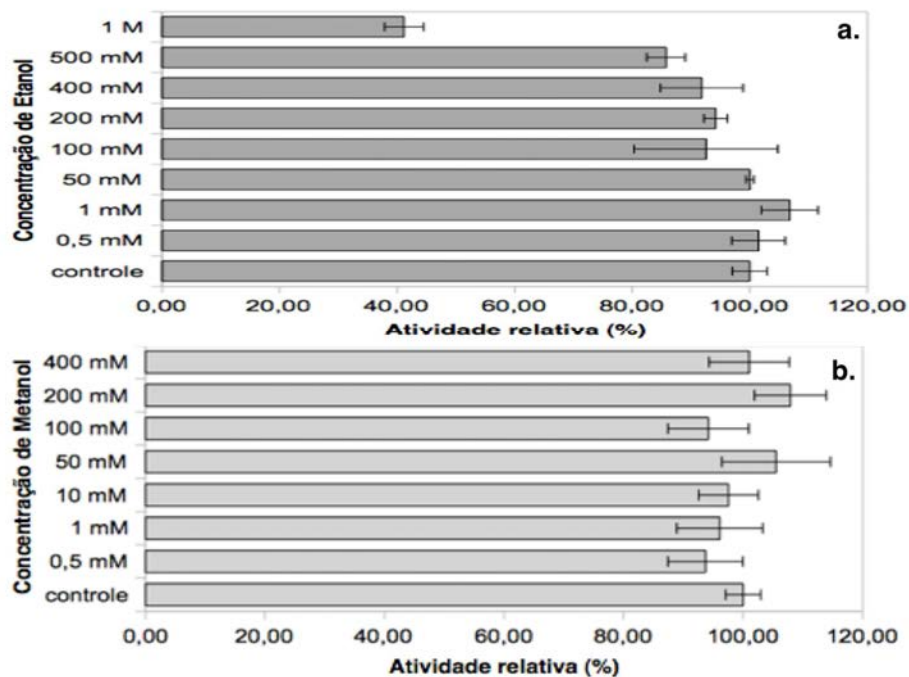


Figura 51. Atividade da enzima  $\beta$ -glicosidase metagenômica na presença de alcoóis de cadeia curta etanol (A) e Metanol (B).

Pequenas concentrações (0,5-50 mM) de etanol e (0,5-200mM) de metanol causaram pouca alteração na atividade relativa da enzima metagenômica Met-Bg, chegando à 106,8% à 1mM de etanol e 108% a 200mM de metanol. Na presença de 400mM de etanol, a enzima deteve 92% da atividade, enquanto à 1M a enzima obteve 41% de atividade residual. Já para o metanol, 400mM representou uma atividade de 101%.

Daroit e colaboradores (2008) observaram um efeito semelhante para baixas concentrações de etanol para uma  $\beta$ -glicosidase fúngica (*Monascus purpureus*): uma atividade relativa de 105-106% à baixas concentrações de etanol (5-50mM) e metanol (5-250mM). Entretanto, a atividade observada à 1M de etanol para esta enzima fúngica foi de 83,4%.

Sobre a influência de outros reagentes, o  $\beta$ -Mercaetanol não afetou a atividade enzimática (1 e 5 mM); enquanto o Dodecil Sulfato de Sódio foi completamente inibitório (0,1-2%).

### E. Determinação dos parâmetros cinéticos para a $\beta$ -glicosidase

Os parâmetros cinéticos correspondentes ao substrato pNP- $\beta$ G foram calculados para a enzima Met-Bg, empregando o software “Graph Pad Prism”. Para tanto, os valores de  $v_0$  para cada concentração de substrato (0,1-10 mM) foram obtidos através das respectivas leituras de reações de 0-60 min. A figura 52 apresenta os dados obtidos para a regressão não linear para cinética enzimática seguindo o modelo de Michaelis-Menten.

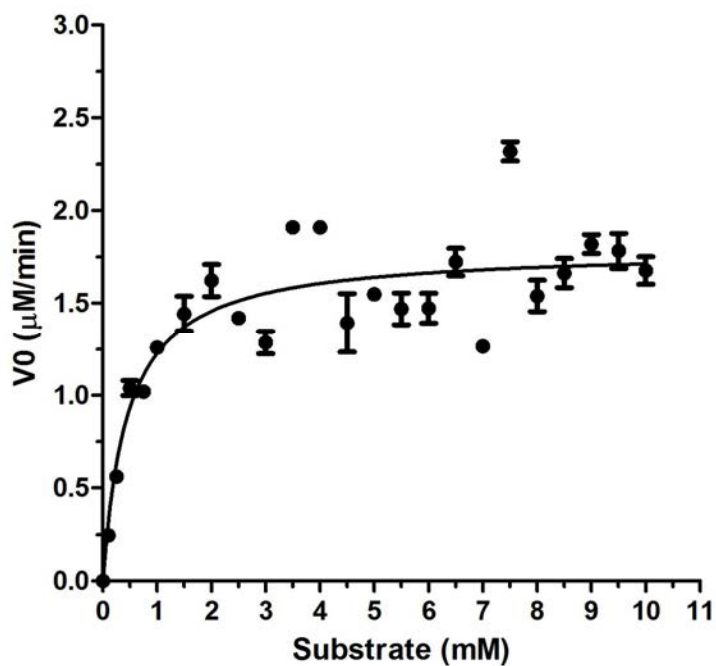


Figura 52. Efeitos da concentração de substrato sobre a velocidade inicial da  $\beta$ -glicosidase metagenômica. A atividade enzimática foi mensurada na presença das concentrações correspondentes de pNP-BG em pH 7.0 em 0-50 min de reação. A. Regressão não linear da velocidade inicial em relação a diferentes concentrações de substrato (0,1-10 mM) foi obtida através do software “Graph Pad Prism”.

Considerando os valores obtidos para os parâmetros cinéticos correspondente a atividade sobre o substrato p-Nitro-fenil- $\beta$ -D-glicopiranosideo segundo o software “Graph Pad Prism”, a velocidade máxima ( $V_{max}$ ) obtida foi de  $1,8 \pm 0,05 \mu$ M/min; enquanto o valor da concentração de substrato necessária para

que a velocidade enzimática atinja metade de seu valor máximo ( $K_m$ ) foi de  $0,46 \pm 0,08$  mM.

#### **4.4. Preparo de materiais para a expressão do cluster de PKSII**

Um dos maiores desafios da expressão heteróloga de “clusters” de PKS é a manipulação e clonagem eficiente de grandes fragmentos de DNA. Ainda que através de métodos clássicos de engenharia (tendo por base amplificação por PCR, ensaios de restrição e ligação) seja possível manipular sequências de 15-20 Kb de extensão, essa prática se mostra pouco eficiente e tediosamente demorada para fragmentos maiores e para realizar a inserção destes dentro dos cromossomos (FU *et al.*, 2008; RIVERO-MULLER, LAJIC & HUHTANIEMI, 2007; STEVES *et al.*, 2010).

Para os casos de expressão homóloga, envolvendo a substituição de vias nativas por outras engenheiradas com adição ou silenciamento (“Knockout”) de um determinado gene, ou troca de regiões bastante conservadas entre organismos próximos, uma alternativa muito interessante é o uso de Recombinação Cruzada (“Crossing-over”). Essa recombinação ocorre graças a regiões comuns entre o cromossomo hospedeiro e o elemento recombinante, o que permite o reconhecimento e permuta de fragmentos por meio de eventos tipo “crossing-over”, de forma que a inserção de um material exógeno se dá sítio específico e não limitada pelo tamanho do fragmento de DNA. Contudo, nem sempre se tem uma região bastante similar no organismo hospedeiro para guiar a inserção deste gene. Em muitas situações, a característica desejável em um bom hospedeiro é justamente não possuir (de forma natural ou artificial) certas vias, facilitando a detecção de mutantes bem sucedidos pela mudança de fenótipo. Nesses casos, mesmo se tratando de permuta entre homólogos, muitas vezes é necessário inserir uma região de homologia para que o fragmento seja integrado ao genoma na posição desejada.

Neste contexto, alguns sistemas baseados em mecanismos tipo RED/ET de integração de Fago- $\lambda$  aliados às grandes construções de engenharia molecular foram elegantemente delineados para tutorar a transposição de grandes insertos (RIVERO-MULLER, LAJIC & HUHTANIEMI, 2007; STEVES *et al.*, 2010). A inserção

sítio específica em regiões de reconhecimento do fago, normalmente conservadas nos genomas de diversas linhagens de *Streptomyces sp.* (como os sítios Attb) é auxiliada pelo uso vetores de integração/conjugação, além de linhagens bacterianas não metilantes que torna o material genético exógeno imune à maquinaria de restrição de DNA linear presente em muitos hospedeiros de *Streptomyces* (Gomez-Escribano, 2012).

Para realizar a expressão homóloga do agrupamento biosintético da PKSII optou-se por utilizar a metodologia de transferência sitio dirigida empregando o sistema RED/ET. Para tanto, o inserto metagenômico foi transferido para o vetor de integração pSET152 empregando os sítios de restrição para EcoRI que o flaqueavam no vetor cosmidial (Fig. 53-54), selecionado em bactéria *E. coli* DH10B e passada para a bactéria não metilante *E. coli* ET12567 pUZ8002.

A suspensão de esporos do hospedeiro final (*S. coelicolor* M1152, fig. 55) foi preparada para a conjugação posterior mediada pela bactéria *E. coli* ET12567 pUZ8002.

Figure 2. pWEB-TNC Cosmid Vector.

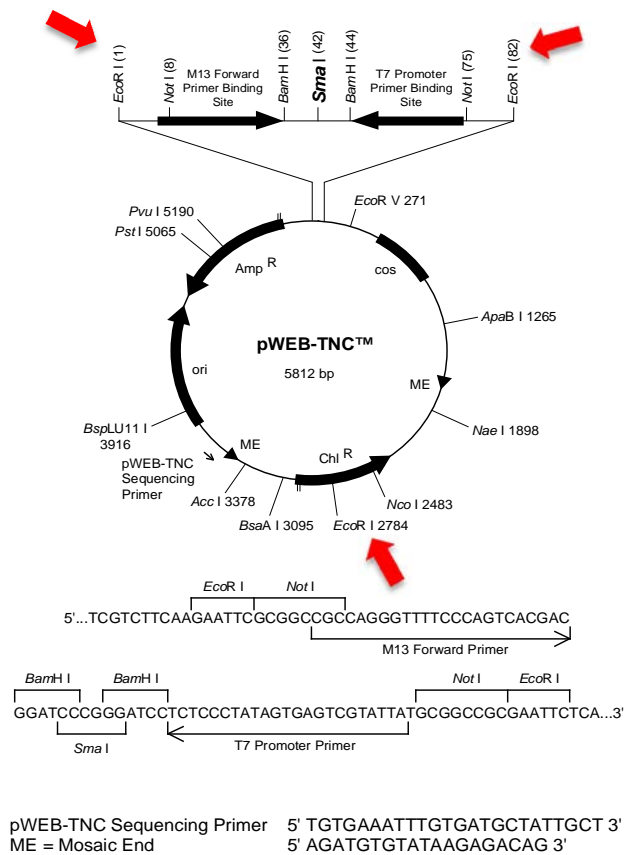


Figura 53. Cosmidio pWEB-TNC, que abriga o inserto metagenômico para o clone B5pI37.

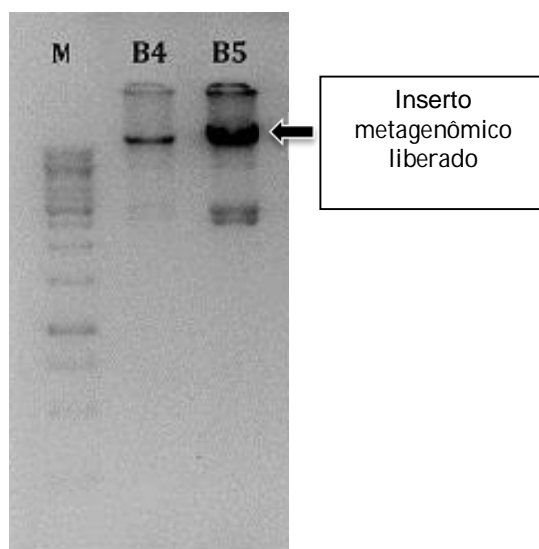


Figura 54. Restrição com enzima EcoRI de dois clones da Biblioteca metagenômica S.E. O clone B5 é o alvo deste trabalho.

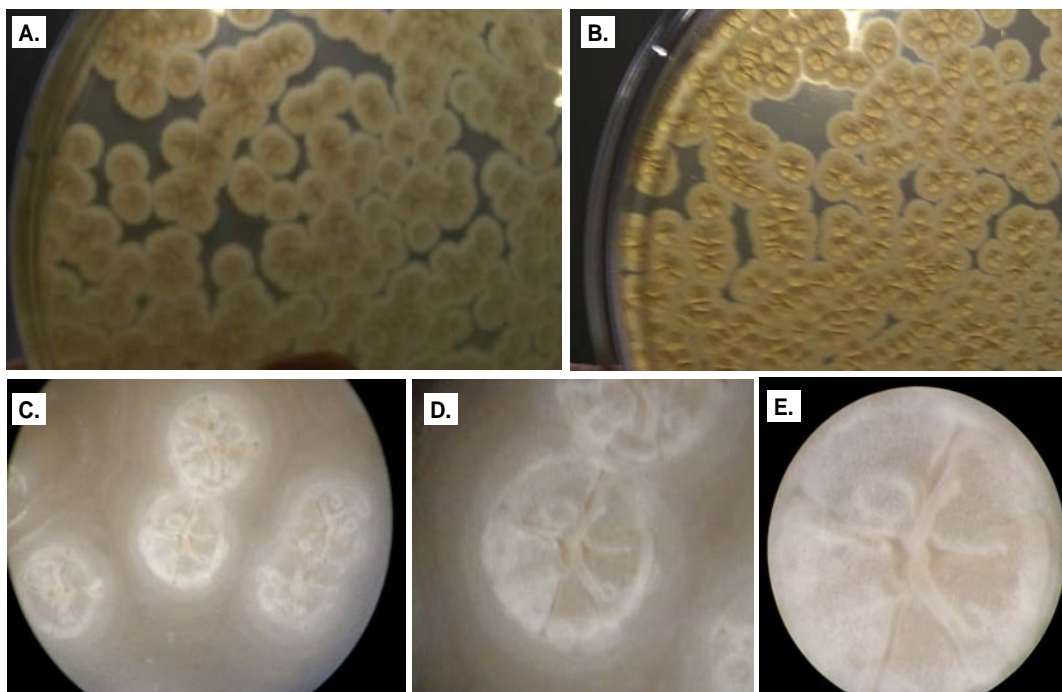


Figura 55. Morfologia da colônia do hospedeiro de expressão *Streptomyces coelicolor* M1152. A: vista das colônias com fonte de iluminação superior. B: vista das colônias com fonte de iluminação inferior. C-E: Vista das colônias em estereomicroscópio nos respectivos aumentos: 12,5x; 25x e 50x.

## 5. CONCLUSÕES E PERSPECTIVAS:

As estratégias utilizadas neste trabalho se mostraram bastantes úteis para o entendimento funcional das enzimas metagenômicas. As análises *in silico* permitiram identificar sítios catalíticos e regiões conservadas que forneceram pistas importantes sobre a classificação destas enzimas.

Dentre as três carboidrases estudadas, as análises *in vitro* foram mais bem sucedidas para a enzima  $\beta$ -glicosidade, enquanto que as análises *in silico* permitiram identificar as possíveis causas das dificuldades enfrentadas para as outras duas: um peptídeo sinal na sequência da celulase e um elevado conteúdo de bases nitrogenadas G/C flanqueando os genes da lacase.

A enzima  $\beta$ -glicosidade foi capaz de atuar em uma ampla faixa de pHs, o que lhe confere versatilidade para aplicação em diferentes processos. A temperatura ótima de 37 °C pode oferecer um importante fator para aplicação tecnológica, uma vez que não demanda um sistema muito dispendioso de aquecimento que possa

aumentar os custos do processo. A capacidade de resistir a diversos íons pode conferir vantagens no processamento de resíduos industriais, ao passo que a sensibilidade a outros íons (ex. cobre, cobalto) pode potencialmente ser burlada através da adição de pequenas concentrações de EDTA ao meio reacional (não afeta a enzima).

Apesar de ainda se encontrar em fase preliminar, tudo indica que a clonagem dos genes metagenômicos em *Streptomyces coelicolor* M1152 será bem sucedida, uma vez que todos os passos preparatórios para preparo clonagem em vetor integrativo em bactéria não metilante de DNA conjugante inter-gênerica foram bem sucedidos.

Uma vez inserido o cluster neste hospedeiro, a expressão dos genes de PKS podem ser evidenciados por estudos de HPLC, halos de inibição e mudança morfológica (principalmente cor). Estudos de deleção gradual de genes poderão ainda auxiliar a identificação dos genes diretamente relacionados à via de PKSII.

## 6. REFERÊNCIAS

ANTONOPOULOS, V. T.; HERNANDEZ, M.; ARIAS, M.E.; MAVRAKOS, E. ; BALL, A.S. The use of extracellular enzymes from *Streptomyces albus* ATCC 3005 for the bleaching of eucalyptus kraft pulp. **Appl Microbiol Biotechnol**, v. 57, p. 92–97, 2001

Askwith C, Eide D, Van Ho A, Bernard PS, Li L, Davis-Kaplan S, Sipe DM, Kaplan J (January 1994). "The FET3 gene of *S. cerevisiae* encodes a multicopper oxidase required for ferrous iron uptake". **Cell** 76 (2): 403–10. doi:10.1016/0092-8674(94)90346-8. PMID 8293473.

Baltz H (2006) Molecular engineering approaches to peptide, polyketide and other antibiotics. **Nat Biotech** 24:1533-1540.

Bento I, Martins LO, Gato Lopes G, Armã©nia Carrondo M, Lindley PF (November 2005). "Dioxygen reduction by multi-copper oxidases; a structural perspective". *Dalton Transactions* (21): 3507–13. doi:10.1039/b504806k. PMID 16234932.

BRACHMANN, A. O. *et al.* A Type II Polyketide Synthase is Responsible for Anthraquinone Biosynthesis in *Photobacterium luminescens*. **ChemBioChem**, Weinheim, v.8, 1721-1728, 2007.

Bradford, M.M. (1976) A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. **Analytical Biochemistry**, v.72, p.248-254.

Brady SF, Chao CJ, Handelsman J, Clard J (2001) Cloning and Heterologous Expression of a Natural Product Biosynthetic Gene Cluster from eDNA. **Org Lett** 3:1981-1984.

Brulc J. M., D. a Antonopoulos, M. E. B. Miller, M. K. Wilson, A. C. Yannarell, E. a Dinsdale, R. E. Edwards, E. D. Frank, J. B. Emerson, P. Wacklin, P. M. Coutinho, B. Henrissat, K. E. Nelson, and B. a White, "Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases.," **Proceedings of the National Academy of Sciences of the United States of America**, vol. 106, no. 6, pp. 1948–53, Feb. 2009.

CARDOSO, E.J.B.N.; FREITAS, S.S. A rizosfera. In: CARDOSO, E.J.B.N.; TSAI, S.M.; NEVES, M.C.P. (Ed.). *Microbiologia do solo*. Campinas: **Sociedade Brasileira de Ciênci do Solo**, 1992. p.41-58.

CASTRO, A.P.; FERNANDES GDA, R.; FRANCO, O.L. (2014) Insights into novel antimicrobial compounds and antibiotic resistance genes from soil metagenomes **Front Microbiol**. v. 5: 489. doi: 10.3389/fmicb.2014.00489. eCollection.

COSTA, S.G. (2013) PURIFICAÇÃO E CARACTERIZAÇÃO BIOQUÍMICA DE B GLICOSIDASES DE *Penicillium Chrysogenum* E SUA APLICAÇÃO NA SUPLEMENTAÇÃO DE COQUETÉIS ENZIMÁTICOS PARA A HIDRÓLISE DE BIOMASSAS. Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Agrícola, para obtenção do título de Magister Scientiae.

DA SILVEIRA, N. J. F. *et al.* Molecular models of protein targets from *Mycobacterium*



*tuberculosis*. **J Mol Model**, v.11, 160-166, 2005.

Dai X., Y. Zhu, Y. Luo, L. Song, D. Liu, L. Liu, F. Chen, M. Wang, J. Li, X. Zeng, Z. Dong, S. Hu, L. Li, J. Xu, L. Huang, and X. Dong, "Metagenomic insights into the fibrolytic microbiome in yak rumen.," **PLoS one**, vol. 7, no. 7, p. e40430, Jan. 2012.

DHARMARAJ, S.; ASHOKKUMAR, B. & DHEVENDARAN, K. Food-grade pigments from *Streptomyces* sp. isolated from the marine sponge *Callyspongia diffusa*. **Food Research International**, v. 42, p. 487–492, 2009.

Duan C., L. Xian, G. Zhao, Y. Feng, H. Pang, X. Bai, J. Tang, Q. Ma, and J. Feng, "Isolation and partial characterization of novel genes encoding acidic cellulases from metagenomes of buffalo rumens," vol. 107, pp. 245–256, 2009.

Enguita F.J., Martins LO, Henriques AO, Carrondo MA (2003) Crystal structure of a bacterial endospore coat component - A laccase with enhanced thermostability properties. **J Biol Chem** 278: 19416–19425.

**FAO**, The state of food and agriculture. BIOFUELS: prospects, risks and opportunities. Roma: Food and Agriculture Organisation of the United Nations., 2008, p. xi+128.

Feng Y., ; C.-J. Duan, H. Pang, X.-C. Mo, C.-F. Wu, Y. Yu, Y.-L. Hu, J. Wei, J.-L. Tang, and J.-X. Feng, "Cloning and identification of novel cellulase genes from uncultured microorganisms in rabbit cecum and characterization of the expressed cellulases.," **Applied microbiology and biotechnology**, vol. 75, no. 2, pp. 319–28, May 2007.

Feng Z, Kallifidas D, Brady SF (2011) Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. **Proc Natl Acad Sci USA** 108:12629-12634.

Feng Z, Kim JH, Brady SF (2010) Fluostatins produced by the heterologous expression of a TAR reassembled environmental DNA derived type II PKS gene cluster. **J Am Chem Soc** 132:11902-11903.

Fernandez-Moreno MA, Caballero JL, Hopwood DA, Malpartida F (1991) The act cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces*. *Cell* 66:769-780.

FLETT, F.; MERSINIAS, V.; SMITH, C.P. 1997. High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA- restricting streptomycetes. **FEMS Microbiology Letters**. 155:223-229.

FU, J. *et al.* Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. **Nucleic Acids Research**, v. 36, n. 17, p. (e113) 1-14, 2008.

FU, J. *et al.* Functional and structural studies of a novel cold-adapted esterase from an Arctic intertidal metagenomic library. **Applied microbiology and biotechnology**, v. 97, n. 9, p. 3965–78, maio. 2013.

Giardina P, Faraco V, Pezzella C, Piscitelli A, Vanhulle S, *et al.* (2010) Laccases: a never-ending story. **Cell Mol Life Sci** 67: 369–385.

**GOMES, E. S.** ANÁLISE *in silico* DE UM NOVO “CLUSTER” DE PKS II METAGENÔMICO. Dissertação de Mestrado apresentado ao programa de Microbiologia Agropecuária – Faculdade de Ciências Agrária e Veterinárias, Universidade Estadual de São Paulo, Jaboticabal, 2011.

**GOMES, E. S.** Utilização da metagenômica como ferramenta para prospecção de novos genes do sistema PKS tipo II. Trabalho apresentado à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Campus de Jaboticabal, para obtenção do título de Bacharel em Ciências Biológicas, 2008.

Gomes, E.S.; Schuch, V.; Lemos, E.G.M (2013) Biotechnology of polyketides: New breath of life for the novel antibiotic genetic pathways discovery through metagenomics. **Brazilian Journal of Microbiology**, v.44 (4), p.1007-1034.

GRAZIANI, S., Vasnier, C. e Daboussi, M. Novel Polyketide Synthase from *Nectria haematococca*. **Applied And Environmental Microbiology**, Washington, v.70, 2984–2988, 2004.

GUSTAFSSON, C., GOVINDARAJAN, S. & MINSHULL, J. Codon bias and heterologous protein expression. **TRENDS in Biotechnology**, Amsterdam, v.22, 2004.

Guyen N. H. N., L. M. Aruset, T. U. Engwetwanit, W. M. Huantong, P. H. Ampicharnchai, V. C. Hampreda, S. T. Anapongpipat, K. J. Irajaroenrat, S. K. R. Akshit, and L. E. Urwilaichitr, “Identification and Characterization of a Cellulase-Encoding Gene from the Buffalo Rumen Metagenomic Library,” vol. 76, no. 6, pp. 1075–1084, 2012.

Hess M., A. Sczyrba, R. Egan, T.-W. Kim, H. Chokhawala, G. Schroth, S. Luo, D. S. Clark, F. Chen, T. Zhang, R. I. Mackie, L. a Pennacchio, S. G. Tringe, A. Visel, T. Woyke, Z. Wang, and E. M. Rubin, “Metagenomic discovery of biomass-degrading genes and genomes from cow rumen.,” **Science** (New York, N.Y.), vol. 331, no. 6016, pp. 463–7, Jan. 2011.

HOPWOOD, D. A. *et al.* Production of ‘hybrid’ antibiotics by genetic engineering. **Nature**, London, v.314, 642-644, 1985.

Howarth R. W. and S. Bringezu, “Biofuels and environmental impacts. Scientific analysis and implications for sustainability,” vol. 9, 2009.

HSIAO, N. e KIRBY, R. Comparative genomics of *Streptomyces avermitilis*, *Streptomyces cattleya*, *Streptomyces maritimus* and *Kitasatospora aureofaciens* using a *Streptomyces coelicolor* microarray system. **Antonie Van Leeuwenhoek**, v. 93, 1–25, 2008.

**IBGE/LSPA (Estado)**, Levantamento sistemático da produção agrícola. Jan2013. Instituto Brasileiro de Geografia e Estatística (IBGE), 2013, p. 83.

Iqbal HA, Feng Z, Brady SF (2012) Biocatalysts and small molecule products from metagenomic studies. **Curr Opin Chem Biol** 16:109–116.

JIANG, C.; MA, G.; LI, S.; HU, T.; CHE, Z.; SHEN, P.; YAN, B. & WU, B. Characterization of a Novel  $\beta$ -glucosidase-Like Activity from a Soil Metagenome. **The Journal of Microbiology**, v. 47, p. 542-548, 2009.

King RW, Bauer JD, Brady, SF (2009) An environmental DNA-derived type II polyketide biosynthetic pathway encodes the biosynthesis of the pentacyclic polyketide erdacin. *Angew*

Chem 48:6257–6261.

LING, L.L.; SCHNEIDER, T.; PEOPLES, A.J.; SPOERING, A.L.; ENGELS, I.; CONLON, B.P.; MUELLER, A.; SCHÄBERLE, T.F.; HUGHES, D.E.; EPSTEIN, S.; JONES, M.; LAZARIDES, L.; STEADMAN, V.A.; COHEN, D.R.; FELIX, C.R.; FETTERMAN, K.A.; MILLETT, W.P.; NITTI, A.G.; ZULLO, A.M.; CHEN, C.; LEWIS, K. (2015) A new antibiotic kills pathogens without detectable resistance. *Nature*. [Epub ahead of print]. doi: 10.1038/nature14098.

LU, L.; CUT, H.-L.; CHEN, Y.-N.; YUANLU, S. Isolation and Identification of *Streptomyces* sp. and Assay of Its Exocellular Water-Soluble Blue Pigments, **Folia Microbiol.**, v.47, p. 493-498, 2002

M. A. F. D. Moraes, "Perspective: Lessons from Brazil," *Nature*, vol. 474, p. S25, 2011.

Mariotto, 2006. Cinética Enzimática. Disponível em:

[http://www.eng.ufsc.br/labs/probio/disc\\_eng\\_bioq/apostilas/Apostila\\_cinetica\\_enzimatica\\_ju.pdf](http://www.eng.ufsc.br/labs/probio/disc_eng_bioq/apostilas/Apostila_cinetica_enzimatica_ju.pdf), data do acesso: janeiro de 2014.

Messerschmidt A, Huber R (January 1990). "The blue oxidases, ascorbate oxidase, laccase and ceruloplasmin. Modelling and structural relationships". *Eur. J. Biochem.* 187 (2): 341–52. doi:10.1111/j.1432-1033.1990.tb15311.x. PMID 2404764.

METSÄ-KETELÄ, M. ; SALO, V.; HALO, L.; HAUTALA, A.; HAKALA, J.; MÄNTSÄLÄ, P.; YLIHONKO, K. An efficient approach for screening minimal PKS genes from *Streptomyces*. **FEMS Microbiology Letters**, Amsterdam, v.180, 1-6, 1999.

Moraes M. Ab. F. D., "Perspective: Lessons from Brazil," *Nature*, vol. 474, p. S25, 2011.

MOREIRA, F. S., SIQUEIRA, J. O. *Microbiologia e Bioquímica do Solo*. Editora UFLA, Lavras, MG, Brasil, 2006.

Nakamura K, Kawabata T, Yura K, Go N (October 2003). "Novel types of two-domain multi-copper oxidases: possible missing links in the evolution". *FEBS Lett.* 553 (3): 239–44. doi:10.1016/S0014-5793(03)01000-7. PMID 14572631.

NASCIMENTO, C.V.; SOUZA, F.H.M.; MASUI, D.C.; LEONE, F.A.; PERALTA, R.M.; JORGE, J.A.; FURRIEL, A.P.M. Purification and Biochemical Properties of a Glucose-Stimulated  $\beta$ -D-Glucosidase Produced by *Humicolagrisea* var. *thermoidea* Grown on Sugarcane Bagasse. **The Journal of Microbiology**, v. 48 (1), p. 53-62, 2010.

Nigam P. S. and A. Singh, "Production of liquid biofuels from renewable resources," *Progress in Energy and Combustion Science*, vol. 37, no. 1, pp. 52–68, Feb. 2011.

NIKODINOVIC, J. & PRIESTLEY, N. D. 2006. A second generation snp-derived *Escherichia coli*-*Streptomyces* shuttle expression vector that is generally transferable by conjugation. **ScienceDirect**, v. 56, p. 223–227, 2006.

Novaenzyme, 2012: <http://www.nnfcc.co.uk/news/novozymes-and-petrobras-partner-to-produce-advanced-biofuels-from-sugarcane-bagasse> acesso em 20/11/12.

ONU, 2011. Seven Billion and growing: the role of population policy in achieving sustainability: Disponível em

[http://www.un.org/esa/population/publications/technicalpapers/TP20113\\_SevenBillionandGrowing.pdf](http://www.un.org/esa/population/publications/technicalpapers/TP20113_SevenBillionandGrowing.pdf), data do acesso em 20/11/12.

Ouzounis C, Sander C (February 1991). "A structure-derived sequence pattern for the detection of type I copper binding domains in distantly related proteins". **FEBS Lett.** 279 (1): 73–8. doi :10.1016/0014-5793(91)80254-Z. PMID 1995346.

PAIXÃO, D.A.A.; DIMITROV, M.R.; PEREIRA, R.M. ; ACORSINI, F. ; VIDOTTI, M.B. ; Lemos E.G.M. (2010). Molecular analysis of the bacterial diversity in a specialized consortium for diesel oil degradation. **Revista Brasileira de Ciência do Solo** (Impresso), v. 34, p. 773-781.

PIEIRU, S. *et al.* 2005 Production of the Potent Antibacterial Polyketide Erythromycin C in *Escherichia coli* . **Applied And Environmental Microbiology**, v. 71, p. 2539–2547

Pope P. B., Denman S. E., M. Jones, S. G. Tringe, K. Barry, S. a Malfatti, a C. McHardy, J.-F. Cheng, P. Hugenholtz, C. S. McSweeney, and M. Morrison, "Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores.," **Proceedings of the National Academy of Sciences of the United States of America**, vol. 107, no. 33, pp. 14793–8, Aug. 2010.

Prosdocimi F, Cerqueira GC, Binneck E, Silva AF, Reis AN, Junqueira ACM, Santos ACF, Nhani-Júnior A, Wust Cl, Camargo-Filho F, Kessedjian JL, Petretski JH, Camargo LP, Ferreira RGM, Lima RP, Pereira RM, Jardim S, Sampaio VS, Folgueras-Flatschart AV (2001) **BIOINFORMÁTICA: Manual do usuário** n29, Biotecnologia Ciência & Desenvolvimento, Brasília, Brazil.

Quintanar L, Stoj C, Taylor AB, Hart PJ, Kosman DJ, et al. (2007) Shall we dance? How a multicopper oxidase chooses its electron transfer partner. **Accounts Chem Res** 40: 445–452. 3.

R. W. Howarth and S. Bringezu, "**Biofuels and environmental impacts**. Scientific analysis and implications for sustainability," vol. 9, 2009.

Reiss R, Ihssen J, Richter M, Eichhorn E, Schilling B, et al. (2013) Laccase versus Laccase-Like Multi-Copper Oxidase: A Comparative Study of Similar Enzymes with Diverse Substrate Spectra. **PLoS ONE** 8(6): e65633. doi:10.1371/journal.pone.0065633.

Rivero-Muller, A.; Lajic, S. & Huhtaniemi, I. Assisted large fragment insertion by Red/ET-recombination (ALFIRE) — an alternative and enhanced method for large fragment recombineering. **Nucleic Acids Research**, v. 35, p. e.78 (1-8), 2007.

RIX, U. *et al.* Modification of post-PKS tailoring steps through combinatorial biosynthesis. **Nat. Prod. Rep.**, v.19, 542-580, 2002.

Roberts SA, Weichsel A, Grass G, Thakali K, Hazzard JT, Tollin G, Rensing C, Montfort WR (March 2002). "Crystal structure and electron transfer kinetics of CueO, a multicopper oxidase required for copper homeostasis in *Escherichia coli*". **Proc. Natl. Acad. Sci. U.S.A.** 99 (5): 2766–71. doi:10.1073/pnas.052710499. PMC 122422.

ROMANO, N.; GIOFFRÉ, A.; SEDE, S.M.; CAMPOS, E.; CATALDI, A.; TALIA, P. (2013) Characterization of cellulolytic activities of environmental bacterial consortia from an

argentinian native forest. **Curr microbiol.** cap. 67(2). p. 138-47. Doi: 10.1007/s00284-013-0345-2.

SAMBROOK, J. e RUSSEL, D.W. **Molecular cloning: a laboratory manual**, 3rd ed. Cold Spring Harbor Laboratory Press, 2001.

SANTOS, C.R.; PAIVA, J.H.; SFORÇA, M.L.; NEVES, J.L.; NAVARRO, R.Z.; COTA, J.; AKAO, P.K.; HOFFMAM, Z.B.; MEZA, A.N.; SMETANA, J.H.; NOGUEIRA, M.L.; POLIKARPOV, I.; XAVIER-NETO, J.; SQUINA, F.M.; WARD, R.J.; RULLER, R.; ZERI, A.C.; MURAKAMI, M.T. Dissecting structure-function-stability relationships of a thermostable GH5-CBM3 cellulase from *Bacillus subtilis* 168. **Biochem J.** v. 441(1), p. 95-104, 2012.

SCHUCH, V. Prospecção de genes envolvidos na biossíntese de antibióticos através da abordagem metagenômica. Dissertação de Mestrado em Microbiologia. Faculdade de Ciências Agrárias e Veterinárias, UNESP. Jaboticabal, 2007.

SHEN, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms **Current Opinion in Chemical Biology**, v.7, 285–295, 2003.

SILVEIRA, E. L. *et al.* Bacterial diversity of soil under eucalyptus assessed by 16S rDNA sequencing analysis. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41,1507-1516, 2006.

SIM, M.; KIM, J. (2015). Metagenome assembly through clustering of next-generation sequencing data using protein sequences. **J Microbiol Methods.** 2015. pii: S0167-7012(15)00003-2. doi: 10.1016/j.mimet.2015.01.002. [Epub ahead of print]

Suzuki S, Kataoka K, Yamaguchi K (October 2000). "Metal coordination and mechanism of multicopper nitrite reductase". *Acc. Chem. Res.* 33 (10): 728–35. doi:10.1021/ar9900257. PMID 110Mann KG, Jenny RJ, Krishnaswamy S (1988). "Cofactor proteins in the assembly and expression of blood clotting enzyme complexes". **Annu. Rev. Biochem.** 57: 915–56. doi:10.1146/annurev.bi.57.070188.004411. PMID 3052293.

Takagi M, Shin-Ya K (2011) New species of actinomycetes do not always produce new compounds with high frequency. **J Antibiot** (Tokyo) 64:699–701.

TAKAGI, M.; SHIN-YA, K. New species of actinomycetes do not always produce new compounds with high frequency. **J. Antibiot.**, v. 64 (10), p. 699 – 701, 2011.

VALLMITJANA, M.; FERRER-NAVARRO, M.; PLANELL, R.; ABEL, M.; AUSÍN, C.; QUEROL, E.; PLANAS, A.; PÉREZ-PONS, J.A. Mechanism of the family 1 beta-glucosidase from *Streptomyces* sp: catalytic residues and kinetic studies. **Biochemistry**, v. 40 (20), p. 5975 – 5982, 2001.

VASUDEVAN P. T. AND M. BRIGGS, "BIODIESEL PRODUCTION-CURRENT STATE OF THE ART AND CHALLENGES.," **JOURNAL OF INDUSTRIAL MICROBIOLOGY & BIOTECHNOLOGY**, VOL. 35, NO. 5, PP. 421–30, MAY 2008.

WALTER, S.; & SCHREMPF, H. Studies of *Streptomyces reticuli* cel-1 (Cellulase) Gene Expression in *Streptomyces* Strains, *Escherichia coli*, and *Bacillus subtilis*. **Applied and Environmental Microbiology**, v. 61, p. 487–494, 1995.

Wang GY, Graziani E, Waters B, Pan W, Li X, McDermott J, Meurer G, Saxena G, Andersen RJ, Davies J (2000) Novel natural products from soil DNA libraries in a Streptomyces host. **Org Lett** 2:2401-2404.

WAWRIK, B. *et al.* Identification of Unique Type II Polyketide Synthase Genes in Soil. **Applied And Environmental Microbiology**, Washington, v.71, 2232-2238, 2005.

WEISSMAN, K. J. (coord.), Introduction to Polyketide Biosynthesis. **Methods Enzymol**, v.459, 2009.

WOOD, P.J.; ERFLE, J.D.; TEATHER, R.M. Use of complex formation between congo red and polysaccharide in detection and assay of polysaccharide hydrolases. **Methods in Enzymology**, v. 160, p.59-74, 1988.

WOOLEY, J.C., GODZIK A., FRIEDBERG, I. (2010). A primer on metagenomics. **PLoS Comput Biol**. v6 (2), p.e1000667. doi: 10.1371/journal.pcbi.1000667.

WULFF, N.A.; CARRER, H.; PASCHOLATI, S.F. (2003). Cloning and Expression of Cellulase Xf-818 of *Xylella fastidiosa* in *Escherichia coli*. **Scientia Agricola**, v.60 (4), p.715-721.

Xia Y., F. Ju, H. H. P. Fang, and T. Zhang, "Mining of novel thermo-stable cellulolytic genes from a thermophilic cellulose-degrading consortium by metagenomics.," **PloS one**, vol. 8, no. 1, p. e53779, Jan. 2013.