UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

BRUNO SILVA ANDRADE

DNA E RNA POLIMERASES DO PLASMÍDEO MITOCONDRIAL DE *MONILIOPHTHORA PERNICIOSA* (STAHEL) AIME & PHILLIPS-MORA: CARACTERIZAÇÃO, EXPRESSÃO, ESTUDO DE INIBIDORES E FILOGENIA MOLECULAR

Feira de Santana, BA
2011

BRUNO SILVA ANDRADE

DNA E RNA POLIMERASES DO PLASMÍDEO MITOCONDRIAL DE *MONILIOPHTHORA PERNICIOSA* (STAHEL) AIME & PHILLIPS-MORA: CARACTERIZAÇÃO, EXPRESSÃO, ESTUDO DE INIBIDORES E FILOGENIA MOLECULAR

Tese apresentada ao Programa de Pós-graduação em Biotecnologia, da Universidade Estadual de Feira de Santana como requisito parcial para obtenção do título de Doutor em Biotecnologia.

Orientador: Prof. Dr. Aristóteles Góes Neto

Feira de Santana, BA
2011

Dedico esta tese a minha mãe, Léa Maria S. Andrade.

AGRADECIMENTOS

Agradeço especialmente ao meu amigo e orientador, **Dr. Aristóteles Góes Neto**, pelo apoio e acompanhamento dado para realização deste trabalho e principalmente por me tornar uma pessoa mais responsável e independente cientificamente, deixando-me construir minhas próprias idéias.

À **Paloma Araújo Bahia**, pelo apoio emocional e afetivo, durante todos esses anos, aturando os meus momentos de "autismo" no período em estive escrevendo esta tese.

Ao meu amigo **Wagner Soares** (UESB), pelas opiniões e sugestões de trabalho, pela parceria em projetos e pelos bons momentos de gargalhadas que demos durante todos esses anos.

Agradeço a todos os professores e alunos do PPGBiotec-UEFS/Fiocruz, que conviveram comigo durante esse período, contribuindo de alguma maneira para a o meu trabalho. Ao secretário **Helton Carneiro**, pela amizade e competência em atender às solicitações diversas feitas durante a execução deste trabalho.

À minha amiga **Catiane Souza**, pelo apoio de sempre, pelas ótimas discussões de trabalho, pela ótima companhia em viagens para congressos e cursos, pelas boas risadas que damos para descontrair.

À minha amiga **Rafaela Galante**, pelos ótimos momentos de discussão e ótima companhia nas viagens.

A todos os integrantes do LAPEM/BIOMOL, especialmente **Rita Terezinha de O. Carneiro** e **Ana Carolina B. Gonçalves**, pela ótima convivência e discussões no laboratório.

Aos meus amigos do PPG - Genética e Biologia Molecular – UESC: **Cristiano Villela Dias**, pela amizade de sempre, pelo apoio e ajuda no trabalho nas minhas idas a UESC, pelas sugestões dadas para realização deste trabalho; **Thyago Hermylly Santana Cardoso** e **Dayane Santos Gomes**, pelo apoio dado à realização dos experimentos em Biologia Molecular.

“Se os fatos não se encaixam na teoria, modifique os fatos”

Albert Einstein

RESUMO

Neste trabalho foram estudados aspectos moleculares e evolutivos das DNA (DPO) e RNA (RPO) polimerases codificadas pelo plasmídeo mitocondrial de *M. perniciosa*. Modelos tridimensionais dessas polimerases foram construídos, utilizando Modelagem por Homologia, seguida de uma simulação por Dinâmica Molecular de 3500 picossegundos. Com as estruturas 3D das enzimas modeladas e validadas, foi realizado um processo de triagem virtual de compostos químicos, utilizando os bancos de dados KEGG, ZINC e PubChem, seguido do Docking Molecular (Autodock Vina) e Dinâmica Molecular MM/PBSA. Então, foram selecionados complexos mais estáveis DPO-Entecavir e RPO-Rifampicina. O estudo da expressão relativa dos genes codificadores das DNA e RNA polimerases foi feito através de RT-qPCR, utilizando-se cDNA diferentes fases de desenvolvimento do *M. perniciosa*. Foi na fase de primórdio em que houve um aumento significativo na atividade desses genes, o que provavelmente está relacionado com um processo de defesa do fungo contra o *T. cacao*. A análise filogenética foi realizada utilizando-se sequências de DNA e proteína das DPO e RPO do plasmídeo de *M. perniciosa* e de polimerases de outros 12 plasmídeos fúngicos, além de sequências de polimerases virais. Foram realizadas análises de Máxima Verossimilhança e Bayesiana, apenas de sequências fúngicas e da combinação entre fungos e vírus. As topologias das árvores filogenéticas resultantes dessas análises corroboram as hipóteses de Transferência Horizontal de Genes (THG) entre polimerases fúngicas e virais, além de delimitar as relações evolutivas das DPO e RPO do plasmídeo de *M. perniciosa*.

Palavras-chave: *Moniliophthora perniciosa*, Mitocôndria, Polimerases, RT-qPCR, Filogenia Molecular

ABSTRACT

In this study we evaluated molecular and evolutionary aspects of DNA (DPO) and RNA (RPO) polymerases encoded by *M. perniciosa* mitochondrial plasmid. Three-dimensional models of these polymerases were constructed using a Homology Modeling approach, followed by Molecular Dynamics simulations of 3500 picoseconds. With modeled and validated 3D structures, we performed a virtual screening process in order to search for chemical compounds, using KEGG, ZINC and PubChem databases, followed by Molecular Docking (AutoDock Vina) and Molecular Dynamics MM/PBSA. Thus, we selected DPO-Entecavir and RPO-Rifampicin as the most stable complexes. The study of the relative expression of DPO and RPO genes was done by RT-qPCR, using cDNA from different developmental stages of *M. perniciosa*. In the phase of primordium, a significant increase in the activity of these genes occured, which is probably related to a self-defense process of the fungus against *T. cacao*. Phylogenetic analyses were performed by using DNA and protein sequences of DPO and RPO from *M. perniciosa* and other 12 fungal polymerases, as well as viral polymerases sequences. Then, Maximum Likelihood and Bayesian analyses were performed by using only fungal polymerase sequences, and the combination of fungal and viral polymerase sequences. Tree topologies results corroborate the hypothesis of Horizontal Gene Transfer (HGT) between fungal and viral polymerases, and delineate the evolutionary relationship of DPO and RPO *M. perniciosa* mitochondrial plasmid.

Keywords: *Moniliophthora perniciosa*, Mitochondria, Polymerases, RT-qPCR, Molecular Phylogeny

SUMÁRIO

	Introdução Geral	10
	Referências	12
	Comparative Modeling of DNA and RNA Polymerases from Moniliophthora perniciosa Mitochondrial Plasmid (Andrade BS, Taranto AG, Góes-Neto A, Duarte AA. Theor Biol Med Model. 2009 Sep 10;6:22.)	14
1.1	Background	15
1.2	Methods	16
1.3	Results and discussion	17
1.4	Conclusions	19
	Acknowledgements	20
	References	20
	Virtual screening reveals a viral-like polymerase inhibitor that complexes with the M. perniciosa DNA polymerase (To be submitted)	31
2.1	Introduction	32
2.2	Methods	34
2.2.1	Ligand Searching	34
2.2.2	Docking Studies	34
2.2.3	Molecular Dynamics of Complex	35
2.3	Results and Discussion	36
2.3.1	Structures and binding energies of DPO complexes from AutoDock Vina	36
2.3.2	Molecular Dynamics MM/PBSA of DPO-Entecavir complex	40
2.4	Conclusions	41
	Acknowledgements	42
	References	42
	The RNA polymerase of Moniliophthora perniciosa mitochondrial plasmid forms a highly stable complex with Rifampicin, a bacterial RNA polymerase inhibitor (To be submitted)	46
3.1	Introduction	47
3.2	Methods	49
3.2.1	Ligand Searching	49
3.2.2	Docking Studies	50
3.2.3	Molecular Dynamics of Complex	51
3.3	Results and Discussion	51
3.3.1	Structures and binding energies of RPO complexes from AutoDock Vina	51
3.3.2	Molecular Dynamics MM/PBSA of RPO-Rifampicin complex	55
3.4	Conclusions	56
	Acknowledgements	57
	References	57
	The activity of DNA and RNA polymerases from Moniliophthora perniciosa mitochondrial plasmid may reveal a self-defense mechanism against oxidative stress (To be submitted)	62
4.1	Introduction	63

4.2	Methods	64
4.2.1	Moniliophthora perniciosa culture	64
4.2.2	RNA extraction and cDNA synthesis	65
4.2.3	Real-time qPCR and data analysis	66
4.3	Results and discussion	66
4.4	Conclusions	69
	Acknowledgements	70
	References	70
	Phylogenetic analysis of DNA and RNA polymerases from Moniliophthora perniciosa (Stahel) Aime & Phillips-Mora mitochondrial plasmid reveals a probable lateral gene transfer (To be submitted)	75
4.1	Introduction	76
4.2	Methods	78
4.3	Results	79
4.4	Discussion	80
4.5	Conclusions	84
	Acknowledgements	86
	References	86
	Conclusão Geral	95
	ANEXOS	97

Introdução Geral

Moniliophthora perniciosa (Stahel) Aime e Phillip-Mora (2005), anteriormente conhecido como *Crinipellis perniciosa* (Singer) Stahel, é um fungo basidiomiceto hemibiotrófico (Marasmiaceae, Agaricales) fungo causador da doença vassoura de bruxa (Witches' Broom Disease - WBD) do cacaueiro (*Theobroma cacao* L.) (Mondego et al, 2008). A WBD está amplamente distribuída nas Americas do Sul e Central, podendo causar perdas de safra de até 90% (Pires et al, 2008). No estado da Bahia, a região sudeste vem sendo severamente afetada pela presence da WBD desde o final da década de 1980 (Mondego et al., 2008).

Desde a conclusão do projeto genoma do *Moniliophthora perniciosa,* diversos trabalhos vêm sendo publicados descrevendo uma série de rotas metabólicas e mecanismos de defesa e interação desse patógeno com o *T. cacao*. Em 2008, Formighieri et al. identificaram a presença de um plasmídeo mitochondrial, linear típico, inserido covalentemente no genoma mitochondrial de *M. perniciosa*, codificando duas polimerases (DNA e RNA polimerases – DPO e RPO) em fases de leitura opostas. Apesar de ser um evento raro em fungos, a presença de um plasmídeo inserido no genoma mitocondrial geralmente interfere no metabolismo energético do hospedeiro, desencadeando processos de senescência ou aumento de sobrevida (Poggeler and Kempken, 2004). Então, é provável também que a ação de DNA e RNA polimerases codificadas por esse tipo de plasmídeo presente no genoma de mitocondrial de *M. perniciosa* esteja envolvida em algum desses processos.

A Bioinformática e a química computacional (ou quimioinformática) estão entre as áreas de maior crescimento dentro dos processos de descoberta de medicamentos (Bleicher et al., 2003). Essas ferramentas são utilizadas para triagem de moléculas *in silico* e tornaram-se componentes cruciais para a descoberta de muitas drogas. A seleção das melhores moléculas candidatas a interagir com seus respectivos alvos protéicos ocorre de acordo com um padrão de

potuações (hits) que obedecem a uma série de características para um encaixe perfeito (Kitchen et al. 2004). O uso do processo de Docking (encaixe) molecular envolve a predição e orientação de um provável inibidor dentro do sítio ativo da molécula-alvo (Kitchen et al. 2004). A triagem vitual de estruturas químicas teve vários êxitos importantes nos últimos anos (Carlson, 2002; Gohlke and Klebe, 2002; Lyne, 2002; Osterberg et al., 2002; Taylor et al., 2002; Zavodszky et al., 2002; Schapira et al., 2003; Cavasotto and Abagyan 2004; Jorgensen, 2004) e é atualmente uma técnica comum no estudo inicial de compostos com potencial farmacológico. A disponibilização de bancos de dados gratuitos, na internet, como o ZINC (Irwin and Shoichet, 2005) e o SEA (Keiser, Roth et al. 2007) facilitam a comparação entre estruturas recém-descobertas em bancada com outras semelhantes que já possuem alguma atividade (Irwin and Shoichet 2005).

Nesta tese foram estudados aspectos moleculares e evolutivos das DNA e RNA polimerases codificadas pelo plasmídeo mitocondrial de *M. perniciosa*. No primeiro capítulo, as estruturas dessas duas enzimas foram elucidadas através de Modelagem Molecular por Homologia e Dinâmica Molecular, onde foram propostas duas estruturas protéicas esteroquimicamente viáveis e semelhantes a outras polimerases da mesma classe. No segundo e terceiro capítulos foram propostos dois inibidores (um para cada polimerase), selecionados e validados através de processos de triagem virtual com posterior validação por Docking e Dinâmica Molecular. O quarto capítulo relaciona a expressão relativa dos genes codificadores das DPO e RPO através de RT-qPCR, em diversas fases de desenvolvimento do fungo, com um mecanismo de defesa do fungo contra o ataque oxidativo do *T. cacao*. No quinto capítulo delimitou-se a filogenia das DPO e RPO do plasmídeo de *M. perniciosa*, com diversas outras codificadas por plasmídeos de Basidiomycota e Ascomycota, além de propor um provável evento de Transferência Horizontal de Genes (THG) nesses grupos em relação às polimerases virais.

Referências

Bleicher KH, Böhm HJ, Müller K, Alanine AI. **Hit and lead generation: beyond high-throughput screening**. Nat Rev Drug Discov. 2003 May;2(5):369-78.

Carlson HA. **Protein flexibility and drug design: how to hit a moving target**. Curr Opin Chem Biol. 2002 Aug;6(4):447-52

Cavasotto CN, Abagyan RA. **Protein flexibility in ligand docking and virtual screening to protein kinases**. J Mol Biol. 2004 Mar 12;337(1):209-25

Gohlke H, Klebe G. A**pproaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors**. Angew Chem Int Ed Engl. 2002 Aug 2;41(15):2644-76.

Irwin JJ, Shoichet BK. **ZINC- A free database of commercially available compounds for virtual screening**. J Chem Inf Model. 2005 Jan-Feb;45(1):177-82.

Jorgensen WL. **The many roles of computation in drug discovery**. Science. 2004 Mar 19;303(5665):1813-8

Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. **Relating protein pharmacology by ligand chemistry**. Nat Biotechnol. 2007 Feb;25(2):197-206.

Kitchen DB, Decornez H, Furr JR, Bajorath J. **Docking and scoring in virtual screening for drug discovery: methods and applications**. Nat Rev Drug Discov. 2004 Nov;3(11):935-49.

Lyne PD. **Structure-based virtual screening: an overview**. Drug Discov Today. 2002 Oct 15;7(20):1047-55.

Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, García O, Thomazella DP, de Oliveira BV, Pires AB, Rio MC, Araújo MR, de Moraes MH, Castro LA, Gramacho KP, Gonçalves MS,

Neto JP, Neto AG, Barbosa LV, Guiltinan MJ, Bailey BA, Meinhardt LW, Cascardo JC, Pereira GA. **A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao**. BMC Genomics. 2008 Nov 18;9:548.

Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS. **Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock**. Proteins. 2002 Jan 1;46(1):34-40.

Pires AB, Gramacho KP, Silva DC, Góes-Neto A, Silva MM, Muniz-Sobrinho JS, Porto RF, Villela-Dias C, Brendel M, Cascardo JC, Pereira GA. **Early development of *Moniliophthora perniciosa* basidiomata and developmentally regulated genes**. BMC Microbiol. 2009 Aug 4;9:158.

Schapira M, Abagyan R, Totrov M. **Nuclear hormone receptor targeted virtual screening**. J Med Chem. 2003 Jul 3;46(14):3045-59.

Taylor RD, Jewsbury PJ, Essex JW. **A review of protein-small molecule docking methods**. J Comput Aided Mol Des. 2002 Mar;16(3):151-66

Zavodszky MI, Sanschagrin PC, Korde RS, Kuhn LA. **Distilling the essential features of a protein surface for improving protein-ligand docking, scoring, and virtual screening**. J Comput Aided Mol Des. 2002 Dec;16(12):883-902.

CHAPTER 1

Comparative Modeling of DNA and RNA Polymerases from *Moniliophthora perniciosa* Mitochondrial Plasmid

Abstract

The filamentous fungus *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora is a hemibiotrophic Basidiomycota that causes witches' broom disease of cocoa (*Theobroma cacao* L.). This disease has resulted in a severe decrease in the Brazilian cocoa production, which changed the position of Brazil in the market from the second largest cocoa exporter to a cocoa importer. Fungal mitochondrial plasmids are usually invertrons encoding DNA and RNA polymerases. Plasmid insertion into host mitochondrial genomes are probably associated to modifications in host generation time, which can be involved in fungal aging. This association suggests activity of polymerases, and these can be used as new targets for drugs against mitochondrial activity of fungi, more specifically against witches' broom disease. DNA and RNA polymerases of *M. perniciosa* mitochondrial plasmid were completely sequenced and their models were carried out by Comparative Homology approach. The sequences of DNA and RNA polymerase showed 25% of identity to 1XHX and 1ARO (pdb code) using BLASTp, which were used as templates. The models were constructed using Swiss PDB-Viewer and refined with a set of Molecular Mechanics (MM) and Molecular Dynamics (MD) in water carried out by AMBER 8.0, both working under the ff99 force fields, respectively. Ramachandran plots were generated by Procheck 3.0 and exhibited models with 97% and 98% for DNA and RNA polymerases, respectively. MD simulations in water showed models with thermodynamic stability after 2000ps and 300K of simulation. This work contributes for the development of new alternatives for controlling the fungal agent of witches' broom disease.

1.1 Background

The filamentous fungus *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora is a hemibiotrophic Basidiomycota (Agaricales, Tricholomataceae) that causes witches' broom disease of cocoa (Theobroma cacao L.). It has been claimed as one of the most important phytopathological problem that has afflicted the Southern Hemisphere in recent decades. In Brazil, this phytopathogen is endemic in the Amazon region [1]. However, since 1989, this fungus was found in the cultivated regions at the state of Bahia, the largest production area in the country. The fungus caused a severe decrease in the Brazilian cocoa production reducing the Brazil from the second largest cocoa exporter to a cocoa importer in just few years [2].

Plasmids are extragenomic DNA or RNA molecules that can independently reproduce in live cells. Their structure can be circular or linear, and include complete protein coding genes, pseudogenes, non-protein coding genes and inverted repetitive elements. The probable known plasmid function in their fungal hosts is related to the change of aging time. Fungal linear mitochondrial plasmids present the same basic structure than those in other organisms, but they also carry viral-like DNA and RNA polymerases (DPO and RPO, respectively) ORFs and have 3' and 5' inverted terminal repeats, also a 5' binding protein. This protein can be involved in both replication and integration processes of these plasmids in the mitochondrial genomes [3, 4]. Interestingly, a linear mitochondrial plasmid with the same typical characteristics carried by the other mitochondrial plasmids was found completely integrated to the *M. perniciosa* mitochondrial genome, by the Witches' Broom Genome Project (http://www.lge.ibi.unicamp.br/vassoura/) [5].

The Φ29 DNA polymerase is in the group α-DNA-polymerases due to its sensitivity to the aphidicolin and specific inhibitors, nucleotides similar to BuAaATP and BuPdGTP [6]. This polymerase is the main replication enzyme of double-strand-DNA viruses from bacteria and

eucaryotes. This is a 66KDa enzyme included in the eucaryotic replicases family [7], with the ability of using a protein as primer in the replication process [8, 9]. The T7 RNA polymerase is a 99KDa single chain viral enzyme that executes a specific-promoter transcription process in vivo and in vitro and is in the singlechain RNA polymerases family. The transcription mechanism carried out by this enzyme shares several similarities with other multichains RNA polymerases [9].

It is generally accepted that the water molecules in the hydration environment around a protein play an important role in its biological activity [10], and it contributes in stabilizing the native state of the protein [11]. In addition, this interaction has long been recognized as a major determinant of chain folding, conformational stability, and internal dynamics of many proteins, and as important to the interactions related to substrate binding, enzyme catalysis, and supramolecular recognition and assembly [12]. Standard Molecular Dynamics approaches measure the conformational space of a protein using atomic interactions from several force fields and including explicitly treated water to reproduce solvent effects [13].

The aim of this work was carried out the homology modeling of both DNA and RNA polymerases from the linear mitochondrial plasmid of *M. perniciosa*. With the accomplishment of this work, these models can be used as new molecular targets to find drugs against the witches' broom disease by de novo design methods [10].

1.2 Methods

After the release of the primary sequences of DNA and RNA polymerases from *M. perniciosa* mitochondrial plasmid, those are available in Witches' broom project database (LGE). 3D models were built by Comparative Modeling approach. Initially, both DNA and RNA polymerases sequences were submitted to BLASTp algorithm [14] restricted to Protein Data Bank (PDB). The found templates were aligned with the protein sequences of both DNA and

RNA polymerases by TCOFFEE [15] to find conserved regions and motifs. The 3D models were constructed using SwissPdb Viewer 3.7 [17] following a standard protocol: (I) load template pdb file; (II) align primary target sequence with template; (III) submit modeling request to Swiss Model Server. Then, the initial models constructed by SwissPdb Viewer, were prepared using LEAP and submitted to SANDER for structure refinement. The model structures were full minimized with 100 steps of steepest descent followed by more 100 steps of conjugate gradient to an RMS gradient of 0.01 kcal/2.71Å in vacuum, and then in water for 200 steps of steepest descent followed by more 200 steps of conjugate gradient to an RMS gradient of 0.01 kcal/2.71Å. Next, MD simulations of refined structures were performed in water using f99 force field at 300 K of temperature during 2000 ps. All MD simulations were carried out without constrain methods. The cutoff value of 14 Å was used for minimization of geometry and MD simulations. LEAP and SANDER are utilities of AMBER 9.0 [18, 19]. Additionally, all calculations were performed without restraints. Time averaged structures were generated by time averaging of simulations from the point a stable trajectory, which were obtained through the end of simulation. The Visual Molecular Dynamics (VMD) software [20] was used to visualize trajectory results produced by SANDER module. Finally, PROCHECK 3.4 [21] and Atomic Non-Local Environment Assessment (ANOLEA) [22, 23] were used to evaluate both DNA and RNA polymerases using Ramachandran plot [24] and energy calculations on a protein chain of each heavy atom in the molecule, respectively [25]. Graphics of RMS x Time were generated by VMD 1.8.6 [26]

1.3 Results and Discussion

Blastp results for both DNA and RNA polymerases of the *M. perniciosa* linear mitochondrial plasmid showed just one reliable template to each enzyme (Table 1). The 1XHX [27] and 1ARO [28] were used as a template DPO and RPO respectively. Although both of them showed a low

identity with targets, it is possible to build useful models for docking studies [10]. The root-mean-squared deviation (RMSD) for Cα between DPO-1XHX and RPO-1ARO are 2,40 Å and 1.84 Å respectively. These values show some differences between models and crystal structures, as one might expect, principally in relation to the number of residues. The models have 543 and 766 residues to DPO and RPO, while crystal structures have 575 and 883 residues for 1XHX and 1ARO, respectively.

In addition, these results address the hypothesis of several authors correlating plasmid sequences to DNA and RNA polymerases of adeviruses and retroviruses sequences [3, 28].

Using 1HXH as a template, the 3D structure of the DNA polymerase was built from the linear mitochondrial plasmid of *M. perniciosa*. This polymerase was classified within the B family of DNA polymerases, which can be found in viruses and cellular organelles. Figure 1 shows that DPO model has transferase features with alpha-beta secondary structure.

This model shows 17 alpha-helices, 36 beta-strands, 57 turns, and 315 hydrogen bonds can be observed in the whole structure. As well as other polymerases from that family, this polymerase showed the three standard domains of the group: Palm, Fingers, and Thumb.

The active site of the DNA polymerase of *M. perniciosa* (Figure 2) carries the conserved motif B represented by Lys380, Leu381, Leu382, Leu383, Asn384, Ser385, Leu386, Tyr387, Gly388, and it is involved in dNTPs selection and template DNA binding activity as described by Truniger et al. [6] in the homologous Φ29 DNA polymerase. These aminoacids are distributed among three domains: Palm, Fingers and Thumb. Other motifs involved with DNA polymerization were found in this polymerase, such as Dx2SLYP (Asp247, Val248, Asn249, Ser250, Leu251, Tyr252, Pro253), YxDTDS (Tyr455, Ser456, Asp457, Thr458, Asp459), Tx2A/GR (Thr309, Asp310, Lys311, Gly312, Tyr313, Arg314) KxY (Lys494, Met495, Tyr496), which were reported in several studies [6, 8, 9, 30, 31, 32, 33].

The active site of the RNA polymerase from *M. perniciosa* (Figure 4) plasmid is formed by aminoacids from two domains: Palm (Asp457 and Asp695) and Fingers (Tyr537 and Lys529). In comparison to the template structure, these aminoacids perform an alignment in the region of the active site, with the aminoacids Asp537 and Asp812 (Palm), and Tyr639 and Lys631 (Fingers) of the template. The presence of these residues (Asp, Tyr, and Lys) in this region is a sign on this group of polymerases, and they are involved with transcriptional process [10, 34, 35].

Both the DNA and RNA polymerases, after submitted to refinement by optimization of geometry and MD simulations, had their structures validated by PROCHECK and ANOLEA (Figure 3). The Ramachandran plot showed that 97% and 98% of residues are within the allowed regions for DPO and RPO, respectively. Almost all residues show negative values of energy (green), whereas few aminoacids obtained positive values of energy (red). It means that most of residues are in favorable environment of energy. In other words, the quality of both main chain and side chain was evaluated showing that the built models had appropriated sterochemical and thermodynamic values. As a result, although the targets and templates proteins showed a low homology identity, the tertiary structure obtained had the same sign of family.

1.4 Conclusions

The great challenge of genome projects is to elucidate new molecular targets, mainly proteins and enzymes. Functional characterization of proteins is one of the most frequent problems in biology. While sequences provide valuable information, the identification of relevant residues inside them is frequently impossible because of their high plasticity, suggesting a construction of 3D models. In the case of enzymes, a similar function can be assumed between two proteins if their sequence identity is above 40%. In addition, polymerases are suitable targets for antiviral drugs [36], which have nucleoside analogs as substrates. Theses inhibitors have development using rational design way. Thus, our findings address to use of fungi polymerases as start points

for drug design against witches' broom disease, following the similar methodologies used for the development of inhibitors of polymerases of virus. Our models are suitable for computer aided-drug design approaches, such as docking, virtual screening, and QM/MM in order to search a new lead compound against the witches' broom disease.

Acknowledgements

State University of Feira de Santana (UEFS); and the scholarship and financial support by FAPESB.

References

1. Aime M C, Phillips-Mora W: **The causal agents of witches' broom and frosty pod rot of cacao (chocolate,Theobroma cacao) form a new lineage of Marasmiaceae**. *Mycologia* 2005, **97**(5):1012-1022
2. Lopes M A: **Estudo molecular de quitinases de Crinipellis perniciosa (Stahel) Singer**. Master Thesis. State University of Santa Cruz, Ilhéus, Bahia, Brazil: 2005.
3. Griffiths A J F: **Natural Plasmids of Filamentous Fungi**. *Microbiol. Rev.* 1995, **59**(4):673 – 685.
4. Jack Kennell and lab co-workers at Saint Louis University [http://pages.slu.edu/faculty/kennellj/index.html]
5. Formighieri E, Tiburcio R A, Armas E D, Medrano F J, Shimo H, Carels N, Góes-Neto A, Cotomacci C, Carazzolle M F, Sardinha-Pinto N, Thomazella D P, Rincones J, Digiampietri L, Carraro D M, Azeredo-Espin A M, Reis S F, Deckmann A C, Gramacho K, Gonçalves M S, Moura Neto J P, Barbosa L V, Meinhardt L W, Cascardo J C, Pereira G A: **The mitochondrial genome of the phytopathogenic basidiomycete**

Moniliophthora perniciosa is 109 kb in size and contains a stably integrated linear plasmid. *Mycol Res.* 2008, **112**(Pt 10):1136-52

6. Truniger V, Lázaro J M, Vega M, Blanco L, Salas M: **Φ29 DNA Polymerase Residue Leu384, Highly Conserved in Motif B of Eukaryiotic Type DNA Replicases, Is Involved in Nucleotide Insertion Fidelity**. *J. Biol. Chem.* 2003, **278**(35):33482 – 33491
7. Koonin E V, Senkevich T G, Dolja V V: **The ancient Virus World and evolution of cells**. *Biol. Direct* 2006, **1**:29
8. Blasco M A, Lazaro J M, Blanco L, Salas M: **Φ29 DNA polymerase active site. Residue Asp249 of conserved amino acid motif Dx2SLYP is critical for synthetic activities**. *J. Biol. Chem.* 1993, **268**(32):24106– 24113
9. Sousa R, Chung Y J, Rose J P, Wang B C: **Structure of bacteriophage T7 RNA polymerase at 3.3 A resolution**. *Nature* 1993, **364**(6438):593 – 599
10. Holtje H D, Sippl W, Rognan D, Folkers G: *Molecular Modeling: Basic principles and applications*. WILEY-VCH, 2003
11. Balasubramanian S, Bandyopadhyay S, Pal S, Bagchi B: **Dynamics of water at the interface of a small protein, enterotoxin.** *Curr. Sci.* 2003, **85**(11):1571-1578
12. Denisov VP, Halle B: **Protein Hydration Dynamics in Aqueous Solution: A Comparison of Bovine Pancreatic Trypsin Inhibitor and Ubiquitin by Oxygen-17 Spin Relaxation Dispersion**. *J. Mol. Biol.* 1995, **245**(5):682–697
13. Zhou1 L, Siegelbaum S A: **Effects of surface water on protein dynamics studied by a novel coarse-grained normal mode approach**. *Biophys. J* 2008, **94**(9):3461-3474
14. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic Acids Res.* 1997, **25**(17):3389-402

15. Notredame C, Higgins D G, Heringa J: **T-Coffee: A novel method for multiple sequence alignments**. *J. M. Biol.* 2000, **302**(1):205-17

16. BiomedCache version 6.0; Beaverton, Fujitsu America Inc. 2003.

17. Guex N, Diemand A, Peitsch M C: **Protein modelling for all**. *Trends Biochem Sci.* 1999, 24(9):364-7

18. Weiner S J, Kollman P A, Case D A, Singh U C, Ghio C, Alagona G, Profeta S, Weiner P: **A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins**. *J. Am. Chem. Soc.* 1984, **106**(3):765 – 784

19. S.J. Weiner et al. J. "An A11 Atom Force Field for Simulations of Proteins and Nucleic Acids", Comput. Chem., 1986, pp. 230 – 252.

20. W. Humphrey, A. Dalke and K. Schulten. "VMD - Visual Molecular Dynamics", J. Molec. Graph., 1996, pp. 33-38. 1996.

21. Laskowski RA, MacArthurm MW, Smith DK, Jones DT, Hutchinson EG, Morris AL, Naylor D, Moss DS, Thornton JM: PROCHECK v.3.0 – Program to check the stereochemistry quality of protein structures – Operating instructions. 1994

22. Melo F, Feytmans E: **Assessing Protein Structures with a Non-local Atomic Interaction Energy**. *J. M. Biol.* 1998, **277**(5):1141-52

23. Melo F, Feytmans E **: Novel knowledge-based mean force potential at atomic level**. *J. M. Biol.* 1997, **267**(1):207-22

24. Ramachandran G N, Ramakrishnan C, Sasisekharan V: **Stereochemistry of polypeptide chain configurations**. *J. Mol. Biol.* 1963, **7**:95-9

25. Melo F, Devos D, Depiereux E, Feytmans E: **ANOLEA: a www server to assess protein structures**. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 1997, **5**:187-90.

26. Humphrey W, Dalke A, Schulten K: **VMD: visual molecular dynamics**. *J. Mol. Graph.*, 1996, pp. 33-38.

27. Kamtekar S, Berman AJ, Wang J, Lázaro JM, de Vega M, Blanco L, Salas M, Steitz TA: **Insights into strand displacement and processivity from the crystal structure of the protein-primed DNA polymerase of bacteriophage phi29**. *Mol.Cell.* 2004, **16**(4):609-18

28. Jeruzalmi D, Steitz T A: **Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme**. *EMBO J.* 1998, **17**(14):4101-13

29. Fukuhara H: **Linear DNA plasmids of yeasts**. *FEMS Microbiol. Lett.* 1995, **131**(1):1 – 9

30. Truniger V, Lázaro J M, Salas M, Blanco L: **Φ29 DNA polymerase requires the N-terminal domain to bind terminal protein and DNA primer substrates**. *J. M. Biol.* 1998, **278**:741 – 755

31. Esteban J A, Salas M and Blanco L: **Fidelity of Φ29 DNA Polymerase: Comparison Between Protein-Primed Initiation and DNA Polymerization**, *J. Biol. Chem.* 1993, **268**:2719–2726

32. Garmendia C, Bernard A, Esteban JA, Blanco L, Salas M: **The Bacteriophage Φ29 DNA Polymerase, a Proofreading Enzyme**, *J. Biol. Chem.* 1992, **267**:2594–2599

33. Eisenbrandt R, Lázaro J M, Salas M, Vega M: **Φ29 DNA Polymerase residues Tyr59, His61 and Phe69 of the high conserved ExoII motif are essential for interaction with the terminal protein**, *Nuc. Acid. Res.* 2002, **30**(6): 1379–1386

34. Bonner G, Patra D, Lafer E M, Sousa R: **Mutations in T7 RNA polymerase that support the proposal for a common polymerase active site structure**, *EMBO J.* 1992, **11**(10):3767 – 3775

35. Cheetham G M T, Jeruzalmi D, Steitz T: **Structural basis for initiation of transcription from an RNA polymerase-promoter complex**, Nature 1999, **399**(6731):80-3.

36. Öberg B: **Rational design of polymerase inhibitors as antiviral drugs**, Antiviral Research 2006, **71**, 90-95.

Figures

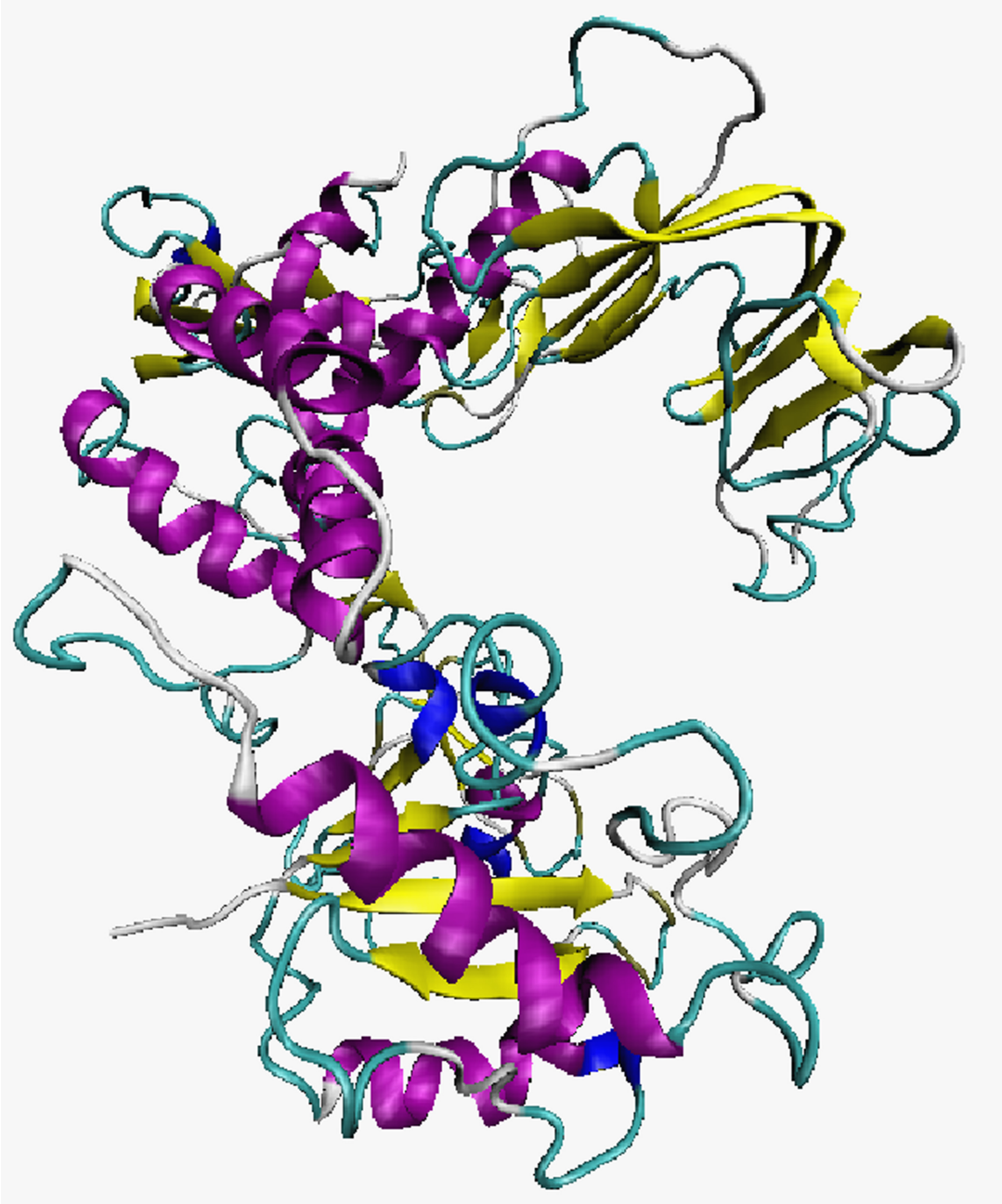


Figure 1.1 - The 3D structure of the DNA polymerase from the *M. perniciosa* mitochondrial plasmid. Magenta: helices; yellow: strands; blue: turns.

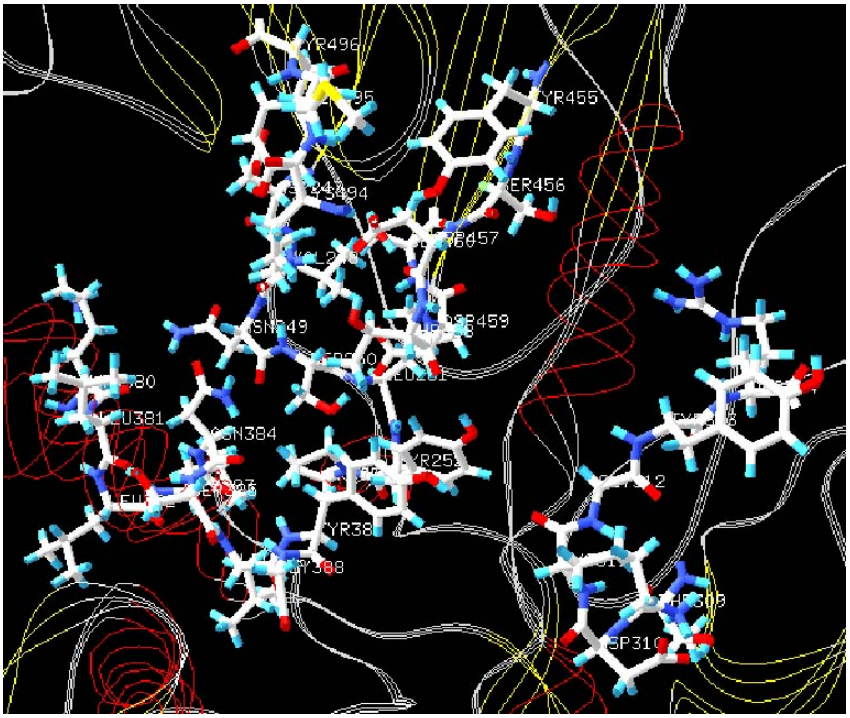


Figure 1.2 - Active site of the DNA polymerase from the *M. perniciosa* mitochondrial plasmid presenting the conserved motif B.

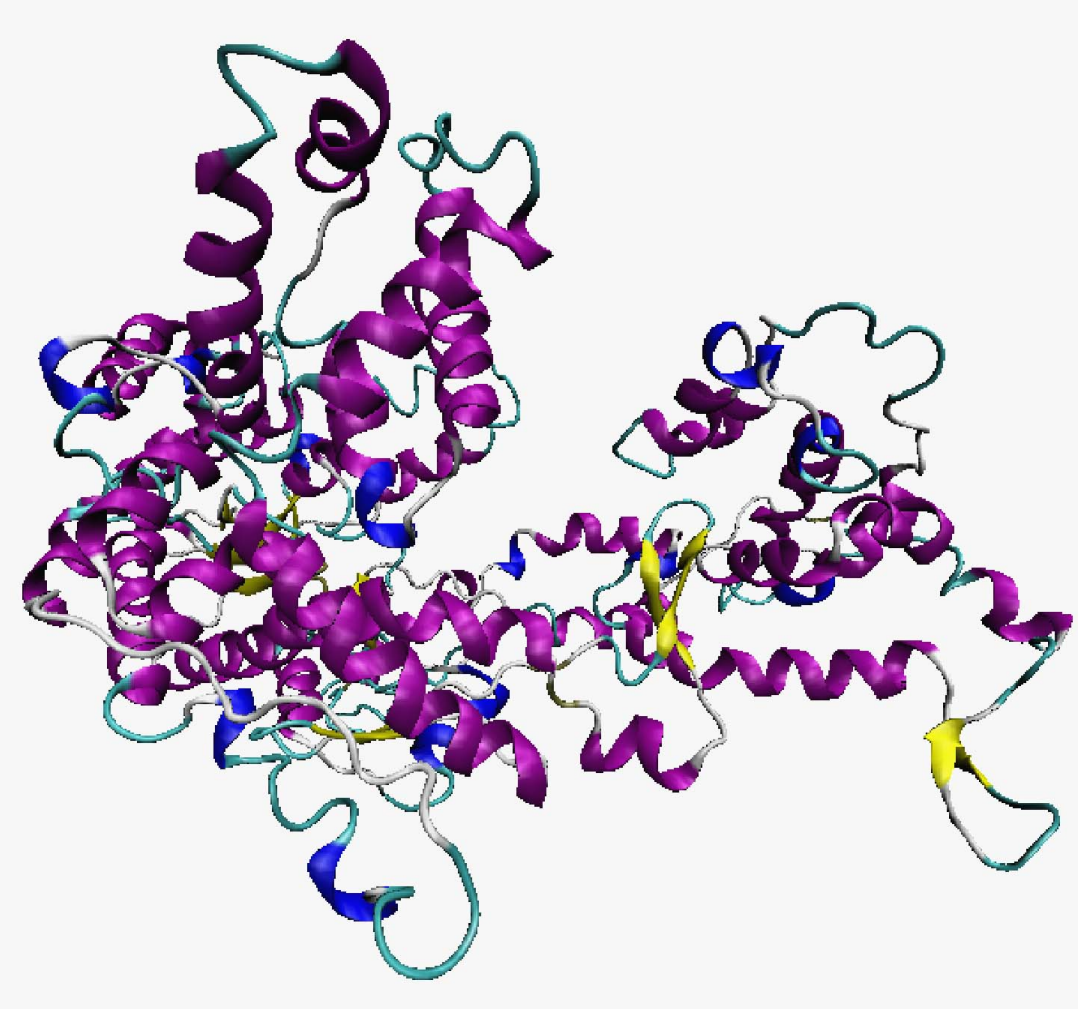


Figure 1.3 - The 3D structure of the RNA polymerase from the *M. perniciosa* mitochondrial plasmid. Magenta: helices; yellow: strands; blue: turn.

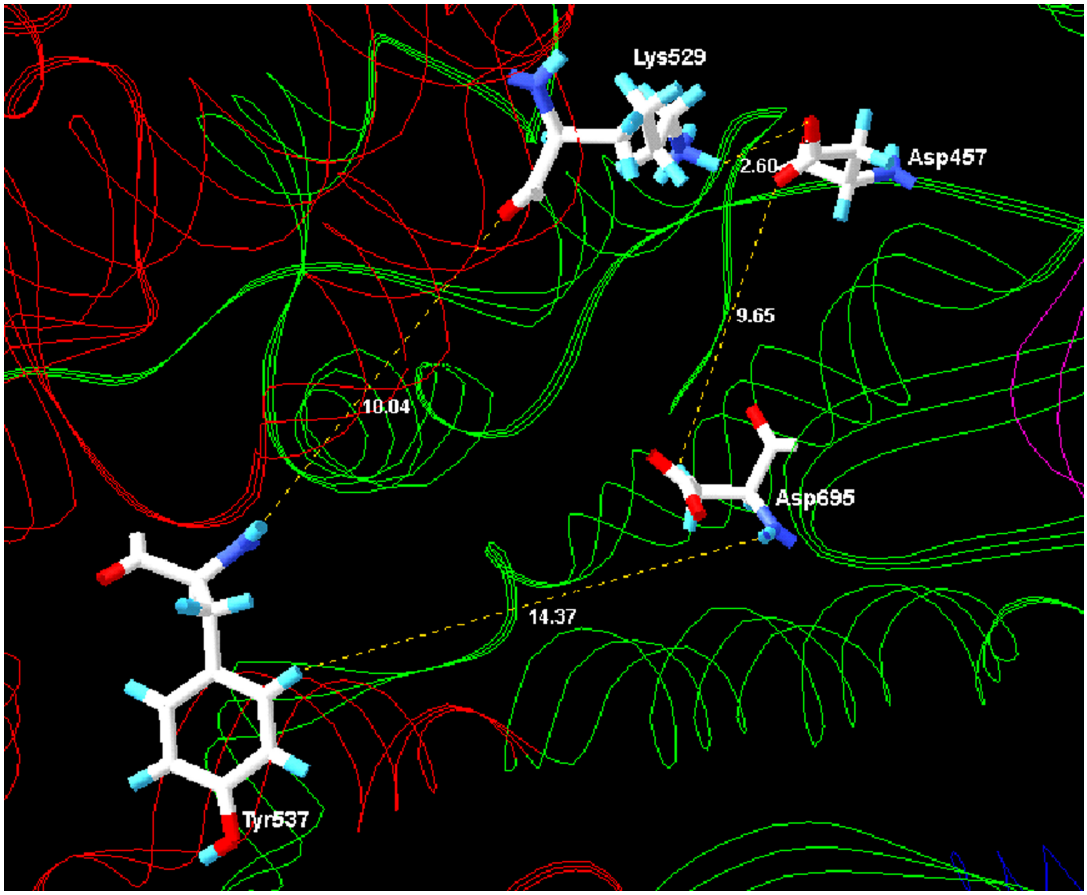


Figure 1.4 - Active site of the RNA polymerase from *M. perniciosa* mitochondrial plasmid formed by two domains: Palm (green) and Fingers (red).

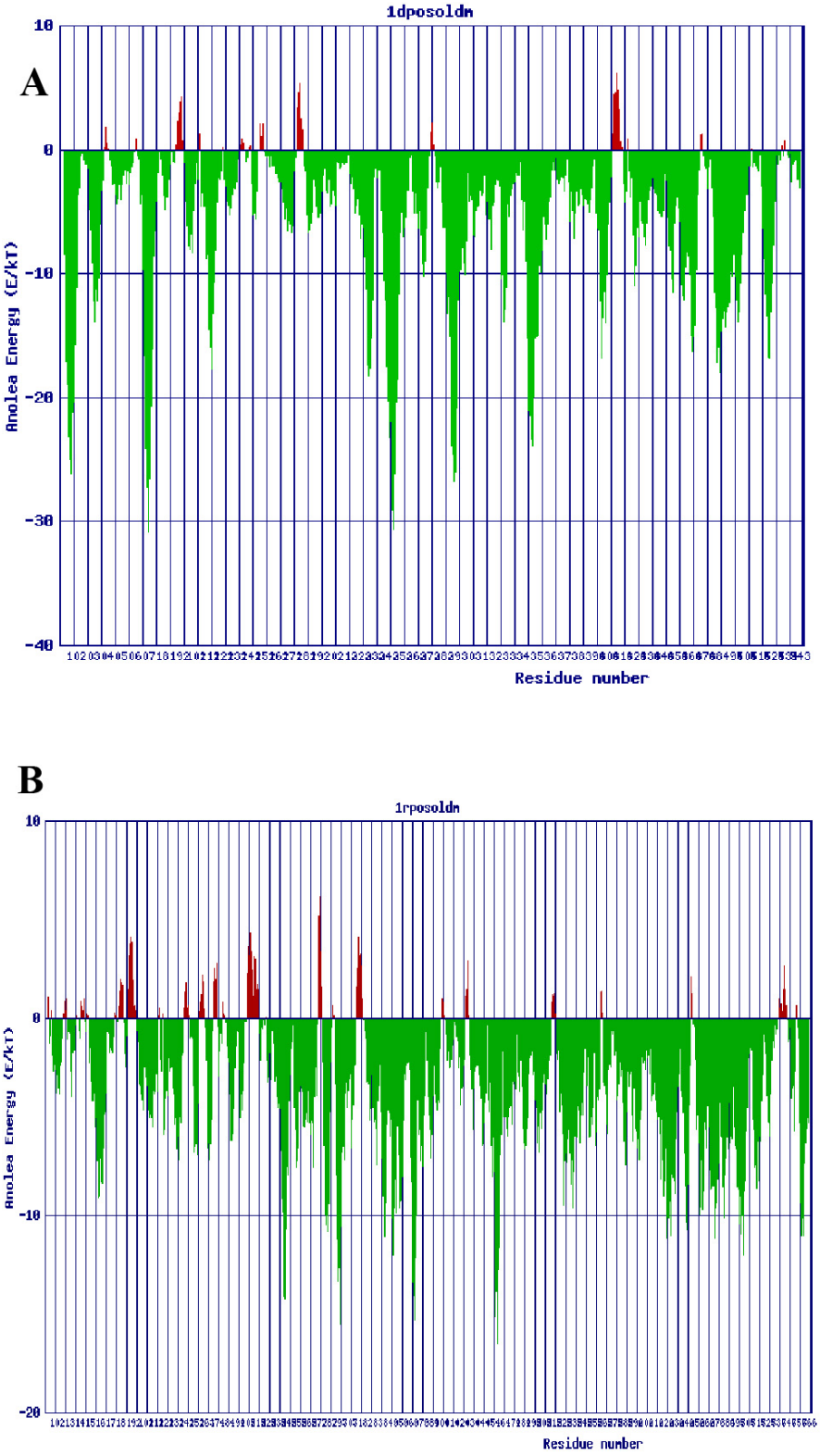


Figure 1.5 - ANOLEA validation of the built model. A) DPO; B) RPO. Green and red means negative and positive values of energy.

Tables

Table 1.1 - Selected templates obtained by Blastp algorithm

	Template	Identity	E-value	Organism	RMS (Å)
DPO	1XHX	32%	8e-06	Phage Φ29	2,40
RPO	1ARO	25%	1e-33	Phage T7	1.84

CHAPTER 2

Virtual screening reveals a viral-like polymerase inhibitor that complex with the *M. perniciosa* DNA polymerase

Abstract

The filamentous fungus *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora is a Basidiomycota that causes witches' broom disease of cocoa (*Theobroma cacao* L.). In some cases, polymerases coded by mitochondrial plasmids may change the aging time of some fungal species. The mitochondrial DNA polymerase of *M. perniciosa* (MpmitDNApol) is classified within the B family of DNA polymerases, which can be found in viruses and cellular organelles. The structure of that polymerase was previously constructed using a homology modeling approach. Using virtual screening processes, accessing KEGG, PubChem and ZINC databases, we selected the 27 best putative nucleoside viral-like polymerase inhibitors to test against MpmitDNApol. Autodock Vina was used to perform docking calculations for each molecule, and to return energy values in several ligand conformations. After, we used Pymol 1.4 to check presence or absence of hydrogen, stereochemistry of chiral carbons, substructure, superstructure, number of rotable bonds, number of rings, number of donor groups, and hydrogen bonding receptors. As a result, we selected the Entecavir Hydrate, a hepatitis B polymerase inhibitor, and then AMBER 10 was used to describe the behavior of polymerase-entecavir complex after a set of 3500 ps of simulation up to 300 K in water. This calculation returned a graph of Potential Energy during the time of simulation, and showed that the ligand remains inside the active site after this time with a final energy of -612587.4214 Kcal/Mol. Therefore, we suggest Entecavir Hydrate as a good inhibitor to be tested *in vitro* and *in vivo* against *M. perniciosa*.

2.1 Introduction

The *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora is a Basidiomycota that causes witches' broom disease of cocoa (*Theobroma cacao* L.). This filamentous fungus carries a linear-type plasmid that encodes viral-like DNA and RNA polymerases (Formighieri et al., 2008). *M. perniciosa* mtDNA polymerase (DPO) structure was previuosly described by Andrade et al. (2009) as 543 amino acid DNA-dependent DNA polymerase within B family of polymerases, which can be found in viruses and cellular organelles. Its active site is distributed among three domains: Palm, Fingers and Thumb, with all the motifs involved in dNTP´s and DNA polymerization.

Entecavir is a guanosine nucleoside analogue, with activity against Hepatitis B virus DNA polymerase (Palumbo, 2009). By competing with the natural substrate deoxyguanosine TP, entecavir-TP functionally inhibits the 3 activities of the viral polymerase: (1) priming of the HBV polymerase, (2) reverse transcription of the negative strand DNA from the pregenomic messenger RNA, and (3) synthesis of the positive strand HBV DNA (Beckebaum et al., 2003). ETV triphosphate (ETV-TP) displays activity against all three synthetic activities of the HBV polymerase, i.e., the unique protein-linked priming activity, RNA-directed first strand DNA synthesis or reverse transcription, and second strand DNA-directed DNA synthesis (Langley et al, 2007; Seifer et al., 1998).

Virtual structure-based screening has become prominent in drug discovery, using protein targets (Hou et al., 2011; Okimoto et al, 2009). Currenty, several free ligands databases are widely available. Searching for molecules that can complex with target proteins may either be done by keywords (eg. KEGG and PubChem databases) as well as using a structure-activity relationship, available in Zinc (Irwin and Shoichet, 2005) and PubChem databases.

One of the most important techniques for receptor-based drug design is Molecular Docking (Hou et al., 2011). Using crystallographic or modeled protein structures, molecular docking is often used for screening compound libraries and predicting the conformation of a protein-ligand complex as well as for calculating its binding affinity (Okimoto et al, 2009). Usually, docking programs such as Autodock Vina (Trott and Olson, 2010) generate multiple protein-ligand conformations by sampling the ligand's probable conformations in the binding pocket of the target protein, using a flexible ligand-rigid receptor docking (Hou et al., 2011; Trott and Olson, 2010). The use of scoring functions in docking calculations by these programs is an attempt to approximate the standard chemical potentials of the system (Trott and Olson, 2010). Autodock Vina uses a force-field-based scoring function approach to estimate the binding affinities by calculating the non-bonded interactions based on traditional force fields, identifying the correct binding pose of a ligand and ranking ligands using the predicted binding affinities (Hou et al., 2011; Trott and Olson, 2010, Wang et al., 2006; Wang et al., 2001). On the other hand, the problems of molecular docking as a screening tool have been widely discussed since the scoring functions are generally inaccurate, neglect the solvent-related terms, as well as the protein flexibility (Okimoto et al, 2009). Coupled Molecular Docking and Molecular Dynamics is a good way to solve this problem, because it can treat both proteins and ligands in a flexible manner, allowing the relaxation of the binding site around the ligand (Hou et al., 2011; Okimoto et al, 2009; Aqvist et al., 2002).

Molecular mechanics/Poisson Boltzmann surface area (MM/PBSA) combines molecular mechanics energy and implicit solvation models, and it is more rigorous than most empirical or knowledged-based scoring functions (Hou et al., 2011; Okimoto et al, 2009; Kollman et al, 2000). It allows for rigorous free-energy decomposition into contributions originating from different groups of atoms or types of interaction (Hou et al., 2011; Hou et al., 2008). In the

MM/PB-SA method, the free energy is calculated using the snapshots of solute molecules obtained from explicit-solvent MD simulation (Okimoto et al, 2009, Kuhn et al., 2005).

The aim of this study was to search a series of likely molecules, available at KEGG, PubChem and Zinc databases, which can form complexes with mitochondrial plasmid DNA polymerase from *M. perniciosa*, and, further, select a potential inhibitor using a coupled Molecular Docking and Molecular Dynamics - MM / PBSA approach.

2.2 Methods

2.2.1 Ligand Searching

Initially, we made a text-based quering in Kegg (http://www.genome.jp/kegg/) and PubChem (http://pubchem.ncbi.nlm.nih.gov/) databases. The selected molecules were those described as inhibitors of DNA polymerases, which had been already tested *in vitro* against other organisms. All 2D structures were copied in Similes format in order to be compared to other Zinc database (http://zinc.docking.org/) molecules. Furthermore, the 3D structures of these molecules were downloaded in mol2 and pdb formats to be used in a Virtual Screening process, which was carried out by Molecular Docking and Molecular Dynamics. According to a protocol described by Irwin and Schoichet (2005), molecules obtained from Zinc database were selected for comparison, following an interval between 95-99% similarity, with the structures found in KEGG and PubChem databases. In addition, selected structures were downloaded in mol2 and pdb formats for subsequent Docking Studies and Molecular Dynamics.

2.2.2 Docking Studies

The downloaded structures from KEGG, PubChem and Zinc databases were first checked in Pymol 1.4 (Schrödinger, LLC, 2011) to evaluate presence or absence of hydrogen, stereochemistry of chiral carbons, substructure, superstructure, number of rotable bonds, number of rings, number of donor groups, and hydrogen bonding receptors.

The molecules of the ligands and the receptor (DPO) were prepared in Autodock Tools 1.5.6 (Sanner, 1999). One added all polar hydrogens for the receptor and Kollman United Atomic Charges were computed (Kollman et al., 1993). For all ligands, we added polar and hydrogens and computed Gasteiger charges. The grid definition, adjusting to the DPO active site, was set up manually, following the recommendations of the program manual (Trott and Olson, 2010; Sanner, 1999). Then, the structures of the ligand and receptor were saved in pqbqt format to be used in docking calculations. Autodock Vina was used to perform Docking Scoring for each ligand-receptor complexes (Trott and Olson, 2010). Before running each Docking calculations, a configuration file was generated with grid size and coordinates information, as well as indicating ligand and receptor files. The reports (log files) for each calculation were analyzed, in order to obtain free-energy (Kcal/mol) values of each ligand conformations with its respective complex. In addition, we used Pymol to verify the number of hydrogen bonds and non-covalent interactions between each different ligand conformations and catalytic residues of DPO, which are involved with recognition and polymerization mechanisms. As a way to optimize the selection of an ideal complex, we selected just one ligand which best fits in the DPO active site, considering all stereochemical aspects previously evaluated and the free-energy results.

2.2.3 Molecular Dynamics of Complex

In this study we used MM/PBSA protocol to calculate the affinity and stability in the interaction of ligand-receptor DPO complex, using AMBER 10 (Case et al., 2008) package. Initially, we used Antechamber program to make ff99 force field recognizes the atom types of both ligand and receptor to avoid errors during the calculations.

Tleap was used to neutralize charges (ff99 force field) and DPO–ligand complex was immersed in a rectangular box of TIP3P water molecules. Following the protocol, we use Sander to carried out a Molecular Dynamics (MD) Equilibrium, and it was restricted to a region of the protein that

contains the active site (amino acids 247-496), according to the following parameters: 1000 cycles of steepest descent and 1000 cycles of conjugate gradient minimization, heating MD for 50 picoseconds (ps), density equilibrium for 50 ps, followed by a Equilibrium Dynamics for 500 ps at constant pressure and 300 K temperature. After the system has reached the equilibrium, we followed with MM/PBSA protocol (Hou et al., 2011; Case et al., 2008; Fogolari et al, 2003). Then, we simulated a total of 3000 ps production step Molecular Dynamics, divided in 3 sets of 1000 ps (prod1, prod2 and prod3), saving coordinates every 10 ps. Furthermore, we used the mm_pbsa.pl script to extract snapshots (without the water) from our production runs and get their trajectories. In addition, we checked complex stability by plotting a Potential Energy x Time (ps) graph, from all the production steps. In a last step, we use Ambpdb to generate a pdb file of the complex, after the last stage of the Molecular Dynamics. This structure was analyzed in PyMOL to verify if the ligand remained in the active site after the whole process.

2.3 Results and Discussion

2.3.1 Structures and binding energies of DPO complexes from AutoDock Vina

After the search conducted in KEGG, Pubchem and Zinc databases, we selected 27 structures that exhibit a good interaction with DPO. These molecules were all classified as nucleosides, as described in PubChem database. A reliable prediction of complex interactions is essential for selecting a potential ligand in virtual screening approaches, and it requires a suitable docking tool that is capable to generate energetic evaluation between ligand-protein, indicating the quality of interaction (Zaheer-ul-Haq et al., 2010). The results of Molecular Docking with AutoDock Vina for different ligand-DPO complexes are presented for the dominating configuration with minimum binding free energy (ΔG) in Table 1. Docking scores returned by Autodock Vina indicate DTP ligand as Top-ranked solution to be a good DPO inhibitor. However, to chose the best DPO inhibitor we analyze other features, instead only top-ranked

scores, such as H-bond donor and H-bond acceptor of each ligand, as well as the capacity at least one conformation of each ligand to bind to amino acids from active site pocket of DPO in the complex (Andrade et al., 2009;). Among the molecules studied, Entecavir Hydrate is that best binds with the amino acids of DPO active site, and it presented a high affinity in docking calculations (Figure 1). This molecule forms three hydrogen bonds with the amino acids Asp 457 and Tyr 496 (Figure 2), belonging to KxY and YxDTDS motifs, respectively, responsible for recognition of nucleotides and polymerization of DNA strands by *M. perniciosa* DPO (Andrade et al., 2009). Similar studies showed the binding efficiency of the Entecavir into the hydrophobic pocket of hepatitis virus DNA polymerase (HBV), in comparison to other potential nucleoside inhibitors (Langley et al., 2007; Walsh et al., 2010). Then, considering the high affinity of Entecavir Hydrate to DPO, it is possible that this inhibitor can block DNA polymerization, which corroborates with previous studies described above.

Entecavir Hydrate seems to be an excellent inhibitor of viral polymerases, since in many studies it is not only capable of inhibiting DNA-dependent DNA (HBV) polymerases but also reverse transcriptase (RNA-dependent DNA polymerase of HIV) (Walsh et al., 2010; Domaoal et al.; Lin et al., 2008; Langley et al., 2007; Lin et al., 2008). The DNA polymerase of *M. perniciosa* mitochondrial plasmid has a viral origin, and it is likely that this fact has influenced for good interaction with the ligand (Andrade et al., 2009).

Table 2.1 Potential inhibitors selected from the KEGG, PubChem and Zinc databases used in docking studies.

Molecule	Affinity (Kcal/mol)	H-Bond Donor	H-Bond Acceptor
Dtp	-7.9	3	2
Entecavir hydrate	-7	5	4
Valaciclovir Hydrochloride	-6.8	3	5
ZINC11679840	-6.7	4	3
Entecavir (ZINC03802690)	-6.6	4	3
ZINC42689357	-6.6	4	3
ZINC14768473	-6.5	5	4
Trifluridine	-6.5	3	8
ZINC00005235	-6.3	4	3
ZINC05157450	-6.3	4	3
Hby	-6.2	1	3
Brivudine	-6.1	3	5
Vidarabine sodium phosphate	-5.8	5	7
Vidarabine_monohydrate	-5.8	5	5
Ganciclovir	-5.8	4	4
Vidarabine phosphate	-5.8	4	4
Vidarabine anhydrous	-5.8	4	4
Penciclovir	-5.8	4	3
Famciclovir	-5.8	1	4
Cytrabine	-5.7	4	5
Cytarabine hydrochloride	-5.6	5	5
Idoxuridine	-5.6	3	5
Penciclovir sodium	-5.6	3	3
Cytarabine ocfosphate hydrate	-5.5	4	8
Ganciclovir sodium	-5.5	3	5
Cidofovir	-5.4	4	6
Aciclovir	-5.2	3	3

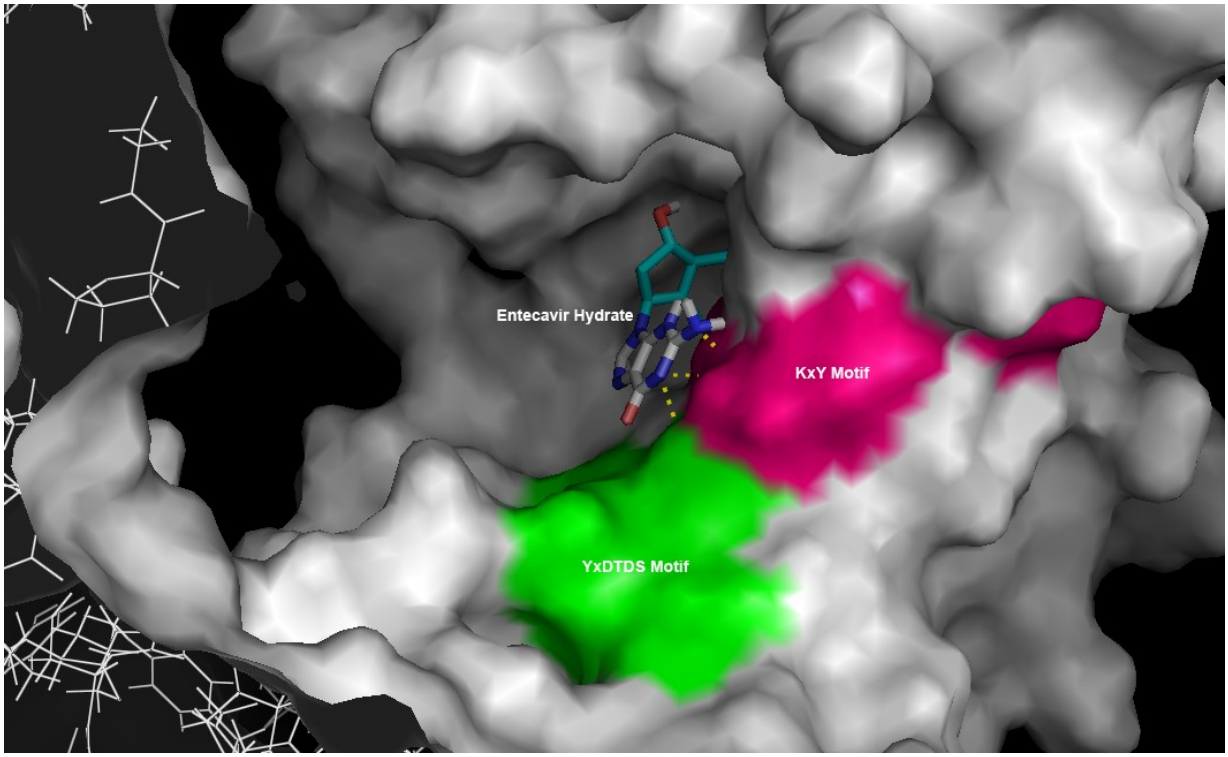


Fig. 2.1 DPO binding pocket showing Entecavir Hydrate interaction. YxDTDS motif in green and KxY motif in dark pink.

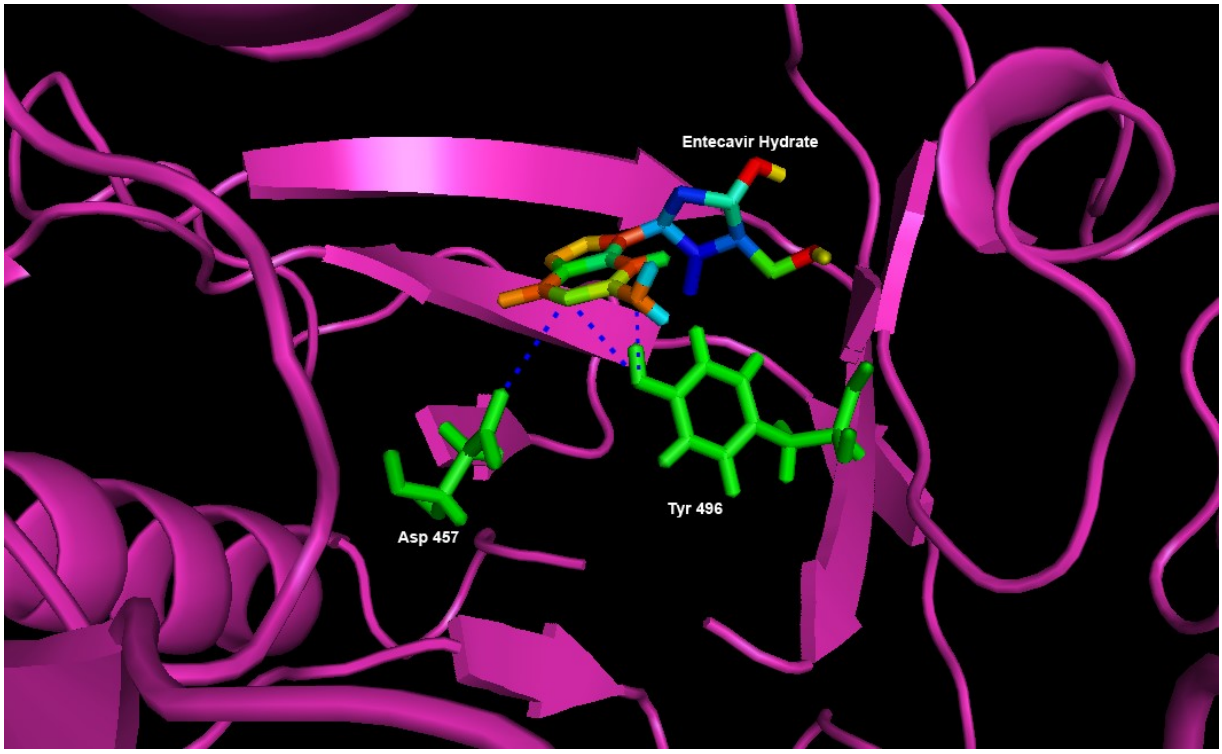


Fig. 2.2. Hydrogen bonds between Asp 457 and Tyr 496 in DPO active site.

2.3.2 Molecular Dynamics MM/PBSA of DPO-Entecavir complex

With Molecular Dynamics results we could evaluate the performance and stability of DPO-Entecavir complex. MM/PBSA simulations are feasible on proteins provide that electronic polarizability is taken into account (Fogolari et al., 2003).

In this paper, we analyzed Molecular Dynamics results regarding potential energy complex, during the simulation process, and its final energy. Regarding the RMSD, we found that even in equilibrium phase before 1000 ps, the convergence on the calculations had been already reached.

As one can see in Figure 3, the graph shows that above 1,000 ps simulation, the complex has already reached a potential energy value near the minimum. Between 1,000 and about 1,600 ps, we found a plateau in the graph, where there is an increase in energy, probably due to some sudden change in the RMSD of the complex during this period. In addition, the potential energy during the plateau shows that the structure of this complex is perfectly feasible to exist. Subsequently, the complex reached minimum energy region again where remained until the end of the simulation. The final energy reached at exactly 3459 ps was -612587.4214 Kcal/Mol.

The pdb of the complex, generated after molecular dynamics, showed that Entecavir Hydrate remains within the active site of DPO after 3500 ps of simulation. Walsh et al. (2010) characterized the catalytic mechanism of HBV DNA polymerase and its process of inhibition by Entecavir Phosphate, with a simulation time of 2000 ps. This was sufficient to show that there is a blockage of DNA polymerization when the complex is formed. Therefore, we use a simulation time greater than the other authors used, showing that this inhibitor remains in the active site of the DPO for a long time, blocking its activity.

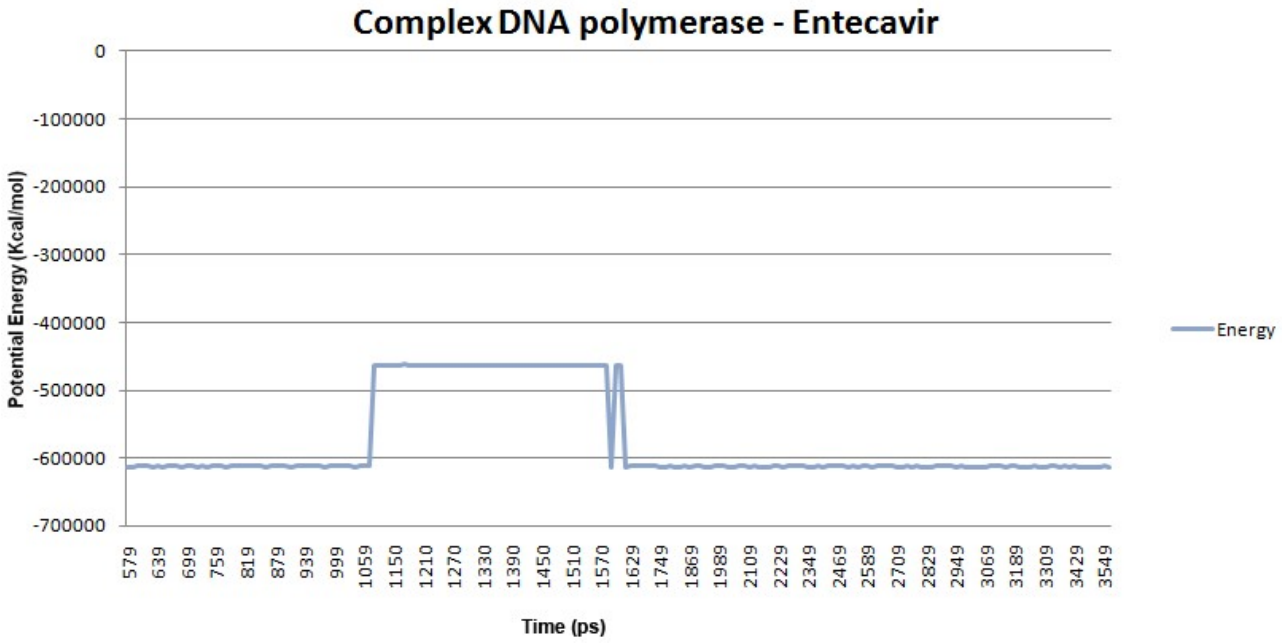


Fig. 2.3. Graph of potential energy of DPO-Entecavir complex DPO during 3.5 ns Molecular Dynamics simulation

2.4 Conclusions

In this article we describe the selection of one nucleoside inhibitor for DNA polymerase of mitochondrial plasmid of *M. perniciosa*, among 27 molecules found in public databases, using a virtual screening approach. An interesting fact in this work is that the ligand which made the top-ranked score in the docking calculations is a viral polymerase inhibitor, previously described in several papers.

The Entecavir Hydrate complex with the amino acids of hydrophobic pocket, which are involved with polymerization mechanism of DPO. This process remains stable during all 3500 ps Molecular Dynamics.

We hope that in a later step, we can use different mechanisms of Biomolecular Simulation to describe the mechanism of inhibition of Entecavir Hydrate to the DPO, and if, within this process, Entecavir Hydrate will acquire different poses or torsions that can effectively interact within the active site of this enzyme.

We cannot discard using other ligands previously described in Table 1 as potential inhibitors of DPO, but the aim of this work was to select the best ligand, which complexew with this enzyme. Thus, Entecavir Hydrate can be tested *in vitro* and *in vivo* against *M. perniciosa* and, we hope it blocks replication processes of DPO in mitochondria.

Acknowledgements

The Graduate program in Biotechnology (PPGBiotec / UEFS-Fiocruz), for logistical support during the execution of this work; The Bioinformatics Laboratory (LAPEM–UEFS), for the infrastructure available to perform calculations; The State University of Feira de Santana.

References

Andrade BS, Taranto AG, Góes-Neto A, Duarte AA. **Comparative modeling of DNA and RNA polymerases from *Moniliophthora perniciosa* mitochondrial plasmid**. Theor Biol Med Model. 2009 Sep 10;6:22.

Aqvist J, Luzhkov VB, Brandsdal BO. **Ligand binding affinities from MD simulations**. Acc Chem Res. 2002 Jun;35(6):358-65.

Beckebaum S, Malagó M, Dirsch O, Cicinnati VR, Trippler M, Lampertico P, Lama N, Treichel U, Gerken G, Broelsch CE. **Efficacy of combined lamivudine and adefovir dipivoxil treatment for severe HBV graft reinfection after living donor liver transplantation**. Clin Transplant. 2003 Dec; 17(6):554-9.

D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvai, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2010), AMBER 10, University of California, San Francisco.

Domaoal RA, McMahon M, Thio CL, Bailey CM, Tirado-Rives J, Obikhod A, Detorio M, Rapp KL, Siliciano RF, Schinazi RF, Anderson KS. **Pre-steady-state kinetic studies establish entecavir 5'-triphosphate as a substrate for HIV-1 reverse transcriptase**. J Biol Chem. 2008 Feb 29;283(9):5452-9.

Fogolari F, Brigo A, Molinari H. **Protocol for MM/PBSA molecular dynamics simulations of proteins**. Biophys J. 2003 Jul;85(1):159-66

Formighieri EF, Tiburcio RA, Armas ED, Medrano FJ, Shimo H, Carels N, Góes-Neto A, Cotomacci C, Carazzolle MF, Sardinha-Pinto N, Thomazella DP, Rincones J, Digiampietri L, Carraro DM, Azeredo-Espin AM, Reis SF, Deckmann AC, Gramacho K, Gonçalves MS, Moura Neto JP, Barbosa LV, Meinhardt LW, Cascardo JC, Pereira GA. **The mitochondrial genome of the phytopathogenic basidiomycete Moniliophthora perniciosa is 109 kb in size and contains a stable integrated plasmid**. Mycol Res. 2008 Oct;112(Pt 10):1136-52.

Hou T, Wang J, Li Y, Wang W. **Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods II: The accuracy of ranking poses generated from docking**. J Comput Chem. 2011 Apr 15;32(5):866-77. doi: 10.1002/jcc.21666.

Hou T, Zhang W, Case DA, Wang W. **Characterization of domain-peptide interaction interface: a case study on the amphiphysin-1 SH3 domain**. J Mol Biol. 2008 Feb 29;376(4):1201-14.

Irwin JJ, Shoichet BK. **ZINC - A free database of commercially available compounds for virtual screening**. J Chem Inf Model. 2005 Jan-Feb;45(1):177-82.

Kollman P. **Free-Energy Calculations: Applications to Chemical and Biochemical Phenomena**. Chem. Rev. 1993. 93: 2395–2417.

Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE 3rd. **Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models**. Acc Chem Res. 2000 Dec;33(12):889-97.

Kuhn B, Gerber P, Schulz-Gasch T, Stahl M. **Validation and use of the MM-PBSA approach for drug discovery**. J Med Chem. 2005 Jun 16;48(12):4040-8.

Langley DR, Walsh AW, Baldick CJ, Eggers BJ, Rose RE, Levine SM, Kapur AJ, Colonno RJ, Tenney DJ. **Inhibition of hepatitis B virus polymerase by entecavir**. J Virol. 2007 Apr;81(8):3992-4001.

Lin PF, Nowicka-Sans B, Terry B, Zhang S, Wang C, Fan L, Dicker I, Gali V, Higley H, Parkin N, Tenney D, Krystal M, Colonno R. **Entecavir exhibits inhibitory activity against human immunodeficiency virus under conditions of reduced viral challenge**. Antimicrob Agents Chemother. 2008 May;52(5):1759-67.

Okimoto N, Futatsugi N, Fuji H, Suenaga A, Morimoto G, Yanai R, Ohno Y, Narumi T, Taiji M. **High-performance drug discovery: computational screening by combining docking and molecular dynamics simulations**. PLoS Comput Biol. 2009 Oct;5(10):e1000528.

Palumbo, E. **Pharmacotherapy of Chronic Hepatitis B with Entecavir**. Clinical Medicine: Therapeutics 2009:1 11–15

Sanner, M. F. **Python: A Programming Language for Software Integration and Development**. J. Mol. Graphics Mod. 1999. 17:57-61.

Seifer M, Hamatake RK, Colonno RJ, Standring DN. **In vitro inhibition of hepadnavirus polymerases by the triphosphates of BMS-200475 and lobucavir**. Antimicrob Agents Chemother. 1998 Dec;42(12):3200-8.

Schrödinger, LLC. **The PyMOL Molecular Graphics System**, Version 1.4.

Trott O, Olson AJ. **AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading**. J Comput Chem. 2010 Jan 30;31(2):455-61.

Walsh AW, Langley DR, Colonno RJ, Tenney DJ. **Mechanistic characterization and molecular modeling of hepatitis B virus polymerase resistance to entecavir**. PLoS One. 2010 Feb 12;5(2):e9195.

Wang W, Donini O, Reyes CM, Kollman PA. **Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions**. Annu Rev Biophys Biomol Struct. 2001;30:211-43.

Wang, J. M.; Hou, T. J.; Xu, X. J. **Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models**. Curr Comput-Aided Drug Des 2006, 2, 287.

Zaheer-ul-Haq, Halim SA, Uddin R, Madura JD. **Benchmarking docking and scoring protocol for the identification of potential acetylcholinesterase inhibitors**. J Mol Graph Model. 2010 Jun;28(8):870-82.

CHAPTER 3

The RNA polymerase of *Moniliophthora perniciosa* mitochondrial plasmid forms a highly stable complex with Rifampicin, a bacterial RNA polymerase inhibitor

Abstract

Moniliophthora perniciosa (Stahel) Aime & Phillips-Mora is the causal agent of witches' broom disease (WBD) in cacao (*Theobroma cacao*). After the mitochondrial genome of this fungus had been completed, an integrated linear-type plasmid that encodes viral-like RNA polymerases it was found. The structure of this polymerase was previously constructed using a homology modeling approach. Using a virtual screening process, accessing KEGG, PubChem and ZINC databases, we selected the eight best probable macrocyclic polymerase inhibitors to test against *M. perniciosa* RNA polymerase (RPO). Autodock Vina was used to perform docking calculations for each molecule, returning energy values in several ligand conformations. Pymol 1.4 was then used to check presence or absence of hydrogen, stereochemistry of chiral carbons, substructure, superstructure, number of rotable bonds, number of rings, number of donor groups, and hydrogen bonding receptors. As a result, we selected the Rifampicin, a bacterial RNA polymerase inhibitor, and then AMBER 10 was used to describe the behavior of the complex RPO-Rifampicin after a set of 3600 ps of simulation up to 300 K in water. This calculation returned a graph of Potential Energy during the time of simulation, which indicated that the ligand remains inside the active site after this time with a final energy of -462620.6888 Kcal/Mol. Therefore, we can accept that Rifampicin could be a good inhibitor to be tested *in vitro* and *in vivo* against *M. perniciosa*.

Keywords: *Moniliophthora perniciosa*, RNA polymerase, Rifampicin, Docking, MM/PBSA

3.1 Introduction

Moniliophthora perniciosa (Stahel) Aime & Phillips-Mora is the causal agent of witches' broom disease (WBD) in cacao (*Theobroma cacao*). After the sequencing of the mitochondrial genome (http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=4751&opt=organelle/) of this fungus, one found an integrated linear-type plasmid that encodes viral-like DNA and RNA polymerases (Formighieri et al., 2008).

RNA polymerase (RNAP) is the enzyme responsible for transcription and is the target, directly or indirectly, of most regulation of this process (Tuske et al., 2005). The RNA polymerase of *M. perniciosa* mitocondrial plasmid (RPO) is a 766 amino acid DNA-dependent RNA polymerase within single chain family of polymerases, which can be found in viruses and cellular organelles (Andrade et al., 2009). Its active site is distributed among two domains: Palm (Asp457 and Asp695) and Fingers (Tyr537 and Lys529), and they are involved with transcriptional processes (Andrade et al., 2009; Cheetham et al., 1999; Bonner et al., 1992). The transcription mechanism carried out by this enzyme shares several similarities with other multichain RNA polymerases (Andrade et al., 2009; Sousa et al., 1993), and it is possible that it has common inhibitors with other polymerases.

Rifampicin is a macrocyclic molecule of the ansamycin family. It contains a methyl-piperazinyliminomethyl side chain at position 3, a cyclopentyl-piperazinyliminomethyl side chain at position 3, and a cyclic spiro-piperidyl side chain at positions 3 and 4 (Ho et al., 2009, Wehrli, 1997). This drug has been used since 1968 against *Mycobacterium tuberculosis*, but is considered a broad-spectrum antibiotic (Mik et al., 2010; Floss and Yu, 2005; Campbell et al., 2001). This molecule has a high capacity to bind and inhibit bacterial DNA-dependent RNA polymerase (RNAP) through its specific interaction with the beta subunit of polymerase

(Aboshkiwa et al., 1995). The essential catalytic core RNAP is evolutionarily conserved among all cellular organisms (Campbell et al., 2001).

Virtual structure-based screening has become prominent in drug discovery, using protein targets (Hou et al., 2011; Okimoto et al, 2009). Several free ligands databases are widely available today. Searching for molecules that can complex with target proteins can either be done by text-based search (eg. KEGG and PubChem databases) as well as using a structure-activity relationship, available in ZINC (Irwin and Shoichet, 2005) and PubChem databases.

The one of the most important techniques for receptor-based drug design is Molecular Docking (Hou et al., 2011). Using crystallographic or modeled protein structures, molecular docking is often used for the screening of compound libraries, and predicting the conformation of a protein-ligand complex as well as calculating its binding affinity (Okimoto et al, 2009). Usually, docking programs such as Autodock Vina (Trott and Olson, 2010) generate multiple protein-ligand conformations by the sampling of the ligand's probable conformations in the binding pocket of the target protein, using a flexible ligand-rigid receptor docking (Hou et al., 2011; Trott and Olson, 2011). The use of scoring functions in docking calculations, by these programs, is an attempt to approximate the standard chemical potentials of the system (Trott and Olson, 2011). Autodock Vina uses a force-field-based scoring function approach to estimate the binding affinities by calculating the non-bonded interactions based on traditional force fields, identifying the correct binding pose of a ligand and ranking ligands using the predicted binding affinities (Hou et al., 2011; Trott and Olson, 2010, Wang et al., 2006; Wang et al., 2001). On the other hand, the problems of molecular docking as a screening tool have been widely discussed, since the scoring functions are usually inaccurate, neglect the solvent-related terms and protein flexibility is ignored (Okimoto et al, 2009). Coupled Molecular Docking and Molecular Dynamics is a good way to solve this problem, because it can treat both proteins and ligands in a

flexible manner, allowing the relaxation of the binding site around the ligand (Hou et al., 2010; Okimoto et al, 2009; Aqvist et al., 2002).

Molecular mechanics/Poisson Boltzmann surface area (MM/PBSA) combines molecular mechanics energy and implicit solvation models, and it is more rigorous than most empirical or knowledged-based scoring functions (Hou et al., 2011; Okimoto et al, 2009, Kollman et al., 2000). It allows for rigorous free-energy decomposition into contributions originating from different groups of atoms or types of interaction (Hou et al., 2008; Hou et al., 2011). In the MM/PB-SA method, the free energy is calculated using the snapshots of solute molecules obtained from explicit-solvent MD simulation (Okimoto et al, 2009, Kuhn et al., 2005).

The aim of this study was to search a series of likely molecules which can form complexes with the RPO from *M. perniciosa*, available at Kegg, PubChem and Zinc databases, and further selecting a potential inhibitor using a coupled Molecular Docking and Molecular Dynamics - MM / PBSA approach.

3.2 Methods

3.2.1 Ligand Searching

Initially, we made a search by keywords, by nucleoside molecules, in KEGG (http://www.genome.jp/kegg/) and PubChem (http://pubchem.ncbi.nlm.nih.gov/) databases. Then, only molecules described as inhibitors of RNA polymerases were selected. All 2D structures were copied in Similes format in order to be compared to other ZINC database (http://zinc.docking.org/) molecules. Furthermore, the 3D structures of these molecules were downloaded in mol2 and pdb formats to be used in a Virtual Screening process, which was carried out by Molecular Docking and Molecular Dynamics. According to a protocol described by Irwin and Schoichet (2005), molecules obtained from Zinc database were selected for comparison, following an interval between 95-99% similarity, with the structures found in

KEGG and PubChem databases. In addition, selected structures were downloaded in mol2 and pdb formats for subsequent Docking Studies and Molecular Dynamics.

3.2.2 Docking Studies

The downloaded structures from Kegg, Pubchem and Zinc databases were first checked in Pymol 1.4 (Schrödinger, LLC, 2011) to evaluate the presence or absence of hydrogen, stereochemistry of chiral carbons, substructure, superstructure, number of rotable bonds, number of rings, number of donor groups, and hydrogen bonding receptors.

The molecules of the ligands and the receptor (RPO) were prepared in Autodock Tools 1.5.6 (Sanner, 1999). All polar hydrogens for the receptor and Kollman United Atomic Charges were added and subsequently computed (Kollman, 1993). For all ligands we added polar and hydrogens and computed Gasteiger charges. The grid definition, adjusting to the RPO active site, was set up manually by following the recommendations of the program manual (Trott and Olson, 2010; Sanner, 1999). Then, the structures of the ligand and receptor were saved in pqbqt format to be used in docking calculations. Autodock Vina was used to perform Docking Scoring for each ligand-receptor complexes (Trott and Olson, 2010). Before running each Docking calculations, a configuration file was generated with grid size and coordinates information, as well as indicating ligand and receptor files. The reports (log files) for each calculation were analyzed in order to obtain free-energy (Kcal/mol) values of each ligand conformations with its respective complex. In addition, we used Pymol to verify the number of hydrogen bonds and non-covalent interactions, between each different ligand conformations and catalytic residues of RPO, which are involved with recognition and polymerization mechanisms. As a way to optimize the choosing of an ideal complex, we selected just one ligand, which best fits in the RPO active site, considering all stereochemical aspects previously evaluated and the free-energy results.

3.2.3 Molecular Dynamics of Complex

In this study we used MM/PBSA protocol to calculate the affinity and stability in the interaction of ligand-receptor RPO complex, using AMBER 10 (Case et al., 2008) package. Initially, we used Antechamber program to make ff99 force field recognizes the atom types of both ligand and receptor, and to avoid errors during the calculations.

Tleap was used to neutralize charges (ff99 force field) and RPO–ligand complex was immersed in a rectangular box of TIP3P water molecules. Following the protocol, we use Sander to carried out a Molecular Dynamics (MD) Equilibrium, which was restricted to a region of the protein that contains the active site (amino acids 457-695), according to the following parameters: 1000 cycles of steepest descent and 1000 cycles of conjugate gradient minimization, heating MD for 50 picoseconds (ps), density equilibrium for 50 ps, followed by a Equilibrium Dynamics for 500 ps at constant pressure and 300 K temperature. After the equilibration of the system we followed with MM/PBSA protocol (Hou et al., 2011; Case et al., 2008; Fogolari et al, 2003). Then, we simulated a total of 3000 ps production step Molecular Dynamics, divided in 3 sets of 1000 ps (prod1, prod2 and prod3), saving coordinates every 10 ps. Furthermore, we used the mm_pbsa.pl script to extract snapshots (without the water) from our production runs and get their trajectories. In addition, we checked complex stability by plotting a Potential Energy x Time (ps) graph, from all production steps. In a last step, we use Ambpdb to generate a pdb file of the complex, after the last stage of the Molecular Dynamics, and this structure was analyzed in Pymol to verify if the ligand remained in the active site after the whole process.

3.3 Results and Discussion

3.3.1 Structures and binding energies of RPO complexes from AutoDock Vina

After research conducted in KEGG, PubChem and ZINC databases, we selected 8 structures that could exhibit a good interaction with RPO, and with different classifications. A reliable

prediction of complex interactions is essential for selecting a potential ligand in virtual screening methodologies, and that requires a proper fit tool that is capable of generating energy assessment of the binding protein, indicating the quality of interaction (Zaheer-ul -Haq et al., 2010). The results of Molecular Docking with AutoDock Vina for different ligand-RPO complexes are presented for the dominating configuration with minimum binding free energy (ΔG) in Table 1. Docking scores returned by Autodock Vina indicates Rifampicin ligand as Top-ranked solution to be a good RPO inhibitor. In addition, we evaluated other features for all screened ligands such as H-bond donor and H-bond acceptor, as well as the capacity for at least one conformation of each ligand to bind to amino acids from active site pocket of RPO in the complex (Andrade et al., 2009). Among the molecules studied, Rifampicin is which best binds to the amino acids of RPO active site and had high affinity energy (Table 1) in docking calculations, for all docking poses. In figure 1, we show that Rifampicin fits inside the hydrophobic pocket of RPO. This molecule forms several hydrogen bonds with amino acids from active site region: one with Ser 397, one with Arg 404, three with Asp 457, one with Tyr 494 and two with Arg 525, as can be seen in figure 2. According to some authors, the amino acid Asp 457 from RPO active site of several organisms is involved with the transcriptional processes (Bonner et al., 1992; Cheetham et al., 1999; Holtje et al., 2003; Andrade et al., 2009). In *Escherichia coli,* Rifampicin binds in a pocket of the RNAP b subunit deep within the DNA/RNA channel and blocks the RNA exit pathway (Campbell et al., 2005). In other study, Campbell et al. (2001) describes a 3.3 Å crystal structure of *Thermus aquaticus* RNA polymerase complexed with Rifampicin, and after biochemical experiments, their results indicate that the predominant effect of Rif is to directly block the path of the elongating RNA transcript at the 5’ end when the transcript becomes either 2 or 3 nucleotides in length (Korzheva et al., 2000).

The use of RNA polymerases as molecular targets for virtual screening is not restricted to prokaryotes. The use of an RNA-dependent RNA polymerase (RdRp) is an attractive target for anti-HCV agents (Lee et al, 2010). However, we could not find any study that specifically dealt with the use of inhibitors of fungal polymerases or polymerases encoded by mitochondrial genes. On the other hand, many authors reported the fact that all cellular RNA polymerases are relatively conserved in their amino acids sequences and catalytic mechanism (Ho et al., 2009; Campbell et al., 2005; Temiakov et al., 2005; Tuske et al., 2005; Campbell et al., 2001),. Thus, in general, the same class of RNA polymerase inhibitors (macrocyclic) acts on different groups of organisms. In addition, it is probable that Rifampicin can act in *vitro* and *in vivo*, inhibiting mitochondrial transcriptional process carried out by RPO and, somehow, blocking the mitochondrial metabolism of *M. perniciosa*.

Table 3.1. Potential RPO inhibitors selected from the KEGG, PubChem and ZINC databases used in docking studies.

Molecule	Affinity (kcal/mol)	H-Bond Donor	H-Bond Acceptor
Rifampicin	-10.4	6	15
Rifapentine	-9	6	15
Rifabutin Mycobutin	-8	5	14
Zinc5220312	-5.2	5	7
Zinc5124992	-5	2	8
Zinc5220316	-4.7	5	7
Zinc5220339	-5.1	5	7
Zinc22173122	-5.5	4	4

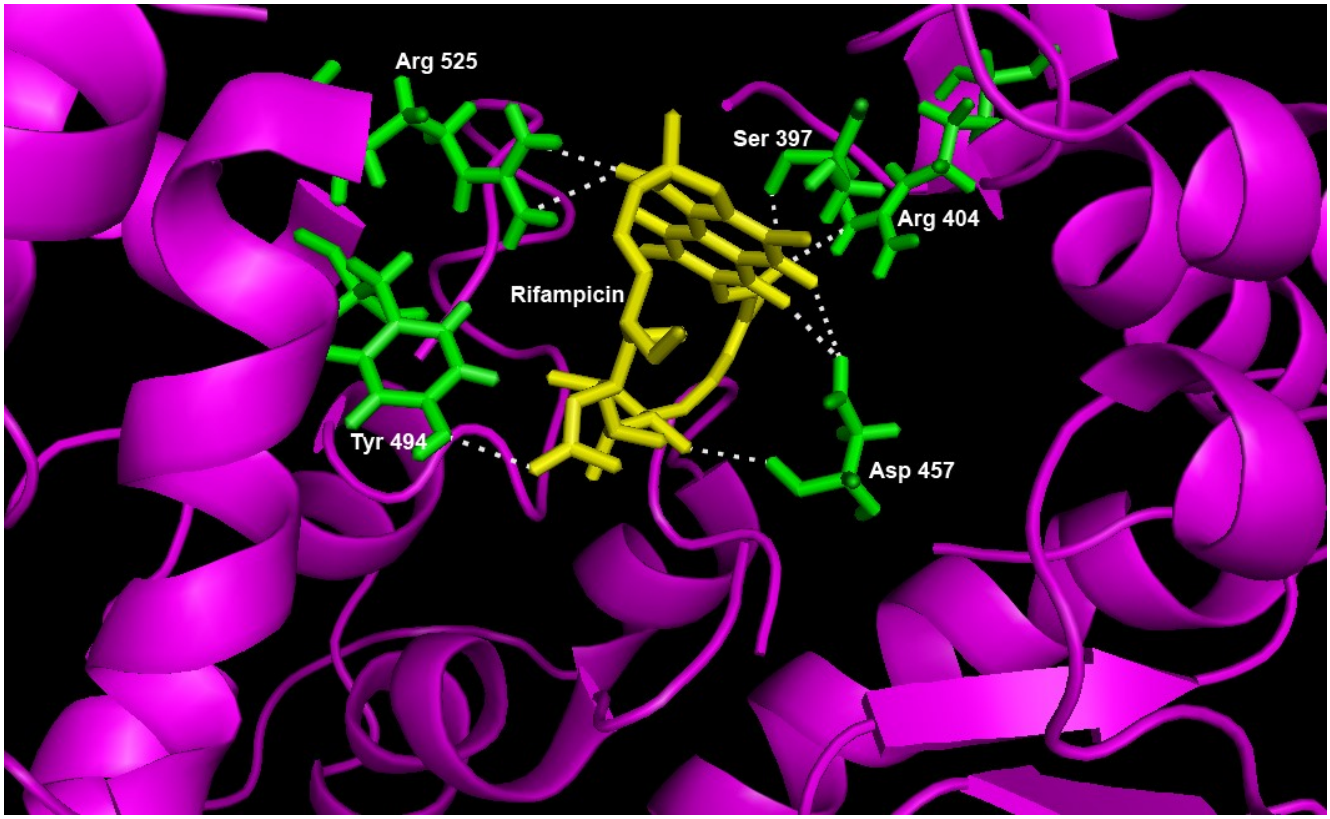


Fig. 3.1. Hydrogen bonds between Rifampicin and Ser 397, Arg 404, Asp 457, Tyr 494 and Arg 525 Asp in RPO active site.

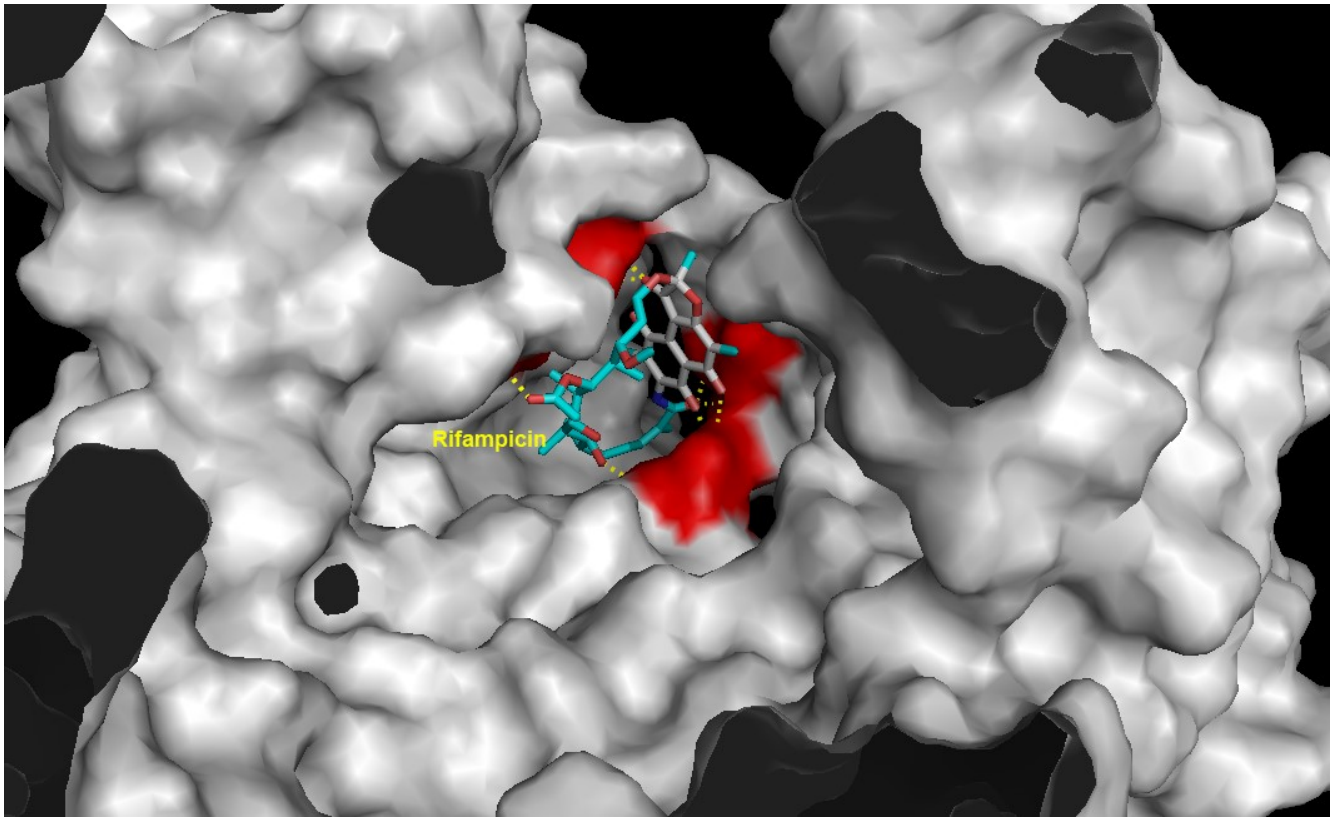


Fig. 3.2. RPO active pocket showing Rifampicin interaction. The catalytic amino acids are in the red region.

3.3.2 Molecular Dynamics MM/PBSA of RPO-Rifampicin complex

Using a Molecular Dynamics approach, we analyzed the performance and stability of RPO-Rifampicin complex. Then, we evaluated the potential energy complex, during the simulation process, and its final energy. As you can see in Figure 3, the graph shows that above 600 ps simulation the complex have already reached a potential energy value near the minimum, and it maintains until the end of simulation process. In addition, the potential energy during the plateau shows that the structure of this complex is perfectly feasible to exist. The final energy reached at exactly 3600 ps was -462620.6888 Kcal/Mol. We also consider the RMSD generated during the heating processes and production, but we noticed that the value has converged before 1000 ps simulation.

The pdb of the complex, generated after molecular dynamics, showed that Rifampicin remains within the active site of RPO after 3600 ps of simulation. Furthermore, we can infer that this time of simulation was sufficient to show that Rifampicin could block *M. perniciosa* RPO activity.

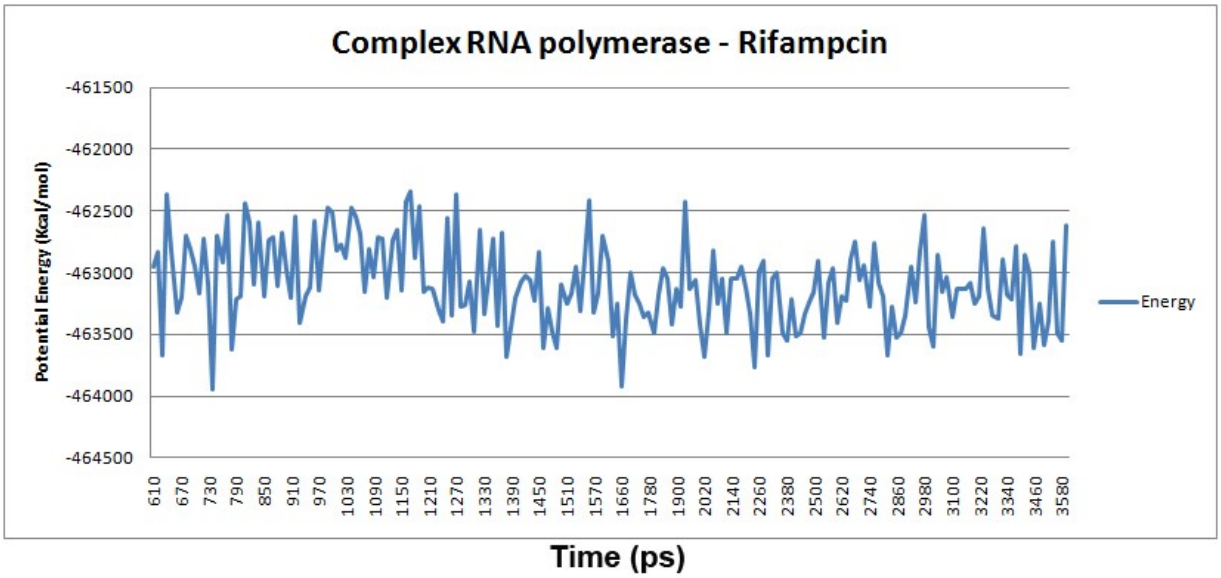


Fig. 3.3. Graph of potential energy of RPO-Rifampicin complex during 3.6 ns Molecular Dynamics simulation

3.4 Conclusions

In this article we describe the selection of one potential inhibitor for RNA polymerase of *M. perniciosa* mitochondrial plasmid, among eight molecules found in public databases, using a virtual screening approach. Althogh Rifampicin is a bacterial RNAPs inhibitor, this molecule forms a very stable complex with RPO, and this may be related to the fact that this type of enzyme is highly conserved among different organisms. The Rifampicin complexes exactly with the amino acids of active pocket, which are directly involved with the transcriptional process of RPO. This integration remains stable during all 3600 ps Molecular Dynamics.

In a further step, we will analyze different mechanisms of Biomolecular Simulation to describe the mechanism of RPO inhibition by Rifampicin, as well as if, in this process, Rifampicin will acquire different conformations that can effectively interact within the active site of this enzyme. We cannot discard using other ligands previously described in Table 1 as potential inhibitors of RPO, as well as other described as macrocyclic polymerase inhibitors. Additionally, Rifampicin was selected because it forms the best ligand-complex interaction. Then, in a future study, we will perform biochemical tests, *in vitro* and *in vivo*, in order to verify if our selected inhibitor can effectively acts against *M. perniciosa*, making replication processes of RPO unfeasible.

Acknowledgements

The Graduate program in Biotechnology (PPGBiotec / UEFS-Fiocruz), for logistical support during the execution of this work; The Bioinformatics Laboratory (LAPEM–UEFS), for the infrastructure available to perform calculations; The State University of Feira de Santana.

References

Aboshkiwa M, Rowland G, Coleman G. **Nucleotide sequence of the *Staphylococcus aureus* RNA polymerase rpoB gene and comparison of its predicted amino acid sequence with those of other bacteria**. Biochim Biophys Acta. 1995 May 17;1262(1):73-8.

Andrade BS, Taranto AG, Góes-Neto A, Duarte AA. **Comparative modeling of DNA and RNA polymerases from *Moniliophthora perniciosa* mitochondrial plasmid**. Theor Biol Med Model. 2009 Sep 10;6:22.

Aqvist J, Luzhkov VB, Brandsdal BO. **Ligand binding affinities from MD simulations**. Acc Chem Res. 2002 Jun;35(6):358-65.

Bonner G, Patra D, Lafer EM, Sousa R. **Mutations in T7 RNA polymerase that support the proposal for a common polymerase active site structure**. EMBO J. 1992 Oct;11(10):3767-75.

Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA. **Structural mechanism for rifampicin inhibition of bacterial RNA polymerase**. Cell. 2001 Mar 23;104(6):901-12.

Campbell EA, Pavlova O, Zenkin N, Leon F, Irschik H, Jansen R, Severinov K, Darst SA. **Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase**.

Cheetham GM, Jeruzalmi D, Steitz TA. **Structural basis for initiation of transcription from an RNA polymerase-promoter complex**. Nature. 1999 May 6;399(6731):80-3. Erratum in: Nature 1999 Jul 1;400(6739):89.

D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvai, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P.A. Kollman (2010), AMBER 10, University of California, San Francisco.

Floss HG, Yu TW. **Rifamycin-mode of action, resistance, and biosynthesis**. Chem Rev. 2005 Feb;105(2):621-32.

Fogolari F, Brigo A, Molinari H. **Protocol for MM/PBSA molecular dynamics simulations of proteins**. Biophys J. 2003 Jul;85(1):159-66

Formighieri EF, Tiburcio RA, Armas ED, Medrano FJ, Shimo H, Carels N, Góes-Neto A, Cotomacci C, Carazzolle MF, Sardinha-Pinto N, Thomazella DP, Rincones J, Digiampietri L, Carraro DM, Azeredo-Espin AM, Reis SF, Deckmann AC, Gramacho K, Gonçalves MS, Moura Neto JP, Barbosa LV, Meinhardt LW, Cascardo JC, Pereira GA. **The mitochondrial genome of**

the phytopathogenic basidiomycete Moniliophthora perniciosa is 109 kb in size and contains a stable integrated plasmid. Mycol Res. 2008 Oct;112(Pt 10):1136-52.

Ho MX, Hudson BP, Das K, Arnold E, Ebright RH . **Structures of RNA polymerase-antibiotic complexes**. Curr Opin Struct Biol. 2009 Dec;19(6):715-23.

Holtje HD, Sippl W, Rognan D, Folkers G: **Molecular Modeling: Basic principles and applications**. WILEY-VCH; 2003.

Hou T, Wang J, Li Y, Wang W. **Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods II The accuracy of ranking poses generated from docking**. J Comput Chem. 2011 Apr 15;32(5):866-77. doi: 10.1002/jcc.21666.

Hou T, Zhang W, Case DA, Wang W. **Characterization of domain-peptide interaction interface: a case study on the amphiphysin-1 SH3 domain**. J Mol Biol. 2008 Feb 29;376(4):1201-14.

Irwin JJ, Shoichet BK. ZINC - **A free database of commercially available compounds for virtual screening**. J Chem Inf Model. 2005 Jan-Feb;45(1):177-82.

Kollman P. **Free-Energy Calculations: Applications to Chemical and Biochemical Phenomena**. Chem. Rev. 1993. 93: 2395–2417.

Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE 3rd. **Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models**. Acc Chem Res. 2000 Dec;33(12):889-97.

Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A, Darst SA. **A structural model of transcription elongation**. Science. 2000 Jul 28;289(5479):619-25.

Kuhn B, Gerber P, Schulz-Gasch T, Stahl M. **Validation and use of the MM-PBSA approach for drug discovery**. J Med Chem. 2005 Jun 16;48(12):4040-8.

Lee JC, Tseng CK, Chen KJ, Huang KJ, Lin CK, Lin YT. **A cell-based reporter assay for inhibitor screening of hepatitis C virus RNA-dependent RNA polymerase**. Anal Biochem. 2010 Aug;403(1-2):52-62.

Mick V, Domínguez MA, Tubau F, Liñares J, Pujol M, Martín R. **Molecular characterization of resistance to Rifampicin in an emerging hospital-associated Methicillin-resistant Staphylococcus aureus clone ST228, Spain**. BMC Microbiol. 2010 Mar 4;10:68.

Okimoto N, Futatsugi N, Fuji H, Suenaga A, Morimoto G, Yanai R, Ohno Y, Narumi T, Taiji M. **High-performance drug discovery: computational screening by combining docking and molecular dynamics simulations**. PLoS Comput Biol. 2009 Oct;5(10):e1000528.

Sanner, M. F. **Python: A Programming Language for Software Integration and Development**. J. Mol. Graphics Mod. 1999. 17:57-61.

Schrödinger, LLC. **The PyMOL Molecular Graphics System**, Version 1.4.

Sousa R, Chung YJ, Rose JP, Wang BC. **Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution**. Nature. 1993 Aug 12;364(6438):593-9.

Temiakov D, Zenkin N, Vassylyeva MN, Perederina A, Tahirov TH, Kashkina E, Savkina M, Zorov S, Nikiforov V, Igarashi N, Matsugaki N, Wakatsuki S, Severinov K, Vassylyev DG. **Structural basis of transcription inhibition by antibiotic streptolydigin**. Mol Cell. 2005 Sep 2;19(5):655-66.

Trott O, Olson AJ. **AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading**. J Comput Chem. 2010 Jan 30;31(2):455-61.

Tuske S, Sarafianos SG, Wang X, Hudson B, Sineva E, Mukhopadhyay J, Birktoft JJ, Leroy O, Ismail S, Clark AD Jr, Dharia C, Napoli A, Laptenko O, Lee J, Borukhov S, Ebright RH, Arnold E. **Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation**.Cell. 2005 Aug 26;122(4):541-52.

Wang W, Donini O, Reyes CM, Kollman PA. **Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions**. Annu Rev Biophys Biomol Struct. 2001;30:211-43.

Wang, J. M.; Hou, T. J.; Xu, X. J. **Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum**. Models Curr Comput-Aided Drug Des 2006, 2, 287.

Wehrli W. **Ansamycins. Chemistry, biosynthesis and biological activity**. Top Curr Chem. 1977;72:21-49.

Zaheer-ul-Haq, Halim SA, Uddin R, Madura JD. **Benchmarking docking and scoring protocol for the identification of potential acetylcholinesterase inhibitors**. J Mol Graph Model. 2010 Jun;28(8):870-82.

CHAPTER 4

The activity of DNA and RNA polymerases from *Moniliophthora perniciosa* mitochondrial plasmid may reveal a self-defense mechanism against oxidative stress

Abstract

Moniliophthora perniciosa (Stahel) Aime and Philips-Mora is a hemibiotrophic basidiomycete (Agaricales – Marasmiaceae), which causes the witches' broom in cocoa (*Theobroma cacao* L.). This pathogen carries a stable integrated invertron-type linear plasmid into its mitochondrial genome, which encodes viral-like DNA and RNA polymerases that can be related to fungal senescence and longevity. After culturing the fungus and obtaining their various stages of development, in triplicate, we carried out total RNA extraction and subsequent cDNA synthesis. In order to analyze DNA and RNA polymerase expression levels, we performed a RT-qPCR for distinct fungal developmental phases. Our results show that DNA and RNA polymerase gene expressions in primordium phase of *M. perniciosa* are related to a potential defense mechanism against *T. cacao* oxidative attack.

Keywords: *M. perniciosa*, Oxidative Stress, Mitochondrial Plasmid, Polymerases

4.1 Introduction

Moniliophthora perniciosa (Stahel) Aime and Philips-Mora (2005) is a hemibiotrophic basidiomycete (Agaricales – Marasmiaceae), endemic to the Amazon basin, which causes the witches' broom disease (WBD) in cocoa (*Theobroma cacao* L.). It is the only cocoa pathogen that grows along with the plant (Aime and Philips-Mora, 2005; Purdy; Schmidt, 1996). This disease is one of the most important phytopathological problems of cacao producing areas of the American continent, and has decimated the Brazilian cacao industry in the last two decades (Griffith et al., 2003). After infection, the pathogen establishes a biotrophic relationship with its host and the monokaryotic mycelium progresses intercellularly in the plant tissues. Four to six weeks later, the hyphae become dikaryotic, develop clamp connections and invade the cells (intracellular mycelium): this stage corresponds to the saprophytic phase of the fungus and ends with the basidiomata and basidiospores formation (Meinhardt et al., 2008).

Besides a series of macroscopic and microscopic changes that occur in *M. perniciosa*, the stages of development show changes in color. Mycelial mats turn light-yellow four days after exposure to air and water, changing to reddish-pink after ten days, and finally becoming dark-reddish pink until the onset of basidiomata development (Pires et al., 2009). The first signal of primordial development is probably the appearance of primary hyphal nodules as well as internal local aggregations on dark pink-reddish mycelium (Pires et al., 2009),

In 2008, Formighieri et al. published the *M. perniciosa* mitochondrial genome and described a stable integrated invertron-type linear plasmid into this genome. This plasmid have the same features than other mitochondrial plasmids previously described (Formighieri et al., 2008), such as the presence of rather large terminal inverted repeats and two large open reading frames on opposite strands (Kempken, 1994). Moreover, the *M. perniciosa* linear-type plasmid also encodes viral-like DNA and RNA polymerases (Formighieri et al., 2008; Poggeler and

Kempken, 2004). Linear plasmids are commonly related to fungal senescence and longevity (Poggeler and Kempken, 2004) due to mitochondrial instability and plasmid insertions into the mitochondrial genome (Bertrand et al. 1985, 1986; Bertrand H., 2000).

Based on *M. perniciosa* mitochondrial genome sequence, some authors described differential expression of mitochondrial genes – mainly associated with oxidative stress and glucose depletion – during the biotrophic phase of the fungus in the plant (Scarpari et al. 2005; Rincones et al., 2008; Pires et al., 2009). In particular, Pires et al. (2009) provided the first description of *M. perniciosa* primordia and basidiomata development, dividing it in six stages according to the mycelium color or the developmental phases: white, yellow, pink, dark pink, primordium and basidiome.

In our study, we analyze, by RT-qPCR, the expression of DNA and RNA polymerase genes of mitochondrial plasmid identified during the fungal development from white mycelium phase to basidiome production. The relation between DNA and RNA polymerase gene expression and *M. perniciosa* potential defense mechanism against *T. cacao* oxidative attack during the infection is then discussed.

4.2 Methods

4.2.1 *Moniliophthora perniciosa* culture

M. perniciosa (isolate 565 VA4 CEPEC / CEPLAC) was grown in artificial system as described by Niella et al. (1999). The fungus was inoculated in triplicate, on plates containing PDA medium (20% potato, 2% dextrose, 2%agar) and incubated at 25°C for 10 to 15 days. After this period, mycelium were inoculated in a solid bran-based medium containing dry broom (37%), rolled oats (9.8%), calcium sulfate (1.5%) and H_2O enough to achieve the saturation point (51.7%). The culture was maintained at 21-25°C under 100% of humidity and 12 h light regime, until the mycelium covered the entire medium (approximately 15 days). The solid bran-based

medium was hung in pots previously autoclaved (adaptation of the broom-maker), under the same conditions of temperature, humidity and light that had been previously described. During the cultivation, the medium were watered two or three times per week on alternate days, using autoclaved water and a sterile syringe (about 10 mL of water to each side of the medium). Before each irrigation, it was necessary to remove excess water, and to carry out a water stress (withdrawal of water for 15 or 20 days) prior to the production of basidiomata. Eight different stages of development of the fungus were studied: (i) white phase: the initial stage of cultivation when the medium had a white mycelium ii) yellow phase, when the mycelium had the first change in color to become yellow, iii) pink phase: when a second color change occur with mycelium featuring a pink and prior to water stress, iv) dark pink: soon after the water stress, v) primordium: when the medium showed the formation of primordia, vi) basidiomata, when the fruiting bodies were fully developed, vii) primary mycelium: monokaryotic growth phase, without clamp connections, and viii) secondary mycelium: dikaryotic growth phase, with clamp connections. These last two phases were obtained by cultivating mycelial starter cultures from the culture collection of the Cocoa Research Center (CEPEC, Ilhéus, Bahia, Brazil) on PDA (Potato Dextrose Agar) for three weeks in the dark, at room temperature. The mycelium was ground in liquid nitrogen and kept at -80°C until use.

4.2.2 RNA extraction and cDNA synthesis

The total RNA extraction was performed from 500 mg of ground material using the RNeasy Midi Kit according to the manufacturer's recommendations (Qiagen). Total RNA was treated with DNase I (Fermentas) and the cDNAs were synthesized using reverse transcriptase RevertAidTMH Minus M-MuLV according to manufacturer's recommendations (Fermentas). Briefly, the reaction containing 10 mL of total RNA, 2.5 mmol.L^{-1} of each dNTP, 20 U of RNase

inhibitor, 0.5 µg of oligo(dT) and 200 U of RevertAidTMH Minus M-MuLV was incubated at 42°C for 1h. The reverse transcriptase was inactivated at 70°C for 10 min.

4.2.3 Real-time qPCR and data analysis

Both DNA and RNA polymerase primers were designed using the IDT Scitools - Realtime PCR (http://www.idtdna.com/Scitools/Applications/RealTimePCR/) (Table 1). Quantitative PCR was performed using SYBRGreen® (Invitrogen) for the detection of fluorescence during amplification, and assays were performed on an ABI PRISM 7500 Sequence Detection System (SDS) coupled to the ABI PRISM 7500 SDS software (Applied Biosystems, Foster City, USA), using standard settings. A 20 µL reaction containing 10 µL of single stranded cDNA, 2 µL of SYBRGreen 1X (Invitrogen), PCR buffer 1X, 200 mM of dNTPs, 3 mM of $MgCl_2$, 1/2 50X Rox, 200 nM of each primer was submitted to qPCR. The thermal cycling conditions were 50°C for 2 min, then 94°C for 10 min, followed by 40 cycles of 94°C for 45 s, 61°C for 35 s and 72°C for 35 s. A dissociation analysis was conducted for all the amplifications to investigate the formation of dimers of primers and hairpins. Melting temperatures of the fragments were determined according to the manufacturer's protocol. No-template reactions were included as negative controls in every plate. The results obtained with the Detection Software (Applied Biosystems, Foster City, USA) were imported into Microsoft Excel.

4.3 Results and discussion

Analysis of the expression of DNA and RNA polymerases mitochondrial plasmid of *M. perniciosa* revealed that these genes were expressed in all phases of the life cycle of the fungus. In the phase dark pink, only the DNA polymerase is expressed. Although there is no prior information about the probable function of these enzymes in the metabolism of *M. perniciosa* (Formighieri, 2009), one knows that these molecules act in the replication mechanism and expression of the mitochondrial genome (Kennel and Cohen, 2003; Kempken, 1994), by

presenting some degree of structural and sequence identity with other fungal plasmid polymerases (Andrade et al., 2009). Thus, it is possible that the activity of these two enzymes is related to the process of increasing the survival of the fungus, as described for pAL2 plasmid of *Podospora anserina* (Maas et al. 2007; Hermanns et al, 1994). In *Neurospora crassa* it was observed that the portion that encodes a DNA polymerase with a functional promoter in pLABELLE plasmid is close to mtDNA promoter, indicating that this polymerase was maintained throughout the process of evolution by selective pressure, and maybe contribute functionally to the mtDNA (Cahan and Kennell, 2005).

The relative expression between the DNA and RNA polymerases at different stages of the fungus is approximately six times higher for DNA polymerase and almost three times higher for RNA polymerase, in primordium (Figure 1). In *M. perniciosa*, the stage of primordium occurs after dark pink stage, when the mycelium hyphae show nodular aggregates in some regions of mycelium (Pires et al., 2009). Developmentally regulated genes related to primordium and basidiomata development have been identified for some Basidiomycota such as *Agaricus bisporus* (De Groot et al., 1997), *Coprinopsis cinerea* (Kues, 2000) and *Pleurotus ostreatus* (Lee et al., 2002), among others. In addition, the rapid increase of fully or partially sequenced genomes and ESTs from fungi already available in databases allows the *in silico* identification of genes possibly involved in these processes (Soanes and Talbot, 2006; Nowrousian and Kück, 2006).

Scarpari et al. (2004) published a detailed biochemical study of *T. cacao* and *M. perniciosa* interaction, with a systematic analysis of the changes in the contents of soluble sugars, amino acids, alkaloids, ethylene, phenolics, tannins, flavonoids, pigments, malondialdehyde, glycerol, and fatty acids in cacao shoots during the development of the infected plant. The alterations in the lipid metabolism, during the infection, showed a highest glycerol content and lipid

peroxidation by the amount of malondialdehyde. These metabolism changes are a common feature of senescence/PCD in plants, and they are involved in the generation of Reative Oxigen Species (ROS) (Blokhina et al., 2003). ROS production (Gratao et al. 2005), is related to lipid peroxidation (Scarpari et al. 2005) and the degradation of calcium oxalate crystals (Ceita et al. 2007), likely triggering mitochondrial recombination events and formation of primordia. It is probable that these events have as mainline effect the change in energy metabolic process of the fungus, with increased expression (up regulation) of cytochrome p450 monooxygenase (Pires et al. 2009). Therefore, it is quite possible that changes in patterns of relative expression of DNA and RNA polymerases inserted in the mitochondrial genome of *M. perniciosa* can be influencing the process of copying and transcription of several mitochondrial genes related to the process of plant-pathogen interaction and, in some way, triggering a mechanism of self-defense metabolites produced by the fungus against *T. cacao*.

Table 4.1. Primers used in Real Time PCR reactions to detect *M. perniciosa* DPO and RPO expression levels. F, forward primer; R, reverse primer. All primers were specifically designed for this study[a]

Primer	Sequence (5'-')	Tm (°C)	Amplicon (bp)
MDPO-F	CAC Tgg Agg Tag TgT ggA TgT TTAC	63	70
MDPO-R	TTg gTT CAC CTg TTg gCA TA	63	
MRPO-F	TTg AAA Aag AAA AgC AAT ACT CAT TTg A	60	247
MRPO-R	AgA TgT TTC TTA TTg TTT gAT CCA CC	60	
ACT1-F*	CCC TTC TAT CgT Cgg TCg T	81	-
ACT1-R*	Agg ATA CCA CgC TTg gAT Tg	81	

[a]All primer sequences are in the 5′–3′ direction. Asterisk (*) indicates transcript used for normalization.

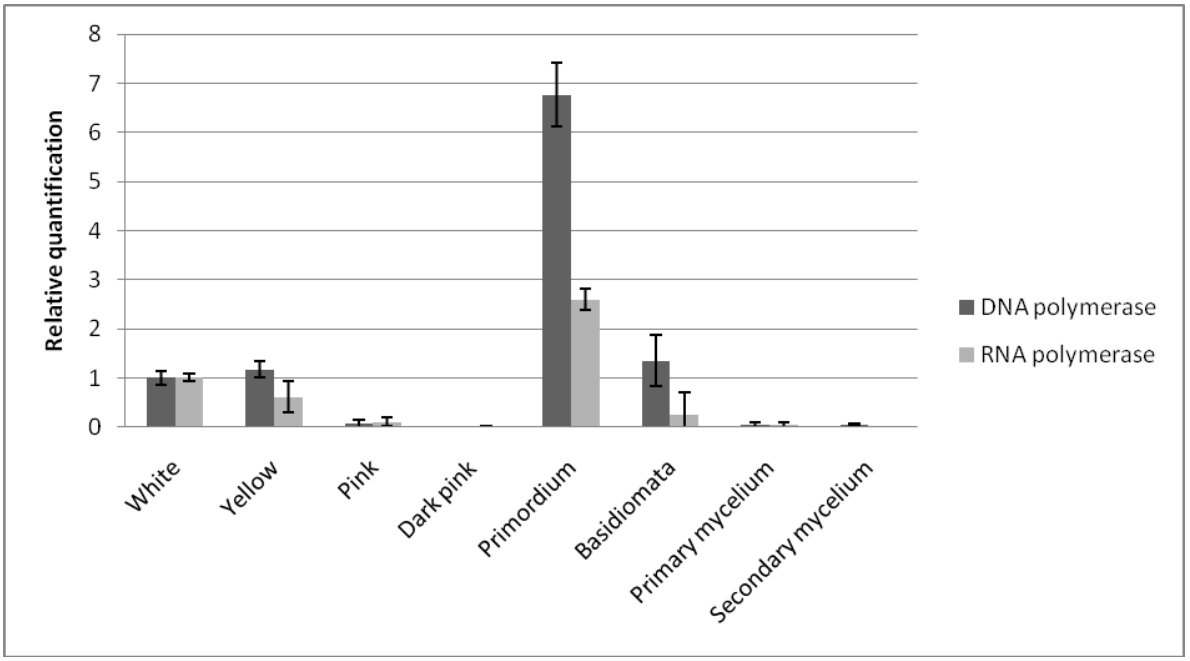


Figure 4.1. RT-qPCR of DPO and RPO genes expressed in different phases during the culture of *M. perniciosa*

4.4 Conclusions

Relative expression of DNA and RNA polymerases mitochondrial plasmid of *M. perniciosa* seems to follow the development process of the fungus during its process of infection in *T. cacao*. Although this work has examined the expression of these genes in different developmental stage of the fungus, we can see that primordium stage presented a clear and statistically significant difference in expression of these polymerases.

The relative increase of polymerases expression in primordium matches with the oxidative damage triggered by the plant. This leads us to relate this process with a defense mechanism of the fungus, when several substances produced by host plant are degraded by the mitochondria of the fungus.

Despite the mitochondrial metabolism is highly dependent on the nuclear genome, it is possible that the polymerases encoded by the integrated plasmid into the mitochondrial genome of *M. perniciosa* are functional on its metabolism. The DNA polymerase present relative expression greater than the RNA polymerase, in the primordium stage, and it is probably related to the

mechanism of replication of mtDNA, in order to prepare the fungus to basidiomata formation, as well as being involved in processes of repairing mtDNA affected by oxidative stress triggered by the plant. Although exhibiting a lower rate of expression of the DNA polymerase, it is likely that RNA polymerase is playing a role in gene transcription of mitochondrial metabolism, increasing the levels of cytochrome oxidase expression (up-regulation) and, thus, helping the maintenance of energy metabolism, s well as preparing the fungus to the process of defense against *T. cacao*. This information may be useful for further studies towards a more complete understanding of the mitochondrial metabolism of *M. perniciosa*, and further environmental controls leading to basidiomata initiation. In addition, new strategies can be provided for an enhanced control of this phytopathogen and for a successful monitoring of witches' broom disease in *T. cacao*.

Acknowledgements
State University of Feira de Santana (UEFS), and The Graduate Program in Biotechnology (PPGBiotec – UEFS/Fiocruz).

References

Aime MC, Phillips-Mora W. **The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae**. Mycologia. 2005 Sep-Oct;97(5):1012-22.

Andrade BS, Taranto AG, Góes-Neto A, Duarte AA. **Comparative modeling of DNA and RNA polymerases from *Moniliophthora perniciosa* mitochondrial plasmid**. Theor Biol Med Model. 2009 Sep 10;6:22.

Bertrand H, 2000. **Role of mitochondrial DNA in the senescence and hypovirulence of fungi and potential for plant disease control**. Annual Review of Phytopathology 38: 397–422.

Bertrand H, Chan BS, Griffiths AJ. **Insertion of a foreign nucleotide sequence into mitochondrial DNA causes senescence in *Neurospora intermedia***. Cell. 1985 Jul;41(3):877-84.

Bertrand H, Griffiths AJ, Court DA, Cheng CK, 1986. **An extrachromosomal plasmid is the etiological precursor of kalDNA insertion sequences in the mitochondrial chromosome of senescent *Neurospora***. Cell 47: 829–837.

Blokhina O, Virolainen E, Fagerstedt KV. **Antioxidants, oxidative damage and oxygen deprivation stress: a review**. Ann Bot. 2003 Jan;91 Spec No:179-94.

Cahan P, Kennell JC. **Identification and distribution of sequences having similarity to mitochondrial plasmids in mitochondrial genomes of filamentous fungi**. Mol Genet Genomics. 2005 Jul;273(6):462-73.

Ceita GO, Macedo JNA, Santos TB, Alemanno L, Gesteira AS, Micheli F, Mariano AC, Gramacho KP, Silva DC, Meinhardt L, Mazzafera P, Pereira GAG, Cascardo JCM. **Involvement of calcium oxalate degradation during programmed cell death in *Theobroma* cacao tissues triggered by the hemibiotrophic fungus *Moniliophthora perniciosa***. Plant Science. 2007; 173:106–117.

De Groot PW, Schaap PJ, Van Griensven LJ, Visser J. **Isolation of developmentally regulated genes from the edible mushroom *Agaricus bisporus***. Microbiology. 1997 Jun; 143 (Pt 6):1993-2001.

Gratão PI, Polle A, Lea PJ, Azevedo RA. **Making the life of heavy metal-stress plants a little easier**. Func. Plant Biol. 2005; 32:481-494.

Griffith GW, Nicholson J, Nenninger A, Birch RN. **Witches' brooms and frosty pods: two major pathogens of cacao**. N Z J Bot. 2003; 41:423-435.

Hermanns J, Asseburg A, Osiewacz HD. **Evidence for a life span-prolonging effect of a linear plasmid in a longevity mutant of *Podospora anserina***. Mol Gen Genet. 1994 May 10;243(3):297-307.

Kempken F, Hermanns J, Osiewacz HD. **Evolution of linear plasmids**. J Mol Evol. 1992 Dec;35(6):502-13.

KEMPKEN F. **Unique features of a linear plasmid of *Ascobolous immerses*: Implications for plasmid evolution in fungi**. Curr. Top. Mol. Genet. 1994; 2: 207-218.

Kües U. Life **history and developmental processes in the basidiomycete *Coprinus cinereus***. Microbiol Mol Biol Rev. 2000 Jun;64(2):316-53.

Lee SH, Kim BG, Kim KJ, Lee JS, Yun DW, Hahn JH, Kim GH, Lee KH, Suh DS, Kwon ST, Lee CS, Yoo YB. **Comparative analysis of sequences expressed during the liquid-cultured mycelia and fruit body stages of *Pleurotus ostreatus***. Fungal Genet Biol. 2002 Mar;35(2):115-34.

Maas MF, Sellem CH, Hoekstra RF, Debets AJ, Sainsard-Chanet A. **Integration of a pAL2-1 homologous mitochondrial plasmid associated with life span extension in *Podospora anserina***. Fungal Genet Biol. 2007 Jul;44(7):659-71.

Meinhardt LW, Rincones J, Bailey B, Aime MC, Griffith GW, Zhang D, Pereira G: ***Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe?** Mol Plant Pathol 2008, 9:577-588.

Niella GR, Castro HA., Silva lHCP, Carvalho JA. **Aperfeiçoamento da metodologia de produção artificial de basidiocarpos de *Crinipellis perniciosa***. 1999. Fitopat. Bras. 4:24:523-527.

Nowrousian M, Kück U. **Comparative gene expression analysis of fruiting body development in two filamentous fungi**. FEMS Microbiol Lett. 2006 Apr;257(2):328-35.

Pires AB, Gramacho KP, Silva DC, Góes-Neto A, Silva MM, Muniz-Sobrinho JS, Porto RF, Villela-Dias C, Brendel M, Cascardo JC, Pereira GA. **Early development of *Moniliophthora perniciosa* basidiomata and developmentally regulated genes**. BMC Microbiol. 2009 Aug 4;9:158.

Pöggeler S, Kempken F. **Mobile genetic elements in mycelial fungi**. In: U. Kück (ed) THE MYCOTA II, Genetics and Biotechnology, 2nd Edition, 2004. Springer Verlag, Heidelberg, New York, Tokyo pp 165-198

Purdy LH, Schmidt RA. **Status of Cacao Witches' Broom: biology, epidemiology, and management**. Annu Rev Phytopathol. 1996;34:573-94.

Rincones J, Scarpari LM, Carazzolle MF, Mondego JM, Formighieri EF, Barau JG, Costa GG, Carraro DM, Brentani HP, Vilas-Boas LA, de Oliveira BV, Sabha M, Dias R, Cascardo JM, Azevedo RA, Meinhardt LW, Pereira GA. **Differential gene expression between the biotrophic-like and saprotrophic mycelia of the witches' broom pathogen *Moniliophthora perniciosa***. Mol Plant Microbe Interact. 2008 Jul;21(7):891-908.

Scarpari LM, Meinhardt LW, Mazzafera P, Pomella AW, Schiavinato MA, Cascardo JC, Pereira GA. **Biochemical changes during the development of witches' broom: the most important disease of cocoa in Brazil caused by *Crinipellis perniciosa***. J Exp Bot. 2005 Mar;56(413):865-77.

Scarpari LM, Meinhardt LW, Mazzafera P, Pomella AW, Schiavinato MA, Cascardo JC, Pereira GA. **Biochemical changes during the development of witches' broom: the most important**

disease of cocoa in Brazil caused by *Crinipellis perniciosa*. J Exp Bot. 2005 Mar;56(413):865-77.

Soanes DM, Talbot NJ. **Comparative genomic analysis of phytopathogenic fungi using expressed sequence tag (EST) collections**. Mol Plant Pathol. 2006 Jan 1;7(1):61-70.

CHAPTER 5

Phylogenetic analysis of DNA and RNA polymerases from *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora mitochondrial plasmid reveals a probable lateral gene transfer

Abstract

The filamentous fungus *Moniliophthora perniciosa* (Stahel) Aime & Phillips-Mora is a hemibiotrophic Basidiomycota that causes witches' broom disease of cacao (*Theobroma cacao* L.). Many fungal mitochondrial plasmids are invertrons, encoding DNA and RNA polymerases with terminal inverted repeats and 5'-linked terminal proteins. The aim of this work was to carry out comparative and phylogenetic analyses of DNA and RNA polymerases for all linear mitochondrial plasmids in fungi; we performed these analyses at both the gene and protein level and assessed the differences between fungal and viral polymerases in order to test the lateral gene transfer (LGT) hypothesis. We analyzed all known mitochondrial plasmids of the invertron type within the fungal clade: five Ascomycota; seven Basidiomycota; and one Chytridiomycota. All phylogenetic analyses showed similar tree topologies regardless of either the method or data set used. DNA and RNA polymerases were probably inserted during different events by LGT into mitochondrial genomes of the 13 fungal host species used in our study. These results are important for a better understanding of the evolutionary relationships among mitochondrial plasmids for the fungal clade.

Key words: *Moniliophthora perniciosa*, Mitochondrial plasmids, Fungi, Molecular Phylogeny

5.1 Introduction

The basidiomycete *Moniliophthora perniciosa* (Stahel) Aime and Philips-Mora [previously known as *Crinipellis perniciosa* (Stahel) Singer] is the causative agent for witches' broom disease of the cacao tree (*Theobroma cacao*), whose seeds are the source of chocolate (Aime and Philips-Mora 2005). It is the most important phytopathological problem of cacao-producing areas in the American continent and has decimated the Brazilian cacao industry (Griffith et al. 2003). *M. perniciosa* is within the order Agaricales, which has few known pathogens (Aime; Phillips-Mora 2005), and is endemic in the Amazon region. It is the only pathogen that develops concurrently with the cocoa plant (Purdy 1996). Several biotypes of *M. perniciosa* have been described, and the biotype C affects *Theobroma cacao* and other species of the genera *Theobroma* and *Herrania* (Malvaceae) (Griffith 1994). The mitochondrial genome (Formighieri et al. 2008) and more recently whole genome (Mondego et al. 2008) of this fungus have been sequenced.

Plasmids are small molecules of DNA (or RNA) that have the ability to replicate inside living cells independently of the host genome (Cahan; Kennel 2005). These structures can sometimes covalently integrate into the genome of the cells or organelles and thus be replicated as part of the host genome (Griffiths 1995). Plasmids were originally discovered in bacteria but similar molecules were subsequently found in eukaryotes, mainly in mitochondria of filamentous fungi (mt plasmids) and plants (Cahan; Kennel 2005). These molecules may have two types of conformation: circular or linear (Griffiths 1995). In fact, most eukaryotic plasmids are linear (Meinhardt et al. 1990). Mt plasmids are generally viewed as being intracellular parasites, and their prevalence, at least in filamentous fungal hosts, relates to their ability to spread horizontally via hiphal anastomosis (Cahan; Kennel 2005). Usually, the impact of introgression or existence of a plasmid on the host phenotype is not clear. However, these molecules may confer some

selective advantage for the host, changing metabolism or mitochondrial division in some aspects. In many physiological studies, no change in rate or pattern of growth was observed, because there are as yet no sensitive tests to identify very small changes in growth, respiration or reproductive rate (Griffiths 1995). On the other hand, a phenomenon known as syndrome of fungal senescence was described in some species of the genus *Neurospora* (Marcou 1961; Chan et al. 1991), which includes progressive loss of mitochondrial function, loss of acrogenic fertility structures and reduction of hyphal growth together with loss of conidial viability and possible death of the mycelium (Meinhardt et al. 1990; Chan et al. 1991). Mitochondrial plasmids are widely distributed in filamentous fungi and share some features, such as the presence of Terminal Inverted Repeats (TIR) and genes coding for DNA and RNA polymerase (DPO and RPO, respectively) (Kempken et al. 1992; Kempken 1994; Cahan; Kennel 2005;). These plasmids are present in different genera and species of Ascomycota and Basidiomycota, including saprophytes and plant pathogens (Giese et al. 2003), particularly in the genus *Neurospora* (Xu et al. 1999). In many cases they are transmitted in the same way as the mitochondria and mitochondrial DNA. In sexual crosses, the maternal relative of the plasmid is transmitted to most or all progeny. According to the theory of endosymbiosis, the mitochondria of fungal plasmids have a prokaryotic origin, and some genes may have been inherited from viral bacteriophages (Griffiths 1995). The degree of relationship among mitochondrial plasmids from different species has been estimated using nucleotide and amino acids sequences of DPO and RPO. The transmission of mitochondrial plasmids during reproduction in fungi has been the subject of many studies (Giese et al. 2003), as introgression in *Neurospora sp.* (Bok et al. 1999) and asexual transmission in *Cryphonectria parasitica* (Murrill) ME Barr (Baidyaroy et al. 2000). Other studies have shown that, in spite of the existence of somatic barriers, the transmission of plasmids has also occurred by anastomosis (Giese et al. 2003). The mitochondrial plasmid from

M. perniciosa has a typical invertron structure and DPO/RPO coding genes in opposite orientations. Furthermore, it appears to be stably integrated into the mitochondrial genome and is probably involved in senescence (Formighieri et al. 2008).

In the present study we analyzed of DNA and protein sequences for DNA and RNA polymerase enzymes encoded by all known types of linear mitochondrial plasmids in fungi, totaling 13 species: seven Ascomycota, five Basidiomycota (including *M. perniciosa*) and one Chitridiomycota. Our aim was to define a phylogeny based on molecular data, testing the hypothesis of lateral transfer of the polymerase gene within the fungal clade and between viruses and fungi.

5.2 Methods

Phylogenetic analyses were performed from all nucleotide and amino acid sequences of DPO and RPO for all known mitochondrial linear plasmids in the fungal clade that had been deposited in the NCBI/EMBL/DDDJ databases. In addition, DPO and RPO protein sequences of viruses were included in this study by means of comparative similarity analysis using BLAST version 2.2.16 (Altschul et al. 1997), because of their probable evolutionary relationships with DPOs and RPOs of mitochondrial fungal plasmids.

DPO and RPO nucleotide sequences for the 13 fungal plasmids were codon-aligned with their corresponding amino acid sequences using PAL2NAL (Suyama et al. 2006). The results of these alignments were edited in BIOEDIT 7.0.5.2 (Hall, 1999), excluding highly variable C-terminal and N-terminal domains in this group of enzymes, and saved in NEXUS format as DPO/RPO contigs. In order to obtain the best evolutionary model for the matrix used in the analysis of maximum likelihood, the nucleotide sequences of the DPO/RPO contigs were submitted to Modeltest (Posada 1998); GTR + G + I was selected as the best evolutionary model. Clade robustness was also assessed using bootstrap proportions (1,000 pseudoreplicates). Phylograms

were generated in PAUP 4.0b10 (Swofford 2002) and then saved and visualized in Treeview 1.6.6 (Page 1996). For these analyses, the outgroup was represented by sequences (DPO and RPO) from the mitochondrial plasmid of *Spizellomyces punctatus* (WJ Koch) DJS Barr, since chytridiomycotan species probably originated before their Ascomycota and Basidiomycota counterparts (James et al. 2006).

The amino acid sequences of fungal and viral DPOs and RPOs were aligned separately using BLOSUM62 matrix (Henikoff 1992) with 1,000 pseudoreplicates of bootstrap proportions in BIOEDIT 7.0.5.2 (Hall 1999). In this case, it was not possible to perform a combined analysis of DPO/RPO contigs using fungal and viral sequences, as the DPO and RPO do not occur concurrently in the same viral species. The phylogeny for fungal and viral DPO and RPO amino acid sequences were defined by Bayesian analysis performed in MRBAYES 3.1.2 (Ronquist 2003) using the evolutionary model MTMAM (Yang et al. 1998). This model was chosen since the mitochondria of both fungi and mammals have similar rates of substitution in their genomes as well as similarities in codon usage when compared to the standard nuclear genome (Osawa 1992). Three independent runs were conducted (each with four chains) for 5 x 10^6 generations, sampling every 1000 generations. In the phylogenetic analyses of viral and fungal DPO nucleotide sequences, bacteriophage Φ29 DPO was used as outgroup, and in the analyses of viral and fungal RPO nucleotide sequences, T7 RPO was used as outgroup.

5.3 Results

There are 13 completely sequenced mitochondrial plasmids from species of the fungal clade showing a linear structure with reversed ends: five in Ascomycota, seven in Basidiomycota and one in Chytridiomycota (Table 1). The size of the plasmids ranged from 1.7 to 8.6 kb, with the mitochondrial plasmid of *Spizellomyces punctatus* showing the smallest size and that of *Neurospora intermedia* the largest. All the studied fungal plasmids exhibited DPO and RPO

genes on opposite strands, except for the *S. punctatus* plasmid which has these genes on the same strand. Comparative sequence analyses using the protein sequences (DPO and RPO) of all fungal mitochondrial plasmids revealed high similarity with the viral DPO and RPO listed in Table 2. Maximum likelihood phylogenetic analysis of DPO/RPO nucleotide contigs revealed a phylogram where ascomycotan sequences were clearly grouped together with basidiomycotan sequences (Fig.1).

The Bayesian majority consensus tree for the amino acid sequences of DNA polymerases from fungi and viruses showed that viral DPOs did not form a monophyletic clade but instead were grouped together with fungal DPO sequences (Fig. 2). Conversely, the Bayesian inference for protein sequences of RNA polymerases from fungi and viruses supported two homogeneous clades: one for fungal RPOs (100%) and another for viral RPOs (74%) (Fig. 3).

5.4 Discussion

According to Griffiths (1995), most plasmids found within species of the fungal clade are linear, using a process similar to the replication protein-primed scenario associated with adenovirus and phage Φ29. This process employs a terminal protein covalently linked to the 5' end of these molecules. According to Sakaguchi (1990), adenovirus and phage Φ29 exhibit features similar to those of DNA plasmids, with a genome containing terminal inverted repeats and terminal proteins covalently linked to the 5' end of the molecules. All 13 plasmids used in our study, except for that from *S. punctatus*, have linear structures very similar to each other, with DNA and RNA polymerases in opposite orientations and terminal inverted repeats, Results of comparative sequence analysis revealed a probable ancestral relationship for the two polymerases encoded by each plasmid and viral polymerases, suggesting that the process of replication for these plasmids is the same as described by Griffiths (1995) for viral polymerases, thus corroborating the data of Sakaguchi (1990).

There are several approaches to ascertaining the origin and distribution of plasmids in mitochondria of fungi. According to Fukuhara (1995), it is very likely that these structures are derived from adenovirus and phages within bacterial ancestors of present-day mitochondria. This hypothesis is supported by data from comparative sequence analysis of the 13 fungal plasmids in our study, since the sequences of DPOs and RPOs are similar to those originated from adenovirus and retrovirus, respectively. Moreover, as described by Fukuhara (1995), these data do not address any hypothesis of how these structures have arisen in the cytoplasm of fungi.

Rosewich and Kistler (2000) reported a many examples of lateral gene transfer (LGT) in fungi, in which mobile genetic elements such as plasmids, transposons and introns must be involved. When we analyzed the data obtained from our phylogeny of individual or combined DPOs and RPOs for the 13 mitochondrial plasmids included here, we found that both polymerases from mitochondrial plasmids of Ascomycota and Basidiomycota do not form monophyletic groups, and therefore do not have a common origin. Another interesting finding is that even the DPO and RPO of species within the same genus (*Neurospora crassa* and *N. intermedia*) have a common evolutionary origin, as suggested by phylogenetic analyses at both the gene and protein level.

A recent molecular phylogeny study conducted by James et al. (2006), based on data from six nuclear genes of 199 species of fungi, strongly suggested that Ascomycota and Basidiomycota are monophyletic while Chytridiomycota is clearly polyphyletic. Thus, it is expected that DNA and RNA polymerases of the 13 fungal mitochondrial plasmids we examined should form monophyletic groups represented by their host fungi Ascomycota and Basidiomycota. The inferred phylogenies of both protein and DNA sequences, however, do not corroborate this assumption.

The DPO phylogeny of the 13 fungal plasmids with their homologous viral polymerases corroborates the relationship of the fungal DNA plasmids with sequences of adenovirus and

bacterial phages, as previously described. The formation of an exclusive fungal monophyletic group with a high statistical support occurred only in the phylogenetic analysis of fungal and viral RNA polymerases. In the phylogenetic analysis of fungal and viral DNA polymerases, Human Adenovirus type 12 and simian DPOs were grouped together with strong support within the same clade, with both ascomycotan and basidiomycotan DPOs in the majority consensus tree.

The results of the DNA polymerase phylogeny for fungal mitochondrial plasmids and viral polymerases were similar to those reported by Rosewich and Kistler (2000). These authors performed a phylogenetic analysis of amino acid sequences for DNA polymerases (11 linear plasmids from fungi and plants, 4 bacteriophages and 8 viruses), which indicated that groups of plasmids with related hosts (plants, fungi, viruses or bacteriophages) are not necessarily closely related.

A likely explanation for the occurrence of non-monophyletic groups that reflect the phylogeny of host species within the analysis of individual plasmids of fungal and viral polymerases is that host DPO and RPO sequences may have originated during a pre-endosymbiotic period, in which viruses and phages had, as hosts, bacteria that subsequently became mitochondria. This hypothesis is suggested by Griffiths (1995) but we must consider it carefully. Furthermore, another possible explanation of the non-monophyly for these host DPO and RPO sequences is a post-endosymbiotic insertion event within the fungal clade carried out by micoviruses, which could have transferred these polymerases into the mitochondrial genomes. This assertion is based on a fact described by Rosewich and Kistler (2000), in which the sequence of a viral RNA polymerase in the mitochondria of the ascomycotan *Ophiostomis new-ulmi* had greater similarity to the RPO of the basidiomycotan *Rhizoctonia solani* than to those of the ascomycotan *Cryphonectria parasitica*.

All phylogenetic trees generated using DNA and RNA polymerases for the 13 plasmids of the fungal clade, as well as for both fungi and viruses, had very similar topologies. In the majority of the consensus trees for both combined (contigs) and separate analysis, *M. perniciosa* plasmid polymerases were more related to *F. velutipes* plasmid polymerases. Interestingly, in most analyses the clade formed by these two Basidiomycota (*M. perniciosa and F. velutipes*) was more close to the clade composed by ascomycotan *Gelasinospora* sp. and *N. intermedia* polymerases than all other groups.

According to Gaag et al. (1998), the transfer of mitochondrial plasmids between species of fungi can occur via mitosis or meiosis. An example of this transfer was documented in *Ascobolus immersus* and *P. pauciseta* using co-culture experiments, with the transfer of mitochondrial plasmid SP12 from the first species to the second. Thus, evidence suggests that the close phylogenetic relationship observed in virtually 100% of the analyses between the clade formed by *M. perniciosa* and *F. velutipes* and the other composed of *Gelasinospora* sp. and *N. intermedia* may be a result of gene transfer at the level of genus or species. This must be considered with caution and further studies are required individually for this more inclusive clade (*M. perniciosa*, *F. velutipes*, *Gelasinospora* sp. and *N. intermedia*). An alternative view to this close relationship is the previously described hypothesis of mycoviruses, in which the same or very similar viruses have infected the mitochondria of fungi as distinct events. Furthermore, mycovirus sequences were always retrieved during the comparative sequence analysis for the selection of the viral polymerases that would be subject to phylogenetic analyses.

The analysis based only on amino acid sequences for fungal and viral RPO plasmids suggest a single origin for these enzymes in fungi. In addition, the retained clades in all retrieved trees, such as those formed by RPOs of *M. perniciosa* and *F. velutipes* plasmids, indicates that the insertion of these sequences into the mitochondrial genomes of the 13 fungal species occurred at

short intervals, whereas the substitution rate of mitochondrial genome is equal for all species used in this study.

According to Formighieri et al. (2008), the integration of the *M. perniciosa* plasmid damaging the mitochondrial genome was a recent event, which may be associated with the development of some biotypes. This process may have occurred by crossing lines that are not evolutionarily related, using the mechanisms of mitosis or meiosis as previously mentioned, or by mycoviruses. This can be observed in all trees since the clade formed by *M. perniciosa* and *F. velutipes* is closer to the clade formed by *Gelasinospora* sp. and *N. intermedia* than to that of any other group. Furthermore, in the Bayesian analysis of amino acids sequences from fungal and viral DPOs, the clade formed by *M. perniciosa* and *F. velutipes* appears more related to viral polymerases. Therefore, the formation of a fungal monophyletic clade for RNA polymerases, when analyzed together with the viral sequences, suggests that the sequences related to RPOs may have been inserted in events other than those that insterted DPOs, in *M. perniciosa* as well as in other fungi.

Considering the assumptions and results that were obtained from analyses of DPOs and RPOs at the protein and DNA level within in fungal clade, and between fungi and viruses, we suggest that there has been a probable lateral transfer of genes encoding DNA and RNA polymerases from the plasmid of *M. perniciosa* to other fungal mitochondrial plasmids, supporting the hypothesis that these sequences may have viral ancestors. It is unclear whether these sequences are from a pre- or post-endosymbiotic period and this issue needs further investigation.

5.5 Conclusions

All phylogenetic analyses (parsimony, distance, maximum likelihood and Bayesian) presented similar tree topologies, using both nucleotide and amino acid sequences in combined (contigs) or separate analyses for the exclusively fungal or fungal and viral data sets. Unlike the fungal DPOs

that did not group together, all fungal RPOs appear as a monophyletic group when compared with viral RPO. Furthermore, our study indicates that all fungal DPOs and RPOs were probably inserted by distinct LGT events into the mitochondrial genomes of the fungi studied here.

Acknowledgements

We would like to acknowledge financial support from FAPESB (Fundação de Amparo à Pesquisa do Estado da Bahia), which provided a M.Sc. scholarship to the first author, as well as the State University of Feira de Santana (UEFS) and its Graduate Program in Biotechnology (PPGBiotec, www.uefs.br/ppgbiotec).

References

AIME MC, PHILLIPS-MORA W (2005) The causal agents of witches' broom and frosty pod rot of cacao (chocolate, *Theobroma cacao*) form a new lineage of Marasmiaceae. Mycologia 97:1012-1022.

ALTSCHUL SF, MADDEN TL, SCHÄFFER AA, ZHANG J, ZHANG Z, MILLER W, LIPMAN DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nuc Acid Res 25:3389-3402.

BAIDYAROY D, GLYNN JM, BERTRAND H (2000) Dynamics of asexual transmission of a mitochondrial plasmid in *Cryphonectria parasitica*. Curr Genet 37:257–267.

BOK J-W, HE C, GRIFFITHS A (1999) Transfer of Neurospora kalilo plasmids among species strains by introgression. Curr Genet 36:275–281.

CAHAN P, KENNEL J C (2005). Identification and distribution of sequences having similarity to mitochondrial plasmids in mitochondrial genomes of filamentous fungi. Mol. Gen. Genomics 273: 462-473.

CHAN BS, COURT DA, VIERULA PJ, BERTRAND H (1991) The kalilo linear senescence-inducing plasmid of *Neurospora* is an invertron and encodes DNA and RNA polymerases. Curr Genet 20:225–237.

FORMIGHIERI EF, TIBURCIO RA, ARMAS ED, MEDRANO FJ, SHIMO H, CARELS N, GÓES-NETO A, COTOMACCI C, CARAZZOLLE MF, SARDINHA-PINTO N, THOMAZELLA DP, RINCONES J, DIGIAMPIETRI L, CARRARO DM, AZEREDO-ESPIN AM, REIS SF, DECKMANN AC, GRAMACHO K, GONÇALVES MS, MOURA-NETO JP, BARBOSA LV, MEINHARDT LW, CASCARDO JC, PEREIRA GA (2008) The mitochondrial genome of the phytopathogenic basidiomycete *Moniliophthora perniciosa* is 109 kb in size and contains a stably integrated linear plasmid. Mycol Res 112:1136-1152.

FUKUHARA H (1995) Linear DNA plasmids of yeasts. FEMS Microbiol Lett 131:1–9.

GIESE H, LYNGKJAER MF, STUMMANN BM, GRELL MN, CHRISTIANSEN SK (2003) Extrachromosomal plasmid-like DNA in the obligate parasitic fungus *Erysiphe graminis f.* sp. Hordei. Mol Genet Gen 269:699–705.

GRIFFITH GW, HEDGER JN (1994) The breeding biotypes of the witches' broom pathogen of cocoa, *Crinipellis perniciosa*. Heredity 72:278–289.

GRIFFITH GW, NICHOLSON J, NENNINGER A, BIRCH RN (2003) Witches’ brooms and frosty pods: two major pathogens of cacao. N Z J Bot 41:423-435.

GRIFFITHS AJF (1995) Natural Plasmids of Filamentous Fungi. Microbiol Rev 59:673-685.

HALL TA (1999) Contribution of Horizontal Gene Transfer to Evolution of *Saccharomyces cerevisae*. Nucl Acids Symp Ser 41:95–98.

HENIKOFF S, HENIKOFF JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89:10915–10919.

JAMES TY, KAUFF F, SCHOCH CL, MATHENY PB, HOFSTETTER V, COX CJ, CELIO G, GUEIDAN C, FRAKER E, MIADLIKOWSKA J, LUMBSCH HT, RAUHUT A, REEB V, ARNOLD AE, AMTOFT A, STAJICH JE, HOSAKA K, SUNG GH, JOHNSON D, O'ROURKE B, CROCKETT M, BINDER M, CURTIS JM, SLOT JC, WANG Z, WILSON

AW, SCHÜSSLER A, LONGCORE JE, O'DONNELL K, MOZLEY-STANDRIDGE S, PORTER D, LETCHER PM, POWELL MJ, TAYLOR JW, WHITE MM, GRIFFITH GW, DAVIES DR, HUMBER RA, MORTON JB, SUGIYAMA J, ROSSMAN AY, ROGERS JD, PFISTER DH, HEWITT D, HANSEN K, HAMBLETON S, SHOEMAKER RA, KOHLMEYER J, VOLKMANN-KOHLMEYER B, SPOTTS RA, SERDANI M, CROUS PW, HUGHES KW, MATSUURA K, LANGER E, LANGER G, UNTEREINER WA, LÜCKING R, BÜDEL B, GEISER DM, APTROOT A, DIEDERICH P, SCHMITT I, SCHULTZ M, YAHR R, HIBBETT DS, LUTZONI F, MCLAUGHLIN DJ, SPATAFORA JW, VILGALYS R (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443:818-822.

KEMPKEN F (1994). Unique features of a linear plasmid of *Ascobolous immerses*: Implications for plasmid evolution in fungi. Curr. Top. Mol. Genet. 2: 207-218.

KEMPKEN F, HERMANS J, OSIEWACZ H D (1992). Evolution of linear plasmids. J. Mol. Evol. 35: 502-513.

MARCOU D (1961) Notion de IongCvitP et nature cytoplasmique du determinant de la senescence chez quelques champignon. Ann Sci Nat Bot 12:653–763.

MEINHARDT F, KEMPKEN F, KAMPER J, ESSER K. (1990) Linear plasmids among eukaryotes: Fundamentals and Applications. Curr Genet 17:89-95.

MONDEGO JM, CARAZZOLLE MF, COSTA GG, FORMIGHIERI EF, PARIZZI LP, RINCONES J, COTOMACCI C, CARRARO DM, CUNHA AF, CARRER H, VIDAL RO, ESTRELA RC, GARCÍA O, THOMAZELLA DP, OLIVEIRA BV DE, PIRES AB, RIO MC, ARAÚJO MR, DE MORAES MH, CASTRO LA (2008) A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao. BMC Genom 9:548.

OSAWA S, JUKES TH, WATANABE K, MUTO A (1992) Recent Evidence for Evolution of the Genetic Code. Microbiol Rev 56:229-264.

PAGE RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358.

POSADA D, CRANDALL K (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818. 1998.

PURDY LH, SCHMIDT RA (1996) Status of cacao witches' broom: biology, epidemiology and management. Annu Rev Phytopathol 34:573-594.

RONQUIST F, HUELSENBECK JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

ROSEWICH UL, KISTLER HC (2000) Role of Horizontal Gene Transfer in The Evolution Of Fungi. Annu Rev Phytopathol 38:325–363.

SAKAGUCHI K (1990) Invertrons, a Class of Structurally and Fuctionally Related Genetic Elements That Includes Linear DNA Plasmids, Transposable Elements, and Genomes of Adeno-Type Viruses. Microbiol Rev 54:66–74.

SUYAMA M, TORRENTS D, BORK, P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nuc Acid Res 34:W609-W612.

SWOFFORD DL (1998) PAUP: Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4. Sunderland: Sinauer Associates.

VAN DER GAAG M, DEBETS AJ, OSIEWACZ HD, HOEKSTRA RF (1998) The dynamics of pAL2-1 homologous linear plasmids in Podospora anserine. Mol Gent Gen 258:521–529.

XU Y, YANG S, TURITSA I, GRIFFITHS A (1999) Divergence of a Linear and a Circular Plasmid in Disjunct Natural Isolates of the Fungus *Neurospora*. Plasmid 42:115–125.

YANG Z, NIELSEN R, HASEGAWA M (1998) Models of Amino Acid Substitution and Applications to Mitochondrial Protein Evolution. Mol Biol 15:1600–1611.

Tables

Table 1. Completely sequenced linear mitochondrial plasmids of the invertron type in fungi.

Plasmid	Fungus spp.	Phylum	Size (bp)	DNA pol. (aa)	RNA pol. (aa)	Access number	Reference
pEM	*Agaricus bitorquis*	Basidiomycota	5810	797	1102	X63075	Robison; Horger 1999
pBgh	*Blumeria graminis f. sp. hordei*	Ascomycota	7965	1062	973	NC_004935	Giese et al., 2003
pClK1	*Claviceps purpurea*	Ascomycota	6752	1063	963	X15648	Oeser; Tudzynski 1989
pMp	*Moniliophthora perniciosa*	Basidiomycota	6743	899	1028	NC_005927	Formighieri et al., 2008
pFV1	*Flammulina velutipes*	Basidiomycota	7363	925	1168	AB028633	Nakai et al., 2000
pG114	*Gelasinospora sp.*	Ascomycota	8231	987	831	L40494	Yuewang et al., 1996
pPK2	*Pichia kluyveri*	Ascomycota	7174	1118	992	Y11606	Blaisonneau et al., 1999
pHC2	*Hebeloma circinans*	Basidiomycota	3229	858	209	Y11504	Bai et al., 1998
pHarbin-3	*Neurospora crassa*	Ascomycota	7050	1021	896	AF133505	Xu et al., 1999
pKalilo	*Neurospora intermedia*	Ascomycota	8642	969	811	X52106	Chan et al.,1991
pMLP2	*Pleurotus ostreatus*	Basidiomycota	7005	900	993	AF355103	Kim et al., 2000
pAL2-1	*Podospora pauciseta*	Ascomycota	8395	1197	948	X60707	Hermanns; Osiewacz, 1992
pSp	*Spizellomyces punctatus*	Chytridiomycota	1775	361	218	AF404303	Forget et al., 2002

Table 2. Viral DNA and RNA polymerases used in phylogenetic analyses.

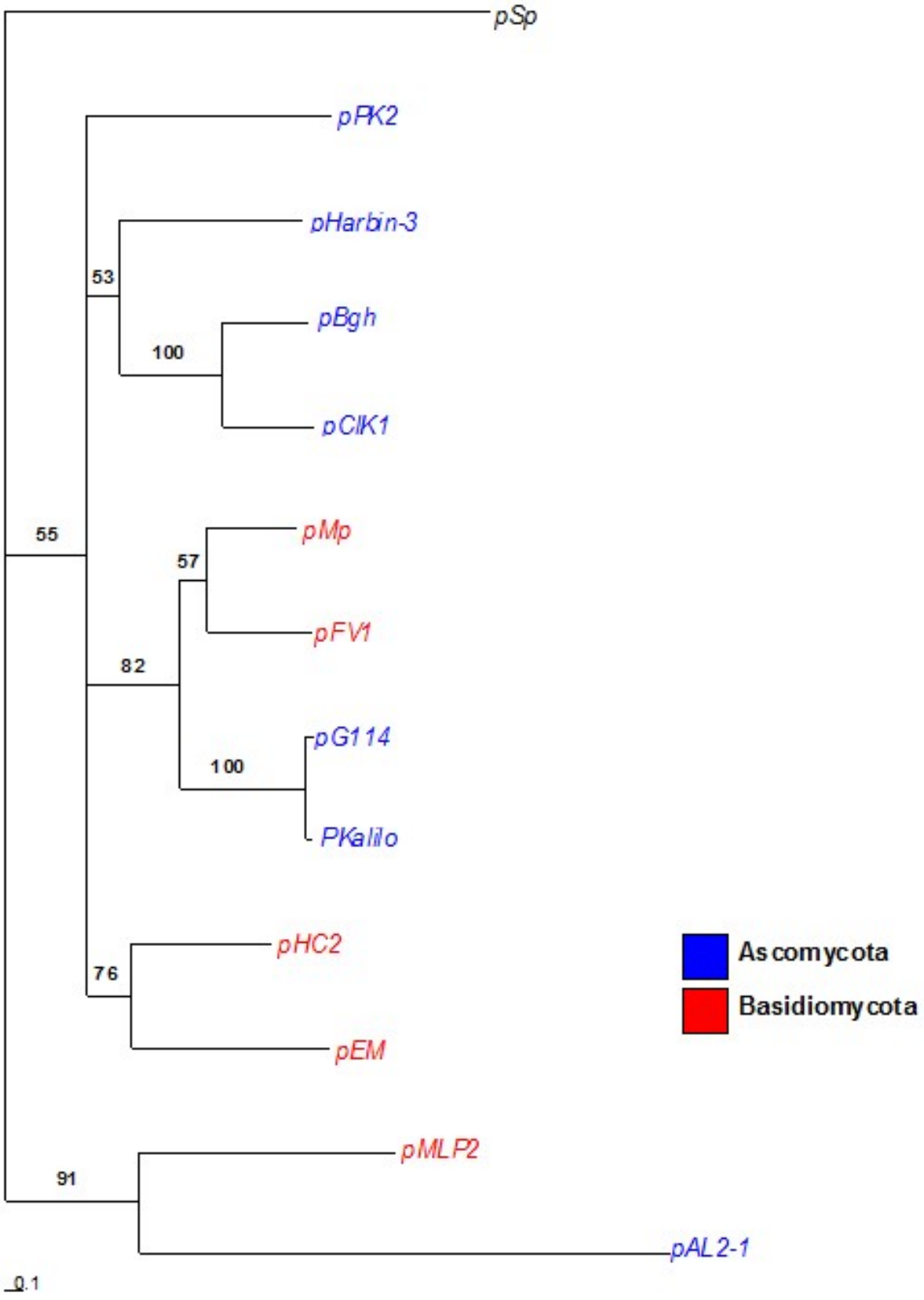
Virus	Type	Pol. type	Size (bp)	Size (aa)	Access number	Reference
Bacteriophage Bam35c	Adenovírus	DPO	2205	735	NP_943751.1	Ravantt et al., 2003
Phage GIL16c	Adenovírus	DPO	2259	753	YP_224103.1	Verheust et al., 2005
Hum. adenovirus 12	Adenovírus	DPO	3567	1189	AP_000112.1	Davison et al., 2003
Bacteriophage Phi29	Adenovírus	DPO	1725	575	1XHX_A	Kamtekar et al., 2004
Simium adenovirus A	Adenovírus	DPO	3507	1169	YP_067908.1	Kovacs et al., 2004
Bacteriophage phiYeO3-12	Retrovírus	RPO	2652	884	NP_052071.1	Pajunen et al., 2001
Bacteriophage T3	Retrovírus	RPO	2652	884	NP_523301.1	Pajunen et al., 2002
Phage K1F	Retrovírus	RPO	2679	893	YP_338094.1	Scholl; Merril, 2005
Phage GH-1	Retrovírus	RPO	2655	885	NP_813747.1	Kovalyova; Kropinski, 2003
BPK11	Retrovírus	RPO	2718	906	P18147	Dietz et al., 1990
Vibriophage VP4	Retrovírus	RPO	2649	883	YP_249577.1	Wang et al., 2005
Phage Berlin	Retrovírus	RPO	2649	883	YP_918986.1	Noeltin et al., 2006
Phage T7	Retrovírus	RPO	2649	833	1CEZ	Dunn; Studier, 1981

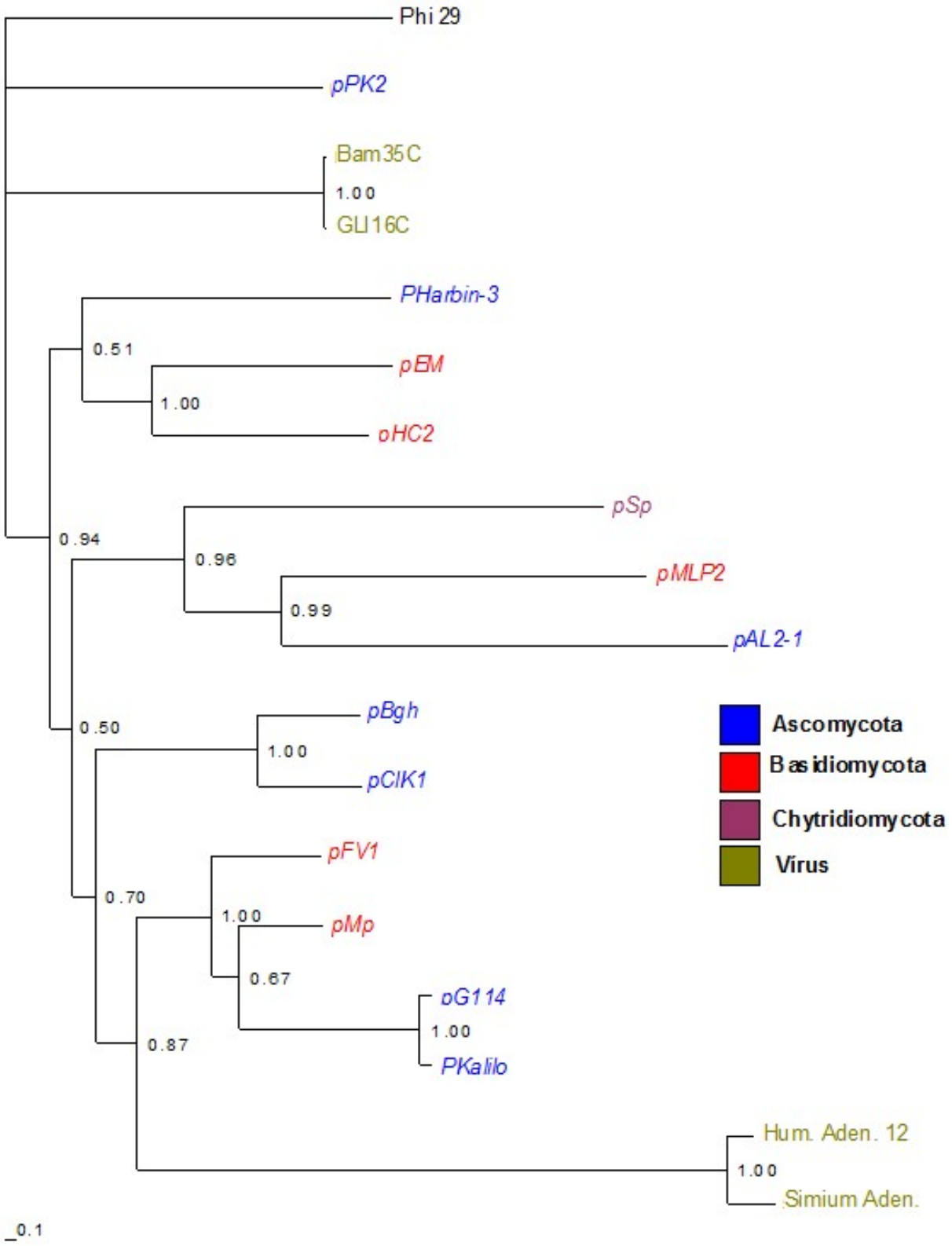
Figure legends

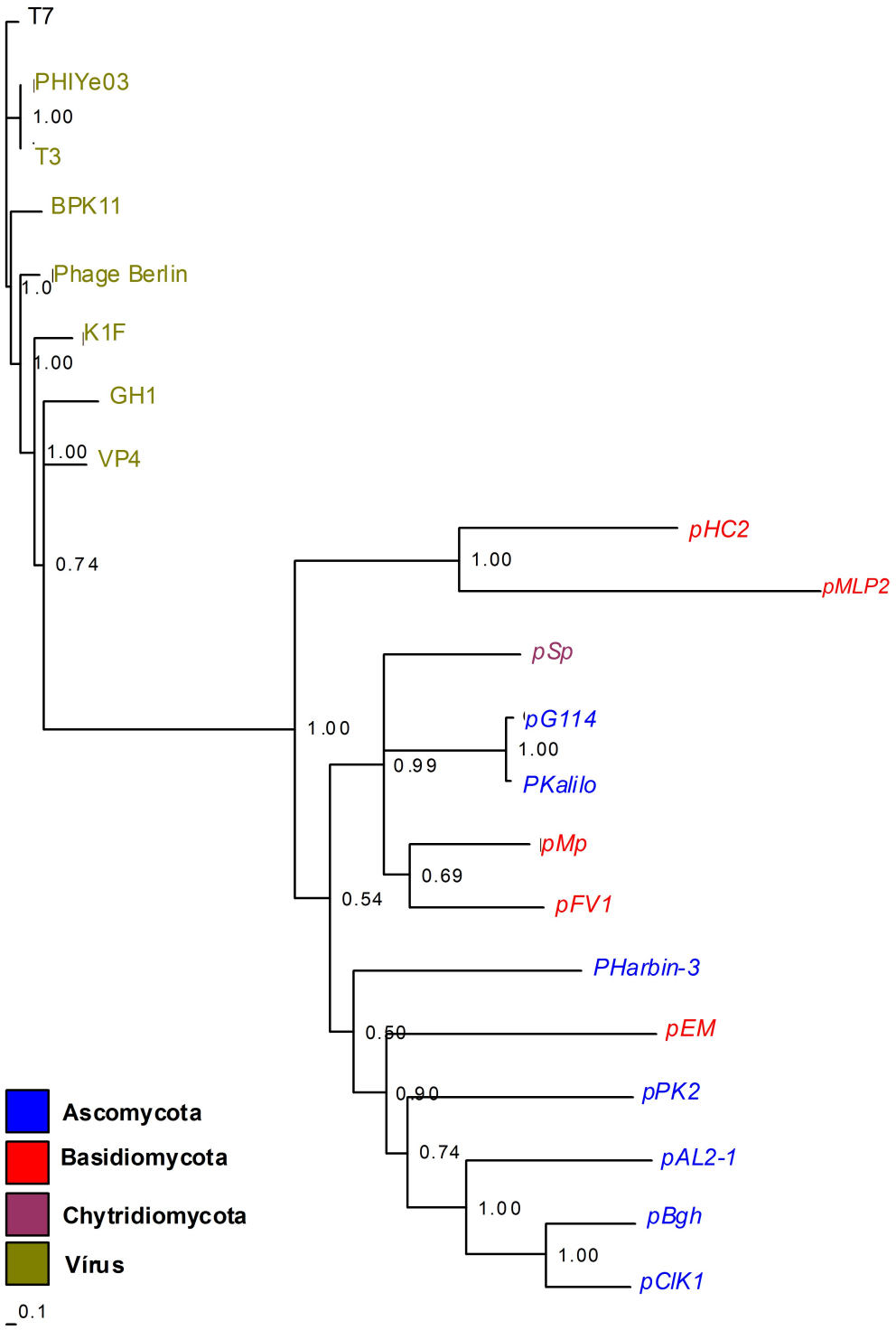
Figure 1. Phylogram (Maximum Likelihood) using nucleotide sequences contigs DPO/RPO of linear mitochondrial plasmids of fungi. Numbers above branches are bootstrap values.

Figure 2. Majority-rule consensus phylogram (Bayesian analysis) of fungal and viral DPO using amino acid sequences.

Figure 3. Majority-rule consensus phylogram (Bayesian analysis) of fungal and viral RPO using amino acid sequences.


pSp
pPK2
pHarbin-3
53
pBgh
100
pClK1
55
pMp
57
pFV1
82
pG114
100
PKalilo
pHC2
76
pEM
pMLP2
91
pAL2-1
0.1
Ascomycota
Basidiomycota


Phi 29
pPK2
Bam35C
1.00
GIL16C
PHarbin-3
0.51
pEM
1.00
pHC2
pSp
0.94
0.96
pMLP2
0.99
pAL2-1
pBgh
0.50
1.00
pCIK1
pFV1
0.70
1.00
pMp
0.67
pG114
1.00
PKalilo
0.87
Hum. Aden. 12
1.00
Simium Aden.
_0.1
Ascomycota
Basidiomycota
Chytridiomycota
Virus


T7
PHIYe03
1.00
T3
BPK11
Phage Berlin
1.0
K1F
1.00
GH1
1.00
VP4
0.74
pHC2
1.00
pMLP2
pSp
1.00
pG114
1.00
0.99
PKalilo
pMp
0.69
0.54
pFV1
PHarbin-3
0.50
pEM
0.90
pPK2
0.74
pAL2-1
1.00
pBgh
1.00
pClK1
Ascomycota
Basidiomycota
Chytridiomycota
Vírus
0.1

Conclusão geral

Nesta tese foram realizados estudos mais aprofundados a respeito das estruturas, expressão relativa, estudo de inibidores e evolução das DNA e RNA polimerases codificadas pelo plasmídeo mitochondrial de *M. perniciosa*. Os resultados obtidos com a determinação das estruturas tridimensionais dessas proteínas mostraram uma íntima relação estrutural e funcional com as polimerases virais que foram utilizadas como moldes para construção da DPO e RPO. A utilização de uma metodologia de triagem virtual de inibidores, acoplada a metodologias de Docking e Dinâmica Molecular, mostrou-se muito eficaz na seleção de bons complexos enzima-inibidor para cada uma das polimerases, indicando que é possível utilizar os inibidores selecionados nestes trabalhos para futuros testes *in vitro* e *in vivo*. No estudo da expressão relativa dos genes codificadores das DPO e RPO, durante diferentes fases de desenvolvimento de *M. perniciosa*, foram obtidos resultados estatisticamente significativos de uma resposta no metabolismo mitocondrial na fase de primórdio, através de um aumento significativo na expressão dessas enzimas justamente quando a planta (*T. cacao*) começa a liberar uma série de metabólitos que desencadeiam um ataque oxidativo contra esse patógeno. Logo, pode-se hipotetizar que, de alguma maneira, essas duas enzimas aceleram o metabolismo mitocondrial, replicando material genético e transcrevendo genes importantes para a manutenção oxidativa do fungo. Em nenhuma das análises filogenéticas realizadas apenas com polimerases fúngicas (agrupadas ou não), ou com polimerases fúngicas e virais, foi verificada a presença de clados de Ascomycota e Basidiomycota. Além disso, as árvores que apresentam polimerases virais mostram agrupamentos entre polimerases de fungos e de vírus. Esse fato é extremamente interessante, já que eventos de THG carreados por plasmídeos são comuns em diversos organismos e é provável que no clado fúngico esses eventos tenham ocorrido para essas polimerases e/ou para outros genes codificados por essas estruturas. Os resultados descritos nesta

tese poderão servir de base para estudos futuros que tratem do mecanismo de resistência do *M. perniciosa* ao ataque de *T. cacao* e outros que visem entender o mecanismo bioquímico de inibição das DPO e RPO, e sua conseqüência, além de delimitar a posição filogenética desse plasmídeo dentro clado fúngico e com seqüências relacionadas.