

## AMPLIAÇÃO DO NATPRODB, UMA BASE VIRTUAL DE ESTRUTURAS QUÍMICAS ORIUNDAS DO SEMI-ÁRIDO BAIANO

**Bruno Cruz de Souza<sup>1</sup>; Genilson Costa Santos<sup>2</sup>; Manoelito Coelho dos Santos Júnior<sup>3</sup>.**

1. Bolsista PROBIC, Graduando em Ciências Farmacêuticas, Universidade Estadual de Feira de Santana, e-mail: bcfarma@gmail.com
2. Participante do projeto “Desenvolvimento de Inibidores para enzimas do fungo *Moniliophthora perniciosa* (Sthael) (Singer) Phillips-Mora por de novo design e por estudos de cinética enzimática”, Laboratório de Modelagem Molecular, Departamento de Saúde, Universidade Estadual de Feira de Santana, e-mail: farm.genilsoncosta@gmail.com.
3. Orientador, Laboratório de Modelagem Molecular, Departamento de Saúde, Universidade Estadual de Feira de Santana, e-mail: mc2500@gmail.com

**PALAVRAS-CHAVE:** Banco de dados, NatProDB, *Moniliophthora perniciosa*.

### INTRODUÇÃO

Até meados da década de 20 o Brasil era o maior produtor de cacau do mundo. No final dos anos 80, ainda ocupava o segundo lugar, atrás apenas da Costa do Marfim, na África. Em 1989 a cultura do cacau brasileira sofreu uma queda em sua produção, o que, em parte, pode ser explicada pelo surgimento e desenvolvimento do fungo *Moniliophthora perniciosa*, responsável pela praga conhecida como vassoura-de-bruxa do cacauero (*Theobroma cacao*) (PEREIRA et al, 1990).

Vários compostos químicos vêm sendo testados com o objetivo de prevenir ou erradicar a vassoura-de-bruxa, porém os resultados não têm sido satisfatórios (SOBERANIS, 2000). Inibidores da biossíntese da parede celular bacteriana como as penicilinas e cefalosporinas, têm apresentado bons resultados no controle de infecções bacterianas. De forma similar, a parede celular dos fungos é um bom alvo para o desenvolvimento de potentes antifúngicos. Na busca por um controle efetivo da vassoura-de-bruxa, escolheu-se a rota metabólica que leva a síntese da quitina, principal componente da parede celular fúngica como alvo para se tentar inibir o desenvolvimento da praga no cacauero (HERSCOVICS; ORLEAN, 1993).

Bancos de dados de pequenas moléculas representam a maior fonte de estudo para interações bioquímicas e desempenham um papel crescente na descoberta novos fármacos. Atualmente o número de bases de compostos tem crescido a cada ano, essas por sua vez incluem conjuntos de compostos químicos, drogas, carboidratos, produtos e substratos de reações enzimáticas, produtos naturais e derivados de produtos naturais. Sendo assim este projeto visa ampliar um banco de dados de estruturas químicas derivadas do semi-árido baiano para otimizar a busca de possíveis inibidores de enzimas do fungo *Moniliophthora perniciosa*. (SONG et al, 2009).

### METODOLOGIA

A primeira etapa do trabalho foi uma extensa busca na literatura por trabalhos que apresentassem em seu escopo compostos químicos extraídos de plantas do semi-árido baiano, para isso foram utilizados os seguintes descritores na pesquisa: extração, semi-árido, baiano, caatinga, produtos naturais. Em seguida as moléculas encontradas foram desenhadas no formato 2D (\*.skc) e 3D (\*.mol) através do software ACD/Labs (ChemSketch) v10.00 e

armazenadas. O armazenamento dos dados foi feito em um programa construído da plataforma Java com apoio do Professor Ângelo Amâncio, Coordenador do Laboratório de Computação de Auto Desempenho - Departamento de Tecnologia, o conjunto de dados (programa e estruturas) recebeu o nome de NatProDB.

Após o desenho estrutural os compostos foram otimizados através de um método semi-empírico que é um tratamento quantitativo de propriedades moleculares com precisão, confiabilidade e custo computacional suficiente. Para essa otimização foi utilizado o método PM6 (STEWART, 2007), implementado no software Gaussian 09W (FRISCH et al, 2009).

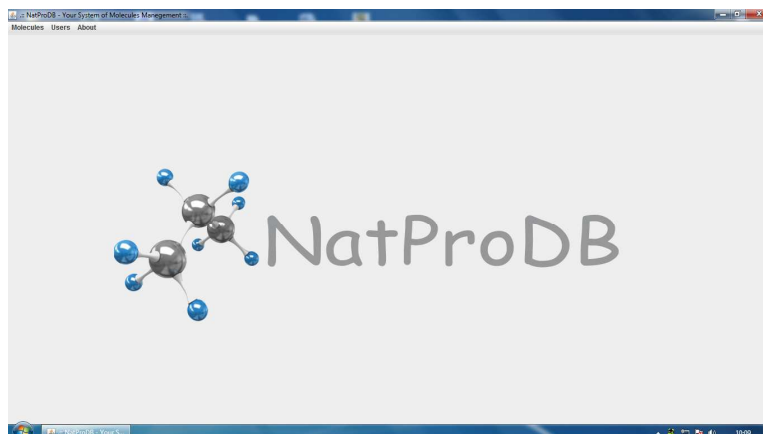
Essas estruturas foram codificadas e organizadas de acordo com a atividade, gerando dessa maneira dois grupos: 1) Com atividade farmacológica; 2) Sem atividade farmacológica. Para a organização e identificação dos compostos foi criado um código alfa numérico constituído por 3 campos como segue: VE/0AAA0/AF. O primeiro campo identifica a origem do composto, sendo VE para vegetal, AN para animal, MI para microorganismos e SI para sintética. O segundo campo identifica os compostos, possuindo dois números, um no início e outro no final, além de 3 letras. Esses dois números crescem variando de 0-9, e as três letras entre os números variam de A-Z. E o terceiro campo serve para identificar se o composto tem atividade farmacológica (AF) ou se não apresentam atividade farmacológica (SF).

Posteriormente utilizou-se o programa Marvin Sketch 5.6.0.2. para calcular propriedades dos compostos como: peso molecular, nomenclatura IUPAC e SMILES. O valor de ClogP foi calculado através do webservice XLOGP3 (CHENG et al, 2007) e os resultados foram tabelados. Todas estas informações foram depositadas em uma tabela padrão no NatProDB.

Após a organização das estruturas e informações referentes a cada uma, as mesmas foram depositadas no Banco de dados de estruturas virtual Zinc<sup>12</sup> (IRWIN; SHOICHET, 2005), que é um banco de dados publico composto por estruturas comercialmente disponíveis ou por bases de dados que são empregadas para ensaios de triagem virtual. O Zinc<sup>12</sup> contém cerca de 21 milhões de compostos prontos para acoplamento molecular em vários programas. O Zinc<sup>12</sup> é mantido pelo *Shoichet Laboratory* no *Department of Pharmaceutical Chemistry da University of California, San Francisco (UCSF)* (IRWIN; SHOICHET, 2005), e atualmente o Laboratório de Modelagem Molecular (LMM) possui uma parceria com o ZINC, no qual algumas estruturas já foram depositadas e encontram-se disponíveis para a comunidade científica.

## RESULTADOS E DISCUSSÃO

A partir da busca na literatura, foram encontrados e catalogados 100 novos compostos, a versão anterior do banco possuía cerca de 456 compostos. A otimização dos compostos foi realizada com o objetivo de minimizar imperfeições nas estruturas desenhadas. Na figura 1 pode-se observar a pagina inicial do NatProDB.



**Figura 1:** Tela de início do NatProDB.

Constatou-se que a maior parte dos compostos foram isolados de plantas da família Leguminosae (20,86%), seguida por Euphorbiaceae (11,51%) e Rutacea (7,55%). No NatProDB há uma diversidade química significativa. Foi possível identificar terpenos, diterpenos, flavonoides, cumarinas, fenilpropanóides dentre outras. A identificação das classes químicas é uma informação importante, pois em muitos casos sabe-se que determinados núcleos de classes químicas possuem ação comprovada frente a alvos biológicos, dessa forma pode-se direcionar estudos para essas classes químicas específicas evitando assim desperdícios de tempo. Identificou-se que as classes químicas mais encontradas foram terpenos com 30,22%, flavonoides com 15,82% e alcaloides com 10,13%

Para cada composto depositado são disponibilizadas informações como peso molecular, fórmula molecular, família, espécie, número de pontos para interações de hidrogênio (HBA e HDB), atividade biológica documentada e as referências (Figura 2)

 The image shows the 'Cadastro de Molécula' (Molecule Registration) window in NatProDB. The form contains the following fields and values:
 

- Code(\*):** VE0JGA0AF
- IUPAC Name(\*):** (2R,3S)-2-(3,4-dihydroxyphenyl)-3,4-dihydro-2H-1- benzopyran-3,5,7-triol
- Name(\*):** (-)Catequina
- Family:** COMBRETACEAE
- Specie:** Terminalia glabrescens Mart.
- Smile(\*):** O[C@H]1CC2=C(O)C=C(O)C=C2O[C@@H]1C1=CC(O)=C(O)C=C1
- HBA:** 5
- HBD:** 5
- Activity Biologic:** Nenhuma informação disponível na literatura.
- Location:** para Bahia, São Paulo, Minas Gerais, Goiás, Mato Grosso do Sul, Norte do Paraná, ocorre em floresta semidecídua.
- File:** C:\Users\usuariol\Desktop\NatProDB\06-12 ~ NatProDB\Estruturas\GVE0JGA0AF.mol
- Molecular Formula(\*):** C15H14O6
- MWT(Da)(\*):** 290.26
- cLogP:** 0.36
- References:** Recife: Associação das Plantas do Nordeste, IMSEAR, 2006. 297p (Instituto do Milênio do Semi-árido; v.6), pag. 178.

 At the bottom right, there are 'OK' and 'Cancel' buttons.

**Figura 2:** Tabela contendo as informações dos compostos depositados no NatProDB.

A flora brasileira continua sendo uma fonte pouco explorada para busca de compostos químicos com potencial de serem utilizados como novos fármacos, fato que pôde ser comprovado com a construção do NatProDB, pois, 39% dos compostos depositados atualmente no NatProDB não apresentaram informações que comprovassem algum tipo de atividade biológica. Em contrapartida 61% dos compostos apresentam algum tipo de atividade biológica, sendo encontradas diversas atividades distribuídas entre os compostos como, por exemplo, analgésica, antiinflamatória, muloscicida, antifúngica entre outras.

Nessa primeira versão o NatProDB possui apenas compostos químicos que foram oriundos de plantas, em versões posteriores pretende-se ampliar o quadro de compostos para outras fontes, como animais e sintéticas.

## CONSIDERAÇÕES FINAIS

A construção da base de dados NatProDB contribuiu significativamente como fonte de informação para a pesquisa e identificação de novos fármacos, pois parte de uma fonte de compostos rica e pouco estudada,. Devido a sua diversidade de compostos, a base de dados, além do intuito inicial de utilizá-la para triagem virtual frente a enzimas envolvidas no ciclo de vida do *M. perniciosa*, vem sendo utilizada em outros ensaios de triagem virtual, envolvendo outros alvos biológicos; assim favorecendo tanto a descoberta de novos compostos que possam ter ação inibitória frente ao fungo *M. perniciosa*, quanto estruturas químicas para outros ensaios.

## REFERÊNCIAS

ACD/ChemSketch Freeware, version 10.00, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2006.

CHENG, T.; ZHAO, Y.; LI, X.; LIN, F.; XU, Y.; ZHANG, X.; LI, Y.; WANG, R.; LAI, L. "Computation of Octanol-Water Partition Coefficients by Guiding an Additive Model with Knowledge", *J. Chem. Inf. Model.* 2007, 47, 2140-2148.

FRISCH, A. E; FRISCH, M. J. Manual Gaussian 03. Pittsburgh, PA: Gaussian, 2003

HERSCOVICS, A.; ORLEAN, P. Glycoprotein biosynthesis in yeast. *FASEB Journal*, n. 7, p. 540-550, 1993.

IRWIN, J. J., SHOICHET B. K., ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.*, 45 (1), p. 177–182, 2005.

PEREIRA, J. L. et al. First occurrence of witches' broom disease in the principal cocoa-growing region of Brazil. *Tropical Agriculture*, v. 67, n. 2, p. 188-189, 1990.

SOBERANIS, W. Increased frequency of phytosanitary pod removal in cocoa (*Theobroma cacao* L.) increases yield economically in eastern Peru. *Crop Protection*, v. 18, p. 667-685, 2000

SONG, C. M.; LIM, S. J.; TONG, J. C. Recent advances in computer-aided drug design. *Briefings in Bioinformatics*. v. 10. n.5. p. 579-591, 2009.