Emprego de algoritmo genético em estudos de QSAR de O-(2-fenóxi)etil-N-aralquilcarbamatos com atividade herbicida.

Marcus T. Scotti[1*] (PG), Vicente P. Emerenciano[1] (PQ), Luciana Scotti[2] (PQ), Renata R. Ishiki[3] (PQ), Hamilton M. Ishiki[3] (PQ). *mscotti@iq.usp.br

1 Instituto de Química - USP – São Paulo, SP; 2 Faculdade de Ciências Farmacêuticas, USP – São Paulo, SP; 3 Universidade do Oeste Paulista – Presidente Prudente – SP.

Palavras Chave: *QSAR, herbicida, algorítimo genético*

Introdução

A interrupção da biossíntese do carotenóide é um importante alvo para se obter a ação herbicida. Vários herbicidas, contendo estruturas químicas diversificadas, são inibidores da fitoene desaturase (PDS) e da ξ-caroteno desaturase (ZDS), ambas envolvidas na biossíntese do carotenóide.

Na tentativa de se compreender melhor quais os fatores (propriedades físico-químicas e/ou descritores topológicos) responsáveis pela inibição da PDS foram selecionadas da literatura[1] uma série de O-(2-fenóxi)etil-N-aralquilcarbamatos, e estudadas através de relações quantitativas estrutura-atividade.

Como variável dependente foi empregado o valor de pI_{50}, responsável pela inibição de 50% da atividade enzimática. As variáveis independentes são os descritores obtidos com o emprego do programa Dragon. As análises estatísticas para a obtenção do melhor modelo foram realizadas através da regressão linear múltipla, com o emprego do algoritmo genético.

Resultados e Discussão

Após a otimização das geometrias dos 37 compostos com o emprego do método semi-empírico AM1, as mesmas foram inseridas no programa DRAGON que automaticamente gerou os valores de 1664 descritores. Estes foram submetidos a um pré-tratamento, no qual foram excluídos os que apresentavam valores constantes, com apenas um valor diferente para a série e os descritores altamente correlacionados (r>0,99). A seguir, as variáveis independentes restantes (descritores) foram submetidas às análises de regressão linear múltipla, obtidas com o emprego de algoritmo genético. Nesta etapa foi utilizado o programa MobyDigs v 1.0. A série foi subdividida em duas: série de treinamento e de teste contendo, respectivamente, 31 e 6 compostos. Ambas as séries abrangeram toda a faixa de valores de pI_{50}.

A melhor equação obtida neste estudo está representada na equação dada a seguir:

pIC_{50} = 1,943 (±0,527) Jhete - 0,340 (±0,078) RDF070v + 0,179 (±0,099) RDF095v + 0,097 (±0,052) Mor02m - 6,910 Dm (±2,011) + 3,851 (±2,980)

(n=31; r^2 = 0,84; Q_{cv}^2 = 0,78; s = 0,26; F = 26,80; next=6; Q_{ext}^2=0,75)

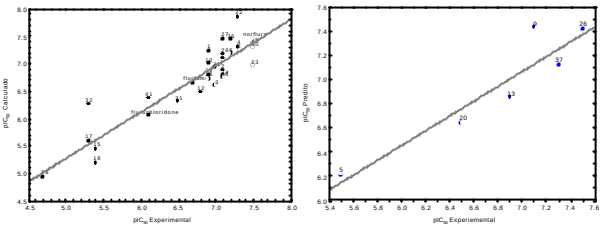

Figura 1. Gráficos dos valores de pIC_{50} calculados (série de treinamento) e preditos (série de teste) em função dos respectivos valores experimentais

A "robustês" e a preditividade da equação selecionada são verificadas pelos altos índices estatísticos tanto de predição interna (0,78) bem como de predição externa (0,75). Os descritores selecionados apresentam contribuições de natureza eletrônicas (Jhete) e estéricas (RDF 070v, RDF095v, Mor02m, Dm).

Conclusões

A utilização do algoritmo genético para a seleção dos descritores capazes de explicar a variância da atividade biológica se mostrou bastante eficiente, o que pode ser observado através dos valores estatísticos, bem como através dos gráficos. Os descritores RDF indicam que dois átomos, situados a distâncias de 7,0Å (RDF070v) e 9,5Å (RDF095v) um do outro, são importantes para a atividade biológica dos compostos. Já o descritor Mor02m indica a importância de ramificações e/ou tamanho das moléculas. O descritor Dm indica a acessibilidade entre os átomos, quanto menos denso (com relação a massa atômica) for a molécula, maior o espaço entre os átomos e maior o valor do índice Dm. O descritor Jhete indica a importância da contribuição do fator eletrônico para a atividade biológica.

Agradecimentos

FAPESP, CAPES, CNPq, UNOESTE

[1] Ohki, S. *et al. J. Agric. Food. Chem.* **2003**, 51, 3049.