

# FERRAMENTA COMPUTACIONAL PARA OBTENÇÃO DE COEFICIENTES DE PARTIÇÃO DE FÁRMACOS

Célio Batista da Silva\*

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Andréa Lucia Braga Vieira Rodrigues\*\*

**RESUMO:** O projeto consiste em desenvolver um sistema computacional que realiza cálculos sobre fragmentações de elementos químicos conduzindo a valores simulados de coeficientes de partição de moléculas orgânicas. A necessidade de calcular o coeficiente de partição justifica-se em descobrir se uma molécula possui maior solução em ambientes lipofílicos [óleo] ou hidrofílicos [água], determinando com maior precisão, onde algum elemento químico [medicamentos, drogas] se localizará em um espaço biológico. O sistema através do processo de fragmentação de moléculas, de acordo com padrões moleculares e regras de validação armazenados em uma base de dados, fragmenta cada molécula desenhada no *software*, busca seu respectivo valor na base de dados e considera validações de acordo com as regras. A soma das contribuições destas frações determina em termos de lipofilicidade, afinidade que uma molécula possui pelas gorduras, o coeficiente de partição logarítmico da molécula final. O projeto visou criar esta ferramenta apoiado em processos de desenvolvimento e engenharia de *software*, como UML, orientação a objetos, MVC, e à linguagem de programação Java. Para a persistência e gerência da base de dados foi utilizado Java com o processo de serialização e um SGBD SQL Server 2005 da Microsoft.

**PALAVRAS-CHAVE:** Tecnologia. Java. Coeficiente de partição. Fármacos. Quimicad.

## COMPUTATIONAL TOOL FOR OBTAINING THE COEFFICIENTS OF COMPOUND PARTITION

**ABSTRACT:** The project consists of developing a computing system that performs calculations on chemical elements fragmentation leading to simulated values of organic compound partition coefficients. The need to calculate this partition coefficient is justified in whether a molecule is more solution in lipophilic [oil] or hydrophilic [water] environments, determining more precisely, where some chemical element [medicines, drugs] will be located in a given biological space. The system, through its compound fragmentation process, identifies each unique molecular sub-fraction, according to molecular patterns and validation rules, stored in a database, fragmenting each molecule drawn in the software, searching their respective value in the database while considering many validation's rules. The sum of the contributions of these fractions determine in terms of lipophilicity, a molecule attribute that states his affinity for fats, the final molecule logarithmic partition coefficient. The project aims to create this tool supported by well oriented development processes as software engineering, UML, object orientation, MVC patterns, and Java programming language. For data persistence and database management will be used Java with the process of serialization and a Microsoft SQL Server 2005 DBMS.

**KEY-WORDS:** Technology. Java. Partition coefficient. Compounds. Quimicad.

\* Estudante do curso de graduação em Sistemas de Informação, Uniso - Sorocaba.  
Apoio: PROBIC – Projetos de Pesquisa com Bolsa-auxílio Institucional – Uniso  
Endereço: Rua Jose de Almeida Teixeira Filho, 831 - Jd. Maria José – Salto / SP CEP 13.321-020  
E-mail: [celio@celiosilva.com](mailto:celio@celiosilva.com)

\*\* Graduada em Engenharia Elétrica pela FACENS (1990), mestrado em Engenharia Elétrica pela USP (1994) e doutorado em Engenharia Elétrica pela USP (2002). Atualmente é prof. titular da Uniso e prof. titular / Coordenador da FACENS. Tem experiência na área de Engenharia Elétrica, com ênfase em Engenharia de Computação.

Recebido em: Março / 2008

Avaliado em: Maio / 2008

## INTRODUÇÃO

A lipofilicidade do fármaco é uma característica físico-química importantíssima que, em muitos casos, dita qual será o tipo de distribuição nos diferentes espaços biológicos do corpo humano. Fármacos lipofílicos conseguem transpor vários tipos de barreiras biológicas, pois estas são constituídas de membranas biológicas, que são igualmente lipofílicas, e acabam sendo distribuídos preferencialmente aos órgãos que possuem composição predominantemente lipofílicas, como o cérebro, o fígado, os tecidos que compõem as reservas corpóreas de lipídeos do abdômen. Já os fármacos de baixa lipofilicidade normalmente possuem acentuada hidrofílicidade e tendem a se acumular em órgãos que são mais irrigados por água, como pulmões, rins, baço, bexiga. A necessidade de obtenção de dados de lipofilicidade de compostos biologicamente interessantes já se manifesta em estudos que datam da virada do século XIX se estendendo até os dias atuais. O comportamento de lipofilicidade das moléculas, principalmente os fármacos, tem sido descrito como “coeficiente de partição” (P).

Coefficiente de partição de uma molécula é uma medida quantitativa que fornece informações a respeito da sua lipofilicidade e hidrofílicidade. Quando a molécula é submetida a um sistema binário de solventes (n-octanol/água) ela se distribui caracteristicamente entre as duas fases e seu P é calculado dividindo-se a concentração desta na fase n-octanol por sua concentração na fase aquosa.

$$P_{\text{oct}} = \frac{[Mo]}{[Ma]}$$

Onde:  $P_{\text{oct}}$  - coeficiente de partição no sistema n-octanol / água

[Mo] - concentração molar da molécula na fase orgânica

[Ma] - concentração molar da molécula na fase aquosa

O P obtido empiricamente vem expresso em termos logarítmicos:

$$\log P_{\text{oct}} = \log Mo - \log Ma$$

A equação logarítmica é especialmente adequada para expressar lipofilicidade e hidrofílicidade uma vez que, se o termo  $\log Mo$  for maior,  $\log P_{\text{oct}}$  dá um valor positivo (lipofilicidade preponderante) ao passo que se o termo  $\log Ma$  for maior,  $\log P_{\text{oct}}$  resulta em valor negativo (hidrofílicidade preponderante). O P pode ser determinado por três diferentes métodos: empiricamente pelo método “shake flask” (SF), por cromatografia líquida de alta eficiência (CLAE) ou através de cálculo.

Os métodos de cálculo de coeficiente de partição são fundamentados no isolamento de subestruturas moleculares [fragmentos moleculares] convenientes que possuam valores ou incrementos de lipofilicidade confiáveis e/ou conhecidos. O  $\log P$  estimado de toda a molécula é calculado somando-se as constantes que são contribuições (em unidades de  $\log P$ ) correspondentes a subunidades moleculares e incluindo-se os termos de correção devidos às interações intramoleculares e intermoleculares intrínsecas da molécula em questão. Os valores empíricos de  $\log P_{\text{oct}}$ , as constantes e termos de correção estão disponíveis em trabalhos originais de seus autores ou em instituições como a POMONA COLLEGE (Medicinal Chemistry Project), que regularmente atualizam seus bancos de dados. No presente trabalho é utilizado o método de cálculo de  $\log P$  desenvolvido por Roelof Rekker-Raimund Mannhold e equipe.

Os valores de P podem ser utilizados para direcionar a biodistribuição de compostos orgânicos em organismos vivos. Pode-se, por exemplo, projetar fármacos capazes de se distribuir em espaços biológicos específicos como o espaço intracerebral: agentes psicotrópicos, anestésicos, radiofármacos para tomografias cerebrais, etc. Quando se projeta um fármaco para ficar retido no espaço intracerebral, este não pode possuir uma lipofilicidade muito pronunciada, pois possuirá tanto a facilidade de entrar como a de sair, ficando dependente do gradiente de concentração de fármaco existente entre os dois lados. Além disso, sua estrutura, apesar de ser lipofílica, deve possuir certo grupo químico que o permita, por exemplo, fazer ligações de hidrogênio ou polarizar-se com o pH próprio do espaço biológico de forma a aumentar sua retenção. A melhor forma de explorar a lipofilicidade dos fármacos consiste em utilizar-se o descritor físico-químico coeficiente de partição.

Inicialmente Rekker utilizou-se de 100 valores experimentais confiáveis de logP de estruturas orgânicas simples, que serviram como série inicial para derivar as constantes fragmentárias por meio da análise de regressão de Free-Wilson continuamente otimizada por ajuste fino por uma ampliação da série de dados. Esta primeira etapa resultou num importante sistema para cálculo de logP baseado em 126 fragmentos. A fragmentação foi efetuada de forma a manter intactos os grupos funcionais possuindo interação direta de ressonância, os fragmentos foram diferenciados segundo a vizinhança alifática e aromática. Uma importante faceta das regressões consiste na detecção de diferenças sistemáticas entre os valores experimentais logP e aqueles calculados a partir da soma dos fragmentos.

Essas diferenças podem ser atribuídas a características químicas das moléculas que, por sua vez, permitem traçar regras de correção para o cálculo de logP. A uma das primeiras regras de correção que foram detectadas, Rekker a correlacionou com uma interação molecular que denominou de “efeito de proximidade” descrevendo configurações moleculares em que dois átomos ou centros eletronegativos separados por 1 ou 2 carbonos saturados. Outras correções foram estendidas para outras interações intramoleculares capazes de modificar o processamento do cálculo de logP tais como condensação de anéis aromáticos, duplas ligações conjugadas, ligação de hidrogênio.

Uma inspeção mais criteriosa dos valores de correção revela que esses consistem de valores múltiplos do valor 0,289 (=CM). A constante CM possui grande utilidade em contornar as diferenças entre os valores experimentais e os calculados propiciando que os últimos possam ser preditivos. Este método passou a ser reconhecido como “original f system”.

Os fatores de correção podem ser aplicados utilizando-se as seguintes regras:

1. Cadeias de hidrocarbonetos alifáticos saturados: + 2 CM - Estas cadeias são compostas somente por hidrogênio e carbonos e não podem possuir ligação dupla ou tripla.

1.1 Para metano: + 1 CM - Um hidrocarboneto alifático saturado composto por 4 hidrogênios e 1 um carbono CH<sub>4</sub>.

2. Anéis de Hidrocarbonetos saturados alifáticos: + 2 CM - Anéis contendo somente hidrocarbonetos saturados alifáticos.

2.1 Para ciclopropano: + 1 CM

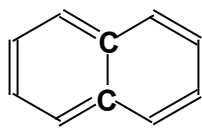
3. Insaturação: a tripla ligação: -1 CM

4. Conjugação estendida de cadeia: + 2 CM C=C-C=O

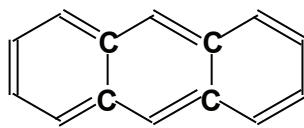
## 5. Hidrocarbonetos aromáticos

5.1 Benzeno: + 1 CM

5.2. Cada sítio de condensação de 2 anéis aromáticos: + 1 CM



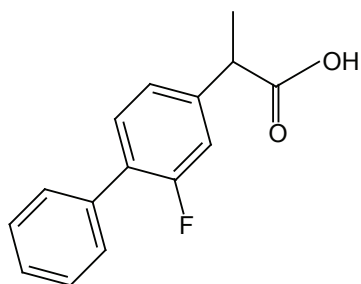
correção=+ 1 CM



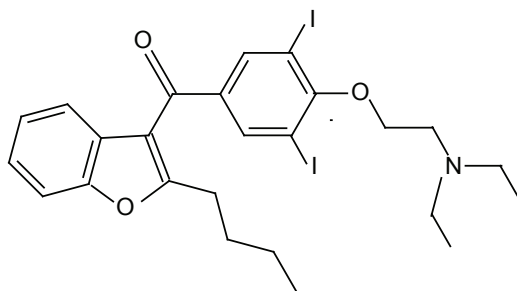
correção = + 2 CM

Cada par de carbonos que está grafado em negrito gera uma correção de + 1 CM.

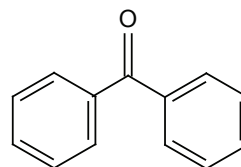
5.3 Conjugação arila-arila: + 1 CM



5.4 Conjugação cruzada: + 1 CM



amiodarona



benzofenona

## 6. Efeitos de proximidade

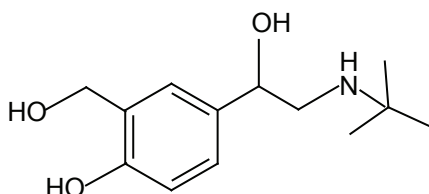
6.1 Um grupo metilênico: Espaçamento de 1 grupo metilênico entre 2 grupos eletronegativos X - CH<sub>2</sub> - X: + 3 CM

6.2 Dois grupos metilênicos: Espaçamento de 2 grupos metilênicos entre 2 grupos eletronegativos X - CH<sub>2</sub> - CH<sub>2</sub> - X: + 2 CM

7. H ligado a grupos eletronegativos: em estruturas como HCOOH o valor do fragmento f H passa a ser acrescido de 0,219, ou seja,  $0,2045 + 0,219 = 0,424$ .

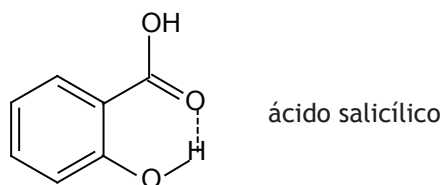
$$f H - X = 0,424$$

8. Eletronegatividade e proximidade de grupos alquila volumosos: grupos alquila volumosos possuindo carbono terciário ligados diretamente a átomos eletronegativos necessitam de correção de -2 CM, ao passo que grupos alquila volumosos possuindo carbono secundário ligados diretamente a átomos eletronegativos necessitam de correção de -1 CM



9. Oxigênio ligado a aromáticos via 1 C: nestas condições ocorre sobreposição entre o átomo de oxigênio e os elétrons do anel fenila, corrigir com +1 CM.

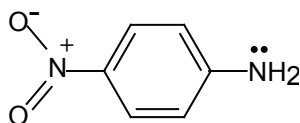
10. Ligação de hidrogênio: arranjos estruturais que propiciam a formação de ligação de hidrogênio intramolecular resulta em acréscimo de lipofilidade de +3 CM.



11. Desacoplamento por ressonância: grupos neutros, como alquilas, ligados em posição orto em relação a outro substituinte que é capaz de efetuar a interação de ressonância (ligado no anel aromático) estarão provocando um desacoplamento da ressonância com relação ao sistema aromático. Neste tipo de arranjo o substituinte aro mático estará sofrendo correção para valores mais alifáticos. A diferença entre os valores aromáticos e alifáticos está relacionada com valores múltiplos de CM, dependendo da intensidade do efeito de ressonância do substituinte.

Substituinte	Correção (regra 11)
OCH3	- 5 CM
COOH, CONH2	- 4 CM
C=O, CONH, NHCO	- 3 CM
NH2	-2 CM

12. Interação de ressonância: a combinação de 2 grupos com nitro, carboxila, ou carbonamida num anel fenílico nas posições para ou meta resulta numa interação de ressonância aumentando logP de 2 CM.



4 nitro anilina

## OBJETIVOS

O escopo deste trabalho foi criar uma ferramenta computacional capaz de processar cálculos para diferentes moléculas orgânicas a partir de método adequado em termos de capacidade preditiva, ou seja, aquele que consiga aproximar correções para apresentar o valor mais próximo possível dos exames realizados em laboratórios. [Preditividade].

## METODOLOGIA

Implementação por meios computacionais do método de Rekker, cujas etapas são: representação da molécula em um ambiente gráfico; validação em termos de valência e em termos das regras de correção do coeficiente de partição, devido às interações intramoleculares; busca de fragmentos armazenados em banco de dados gerando a fragmentação em subunidades que possuam valores confiáveis; cálculo do valor de coeficiente de partição; exibição dos fragmentos, regras aplicadas e do valor calculado.

Implementação de uma ferramenta, QuimiCad, que permite a edição molecular em ambiente gráfico, escrita e leitura de fragmentos em banco de dados (SQL Server), fragmentação com destaque dos elementos por meio de cores, cálculo do coeficiente de partição e aplicação de regras de validação. O *software* foi desenvolvido em Java com utilização de técnicas como orientação a objetos, UML, padrões de projeto como MVC e manipulação dos elementos em formato de objetos.

A aplicação contém elementos construídos como APIs que possibilitam o desenvolvimento e a reutilização do código gerado anteriormente. As APIs foram construídas como componentes para assim aumentar a produtividade.

O método de fragmentação utilizado pelo programa baseia-se no conceito de algoritmos de busca e comparação dentro do ambiente molecular. O ambiente molecular, tanto fragmento quanto molécula, é composto de objetos do tipo “Elemento”, “Ligação”, “ElementoAtômico” e “Molécula”. Estes elementos estão interligados entre si, de maneira a simular as ligações entre elementos químicos do mundo real. Para realização desta tarefa é necessária uma seqüência de passos buscando-se e comparando-se todos os elementos e ligações possíveis encontrados dentro de um fragmento ou molécula. Primeiro é necessário obter-se um fragmento, conjunto incompleto de elementos, o qual será utilizado como fonte para busca em um elemento maior, um conjunto completo de elementos, uma “Molécula”. O usuário desenha uma molécula, na ferramenta QuimiCad. O Quimicad, para fragmentar este elemento, buscará fragmentos na base de dados existentes. Após o retorno do fragmento da base de dados, que será utilizado na comparação com a molécula desenhada pelo usuário, será necessário que o sistema inicie o algoritmo de busca e comparação.

Utilizando-se da UML é possível escrever toda a documentação do programa, em cada fase, seguindo desde a análise até o desenvolvimento. Para a análise foram utilizados casos de uso, para o desenvolvimento do programa foi utilizado diagramas de classes.

A orientação a objetos está presente no projeto todo. Para descrever o ambiente molecular foram utilizados objetos que mapeiam o mundo químico real para o mundo químico computacional. Com a utilização de objetos que representem os elementos, ligações e o ambiente molecular foi possível representar um ambiente real no âmbito computacional. Para o projeto foi utilizado o padrão de projeto MVC [Model , View, Controller] que visa distribuir as responsabilidades de cada parte do programa. Com esta distribuição foi possível criar-se APIs [Application programming interface] que visam a reutilização de código, portátil e adaptável. A parte Model [modelagem] visa estabelecer a representação dos elementos químicos reais para o mundo computacional. A parte Controller [controle] visa estabelecer, implementar e realizar todos os cálculos necessários para obtenção do coeficiente de Partição. A parte View [visão] é responsável por gerar uma interface simples com o usuário para entendimento entre o homem e a máquina.

Serialização - com esta característica de Java foi possível converter-se facilmente os elementos para posterior persistência em bancos de dados. Serializar significa transformar um objeto ou informação na memória do computador em sua representação binária e vice-versa. Com isto, foi possível transformar os elementos em binários, persisti-los no SGBD e recuperá-los quando necessário, deserializando para exibição para o usuário.

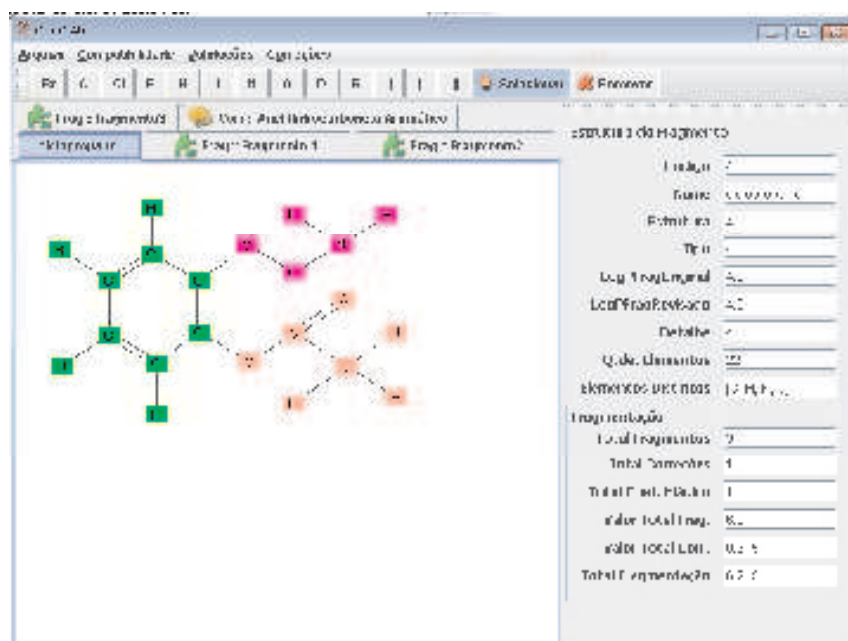
## DESENVOLVIMENTO

Com o QuimiCAD é possível realizar edição de moléculas em ambientes gráficos facilmente, apenas selecionando as moléculas necessárias e adicionando-as no espaço molecular. Em seguida pode-se adicionar ligações a estas moléculas representado-as fielmente às representações moleculares realizadas por químicos.

Com o sistema de persistência e fragmentação por meio de serialização é possível armazenar quantos fragmentos forem necessários no SGBD para posterior recuperação. O limite de moléculas é somente imposto pelo total de memória que um computador pode disponibilizar.

## Resultados

Ferramenta QuimiCAD.



## Considerações Finais

O método computacional é importante à medida que automatiza a inclusão e o processamento de vários passos envolvidos nos cálculos de P das moléculas com maior grau de complexidade estrutural. Permitirá estudos de centenas de moléculas, possibilitando, futuramente, corrigir os desvios e re-dimensionar os cálculos para otimizar sua característica preditiva. Poderá ser aplicado em diversos campos das ciências da saúde e ciências químicas, auxiliando no monitoramento da poluição ambiental, aplicações em estudos de CLAE, na pesquisa de patologias e na criação de novos fármacos mais eficazes e com efeitos colaterais reduzidos.

## REFERÊNCIAS

DEITEL, H. M.; DEITEL, P. J. **Java: como programar**. 3. ed. Porto Alegre: Bookman, 2001.

GARGADENNEC, S.; BURGOT, G.; BURGOT, J.L.; MANNHOLD, R.; REKKER, R.F. Log P Estimation of 1, 2-dithiole-3-thiones and 1,2-dithiole-3-ones. A comparison of -experimental and calculative Approaches. **Pharm. Res.**, v. 22, n. 6, p. 875-88, 2005.

GUNDERLOY, M.; JORDEN, J. L. **Dominando SQL server 2000**. São Paulo: Makron Books, 2001.

INPUT OF Structures by means of SMILES strings. Disponível em: <http://cactus.nci.nih.gov/services/gifcreator/smiles.html>. Acesso em: jul. 2006.

MANNHOLD, R.; REKKER, R. F. The hydrophobic fragmental constant approach for calculating log in octanol/water and aliphatic hydrocarbon/ water systems. **Persp. Drug Disc. Des**, v.18, p. 1-18, 2000.

REKKER, R. F. **The hydrophobic fragmental constant, its derivation and application: a means of characterizing membrane systems**. Amsterdam: Elsevier, 1977. v. 1

SMILES Tutorial. Disponível em: [http://www.epa.gov/medatwrk/Prods\\_Pubs/smiles.htm](http://www.epa.gov/medatwrk/Prods_Pubs/smiles.htm). Acesso em: jul. 2006.

THEORY smiles. Disponível em: <http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>. Acesso em: jun. 2006.