

## MODELOS CLÁSSICOS DE ESTUDO QUANTITATIVO DAS RELAÇÕES ENTRE ESTRUTURA QUÍMICA E ATIVIDADE BIOLÓGICA

Anderson Coser Gaudio\*

Departamento de Física - Centro de Ciências Exatas - Universidade Federal do Espírito Santo - Av. Fernando Ferrari s/n - Campus Universitário Goiabeiras - 29060-900 - Vitória - ES

Recebido em 27/10/94; aceito em 16/2/96

**CLASSICAL MODELS OF QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDIES.** The main contributions for the development of the quantitative classical models that relate the structural properties to the activity of bioactive compounds are reported. Starting approximately from the middle of the last century (the Crum-Brown and Fraser work), the Overton and Meyer, Ferguson, Zahradnik, Bruice, Hansen, Free-Wilson, Hansch-Fujita, and Kubinyi's works, among others, are discussed. The models due to Free-Wilson and Hansch-Fujita are discussed in more details.

**Keywords:** structure-activity relationships; history of QSAR; drug design.

### I. INTRODUÇÃO

Já faz algum tempo que duas siglas vêm aparecendo com bastante frequência nos congressos nacionais das áreas de química e bioquímica: SAR e QSAR. Os termos SAR-*Structure-Activity Relationships* e QSAR-*Quantitative Structure-Activity Relationships* designam o ramo de conhecimento que lida com as relações entre estrutura química e atividade biológica/toxicológica de fármacos. De forma bastante simplificada, SAR compreende as relações estrutura-atividade de ordem qualitativa e semi-quantitativa. Em QSAR, as relações estrutura-atividade são representadas através de modelos matemáticos bem definidos. Com este trabalho, pretende-se discutir os principais modelos clássicos de estudo QSAR, dar o histórico de seu desenvolvimento e arrolar exemplos de aplicações.

### II. HISTÓRICO

O primeiro trabalho na área de relações estrutura-atividade de que se tem notícia deve-se a Crum-Brown e Fraser<sup>1</sup>. Em trabalho que data de 1868, eles descreveram o resultado obtido após sucessivas injeções de derivados da morfina, estricnina e atropina em cobaias. Com base na análise desses resultados, propuseram que a ação fisiológica,  $\phi$ , de determinada molécula, era função de sua constituição química, C, ou seja,

$$\phi = f(C) \quad (1)$$

Segundo Tute<sup>2</sup>, "isto foi realmente imenso ato de fé, pois naquele momento nenhum número poderia ter sido atribuído a quaisquer dos lados da Equação (1), e eram conhecidas somente as fórmulas empíricas, certamente não as estruturas, de seus derivados". Em verdade, nem mesmo hoje conseguimos aplicar a Eq. (1), ou seja, ainda não é possível relacionar matematicamente a estrutura química das substâncias à atividade biológica. O que se tem feito é relacionar o efeito de pequenas variações da estrutura química às respectivas variações na atividade biológica.

Em 1893, Richet<sup>3</sup> propôs que a toxicidade de alguns álcoois, éteres e cetonas era inversamente proporcional à sua solubilidade em água. De acordo com Rekker<sup>4</sup>, "a importância dessa

descoberta não poderia ter sido completamente compreendida antes de terem sido corretamente estabelecidas as relações entre comportamento lipofílico e solubilidade; e, de fato, nós temos de homenagear Richet por ter esclarecido a primeira relação entre toxicidade e lipofilicidade".

Na virada do século, Overton<sup>5-7</sup> e Meyer<sup>8</sup>, independentemente, estudaram o efeito narcótico de vários compostos não ionizados e verificaram sua correlação com o coeficiente de partição lipídio/água. Foi constatado que o efeito narcótico era intensificado com o aumento do coeficiente de partição clorofórmio/água do fármaco testado. No entanto, a atividade biológica não aumentava indefinidamente; quando o coeficiente de partição atingia certo patamar a atividade narcótica não mais aumentava, sofrendo, a partir daí, contínuo decréscimo. A quebra dessa relação linear foi interpretada como sendo ocasionada pelos compostos excessivamente solúveis em lipídios que, sendo praticamente insolúveis em água, não conseguiriam atingir os tecidos-alvo por não poderem se dissolver no meio aquoso responsável pelo seu transporte. Sobre Overton, afirma Rekker<sup>4</sup> que "pioneira foi sua afirmação de que a narcose é devida a mudanças físicas efetuadas nos constituintes lipídicos das células".

Em 1939, Ferguson<sup>9</sup> estudou cuidadosamente o comportamento de diversas propriedades (solubilidade em água, partição, capilaridade e pressão de vapor) em relação à atividade tóxica, ao longo de séries homólogas e de séries correlatas de compostos. Segundo Ferguson, nos casos em que a atividade tóxica é consequência de processo físico, deve haver equilíbrio termodinâmico entre a exobiofase, na qual são realizadas as medidas da concentração do fármaco, e a biofase, local da ação primária do fármaco, em que se encontra o receptor farmacológico e que não é acessível às medidas experimentais. Nesses casos, a forma adequada de avaliar a atividade tóxica do fármaco é através da medida do seu potencial químico. Como nessas situações existe equilíbrio químico entre a biofase e a exobiofase, o potencial químico nestes dois ambientes deve ser o mesmo. Portanto, efetuando-se a medida do potencial químico na fase externa, na realidade estaremos determinando o potencial químico no local de ação do composto<sup>10,11</sup>.

Em 1937/1940, Hammett<sup>12,13</sup> verificou que as constantes de equilíbrio da ionização de ácidos benzóicos *meta* e *para* substituídos ( $k_X$ ) e do ácido benzóico não-substituído ( $k_H$ ) estavam relacionadas pela Eq. (2),

$$\log(k_X / k_H) = \rho \sigma \quad (2)$$

\*E-mail: anderson@cce.ufes.br

em que a constante  $\sigma$  mede o efeito eletrônico do substituinte X sobre o processo de ionização e  $\rho$  é uma constante que indica a susceptibilidade da reação ao efeito eletrônico exercido pelo substituinte X. A constante  $\sigma$  é característica do substituinte e independente da reação, enquanto que  $\rho$  é característica da reação e independente do substituinte. Hammett adotou a reação de ionização dos derivados do ácido benzóico como padrão, à qual atribuiu o valor de  $\rho = 1,00$ . Isto implicou no valor de  $\sigma_H = 0,00$  e, a partir daí, foi possível a determinação do valor de  $\sigma$  para inúmeros substituintes.

Reescrevendo a Eq. (2) na forma da Eq. (3),

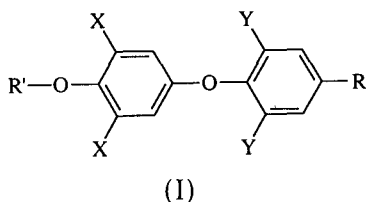
$$\log k_X = \rho\sigma + \log k_H \quad (3)$$

podemos notar que se construirmos um gráfico de  $\log k_X$  em função dos valores de  $\sigma$ , o resultado será uma reta cuja declividade é igual a  $\rho$  e cujo intercepto é igual à atividade do composto não-substituído.

A equação de Hammett é exemplo de *relação linear de energia livre*<sup>14</sup>. Segundo Tute<sup>15</sup>, "ela é preditiva no sentido de que o conhecimento da constante  $\sigma$  possibilitará que se faça a estimativa da reatividade de um composto cuja reatividade é desconhecida. A equação de Hammett também pode funcionar como ferramenta de classificação de certo tipo de reação: se a equação é satisfeita podemos dizer que o efeito eletrônico do substituinte é importante para o mecanismo da reação". Embora estejamos falando de reações químicas no sentido mais geral da expressão, a constante de Hammett foi e continua sendo amplamente utilizada em estudos de relações estrutura-atividade, pois, nesse aspecto, a interação fármaco-receptor também é considerada como reação química, embora com aspectos bem particulares<sup>16</sup>.

O trabalho de Hammett constituiu-se em grande avanço pelo fato de que, pela primeira vez, alguém se utilizou de constante de substituinte para descrever uma propriedade estrutural. A partir daí muitas outras propriedades estruturais seriam descritas em termos de constante de substituinte<sup>17-19</sup>.

Em 1956, Bruice e colaboradores<sup>20</sup> estudaram a atividade tireoidomimética de uma série de 47 análogos da tiroxina (I).



Para racionalizar as relações estrutura-atividade, eles formularam a Eq. (4),

$$\log AB = k\Sigma f + c, \quad (4)$$

em que AB é a atividade biológica medida em porcentagem de atividade tireoidomimética,  $f$  representa o conjunto de constantes de substituintes dos grupos X, Y e OR', Eq. (5), e  $k$  e  $c$  são constantes de ajuste.

$$\Sigma f = f_X + f_Y + f_{OR'} \quad (5)$$

As constantes representadas por  $f$  denotam contribuições intrínsecas dos diversos substituintes para a atividade biológica global. Cada valor de  $f$  traz embutido em si contribuições de diversas origens. Assim, os autores foram capazes de identificar duas contribuições de  $f$  nas correlações obtidas: a contribuição do substituinte para o momento dipolar total é a capacidade doadora ou receptora de elétrons. Fato curioso é a origem das constantes  $f$ . Não foram fornecidos os detalhes

matemáticos de sua determinação. Bruice e colaboradores<sup>20</sup> revelaram apenas que se tratam de parâmetros totalmente empíricos e que foram selecionados de forma similar à constante  $\sigma$  de Hammett. É louvável que os autores tenham sido capazes de interpretar as contribuições individuais das constantes de substituintes em termos de propriedades físico-químicas. Eles concluíram que quando os grupos X, Y e OR' são receptores de elétrons, ocorre diminuição da atividade biológica, enquanto que esta é intensificada na medida em que os grupos X e Y são capazes de formar pontes de hidrogênio. Também foi sugerida a possível influência do pKa a partir do comportamento dos valores de  $f$  para o substituinte R. Um ponto relevante é que este trabalho foi o primeiro a apresentar a utilização da medida da atividade biológica em escala logarítmica.

Entre 1960/1962, Zahradník e colaboradores<sup>21-24</sup> propuseram a Eq. (6) com o objetivo de correlacionar a intensidade da atividade biológica, em termos relativos, numa série homóloga de compostos do tipo R-X, em que R é cadeia hidrocarbônica alifática e X é grupo funcional.

$$\log(\tau_i / \tau_{Et}) = \alpha\beta \quad (6)$$

O termo  $\log(\tau_i / \tau_{Et})$  corresponde à atividade biológica relativa, no qual  $\tau_i$  representa a *eficácia biológica*, expressa em termos de concentração molar do  $i$ -ésimo membro da série homóloga, e  $\tau_{Et}$  é a concentração molar do homólogo etila, tomado como padrão. A semelhança guardada entre as Eqs. (6) e (2) resulta em propriedades similares entre as constantes  $\alpha$  e  $\rho$  e entre  $\beta$  e  $\sigma$ . Na Eq. (6),  $\alpha$  caracteriza a sensibilidade do sistema biológico e  $\beta$  é constante característica do substituinte R; note que, sendo a série homóloga, X permanece constante e o que varia é o comprimento da cadeia R. A constante  $\beta$  é independente tanto do grupo funcional X quanto do sistema biológico utilizado. Zahradník construiu uma escala para os valores de  $\beta$  a partir da adoção de R = Et como referência, cujo valor arbitrário foi  $\beta = 0,00$ . O grupo químico escolhido como referencial foi X = OH, cuja constante  $\alpha$  foi arbitrariamente feita igual a 1,00. Considerando o efeito tóxico de álcoois em camundongos, avaliado em termos de LD<sub>50</sub>, Zahradník pôde gerar valores de  $\beta$  para 25 grupos alifáticos e valores de  $\alpha$  para 39 sistemas biológicos diferentes. Segundo Hansch<sup>25</sup>, o ajuste dos valores experimentais à Eq. (6) foi bom, especialmente "onde efeitos eletrônicos e estéricos altamente específicos não são críticos". Hoje sabemos que  $\beta$  representa a variação relativa de energia livre associada à transferência de moléculas da fase aquosa para a fase orgânica. Em outras palavras,  $\beta$  mede a mesma propriedade da constante  $\pi$  de Hansch-Fujita<sup>17</sup> e, portanto, essas constantes estão altamente correlacionadas. A constante  $\alpha$  indica a importância relativa desse efeito de transferência de fase para o sistema biológico estudado.

Em 1962, Hansen<sup>26</sup> sugeriu a possibilidade de utilização da equação de Hammett como base para a investigação da natureza das interações fármaco-receptor. Ele imaginou que, se muitas substâncias exercem sua ação sobre biomacromoléculas, ou seja, através de alguma forma de interação química entre fármaco e receptor, deveria ser possível adaptar a equação de Hammett para detectar tais interações. Para comprovar essa hipótese, foram utilizados dados sobre a inibição do crescimento de diversos microrganismos. Hansen supôs que o mecanismo de inibição do crescimento se daria como resultado da interação bimolecular entre o inibidor e uma das enzimas responsáveis pelo processo de crescimento celular, provocando sua destruição. Portanto, a razão de destruição da enzima (E) pode ser representada por

$$dE/dt = k [E] [I] \quad (7)$$

em que  $k$  é a constante cinética,  $[E]$  e  $[I]$  são as concentrações molares da enzima e da substância inibidora, respectivamente.

Trabalhando-se com a mesma concentração enzimática, foram determinadas as concentrações de diferentes inibidores capazes de produzir a mesma razão de destruição enzimática. Isto equivale a fazer  $dE/dt$  constante e permite efetuar o seguinte rearranjo na Eq. (7):

$$\log k + \log[I] = \log(dE/dt) - \log[E] \quad (8)$$

Substituindo na Eq. (8) o valor de  $\log k$  da Eq. (3), obtemos

$$-\log[I] = \rho \sigma + a \quad (9)$$

em que

$$a = \log k_H + \log[E] - \log(dE/dt) \quad (10)$$

A Eq. (9) ficou conhecida como *equação biológica de Hammett*, devido à semelhança guardada com a Eq. (3). O próprio autor<sup>26</sup> reconheceu que dificilmente se encontraria um conjunto de substâncias, com substituintes bem variados, que pudessem ser descritos pela Eq. (9). Isto se deve às diversas suposições simplificadoras envolvidas em sua dedução<sup>15</sup>. Em primeiro lugar, nem toda interação enzima-inibidor obedece à cinética de reação bimolecular. A influência do transporte do inibidor através das biomembranas pode interferir na correlação entre  $\sigma$  e a atividade biológica. Presumiu-se que os substituintes presentes nas moléculas do inibidor exercem sua ação exclusivamente através de efeito eletrônico e este nem sempre pode ser o caso. Também se presumiu que apenas uma enzima é afetada por todo o grupo de substâncias testadas e isto pode não ocorrer ao longo da série de inibidores. Por último, foram ignoradas as contribuições devidas às variações de entropia relativas às diferentes orientações do inibidor na superfície da enzima<sup>15</sup>. Com todas essas limitações, nem mesmo Hansen esperou obter bom ajustamento dos resultados experimentais à sua teoria. No entanto, a abordagem de Hansen marcou importante estágio no estudo das relações estrutura-atividade.

O ano de 1964 assinalou o desenvolvimento do estudo das relações estrutura-atividade. Naquele ano surgiram os principais modelos clássicos de estudo QSAR. Vamos analisar cada um deles com algum pormenor. Sobre o histórico do desenvolvimento dos modelos utilizados em QSAR, podem ser consultados alguns artigos de revisão<sup>4,15,27</sup>.

### III. MODELO DE FREE-WILSON

O modelo de Free e Wilson<sup>28</sup> é baseado na hipótese de que, numa série de compostos que diferem entre si pela presença de substituintes em certas posições da molécula, a contribuição dos substituintes para a atividade biológica é aditiva e depende somente do tipo e da posição do substituinte. O modelo de Free-Wilson pode ser expresso pela Eq. (11)

$$\log AB_i = \sum a_{jk} X_{jk} + \mu \quad (11)$$

em que  $AB_i$  é a atividade biológica do  $i$ -ésimo composto e  $a_{jk}$  é a contribuição para a atividade biológica do substituinte  $X_k$ , em cada posição  $j$ .  $X_{jk}$  possui valor igual a 1 se o substituinte  $X_k$  está presente na posição  $j$ , caso contrário terá valor igual a zero.

A versão original do modelo de Free-Wilson<sup>28</sup>, também conhecida como *modelo clássico de Free-Wilson*, traz em seu bojo problemas matemáticos que muito limitaram sua aplicação. Vamos supor que o conjunto-teste possua  $n$  compostos, que difiram entre si pela presença, em certas posições, de  $m$  grupos substituintes. Isso resultará no aparecimento de  $n$  equações a  $m$  incógnitas mais o termo constante  $\mu$ , a média global dos valores da atividade biológica. No entanto este sistema de equações não pode ser resolvido devido à dependência linear presente. Uma vez que  $\mu$  é a média de todos os valores da

atividade biológica (Eq. (12)), a soma de todas as contribuições grupais  $a_{jk}$  precisa ser zero<sup>29</sup>.

$$\mu = \frac{1}{n} \sum \log AB_i \quad (12)$$

Assim, Free e Wilson definiram *equações de simetria*, para cada posição de substituição, baseados na suposição arbitrária de que a soma das contribuições em cada posição de substituição é zero.

$$\sum_j a_{jk} X_{jk} = 0 \quad (13)$$

(para cada valor de  $j$ , ou seja, cada posição de substituição)

Isto torna possível a resolução do sistema de equações, às custas de algum trabalho manual.

O processo de cálculo envolvido no modelo de Free-Wilson pode ser melhor compreendido através de um exemplo<sup>29-32</sup>. A tabela 1 mostra a estrutura e a atividade antiadrenérgica de 22 derivados da *N,N*-dimetil- $\alpha$ -bromofenetilamina<sup>33</sup>. A parte central da tabela 1 mostra a matriz estrutural dos compostos utilizados. Os espaços em branco representam zeros, que foram omitidos para maior clareza. Podemos notar que o número 1 corresponde à presença de um dado substituinte na posição *meta* ou *para* do composto-base.

A matriz estrutural da tabela 1 corresponde a um sistema de 22 equações a 12 incógnitas, mais o termo constante,  $\mu = 8,696$ , que é a média global dos valores da atividade biológica. Esse sistema não pode ser resolvido devido à dependência linear presente. A dependência linear tem origem no fato de que, em cada linha da matriz estrutural, a soma de  $X_{jk}$  em *meta* e *para* é sempre 1, ou seja, em cada composto há sempre um e somente um substituinte em cada uma das posições *meta* e *para*<sup>29</sup>.

Chegamos ao ponto em que, segundo Free e Wilson, há necessidade da introdução de equações de simetria na forma da Eq. (13). No caso da tabela 1, as equações de simetria serão:

$$\text{posição meta:} \\ 6 a_H + a_F + 4 a_{Cl} + 5 a_{Br} + a_I + 5 a_{Me} = 0 \quad (14)$$

$$\text{posição para:} \\ 6 a_H + 4 a_F + 4 a_{Cl} + 4 a_{Br} + a_I + 3 a_{Me} = 0 \quad (15)$$

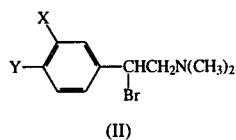
Observe que os coeficientes de  $a_k$  nas Eqs. (14) e (15) correspondem à soma dos valores de  $X_{jk}$  que aparecem na última linha da tabela 1. Utilizando-se as Eqs. (14) e (15), podemos representar cada um dos coeficientes  $a_k$  em função dos demais. Representando  $a_H$ , obtemos

$$\text{posição meta:} \\ a_H = - 1/6 a_F - 2/3 a_{Cl} - 5/6 a_{Br} - 1/6 a_I - 5/6 a_{Me} \quad (16)$$

$$\text{posição para:} \\ a_H = - 2/3 a_F - 2/3 a_{Cl} - 2/3 a_{Br} - 1/6 a_I - 1/2 a_{Me} \quad (17)$$

Podemos agora construir a tabela 2 eliminando da tabela 1 as colunas referentes a *meta*-H e *para*-H, e substituindo os zeros das linhas em que aparece o substituinte H pelos coeficientes do segundo membro das Eqs. (16) e (17). Procedendo dessa forma, a soma dos valores de  $X_{jk}$  torna-se zero, o que pode ser visto na última linha da tabela 2. Os termos da matriz estrutural da tabela 2 não mais apresentam dependência linear, que foi eliminada pelo procedimento acima. Podemos agora construir um modelo multidimensional linear através de regressão linear múltipla, em que a variável dependente é a atividade biológica,  $\log(1/ED_{50})$ , e as variáveis independentes são os valores de  $X_{jk}$ , presentes na tabela 2. Fazendo isso, obtemos a Eq. (18).

**Tabela 1.** Atividade antiadrenérgica das *N,N*-dimetil- $\alpha$ -bromofenetilaminas. Estrutura química, matriz estrutural e valores da atividade biológica<sup>29-32</sup>.



No.	Estrutura		Substituintes <i>meta</i> X					Substituintes <i>para</i> Y					log (1/ED <sub>50</sub> <sup>a</sup> ) obs.		
	X	Y	H	F	Cl	Br	I	Me	H	F	Cl	Br		I	Me
1	H	F	1							1					8,16
2	H	Cl	1								1				8,68
3	H	Br	1									1			8,89
4	H	I	1										1		9,25
5	H	Me	1											1	9,30
6	F	H		1					1						7,52
7	Cl	H			1				1						8,16
8	Br	H				1			1						8,30
9	I	H					1		1						8,40
10	Me	H						1	1						8,46
11	Cl	F			1					1					8,19
12	Br	F				1				1					8,57
13	Me	F						1		1					8,82
14	Cl	Cl			1							1			8,89
15	Br	Cl				1						1			8,92
16	Me	Cl						1				1			8,96
17	Cl	Br			1								1		9,00
18	Br	Br				1							1		9,35
19	Me	Br						1					1		9,22
20	Me	Me						1						1	9,30
21	Br	Me				1								1	9,52
22	H	H	1						1						7,46
$\Sigma$			6	1	4	5	1	5	6	4	4	4	1	3	191,32

<sup>a</sup> ED<sub>50</sub> = dose eficaz para 50% da população.

**Tabela 2.** Atividade antiadrenérgica das *N,N*-dimetil- $\alpha$ -bromofenetilaminas. Estrutura química, matriz estrutural e valores da atividade biológica<sup>31,32</sup>.

No.	Estrutura		Substituintes <i>meta</i> X					Substituintes <i>para</i> Y					log (1/C) obs.	log (1/C) calc. <sup>a</sup>	res.	
	X	Y	F	Cl	Br	I	Me	F	Cl	Br	I	Me				
1	H	F	-1/6	-2/3	-5/6	-1/6	-5/6	1						8,16	8,161	-0,001
2	H	Cl	-1/6	-2/3	-5/6	-1/6	-5/6		1					8,68	8,589	0,091
3	H	Br	-1/6	-2/3	-5/6	-1/6	-5/6			1				8,89	8,841	0,049
4	H	I	-1/6	-2/3	-5/6	-1/6	-5/6				1			9,25	9,250	0,000
5	H	Me	-1/6	-2/3	-5/6	-1/6	-5/6					1		9,30	9,077	0,223
6	F	H	1					-2/3	-2/3	-2/3	-1/6	-1/2		7,52	7,520	0,000
7	Cl	H		1				-2/3	-2/3	-2/3	-1/6	-1/2		8,16	8,028	0,132
8	Br	H			1			-2/3	-2/3	-2/3	-1/6	-1/2		8,30	8,255	0,045
9	I	H				1		-2/3	-2/3	-2/3	-1/6	-1/2		8,40	8,400	0,000
10	Me	H					1	-2/3	-2/3	-2/3	-1/6	-1/2		8,46	8,275	0,185
11	Cl	F		1				1						8,19	8,368	-0,178
12	Br	F			1			1						8,57	8,595	-0,025
13	Me	F					1	1						8,82	8,615	0,205
14	Cl	Cl		1					1					8,89	8,796	0,094
15	Br	Cl			1				1					8,92	9,023	-0,103
16	Me	Cl					1		1					8,96	9,043	-0,083
17	Cl	Br		1						1				9,00	9,048	-0,048
18	Br	Br			1					1				9,35	9,275	0,075
19	Me	Br					1			1				9,22	9,295	-0,075
20	Me	Me					1						1	9,30	9,531	-0,231
21	Br	Me			1								1	9,52	9,511	0,009
22	H	H	-1/6	-2/3	-5/6	-1/6	-5/6	-2/3	-2/3	-2/3	-1/6	-1/2		7,46	7,821	-0,361
$\Sigma$			0	0	0	0	0	0	0	0	0	0				

<sup>a</sup> Valores calculados através das Eqs. (19) e (21).

$\log(1/ED_{50}) = -0,553 (\pm 0,45) m-F - 0,045 (\pm 0,20) m-Cl + 0,182 (\pm 0,17) m-Br + 0,327 (\pm 0,45) m-I + 0,202 (\pm 0,17) m-Me - 0,283 (\pm 0,20) p-F + 0,144 (\pm 0,20) p-Cl + 0,397 (\pm 0,20) p-Br + 0,806 (\pm 0,45) p-I + 0,633 (\pm 0,24) p-Me + 8,696 (\pm 0,09)$

$n = 22 \quad r = 0,969 \quad s = 0,194 \quad F = 16,99 \quad (18)$

Na Eq. (18) e nas demais equações oriundas da análise de regressão linear,  $n$  é o número de objetos analisados, ou seja, compostos,  $r$  é o coeficiente de correlação,  $s$  é o desvio padrão e  $F$  é o teste de significância estatística de Fischer<sup>34</sup>. Os números entre parênteses correspondem sempre ao intervalo de confiabilidade de 95% dos respectivos coeficientes.

Pode-se notar que algumas variáveis presentes na Eq. (18), como por exemplo  $m-Cl$ ,  $m-I$  e  $p-Cl$ , apresentam coeficientes cujos intervalos de confiabilidade incluem o valor zero. Isso não significa que essas variáveis tenham de ser eliminadas da Eq. (18). Significa apenas que as variáveis  $m-Cl$ ,  $m-I$  e  $p-Cl$  apresentam coeficientes cujos valores não podem ser diferenciados do valor do coeficiente do substituinte de referência, que usualmente é H. Caso o substituinte de referência seja mudado, alguns coeficientes que antes não eram significativos podem tornar-se significativos e vice-versa.

O termo constante da Eq. (18) é a média global dos valores de  $\log(1/ED_{50})$  da tabela 1. Há dois procedimentos para conhecermos os valores de  $a_H$  para *meta-H* e *para-H*. O primeiro é substituir os valores de  $a_k$ , obtidos pela Eq. (18), nas Eqs. (16) e (17). Dessa forma, obtemos  $a_{m-H} = -0,252$  e  $a_{p-H} = -0,623$ . No entanto, esse procedimento não fornece o intervalo de confiabilidade de  $a_H$ . O segundo procedimento consiste em escolher dois outros substituintes, que não *meta-H* e *para-H*, removê-los da tabela 1, reconstruir as Eqs. (16) e (17) em termos dos substituintes escolhidos, substituir os novos valores de  $a_k$  na tabela 1 e fazer nova regressão linear múltipla. Este procedimento fornece  $a_{m-H} = -0,252 (\pm 0,16)$  e  $a_{p-H} = -0,623 (\pm 0,17)$ .

Pelo fato de o modelo de Free-Wilson considerar todas as possíveis variações estruturais que ocorrem com os substituintes presentes, exceto interações entre os próprios substituintes (termos cruzados), o coeficiente de correlação obtido através da análise de Free-Wilson representa o valor limite superior de correlação que pode ser obtido utilizando-se modelo aditivo e não-parabólico, como por exemplo o modelo de Hansch em sua versão não-parabólica (ver Eq. (21) adiante)<sup>30</sup>.

O modelo de Free-Wilson apresenta algumas limitações para sua aplicação<sup>16,29</sup>: (a) os compostos da série devem possuir a mesma estrutura fundamental, caso contrário as diferenças devem ser consideradas através de parâmetro extra; (b) no mínimo duas posições da molécula devem sofrer substituição ao longo da série para que seja possível incluir mais compostos do que variáveis; (c) os substituintes presentes em diferentes posições não devem interagir entre si, sob pena de tornar falso o conceito de aditividade; (d) o modelo é incapaz de gerar previsões acerca da atividade biológica de derivados que não estejam incluídos na série inicial.

Por outro lado, o modelo de Free-Wilson é ferramenta simples e eficaz, principalmente quando utilizado no estágio inicial do processo de otimização estrutural. O modelo lida com o conceito intuitivo de que cada substituinte acrescentado à estrutura química-base contribui para a atividade biológica global com uma fração que é característica do próprio substituinte.

Nas primeiras aplicações do modelo de Free-Wilson surgiram controvérsias a respeito da escala em que deveriam ser apresentados os valores da atividade biológica. Não havia certeza se a melhor escala era a linear ou a escala logarítmica. Hoje sabemos que o conceito de aditividade só é verdadeiro quando utilizamos a escala logarítmica de atividade biológica. Além disso, como o modelo utiliza a regressão linear múltipla para determinar a contribuição de cada substituinte, é necessário que os erros presentes nas medidas dos valores da atividade

biológica apresentem distribuição normal. Somente a escala logarítmica apresenta distribuição aproximadamente normal desses erros, o que não ocorre na escala linear<sup>29</sup>. Por último, pode ser demonstrado que somente a escala logarítmica de atividade biológica é parâmetro linearmente relacionado com a energia livre<sup>35</sup>.

#### IV. MODELO DE FUJITA-BAN

O modelo de Fujita-Ban<sup>36</sup> consiste, na verdade, na modificação do modelo de Free-Wilson de forma a simplificá-lo. Fujita e Ban interpretaram  $\mu$  como sendo o valor calculado da atividade biológica do composto não-substituído, que chamaremos  $\mu_H$ . Além disso, relacionaram todas as contribuições grupais ao hidrogênio, na mesma posição de substituição. Isso eliminou a necessidade das equações de simetria e simplificou o processo de cálculo. A equação modificada passou a ser

$$\log AB_i = \sum b_{jk} X_{jk} + \mu_H \quad (19)$$

em que  $b_{jk}$  é a contribuição para a atividade biológica do substituinte  $X_k$  na posição  $j$ , tomada em relação ao hidrogênio e  $\mu_H$  é o valor calculado da atividade biológica para o composto não-substituído (R-H).

Para aplicar o modelo de Fujita-Ban aos dados da tabela 1 basta eliminar as colunas referentes a *meta-H* e *para-H* e em seguida proceder à regressão linear múltipla diretamente. Obtemos assim a Eq. (20).

$\log(1/C) = -0,301 (\pm 0,50) m-F - 0,207 (\pm 0,29) m-Cl + 0,434 (\pm 0,27) m-Br + 0,579 (\pm 0,50) m-I + 0,454 (\pm 0,27) m-Me + 0,340 (\pm 0,30) p-F + 0,768 (\pm 0,30) p-Cl + 1,020 (\pm 0,30) p-Br + 1,429 (\pm 0,50) p-I + 1,256 (\pm 0,33) p-Me + 7,821 (\pm 0,27)$

$n = 22 \quad r = 0,969 \quad s = 0,194 \quad F = 16,99 \quad (20)$

Naturalmente os valores de  $b_{jk}$  modificaram-se em relação a  $a_{jk}$ , porém a qualidade estatística da Eq. (20) é a mesma da Eq. (18). O termo constante que aparece na Eq. (20) corresponde ao valor calculado da atividade biológica do composto não-substituído, R-H, como pode ser conferido na tabela 2.

Como foi salientado por Kubinyi<sup>29</sup>, "as relações entre a modificação de Fujita-Ban e a análise clássica de Free-Wilson são tão simples como a relação entre as escalas de temperatura Celsius e Kelvin. Somente a escolha do ponto de referência determina a magnitude dos valores". Outras propriedades e limitações podem ser vistas em excelentes artigos de revisão<sup>29,32</sup>.

Diversos exemplos de aplicação dos modelos de Free-Wilson e Fujita-Ban em química medicinal podem ser encontrados na literatura. Podemos citar os seguintes: atividade antibacteriana<sup>28,37-39</sup>, atividade anti-inflamatória<sup>40</sup>, atividade antimalárica<sup>41-43</sup>, atividade antitumoral<sup>44</sup>, atividade antivirótica<sup>45</sup>, atividade hipoglicêmica<sup>46</sup> e atividade tireoidomimética<sup>30</sup>.

Antes de abordar o modelo de Hansch-Fujita, é interessante relatar o modelo proposto por Bocek e Kopecky<sup>47,48</sup>. Este é um modelo que pode ser considerado semelhante ao de Free-Wilson<sup>30</sup>, diferindo apenas pela inclusão de termos cruzados do tipo  $e_{xy}$ . O termo  $e_{xy}$  é utilizado para detectar possíveis interações entre os substituintes X e Y localizados em posições de substituição próximas. Embora o modelo Bocek-Kopecky seja bastante citado em artigos de revisão, não há muita utilidade prática para ele, uma vez que seu uso implica na manipulação de grande número de variáveis<sup>30</sup>.

#### V. MODELO DE HANSCH-FUJITA

O modelo de Hansch-Fujita<sup>49</sup> para o estudo das relações entre estrutura química e atividade biológica é o modelo clássico mais utilizado e difundido. Este modelo se aplica ao estudo de séries de compostos que possuem estruturas químicas

semelhantes, porém diferindo entre si por um ou mais grupos substituintes em posições definidas. Segundo Hansch, a atividade biológica observada é o resultado da contribuição de diversos fatores, que se comportam de maneira independente. Desde o instante da introdução do fármaco no corpo de um indivíduo até o aparecimento do efeito farmacológico, ocorrem vários fenômenos. As moléculas do fármaco devem ser absorvidas, transportadas através de inúmeras barreiras constituídas por biomembranas, muitas vezes são conjugadas a proteínas, desconjugadas, solvatadas e dessolvatadas, até que finalmente possam interagir com o receptor. A partir daí, certas propriedades moleculares intrínsecas do fármaco e do receptor irão determinar a natureza das interações que irão ocorrer entre os dois. O modelo de Hansch-Fujita tenta caracterizar e isolar as contribuições de cada um dos fatores físico-químicos responsáveis pela atuação farmacológica do fármaco.

Os fatores físico-químicos que governam o transporte e as interações do fármaco com o sítio receptor são normalmente classificados em componentes lipofílicos, eletrônicos e estéricos. Cada contribuição é representada por uma constante de substituinte e a atividade biológica do conjunto de compostos é ajustada ao modelo linear multidimensional do tipo

$$\log (1/C) = a \log P + b \sigma + c E_s + d \quad (21)$$

em que a dependência da atividade biológica em relação à lipofilicidade é linear<sup>50</sup>, ou do tipo

$$\log (1/C) = -a (\log P)^2 + b \log P + c \sigma + d E_s + e \quad (22)$$

em que a dependência da atividade biológica em relação à lipofilicidade obedece ao modelo parabólico ou quadrático<sup>51,52</sup>. Nas Eqs. (21) e (22), P é o coeficiente de partição 1-octanol/água,  $\sigma$  é a constante eletrônica de Hammett<sup>12,13</sup>,  $E_s$  é a constante estérica de Taft<sup>18,53</sup> e C é a concentração molar do composto que produz resposta biológica padrão (IC<sub>50</sub>, ED<sub>50</sub>, LD<sub>50</sub>, etc.). As constantes a, b, c, d e e são os coeficientes obtidos pela análise de regressão linear múltipla. Além dos parâmetros citados cima, muitos outros já foram descritos e utilizados. Entre esses, incluem-se: parâmetro lipofílico  $\pi$ <sup>17,54-62</sup>, refratividade molar (RM)<sup>16,58,60,61,63-66</sup>, constantes derivadas de  $\sigma$  ( $\sigma^+$ ,  $\sigma^-$ ,  $\sigma^*$ , etc.)<sup>58,63,67-69</sup>, constantes eletrônicas de Swain e Lupton, que descrevem efeitos de campo e ressonância (F e R)<sup>19,61,66,70,71</sup>, pKa<sup>72,73</sup>, entropia<sup>74,75</sup>, etc. Algumas discussões sobre o uso de parâmetros físico-químicos em QSAR<sup>25,76,77</sup> bem como artigos que tratam de parâmetros lipofílicos<sup>78-82</sup>, eletrônicos<sup>19,83-85</sup> e estéricos<sup>86,87</sup> estão disponíveis na literatura.

As Eqs. (21) e (22) são ditas *relações lineares de energia livre*, pois os parâmetros que aparecem no segundo membro dessas equações são constantes que estão relacionadas a algum

estado de equilíbrio químico particular: log P, equilíbrio de partição 1-octanol/água;  $\sigma$ , equilíbrio de ionização de ácidos benzóicos;  $E_s$ , equilíbrio de esterificação de ácidos acetoacéticos. Apesar de o modelo de Hansch-Fujita, em sua versão original, ser uma relação linear de energia livre, isso não impediu que parâmetros não diretamente relacionados com energia livre fossem incluídos nas correlações com a atividade biológica. É o caso do uso das constantes estéricas de Verloop<sup>61,62,86,88-92</sup>, da polarizabilidade<sup>93,94</sup>, do volume molecular<sup>64,89,92,95-98</sup>, do momento dipolar<sup>99,100</sup>, dos ângulos torsionais<sup>74,75,101,102</sup>, da energia de orbitais de fronteira<sup>92,100,103</sup>, das cargas atômicas líquidas<sup>92,96,98,100,101,104-106</sup>, dos índices de reatividade derivados da teoria dos orbitais de fronteira<sup>92,94</sup>, da energia de rotação interna<sup>92</sup>, do deslocamento de bandas espectrais<sup>104</sup>, do número de átomos de carbono na cadeia do substituinte<sup>107</sup>, do raio de van der Waals<sup>68,108,109</sup>, etc.

Em relação às Eqs. (21) e (22), podemos dizer que os modelos lineares são mais adequados quando a atividade biológica é medida *in vitro*, como por exemplo, numa reação enzima-inibidor, ou em peças anatômicas simples e isoladas, como pequenas extensões de parede intestinal, fibras musculares, artérias, etc. A dependência não-linear da atividade biológica em relação à lipofilicidade é observada quando trabalhamos com o animal-teste vivo ou com peças anatômicas mais complexas. As tabelas 3 e 4 mostram exemplos de estudos QSAR em que foram utilizados os modelos lineares e parabólicos, respectivamente.

No caso de sistemas biológicos complexos, relações lineares entre lipofilicidade e atividade biológica somente são aplicáveis quando o conjunto de moléculas testadas possui extensão limitada de lipofilicidade. Se a lipofilicidade excede determinado limite, será observado decréscimo da atividade biológica<sup>79,81</sup>. As numerosas equações lineares obtidas para as relações atividade-lipofilicidade não são argumentos contra essa afirmação: elas apenas indicam que aquele valor limite encontra-se além da extensão de lipofilicidade incluídas nessas análises. Se não houver a lipofilicidade ótima em cada uma das séries, os compostos poderão apresentar atividade biológica infinita, caso sua lipofilicidade seja suficientemente grande<sup>79</sup>.

De forma geral, a dependência da atividade biológica em relação à lipofilicidade pode ser representada pelo gráfico da figura 1. O conjunto de pontos nas regiões A e C podem ser muito bem representados por modelos lineares, enquanto que os pontos na região B podem ser bem representados por modelo não-linear, como o parabólico. Quando tentamos ajustar o modelo parabólico a todo o conjunto de pontos no gráfico da figura 1, estamos representando bem a região B e mal as regiões A e C, pois ambos os lados da parábola são linhas curvas. Para corrigir essa ligeira inadequação do modelo parabólico,

**Tabela 3.** Exemplos de dependência linear da atividade biológica em relação à lipofilicidade<sup>50</sup>.

		log (1/C) = a log P + b <sup>a</sup>				
Tipo de Atividade	Compostos	a	b	n	r	s
MIC, nervo ciático de rã	ROH	1,050 (0,10)	0,258 (0,21)	8	0,995	0,142
Hemólise no cão	RNH <sub>2</sub> .HCl	0,953 (0,32)	1,625 (0,46)	5	0,984	0,318
Hemólise no homem	ROSO <sub>3</sub> <sup>-</sup> Na <sup>+</sup>	1,073 (0,27)	1,912 (0,42)	4	0,997	0,139
Conversão do citocromo P-450 a P-420, fígado de coelho	Anilinas	0,666 (0,14)	0,338 (0,19)	7	0,984	0,079
MIC, <i>B. subtilis</i>	Fenil metacrilatos	0,617 (0,11)	1,530 (0,45)	10	0,976	0,204
MED, hipnose em camundongo	Uréias	0,546 (0,09)	2,418 (0,12)	23	0,943	0,116
Desnaturação, DNA do fago T4	Amidas	0,412 (0,06)	0,451 (0,06)	5	0,997	0,037
I <sub>50</sub> , reação de Hill, cloroplasto	Carbamatos	0,850 (0,32)	0,540 (0,98)	9	0,919	0,213

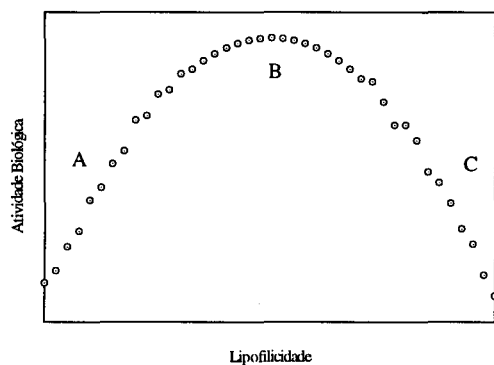
<sup>a</sup> a, b = coeficientes de ajuste; n = número de compostos analisados; r = coeficiente de correlação; s = desvio padrão; os números entre parênteses são os intervalos de 95% de confiabilidade dos respectivos coeficientes de ajuste; MIC = concentração inibitória mínima; MED = dose efetiva mínima; I<sub>50</sub> = concentração molar que produz 50% de inibição.

**Tabela 4.** Exemplos de aplicação do modelo parabólico a sistemas em que ocorre dependência não-linear da atividade biológica em relação à lipofilicidade<sup>52</sup>.

$\log (1/C) = a (\log P)^2 + b \log P + c^a$								
Tipo de Atividade	Compostos	a	b	c	n	r	s	
MKC, <i>D. pneumoniae</i>	RCHBrCOO <sup>-</sup>	-0,20 (0,12)	1,21 (0,71)	2,88 (0,92)	8	0,893	0,596	
MHC, hemácia de pombo	$\alpha$ -Monoglicérides	-0,24 (0,04)	1,98 (0,37)	0,17 (0,72)	8	0,989	0,107	
I <sub>50</sub> , mitocôndria do fígado de rato	Alquilguanidinas	-0,09 (0,05)	0,27 (0,18)	2,77 (0,21)	10 <sup>b</sup>	0,989	0,180	
MKC, <i>S. typhosa</i>	RN <sup>+</sup> Me <sub>3</sub>	-0,17 (0,10)	0,89 (0,24)	2,84 (0,53)	7	0,986	0,305	
MIC, <i>S. aureus</i>	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub> N <sup>+</sup> RMe <sub>2</sub>	-0,21 (0,04)	0,90 (0,14)	4,80 (0,14)	12	0,979	0,158	
MIC, <i>M. tuberculosis</i>	Aminopiridinas e anilinas	-0,57 (0,14)	2,73 (0,70)	2,22 (0,75)	20	0,898	0,474	
LD <sub>50</sub> , camundongo	Éteres	-0,29 (0,18)	1,12 (0,44)	1,85 (0,23)	25	0,864	0,208	
Inibição de crescimento, <i>A. niger</i>	5-Alquil-8-OH-quinolinas	-0,13 (0,03)	1,20 (0,31)	-1,84 (0,64)	11	0,955	0,106	

<sup>a</sup> Abreviações iguais às da Tab. 3, mais c = coeficiente de ajuste; MKC = concentração mortal mínima; MHC = concentração mínima para hemólise; LD<sub>50</sub> = dose letal para 50% da população.

<sup>b</sup> Foi excluído o derivado isopropil da Tabela III-12 da referência 52, para efeito de comparação com as equações bilineares da Tab. 5.



**Figura 1.** Dependência não-linear da atividade biológica em relação à lipofilicidade. As regiões A e C, isoladamente, podem ser muito bem representadas por modelos lineares. O modelo bilinear em geral é superior ao modelo parabólico na representação do conjunto total de pontos das regiões A, B e C.

Kubinyi<sup>110</sup> propôs o *modelo bilinear*, que é representado na Eq. (23),

$$\log (1/C) = a \log P - b \log (\beta P + 1) + c \quad (23)$$

em que a, b, c, e  $\beta$  são constantes de ajuste. Quando representada graficamente, a Eq. (23) apresenta as regiões ascendente e descendente lineares (regiões A e C) e o ápice redondo (região B). Assim, o modelo bilinear é geralmente superior ao modelo parabólico na representação da dependência não-linear da atividade biológica em relação à lipofilicidade. Exemplos de relações bilineares são apresentados na tabela 5. Compare a qualidade de ajuste das equações bilineares da tabela 5 com as equações parabólicas da tabela 4.

As Eqs. (21) e (22) têm origem empírica. No entanto, Hansch propôs um modelo, chamado *random walk* (caminhada aleatória)<sup>49</sup>, que foi capaz de justificar a estrutura dessas equações. O modelo bilinear é matematicamente mais rigoroso e baseia-se na hipótese multicompartmental<sup>110</sup>.

**Tabela 5.** Exemplos de aplicação do modelo bilinear a sistemas em que ocorre dependência não-linear da atividade biológica em relação à lipofilicidade<sup>32</sup>.

$\log (1/C) = a \log P + b \log (\beta P + 1) + c^a$								
Tipo de Atividade	Compostos	a	b	c	log $\beta$	n	r	s
MKC, <i>D. pneumoniae</i>	RCHBrCOO <sup>-</sup>	0,79 (0,62)	1,96 (1,22)	2,77 (0,91)	-3,60	8	0,934	0,530
MHC, hemácias de pombo	$\alpha$ -Monoglicérides	0,83 (0,14)	1,65 (0,15)	1,52 (0,35)	-4,10	8	0,998	0,049
I <sub>50</sub> , mitocôndria do fígado de rato	Alquilguanidinas	0,83 (0,11)	0,93 (0,43)	3,50 (0,34)	0,59	10	0,997	0,103
MKC, <i>S. typhosa</i>	RN <sup>+</sup> Me <sub>3</sub>	1,00 (0,23)	1,79 (1,14)	2,51 (0,33)	-2,31	7	0,977	0,169
MIC, <i>S. aureus</i>	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub> N <sup>+</sup> RMe <sub>2</sub>	1,05 (0,19)	1,51 (0,19)	4,76 (0,12)	-1,44	12	0,993	0,100
MIC, <i>M. tuberculosis</i>	Aminopiridinas e anilinas	2,16 (0,50)	3,62 (0,81)	2,27 (0,66)	-2,05	20	0,922	0,431
LD <sub>50</sub> , rato	Éteres	2,64 (1,28)	2,51 (1,46)	2,50 (0,29)	-0,08	25	0,875	0,205
Inibição de crescimento, <i>A. niger</i>	5-Alquil-8-OH-quinolinas	0,38 (0,11)	0,93 (0,24)	-0,60 (0,36)	-5,13	11	0,977	0,082

<sup>a</sup> Abreviações iguais às da Tab. 4, mais c = coeficiente de ajuste;  $\beta$  = termo bilinear de ajuste.

Existem diversas hipóteses que tentam explicar a dependência não linear entre a atividade biológica e a lipofilicidade<sup>16,35,79</sup>. É possível que todas elas contribuam em maior ou menor extensão para esse tipo de comportamento. A hipótese que mais deve contribuir, porém, está relacionada à dependência da solubilidade aquosa na lipofilicidade; se para congêneres lipofílicos a solubilidade em água diminui mais rápido do que aumenta a atividade, irá ocorrer quebra repentina na relação linear entre a atividade biológica e a lipofilicidade. Isto acontece no instante em que a solubilidade do fármaco na fase aquosa for tão pequena que ele não mais será capaz de caminhar através das sucessivas barreiras do organismo, constituídas naturalmente por meios aquosos e lipofílicos. Nesse ponto, o fármaco ficará completamente retido nas primeiras barreiras lipofílicas e não poderá chegar ao local primário de sua ação<sup>79</sup>.

Existem três hipóteses que regem a análise correlacional de Hansch-Fujita<sup>16,111</sup>: (a) as propriedades moleculares relacionadas com a atividade biológica podem ser separadas e quantificadas. Essas propriedades moleculares, por hipótese, são representáveis por parâmetros físico-químicos experimentais ou calculados; (b) a atividade biológica de interesse pode ser medida quantitativamente. Está implícito que todos os compostos estudados possuem o mesmo mecanismo de ação, o que pode ser evidenciado pelo grau de paralelismo das curvas dose-resposta de cada um dos compostos; (c) a relação entre a estrutura química, representada pelos parâmetros físico-químicos, e a atividade biológica pode ser descrita por equação matemática simples.

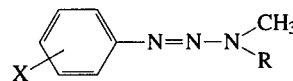
Essas hipóteses são mais realísticas quando o conjunto de compostos possuir as seguintes características<sup>16</sup>: (a) os compostos devem ser análogos estruturais que agem pelo mesmo mecanismo de ação; (b) deve haver disponibilidade de parâmetros físico-químicos confiáveis para todos os compostos; (c) o conjunto-teste deve ser grande o bastante para que os métodos estatísticos sejam aplicáveis; (d) a variância da atividade biológica no conjunto-teste deve ser apreciável, usualmente maior do que duas unidades de log (1/C); (e) os parâmetros físico-químicos devem possuir apreciável variação ao longo do conjunto de compostos, mas a colinearidade entre os parâmetros precisa ser evitada<sup>112</sup>.

O modelo de Hansch-Fujita apresenta diversos pontos fortes e também alguns pontos fracos<sup>16</sup>: (a) a característica mais importante desse modelo é que ele é baseado na suposição de que as interações que ocorrem entre duas moléculas pequenas são semelhantes às que ocorrem entre uma molécula pequena (fármaco) e uma biomolécula (receptor). Como exemplo, pode-se citar o uso do coeficiente de partição 1-octanol/água como forma de modelar o comportamento da molécula do fármaco frente às estruturas lipofílicas encontradas no organismo e no meio aquoso circundante. Uma consequência direta da suposição acima é que podemos utilizar dados extraídos a partir de modelos simples para prever o comportamento de sistemas biológicos complexos. Além disso, o modelo proporciona alguma compreensão sobre o comportamento químico do fármaco no sistema biológico; (b) o modelo permite a elaboração de previsões quantitativas cujos intervalos de confiança são estatisticamente definidos; (c) é relativamente simples de usar e barato; (d) ao contrário do modelo de Free-Wilson, o modelo de Hansch-Fujita permite fazer previsões da atividade biológica de compostos não incluídos na série inicial.

Entretanto, o modelo de Hansch-Fujita apresenta algumas limitações<sup>16</sup>: (a) devem haver parâmetros físico-químicos disponíveis para todos os substituintes da série. Muitas vezes um composto precisa ser eliminado da série porque um de seus parâmetros não foi localizado (valores de diversos parâmetros físico-químicos para muitos grupos substituintes podem ser obtidos na literatura<sup>113-118</sup>) ou não pôde ser adequadamente avaliado; (b) é necessário incluir número relativamente elevado de compostos a fim de que se possa obter equações confiáveis e que todos os possíveis preditores sejam testados.

A regra geral é que deve haver cerca de cinco compostos para cada variável incluída no modelo a fim de evitar a correlação por coincidência (*chance correlation*)<sup>112,119</sup>; (c) a aplicação da abordagem extratermodinâmica requer algum conhecimento estatístico, campo em que os químicos medicinais nem sempre se sentem à vontade; (d) a mais séria limitação da abordagem extratermodinâmica é que os modelos construídos a partir das interações entre moléculas pequenas são imperfeitos para o sistema biológico. Alguns bons artigos introdutórios à abordagem extratermodinâmica<sup>120-124</sup> e de revisão<sup>15,35</sup>, bem como livros<sup>16,125,126</sup> estão disponíveis.

Como exemplo da aplicação do modelo de Hansch-Fujita, vamos considerar a pesquisa de derivados triazínicos com atividade anticâncer. Hansch e colaboradores<sup>63</sup> estudaram a atividade anticâncer de 64 derivados da 1-(X-fenil)-3-R-3-metil-triazina, (III),



(III)

em que X e R representam grupos químicos variados. A Eq. (24) melhor representou a atividade de 61 compostos do grupo original.

$$\log (1/C_{140}) = 0,100 (\pm 0,08) \log P - 0,042 (\pm 0,02) (\log P)^2 - 0,312 (\pm 0,11) \sum \sigma^+ - 0,178 (\pm 0,08) MR_{(2,6)} + 0,391 (\pm 0,18) E_{s(R)} + 4,124 (\pm 0,27)$$

$$n = 61 \quad r = 0,836 \quad s = 0,191 \quad F = 25,52 \quad \log P_{\text{ótimo}} = 1,18 \quad (24)$$

Na Eq. (24),  $C_{140}$  é a concentração molar capaz de produzir  $T/C = 140$ , em que T é o tempo de vida de animais-teste inoculados com células leucêmicas e tratados com o fármaco, e C é o tempo de vida de animais-teste, usados como controle, inoculados com células leucêmicas e que não receberam o fármaco.  $\sum \sigma^+$  é a soma das constantes  $\sigma^+$  dos substituintes X,  $MR_{(2,6)}$  é a soma das refratividades molares dos substituintes X presentes nas posições *orto* do grupo fenil e  $E_{s(R)}$  é a constante de Taft do grupo R.

A Eq. (24) e também outras equações obtidas a partir da análise de diferentes derivados triazínicos<sup>63</sup> confirmam que o log P ideal desses compostos contra leucemia gira em torno de 1. Isso impede que se possa aumentar a atividade antileucêmica das triazinas por manipulação do parâmetro lipofílico. O sinal negativo do coeficiente de  $MR_{(2,6)}$  sugere que grupos grandes ligados na posição *orto* do grupo fenil diminuam a atividade.

O sinal negativo do coeficiente de  $\sum \sigma^+$  indica que grupos X capazes de ceder elétrons via ressonância aumentam a atividade. No entanto, a velocidade de hidrólise dos compostos derivados de (III) é correlacionada com o efeito eletrônico do substituinte X, conforme foi demonstrado por Kolar e Preussmann<sup>127</sup>. A Eq. (25) foi obtida por Hansch e colaboradores<sup>63</sup> com base nos dados de Kolar e Preussmann para a reação  $X-C_6H_4N=NN(CH_3)_2 + H_2O \rightarrow X-C_6H_4N_2^+ + HN(CH_3)_2$ .

$$\log (K_X/K_H) = -4,42 (\pm 0,29) \sigma - 0,016 (\pm 0,13)$$

$$n = 14 \quad r = 0,995 \quad s = 0,171 \quad (25)$$

$K_X$  e  $K_H$  são as constantes de hidrólise para os derivados de (III) que apresentam anel aromático substituído e não-substituído, respectivamente. Uma vez que a velocidade de hidrólise é enormemente intensificada pela presença de grupos elétron-doadores, de acordo com o sinal negativo do coeficiente de  $\sigma$ , é praticamente impossível ir além de 4-metoxi, cuja presença em (III) proporciona meia-vida de 12 minutos, no uso de grupos elétron-doadores.



Outro problema que deve ser atacado é a toxicidade dos fármacos em estudo. Paralelamente à maximização da atividade biológica, deve-se minimizar a toxicidade. No presente exemplo, Hansch e Hatheway<sup>128</sup> estudaram a toxicidade de 11 derivados de (III) e obtiveram a Eq. (26).

$$\log(1/LD_{50}) = -0,0241 (\pm 0,013) (\log P)^2 - 0,264 (\pm 0,16) \sigma^+ + 3,490 (\pm 0,12)$$

$$n = 11 \quad r = 0,913 \quad s = 0,110 \quad F = 20,06 \quad (26)$$

LD<sub>50</sub> é a concentração molar capaz de matar 50% da população-teste. Comparando-se as Eqs. (24) e (26), nota-se que o mesmo efeito eletrônico que intensifica a atividade também intensifica a toxicidade dos derivados de (III). Hansch e colaboradores concluíram não ser possível separar características estruturais para a atividade e toxicidade e desaconselharam futuros esforços no sentido de sintetizar novos derivados de (III).

No exemplo acima, embora a busca por fármacos mais eficazes tenha sido frustrada, há um aspecto muito importante que deve ser ressaltado. O número de possíveis combinações de compostos que podem ser sintetizados, variando-se X e R em (III), pode facilmente chegar à casa dos milhões. A síntese aleatória de derivados de (III) na busca de um exemplar que apresentasse boa atividade e baixa toxicidade, certamente consumiria anos de trabalho, muito dinheiro e, com muita possibilidade, resultaria em malogro. No entanto, após poucas dezenas de sínteses, o modelo de Hansch foi capaz de identificar a inviabilidade dos derivados de (III) como fármacos no combate à leucemia.

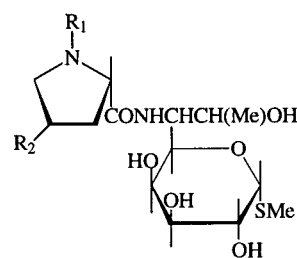
Existe farta literatura com aplicações do modelo de Hansch-Fujita. Podemos citar alguns exemplos: interação hapteno-anticorpo<sup>129</sup>, inibição da diidrofolato redutase<sup>130,131</sup>, atividade anti-inflamatória<sup>60</sup>, inibição da uroquinase<sup>66</sup>, atividade anti-hipertensiva<sup>91</sup>, atividade citostática<sup>99</sup>, inibição da tiopurina metiltransferase<sup>132</sup>, atividade antitumoral<sup>70,133</sup>, atividade anticâncer<sup>63</sup> e inibição da DNA polimerase<sup>65</sup>.

## VI. ABORDAGEM MISTA

Hoje sabemos que o modelo de Free-Wilson e o modelo linear de Hansch-Fujita, Eq. (21), são teoricamente interrelacionados e numericamente equivalentes<sup>29,134</sup>. Também sabemos que o modelo de Bocek-Kopecky é matematicamente equivalente ao modelo parabólico de Hansch-Fujita, Eq. (22)<sup>29,30,134</sup>. A partir do conhecimento dessas interrelações, Kubinyi<sup>39</sup> desenvolveu o que chamou de *abordagem mista*, baseada nos modelos de Free-Wilson e Hansch-Fujita. A abordagem mista é a combinação desses dois modelos e faz uso das vantagens de cada um deles, ampliando a aplicabilidade das análises de Hansch-Fujita e Free-Wilson. Segundo Kubinyi<sup>39</sup> existem duas situações especiais em que o emprego da abordagem mista pode ser especialmente favorável. A primeira situação é quando existe pequeno número de substituintes diferentes numa posição definida da molécula, o que frequentemente torna impossível obter correlações significativas pela análise de Hansch-Fujita; ou no caso de equações de Hansch-Fujita, obtidas a partir de contribuições de grupo oriundas da análise de Free-Wilson<sup>36,37</sup> que incluem grande número de determinações unitárias (*single points determination*)<sup>29</sup>. Essas equações não são muito confiáveis porque as contribuições de grupo derivadas de grande número de compostos e contribuições de grupo incluindo o erro experimental de um único composto possuem o mesmo peso. A segunda aplicação da abordagem mista é o uso de variáveis indicadoras (*indicator variables* ou *dummy variables*) na análise de Hansch-Fujita<sup>135</sup>. Variáveis indicadoras são variáveis especiais capazes de assumir apenas dois valores diferentes: 0 e 1. O valor 1 indica a presença de certo evento, enquanto que o valor 0 indica sua ausência. As variáveis indicadoras são normalmente utilizadas para detectar

propriedades estruturais específicas (impedimento estérico ou mudanças conformacionais), interações específicas entre substituintes (pontes de hidrogênio) e para detectar todas as mudanças nas contribuições polares, eletrônicas e estéricas no caso de um grupo X que é substituído por um grupo Y<sup>16</sup>. Hansch e Yoshimoto<sup>135</sup> interpretaram a utilização de variáveis na análise de Hansch-Fujita como sendo a "abordagem extratermodinâmica auxiliada pelo que é, às vezes, chamado de modelo de Free-Wilson". Eles definiram a abordagem de Free-Wilson como "relacionamento quantitativo entre estrutura e atividade formulado em termos de variáveis indicadoras". Segundo Kubinyi<sup>39</sup>, o uso de certos tipos de variáveis indicadoras na análise de Hansch-Fujita é o resultado da aplicação da abordagem mista.

Como exemplo da abordagem mista, podemos citar o trabalho de Martin e Lynn<sup>136</sup>, em que foi estudada a atividade antimicrobiana de lincomicinas N-alquil-substituídas, (IV), em *Sarcina lutea*<sup>137</sup>.



(IV)

Na estrutura (IV), R<sub>1</sub> é um grupo metil (14 compostos) ou etil (11 compostos) e R<sub>2</sub> é um substituinte alquil. O modelo parabólico descreveu muito bem a atividade dos derivados N-Metil e N-Etil, de acordo com as Eqs. (27) e (28), que foram recalculadas a partir dos dados da tabela III da referência 136,

N-Me  
 $\log(1/C) = -0,274(\pm 0,064)\pi^2 + 1,458(\pm 0,29)\pi - 0,238(\pm 0,31)$   
 $n = 14 \quad r = 0,961 \quad s = 0,155 \quad F = 68,27 \quad \pi_{\text{ótimo}} = 2,66 \quad (27)$

N-Et  
 $\log(1/C) = -0,284(\pm 0,073)\pi^2 + 1,731(\pm 0,40)\pi - 1,102 (\pm 0,53)$   
 $n = 11 \quad r = 0,963 \quad s = 0,168 \quad F = 51,75 \quad \pi_{\text{ótimo}} = 3,05 \quad (28)$

em que C é a concentração molar do composto. Logo verificou-se que os derivados *trans* apresentavam maior atividade do que os derivados *cis*. Portanto, foi possível descrever melhor a atividade desses compostos criando-se uma variável indicadora da configuração *trans*, T, cujo valor é 1 para os derivados *trans* e 0 para os *cis*. As equações recalculadas são:

N-Me  
 $\log(1/C) = -0,263 (\pm 0,049) \pi^2 + 1,402 (\pm 0,23) \pi + 0,196 (\pm 0,14)$   
 $T - 0,282 (\pm 0,24)$   
 $n = 14 \quad r = 0,981 \quad s = 0,117 \quad F = 83,82 \quad \pi_{\text{ótimo}} = 2,67 \quad (29)$

N-Et  
 $\log(1/C) = -0,271 (\pm 0,056) \pi^2 + 1,642 (\pm 0,31) \pi + 0,216 (\pm 0,18)$   
 $T - 1,070 (\pm 0,41)$   
 $n = 11 \quad r = 0,983 \quad s = 0,124 \quad F = 65,77 \quad \pi_{\text{ótimo}} = 3,03 \quad (30)$

Devido à semelhança estrutural entre as Eqs. (29) e (30), é possível combiná-las introduzindo nova variável indicadora, Et, cujo valor 1 assinala a presença do grupo N-Etil e 0 a presença de N-Metil.

$$\log (1/C) = -0,242 (\pm 0,043) \pi^2 + 1,388 (\pm 0,22) \pi + 0,232 (\pm 0,14) T - 0,203 (\pm 0,14) Et - 0,413 (\pm 0,26)$$

$$n = 25 \quad r = 0,961 \quad s = 0,162 \quad F = 60,30 \quad \pi_{\text{ótimo}} = 2,86 \quad (31)$$

Dessa forma, é possível analisar, na mesma equação, compostos que apresentam características estruturais diferentes.

Outros exemplos do uso de variáveis indicadoras em QSAR: inibição do complemento<sup>54,55</sup>, inibição da anidrase carbônica<sup>138,139</sup>, inibição da quimotripsina<sup>140-142</sup>, atividade antifúngica<sup>143</sup>, atividade antimalárica<sup>56</sup>, atividade antialérgica<sup>144</sup> e inibição da diidrofolato redutase<sup>57,145,146</sup>.

## VII. CONSIDERAÇÕES FINAIS

Os modelos clássicos de estudo QSAR foram concebidos num momento histórico em que a visualização do evento da interação fármaco-receptor não era possível. Na equação

$$\text{Fármaco} + \text{Receptor} = \text{Atividade Biológica}$$

somente eram conhecidos os termos Fármaco e Atividade Biológica. Não havia meios confiáveis de concluir-se o que quer que fosse a respeito da geometria molecular bioativa do fármaco e também sobre a estrutura do receptor farmacológico.

Com o surgimento de computadores de alta performance para cálculo numérico e visualização gráfica, ocorreu verdadeira revolução na abordagem do problema de relações estrutura-atividade, o que hoje denominamos QSAR-3D. Hoje, pelo menos no caso de enzimas, é possível visualizar as interações que ocorrem entre os grupos químicos da molécula do fármaco e as diversas cadeias laterais dos aminoácidos que formam a estrutura do receptor farmacológico. O próprio Corwin Hansch viu a confirmação da validade de seu modelo ante a imagem gráfica de um inibidor no sítio ativo de uma enzima: "Vendo a beleza da região hidrofóbica em vermelho na figura I demandada pela equação 2 (presentes no corpo do artigo<sup>139</sup>), após tantos anos imaginando qual o verdadeiro significado dos nossos parâmetros QSAR no plano molecular, sentimos enorme sensação de alívio"<sup>139</sup>.

## AGRADECIMENTOS

Gostaria de deixar registrado meus sinceros agradecimentos pela ajuda e apoio recebidos dos seguintes pesquisadores e amigos: Dr. Yuji Takahata, orientador de minha dissertação de mestrado que inclui parte deste trabalho, Dr. Andrejus Korolkovas e Dra. Maria Inês Nogueira de Camargo Harris, que dispuseram-se a ceder horas de seu lazer na revisão deste texto, Dra. Antonia Tavares do Amaral, em cujo laboratório iniciei o presente trabalho de revisão, e Dr. Hugo Kubinyi, pelas opiniões emitidas e pela gentil cessão de seu programa BILIN, utilizado para reproduzir as equações de regressão apresentadas no corpo deste artigo. Gostaria de agradecer também o apoio financeiro do CNPq e CAPES.

## REFERÊNCIAS

1. Crum-Brown, A.; Fraser, T. R. *Trans. Roy. Soc. Edinburgh* **1868-1869**, 25, 151, 693.
2. Tute, M. S. In *Physical Chemical Properties of Drugs*; Yalkowsky, S.H.; Sinkula, A. A.; Valvani, S.C., Eds.; Marcel Dekker: New York, 1980; p 141.
3. Richet, C. C. R. *Seances Soc. Biol.* **1893**, 9, 775.
4. Rekker, R. F. *Quant. Struct.-Act. Relat.* **1992**, 11, 195.
5. Overton, E. *Z. Phys. Chem., Stoechiom. Verwandtschaftsl.* **1897**, 22, 189.
6. Overton, E. *Vierteljahrsschr. Naturforsch. Ges. Zuerich.* **1899**, 44, 88.
7. Overton, E. In *Studien über die Narcose*; Fischer G., Ed.; Jena, 1901.
8. Meyer, H. *Arch. Exp. Pathol. Pharmacol.* **1899**, 42, 109.
9. Ferguson, J. *Proc. Royal Soc. London, Ser. B* **1939**, 127, 387.
10. Korolkovas, A. *Fundamentos de Farmacologia Molecular*; Universidade de São Paulo: São Paulo, 1974.
11. Korolkovas, A. *Essentials of Medicinal Chemistry*; 2nd ed.; Wiley-Interscience: New York, 1988.
12. Hammett, L. P. *J. Am. Chem. Soc.* **1937**, 59, 96.
13. Hammett, L. P. *Physical Organic Chemistry*; McGraw-Hill: New York, 1940.
14. Ritchie, C. D.; Sager, W. F. *Prog. Phys. Org. Chem.* **1964**, 2, 323.
15. Tute, M. S. *Adv. Drug. Res.* **1971**, 6, 1.
16. Martin, Y. C. *Quantitative Drug Design: A Critical Introduction*; Marcel Dekker: New York, 1978.
17. Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, 86, 5175.
18. Taft, R. W. *J. Am. Chem. Soc.* **1952**, 74, 3120.
19. Swain, C. G.; Lupton, Jr., E. C. *J. Am. Chem. Soc.* **1968**, 90, 4328.
20. Bruce, T. C.; Kharasch N.; Winzler, R. J. *Arch. Biochem. Biophys.* **1956**, 62, 305.
21. Zahradník, R.; Chvapil, M. *Experientia* **1960**, 16, 511.
22. Zahradník, R. *Experientia* **1962**, 18, 534.
23. Zahradník, R. *Arch. int. Pharmacodyn.* **1962**, 125, 311.
24. Chvapil, M.; Zahradník, R.; Cmuchalova, B. *Arch. int. Pharmacodyn.* **1962**, 135, 330.
25. Hansch, C. *Ann. Rep. Med. Chem.* **1966**, 34, 347.
26. Hansen, O. R. *Acta Chem. Scand.* **1962**, 16, 1593.
27. van der Waterbeemd, H. *Quant. Struct.-Act. Relat.* **1992**, 11, 200.
28. Free, S. M. Jr.; Wilson, J. W. *J. Med. Chem.* **1964**, 7, 395.
29. Kubinyi, H. *Quant. Struct.-Act. Relat.* **1988**, 7, 121.
30. Kubinyi, H.; Kehrhahn, O.-H. *J. Med. Chem.* **1976**, 19, 578.
31. Kubinyi, H.; Kehrhahn, O.-H. *J. Med. Chem.* **1976**, 19, 1040.
32. Kubinyi, H. *The Free-Wilson Method and its Relationship to the Extrathermodynamic Approach*, In *Comprehensive Medicinal Chemistry. The Rational Design, Mechanistic Study, and Therapeutic Application of Chemical Compounds*; Hansch, C.; Sammes, P. G.; Taylor, J. B., Eds.; Pergamon: New York, 1990; Vol. 4, p 589.
33. Graham, J. P. D.; Karrar, M. A. *J. Med. Chem.* **1963**, 6, 103.
34. Draper, N. R.; Smith, H. *Applied Regression Analysis*; John Wiley & Sons: New York, 1981.
35. Fujita, T. *The Extrathermodynamic Approach in Drug Design*, In *Comprehensive Medicinal Chemistry. The Rational Design, Mechanistic Study, and Therapeutic Application of Chemical Compounds*; Hansch, C.; Sammes, P. G.; Taylor, J. B., Eds.; Pergamon: New York, 1990; Vol. 4, p 497.
36. Fujita, T.; Ban, T. *J. Med. Chem.* **1971**, 14, 148.
37. Cammarata, A.; Yau, S. J. *J. Med. Chem.* **1970**, 13, 93.
38. Martin, Y. C.; Jones, P. H.; Perun, T. J.; Grundy, W. E.; Bell, S.; Bower, R.R.; Shipkowitz, N. L. *J. Med. Chem.* **1972**, 15, 635.
39. Kubinyi, H. *J. Med. Chem.* **1976**, 19, 587.
40. Buckler, R. T. *J. Med. Chem.* **1972**, 15, 578.
41. Hudson, D. R.; Bass, G. E.; Purcell, W. P. *J. Med. Chem.* **1970**, 13, 1184.
42. Craig, P. N. *J. Med. Chem.* **1972**, 15, 144.
43. Craig, P. N.; Hansch, C. *J. Med. Chem.* **1973**, 16, 661.
44. Purcell, W. P.; Clayton, J. M. *J. Med. Chem.* **1968**, 11, 199.
45. Thomas, J.; Berkoff, C. E.; Flag, W. B.; Gallo, J. J.; Haff, R. F.; Pinto, C. A.; Pellerano, C.; Savini, L. *J. Med. Chem.* **1975**, 18, 245.

46. Smithfield, W. R.; Purcell, W. P. *J. Med. Chem.* **1967**, *56*, 577.
47. Bocek, K.; Kopecký, J.; Krivucova, M. Vlachova, D. *Experientia* **1964**, *20*, 667.
48. Kopecký, J.; Bocek, K.; Vlachova, D. *Nature* **1965**, *207*, 981.
49. Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* **1964**, *86*, 1616.
50. Hansch, C.; Dunn, III, W. J. *J. Pharm. Sci.* **1972**, *61*, 1.
51. Hansch, C.; Stewart, A. R.; Anderson, S. M.; Bentley, D. *J. Med. Chem.* **1967**, *11*, 1.
52. Hansch, C.; Clayton, J. M. *J. Pharm. Sci.* **1973**, *62*, 1.
53. Taft, R. W. *Steric Effects in Organic Chemistry* Newman, M. S., Ed.; Wiley: New York, 1956; p 556.
54. Silipo, C.; Hansch, C. *J. Med. Chem.* **1976**, *19*, 62.
55. Hansch, C.; Yoshimoto, M.; Doll, M. H. *J. Med. Chem.* **1976**, *19*, 1089.
56. Kim, K. H.; Hansch, C.; Fukunaga, J. Y.; Steller, E. E.; Jow, P. Y. C.; Craig, P.N.; Page, J. *J. Med. Chem.* **1979**, *22*, 366.
57. Hathaway, B. A.; Guo, Z.; Hansch, C.; Delcamp, T. J.; Susten, S. S.; Freisheim, J. H. *J. Med. Chem.* **1984**, *27*, 144.
59. Lessassie, C. D.; Fang, Z.-X.; Li, R.-L.; Hansch, C.; Debnath, G.; Klein, T. E.; Langridge, R.; Kaufman, B. T. *J. Med. Chem.* **1989**, *32*, 1895.
60. Kim, K.H.; Martin, Y. C.; Norris, B.; Young, P. R.; Carter, G. W.; Haviv, F.; Walters, R. L. *J. Pharm. Sci.* **1990**, *79*, 609.
61. Caliendo, G.; Cirino, G.; Greco, G.; Novellino, E.; Perissutti, E.; Pinto, A.; Santagada, V.; Silipo, C.; Sorrentino, L. *Eur. J. Med. Chem.* **1994**, *29*, 381.
62. Bordi, F.; Mor, M.; Morini, G.; Plazzi, P.V.; Silva, C.; Vitali, T. *Fármaco* **1994**, *49*, 153.
63. Hatheway, G. J.; Hansch, C.; Kim, K. H.; Milstein, S. R.; Schmidt, C. L.; Smith, R. N.; Quinn, F. R. *J. Med. Chem.* **1978**, *21*, 563.
64. Carotti, A.; Casini, G.; Hansch, C. *J. Med. Chem.* **1984**, *27*, 1427.
65. Wright, G. E.; Gambino, J. J. *J. Med. Chem.* **1984**, *27*, 181.
66. Yang, H.; Henkin, J.; Kim, K. H.; Greer, J. *J. Med. Chem.* **1990**, *33*, 2956.
67. Fujita, T. *J. Med. Chem.* **1973**, *16*, 923.
68. Unger, S. H.; Hansch, C. *J. Med. Chem.* **1973**, *16*, 745.
69. Panthananickal, H.; Hansch, C.; Leo, A.; Quinn, F. R. *J. Med. Chem.* **1978**, *21*, 16.
70. Dunn III, W. J.; Greenberg, M. J.; Callejas, S. S. *J. Med. Chem.* **1976**, *19*, 1299.
71. Wilkerson, W. W.; Galbraith, W.; Gaus-Brangs, K.; Grubb, M.; Hewes, W. E.; Jaffee, B.; Kenney, J. P.; Kerr, J.; Wong, N. *J. Med. Chem.* **1994**, *37*, 988.
72. Hansch, C.; Kiehs, K.; Lawrence, G. L. *J. Am. Chem. Soc.* **1965**, *87*, 5770.
73. Yasuda, Y.; Tochikubo, K.; Hachisuka, Y.; Tomida, H.; Ikeda, K. *J. Med. Chem.* **1982**, *25*, 315.
74. de Compadre, R. L. L.; Pearlstein, R. A.; Hopfinger, A. J.; Seydel, J. K. *J. Med. Chem.* **1987**, *30*, 900.
75. de Compadre, R. L. L.; Pearlstein, R. A.; Hopfinger, A. J.; Seydel, J. K. *J. Med. Chem.* **1988**, *31*, 2315.
76. Hansch, C. *Ann. Rep. Med. Chem.* **1967**, *35*, 348.
77. Purcell, W. P.; Clayton, J. M. *Ann. Rep. Med. Chem.* **1968**, *36*, 314.
78. Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525.
79. Kubinyi, H. *Prog. Drug Res.* **1979**, *23*, 97.
80. Ghose, A. K.; Crippen, G. M. *J. Chem. Inf. Comput. Sci.* **1987**, *27*, 21.
81. Taylor, P. J. *Hydrophobic Properties of Drugs*, In *Comprehensive Medicinal Chemistry. The Rational Design, Mechanistic Study, and Therapeutic Application of Chemical Compounds*; Hansch, C.; Sammes P. G.; Taylor, J. B., Eds., Pergamon: New York, 1990; Vol. 4, p 241.
82. Dearden, J. C. *Molecular Structure and Drug Transport*, In *Comprehensive Medicinal Chemistry. The Rational Design, Mechanistic Study, and Therapeutic Application of Chemical Compounds*, Hansch, C.; Sammes, P. G.; Taylor, J. B., Eds.; Pergamon: New York, 1990; Vol. 4, 375
83. Topsom, R. D.; *Progr. Phys. Org. Chem* **1976**, *12*, 1.
84. Reynolds, W. F.; *Progr. Phys. Org. Chem.* **1983**, *14*, 165.
85. Bowden, K. *Electronic Effects in Drugs*, In *Comprehensive Medicinal Chemistry. The Rational Design, Mechanistic Study, and Therapeutic Application of Chemical Compounds*; Hansch, C.; Sammes, Taylor, J. B., Eds.; Pergamon: New York, 1990; Vol. 4, p 205.
86. Unger, S. H.; Hansch, C.; *Progr. Phys. Org. Chem.* **1976**, *12*, 91.
87. Silipo, C.; Vittoria, A. *Three-Dimensional Structure of Drugs*, In *Comprehensive Medicinal Chemistry. The Rational Design, Mechanistic Study, and Therapeutic Application of Chemical Compounds*; Hansch, C.; Sammes, P. G.; Taylor, J. B., Eds.; Pergamon: New York, 1990; Vol. 4, p 153.
88. Verloop, A.; Hoogenstraaten, W.; Tipker, J. *Development and Application of a New Steric Substituent Parameter in Drug Design*, In *Drug Design*; Ariens, E. J., Ed.; Academic Press: New York, 1976; Vol. 7, p 165.
89. Rodenkirchen, R.; Bayer, R.; Steiner, R.; Bossert, F.; Meyer, H.; Möller, E.; *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1979**, *310*, 69.
90. Mahmoudian, M.; Richards, W. G.; *J. Pharm. Pharmacol.* **1986**, *38*, 272.
91. Coburn, R. A.; Wierzba, M.; Suto, M. J.; Solo, A. J.; Triggler, A. M.; Triggler, D. J.; *J. Med. Chem.* **1988**, *31*, 2103.
92. Gaudio, A. C.; Korolkovas, A.; Takahata, Y.; *J. Pharm. Sci.* **1994**, *83*, 1110.
93. Hansch, C.; Kutter, E.; Leo, A.; *J. Med. Chem.* **1969**, *12*, 746.
94. Abdul-Ahad, P. G.; Blair, T.; Webb, G. A.; *Int. J. Quantum Chem.* **1980**, *17*, 821.
95. Hopfinger, A.; *J. Med. Chem.* **1981**, *24*, 818.
96. Orita, Y.; Ando, A.; Yamabe, S.; Nakanishi, T.; Arakawa, Y.; Abe, H.; *Arzneim.-Forsch./Drug Res.* **1983**, *33*, 688.
97. Mannhold, R.; Bayer, R.; Ronsdorf, M.; Martens, L.; *Arzneim.-Forsch./Drug Res.* **1987**, *37*, 419.
98. Costa, M. C. A.; Barata, L. E. S.; Takahata, Y.; *J. Mol. Struct. (Theochem)*, no prelo.
99. Chan, C. L.; Lien, E. J.; Tokes, Z. A.; *J. Med. Chem.* **1987**, *30*, 509.
100. Clare, B. W.; *J. Med. Chem.* **1990**, *33*, 687.
101. Bass, G. E.; Hudson, D. R.; Parker, J. E.; Purcell, W. P.; *J. Med. Chem.* **1971**, *14*, 275.
102. Seidel, W.; Meyer, H.; Born, L.; Kazda, S.; Dompert, W. *QSAR Strategies Des. Bioact. Compd., Proc. Eur. Symp. Quant. Struct.-Act. Relat.*, 5<sup>th</sup> ed.; Seydel, J. K., Ed.; VCH: Weinheim, 1985; p 366.
103. Mehler, E. L.; Gerhards, J.; *Mol. Pharmacol.* **1987**, *31*, 284.
104. De Benedetti, P. G.; Iarossi, D.; Menziani, C.; Caiolfa, V.; Frassinetti, C.; *J. Med. Chem.* **1987**, *30*, 459.
105. Mahmoudian, M.; Richards, W. G.; *J. Sci. I.R. Iran* **1990**, *1*, 261.
106. Ramos, M. N.; Neto, B. B.; *J. Comput. Chem.* **1990**, *11*, 569.
107. Moriguchi, I.; Komatsu, K.; Matsuchita, Y.; *J. Med. Chem.* **1980**, *23*, 20.
108. Kutter, E.; Hansch, C.; *J. Med. Chem.* **1969**, *12*, 647.

109. Cammarata, A.; *J. Med. Chem.* **1972**, *15*, 573.
110. Kubinyi, H.; *J. Med. Chem.* **1977**, *20*, 625.
111. Jurs, P. C.; Stouch, T. R.; Czerwinski, M.; Narvaez, J. N.; *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 296.
112. Craig, P. N.; *J. Med. Chem.* **1971**, *14*, 680.
113. Iwasa, J.; Fujita, T.; Hansch, C.; *J. Med. Chem.* **1965**, *8*, 150.
114. Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J.; *J. Med. Chem.* **1973**, *16*, 1207.
115. Norrington, F. E.; Hyde, R. M.; Williams, S. G.; Wootton, R.; *J. Med. Chem.* **1975**, *18*, 604.
116. Hansch, C.; Rockwell, S. D.; Jow, P. Y. C.; Leo, A.; Steller, E. E.; *J. Med. Chem.* **1977**, *20*, 304.
117. Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley: New York, 1979.
118. Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR: Hydrophobic, Electronic, and Steric Constants*; ACS Professional Reference Book: New York, 1995; Vol. 2.
119. Topliss, J. G.; Costello, R. J.; *J. Med. Chem.* **1972**, *15*, 1066.
120. Hansch, C.; *Acc. Chem. Res.* **1969**, *2*, 232.
121. Neely, W. B.; *Chemtech* **1973**, *3*, 573.
122. Redl, G.; Cramer, D.; Berkoff, C. E.; *Chem. Soc. Rev.* **1974**, *3*, 273.
123. Martin, Y. C.; *J. Med. Chem.* **1981**, *24*, 229.
124. Dunn III, W. J.; *Chemom. Intell. Lab. Sys.* **1989**, *6*, 181.
125. Kubinyi, H. *QSAR: Hansch Analysis and Related Approaches*; Mannhold, R.; Krosgaard-Larsen, P.; Timmerman, H., Eds.; VCH: Weinheim, 1993.
126. Hansch, C.; Leo, A.; Hoekman, D. *Exploring QSAR: Fundamentals and Application in Chemistry and Biology*; ACS Professional Reference Book: New York, 1995; Vol. 1.
127. Kolar, G. F.; Preussmann, Z. *Naturforsch. B* **1971**, *26*, 950.
128. Hansch, C.; Hatheway, G. J.; Quinn, F. R.; Greenberg, N.; *J. Med. Chem.* **1978**, *21*, 574.
129. Kutter, E.; Hansch, C.; *Arch. Biochem. Biophys.* **1969**, *135*, 126.
130. Li, R.-L.; Hansch, C.; Kaufman, B. T.; *J. Med. Chem.* **1982**, *25*, 435.
131. Selassie, C. D.; Li, R.-L.; Poe, M.; Hansch, C.; *J. Med. Chem.* **1991**, *34*, 46.
132. Ames, M. M.; Selassie, C. D.; Woodson, L. C.; Van Loon, J. A.; Hansch, C.; Weinshilboum, R. M.; *J. Med. Chem.* **1986**, *29*, 354.
133. Panthanickal, H.; Hansch, C.; Leo, A.; *J. Med. Chem.* **1979**, *22*, 1267.
134. Singer, J. A.; Purcell, W. P.; *J. Med. Chem.* **1967**, *10*, 1000.
135. Hansch, C.; Yoshimoto, M.; *J. Med. Chem.* **1974**, *17*, 1160.
136. Martin, Y. C.; Lynn, K. R.; *J. Med. Chem.* **1971**, *14*, 1162.
137. Magerlein, B. J.; Birkenmeyer, R. D.; Kagan, F.; *J. Med. Chem.* **1967**, *10*, 355.
138. Hansch, C.; McClarin, J.; Klein, T.; Langridge, R.; *Mol. Pharmacol.* **1985**, *27*, 493.
139. Hansch, C.; Klein, T. E.; *Acc. Chem. Res.* **1986**, *19*, 392.
140. Hansch, C.; Grieco, C.; Silipo, C.; Vittoria, A.; *J. Med. Chem.* **1977**, *20*, 1420.
141. Grieco, C.; Hansch, C.; Silipo, C.; Smith, R. N.; Vittoria, A.; Yamada, K.; *Arch. Biochem. Biophys.* **1979**, *194*, 552.
142. Silipo, C.; Hansch, C.; Grieco, C.; Vittoria, A.; *Arch. Biochem. Biophys.* **1979**, *194*, 552.
143. Dittmar, W.; Druckrey, E.; Urbach, H.; *J. Med. Chem.* **1974**, *17*, 753.
144. Ford, R. E.; Knowles, P.; Lunt, E.; Marshall, S. M.; Penrose, A. J.; Ramsden, C. A.; Summers, A. J. H.; Walker, J. L.; Wright, D. E.; *J. Med. Chem.* **1986**, *29*, 538.
145. Hansch, C.; Hatheway, B. A.; Guo, Z.; Selassie, C. D.; Dietrich, S. W.; Blaney, J. M.; Langridge, R.; Volz, K.W.; Kaufman, B. T.; *J. Med. Chem.* **1984**, *27*, 129.
146. Blaney, J. M.; Hansch, C.; Silipo, C.; Vittoria, A.; *Chem. Rev.*, **1984**, *84*, 333.