

PRÁCTICA II.1

RELAÇÕES QUANTITATIVAS ESTRUTURA– ATIVIDADE: FUNDAMENTOS E APLICAÇÃO DA ANÁLISE DE HANSCH

QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIPS: FUNDAMENTALS AND APPLICATION OF THE HANSCH ANALYSIS

Leoberto Costa Tavares e Elizabeth Igne Ferreira

*Faculdade de Ciências Farmacêuticas; Universidade de São Paulo,
São Paulo, Brasil*

E-mail: hajudan@usp.br

A ação terapêutica de fármacos resulta de interações destes com sistemas biológicos e é dependente de fatores relacionados com sua estrutura química e, conseqüentemente, de suas propriedades físico-químicas. Estes fatores, sejam eles de caráter eletrônico, hidrofóbico ou estérico, influenciam na interação do fármaco com a biofase e na sua distribuição nos compartimentos que compõem o sistema biológico. Assim, dois fármacos com estrutura química semelhantes, diferenciando-se apenas por um átomo ou posição que este ocupa na molécula, podem apresentar diferenças quanto às suas propriedades físico-químicas e, conseqüentemente, quanto à atividade biológica, tanto do ponto de vista quantitativo como qualitativo.

As propriedades físico-químicas de compostos bioativos, reflexo de sua estrutura química, podem ser descritas quantitativamente se forem expressas por meio de parâmetros físico-químicos ou descritores estruturais adequados ^{20, 21, 23, 26}. Assim, com base neste conceito, foi possível desenvolver nova área do conhecimento, que se preocupa com o estudo das relações entre a estrutura química e a atividade biológica ^{20, 23, 26, 32, 51, 57}. Estas relações, atualmente, são expressas por modelos matemáticos, que correlacionam descritores estruturais de séries de compostos análogos com a atividade biológica por eles desempenhada ^{23, 25, 26, 28, 39, 56}.

Estudos de *QSAR (Quantitative Structure-Activity Relationships)* vêm sendo, progressivamente, desenvolvidos durante os últimos 40 anos. Hoje, são amplamente aplicados para descrever quantitativamente as relações entre a estrutura química de moléculas e a sua atividade biológica, visando à identificação de valores ótimos para determinadas propriedades físico-químicas e, por meio deles, fundamentar o planejamento de novas substâncias candidatas a fármaco, que possuam perfil mais adequado às necessidades atuais ^{06,13, 20, 21, 22, 26, 29, 51, 52, 54, 56, 58}.

1-PROPRIEDADES FÍSICO-QUÍMICAS QUE CONDICIONAM A AÇÃO DE FÁRMACOS

Entre as propriedades físico-químicas mais importantes para o estabelecimento da atividade biológica de fármacos citam-se a distribuição eletrônica, a hidrofobicidade e a estereoquímica da molécula. Cada uma destas propriedades contribui em maior ou menor intensidade para o estabelecimento da resposta biológica e podem ser representadas por descritores estruturais, que expressam de forma quali e quantitativa a sua influência na atividade biológica destes compostos ^{20, 21, 23, 26, 29, 31, 33}.

1.1- DESCRITORES ESTRUTURAIS DE CARÁTER ELETRÔNICO

Desde o estabelecimento da equação de HANSCH ¹⁹, em 1964, muitos trabalhos têm reconhecido que os efeitos eletrônicos transmitidos por grupos substituintes, não raro, influenciam a atividade biológica de fármacos ^{01,05, 10, 12, 17, 25, 35, 37, 41, 43, 46}. O entendimento da natureza destes efeitos bem como o seu

dimensionamento são de importância fundamental para o entendimento das relações quantitativas entre a estrutura química e a atividade de compostos bioativos e envolve, necessariamente, o conhecimento da equação de Hammett^{15, 20, 23, 26}.

1.1.1- EQUAÇÃO DE HAMMETT

Desde o início do século passado vários pesquisadores^{22, 25, 56} vêm tentando encontrar um modelo matemático que correlacione a estrutura química com a atividade biológica de fármacos, mas foi somente a partir de 1937, com base em conhecimentos na área de físico-química orgânica, que os avanços neste campo passaram a ser realmente significativos^{14, 15, 16, 17, 51, 56}.

Hammett em seu clássico trabalho *Physical Organic Chemistry*¹⁵, publicado em 1937, estudou a ionização de ácidos benzóicos *meta* ou *para*-substituídos em água a 25 °C, observando a ocorrência de relação linear entre os logaritmos da constante de ionização do ácido benzóico substituído, Log Ka, e o logaritmo da constante de equilíbrio para a hidrólise do éster etílico-funcionalizado correspondente, Log K. Em seguida, Hammett definiu a constante σ para representar o efeito eletrônico do substituinte donde Ka é a constante de dissociação do ácido benzóico. A equação 1 representa esta relação, que passou a ser conhecida como equação de Hammett^{15, 17, 20, 22, 23, 26}.

$$\sigma = \text{Log} (K/K_0)$$

ou

$$\sigma = \text{Log}K - \text{Log} K_0$$

ou ainda

$$\text{Log}K = \rho\sigma + \text{Log} K_0 \quad (\text{Eq.1})$$

Em que:

Kx é a constante de ionização do ácido benzóico *p* ou *m*-substituído;

K_o é a constante de ionização do ácido benzóico não substituído;

σ é a constante de grupo;

ρ é a constante de reação

Relações logarítmicas entre constantes de equilíbrio ou de velocidade de reações químicas são essencialmente relações lineares de energia livre, uma vez que variações entre a energia livre padrão de reações é diretamente proporcional à energia livre de ativação^{07, 08, 09, 58}.

A constante sigma, σ , definida pela equação de Hammett, passou a ser conhecida como constante de grupo, uma vez que mede a influência eletrônica de um dado substituinte, independentemente da reação ou da molécula a que ele está ligado^{15, 16, 21, 57, 58}. Vale salientar que a constante σ de Hammett sofre influência da posição que o substituinte ocupa na molécula e é de caráter aditivo^{20, 22, 23, 26, 32}.

O valor absoluto de σ reflete a grandeza dos efeitos indutivos e de ressonância exercidos pelo grupo substituinte no centro de reação ou na propriedade físico-química medida^{20, 22, 23, 26}. Valores positivos são observados em substituintes que atraem elétrons e valores negativos, em substituintes que repelem elétrons^{21, 24}.

O coeficiente angular da equação de Hammett corresponde à constante ρ , conhecida como constante de reação. Sua grandeza, em módulo, mede a suscetibilidade da reação ou da propriedade medida ao efeito polar exercido pelo substituinte e depende da natureza da reação que a definiu. Valores positivos de ρ são observados em reações favorecidas por grupos que atraem elétrons e valores negativos, em reações favorecidas por grupos que repelem elétrons^{14, 15, 20, 23, 26}.

1.1.1.1- Equação de Hammett expandida

A equação de Hammett foi revisada, discutida e criticada por diversos autores^{09, 11, 16, 48, 52} resultando no surgimento de séries de constantes que diferem dos valores de σ definidos originalmente por Hammett.

Taft e Lewis ⁴⁵ foram os primeiros autores a sugerir a possibilidade de se conhecer separadamente as contribuições dos efeitos indutivo e de ressonância exercidos pelos substituintes em posição *meta* ou *para* em anéis benzênicos e isto foi possível por meio da aplicação das equações 2 e 3.

$$\sigma_p = \sigma_I + \sigma_R \quad (\text{Eq. 2})$$

$$\sigma_m = \sigma_I + \alpha\sigma_R \quad (\text{Eq. 3})$$

Em que:

σ_p e σ_m são valores da constante σ de Hammett para substituintes em posição *para* e *meta* no anel benzênico, respectivamente;

σ_I e σ_R são contribuições dos efeitos indutivo e de ressonância respectivamente;

α é um fator de correção para o efeito de ressonância a partir da posição *meta*.

A grande vantagem do desmembramento do valor original de σ em seus componentes indutivo e de ressonância, representados por σ_I e σ_R , respectivamente, é a possibilidade de emprego da equação de Hammett expandida. Desta forma, é possível expressar as contribuições individuais dos efeitos indutivo e de ressonância para a velocidade ou posição de equilíbrio da reação ou para a propriedade medida ^{17, 18, 20, 23, 24, 32}.

A equação 4 representa a equação de Hammett expandida, considerando o desmembramento do efeito polar exercido pelo grupo substituinte em seus componentes indutivo e de ressonância, bem como suas respectivas influências sobre a velocidade ou a posição de equilíbrio da reação ou propriedade medida.

$$\text{Log } K = \rho_I\sigma_I + \rho_R\sigma_R + \text{Log } K_o \quad (\text{Eq. 4})$$

Em que:

Log K é a constante de equilíbrio do composto substituído;

Log K_o é a constante de equilíbrio do composto não substituído;

- ρ_I é o coeficiente angular relativo ao efeito indutivo;
- σ_I é a contribuição do efeito indutivo;
- ρ_R é o coeficiente angular relativo ao efeito de ressonância;
- σ_R é a contribuição do efeito de ressonância

Esta equação permite calcular isoladamente a susceptibilidade da reação ou propriedade físico-química medida aos efeitos indutivo e de ressonância de substituintes e tem sido aplicada, com êxito, a vários sistemas em que a equação de Hammett original não fornece boa correlação^{09, 18, 20, 32}.

1.1.1.2 Correções da constante sigma de Hammett

Swain e Lupton⁴⁴, em 1968, revendo a constante σ de Hammett, propuseram a decomposição do efeito eletrônico de grupos substituintes no efeito indutivo e no efeito de ressonância. Estes autores consideraram o efeito indutivo ou de campo, \mathfrak{F} , de "field", combinado linearmente ao efeito de ressonância, \mathfrak{R} , de "resonance", equivalentes ao valor da constante sigma de Hammett^{21, 26}. A proposição de Swain e Lupton é expressa pela equação 5.

$$\sigma = f\mathfrak{F} + r\mathfrak{R} \quad (\text{Eq. 5})$$

Em que:

- σ é a constante de grupo de Hammett;
- \mathfrak{F} representa os efeitos indutivo e de campo;
- \mathfrak{R} representa o efeito de ressonância;
- f e r são coeficientes empíricos e independentes.

O estudo de Swain e Lupton foi feito com base em duas premissas: a primeira considera que o efeito de *para*-substituintes na ionização do ácido biciclo-(2,2,2)-octano-1-carboxílico resulta inteiramente da ação de efeito indutivo do substituinte, uma vez que o efeito de ressonância não se transmite por sistemas saturados; a segunda premissa considera que o efeito eletrônico transmitido pelo grupo substituinte trimetilamônio, $-\text{N}(\text{CH}_3)_3$, não envolve efeito

de ressonância, podendo-se, portanto, considerar este efeito como sendo igual a zero. Desta forma, foi possível calcular o valor de \mathfrak{R} a partir da equação 6.

$$\sigma_p = 0,56\mathfrak{S} + \mathfrak{R} \quad (\text{Eq. 6})$$

Em que:

σ_p é a constante sigma de Hammett na posição *para*;

\mathfrak{S} são os efeitos indutivo e de campo;

\mathfrak{R} é o efeito de ressonância e;

0,56 é o valor de α para o sistema de reações utilizado por Swain e Lupton.

Hansch ^{18, 20} reexaminou os valores e a abordagem de Swain e Lupton e considerou que os mesmos não estavam corretamente dimensionados. Este autor propôs, então, a utilização do fator correspondente ao valor de α para a ionização de ácidos benzóicos em etanol/água 50% v/v a 25 °C, redefinindo, desta forma, a equação que calcula os valores de \mathfrak{S} e \mathfrak{R} . A proposição de Hansch pode ser resumida nas equações 7 e 8.

$$\mathfrak{S} = 1,396\sigma_m - 0,373\sigma_p - 0,009 \quad (\text{Eq. 7})$$

$$n = 14 \quad r = 0,992 \quad s = 0,0042$$

$$\mathfrak{R} = \sigma_p - 0,921 \mathfrak{S} \quad (\text{Eq. 8})$$

Apesar de criticados os descritores eletrônicos \mathfrak{S} e \mathfrak{R} propostos por Swain e Lupton têm sido empregados com sucesso em diversos trabalhos, principalmente aqueles envolvidos em estudos de QSAR ^{20, 22, 24, 32}. Apresentam-se na tabela 1 valores corrigidos da constante σ de Hammett e constantes derivadas bem como os valores de \mathfrak{S} e \mathfrak{R} de Swain e Lupton.

TABELA 1

Valores selecionados de constantes de grupo de substituintes
freqüentemente presentes em fármacos

X	σ_p	σ_m	σ_I	σ_{R^*}	\mathfrak{S}	\mathfrak{R}
NO ₂	0,78	0,71	0,64	0,14	0,65	0,13
F	0,06	0,34	0,52	-0,46	0,45	-0,39
Br	0,23	0,39	0,44	-0,21	0,45	-0,22
Cl	0,23	0,37	0,47	-0,23	0,42	-0,19
H	0,00	0,00	0,00	0,00	0,00	0,00
CH ₃	-0,17	-0,07	-0,04	-0,13	0,01	-0,18
C ₂ H ₅	-0,15	-0,07	-0,01	-0,14	0,00	-0,15
OH	-0,37	0,12	0,29	-0,66	0,33	-0,70
OCH ₃	-0,27	0,12	0,27	-0,54	0,29	-0,56
OC ₂ H ₅	-0,24	0,10	0,28	-0,52	0,26	-0,50
COCH ₃	0,50	0,38	0,30	0,15	0,33	0,17
CF ₃	0,54	0,43	0,40	0,14	0,38	0,16
CN	0,66	0,56	0,53	0,13	0,51	0,15
NH ₂	-0,66	-0,16	0,12	-0,78	0,08	-0,74
N(CH ₃) ₂	-0,83	-0,16	0,06	-0,89	0,15	-0,98
SO ₂ NH ₂	0,60	0,53	0,46	0,14	0,49	0,11

Em que:

X é o grupo substituinte;

σ_p , σ_m , σ_I e σ_{R^*} - são constantes de grupo derivadas do de σ

Hammett;

\mathfrak{S} e \mathfrak{R} são as constantes de grupo propostas por Swain e

Lupton;

* Valores referentes à posição *para* em relação ao centro de reação.

Fonte: HANSCH, C., LEO, A., HOEKMAN, D. *Exploring QSAR: hydrophobic, eletronic and*

steric constants. Washington: ACS Professional Reference

Book, 1995. v.2, 348p.

1.1.2 - OUTROS DESCRITORES ESTRUTURAIS DE CARÁTER ELETRÔNICO

Vários outros descritores estruturais de caráter eletrônico podem ser definidos por meio da observação de variações em propriedades físico-químicas de compostos químicos sensíveis ao efeito eletrônico de grupos substituintes. Parâmetros espectroscópicos como frequência de absorção na região do infravermelho, deslocamentos químicos de RMN-¹H e RMN-¹³C, além de constantes de ionização e variação no momento dipolar, são alguns dos parâmetros que podem refletir a influência eletrônica de grupos substituintes sobre as propriedades de um determinado composto químico. Parâmetros como frequência de absorção na região do infravermelho e deslocamentos químicos de RMN-¹H e RMN-¹³C têm sido empregados com sucesso como descritores estruturais de caráter eletrônico em estudos de QSAR^{20, 23, 26, 32, 46, 52}.

1.2 - HIDROFOBICIDADE E DESCRITORES ESTRUTURAIS RELACIONADOS

A resposta biológica produzida por compostos bioativos está diretamente relacionada com os mecanismos de absorção, distribuição e atividade intrínseca^{08, 23, 26, 58}. Os processos de absorção e distribuição destes compostos são regulados basicamente pela sua hidrofobicidade, uma vez que é necessário que se dissolvam e atravessem as membranas biológicas que formam os tecidos e sistemas multi-compartimentados dos seres vivos, até que alcancem seus respectivos sítios de ação^{25, 26, 47, 58, 59}.

Os clássicos trabalhos de Overton³⁶ e de Meyer³³, juntamente com as inúmeras correlações encontradas entre a atividade biológica e a hidrofobicidade de compostos bioativos, demonstraram claramente que esta propriedade desempenha papel fundamental na expressão da bioatividade destes compostos, embora esta seja apenas uma das várias propriedades físico-químicas que, atuando conjuntamente, determinam a atividade biológica destes compostos^{08, 20, 23, 26, 58, 59}.

É importante reconhecer que a hidrofobicidade de um fármaco determina a extensão e a velocidade de sua absorção e distribuição, a sua capacidade de ligação ao sítio receptor, bem como a sua biotransformação e excreção. O entendimento destas etapas envolve, por um lado, o conhecimento da estrutura, função e propriedades das membranas biológicas e, por outro lado, o conhecimento da estrutura e das propriedades físico-químicas do fármaco, como o tamanho e forma da molécula, seu grau de ionização e lipossolubilidade relativa de suas formas neutra e ionizada^{05, 08, 25, 47, 58, 59}.

Um dos modelos de membrana biológica adotado em estudos de QSAR é aquele conhecido por mosaico fluido. Este modelo representa a membrana biológica como sendo constituída por bicamadas contínuas de lipídeos polares com proteínas dispersas por toda a sua extensão. Estas membranas apresentam-se como estruturas fluidas, flexíveis, permeáveis à água e impermeáveis a íons e moléculas muito polares. Apresenta-se na figura 1 o modelo de mosaico fluido para membranas biológicas.

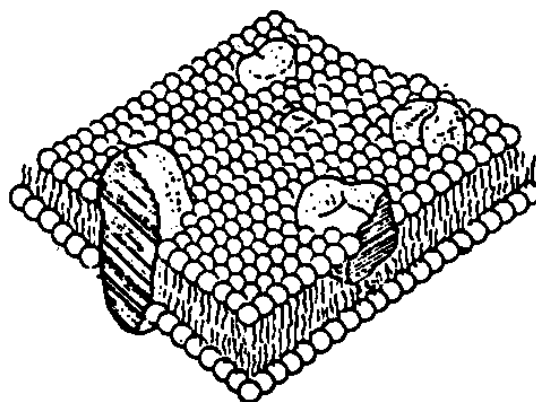


Figura 1: Modelo de mosaico fluido proposto para representar as membranas biológicas

Fonte: Taylor J. B. Kennevell, P. D. *Introductory medicinalchemistry*, New York, Wiley, 1981.

Muitos fármacos são ácidos ou bases fracas, que, quando em solução, apresentam equilíbrio entre suas formas neutra e ionizada. A forma neutra é

geralmente mais lipossolúvel e se difunde através das membranas biológicas por processo passivo, que depende do gradiente de concentração do sistema e do coeficiente de partição da molécula. A forma ionizada é absorvida por processo ativo e tem sua distribuição condicionada ao seu pKa e ao pH do meio^{08, 23, 26, 58, 59}.

1.2.1- COEFICIENTE DE PARTIÇÃO

O coeficiente de partição de uma determinada espécie química é definido como sendo a razão entre as concentrações que se estabelecem nas condições de equilíbrio de uma substância química, quando dissolvida em sistema constituído por uma fase orgânica e uma fase aquosa, e está associado à mudança de energia livre provocada pela substância sobre o equilíbrio termodinâmico do sistema^{08, 20, 25, 26, 31, 47}. Esta relação pode ser expressa pela equação 9.

$$P = [\text{orgânica}]/[\text{aquosa}] \quad (\text{Eq. 9})$$

Em que:

P é o coeficiente de partição do composto analisado;

[orgânica] é a concentração do composto na fase orgânica nas condições de equilíbrio;

[aquosa] é a concentração da substância a fase aquosa nas condições de equilíbrio.

Considerando-se que o coeficiente de partição é um dos parâmetros físico-químicos mais amplamente utilizados em estudos de QSAR^{01, 05, 23, 25, 32, 58, 59} e que sua determinação é freqüentemente necessária, especialistas optaram por definir um sistema de solventes preferencial para sua determinação, embora vários sistemas de solventes possam ser empregados^{02, 42}. Assim, o sistema *n*-octanol/tampão fosfato pH 7,4 é considerado o sistema preferencial para a determinação do coeficiente de partição visando à aplicação em estudo de QSAR. As principais vantagens do uso do *n*-octanol como fase orgânica na determinação do coeficiente de partição são as seguintes^{23, 25}.

- possui ampla capacidade de dissolução frente a diferentes compostos químicos;
- seu grupo hidroxila pode agir tanto como doador como receptor de elétrons na formação de ligações de hidrogênio;
- embora imiscível em água o *n*-octanol tem capacidade de dissolver até 2,3 *M* de água sob condições de equilíbrio;
- as ligações de hidrogênio de molécula solvatada não necessitam ser quebradas durante sua transferência da fase orgânica para a fase aquosa. Deste modo, os coeficientes de partição determinados em *n*-octanol/tampão fosfato refletem apenas as interações hidrofóbicas;
- não é volátil em temperatura ambiente;
- **é adequado para medição direta na região do UV, uma vez que sua absorção ocorre em comprimento de onda muito abaixo da faixa em que absorve a maioria dos fármacos,**
- é quimicamente estável e disponível comercialmente.

Entre os diversos métodos disponíveis para a determinação de coeficientes de partição citam-se o método de *shake-flask*^{23, 25, 26, 31} que, em princípio, é bastante simples e se fundamenta na dissolução de um composto químico em sistema bifásico formado por um solvente polar e um solvente apolar.

Cromatografia em camada delgada, CCD, e cromatografia líquida de alta eficiência, *HPLC*, têm sido bastante utilizadas na determinação dos parâmetros R_f e K' , respectivamente. Estas técnicas têm sido aplicadas, com sucesso, na obtenção de coeficientes de partição de compostos com solubilidade consideravelmente maior em uma das fases. O valor R_f se relaciona com R_m pela equação 10 e R_m se relaciona com o coeficiente de partição por meio da equação 11^{23, 25, 26, 30}.

$$R_m = \text{Log} [1/R_f - 1] \quad (\text{Eq. 10})$$

$$\text{Log } P = a \cdot R_m + b \quad (\text{Eq. 11})$$

Em que:

R_m é a distância percorrida em relação a um composto considerado como padrão;

Rf é a razão entre a distância percorrida pelo composto e a distância percorrida pelo eluente;
 Log P é o coeficiente de partição do composto analisado;
 b é o coeficiente linear da equação.

O tempo de retenção em cromatografia líquida de alta eficiência, *HPLC*, é quantitativamente descrito pelo fator de capacidade, K' , dado pelo tempo de retenção normalizado, ou seja:

$$K' = (t_r - t_o)/t_o \quad (\text{Eq. 12})$$

Em que:

K' é o fator de capacidade;
 t_r é o tempo de retenção ;
 t_o é o tempo de eluição do solvente.

Relação entre Log K' e Log $P_{n\text{-oct/tampão}}$ é possível ser observada aplicando-se a equação de Collander⁰². Contudo, a hidrofobicidade expressa em termos de K' é de natureza relativa, uma vez que é necessária a utilização de compostos de referência^{25, 30, 31, 34}.

Determinação do coeficiente de partição utilizando ensaios cromatográficos, como CCD ou *HPLC*, permitem certo grau de impureza no soluto, enquanto que o método de *shake-flask* exige substâncias puras para determinação confiável do respectivo coeficiente de partição^{25, 26, 30, 31}.

1.2.2 - O PARÂMETRO π DE HANSCH

Hansch^{18, 23} definiu, analogamente ao parâmetro σ de Hammett, o parâmetro π , que representa a contribuição hidrofóbica de um determinado substituinte, como sendo a relação logarítmica entre o coeficiente de partição de um composto substituído e o coeficiente de partição de seu análogo não substituído. Esta relação pode ser representada pela equação 13.

$$\pi_x = \text{Log } P_x - \text{Log } P_H \quad (\text{Eq. 13})$$

Em que:

π_x é o parâmetro que reflete a contribuição hidrofóbica do grupo substituinte, X;

P_x é o coeficiente de partição do composto substituído;

P_H é o coeficiente de partição do composto não substituído.

O parâmetro π de Hansch tem sido amplamente utilizado em estudos de QSAR como medida da contribuição hidrofóbica de grupos substituintes, uma vez que mede a contribuição individual destes grupos para o coeficiente de partição da molécula como um todo. A posição do grupo substituinte influi significativamente no valor da constante de hidrofobicidade π , sendo possível observar diferentes valores para um mesmo grupo substituinte em função de sua posição na molécula^{08, 20, 23, 26, 32, 58, 59}. Na tabela 2 apresentam-se valores de π de Hansch para grupos substituintes em função de sua posição em anéis benzênicos substituídos.

TABELA 2

Valores selecionados do parâmetro de π Hansch relativo a grupos substituintes em sistemas benzênicos dissustituídos

X	π	X	π
NO ₂	-0,28	OCH ₃	-0,02
F	0,14	OC ₂ H ₅	0,38
Cl	0,71	COCH ₃	-0,55
Br	0,86	CF ₃	0,88
H	0,00	CN	-0,57
CH ₃	0,56	NH ₂	-1,23
C ₂ H ₅	1,02	N(CH ₃) ₂	0,18
OH	-0,67	SO ₂ NH ₂	-1,82

Em que:

X é o grupo substituinte e;

π são os valores da constante de hidrofobicidade π de Hansch

Fonte: HANSCH, C., LEO, A., HOEKMAN, D. *Exploring QSAR: hydrophobic, eletronic and*

steric constants. Washington: ACS Professional

Reference Book, 1995. v.2, 348p.

1.2.3 - A CONSTANTE f DE REKKER

A análise de QSAR se apóia na validade do conceito de aditividade^{17, 20, 23, 25, 26, 29, 42}, que assume que todas as sub-estruturas de um composto bioativo contribuem, de modo aditivo e sob influência de sua vizinhança, para a hidrofobicidade do composto.

Rekker^{34, 39} considerando este conceito, observou que a aplicação do parâmetro π de Hansch em sistemas contendo cadeias alifáticas resulta em falhas no princípio de aditividade^{25, 26}. Por esta razão, este autor revendo a contribuição hidrofóbica de grupos substituintes individuais, definiu o parâmetro de hidrofobicidade f por meio da equação 14.

$$\text{LogP} = \sum a.f \quad (\text{Eq. 14})$$

Em que:

LogP é o coeficiente de partição do composto;

f é a constante fragmentar e expressa a contribuição individual de cada sub-estrutura do composto;

a é a frequência com que um dado fragmento ocorre na estrutura analisada

A diferença fundamental entre os significados de π e de f é que π se refere a contribuições relativas à substituição do átomo de hidrogênio pelo substituinte X, enquanto que f se refere a contribuições absolutas de substituintes e de sub-estruturas para a lipofilicidade da molécula como um todo. Desta forma,

para substituintes em sistemas aromáticos os valores de π e de f são igualmente empregados, enquanto que para sistemas alifáticos apenas o emprego dos valores de f são adequados.

Assumindo-se que o conceito de aditividade pode ser aplicado e considerando-se as contribuições individuais de cada grupo substituinte ou sub-estrutura para o coeficiente de partição de moléculas como um todo, foi possível o desenvolvimento de métodos matemáticos que permitem a previsão do valor do coeficiente de partição. Estes métodos apresentam excelente concordância com valores experimentais^{01, 05, 23, 26, 30, 40, 47, 53}. Discordâncias ocorrem, entretanto, em compostos que apresentam anéis aromáticos justapostos ou heterocíclicos e em compostos iônicos. Entre os diversos aplicativos disponíveis para esta finalidade destaca-se o *CLogP Program*, comercializado pela *Biobyte Co.*

1.3 - ESTEREOQUÍMICA E DESCRITORES ESTRUTURAIS RELACIONADOS

O estudo do aspecto tridimensional de compostos bioativos, bem como de seus receptores no sistema biológico é muitas vezes decisivo para a compreensão dos mecanismos envolvidos no estabelecimento da atividade biológica desempenhada por estes compostos^{20, 23, 28, 29}.

O conceito atual da teoria de ação de fármacos se baseia no princípio de ajuste entre a estrutura molecular do fármaco e seu receptor e considera, entre outras propriedades, a flexibilidade presente nas biomacromoléculas envolvidas. Assim, a estereoquímica de compostos bioativos governa os princípios de reconhecimento e discriminação molecular, sendo a estereoseletividade um aspecto importante nesta abordagem^{20, 23, 32, 58, 59}.

A Análise de Hansch, também conhecida como *QSAR-2D*, apesar de ser metodologia bastante eficaz na investigação das relações quantitativas entre a estrutura química e a atividade de compostos bioativos, apresenta especial limitação em relação à abordagem do aspecto tridimensional das relações fármaco-receptor. Isto ocorre, por um lado, em função do próprio modelo matemático utilizado, a Equação de Hansch, que sendo de caráter bidimensional não se mostra adequado para a exploração de aspectos

estereoquímicos da interação fármaco-receptor e, por outro lado, os descritores estruturais relacionados com o aspecto tridimensional não expressam adequadamente esta propriedade em duas dimensões. Citam-se, a seguir, alguns dos descritores estruturais propostos para expressar o efeito estérico de grupos substituintes.

A primeira tentativa para a obtenção de um parâmetro estérico que pudesse ser utilizado em estudos de Análise de Hansch foi feita por Taft ⁴⁵. Este autor considerou a hidrólise ácida de acetatos *alfa*-substituídos como modelo de reação para a definição do parâmetro *Es*. Este parâmetro foi, então, definido como sendo a relação logarítmica entre as constantes de hidrólise dos correspondentes análogos substituído e não substituído. A utilização deste parâmetro tem sido fortemente criticada, uma vez que outros fatores também influenciam as reações de hidrólise tomadas como padrão ^{17, 18, 20, 23, 26, 29, 32}.

Outro parâmetro estérico utilizado é o volume de van der Waals, *V_w*, que está relacionado com o relevo de uma determinada estrutura ou sub-estrutura molecular. Este parâmetro pode ser definido como sendo o volume impenetrável por outra molécula, obedecendo a lei da ação das massas. É importante reconhecer que as coordenadas tridimensionais que definem determinada região da molécula podem ser alteradas em função do estado físico da mesma, resultando, portanto, em restrição no emprego deste parâmetro em estudos de *QSAR* ^{20, 23, 26}.

O peso molecular, *PM*, embora seja indiretamente relacionado com o volume da molécula, não exprime convenientemente o seu perfil tridimensional, sendo, apesar disto, também empregado como parâmetro estérico em estudos de *QSAR* ^{20, 23, 26}.

Verloop e colaboradores ⁵⁵, em 1976, desenvolveram um conjunto de parâmetros que se mostram mais adequados na expressão do aspecto tridimensional de grupos substituintes ou de sub-estruturas. Estes parâmetros, denominados esterimol, foram obtidos pela medida do tamanho do substituinte considerando cinco diferentes direções. Estas distâncias foram calculadas no aplicativo *Sterimol Program*, que simula o modelo tridimensional de estruturas ou sub-estruturas moleculares. O padrão direcional proposto por Verloop na

obtenção dos parâmetros esterimol está representado esquematicamente na figura 2.

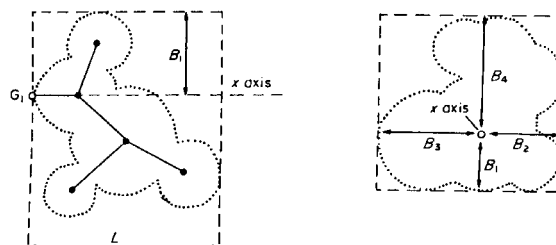


Figura 2: Padrão direcional na determinação dos parâmetros esterimol

Fonte: Verloop, A., Hoogenstraaten, W., Tipker, J.
Drug Design. v. VII, In:Ariens. E.J. Academic
Press, New York, 1976. p.165-207

1.4- DESCRITORES ESTRUTURAIS DE CARÁTER MISTO

Vários outros parâmetros estruturais são descritos na literatura ^{21, 24, 26}. Entre eles podemos citar a refratividade molar, MR, e a conectividade molecular, χ . Estes parâmetros até recentemente eram pouco empregados, mas, atualmente, são encontrados com certa frequência em estudo de QSAR.

1.4.1- REFRATIVIDADE MOLAR

O parâmetro de refratividade molar, MR, expressa uma propriedade físico-química de caráter constitutiva e aditiva, sendo, portanto, extremamente dependente da estrutura do composto químico ^{20, 21, 23, 26}. Com base nesta propriedade, os valores de MR relativos a sub-estruturas moleculares podem ser determinados experimentalmente pelo emprego da relação de Lorentz-Lorenz ^{23, 26}, descrita na equação 15.

$$M_R = \frac{n^2 - 1}{n^2 + 2} \times \frac{M_w}{d} \quad (\text{Eq. 15})$$

Em que:

n é o índice de refração do composto a 20 °C;

d é a sua densidade a 20 °C;

PM é o peso molecular do composto.

Dunn ²³, em 1977, sugeriu a existência de uma relação entre o valor de MR e a conformação tridimensional da estrutura molecular de fármacos. Assim, a refratividade molar estaria diretamente relacionada ao ajuste da molécula do fármaco a seu sítio receptor.

Por outro lado, o valor de MR de um determinado grupo substituinte pode ser relacionado, também, com a sua lipofilicidade expressa pela constante π de Hansch ^{20, 23, 26}. Esta relação resulta em correspondência altamente significativa de valores, uma vez que ambos dependem, em certa extensão, do volume molar, PM/d, do substituinte. Assim sendo, os valores de MR também podem expressar o caráter lipofílico de grupos substituintes e deles depende diretamente, sendo possível haver colinearidade ou ortogonalidade entre eles. Encontram-se descritos na tabela 3 valores de refratividade molar de grupos substituintes comumente presentes em séries de compostos utilizadas em estudos de QSAR.

TABELA 3

Valores selecionados de refratividade molar de grupos substituintes freqüentemente presentes em fármacos

X	MR	X	MR
NO ₂	0,74	OCH ₃	0,79
F	0,09	OC ₂ H ₅	1,25
Br	0,89	COCH ₃	1,12
Cl	0,60	CF ₃	0,50
H	0,10	CN	0,63
CH ₃	0,56	NH ₂	0,54
C ₂ H ₅	1,03	N(CH ₃) ₂	1,55
OH	0,28	SO ₂ NH ₂	1,23

Em que:

X é o grupo substituinte e

MR é a refratividade molar.

Fonte: HANSCH, C., LEO, A.,
HOEKMAN, D. **Exploring QSAR:**
hydrophobic, eletronic and steric
constants. Washington: ACS
Professional Reference Book,
1995. v.2, 348p.

1.4.2 - CONECTIVIDADE MOLECULAR

A conectividade molecular, χ , é um parâmetro que descreve a estrutura molecular em termos topológicos e se baseia na relação entre o grau de ramificação de compostos químicos e seu coeficiente de partição^{21, 23, 26, 59}.

A determinação do índice de conectividade molecular de compostos químicos é feita por meio da dissecção da estrutura molecular em sub-estruturas, tendo cada uma delas sua contribuição determinada por métodos matemáticos auxiliados por computador. Por outro lado, a conectividade molecular, por ser um parâmetro artificial e empírico, apresenta grande probabilidade de estabelecer correlações com parâmetros relacionados com a hidrofobicidade bem como com a refratividade molar^{23, 26}.

As boas correlações existentes entre a conectividade molecular e o coeficiente de partição e a conectividade molecular e a refratividade molar demonstram que este parâmetro pode representar adequadamente o caráter de hidrofobicidade de grupos substituintes, bem como as interações polares intramoleculares. A conectividade molecular tem sido utilizada com relativa freqüência em estudos de *QSAR*, a despeito de haver controvérsias a respeito da interpretação das correlações com ela obtidas^{20, 23, 26, 32, 59}.

1.5 - PARÂMETROS QUE EXPRESSAM A ATIVIDADE BIOLÓGICA

A qualidade das relações lineares de energia livre está diretamente relacionada com a natureza e precisão dos descritores estruturais considerados. Em linhas gerais, pode-se dizer que em estudos de *QSAR* o significado estatístico das correlações obtidas será tanto maior quanto mais precisas e exatas forem as medidas de atividade biológica, uma vez que os descritores estruturais, em geral, podem ser medidos com alto grau de precisão.

Os parâmetros biológicos expressam, geralmente, a atividade de compostos biologicamente ativos em termos de potência^{08, 14, 20, 23, 26, 58, 59}. Esta, por sua

vez, pode ser definida como sendo o inverso da dose ou concentração necessária para a obtenção de uma determinada resposta biológica. Desta forma, quanto menor for a dose necessária para a obtenção de determinada resposta biológica, maior será a potência do composto analisado.

A estimativa da potência de compostos bioativos é obtida por meio de ensaios biológicos, cuja aplicação se estende desde a seleção de novos compostos químicos de interesse terapêutico até a avaliação de sua eficácia, incluindo a análise de toxicidade. Ensaios biológicos podem ser classificados em ensaios qualitativos, ensaios quantais e ensaios quantitativos^{08, 32, 57}

Os ensaios qualitativos visam à seleção e identificação de substâncias químicas dotadas de alguma atividade biológica de interesse em misturas ou em extratos de tecidos vegetais ou animais. A importância deste tipo de ensaio biológico está vinculada à pesquisa de novos compostos químicos que possam despertar algum interesse farmacológico, bem como à identificação de contaminantes bioativos em formulações farmacêuticas.

Os ensaios quantais baseiam-se no aparecimento ou não de uma determinada resposta em sistemas biológicos tratados com um composto bioativo, sem, no entanto, quantificar o efeito observado. Por outro lado, é possível determinar a dose mínima do composto analisado que desencadeia este efeito. Exemplos típicos da aplicação desta espécie de ensaio são as determinações de dose eficaz 50%, ED₅₀, dose letal 50%, DL₅₀, e concentração mínima inibitória, CIM. O emprego de ensaios quantitativos visa à medida do efeito ou resposta desencadeada por um determinado composto bioativo frente a uma dose previamente estabelecida do mesmo, sendo a resposta biológica proporcional à dose utilizada^{08, 23, 26, 33, 58}. Neste tipo de ensaio observam-se nas curvas dose-resposta que apenas a faixa situada entre 20 e 80% do efeito máximo é função linear do logaritmo da dose, uma vez que o perfil desta curva apresenta aspecto sigmóide. Apesar desta limitação, curvas dose-resposta são aceitas para finalidades práticas dentro dos limites desta faixa.

1.5.1 - CARACTERÍSTICAS DE UM ENSAIO BIOLÓGICO IDEAL

O planejamento de ensaios biológicos visando à determinação da potência de compostos bioativos se baseia, geralmente, na análise de curvas que

representam determinada resposta padrão em função de diferentes doses do composto^{08, 23, 26, 32}. Ensaio biológico fundamentado na observação de determinada resposta biológica frente a uma única dose podem resultar em erros, como por exemplo a observação de inatividade do composto analisado na dose analisada. Este problema ocorre quando a dose está abaixo do limiar necessário para o desencadeamento da resposta biológica. Assim, infere-se que diferentes compostos podem desencadear a resposta biológica máxima, mas podem diferir drasticamente em termos de potência^{08, 26, 32, 58}.

A potência de compostos bioativos também pode ser determinada por meio do estabelecimento de curvas tempo-resposta. Esta análise é particularmente adequada para a avaliação da atividade de compostos administrados por via oral, nos quais o tempo necessário para o desencadeamento da resposta biológica varia para cada análogo a um composto padrão, ou ainda para compostos com tempos de duração de ação diferenciados. Neste tipo de ensaio, cada análogo é administrado em várias doses e a resposta biológica é observada em diferentes intervalos de tempo. Constrói-se, então, com estes dados e para cada dose administrada, curva da resposta em função do tempo e se calcula a área sob a curva. A seguir, constrói-se um gráfico considerando os valores das áreas sob as curvas em função do logaritmo das respectivas doses. A potência relativa dos vários análogos passa a ser, portanto, o inverso da dose necessária para produzir uma determinada resposta biológica^{08, 26, 58}.

1.5.2 - PARÂMETROS BIOLÓGICOS EMPREGADOS EM ESTUDOS DE QSAR

A determinação da atividade de compostos bioativos para emprego em estudos de QSAR exige a quantificação da dose ou efeito desencadeado pelo composto testado. Com relação à dose, encontram-se na literatura, com freqüência, estudos envolvendo a potência de compostos bioativos, expressa por meio do logaritmo da concentração molar^{23, 26}. Entre os parâmetros biológicos mais freqüentemente utilizados em estudos de QSAR encontram-se a dose eficaz 50%, DE₅₀, a dose letal 50%, DL₅₀, e a concentração inibitória mínima, CIM. Em estudos de QSAR também é freqüente avaliar a potência de compostos bioativos mediante ensaios quantitativos, ou seja, por meio da determinação de curvas dose-resposta ou curvas tempo-resposta. Este

procedimento é particularmente freqüente quando se deseja obter informações a respeito do perfil farmacodinâmico dos compostos pesquisados^{08, 23, 26, 32, 58, 59}.

2. ANÁLISE DE HANSCH

A interação de um fármaco com seu receptor pode, em princípio, ser descrita em termos de energias envolvidas nesta interação. A atividade biológica estará, assim, associada às mudanças de energia livre que ocorrem nos processos de absorção, distribuição e biotransformação ou à própria interação fármaco-receptor. Assim, a resposta biológica pode ser expressa, portanto, como sendo uma função complexa envolvendo as características estruturais de um determinado fármaco e de sua concentração no sítio receptor^{08, 23, 26, 58, 59}.

O estabelecimento de relações quantitativas estrutura – atividade, QSAR, assume que a atividade biológica está relacionada à variação de energia livre, ΔG , envolvida na interação fármaco-receptor^{07, 08, 23, 26, 32, 58, 59}. Essa variação pode ser decomposta em termos independentes e aditivos relacionados com as propriedades físico-químicas de um determinado composto biologicamente ativo. Em uma série de compostos análogos interagindo com o sistema biológico, as variações de energia livre envolvidas podem ser expressas pela equação 16:

$$\Delta G = a(\Delta GE) + b(\Delta GH) + c(\Delta GS) + R \quad (\text{Eq. 16})$$

Em que:

ΔG é a variação de energia livre;

a, b, c correspondem à contribuição de cada termo e;

ΔGE , ΔGH e ΔGS são variações de energia livre associadas às propriedades físico-químicas que influenciam a atividade biológica

Segundo o modelo proposto por Hansch¹⁹, os termos da equação 16 podem ser expressos por parâmetros que medem a influência de cada uma das propriedades físico-químicas envolvidas na determinação da atividade biológica, como demonstrado na equação 17.

$$AB = a\sigma + b\pi + cEs + d \quad (\text{Eq. 17})$$

Em que:

AB é a atividade biológica;

σ é a constante de grupo substituinte;

π é a constante de hidrofobicidade;

Es é a constante estérica ;

a, b, c expressam a contribuição individual de cada propriedade físico-química para a atividade biológica;

d é uma constante.

A equação 17 pode ser simples ou multiparamétrica e cada um dos termos pode ser linear ou quadrático, devendo-se sempre evitar a utilização de parâmetros colineares^{23, 26, 38}.

A Análise de Hansch correlaciona a atividade biológica com propriedades físico-químicas expressas por descritores estruturais, por meio de regressão linear, não-linear ou linear múltipla^{20, 26}. Como praticamente todos os parâmetros estruturais utilizados em Análise de Hansch são valores relacionados com as variações lineares de energia livre, isto é, derivam de constantes de equilíbrio ou de velocidade, surgiram sinônimos para a Análise de Hansch como Método de Energia Livre Linear ou Método Extratermodinâmico e, mais recentemente, QSAR-2D, já que considera apenas duas dimensões, ou seja, atividade biológica em função de propriedades físico-químicas.

Hansch e Fujita¹⁹, em 1964, combinaram diferentes descritores estruturais em uma equação linear, equação 18, em que C é a concentração molar que produz determinado efeito biológico, Log 1/C é a potência biológica e a, ρ e b são coeficientes determinados pela análise de regressão para o conjunto de dados.

$$\text{Log } 1/C = a \text{ Log } P + \rho\sigma + b \quad (\text{Eq. 18})$$

O modelo linear, equação 18, foi estendido para o modelo parabólico, equação 19, pela introdução de termo quadrático relacionado à hidrofobicidade²⁰. A idéia envolvida neste modelo matemático é que moléculas muito hidrofílicas não são capazes de atravessar barreiras muito hidrofóbicas ou vice-versa. O modelo parabólico expressa o aumento da hidrofobicidade em relação ao

aumento da atividade biológica até alcançar a hidrofobicidade ótima, a partir do qual a atividade tende a decrescer, de acordo com o perfil da curva parabólica.

$$\text{Log } 1/C = a (\text{Log}P)^2 + b \text{Log } P + \rho\sigma + c \quad (\text{Eq. 19})$$

O modelo bilinear desenvolvido por Kubinyi²⁷, em 1977, equação 20, tem mostrado, em alguns casos, melhor ajuste aos dados experimentais na expressão de relações quantitativas entre a estrutura química e a atividade biológica quando comparado ao modelo parabólico. Isto se deve ao termo bilinear, β , que assume caráter probabilístico e descreve o movimento do composto em sistemas multicompartimentados, como o sistema biológico, no qual a potência do composto analisado é fortemente dependente de seu coeficiente de partição, P .

$$\text{Log } 1/C = a \text{Log } P - b \text{Log } (\beta P + 1) + c \quad (\text{Eq. 20})$$

Com a introdução de modelos matemáticos multiparamétricos tornou-se possível descrever dependências complexas da atividade biológica em relação às propriedades físico-químicas de compostos bioativos. Nas últimas três décadas a atividade biológica de muitos desses compostos tem sido correlacionada, com sucesso, com diversas combinações de descritores estruturais empregando-se os modelos linear, parabólico e bilinear, estabelecendo quali e quantitativamente as relações entre a estrutura química e a atividade de compostos bioativos.

2.1- ESTRATÉGIA PARA ANÁLISES DE HANSCH

A metodologia mais adequada para a escolha de grupos substituintes visando o delineamento de série de compostos análogos para o estudo de QSAR por Análise de Hansch é o Diagrama de Craig⁰³. Este método está fundamentado na análise da intercorrelação de parâmetros físico-químicos e se baseia na construção de diagramas, que relacionam os descritores estruturais analisados em uma distribuição cartesiana.

Craig, em 1971, com base nos resultados de estudos estatísticos e na necessidade de selecionar grupos de substituintes apropriados para estudos de QSAR, propôs um diagrama em duas dimensões onde as coordenadas são descritores estruturais independentes. O Diagrama de Craig é freqüentemente

utilizado e fundamenta muitas das mais avançadas técnicas de planejamento de séries para o estudo de QSAR^{20, 28, 32, 49, 50}. A figura 4 mostra o Diagrama de Craig para os descritores estruturais π e σ .

A escolha adequada de grupos substituintes na definição de séries para o estudo da atividade biológica em função de dois descritores estruturais^{23, 26} deve ser feita buscando a maior dispersão possível no diagrama, isto é, deve-se selecionar substituintes em quadrantes diferentes incluindo os grupos $-\text{CH}_3$, $-\text{NO}_2$, $-\text{Cl}$ ou $-\text{Br}$, $-\text{H}$, $-\text{NH}_2$ ou $-\text{N}(\text{CH}_3)_2$. Se, contudo, o objetivo é analisar a influência de uma única propriedade físico-química sobre o efeito biológico analisado, deve-se escolher substituintes situados nas proximidades de uma linha imaginária paralela ao eixo da propriedade a ser analisada. Na escolha dos substituintes pelo Diagrama de Craig, deve-se evitar a colinearidade entre parâmetros evitando, assim, a obtenção de correlações não significativas.

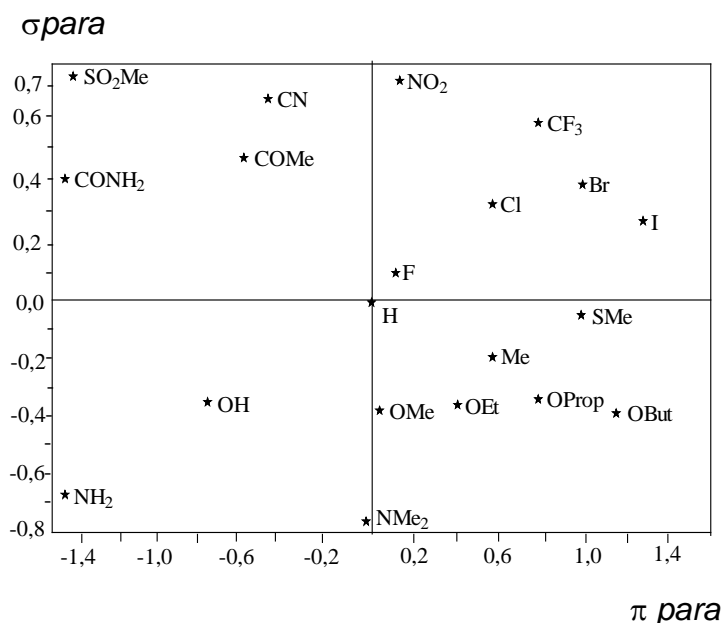


Figura 4: Diagrama de intercorrelação σ X π em para-substituição aromática

Fonte: CRAIG, P. N. *J. Med. Chem.*, 14(8): 680-4, 1971.

3- ESTRATÉGIAS EMPREGADAS NA OTIMIZAÇÃO DA ATIVIDADE BIOLÓGICA DE FÁRMACOS

O planejamento adequado de variações na estrutura de um composto bioativo pode resultar em derivados com maior interesse terapêutico, seja por apresentar maior atividade, menor toxicidade ou, ainda, por adquirir características farmacotecnicamente mais adequadas^{03, 04, 28}.

A definição de uma série adequada ao estudo de *QSAR* envolve a identificação de um composto que possua apenas algumas das propriedades biológicas esperadas para um melhor desempenho biológico. A otimização das propriedades biológicas deste composto, denominado composto líder ou padrão, é alcançada a partir de estudos sistemáticos. O estudo de *QSAR* em uma série de compostos análogos, considerando a abordagem de Hansch, envolve o planejamento estratégico representado na figura 3.

A análise das correlações entre descritores estruturais e a atividade biológica desempenhada em uma série de compostos análogos permite, dentro de certos limites^{23, 26, 32} a previsão da atividade biológica de análogos ainda não disponíveis, por meio da interpolação ou extrapolação de valores na correlação obtida. Previsões interpolativas são freqüentemente utilizadas na avaliação da eficiência do modelo proposto pela correlação, dentro dos limites de confiança estatísticos.

A Análise de Hansch em séries adequadamente planejadas e com valor preditivo confiável^{20, 26, 32} permite a previsão quantitativa da atividade biológica de análogos a um composto líder ou padrão e, conseqüentemente, o planejamento de novos análogos com grande economia de tempo e investimento.

A decisão de continuar ou interromper o programa de síntese em um estudo de *QSAR*, pode ser auxiliada pela análise da equação obtida a partir de dados parciais. Assim, a introdução de novos análogos na série pode ser interrompida quando, por exemplo, o composto de maior atividade já se encontra incluído entre os análogos testados.

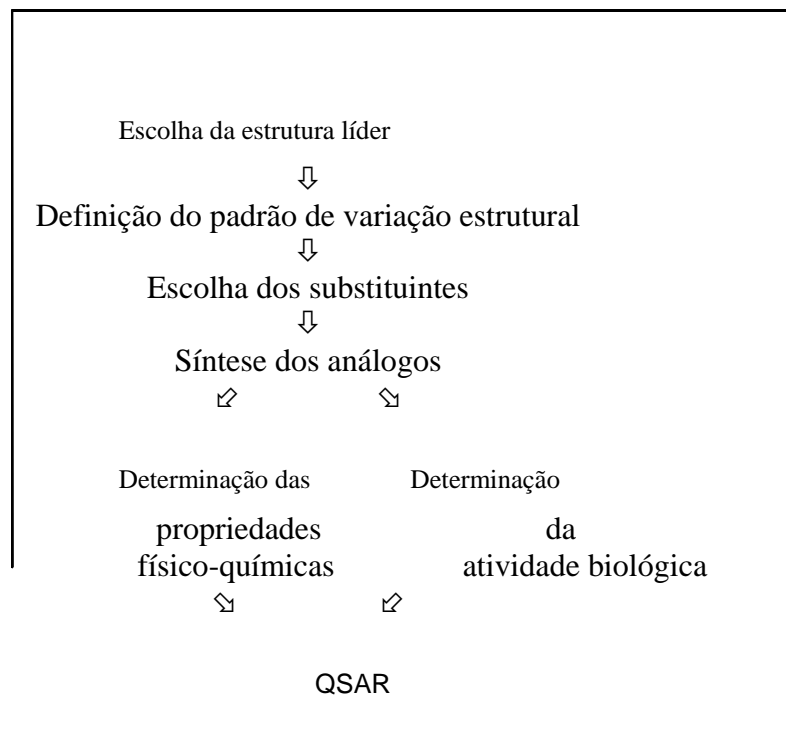


Figura 3: Estratégia para o estudo de QSAR pela aplicação de Análise de Hansch.

O valor preditivo das correlações obtidas em estudos de QSAR somente é válido para compostos com estrutura análoga aos componentes da série analisada e dentro da faixa de valores dos descritores estruturais utilizados na definição da correlação. Segundo Martin ³², a série ideal de análogos a um composto padrão não é aquela constituída, necessariamente, pelo menor número de compostos, mas aquela que possa assegurar que o composto mais ativo da série esteja incluído ou seja sugerido pela análise. Por outro lado, o não estabelecimento de uma relação de energia livre com os dados obtidos sugere que a série não foi bem delineada ou que outros parâmetros devem ser considerados na correlação.

O estudo das relações entre a estrutura química e a atividade biológica de fármacos tem contribuído significativamente com geração de novos conhecimentos acerca da contribuição quali e quantitativa de propriedades físico-químicas que influenciam a ação de fármacos. Por outro lado, também tem contribuído de forma importante com a otimização da performance de

fármacos já existentes, conhecimentos estes que estão sendo cada vez mais utilizados ou absorvidos pela indústria farmacêutica. Ressalta-se que a perspectiva do estudo e desenvolvimento na área de QSAR é bastante ampla, especialmente na área de planejamento de novos fármacos, em que há, ainda, muito a ser explorado.

4- APLICAÇÃO DE SOFTWARES NO ESTUDO DE QSAR POR ANÁLISE DE HANSCH

O estabelecimento de relações quantitativas entre a estrutura química e a atividade de compostos bioativos por Análise de Hansch, *QSAR-2D*, envolve a definição de equações matemáticas simples ou multiparamétricas envolvendo correlações entre a atividade biológica e descritores estruturais, com base nos modelos linear, parabólico e bilinear. A definição destas equações, a partir do conjunto de dados experimentais é, normalmente, obtida com o auxílio pacotes computacionais. As equações lineares ou parabólicas podem ser obtidas com a aplicação da maioria das planilhas de cálculo disponíveis no mercado, no entanto a obtenção de equações bilineares só pode ser feita com a utilização de *softwares* mais sofisticados ou desenvolvidos para esta finalidade.

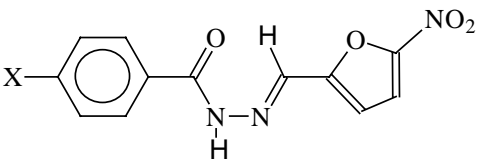
Após a formulação do modelo bilinear Kubinyi observou grande dificuldade entre os pesquisadores para a obtenção de correlações com base neste modelo, assim, visando solucionar o problema, este autor, em 1994, desenvolveu um *software* a que chamou de *Billin Program*, especialmente adaptado para o estudo de QSAR por Análise de Hansch, facilitando grandemente a obtenção de correlações envolvendo os modelos linear, parabólico e bilinear permitindo, inclusive, a obtenção e visualização amigável dos parâmetros estatísticos, depreendidos da correlação, considerados importantes para a avaliação da adequação do modelo aos dados experimentais. Por outro lado e considerando o grande número de correlações que devem ser investigadas em uma Análise de Hansch, o *Billin Program* permite a comparação sistemática entre correlações obtidas considerando os diferentes modelos matemáticos, resultando em racionalização na pesquisa da melhor correlação com grande economia de tempo.

Este aplicativo foi atualizado em 1998 e disponibilizado pelo autor para uso irrestrito pela comunidade científica brasileira juntamente com seu tutorial e

alguns exemplos de aplicação. Recomendamos, no entanto, o conjunto de dados apresentados na tabela 4, obtidos no Laboratório de Planejamento e Desenvolvimento de Fármacos do Departamento de Tecnologia Bioquímico-Farmacêutica da Faculdade de Ciências Farmacêuticas da Universidade de São Paulo, sob nossa coordenação, referente à potência antimicrobiana de derivados 5-nitro-2-tiofilidênicos frente a *Staphylococcus aureus*.

TABELA 4

Potência antimicrobiana de 5-nitro-2-tiofilidenos benzidrazidas 4-X-substituídas e valores descritores estruturais relacionados

								
X	Log1/ μM	π	CLogP	σ_p	σ_I	σ_{R^*}	\mathfrak{J}	\mathfrak{R}
NO ₂	-0,70	-0,28	2,05	0,78	0,64	0,14	0,65	0,13
Cl	-0,74	0,71	2,91	0,23	0,47	-0,24	0,42	-0,19
H	-1,22	0,00	2,10	0,00	0,00	0,00	0,00	0,00
CH ₃	-1,13	0,56	2,60	-0,17	-0,04	-0,13	0,01	-0,18
OCH ₃	-0,91	-0,02	2,41	-0,27	0,27	-0,54	0,29	-0,56
OC ₂ H ₅	-0,97	0,38	2,93	-0,24	0,28	-0,52	0,26	-0,50
CF ₃	-0,75	0,88	3,15	0,54	0,40	0,14	0,38	0,16
CN	-0,85	-0,57	1,76	0,66	0,53	0,13	0,51	0,15
NH ₂	-0,98	-1,23	1,55	-0,66	0,12	-0,78	0,08	-0,74
SO ₂ NH ₂	-1,50	-1,82	0,98	0,60	0,46	0,14	0,49	0,11

Em que:

X é o grupo substituinte;

π é a constante de hidrobicidade de Hansch;

CLogP é o coeficiente de partição calculado pelo CLogP Program, versão 4.0, Biobyte Co.

σ_p , σ_I , σ_{R^*} , \mathfrak{J} e \mathfrak{R} são as constantes de grupo relativas ao efeito eletrônico;

Fonte: HANSCH, C., LEO, A., HOEKMAN, D. *Exploring QSAR: hydrophobic, eletronic and steric constants*. Washington: ACS Professional Reference Book, 1995. v.2, 348p.

O estudo correlacional dos dados da tabela 4, utilizando o *Billin Program*, versão 1998, considerando os modelos linear, parabólico e bilinear, resultou em equações estatisticamente válidas e com capacidade preditiva confiável, demonstrando a importância e aplicabilidade da Análise de Hansch, *QSAR-2D*, na identificação da influência quali e quantitativa de propriedades físico-químicas na atividade biológica de substâncias candidatas a fármaco, com aproveitamento do conhecimento para o planejamento de análogos com melhor perfil farmacológico.

5- BIBLIOGRAFIA

1. P. Buchwald, N. Bodor. *Proteins: Struct., Funct., Genet.* **30**, 86 (1998).
2. R. Collander *Acta Chem. Scand.* **5**, 774 (1951).
3. P. N. Craig. *J. Med. Chem.* **14**, 680 (1971).
4. F. Darvas. *J. Med. Chem.* **17**, 799 (1974).
5. J. DeVillers, D. Domine, C. Guillon, W. Karcher. *J. Pharm. Sci.* **87**, 1086 (1998).
6. S. W. Dietrich. “Chemical information computing systems in drug discovery”, in *Burger’s Medicinal Chemistry and Drug Discovery: Principles and Practice*, Vol. 1, 5th ed., M. E. Wolff, pp. 415–496, John Wiley, New York (1995).
7. (a) J. Ferguson. *Proc. R. Soc., Ser. B* **127**, 387 (1939); (b) M. C. Richet. *Compt. Rend. Soc. Biol. (Paris)* **45**, 775 (1893); (c) Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 4, VCH, New York (1993).
8. W. O. Foye, T. L. Lemke, D. A. Williams. *Principles of Medicinal Chemistry*, Williams & Wilkins, 4th ed., Baltimore (1995).
9. T. Fujita. “The extrathermodynamic approach to drug design”, in *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study & Therapeutic Application of Chemical Compounds*, Vol. 4, C. Hansch, P. G. Sammes, J. B. Taylor (Eds.), Chap. 21.1, pp. 497–560, Pergamon Press, Oxford (1990).
10. T. Fujita, T. Ban. *J. Med. Chem.* **14**, 148 (1971).
11. T. Fujita, J. Iwasa. *J. Am. Chem. Soc.* **86**, 5175 (1964).
12. R. D. Gillion, W. P. Purcell, T. R. Bosin. *Eur. J. Med. Chem.* **12**, 187 (1977).
13. A. Gringauz. *Medicinal Chemistry: How Drugs Act and Why*, Wiley-VCH, New

- York (1997).
14. L. P. Hammett. *Physical Organic Chemistry*, McGraw-Hill, New York (1940).
 15. T. L. Hammett. *J. Am. Chem. Soc.* **59**, 963 (1937).
 16. C. K. Handoch, E. A. Meyer, B. J. Yager. *J. Am. Chem. Soc.* **83**, 4211 (1961).
 17. C. Hansch. *Acc. Chem. Res.* **2**, 232 (1969).
 18. (a) C. Hansch, A. Leo. *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley, New York (1976); (b) Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, VCH, New York (1993).
 19. C. Hansch, T. Fujita. *J. Am. Chem. Soc.* **86**, 1616 (1964).
 20. C. Hansch, A. Leo. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*, American Chemical Society, Washington, DC (1995).
 21. C. Hansch, A. Leo, D. Hoekman. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*, Vol. 2, ACS Professional Reference Book, American Chemical Society, Washington, DC (1995).
 22. C. Hansch, A. Leo, R. W. Taft. *Chem. Rev.* **91**, 165 (1991).
 23. C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, Pergamon Press, Oxford (1990).
 24. C. Hansch, A. Leo. Substituent constants for correlation analysis in chemistry and biology, John Wiley, New York (1976). Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 35, VCH, New York (1993).
 25. H. Kubinyi. *Prog. Drug Res.* **23**, 97 (1979).
 26. H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, Wiley-VCH, New York (1993).
 27. H. Kubinyi. *J. Med. Chem.* **20**, 625 (1977).
 28. H. Kubinyi. *Pharmazie* **50**, 647 (1995).
 29. H. Kubinyi. "The quantitative analysis of structure-activity relationship", in *Burger's Medicinal Chemistry and Drug Discovery: Principles and Practice*, 5th ed., Vol. 1, M. E. Wolff, pp. 497–572, John Wiley, New York (1995).
 30. W. J. Lambert. *J. Chromatogr., A* **656**, 469 (1993).
 31. A. J. Leo. "Methods of calculating partition coefficients", in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, chap. 18.7, Pergamon Press, Oxford (1990).

32. Y. C. Martin. *Quantitative Drug Design: A Critical Introduction*, Medicinal Research Series, Vol. 8, Marcel Dekker, New York (1978).
33. H. Meyer. *Arch. Exp. Path. Pharm.* **42**, 109 (1899); Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 4, VCH, New York (1993).
34. G. G. Nys, R. F. Rekker. *Chim. Ther.* **8**, 521 (1973).
35. G. Oros, T. Cserhati, A. Vrbanova. *Biochem. Mol. Biol. Int.* **47**, 79 (1999).
36. E. Overton. *Studien über die Narkose, zugleich ein Beitrag zur allgemeinen Pharmakologie*, G.Fischer, Jena, 1901; English translation by R. L. Lipnich (Ed.), *Studies on Narcosis, Charles Ernest Overton*, Chapman and Hall, London (1991). Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 4, VCH, New York (1993).
37. W. C. Patt, J. J. Edmons, J. T. Repine, K. A. Barryman, B. R. Reisdorf, C. Lee, M. S. Plummer, A. Shalvirpour, S. J. Haleel, J. A. Keiser, M. A. Flyman, K. M. Welch, E. E. Reynolds, R. Rubin, B. Tobias, H. Hallak, A. M. Dolvet. *J. Med. Chem.* **40**, 1063 (1997).
38. M. A. Pleiss, S. H. Unger. "The design of test series and the significance of QSAR relationships", in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 561–569, Pergamon Press, Oxford (1990).
39. R. F. Rekker, R. Mannhold. *Calculation of Drug Lipophilicity: The Hydrophobic Fragmental Constant Approach*, Wiley-VCH, Weinheim (1992).
40. M. C. Richet. *Compt. Rend. Soc. Biol. (Paris)* **45**, 775-776 (1893). Apud: Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, VCH, New York (1993).
41. J. D. Roberts, W. T. Moreland. *J. Am. Chem. Soc.* **75**, 2167 (1953).
42. P. Seiler. *Eur. J. Med. Chem.* **9**, 473 (1974).
43. B. A., Steinbaugh, H. W. Hamilton, V. Prasad, K. S. Para, P. J. Tummino, D. Ferguson, A. Lunney, C. J. Blankley. *Bioorg. Med. Chem. Lett.* **6**, 1099 (1996).
44. C. G. Swain, E. C. Lupton. *J. Am. Chem. Soc.* **90**, 4328 (1968).
45. R. W. Taft. *J. Am. Chem. Soc.* **74**, 3120 (1952).
46. L. C. Tavares. *Rev. Farm. Quim.* **30**, 53 (1997).
47. P. J. Taylor. "Hydrophobic properties of drugs", in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 241–294,

- Pergamon Press, Oxford (1990).
48. J. G. Topliss, Y. C. Martin. *Drug Design*, Vol. 5, pp. 1–21, Academic Press, New York (1975).
 49. J. G. Topliss. *J. Med. Chem.* **2**, 463 (1977).
 50. J. G. Topliss. *J. Med. Chem.* **15**, 1006 (1972).
 51. M. S. Tute. “History and objectives of quantitative drug design”, in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 1–31, Pergamon Press, Oxford (1990).
 52. M. S. Tute. *Adv. Drug Res.* **6**, 1 (1971).
 53. S. H. Unger. *J. Med. Chem.* **24**, 262 (1981).
 54. M. C. Venuti. “The role of recombinant DNA technology in medicinal chemistry and drug discovery”, in WOLFF, M.E. *Burger’s medicinal chemistry*. 5.ed. New York: John Wiley, 1995. v.1, cap.16, p.661-696.
 55. A. Verloop, W. Hoogenstraaten, J. Tipker. “Development and application of new steric substituent parameters in drug design”, in *Drug Design*, Vol. 7, E. J. Ariens, pp. 165–207, Academic Press, New York (1967).
 56. H. V. Waterbeend. *Quant. Struct.-Act. Relat.* **6**, 200 (1992).
 57. D. Weininger, J. L. Weininger. “Chemical structure and computers”, in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 59–82, Pergamon Press, Oxford (1990).
 58. D. G. Wermuth. *The Practice of Medicinal Chemistry*, Academic Press, San Diego (2000).
 59. M. E. Wolf. *Burger’s Medicinal Chemistry: Principles and Practice*, 5th ed., pt. 1, John Wiley, New York (1995).

Leoberto Costa Tavares e Elizabeth Igne Ferreira

hajudan@usp.br

PRACTICE II.1

QUANTITATIVE STRUCTURE–ACTIVITY RELATIONSHIPS: FUNDAMENTALS AND APPLICATION OF THE HANSCH ANALYSIS

Leoberto Costa Tavares and Elizabeth Igne Ferreira

*Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo,
Brasil*

E-mail: hajudan@usp.br

The therapeutic action of a drug is a result of interactions with biological systems and depends on the chemical structure of the drug, as well as on its physicochemical properties. Therefore, electronic, hydrophobic and steric properties influence the drug interaction with the biophase and distribution among biological compartments. In this respect, two drugs with similar chemical structure differing only by one atom or the position it occupies on the molecule may present different physicochemical properties and, consequently, qualitative and/or quantitative differences in biological activity.

The physicochemical properties of a bioactive compound may be quantitatively described by means of physicochemical parameters or adequate structural descriptors [20,21,23,26]. Based on this concept, a new area of knowledge, dealing with the study of chemical structures and their relationships with biological activity, was developed [20,23,26,32,51,57]. At present, these relationships are described by mathematical models correlating structural descriptors of series of analog compounds with the biological activity presented by the compounds [23,25,26,28,39,56].

Quantitative structure–activity relationships or QSAR studies have been progressively developed during the last 40 years. Today, these studies are widely applied to quantitatively describe relationships between the chemical structure of a drug molecule and its biological activity, aiming at defining optimal values for some physicochemical properties of the molecule and providing the fundamentals for planning new substances as drug candidates, according to current needs [6,13,20–22,26,29,51,52,54,56,58].

1. PHYSICOCHEMICAL PROPERTIES CONDITIONING DRUG ACTION

Electronic distribution, hydrophobicity, and stereochemistry are some of the most important physicochemical properties establishing the biological activity of drugs. Each property may contribute at some extension to the biological response, and can be represented by structural descriptors that exert qualitative and quantitative influence on the biological activity of the compounds [20,21,23,26,29,31,33].

1.1 STRUCTURAL DESCRIPTORS OF ELECTRONIC CHARACTER

Since the establishment of Hansch's equation [19] in 1964, numerous works recognizing the influence of electronic effects resulting from substituent groups on the biological activity of drugs have been published [1,5,10,12,17,25,35,37,41,43,46]. Comprehension of the nature of these effects, as well as dimensioning them, is of paramount importance to understand the quantitative relationships between chemical structure and activity of drugs and, necessarily, involves knowledge of Hammett's equation [15,20,23,26].

1.1.1 Hammett's equation

Since the beginning of the last century, researchers [22,25,56] have been searching for a mathematical model to correlate chemical structure and biological activity of drugs. However, significant development in this field has only gained force in 1937, based on organic physical-chemistry knowledge [14–17,51,56].

In the classic work *Physical Organic Chemistry* [15], published in 1937, Hammett studied the ionization of benzoic acid *meta*- or *para*-substituted, in aqueous solution at 25 °C, and observed a linear relationship between the logarithm of the ionization constant of substituted benzoic acid, $\log K_a$, and the logarithm of the rate constant for hydrolysis of the correspondingly substituted ethyl ester, $\log k$. Hammett then defined a constant σ to represent the electronic effect of the substituent, where K_o is the acid dissociation constant of benzoic acid, itself. Equation 1 represents this relationship, and is known as Hammett's equation [15,17,20,22,23,26].

$$\sigma = \log (K/K_0)$$

or

$$\sigma = \log K - \log K_0$$

or even

$$\log K = \rho\sigma + \log K_0 \tag{1}$$

where:

K_x is the ionization constant of *p*- or *m*-substituted benzoic acid;

K_0 is the ionization constant of nonsubstituted benzoic acid;

σ is the group constant; and

ρ is the reaction constant.

Considering that variations on the standard free energy of reactions are proportional to the free energy of activation, logarithm relationships between equilibrium constants or chemical reaction rates are, essentially, linear relationships of free energy [7–9,58].

The σ constant defined by Hammett's equation is also known as group constant. It measures the electronic influence of a given substituent, independently of the reaction or molecule at which the group is attached [15,16,21,57,58]. One should note that the σ of Hammett is influenced by the substituent position at the molecule, and represents an additive effect [20,22,23,26,32].

The absolute value of σ reflects the importance of inductive and resonance effects caused by the substituent group on the center of reaction or on the physicochemical property evaluated [20,22,23,26]. Positive values of σ are observed for substituent groups withdrawing electrons, whereas negative values are observed for electron-donating substituent groups [21,24].

The angular coefficient of Hammett's equation corresponds to the ρ constant, also known as reaction constant. The absolute value of the ρ constant measures the susceptibility of the reaction or property to the polar effect of the substituent, and depends on the nature of the reaction defining it. Positive values of ρ are observed for reactions favored by substituent groups withdrawing electrons, and negative values are observed for reactions favored by electron-donating substituent groups [14,15,20,23,26].

1.1.1.1 Expanded Hammett's equation

Hammett's equation was revised, discussed, and criticized by numerous authors [9,11,16,48,52], resulting in a series of constants differing from the original values of σ defined by Hammett.

The first authors to suggest the possibility of quantifying separately inductive and resonance effects contributions by *meta*- or *para*-substituents on benzene rings were Taft and Lewis [45]. This was possible by means of eqs. 2 and 3.

$$\sigma_p = \sigma_I + \sigma_R \quad (2)$$

$$\sigma_m = \sigma_I + \alpha\sigma_R \quad (3)$$

where:

σ_p and σ_m are values of the σ of Hammett for *meta*- or *para*-substituents in benzene ring, respectively; σ_I and σ_R are contributions of the inductive and resonance effects, respectively; and α is a correction factor for the resonance effect from the *meta*-position.

The greatest advantage of separating the original σ value on its inductive and resonance components, represented by σ_I and σ_R , respectively, is the possibility to employ the expanded Hammett's equation. Therefore, it is possible to write individual contributions to the inductive and resonance effects for the rate or equilibrium position of the reaction or property measured [17,18,20,23,24,32].

Equation 4 represents the expanded Hammett's equation considering the polar effect of the substituent group separated on its inductive and resonance components, as well as their respective influences on the rate or equilibrium position of the reaction or property measured.

$$\log K = \rho_I \sigma_I + \rho_R \sigma_R + \log K_0 \quad (4)$$

where:

$\log K$ is the equilibrium constant of the substituted compound;

$\log K_0$ is the equilibrium constant of the nonsubstituted compound;

ρ_I is the angular coefficient relative to the inductive effect;

σ_I is the contribution due to inductive effect;

ρ_R is the angular coefficient relative to the resonance effect; and

σ_R is the contribution due to resonance effect.

With the help of this equation, it is possible to calculate separately the susceptibility of the reaction or physicochemical property measured to inductive and resonance effects of substituents. It has been employed with success to various systems in which the original Hammett's equation cannot provide good correlation [9,18,20,32].

1.1.1.2 Corrections of the constant σ of Hammett

In 1968, Swain and Lupton [44] revised the constant σ of Hammett and proposed a decomposition of the electronic contribution of substituent groups on inductive and resonance effects. These authors considered the combination of the inductive or field effect, \mathfrak{F} , with the resonance effect, \mathfrak{R} , equivalent to the σ of Hammett [21,26]. The correction proposed by Swain and Lupton is represented by eq. 5.

$$\sigma = f\mathfrak{F} + r\mathfrak{R} \quad (5)$$

where:

σ is the group constant of Hammett;
 \mathfrak{F} represents inductive and field effects;
 \mathfrak{R} represents the resonance effect;
 f and r correspond to independent empirical coefficients.

Swain and Lupton studies were based on two premises: (i) the effect of *para*-substituents on the ionization of bicyclo-(2,2,2)-octane-1-carboxylic acid is a result of the inductive effect, once resonance effect cannot be transmitted by saturated systems; and (ii) the electronic effect resulting from the substituent trimethylammonium, $-\text{N}(\text{CH}_3)_3$, does not involve resonance effect and, therefore, can be considered as zero. Based on these statements, it was possible to calculate \mathfrak{R} employing eq. 6.

$$\sigma_p = 0.56\mathfrak{F} + \mathfrak{R} \quad (6)$$

where:

σ_p is the constant σ of Hammett at the *para*-position;
 \mathfrak{F} represents inductive and field effects;
 \mathfrak{R} is the resonance effect; and
 0.56 is the value of α for the system of reactions used by Swain and Lupton.

Hansch [18,20] reevaluated the treatment given by Swain and Lupton to the σ of Hammett and concluded that it was not correctly dimensioned. This author proposed the utilization of a factor corresponding to α for the ionization of benzoic acid in ethanol/water 50 % v/v at 25 °C. In this sense, Hansch redefined the equation to calculate \mathfrak{F} and \mathfrak{R} values, and his proposition can be summarized by eqs. 7 and 8.

$$\mathfrak{F} = 1.396\sigma_m - 0.373\sigma_p - 0.009 \quad (7)$$

$$n = 14 \quad r = 0.992 \quad s = 0.0042$$

$$\mathfrak{R} = \sigma_p - 0.921 \mathfrak{F} \quad (8)$$

In spite of the criticism, the electronic descriptors \mathfrak{F} and \mathfrak{R} of Swain and Lupton have been successfully employed in numerous works, especially those related to QSAR studies [20,22,24,32]. Table 1 presents corrected values of the constant σ of Hammett and its derived constants, as well as values of \mathfrak{F} and \mathfrak{R} of Swain and Lupton.

Table 1 Selected values of constants for substituent groups frequently present in drugs.

X	σ_p	σ_m	σ_I	σ_R^*	\mathfrak{S}	\mathfrak{R}
NO ₂	0.78	0.71	0.64	0.14	0.65	0.13
F	0.06	0.34	0.52	-0.46	0.45	-0.39
Br	0.23	0.39	0.44	-0.21	0.45	-0.22
Cl	0.23	0.37	0.47	-0.23	0.42	-0.19
H	0.00	0.00	0.00	0.00	0.00	0.00
CH ₃	-0.17	-0.07	-0.04	-0.13	0.01	-0.18
C ₂ H ₅	-0.15	-0.07	-0.01	-0.14	0.00	-0.15
OH	-0.37	0.12	0.29	-0.66	0.33	-0.70
OCH ₃	-0.27	0.12	0.27	-0.54	0.29	-0.56
OC ₂ H ₅	-0.24	0.10	0.28	-0.52	0.26	-0.50
COCH ₃	0.50	0.38	0.30	0.15	0.33	0.17
CF ₃	0.54	0.43	0.40	0.14	0.38	0.16
CN	0.66	0.56	0.53	0.13	0.51	0.15
NH ₂	-0.66	-0.16	0.12	-0.78	0.08	-0.74
N(CH ₃) ₂	-0.83	-0.16	0.06	-0.89	0.15	-0.98
SO ₂ NH ₂	0.60	0.53	0.46	0.14	0.49	0.11

where:

X is the substituent group;

σ_p , σ_m , σ_I , and σ_R^* are the group constants derived from the σ of Hammett; and

\mathfrak{S} and \mathfrak{R} are the group constants suggested by Swain and Lupton;

*Values related to the *para*-position in relation to the center of reaction.

Source: C. Hansch, A. Leo, D. Hoekman. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*, Vol. 2, ACS Professional Reference Books, American Chemical Society, Washington, DC (1995).

1.1.2 Other structural descriptors of electronic character

Many other structural descriptors of electronic character can be defined by means of variations on physicochemical properties of chemical compounds sensitive to the electronic effect of substituent groups. Spectroscopic parameters such as IR absorption frequency, NMR-¹H and NMR-¹³C chemical displacements, ionization constants, and dipolar moment are some of the parameters reflecting the electronic influence of substituent groups on the properties of a given compound. Some of these parameters, for example, IR absorption frequency and NMR-¹H and NMR-¹³C chemical displacements, have been successfully employed as structural descriptors of electronic character in QSAR studies [20,23,26,32,46,52].

1.2 Hydrophobicity and related structural descriptors

The biological response generated by bioactive compounds is intrinsically related to the mechanisms of absorption, distribution, and intrinsic activity [8,23,26,58]. The absorption and distribution processes are basically regulated by the hydrophobicity of the compounds. To exert biological activity, these compounds must dissolve and cross biological membranes that constitute tissues and multicompartimental systems of living beings, and finally reach the site of action [25,26,47,58,59].

The classical works of Overton [36] and Meyer [33], together with numerous correlations found between biological activity and hydrophobicity, clearly demonstrate the importance of this property on expressing the bioactivity of compounds. Nevertheless, hydrophobicity is only one of the physicochemical acting properties orchestrated to determine the biological activity of these compounds [8,20,23,26,58,59].

It is important to consider that the hydrophobicity determines the rate and extension of absorption and distribution of a drug, its binding capacity to the receptor site, as well as biotransformation and excretion. To understand these steps, it is necessary, from one side, the knowledge of biological membranes structure, functions, and properties, and from the other side, the knowledge of the drug structure and physicochemical properties such as molecular size and shape, ionization degree, and relative liposolubility of its ionized and neutral forms [5,8,25,47,58,59].

One of the biological membrane models adopted in QSAR studies is the fluid mosaic model that represents the biological membrane as a continuous bilayer of polar lipids embedded with flexible proteins in all its extension. According to this model, the membranes can be seen as fluid and flexible structures, permeable to water and impermeable to ions and polar molecules. Figure 1 presents the fluid mosaic model of biological membranes.

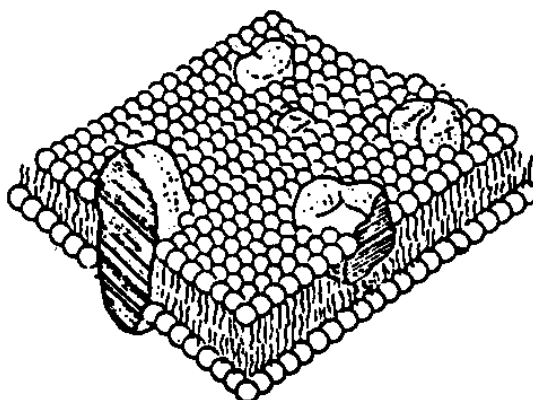


Fig. 1 Fluid mosaic model proposed to represent biological membranes. Source: J. B. Taylor, P. D. Kennevell. *Introductory Medicinal Chemistry*, John Wiley, New York (1981).

Many of the therapeutic drugs available are weak acids or weak bases. In solution, these drugs are found as an equilibrium between ionized and neutral forms. The last one usually presents stronger

liposolubility and may diffuse through biological membranes passively, depending on the concentration gradient of the system and on the molecule partition coefficient. The ionized form, on the other hand, may be actively absorbed and distributed according to its pK_a and the pH of the medium [8,23,26,58,59].

1.2.1 Partition coefficient

The partition coefficient of a chemical compound can be defined as the ratio between the substance concentration in an organic phase and the substance concentration in an aqueous phase of a system at equilibrium conditions. It is associated with the change in free energy generated by the substance on the thermodynamic equilibrium of the system [8,20,25,26,31,47]. The partition coefficient can be represented by eq. 9.

$$P = [\text{organic}]/[\text{aqueous}] \quad (9)$$

where:

P is the partition coefficient of the compound;

[organic] is the concentration of compound in the organic phase at equilibrium conditions; and

[aqueous] is the concentration of compound in the aqueous phase at equilibrium conditions.

The partition coefficient is one of the most widely physicochemical parameters employed in QSAR studies [1,5,23,25,32,58,59], and, therefore, its determination is frequently required. Considering the importance of this parameter, a solvent system to be preferentially used in partition coefficient determination was established. Nevertheless, numerous solvent systems find application in partition coefficient determination [2,42]. In this sense, the system *n*-octanol/phosphate buffer pH 7.4 is considered preferential in obtaining partition coefficient values for QSAR studies. The main advantages of *n*-octanol as organic phase are as follows [23,25].

- It possesses good solubilization power for various chemical compounds.
- The hydroxyl group of *n*-octanol can act both as electron donor or acceptor in hydrogen bonds.
- In spite of the immiscibility with water, *n*-octanol can dissolve up to 2.3 M of water under equilibrium conditions.
- The hydrogen bonds between the molecule and *n*-octanol may not be disrupted when transferring the molecule from the organic phase to the aqueous phase. Therefore, partition coefficient values measured in *n*-octanol/phosphate buffer systems reflect only hydrophobic interactions.
- It is not volatile at room temperature.
- It does not interfere with UV determinations, since it absorbs in a wavelength far lower than most parts of the drugs.
- It is chemically stable and commercially available.

One of the most common techniques for partition coefficient determination refers to the *shake-flask* method [23,25,26,31]. It is based on the simple principle of dissolving a chemical compound in a biphasic system constituted of a polar solvent and a nonpolar solvent.

Thin layer chromatography (TLC) and high-performance liquid chromatography (HPLC) are techniques that have been successfully employed in partition coefficient determinations of compounds presenting increased solubility in one of the phases of the system. TLC and HPLC have been used for determination of the parameters R_f and K' , respectively. The R_f value is correlated to R_m through eq. 10 and R_m is correlated to the partition coefficient through eq. 11 [23,25,26,30].

$$R_m = \log [1/ R_f - 1] \quad (10)$$

$$\log P = a. R_m + b \quad (11)$$

where:

R_m is the chromatographic distance reached by the studied compound relative to a standard compound;

R_f is the ratio between the chromatographic distance reached by the compound and the chromatographic distance reached by the eluent;

$\log P$ is the partition coefficient of the studied compound; and

b is the linear coefficient of the equation.

In HPLC, the retention time is quantitatively described by the capacity factor, K' , which, in turn, is given by normalization of the retention time. Accordingly:

$$K' = (t_r - t_0)/t_0 \quad (12)$$

where:

K' is the capacity factor;

t_r is the retention time; and

t_0 is the time of solvent elution.

A relationship between $\log K'$ and $\log P_{n\text{-oct}/\text{buffer}}$ can be observed applying the Collander's equation [2]. However, hydrophobicity given in terms of K' is only relative, since it implies the use of reference standards [25,30,31,34].

Determinations of partition coefficient by means of chromatographic techniques such as TLC or HPLC allow for certain degree of impurity of the solute. The *shake-flask* method, on the other hand, requires pure substances in order to provide reliable partition coefficient results [25,26,30,31].

1.2.2 The parameter π of Hansch

In analogy to the σ of Hammett, Hansch [18,23] has defined a new parameter, π , representing the

hydrophobic contribution of a substituent as the logarithm relationship between the partition coefficient of a substituted compound and the partition coefficient of the nonsubstituted analog. This relationship is represented by eq. 13.

$$\pi_x = \log P_x - \log P_H \quad (13)$$

where:

π_x is the parameter reflecting the hydrophobic contribution of the substituent group, X;

P_x is the partition coefficient of the substituted compound; and

P_H is the partition coefficient of the nonsubstituted compound.

The parameter π of Hansch has been widely employed in QSAR studies to quantify the hydrophobic contribution of substituent groups, and it measures the individual contribution of these groups to the partition coefficient of the molecule. The substituent group position influences significantly the value of π , and different values can be observed for the same substituent group depending on its position on the molecule [8,20,23,26,32,58,59]. Table 2 presents values of the parameter π of Hansch for substituent groups as a function of the position on substituted benzene rings

Table 2 Selected values of the parameter π of Hansch for substituent groups on bis-substituted benzene rings.

X	π	X	π
NO ₂	-0.28	OCH ₃	-0.02
F	0.14	OC ₂ H ₅	0.38
Cl	0.71	COCH ₃	-0.55
Br	0.86	CF ₃	0.88
H	0.00	CN	-0.57
CH ₃	0.56	NH ₂	-1.23
C ₂ H ₅	1.02	N(CH ₃) ₂	0.18
OH	-0.67	SO ₂ NH ₂	-1.82

where:

X is the substituent group and

π correspond to the values of the hydrophobicity constant π of Hansch.

Source: C. Hansch, A. Leo, D. Hoekman. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*, Vol. 2, ACS Professional Reference Books, American Chemical Society, Washington, DC (1995).

1.2.3 The constant *f* of Rekker

The QSAR analysis is based on the additive concept [17,20,23,25,26,29,42], which assumes that all

structures of a bioactive compound may contribute, under influence of its neighborhood, to the hydrophobic character of the compound.

Rekker [34,39] has observed some errors on the additive concept for the π of Hansch in systems presenting aliphatic chains [25,26]. Based on his observations, this author revised the hydrophobic contribution of individual substituent groups and defined the hydrophobicity parameter f by means of eq. 14.

$$\log P = \sum a.f \quad (14)$$

where:

$\log P$ is the partition coefficient of the compound;

f is the fragment constant and expresses the individual contribution of each sub-structure of the compound;

a is the frequency of a fragment on the structure analyzed.

The most important difference between π and f is that π refers to relative contributions of replacing H by a substituent group X , whereas f is related to absolute contributions of the substituent groups and substructures to the lipophilic character of the whole molecule. Therefore, for aromatic systems, the π and f values of the substituent groups are employed, while for aliphatic systems only f values are adequate.

Based on the additive concept and considering individual contributions of each substituent group or substructure to the molecule partition coefficient, mathematical models were developed to estimate this parameter. Partition coefficients calculated by these models usually are in good quantitative agreement with experimentally obtained ones [1,5,23,26,30,40,47,53]. Nevertheless, discrepancies may be present in compounds possessing juxtaposed aromatic rings, heterocyclic rings, or in ionic compounds. There are numerous software packages available to estimate partition coefficient, for example, the CLogP Program, from the company Biobyte.

1.3 Stereochemistry and structural descriptors

Knowledge of the three-dimensional features of bioactive compounds, as well as of its receptors in the biological system, is fundamental to comprehend mechanisms involved on the biological activity of the compounds [20,23,28,29].

The theory of drug action is currently based on the adjustment between drug molecule and its receptor and, therefore, flexibility is one of the most important properties of the biomacromolecules involved. Stereochemistry of bioactive compounds governs the recognition and discrimination of molecules and stereoselectivity is an important aspect in this matter [20,23,32,58,59].

The Hansch Analysis, also known as QSAR-2D, is a powerful tool to investigate quantitative relationships between chemical structure and biological activity. However, it presents some limitations regarding the three-dimensional aspects of drug–receptor interactions. One of the reasons for this is the Hansch's equation by itself with its bi-dimensional character. Another consideration is that structural descriptors related to three-dimensional features cannot be well represented by only two dimensions. Some of the structural descriptors suggested to represent the steric effect of substituent groups are introduced subsequently.

The first attempt to develop a steric parameter adequate to the Hansch Analysis was made by Taft [45]. This author has considered the acid hydrolysis of *alpha*-substituted acetates as a model reaction to define the parameter E_s . This parameter was then defined as the logarithm relationship between the hydrolysis constants of the substituted analogs and the hydrolysis constants of the corresponding non-substituted acetates. Since other factors may also influence the hydrolysis reactions taken as standard, the use of the parameter E_s has been strongly criticized [17,18,20,23,26,29,32].

Another steric parameter of great importance is the van der Waals volume, V_w , related to the relief of a given molecular structure or substructure. The V_w can be defined as the surrounding volume of a molecule which is impenetrable by other molecules, according to the law of mass action. It is important to keep in mind that the three-dimensional coordinates of a region on the molecule may change according to its physical state. Therefore, this parameter should be used with restriction in QSAR studies [20,23,26].

In spite of the indirect relationship between molecular weight, M_w , and molecular volume, the first one does not express conveniently three-dimensional features of the molecule. Nevertheless, the molecular weight is employed as steric parameter in QSAR studies [20,23,26].

In 1976, Verloop and collaborators [55] have developed a set of parameters more adequate to express three-dimensional aspect of substituent groups or substructures. These parameters, known as Sterimol, are based on measurements of the substituent group size in five different directions. The distances obtained were calculated by the Sterimol program, which simulates three-dimensional models of molecular structures or substructures. The standard directions suggested by Verloop for obtaining Sterimol parameters are schematically represented in Fig. 2.

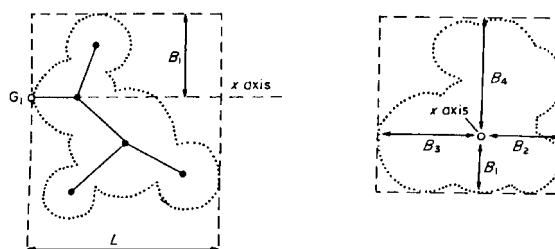


Fig. 2 Standard directions for the determination of Sterimol parameters. Source: A. Verloop, W. Hoogenstraaten, J. Tipker. In *Drug Design*, Vol. II, E. J. Ariëns (Ed.), pp. 165–207, Academic Press, New York (1976).

1.4 Mixed structural descriptors

There are a considerable number of other structural descriptors described in literature [21,24,26]. Among them, we can mention the molar refractivity, M_R , and the molar connectivity, χ . Until recently, these parameters were seldom employed, but currently they appear very frequently in QSAR studies.

1.4.1 Molar refractivity

The molar refractivity, M_R , expresses a constitutive and additive physicochemical property and, therefore, depends on the structure of chemical compound [20,21,23,26]. Values of M_R related to molecular substructures may be experimentally determined with the help of the Lorentz–Lorenz relationship [23,26], described by eq. 15.

$$M_R = \frac{n^2 - 1}{n^2 + 2} \times \frac{M_W}{d} \quad (15)$$

where:

n is the refraction index of the compound at 20 °C;

d is the density of the compound at 20 °C;

M_W is the molecular weight of the compound.

In 1977, Dunn [23] suggested the existence of a relationship between M_R value and the three-dimensional conformation of the molecular structure of a drug. Accordingly, the molar refractivity would be correlated to the adjustment of the drug molecule to the receptor.

On the other side, the M_R value of a given substituent group can be correlated to its lipophilicity, expressed by the constant π of Hansch [20,23,26]. A high correlation between the values of M_R and π is observed, since both parameters depend, at some extension, on the substituent molar volume, M_W/d . Therefore, M_R values may also express the lipophylic character of substituent groups, and both colinear and orthogonal relationships may be observed. Table 3 gives molar refractivity values of common substituent groups in QSAR studies.

Table 3 Selected values of molar refractivity of substituent groups frequently present in drugs

X	M_R	X	M_R
NO ₂	0.74	OCH ₃	0.79
F	0.09	OC ₂ H ₅	1.25
Br	0.89	COCH ₃	1.12
Cl	0.60	CF ₃	0.50
H	0.10	CN	0.63
CH ₃	0.56	NH ₂	0.54
C ₂ H ₅	1.03	N(CH ₃) ₂	1.55
OH	0.28	SO ₂ NH ₂	1.23

where:

X is the substituent group and

M_R corresponds to the molar refractivity.

Source: C. Hansch, A. Leo, D. Hoekman. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*, Vol. 2, ACS Professional Reference Books, American Chemical Society, Washington, DC (1995).

1.4.2 Molecular connectivity

The molecular connectivity, χ , describes molecular structure in terms of topology. It is based on relationships between the degree of ramification of a chemical compound and partition coefficients [21,23,26,59].

In order to determine molecular connectivity, the chemical structure is broken into substructures, each one having its contribution quantified by mathematical models, with computer help. Since it is an artificial and empiric parameter, the molar connectivity presents great probability of correlations with hydrophobicity parameters, as well as with molar refractivity [23,26].

The good correlations between molecular connectivity and partition coefficient and between molecular connectivity and molar refractivity indicate that this parameter can represent adequately the hydrophobic character of substituent groups, as well as intramolecular polar interactions. In spite of the controversial interpretation of the correlations generated, molecular connectivity has been frequently employed in QSAR studies [20,23,26,32,59].

1.5 Parameters expressing biological activity

The quality of the free energy linear relationships is a function of nature and precision of the structural descriptors employed. In QSAR studies, the more precise and exact the biological activity

measurements, the higher the statistical meaning of correlations obtained, and structural descriptors usually can be precisely measured.

Biological parameters express the potency of biologically active compounds [8,14,20,23,26,58,59] that can be defined as the inverse of dose or concentration needed to reach a given biological response. Therefore, the lower the dose required for biological response, the higher the potency of the compound.

The potency of bioactive compounds can be estimated based on biological tests. This test also finds application in screenings for new drug candidates as well as efficacy evaluation, including toxicity analysis. Biological tests can be classified into qualitative, quantal, and quantitative tests [8,32,57].

Qualitative tests aim at selecting and identifying chemical compounds presenting some biologic activity in mixtures or vegetal or animal extract tissues. This variant of biological test is important in searching new compounds of pharmacological interest, as well as in identifying bioactive contaminants in pharmaceutical preparations.

Quantal tests are based on a positive/negative answer of a biological system to a bioactive compound, without quantifying the observed effect. Nevertheless, it is possible to determine the minimum dose of the compound. Classical examples of quantal tests correspond to the 50 % efficacy dose, ED_{50} , 50 % lethal dose, DL_{50} , and minimal inhibition concentration, MIC.

The use of quantitative tests aims at measuring biological response resulting from a pre-established dose of a given compound, so that the response is proportional to the dose [8,23,26,33,58]. In this kind of test, sigmoidal dose–response curves are obtained, where only the region corresponding to a range from 20 to 80 % of the maximum effect is linearly related to the logarithm of the dose. In spite of this limitation, dose–response curves are well accepted for practical purposes within the limits of linearity.

1.5.1 Characteristics of an ideal biological test

Planning of biological tests for the determination of bioactive compound potency usually is based on the analysis of standard response curves as a function of the dose [8,23,26,32]. The use of only one dose to study a given biological response may result in errors such as the non-observation of activity for the analyzed dose of the compound. This error is observed when the studied dose is below the minimum threshold required for biological response. Therefore, one can infer that different compounds may cause maximum biological response, but differing drastically in potency [8,26,32,58].

Potency of bioactive compounds may also be measured by means of dose–response curves. This approach is particularly adequate to evaluate the activity of orally administered compounds, in which the time required to initiate the biological response varies for each analog relative to the standard compound. Another application refers to compounds with different periods of action. In this kind of test, each analog is administered in different doses and the biological response observed at different

time intervals. For each dose, a curve of biological response as a function of time is obtained and the area under the curve is calculated. Subsequently, a graph considering the values of the areas obtained as a function of the logarithm of the corresponding doses is obtained. Therefore, the relative potency of the analogs is the inverse of the dose required to promote a biological response [8,26,58].

1.5.2 Biological parameters employed in QSAR studies

The determination of bioactive compound activity for QSAR studies requires quantification of the dose or effect generated by the compound under test. Concerning the dose, studies with the potency expressed by means of logarithm of molar concentration are frequently found in the literature [23,26]. Among the biological parameters most employed in QSAR studies, one can name the 50 % efficacy dose, ED₅₀, the 50 % lethal dose, LD₅₀, and the minimum inhibition concentration (MIC). In QSAR studies, it is also common to evaluate the potency of bioactive compounds by quantitative tests, in other words, applying dose–response or time–response curves. This procedure is particularly employed to obtain information on the pharmacodynamic profile of compounds [8,23,26,32,58,59].

2. THE HANSCH ANALYSIS

The drug–receptor interaction can be described in terms of energy contributions. In this sense, biological activity would be associated to changes in free energy taking place in the absorption, distribution and biotransformation processes, or even to the drug–receptor interaction by itself. Therefore, biological response can be expressed as a complex function involving structural features of a drug and its concentration in the receptor site [8,23,26,58,59].

To establish quantitative structure-activity relationships, QSAR, biological activity is assumed to be correlated to the change in free energy, ΔG , associated to drug–receptor interaction [7,8,23,26,32,58,59]. This change in free energy can be divided into independent and additive terms related to the physicochemical properties of a given compound. For a series of analog compounds interacting with the biological system, the changes in free energy can be represented by eq. 16.

$$\Delta G = a(\Delta GE) + b(\Delta GH) + c(\Delta GS) + R \quad (16)$$

where:

ΔG is the free energy change;

a, b, c correspond to the contributions of each term e;

ΔGE , ΔGH , and ΔGS are free energy changes associated to physicochemical properties that influence biological activity.

According to the Hansch [19] model, the terms in eq. 16 can be represented by parameters
<www.iupac.org/publications/cd/medicinal_chemistry/>

measuring the influence of each physicochemical property involved on biological activity determination, as demonstrated by eq. 17.

$$BA = a\sigma + b\pi + cE_s + d \quad (17)$$

where:

BA corresponds to biological activity;

σ is the substituent group constant;

π is the hydrophobicity constant;

E_s is the steric constant;

a, b, c represent the individual contribution of each physicochemical property to the biological activity;

d is a constant.

Equation 17 may be simple or multiparametric, and each one of the terms may be linear or quadratic; collinear parameters must be avoided [23,26,38].

The Hansch analysis correlates biological activity with physicochemical properties expressed by structural descriptors, by means of linear, nonlinear, or linear multiple regression [20,26]. Since practically all structural parameters employed in the Hansch analysis are related to linear variations of free energy (they are derived from equilibrium or rate constants), synonyms to this method such as the Linear Free Energy Method or the Extrathermodynamics Method have been used. More recently, the Hansch analysis has also been designated QSAR-2D, since it considers only two dimensions: biological activity as a function of physicochemical properties.

In 1964, Hansch and Fujita [19] combined different structural descriptors on a linear equation, eq. 18, where C is the molar concentration generating biological effect, $\log 1/C$ is the biological potency, and a, ρ , and b are coefficients given by the regression analysis of the set of data.

$$\log 1/C = a \log P + \rho\sigma + b \quad (18)$$

The linear model expressed by eq. 18 was extended to a parabolic model, given by eq. 19, by the introduction of a quadratic term related to hydrophobicity [20]. The rationale for this model is that pronounced hydrophilic molecules are not able to cross very hydrophobic biological barriers, or vice-versa. The parabolic model expresses the increase in hydrophobicity relative to the increase in biological activity until it reaches an optimum of hydrophobicity, then activity tends to decrease according to the parabolic profile.

$$\log 1/C = a (\log P)^2 + b \log P + \rho\sigma + c \quad (19)$$

When compared to the parabolic model, the bilinear model developed by Kubinyi [27] in 1977, eq. 20, may demonstrate a better adjustment to experimental data to express QSARs. This observation
www.iupac.org/publications/cd/medicinal_chemistry/

results from the bilinear term, β , that assumes probabilistic character and describes the compound movement in multicompartiment systems, such as the biological system where the potency of the analyzed compound is strongly dependent on the partition coefficient, P .

$$\log 1/C = a \log P - b \log (\beta P + 1) + c \quad (20)$$

The introduction of multiparametric mathematical models has made possible to describe complex relationships between biological activity and the physicochemical properties of a bioactive compound. In the last three decades, the biological activity of numerous compounds has been successfully correlated to structural descriptors by linear, parabolic and bilinear models, establishing qualitative and quantitative relationships between structure and activity of compounds.

2.1 Strategies for the Hansch analysis

The most adequate method to delineate substituent groups for a QSAR study by Hansch analysis of a series of analog compounds is the Craig's diagram [3]. This method is based on the analysis of physicochemical parameters intercorrelation and involves the construction of diagrams with structural descriptors on a Cartesian distribution.

Based on statistical studies and on the need to select adequate substituent groups for QSAR studies, in 1971 Craig proposed a two-dimensional diagram with the coordinates corresponding to independent structural descriptors. Craig's diagram is frequently employed and provides foundations to many of the most advanced techniques of planning series for QSAR studies [20,28,32,49,50]. Figure 4 represents Craig's diagram for the structural descriptors π and σ .

The choice of adequate substituent groups when defining a series for the study of biological activity as a function of structural descriptors [23,26] must be based on the most dispersed selection possible on the diagram. In other words, substituents present in different quadrants must be selected, including -CH₃, -NO₂, -Cl or -Br, -H, -NH₂ or -N(CH₃)₂. However, when analyzing the influence of a given physicochemical property on the biological effect, substituents must be selected according to an imaginary line drawn parallel to the axis of the property studied. When choosing substituents on Craig's diagram, collinear parameters and, therefore, nonsignificant correlations should be avoided.

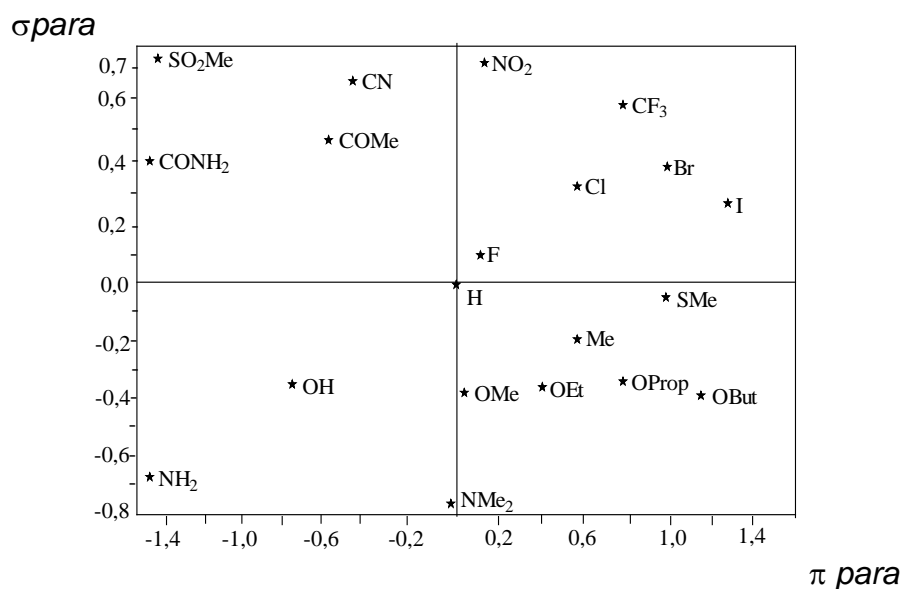


Fig. 4 Intercorrelation diagram for σ X π at aromatic *para*-substitution. Source: P. N. Craig. *J. Med. Chem.* **14**, 680 (1971).

3. STRATEGIES FOR OPTIMIZING THE BIOLOGICAL ACTIVITY OF DRUGS

Therapeutically more interesting derivatives of a drug can be obtained with adequate design of the variations on the structure of the bioactive compound. Some of the advantages a new derivative can offer are higher activity, lower toxicity and better pharmacotechnical properties [3,4,28].

The definition of adequate series for a QSAR study requires identification of a compound presenting only some of the biological properties that may result in a better biological performance. The optimization of this lead or standard compound properties is a result of systematic studies. The QSAR study of a series of analog compounds based on the Hansch approach involves the strategic planning represented in Fig. 5.

The analysis of correlations between structural descriptors and biological activity for a series of analog compounds allows, at some extension [23,26,32], the prediction of non-available analogs biological activity, by means of interpolation or extrapolation of values obtained in the correlation. Interpolative predictions are frequently employed on the evaluation of the correlation model efficacy, under statistical confidence limits.

The Hansch analysis of well-planned series with reliable predictive value [20,26,32] allows quantitative estimation of biological activity of analogs and, consequently, the design of new analogs with gain of time and investment

When performing a QSAR study, the decision to follow or interrupt the synthesis program may be done with the help of an equation based on partial data. The introduction of new analogs on the series may be interrupted, for example, when the compound of higher activity is already included on the group of analogs tested.

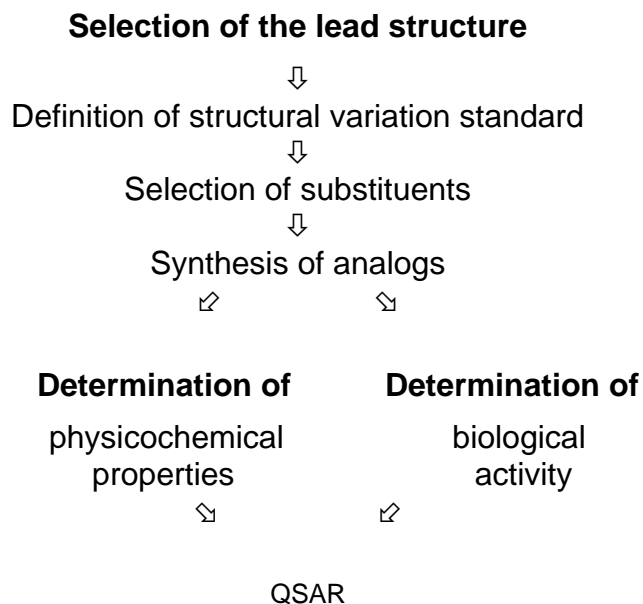


Fig. 5 Strategies for QSAR study applying the Hansch analysis.

Correlations obtained in QSAR studies can only be used in predictions for compounds structurally analogous to the components of the series studied and within the range of values employed. According to Martin [32], an ideal series of analogs is not necessarily the one with the lowest number of compounds, but it should ensure that the most active compound of the series is included or suggested by the analysis. On the other hand, the absence of a free energy relationship with the data obtained suggests that the series is not well delineated or that other parameters must be considered for correlation.

Studies of relationships between chemical structure and biological activity of drugs have contributed significantly to the generation of knowledge of qualitative and quantitative contributions by physicochemical properties to drug action. QSAR studies have also contributed to optimize the performance of drugs already available, generating knowledge each day that is absorbed and employed by the pharmaceutical industry. The development of QSAR field presents very good perspectives, especially on the planning of new drugs, still not fully explored.

4. SOFTWARES APPLIED TO QSAR STUDIES EMPLOYING THE HANSCH ANALYSIS

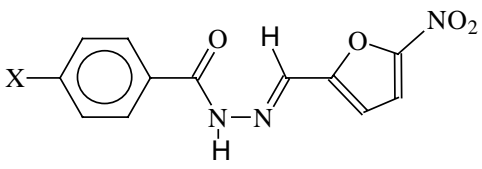
In order to establish quantitative relationships between chemical structure and biological activity by means of Hansch analysis, QSAR-2D, simple or multiparametrical equations, based on linear, parabolic, and bilinear models, correlating structural descriptors and biological activity are necessary. Usually, computational packs are employed to get these equations from a set of experimental data. Linear and parabolic equations may be obtained with the help of most of the software engineering tools available on marketing. Bilinear equations, on the other hand, can only be obtained employing

more sophisticated software or the ones developed specially for this kind of analysis.

Since the bilinear model proposed by Kubinyi, researchers have found some difficulties in obtaining correlations with this model. Aiming at solving this problem, the same author developed, in 1994, the software Billin program. It was specially adapted to QSAR studies by Hansch analysis and has facilitated considerably the generation of linear, parabolic, and bilinear model correlations. The Billin program even allows generation and friendly visualization of statistical parameters resulting from the correlation, useful on adequate evaluation of the experimental data modeling. Considering the high number of correlations that need to be investigated in a Hansch analysis, this program allows systematic comparison of correlations obtained with different mathematical models, establishing a rationale for the correlation choice and saving time.

The Billin program was reviewed in 1998 and altogether with a tutorial guide and a few application examples, it was offered by the author to the Brazilian scientific community for unrestricted use. Nevertheless, we recommend the set of data presented in Table 4, referring to antimicrobial potency of 5-nitro-2-thiophylidene derivatives against *Staphylococcus aureus*. This set of data was generated on the Laboratory of Drug Design and Development, Department of Biochemical and Pharmaceutical Technology of the Faculty of Pharmaceutical Sciences, University of São Paulo, under our supervision.

Table 4 Antimicrobial potency of 5-nitro-2-thiophylidene benzhydrazides 4-X-substituted and values of structural descriptors related.

								
X	log1/C(μ M)	π	CLogP	σ_p	σ_1	σ_R^*	ζ	\mathfrak{R}
NO ₂	-0.70	-0.28	2.05	0.78	0.64	0.14	0.65	0.13
Cl	-0.74	0.71	2.91	0.23	0.47	-0.24	0.42	-0.19
H	-1.22	0.00	2.10	0.00	0.00	0.00	0.00	0.00
CH ₃	-1.13	0.56	2.60	-0.17	-0.04	-0.13	0.01	-0.18
OCH ₃	-0.91	-0.02	2.41	-0.27	0.27	-0.54	0.29	-0.56
OC ₂ H ₅	-0.97	0.38	2.93	-0.24	0.28	-0.52	0.26	-0.50
CF ₃	-0.75	0.88	3.15	0.54	0.40	0.14	0.38	0.16
CN	-0.85	-0.57	1.76	0.66	0.53	0.13	0.51	0.15
NH ₂	-0.98	-1.23	1.55	-0.66	0.12	-0.78	0.08	-0.74
SO ₂ NH ₂	-1.50	-1.82	0.98	0.60	0.46	0.14	0.49	0.11

where:

X is the substituent group;

π is the hydrophobicity constant of Hansch;

CLogP is the partition coefficient calculated by the CLogP Program, version 4.0, Biobyte Co.

σ_p , σ_1 , σ_R , \mathfrak{S} , and \mathfrak{R} are the group constants related to electronic effects;

Source: C. Hansch, A. Leo, D. Hoekman. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*, Vol. 2, ACS Professional Reference Books, American Chemical Society, Washington, DC (1995).

The correlation study of the data presented in Table 4, with the help of the Billin program, version 1998, and considering linear and parabolic bilinear models, resulted in statistically significant equations to estimate potency. Therefore, it is clear the importance and applicability of the Hansch analysis, QSAR-2D, on identifying qualitative and quantitative influences of physicochemical properties on the biological activity of drug candidates, with use of the generated knowledge for the design of new analogs with better pharmacological profile.

REFERENCES

1. P. Buchwald, N. Bodor. *Proteins: Struct., Funct., Genet.* **30**, 86 (1998).
2. R. Collander *Acta Chem. Scand.* **5**, 774 (1951).
3. P. N. Craig. *J. Med. Chem.* **14**, 680 (1971).
4. F. Darvas. *J. Med. Chem.* **17**, 799 (1974).
5. J. DeVillers, D. Domine, C. Guillon, W. Karcher. *J. Pharm. Sci.* **87**, 1086 (1998).
6. S. W. Dietrich. "Chemical information computing systems in drug discovery", in *Burger's Medicinal Chemistry and Drug Discovery: Principles and Practice*, Vol. 1, 5th ed., M. E. Wolff, pp. 415–496, John Wiley, New York (1995).
7. (a) J. Ferguson. *Proc. R. Soc., Ser. B* **127**, 387 (1939); (b) M. C. Richet. *Compt. Rend. Soc. Biol. (Paris)* **45**, 775 (1893); (c) Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 4, VCH, New York (1993).
8. W. O. Foye, T. L. Lemke, D. A. Williams. *Principles of Medicinal Chemistry*, Williams & Wilkins, 4th ed., Baltimore (1995).
9. T. Fujita. "The extrathermodynamic approach to drug design", in *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study & Therapeutic Application of Chemical Compounds*, Vol. 4, C. Hansch, P. G. Sammes, J. B. Taylor (Eds.), Chap. 21.1, pp. 497–560, Pergamon Press, Oxford (1990).
10. T. Fujita, T. Ban. *J. Med. Chem.* **14**, 148 (1971).
11. T. Fujita, J. Iwasa. *J. Am. Chem. Soc.* **86**, 5175 (1964).
12. R. D. Gillion, W. P. Purcell, T. R. Bosin. *Eur. J. Med. Chem.* **12**, 187 (1977).

13. A. Gringauz. *Medicinal Chemistry: How Drugs Act and Why*, Wiley-VCH, New York (1997).
14. L. P. Hammett. *Physical Organic Chemistry*, McGraw-Hill, New York (1940).
15. T. L. Hammett. *J. Am. Chem. Soc.* **59**, 963 (1937).
16. C. K. Handoch, E. A. Meyer, B. J. Yager. *J. Am. Chem. Soc.* **83**, 4211 (1961).
17. C. Hansch. *Acc. Chem. Res.* **2**, 232 (1969).
18. (a) C. Hansch, A. Leo. *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley, New York (1976); (b) Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, VCH, New York (1993).
19. C. Hansch, T. Fujita. *J. Am. Chem. Soc.* **86**, 1616 (1964).
20. C. Hansch, A. Leo. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*, American Chemical Society, Washington, DC (1995).
21. C. Hansch, A. Leo, D. Hoekman. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*, Vol. 2, ACS Professional Reference Book, American Chemical Society, Washington, DC (1995).
22. C. Hansch, A. Leo, R. W. Taft. *Chem. Rev.* **91**, 165 (1991).
23. C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, Pergamon Press, Oxford (1990).
24. C. Hansch, A. Leo. Substituent constants for correlation analysis in chemistry and biology, John Wiley, New York (1976). Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 35, VCH, New York (1993).
25. H. Kubinyi. *Prog. Drug Res.* **23**, 97 (1979).
26. H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, Wiley-VCH, New York (1993).
27. H. Kubinyi. *J. Med. Chem.* **20**, 625 (1977).
28. H. Kubinyi. *Pharmazie* **50**, 647 (1995).
29. H. Kubinyi. "The quantitative analysis of structure-activity relationship", in *Burger's Medicinal Chemistry and Drug Discovery: Principles and Practice*, 5th ed., Vol. 1, M. E. Wolff, pp. 497–572, John Wiley, New York (1995).
30. W. J. Lambert. *J. Chromatogr., A* **656**, 469 (1993).
31. A. J. Leo. "Methods of calculating partition coefficients", in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, chap. 18.7, Pergamon Press, Oxford (1990).
32. Y. C. Martin. *Quantitative Drug Design: A Critical Introduction*, Medicinal Research Series, Vol.

- 8, Marcel Dekker, New York (1978).
33. H. Meyer. *Arch. Exp. Path. Pharm.* **42**, 109 (1899); Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 4, VCH, New York (1993).
34. G. G. Nys, R. F. Rekker. *Chim. Ther.* **8**, 521 (1973).
35. G. Oros, T. Cserhati, A. Vrbanova. *Biochem. Mol. Biol. Int.* **47**, 79 (1999).
36. E. Overton. *Studien über die Narkose, zugleich ein Beitrag zur allgemeinen Pharmakologie*, G.Fischer, Jena, 1901; English translation by R. L. Lipnich (Ed.), *Studies on Narcosis, Charles Ernest Overton*, Chapman and Hall, London (1991). Apud: H. Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, p. 4, VCH, New York (1993).
37. W. C. Patt, J. J. Edmons, J. T. Repine, K. A. Barryman, B. R. Reisdorf, C. Lee, M. S. Plummer, A. Shalvirpour, S. J. Haleel, J. A. Keiser, M. A. Flyman, K. M. Welch, E. E. Reynolds, R. Rubin, B. Tobias, H. Hallak, A. M. Dolvet. *J. Med. Chem.* **40**, 1063 (1997).
38. M. A. Pleiss, S. H. Unger. "The design of test series and the significance of QSAR relationships", in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 561–569, Pergamon Press, Oxford (1990).
39. R. F. Rekker, R. Mannhold. *Calculation of Drug Lipophilicity: The Hydrophobic Fragmental Constant Approach*, Wiley-VCH, Weinheim (1992).
40. M. C. Richet. *Compt. Rend. Soc. Biol. (Paris)* **45**, 775-776 (1893). Apud: Kubinyi. *QSAR: Hansch Analysis and Related Approaches*, VCH, New York (1993).
41. J. D. Roberts, W. T. Moreland. *J. Am. Chem. Soc.* **75**, 2167 (1953).
42. P. Seiler. *Eur. J. Med. Chem.* **9**, 473 (1974).
43. B. A., Steinbaugh, H. W. Hamilton, V. Prasad, K. S. Para, P. J. Tummino, D. Ferguson, A. Lunney, C. J. Blankley. *Bioorg. Med. Chem. Lett.* **6**, 1099 (1996).
44. C. G. Swain, E. C. Lupton. *J. Am. Chem. Soc.* **90**, 4328 (1968).
45. R. W. Taft. *J. Am. Chem. Soc.* **74**, 3120 (1952).
46. L. C. Tavares. *Rev. Farm. Quim.* **30**, 53 (1997).
47. P. J. Taylor. "Hydrophobic properties of drugs", in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 241–294, Pergamon Press, Oxford (1990).
48. J. G. Topliss, Y. C. Martin. *Drug Design*, Vol. 5, pp. 1–21, Academic Press, New York (1975).
49. J. G. Topliss. *J. Med. Chem.* **2**, 463 (1977).
50. J. G. Topliss. *J. Med. Chem.* **15**, 1006 (1972).
51. M. S. Tute. "History and objectives of quantitative drug design", in C. Hansch, P. G. Sammes, J.

- B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 1–31, Pergamon Press, Oxford (1990).
52. M. S. Tute. *Adv. Drug Res.* **6**, 1 (1971).
53. S. H. Unger. *J. Med. Chem.* **24**, 262 (1981).
54. M. C. Venuti. “The role of recombinant DNA technology in medicinal chemistry and drug discovery”, in WOLFF, M.E. *Burger’s medicinal chemistry*. 5.ed. New York: John Wiley, 1995. v.1, cap.16, p.661-696.
55. A. Verloop, W. Hoogenstraaten, J. Tipker. “Development and application of new steric substituent parameters in drug design”, in *Drug Design*, Vol. 7, E. J. Ariens, pp. 165–207, Academic Press, New York (1967).
56. H. V. Waterbeend. *Quant. Struct.-Act. Relat.* **6**, 200 (1992).
57. D. Weininger, J. L. Weininger. “Chemical structure and computers”, in C. Hansch, P. G. Sammes, J. B. Taylor. *Comprehensive Medicinal Chemistry: The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds*, Vol. 4, pp. 59–82, Pergamon Press, Oxford (1990).
58. D. G. Wermuth. *The Practice of Medicinal Chemistry*, Academic Press, San Diego (2000).
59. M. E. Wolf. *Burger’s Medicinal Chemistry: Principles and Practice*, 5th ed., pt. 1, John Wiley, New York (1995).

Leoberto Costa Tavares e Elizabeth Igne Ferreira

hajudan@usp.br