

Universidade Federal de Santa Catarina

Curso de Pós-Graduação em Química

Anna Celia Silva Arruda

**AMPLIAÇÃO E APLICAÇÃO DO MÉTODO SEMI-EMPÍRICO TOPOLÓGICO (I_{ET})
EM MODELOS QSRR/QSPR/QSAR PARA COMPOSTOS ALIFÁTICOS
HALOGENADOS E CICLOALCANOS**

Florianópolis

2008

ANNA CELIA SILVA ARRUDA

**AMPLIAÇÃO E APLICAÇÃO DO MÉTODO SEMI-EMPÍRICO TOPOLÓGICO (I_{ET})
EM MODELOS QSRR/QSPR/QSAR PARA COMPOSTOS ALIFÁTICOS
HALOGENADOS E CICLOALCANOS**

Tese submetida ao Curso de Pós-Graduação em Química da Universidade Federal de Santa Catarina, como requisito parcial para a obtenção do grau de Doutora em Química

Área de concentração: Química Analítica

Orientadora: Prof^a Dra. Vilma Edite Fonseca Heinzen

Co-orientador: Prof^o Dr. Rosendo Augusto Yunes

Florianópolis

2008

**À minha filha Caroline, que ela acredite nos seus
sonhos e tenha a tenacidade de persegui-los.**

AGRADECIMENTOS

◆ Aos orientadores, Profa. Dra. Vilma Edite Fonseca Heinzen e Prof. Dr. Rosendo Augusto Yunes, pela valiosa orientação acadêmica e pessoal, cujos exemplos de vida valem a pena serem seguidos.

◆ Aos pais, Célio e Gecy Arruda, pela dedicação ao longo da minha vida.

◆ Aos meus cúmplices, Guto e Carol, que contribuíram de forma primária para a realização deste trabalho.

◆ A querida amiga Rosane, pela sua generosidade.

◆ Aos queridos amigos Berenice, Sandra, Anderson, Renata, Érica e Ledilege pelo coleguismo durante a realização desse trabalho.

◆ Aos professores do Curso de Pós Graduação em Química da Universidade Federal de Santa Catarina por dividir a sua experiência.

◆ Aos funcionários Graça e Jadir, pela competência e gentileza a mim dispensada a mais de uma década.

◆ A Universidade de Caxias do Sul e CAPES, pelo suporte financeiro que permitiu a realização desta Tese.

SUMÁRIO

RESUMO.....	9
ABSTRACT.....	11
Índice de Figuras.....	13
Índice de Tabelas.....	15
Índice de Símbolos e Abreviações.....	18
1 -INTRODUÇÃO.....	23
1.1 - Considerações Gerais.....	23
1.2 - Histórico.....	26
2 - OBJETIVOS	29
3 - FUNDAMENTOS TEÓRICOS.....	31
3.1 - Correlações Quantitativas entre Estrutura - Propriedade/Atividade.....	31
3.2 - Modelo de Correlação Quantitativa entre Estrutura-Propriedade/Atividade.....	33
3.3 - Descritores Moleculares.....	39
3.3.1 - Grafos Moleculares.....	42
3.3.2 - Descritores Topológicos.....	44
3.3.3 - Descritores Físicos.....	53
3.3.4 - Descritores Químico-Quânticos.....	55
3.4 - Propriedades Moleculares e Aplicações da Correlação Quantitativa entre Estrutura-Propriedade.....	59
3.4.1 - Retenção Gás Cromatográfica.....	59
3.4.2 - Propriedades Físico-Químicas.....	62

3.5	- Atividades Biológicas e Toxicidade - Aplicações das Correlações Quantitativas entre Estrutura-Atividade.....	64
3.6	- Teoria do Método Semi-Empírico Topológico (I_{ET}).....	68
3.7	- Classe de Compostos Orgânicos Investigados.....	73
3.7.1	- Hidrocarbonetos Alifáticos Halogenados.....	73
3.7.2	- Cicloalcanos.....	76
4	- METODOLOGIA.....	79
4.1	- Índices de Retenção de Kováts para Compostos Alifáticos Halogenados e Cicloalcanos.....	79
4.2	- Propriedades Físico-Químicas, Termodinâmicas e Atividades Biológicas.....	83
4.3	- Cálculo do Índice Semi-Empírico Topológico para Compostos Alifáticos Halogenados.....	84
4.3.1	- Cálculo dos Valores dos Fragmentos C-X e X-C-X.....	88
4.3.2	- Cálculo dos Valores de C_i para Átomos de Carbono Terciários (>CH-) em Posições α , β e γ e >CH-X (átomo de halogênio diretamente ligado ao átomo de carbono terciário) em Posições α , β e γ em Relação ao Primeiro Fragmento C-X.....	89
4.4	- Cálculo do Índice Semi-Empírico Topológico para Cicloalcanos.....	90
4.5	- Métodos Estatísticos.....	91
4.5.1	- Desenvolvimento dos Modelos de QSRR/QSPR/QSAR por Regressão Linear Simples.....	91
4.5.2	- Avaliação do Potencial de Prognóstico dos Modelos de QSRR/QSPR/QSAR.....	91
4.5.3	- Validação Externa dos Resultados.....	92
5	- RESULTADOS E DISCUSSÃO.....	93
5.1	- Aplicação do I_{ET} em Correlações Quantitativas entre Estrutura – Retenção Cromatográfica para Compostos Alifáticos Halogenados.....	93
5.2	- Aplicação do I_{ET} em Correlação Quantitativa entre Estrutura – Retenção Cromatográfica para Cicloalcanos.....	114

5.3 - Aplicação do I_{ET} em Correlações Quantitativas entre Estrutura – Propriedades Físico-Químicas e Termodinâmicas de Compostos Alifáticos Halogenados.....	126
5.4 - Aplicação do I_{ET} em Correlações Quantitativas entre Estrutura – Atividade Biológica de Compostos Alifáticos Halogenados.....	141
6 - CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES PARA TRABALHOS FUTUROS.....	149
6.1 - Conclusões.....	149
6.2 - Recomendações para Trabalhos Futuros.....	151
7 – REFERÊNCIAS.....	152
APÊNDICE.....	177
Apêndice A.....	177

RESUMO

Este estudo foi desenvolvido para avaliar a capacidade de prognóstico do índice semi-empírico topológico (I_{ET}) em estimar a retenção cromatográfica de compostos alifáticos halogenados e cicloalcanos. Também foram desenvolvidos modelos de correlação quantitativa entre estrutura propriedade/atividade (QSPR/QSAR) para prever várias importantes propriedades físico-químicas, termodinâmicas e atividades tais como: ponto de fusão, PF ($^{\circ}C$), ponto de ebulição, PE ($^{\circ}C$), coeficiente de partição octanol/água, $\log P$, densidade, D (g/cm^3), índice de refração, n_D , volume molecular, v (A^3), coeficiente de absorção de Ostwald gás-hexadecano, $\log \gamma_{hd}$, equilíbrio vapor-líquido, $\log \rho^{gas}$ (mol/dm^3), solubilidade aquosa, $\log \rho^w$, volume de retenção específico sob 373K, $\log V_w$ (cm^3/g), diferença de entalpia padrão parcial molar de solvatação, $\Delta_{sol}H^{\circ}$, (J/mol); diferença de entropia padrão parcial molar de solvatação, $\Delta_{sol}S^{\circ}$, ($J/mol \cdot K$) e calor de formação, ΔH_f° ($kcal/mol$) de hidrocarbonetos alifáticos halogenados. Os valores dos índices de retenção experimentais (IR_{Exp}) de 141 hidrocarbonetos alifáticos halogenados foram usados para a construção do modelo da correlação quantitativa entre estrutura retenção (QSRR) com o I_{ET} . A regressão linear simples entre o IR_{Exp} e o I_{ET} mostrou boa qualidade estatística, com coeficiente de determinação de $r^2=0,9995$, desvio padrão de $SD=8$ e coeficiente de validação cruzada "leave-n-out" de $r^2_{cv}=0,999$. Bons resultados foram alcançados entre os valores experimentais de ponto de fusão, PF ($^{\circ}C$), e ponto de ebulição, PE ($^{\circ}C$), com o I_{ET} através de equações polinomiais quadráticas. A partir do modelo de QSPR obtido entre o PE ($^{\circ}C$) com o I_{ET} para 86 compostos ($r^2=0,9971$; $SD=4,2$; $r^2_{cv}=0,997$), foram calculados os valores para um grupo externo de 24 compostos ($r^2=0,9931$; $SD=7,6$). Uma boa correlação entre o PF ($^{\circ}C$) e o I_{ET} para 43 compostos foi obtida

($r^2=0,9865$; $SD=6,1$; $r^2_{cv}=0,985$) e os valores calculados de 8 compostos não incluídos no modelo inicial também mostraram muito boa correlação com os valores experimentais ($r^2=0,9903$; $SD=4,1$). Os valores de log P estimados através do Método Semi-Empírico Topológico foram comparados com aqueles obtidos com o Método de Contribuição dos Fragmentos, mais comumente usado. Para 19 compostos, as correlações obtidas entre os valores calculados e os experimentais com esses dois métodos foram de $r^2=0,9871$ e $r^2=0,9750$, respectivamente para os Métodos Semi-Empírico Topológico e Contribuição dos Fragmentos. Esses resultados mostram a capacidade de prognóstico do I_{ET} para propriedades físico-químicas e termodinâmicas. A habilidade de prognóstico da retenção cromatográfica pelo I_{ET} também foi verificada usando fases estacionárias com diferentes polaridades. O modelo de QSRR combinado incluiu um parâmetro (P_R) que representa a polaridade de cada fase. O gráfico do IR_{Cal} em relação ao IR_{Exp} mostrou a adequação do modelo ($r^2=0,9984$; $SD=18$). Resultados satisfatórios foram encontrados aplicando o I_{ET} para estimar a retenção cromatográfica de 48 cicloalcanos ($r^2=0,9905$; $SD=7$; $r^2_{cv}=0,997$) e o ponto de ebulição ($N=33$; $r^2_{cv}=0,988$). Esse método permitiu retirar informações sobre as características estruturais, eletrônicas e geométricas das moléculas que estão influenciando no processo de retenção cromatográfico e a distinção entre isômeros cis/trans dos compostos estudados. Os resultados obtidos demonstraram o potencial do I_{ET} para ser aplicado e expandido em modelos de QSRR/QSPR/QSAR para compostos alifáticos halogenados e cicloalcanos.

Palavras-chave: QSRR, QSPR, QSAR, Hidrocarbonetos Alifáticos Halogenados, Cicloalcanos, Método Semi-Empírico Topológico, Índice de Retenção Cromatográfico

ABSTRACT

This study was carried out to show the predictive capacity of the semi-empirical topological index (I_{ET}) in estimating the chromatographic retention of halogenated aliphatic and cycloalkanes compounds. Also were developed models of quantitative structure-properties/activities relationship (QSPR/QSAR) to predict several important physicochemical, thermodynamics properties and activity such: melting point, Mp ($^{\circ}C$), boiling point, Bp ($^{\circ}C$), octanol/water partition coefficient, $\log P$, density, D (g/cm^3), refraction index, n_D , molecular volume, v (A^3), gas-hexadecane Ostwald absorption coefficients, $\log \gamma_{hd}$, vapor-liquid equilibrium, $\log \rho^{gas}$ (mol/dm^3), water solubility, $\log \rho^w$, specific retention volumes at 373K, $\log V_w$ (cm^3/g), partial molar standard enthalpy of the solvation, $\Delta_{sol}H^{\circ}$ (J/mol) and partial molar standard entropy of the solvation, $\Delta_{sol}S^{\circ}$ (J/mol), heats of formation, ΔH_f° ($kcal/mol$) of halogenated aliphatic hydrocarbons. The values of experimental retention index (RI_{Exp}) of 141 aliphatic halogenated hydrocarbons were used in the building of model of quantitative structure-retention relationship (QSRR) with the I_{ET} . The simple linear regression between RI_{Exp} and the I_{ET} exhibits good statistical quality, with a coefficient of determination of $r^2=0.9995$, standard deviation of $SD=8$ and leave-n-out cross-validation correlation coefficient of $r_{cv}^2=0.999$. Fine results between experimental melting, Mp ($^{\circ}C$), and boiling Bp ($^{\circ}C$) points, and I_{ET} were achieved through quadratic polynomial equations. From de model of QSPR obtained between the Bp ($^{\circ}C$) with I_{ET} for 86 compounds ($r^2=0.9971$; $SD=4.2$; $r_{cv}^2=0.997$), the values of 24 compounds not included in the model were calculated ($r^2=0.9931$; $SD=7.6$). The correlation between Mp ($^{\circ}C$) and I_{ET} for 43 compounds was good ($r^2=0.9865$; $SD=6.1$; $r_{cv}^2=0.985$), and calculated values of 8 compounds not included in the initial model also showed very good correlation with

experimental values ($r^2=0.9903$; $SD=4.1$). The $\log P$ values estimated through of Semi-Empirical Topological Method were compared with those obtained with the Fragment Addition Method, more usually used. For 19 compounds, the correlations obtained between the calculated $\log P$ values and experimental using these two methods went of $r^2=0.9871$ and $r^2=0.9750$, respectively for Semi-Empirical Topological and Fragment Addition Methods. These results show the predictive capacity of the I_{ET} for these physicochemical and thermodynamic properties. The chromatographic retention predictive ability of I_{ET} also was verified in stationary phase of different polarities. The combined model of QSRR included a parameter (P_R) that represents the polarity of each phase. The fit of RI_{Cal} versus RI_{Exp} showed the model adequation ($r^2=0.9984$; $SD=18$). Satisfactory results were obtained using the I_{ET} for estimation the chromatographic retention of 48 cycloalkanes ($r^2=0.9905$; $SD=7$; $r^2_{cv}=0.997$) and the boiling point ($N=33$; $r^2_{cv}=0.988$). This method allowed withdraw information on the structural, electronic and geometric characteristics of the molecules that are influencing in the chromatographic retention process and the distinction between isomers cis/trans of the compound studied. The results obtained showed the power of I_{ET} for application and expanded in QSRR/QSPR/QSAR models for halogenated aliphatic and cycloalkanes compounds.

Keywords: QSRR, QSPR, QSAR, Halogenated Aliphatic Compounds, Cycloalkanes, Semi-Empirical Topological Method, Retention Index

ÍNDICE DE FIGURAS

- Figura 1:** A molécula do cicloexano representada como fórmula estrutural (a) e grafo molecular enumerado (b) onde foram omitidos os átomos de hidrogênio. Os vértices correspondem aos átomos e as linhas as ligações químicas..... 43
- Figura 2:** Gráfico da correlação linear simples entre os índices de retenção calculados (IR_{Cal}) e os índices de retenção experimentais (IR_{Exp}) para compostos alifáticos halogenados em fase estacionária de baixa polaridade, empregando a Equação 9..... 103
- Figura 3:** Gráfico dos valores dos erros residuais em função dos índices de retenção calculados para compostos alifáticos halogenados em fase estacionária de baixa polaridade, empregando a Equação 9..... 104
- Figura 4:** Correlações lineares entre os índices de retenção experimentais (IR_{Exp}) e os índices semi-empíricos topológicos (I_{ET}) de compostos alifáticos halogenados em fases estacionárias DB-1, DB-35, DB-17, DB-608, DB-210, DB-225, DB-Wax e HP-Basic Wax..... 109
- Figura 5:** Diagrama: (a) da correlação linear simples entre os índices de retenção calculados (IR_{Cal}) em função dos índices de retenção experimentais, e (b) dos valores dos erros residuais em função dos índices de retenção calculados (IR_{Cal}),

para compostos alifáticos halogenados em fases estacionárias de diferentes polaridades, empregando a Equação 12..... 113

Figura 6: Gráfico de: a) Regressão linear simples entre os índices de retenção experimentais (IR_{Exp}) e o I_{ET} ; b) Valores dos erros residuais em função dos índices de retenção calculados (IR_{Cal}) para 48 cicloalcanos em fases estacionárias esqualano e dimetisiloxano..... 122

Figura 7: Diagrama de: (a) Correlação entre os valores de ponto de ebulição normal (PE) e o I_{ET} ; b) Ponto de ebulição calculado com a Equação 14 em função do ponto de ebulição normal experimental..... 125

Figura 8: Calor de formação (ΔH_f^0) como uma função do índice semi-empírico topológico para 1-halo-n-alkanos..... 137

Figura 9: Calor de formação (ΔH_f^0) como uma função do índice semi-empírico topológico modificado para 1-halo-n-alkanos..... 139

Figura 10: Correlação entre a toxicidade e o I_{ET} para 1- bromoalcanos.... 145

Figura 11: Correlação entre a toxicidade e o I_{ET} para compostos halifáticos clorados mono e di-substituídos..... 148

ÍNDICE DE TABELAS

- Tabela 1:** Lista de índices topológicos selecionados..... 48
- Tabela 2:** Lista das características desejadas para os índices topológicos proposta por Randic citada no artigo publicado por Mihalic e Trinajstic..... 51
- Tabela 3:** Valores de C_i e $\log C_i$ para os átomos de carbono primário, secundário, terciário e quaternário de alcanos..... 71
- Tabela 4:** Valores de índice de retenção experimental, IR_{Exp} (Apiezon–L 130°C); índice semi-empírico topológico, I_{ET} ; ponto de fusão, PF (°C); ponto de ebulição, PE (°C), e coeficiente de partição octanol/água, $\log P$, experimental e calculado pelos métodos CNDO e Contribuição dos Fragmentos 177
- Tabela 5:** Valores de índice de retenção experimental, IR_{Exp} (esqualano a 70°C e 120°C e dimetilsilicona a 60°C) e ponto de ebulição normal, PE (°C), para cicloalcanos..... 81
- Tabela 6:** Dados experimentais de atividade biológicas, propriedades físico-químicas e termodinâmicas para os compostos estudados: Toxicidade para o *Tetrahymena pyriformis*, $\log EC_{50}$, (mg/ dm³); Toxicidade ao ciliato *Tetrahymena pyriformis*, $\log IGC^{-1}_{50}$, (mg/L); Densidade, (g/cm³); Índice de refração, nD; Volume de retenção específico a 373K, V_w (cm³/g); Volume molecular, v , (Å³); Coeficiente

de absorção de Ostwald gás-hexadecano, γ_{hd} ; Equilíbrio líquido-vapor, $\rho^{gás}$; Solubilidade aquosa, ρ^w ; Diferença de entalpia padrão molar de solvatação, $\Delta_{sol}H^0$, (J/mol); Diferença de entropia padrão molar de solvatação, $\Delta_{sol}S^0$, (J/mol·K); Calor de formação, (kcal/mol)..... 85

Tabela 7: Valores de C-X, onde X = átomo de Cl, Br ou I, atribuído para os compostos alifáticos halogenados e valores de C_i atribuídos aos átomos de carbono primário, secundário e terciário em posições α , β e γ em relação ao grupos C-X ou X-C-X..... 96

Tabela 8: Verificação da validade estatística do modelo QSRR para compostos alifáticos halogenados..... 101

Tabela 9: Valores dos índices de retenção experimentais (IR) de compostos alceno halogenados medidos em fases estacionárias de diferentes polaridades: DB-1 (IR_{BD-1}), DB-35 (IR_{BD-35}), DB-17 (IR_{BD-37}), DB-608 (IR_{BD-608}), DB-210 (IR_{BD-210}), DB-225 (IR_{BD-225}), DB-Wax (IR_{BD-Wax}) e HP-B/Wax ($IR_{HP-B/Wax}$)..... 106

Tabela 10: Escala de polaridade da fase estacionária, coeficientes de regressão para a equação linear simples ($IR_{Exp} = \beta_0 + \beta_1|ET$), desvios padrões, coeficientes de determinação e de validação cruzada obtidos em cada fase estacionária para compostos alifáticos monohalogenados..... 108

Tabela 11: Valores de C_i para átomos de carbono primários, secundários e terciários das ramificações, átomos de carbono dos anéis com 5 a 12 átomos de carbono e átomos de carbono terciários e quaternários dos anéis com isomeria cis/trans de cicloalcanos calculados para fases estacionárias de baixa polaridade..... 117

Tabela 12: Valores dos índices semi-empírico topológicos (I_{ET}) calculados para cicloalcanos..... 118

Tabela 13: Verificação da validade estatística do modelo QSRR para cicloalcanos..... 121

Tabela 14: Correlação entre as propriedades físico-químicas, termodinâmicas e atividades biológicas dos compostos alifáticos halogenados estudados (coeficiente de correlação e número de compostos)..... 128

Tabela 15: Equações lineares e não lineares (polinomial quadrática) para as propriedades físico-químicas e termodinâmicas estudadas e os respectivos parâmetros estatísticos, empregando o índice semi-empírico topológico (I_{ET}).. 129

Tabela 16: Parâmetros estatísticos para a correlação entre os valores de log P calculado versus experimental pelo Método Semi-Empírico Topológico, Contribuição dos Fragmentos e CNDO..... 141

ÍNDICE DE SÍMBOLOS E ABREVIACÕES

QSPR = Correlações Quantitativas entre Estrutura – Propriedade

QSAR = Correlações Quantitativas entre Estrutura – Atividade

QSRR = Correlações Quantitativas entre Estrutura – Retenção

k_x = constantes de equilíbrio da ionização do ácido benzóico meta e para substituídos

k_H = ácido benzóico não substituído

e = parâmetro eletrônico de Hammett

ρ = constante de proporcionalidade

IB = índices Ivanciuc-Balabam

RLM = regressão linear múltipla

RMN = ressonância magnética nuclear

IV = radiação na região do infravermelho

UV = radiação na região do ultravioleta

HTSA = “high-throughput screening”

PLS = regressão por mínimos quadrados parciais

PCR = regressão por componentes principais

W = índice de Wiener

D = distância entre dois pontos *i* e *j*

C_{ij} = peso entre os vértices *i* e *j*

Z_i^v = número de elétrons de valência do átomo i

h_i = número de átomos de hidrogênio ligados ao átomo i

Z_i = número atômico do átomo i

M = número de ligações no grafo molecular G

u = número característico para um grafo policíclico G , que é igual ao número de ligações que devem ser removidas para transformar este em um gráfico acíclico

$e_i (i = 1, 2, N)$ = representa os elementos da matriz de ordem N : $V[A+D] = [e_i, e_i, e_N]$

V = uma matriz grau do vértice obtida pela soma das colunas ou linhas da matriz adjacência

A = matriz adjacência

D = matriz distância de um grafo molecular G

AM1 = modelo de Austin

PM3 = modelo paramétrico 3

EHT = teoria estendida de Hückel

q^+/q^- = densidade de cargas

μ = momento de dipolo

HOMO = orbital ocupado de maior energia

LUMO = orbital vazio de menor energia

GAP = intervalo entre estes orbitais

PI = potencial de ionização

AE = afinidade eletrônica da molécula

η = dureza e moleza molecular

α' = polarizabilidade

IR = índice de retenção de Kováts

IR_{Exp} = índice de retenção experimental

IR_{Cal} = índice de retenção calculado

t'_{Rz} = tempo de retenção ajustado do composto padrão homólogo com número de carbonos z

t'_{R(z+1)} = tempo de retenção ajustado do composto padrão homólogo com número de carbonos z+1

t'_{Ri} = tempo de retenção ajustado da substância i

CG = cromatografia gasosa

MS = espectrometria de massa

IGC₅₀ = concentração que inibe 50% do crescimento dessa população

EC₅₀ = inibição da bioluminescência de *Photobacterium phosphoreum*, expressa como concentração efetiva

u.i. = unidade de retenção

C_i = fragmentos

C_i = valor atribuído ao fragmento e a cada átomo de carbono e/ou fragmento i na molécula

δ_i = soma do logaritmo dos valores de cada átomo de carbono adjacente (C₁, C₂, C₃ e C₄) e/ou o logaritmo do valor do fragmento.

I_{ET} = Índice Semi-Empírico Topológico

H_{CX} = hidrocarbonetos alifáticos halogenados

I_{Es} = energias de ionização adiabáticas

BTX = benzeno, tolueno e xileno

GCxGC = cromatografia gasosa bidimensional abrangente

IR_{metilsilicona} = índice de retenção medido em fase estacionária metilsilicona

IR_{Apiezon-L 75°C} = índice de retenção medido em fase estacionária Apiezon a 75°C

IR_{Apiezon 125°C} = índice de retenção medido em fase estacionária Apiezon a 125°C

PE (°C) = ponto de ebulição normal

PE_{Cal} (°C) índice de retenção calculado

PF (°C) = ponto de fusão

D (g/cm³) = densidade

n_D = índice de refração

log P = coeficiente de partição octanol-água

v (Å³) = volume molecular

log γ_{hd} = coeficiente de absorção de Ostwald gás-hexadecano

log ρ^{gas} = equilíbrio líquido-vapor

log ρ^w = solubilidade aquosa

log V_w (cm³/g) = volume de retenção específico a 373 K

Δ_{sol}H^o (J/mol) = diferença de entalpia padrão parcial molar de solvatação

Δ_{sol}S^o (J/mol·K) = diferença de entropia padrão parcial molar de solvatação

ΔH_f° (kcal/mol) = calor padrão de formação

$\Delta_{sol}G^\circ$ (J/mol) diferença de energia livre padrão parcial molar de solvatação

C-X e X-C-X = fragmento onde X representa o átomo de cloro, bromo ou iodo

α = posição adjacente alfa ao primeiro grupo halogenado

β = posição adjacente beta ao primeiro grupo halogenado

γ = posição adjacente gama ao primeiro grupo halogenado).

-CH₃= átomos de carbono primário

-CH₂= secundário

>CH= terciário

>C<= quaternário

r^2 = coeficiente de determinação

r = coeficiente de correlação

SD = desvio padrão

r^2_{cv} = coeficiente de validação cruzada

$^1X^v$ = índice de conectividade de Kier e Hall

P_R = polaridade de retenção

β_0 = coeficiente linear

β_1 = coeficiente angular

1 - INTRODUÇÃO

1.1 - Considerações Gerais

A procura de regularidades entre as propriedades físicas, químicas e biológicas de uma substância relacionada com sua estrutura molecular tem sido, durante muito tempo, a meta de vários cientistas. Além de ser uma das idéias mais úteis na organização de um grande número de informações químicas, a elucidação, em detalhes, de como essas propriedades são determinadas pela sua estrutura, oferece uma excelente oportunidade para desenvolver uma nova teoria que possibilite explicar o fenômeno observado e explorar o mecanismo das propriedades físicas/químicas/biológicas dos compostos.¹⁻¹⁰ Exemplos bem estabelecidos do ponto de vista químico serão apresentados a seguir:

As regularidades nas posições relativas dos tempos de retenção de hidrocarbonetos são bem conhecidas há muito tempo e continuam a despertar o interesse de muitos químicos quanto a elucidação dos fatores sub-estruturais envolvidos.⁴

A solubilidade dos alcoóis decresce com o aumento da cadeia carbônica, porque diminui a habilidade dos grupos hidroxila em formar interações do tipo ligações de hidrogênio com a água.⁸

O aumento dos pontos de ebulição de compostos orgânicos depende de suas interações intermoleculares (forças de London) e da diferença da função de partição molecular interna na fase gasosa e na fase líquida na temperatura de ebulição.¹¹

A origem das previsões citadas anteriormente fundamenta-se na estrutura molecular dos compostos orgânicos. Essas e muitas outras propriedades das moléculas podem ser reunidas em um banco de dados, e servir de subsídio para novas previsões, tendo por base a capacidade do químico em estabelecer relações válidas entre a estrutura e a propriedade. Entretanto, muitas irregularidades que surgem também são inerentes a estrutura, e revelam sutilezas interessantes e úteis.

O primeiro modelo sobre estrutura - propriedade foi sugerido por Crum Brown e Fraser em 1868.^{6,7} Cros.⁶, em seu trabalho de Tese apresentado no ano de 1863, demonstrou a primeira relação sobre estrutura – atividade para a toxicidade de etil e amil alcoóis

Martin (apud KALISZAN,1987)¹², em 1949, apresentou os primeiros estudos sobre correlações estrutura molecular – retenção cromatográfica.

Nas áreas de físico – química orgânica, bioquímica e biologia, essas relações são imprescindíveis porque possibilitam transferir relações já generalizadas para associações quantitativas, as quais podem ser expressas na forma de equações algébricas conhecidas como Correlações Quantitativas entre Estrutura – Propriedade (QSPR), Estrutura – Atividade (QSAR) e Estrutura – Retenção (QSRR).¹³

O acelerado desenvolvimento de métodos computacionais, e a possibilidade de aplicar técnicas de tratamento de dados mais complexos, tornaram os estudos da QSPR/QSAR/QSRR mais atraentes, uma vez que possibilitam prognosticar uma propriedade de interesse numa determinada molécula, bem como obter informações sobre o ambiente molecular, eliminando uma dispendiosa e lenta etapa prévia de medidas experimentais, com economia de tempo e custos financeiros.¹⁴

O elevado número de artigos publicados a partir da década de 60 mostra o grande interesse despertado neste tema, o qual continua atraindo a atenção dos químicos nos tempos atuais. Pesquisadores como Randić⁴, Kier e Hall¹⁵, Ferreira¹⁶, Héberger¹⁷, Balaban¹⁸, entre outros que serão citados no desenvolvimento do texto, têm contribuído de forma efetiva nessa área de conhecimento.¹⁵⁻¹⁸

A necessidade crescente de dispor de dados confiáveis para a otimização de processos industriais, a síntese de novos medicamentos com atividades específicas, avaliar a persistência de contaminantes ambientais, entre outros exemplos, justifica plenamente o esforço desses numerosos pesquisadores em desenvolver modelos da QSPR/QSAR/QSRR.¹⁹

Este trabalho está organizado para mostrar o histórico e fundamentos teóricos da correlação quantitativa entre estrutura-propriedade/atividade, os vários tipos de descritores adotados para construir os modelos, a teoria do método semi-empírico topológico, o vasto campo de aplicação dos modelos da QSPR/QSAR/QSRR destacando os compostos alifáticos halogenados e cicloalcanos, pretendendo levar o leitor a fazer uma conexão entre propriedades experimentais e a estrutura molecular.

A metodologia aplicada no Método Semi-Empírico Topológico e o desenvolvimento do descritor molecular denominado índice semi-empírico topológico para compostos alifáticos halogenados e cicloalcanos será descrita detalhadamente, a fim de explicar as considerações observadas para atribuir os valores aos diferentes fragmentos da molécula, bem como o método estatístico adotado para validar os resultados.

O índice semi-empírico topológico será aplicado na construção de simples modelos da QSRR/QSPR/QSAR com algumas propriedades/atividades de

interesse ambiental reunidas a partir da literatura e a discussão abordará a eficácia do modelo matemático obtido e a busca de explicações para o mecanismo responsável pelo comportamento observado em cada parâmetro destacado nesse estudo.

Por fim, serão expostas as conclusões resultantes dessa pesquisa, destacando os fatores estruturais que afetam os valores experimentais e a capacidade preditiva do índice semi-empírico topológico para as propriedades/atividades estudadas.

1. 2- Histórico

Para encontrar relações quantitativas entre estrutura – propriedades/atividades, vários métodos de modelagem matemática vêm sendo aplicados.

Um dos esforços mais antigos, reconhecido para encontrar a regularidade entre estrutura – propriedade resultou no trabalho do padre Jesuíta Croatian, cientista e do filósofo Rugjer Josip Böskovic, os quais introduziram a idéia de representar os átomos como pontos no espaço.⁶ A grande variedade de arranjos, correspondentes às diferentes substâncias, pode ser considerada como o primeiro modelo topológico da estrutura da matéria.

Richet em 1893 (apud GAUDIO,1996)⁷, ao propor que a toxicidade de alguns alcoóis, éteres e cetonas é inversamente proporcional à sua solubilidade em água, indicou a primeira relação entre toxicidade e lipofilicidade.

No início do século XX, Overton e Meyer (apud GAUDIO,1996)⁷, independentemente, estudaram o efeito narcótico de vários compostos não ionizados e verificaram sua correlação com o coeficiente de partição óleo/água.

Em 1939, Ferguson (apud GAUDIO,1996)⁷ estudou o comportamento de diversas propriedades físicas e químicas em relação à atividade tóxica de uma série de compostos (homólogos e correlatos), fornecendo uma interpretação termodinâmica da atividade tóxica, sugerindo que esta pode ser determinada através da medida do seu potencial químico. Atualmente, estudos de QSAR adotam descritores químico-quânticos calculados usando programas computacionais.²⁰

O trabalho de Hammett, em 1940 (apud GAUDIO,1996)⁷, utilizou pela primeira vez uma reação química para descrever uma propriedade estrutural. Ao propor a Equação 1 para descrever a ionização do ácido benzóico em água, verificou que as constantes de equilíbrio da ionização do ácido benzóico meta e para substituídos (k_X) e do ácido benzóico não substituído (k_H), estavam relacionados pelas constantes σ , parâmetro eletrônico de Hammett e ρ , constante de proporcionalidade:

$$\log k_X = \log k_H + \rho\sigma \quad (1)$$

A constante σ representa o efeito eletrônico ou polar do substituinte Y sobre o processo de ionização. A constante ρ é uma constante que indica a susceptibilidade da reação ao efeito eletrônico exercido pelo substituinte Y. Hammett adotou a reação de ionização dos derivados do ácido benzóico como padrão, a qual atribuiu o valor de $\rho = 1,00$. Isso implicou no valor de $\sigma_{\text{não substituído}} =$

0,00. A partir dessas considerações, foi possível a determinação do valor de σ para inúmeros substituintes. A possibilidade de prognosticar a constante σ permitiu estimar a reatividade de um composto que ainda não foi determinada experimentalmente.

A equação de Hansch²¹ marcou um importante estágio na área de QSAR, sendo que o modelo clássico mais utilizado e difundido em estudos de QSAR é o de Hansch – Fujita. Esse método correlaciona a atividade biológica de um composto com propriedades hidrofóbicas, eletrônicas e estéricas que representam a estrutura química da molécula.^{22,23} O modelo produzido tenta caracterizar e isolar as contribuições de cada um dos fatores físico-químicos indicados, responsáveis pela atividade biológica ou pela atuação farmacológica no caso dos fármacos, fornecendo assim uma explicação para o efeito de cada parâmetro no modelo. Existem vários exemplos na literatura de aplicações do modelo de Hansch-Fujita apresentados em Kubinyi.²⁴

Segundo Katritzky *et al.*²⁵, os modelos de QSAR já se tornaram uma ferramenta essencial em química médica. As maiores companhias farmacêuticas têm investido substancialmente na elucidação do efeito da estrutura sobre uma determinada propriedade biológica, principalmente de compostos medicinais ativos.

2 - OBJETIVOS

O objetivo geral deste trabalho é ampliar o Índice Semi-Empírico Topológico (I_{ET}) para prever a retenção cromatográfica (QSRR) para compostos alifáticos halogenados e cicloalcanos e aplicar o I_{ET} para estabelecer úteis e significativas correlações quantitativas entre a estrutura molecular-propriedade/atividade (QSPR/QSAR).

Como objetivo específico pretende-se:

- Desenvolver o I_{ET} para prever a retenção cromatográfica em fase estacionária de baixa polaridade de compostos alifáticos halogenados e cicloalcanos;
- Avaliar o potencial de prognóstico do I_{ET} para a retenção cromatográfica de compostos alifáticos halogenados em fases estacionárias de diferentes polaridades;
- Retirar informações estruturais das moléculas estudadas, bem como sobre o mecanismo de retenção dessas moléculas;
- Aplicar o I_{ET} para prever propriedades físico-químicas, termodinâmicas e atividades biológicas relevantes de compostos alifáticos halogenados e cicloalcanos;

- Verificar a habilidade de predição dos modelos de QSRR/QSPR/QSAR obtidos através de parâmetros estatísticos, de testes de validação externos e internos e do procedimento de validação cruzada (r^2_{cv}).

3 - FUNDAMENTOS TEÓRICO

3. 1 - Correlações Quantitativas entre Estrutura - Propriedade/Atividade

A hipótese fundamental em estudo de QSPR/QSAR é que qualquer propriedade de uma molécula está relacionada com sua estrutura química, quer nos aspectos estéreo, eletrônicos ou topológicos.

As relações estrutura e propriedade quantificam a conectividade entre a estrutura e as propriedades de uma série de compostos. Essas relações são modelos matemáticos simples que podem ser construídos usando métodos estatísticos clássicos e técnicas de reconhecimento de padrões, isto é, métodos estatísticos multivariados e redes neurais artificiais. Alguns aspectos teóricos e práticos devem ser observados para a proposição, validação e análise dos modelos matemáticos que correlacionam estrutura química e atividade biológica para a interpretação apropriada dos modelos criados.^{26,27}

Quando é estabelecida a correlação entre a estrutura e propriedade/atividade, os modelos de QSPR/QSAR geralmente oferecem a melhor resposta para entender de que maneira as estruturas influenciam uma propriedade em particular e permitem selecionar a estrutura de novos compostos com propriedades pré-determinadas.¹¹

Em pesquisas de QSPR, a estrutura química de um composto deve ser transformada em uma série de descritores numéricos que representam as características mais relevantes para uma dada propriedade/atividade e, ao mesmo tempo, correlacionar bem com valores medidos experimentalmente.²⁸

A representação numérica da estrutura que descreve a molécula como um todo ou uma parte da mesma é denominado descritor molecular. Essa representação é útil uma vez que as propriedades moleculares também são registradas com números simples. Portanto, QSPR/QSAR é um modelo de calibração entre dois conjuntos numéricos que estão relacionados por uma expressão algébrica. Nessa expressão, as variáveis independentes são os descritores moleculares que descrevem a estrutura da molécula sob estudo, e a variável dependente é a propriedade/atividade de interesse como, por exemplo, a retenção cromatográfica, ponto de ebulição ou toxicidade (atividade biológica).⁶

A obtenção de correlações significativas depende da habilidade do descritor utilizado para quantificar exatamente as características moleculares microscópicas relacionadas as propriedades macroscópicas.²⁹ Geralmente, os descritores empregados nesses estudos são obtidos a partir de considerações teóricas, empíricas ou derivados de características experimentais disponíveis na literatura.³⁰ Katrizky e Fara²² em seu artigo apresentam uma classificação atualizada dos descritores moleculares, bem como o estado da arte dos modelos de QSPR/QSAR devido ao desenvolvimento de novos programas computacionais mais práticos e executáveis. Os índices topológicos, derivados da teoria dos grafos de compostos orgânicos, descrevem a conectividade entre os átomos de um composto e estão entre os descritores moleculares mais difundidos na literatura.³¹⁻³⁵

Com a crescente necessidade de dados disponíveis para a otimização de processos nas indústrias químicas, farmacêuticas e de alimentos, a produção de modelos práticos e significativos de QSPR para estimar diferentes propriedades de compostos, bem como obter explicações sobre o fenômeno físico-químico envolvido nessas propriedades continuam em pleno desenvolvimento.³⁵

Uma vez estabelecida uma correlação entre estrutura molecular e propriedade/atividade, compostos com propriedades de interesse, incluindo aqueles ainda não sintetizados, podem ser prognosticados rapidamente. É então possível selecionar compostos mais promissores para sintetizar e testar em laboratórios, com redução de custos e economia de tempo.²⁰ Analogamente, o desenvolvimento de modelos de QSAR pode ser usado para a predição em uma grande variedade de propriedades biológicas com aplicação em ciências biomédicas, toxicológicas e ambientais.

A possibilidade de acelerar o processo de desenvolvimento de novas moléculas para os mais diversos usos tais como fármacos, aditivos e polímeros é especialmente atraente na área Química, uma vez que os profissionais estão cada vez mais preocupados com a manutenção e o aumento da qualidade ambiental.

Nesse sentido, espera-se que as investigações em QSPR/QSAR continuem a aumentar, em função do seu uso como uma ferramenta em estudos acadêmicos e processos industriais.

3. 2- Modelos de Correlação Quantitativa entre Estrutura – Propriedade/Atividade

O objetivo fundamental dos estudos de QSPR/QSAR é prognosticar propriedades físicas, químicas e biológicas utilizando descritores moleculares simples, e também obter alguma informação sobre o comportamento molecular.³⁵ Para esse fim, vêm sendo desenvolvidos numerosos descritores moleculares, calculados somente a partir da estrutura ou por meio de simulação computacional. Independentemente da forma como foi calculado, cada descritor molecular associa

um número simples a uma substância química, e ordena um grupo de compostos de acordo com um valor numérico de uma propriedade específica. Se a variável independente (descritor molecular experimental ou calculado teoricamente) está relacionada matematicamente (linearmente ou não linearmente) com a magnitude da propriedade em estudo (variável dependente), de acordo com critérios estatísticos previamente estabelecidos, o modelo de QSAR/QSPR é válido.³⁶

Tradicionalmente as relações quantitativas estrutura-atividade eram desenvolvidas utilizando a metodologia de Hansch e Leo, onde os compostos eram caracterizados por suas propriedades físico – químicas.³⁷ O paradigma foi quebrado quando Kier e Hall³⁸ sugeriram o uso de descritores para caracterizar topologicamente as moléculas. A construção do índice de conectividade molecular de primeira ordem, ${}^1\chi$, foi correlacionada linearmente com o ponto de ebulição de alcanos de cadeia curta. Na ocasião, 1975, foi assumido que a relação entre as propriedades investigadas e o índice de conectividade era linear.

A expansão dos índices de conectividade para ordens maiores, ${}^n\chi$, no mesmo ano, permitiram aplicar a análise de regressão linear múltipla para a caracterização topológica dos compostos sob estudo em QSAR/QSPR. Os autores mostraram que a introdução de um segundo índice de conectividade na equação, relacionada com a densidade, melhora significativamente a correlação ($r=0,903$ para $r=0,980$). Uma vez que os índices topológicos geralmente resultam em boas correlações lineares, vários autores continuam usando correlações lineares simples.^{4,39,40}

Existem diferentes metodologias para o desenvolvimento de QSPRs. Neste estudo será adotada a citada por Mihalic e Trinajstić.⁶ As etapas envolvidas na construção do modelo de QSPR estão descritas a seguir:

Etapa 1 – Obtenção dos dados

Primeiramente foram obtidos os dados de retenção cromatográfica experimentais (índice de retenção) da literatura de compostos alifáticos halogenados e cicloalcanos que apresentam estruturas e modo de ações similares, bem como de propriedades físicas, químicas, termoquímicas e atividades biológicas de interesse para os grupos estudados.³⁶⁻³⁸

Este grupo de moléculas também pode ser denominado de grupo de treinamento. Nesta etapa inicial é muito importante obter dados confiáveis e precisos, pois da sua qualidade dependerão todas as etapas subseqüentes;

Etapa 2 – Caracterização da estrutura molecular

Nesta etapa foi desenvolvido o descritor molecular, denominado índice semi-empírico topológico, para os compostos alifáticos halogenados;

Etapa 3 - Seleção de um grupo de treinamento representativo

Dois grupos de números, os índices de retenção (coeficiente de partição, ponto de ebulição, etc.) e o índice semi-empírico topológico foram correlacionados através de métodos estatísticos resultando numa expressão algébrica aceitável. O modelo obtido foi avaliado estatisticamente bem como sua qualidade;

Etapa 4 – Habilidade de previsão do modelo

Predições para os valores das propriedades moleculares de interesse das espécies que não fazem parte do grupo inicial tomado no modelo QSPR foram realizadas. Outros testes estatísticos podem ser utilizados que também estimam a

habilidade de predição da equação (modelo) obtida como o procedimento de validação cruzada, que segue o esquema “leave-n-out”;

Etapa 5 – Validação do modelo

Através de determinação experimental também é possível confirmar os prognósticos realizados pela QSPR das moléculas desconhecidas. Caso os testes realizados nesta etapa não apresentem resultados satisfatórios, o modelo deve ser revisado e o procedimento repetido (etapa 3) até a obtenção de um modelo final adequado;

Etapa 6 – Aplicação do modelo

Se os testes realizados confirmam os prognósticos, o modelo QSPR é aceito na sua forma final com seus dados estatísticos.

Mihalic e Trinajstić⁶ propuseram que a qualidade de um modelo de QSPR pode ser convenientemente medida pelo coeficiente de correlação (r) e pelo desvio padrão (SD). Eles propuseram que um bom modelo de QSPR deve ter $r > 0,99$, enquanto SD depende da propriedade estudada.

Vários autores têm seguido esta orientação: Estrada⁴¹ considerou que o seu índice topológico relacionado com a ligação adjacente, análogo ao índice de ramificação de Randić, descreveu bem o volume molar de composto alifáticos orgânicos contendo heteroátomos em suas estruturas ($r > 0,99$); Duchowicz e Castro⁴² também consideraram bons os resultados obtidos por regressão linear múltipla para pontos de ebulição de alquil álcool.

Ivanciuc⁴³ comparou os índices estatísticos de equações de regressão mono e bi-paramétricas obtidos para modelos de QSPR usando seis propriedades dos alcanos. Todas as correlações bi-paramétricas contêm o número de átomos de carbono (n) e índices Ivanciuc-Balabam (IB). Os resultados indicaram que o uso de correlações bi-paramétricas levou aos melhores resultados.

As melhores correlações obtidas com modelos de regressão linear múltipla de n-parâmetros, e a possibilidade de obter informações sobre partes específicas das moléculas que estão influenciando a propriedade/atividade estudada, têm engajado muitos investigadores.^{8,13,44-47} Também foram conduzidas pesquisas que comparam o potencial de vários tipos de descritores moleculares relacionados com diferentes propriedades físico-químicas.¹

Atualmente um grande número de novos índices topológicos está disponível na literatura, incluindo programas que efetuam os cálculos diretamente. O programa ADAPT de Jurs⁴⁸ e o programa SPARC desenvolvido pela Agência de Proteção Ambiental dos Estados Unidos foram os primeiros a contribuir para o grande avanço nas áreas de QSAR/QSPR.

O programa CODESSA combina uma grande variedade de descritores clássicos não empíricos (descriptor derivado diretamente da estrutura dos compostos), junto com novos descritores químico-quânticos e combinados. Esse programa também realiza o tratamento estatístico dos dados adotando técnicas padrão e avançadas, tais como regressão múltipla e não linear, análise de fatores, e métodos heurísticos para o desenvolvimento de QSPR.²⁴

Segundo Todeschini e Consonni⁴⁸, o grande interesse da comunidade científica nos descritores moleculares pode ser avaliado pelo grande número de descritores propostos atualmente. São mais de 2000 descritores definidos e facilmente calculados através de programas computacionais. Com tantos dados

disponíveis, autores como Pompe²⁶ utilizaram modelos de regressão linear múltipla (RLM), retropropagação de erro e redes neurais para selecionar, entre 68 descritores calculados com o programa CODESSA, 16 descritores para descrever a estrutura molecular de 381 compostos e construir o modelo de predição de dados de retenção.

Katritzki^{11,29,35} tem publicado vários trabalhos para demonstrar o potencial do programa CODESSA, e também sugere a transição dos tratamentos familiares de uma dimensão da QSAR/QSPR, onde é estudada a variação de uma simples propriedade com a estrutura, para um tratamento multidimensional. Isso implica no estudo simultâneo de muitos descritores e o tratamento combinado entre a QSPR e PCA para a seleção efetiva das variáveis. Desta forma, um grande número de parâmetros que descrevem um composto pode ser medido ou computado, sem exigir um conhecimento prévio sobre a função do parâmetro na propriedade. Esses parâmetros são posteriormente selecionados usando PCA, que implica na linearidade do modelo, ou algoritmos genéricos, que também podem levar para cálculos não lineares.

Em 1990, métodos de regressão baseados em redes neurais ofereceram uma nova possibilidade para otimizar os modelos de QSPR, permitindo a exploração de relações não lineares para a modelagem de novas relações estrutura-propriedade. Como exemplo pode-se citar Wessel *et al.*⁴⁸, que adotam redes neurais para otimizar o prognóstico de pontos de ebulição normal de hidrocarbonetos baseados em 6 descritores.

O progresso substancial de métodos de derivação de relações lineares e não-lineares, e novos e mais exatos descritores moleculares, possibilitaram a realização de muitas investigações de QSPR/QSAR para prognosticar as mais

diversas propriedades/atividades de compostos químicos e materiais de interesse industrial.

Robustos métodos computacionais permitiram a aplicação de sofisticados métodos quimiométricos em QSAR/QSPR, preparando este campo de investigação para enfrentar novos desafios em um futuro próximo. A futura direção na metodologia aponta para a necessidade de prognosticar propriedades de sistemas multicomponentes complexos e massas moleculares mais elevadas. Descritores moleculares específicos estão sendo desenvolvidos para este fim, envolvendo química quântica e mecânica estatística.²⁹ Os recentes avanços nessa área fornecem os pré-requisitos necessários e estimulam o desenvolvimento de novas metodologias.¹¹

3. 3- Descritores Moleculares

Está bem fundamentado que muitas propriedades físico-químicas ou funções termodinâmicas de um composto dependem da sua estrutura química. O termo estrutura inclui os aspectos topológicos, eletrônicos e geométricos. Uma vez que muitas propriedades moleculares são medidas e os dados são registrados como valores numéricos, para a construção de modelos de QSPR/QSAR, a estrutura molecular também deve ser apresentada da mesma maneira. Portanto, são necessários dois conjuntos numéricos: um que representa a propriedade/atividade e o outro que representa a estrutura molecular. Desta forma, pode-se estabelecer uma relação quantitativa entre os dois conjuntos de números através de uma equação algébrica.⁶

Até a década de 60, a maioria das determinações químicas era realizada através de técnicas simples, como análise gravimétrica e titulométrica. A necessidade de detectar compostos químicos em concentrações cada vez menores, e o desenvolvimento de suportes eletrônicos para automatização instrumental, possibilitaram o surgimento de técnicas instrumentais de análise realizadas através da utilização de sofisticados equipamentos, como os métodos espectrais (ressonância magnética nuclear - RMN, infravermelho - IV, ultravioleta - UV) e os métodos cromatográficos.

A aplicação dessas técnicas em muitos laboratórios tem permitido a determinação sistemática de várias propriedades físico-químicas e funções termodinâmicas de um grande número de compostos orgânicos. Uma empresa farmacêutica média que realize "high-throughput screening" (HTSA) pode produzir resultados de dezenas de milhares de experiências por dia. O "Chemical Abstracts" adiciona atualmente cerca de 900.000 novos compostos por ano à sua base de dados. A grande diversidade de informações já catalogadas permite a construção de Bases de Dados contendo informações corretas, precisas e de fácil acesso. A consulta a esses Bancos de Dados por químicos e cientistas interessados em obter uma propriedade/atividade em particular é rápida e fácil.⁵⁰

Para a descrição da estrutura química na forma de um número ou um conjunto de números foram desenvolvidos muitos termos matemáticos denominados de descritores moleculares. Muitos descritores refletem propriedades moleculares simples, e assim eles podem fornecer algum entendimento da característica estrutural intrínseca relacionada com a natureza físico-química da propriedade/atividade sob consideração.⁴⁷

Os descritores moleculares podem ser derivados de maneira empírica ou não empírica. Os descritores empíricos correspondem às propriedades físicas, e

os descritores não empíricos abrangem os índices topológicos e químico-quânticos (eletrônicos e geométricos).³⁶ Descritores derivados somente a partir da conectividade e composição da estrutura são indicados como índices topológicos. Descritores derivados da geometria molecular tri-dimensional são chamados de geométricos e aqueles calculados a partir da distribuição de carga parcial são denominados eletrônicos. Os descritores combinados levam em consideração simultaneamente a estrutura eletrônica e a conectividade dos átomos na molécula, ou a geometria molecular e a estrutura eletrônica.^{1,5,30,36}

Anker e Jurs⁵¹ consideram como descritores físicos aqueles obtidos diretamente a partir de medidas experimentais. Entretanto, a estrutura molecular de um composto pode ser descrita por uma série de parâmetros físico-químicos (hidrofóbicos, estéreos ou eletrônicos), obtidos a partir de medidas experimentais ou calculados por métodos químico-quânticos, como por exemplo, o coeficiente de partição octanol-água.

O maior desafio está em selecionar o descritor molecular apropriado para ser empregado no modelo de QSPR/QSAR a fim de obter uma correlação significativa sob o ponto de vista estatístico, e de fácil interpretação estrutural.¹³

Para obter excelentes modelos de QSPR/QSAR, freqüentemente é necessário escolher um descritor dentre um conjunto de muitos descritores estruturais usando vários métodos de otimização, incluindo métodos de análise multivariada clássicos, como regressão univariada e multivariada, regressão por mínimos quadrados parciais (PLS), regressão por componentes principais (PCR), redes neurais, algoritmos genéricos, entre outros.^{19,47,52-54}

3. 3. 1 - Grafos Moleculares

Muitos modelos de QSPR/QSAR foram desenvolvidos usando vários descritores numéricos da estrutura química. Em química teórica, a estrutura química de uma molécula pode ser representada como um grafo molecular. Grafo molecular são grafos químicos que representam a constituição das moléculas usando diferentes convenções. Na representação das moléculas num grafo molecular, a ligação química dos átomos é considerada como sua característica mais importante. Portanto, a teoria dos grafos em química é um conceito associado mais à topologia molecular do que à geometria, e trata da análise de todas as conseqüências da conectividade em um sistema químico.^{6,9}

Características geométricas como comprimentos de ligação, ângulos de ligação ou configurações tridimensionais, não são levadas em consideração. No grafo molecular os vértices, ou pontos, correspondem aos átomos, e as linhas, ou arestas, representam as ligações covalentes entre os átomos.⁴³ Os átomos de hidrogênio e suas ligações são omitidos representando-se apenas o esqueleto molecular.^{6,55} A primeira etapa para a determinação de um índice consiste em transformar a estrutura química de uma molécula em um grafo molecular.

A Figura 1 mostra a fórmula estrutural da molécula do cicloexano e o correspondente diagrama do grafo molecular no qual foi atribuído um número específico para cada átomo de carbono.

Uma vez que o grafo é definido, os vértices são aleatoriamente numerados e a “matriz topológica” é construída. Dois grafos G_1 e G_2 são isomórficos se existe uma correspondência direta entre seus vértices e suas linhas. Um grafo G invariante detém uma quantidade associada com G que tem o mesmo valor para qualquer grafo que é isomórfico com G . Dois vértices i e j de um grafo G são

adjacentes se existe uma linha de união entre eles. Duas linhas de G são adjacentes se elas têm um ponto em comum. A valência ou grau de um vértice de G é um número de linhas ligadas a ele. Um caminho é uma seqüência alternada de pontos e linhas que inicia com um ponto e termina com outro ponto. Um percurso é um caminho onde nenhum ponto ocorre mais do que uma vez. A distância entre dois pontos é o número de linhas no percurso mais curto ligando os dois pontos.

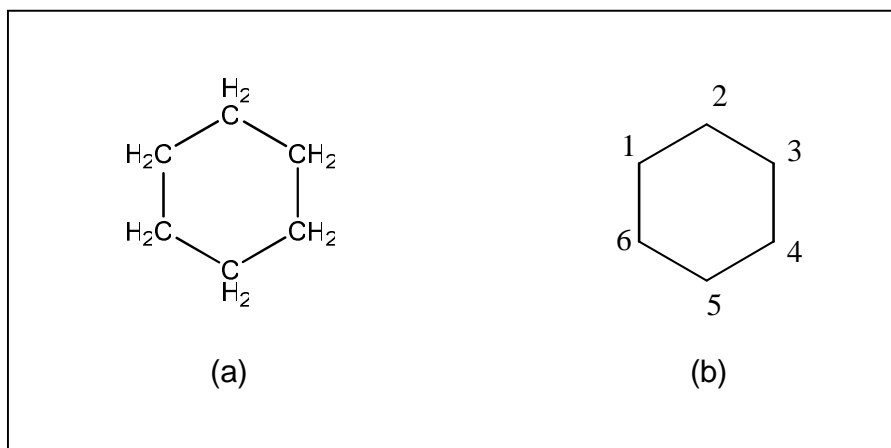


Figura 1. A molécula do cicloexano representada como fórmula estrutural (a) e grafo molecular enumerado (b) onde foram omitidos os átomos de hidrogênio. Os vértices correspondem aos átomos e as linhas as ligações químicas.

Adotando um grafo molecular, a estrutura química de um composto pode ser expressa por meio de diferentes matrizes, polinômios, espectros, momentos espectral, seqüência da contagem das distâncias, caminhos e percursos, que

originam diferentes índices topológicos. O principal objetivo dessas considerações consiste em encontrar uma maneira para representar as estruturas como número simples, ou seja, transformar a informação contida no grafo molecular em atributos numéricos.

A teoria dos grafos pode ser aplicada em química nas mais diversas áreas, como síntese, polímeros, termoquímica, cinética química, equilíbrio de fases, análise espectral, entre outras.⁹

3. 3. 2 - Descritores Topológicos

Os índices topológicos são descritores numéricos que caracterizam a estrutura química e assim possibilitam estabelecer relações quantitativas entre a estrutura e propriedades físicas, químicas e biológicas.

A análise topológica de uma molécula inicia com a sua representação em um grafo molecular que reflete a conexão entre seus átomos.²⁸ A série de átomos e conexões no grafo molecular contém informação estrutural que deve ser transformada em um índice numérico chamado de índice topológico, que posteriormente poderá ser correlacionado com propriedades experimentais de uma Base de Dados.

Um índice topológico é um descritor numérico da estrutura química baseado em certas características topológicas do grafo molecular e caracteriza o tamanho e a forma molecular.⁴³

Quando um número simples, tal como um índice topológico é utilizado para a representação da estrutura molecular, é esperada uma considerável perda de informação, porque a estrutura tridimensional da molécula é descrita por um índice

topológico unidimensional. No entanto, índices derivados do conceito de distância topológica, que representam simplesmente o número de ligações (índice de Wiener) ou número de pares de linhas adjacentes (índice de Platt) em um grafo molecular são exemplos de descritores estruturais simples, mas que são capazes de fazer correlações satisfatórias com uma série de propriedades, indicando que muitas informações estruturais relevantes ficam retidas no índice topológico.^{5,28}

Entretanto, os índices topológicos simples são aqueles que não diferenciam o tipo de átomo no grafo molecular, o que limita o seu campo de aplicação em função da falta de informação sobre o heteroátomo e características estereoquímicas das moléculas.⁴² Atualmente, vários índices topológicos têm incluído a diferenciação de heteroátomos adotando considerações empíricas e não empíricas.^{42,56}

Apesar de existirem outras classes de descritores estruturais, tais como geométricos, eletrônicos ou descritores quânticos, os índices topológicos são geralmente utilizados para a codificação, ordenamento e pesquisa de estruturas químicas em Bases de Dados e para solucionar problemas de grafos isomórficos.³²

O excelente desempenho desses índices, quando comparado às outras classes de descritores estruturais, pode ser atribuído: (i) a facilidade em calcular esses índices a partir do grafo molecular; (ii) as excelentes correlações obtidas com várias propriedades/atividades; (iii) a fácil interpretação dos resultados, porque freqüentemente fornecem uma maneira simples de medir a ramificação molecular, forma, tamanho, semelhanças e diferenças moleculares.^{27,55}

Apesar de ser possível executar cálculos quânticos semi-empíricos para moléculas com mais do que uma centena de átomos, as similaridades e

diversidades dos compostos químicos em Bases de Dados são caracterizadas quase que exclusivamente por descritores derivados do grafo molecular.

Em 1995, Estrada⁴¹ já computava mais do que 120 índices topológicos definidos na literatura, mas somente alguns têm sido extensivamente adotados em investigações QSPR/QSAR. De acordo com Katritzky e Gordeeva¹ e Icardo *et al.*³³ os índices mais proeminentes foram desenvolvidos por Wiener, Gordon e Scantlebury, Hosoya, Balaban, Randic e Bonchev. Contudo, estudos de Kier e Hall⁵⁶ têm revelado a efetividade dos índices de conectividade molecular, índices topológicos geométricos, de carga e eletrotológico em métodos de QSPR/QSAR.

Com o intuito de mostrar a idéia original desses autores em uma seqüência cronológica de publicações, a Tabela 1 apresenta vários índices topológicos bem conhecidos que serviram de base para correlacionar diferentes propriedades/atividades com vários conjuntos de compostos químicos. Icardo *et al.*³³ apresentam uma sintética revisão sobre os fundamentos da topologia molecular e definição dos índices topológicos.

De fato, qualquer matriz bem definida do grafo associado com uma molécula pode gerar um novo índice topológico.^{57,58} Para ilustrar a diversidade de invariantes matemáticos existentes para a caracterização das moléculas, pode-se consultar Randic³⁵, que mostra três maneiras relativamente simples e elegantes de cálculo do índice de Wiener (W), onde cada tipo pode ser usado para fazer uma definição matemática equivalente de W . No entanto, o problema da interpretação dos índices topológicos ainda persiste, apesar de vários autores estarem empenhados para encontrar explicações para o significado físico-químico desses índices.^{32,40,59} Esse problema não deve ser confundido com o fato de que a

construção de muitos índices topológicos está bem fundamentada em definições matemáticas relativamente sofisticadas.

Segundo Randić⁶⁰, os melhores descritores são aqueles que apresentam simplicidade estrutural, porque ajudam na interpretação e explicação estrutural relacionado a propriedade. A Tabela 2 enumera as características desejadas para os índices topológicos propostas por Randić.

A escolha e a concepção dos descritores moleculares irão definir o sucesso da investigação estrutura-propriedade/atividade. Para a seleção, freqüentemente são escolhidos aqueles que resultam na melhor regressão obtida de acordo com os critérios estatísticos adotados.

Um procedimento de notável importância que permite extrair e eliminar um descritor que duplica informação já contida em outro descritor previamente selecionado é conhecido como ortogonalização. O índice de conectividade molecular ortogonal é derivado a partir do correspondente índice de conectividade normal. Esse processo permite a construção de equações de regressão mais estáveis (isto é, os coeficientes dos descritores individuais são insensíveis à inclusão ou exclusão de outros descritores na regressão), a derivação do descritor dominante e ao aumento da confiabilidade da relação. A descrição do procedimento de ortogonalização pode ser consultada em Randić.⁶⁰

Tabela 1. Lista de índices topológicos selecionados

Autor	Ano	Símbolo	Interpretação Estrutural
Wiener	1947	W	<p>Primeiro índice estrutural de natureza topológica.⁴⁰ Fundamenta-se no conceito topológico de distância, a qual é definida como a soma das distâncias topológicas entre todos os pares possíveis de vértices do grafo químico. Entendendo-se por distância topológica o número de segmentos entre dois vértices quaisquer.²⁸ O número de Wiener é calculado com a metade da soma das menores distâncias entre dois pontos <i>ij</i> da matriz distância.</p> $W = 1/2 \sum_{ij} (D_{ij})$ <p>onde, D = distância entre dois pontos <i>i</i> e <i>j</i> <i>i, j</i> = dois pontos <i>i</i> e <i>j</i></p> <p>Esse índice, particularmente, é alto para as moléculas estendidas e baixo para as compactas.⁶</p>
Platt	1947-1952	P	<p>Índice que representa o número de pares de linhas adjacentes.³⁸</p> <p>Este índice associa a topologia com a estabilidade química das moléculas.⁴⁰</p>
Hosoya	1971	Z	<p>Hosoya introduziu o termo topológico para os descritores moleculares baseados no grafo molecular. Índice definido com a soma de todas as ligações não adjacentes em um grafo molecular.³⁸</p>
Randic	1975	χ	<p>A partir da publicação de Randic ocorreu um aumento substancial no número de índices divulgados na literatura. Randic introduziu o conceito do grau de vértice, δ, definido como o número de vértices vizinhos a um vértice específico. Por exemplo, para o grupo metila que possui apenas um vizinho, $\delta = 1$; o grupo metileno tem dois vizinhos, $\delta = 2$. Randic também propôs um peso específico para a linha entre os vértices <i>i</i> e <i>j</i>, definindo-o por $(\delta_i \delta_j)^{-1/2}$, que corresponde a um número que codifica as ramificações de esqueletos de alcanos. O índice de ramificação é definido como a soma dos pesos das linhas, para todas as linhas em um grafo molecular.²⁸ Expressando matematicamente, cada linha entre os vértices <i>i</i> e <i>j</i> é caracterizado por um número definido como:</p> $C_{ij} = (\delta_i \delta_j)^{-1/2}$ <p>onde, C_{ij} = peso entre os vértices <i>i</i> e <i>j</i></p> <p>O índice de ramificação ou de conectividade molecular é obtido fazendo-se a soma de todos os C_{ij}³⁸</p> $\chi = \sum_{ij} C_{ij} \sum (\delta_i \delta_j)^{-1/2}$

Tabela 1. Continuação

Autor	Ano	Símbolo	Interpretação Estrutural
Kier, Murray Radic Hall	1976	${}^m\chi$	<p>Re-nomearam o índice de ramificação para índice de conectividade molecular de primeira ordem (${}^1\chi$). Generalizaram o índice de Randić³⁶ com o objetivo de alcançar melhor precisão para descrever as propriedades físico-químicas selecionadas.</p> <p>O índice ${}^m\chi$ é derivado da decomposição do grafo molecular em subgrafos de ligações, com distâncias de comprimento 1, 2, ..., m, que representam uma característica distinta da estrutura molecular. Definiram o índice ${}^m\chi$ em várias ordens, em função do número de ligações consideradas em cada caso:³⁸</p> <p>átomo isolado</p> ${}^0\chi = \sum_{ij} (\delta_i)^{-1/2}$ <p>uma ligação, corresponde ao índice de conectividade molecular</p> ${}^1\chi = \sum_{ij} (\delta_i, \delta_j)^{-1/2}$ <p>dois fragmentos de ligação, corresponde ao índice de conectividade molecular de segunda ordem</p> ${}^2\chi = \sum_{ij} (\delta_i, \delta_j, \delta_k)^{-1/2}$ <p>Para ordens maiores do que 2 foram definidos novos índices de conectividade que levam em consideração os tipos de subestruturas consideradas. Desta forma, pode-se ter seqüências do tipo linha, ramo, linha/ramo ou anel, dependendo da molécula apresentar ramificações ou ciclos.²⁸</p>
Kier Hall	1976	${}^m\chi^v$	<p>Índice de conectividade de valência. Os índices de conectividade, como definidos acima, não conseguem distinguir entre diferentes átomos nem entre diferentes hibridizações. O índice ${}^m\chi^v$ foi modificado para melhor discriminar compostos que possuem ligações múltiplas, heteroátomos e estados de valência. Kier e Hall sugeriram a definição do grau do vértice (d) não mais em função do número de ligações vizinhas ao vértice, mas em função da valência do átomo representado pelo vértice. O grau do vértice foi definido pela fórmula:^{3,36}</p> $\delta_i^v = \frac{Z_i^v - h_i}{Z_i - Z_i^v - 1}$ <p>onde, Z_i^v = número de elétrons de valência do átomo i h_i = número de átomos de hidrogênio ligados ao átomo i Z_i = número atômico do átomo i</p>

Tabela 1. Continuação

Autor	Ano	Símbolo	Interpretação Estrutural
Balaban	1982	J	<p>Índice que apresenta a menor degenerescência, isto é, um menor número de grafos moleculares apresentam o mesmo valor de índice topológico. Mistura os conceitos de topologia e conectividade. O índice é definido como a soma das distâncias médias no grafo molecular. O par (ij) tem peso $1/\sqrt{D_{ij}}$</p> $J = \frac{M}{u+1} \sum_{ij} (D_{ij})^{-1/2}$ <p>onde, M = número de ligações no grafo molecular G u = número característico para um grafo policíclico G, que é igual ao número de ligações que devem ser removidas para transformar este em um gráfico acíclico. $u = 0$ grafo acíclico e $u = 1$ grafo monocíclico³⁷</p>
Kier	1985	^m K	<p>Índice da forma molecular - Índice Kappa. Desenvolvido com o objetivo de quantificar a forma molecular. Os valores Kappa são derivados da contagem de um, dois e três fragmentos de ligação.²⁸</p> <p>¹K índice de ordem 1, informa sobre a ciclicidade das estruturas ²K de ordem 2, informa sobre a densidade espacial ³K de ordem 3, informa sobre ramificações centrais</p>
Schultz	1989	MTI	<p>Índice que reflete a topologia molecular e o estado de valência dos átomos definido como:²⁸</p> $MTI = \sum_{i=1}^N e_i$ <p>onde, $e_i (i = 1, 2, N)$ representa os elementos da matriz de ordem N: $V[A+D]$ $= [e_i, e_i, e_N]$</p> <p>onde, V é uma matriz grau do vértice obtida pela soma das colunas ou linhas da matriz adjacência, A é matriz adjacência e D é a matriz distância de um grafo molecular G.²⁸</p>
Kier Hall	1990		<p>Índice do estado eletrotológico. São valores numéricos calculados para cada átomo da molécula que contêm informações sobre o ambiente topológico dos átomos e sobre a interação eletrônica devido a outros átomos da molécula.²⁸</p>
Randic	1990	¹ χ^f	<p>Índice de conectividade variável. Randic definiu um descritor molecular em função de variáveis simples $f(x, y, \dots)$, de acordo com os tipos de heteroátomos presentes na molécula, ou atribuir diferentes pesos para um mesmo tipo de átomo em diferentes ambientes moleculares.³⁶</p>
Balaban Randic	1998 2000	IB p_k/w_k	<p>Representa uma extensão do índice J.⁴³ Quociente percurso/caminho³⁶</p>

Tabela 2. Lista das características desejadas para os índices topológicos proposta por Randic citada no artigo publicado por Mihalic e Trinajstic.⁶

NÚMERO	CARACTERÍSTICAS
1	Interpretação estrutural direta
2	Boa correlação com pelo menos uma propriedade molecular
3	Boa discriminação de isômeros
4	Bem definido
5	Generalizável
6	Independente linearmente
7	Simplicidade
8	Não baseado em propriedades físico-químicas
9	Não relacionado superficialmente com outros índices
10	Eficiência de interpretação
11	Baseado em conceitos estruturais familiares
12	Correta dependência de tamanho
13	Mudança gradual com a gradual variação da estrutura

Fonte: MIHALIC, Z.; TRINAJSTIC, N. *J. Chem. Ed.*, v. 69, p. 701-12, 1992.

Como exemplo, Randić⁶⁰ comparou a análise de regressão múltipla para a refração molar de 35 alquilbenzenos baseado nos índices de conectividade de ordens 1 a 4 e o correspondente descritor molecular ortogonalizado, e confirmou a estabilidade dos coeficientes nas equações de regressão onde foram usados os descritores ortogonalizados.

Wessel e Jurs⁴⁸ adotaram o procedimento de ortogonalização relatado por Randić para criar um modelo de seis variáveis, selecionadas entre 81 variáveis, para o ponto de ebulição normal de alcanos. Índices de conectividade ortogonalizados também foram propostos e testados com sucesso para aminoácidos, e uma discussão mais detalhada pode ser acompanhada nos artigos de Pogliani (2000) e Randić (2002).^{37,60}

No entanto, índices topológicos simples podem conter informações interessantes para a propriedade estudada.^{36,41} Considerando para alcanos, índices simples e diretos tais como o número de átomos de carbono, provem uma medida direta do tamanho molecular; o número de grupos metila terminais pode ser considerado como uma medida de ramificação ou compactação, e os índices de conectividade e conectividade de valência representam uma medida do tamanho da molécula.³⁷

A pesquisa para aumentar a área de aplicação de muitos índices topológicos bem estabelecidos tem levado a modificações para incluir características moleculares tri-dimensionais,³⁷ bem como representar moléculas contendo anéis, ligações múltiplas e heteroátomos.^{41,44,46,58}

Em relação a pobre discriminação de isômeros, vários autores vêm dando contribuições significativas. Dentre estes, podemos citar Cao e Yuan⁶¹ que derivaram um novo índice topológico a partir da matriz de distância modificado para discriminar isômeros cis/trans de cicloalcanos.

Heizen *et al.*⁶² desenvolveram o índice semi-empírico topológico (I_{ET}) para a predição da retenção cromatográfica de alcanos e alcenos lineares isômeros cis- e trans-. Esse índice mostrou boa capacidade para a discriminação de isômeros para um conjunto de alcenos lineares com C₅-C₁₄ átomos de carbono, contribuindo para resolver um ponto fraco dos índices topológicos.

Com o desenvolvimento de novos e mais sofisticados índices e métodos, suportados pelas ferramentas computacionais, o estudo das correlações estrutura-propriedade/atividade moleculares podem ser expandidos e aplicados em importantes áreas da química, farmacologia/toxicologia. Segundo Hall e Kier: “Nós estamos somente abrindo uma torneira do grande reservatório de informações”.³⁸

3. 3. 3 - Descritores Físicos

A inclusão dos descritores físicos neste trabalho deve-se a importância histórica dos estudos QSAR tradicionais, baseados no uso de propriedades físicas para caracterizar as moléculas. Propriedades físicas de uma molécula como coeficiente de partição octanol-água, refração molar e ponto de ebulição dominavam até 34 anos atrás as relações estrutura-atividade.³⁶ Esse conjunto de descritores moleculares é empírico e os dados experimentais são facilmente acessíveis na literatura.

Randic³⁵ considera que os descritores físicos não oferecem uma relação estrutura-propriedade, mas representam uma mistura de relações propriedade-propriedade e estrutura-atividade. Entretanto, as correlações propriedade-propriedade não devem ser desencorajadas, porque podem mostrar se as propriedades consideradas dependem da mesma característica estrutural ou não.

Contudo, atenção especial deve ser dada a possíveis colinearidades entre as propriedades quando se aplica RLM, o que pode apontar para dados inconsistentes.

Por outro lado, descritores baseados na mecânica quântica proporcionam uma descrição mais exata e detalhada dos efeitos eletrônicos e geométricos do que os métodos empíricos. Portanto, os descritores físicos, que não são derivados diretamente a partir da estrutura, podem ser muito úteis e complementares em combinação com outros tipos de descritores.

Romanelli *et al.*⁸ propuseram vários modelos de QSPR para prognosticar a solubilidade aquosa de alcoóis alifáticos baseados em descritores físicos. Os autores selecionaram o logaritmo do coeficiente de partição octanol-água, utilizado como uma medida da hidrofobicidade; polarizabilidade; volume molecular ligado a área acessível do solvente (SAG); massa; refratividade molar e volume. Os descritores derivados de métodos químico-quânticos estão relacionados com a estrutura molecular e complementam estudos prévios baseados em índices topológicos. Segundo os autores, em muitos casos é difícil dar explicação física para o comportamento molecular por uma representação puramente topológica da estrutura. Portanto, mais atenção deve ser dada para o desenvolvimento e aplicação de descritores físico-químicos, que podem ajudar no entendimento da estrutura molecular.

Os autores também aplicaram propriedades no estudo QSAR de toxicidade de álcool saturados, objetivando empregar descritores moleculares geométricos, eletrônicos e propriedades físicas.¹³

3. 3. 4 - Descritores Químico-Quânticos

Parâmetros estruturais eletrônicos são funções das cargas atômicas parciais de uma molécula, e eram calculados pelo método da carga atômica e o método estendido de Hückel. O desenvolvimento de novos programas computacionais como o sistema ADAPT e MOPAC, e de novos métodos semi-empíricos como modelo de Austin (AM1), modelo paramétrico 3 (PM3) e a teoria estendida de Hückel (EHT) têm permitido efetuar cálculos químico-quânticos mais rapidamente. Esta facilidade possibilitou empregar parâmetros eletrônicos como descritores moleculares em vários estudos de QSAR/QSPR. Esses métodos têm sido desenvolvidos com a divisão matemática da teoria do orbital molecular, mas baseados em simplificações e aproximações. Dados experimentais dos átomos e sistemas de protótipos moleculares têm sido freqüentemente usados para estimar valores das quantidades usadas nos cálculos como parâmetros, razão pela qual o procedimento é denominado de métodos semi-empíricos.

O artigo de Katritzky e Fara²² apresenta as várias classes dos descritores moleculares, de acordo com o método de cálculo. O trabalho de Karelson *et al.*³⁰ apresenta uma revisão de aplicações de descritores químico-quânticos no desenvolvimento de estudos de QSAR/QSPR com propriedades químicas, físicas, bioquímicas e farmacológicas de compostos. Os descritores químico-quânticos se dividem em:

(i) **Descritores Geométricos:** São descritores tridimensionais que caracterizam a molécula como um todo, tamanho e forma, e permitem a diferenciação de heteroátomos⁵⁸ entre moléculas que são topologicamente muito similares.⁵¹ É através da comparação entre o tamanho molecular de dois

compostos similares que podemos estimar qual terá o maior índice de retenção numa análise cromatográfica, por exemplo, ou qual terá o maior ponto de ebulição.

Os descritores geométricos são derivados somente a partir da geometria molecular 3D e da natureza dos grupos ligados em uma molécula⁵⁸, e são fundamentais no estudo entre estrutura-atividade. Alguns exemplos serão citados a seguir.^{1,50}

O volume molecular pode ser estimado por meio de cálculos à partir do raio de van der Waals dos átomos individuais.

A área superficial de um átomo em uma molécula corresponde à quantidade de área superficial daquele átomo que não apresenta intersecção com a área superficial de nenhum outro átomo.

A ovalidade molecular é igual a razão entre a área superficial e a superfície esférica mínima e indica o grau de desvio esférico da estrutura molecular.

(ii) Descritores Eletrônicos: Esses descritores refletem basicamente alguma característica global da distribuição de carga parcial e podem dar alguma interpretação físico-química para o modelo com o qual se relacionam.¹

Em princípio, a teoria química quântica deve ser capaz de fornecer descrições quantitativas precisas da estrutura eletrônica da molécula e suas propriedades físico-químicas. Portanto, os parâmetros eletrônicos podem ser calculados por métodos químico-quânticos conhecidos como semi-empíricos. Nesse procedimento as quantidades usadas nos cálculos como parâmetros foram estimadas a partir de dados experimentais atômicos e protótipos de sistemas moleculares. Os descritores obtidos representam o ambiente eletrônico de cada molécula.⁵¹ Geralmente os descritores são funções da distribuição de cargas

atômicas parciais σ e π atribuídas para cada átomo na estrutura, e das energias do orbital molecular.³⁶

Como vários parâmetros podem ser descritos para cada molécula, é gerado um grande número de descritores. Alguns parâmetros da estrutura eletrônica molecular utilizada como descritores eletrônicos e o que representam serão citados a seguir.^{50,63,64}

A densidade de cargas (q^+/q^-) de uma estrutura molecular descreve a distribuição dos elétrons, que é responsável pelo comportamento químico de cada espécie. A carga atômica de uma molécula pode ser calculada em termos de distribuição de carga atômica ao longo de todos os centros atômicos pelo método de Mülliken para cargas parcionadas.

O momento de dipolo (μ) é um parâmetro relacionado à distribuição de cargas numa molécula, e tem sido largamente empregado na explicação de propriedades físico-químicas e reatividade molecular. O momento de dipolo deve ser zero para moléculas neutras e diferente de zero para moléculas com ligações covalentes entre átomos diferentes. A polarização originada pode ser estimada em função da geometria molecular, e pode oferecer indicações a respeito da reatividade ou das forças intermoleculares que determina o estado de agregação das mesmas. O momento dipolar total de uma molécula é o vetor soma dos momentos de dipolo das ligações individuais e costuma ser expresso em Debye (D).

Energia do HOMO, Energia do LUMO e GAP. Estes parâmetros estão relacionados com as energias dos orbitais de fronteira de uma molécula: o orbital ocupado de maior energia (HOMO), o orbital vazio de menor energia (LUMO) e o intervalo entre estes orbitais (GAP). O HOMO e o LUMO podem ser usados, respectivamente, como uma medida da capacidade molecular para doar ou aceitar

um par de elétrons, ou seja, estão relacionados ao potencial de ionização (PI) e à afinidade eletrônica (AE) da molécula. O GAP entre eles expressa a energia necessária para excitar um elétron a partir do HOMO para o LUMO.

Dureza e Moleza Molecular (η). A média da separação entre os orbitais de fronteira, ou seja, o intervalo entre as energias do HOMO e do LUMO, é a medida da dureza de um átomo. Quando este intervalo é pequeno, uma pequena quantidade de energia é capaz de elevar um elétron do HOMO para o LUMO. Isto significa que a distribuição de elétrons é facilmente rearranjada por uma perturbação, então se costuma dizer que a molécula é mole (pequeno intervalo entre HOMO e LUMO). Quando ocorre o contrário, e é preciso uma maior energia para perturbar o sistema, a molécula é dura.

A polarizabilidade (α') é a medida da susceptibilidade da molécula ser polarizada quando sofre uma perturbação externa, e é fornecida pela constante de proporcionalidade α' entre o momento de dipolo e o campo elétrico.

O desenvolvimento de algoritmos eficientes tem ajudado na rotina de cálculos mecânico-quântico. Novos métodos semi-empíricos resultam em valores realísticos de grandezas moleculares obtidas por modelos da mecânica quântica em tempos razoáveis de computação. Assim a química quântica é uma nova e atrativa fonte de descritores que pode, a princípio, expressar todas as propriedades eletrônicas e geométricas de uma molécula e suas interações com o ambiente, fornecendo uma visão mais detalhada e acurada dos efeitos eletrônicos que os métodos empíricos.

3. 4 - Propriedades Moleculares e Aplicações da Correlação Quantitativa entre Estrutura-Propriedade

Um conceito considerado fundamental na química é que as características de uma molécula são responsáveis por suas propriedades. Assim sendo, pesquisas para encontrar regularidades entre propriedades (físicas, químicas e biológicas) de compostos que dependem de suas estruturas moleculares, continuam sendo assunto de investigações na área de correlação quantitativa estrutura-propriedade/atividade (QSPR/QSAR).^{22,29}

A seguir serão apresentadas algumas propriedades que supostamente são determinadas pelas forças intermoleculares e, portanto, podem ser adequadamente prognosticadas através de correlações estrutura-propriedade.

3. 4. 1 - Retenção Gás Cromatográfica

A cromatografia gasosa é muito utilizada na indústria petroquímica porque permite a separação de misturas complexas de hidrocarbonetos, tais como os compostos aromáticos, que recebem atenção especial em razão do valor agregado que apresentam.⁶⁵

Entretanto, a retenção cromatográfica é um processo muito complexo, influenciado por diferentes forças de interações intermoleculares, incluindo as forças de dispersão (ou forças de London), orientação (dipolo-dipolo ou forças de Keesom), indução (dipolo-dipolo induzido ou forças de Debye) e do tipo doador-receptor de elétrons em interações de complexação, ligações de hidrogênio e partição gás/líquido do soluto.^{66,67}

As forças associadas com estas interações podem estar relacionadas com as propriedades topológicas, geométricas e eletrônicas destas moléculas.¹

Essas interações determinam os valores da volatilidade relativa dos solutos bem como dos coeficientes de partição sendo, portanto, responsáveis pelos tempos de retenção e as separações que ocorrem nas colunas cromatográficas.

As regularidades dos tempos de retenção com o número de átomos de carbono em alcanos lineares levaram Kováts a propor uma escala de medidas denominada índice de retenção. O índice de retenção de Kováts solucionou o problema de uniformizar a expressão de dados de retenção³⁶, e foi estendido para moléculas com estruturas mais complexas⁵, sendo geralmente recomendado para a identificação de compostos orgânicos.^{68,69}

O índice de retenção de Kováts (IR) provavelmente é o parâmetro de retenção mais exato para expressar a separação cromatografia.⁷⁰ Esse índice é muito conveniente para ser adotado no lugar dos tempos de retenção porque relaciona o tempo de retenção de um soluto com os tempos de retenção dos alcanos lineares, usados como padrões.

O índice de retenção de Kováts calculado para condições isotérmicas é obtido pela Equação 2:

$$IR = 100z + 100 \frac{\log t'_{Ri} - \log t'_{Rz}}{\log t'_{R(z+1)} - \log t'_{Rz}} \quad (2)$$

onde, IR é o índice de retenção de Kováts característico da substância i cromatografada em uma dada coluna (fase estacionária) a uma dada temperatura;

t'_{Rz} é o tempo de retenção ajustado do composto padrão homólogo com número de carbonos z ;

$t'_{R(z+1)}$ é o tempo de retenção ajustado do composto padrão homólogo com número de carbonos $z+1$;

t'_{Ri} é o tempo de retenção ajustado da substância i .

Vários pesquisadores vêm realizando estudos para correlacionar e prever índices de retenção de diversos tipos de compostos, tais como compostos odoríferos, olefinas, drogas, bifenilas policloradas, usando parâmetros físico-químicos ou descritores para caracterizar a molécula. Além da proposta de predição, extremamente útil para a identificação de compostos em química analítica, muita atenção tem sido dada a obtenção de informação estrutural da molécula a partir do índice de retenção.^{4,70-73}

A existência de correlações quantitativas entre a estrutura e retenção cromatográfica torna possível o prognóstico de índices de retenção de compostos através de suas estruturas químicas, ou permite o prognóstico de estruturas a partir de índices de retenção. Os modelos QSRR têm demonstrado ser uma ferramenta adequada em processos cromatográficos, uma vez que permitem obter modelos simples para explicar o comportamento de solutos em determinados sistemas cromatográficos.⁷⁴

Modelos matemáticos para a predição de índices de retenção adotando índices topológicos e diferentes fases estacionárias também têm sido foco de estudo.^{75,76}

Duas importantes revisões de estudos de QSRR, que serão de grande ajuda em estudos nessa área de conhecimento, foram publicadas recentemente por Kaliszan (2007)⁷¹ e Héberger (2007)¹⁷.

A combinação da metodologia de QSRR com a experiência e a intuição dos usuários de cromatografia a gás pode facilitar o desenvolvimento de métodos analíticos, permitindo um entendimento do mecanismo de retenção e, eventualmente, auxiliar na identificação de um pico através apenas dos dados de cromatografia gasosa (CG).⁷⁷ Portanto, algum conhecimento das características de retenção é imprescindível para complementar os dados do espectro de massas. Porém, quando uma biblioteca virtual não contém dados de retenção cromatográficos, uma predição automática do índice de retenção a partir da estrutura química pode ser de grande ajuda para a identificação a “priori” de um composto não conhecido.⁷⁸

3. 4. 2 - Propriedades Físico-Químicas

A habilidade para modelar teoricamente propriedades físico-químicas e termodinâmicas parece uma proposta atraente em função da falta de informações experimentais para um grande número de compostos. Os estudos de QSPR também podem fornecer informações importantes para o entendimento das características estruturais intrínsecas das moléculas relacionadas a essas propriedades.

Ponto de Ebulição

O ponto de ebulição é uma importante propriedade físico-química de compostos orgânicos usualmente utilizado para testar a eficácia de índices topológicos.⁴⁴ A temperatura de ebulição é uma propriedade física fortemente dependente das interações intermoleculares (ou forças de coesão) e da massa molecular e, portanto, está diretamente relacionado a estrutura química da molécula.¹¹ Buchwald e Bodor⁷⁸ derivaram um modelo simplificado fundamentado em considerações teóricas, que pode dar razoável descrição de uma variedade de propriedades relacionadas a interações intermoleculares (ponto de ebulição, entalpia de vaporização, pressão de vapor, tensão superficial, coeficiente de partição e solubilidade em água) para um grande grupo de líquidos orgânicos, excluindo aqueles capazes de formar ligações de hidrogênio.

Coefficiente de Partição Octanol/água

O coeficiente de partição octanol/água é uma propriedade de compostos orgânicos muito utilizada como uma medida da hidrofiliicidade/lipofiliicidade de uma molécula, e está relacionada ao processo de transporte através das membranas das células. Katritzky *et al.*¹¹ apresentam uma boa revisão de vários métodos utilizados para estimar o coeficiente de partição octanol/água a partir da estrutura química.

Volume de Retenção Específico

O conhecimento do volume de retenção específico, facilmente calculado a partir dos dados de retenção, permite calcular dados de equilíbrio líquido-vapor com base em cromatografia líquido-gás, tais como entalpia padrão molar de

solvatação e entropia padrão molar de solvatação. Dessa maneira, a energia livre de Gibbs é facilmente calculada, através da conhecida equação de estado que relaciona essa três propriedades termodinâmicas.⁷⁹

Solubilidade em Água

A solubilidade de um composto é definida como a quantidade de soluto dissolvido em uma solução saturada sob condições de equilíbrio.⁸⁰ A solubilidade em água é uma medida direta da hidrofobicidade, isto é, a tendência da água para excluir a substância a partir da solução. Essa é uma das propriedades físicas mais importantes, e tem numerosas aplicações para a predição do comportamento de reagentes químicos nos compartimentos ambientais ou utilizados em aplicações farmacêuticas. Vários exemplos de correlações-ponto de ebulição são encontrados na literatura.¹¹

3. 5 - Atividades Biológicas e Toxicidade - Aplicações das Correlações Quantitativas entre Estrutura-Atividade

As indústrias químicas e farmacêuticas vêm produzindo reagentes químicos e medicamentos em quantidades crescentes, sendo que alguns têm volume de produção maior do que um milhão de toneladas.⁸¹

A grande parte desses reagentes é liberada para o ambiente e devido as suas propriedades físico-químicas, tal como solubilidade em água, pressão de vapor e coeficiente de partição entre o solo ou sedimento e água, eles se distribuem entre os vários compartimentos ambientais.

O principal agravante é que existem suspeitas de que muitos desses reagentes são tóxicos e que pouco ou nenhum dado empírico está disponível na literatura.⁸² Assim sendo, existe a necessidade da avaliação potencial dos efeitos desses produtos químicos para o ambiente e a saúde humana.

Por outro lado na indústria farmacêutica, antes de entrar no mercado, todos os medicamentos são avaliados quanto a sua toxicidade.

A avaliação da toxicidade consiste na determinação do potencial de uma dada substância atuar como tóxica, as condições em que esse potencial se evidencia e identificar o mecanismo de ação tóxica.

A avaliação de risco é uma avaliação quantitativa da probabilidade de ocorrerem efeitos negativos no organismo vivo sob determinadas condições. Em função da quantidade de informações disponíveis, existem inúmeros bancos de dados listando os mais diferentes tipos de substâncias químicas e suas atividades farmacológicas.⁸³

Tendo em vista a necessidade de aumentar a eficiência de pesquisas e reduzir custos e tempo, bem como obter dados para prognosticar os efeitos sobre a saúde humana ou ambiental, o desenvolvimento de correlações quantitativas entre estrutura e atividade (QSAR) vem sendo usado como uma ferramenta de trabalho com credibilidade científica.⁸⁴

Investigações toxicológicas baseadas em QSAR permitem prognosticar o potencial tóxico de produtos químicos ainda não testados e a interpretação do mecanismo de ação, o qual pode ser reunido em duas categorias gerais: narcose (ou toxicidade não específica reversível) e toxicidade específica irreversível. Cerca de 70% dos reagentes químicos industriais apresentam ação tóxica do tipo narcose e a toxicidade específica mais comum ocorre através de interações eletro(núcleo)filica^{85,86}

Para desenvolver modelos de QSAR baseados na toxicidade, três componentes são necessários: dados confiáveis de atividade biológica/toxicidade para um grupo de compostos, dados de propriedade/estrutura molecular para cada composto do grupo e métodos estatísticos que relacionam a toxicidade e a estrutura, definidas como a variável dependente e independente, respectivamente.

A seleção de estruturas análogas tem seguido a regra na qual compostos de uma mesma classe estrutural devem apresentar toxicidades similares.⁸⁴

Para a seleção do descritor, é importante que esse tenha uma interpretação física que possibilite obter o melhor entendimento do mecanismo envolvido na resposta tóxica.⁸⁷

O mecanismo de ação tóxico é o princípio básico para determinar tanto a classe de compostos químicos apropriado para a inclusão na seleção, quanto o descritor molecular potencialmente relevante para o modelo de ação tóxica.⁸⁸

Os modelos estatísticos usados para o desenvolvimento de relações estrutura-toxicidade são de dois tipos: correlação e reconhecimento de padrão.

O método de correlação mais comum é a análise de regressão, provavelmente porque leva a um resultado de fácil entendimento e também fornece informações sobre o mecanismo básico da toxicidade.

A utilização de técnicas de QSAR em toxicologia ambiental e saúde humana tem sido o tópico de inúmeras revisões, incluindo a apresentação de conceitos básicos para o desenvolvimento dos modelos, mecanismos de ação tóxica e aplicações.^{81,82,88,89}

Cronim e Shultz⁸⁵ obtiveram correlações entre estrutura-atividade com alta significância estatística, adotando dados de toxicidade disponíveis na literatura. Esse resultado demonstra a robustez do método de QSAR, uma vez que as

medidas foram realizadas em diferentes condições experimentais, o que contribuiu para o decréscimo do ajuste estatístico.

Miller *et al.*¹⁶ em 2005 apresentaram o primeiro estudo da QSAR adotando descritores químico-quânticos para a predição do potencial de sensibilização da pele de 87 substâncias químicas. Segundo os autores, o modelo também pode ser usado para reduzir o número de animais submetidos a testes porque permite estimar a concentração requerida para dados de taxa de risco.

Akers *et al.*⁸⁶ desenvolveram modelos de QSAR de toxicidade para o ciliato *Tetrahymena pyriformis*. A concentração que inibe 50% do crescimento dessa população, IGC₅₀, foi determinada para alcanos halogenados, alcoóis halogenados e nitrilas halogenadas. Para essa classe de reagentes químicos é esperado uma ação tóxica segundo o mecanismo eletro(nucleo)fílico por causa da presença de átomos ou substituintes ricos em elétrons. Nesse estudo, o coeficiente de partição octanol/água foi usado como termo hidrofóbico e o E_{LUMO} como uma medida da eletrofilicidade. Dois modelos dependentes da hidrofobicidade foram obtidos: um para alcano halogenados e um segundo para os alcoóis halogenados. A análise de regressão para dados de inibição de crescimento em alcano halogenados indicou que esses compostos basicamente apresentam atividade narcótica (toxicidade não específica), e o mecanismo de toxicidade é largamente dependente da hidrofobicidade da molécula. Para as nitrilas halogenadas, não foi observado uma relação entre estrutura-atividade dependente da hidrofobicidade.

Bláha *et al.*⁹⁰ construíram um modelo de QSAR adotando descritores físico-químicos e utilizando métodos de análise multivariada para estimar a toxicidade exata de compostos alifáticos halogenados saturados e insaturados, determinada pela inibição da bioluminescência de *Photobacterium phosphoreum*, expressa como concentração efetiva, EC₅₀. Nesse modelo, foi encontrado que o tamanho

das moléculas é o parâmetro mais importante para a toxicidade de compostos haloalifáticos saturados, e surpreendentemente a hidrofobicidade é o parâmetro menos importante para a descrição da toxicidade desses compostos. Moléculas maiores apresentaram maior toxicidade quando comparadas com moléculas menores. Essa característica pode estar relacionada ao acúmulo de haloalcanos na membrana biológica ou a formação de ligações com biomacromoléculas.

3. 6 - Teoria do Método Semi-Empírico Topológico (I_{ET})

O grupo de pesquisa coordenado pelos Professores Dr. Rosendo Augusto Yunes e Dra. Vilma E. Fonseca Heinzen iniciou seus estudos na área de correlação entre estrutura e retenção cromatográfica (QSRR) para auxiliar a determinação estrutural das substâncias que se encontram em pequenas concentrações e que são isoladas e determinadas por métodos já bem conhecidos (CG-MS, RMN, etc.). Os autores publicaram uma série de artigos empregando descritores topológicos, eletrônicos e/ou geométricos.^{74,91-97}

Os descritores topológicos, derivados da teoria do grafo molecular, codificam a informação encontrada na estrutura molecular, tais como número de átomos, suas conectividades, número de insaturações (se contém) e a presença de heteroátomos. Contudo, nenhum índice topológico, até aquele momento, havia sido capaz de diferenciar estruturas isoméricas cis- e trans- de alcenos.⁶² Por outro lado, o princípio da aditividade do índice de retenção proposto por Kováts não é observado quando se aumenta o número de ramificações do átomo de carbono em uma molécula.

Esses fatos levaram Heinzen e Yunes⁶² a propor o Método Semi-Empírico Topológico e a formular o descritor molecular denominado índice semi-empírico topológico (I_{ET}), reunindo considerações topológicas e empíricas.

Os autores basearam-se na suposição de que a retenção de um átomo de carbono, que é causada por interação com a fase estacionária e seu peso molecular, é diminuída pelos efeitos estéreos de seus vizinhos, seja átomos de carbono ou grupos funcionais que também afetam a retenção cromatográfica de moléculas ramificadas.⁹⁸

O Método Semi-Empírico Topológico tem por objetivo auxiliar no problema da predição do índice de retenção, retirar informações sobre as características estruturais da molécula e o mecanismo de retenção cromatográfico.

Para o desenvolvimento do índice semi-empírico topológico foram considerados os seguintes princípios básicos:⁶²

- i. A retenção das moléculas do soluto é devida ao número de átomos de carbono e às interações específicas de cada átomo da molécula com a fase estacionária. A interação dos átomos de carbono com a fase estacionária é determinada por suas propriedades elétricas e reduzida pelo impedimento estéreo causado pelos átomos de carbono vizinhos. Desta forma, tanto os fatores eletrostáticos como os estéreos estão envolvidos neste fenômeno;
- ii. Os valores atribuídos aos átomos de carbono e grupos funcionais foram baseados no comportamento cromatográfico experimental das moléculas em estudo.

- iii. A representação das moléculas é baseada na teoria dos grafos, onde os átomos de carbono são considerados vértices do grafo molecular, sendo omitidos os átomos de hidrogênio.⁹

O método de Kováts considera que o índice de retenção cromatográfica de uma molécula deve-se a soma das contribuições de cada átomo de carbono (100 u.i. para cada) e para n-alcenos a relação entre o índice de retenção e o número de carbonos é linear. No entanto, essas considerações falham quando aplicadas para hidrocarbonetos ramificados. Como exemplo, pode-se citar o 2-metilpropano. O índice de retenção esperado para esse composto é de 400 u.i., entretanto o valor experimental foi de 365,7 u.i., mostrando que átomos de carbono secundário e terciário possuem um valor menor do que 100 u.i.⁹⁹

Nesse sentido, Heinzen e Yunes⁶² separam os átomos de carbono em primários, secundários, terciários e quaternários e atribuíram valores numéricos que consideram o decréscimo da contribuição de cada átomo para a retenção cromatográfica devido a efeitos estéreos (ver Tabela 3). A determinação do grau ou valor do efeito estéreo presente em cada átomo de carbono depende também do tamanho do substituinte e não somente do seu grau de substituição. Como as tendências dos desvios da linearidade são logarítmicas, é também necessário adicionar o valor do logaritmo do(s) átomo(s) de carbono adjacente(s) de acordo com a sua característica como carbono primário, secundário, terciário ou quaternário. O logaritmo destes valores, denominados de fragmentos C_i , dá uma correlação linear com o número de átomos de carbono ligados ao grupo metila e aos átomos de carbono secundário, terciário e quaternário.

Este método considera que o índice de retenção de alcanos é o resultado da interação de cada átomo de carbono com a fase estacionária e que essas

interações são decorrentes das forças de dispersão e reduzidas pelo efeito estéreo da vizinhança. Moléculas que apresentam heteroátomos e/ou insaturações também envolvem forças dipolo-dipolo induzido. A influência dos efeitos conformacionais sobre as forças intermoleculares tornou difícil a previsão destes efeitos sobre a retenção, baseando-se somente em considerações teóricas que se mostraram válidas para os alcanos. Desse modo, os autores atribuíram valores aos átomos de carbono da dupla ligação de alcenos baseado também nos valores experimentais, cálculos através de uma aproximação numérica.⁶²

Tabela 3 – Valores de C_i e $\log C_i$ para os átomos de carbono primário, secundário, terciário e quaternário de alcanos.

Fragmento	Valor	C_i	$\log C_i$
-CH ₃	100	1,0	-
-CH ₂ -	90	0,9	- 0,0457
-CH<	80	0,8	- 0,0969
>C<	70	0,7	- 0,1549

Fonte: HEINZEN, V. E. F. *et al.* **J. Chromatogr. A**, v. 849, p.495-506, 1999.

Sendo assim, foram obtidos valores para cada fragmento molecular e o I_{ET} pode ser expresso através da Equação 3:

$$I_{ET} = \Sigma (C_i + \delta_i)$$

$$\delta_i = \log C_1 + \log C_2 + \dots \quad (3)$$

sendo, C_i = valor atribuído ao fragmento e a cada átomo de carbono e/ou fragmento i na molécula;

δ_i = soma do logaritmo dos valores de cada átomo de carbono adjacente (C_1, C_2, C_3 e C_4) e/ou o logaritmo do valor do fragmento.

Junkes *et al.* estenderam o índice semi-empírico topológico em estudos QSRR para diversas classes de compostos orgânicos, tais como alcanos metil ramificados produzidos por insetos⁹⁹, alcenos ramificados¹⁰⁰, ésteres¹⁰¹, aldeídos e cetonas em fases estacionárias de diferentes polaridades⁷⁵ e alcoóis saturados¹⁰². O I_{ET} possibilitou retirar informação sobre as características estruturais e as interações responsáveis pelo mecanismo de adsorção-dessorção do soluto na fase estacionária e também permitiu obter um modelo geral de QSRR para um variado grupo de compostos pertencentes a diferentes grupos funcionais.¹⁰³

Os excelentes resultados obtidos com apenas um descritor comparado com métodos que usam vários descritores em correlações múltiplas, encorajaram os autores para empregar o I_{ET} em estudos de QSPR/QSAR de oxo compostos¹⁰⁴, ésteres¹⁰¹ e alcoóis saturados¹⁴. Foram selecionadas algumas propriedades físicas e químicas representativas (ponto de ebulição, coeficiente de partição octanol/água, entre outras), e interessantes atividades biológicas (odor, atividade narcótica e toxicidade).

A simplicidade dos modelos que utilizaram correlações lineares simples também sugere que o método é adequado para auxiliar na identificação de compostos orgânicos em trabalho de rotina laboratorial. Portanto, os resultados obtidos com a aplicação desse método permitem, segundo os autores, que o mesmo possa ser estendido para desenvolver modelos de QSPR/QSAR de grupos de compostos orgânicos ainda não estudados.

3. 7- Classe de Compostos Orgânicos Investigados

3. 7. 1 - Hidrocarbonetos Alifáticos Halogenados

Os hidrocarbonetos alifáticos halogenados (HCX) são amplamente utilizados em atividades domésticas e industriais como solventes orgânicos, pesticidas, ou em várias aplicações de processos técnicos. Desde que esses compostos e seus derivados freqüentemente são hidrofóbicos e persistentes, o uso indiscriminado e o descarte inapropriado resulta no transporte e bioacumulação dos HCX nos vários compartimentos ambientais.^{105,106} Vários pesquisadores^{107,108} estão interessados na química dos compostos halogenados uma vez que foi identificado que esses compostos não são somente fabricados pelo homem, mas também sintetizados na natureza por um grande número de diferentes organismos.¹⁰⁹

No entanto, para o estudo da dinâmica destes compostos no meio ambiente é necessário a medida experimental de várias propriedades físicas, químicas e

biológicas, uma vez que a literatura é escassa em relação a estes dados devido as dificuldades experimentais e do alto custo envolvido.⁷⁷

Nesse contexto, os haloalcanos apresentam uma série de propriedades interessantes e constituem uma classe de compostos orgânicos atrativa para aplicações em estudos de QSRR/QSPR/QSAR.

A separação e identificação dos HCX em fontes sintéticas e naturais vêm sendo realizada satisfatoriamente em cromatografia gasosa (CG).^{107,110,111} Em alguns casos, quando a cromatografia gasosa combinada com espectrometria de massas (CG-MS) é usada para as análises, a tarefa de identificação nem sempre pode ser feita somente com base nos dados de MS uma vez que compostos isoméricos geralmente apresentam espectros de massa muito similares. Portanto, algum conhecimento das características de retenção é requerido para complementar os dados do espectro de massas. Contudo, quando livrarias virtuais não contêm dados de retenção cromatográfica, o prognóstico dos índices de retenção a partir da estrutura química pode ser de grande ajuda para a identificação de um composto ainda não medido experimentalmente.

Para a proposta de previsão, correlações quantitativas estrutura-retenção entre dados de retenção de cromatografia gasosa e diferentes descritores moleculares são inquestionavelmente de grande importância em química moderna e bioquímica.

QSPR/QSRR tem sido adotado para obter modelos matemáticos simples e de fácil construção para explicar e prognosticar o comportamento cromatográfico de compostos alifáticos halogenados adotando descritores derivados diretamente do grafo molecular. Embora o índice topológico desenvolvido por Kier e Hall^{112,113} seja o descritor mais conhecido, Ivanciuc e Ivanciuc⁷⁷ têm usado QSRR para

determinar o índice de retenção para 207 compostos halogenados adotando descritores computados a partir da estrutura química.

Sabljić¹¹⁴ estudou a habilidade do modelo de conectividade molecular para prognosticar o valor do índice de retenção de mono- e dicloroalcanos usando como parâmetro estatístico o coeficiente de correlação e a ordem correta de eluição como critério de ajuste dos dados. Os resultados permitiram relacionar a retenção cromatográfica de cloroalcanos com suas características estruturais.

Modelos de QSPR para prognosticar o ponto de ebulição (PE) para um grande número de alcanos halogenados e polihalogenados com 1 a 4 átomos de carbono foram desenvolvidos satisfatoriamente por Balaban *et al.*¹¹⁵ e Ivanciuc *et al.*^{19,77}

Garkani-Nejad *et al.*¹¹⁶ estudaram os dados cromatográficos de 846 compostos com uma grande diversidade de estruturas (drogas, pesticidas, poluentes e seus metabólitos) usando métodos de QSPR para desenvolver modelos de identificação de substâncias importantes em química forense analisadas por CG-MS.

A habilidade para modelar teoricamente propriedades termoquímicas de compostos halogenados é de grande importância devido a falta de informações experimentais, especialmente em moléculas contendo átomos de bromo e iodo. No entanto, o índice de retenção de Kováts de uma substância permite calcular parâmetros termodinâmicos de solvatação e interação molecular soluto-solvente tais como: Diferença de entalpia padrão parcial molar de solvatação, $\Delta_{\text{sol}}H^{\circ}$ (J/mol), diferença de entropia padrão parcial molar de solvatação, $\Delta_{\text{sol}}S^{\circ}$ (J/mol·K) e diferença de energia livre padrão parcial molar de solvatação, $\Delta_{\text{sol}}G^{\circ}$ (J/mol). Nesse sentido, modelos de QSRRs podem auxiliar na previsão de propriedades termoquímicas ainda não disponíveis na literatura.¹¹⁷

Luo e Pacey¹¹⁸ apresentaram um método para estimar as energias de ionização adiabáticas (IEs) e calor de formação de hidrocarbonetos e haloalcanos, com o objetivo de fornecer uma boa base para o entendimento de reações em fase gasosa de cátions de hidrocarbonos e seus derivados halogenados.

Buchward e Bodor⁷⁸ obtiveram um modelo simples a partir de considerações a nível molecular, unificado, onde o volume molecular pode dar razoável descrição de uma variedade de propriedades relacionadas com interações intermoleculares (ponto de ebulição, entalpia de vaporização, pressão de vapor, tensão superficial, partição e solubilidade) para um grande grupo de líquidos orgânicos, incluindo haloalcanos, os quais formam associações ou não contém substituintes fortemente polares.

3. 7. 2 - Cicloalcanos

Os cicloalcanos, conhecidos na indústria do petróleo por naftenos. Mais de 25% do petróleo bruto é constituído de compostos naftênicos e o cicloexano e seus derivados correspondem cerca de 5% em volume da composição da gasolina. Apresentam várias aplicações das quais se destacam a sua utilização como solvente e matéria-prima para a produção de outros importantes compostos orgânicos, destacando-se os hidrocarbonetos aromáticos, conhecidos como BTX (benzeno, tolueno e xileno). Na aplicação como solvente, o cicloexano é bem aceito no processo de desidratação de álcool substituindo o benzeno, mais tóxico em relação ao cicloexano. Também é utilizado como solvente da borracha crua, éteres de celulose, extração de óleos essenciais e na indústria de plásticos. É utilizado ainda na síntese de compostos de corrosão, inseticidas, adoçantes e corantes orgânicos, entre outros.^{119,120}

Devido a presença de hidrocarbonetos no petróleo, esses compostos são utilizados como indicadores de poluição, principalmente devido ao derramamento dessas substâncias em acidentes ambientais. No entanto, deve-se ressaltar que os hidrocarbonetos não existem apenas nos combustíveis fósseis, eles ocorrem também como produtos de biossíntese da maioria das plantas e animais ^x e são encontrados em diversas, tais como ar, solo e água.¹²¹

Os compostos são apolares, insolúveis em água e dissolvem-se em tetracloreto de carbono. São menos densos que a água e apresentam pontos de fusão e de ebulição que tendem a aumentar com o peso molecular, embora isômeros constituídos de moléculas mais compactas (com forças de London mais fracas) geralmente se fundem e fervem a temperaturas mais baixas do que os isômeros de cadeias mais extensas. Geralmente são compostos pouco reativos, com exceção do ciclopentano e ciclobutano, que tendem a sofrer reações com quebra da cadeia carbônica. Nestes compostos os ângulos das ligações no anel não correspondem ao ângulo tetraédrico ideal e conseqüentemente a sobreposição dos orbitais não é muito eficiente, diminuindo a estabilidade.

Para a análise de amostras petroquímicas, amplas pesquisas vêm sendo realizadas para desenvolver métodos analíticos mais avançados para a identificação e caracterização físico-química dos seus constituintes. Essas investigações são plenamente justificadas uma vez que as amostras petroquímicas são conhecidas como as mais complexas em Química Analítica. O número de compostos, somente no destilado intermediário, foi estimado acima de um milhão. As técnicas mais utilizadas na análise dessas substâncias são as cromatográficas e as espectrográficas. Dentre essas técnicas, a cromatografia gasosa é a mais popular.^{69, 122-125}

von Mühlen *et al.*¹²³ publicaram estudos de GCxGC (cromatografia gasosa bidimensional abrangente) para a caracterização de amostras petroquímicas e uma excelente revisão abordando biomarcadores e poluentes ambientais em água, ar e solo.

Neste trabalho, os cicloalcanos foram escolhidos como grupo de treinamento para construir modelos de QSRR/QSPR em função da sua importância industrial e de suas interessantes propriedades químicas.

4 - METODOLOGIA

4. 1- Índices de Retenção de Kováts para Compostos Alifáticos Halogenados e Cicloalcanos

Os dados cromatográficos dos índices de retenção de Kováts (IR) de 141 hidrocarbonetos alifáticos halogenados (clorados, bromados e iodados, mono ou dissubstituídos) medidos em diferentes fases estacionárias com baixa polaridade (OV-1, Apiezon-L e metilsilicona) em diferentes temperaturas foram obtidos da literatura^{111,114,126-128} e transferidos para Apiezon-L sob 130°C como mostrado em:

(i) Os valores de IR medidos em fase estacionária metilsilicona foram extrapolados para Apiezon-L sob 130°C conforme a Equação (4):

$$IR_{\text{metilsilicona}} = -320,4312(\pm 14,6403) + 0,9814(\pm 0,0156) RI_{\text{Apiezon-L } 130^{\circ}\text{C}}$$

$$N = 9; r^2 = 0,9982; SD = 5,9 \quad (4)$$

(ii) Os valores de IR medidos em fase estacionária OV-1, Apiezon-L sob 75°C e 125°C foram transferidos para Apiezon-L sob 130°C de maneira similar de acordo com as in Equações (5), (6) e (7), respectivamente e estão apresentados na Tabela 4. Os coeficientes de correlação entre as diferentes fases estacionárias foram excelentes com valores maiores do que 0,99.

$$IR_{OV-1} = 25,8169(\pm 9,1996) + 0,97223(\pm 0,0078) IR_{Apiezon-L 130^{\circ}C}$$

$$N = 26; r^2 = 0,9985; SD = 7,8 \quad (5)$$

$$IR_{Apiezon-L 75^{\circ}C} = 110,0848(\pm 18,9657) + 0,9530 (\pm 0,0200) IR_{Apiezon-L 130^{\circ}C}$$

$$N = 6; r^2 = 0,9982; SD = 5,4 \quad (6)$$

$$IR_{Apiezon-L 125^{\circ}C} = 125,4980(\pm 24,3542) + 0,9342 (\pm 0,0277) IR_{Apiezon-L 130^{\circ}C}$$

$$N = 9; r^2 = 0,9939; SD = 11,8 \quad (7)$$

Os dados cromatográficos dos índices de retenção de Kováts (IR) de 48 cicloalcanos fases estacionárias esqualano (70°C e 120°C) e dimetilsilicona (60°C) listados na Tabela 5 foram obtidos da literatura^{26,129,130}

O I_{ET} pode ser considerado como uma invariante usando fases estacionárias de baixa polaridade onde a relação entre o IR de Kováts e a temperatura são predominantemente lineares.¹³¹⁻¹³² Considerando a predominância da relação linear entre IR e temperatura em fases estacionárias não polar e de baixa polaridade, o índice semi-empírico topológico pode ser aplicado em diferentes intervalos de temperaturas em fases estacionárias de baixa polaridade. A influência da temperatura sobre a qualidade da regressão linear entre IR e I_{ET} pode ser verificada em prévias publicações.⁹⁹⁻¹⁰¹

Tabela 5. Valores de índice de retenção experimental, IR_{Exp} (esqualano a 70°C e 120°C e dimetilsilicona a 60°C) e ponto de ebulição normal, PE (°C), para cicloalcanos.

Nº	Compostos	(continua)	
		IR_{Exp}	PE (°C)
1	Ciclopentano	566 ¹	49,3
2	Metilciclopentano	628 ¹	71,8
3	Etilciclopentano	734 ¹	103,5
4	Isopropilciclopentano	812 ¹	126,5
5	Propilciclopentano	830 ¹	131
6	Butilciclopentano	929 ³	156,6
7	1,1-dimetilciclopentano	673 ¹	87,5
8	1,1,2-trimetilciclopentano	763 ¹	114
9	1,1,3-trimetilciclopentano	724 ¹	104,9
10	cis-1,2-dimetilciclopentano	721 ¹	99,5
11	trans-1,2-dimetilciclopentano	689 ¹	99,9
12	cis-1,3-dimetilciclopentano	683 ¹	90,8
13	trans-1,3-dimetilciclopentano	687 ¹	91,7
14	trans-3-etil-1-metilciclopentano	788 ³	-
15	cis-3-etil-1-metilciclopentano	790 ³	-
16	trans-2-etil-1-metilciclopentano	792 ³	-
17	cis-2-etil-1-metilciclopentano	820 ³	-
18	1-etil-1-metilciclopentano	794 ³	121,6
19	Cicloexano	663 ¹	80,73
20	Metilcicloexano	726 ¹	100,93
21	Etilcicloexano	834 ¹	131,9
22	Isopropilcicloexano	918 ¹	154,8
23	n-propilcicloexano	926 ¹	156,7
24	1,1-dimetilcicloexano	787 ¹	119,6

Tabela 5.
Continuação

Nº	Compostos	IR_{Exp}	PE(°C)
25	1,1,2-trimetilcicloexano	879 ³	145,2
26	1,1,3-trimetilcicloexano	840 ¹	136,6
27	cis-1,2-dimetilcicloexano	829 ¹	129,8
28	trans-1,2-dimetilcicloexano	802 ¹	123,5
29	cis-1,3-dimetilcicloexano	785 ¹	120,1
30	trans-1,3-dimetilcicloexano	806 ¹	124,5
31	cis-1,4-dimetilcicloexano	806 ¹	124,4
32	trans-1,4-dimetilcicloexano	798 ¹	119,4
33	Cicloeptano	786 ¹	118,4
34	Metilcicloeptano	859 ¹	-
35	etilcicloeptano	965 ¹	-
36	1,1-dimetilcicloeptano	912 ¹	-
37	trans-1,2-dimetilcicloeptano	936 ¹	-
38	cis-1,2-dimetilcicloeptano	956 ¹	-
39	trans-1,3-dimetilcicloeptano	922 ¹	-
40	cis-1,3-dimetilcicloeptano	916 ¹	-
41	trans-1,4-dimetilcicloeptano	917 ¹	-
42	cis-1,4-dimetilcicloeptano	921 ¹	-
43	Ciclooctano	915 ¹	149
44	Metilciclooctano	984 ¹	-
45	Ciclononano	1034 ¹	178,4
46	Ciclodecano	1138 ²	202
47	Cicloundecano	1223 ²	-
48	Ciclododecano	1312 ²	247

¹ Valores da referência [26]; ² Valores da referência [129]; ³ Valores da referência [130].

4. 2- Propriedades Físico-Químicas, Termodinâmicas e Atividades Biológicas

O ponto de ebulição normal, PE (°C), e ponto de fusão, PF(°C), a densidade, D (g/cm³), e o índice de refração, nD, usados para desenvolver os modelos estrutura-propriedade foram obtidos da literatura.¹³³

Os valores experimentais para o coeficiente de partição octanol-água, log P, o descritor universalmente aceito como medida da hidrofobicidade (ou lipofilicidade) de compostos orgânicos, também foram obtidos da literatura.^{78,86,90} Valores de log P também foram calculados segundo o método baseado na aditividade de Contribuição de Fragmentos e o método CNDO para serem comparados com os valores obtidos com o método I_{ET}.¹³⁴ Esses dados estão na Tabela 4 do Apêndice A.

O volume molecular, v (Å³), coeficiente de absorção de Ostwald gás-hexadecano, log γ_{hd} , equilíbrio líquido-vapor, log ρ^{gas} , e solubilidade aquosa, log ρ^w , foram obtidos no artigo publicado por Buchwald and Bodor⁷⁸. O interesse nessa variedade de propriedades é porque elas estão diretamente relacionadas com as forças intermoleculares. Portanto, o teste mais direto de qualquer modelo que permita tirar informações sobre forças intermoleculares envolve o prognóstico dessas propriedades.⁷⁸ Os dados exatos obtidos para o equilíbrio líquido-vapor por cromatografia líquido-gás em fases estacionárias de alcanos não voláteis com estruturas bem definidas sob altas temperaturas são apropriados para a determinação de parâmetros termodinâmicos de solvatação e interações moleculares soluto-solvente e para prognosticar o coeficiente de atividade de solutos em solventes voláteis sob temperatura ambiente.

O volume de retenção específico a 373 K, $\log V_w$ (cm^3/g), diferença de entalpia padrão parcial molar de solvatação, $\Delta_{\text{sol}}H^\circ$ (J/mol), diferença de entropia padrão parcial molar de solvatação, $\Delta_{\text{sol}}S^\circ$ (J/mol·K) foram compilados de Dallos.¹¹⁷

O calor padrão de formação, ΔH_f° (kcal/mol), foi obtido no trabalho de Luo e Pacey.¹¹⁸ O conhecimento de dados termoquímicos de íons na fase gás fornece uma boa base para o entendimento de reações gasosas de quebra heterolítica de halocompostos produzindo hidrocarbonetos (cátions) e seus halogênios (ânions). A avaliação da toxicidade do potencial mutagênico, concentração de um composto que provoca a morte de 50% da população ensaiada de *Tetrahymena pyriformis*, $\log EC_{50}$, (mg/dm^3) foi compilada de Bláha *et al.*⁹⁰ e a concentração do reagente que inibe 50% do crescimento da população do ciliato *Tetrahymena pyriformis*, $\log IGC^{-1}_{50}$, (mg/dm^3) de Akers *et al.*⁸⁶. Os compostos e os valores experimentais de todas as propriedades físico-químicas/termodinâmicas e atividades estão listadas na Tabela 6.

4. 3- Cálculo do Índice Semi-Empírico Topológico para Compostos Alifáticos Halogenados

O desenvolvimento do I_{ET} para alcanos, alcenos, ésteres, aldeídos, cetonas e alcoóis lineares e ramificados, foram descritos em trabalhos anteriores.^{14,62,75,97-104} As mesmas considerações que foram observadas durante o desenvolvimento do Método Semi-Empírico Topológico para prognosticar os índices de retenção dos grupos de compostos citados anteriormente foram empregadas nesse trabalho.

Na investigação para os hidrocarbonetos alifáticos halogenados, as moléculas foram representadas por grafos moleculares omitindo-se os átomos de hidrogênio, baseado na teoria dos grafos, onde os átomos de carbono (C_i) são os vértices do grafo. Em função da elevada eletronegatividade dos átomos de halogênio, a presença e posição do heteroátomo na cadeia carbônica influenciam significativamente as interações entre o soluto e a fase estacionária.

Por esse motivo, e para seguir as mesmas considerações ou convenções adotadas para os grupos funcionais em trabalhos anteriores, a contribuição para a retenção cromatográfica foi considerada para o fragmento C-X e X-C-X, onde X representa o átomo de cloro, bromo ou iodo.^{75,97,101-104} Cada fragmento também é representado por simples símbolo C_i e é um vértice do grafo.

Esse modo simples de representar a molécula tem permitido obter as melhores aproximações numéricas, capazes de refletir o comportamento cromatográfico dos compostos previamente investigados. Dessa forma, o índice semi-empírico topológico (I_{ET}) é expresso através da Equação 3, apresentada novamente a seguir:

$$I_{ET} = \sum (C_i + \delta_i)$$

$$\delta_i = \log C_1 + \log C_2 + \dots \quad (3)$$

onde, para compostos alifáticos halogenados, C_i é o valor atribuído para cada átomo de carbono i e/ou para o fragmento C-X ou X-C-X na molécula; δ_i é a soma i do logaritmo dos valores de cada átomo de carbono (C_1 , C_2 , C_3 e C_4) e/ou o logaritmo do valor do fragmento C-X e X-C-X adjacente a C_i .

Os valores atribuídos aos átomos de carbono e aos grupos funcionais (C_i) para os hidrocarbonetos alifáticos halogenados foram calculados por aproximações numéricas, baseados em dados de índices de retenção de Kováts experimentais e suportados por considerações teóricas, apresentadas a seguir:

4. 3. 1 - Cálculo dos Valores dos Fragmentos C-X e X-C-X

(i) A partir dos valores de cada índice de retenção experimental de Kováts dos compostos halogenados, um valor aproximado de I_{ET} (I_{ETa}) é calculado através da Equação 8, previamente obtida para alcanos lineares contendo de 3 a 10 átomos de carbono⁶²:

$$IR_{exp} = 123,6871 (\pm 0,0007) I_{ETa} - 47,3557 (\pm 0,0044)$$

$$N = 8; r^2 = 1,0; SD = 0,0039 \quad (8)$$

(ii) Os compostos halogenados são agrupados segundo suas similaridades estruturais, isto é: tipo de átomo de halogênio (ligado ao átomo de carbono primário, secundário ou terciário), e átomos monohalogenados e dihalogenados diretamente ligados aos átomos de carbono. Os halocompostos secundários e terciários também foram agrupados de acordo com a posição do grupo na cadeia carbônica (2, 3, 4 ou 5). Moléculas com um segundo fragmento C-X na cadeia carbônica foram agrupadas de acordo com a posição adjacente ao primeiro grupo halogenado, isto é, posição α , β e γ (α = posição adjacente alfa ao primeiro grupo halogenado; β = posição adjacente beta ao primeiro grupo halogenado; γ = posição adjacente gama ao primeiro grupo halogenado).

(iii) Na próxima etapa, os valores de C_i para átomos de carbono primário ($-\text{CH}_3$), secundário ($-\text{CH}_2-$) e terciário ($>\text{CH}-$) previamente atribuídos por Yunes *et al.*⁶² (Tabela 3), foram usados para calcular os valores dos fragmentos (C_i) C-X e X-C-X para cada grupo de átomos monohalogenado e dihalogenado diretamente ligado ao átomo de carbono, usando o valor previamente calculado por I_{ETA} .

4. 3. 2 - Cálculo dos Valores de C_i para Átomos de Carbono Terciários ($>\text{CH}-$) em Posições α , β e γ e $>\text{CH-X}$ (átomo de halogênio diretamente ligado ao átomo de carbono terciário) em Posições α , β e γ em Relação ao Primeiro Fragmento C-X

(i) Para calcular os valores de C_i de $>\text{CH}-$ e $>\text{CH-X}$, o valor do I_{ETA} aproximado (item 1(i)), os valores atribuídos aos átomos de carbono primário ($-\text{CH}_3$), secundário ($-\text{CH}_2$) e terciário ($>\text{CH}-$) por Heinzen *et al.*⁶², juntamente com os valores previamente obtidos para o fragmento C-X (item 1(iii)) foram usados na Equação 3. Assim, diferentes valores medianos foram obtidos para os átomos de carbono terciário nas posições α , β e γ ou para $>\text{CH-X}$ em posições α , β e γ em relação ao fragmento C-X (vértice do grafo molecular).

Presume-se que o índice semi-empírico topológico codifica informações intrínsecas as características estruturais tais como a presença e posição de heteroátomos nas moléculas, as quais são responsáveis pela retenção entre essas moléculas e a fase estacionária.

4. 4- Cálculo do Índice Semi-Empírico Topológico para Cicloalcanos

No desenvolvimento do índice semi-empírico topológico para cicloalcanos com e sem substituições no anel, foram empregadas as mesmas considerações observadas para os hidrocarbonetos.⁶² A Equação 8 foi utilizada para o cálculo do I_{ETa} para cada composto a partir de seus IRs experimentais. Para as ramificações, foram fixados os valores atribuídos aos fragmentos C_i listados na Tabela 3 para os átomos de carbono primário (-CH₃), secundário (-CH₂) e terciário (>CH-).⁶² Os valores atribuídos ao fragmento C_i para átomos de carbono secundário (-CH₂-), terciário (>CH-) e quaternário (>C<) da cadeia cíclica foram determinados a fim de obter a melhor aproximação numérica capaz de refletir o comportamento cromatográfico destes compostos. Para a obtenção dos valores, foram consideradas as etapas descritas a seguir:

- (i) Os cicloalcanos foram divididos em dois grupos: com e sem ramificações.
- (ii) Adotando a Equação 3, foi calculado o valor do fragmento C_i que representa o grupo -CH₂- da cadeia cíclica não ramificada.
- (iii) Fixando o valor de -CH₂- calculado anteriormente, e os valores de I_{ETa} , foram determinados os valores dos fragmentos >CH- e >C< dos cicloalcanos monosubstituídos.
- (iv) Os cicloalcanos dissustituídos foram agrupados de acordo com a posição das ramificações em α , β ou γ , em relação aos fragmentos >CH- e >C< com valores determinados em (iii). O valor de C_i de cada carbono substituído foi estimado com o objetivo de diferenciar isômeros -cis/trans- de acordo com o comportamento cromatográfico observado.

4. 5- Métodos Estatísticos

4. 5. 1 - Desenvolvimento dos Modelos de QSRR/QSPR/QSAR por Regressão Linear Simples

Os programas computacionais Origin e Bilin foram usados para as análises de regressão entre o Índice semi-empírico topológico, considerado como variável independente, e a propriedade/atividade, como variável dependente.^{135,136} Para testar a qualidade da equação de regressão, serão usados como parâmetros estatísticos o coeficiente de determinação (r^2), o coeficiente de correlação (r) e o desvio padrão (SD).

4. 5. 2 - Avaliação do Potencial de Prognóstico dos Modelos de QSRR/QSPR/QSAR

Para estimar se o modelo apresenta excelente habilidade de prognóstico, bem como a validação interna dos modelos construídos, foi executado o teste de validação cruzada, usando o método "leave-n-out".¹³⁷ Nesse método, o modelo foi sistematicamente construído deixando, a cada rodada, um ou mais componentes fora, de acordo com o número total de compostos. A estabilidade do modelo é indicada pela comparação entre o coeficiente de validação cruzada (r^2_{CV}) e o coeficiente de determinação. O julgamento da exatidão do modelo é baseado na

comparação entre os valores calculados com o modelo obtido e os valores do conjunto de compostos de treinamento, obtendo assim uma medida do ajuste dos dados sobre uma curva. Para verificar se os resultados calculados apresentam erros sistemáticos ou se estão aleatoriamente distribuídos, é observada a distribuição dos valores residuais versus os valores calculados.

4. 5. 3 - Validação Externa dos Resultados

A estabilidade externa do modelo foi verificada através do método de validação externa, no qual o conjunto de dados original foi sistematicamente dividido em três subgrupos, e cada subgrupo é prognosticado usando os outros dois subgrupos como grupo treinamento ou um conjunto externo de compostos que não participou do modelo inicial. O modelo proposto será válido somente se apresentar alta estabilidade estatística.³⁹

5 - RESULTADOS E DISCUSSÃO

5. 1 - Aplicação do I_{ET} em Correlações Quantitativas entre Estrutura – Retenção Cromatográfica para Compostos Alifáticos Halogenados

A investigação do comportamento de retenção cromatográfico, o qual é determinado por diferentes interações entre a fase estacionária e a molécula do soluto é um problema muito interessante em estudos de QSRR.

Solutos polares, tais como hidrocarbonetos halogenados, podem contribuir significativamente para essas investigações, porque as forças indutivas e dispersivas são um fator muito importante na interação da molécula em cromatografia líquido-gás.

Pesquisando o comportamento cromatográfico em fases estacionárias apolares e de baixa polaridade, foi desenvolvido o índice semi-empírico topológico. A teoria e o desenvolvimento do I_{ET} está baseada nas seguintes suposições: o comportamento cromatográfico de uma molécula é causado predominantemente pelo número e interações dos átomos da molécula com a fase estacionária. Essas interações são determinadas pelas suas propriedades elétricas e pelo efeito estéreo sobre essas interações devido a outros átomos ligados entre si.

Assim, valores específicos são atribuídos para os átomos de carbono e o grupo funcional da molécula, observando o comportamento cromatográfico experimental e suportado por deduções teóricas. Esse índice foi satisfatoriamente usado como um simples descritor em uma variedade de regressões lineares

simples para desenvolver modelos de QSRR/QSPR/QSAR para alcanos, alcenos, ésteres, aldeídos, cetonas e alcoóis lineares e ramificados.^{14,62,75,97-104}

Um dos fatores mais importantes na retenção cromatográfica de compostos em fase estacionária não polar ou de baixa polaridade são as forças de dispersão entre a superfície do soluto e a fase estacionária, as quais estão relacionadas com fatores estéreos, eletrostático e tamanho molecular. Para os hidrocarbonetos alifáticos halogenados, também é necessário considerar o momento de dipolo permanente do grupo halogenado, o qual pode aumentar as interações dipolo-dipolo induzido com a fase estacionária.

Os átomos dos halogênios possuem elevada eletronegatividade e retiram elétrons pelo efeito indutivo. Também é conhecido que o tamanho do átomo aumenta na ordem do cloro, bromo e iodo, enquanto que a eletronegatividade decresce, assim como o momento de dipolo permanente das moléculas. As contribuições dos átomos de halogênio para a retenção cromatográfica estão indicadas pelos valores do fragmento $-\text{CH}_2\text{-X}$ na Tabela 7. Esses valores aumentam na ordem $\text{CH}_2\text{-Cl}$, $\text{CH}_2\text{-Br}$ e $\text{CH}_2\text{-I}$ e refletem o efeito de cada átomo de halogênio sobre as forças intermoleculares. O fragmento $\text{CH}_2\text{-I}$ quando comparado com os outros fragmentos, apresenta o maior tamanho molecular e polarizabilidade, indicando que as forças de dispersão são mais importantes do que as forças de indução no processo de separação cromatográfico dos compostos 1-halo-n-alcanos.

A fim de obter informações sobre o mecanismo de retenção de moléculas monohalogenadas, na Tabela 7, pode ser visto que os valores para o fragmento C-X diminuem na ordem de alcanos halogenados primário, secundário e terciário, e também quando o átomo de halogênio move-se para dentro da cadeia carbônica. Esses resultados são similares àqueles obtidos para cetonas alcenos,

alcanos metil ramificados, ésteres, aldeídos, e alcoóis em estudos anteriores^{75,99-102} e pode ser atribuído ao impedimento estéreo do lado alifático da cadeia sobre o grupo funcional, reduzindo significativamente a contribuição do heteroátomo para retenção cromatográfica. Um comportamento similar foi observado para compostos halogenados com dois átomos de cloro ligados diretamente no átomo de carbono secundário. Essa tendência indica que um importante comportamento geral foi encontrado para essas classes de compostos, independentemente do tipo de heteroátomo ou força de interação específica que o soluto pode fazer durante o processo de eluição. Pode-se supor que diferentes grupos funcionais atuam como uma ramificação na cadeia carbônica, reduzindo a interação da molécula com a fase estacionária. Portanto, os resultados sugerem que o fator estéreo é mais importante para o índice de retenção de uma molécula do que os efeitos polares.

Analisando a natureza e a variedade dos grupos funcionais é de se esperar que para cetonas, aldeídos e alcoóis o momento de dipolo permanente e possíveis ligações de hidrogênio influenciem na retenção cromatográfica. Similarmente o efeito indutivo doador de grupos metilas e receptor dos halogênios também devem ser importantes no processo de separação.

Como observado previamente⁶², devido a redução da interação entre os átomos de carbono terciário e quaternário com a fase estacionária, valores diferentes foram atribuídos para o átomo de carbono terciário nas posições α , β e γ ligados no fragmento C-X. Esses valores estão relacionados com a maneira na qual os fragmentos C-X estão conectados na molécula e explicam satisfatoriamente o comportamento cromatográfico observado.

Segundo Peng⁶⁶, em alcanos com cadeia ramificada o efeito de um grupo substituinte pode se estender para os átomos de carbono α , β e γ adjacentes e pode ser medido através da distribuição da densidade eletrônica na molécula.

Tabela 7. Valores de C-X, onde X = átomo de Cl, Br ou I, atribuído para os compostos alifáticos halogenados e valores de C_i atribuídos aos átomos de carbono primário, secundário e terciário em posições α , β e γ em relação ao grupos C-X ou X-C-X.

Fragmento	Posição do fragmento	C_i (X=Cl)	C_i (X=Br)	C_i (X=I)
—CH ₂ —X	-	2,7100	3,2882	3,8845
	2	1,9287	2,4207	2,9177
>CH—X	3	1,9344	2,3571	2,8996
	4	1,8614	2,3236	2,8251
	5	1,8339	-	-
	2	1,4280	1,7847	2,2878
>C—X	3	1,5548	1,9669	2,3810
	α^*	0,7094	0,6800	0,7273
>CH—	β^*	0,7106	0,6838	0,7201
	γ^*	0,7146	0,7035	0,7215
	α^*	1,5374	1,9253	-
>CH—X	β^*	1,6949	2,1033	-
	γ^*	1,8418	2,1516	-
X >CH— X		3,5900	3,5226	-
X X >C< X	2	2,6502	-	-
	3	2,7602	-	-
—CH ₃	-	1,0000	-	-
—CH ₂ —	-	0,9000	-	-

* em relação ao fragmento C-X ou X-C-X.

Sabljić¹¹³ também observou que cloro e grupo metila têm efeitos opostos sobre os valores do índice de retenção. Isso indica que os átomos de carbono terciário acomodam melhor as suas cargas e conseqüentemente estão mais estabilizados, necessitando de menor contribuição dos elétrons dos halogênios, e interagem menos com a fase estacionária.

Os valores atribuídos para a introdução de um segundo átomo de Cl ou Br nas posições α , β e γ com respeito ao fragmento $>CH-X$ estão listados na Tabela 7. Esses valores mostram que a contribuição decresce na seguinte ordem: γ , β e α . Essa falha do princípio de aditividade pode ser atribuída a forte interação entre o efeito indutivo dos átomos de halogênios quando eles estão ligados em átomos de carbonos vizinhos entre si e foi observado, também, em outras propriedades físico químicas. Nesse grupo de compostos estão incluídas moléculas que apresentam átomos de halogênios substituídos em extremidades opostas da cadeia carbônica. Essas moléculas são apolares em função dos efeitos indutivos em sentido contrário que surgem no interior da molécula.

Acompanhando o comportamento experimental dessas moléculas, pode-se observar que a contribuição de cada fragmento CH_2-X para os índices de retenção é fortemente dependente do tamanho do átomo e do comprimento da cadeia carbonada. Por exemplo, para o iodo e bromo foi observado que cada grupo contribui independentemente em moléculas a partir de três átomos de carbono, ou seja, cada extremidade está separada por apenas um grupo $-CH_2-$.

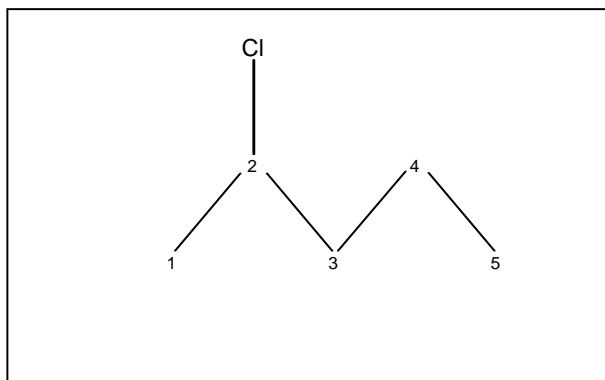
No entanto, para moléculas com um átomo de cloro substituído em cada extremidade da cadeia carbônica comportamento semelhante só foi observado em moléculas a partir de cinco átomos de carbono. Como visto anteriormente, um sistema carregado eletricamente se estabiliza quando a carga é dispersada ou

deslocada. Pode-se supor que grupos mais ramificados estão mais estabilizados e contribuem menos para a retenção do que grupos menos ramificados.

Essas discussões buscam explicar o fenômeno experimental observado e sugerem que as forças de dispersão determinam o mecanismo de interação de moléculas halogenadas apolares.

Os compostos halogenados revelaram uma grande versatilidade em suas interações com a fase estacionária apolar. Essa versatilidade é uma característica bem conhecida dos compostos alifáticos halogenados com diferentes estruturas moleculares, os quais podem ser desalogenados por um mecanismo de substituição nucleofílica monomolecular em função de fatores eletrônicos ou bimolecular devido a efeitos estéreos. Talvez essa seja uma das razões que levaram Ivanciuc e Ivanciuc⁷⁷ a considerar que os halogênios são particularmente difíceis para caracterizar somente com descritores estruturais. A inclusão de um descritor relacionado ao átomo de halogênio foi necessária para obter o melhor modelo QSRR em seu trabalho. A teoria e o desenvolvimento do I_{ET} levam em consideração não somente considerações teóricas, mas também aproximações empíricas. A inclusão do parâmetro empírico para calcular cada fragmento foi capaz de refletir o efeito do ambiente estrutural e codificar o efeito de cada heteroátomo.

A molécula do 2-cloropentano é dada como um exemplo do cálculo do índice semi-empírico topológico, adotando a Equação 3 e os valores listados na Tabela 7.



$$C_1 (-CH_3) = C_1 + \log C_2 = 1,0 + \log 1,9287 = 1,2853$$

$$C_2 (>CH-Cl) = C_2 + \log C_1 + \log C_3 = 1,9287 + \log 1,0 + \log 0,9 = 1,8829$$

$$C_3 (-CH_2-) = C_3 + \log C_2 + \log C_4 = 0,9 + \log 1,9287 + \log 0,9 = 1,1395$$

$$C_4 (-CH_2-) = C_4 + \log C_3 + \log C_5 = 0,9 + \log 0,9 + \log 1,0 = 0,8542$$

$$C_5 (-CH_3-) = C_5 + \log C_4 = 1,0 + \log 0,9 = 0,9542$$

$$I_{ET} = 1,2853 + 1,8831 + 1,1395 + 0,8542 + 0,9542 = \mathbf{6,1161}$$

A regressão linear simples obtida entre o índice de retenção experimental (IR_{EXP}) e o índice semi-empírico topológico (I_{ET}) para 141 hidrocarbonetos alifáticos halogenados é de excelente qualidade (Equação 9).

$$IR_{Exp} = -56,8944 (\pm 2,1311) + 124,7788 (\pm 0,2324) I_{ET}$$

$$N = 141; r^2 = 0,9995; SD = 8; r_{cv}^2 = 0,999 \quad (9)$$

No diagrama correspondente ao índice de retenção calculado (IR_{Cal}) versus o IR_{Exp} para compostos alifáticos halogenados em fase estacionária de baixa polaridade empregando a Equação 9, Figura 2, é possível visualizar o excelente ajuste dos dados sobre a curva. Em muitos casos o etano e o propano foram freqüentemente excluídos para melhorar o ajuste dos dados.¹³⁸ Nesse estudo, somente o etano não foi bem modelado com o I_{ET} . Provavelmente para esses compostos o efeito estéreo é menos importante do que a polarizabilidade no processo de retenção cromatográfico.

Para evidenciar que não ocorreram resultados tendenciosos, o gráfico dos valores dos erros residuais em função dos índices de retenção calculados (IR_{Cal}), Figura 3, mostram que os valores residuais estão aleatoriamente distribuídos em relação aos índices de retenção calculados.

O modelo tem alta estabilidade interna como indicado pelo coeficiente de correlação de validação cruzada, r^2_{cv} , de 0,999. A ordem de eluição da maioria dos compostos está correta demonstrando a alta capacidade de predição.

Com esta equação foi possível calcular o índice de retenção para 29 compostos que não foram incluídos no modelo (ver os valores de IR para os últimos 29 compostos na Tabela 4, Apêndice A).

A fim de verificar a validade estatística do modelo, a série completa dos dados foi sistematicamente dividida em três subgrupos (usando os compostos números 1, 4, 7, etc. da Tabela 4) e cada subgrupo foi prognosticado usando os outros dois subgrupos como grupo de treinamento. Os resultados mostraram uma qualidade de treinamento média de $r^2 = 0,9995$, e uma qualidade de predição média de $r^2 = 0,9995$, indicando que o modelo proposto tem uma alta validade, estabilidade estatística e robustez (Tabela 8).

Tabela 8. Verificação da validade estatística do modelo QSRR para compostos alifáticos halogenados.

Grupos de Treinamento	r^2^*	Grupo de Teste	r^2^*
1 e 2	0,9996	3	0,9995
1 e 3	0,9995	2	0,9995
2 e 3	0,9995	1	0,9996
Média	0,9995	Média	0,9995

* r^2 = coeficiente de determinação

Em um importante estudo, Ivanciuc e Ivanciuc⁷⁷ desenvolveram modelos QSRRs para estudar o índice de retenção cromatográfico de 207 compostos halogenados adotando descritores teóricos computados a partir da estrutura química e calculados com o programa CODESSA. As melhores regressões lineares simples foram obtida adotando os índices de Kier e Hall, $^1X^V$ e $^oX^V$ ($r=0,887$ e $0,854$ respectivamente).

Esses resultados foram melhores do que aqueles obtidos com descritores eletrostáticos ou quânticos mais complexos. Os melhores resultados foram obtidos com seis descritores estruturais com coeficientes de correlação entre 0,994 e 0,993. Pode ser observado que em todos os modelos, o índice de conectividade de Kier e Hall, $^1X^V$, estava presente.

O modelo obtido usando um descritor topológico simples, I_{ET} , obteve bons resultados (Equação 9), e foi capaz de codificar características estruturais do soluto que exercem um papel importante nas interações entre o soluto e a fase estacionária.

O prognóstico da retenção cromatográfica ou propriedades físico-químicas pode ser obtido para compostos que estão quimicamente relacionados com aqueles para os quais o modelo foi desenvolvido. O modelo de QSRR foi desenvolvido com um grupo de compostos que possuem três ou mais átomos de carbono na cadeia. Nesse modelo QSRR moléculas halogenadas pequenas e ou altamente substituídas não foram incluídas, porque esses compostos não são bem codificados pelo I_{ET} . Esse comportamento pode ser atribuído ao impedimento estéreo que provavelmente aumenta devido a conformação das moléculas do soluto, diminuindo as interações soluto-fase estacionária.

O processo de retenção cromatográfica é influenciado por diferentes forças de interações intermoleculares, incluindo as forças de dispersão (ou forças de London), orientação (dipolo-dipolo ou forças de Keesom), indução (dipolo-dipolo induzido ou forças de Debye) e do tipo doador-receptor de elétrons em interações de complexação, ligações de hidrogênio e partição gás/líquido do soluto.⁶⁶

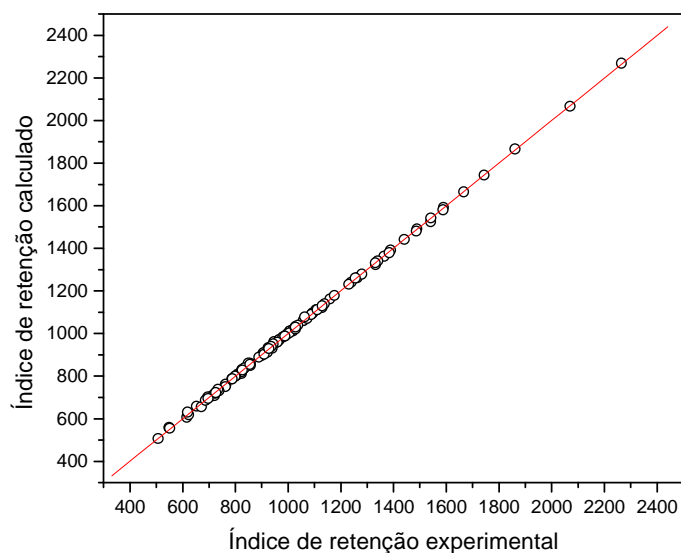


Figura 2. Gráfico da correlação linear simples entre os índices de retenção calculados (IR_{cal}) e os índices de retenção experimentais (IR_{exp}) para compostos alifáticos halogenados em fase estacionária de baixa polaridade, empregando a Equação 9.

Os tipos de interações moleculares favorecidos durante o processo de retenção estão relacionados principalmente com a polaridade da fase estacionária e a característica físico-química do soluto. A influência da polaridade da fase estacionária sobre o índice de retenção de alcoóis, aldeídos e cetonas alifáticas foram objeto de estudos de Junkes *et al.*^{75,102}, demonstrando o interesse dos Professores Yunes e Heinzen sobre esse tema.

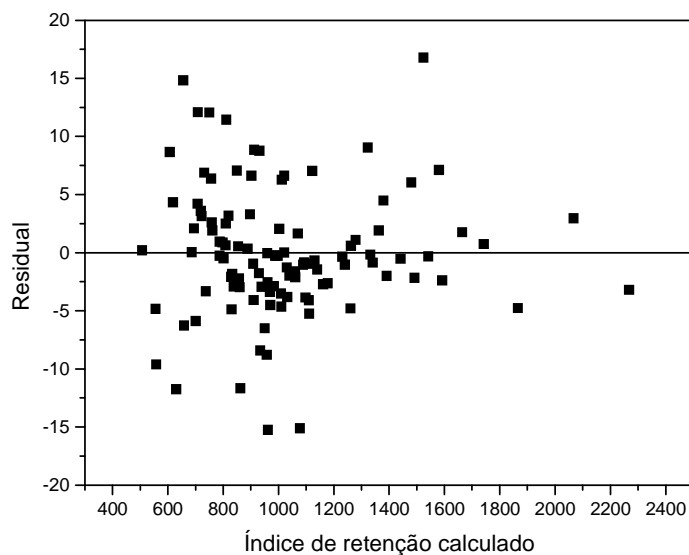


Figura 3. Gráfico dos valores dos erros residuais em função dos índices de retenção calculados para compostos alifáticos halogenados em fase estacionária de baixa polaridade, empregando a Equação 9.

No presente trabalho, o modelo de QSRR foi desenvolvido em coluna com fase estacionária não polar e demonstrou a habilidade do I_{ET} em prognosticar a retenção cromatográfica da série de compostos estudados. Entretanto, torna-se importante elucidar como o aumento da polaridade da fase estacionária alteraria as interações moleculares bem como o modelo obtido.

Nesse sentido, com o objetivo de verificar a habilidade de previsão do I_{ET} em fases estacionárias de diferentes polaridades e examinar a influência da polaridade da fase estacionária sobre a retenção cromatográfica, foram construídos oito modelos de QSRR para uma série de compostos alifáticos monohalogenados adotando índices de retenção obtidos em colunas

cromatográficas com fases estacionárias de diferentes polaridades, separadamente. A Tabela 9 apresenta os valores dos índices de retenção experimentais e as respectivas fases estacionárias. As fases estacionárias utilizadas em diferentes temperaturas apresentam polaridades que variam desde a menos polar DB-1 até a mais polar DB-Wax na seguinte ordem¹³⁹:

DB-1 (100% dimetilpolisiloxano) < DB-35 (35% difenil-65% dimetilpolisiloxano) < DB-17 (50% difenil-50% dimetilpolisiloxano) \cong DB-608 (50% difenil-50% dimetilpolisiloxano) (diferem na espessura do filme na coluna) < DB-210 (50% 3,3,3-trifluorpropil-50% metilpolisiloxano) < DB-225 (25% 3-cianopropil-25% fenil-50% dimetilpolisiloxano) < DB-Wax (polietileno glicol) \cong HP-Basic Wax (polietileno glicol modificado).

As correlações lineares simples entre os índices de retenção experimentais e o Índice Semi-Empírico Topológico para compostos alifáticos monohalogenados nas diversas fases estacionárias investigadas foram de boa qualidade, de acordo com os dados estatísticos apresentados na Tabela 10 (coeficientes de regressão β_0 e β_1 , coeficiente de determinação r^2 , desvio padrão SD e a equação linear simples $IR_{EXP} = \beta_0 + \beta_1 I_{ET}$). De acordo com Peng⁶⁶, a equação linear deve separar a contribuição molecular e a contribuição da interação do soluto com a fase estacionária e são parâmetros importantes para prognosticar a retenção cromatográfica, uma vez que as interações dependem da natureza dos grupos funcionais. A variação linear do índice de retenção dos compostos monoalogenados com o aumento da polaridade da fase estacionária pode ser acompanhada na Figura 4. Como esperado, os índices de retenção dos compostos aparecem deslocados para valores maiores como o aumento da polaridade da fase estacionária.

Tabela 9. Valores dos índices de retenção experimentais (IR) de compostos alceno halogenados medidos em fases estacionárias de diferentes polaridades: DB-1 (IR_{DB-1}), DB-35 (IR_{DB-35}), DB-17 (IR_{DB-17}), DB-608 (IR_{DB-608}), DB-210 (IR_{DB-210}), DB-225 (IR_{DB-225}), DB-Wax (IR_{DB-Wax}) e HP-B/Wax (IR_{HP-B/Wax}).

(continua)

Nº	Compostos	I_{ET}	IR _{DB-1}	IR _{DB-35}	IR _{DB-17}	IR _{DB-608}	IR _{DB-210}	IR _{DB-225}	IR _{DB-Wax}	IR _{HP-B/Wax}
1	1-clorobutano	5,7599	-	-	-	-	-	-	-	832
2	1-cloropentano	6,5684	739	811	829	823	880	906	946	941
3	1-clorohexano	7,3769	844	914	933	929	984	1008	1047	1049
4	1-cloroheptano	8,1854	948	1017	1036	1036	1088	1113	1152	1155
5	1-clorooctano	8,9939	1051	1121	1138	1138	1193	1220	1257	1260
6	1-clorononano	9,8024	1154	1224	1241	1241	1296	1327	1363	1364
7	1-clorodecano	10,6108	1257	1327	1343	1343	1401	1434	1469	1468
8	1-clorotetradecano	13,8400	1668	1740	1757	1757	1823	1860	1890	1886
9	1-clorohexadecano	15,4600	1873	1945	1964	1964	2031	2079	2104	2099
10	1-clorooctadecano	17,0800	2079	2152	2172	2173	2244	2293	2312	2307
11	1-bromobutano	6,4221	719	773	813	813	835	898	948	944
12	1-bromopentano	7,2306	823	902	923	920	954	1001	1055	1052
13	1-bromohexano	8,0391	927	1005	1026	1027	1059	1105	1160	1156
14	1-bromoheptano	8,8476	1031	1109	1130	1130	1165	1213	1266	1262
15	1-bromooctano	9,6561	1138	1216	1235	1234	1271	1322	1373	1368
16	1-bromononano	10,4646	1240	1314	1339	1343	1381	1439	1478	1476
17	1-bromodecano	11,2730	1344	1426	1444	1446	1483	1544	1583	1579
18	1-bromoundecano	12,0815	1448	1531	1549	1550	1589	1653	1692	1687
19	1-bromododecano	12,8900	1552	1636	1655	1655	1689	1764	1800	1793

Tabela 9.
Continuação

Nº	Compostos	<i>I</i>_{ET}	IR_{DB-1}	IR_{DB-35}	IR_{DB-17}	IR_{DB-608}	IR_{DB-210}	IR_{DB-225}	IR_{DB-Wax}	IR_{HP-B/Wax}
20	1-bromotridecano	13,6985	1652	1731	1762	1768	1814	1878	1905	1900
21	1-bromotetradecano	14,5070	1758	1837	1867	1876	1924	1989	2013	2007
22	1-bromohexadecano	16,1240	1965	2055	2079	2078	2128	2201	2258	2220
23	1-bromooctadecano	17,7409	2172	2265	2291	2292	2340	2405	2465	2436
24	1-iodopropano	6,2823	700	795	826	821	-	-	971	-
25	1-iodobutano	7,0908	798	900	929	923	920	1000	1062	1063
26	1-iodopentano	7,8923	903	1004	1032	1031	1026	1109	1164	1166
27	1-iodohexano	8,7078	1008	1110	1137	1134	1132	1218	1275	1272
28	1-iodoheptano	9,5163	1114	1218	1243	1241	1241	1330	1384	1381
29	1-iodooctano	10,3248	1221	1324	1350	1348	1349	1441	1493	1487
30	1-iodononano	11,1332	1327	1431	1457	1455	1457	1552	1600	1596
31	1-iododecano	11,9417	1433	1538	1563	1562	1562	1663	1706	1705
32	1-iodoundecno	12,7502	1539	1646	1671	1670	1671	1774	1816	1813
33	1-iodododecano	13,5587	1644	1753	1778	1778	1779	1885	1926	1922
34	1-iodohexadecano	16,7926	2064	2180	2209	2208	2213	2332	2362	-
35	1-iodooctadecano	18,4096	2272	2390	2422	2421	2427	2548	2576	-

Tabela 10. Escala de polaridade da fase estacionária, coeficientes de regressão para a equação linear simples ($IR_{Exp} = \beta_0 + \beta_1 I_{ET}$), desvios padrões, coeficientes de determinação e de validação cruzada obtidos em cada fase estacionária para compostos alifáticos monohalogenados.

Nº	FASES ESTACIONÁRIAS	P_R^*	β_0^*	β_1^*	N*	SD*	r^2^*	$r^2_{cv}^*$
1	DB-1	6,98	-108,6971	128,7742	34	5	0,9999	0,9999
2	DB-35	25,36	-38,4687	130,2094	34	18	0,9985	0,9983
3	DB-17	27,40	-15,809	130,2948	34	21	0,9980	0,9977
4	DB-608	30,11	-20,4608	130,6703	34	20	0,9982	0,9979
5	DB-210	47,18	-2,9608	131,8434	33	9	0,9997	0,9996
6	DB-225	56,41	24,1954	135,2109	33	21	0,9981	0,9977
7	DB-Wax	70,78	92,8091	133,0092	34	27	0,9968	0,9963
8	HP-B/Wax	71,91	104,9492	131,0525	32	24	0,9970	0,9966
9	Combinada ^a	-	-	-	268	18	0,9984	-

* P_R = escala de polaridade da fase estacionária; β_0 = coeficiente linear; β_1 = coeficiente angular; N= número de compostos; SD= desvio padrão; r^2 = coeficiente de determinação; r^2_{cv} = coeficiente de validação cruzada; ^a equação combinada

As curvas para as fases estacionárias DB-17 e DB-608, bem como para HP-Wax e HP-Basic Wax ficaram sobrepostas, confirmando a semelhança entre suas polaridades (Figura 4). Na Tabela 10 pode-se observar que o coeficiente angular, β_1 , varia pouco e oscila em torno de um valor médio igual a 131,3831, enquanto que o coeficiente linear, β_0 , aumenta a medida que aumenta a polaridade da fase estacionária. Esses resultados parecem indicar que a constante “ β_0 ” reflete o grau de influência da propriedade da fase estacionária sobre o índice de retenção, enquanto que a constante “ β_1 ” é pouco influenciada pela polaridade da fase estacionária. Resultados similares foram obtidos por Peng¹³⁹ na análise da influência da polaridade da fase estacionária sobre o índice de retenção de 1-halo-n-alcenos. No seu trabalho, Peng¹³⁹ estudou a correlação de cada halogênio separadamente para chegar a esta constatação.

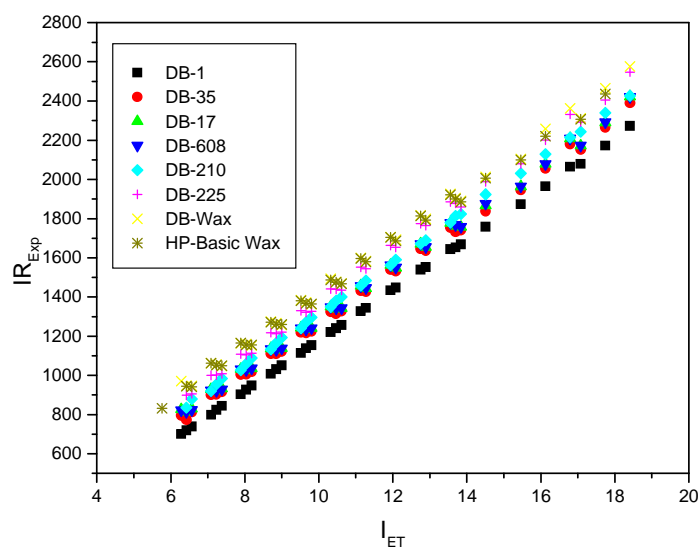


Figura 4. Correlações lineares entre os índices de retenção experimentais (IR_{exp}) e os índices semi-empíricos topológicos (I_{ET}) de compostos alifáticos halogenados em fases estacionárias DB-1, DB-35, DB-17, DB-608, DB-210, DB-225, DB-Wax e HP-Basic Wax.

No presente trabalho, o I_{ET} permitiu estabelecer um único modelo que ajusta adequadamente as moléculas com substituintes cloro, bromo e iodo, facilitando a obtenção de informações sobre a polaridade da fase estacionária e sua interação com o soluto sem a interferência da eletronegatividade do átomo de halogênio. A estrutura da fase estacionária é um fator importante para a retenção do soluto. A fase líquida DB-1 contém ligações Si-O e Si-C, enquanto que as fases DB-35, DB-17, DB-608, DB-210 e DB-225 incorporam grupos funcionais polarizáveis. As fases líquidas DB-Wax e HP-Basic Wax são as mais polares devido a série (C-C-O-) do polietileno glicol. O valor de " β_o " varia de acordo com a extensão das interações entre a estrutura química do soluto e as diferentes moléculas de polisiloxano na coluna cromatográfica.

Portanto, para a fase DB-1 as interações predominantes são de indução (dipolo-dipolo induzido ou forças de Debye), refletida pelo menor valor de " β_o ". O aumento do valor de " β_o " mostra que as orientações (dipolo-dipolo ou forças de Keesom) são as interações dominantes no processo de retenção. Esses resultados permitiram descrever o comportamento das interações intermoleculares, bem como quantificar essas interações. Portanto, os valores " β_o " e " β_1 " são parâmetros importantes para julgar o comportamento de retenção.

Para estabelecer o grau de polaridade das fases estacionárias, foi calculada a polaridade de retenção, P_R , segundo a escala de polaridade apresentada por Budahegyi *et al.*¹³² aplicando o conceito de Rohrschneider¹⁴⁰ e o sistema de McReynolds¹⁴⁰. Na tabela de McReynolds não constam as constantes correspondentes as fases estacionárias adotadas nessa investigação, portanto para caracterizar as polaridades foram escolhidas fases estacionárias consideradas equivalentes¹⁴¹: DB-1 \cong SE-30; DB-35 \cong SPB-35; DB-17 \cong SPB-50; DB-608 \cong OV-

17; DB-210 \cong SP-2401; DB-225 \cong OV-225; DB-Wax \cong SUPELCOWAX 10; HP-Basic Wax \cong Carbowax 20M.

Para investigar essas ponderações, foi obtida a equação 11 que correlaciona o coeficiente linear " β_0 " e as polaridades relativas, P_R , de acordo com a Tabela 10:

$$\beta_0 = - 117,3973 + 2,8999 P_R$$

$$N = 8; r^2 = 0,9429; SD = 18 \quad (11)$$

Os parâmetros estatísticos encontrados para a Equação 11 foram menores em qualidade do que aqueles alcançados nos trabalhos anteriores para alcoóis¹⁰², aldeídos e cetonas alifáticas⁷⁵. O menor desempenho pode ser atribuído a escolha da fase estacionária de polaridade similar que foi utilizada para o cálculo de P_R , uma vez que o critério de similaridade considera equivalentes fases estacionárias com ± 4 unidades nos valores de McReynolds.

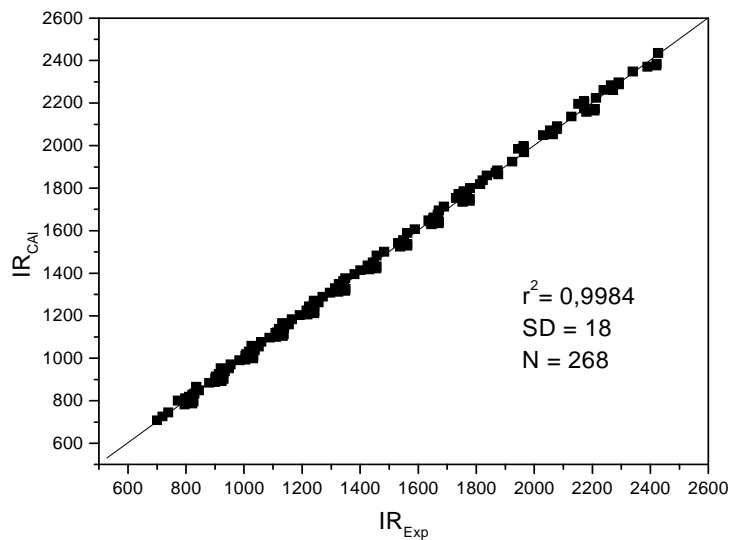
Como o I_{ET} foi desenvolvido em fases estacionárias de baixa polaridade, um parâmetro que representa a propriedade da fase estacionária e a polaridade de retenção foi acrescentado e um modelo único foi gerado para fases estacionárias com diferentes polaridades, de acordo com a Equação combinada 12:

$$IR_{Cal} = - 117,3973 + 2,8999 P_R + 131,3831 I_{ET} \quad (12)$$

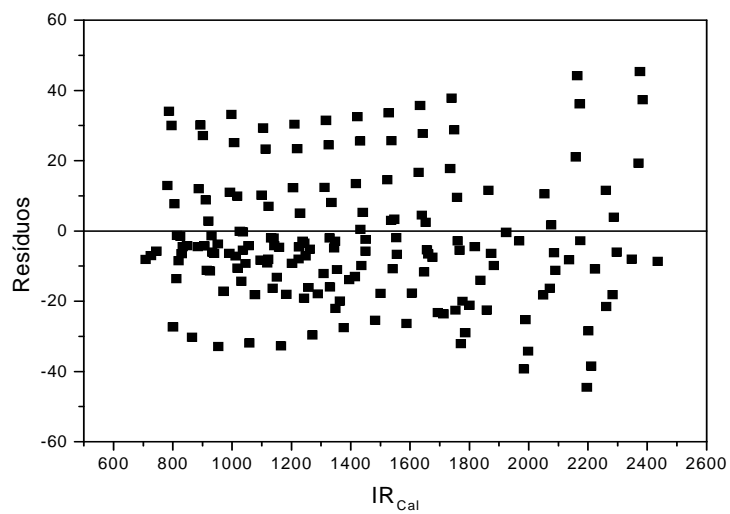
A Equação 12 foi empregada com sucesso para prognosticar os índices de retenção dos compostos alifáticos halogenados nas diferentes fases

estacionárias, como pode ser observado (Figura 5) através da concordância entre os valores calculados em função dos valores experimentais. Os parâmetros estatísticos corroboram com esse resultado ($r^2=0,9984$; $SD=18$; $N=268$). A dispersão dos resíduos em função do índice de retenção calculado mostra que o método não apresenta tendência, visto que os dados estão aleatoriamente distribuídos (Figura 5).

No entanto, o desvio padrão do método indica que esse modelo não pode ser convenientemente aplicado para a identificação dos compostos objeto deste estudo.



(a)



(b)

Figura 5. Diagrama: (a) da correlação linear simples entre os índices de retenção calculados (IR_{Cal}) em função dos índices de retenção experimentais, e (b) dos valores dos erros residuais em função dos índices de retenção calculados (IR_{Cal}), para compostos alifáticos halogenados em fases estacionárias de diferentes polaridades, empregando a Equação 12.

5. 2- Aplicação do I_{ET} em Correlação Quantitativa entre Estrutura – Retenção Cromatográfica para Cicloalcanos

O ambiente eletrônico das moléculas é um fator importante para as interações soluto-soluto, soluto-fase estacionária e soluto-fase móvel, as quais são responsáveis pelo processo de separação cromatográfica.²⁶ Em fases estacionárias apolares, dois tipos de interações contribuem para a sua interação com o soluto: forças de indução e de dispersão. Para os cicloalcanos, as forças de interação predominantes entre o soluto e a fase estacionária apolar são do tipo dispersiva. Forças de dispersão estão relacionadas com o fator estéreo, tamanho molecular e ramificação da cadeia, já as forças de indução estão relacionadas com o momento de dipolo permanente. Observando o comportamento geral dos índices de retenção cromatográficos de 48 cicloalcanos medidos em fases estacionárias de baixa polaridade, foram atribuídos valores aos fragmentos C_i dessas moléculas. Os valores de C_i , listados na Tabela 11, foram empregados no cálculo do Índice Semi-Empírico Topológico para cicloalcanos (ver Tabela 12). A interação de cada fragmento da molécula com a fase estacionária contribui para retenção cromatográfica.

Valores distintos de C_i foram atribuídos aos grupos metileno ($-CH_2-$) das cadeias cíclicas. Para anéis com 5 a 10 átomos de carbono o valor foi de 0,9712 e para anéis com 11 e 12 átomos de carbono o valor foi menor (0,9486). Os anéis com 5 e 6 átomos de carbono são abundantes na natureza, porque são suficientemente grandes para não ter tensão angular. A medida que aumenta a cadeia cíclica, os anéis não são planares, mas pregueados, de modo que cada ângulo de ligação do átomo de carbono sp^3 continue igual a $109,5^\circ$.¹²⁰ Anéis maiores estão mais aglutinados e a conformação das moléculas tem um papel

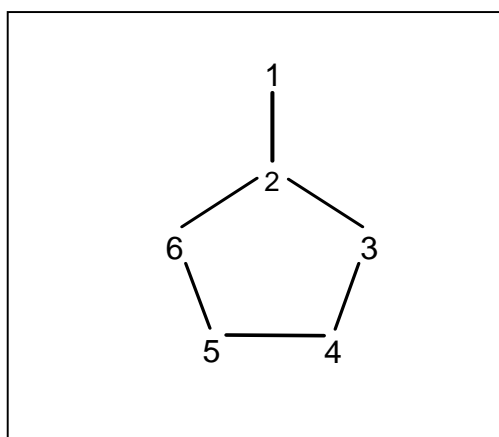
importante nas interações soluto-fase estacionária, a qual foi codificada pelo menor valor de C_i atribuído ao grupo $-CH_2-$ do cicloundecano e ciclododecano.

Para os cicloalcanos ramificados, foram considerados valores distintos para átomos de carbono terciário e quaternário (Tabela 11). Os valores representam o impedimento estéreo sobre esses átomos de carbono. Como esperado, o carbono terciário contribui mais para o IR do que a carbono quaternário, e grupos etila vizinhos diminuem ainda mais a interação do carbono com a fase estacionária.

Os cicloalcanos substituídos com isomeria cis/trans apresentam interessantes características estruturais na estrutura molecular, as quais foram reveladas pelos respectivos valores dos fragmentos, e estão de acordo com o ponto de ebulição experimental deste tipo de compostos. Na Tabela 11, pode-se acompanhar que o isômero cis- com substituição α e γ contribui mais para o IR do que o respectivo isômero trans-. Em isômeros trans- os grupos metila se encontram em lados opostos da molécula e na mesma distância um do outro, provavelmente compensando o efeito indutivo de cada grupo metila. Ao contrário, para isômeros trans- com substituição β , a contribuição para o IR diminui, quando comparada ao isômero cis-, indicando que grupos metila em lados opostos na molécula não tem seus efeitos anulados, predominando o impedimento estéreo. O mesmo comportamento é observado para o ponto de ebulição. Isto sugere que não somente a posição do substituinte, mas também a direção espacial relativa dos grupos tem efeito sobre o ponto de ebulição e para a retenção cromatográfica.⁶¹ Considerando a Tabela 11, pode-se observar que o valor de C_i para a substituição α é maior do que β e γ , ao contrário do esperado considerando o efeito estéreo do carbono vizinho, apontando para a sua maior afinidade com a fase estacionária.

As diferentes configurações geométricas resultaram em vários valores de C_i , os quais codificam as características estruturais que tem alguma influência sobre o processo de retenção cromatográfica.

Adotando os valores dos fragmentos C_i listados na Tabela 11, a molécula do metilciclopentano é utilizada com exemplo de cálculo do índice semi-empírico topológico.



$$C_1 (-\text{CH}_3) = C_1 + \log C_2 = 1,0 + \log 0,800 = 0,9031$$

$$C_2 (>\text{CH}-) = C_2 + \log C_1 + \log C_3 + \log C_6 = 0,8000 + \log 1,0 + \log 0,9712 + \log 0,9712 = 0,7746$$

$$C_3 (-\text{CH}_2-) = C_3 + \log C_2 + \log C_4 = 0,9712 + \log 0,800 + \log 0,9712 = 0,8616$$

$$C_4 (-\text{CH}_2-) = C_4 + \log C_3 + \log C_5 = 0,9712 + \log 0,9712 + \log 0,9712 = 0,9458$$

$$C_5 (-\text{CH}_2-) = C_5 + \log C_4 + \log C_6 = 0,9712 + \log 0,9712 + \log 0,9712 = 0,9458$$

$$C_6 (-\text{CH}_2-) = C_6 + \log C_2 + \log C_5 = 0,9712 + \log 0,800 + \log 0,9712 = 0,8616$$

$$I_{\text{ET}} = 0,9031 + 0,7746 + 0,8616 + 0,9458 + 0,9458 + 0,8616 = 5,2925$$

Tabela 11 - Valores de C_i para átomos de carbono primários, secundários e terciários das ramificações, átomos de carbono dos anéis com 5 a 12 átomos de carbono e átomos de carbono terciários e quaternários dos anéis com isomeria cis/trans de cicloalcanos calculados para fases estacionárias de baixa polaridade.

Fragmento	Posição do fragmento	C_i
$-\text{CH}_3^a$	-	1,0
$-\text{CH}_2^b$	-	0,9
$>\text{CH}^-^c$	-	0,8
$-\text{CH}_2^-^d$	-	0,9712
$-\text{CH}_2^-^e$	-	0,9486
$>\text{CH}^-^f$	-	0,8000
$>\text{CH}^-^g$ cis	α	0,8700
trans		0,7850
$>\text{CH}^-^g$ cis	β	0,7600
trans		0,7700
$>\text{CH}^-^g$ cis	γ	0,7700
trans		0,7600
$>\text{C}^h$	-	0,6500
$>\text{C}^i$	α	0,6900
$>\text{C}^i$	β	0,6000

^a metil; ^b etil; ^c isopropil; ^d carbono secundário do anel com 5 a 10 átomos de carbono; ^e carbono secundário do anel com 11 e 12 átomos de carbono; ^f carbono terciário do anel com 5 a 7 átomos de carbono; ^g carbono terciário na posição α , β ou γ em relação ao fragmento $>\text{CH}^-$; ^h carbono quaternário do anel com 5 a 7 átomos de carbono; ⁱ carbono quaternário do anel na posição α ou β em relação ao fragmento $>\text{C}^h$

Tabela 12. Valores dos índices semi-empírico topológicos (I_{ET}) calculados para cicloalcanos.

(continua)

Nº	Compostos	I_{ET}
1	Ciclopentano	4,7291
2	Metilciclopentano	5,2925
3	Etilciclopentano	6,0500
4	Isopropilciclopentano	6,8018
5	Propilciclopentano	6,8584
6	Butilciclopentano	7,7638
7	1,1-dimetilciclopentano	5,6849
8	1,1,2-trimetilciclopentano	6,3961
9	1,1,3-trimetilciclopentano	6,0593
10	cis-1,2-dimetilciclopentano	6,0353
11	trans-1,2-dimetilciclopentano	5,8163
12	cis-1,3-dimetilciclopentano	5,7492
13	trans-1,3-dimetilciclopentano	5,7762
14	trans-3-etil-1-metilciclopentano	6,5847
15	cis-3-etil-1-metilciclopentano	6,5575
16	trans-2-etil-1-metilciclopentano	6,6248
17	cis-2-etil-1-metilciclopentano	6,8438
18	1-etil-1-metilciclopentano	6,6372
19	Cicloexano	5,6749
20	Metilcicloexano	6,2384
21	Etilcicloexano	7,0468
22	Isopropilcicloexano	7,7476
23	n-propilcicloexano	7,8553
24	1,1-dimetilcicloexano	6,6307

Tabela 12.
Continuação

Nº	Compostos	I_{ET}
25	1,1,2-trimetilcicloexano	7,3379
26	1,1,3-trimetilcicloexano	7,0178
27	cis-1,2-dimetilcicloexano	6,9811
28	trans-1,2-dimetilcicloexano	6,7621
29	cis-1,3-dimetilcicloexano	6,6950
30	trans-1,3-dimetilcicloexano	6,7220
31	cis-1,4-dimetilcicloexano	6,7220
32	trans-1,4-dimetilcicloexano	6,6950
33	Cicloeptano	6,6207
34	Metilcicloeptano	7,1842
35	etilcicloeptano	7,9927
36	1,1-dimetilcicloeptano	7,5766
37	trans-1,2-dimetilcicloeptano	7,7080
38	cis-1,2-dimetilcicloeptano	7,9269
39	trans-1,3-dimetilcicloeptano	7,6678
40	cis-1,3-dimetilcicloeptano	7,6408
41	trans-1,4-dimetilcicloeptano	7,6408
42	cis-1,4-dimetilcicloeptano	7,6678
43	Ciclooctano	7,5665
44	Metilciclooctano	8,1300
45	Ciclononano	8,5124
46	Ciclodecano	9,4582
47	Cicloundecano	10,0272
48	Ciclododecano	10,7296

Pode-se observar que o Método Semi-Empírico Topológico também pode ser aplicado para prognosticar o índice de retenção de cicloalcanos, uma vez que permite obter informações sobre o comportamento da retenção cromatográfica e distinguir os isômeros cis- e trans- adotando um único descritor. Neste sentido, esse trabalho reporta o modelo de QSRR obtido para cicloalcanos em diferentes fases estacionárias apolares e alta qualidade dos parâmetros estatísticos:

$$\text{IR}_{\text{Exp}} = -37,5238 (\pm 6,5487) + 125,04264 (\pm 0,91726) \text{I}_{\text{ET}}$$
$$N = 48; r^2 = 0,9975; SD = 7; r^2_{\text{cv}} = 0,997 \quad (13)$$

A Equação 13 mostra que o coeficiente de determinação (r^2) ficou acima de 0,99 e que o modelo explica 99% da variância nos valores dos índices de retenção experimentais destes compostos. O procedimento de validação cruzada ($r^2_{\text{cv}} = 0,997$), utilizado para provar a estabilidade e validade interna do modelo, mostra a boa capacidade preditiva do modelo de QSRR para cicloalcanos. O ajuste dos dados entre o índice de retenção experimental de cada composto da série de treinamento e o respectivo I_{ET} (Figura 6), permite visualizar a adequação do modelo para estimar o índice de retenção de cicloalcanos em fases estacionárias de baixa polaridade. O gráfico da dispersão dos erros residuais em função dos índices de retenção calculados (IR_{Cal}) para 48 cicloalcanos em fases estacionárias esqualano e dimetisiloxano (Figura 6) evidencia que o modelo não apresenta tendências, uma vez que os pontos estão aleatoriamente distribuídos em torno de zero.

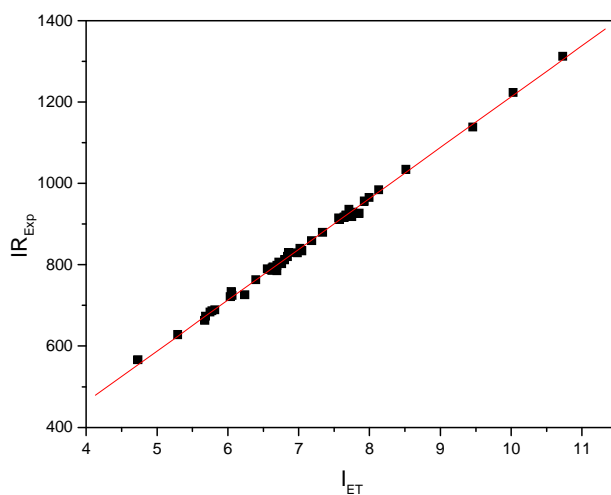
A estabilidade externa do modelo obtido foi testada dividindo sistematicamente os 48 compostos em 3 diferentes subgrupos (usando os

compostos nºs 1-2, 2-3, 1-3, etc. (Tabela 13) do conjunto de dados originais e cada subgrupo foi prognosticado usando os outros dois subgrupos como grupos de treinamento. Os valores de índice de retenção estimados para cada subgrupo (grupo de teste) foram obtidos através do modelo construído com os dois subgrupos (grupos de treinamento). Os resultados obtidos estão apresentados na Tabela 13 com uma qualidade de treinamento média ($r^2 = 0,9976$) e qualidade de predição média ($r^2 = 0,9923$), indicando a robustez do modelo proposto. Dimov¹²⁴ obteve um coeficiente de determinação de 0,97 para o IR de 30 cicloalcanos e derivados com 5 a 9 átomos de carbonos no anel, inferiores aqueles obtidos na Equação 13. Firpo *et al.*¹⁴² apresentam três equações resultantes de correlações múltiplas para calcular o IR de 30 cicloalcanos com descritores empírico químico quânticos. Os coeficientes de determinação apresentaram valores entre 0,9755 e 0,9949.

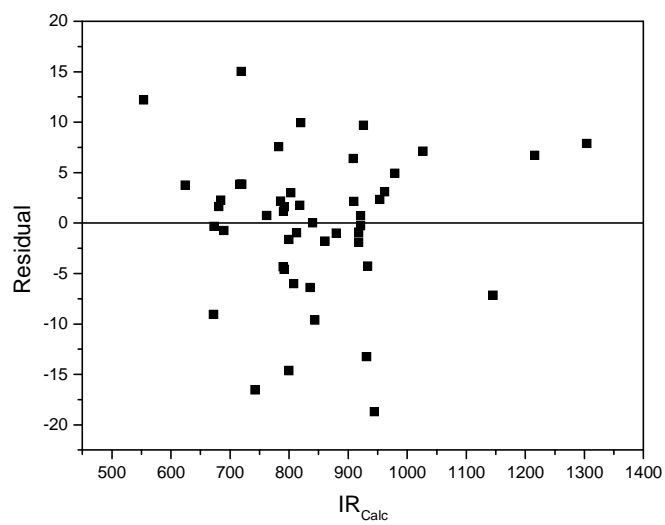
Tabela 13. Verificação da validade estatística do modelo QSRR para cicloalcanos.

Grupos de Treinamento	r^{2*}	Grupo de Teste	r^{2*}
1 e 2	0,9972	3	0,9829
1 e 3	0,9980	2	0,9967
2 e 3	0,9975	1	0,9975
Média	0,9976	Média	0,9923

* r^2 = coeficiente de determinação



(a)



(b)

Figura 6. Gráfico de: a) Regressão linear simples entre os índices de retenção experimentais (IR_{Exp}) e o I_{ET} ; b) Valores dos erros residuais em função dos índices de retenção calculados (IR_{Calc}) para 48 cicloalcanos em fases estacionárias esqualano e dimetilsiloxano.

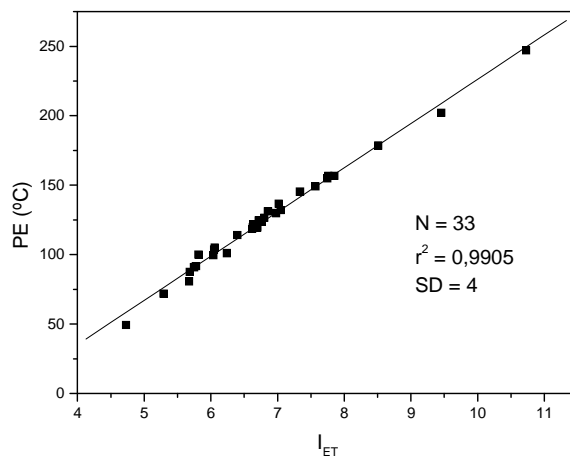
A qualidade do modelo de QSRR encontrado para o grupo de compostos alifáticos de cadeia cíclica nos estimula a propor um modelo de QSPR para prognosticar o ponto de ebulição normal dos cicloalcanos que participaram da série de treinamento. Os excelentes resultados mostraram o I_{ET} possui potencial suficiente para aplicação em estudos mais aprofundados envolvendo propriedades/atividades, importantes principalmente sob o ponto de vista ambiental, uma vez que a aquisição de dados é alcançada indiretamente, sem a necessidade de medidas experimentais. É de conhecimento geral que para a construção de modelos QSRR/QSPR/QSAR é necessário um grande número de dados experimentais e também que muitos desses dados estão disponíveis na literatura. Cabe aos pesquisadores selecionar as fontes mais confiáveis e continuar a enriquecer essa área de investigação.

O modelo de QSPR obtido, empregando um único descritor molecular, apresentado na Equação 14, para estimar o PE de cicloalcanos, apresentou parâmetros estatísticos satisfatórios ($r^2=0,9905$; $SD=4$; $r^2_{cv}=0,988$):

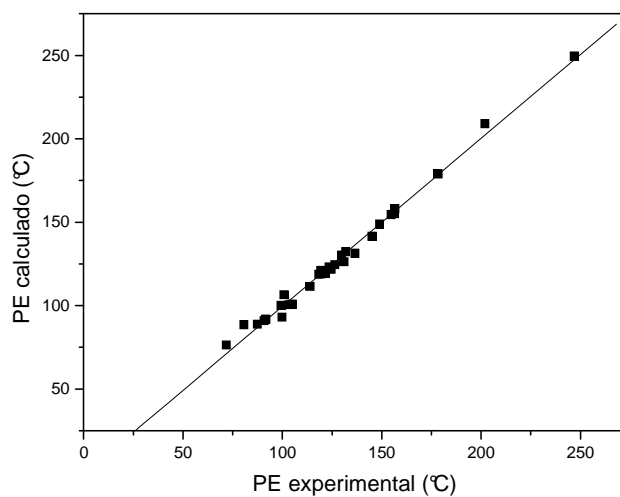
$$\begin{aligned} PE &= -92,29765 (\pm 3,8785) + 31,85815 (\pm 0,5617) I_{ET} \\ N &= 33; r^2 = 0,9905; SD = 4; r^2_{cv} = 0,988 \end{aligned} \quad (14)$$

A Figura 7 mostra o bom ajuste dos dados entre os valores de ponto de ebulição normal (PE) e o I_{ET} , bem como a concordância entre os valores experimentais e estimados (calculado com a Equação 14) de PE, demonstrando a habilidade de previsão do I_{ET} em estudos QSPR adotando dados reportados na literatura. Comparamos os resultados com aqueles obtidos por outros autores que empregaram apenas um descritor molecular derivado por diferentes métodos.

Toropov *et al.*¹⁴³ estimou o PE de 45 cicloalcanos usando um descritor calculado segundo o grafo da camada de valência e obteve o modelo com qualidade similar ao Método Semi-Empírico Topológico ($r=0,9960$; $SD=3,980$). Cao e Yuan⁶¹ modificaram a matriz de distância para distinguir 53 isômeros cis/trans de cicloalcanos e obtiveram resultados que também podem ser considerados similares ao nosso ($r=0,9961$; $SD=3,31$). Spicksma¹³⁰ estimou o PE de 36 naftenos com 5 e 6 átomos na cadeia cíclica e obteve baixo desvio padrão ($SD=0,72$) adotando o modelo de Kovats-Wilson. Este modelo não se mostrou adequado para estimar o PE de anéis com 7 a 9 átomos de carbono.



(a)



(b)

Figura 7. Diagrama de: (a) Correlação entre os valores de ponto de ebulição normal (PE) e o I_{ET} ; b) Ponto de ebulição calculado com a Equação 14 em função do ponto de ebulição normal experimental.

5.3 - Aplicação do I_{ET} em Correlações Quantitativas entre Estrutura – Propriedades Físico-Químicas e Termodinâmicas de Compostos Alifáticos Halogenados

Em trabalhos anteriores, o nosso grupo estendeu o I_{ET} em estudos de QSPR/QSAR para uma série de importantes propriedades físico-químicas, atividades biológicas e toxicidades de alcoóis.^{14,97,104} Muitas dessas propriedades foram bem modeladas ($r > 0,98$) empregando esse descritor molecular.

No presente trabalho, o I_{ET} foi empregado como único descritor baseado na estrutura molecular em estudos de propriedades físico-químicas, termodinâmicas e atividades biológicas para uma série de compostos alifáticos halogenados. De acordo com Ivanciuc e Ivanciuc⁷⁷, poucas propriedades experimentais tem sido determinadas para haloalcanos, o que reforça a necessidade de implementar modelos de QSPR/QSAR.

Os compostos e as propriedades/atividades biológicas adotadas para a construção dos modelos de QSPR/QSAR estão nas Tabelas 4 e 6.

A motivação para esse estudo surgiu da suposição de que propriedades que são determinadas por forças intermoleculares podem ser adequadamente modeladas pelo índice semi-empírico topológico. O coeficiente de determinação, desvio padrão e de validação cruzada foram adotados como critérios estatísticos para o julgamento da qualidade e força preditiva durante o desenvolvimento dos modelos.

Em um primeiro momento, pode-se esperar que propriedades/atividades que relacionam linearmente com o índice semi-empírico topológico podem estar correlacionadas entre si. Neste caso, os modelos não permitiram obter

informações sobre as características das moléculas que estão contribuindo para a propriedade/atividade. A ocorrência de intercorrelações entre as propriedades/atividades adotadas nesse estudo foi investigada (Tabela 14).

Mihalic *et al.* (apud Pogliani, 2000)³⁷ sugerem um interessante critério para examinar a colinearidade. Eles propõem que aquelas com $r \geq 0,98$ devem ser consideradas como fortemente intercorrelacionadas.

Como pode ser observado na Tabela 14, a maioria das propriedades/atividades estão correlacionadas entre si, com exceção das propriedades $\log p^w$ e $\log EC_{50}$ ($r = -0,1101$) e ΔH_f° , que foram pouco correlacionadas com todas as outras propriedades físico-químicas ($r < 0,70$). No entanto, é interessante observar que o ΔH_f° apresentou coeficientes de correlação maior do que 0,99 com as atividades biológicas estudadas.

A Tabela 15 resume os modelos QSPR obtidos para as propriedades físico-químicas e termodinâmicas usando equações de regressão linear e polinomial quadráticas com um único descritor molecular, I_{ET} , a qualidade estatística e a validação dos modelos. Esses resultados serão discutidos a seguir.

Tabela 14. Correlação entre as propriedades físico-químicas, termodinâmicas e atividades biológicas dos compostos alifáticos halogenados estudados (coeficiente de correlação e número de compostos)

	log EC ₅₀	log (IGC ₅₀ ⁻¹)	log P	PE	log Vw	v	log γ _{hd}	log ρ ^{gas}	log ρ ^w	Δ _{sol} H ^o	Δ _{sol} S ^o	ΔH _f ^o	
log EC ₅₀ ^a	1,0000												
log (IGC _b ⁻¹)		1,0000											
log P ^c	-0,9806 (3)	0,9964 (4)	1,0000										
PE ^d	-0,9807 (5)	0,9996 (3)	0,9900 (14)	1,0000									
log Vw ^e	(2)	0,9987 (3)	0,9813 (7)	0,9994 (6)	1,0000								
v ^f	(2)	0,9974 (3)	0,9980 (11)	0,9918 (11)	0,9936 (5)	1,0000							
log γ _{hd} ^g	-0,9200 (3)	0,9998 (3)	0,9924 (10)	0,9920 (11)	0,9987 (5)	0,9964 (10)	1,0000						
log ρ ^{gas h}	0,9818 (3)	-0,9983 (3)	-0,9932 (10)	-0,9979 (11)	-0,9995 (5)	-	0,9979 (10)	-0,9950 (11)	1,0000				
log ρ ^{w i}	-0,1101 (3)	-0,9988 (3)	-0,9983 (10)	-0,9434 (11)	-0,991 (5)	-	0,9974 (10)	-0,9762 (11)	0,9550 (11)	1,0000			
Δ _{sol} H ^{o j}	(2)	-0,9981 (3)	-0,9842 (8)	-0,9980 (7)	-0,9989 (12)	-	0,9974 (6)	-0,9997 (6)	0,9995 (6)	0,9944 (6)	1,0000		
Δ _{sol} S ^{o k}	(2)	-0,9973 (3)	-0,9859 (8)	-0,9916 (7)	-0,9950 (12)	-	0,9978 (6)	-0,9976 (6)	0,9966 (5)	0,9944 (6)	0,9989 (13)	1,000 0	
ΔH _f ^{o l}	0,9908 (4)	-0,9957 (4)	-0,5262 (11)	-0,1950 (14)	-0,5628 (9)	-	0,6932 (8)	-0,6532 (8)	0,6514 (8)	0,6953 (8)	0,6239 (9)	0,685 1 (9)	1,000 0

^a toxicidade ao *Tetrahymena pyriformis*, (mg/dm³); ^b Toxicidade ao ciliato *Tetrahymena pyriformis*, (mg/dm³);
^c Coeficiente de partição octanol/água; ^d Ponto de ebulição normal, (°C); ^e Volume de retenção específico a 373K, (cm³/g); ^f Volume molecular, (Å³); ^g Coeficiente de absorção de Ostwald gás-hexadecano; ^h Equilíbrio líquido-vapor; ⁱ Solubilidade aquosa; ^j Diferença de entalpia padrão parcial molar de solvatação, (J/mol);
^k Diferença de entropia padrão parcial molar de solvatação, (J/mol·K); ^l Calor padrão de formação, (kcal/mol).

Tabela 15. Equações lineares e não lineares (polinomial quadrática) para as propriedades físico-químicas e termodinâmicas estudadas e os respectivos parâmetros estatísticos, empregando o índice semi-empírico topológico (I_{ET}).

Propriedade	Equação de regressão	Parâmetros estatísticos			
		N ⁿ	r ² ^o	SD ^p	r ² _{cv} ^q
D ^a	$-0,915(\pm 70,005) - 1,88 \times 10^{-4}(\pm 4,4 \times 10^{-5}) I_{ET}^2 - 6,28 \times 10^{-3}(\pm 9,8 \times 10^{-4}) I_{ET}$	13	0,990	0,001	0,974
nD ^b	$-1,317(\pm 0,012) - 6,27 \times 10^{-4}(\pm 1,3 \times 10^{-4}) I_{ET}^2 + 1,83 \times 10^{-2}(\pm 2,6 \times 10^{-3}) I_{ET}$	27	0,978	0,004	0,948
PF ^c	$-256,9390(\pm 15,1450) - 0,6650(\pm 0,1260) I_{ET}^2 + 28,0478(\pm 2,9101) I_{ET}$	43	0,9865	6,1	0,985
PE ^d	$-150,6077(\pm 8,6323) - 1,0085(\pm 0,0834) I_{ET}^2 + 46,4810(\pm 1,7763) I_{ET}$	86	0,9971	4,2	0,997
log V _w ^e	$-0,5982(\pm 0,0430) + 0,3997(\pm 0,0053) I_{ET}$	13	1,0000	0,0158	0,999
v ^f	$-17,76(\pm 7,10) + 16,38(\pm 0,92) I_{ET}$	11	0,9940	2,83	0,991
log γ_{hd} ^g	$-0,89(\pm 0,40) + 0,62(\pm 0,06) I_{ET}$	11	0,9860	0,14	0,983
log ρ^{gas} ^h	$1,28(\pm 0,13) - 0,63(\pm 0,02) I_{ET}$	11	0,9980	0,05	0,997
log ρ^w ⁱ	$2,35(\pm 1,39) - 0,75(\pm 0,19) I_{ET}$	11	0,8968	0,48	0,866
$\Delta_{sol}H^0$ ^j	$6287(\pm 906) - 5379(\pm 110) I_{ET}$	13	1,0000	317	0,999
$\Delta_{sol}S^0$ ^k	$3,482(\pm 2,83) - 6,789(\pm 0,34) I_{ET}$	13	0,9940	0,990	0,992
ΔH_f^0 ^l	$-0,85(\pm 0,23) - 6,26(\pm 0,04) I_{ET}$	15	1,000	0,17	1,000
log P ^m	$-0,9841(\pm 0,2423) + 0,6072(\pm 0,0336) I_{ET}$	19	0,9871	0,13	0,984

^a Densidade, (g/cm³); ^b Índice de refração; ^c Ponto de fusão, (°C); ^d Ponto de ebulição normal, (°C); ^e Volume de retenção específico a 373 K, (cm³/g); ^f Volume molecular, (Å³); ^g Coeficiente de absorção de Ostwald gás-hexadecano; ^h Equilíbrio líquido-vapor; ⁱ Solubilidade aquosa; ^j Diferença de entalpia padrão parcial molar de solvatação, (J/mol); ^k Diferença de entropia padrão parcial molar de solvatação, (J/mol·K); ^l Calor padrão de formação, (kcal/mol); ^m Coeficiente de partição octanol/água; ⁿ Numero de compostos; ^o Coeficiente de determinação; ^p Desvio padrão; ^q Coeficiente de validação cruzada.

Como pode ser observado na Tabela 15, quatro propriedades resultaram em modelos de QSRP segundo uma equação não linear com o I_{ET} , a saber, densidade, índice de refração, ponto de fusão e ponto de ebulição. Informações relacionadas ao estado físico na tabela periódica mostram que os pontos de fusão e ebulição tendem a aumentar com a densidade, assim como o índice de refração também está relacionado com a densidade ótica. Esse comportamento pode indicar que este grupo de propriedades apresenta uma característica similar, a qual está contribuindo predominantemente para o valor dessas propriedades e não influencia nas propriedades relacionadas linearmente com o I_{ET} .

Densidade

A densidade, D , é uma propriedade físico-química freqüentemente utilizada para a caracterização e identificação de compostos e também pode ser usada para estimar a pressão crítica.¹¹ O modelo obtido segue uma equação polinomial quadrática com bons parâmetros estatísticos (Tabela 15) ($N=13$, $r^2=0,990$, $SD=0,001$, $r^2_{cv}=0,974$).

Índice de refração

O índice de refração, n_D , é considerado uma das propriedades óticas mais importantes e exibe boa correlação com propriedades ótica, elétrica e magnética, especialmente com a polarizabilidade.¹¹ O modelo QSPR desenvolvido com o I_{ET} também seguiu uma equação quadrática com parâmetros estatísticos levemente inferiores aos obtidos com a densidade, mas ainda aceitáveis para proposta de predição. Estas medições são realizadas em equipamentos chamados interferômetros e a incerteza da medição fica em torno de 10%, o que pode ter

contribuído para decréscimo dos resultados estatísticos (Tabela 15) (N=27, $r^2=0,978$, SD=0,004, $r^2_{cv}=0,948$).¹⁴⁴

Ponto de fusão

O ponto de fusão, PF, é um dos métodos de identificação e testes mais antigos para substâncias orgânicas. O ponto de fusão tem sido aplicado em ciências bioquímicas e ambientais, devido a sua relação com a solubilidade dos compostos.¹¹

Para modelar o ponto de fusão, PF, e o ponto de ebulição normal, PE, foi encontrado que a relação mais apropriada é um modelo não linear usando uma equação quadrática polinomial (Tabela 15). Para o PF os altos valores do coeficiente de determinação ($r^2 = 0,9865$) e o coeficiente de validação cruzada ($r^2_{cv}=0,985$) mostram a alta qualidade do modelo. Este modelo de ajuste explica 98,65% da variação total no PF obtido para 43 compostos sendo que 8, não incluídos no modelo inicial, foram prognosticados com esta equação e a correlação entre o PF calculado versus o PF experimental levou a bons resultados (N=43, $r^2 = 0,9903$; SD = 4,1). A transferência da fase sólida para líquida está relacionada a forma, tamanho, simetria relacionada ao empacotamento do cristal e interações intermoleculares de transferência de cargas e dipolo-dipolo.¹¹

Ponto de ebulição normal

Segundo Ren¹³⁸, o ponto de ebulição normal é uma propriedade física universal e precisamente medida para compostos de baixo peso molecular. Para moléculas neutras, o PE é o resultado de forças de atração dipolo-dipolo, de van der Waals e em algumas situações, das ligações de hidrogênio. Os haloalcanos não

têm habilidade para formar ligações de hidrogênio, portanto, é esperado que as forças de dispersão predominem entre as interações de van der Waals. A energia de interação de van der Waals decresce com o inverso da força de separação, portanto, é esperado que moléculas com arranjos de empacotamento mais denso apresentem valores de PE maiores em relação as moléculas menos densamente empacotadas.³ Como o I_{ET} explica satisfatoriamente o comportamento cromatográfico observado para esses compostos, e como o PE está intimamente relacionado com a separação cromatográfica, pode-se pressupor que essa propriedade será bem modelada pelo descritor semi-empírico topológico.

O modelo QSPR obtido para o ponto de ebulição experimental de 86 compostos mostrou muito bom coeficiente de determinação ($r^2 = 0,9971$) e capacidade preditiva ($r^2_{cv} = 0,997$) (Tabela 15). Neste estudo, o modelo QSPR obtido através de uma regressão não-linear, usando uma equação polinomial quadrática, pode explicar 99,71% da variação nos valores experimentais e tem um SD de 4,2. A partir da equação obtida para o PE, os valores teóricos para 24 compostos os quais não foram incluídos no modelo inicial foram calculados e então relacionados com os valores experimentais para os 24 compostos dando parâmetros estatísticos muito bons ($r^2 = 0,9931$; SD = 7,6), mostrando a estabilidade externa do modelo.

A fim de comparar o desempenho do I_{ET} com outros diferentes descritores moleculares, foi selecionado o trabalho de Balaban *et al.*¹¹⁵ para um conjunto de 532 alcanos halogenados com um a quatro carbonos na cadeia. O melhor modelo foi obtido para uma regressão com seis índices topológicos ($r^2=0,97$ e SD= 10,94). Para uma série de 44 halometanos com dois ou mais átomos de halogênios, a melhor correlação obtida para quatro parâmetros foi $r^2 = 0,99$ e SD = 5,26. Yasuhara *et al.*¹¹¹ correlacionaram o PE com o índice de retenção de 170 compostos halogenados e obtiveram $r = 0,996$. Kier e Hall¹¹² investigaram a

correlação entre pontos de ebulição de 24 alcanos monohalogenados (com dois a cinco carbonos na cadeia) com dois índices de conectividade molecular e obtiveram $r = 0,992$ e $SD = 4,8$, para 17 compostos com três a seis carbonos, a melhor correlação com três índices de conectividade molecular foi $r=0,998$ e $SD=2,57$.

Os químicos afirmam que na natureza a maior parte das reações ocorre em meio aquoso. No entanto, o comportamento dos átomos nas fases gasosa e sólida segue leis bem estabelecidas, enquanto que a fase líquida e superfícies merecem mais atenção. Portanto, nesse trabalho foram selecionadas seis propriedades que consideram o tamanho molecular e refletem as interações intermoleculares para demonstrar a capacidade do I_{ET} para prever essas importantes propriedades. As propriedades são: Volume de retenção específico a 373 K, volume molecular, coeficiente de absorção de Ostwald gás-hexadecano, equilíbrio líquido-vapor e solubilidade aquosa.

O segundo objetivo foi avaliar a aplicabilidade do modelo QSPR para prognosticar parâmetros termodinâmicos de solvatação e interações moleculares soluto-solvente. As propriedades são: Diferença de entalpia e entropia padrão parcial molar de solvatação e calor de formação.

Volume de retenção específico

O volume de retenção específico a 373 K, V_w , (cm^3/g) calculado para 13 compostos halogenados por Dallos¹¹⁷ foi usado no modelo QSPR para verificar a viabilidade do I_{ET} em modelar essa propriedade para 1-cloro e bromoalcanos baseado na estrutura molecular. Os resultados na Tabela 15 ($N=13$, $r^2=1,0000$, $SD=0,0158$, $r^2_{cv}=0,999$) demonstraram a excelente habilidade do modelo para

predizer o V_w para essa série de compostos. O V_w está relacionado com o coeficiente de atividade do soluto em diluição infinita.

Volume molecular

Em líquidos simples, a energia de ligação e o trabalho de interação de uma molécula com esse ambiente são linearmente relacionados com o volume molecular. No presente estudo, o volume molecular, v , (Å^3) foi modelado adequadamente pelo I_{ET} segundo os parâmetros estatísticos ($N=11$, $r^2=0,9940$, $SD=2,8325$, $r^2_{CV}=0,991$), indicando a sua capacidade em reter informação referente a efeitos estéreos como o potencial para prever dados termodinâmicos.

Coefficiente de absorção de Ostwald gás-hexadecano e equilíbrio líquido-vapor

O coeficiente de absorção de Ostwald é definido como a razão da densidade entre o líquido e o gás no equilíbrio, ambos em concentração molar.⁷⁸ A partição gás-hexadecano está diretamente relacionada a transferência gás-líquido e descreve o equilíbrio entre o líquido puro e o seu vapor (pressão de vapor) se esse é considerado simplesmente como a solvatação da molécula pelo seu próprio líquido.⁷⁸ A habilidade do modelo simular adequadamente essa propriedade é indicada pela qualidade dos parâmetros estatísticos para a partição gás-hexadecano, γ_{hd} , ($N=11$, $r^2=0,9860$, $SD=0,1399$, $r^2_{CV}=0,983$) e para o equilíbrio líquido-vapor, $\rho^{gás}$, ($N=11$, $r^2=0,9980$, $SD=0,0460$, $r^2_{CV}=0,997$). Buchwald e Bodor⁷⁸ usaram uma regressão linear simples para correlacionar a partição gás-hexadecano com o volume molecular de 49 compostos orgânicos com estruturas

simples e obtiveram parâmetros estatísticos inferiores aos obtidos nesse trabalho ($r^2=0,943$ e $SD=0,264$).

Solubilidade aquosa

Os modelos de QSPR são especialmente importantes em estudos ambientais, onde os compostos de interesse são tóxicos, carcinogênicos ou indesejáveis por alguma outra razão.¹⁴⁵

Uma vez que muitos compostos halogenados largamente utilizados em processos industriais têm baixa solubilidade em água, e constituem um sério problema em tratamento de águas residuais e disposição de lodo,¹⁴⁶ propostas de tratamento desses resíduos em meio não convencional, que assegurem a sua transformação em derivados menos agressivos ao ambiente, são indubitavelmente de grande utilidade.^{105,146} De acordo com os resultados obtidos ($N=11$, $r^2=0,8968$, $SD=0,481$, $r^2_{CV}=0,866$), aparentemente a solubilidade aquosa, $\log p^w$, não foi modelada adequadamente pelo I_{ET} . Nos compostos investigados, as forças de dispersão e orientação são dominantes e o I_{ET} foi desenvolvido para refletir esse ambiente. Porém, a habilidade da água em formar ligações de hidrogênio não tem semelhança com as forças intermoleculares dos halogênios estudados, por conseguinte, o modelo não foi capaz de refletir essas forças intermoleculares específicas em uma extensão razoável. Buchwald e Bodor⁷⁸ obtiveram melhores resultados ($r^2=0,966$ e $SD=0,221$) para o modelo baseado no tamanho molecular de 68 compostos que não possuíam substituintes fortemente polar ou formadores de ligações de hidrogênio (sem N ou O).

Diferença de entalpia e entropia padrão parcial molar de solvatação

O conhecimento de propriedades termodinâmicas de equilíbrio líquido-vapor pode fornecer uma boa base para o entendimento de reações em fase gasosa de compostos halogenados.

Para testar a capacidade do I_{ET} para prever parâmetros termodinâmicos para a solvatação de uma substância em coluna de baixa polaridade, propriedades de equilíbrio gás-líquido tal como a diferença de entalpia e entropia padrão molar parcial de solvatação foram consultadas na literatura.¹¹⁷

Pode-se observar na Tabela 15 que o I_{ET} apresenta correlações altamente significativas com a diferença de entalpia e entropia padrão molar parcial de solvatação. Para os modelos obtidos, a comparação entre o coeficiente de validação cruzada e de determinação (para $\Delta_{sol}H^{\circ}$ N=13, $r^2=1,0000$, $r^2_{CV} = 0,999$ e $\Delta_{sol}S^{\circ}$ N=13, = 0,9940, $r^2_{CV} =0,992$), indicam a alta estabilidade da equação de regressão. A investigação de Dallos¹¹⁷ mostrou que a dependência de $\Delta_{sol}H^{\circ}$ e $\Delta_{sol}S^{\circ}$ do tamanho molecular indica que as interações entre soluto e solvente são proporcional a polarizabilidade da molécula do soluto num solvente parafínico. Provavelmente esses fatores foram bem modelados pelo I_{ET} como único descritor.

Calor de formação

Atualmente é aceito que muitos halocarbonos clorados e alguns bromados possuem origem natural e também antropogênica. A produção natural de halocarbonos está fortemente relacionada à fotólise¹¹⁰ e mais do que 3500 halocompostos são produzidos biologicamente, alguns dos quais em grande quantidade.^{109,110} Uma vez que as reações químicas se caracterizam por certo arranjo de átomos nos reagentes e um diferente arranjo dos mesmos átomos nos

produtos, a adequada análise termodinâmica do sistema pode fornecer valiosas informações sobre as condições adequadas para a obtenção dos produtos desejados.¹⁴⁶ Nesse contexto, a habilidade para modelar teoricamente a termoquímica dos compostos halogenados apresenta-se como uma alternativa de grande importância devido a carência de informações experimentais, especialmente de moléculas contendo átomos de bromo e iodo.¹⁰⁸

As correlações lineares simples entre o índice semi-empírico topológico e o calor de formação para 1-halo-n-alcanos, mostradas na Figura 8, revelaram essencialmente linhas paralelas.

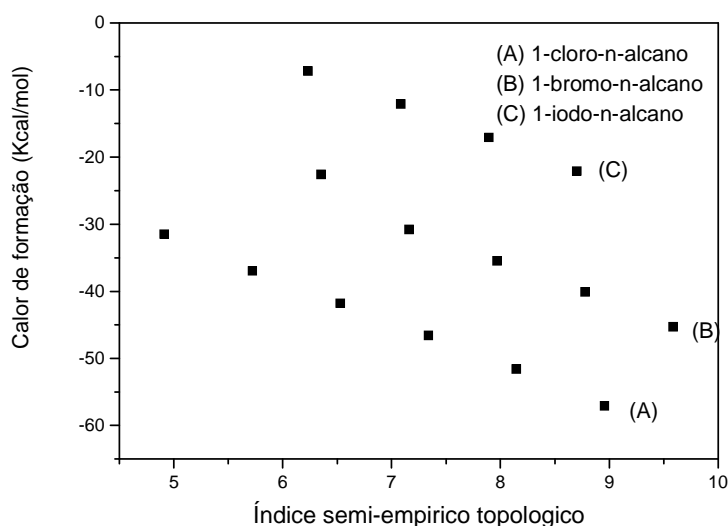


Figura 8. Calor de formação (ΔH_f°) como uma função do índice semi-empírico topológico para 1-halo-n-alcanos.

Como pode ser observado na Figura 8, a inclinação das curvas muda de acordo com o tamanho dos átomos. A Figura 8 também mostra que o calor de formação é dependente do tamanho dos átomos, uma vez que a inclinação de ΔH_f° versus o I_{ET} é essencialmente a mesma para as moléculas alifáticas contendo apenas um heteroátomo, seja Cl, Br ou I. Essa observação indica a similaridade das interações entre as estruturas desses compostos e permite a possibilidade de generalizar o modelo. Para ajustar todos os dados em apenas uma correlação linear simples (Figura 8) foi gerada uma nova equação entre o calor de formação e índice semi-empírico topológico modificado (I_{ET}^*). Para calcular o I_{ET}^* para os compostos substituídos com Br e I, foi determinada a inclinação média das 3 equações previamente encontradas e essa inclinação foi substituída nas respectivas equações. O valor do novo intercepto (β_1) usado em cada equação foi calculado subtraindo o intercepto obtido na equação dos compostos clorados menos o intercepto obtido inicialmente em cada equação. Dessa maneira, novos valores de I_{ET} , menores do que aqueles previamente encontrados para os compostos com Br e I, foram calculados e foi gerado um novo modelo (Figura 9). Os excelentes dados estatísticos entre o calor de formação e o I_{ET}^* ($N=15$, $r^2=1,0000$, $SD=0,1675$, $r^2_{CV}=1,0000$) demonstraram a força preditiva desse índice para estimar o ΔH_f° baseado na informação da estrutura molecular.

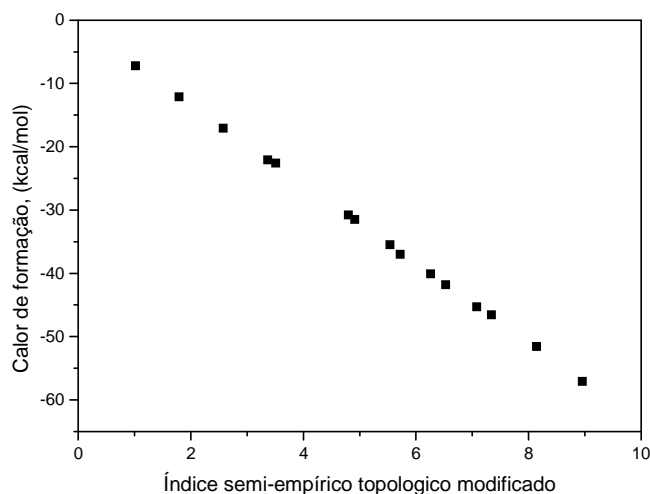


Figura 9. Calor de formação (ΔH_f°) como uma função do índice semi-empírico topológico modificado para 1-halo-n-alkanos.

Coeficiente de partição octanol/água

A natureza lipofílica e a persistência dos haloalcanos contribuem com a bioacumulação desses compostos nas membranas biológicas ou em ligações com biomacromoléculas.⁹⁰

O coeficiente de partição octanol/água, $\log P$, é um descritor universal do biotransporte via penetração da membrana celular⁸⁶ e tem recebido muita atenção em estudos de QSPR/QSAR.^{78,86,90,148} Nesse sentido, o $\log P$ foi a propriedade selecionada para este estudo e sua determinação tem sido recomendada pelo programa de evolução de compostos perigosos (OECD).¹⁴⁸ Os valores experimentais de $\log P$ para 19 compostos halogenados considerados nesse

estudo encontram-se na Tabela 4, sendo que dois desses compostos, os quais foram incluídos na construção do modelo inicial com 141 compostos, foram prognosticados a partir deste modelo. A concordância entre o coeficiente de determinação ($r^2=0,9871$) e o coeficiente de validação cruzada ($r^2_{cv}=0,984$) mostram a qualidade preditiva do modelo de QSPR (Tabela 14).

O Método Semi-Empírico Topológico foi comparado com dois outros métodos computacionais empregados para estimar teoricamente o log P: "Complete Neglect Of Differential Overlap" (CNDO) e de Contribuição dos Fragmentos.¹³⁴ Os valores de log P calculados a partir dos Métodos CNDO e de Contribuição dos Fragmentos estão listados na Tabela 4. Como pode ser observado na Tabela 4, o Método CNDO não pode ser adotado para o cálculo de compostos contendo átomos com número atômico maior do que 18 e o método de Contribuição dos Fragmentos têm um baixo grau de degenerescência e não discrimina vários isômeros de posição (isto é, compostos nº 5, 18, 23, 26 e 6, 19, 24, 27).

Os parâmetros estatísticos para as regressões lineares simples entre os valores de log P calculado e experimental para os três métodos citados estão apresentadas na Tabela 16. Os parâmetros estatísticos (r^2 e SD) para o Método Semi-Empírico Topológico mostraram uma melhor ordem quando comparados com os Métodos Contribuição dos Fragmentos e CNDO (Tabela 16). Nesse sentido, o I_{ET} pode ser considerado como uma ferramenta útil para calcular esta propriedade, atingindo alto grau de degenerescência e boa qualidade estatística. O log P de 39 haloalcanos, 48 compostos aromáticos e alquiaromáticos e 31 haloaromáticos (excluindo compostos capazes de formar ligações de hidrogênio) foram correlacionados com o volume molecular por Buchwar e Bodor⁷⁸. A equação apresentou $r^2 = 0,973$ e $SD = 0,242$.

Tabela 16. Parâmetros estatísticos para a correlação entre os valores de log P calculado versus experimental para o Método Semi-Empírico Topológico, Contribuição dos Fragmentos e CNDO.

Propriedade	Método	Parâmetros estatísticos			
		N^a	r^2^b	SD^c	$r^2_{cv}^d$
log P	Semi-Empírico Topológico	19	0,9871	0,13	0,984
log P	Contribuição dos Fragmentos	19	0,9750	0,17	0,971
log P	CNDO	10	0,8918	0,21	0,831

^a Número de compostos; ^b Coeficiente de determinação; ^c Desvio padrão; ^d Coeficiente de validação cruzada.

5.4 - Aplicação do I_{ET} em Correlações Quantitativas entre Estrutura-Atividade Biológica de Compostos Alifáticos Halogenados

O objetivo primordial das Correlações Quantitativas entre Estrutura-Atividade (QSAR) é desenvolver modelos com um grupo de compostos, denominado de grupo de treinamento, e utilizar esses modelos para o prognóstico da atividade biológica bem como a interpretação do mecanismo de ação de compostos estruturalmente relacionados.⁸⁵

Para testar a toxicidade de compostos químicos orgânicos, o método tradicional tem sido selecionar um grupo de compostos estruturalmente similares que contém um grupo funcional – isto é, álcool, éster, halogênio, etc. e histórico de toxicidade.

A contribuição do grupo funcional à reatividade química geralmente é constante e permite sugerir modelos de mecanismo de ação tóxica baseados no comportamento típico desses grupos nas reações químicas.

Por outro lado, o aumento da cadeia carbônica contribui diretamente para a hidrofobicidade (*biouptake*) e conseqüentemente para a toxicidade.

Desse modo, geralmente os modelos preditivos consistem em regressões múltiplas entre a toxicidade e propriedades que representam a hidrofobicidade e a reatividade.⁸⁶

Entre os bioensaios adotados para a avaliação da toxicidade está a irritação ocular, avaliação do potencial mutagênico, LC_{50} , que é a concentração de um composto que provoca a morte de 50% da população ensaiada e IGC_{50} , concentração do reagente que inibe 50% do crescimento da população.

O logaritmo da constante de partição octanol/água é o descritor molecular universal de hidrofobicidade.⁸⁶

Não existe um consenso sobre um descritor universal de interação molecular, mas descritores físico-químicos como a refratividade molar e volume molecular vêm sendo adotados como parâmetro de efeito estéreo-eletrônicos, enquanto que propriedades do orbital molecular (HOMO e LUMO) são considerados em mecanismos onde ocorre a formação de ligações covalente.^{86,87}

Idealmente os modelos de QSAR devem ser simples, facilmente interpretáveis e com mecanismos de ação compreensíveis.⁸⁹ Os fundamentos do

mecanismo de ação tóxica servem como guia primário para a escolha da classe de compostos e dos descritores mais promissores para estabelecer correlações significativas com o parâmetro de controle que representa a toxicidade.⁸⁸

A toxicidade pode ser reunida em três grandes categorias, as quais se subdividem em vários mecanismos de ação.^{88,149}

- Toxicidade não específica reversível, também denominada como narcose, narcose não reversível ou narcose não covalente;

- Toxicidade específica não reversível ou reatividade covalente, relacionada a habilidade de um compostos em formar ligações covalentes com o sistema biológico, devido a interações eletro(núcleo)fílicas com alguma macromolécula;

- Toxicidade funcional mediada pelo receptor, resultado de interações não covalentes, que ocorrem na interface segundo um mecanismo tipo chave-fechadura.

A maior parte dos reagentes orgânicos industriais apresenta toxicidade aguda segundo o mecanismo narcose.⁸⁶ Narcose é o resultado do efeito cumulativo de compostos nas células das membranas, o qual perturba suas funções normais pela interrupção das forças de van de Walls entre o lipídio e/ou proteínas que compõem a membrana. Narcose é o mecanismo de ação tóxico mais básico e largamente dependente da hidrofobicidade e, portanto, quantificado pelo log P. Esse modo de ação inclui vários mecanismos de ação, tal como narcose não polar ou narcose polar, entre outros.¹⁴⁹

Compostos alifáticos halogenados têm importantes usos industriais, no entanto, muitos haloalcanos são altamente tóxicos e uma ameaça ao meio ambiente e a saúde humana, uma vez que incluem compostos carcinogênicos e patogênicos.⁸⁷

Os haloalcanos tendem a ser lipossolúveis – eles são parcionados entre o sangue e o tecido adiposo, podendo ficar acumulados por um longo tempo. A ligação carbono-halogênio é polar e a polaridade desta ligação governa o mecanismo de ação dos halogênios, que podem ser de dois tipos: reações de substituição nucleofílica e formação de radicais livres. O carbono ao qual o halogênio está ligado é parcialmente positivo, tornando-o suscetível a reações de substituição nucleofílica, nas quais um nucleófilo substitui o átomo de halogênio segundo o mecanismo S_{N1} ou S_{N2} , onde S representa substituição, N representa nucleofílico e 1-2 indica a molecularidade da etapa determinante da velocidade de reação.¹⁵⁰ A eletrofilicidade também é um fator importante que determina a relativa propensão dos haloalcanos sofrerem transformações metabólicas redutivas e a geração de radicais livres. A quebra homolítica da ligação carbono-halogênio é a chave para a produção de radicais livres no metabolismo.⁸⁷ Conforme exposto até aqui, pode-se deduzir que os haloalcanos são compostos promissores para estabelecer modelos de QSAR.

Várias investigações tem mostrado que os haloalcanos atuam como narcose não polar, portanto, a discussão ficará limitada a esse tipo de toxicidade e será considerado que parâmetros que representam potencial reatividade covalente são desnecessários.⁸⁶

Objetivando adotar o método semi-empírico topológico em estudos de QSAR de compostos alifáticos halogenados e obter informações sobre o mecanismo de ação, a toxicidade foi expressa em termos de $\log LC_{50}$ e $\log IGC^{-1}_{50}$ e o I_{ET} foi adotado como o único descritor molecular. Os dados para os compostos e atividades avaliadas estão na Tabela 6. O I_{ET} está baseado na suposição de que o comportamento cromatográfico de uma molécula é predominantemente o resultado do número e interações dos átomos da molécula com a fase estacionária. Essas interações são determinadas por suas propriedades elétricas

e pelo impedimento estéreo dessas interações por outros átomos ligados a ele.¹⁰³ Portanto, o I_{ET} reúne o termo hidrofóbico e estéreo.

A Figura 10 mostra a relação entre a toxicidade para o *Tetrahymena pyriformis* ($\log IGC^{-1}_{50}$) e o I_{ET} . Embora esta apresente somente uma correlação com somente quatro pontos, o ajuste dos dados sobre a linha de base indica apenas um mecanismo de ação para esse grupo de compostos. Autores como Aptula¹⁵¹ e Akers *et al.*⁸⁶ também desenvolveram modelos de QSAR com apenas quatro pontos, evidenciando a falta de dados disponíveis na literatura e a necessidade de modelos preditivos.

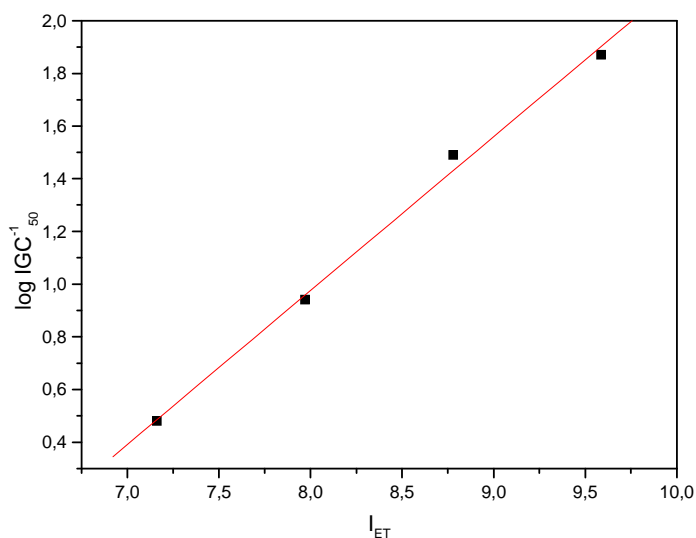


Figura 10. Correlação entre a toxicidade e o I_{ET} para 1- bromoalcanos.

Para os 1-bromoalcanos, uma excelente relação linear foi encontrada entre o $\log \text{IGC}^{-1}_{50}$ e o I_{ET} :

$$\log \text{IGC}^{-1}_{50} = -3,70(\pm 1,00) + 0,58(\pm 0,12) I_{\text{ET}}$$

$$N = 4; r^2 = 0,9960; SD = 0,05; r^2_{\text{cv}} = 0,982 \quad (15)$$

Os resultados estatísticos representam satisfatoriamente o modelo de QSAR e a inclinação positiva mostra a importância do tamanho da molécula para a toxicidade de 1-bromoalcanos. Para obter indícios sobre o mecanismo de ação, a colinearidade entre o $\log \text{IGC}^{-1}_{50}$ e $\log P$ ($r = -0,9806$) (Tabela 13) indicaram que esses compostos atuam com um mecanismo consistente com a narcose não polar.

Contudo, deve-se ressaltar que esse tipo de mecanismo é geral, e não foi possível obter mais informações sobre um tipo de mecanismo específico. Tendo em mente que a toxicidade das moléculas é governada pelas suas propriedades, as quais são determinadas pela sua estrutura química, deve-se reconhecer que o atual nível de conhecimento dessas relações não significa que o conhecimento da estrutura molecular sozinho é suficiente para prognosticar o mecanismo de ação e a potencialidade de um composto.

Outro fator que deve ser considerado é que os dados consultados na literatura foram obtidos em diferentes condições experimentais, o que contribui para o decréscimo estatístico do ajuste desses dados. Esforços consideráveis vêm sendo feitos para padronizar os ensaios ecotoxicológicos.⁸⁶

No entanto, Cronin e Schultz⁸⁵ (1998) mostraram que é possível desenvolver modelos de QSAR usando dados da literatura e obter, a partir da estrutura, informações sobre a toxicidade do modo de ação de compostos.

Dados de log EC₅₀ para a toxicidade aguda do *Photobacterium phosphoreum* estão disponíveis na literatura para um número limitado de compostos organoclorados halifáticos com estruturas mais diversificadas em relação aquelas usadas no modelo anterior e foram usados para estabelecer a correlação entre o log EC₅₀ com o I_{ET}:

$$\log EC_{50} = 5,3570(\pm 0,98) - 0,4805(\pm 0,15) I_{ET}$$
$$N = 5; r^2 = 0,9741; SD = 0,1292; r^2_{cv} = 0,945 \quad (16)$$

A análise de regressão revela a forte correlação entre o log EC₅₀ e o I_{ET} (r²=0,9741) e a inclinação negativa indica a importância do tamanho da molécula para a toxicidade de compostos orgânicos clorados mono e di-substituídos. É importante ressaltar que o I_{ET} codifica o tamanho da molécula, bem como as ramificações. O pequeno tamanho das moléculas aparentemente não favoreceu a penetração celular. Esta observação está de acordo com o modelo de QSAR para haloalcanos desenvolvido por Bláha e Némee⁹⁰, onde foi encontrado como parâmetro mais importante para a toxicidade o tamanho das moléculas. Moléculas maiores aparentemente são mais tóxicas quando comparadas com as menores. Segundo os autores, essa característica dos compostos pode estar relacionada ao acúmulo dos haloalcanos nas membranas biológicas ou ligações com biomacromoléculas, tais como proteínas, enzimas ou DNA.

Os valores para os parâmetros estatísticos obtidos com esse modelo são similares em ordem e qualidade a aqueles obtidos para o $\log \text{IGC}^{-1}_{50}$ com o I_{ET} , indicando que para alcanos halogenados, sejam clorados ou bromados, o mecanismo de ação depende do tamanho da molécula. Também se pode observar um decréscimo no coeficiente de correlação e no ajuste dos dados (Figura 11), indicando uma mudança no mecanismo de ação. A introdução de um segundo átomo de cloro na cadeia carbônica aumenta a probabilidade de alterações da reatividade química.

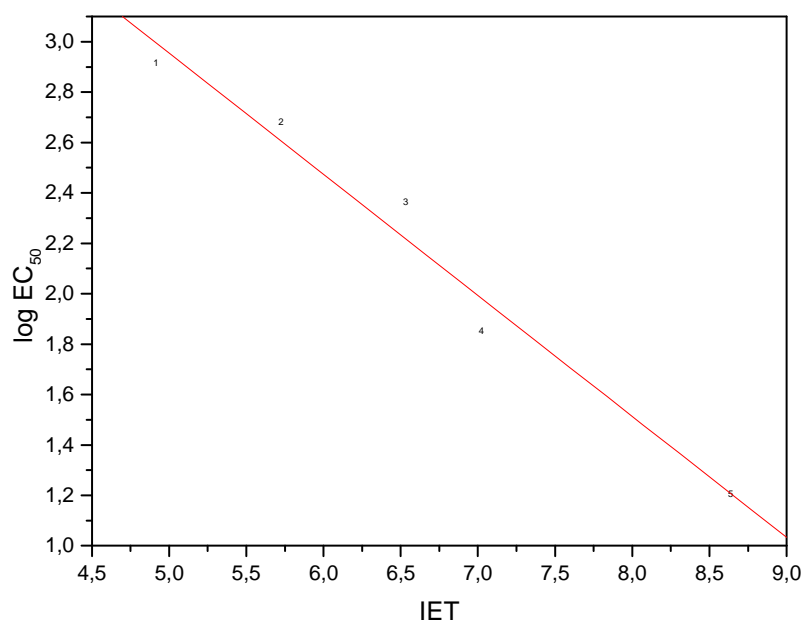


Figura 11. Correlação entre a toxicidade e o I_{ET} para compostos alifáticos clorados mono e di-substituídos.

6 - CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES PARA TRABALHOS FUTUROS

6.1 - Conclusões

A correlação entre os valores dos índices de retenção experimentais e I_{ET} para os compostos alifáticos halogenados mostrou a boa qualidade do modelo QSRR ($r^2 = 0,9995$ e $SD = 8$) e a alta capacidade preditiva dos índices de retenção para esses compostos usando o Método Semi-Empírico Topológico ($r^2_{cv} = 0,999$). O método também permitiu retirar informações sobre as características da molécula que estão influenciando no processo de retenção cromatográfica.

O modelo empregado para desenvolver o I_{ET} usou o índice de retenção de 141 compostos alifáticos halogenados e a validação externa foi executada usando as propriedades de outros compostos que não foram incluídos no modelo inicial.

A eficiência desse índice foi verificada através da alta qualidade dos modelos obtidos, demonstrando a viabilidade em adotar o I_{ET} na modelagem de importantes propriedades destes compostos, baseado na informação da estrutura molecular.

A habilidade de prognóstico do I_{ET} também foi verificada em fases estacionárias de diferentes polaridades através do modelo combinado incluindo um parâmetro (P_R) que representa a polaridade de cada fase. O ajuste dos IR_{Cal} em relação aos IR_{Exp} mostra a adequação do modelo ($r^2=0,9984$; $SD=18$).

Resultados satisfatórios também foram obtidos aplicando o I_{ET} para estimar a retenção cromatográfica de 48 cicloalcanos ($r^2=0,9905$; $SD=7$; $r^2_{cv}=0,997$). O

descriptor refletiu características estruturais, eletrônicas e geométricas das moléculas e permitiu a distinção entre isômeros cis/trans dos compostos estudados. Tais resultados mostraram o potencial do I_{ET} em modelos de QSRR para compostos que apresentam cadeia em anel.

O I_{ET} também foi aplicado satisfatoriamente em estudos de QSPR para diferentes propriedades determinadas por forças intermoleculares para halocompostos e cicloalcanos. Os resultados estatísticos representam a qualidade dos modelos e permitem prognosticar parâmetros termodinâmicos úteis para auxiliar na explicação das interações moleculares entre o soluto e a fase estacionária.

Finalmente, modelos de QSAR empregando o I_{ET} , foram testados para hidrocarbonetos alifáticos halogenados e duas atividades biológicas, $\log IGC^{-1}_{50}$ (concentração que inibe 50% do crescimento dessa população) e $\log EC_{50}$ (inibição da bioluminescência de *Photobacterium phosphoreum*, expressa como concentração efetiva). Os resultados estatísticos representam satisfatoriamente os modelos QSAR ($r^2=0,9960$; $r^2_{cv}=0,982$ e $r^2=0,9741$; $r^2_{cv}=0,945$, respectivamente) revelando o potencial de aplicação do I_{ET} em estudos de estrutura-atividade.

Os resultados obtidos indicam que o I_{ET} pode ser empregado para prognosticar a retenção cromatográfica dos grupos de compostos investigados, podendo ser utilizado como um método complementar para auxiliar a identificação destes. Este descritor molecular mostrou ter potencial suficiente para ser aplicado em estudos de QSPR/QSAR para prognosticar outras propriedades físico-químicas e atividades biológicas relacionadas.

A maior vantagem em adotar o Método Semi-Empírico Topológico consiste na possibilidade de prognosticar a retenção cromatográfica, um importante parâmetro experimental largamente utilizado em química analítica. O método

adota um único descritor em uma correlação linear simples e os modelos QSRR/QSPR/QSAR desenvolvidos apresentam alta qualidade estatística. Portanto, os resultados obtidos neste trabalho demonstram o potencial do I_{ET} para ser utilizado como uma ferramenta auxiliar na caracterização e identificação de compostos alifáticos halogenados e cicloalcanos. Também é importante salientar que o índice de retenção permitiu obter informações sobre a estrutura molecular e o mecanismo de retenção dos compostos estudados.

6. 2 – Recomendações para Trabalhos Futuros

Os resultados encorajam a aplicação do índice semi-empírico topológico em futuros estudos de QSAR para grupos de compostos orgânicos que apresentem propriedades/atividades de interesse para a química e demais áreas.

7 - REFERÊNCIAS

1. KATRITZKY, A. R., GORDEEVA, E. V. Traditional topological indices vs electronic, geometrical, and combined molecular descriptors in QSAR/QSPR research. **J. Chem. Inf. Comput. Sci.**, v. 33, p.835-857, 1993.
2. NELSON, S. D., SEYBOLD, P. G. Molecular structure-property relationships for alkenes, Molecular structure-property relationships for alkenes. **J. Mol. Graphics Modell.**, v. 20, 36-53, 2001.
3. SEYBOLD, P. G., MAY, M.; BAGAL, U. A. Molecular structure-property relationships. **J. Chem. Educ.**, v.64, p.575-582, 1987.
4. RANDIC, M. The structural origin of chromatographic retention data. **Journal of Chromatography**, v. 161, p.1-14, 1978.
5. RANDIC, M. Generalized molecular descriptors. **J. Math. Chem.**, v. 7, p.155-168, 1991.
6. MIHALIC, Z., TRINAJSTIC, N. A graph-theoretical approach to structure-property relationships. **J. Chem. Educ.**, v. 69, p.701-712, 1992.

7. GAUDIO, A. C. Modelos clássicos de estudo quantitativo das relações entre estrutura química e atividade biológica. *Química Nova*, v. 19, p.278-289, 1996.
8. ROMANELLI, G. P., MARTINO, C. M., CASTRO, E. A. Modelling the solubility of aliphatic alcohols via molecular descriptors, 2001.
Disponível em: <http://preprint.chemweb.com/physchem/0101002>.
Acesso em: 15/09/2002.
9. HANSEN, P. J., JURIS, P. C. Chemical applications of graph theory. *J Chem. Educ.*, v. 65, n. 7, p.574-580, 1988.
10. FERREIRA, M. M. Multivariate QSAR. *J. Braz. Chem. Soc.*, v. 13, p.742-753, 2002.
11. KATRITZKY, A. R.; MARAN, V., LOBANOV, V.C., KARELSON, M. Structurally diverse quantitative structure-property relationship correlation of technologically relevant physical properties. *J. Chem. Inf. Comput. Sci.*, v. 40, p.1-18, 2000.
12. KALISZAN, R. **Quantitative Structure-Chromatographic Retention Relationship**. 1^o ed., John Wiley & Sons, New York, 1987.

13. ROMANELLI, G. P., CAFFERATA, L. F. R., CASTRO, E. A. An improved QSAR study of toxicity of saturated alcohols. **J. Mol. Struct. (Theochem)**, v. 504, p.261-265, 2000.
14. JUNKES, B. S.; et al, Semi-empirical topological index: Application to QSPR/QSAR modeling. **Journal of Molecular Modeling**, v. 11, p. 128-134, 2005.
15. KIER, L. B., HALL, L. H. Structure-Activity studies on odor molecules using molecular connectivity. **J. Theor. Biol.**, v. 67, p.585-595, 1977.
16. FERREIRA, M. M. C. Polycyclic aromatic hydrocarbons: a QSPR study. **Chemosphere**, v. 44, p. 125-146, 2001.
17. HÉBERGER, K. Quantitative structure-(chromatographic) retention relationships. **J. Chromatogr. A.**, v. 1158, 273-305, 2007.
18. BALABAN, A. T. Can topological indices transmit information on properties but not on structures? **J. Comput.-Aided Mol. Design**, v. 19, p.651-660, 2005.
19. IVANCIUC, O., IVANCIUC, T., BALABAN, A. T. Quantitative structure-property relationship study of normal boiling for halogen-/oxygen-/sulfur-containing organic compounds using the CONDESSA program. **TETRAHEDRON**, v. 54, p.9129-9142, 1998.

20. MILLER, M. D. et al. Quantum mechanical structure-activity relationship analyses for skin sensitization. *J. Chem. Inf. Model.*, v. 45, p. 924-29, 2005.
21. HANSCH, C., FUJITA, T. ρ , σ , π Analysis. A method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc.*, v. 86, p.1616-26, 1964.
22. KATRITZKY, A. R., FARA, D. C. How chemical structure determines physical, chemical and technological properties: An overview illustrating the potential of quantitative structure-property relationships for fuels science. *Energy & Fuels*, v. 19, p.922-935, 2005.
23. HANSCH, C., LEO, A., HOEKMAN., H. *Exploring QSAR*, Vol. I e II, ACS Professional Reference Book, New York, 1995.
24. KUBINYI, H. **QSAR: Hansch Analysis and Related Approaches**. Vol. 1 in *Methods and Principles in Medicinal Chemistry*, VCH, Weinheim, 1993.
25. KATRITZKY, A. R. et al. QSPR Correlation and predictions of GC retention index for methyl-branched hydrocarbons produced by insects. *Anal. Chem.*, v. 72, p. 101-109, 2000.

26. POMPE, M., NOVIC, M. Prediction of gas-chromatographic retention indices using topological descriptors. *J. Chem. Inf. Comput. Sci.*, v. 39, p.59-67, 1999.
27. GAUDIO, A. C., ZANDONADE, E. Proposição, validação e análise dos modelos que correlacionam estrutura química e atividade biológica. *Química Nova*, v. 24, n. 5, p.658-671, 2001.
28. DAS NEVES, P. J., DA COSTA, J. B. N., NDIYAE, P. M. TOP – Um programa de cálculo de descritores topológicos para uso em correlações entre estrutura e atividade. *Química Nova*, v. 21, p.709-713, 1998.
29. KATRITZKY, A. R. et al. A general treatment of solubility. 3. Principal component analysis (PCA) of the solubility of diverse solutes in diverse solvents. *J. Chem. Inf. Model.*, v. 45, p. 913-923, 2005.
30. KARELSON, M., LOBANOV, V. S., KATRITZKY, A. R. Quantum-Chemical descriptors in QSAR/QSPR studies. *Chem. Rev.*, v. 96, p.1027-1043, 1996.
31. TODESCHINI, R., CAZAR, R., COLLINA, E. The chemical meaning of topological indices. *Chemom. Intell. Lab. Syst.*, v. 15, p.51-59, 1992.

32. STANKEVICH, I. V., SKOVORTSOVA, M. I., ZEFIROV, N. S. On a quantum chemical interpretation of molecular connectivity indices for conjugated hydrocarbons. **J. Mol. Struct. (Theochem)**, v. 342, p.173-179, 1995.
33. ICARDO, M. C. et al. Molecular connectivity as a relevant new tool for predicting analytical behavior: A survey of chemiluminescence and chromatography. **Trends in Analytical Chemistry**, v. 24, n. 8, 2005.
34. ESTRADA, E., MOLINA, E. Novel local (fragment-based) topological molecular descriptors for QSPR/QSAR and molecular design. **J. Mol. Graphics Modell.**, v.20, 54-64, 2001.
35. KATRITZKY, A. R., PETRUKHIN, R., TATHAM, D. Interpretation of quantitative structure-property and –activity relationships. **J. Chem. Inf. Comput. Sci.**, v. 41, p. 679-685, 2001.
36. RANDIC, M. The connectivity index 25 years after. **J. Mol. Graphics Modell.**, v.20, p.19-35, 2001.
37. POGLIANI, L. From molecular connectivity indices to semi empirical connectivity terms: recent trends in graph theoretical descriptors. **Chem. Rev.**, v. 100, p.3827-3858, 2000.

38. HALL, L., H., KIER, L. B. Issues in representation of molecular structure: the development of molecular connectivity. *J. Mol. Graphics Modell.*, v. 20, p.4-18, 2001.
39. BASAK, S. C., NIKOLIC, S., TRINAJSTIC, N. QSPR modeling: Graph connectivity indices versus line graph connectivity indices. *J. Inf. Comput. Sci.*, v. 40, p.927-933, 2000.
40. LABANOWSKY, J. K., MOTOC, I., DAMMKOEHLER, R. A. The physical meaning of topological indices. *Comput. Chem.*, v. 15, p.47-53, 1991.
41. ESTRADA, E. Edge adjacency relationships in molecular graphs containing heteroatoms: A new topological index related to molar volume. *J. Chem. Inf. Comput. Sci.*, v. 35, p. 701-707, 1995.
42. DUCHOWICZ, P., CASTRO, E. A. QSPR modeling of normal boiling point of aldehydes, ketones, and esters by means of nearest neighboring codes correlation weighting, 2002. Disponível em: <http://preprint.chemweb.com/physchem/0203009>. Acesso em: 25/04/2002.
43. IVANCIUC, O., IVANCIUC, T., CARBOL-BASS, D., BALABAN, A. T. Optimum structural descriptors derived from Ivanciuc-Balaban operator. *Internet Electron. J. Mol. Des.*, v. 1, p. 319-331, 2002. Disponível em: <http://www.biochempress.com>. Acesso em: 10/10/2002.

44. REN, B. Application of novel atom-type AI topological indices to QSPR studies of alkanes. **Comput. & Chem.**, v. 26, p.357-369, 2002.
45. REN, B. Novel atom type al indices for QSPR studies of alcohols. **Comput. & Chem.**, v. 26, p.223-235, 2002.
46. REN, B. Application of novel atom-type al topological indices in the structure-property correlations. **J. Mol. Struct. (Theochem)**, v. 586, p.137-148, 2002.
47. LIU, S., CAI, S., CAO, C., LI, Z. Molecular electronegative distance vector (MEDV) related to 15 properties of alkanes. **J. Inf. Comput. Sci.**, v. 40, p.1337-1348, 2000.
48. WESSEL, M. D., SUTTER, J. M., JURIS, P. C. Prediction of reduced ion mobility constants of organic compounds from molecular structure. **Anal. Chem.**, v. 68, p.4237-4243, 1996.
49. TODESCHINI, R., CONSONNI, V., **Handbook of molecular descriptors**. VCH, Wiley, 2000.
50. RIBEIRO, F. A. L. **Aplicação de métodos de análise multivariada no estudo de hidrocarbonetos policíclicos aromáticos**. 2001. Tese

(Mestre em Química) – Programa de Pós-Graduação em Físico-Química, UNICAMP, Campinas.

51. ANKER, L. S., JURIS, P. C. Quantitative structure-retention relationship studies of odor-active aliphatic compounds with oxygen-containing functional groups. *Anal. Chem.*, v. 62, p.2676-2684, 1990.
52. BEEBE, K. R., PELL, R. J., SEASHOLTZ, M. B. **Chemometrics: a practical guide**. 1^o ed., John Wiley & Sons, New York, 1998.
53. LIVINGSTONE, D. **Data Analysis for Chemists. Applications to QSAR and Chemical Product Design**. Oxford University Press Inc., New York, 1995.
54. IVANCIUC, O., IVANCIUC, T., BALABAN, A. T. Quantitative structure-property relationship for the normal boiling temperature of acyclic carbonyl compounds. *Internet Electron. J. Molec. Des.*, v. 1, 2002, p.252-268. Disponível em: <http://www.biochempress.com>. Acesso em: 21/09/2002.
55. BONCHEV, D. Topological order in molecules 1. Molecular branching revisited. *J. Mol. Struct. (Theochem)*, v. 366, p. 137-156, 1995.

56. HALL, L. H., MOHNEY, B., KIER, L. B. The electrotopological state: structure information at the atomic level for molecular graphs. **J. Chem. Inf. Comput. Sci.**, v. 31, p.76-82, 1991.
57. GUEVARA, N. Fragmental graphs. A novel approach to generate a new family of descriptors. Applications to QSPR studies. **J. Mol. Struct. (Theochem)**, v. 493, p.29-36, 1999.
58. REN, B. Novel atomic-level-based AI topological descriptors: Applications to QSPR/QSAR modeling. **J. Chem. Inf. Comput. Sci.**, v. 42, p. 858-868, 2002.
59. ESTRADA, E. The structural interpretation of randic index. **Internet Electron. J. Molec. Des.**, v. 1, p.360-366, 2002. Disponível em: <http://www.biochempress.com>. Acesso em: 17/10/2002.
60. RANDIC, M. Chemical structure-what is "she"? **Symposium on Graph Theory in Chemistry**, v. 69, n. 9, p. 713-718, 1992.
61. CAO, C., YUAN, H. A modified distance matrix to distinguish cis/trans isomers of cicloalkanes. **Internet Electron. J. Mol. Des.**, v. 1, p. 319-331, 2002. Disponível em: <http://www.biochempress.com>. Acesso em: 10/11/2002.
62. HEINZEN, V. E. F., SOARES, M. F., YUNES, R. A. Semi-empirical topological method for prediction of the chromatographic retention of

- cis*- and *trans*- alkene isomers and alkanes. **J. Chromatogr. A**, v. 849, p. 495-506, 1999.
63. RIBEIRO, F. A. L., FERREIRA, M. M. C. QSAR modelo f the phototoxicity of polycyclic aromatic hydrocarbons. **J. Mol. Struct. (Theochem)**, v. 719, p.191-200, 2005.
64. BORGES, E. G. **Estudo das relações entre a estrutura química e a atividade biológica em compostos anti-úlceras**. 1998. Tese (Mestrado em Química) – Programa de Pós-Graduação em Físico-Química, UNICAMP, Campinas.
65. WOLOSZYN, T. F., JURIS, P. C. Prediction of gas chromatographic retention data for hydrocarbons from naphtas. **Anal. Chem.**, v. 65, p.582-7, 1993.
66. PENG, C. T. Retrieval of structure information from retention index. **J. Chromatogr. A**, v. 678, 189-200, 1994.
67. CIOLA, R. **Introdução à cromatografia em fase gasosa**. Edgard Blücher, São Paulo, 1973.
68. HARANGI, J. Retention index calculation without *n*-alkanes-the virtual carbon number. **J. Chromatogr. A**, v. 993, 187-195, 2003.
69. SKRBIÉ, B. D. Unified retention concept – statistical treatment of Kováts retention index. **J. Chromatogr. A**, v. 765, p. 257-64, 1997.

70. ZHANG, X., LU, P. Unified equation between Kováts indices on different stationary phases for select types of compounds. **J. Chromatogr. A**, v. 731, p. 187-199, 1996.
71. KALISZAN, R. QSRR Quantitative structure-(chromatographic) retention relationships. **Chem. Rev.**, v. 107, n. 7, p.3212-18, 2007.
72. CHEN, S., LIU, H., WANG, Z. Study of quantitative structure-retention relationship (QSRR) of gas chromatography for polychlorinated dibenzothiophenes on non-polar columns. **QSAR Comb. Sci.**, v. 26, n.8, p.889-96, 2007.
73. REN, Y., LIU, H., YAO, X., LIU, M. An accurate QSRR model for the prediction of the GCxGC-TOFMS retention time of polychlorinated biphenyl (PCB) congeners. **Anal. Bioanal. Chem.**, v. 388, n. 1, p.165-172, 2007.
74. SOARES, M. F., BOING, C. R., HEINZEN, V. E. F., YUNES R. A. Aplicação da teoria de QSRR na interpretação da retenção cromatográfica de uma série de acetofenonas. **Anais Associação Brasileira de Química**, v. 49, p.24-30, 2000.
75. JUNKES, B. S. et al. Application of novel semi-empirical topological index in the QSRR of aliphatic ketones and aldehydes on stationary phases of different polarity. **J. Braz. Chem. Soc.**, v. 15, p. 183-189, 2004.

76. LIU, F., LIANG, Y., CAO, C., ZHOU, N. QSPR study of CG retention indices for saturated esters on seven stationary phases based on novel topological indices. *Talanta*, v. 72, n. 4, p.1307-15, 2007.
77. IVANCIUC, T., IVANCIUC, O. Quantitative structure-retention relationship study of gas chromatographic retention indices for halogenated compounds. *Internet Electron. J. Mol. Des.*, v. 1, p.94-107, 2002. Disponível em: <http://www.biochempress.com>. Acesso em: 29/04/2002.
78. BUCHWALD, P., BODOR, N. Simple model for nonassociative organic liquids and water. *J. Am. Chem. Soc.*, v. 122, p. 10671-10679, 2000.
79. DIMOV N., MOSKOVKINA, M. New aspects of quantitative structure-retention relationships in chromatography. *Chromatographia*, v. 552, p. 59-69, 1991.
80. CHENG, A., MERZ Jr., K. M. M. Prediction of aqueous solubility of a diverse set of compounds using quantitative structure-property relationships. *J. Med. Chem.*, v. 46, p. 3572-3580, 2003.
81. WALKER, J. D. Applications of QSARs in toxicology: a US Government perspective, *J. Mol. Struct. (Theochem)*. v. 622, p.167-184, 2003.

82. SCHULTZ, T. W., CRONIN, M. T. D., NETZEVA, T. I. The present status of QSAR in toxicology, *J. Mol. Struct. (Theochem)*, v. 622, p.23-38, 2003.
83. LEITÃO, A., MONTANARI, C. A., DONNICI, C. L. Sobre o uso de métodos quimiométricos em QSPR. *Química Nova*, v. 23, n. 2, p. 178-184, 2000.
84. BRADBURY, S. P. Quantitative structure-activity relationships and ecological risk assessment: an overview of predictive aquatic toxicology research. *Toxicol. Lett.*, v. 79, p.229-237, 1995.
85. CRONIN, M. T. D., SCHULTZ, T. W. Structure-toxicology relationships for three mechanisms of action of toxicology to *Vibrio fischeri*. *Ecotoxicol. Environ. Saf.*, v. 39, p. 65-69, 1998.
86. AKERS, K. S., SINKS, G. D., SCHULTZ, T. W. Structure-toxicity relationships for selected halogenated aliphatic chemicals. *Environ. Toxicol. Pharmacol.*, v. 7, p. 33-39, 1999.
87. TROHALAKI, S., GIFFORD, E., PACHTER, R. Improved QSARs for predictive toxicology of halogenated hydrocarbons, *Comput. Chem.*, v. 24, p.421-427, 2000.

88. SCHULTZ, T. W., CRONIN, M. T. D., WALKER, J. D., APTULA, A. O., The present status of QSAR in toxicology, **J. Mol. Struct. (Theochem)**. v. 622, p.1-22, 2003.
89. CRONIN, M. T. D., SCHULTZ, T. W. Pitfalls in QSAR, **J. Mol. Struct. (Theochem)**. v. 622, p.39-51, 2003.
90. BLÁHA, L., DAMBORSKY, J., NÉMEE, M. QSAR for acute toxicity of saturated and unsaturated halogenated aliphatic compounds. **Chemosphere**, v. 36, n. 6, p. 1345-1355, 1998.
91. HEINZEN, V. E. F., YUNES, R. A. Relationship between gas chromatographic retention indices and molecular connectivity indices of chlorinated pesticides and structurally related compounds. **J. Chromatogr. A.**, v. 598, p.243-250, 1992.
92. ARRUDA, A. C., HEINZEN, V. E. F., YUNES, R. A. Relationship between Kováts retention indices and molecular connectivity indices of tetralonas, coumarins and structurally related compounds. **J. Chromatogr. A.**, v. 630, p.251-256, 1993.
93. HEINZEN, V. E. F., YUNES R. A. Correlation between gas chromatographic retention indices of linear alkylbenzene isomers and molecular connectivity indices. **J. Chromatogr. A.**, v. 654, p.83-89, 1993.

94. HEINZEN, V. E. F., YUNES, R. A. Using topological indices in the prediction of chromatographic retention indices of linear Alkylbenzene isomers. **J. Chromatogr. A.**, v. 719, p.462-467, 1996.
95. SOARES, M. F., MONACHE, F. D., HEINZEN, V. E. F., YUNES, R. A. Prediction of gas chromatographic retention indices of coumarins. **J. Braz. Chem. Soc.**, v. 10, p.189-196, 1999.
96. HEINZEN, V. E. F., CECHINEL FILHO, V., YUNES, R. A. Correlation of activity of 2-(X-benzyloxy)-4,6-dimethoxyacetophenones with topological indices and with the Hansch equation. **IL Farmaco**, v. 54, p.125-129, 1999.
97. AMBONI, R. D. M. C., JUNKES, B. S., YUNES, R. A., HEINZEN, V. E. F. Quantitative structure-odor relationships of aliphatic esters using topological indices. **J. Agric. Food Chem.**, v. 48, p.3517-3521, 2000.
98. AMBONI, R. **Estudo de correlação quantitativa entre estrutura e propriedade (QSPR) usando descritores topológicos para compostos carbonílicos alifáticos.** 2001. Tese (Doutorado em Química) – Curso de Pós-Graduação em Química, UFSC, Florianópolis.
99. JUNKES, B. S., AMBONI, R. D. M. C., YUNES, R. A., HEINZEN, V. E. F. Quantitative structure-retention relationships (QSRR), using the optimum

semi-empirical topological index, for methyl-branched alkanes produced by insects. *Chromatographia*, v. 55, p.707-715, 2002.

100. JUNKES, B. S., AMBONI, R. D. M. C., HEINZEN, V. E. F., YUNES, R. A. Use of a semi-empirical topological method to predict the chromatographic retention of branched alkenes. *Chromatographia*, v. 55, p.75-81, 2002.
101. AMBONI, R. D. M. C., JUNKES, B. S., YUNES, R. A., HEINZEN, V. E. F. Semi-empirical topological method for prediction of the chromatographic retention of esters. *J. Mol. Struct. (Theochem)*, v.579, p.53-62, 2002.
102. JUNKES, B. S., AMBONI, R. D. M. C., YUNES, R. A., HEINZEN, V. E. F. Prediction of chromatographic retention of saturated alcohols on stationary phases of different polarity applying the novel semi-empirical topological index. *Anal. Chim. Acta.*, v. 477, p.29-39, 2003.
103. JUNKES, B. S., AMBONI, R. D. M. C., YUNES, R. A., HEINZEN, V. E. F. Semi-Empirical topological index: a novel molecular descriptor for quantitative structure-retention relationship studies. *Internet Electron. J. Molec. Des.*, v. 2, p. 33-49, 2003. Disponível em: <<http://www.biochempress.com>>. Acesso em: 23/07/2004.
104. AMBONI, R. D. M. C., JUNKES, B. S., YUNES, R. A., HEINZEN, V. E. F. Quantitative structure-property relationship study of chromatographic

retention indices and normal boiling points for oxo compounds using the semi-empirical topological method. **J. Mol. Struct. (Theochem)**, v. 586, p.71-80, 2002.

105. ERABLE, B., MAUGARD, T., GOUBET, I., LAMARE, S., LEGOY, M. D. Biotransformation of halogenated compounds by lyophilized cells of *Rhodococcus erythropolis* in a continuous solid-gas biofilter. **Process Biochem.**, v. 40, p.45-51, 2005.
106. ZHAO, H., CHEN, J., QUAN, X., YANG, F., PEIJNENBURG, W. J. G. M. Quantitative structure-property relationship study on reductive dehalogenation of selected halogenated aliphatic hydrocarbons in sediment slurries. **Chemosphere**, v. 44, p.1557-63, 2001.
107. GROSSI, V., RAPHEL, D. Long-chain (C₁₉-C₂₉) 1-chloro-*n*-alkanes in leaf waxes of halophytes of the Chenopodiaceae. **Phytochemistry**, v. 63, p.693-8, 2003.
108. FELLER, D., PETERSON, K. A., JONG, W. A., DIXON, A. D. Performance of coupled cluster theory in thermochemical calculations of small halogenated compounds. **J. Chem. Phys.**, v. 118, n. 8, p.3510-22, 2003.
109. van PÉE, K., UNVERSUCHT, S. Biological dehalogenation and halogenation reactions. **Chemosphere**, v. 52, p.299-312, 2003.

110. ABRAHAMSSON, K., BERTILSSON, S., CHIERICI, M., FRANSSON, A., FRONEMAN, P. W., LORÉN, A., PAKHOMOV, E. A. Variations of biochemical parameters along a transect in the southern ocean, with special emphasis on volatile halogenated organic compounds. *Deep-Sea Research II*, v. 51, p.2745-56, 2004.
111. YASUHARA, A., MORITA, M., FUWA, K. Temperature-programmed retention indices of 221 halogenated organic compounds with 1-bromoalkanes as references. *Journal of Chromatography*, v. 328, p. 35-48, 1985.
112. KIER, L. B., HALL, L. H. *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.
113. KIER, L. B., HALL, L. H. *Molecular Connectivity in Structure-activity analysis*, Wiley, New York, 1986.
114. SABLJIC, A. Calculation of retention indices by molecular topology chlorinated alkanes. *Journal of Chromatography*, v. 314, p. 1-12, 1984.
115. BALABAN A. T., JOSHI N., KIER L. B., HALL L. H. Correlations between chemical structure and normal boiling points of halogenated alkanes C1-C4. *J. Chem. Inf. Comput. Sci.*, v. 32, p.233-7, 1992.

116. GARKANI-NEJAD, Z. et al. Prediction of gas chromatographic retention indices of a diverse set of toxicologically relevant compounds. **J. Chromatogr. A**, v. 1028, p. 287-295, 2004.
117. DALLOS, A., KRESZ, R., KOVÁTS, E. Solvation properties and limiting activity coefficients of halogenated hydrocarbons in C₇₈H₁₅₈ branched saturated alkane solvent. **Fluid Phase Equilib.**, v. 210, p.57-67, 2003.
118. LUO, Y., PACEY, P. D., Effects of alkyl substitution on ionization energies of alkanes and haloalkanes and on heats of formation of their molecular cations Part 2.* Alkanes and chloro-, bromo- and iodoalkanes, **International Journal of Mass Spectrometry and Ion Processes**, v. 112, p.63-77, 1992.
119. SOLOMONS, T. W. G. **Química Orgânica**. Vol I, 6^a ed., Livros Técnicos e Científicos, Rio de Janeiro, 1996.
120. MORRISON, R., BOYD, R. **Química Orgânica**. 6^a ed., Fundação Calouste Gulbenkian, , Lisboa, 1972.
121. CETESB, Companhia de Tecnologia de Saneamento Ambiental. **Acidentes Ambientais**. Disponível em: http://www.cetesb.sp.gov.br/emergencia/acidentes/vazamento/oleo/a_fisicos.asp Acesso em 15/12/2007.

122. RAMONI, M. C. **Estudos da fotodegradação de petróleo por ressonância paramagnética eletrônica (RPE)**. 2000. Tese (Mestrado em Física), Curso de Pós-Graduação em Física, UEL, Londrina.
123. von MÜHLEN, C., ZINI, C. A., CARAMÃO, E. B., MARRIOTT, P. J. Caracterização de amostras petroquímicas e derivados utilizando cromatografia gasosa bidimensional abrangente (GCxGC). **Química Nova**, v. 29, n. 4, p.765-75, 2006.
124. DIMOV, N., PAPAZOVA, D., Calculation of the retention indices of C₅-C₉ cycloalkanes on squalane. **Journal of Chromatography**, v.148, p.11-5, 1978.
125. ESTEL, D., MOHNKE, M., BIERMANS, F., ROTZSCHE, H. The analysis of C₄-C₁₁ hydrocarbons in naphtha and reformat with a new apolar fused silica column. **J. High Resol. Chromatogr.**, v. 18, p.403-11, 1995.
126. **The Sadtler standard gas chromatography retention index library**. Sadtler Research Laboratories, Philadelphia, PA, 1985.
127. CASTELLO, G. Gas chromatographic identification of alkyl radicals formed in plasma radiofrequency discharges by using iodine as a scavenger. **Journal of Chromatography**, v. 303, p.61-66, 1984.

128. CASTELLO, G., D'AMATO, G. Gas chromatography separation and identification of linear and branched-chain alkyl bromides. *Journal of Chromatography*, v. 324, p.363-371, 1985.
129. ENGEWALD, W., EPSCH, K., GRAEFE, J., WELSCH, T. Molekülstruktur und retentionsverhalten. II. Retentionsverhalten cycloaliphatischer kohlenwasserstoffe bei der gas- adsorptions- und gas- verteilungschromatographie. *Journal of Chromatography*, v. 91, p.623-31, 1983.
130. SPIEKSMAN, W., Determination of vapor liquid equilibrium from the Kováts retention index on dimethylsilicone using the Wilson Mixing Model. *J. High Resol. Chromatogr.*, v. 22, p.565-88, 1999.
131. ETTRE, L. S., BILLED, K. J. Eine method zur charakterisierung von gaschromato-graphischem trennflussigkeiten. *Journal of Chromatography*, v. 30, p.1-11, 1967.
132. BUDAHEGYI, M. V., LOMBOSI, E. R., LOMBOSI, T. S., MÉSZÁROS, S. Y., NYIREDY, Sz., TARJÁN, G., TIMAR, I., TAKÁCS, J. M. Twenty-fifth anniversary of the retention index system in gas - liquid chromatography. *Journal of Chromatography*, v. 271, p.213-307, 1983.
133. Lide DR. (Ed.) *Handbook of Chemistry and Physics*, 85^a ed., CRC Press, New York, 2004.

134. Quinn JA. **Molecular drawing and modeling software**, version 6 1 0, ChemSW, Inc, Fairfield, CA 94585-4019, USA, 2006. Disponível em: <http://www.chemsw.com/> Acesso em 05/06/2007.
135. MicroCal Origin version 5.0, serial number G43S5-9478-7055154, 1997.
136. Kubinyi H, Mannhold R, Gaard-Larsen PK, Timmerman H (Eds), **QSAR: Hansch Analysis and Related Approaches**, Wiley-VCH, Weinheim, 1993.
137. van de Waterbeemd H. **Chemometric Methods in Molecular Design**, Wiley VCH, Weinheim, 1995.
138. REN, B. A new topological index for QSPR of alkanes, **J. Chem. Inf. Comput. Sci.**, v. 39, p.139-43, 1999.
139. PENG, C. T. Prediction of retention indices. V. Influence of electronic effects and column polarity on retention index. **J. Chromatogr. A**, v. 903, p.117-143, 2000.
140. McREYNOLDS, W. O. Characterization of some liquid phases. **J. Chromatogr. Sci.**, v.8, p.685-91, 1970.

141. SUPELCO Bulletin 880: **The retention index system in gas chromatography: McReynolds constants**. Sigma Aldrich Co. Disponível em www.sigma-aldrich.com Acesso em 02/02/2008.
142. FIRPO, G., GRASSIOT, M. MARTIN, M. CARBO, R., GUARDINO. X., ALBAIGÈS, J. Empirical quantum chemical approach to structure-gas chromatographic retention index relationship. II. Cycloalkane derivatives. **Journal of Chromatography** n. 117, p.105-116, 1976.
143. TOROPOV, A. A., NESTEROV, I. V., NABIEV, O. M. QSPR modeling of cycloalkanes properties by correlation weighting of extended graph valence shells. **J. Mol. Struct. (Theochem)**, v. 637, p.37-42, 2003.
144. ONO, L. K. **Determinação da Relação entre o Índice de Refração do ar e a Pressão através do Interferômetro de Michelson**. USP – Instituto de Física – FGE 327 – Introdução à Óptica disponível em <http://socrates.if.usp.br/~lkono/grad/fge327/exp5.pdf> Acesso em 12/01/2008
145. DELGADO, E., Predicting aqueous solubility of chlorinated hydrocarbons from molecular structure. **Fluid Phase Equilib.**, v. 199, p.101-7, 2002.
146. HSIEH, H., MUKHERJEE, S. A QSAR model for desorption of halogenated aliphatics from biosolids. **Advances in Environmental Research**, v.7, p.511-20, 2003.

147. DUCHOWICZ, P., CASTRO, E. A., CHEN, P. Improved correlation between theoretical and experimental determination of heat of formation of some aliphatic nitro compounds. **Journal of the Korean Chemical Society**. V. 47, n. 2, p.89-95, 2003.
148. SILVA, L. R., FERREIRA, M. M. C., Estudo do coeficiente de partição octanol-água de bifenilas policloradas (PCBs) utilizando parâmetros topológicos. **Química Nova**, v. 26, n. 3, 2003. Disponível em http://www.scielo.br/scielo.php?pid=S0100-40422003000300005&script=sci_arttext
Acesso em 15/01/2008.
149. BERMÚDEZ-SALDAÑA, J. M., ESCUDER-GILABERT, I., MEDINA-HERMÁNDEZ, M. J., VILLANUEVA-CAMAÑAS, R. M., SALVADOR, S. Chromatographic evaluation of the toxicity in fish pesticides. **J. Chromatogr. B**, v. 814, p.115-125, 2005.
150. ATKINS, P., JONES, L. **Princípios de Química. Questionando a vida moderna e o meio ambiente**. Bookman, Porto Alegre, 2001.
151. APTULA, A. O., ROBERTS, D. W., CRONIN, M. T. D., SCHUTZ, T. W. Chemistry-toxicity relationships for the effects of di- and trihydroxybenzenes to *Tetrahymena pyriformis*, **Chem. Res. Toxicol.**, v. 18, p.844-54, 2005.

APÊNDICE
Apêndice A

Tabela 4. Valores de índice de retenção experimental, IR_{Exp} , (Apiezon-L 130°C); índice semi-empírico topológico, I_{ET} ; ponto de fusão, PF (°C); ponto de ebulição, PE (°C), e coeficiente de partição octanol/água, $\log P$, experimental e calculado pelos métodos CNDO e Contribuição dos Fragmentos.

(continua)

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
1	1-cloropropano	549 ^a	4,9515	-122,9	46,5	2,04	1,25	1,99
2	1-clorobutano	653 ^a	5,7599	-123,1	78,4	2,64	1,72	2,52
3	1-cloropentano	762 ^b	6,5684	-99,0	108,4	3,05	2,18	3,05
4	1-clorohexano	857 ^b	7,3769	-94,0	135,1	3,54	2,64	3,58
5	1-cloroheptano	958 ^b	8,1854	-69,5	160,4	4,15	3,10	4,11
6	1-clorooctano	1059 ^b	8,9939	-57,8	183,5	-	3,54	4,64
7	1-clorononano	1159 ^b	9,8024	-39,4	205,2	-	3,99	5,17
8	1-clorodecano	1263 ^a	10,6108	-31,3	225,9	-	4,43	5,70
9	1-cloroundecano	1365 ^a	11,4193	-17,0	242	-	4,87	6,23
10	1-clorotetradecano	1666 ^a	13,8400	4,9	296,8	-	6,16	7,81
11	1-clorohexadecano	1861 ^a	15,4600	17,9	326,6	-	7,02	8,87
12	1-clorooctadecano	2070 ^a	17,0800	28,6	352	-	7,87	9,93
13	1-cloroicicosano	2264 ^a	18,7000	-	-	-	8,71	10,99

Tabela 4.
Continuação

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
14	2-cloropropano	507 ^a	4,4992	-	35,7	1,9	1,20	1,99
15	2-clorobutano	616 ^a	5,3077	-131,3	68,2	2,33	1,67	2,52
16	2-cloropentano	712 ^b	6,1162	-	97,0	2,50	2,14	3,05
17	2-clorohexano	809 ^b	6,9248	-	122,5	-	2,60	3,58
18	2-cloroheptano	908 ^b	7,7332	-	-	-	3,05	4,11
19	2-clorooctano	1006 ^b	8,5417	-	172	-	3,50	4,64
20	2-clorononano	1106 ^b	9,3501	-	-	-	3,94	5,17
21	3-cloropentano	721 ^b	6,1244	-105	97,5	-	2,14	3,05
22	3-clorohexano	812 ^b	6,9330	-	123	-	2,59	3,58
23	3-cloroheptano	906 ^b	7,7414	-	144	-	3,04	4,11
24	3-clorooctano	1006 ^b	8,5499	-	-	-	3,49	4,64
25	3-clorononano	1106 ^b	9,3584	-	-	-	3,94	5,17
26	4-cloroheptano	900 ^b	7,6350	-	144	-	3,04	4,11
27	4-clorooctano	997 ^b	8,4435	-	-	-	3,49	4,64
28	4-clorononano	1094 ^b	9,2520	-	-	-	3,93	5,17
29	5-clorononano	1092 ^b	9,2116	-	-	-	3,93	5,17
30	1-cloro-2-metilpropano	623 ^a	5,4050	-130,3	68,5	2,45	1,73	2,39
31	1-cloro-2-metilpentano	823 ^b	7,0220	-	-	-	2,65	3,45
32	1-cloro-2-metilhexano	921 ^b	7,8305	-	-	-	3,10	3,98

Tabela 4.
Continuação.

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
33	1-cloro-2-metilheptano	1019 ^b	8,6390	-	-	-	3,35	4,51
34	1-cloro-3-metilbutano	723 ^a	6,2169	-104,4	98,2	-	2,19	2,92
35	1-cloro-4-metilpentano	823 ^b	7,0367	-	-	-	2,65	3,45
36	2-cloro-2-metilpropano	551 ^a	4,8861	-	50,9	-	1,59	2,39
37	2-cloro-2-metilbutano	671 ^a	5,7007	-	85,6	-	2,05	2,92
38	2-cloro-2-metipentano	763 ^b	6,5091	-	-	-	2,50	3,45
39	2-cloro-2-metilhexano	855 ^b	7,3176	-	-	-	2,96	3,98
40	2-cloro-2-metilheptano	949 ^b	8,1261	-	-	-	3,40	4,51
41	3-cloro-2-metilpentano	763 ^b	6,5698	-	-	-	2,55	3,45
42	3-cloro-2-metilhexano	850 ^b	7,3865	-	-	-	3,00	3,98
43	3-cloro-2-metilheptano	947 ^b	8,1953	-	-	-	3,46	4,51
44	3-cloro-6-metilheptano	966 ^b	8,2097	-	-	-	3,49	4,51
45	1,1-dicloropentano	889 ^b	7,5705	-	140	-	2,61	3,36
46	1,1-diclorohexano	989 ^b	8,3790	-	164	-	3,06	3,89
47	1,2-dicloropropano	695 ^a	6,0539	-	96,4	-	1,54	1,99
48	1,2-diclorobutano	801 ^a	6,8624	-	124,1	-	1,98	2,52
49	1,2-dicloropentano	909 ^b	7,6709	-	148,3	-	2,43	3,04
50	1,2-diclorohexano	1005 ^b	8,4794	-	173	-	2,87	3,57
51	1,3-diclorobutano	826 ^a	7,1046	-	134	-	1,96	2,24
52	1,3-dicloropentano	940 ^b	7,9131	-	-	-	2,39	2,76

Tabela 4.
Continuação.

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
53	1,3-diclorohexano	1028 ^b	8,7211	-	-	-	2,83	3,29
54	1,4-diclorobutano	895 ^a	7,4699	-	-	-	2,11	2,24
55	1,4-dicloropentano	959 ^b	8,1322	-	-	-	2,35	2,76
56	1,4-diclorohexano	1058 ^b	8,9407	-	-	-	2,78	3,29
57	1,5-dicloropentano	1028 ^b	8,7114	-	179	-	2,45	2,76
58	1,5-diclorohexano	1072 ^b	9,0677	-	-	-	2,82	3,29
59	1,6-diclorohexano	1129 ^b	9,5199	-	204	-	2,83	3,29
60	1,8-diclorooctano	1332 ^a	11,1368	-	241	-	3,86	4,35
61	1,9-diclorononano	1435 ^c	11,9453	-	260	-	4,19	4,88
62	1,10-diclorodecano	1541 ^a	12,7538	-	-	-	4,72	5,41
63	2,2-dicloropropano	619 ^a	5,4967	-	69,3	-	1,67	2,31
64	2,2-diclorobutano	738 ^a	6,3052	-	104	-	2,11	2,84
65	2,2-dicloropentano	831 ^b	7,1217	-	-	-	2,40	3,36
66	2,2-diclorohexano	925 ^b	7,9302	-	-	-	2,84	3,89
67	2,3-diclorobutano	763 ^a	6,4102	-	119	-	2,09	2,52
68	3,3-dicloropentano	857 ^b	7,2591	-	-	-	2,51	3,36
69	3,3-diclorohexano	944 ^b	8,0675	-	-	-	2,95	3,89
70	1-bromopropano	632 ^a	5,6136	-110,3	71,1	-	-	2,13

Tabela 4.
Continuação

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
71	1-bromobutano	734 ^a	6,4221	-112,6	101,6	2,75	-	2,66
72	1-bromopentano	836 ^a	7,2306	-88,0	129,8	3,37	-	3,19
73	1-bromohexano	938 ^a	8,0391	-83,7	155,3	3,80	-	3,72
74	1-bromoheptano	1040 ^a	8,8476	-56,1	178,9	4,36	-	4,25
75	1-bromooctano	1142 ^a	9,6561	-55,0	200,8	4,89	-	4,78
76	1-bromononano	1244 ^a	10,4646	-29,0	221,4	-	-	5,31
77	1-bromodecano	1345 ^a	11,2730	-29,2	240,6	6,00	-	5,84
78	1-bromoundecano	1447 ^a	12,0815	-9,7	258,8	--	--	6,37
79	1-bromododecano	1549 ^a	12,8900	-9,5	276	-	-	6,89
80	1-bromotridecano	1651 ^a	13,6985	6,2	292	-	--	7,42
81	1-bromotetradecano	1753 ^a	14,5070	5,6	307	-	-	7,95
82	1-bromohexadecano	1957 ^a	16,1240	18	336	-	-	9,01
83	1-bromoheptadecano	2059 ^a	16,9324	29,6	349	-	-	9,54
84	1-bromooctadecano	2161 ^a	17,7409	28,2	362	-	-	10,07
85	1-bromoeicosano	2364 ^a	19,3579	38	-	-	-	11,13
86	2-bromobutano	689 ^a	5,9971	-112,6	91,3	2,66	-	2,66
87	2-bromopentano	786 ^a	6,8056	-95,5	117,4	-	-	3,19
88	2-bromohexano	885 ^e	7,6140	-	143	-	-	3,72
89	2-bromoheptano	986 ^e	8,4225	-	166	-	-	4,25

Tabela 4.
Continuação

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
90	2-bromooctano	1093 ^c	9,2310	-	188,5	-	-	4,78
91	2-bromoundecano	1402 ^c	11,6565	-	-	-	-	6,37
92	2-bromododecano	1505 ^c	12,4650	-	-	-	-	6,89
93	2-bromotridecano	1609 ^c	13,2734	-	-	-	-	7,42
94	3-bromopentano	779 ^e	6,7188	-	118,6	-	-	3,19
95	3-bromohexano	884 ^e	7,5273	-	142	-	-	3,72
96	3-bromoheptano	988 ^e	8,3358	-	-	-	-	4,25
97	4-bromoheptano	978 ^e	8,2899	-	161	-	-	4,25
98	1-bromo-2-metilbutano	806 ^e	6,7912	-	119	-	-	3,06
99	1-bromo-2-metilpentano	880 ^e	7,5997	-	-	-	-	3,59
100	1-bromo-3-metilbutano	794 ^e	6,8022	-	120,4	-	-	3,06
101	1-bromo-4-metilpentano	901 ^a	7,6674	-	145	-	-	3,59
102	2-bromo-2-metilpropano	633 ^a	5,5394	-	73,3	-	-	2,53
103	2-bromo-2-metilbutano	750 ^a	6,3479	-	108	-	-	3,06
104	2-bromo-2-metilpentano	826 ^e	7,1564	-	-	-	-	3,59
105	2-bromo-2-metilhexano	942 ^a	7,9649	-	-	-	-	4,12
106	3-bromo-3-metilpentano	876 ^e	7,4652	-	-	-	-	3,59
107	1,2-dibromopropano	851 ^a	7,2995	-	141,9	-	-	2,27
108	1,2-dibromobutano	960 ^a	8,1079	-65,4	166,3	-	-	2,80

Tabela 4.
Continuação

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	Bp (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
109	1,3-dibromobutano	987 ^a	8,3627	-	174	-	-	2,52
110	1,4-dibromopentano	1113 ^a	9,2393	-	-	-	-	3,04
111	1,5-dibromopentano	1178 ^a	10,0356	-39,5	222,3	-	-	3,04
112	1,6-dibromohexano	1310 ^a	10,8443	-	245,5	-	-	3,57
113	1,7-dibromoheptano	1397 ^c	11,6527	-	263	-	-	4,10
114	1,8-dibromooctano	1519 ^a	12,4612	-	271	-	-	4,63
115	1,9-dibromononano	1607 ^c	13,2697	-	285	-	-	5,16
116	2,2-dibromopropano	771 ^a	6,6163	-	-	-	-	2,59
117	1-iodopropano	719 ^a	6,2823	-101,3	102,5	-	-	2,52
118	1-iodobutano	819 ^a	7,0908	-103	130,5	-	-	3,05
119	1-iodopentano	928 ^a	7,8923	-85,6	157,0	-	-	3,58
120	1-iodohexano	1032 ^a	8,7078	-74,2	181,3	-	-	4,11
121	1-iodoheptano	1132 ^a	9,5163	-48,2	204,0	4,70	-	4,64
122	1-iodooctano	1236 ^a	10,3248	-45,7	225,1	-	-	5,17
123	1-iodononano	1342 ^c	11,1332	-20	245,0	-	-	5,70
124	2-iodopropano	670 ^a	5,8478	-	89,5	-	-	2,52
125	2-iodobutano	789 ^d	6,6563	-	120,1	-	-	3,05
126	2-iodopentano	873 ^d	7,4678	-	-	-	-	3,58

Tabela 4.
Continuação

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
127	2-iodohexano	973 ^d	8,2732	-	-	-	-	4,11
128	2-iodoheptano	1076 ^d	9,0817	-	-	-	-	4,64
129	3-iodopentano	876 ^d	7,4412	-	145,5	-	-	3,58
130	3-iodohexano	972 ^d	8,2497	-	-	-	-	4,11
131	3-iodoheptano	1071 ^d	9,0582	-	-	-	-	4,64
132	4-iodoheptano	1061 ^d	8,9611	-	-	-	-	4,64
133	1-iodo-2-metilbutano	896 ^d	7,5948	-	-	-	-	3,45
134	1-iodo-2-metilpentano	988 ^d	8,4032	-	-	-	-	3,98
135	1-iodo-3-metilbutano	884 ^a	7,5746	-	147	-	-	3,45
136	1-iodo-3-metilpentano	995 ^d	8,3831	-	-	-	-	3,98
137	1-iodo-4-metilpentano	990 ^d	8,3870	-	-	-	-	3,98
138	2-iodo-2-metilpropano	715 ^a	6,3661	-	-	-	-	2,92
139	2-iodo-2-metilbutano	845 ^a	7,1745	-	-	-	-	3,45
140	2-iodo-2-metilpentano	960 ^d	7,9830	-	-	-	-	3,98
141	3-iodo-3-metilpentano	958 ^d	8,1282	-	-	-	-	3,98
142	1-clorododecano	1469 ^f	12,2278	-9,3	263,2	-	5,30	6,75
143	1-cloro-2-metilbutano	718 ^f	6,2135	-	100	2,52	2,19	2,92
144	1-cloro-6-metilheptano	1052 ^f	8,8862	-	-	-	3,55	4,51
145	2-cloro-3-metilbutano	662 ^f	5,7613	-	91,5	-	2,12	2,92

Tabela 4.
Continuação

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
146	2-cloro-4-metilpentano	763 ^f	6,5732	-	113	-	2,60	3,45
147	2-cloro-6-metilheptano	1001 ^f	8,4797	-	-	-	3,50	4,51
148	3-cloro-3-metilpentano	785 ^f	6,7468	-	-	-	2,50	3,45
149	4-cloro-2-metilheptano	953 ^f	8,0920	-	-	-	3,34	4,51
150	1,1-dicloropropano	686 ^f	5,9536	-	88,1	-	1,72	2,31
151	1,1-diclorobutano	787 ^f	6,7621	-	113,8	-	2,17	2,84
152	1,1-dicloroheptano	1090 ^f	9,1875	-	187	-	3,49	4,42
153	1,1-diclorooctano	1190 ^f	9,9960	-	208	-	3,53	4,95
154	1,1-diclorononano	1291 ^f	10,8045	-	228	-	4,37	5,48
155	2,3-dicloropentano	844 ^f	7,2187	-	139	-	2,52	3,04
156	1-bromopentadecano	1854 ^f	15,3155	19	322	-	-	8,48
157	2-bromopropano	591 ^f	5,1886	-	59,5	2,14	-	2,13
158	1-bromo-2-metilpropano	731 ^f	6,3177	-	-	-	-	2,53
159	2-bromo-3-metilbutano	737 ^f	6,3661	-	115	-	-	3,06
160	2-bromo-4-metilpentano	840 ^f	7,1857	-94	131	-	-	3,59
161	1,2-dibromopentano	1056 ^f	8,9164	-	184	-	-	3,32
162	1,2-dibromoheptano	1257 ^f	10,5334	-	228	-	-	4,38
163	1,4-dibromobutano	1094 ^f	9,2273	-	197	-	-	2,52
164	1-iododecano	1433 ^f	11,9417	-16,3	263,7	-	-	6,23

Tabela 4.
Continuação

Nº	Compostos	IR_{Exp}	I_{ET}	PF (°C)	PE (°C)	$\log P$	$\log P$ (CNDO)	$\log P$ (Contribuição dos Fragmentos)
165	1-iodoundecano	1534 ^f	12,7502	2,0	282	-	-	6,76
166	1-iodododecano	1635 ^f	13,5587	3,0	298,2	-	-	7,28
167	1-iodohexadecano	2038 ^f	16,7926	24,7	357	-	-	9,40
168	1-iodooctadecano	2240 ^f	18,4096	34	383	-	-	10,46
169	2-iodooctano	1119 ^f	9,4252	-	210	-	-	5,17
170	1-iodo-2-metilpropano	716 ^f	6,1969	-	121,1	-	-	2,93

^aValores extrapolados da referência [111]; ^bValores da referência [114]; ^cValores extrapolados da referência [126]; ^dValores extrapolados da referência [127]; ^eValores extrapolados da referência [128]; ^fÍndice de retenção calculado com a Equação 9.