



MARIA CAROLINA GUIMARÃES

**RELATIONSHIPS BETWEEN
CONFORMATIONAL ANALYSIS AND QSAR:
THE CASES FOR THE VOLATILE ANESTHETIC
ISOFLURANE AND FOR A XYLULOSE KINASE
ENZYME LIGAND**

**LAVRAS – MG
2017**

MARIA CAROLINA GUIMARÃES

**RELATIONSHIPS BETWEEN CONFORMATIONAL ANALYSIS AND
QSAR: THE CASES FOR THE VOLATILE ANESTHETIC ISOFLURANE
AND FOR A XYLULOKINASE ENZYME LIGAND**

Tese apresentada ao Departamento de Química da Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, área de concentração em Química/Bioquímica, para a obtenção do título de Doutora.

Prof. Dr. Matheus Puggina de Freitas
Orientador

**LAVRAS – MG
2017**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha
Catalográfica da Biblioteca Universitária da UFLA, com dados
informados pelo(a) próprio(a) autor(a).**

Guimarães, Maria Carolina.

Relationships between conformational analysis and QSAR: the cases
for the volatile anesthetic isoflurane and for a xylulokinase enzyme
ligand / Maria Carolina Guimarães. - 2017.

104 p.

Orientador(a): Matheus Puggina de Freitas.

Tese (doutorado) - Universidade Federal de Lavras, 2017.

Bibliografia.

1. Anestésicos voláteis. 2. DFX. 3. Análise conformacional. I. de
Freitas, Matheus Puggina. . II. Título.

MARIA CAROLINA GUIMARÃES

**RELATIONSHIPS BETWEEN CONFORMATIONAL ANALYSIS AND
QSAR: THE CASES FOR THE VOLATILE ANESTHETIC ISOFLURANE
AND FOR A XYLULOKINASE ENZYME LIGAND**

**RELAÇÕES ENTRE ANÁLISE CONFORMACIONAL E QSAR: O CASO
DO ANESTÉSICO VOLÁTIL ISOFLURANO E UM LIGANTE DA ENZIMA
XILULOQUINASE**

Tese apresentada ao Departamento de
Química da Universidade Federal de
Lavras, como parte das exigências do
Programa de Pós-Graduação em
Agroquímica, área de concentração em
Química/Bioquímica, para a obtenção
do título de Doutora.

APROVADA em 29 de setembro de 2017.

Dra. Melissa Soares Caetano UFOP

Dra. Maria das Graças Cardoso UFLA

Dra. Elaine Fontes Ferreira da Cunha UFLA

Dra. Daiana Teixeira Mancini UFLA

Prof. Dr. Matheus Puggina de Freitas
Orientador

**LAVRAS – MG
2017**

AGRADECIMENTOS

Agradeço, em primeiro lugar, a Deus e a Nossa Senhora, por iluminarem meus caminhos e serem a força necessária que me impulsiona a prosseguir.

A toda a minha família, em especial aos meus amados pais, Regina e Márcio; ao meu irmão, Lucas e ao meu marido, Cristiano, por serem minha base e estarem sempre presentes me apoiando, incondicionalmente, em todas as minhas decisões e projetos.

Agradeço ao meu orientador, Matheus, por todos os ensinamentos ao longo da minha trajetória acadêmica, paciência e disponibilidade de sempre que em muito contribuíram para o meu crescimento profissional.

Aos membros da banca, pelas considerações que, certamente, em muito contribuirão para este trabalho.

Aos professores e amigos do Laboratório de Modelagem Molecular, de modo especial a Mariene e Josué, que contribuíram diretamente com este trabalho, por todo o apoio e a amizade. Aos demais amigos, Stephen, Estella, Fátima, Letícia Garcia, Letícia Assis, Dani, Giovanna, Laize, Tamiris, Maíra e Lívia, pela amizade e colaboração, sobretudo nos momentos de dificuldade.

À Universidade Federal de Lavras, pela oportunidade de realizar este doutorado, bem como à CAPES, pelo financiamento dos recursos necessários ao desenvolvimento do mesmo.

A todos o meu agradecimento e gratidão!

RESUMO

A conformação bioativa das moléculas de um conjunto de dados usado em *Tridimensional Quantitative Structure-Activity Relationships*, QSAR-3D, é, geralmente, desconhecida. Nesse caso, o uso de descritores moleculares obtidos a partir de estruturas químicas otimizadas em um ambiente livre do receptor pode induzir a erros de predição do modelo e de interpretação de seus resultados. Atualmente, a metodologia QSAR-3D tem sido a mais utilizada, embora considere apenas o fato de que a estabilidade da molécula isolada é governada por interações intramoleculares, diferente do que acontece no meio biológico. Diante disso, o presente trabalho foi realizado com a finalidade avaliar quão imprescindíveis são os descritores 3D na produção de modelos QSAR com poder de predição elevado, por meio da comparação com uma técnica QSAR genuinamente bidimensional, a qual tem se mostrado altamente preditiva e de baixo custo computacional, no que se refere à memória e ao espaço em disco. Assim, após a análise conformacional do anestésico inalatório isoflurano e a comparação entre suas conformações preferenciais e bioativas, disponíveis no *Protein Data Bank* (PDB), foi construído um modelo *Multivariate Image Analysis applied to QSAR* (MIA-QSAR), essencialmente 2D, para uma série de seus derivados disponíveis na literatura (MEHDIPOUR; HEMMATEENEJAD; MIRI, 2007). Por fim, foram realizados estudos acerca do papel que a ligação de hidrogênio intramolecular presente na estrutura química da 5-deóxi-5-flúor-D-xilulose (DFX), desempenha sobre a sua bioconformação, como ligante da enzima xiluloquinase. Os estudos demonstraram que a metodologia 2D pode ser empregada com o mesmo grau de confiabilidade e predição das técnicas 3D. Além disso, este trabalho demonstrou que interações intramoleculares não governam a estabilidade da bioconformação das moléculas em estudo, a qual é regida pelo encaixe induzido da enzima.

Palavras-chave: Anestésicos voláteis. DFX. Análise conformacional. MIA-QSAR. Descritores moleculares.

ABSTRACT

The bioactive conformation of molecules used in 3D-QSAR (Tridimensional Quantitative Structure-Activity Relationships) is usually unknown. In this case, the use of molecular descriptors obtained from chemical structures optimized in a receptor-free environment can induce errors either in the prediction or interpretation of the QSAR models. In recent years, 3D-QSAR has been the most explored QSAR methodology, despite considering only the interactions governing the stability of a molecule in the isolated state, rather than in the biological environment. Therefore, this work aimed to evaluate how valuable 3D descriptors are in providing high performance QSAR models, by comparing the results with a bidimensional QSAR technique, which has been shown to be highly predictive and requires a low computation cost related to memory and disk space. The conformational analysis of the inhalational anesthetic isoflurane was theoretically carried out and the stable conformations were compared to the bioactive structures available in the Protein Data Bank (PDB). Also, a bidimensional MIA-QSAR (*Multivariate Image Analysis applied to QSAR*) modeling was performed for a series of fluorinated anesthetics similar to isoflurane. In another study, the role of an intramolecular hydrogen bond in 5-deoxy-5-fluoro-D-xylulose (DFX) on its bioconformation was theoretically analyzed. The outcomes indicate that a bidimensional QSAR methodology can be employed with reliability and similar prediction power than tridimensional techniques. In addition, this work showed that intramolecular interactions do not govern the stability of the bioconformation of the studied molecules, which is rather driven by an enzyme induced-fit.

Keywords: Volatile anesthetics. DFX. Conformational analysis. MIA-QSAR. Molecular descriptors.

LISTA DE FIGURAS

PRIMEIRA PARTE

- Figura 1 - Métodos de cálculo clássicos e quânticos utilizados em modelagem molecular. 18
- Figura 2 - Representação da cavidade do soluto e cargas de polarização na interação soluto-solvente. 25
- Figura 3 - (A) Hiperconjugação $\sigma_{CH} \rightarrow \sigma_{CF}^*$ responsável pelo efeito *gauche* no 1,2-difluór-etano; (B) Repulsão dipolar responsável pelo efeito anomérico em 2-X-tetraidropiranos; (C) Hiperconjugação $n_O \rightarrow \sigma_{CX}^*$ responsável pelo efeito anomérico em 2-X-tetraidropiranos. 27
- Figura 4 - Gráfico molecular de um dos confôrmeros do ácido aminofluorobenzoico. 31
- Figura 5 - Sobreposição das estruturas formando o arranjo tridimensional desdobrado na matriz X de descritores MIA-QSAR. 35
- Figura 6 - Sobreposição das estruturas desenhadas considerando o raio de Van der Waals, com diferentes cores na representação atômica e obtenção da matriz X de descritores *aug*-MIA-QSAR. 36
- Figura 7 - Enantiômeros do isoflurano. 41
- Figura 8 - (A) Isoflurano ligado à hélice 4 da apoferritina (código PDB: 3F78) e (B) isoflurano ligado à integrina LFA-1 (código PDB: 1XZ3). 42
- Figura 9 - (A) Estrutura química da 5-deóxi-5-flúor-D-xilulose e a respectiva estrutura bioativa (códigos PDB: 5FX); (B) enzima xiluloquinase (código PDB: 4BC5). 43

SEGUNDA PARTE

ARTIGO 1

- Figure 1 - Optimized structures of isoflurane at the ω B97X-D/6-311++g(d,p) level (gas phase and explicit water) and the experimental bioconformations 3F78 and 1XZ3. 59
- Figure 2 - Superimposed chemical structures used to generate the MIA descriptors with atoms colored according to the corresponding Pauling's electronegativity. 64
- Figure 3 - Plot of actual vs. predicted pMAC obtained from the MIA-QSAR model. 67

Figure 4 - PCA plots for the mean-centered data of the 25 anesthetic haloethers.....	68
--	----

ARTIGO 2

Figure 1 - Chemical structure of 5-deoxy-5-fluoro-D-xylulose (DFX) and its corresponding bioactive conformation (Protein Data Bank code: 5FX).....	78
Figure 2 - Crystal structure and the optimized geometries of DFX in the gas phase and implicit water. The corresponding hydrogen bonds obtained by NBO analysis ($n_{O/F} \rightarrow \sigma^*_{OH}$, in kcal mol ⁻¹) are shown.....	83
Figure 3 - Conformers 5FX , 1g and 1w docked inside the active site of xylulokinase. Hydrogen bonds with amino acid residues are shown, predominantly in 5FX	92

LISTA DE TABELAS

SEGUNDA PARTE

ARTIGO 1

Table 1 -	Relative electronic/Gibbs free energies (kcal mol ⁻¹ , % Gibbs population in parenthesis) for the conformers of isoflurane. ^a	60
Table 2 -	Anomeric interactions and relative Lewis and non-Lewis contributions (kcal mol ⁻¹) for the conformers of isoflurane, in the gas phase at the ω B97X-D/6-311++g(d,p) level.	61
Table 3 -	Atomic properties (a.u.) obtained by QTAIM to characterize hydrogen bonds in isoflurane.	62
Table 4 -	Data set of anesthetic haloethers and the corresponding experimental pMAC values.	65
Table 5 -	Correspondence of Pauling's electronegativity with atomic colors and respective pixel values (as a combination of RGB components) used in the MIA-QSAR model.	66
Table 6 -	Reduced matrix with the selected descriptors and their approximate coordinates in the images.	69

ARTIGO 2

Table 1 -	Relative energies (kcal mol ⁻¹), conformer Gibbs populations (%) and dihedral angles (degrees) for the conformations of DFX in the gas phase and implicit water, optimized at the ω B97X-D/6-311++g(d,p) level.	81
Table 2 -	QTAIM parameters ^a used to characterize hydrogen bonds in DFX.	84
Table 3 -	Hyperconjugative interactions between antiperiplanar donor and acceptor orbitals (kcal mol ⁻¹) for the conformers of DFX.	87
Table 4 -	Intermolecular (ligand-protein) interactions (kcal mol ⁻¹) for the main optimized conformers of DFX, obtained by docking studies. ^a	91
Table S1 -	Optimized conformations for DFX in the gas phase at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol ⁻¹) and key dihedral angles.	98
Table S2 -	Optimized conformations for DFX in implicit water at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol ⁻¹) and key dihedral angles.	101

LISTA DE ABREVIATURAS

PRIMEIRA PARTE

X	Átomo eletronegativo genérico
:B	Base de Lewis
E_{HIPER}	Energia de hiperconjugação
E_{LEWIS}	Energia de Lewis
E_{TOTAL}	Energia total
$q_{\text{(H)}}$	Carga atômica
E	Energia
$M_{1\text{(H)}}$	Primeiro momento de dipolo
$V_{\text{(H)}}$	Volume atômico
DFX	5- deóxi-5-flúor-D-xilulose

SEGUNDA PARTE

ARTIGO 1

r^2	Determination coefficient
r_{test}^2	Determination coefficient for the external validation
q^2	Determination coefficient for the cross validation
$r_{\text{y_rand}}^2$	Determination coefficient for the y_randomization test
c_r^2	Difference between randomized/calibrated determination coefficient
	Pauling's electronegativity
pMAC	Anesthetic activity
Log P	Logarithm of the octanol/water partition coefficient

ARTIGO 2

V^2_{P}	Laplacian of the electron density
$H_{\text{(r)}}$	Total energy density at the critical point
G°_{rel}	Relative Gibbs free energy
E_{rel}	Relative energy
E_{NL}	Non-Lewis energy
$\rho_{\text{(r)}}$	Electron density
$\Delta G^{\circ}_{\text{rel}}$	Relative Gibbs free energy difference

LISTA DE SIGLAS

PRIMEIRA PARTE

2D	Segunda dimensão
3D	Terceira dimensão
4D	Quarta dimensão
5D	Quinta dimensão
6D	Sexta dimensão
7D	Sétima dimensão
QSAR	Quantitative Structure-Activity Relationships
MIA-QSAR	Multivariate Image Analysis Applied to QSAR
<i>aug</i> -MIA-QSAR	Augmented MIA-QSAR
PDB	Protein Data Bank
LFA-1	Lymphocyte Function-Associated Antigen-1
NBO	Natural Bond Orbital
QTAIM	Quantum Theory of Atoms in Molecules
TOM	Teoria dos orbitais moleculares
MM	Mecânica Molecular
DFT	Density Functional Theory
STO	Função de base do tipo Slater
GTO	Função de base do tipo gaussiana
PCM	Polarized continuum model
LH	Ligação de hidrogênio
RMN	Ressonância magnética nuclear
BP	Bond Path
BCP	Bond Critical Point
RGB	Red-green-blue
MVD	Molegro virtual Docker
MD	Descritores moleculares
DMSO	Dimetilsulfóxido

SEGUNDA PARTE

ARTIGO 1

PLS	Partial Least Square
RMSE	Root Mean Square Error
PCA	Principal Component Analysis

SUMÁRIO

	PRIMEIRA PARTE	13
1	INTRODUÇÃO GERAL.....	13
2	JUSTIFICATIVA	17
3	REFERENCIAL TEÓRICO.....	18
3.1	A química computacional aplicada ao estudo de compostos bioativos	18
3.2	Análise conformacional	21
3.2.1	Métodos de estruturas eletrônicas.....	22
3.2.2	O funcional ωB97X-D e funções de base.....	23
3.2.3	Efeitos do solvente.....	24
3.2.4	Interações moleculares	26
3.3	Cálculos de orbitais.....	29
3.4	QTAIM	30
3.5	QSAR	32
3.6	MIA-QSAR.....	34
3.7	Ancoramento molecular ou <i>docking</i>	37
3.8	Análise qualitativa dos dados.....	38
3.9	Compostos organofluorados	39
	REFERÊNCIAS.....	45
	SEGUNDA PARTE - ARTIGOS.....	54
	ARTIGO 1 - Is conformation a fundamental descriptor in QSAR? A case for halogenated anesthetics	54
	ARTIGO 2 - Is the bioconformation of 5-deoxy-5-fluoro-D-xylulose affected by intramolecular hydrogen bond?.....	74

PRIMEIRA PARTE

1 INTRODUÇÃO GERAL

A química computacional corresponde a um ramo interdisciplinar da química que utiliza o potencial oferecido pela tecnologia de hardwares e softwares para realizar estudos de propriedades moleculares que se estendem desde as estruturas eletrônicas e conformações, até os diferentes tipos de energia, dinâmica e reatividade de sistemas (FERNANDES, 2011), chegando ao planejamento de fármacos, à química ambiental e, de forma complementar, à ciência dos materiais (ALLOUCHE, 2011; ORTOLAN, 2014).

Por meio de cálculos teóricos computacionais, a análise conformacional permite a identificação das várias conformações ou arranjos moleculares temporários possíveis para um determinado composto e calcula as diferentes energias de estabilização (SOLOMONS; FRYHLE, 2009), além de possibilitar o estudo das interações que ocorrem entre ligante e receptor (FRAGA, 2001), despertando, cada vez mais, o interesse de químicos e bioquímicos.

É certo que a conformação e as propriedades físico-químicas governam, direta ou indiretamente, as respostas biológicas dos compostos químicos. Assim, os modelos QSAR apresentam-se como uma alternativa prática na identificação das propriedades relacionadas com a atividade exercida ao tratar da relação quantitativa entre estrutura-atividade, expressa por modelos matemáticos (MONTANARI, 2011).

Em QSAR, as características da estrutura química são codificadas por meio dos chamados descritores moleculares e a complexidade do problema aumenta de acordo com a quantidade e a diversidade dos descritores, o que exige métodos mais sofisticados, como os multidimensionais, para o tratamento dos dados (MONTANARI, 2011). Dentre eles, o QSAR-3D é o

mais popular e se baseia no uso de descritores moleculares gerados a partir de estruturas tridimensionais alinhadas no interior de uma caixa virtual 3D ou *grid cell* (CRAMER III; PATTERSON; BUNCE, 1988; KLEBE; ABRAHAM; MIETZNER, 1994).

Com a incorporação da liberdade conformacional aos modelos 3D, foi possível obter o QSAR-4D (HOPFINGER et al., 1997). Ao levar em consideração a representação do receptor com a simulação do encaixe induzido, chegou-se ao QSAR-5D (VEDANI; DOBLER, 2002). A possibilidade de avaliar diferentes modelos de solvatação simultaneamente deu origem ao QSAR-6D (VEDANI; DOBLER; LILL, 2005), enquanto o modelo baseado no alvo e no receptor real representou a sétima dimensão (QSAR-7D) (POLANSKI, 2009).

Embora o QSAR-3D seja o mais utilizado, ele requer uma exaustiva varredura conformacional, seguida de rigorosas regras para o alinhamento tridimensional de moléculas (CRAMER III; PATTERSON; BUNCE, 1988; KLEBE; ABRAHAM; MIETZNER, 1994) e ainda apresenta a desvantagem de desconsiderar a conformação bioativa e otimizar o conjunto de dados em um ambiente livre do receptor, diferente do que acontece em meio biológico.

Em contrapartida, a análise multivariada de imagens aplicada em QSAR (MIA-QSAR) (FREITAS; BROWN; MARTINS, 2005), assim como métodos QSAR que utilizam descritores moleculares 2D, tem se mostrado satisfatoriamente preditiva há mais de uma década de seu desenvolvimento, em razão da simplicidade dos cálculos, da rapidez e do baixo custo computacional.

Atualmente, por considerar informações que descrevam melhor as estruturas químicas, tais como a representação dos diferentes átomos como esferas coloridas com tamanhos proporcionais aos seus respectivos raios de van der Waals (NUNES; FREITAS, 2013) e cores sólidas na representação atômica (DUARTE et al., 2015), os modelos MIA-QSAR têm apresentado resultados

cada vez mais comparáveis aos métodos QSAR multidimensionais mais modernos e bioquimicamente interpretáveis.

O presente trabalho foi desenvolvido com a finalidade de testar a hipótese de que as conformações obtidas em um ambiente livre do receptor devem ser utilizadas com cautela em metodologias QSAR, independente da dimensionalidade do método escolhido.

Por apresentarem diversas aplicações industriais e, dessa forma, estarem presentes em nosso cotidiano, os compostos orgânicos contendo flúor, conhecidos como organofluorados (BOECHAT; PINTO; BASTOS, 2015), foi a classe escolhida para esse estudo.

Assim, o anestésico inalatório isoflurano [2-cloro-2-(diflúor-metóxi)-1,1,1-(trifluoroetano)] foi analisado conformacionalmente, por meio de cálculos teóricos, a fim de ter suas geometrias otimizadas comparadas às estruturas bioativas disponíveis no banco de dados PDB, as quais se ligam à apoferritina (código PDB: 3F78) (LIU; PATRICK; RODERIC, 2005) e à integrina LFA-1 (PDB: 1XZ3) (ZHANG et al., 2009).

Alguns derivados do isoflurano, obtidos da literatura (MEHDIPOUR; HEMMATEENEJAD; MIRI, 2007), foram utilizados na construção de um modelo MIA-QSAR, com o objetivo de comprovar que descritores bidimensionais são capazes de estimar propriedades biológicas igualmente bem, em relação aos descritores 3D, e que eles não são imprescindíveis para que se produzam modelos QSAR bem interpretáveis e com poder de predição considerável.

Em um segundo momento, foram realizados estudos sobre o papel que a ligação de hidrogênio intramolecular na 5-deóxi-5-flúor-D-xilulose (DFX) desempenha sobre a sua estabilidade bioconformacional. Assim, foi possível avaliar o impacto dessa interação em suas conformações preferenciais. Suas principais geometrias otimizadas foram estudadas por meio da análise dos

orbitais naturais de ligação (NBO, do inglês *natural bond orbital*) e por meio da teoria quântica dos átomos em moléculas (QTAIM, do inglês *quantum theory of atoms in molecules*), para a identificação das ligações de hidrogênio intramoleculares e compreensão do seu efeito na conformação livre. Além disso, tais geometrias foram comparadas à estrutura bioativa, de código PDB: 5FX e ancoradas no sítio ativo da enzima xiluloquinase (código PDB: 4BO5) por meio dos estudos de ancoramento molecular ou *docking*, e as interações intermoleculares possíveis, existentes entre ligante e receptor, foram avaliadas e comparadas com as que ocorrem entre a 5FX e a xiluloquinase, descritas na literatura (BUNKER et al., 2013).

Por fim, o presente estudo possibilitou a compreensão do efeito das forças intra e intermoleculares que operam nas conformações preferenciais de duas moléculas organofluoradas com propriedades biológicas conhecidas e reforçou a hipótese de que o uso, em QSAR, de informações sobre conformações obtidas em um ambiente livre do receptor pode ocasionar má interpretação de resultados químicos e biológicos.

2 JUSTIFICATIVA

A execução deste estudo é justificada pelo fato de a discussão sobre a importância de se levar em consideração ou não descritores tridimensionais em estudos QSAR ser escassa. Uma vez demonstrado que varreduras conformacionais e o alinhamento tridimensional de moléculas não são indispensáveis na obtenção de modelos QSAR satisfatoriamente preditivos, poder-se-ão utilizar métodos bidimensionais, com a certeza de que os resultados de previsão terão qualidade comparável à dos métodos QSAR multidimensionais. Certamente, isso irá acelerar o processo de desenho racional de novos fármacos e compostos bioativos.

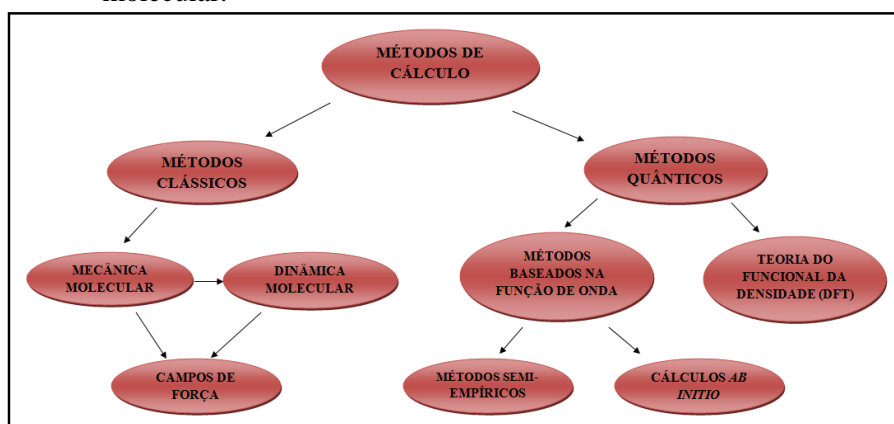
3 REFERENCIAL TEÓRICO

3.1 A química computacional aplicada ao estudo de compostos bioativos

A química computacional tem colaborado, de forma significativa, para o desenvolvimento de uma grande variedade de técnicas de cálculos e simulações dedicadas à resolução de problemas químicos, bioquímicos, tecnológicos e industriais (FERNANDES, 2011). Esses cálculos podem ser aplicados na análise e na interpretação de dados experimentais e, até mesmo, na previsão de fenômenos químicos que não poderiam ser obtidos experimentalmente (ARROIO et al., 2005).

Por meio dos diferentes métodos de cálculos disponíveis atualmente é possível o cálculo de propriedades moleculares (energia, dinâmica e reatividade de sistemas moleculares), fundamentado em teorias e modelos da mecânica clássica e quântica, do eletromagnetismo e da termodinâmica estatística (FERNANDES, 2011).

Figura 1 - Métodos de cálculo clássicos e quânticos utilizados em modelagem molecular.



Fonte: Adaptado de Campos (2013).

De forma geral, os métodos de cálculo computacionais típicos são baseados na mecânica molecular e na estrutura eletrônica, da qual derivam os métodos baseados na função de onda e a teoria do funcional da densidade (DFT, do inglês *density functional theory*), conforme ilustrado na Figura 1.

A mecânica molecular (MM) é adequada para lidar com sistemas contendo um grande número de átomos, tais como as proteínas. Os métodos de mecânica molecular desconsideram o movimento dos elétrons e consistem no cálculo da energia de uma molécula ou sistema como uma função apenas das posições nucleares (COELHO et al., 1999). Já a dinâmica molecular pode ser empregada tanto para sistemas atômicos quanto para sistemas macromoleculares e fornece informações sobre o comportamento dinâmico microscópico, dependente do tempo, dos átomos que compõem um sistema, simulado por meio da integração numérica da equação de movimento de Newton (NAMBA; SILVA; SILVA, 2008), que pode ser descrita para um sistema atômico simples pelas Equações 1 e 2.

$$\mathbf{F}_i(t) = m_i \mathbf{a}_i \quad (1)$$

$$\mathbf{F}_i(t) = \frac{\partial V(\mathbf{r}_i)}{\partial \mathbf{r}_i} \quad (2)$$

em que \mathbf{F}_i representa a força que atua sobre cada partícula do sistema em um instante de tempo t , enquanto \mathbf{a}_i está relacionada à aceleração do átomo i de massa m_i (NAMBA, SILVA; SILVA, 2008).

A força que irá atuar sobre cada átomo é definida pelos campos de força que correspondem às funções matemáticas que descrevem o átomo por meio das forças intra (soma dos comprimentos e ângulos de ligação e diedros) e intermoleculares (interações de van der Waals e de Coulomb), que são incluídas na mecânica e na dinâmica molecular (NAMBA; SILVA; SILVA, 2008).

Os métodos baseados na função de onda, ou seja, na resolução da equação de Schrödinger independente do tempo (Equação 3), são divididos em *ab initio* e semiempíricos. O primeiro é indicado para moléculas menores e, apesar de ser mais preciso e não necessitar de dados armazenados, requer grande capacidade de memória e tempo de cálculo do computador (CARVALHO et al., 2003). O segundo inclui parâmetros empíricos ou previamente calculados com a intenção de reduzir algumas integrais e facilitar o cálculo (SANTANNA, 2009).

$$\hat{H}\psi = E\psi \quad (3)$$

Já a teoria do funcional da densidade (DFT), incluída entre os métodos quânticos, considera que a energia de um conjunto de elétrons sob a influência de um campo externo é um funcional único da densidade eletrônica (SANTANNA, 2009), ou seja, apresenta as vantagens de incluir nos cálculos, de forma implícita, a correlação eletrônica e de exigir uma menor demanda computacional, quando comparada aos métodos *ab initio*.

O estudo sobre as preferências conformacionais de compostos químicos e bioativos por meio de cálculos computacionais tem despertado o interesse de químicos e bioquímicos (SMITH et al., 2007), em que se destacam os estudos de análise conformacional.

Além disso, a modelagem molecular tem se firmado como uma importante ferramenta não somente no processo de descoberta e desenvolvimento de fármacos e compostos bioativos, como também na otimização de protótipos já existentes, por meio da determinação das relações entre estrutura-atividade molecular (QSAR), relacionada à análise e ao tratamento da informação química por meio da quimiometria (FERNANDES, 2011; RODRIGUES, 2001).

3.2 Análise conformacional

A investigação de propriedades macro e microscópicas de compostos com um ou mais graus de liberdade envolve abordagens computacionais de análise conformacional, a fim de, posteriormente, serem aplicadas em metodologias como QSAR e estudos de ancoramento molecular ou *docking*.

A análise conformacional corresponde a um ramo da stereoquímica dedicado ao estudo dos diferentes arranjos espaciais que os átomos em uma molécula podem assumir quando uma ou mais ligações químicas simples sofrem rotação; aos diferentes arranjos tridimensionais da molécula dá-se o nome de confôrmeros (SOLOMONS; FRYHLE, 2009). Por meio de cálculos computacionais podem-se prever as geometrias dos confôrmeros estáveis, bem como suas energias e estados de transição, o que pode ser benéfico para a compreensão de mecanismos de reação e identificação de interações da molécula com sítios ativos enzimáticos (SMITH et al., 2007).

Durante a escolha e a aplicação do método computacional, é importante considerar as propriedades que se pretende avaliar, bem como a precisão desejada nas análises e a capacidade computacional disponível para a realização dos cálculos (SANTANNA, 2009). Vale ressaltar que as diferentes abordagens de cálculo compreendem diferentes métodos, funcionais e funções de base que impactam diretamente na qualidade dos resultados, no tempo de cálculo e no custo computacional (memória e espaço em disco).

Além das interações intramoleculares, as preferências conformacionais podem ser influenciadas pelo efeito dos solventes. Por essa razão, eles devem ser levados em consideração durante a análise de um equilíbrio conformacional, o que pode ser feito de forma explícita (em que as moléculas do solvente são incluídas no cálculo, aumentando o custo computacional) ou implícita (o soluto é inserido em uma cavidade cujo exterior apresenta uma constante dielétrica,

correspondente à do solvente que se deseja analisar) (TOMASI; MENNUCCI; CAMMI, 2005).

3.2.1 Métodos de estruturas eletrônicas

Dentre os métodos de cálculo, a Teoria do Funcional da Densidade (DFT), desenvolvida por Hohenberg e Kohn (1964) e a Teoria da Perturbação de Segunda Ordem de Möller-Plesset (MÖLLER; PLESSET, 1934) apresentam como variável fundamental a correlação eletrônica, ou seja, a relação existente entre os elétrons em movimento e, de forma geral, fornecem resultados congruentes (MARTINS, 2010).

Enquanto os métodos *ab initio* (do latim “do início”), como o de perturbação MP2, não utilizam parâmetros experimentais e preveem propriedades de sistemas atômicos e moleculares por meio das leis da mecânica quântica e de algumas constantes físicas, como a massa e a carga do elétron, a velocidade da luz e a constante de Planck (ALCÁCER, 2007), a teoria do funcional da densidade pode ser aplicada ao estudo de moléculas maiores do que aquelas estudadas por meio de métodos *ab initio*, pois seus cálculos para correlação eletrônica são simplificados e, aliada ao avanço computacional, tem permitido o estudo de sistemas cada vez mais complexos (MARQUES; BOTTI, 2006).

A DFT vem se destacando por utilizar a densidade eletrônica, descrita por meio de funções de troca e correlação, na determinação de sistemas com a vantagem de exigir baixa demanda computacional, em comparação ao MP2, por exemplo (COHEN; MORI-SANCHEZ; YANG, 2012; KADUK; KOWALCZYK; VOORHIS, 2012; ORTOLAN, 2014). Em razão de exigir uma menor demanda computacional, a DFT possibilita, em alguns casos, maior

concordância com dados experimentais, uma vez que conjuntos de bases mais sofisticados podem ser empregados (SILLA, 2013).

De forma geral, é necessário ter coerência na escolha do método de cálculo a ser utilizado, bem como dos funcionais e funções de base, uma vez que não há disponível uma única metodologia que seja aplicável a todas as moléculas. Esse cuidado na escolha irá garantir a precisão necessária ao estudo da propriedade de interesse no sistema, além da economia de tempo e custo computacional.

3.2.2 O funcional ω B97X-D e funções de base

Segundo a DFT, a correlação eletrônica é calculada utilizando-se os funcionais de densidade eletrônica. Dentre eles, o funcional ω B97X-D, desenvolvido por Chai e Head-Gordon (2008), fornece melhores resultados para átomos com excitação eletrônica e transferência de carga ao considerar os efeitos de dispersão.

De acordo com a Teoria do Orbital Molecular (TOM), as funções de base correspondem ao conjunto de funções matemáticas que representam os orbitais moleculares e são divididas em orbitais atômicos do tipo *slater* (STO, do inglês *slater-type atomic orbitals*) e GTO, do tipo gaussiana. As GTOs surgiram como alternativa ao uso das STOs, que contêm integrais exponenciais que nem sempre apresentam soluções analíticas (CRAMER, 2004; SILLA, 2013).

As funções atômicas do tipo gaussianas, representadas pela letra G, apresentam várias combinações possíveis, tais como as funções contraídas de Pople, que podem ser do tipo 3-21G, 6-31G, 6-31+G, 6-311++G(d,p), entre outras. Segundo a notação de Pople, os números indicam o número de funções primitivas, sendo que os orbitais mais próximos do núcleo estão à esquerda do hífen, enquanto os de valência estão à sua direita. Podem ser incluídas funções

difusas (representadas por + ou ++), que contribuem para uma melhor descrição espacial dos orbitais ocupados. As funções de polarização (tais como as “d,p” do exemplo acima), auxiliam na descrição da distorção da nuvem eletrônica para átomos pesados e para o hidrogênio, respectivamente (MORGON; CUSTODIO, 2001).

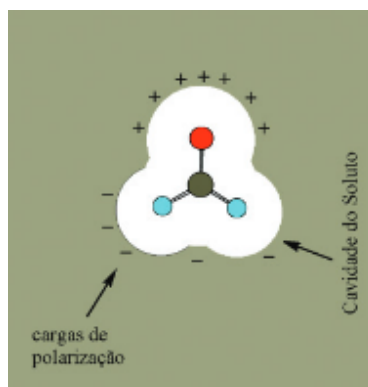
3.2.3 Efeitos do solvente

Os métodos teóricos, a princípio, calculam as geometrias das moléculas de forma isolada, desconsiderando as perturbações causadas por suas vizinhanças e outros sistemas (TOMASI; MENNUCCI; CAMMI, 2005). Porém, sabe-se que o meio pode influenciar diversas propriedades de moléculas, bem como o curso e, conseqüentemente, o produto de uma reação química (PLIEGO JÚNIOR, 2006; REICHARDT, 1988).

No isomerismo conformacional, o efeito dos solventes, implícito e explícito, é levado em consideração. O solvente implícito é caracterizado por meio de uma constante dielétrica, enquanto o solvente explícito é representado como moléculas individuais, inseridas no cálculo da mesma forma que o soluto (TOMASI; MENNUCCI; CAMMI, 2005).

Dentre as metodologias descritas para a modelagem do meio de solvatação, destaca-se o modelo de polarização contínua PCM, do inglês *polarizable continuum model* (CANCES; MENNUCCI, 1998), que define a cavidade ocupada pelo soluto como a união de esferas que circundam os átomos da molécula e que apresenta em seu interior a distribuição de cargas do soluto (Figura 2), a fim de descrever essa interação (TOMASI; MENNUCCI; CAMMI, 2005).

Figura 2 - Representação da cavidade do soluto e cargas de polarização na interação soluto-solvente.



Fonte: Adaptado de Pliego Júnior (2006).

Ao contrário do modelo de reação autoconsistente (SCRFF, do inglês *self-consistence reaction field*), em que o soluto é disposto no interior de uma cavidade esférica ou elipsoide e a densidade de carga da molécula é considerada como sendo rígida, no PCM a nuvem eletrônica sofre deformação na presença das cargas de polarização do dielétrico (PLIEGO JÚNIOR, 2006).

Os diferentes tipos de solventes disponíveis são classificados em três categorias principais, de acordo com suas polaridades (polares ou apolares) e capacidade de ligações de hidrogênio com o soluto (próticos e apróticos) (BÖES, 2005; REICHARDT, 1990). Os solventes apolares apróticos incluem os hidrocarbonetos, seus derivados halogenados, aminas terciárias e o dissulfeto de carbono e, além da incapacidade de formar ligação de hidrogênio, apresentam baixos momentos de dipolo. Os solventes polares próticos são formadores de ligações de hidrogênio (apresentam átomos de hidrogênio ligados a um elemento eletronegativo) e são representados por água, amônia líquida, alcoóis, ácidos carboxílicos, aminas e algumas amidas. Já os solventes polares apróticos não são doadores de ligações de hidrogênio e a presença de pares de elétrons não ligados faz com que solvatem os cátions mais fortemente. Nesse grupo estão incluídos o

nitrometano, o dimetilsufóxido (DMSO), a *N,N*-dimetilformamida, a acetonitrila e o nitrobenzeno, dentre outros (BÖES, 2005; REICHARDT, 1990).

As estruturas apresentadas no presente trabalho, além de serem estudadas na fase gasosa, foram otimizadas levando-se em consideração o efeito dos solventes cicloexano (C₆H₁₂), DMSO e água, por meio do modelo de solvatação implícito PCM, além da inclusão de uma molécula de água de forma explícita.

3.2.4 Interações moleculares

Além do efeito do solvente, as interações moleculares exercem forte influência na estereoquímica dos compostos e podem ser classificadas como intra ou intermoleculares. As interações intramoleculares tornam alguns arranjos moleculares energeticamente mais favoráveis que outros, ou seja, afetam a estabilidade conformacional (DUARTE; FREITAS, 2009). Já as interações intermoleculares refletem na afinidade e na especificidade entre ligante e receptor, por exemplo (FOYE; LEMKE; WILLIAMS, 1995; FRAGA, 2001; GRINGAUZ, 1997; TAYLOR; KENNEWELL, 1981; WOLFF, 1995).

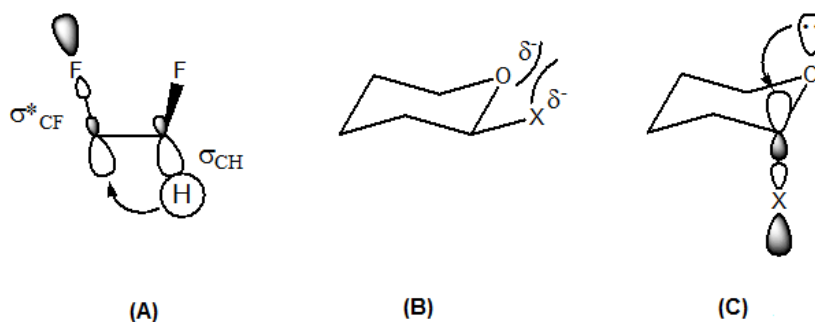
Essas interações podem ser de origem clássica (estéricas e eletrostáticas) e quântica (hiperconjugação). O efeito estérico diz respeito à repulsão espacial entre os átomos em uma molécula, interferindo na estabilidade dos conformeros (LIU; GOVIND, 2008). O efeito eletrostático, por sua vez, está relacionado às forças de atração ou repulsão eletrostáticas, que são resultantes entre dois dipolos e/ou íons de cargas opostas e é diretamente dependente da constante dielétrica do meio e da distância entre as cargas (FRAGA, 2001).

Já o efeito de hiperconjugação é importante para o entendimento das propriedades moleculares e reatividade e, de forma geral, provém da interação entre orbitais moleculares, sendo que um deles corresponde a uma ligação σ . O

orbital ocupado de maior energia (HOMO) transfere densidade eletrônica para o orbital vacante de menor energia (LUMO) (ALABUGIN et al., 2011). Tais efeitos podem governar a estabilidade de uma estrutura tridimensional responsável por uma determinada propriedade molecular, como atividade biológica.

Entre as interações intramoleculares que afetam a estabilidade conformacional, o bem conhecido efeito anomérico tem grande contribuição de interações hiperconjugativas; esse efeito refere-se à preferência pela geometria, em anéis piranosídicos 2-substituídos, que permite a interação entre um dos pares de elétrons livres do átomo de oxigênio com o orbital σ^* vizinho (ANDERSON; SEEP, 1968). Embora sua origem seja controversa, assim como para o efeito *gauche*, pelo menos duas hipóteses são utilizadas para explicá-lo, que são por meio de repulsão dipolar (MO, 2010) e por efeitos hiperconjugativos (Figura 3) (FREITAS, 2013).

Figura 3 - (A) Hiperconjugação $\sigma_{CH} \rightarrow \sigma^*_{CF}$ responsável pelo efeito *gauche* no 1,2-difluór-etano; (B) Repulsão dipolar responsável pelo efeito anomérico em 2-X-tetraidropiranos; (C) Hiperconjugação $n_O \rightarrow \sigma^*_{CX}$ responsável pelo efeito anomérico em 2-X-tetraidropiranos.



Fonte: Buissonneaud, Mourik e O'Hagan (2010), Goodman, Gu e Pophristic (2005) e Souza e Freitas (2011).

As ligações de hidrogênio (LH) são extensamente estudadas na literatura e, embora ainda não haja uma definição absoluta para esse tipo de interação devido à sua natureza (eletrostática ou covalente), complexidade e possível participação de grupos não usuais, de forma geral, sua origem está relacionada com a atração eletrostática entre heteroátomos eletronegativos, tais como o flúor, o oxigênio, o nitrogênio e, até mesmo, o cloro ou o bromo e o átomo de hidrogênio de ligações O-H e N-H (DUARTE; FREITAS, 2009; ROCHA, 2001).

Assim, o átomo de hidrogênio que está ligado a outro muito eletronegativo fica com uma carga parcial significativamente positiva, enquanto o heteroátomo eletronegativo fica com a carga parcial negativa, sendo a ligação de hidrogênio o resultado de suas polarizações (ROCHA, 2001). Além disso, as ligações de hidrogênio podem ser classificadas como intramolecular, ou seja, ocorre entre átomos de uma mesma molécula e pode interferir no isomerismo conformacional, e ligação de hidrogênio intermolecular, que ocorre entre moléculas e reflete na organização dos agregados moleculares (DUARTE; FREITAS, 2009).

As ligações de hidrogênio são extremamente importantes na maior parte dos processos de interesse biológico, ainda mais quando se considera que a maior parte das interações entre moléculas biológicas acontece em solução aquosa que, muitas vezes, atuam como pontes para esse tipo de ligação (MONTANARI, 2011).

Moléculas orgânicas fluoradas têm sido objeto de investigações, uma vez que a porção C-F pode atuar como acceptora de prótons em ligações de hidrogênio do tipo F...H (SILLA, 2013). Além disso, o átomo de flúor é magneticamente ativo, sendo adequado, por exemplo, para estudos de ressonância magnética nuclear (RMN), o que fornece informações estruturais

por meio dos parâmetros deslocamento químico e constante de acoplamento (MACOMBER, 1997; SILVERSTEIN; WEBSTER; KIMBLE, 2006).

De forma análoga à ligação de hidrogênio, tem-se a ligação de halogênio, que trata da interação não covalente entre um átomo de halogênio X (F, Cl, Br e I) e uma molécula com sítio negativo, como um par de elétrons isolado de uma base de Lewis (X...:B) ou outros com características estruturais e espectroscópicas semelhantes às da ligação de hidrogênio (CLARK et al., 2006).

Na detecção e na caracterização das interações moleculares podem ser utilizados parâmetros clássicos obtidos de dados espectroscópicos, bem como de ferramentas teóricas, tais como as metodologias computacionais QTAIM (do inglês *quantum theory of atoms in molecules*) e NBO (do inglês *natural bond orbital*).

3.3 Cálculos de orbitais

A análise dos orbitais naturais de ligação (NBO) é largamente empregada e consiste na avaliação dos efeitos de deslocalização eletrônica, com base em átomos ou grupos doadores e aceptores de elétrons. Assim, é feita a interpretação de uma função de onda molecular de vários elétrons em termos dos pares de elétrons ligantes (WEINHOLD; LANDIS, 2012).

Os orbitais moleculares obtidos por algum método de cálculo são convertidos em orbitais localizados, que têm a máxima ocupação, ou seja, uma aproximação da estrutura de Lewis. Como resultado (Equação 1) tem-se a energia estabilizante de hiperconjugação (E_{Hiper}), que é o resultado da diferença da energia total eletrônico-nuclear (E_{Total}), antes de ser feita a transformação de seus orbitais para NBOs e a energia de Lewis (E_{Lewis}), que representa contribuições estéricas e eletrostáticas (CORMANICH, 2011).

$$E_{\text{Hiper}} = E_{\text{Total}} - E_{\text{Lewis}} \quad (4)$$

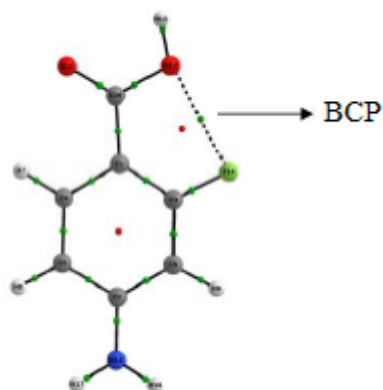
Além de sua aplicação em estudos dos efeitos de origem clássica ou quântica, os cálculos NBO são úteis na identificação de ligações de hidrogênio, considerando a doação de elétrons do acceptor de prótons para o orbital σ_{X-H}^* (doador de próton), em que X corresponde a um átomo eletronegativo. De forma geral, o método permite avaliar a ocorrência de interações intramoleculares mais específicas, que podem ajudar na determinação da preferência conformacional, segundo critérios que envolvem a geometria e a energia de cada confômero.

3.4 QTAIM

A teoria quântica dos átomos em moléculas (QTAIM), desenvolvida por Bader (1991), é utilizada na interpretação de calores de reação e tensão angular de cicloalcanos, constante de acidez de ácidos orgânicos; no estudo do comportamento de átomos, quando formam ligações covalentes e na identificação de ligações de hidrogênio (FONSECA et al., 2012).

Resumidamente, por serem grandezas vetoriais, todas as trajetórias dos vetores gradientes da densidade eletrônica têm direção e sentido que, na prática, são bem definidos por um ponto específico no espaço, denominado atrator. As trajetórias do gradiente da densidade eletrônica se organizam para descrever um sistema químico em um ponto localizado entre dois atratores chamado ponto crítico de ligação (*bond critical point*, BCP), sendo as duas trajetórias $\nabla\rho(\mathbf{r})$ formadas a partir de um único BCP denominadas linhas de ligação (*bond path*, BP) e configuram uma ligação química (OLIVEIRA et al., 2009).

Figura 4 - Gráfico molecular de um dos confôrmeros do ácido aminofluorobenzoico.



Fonte: Adaptado de Silla et al. (2013).

Koch e Popelier (1995) introduziram alguns critérios baseados na QTAIM para a identificação de ligações de hidrogênio, sintetizados a seguir.

- a) Deve haver a presença de um BCP e um BC entre o átomo de hidrogênio e o receptor de hidrogênio que participam da interação.
- b) O valor de densidade eletrônica $HBCP$ observado na ligação de hidrogênio deve estar no intervalo de 0,002-0,004 ua.
- c) O laplaciano $\nabla^2 \rho_{H...Y}$ correspondente à densidade eletrônica deve estar no intervalo de 0,024-0,139 ua.
- d) Interpenetração, ou seja, uma ‘penetração mútua’ entre o átomo de hidrogênio e o átomo receptor de hidrogênio na LH deve ser observada.
- e) O átomo de hidrogênio deve perder carga atômica $q(H)$ e, em consequência dessa perda, deve aumentar em energia E e decrescer seus valores de seu primeiro momento de dipolo $M_1(H)$ e de volume atômico $V(H)$.

Por meio desses critérios é possível afirmar que uma interação que não obedeça a um ou mais critérios não pode ser classificada como ligação de hidrogênio e é pouco provável que seja importante no equilíbrio conformacional. Vale ressaltar que os critérios de Popelier para alguma estrutura devem ser mensurados, utilizando, como referência, alguma estrutura análoga ou confôrmero que não apresentem ligação de hidrogênio.

3.5 QSAR

O planejamento de fármacos e compostos bioativos de forma racional, com o auxílio de métodos computacionais, apresenta o potencial de contornar o problema gerado pelo elevado tempo e o alto custo relacionado aos métodos convencionais (AFANTITIS et al., 2006; MARTINS, 2010).

Diante disso, a relação quantitativa entre estrutura-atividade (QSAR, do inglês *quantitative structure-activity relationships*) vem se destacando em química computacional como uma eficiente técnica que considera, simultaneamente, a estrutura química e sua respectiva atividade, obtendo relações quantitativas entre elas (MONTANARI, 2011).

Tais relações são basicamente obtidas por meio da transformação da informação química de determinado composto, mais relevante para uma dada propriedade, em uma série de descritores numéricos, a fim de que possam ser correlacionados com a atividade biológica que se deseja otimizar (NEVESA et al., 1998).

Embora as técnicas utilizadas em QSAR tenham tido início em 1863, quando Cross observou que a toxicidade de álcoois em mamíferos era inversamente proporcional à sua solubilidade em água (MARTINS, 2010), foi a partir da década de 1960, com o trabalho de Hansch e Fujita (1964), que os

modelos QSAR começaram a ser utilizados em estudos que correlacionavam propriedades estruturais com a atividade exercida (MONTANARI, 2011).

Em 1988, as técnicas QSAR passaram por uma considerável transformação com a inserção de parâmetros moleculares tridimensionais, recebendo o nome de QSAR-3D. Introduzida por Cramer, a análise comparativa do campo molecular CoMFA, do inglês *comparative molecular field analysis*, é baseada na representação das moléculas em sua forma tridimensional, alinhadas no interior de uma caixa 3D virtual (CRAMER III; PATERSON; BUNCE, 1988; MARTINS, 2010).

Com o passar do tempo e diante das limitações e necessidades de melhorias na qualidade das previsões dos modelos QSAR-3D, as técnicas QSAR foram se tornando cada vez mais multidimensionais. O QSAR-4D foi desenvolvido por Hopfinger et al. (1997) e incorpora a liberdade conformacional aos modelos 3D por meio de simulação de dinâmica molecular (DAMALE et al., 2014). Além disso, permite considerar a geometria do receptor e incorporar tal informação ao método, recebendo o nome de QSAR-4D dependente do receptor (PAN; TSEN; HOPFINGER, 2003).

Ao permitir a representação da topologia do ligante ao sítio ativo do receptor, Vedani e Dobler (2002) desenvolveram o QSAR-5D, que ganhou mais uma dimensão ao possibilitar uma avaliação simultânea de diferentes modelos de solvatação (VEDANI; DOBLER; LILL, 2005). A necessidade de inclusão de modelos baseados na dimensão do receptor e na interação virtual com seu ligante levou ao desenvolvimento do chamado QSAR-7D (POLANSKI, 2009).

É importante ressaltar que a incorporação de novas dimensões ao método QSAR não torna o QSAR clássico (1D e 2D) ultrapassado. Ao contrário, descritores moleculares 2D não têm se mostrado inferiores aos tridimensionais, além de apresentarem a vantagem de serem mais convenientes e de exigirem

cálculos mais simples (CORMANICH; NUNES; FREITAS, 2012; TIAN; ZHOU; LI, 2007).

3.6 MIA-QSAR

Os avanços nas metodologias QSAR possibilitaram a aplicação de uma variedade de descritores multidimensionais na geração dos modelos (PINHEIRO et al., 2008). Contudo, os descritores moleculares 2D têm se mostrado comparáveis aos demais, sendo extremamente potentes quanto à conveniência, ao tempo e à simplicidade dos cálculos (CORMANICH; NUNES; FREITAS, 2012; THIAN; ZHOU; LI, 2007).

Dentre as metodologias QSAR bidimensionais atualmente disponíveis, destaca-se a análise multivariada de imagens aplicada em QSAR (MIA-QSAR do inglês *multivariate image analysis applied to QSAR*) (FREITAS; BROWN; MARTINS, 2005). Trata-se de um método essencialmente 2D que utiliza imagens de estruturas químicas de uma série congênere de compostos para a descrição de suas propriedades químicas, físicas e/ou biológicas (FREITAS, 2007).

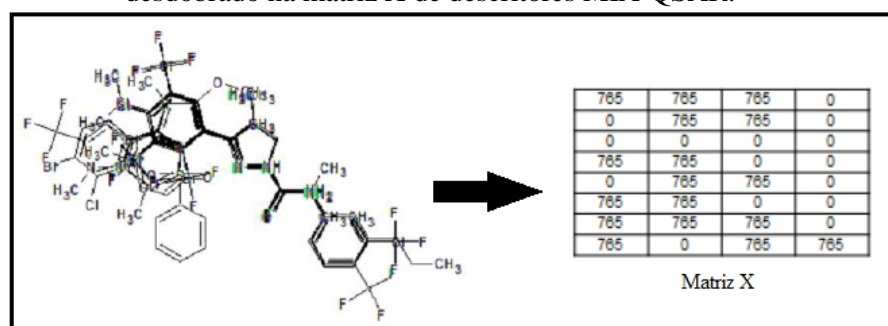
Tais estruturas, após serem sistematicamente desenhadas por meio de programas apropriados, são sobrepostas, formando um arranjo tridimensional e convertidas em *pixels* de imagens que darão origem a uma matriz X de descritores moleculares, a qual será, então, correlacionada com as respectivas atividades na geração do modelo QSAR (FREITAS; BROWN; MARTINS, 2005).

Mudanças na posição dos substituintes ou modificações estruturais causam alterações nas coordenadas dos *pixels* da imagem e são essas alterações que explicam as variações nas atividades exercidas ou nas propriedades moleculares (CORMANICH; NUNES; FREITAS, 2012).

Bons modelos QSAR estão diretamente relacionados a fatores como a escolha de uma série de compostos químicos que apresentem certa similaridade estrutural e dados químicos ou biológicos de qualidade, ou seja, experimentalmente conhecidos e obtidos conforme um mesmo protocolo de ensaio. Além disso, a escolha dos descritores e dos métodos estatísticos adotados reflete na qualidade da predição alcançada (MARTINS, 2010; MONTANARI, 2011).

Inicialmente, os modelos MIA-QSAR eram construídos a partir de imagens em preto e branco, desenhadas por meio de programas em duas dimensões, tais como o ChemDraw e o ChemSketch, e seus *pixels* eram tratados numericamente como binários, baseados no sistema de cores *Red-Green-Blue* (RGB), em que o *pixel* preto corresponde ao dígito 0 (ausência de cor) e representa os átomos e as ligações químicas das moléculas e os *pixels* brancos são identificados pelo número 765, obtido pela soma do valor correspondente a cada cor do sistema RGB, ou seja, 255, e representam os espaços em branco presentes no desenho das estruturas químicas. Os *pixels* obtidos, após a exclusão dos espaços em branco e das porções congruentes a todas as estruturas do conjunto de dados sobrepostas, formam a matriz de descritores (Figura 5), utilizada na construção do modelo (FREITAS; BROWN; MARTINS, 2005).

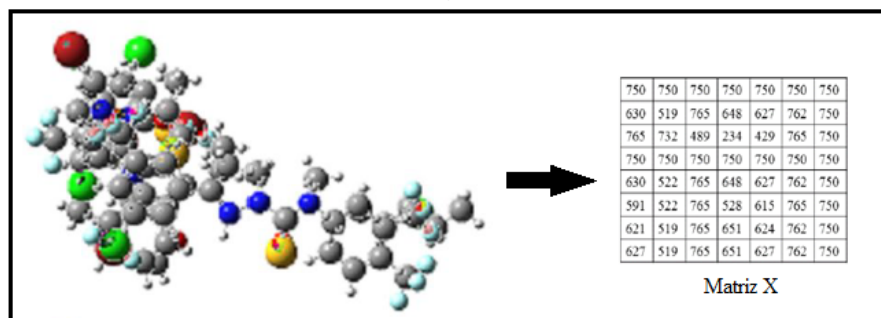
Figura 5 - Sobreposição das estruturas formando o arranjo tridimensional desdobrado na matriz X de descritores MIA-QSAR.



Fonte: Adaptado de Duarte et al. (2015).

Embora apresentassem resultados satisfatórios, a utilização de um sistema binário de *pixels* para representar as imagens das estruturas químicas apresentava, como desvantagem, o fato de não considerar informações, tais como o raio atômico e a natureza dos substituintes, comprometendo a interpretação do modelo. Tal limitação motivou a incorporação de informações que descrevessem melhor as estruturas químicas, dando origem a uma nova abordagem, denominada *aug*-MIA-QSAR (*aug* do inglês *augmented*), em que os diferentes átomos são representados por círculos com tamanhos proporcionais aos respectivos raios de van der Waals e em diferentes cores (Figura 6) (NUNES; FREITAS, 2013).

Figura 6 - Sobreposição das estruturas desenhadas considerando o raio de Van der Waals, com diferentes cores na representação atômica e obtenção da matriz X de descritores *aug*-MIA-QSAR.



Fonte: Adaptado de Duarte et al. (2015).

Nessa abordagem, o aumento na quantidade de *pixels* gerados está relacionado ao tamanho molecular e à inclusão de cores (ainda de acordo com sistema de cores RGB), além dos *pixels* pretos e brancos, e permite diferenciar os átomos e explicar variações nas propriedades atômicas, melhorando a capacidade preditiva do modelo (BARIGYE; FREITAS, 2016).

3.7 Ancoramento molecular ou *docking*

As técnicas QSAR-3D não levam em consideração a estrutura do receptor, uma vez que, na essência, apenas o ligante é utilizado nesse tipo de abordagem.

Nesse sentido, destacam-se os métodos de modelagem molecular baseados na estrutura da proteína, tais como os estudos de ancoramento molecular, ou *docking*, que correspondem a uma das mais importantes técnicas de investigação de compostos capazes de interagir com determinado sítio molecular (MONTANARI, 2011).

Os programas de docagem, tais como o Molegro Virtual Docker (MVD), calculam a interação ligante-receptor por meio de suas estruturas tridimensionais obtidas por modelos de comparação ou homologia, bem como por experimentos de cristalografia de raios X ou espectroscopia de ressonância magnética nuclear (RMN) (THOMSEN; CHRISTENSEN, 2006).

Os valores da função de desempenho do ancoramento E_{score} são definidos pela soma das energias intra e intermoleculares, conforme Equação 5.

$$E_{score} = E_{inter} + E_{intra} \quad (5)$$

em que a energia de interação entre o ligante e a proteína é representada pela E_{inter} , que usa um parâmetro para a aproximação do termo estérico (van der Waals) entre átomos e outro para a ligação de hidrogênio, e o termo E_{intra} corresponde à energia interna do ligante que considera a energia de ligação e torção do átomo. Em suma, são funções utilizadas para ancorar automaticamente uma molécula flexível em uma molécula molde ou proteína (RAMALHO et al., 2009; THOMSEN; CHRISTENSEN, 2006).

Em conjunto com as metodologias QSAR, os estudos de *docking* podem colaborar na validação e na interpretação de dados bioquímicos ao avaliar a

potencialidade do ligante, predizer sua orientação e conformação no sítio de ligação, sendo essenciais na compreensão do mecanismo de interação de um fármaco ou compostos bioativos com o seu alvo biológico (MONTANARI, 2011).

3.8 Análise qualitativa dos dados

Por meio da análise qualitativa dos dados é possível identificar a contribuição das diferentes variáveis para o modelo construído e compreender as informações que elas codificam.

Uma vez que é comum a geração de uma grande quantidade de descritores em modelos QSAR, procedimentos de seleção de variáveis podem ser úteis para a obtenção das informações mais relevantes por meio da identificação e da exclusão das variáveis redundantes e com baixa variância e/ou alta correlação (FERREIRA; MONTANARI; GÁUDIO, 2002).

Neste processo pode ser utilizada, de acordo com a literatura, uma variedade de métodos para a seleção de variáveis que vão desde filtros que funcionam como um pré-processamento, envoltórios (*wrappers*) e métodos embutidos (*embedded methods*) (GUYON; ELISSEFF, 2013), até os métodos classificados como supervisionados e não supervisionados (URIAS et al., 2015). Os métodos supervisionados selecionam variáveis de acordo com critérios pré-definidos, como valores de atividade (variável dependente Y) e incluem a regressão por mínimos quadrados parciais (PLS, do inglês *partial least squares*) e os coeficientes de correlação de Pearson, dentre outros. Já os métodos não supervisionados utilizam apenas as variáveis independentes (X) e, de forma geral, incluem a análise de agrupamento hierárquico, a análise de componentes principais (PCA, do inglês *Principal Component Analysis*) e a entropia de

Shannon, por exemplo (KLEBE, ABRAHAM; MIETZNER, 1994; URIAS et al., 2015).

No presente trabalho, a matriz de descritores apresentou alta dimensionalidade e, por esse motivo, foram empregados procedimentos como a entropia de Shannon, classificado como um método não supervisionado que reduz o número de variáveis de acordo com a sua relevância para a atividade exercida. Em seguida, foi aplicado o coeficiente de correlação (r), em que, para cada conjunto de variáveis com um coeficiente de correlação 0,98, uma variável foi retirada (GUIMARÃES et al., 2016).

Após a seleção de variáveis, foi aplicada a análise de componentes principais (PCA), que é uma ferramenta de reconhecimento de padrões que classifica as amostras de acordo com a similaridade. Nesse método, um novo conjunto de variáveis (eixo) é gerado e recebe o nome de componentes principais (PC), que são completamente não correlacionadas e ortogonais entre si, ou seja, a informação contida em uma delas não está presente em outra (MONTANARI, 2011).

Como as PCs são ortogonais, é possível examinar as relações entre amostras (*scores*) e variáveis (*loadings*) por meio dos gráficos gerados, além de estimar a influência de cada variável em cada amostra (PANERO et al., 2009). No presente estudo, foi realizada a seleção das variáveis de maior relevância no modelo QSAR construído e, por meio da PCA, foi possível classificar o conjunto de amostras de acordo com a intensidade da atividade anestésica exercida (atividade baixa, média e alta).

3.9 Compostos organofluorados

O flúor é o elemento químico mais eletronegativo entre os elementos conhecidos, com menor raio atômico, nove prótons, nove elétrons e massa

atômica de 18,99840 u (BOECHAT; PINTO; BASTOS, 2015; BRANCO et al., 2015).

Crescentes investimentos em P&D, nas últimas décadas, visam explorar as propriedades particulares do flúor, como sua alta eletronegatividade, que resulta em significantes alterações das características eletrônicas do composto e reatividade de grupos vizinhos, pequeno raio de van der Waals (que causa pouco impacto estérico), maior resistência da ligação C-F, quando comparada à de C-H, resultando no aumento da estabilidade química e térmica da molécula, além da maior lipofilicidade dos compostos que contêm flúor, comparada à de análogos não fluorados (BRANCO et al., 2015; SHIMIZU; HIYAMA, 2005).

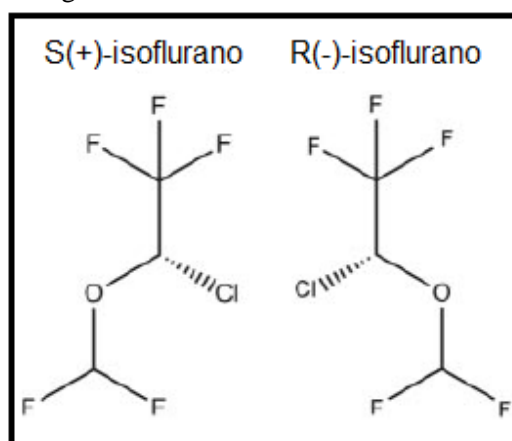
Atualmente, compostos orgânicos contendo flúor, conhecidos como organofluorados, apresentam diversas aplicações industriais e estão presentes em nosso cotidiano, como em gases refrigerantes, extintores de incêndio, solventes, surfactantes, propelentes em aerossóis, bateria de lítio recarregável, polímeros de alta resistência química e térmica, cristais líquidos, radiofármacos em tomografia por emissão de pósitrons, agroquímicos e, principalmente, em fármacos (BOECHAT; PINTO; BASTOS, 2015).

O isoflurano [2-cloro-2-(diflúor-metóxi)-1,1,1-(trifluoroetano)] é um éter halogenado utilizado como anestésico inalatório por vaporização, geralmente em conjunto com o oxigênio, desde a década de 1980 (CHIAO; ZUO, 2014). Considerado um anestésico seguro indicado na indução e na manutenção da anestesia geral, apresenta-se como um líquido estável, não inflamável e incolor, sem aditivos ou conservantes químicos e com odor pungente, fato que tem restringido a sua aplicação em pacientes pediátricos (RENAME, 2010).

De forma semelhante ao desflurano e ao enflurano, o isoflurano apresenta um carbono assimétrico e existe como uma mistura de seus enantiômeros (+) e (-) (Figura 7). O (+)-isoflurano é cerca de 50% a 60% mais

potente que seu enantiômero (MEHDIPOUR; HEMMATEENEJAD; MIRI, 2007). Tal fato destaca a necessidade de um entendimento detalhado das interações em nível molecular e o conhecimento de suas propriedades conformacionais (HERMANN; MACK; OBERHAMMER, 2000).

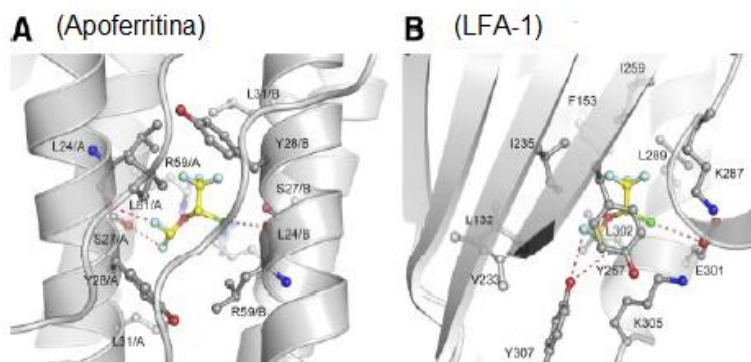
Figura 7 - Enantiômeros do isoflurano.



Fonte: Zhang et al. (2009).

Assim como os demais anestésicos inalatórios, o isoflurano não apresenta mecanismo de ação totalmente esclarecido, apesar de inúmeras pesquisas já terem sido realizadas nesse sentido. Sua atuação foi identificada sobre uma variedade de alvos moleculares, tais como o sítio alostérico da integrina LFA-1 (*lymphocyte function associated antigen-1*) (ZHANG et al., 2009) e a hélice 4 da apoferritina (LIU; PATRICK; RODERIC, 2005) (Figura 8), além dos receptores GABA, de glutamato e de glicina (LESARRI et al., 2011), fato que reforça os indícios da ausência de um local de ação específico.

Figura 8 - (A) Isoflurano ligado à hélice 4 da apoferritina (código PDB: 3F78) e (B) isoflurano ligado à integrina LFA-1 (código PDB: 1XZ3).

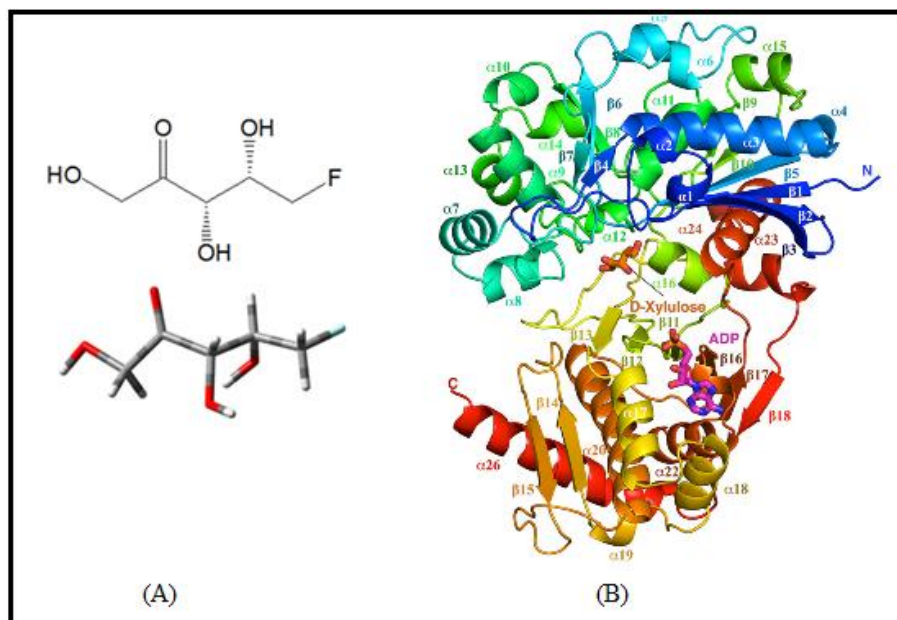


Fonte: Zhang et al. (2009).

Embora o equilíbrio conformacional do isoflurano tenha sido estudado anteriormente (HÉNIN et al., 2010; HERMANN; MACK; OBERHAMMER, 2000; MICHALSKA et al., 2007; PFEIFFER; MACK; OBERHAMMER, 1998) e que até mesmo seu efeito anomérico tenha sido avaliado em fase gasosa (LESARRI et al., 2011), uma análise minuciosa dos efeitos operantes na fase gasosa e em solução, bem como a comparação com suas estruturas bioativas, ainda não foi realizada, fato que justifica o desenvolvimento do presente trabalho.

Outro composto organofluorado utilizado nesse estudo é a 5-deóxi-5-flúor-D-xilulose (DFX) (Figura 9). Sua forma bioativa, disponível no *Protein Data Bank* com registro 5FX, liga-se à enzima xiluloquinase (código PDB: 4BC5), a qual catalisa a fosforilação dependente de ATP da D-xilulose para produzir a xilulose-5-fosfato, que é um regulador chave do metabolismo da glicose e da lipogênese (BUNKER et al., 2013).

Figura 9 - (A) Estrutura química da 5-deóxi-5-flúor-D-xilulose e a respectiva estrutura bioativa (códigos PDB: 5FX); (B) enzima xiluloquinase (código PDB: 4BC5).



Fonte: Bunker et al. (2013).

Como ligante livre, a DFX apresenta suas conformações preferenciais regidas por interações intramoleculares, tais como ligações de hidrogênio e interações hiperconjugativas, que ainda não foram estudadas. Diante disso, no presente trabalho foram desenvolvidos estudos conformacionais da DFX em sua forma livre da enzima xiluloquinase, a fim de elucidar os efeitos das interações intramoleculares em suas conformações preferenciais e compará-las à sua forma bioativa disponível no PDB.

As estruturas otimizadas foram ancoradas no sítio ativo da enzima xiluloquinase, por meio do programa Molegro Virtual Docker (MVD) (THOMSEN; CHRISTENSEN, 2006), para identificar as possíveis ligações intermoleculares e compará-las com as interações presentes na ligação entre a 5FX e a enzima xiluloquinase. Assim, foi possível reforçar a hipótese de que as

conformações obtidas em um ambiente livre do receptor devem ser utilizadas com cautela em métodos QSAR, uma vez que nem sempre coincidem com as formas bioativas. Por fim, a compreensão da interação ligante-proteína poderá colaborar no desenho racional de compostos congêneres, sendo útil em estudos futuros.

REFERÊNCIAS

- AFANTITIS, A. et al. A novel simple QSAR model for the prediction of anti-HIV activity using multiple linear regression analysis. **Molecular Diversity**, New York, v. 10, p. 405-414, 2006.
- ALABUGIN, I. V. et al. Hyperconjugation. **Computacional Molecular Science**, New York, v. 1, p. 109-141, Feb. 2011.
- ALCÁCER, L. **Introdução à química quântica computacional**. Rio de Janeiro: IST, 2007. 324 p.
- ALLOUCHE, A. Gabedit: a graphical user interface for computational chemistry softwares. **Journal of Computational Chemistry**, New York, v. 32, p. 174-182, Jan. 2011.
- ANDERSON, C. B.; SEPP, D. T. Conformational and the anomeric effect in 2-oxy-substituted tetrahydropyrans. **Tetrahedron**, Philadelphia, v. 24, n. 4, p. 1707-1716, 1968.
- ARROIO, A. et al. O ensino de química quântica e o computador na perspectiva de projetos. **Química Nova**, São Paulo, v. 28, n. 2, p. 360-363, 2005.
- BADER, R. F. W. A quantum theory of molecular structure and its applications. **Chemical Reviews**, Washington, v. 91, n. 5, p. 893-928, July 1991.
- BARIGYE, S. J.; FREITAS, M. P. Ten years of the MIA-QSAR strategy: historical development and applications. **International Journal of Quantitative Structure-Property Relationships**, New Delhi, v. 1, n. 1, p. 64-77, Jan./June 2016.
- BOECHAT, N.; PINTO, A. da C.; BASTOS, M. M. Métodos seletivos de fluoração de moléculas orgânicas. **Química Nova**, São Paulo, v. 38, n. 10, p. 1323-1338, 2015.
- BÖES, E. S. **Estudo teórico da solvatação de ânions em solventes orgânicos**. Porto Alegre: Ed. UFRGS, 2005.
- BRANCO, F. S. C. et al. Ressonância magnética nuclear de substâncias organofluoradas: um desafio no ensino de espectroscopia. **Química Nova**, São Paulo, v. 38, n. 9, p. 1237-1246, 2015.

BUISSONNEAUD, D. Y.; MOURIK, T. van; O'HAGAN, D. A DFT study on the origin of the fluorine gauche effect in substituted fluoroethanes.

Tetrahedron, Philadelphia, v. 66, n. 12, p. 2196-2202, Mar. 2010.

BUNKER, R. D. et al. Structure and function of human Xylulokinase, an enzyme with important roles in carbohydrate metabolism. **Journal of Biological Chemistry**, Bethesda, v. 288, p. 1643-1652, 2013.

CAMPOS, A. F. C. **Introdução à modelagem molecular**. Brasília, DF: Ed. UnB, 2013.

CANCES, E.; MENNUCCI, B. New applications of integral equations methods for solvation continuum models: ionic solutions and liquid crystals. **Journal of Mathematical Chemistry**, Marne-la-Vallée, v. 23, n. 3, p. 309-326, Feb. 1998.

CARVALHO, I. et al. Introdução a modelagem molecular de fármacos no curso experimental de química farmacêutica. **Química Nova**, São Paulo, v. 26, p. 428-438, 2003.

CHAI, J. D.; HEAD-GORDON, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. **Physical Chemistry Chemical Physics**, London, v. 10, p. 6615-6620, Sept. 2008.

CLARK, T. et al. Halogen bonding: the σ -role. **Journal of Molecular Modeling**, Washington, v. 13, p. 291-296, Aug. 2006.

CHIAO, S.; ZUO, Z. A Double-edged sword: volatile anesthetic effects on the neonatal brain. **Brain Sciences**, Basel, v. 4, n. 2, p. 273-294, 2014.

COELHO, L. W. et al. Aplicações de mecânica molecular em química inorgânica. **Química Nova**, São Paulo, v. 22, p. 396-404, 1999.

COHEN, A. J.; MORI-SÁNCHEZ, P.; YANG, W. Challenges for density functional theory. **Chemical Reviews**, Washington, v. 112, n. 1, p. 289-320, 2012.

CORMANICH, R. A. **Análise conformacional de alguns aminoácidos e ésteres metílicos de aminoácidos**. Campinas: Ed. UNICAMP, 2011.

CORMANICH, R. A.; NUNES, C. A.; FREITAS, M. P. Desenhos de estruturas químicas correlacionam-se com propriedades biológicas - MIA-QSAR. **Química Nova**, São Paulo, v. 35, n. 6, p. 1157-1163, 2012.

CRAMER, C. J. **Essentials of computational chemistry: theories and models**. 2nd ed. Chichester: Wiley, 2004.

CRAMER III, R. D.; PATTERSON, D. E.; BUNCE, J. D. Comparative molecular field analysis (CoMFA): I., effect of shape on binding of steroids to carrier proteins. **Journal of the American Chemical Society**, Easton, v. 110, p. 5959-5967, 1988.

DAMALE, M. G. et al. Recent advances in multidimensional QSAR(4D-6D): a critical review. **Mini Reviews in Medicinal Chemistry**, Sharjah, v. 14, n. 1, p. 35-55, Jan. 2014.

DUARTE, C. J.; FREITAS, M. P. Hydrogen bonding and stereoelectronic effects in the conformational isomerism of trans-2-bromocyclohexanol. **Journal of Molecular Structure**, Amsterdam, v. 930, p. 135-139, July 2009.

DUARTE, M. H. et al. Exploring MIA-QSARs for antimalarial quinolon-4(1H)-imines. **Combinatorial Chemistry & High Throughput Screening**, Beijing, v. 18, p. 208-216, 2015.

FERNANDES, F. M. S. S. Perspectivas da química computacional. **Química**, Lisboa, v. 123, p. 47-53, out./dez. 2011.

FERREIRA, M. M. C.; MONTANARI, C. A.; GÁUDIO, A. C. Seleção de variáveis em QSAR. **Química Nova**, São Paulo, v. 25, p. 439-448, 2002.

FONSECA, T. A. O. et al. F...HO intramolecular hydrogen bond as the main transmission mechanism for $^1h/f, H(O)$ coupling constant in 2'-fluoroflavonol. **Magnetic Resonance in Chemistry**, Chichester, v. 50, n. 8, p. 551-556, Aug. 2012.

FOYE, W. O.; LEMKE, T. L.; WILLIAMS, D. A. **Principles of medicinal chemistry**. Baltimore: Lea & Febiger, 1995.

FRAGA, C. A. M. Razões da atividade biológica: interações micro- e biomacromoléculas. **Cadernos Temáticos de Química Nova na Escola**, São Paulo, n. 3, p. 33-42, maio 2001.

FREITAS, M. P. The anomeric effect on the basis of NBO analysis. **Organic & Biomolecular Chemistry**, Cambridge, v. 11, n. 17, p. 2885-2890, Mar. 2013.

FREITAS, M. P. Multivariate QSAR: from classical descriptors to new perspectives. **Current Computer-Aided Drug Design**, Oxford, v. 3, n. 4, p. 235-239, Apr. 2007.

FREITAS, M. P.; BROWN, S. D.; MARTINS, J. A. MIA-QSAR: a simple 2D image-based approach for quantitative structure-activity relationship analysis. **Journal of Molecular Structure**, Amsterdam, v. 738, n. 1, p. 149-154, Mar. 2005.

GOODMAN, L.; GU, H.; POPHRISTIC, V. Gauche effect in 1,2-difluoroethane: hyperconjugation, bent bonds, steric repulsion. **The Journal of Physical Chemistry A**, Washington, v. 109, n. 6, p. 1223-1229, 2005.

GRINGAUZ, A. **Introduction to medicinal chemistry: how drug act and why**. Weinheim: Wiley-VCH, 1997.

GUIMARÃES, M. C. et al. Is conformation a fundamental descriptor in QSAR?: a case for halogenated anesthetics. **Beilstein Journal of Organic Chemistry**, Frankfurt, v. 12, p. 760-768, 2016.

GUYON, I.; ELIESSEFF, A. An introduction to variable and feature selection. **The Journal of Machine Learning Research**, Edinburg, v. 3, p. 1157-1182, Mar. 2013.

HANSCH, C.; FUJITA, T. p - σ - π Analysis: a method for the correlation of biological activity and chemical structure. **Journal of the American Chemical Society**, Easton, v. 86, n. 8, p. 1616-1626, Nov. 1964.

HÉNIN, J. et al. An atomistic model for simulations of the general anesthetic isoflurane. **The Journal of Physical Chemistry B**, Uppsala, v. 114, n. 1, p. 604-612, 2009.

HERMANN, A.; MACK, H. G.; OBERHAMMER, H. Conformations and structures of desflurane and isoflurane. **Journal of Fluorine Chemistry**, Amsterdam, v. 101, n. 2, p. 223-231, Feb. 2000.

HOHENBERG, P.; KOHN, W. Inhomogeneous electron gas. **Physical Review B**, College Park, v. 136, n. 3, p. 864-871, Sept. 1964.

HOPFINGER, A. et al. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. **Journal of the American Chemical Society**, Washington, v. 119, n. 43, p. 10509-10524, Oct. 1997.

KADUK, B.; KOWALCZYK, T.; VOORHIS, T. V. Constrained density functional theory. **Chemical Reviews**, Washington, v. 112, n. 1, p. 321-370, 2012.

KLEBE, G.; ABRAHAM, U.; MIETZNER, T. J. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. **Medicinal Chemistry**, Washington, v. 37, n. 24, p. 4130-4146, Nov. 1994.

KOCH, U.; POPELIER, P. L. A. Characterization of C-H-O hydrogen bonds on the basis of the charge density. **Journal of Physical Chemistry**, Ithaca, v. 99, n. 24, p. 9747-9754, June 1995.

LESARRI, A. et al. Structural evidence of anomeric effects in the anesthetic isoflurane. **Physical Chemistry Chemical Physics**, Cambridge, v. 13, p. 6610-6618, 2011.

LIU, R.; PATRICK, J. L.; RODERIC, G. E. Structural basis for high affinity volatile anesthetic binding in a natural 4-helix bundle protein. **The FASEB Journal**, New York, v. 19, p. 567-576, Apr. 2005.

LIU, S. B.; GOVIND, N. Toward understanding the nature of internal rotation barriers with a new energy partition scheme: ethane and *n*-butane. **Journal of Physical Chemistry A**, Washington, v. 112, p. 6690-6699, July 2008.

MACOMBER, R. S. **A complete introduction to modern nmr spectroscopy**. New York: Wiley-Interscience, 1997. 400 p.

MARQUES, M. A. L.; BOTTI, S. O que é e para que serve a teoria dos funcionais da densidade? **Gazeta de Física**, Lisboa, v. 29, n. 4, p. 10-15, 2006.

MARTINS, J. P. A. **Química computacional aplicada a QSAR**. São Paulo: CENAPAD, 2010. Disponível em:
<<http://www.cenapad.unicamp.br/servicos/treinamentos/apostilas/apostila-QuimComp.pdf>>. Acesso em: 10 abr. 2017.

MEHDIPOUR, A. R.; HEMMATEENEJAD, B.; MIRI, R. QSAR studies on the anesthetic action of some polyhalogenated ethers. **Chemical Biology & Drug Design**, Oxford, v. 69, n. 5, p. 362-368, May 2007.

MICHALSKA, D. et al. Theoretical and experimental studies of enflurane. infrared spectra in solution. in low-temperature argon matrix and blue shifts resulting from dimerization. **The Journal of Physical Chemistry B**, Washington, v. 111, n. 42, p. 12228-12238, Oct. 2007.

MO, Y. Computacional evidence that hyperconjugative interactions are not responsible for the anomeric effect. **Nature Chemistry**, London, v. 2, n. 8, p. 666-671, July 2010.

MÖLLER, C.; PLESSET, M. S. Note on an approximation treatment for many-electron systems. **Physical Review**, New York, v. 46, p. 618-622, 1934.

MONTANARI, C. A. **Química medicinal: métodos e fundamentos em planejamento de fármacos**. São Paulo: EDUSP, 2011. 710 p.

MORGON, N. H.; CUSTÓDIO, R. **Funções de base: o ajuste variacional: texto didático para offsite ChemKeys**. Campinas: Chemkeys, 2001. Disponível em: <<http://chemkeys.com/br/2001/02/18/funcoes-de-base-o-ajuste-variacional>>. Acesso em: 13 ago. 2017.

NAMBA, A. M.; SILVA, V. B.; SILVA, C. H. T. P. Molecular dynamics: theory in applications in drug design. **Eclética Química**, Araraquara, v. 33, n. 4, p. 13-24, 2008.

NEVESA, P. J. D. et al. Top: um programa de cálculo de descritores topológicos para uso em correlações entre estrutura e atividade. **Química Nova**, São Paulo, v. 21, n. 6, p. 709-713, jul. 1998.

NUNES, C.; FREITAS, M. P. Introducing new dimensions in MIA-QSAR: a case for chemokine receptor inhibitors. **European Journal of Medicinal Chemistry**, Paris, v. 62, p. 297-300, 2013.

OLIVEIRA, B. G. et al. Uma comprovação químico-quântica sobre a formação de ligações de hidrogênio secundárias em sistemas heterocíclicos intermoleculares. **The Electronic Journal of Chemistry**, Valparaíso, v. 1, p. 167-182, Mar. 2009.

ORTOLAN, A. O. **Apostila de práticas de química computacional**. Curitiba: Ed. UFPR, 2014.

PAN, D.; TSENG, Y.; HOPFINGER, A. J. Quantitative structure-based design: formalism and application of receptor-dependent RD-4D-QSAR analysis to a set of glucose analogue inhibitors of glycogen phosphorylase. **Journal of Chemical Information and Computer Sciences**, Washington, v. 43, n. 5, p. 1591-1607, 2003.

PANERO, F. S. et al. Aplicação da análise exploratória de dados na discriminação geográfica do quiabo do Rio Grande do Norte e Pernambuco. **Eclética Química**, Araraquara, v. 34, n. 3, p. 33-40, 2009.

PFEIFFER, A.; MACK, H. G.; OBERHAMMER, H. Enflurane: structure and conformational properties. **Journal of the American Chemical Society**, Easton, v. 120, n. 25, p. 6384-6388, 1998.

PINHEIRO, J. R. et al. Novel anti-HIV cyclotriazadisulfonamide derivatives as modeled ligand-and receptor approaches. **Bioorganic & Medicinal Chemistry**, Oxford, v. 16, n. 4, p. 1683-1690, Feb. 2008.

PLIEGO JÚNIOR, J. R. Modelos contínuos do solvente: fundamentos. **Química Nova**, São Paulo, v. 29, n. 3, p. 535-542, 2006.

POLANSKI, J. Receptor dependent multidimensional QSAR for modeling drug-receptor interactions. **Current Medicinal Chemistry**, Sharjah, v. 16, n. 25, p. 3243-3257, Sept. 2009.

RAMALHO, T. C. et al. Construction and assessment of reaction models of class I EPSP synthase: molecular docking and density functional theoretical calculations. **Journal of Biomolecular Structure and Dynamics**, New York, v. 27, n. 2, p. 195-208, Oct. 2009.

REICHARDT, C. **Solvents and solvent effects in organic chemistry**. New York: VCH, 1990.

REICHARDT, C. **Solvents and solvent effects in organic chemistry**. Weinheim: VHC, 1988.

RENAME. **Formulário terapêutico nacional**. 2. ed. Brasília, DF, 2010.

ROCHA, W. R. Interações intermoleculares. **Cadernos Temáticos de Química Nova na Escola**, São Paulo, n. 4, p. 31-36, maio 2001.

RODRIGUES, C. R. Modelagem molecular. **Cadernos Temáticos de Química Nova na Escola**, São Paulo, n. 3, p. 43-49, maio 2001.

SANTANNA, C. M. R. Métodos de modelagem molecular para estudo e planejamento de compostos bioativos. **Revista Virtual de Química**, Niterói, v. 1, n. 1, p. 49-57, 2009.

SHIMIZU, M.; HIYAMA, T. Modern synthetic methods for fluorine-substituted target molecules. **Angewandte Chemie**, Weinheim, v. 44, n. 2, p. 214-231, Dec. 2005.

SILLA, J. M. **Análise conformacional e interações intramoleculares em ácidos aromáticos fluorados e derivados**. Lavras: Ed. UFLA, 2013.

SILLA, J. M. et al. Conformational analysis and intramolecular interactions in aminofluorobenzoic acids. **The Journal of Physical Chemistry**, Washington, v. 117, n. 7, p. 1659-1664, Jan. 2013.

SILVERSTEIN, R. M.; WEBSTER, F. X.; KIEMLE, D. J. **Spectrometric identification of organic compounds**. 7th ed. Rio de Janeiro: LTC, 2006. 490 p.

SMITH, J. et al. Structural basis for activity of the RSK specific inhibitor, SLO 101. **Bioorganic & Medicinal Chemistry**, Oxford, v. 15, n. 14, p. 5018-5034, July 2007.

SOLOMONS, T. W.; FRYHLE, C. B. **Química orgânica**. 9. ed. Rio de Janeiro: LTC, 2009. 675 p.

SOUZA, R. F.; FREITAS, M. P. Conformational analysis and intramolecular interactions in 2-haloethanols and their methyl ethers. **Computational and Theoretical Chemistry**, Amsterdam, v. 964, n. 1, p. 155-159, 2011.

TAYLOR, J. B.; KENNEWELL, P. D. **Introductory medicinal chemistry**. New York: J. Wiley, 1981.

THOMSEN, R.; CHRISTENSEN, M. H. MolDock: a new technique for high-accuracy molecular docking. **Journal of Medicinal Chemistry**, Washington, v. 49, p. 3315-3321, 2006.

TIAN, F.; ZHOU, P.; LI, Z. A novel atom-pair hologram (APH) and its application in peptide QSARs. **Journal of Molecular Structure**, Amsterdam, v. 871, n. 1/3, p. 140-148, Dec. 2007.

TOMAS I, J.; MENNUCCI, B.; CAMMI, R. Quantum mechanical continuum solvation models. **Chemical Reviews**, Washington, v. 105, n. 8, p. 2999-3093, July 2005.

URIAS, R. W. P. et al. IMMAN: free software for information theory-based chemometric analysis. **Molecular Diversity**, Dordrecht, v. 19, n. 2, p. 305-319, May 2015.

VEDANI, A.; DOBLER, M. 5D-QSAR: the key for simulating induced fit? **Journal of Medicinal Chemistry**, Washington, v. 45, n. 11, p. 2139-2149, May 2002.

VEDANI, A.; DOBLER, M.; LILL, M. A. Combining protein modeling and 6D-QSAR: simulating the binding of structurally diverse ligands to the estrogen receptor. **Journal of Medicinal Chemistry**, Washington, v. 48, n. 11, p. 3700-3703, June 2005.

WEINHOLD, F.; LANDIS, C. R. **Discovering chemistry with natural bond orbitals**. New Jersey: J. Wiley, 2012. 336 p.

WOLFF, M. E. **Burger's medicinal chemistry and drug discovery**. New York: J. Wiley, 1995.

ZHANG, H. et al. Crystal structure of isoflurane bound to integrin LFA-1 supports a unified mechanism of volatile anesthetic action in the immune and central nervous systems. **The FASEB Journal**, Bethesda, v. 23, p. 2735-2740, Aug. 2009.

SEGUNDA PARTE - ARTIGOS

ARTIGO 1

**Is conformation a fundamental descriptor in QSAR? A case for halogenated
anesthetics**

Maria C. Guimarães, Mariene H. Duarte, Josué M. Silla and Matheus P. Freitas*

Address: Department of Chemistry, Federal University of Lavras, P. O. Box
3037, 37200-000, Lavras, MG, Brazil

Email: matheus@dqi.ufla.br

* Corresponding author

(Artigo publicado no *Beilstein Journal of Organic Chemistry*)

Abstract

An intriguing question in 3D-QSAR lies on which conformation(s) to use when generating molecular descriptors (MD) for correlation with bioactivity values. This is not a simple task because the bioactive conformation in molecule data sets is usually unknown and, therefore, optimized structures in a receptor-free environment are often used to generate the MD's. In this case, a wrong conformational choice can cause misinterpretation of the QSAR model. The present computational work reports the conformational analysis of the volatile anesthetic isoflurane (2-chloro-2-(difluoromethoxy)-1,1,1-trifluoro-ethane) in the gas phase and also in polar and nonpolar implicit and explicit solvents to show that stable minima (ruled by intramolecular interactions) do not necessarily coincide with the bioconformation (ruled by enzyme induced fit). Consequently, a QSAR model based on two-dimensional chemical structures was built and exhibited satisfactory modeling/prediction capability and interpretability, then suggesting that these 2D MD's can be advantageous over some three-dimensional descriptors.

Keywords

Isoflurane; conformational analysis; QSAR; volatile anesthetics; theoretical calculations

Resumo

Uma intrigante pergunta em QSAR-3D está em qual ou quais conformações usar na geração de descritores moleculares (MD) a serem correlacionados com os valores de atividade biológica. Essa não é uma tarefa simples, pois a conformação bioativa em conjuntos de dados moleculares, geralmente, é desconhecida e, portanto, as estruturas otimizadas em um ambiente livre do receptor são frequentemente utilizadas para gerar os descritores. Nesse caso, uma escolha conformacional incorreta pode causar má interpretação do modelo QSAR. O presente trabalho computacional trata da análise conformacional do anestésico volátil isoflurano [2-cloro-2- (diflúor-metóxi) -1,1,1-trifluoro-etano)] em fase gasosa e também na presença dos solventes implícitos e explícitos polares e não polares, a fim de mostrar que os mínimos de energia (governados por interações intramoleculares) não coincidem, necessariamente, com a conformação bioativa (governada pelo ajuste induzido pela enzima). Consequentemente, um modelo QSAR baseado em estruturas químicas bidimensionais foi construído e exibiu capacidade satisfatória de interpretação e modelagem - predição, sugerindo que descritores moleculares 2D podem ser vantajosos em relação a alguns descritores tridimensionais.

Palavras-chave

Isoflurano; análise conformacional; QSAR; anestésicos voláteis; cálculos teóricos.

Introduction

Quantitative structure-activity relationship (QSAR) studies try to find a correlation between chemical structures and the corresponding bioactivities by means of molecular descriptors (MD's). In this way, molecular architecture and substitution patterns in a series of congeneric molecules are described by calculable or empirical data having some relationship with biological activity, becoming the technique useful for understanding the action mechanism of related drugs, as well as to drive synthesis of new drug like compounds.

Since the milestone work by Hansch and Fujita [1], a variety of MD's have been developed to improve the correlation of chemical structure with bioactivity, ranging from hydrophobicity to tridimensional descriptors [2]. Indeed, the most popular QSAR methods are based on molecular descriptors generated from 3D molecular structures, such as the widely used techniques based on molecular field analysis [3, 4]. The problem with most 3D-QSAR methods is that the bioactive conformation of the compounds in a data set is usually unknown and, therefore, geometry optimization is carried out in a receptor-free environment to generate the molecular structure and, consequently, the 3D MD's. While molecular conformation in vacuum is governed by intramolecular interactions, the bioconformation is ruled by enzyme induced-fit; consequently, optimized and bioactive geometries are probably different and, to obtain insight on the action mechanism of a drug and substituent effects, MD's should not be generated over geometries optimized in a receptor-free environment. Efforts have been made to attenuate the drawback of using a possible wrong conformation, *e.g.* by using average conformations [5], but the risk of chemical-biological misinterpretation remains. Receptor-dependent QSAR methods have also been developed [6], but the extensive procedure of docking structures into an enzyme active site to generate 3D chemical structures for use in QSAR modeling sounds like work replication, since the docking

methodology itself provides intermolecular energies and docking scores that correlate with bioactivity. On the other hand, despite not encoding conformational information, 2D-QSAR can incorporate other stereochemical properties and also account for group sizes, substituent type and 2D shape of molecules. These have shown to be sufficient parameters to obtain satisfactory correlation with bioactivity data and valuable understanding on structural requirements for drug development.

Multivariate Image Analysis applied to QSAR (MIA-QSAR) is a genuine 2D method, since descriptors are pixels of molecular projections drawn as black and white wireframes [7] and, more recently, as colored 2D-images (chemical structures) with spheres representing atoms with sizes proportional to the respective van der Waals radii [8]. The variance in the atom color and coordinate of pixels in the images (the structural variance) explains the changes in the bioactivities block. Thus, the MIA-QSAR method is an appropriate approach to probe the validity of 2D-QSAR methods when the molecules in a data set undergo rotational isomerization.

In order to test our hypothesis, the biological activities of a set of volatile halogenated anesthetics were modeled using the recent version of the MIA-QSAR method. In addition, the suitability of applying 3D information obtained from high level calculations (in a receptor-free environment) in QSAR modeling was evaluated using a comparative study of the optimized and bioactive conformations of the fluorinated anesthetic isoflurane, which binds to a 4-helix bundle protein (apoferritin) [9] and to the integrin LFA1 enzyme [10].

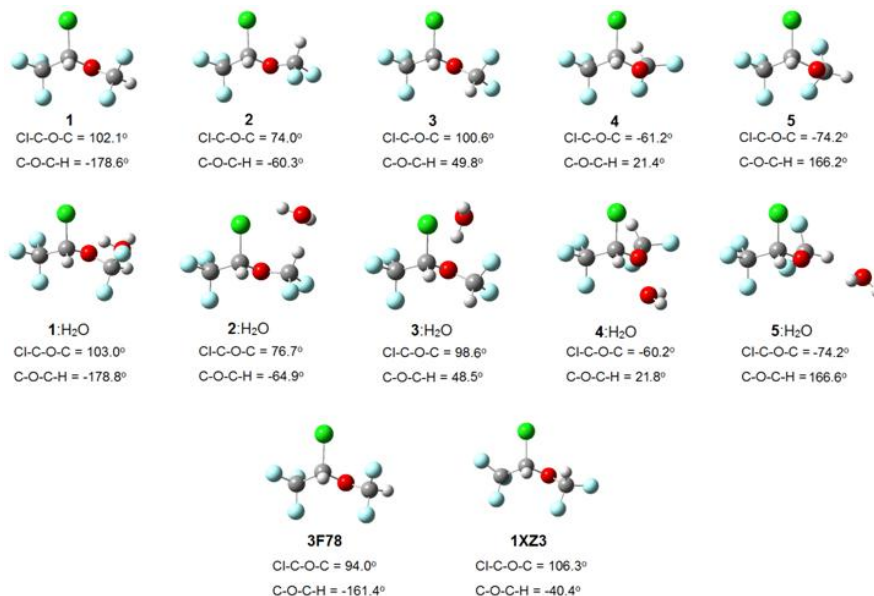
Results and Discussion

Conformational analysis

2-Chloro-2-(difluoromethoxy)-1,1,1-trifluoro-ethane (isoflurane) undergoes rotational isomerization around two dihedral angles (Cl-C-O-C and

C-O-C-H) and, considering three limit orientations for each of them (*gauche*, *gauche'* and *anti*), nine conformations are possible for isoflurane. However, geometry optimization at the MP2/6-311++g(d,p) and ω B97X-D/6-311++g(d,p) (a DTF method which includes dispersion effects) levels converged to five energy minima for the gas phase, implicit solvents (cyclohexane, DMSO and water, using the polarizable continuum model), and using one explicit water molecule as solvent to mimic a physiological medium (Figure 1).

Figure 1 - Optimized structures of isoflurane at the ω B97X-D/6-311++g(d,p) level (gas phase and explicit water) and the experimental bioconformations 3F78 and 1XZ3.



The results for the gas phase are in agreement with previous calculations [11] and microwave experiments [12], where five conformations were found, but only three could be experimentally detected, due to their lower energies. These correspond to conformers **1**, **2** and **3** of Table 1, the most stable ones at the gas phase and implicit solution. It is worth mentioning that, according to our calculations, the solvent plays little effect on the conformer populations, such as

in enflurane [13], suggesting that intramolecular interactions govern the conformational equilibrium in a biological-free environment. Indeed, Lesarri et al. [12] pointed out that anomeric effects owing to donor-acceptor ($LP \rightarrow \sigma^*$) interactions are responsible for the conformational preference of isoflurane. Despite the contribution from specific hyperconjugative interactions representing anomeric effects (mainly $LP_O \rightarrow \sigma^*_{CF}$, Table 2), our natural bond orbital (NBO) analysis indicates that the main conformers **1** and **2** are less favored by electronic delocalization than **3**, **4** and **5**. Since the overall energy of a system is a composition of repulsive Lewis (steric and electrostatic) and attractive non-Lewis-type (hyperconjugation) contributions, the main factor governing the conformational stabilization of isoflurane comes from the more classical steric and dipolar interactions, likewise sevoflurane reported earlier [14].

Table 1 - Relative electronic/Gibbs free energies (kcal mol^{-1} , % Gibbs population in parenthesis) for the conformers of isoflurane.^a

Medium	1	2	3	4	5
DFT					
Gas	0.0/0.0 (64)	0.7/0.4 (32)	2.5/1.9 (3)	2.3/2.4 (1)	5.3/6.2 (0)
C ₆ H ₁₂	0.0/0.0 (59)	0.7/0.3 (36)	2.0/1.6 (4)	2.3/2.6 (1)	5.3/6.2 (0)
DMSO	0.0/0.0 (37)	0.4/0.2 (31)	0.8/0.1 (31)	2.3/2.6 (1)	5.1/6.0 (0)
H ₂ O	0.0/0.0 (37)	0.4/0.2 (26)	0.7/0.0 (37)	2.3/2.5 (0)	5.1/6.0 (0)
H ₂ O _{explicit}	0.0/0.0 (85)	1.0/1.1 (14)	4.6/3.8 (0)	1.7/2.5 (1)	6.3/6.8 (0)
MP2					
Gas	0.0/0.0 (71)	0.8/0.6 (26)	2.5/1.9 (3)	2.8/3.0 (0)	5.5/6.4 (0)
C ₆ H ₁₂	0.0/0.0 (68)	0.8/0.6 (25)	2.1/1.3 (8)	2.9/3.2 (0)	5.4/6.1 (0)
DMSO	0.0/0.0 (51)	0.4/0.3 (31)	0.9/0.6 (18)	2.8/3.1 (0)	5.3/6.0 (0)
H ₂ O	0.0/0.0 (54)	0.8/0.5 (23)	0.8/0.5 (23)	2.8/3.1 (0)	5.3/6.0 (0)
H ₂ O _{explicit}	0.0/0.0 (88)	1.2/1.2 (12)	4.6/3.8 (0)	2.2/3.0 (0)	6.5/6.9 (0)

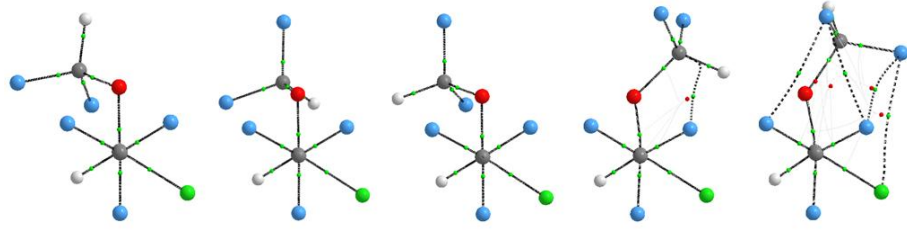
^a The converged geometries did not exhibit imaginary frequency and the standard Gibbs free energies were theoretically calculated under conditions 1.00 atm and 298.15 K, including electronic and thermal corrections.

Table 2 - Anomeric interactions and relative Lewis and non-Lewis contributions (kcal mol⁻¹) for the conformers of isoflurane, in the gas phase at the ω B97X-D/6-311++g(d,p) level.

Conformer	$n_{\text{O}} \rightarrow \sigma_{\text{CF}}^*$	$n_{\text{O}} \rightarrow \sigma_{\text{CCl}}^*$	$n_{\text{O}} \rightarrow \sigma_{\text{CCF}_3}^*$	$n_{\text{O}} \rightarrow \sigma_{\text{C}_2\text{H}}^*$	$n_{\text{O}} \rightarrow \sigma_{\text{C}_1\text{H}}^*$	E_{Lewis}	$E_{\text{non-Lewis}}$
1	15.5+2.4+15.9+2.6	17.6	4.2	3.9+1.1	3.7	0.0	0.3
2	15.4+2.4+7.3	1.0+21.3	3.4+0.8	5.4+2.7	0.6+7.7	0.4	0.0
3	7.0+0.8+20.1	19.2	1.8+4.9	4.4+1.0	2.3+5.1	10.9	8.7
4	6.1+5.3+18.6	3.1+18.1	2.1+9.4	4.6	3.3+1,7	6.9	4.9
5	2.0+17.4+3.6+11.6	1.1+19.1	3.8+6.4	4.4+0.6	4.0	18.1	13.1

It has been proposed that intramolecular hydrogen bond CH...FC drives the conformational preference of enflurane [15]. In order to check if such an interaction operates in isoflurane, QTAIM (Quantum Theory of Atoms in Molecules) calculations were performed for **1-5**. The molecular graphs of Table 3 do not exhibit F...H or Cl...H bond paths (only weak nonbonding interactions are observed for the high energy conformers **4** and **5**). According to Koch and Popelier [16], the following QTAIM parameters, obtained by integration over the atomic basins of the hydrogen atoms participating in hydrogen bonds, should be observed to characterize hydrogen bonds along with a bond path: loss of atomic charge (q), decreased first dipole moment (M_I), decreased atomic volume (V) and increased atomic energy (E). Since these criteria do not vary significantly among the conformers (either when H approaches to Cl and F or not), we can confirm that hydrogen bonding is not a decisive interaction for the conformational equilibrium of isoflurane.

Table 3 - Atomic properties (a.u.) obtained by QTAIM to characterize hydrogen bonds in isoflurane.



Conformer	q (H)	M_I (H)	V (H)	E (H)
1 _{H(CCl)/H(CF₂)}	+0.126/+0.126	+0.125/+0.127	+39.37/+41.03	-0.583/-0.589
2 _{H(CCl)/H(CF₂)}	+0.122/+0.112	+0.127/+0.133	+40.11/+41.94	-0.584/-0.591
3 _{H(CCl)/H(CF₂)}	+0.097/+0.093	+0.135/+0.136	+42.15/+42.84	-0.592/-0.599
4 _{H(CCl)/H(CF₂)}	+0.128/+0.119	+0.127/+0.127	+40.80/+38.93	-0.584/-0.591
5 _{H(CCl)/H(CF₂)}	+0.123/+0.118	+0.128/+0.129	+40.66/+41.30	-0.588/-0.593

Using an explicit water model of solvation, where specific solute-solvent interactions take place, conformer **1** is again the most stable form, suggesting that implicit solvation describes satisfactorily the actual conformational isomerism. However, is there an environment capable of overcoming intramolecular interactions and then changing the conformational preference of isoflurane? While comparison of **1** with the bioconformation of isoflurane bound to the integrin LFA1 enzyme (PDB code: 3F78 [10]) reveals small differences in the Cl-C-O-C and C-O-C-H dihedral angles ($\pm 8^\circ$ and 17° , respectively), the isoflurane structure bound to apoferritin (PDB code: 1XZ3 [9]) does not match any optimized conformer (Figure 1). So, does it make sense to use enzyme-free optimized conformations to obtain biochemical insights from 3D-QSAR? The following MIA-QSAR modeling can help to elucidate this question and to support the use of 2D approaches with the aim at obtaining biochemical information.

MIA-QSAR of anesthetic haloethers

MIA-QSAR is a genuine 2D technique in the sense that it takes into account 2D projections of chemical structure images to generate MD's. Each compound of a congeneric series is aligned one another by a congruent motif and, therefore, the variable substructure along the series explains the variance in the bioactivities block. The variance in the chemical structures is captured by the different coordinates of the pixels composing the images (Figure 2 shows the superimposed images used in this work to illustrate the structural variance). Pixels also vary in color, depending on the atom to which it refers. Since pixel values are a summation of RGB (red-green-blue) components and each channel is numerically equivalent to 255, the whole spectrum of colors can vary from 0 (black) to 765 (white). The pixel value (atom color) can be managed according to atomic properties and, therefore, each different atom in the series of 25

haloethers of Table 4 [17,18] was colored according to the respective Pauling's electronegativity (ϵ) scale, since polar properties should help to modulate their mode of interaction with an enzyme (see Table 5 for correspondence of ϵ with pixel values and approximate colors). Another useful parameter ruling bioactivity is the steric effect, which was accounted for by representing atoms in the molecules as colored spheres with sizes proportional to the corresponding van der Waals radii. The molecules were constructed using the GaussView program [19] and each 2D projection (each image, molecule) with $m \times l$ dimension was unfolded to a row vector. Combination of the 25 images yielded a data matrix, which was regressed against the bioactivities column vector (pMAC, the negative logarithm of the partial pressure capable of suppressing the movement in response to noxious stimuli in 50% of rats) using partial least squares (PLS) regression. The MIA-QSAR model was constructed over 19 training set compounds and externally validated using the remaining 6 compounds (**7**, **8**, **9**, **15**, **17** and **19**), which were chosen using Kennard-Stone sampling.

Figure 2 - Superimposed chemical structures used to generate the MIA descriptors with atoms colored according to the corresponding Pauling's electronegativity.

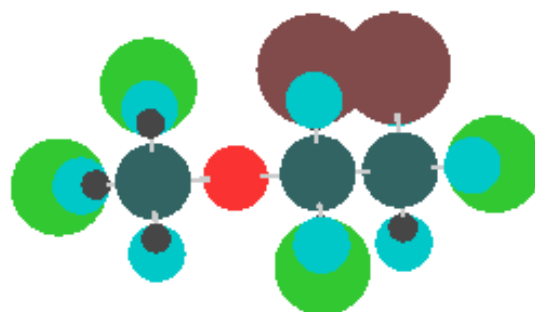
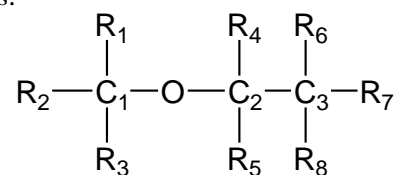


Table 4 - Data set of anesthetic haloethers and the corresponding experimental pMAC values.



Cpd	Chemical formula	Structural formula	pMAC
1	C ₃ Cl ₃ F ₅ O	CClF ₂ OCCL ₂ F	0.88
2	C ₃ Cl ₃ F ₅ O	CClF ₂ OCF ₂ CCl ₂ F	0.73
3	C ₃ Cl ₃ F ₅ O	CCl ₂ FOCF ₂ CClF ₂	0.74
4	C ₃ ClF ₆ HO	CClF ₂ OCFHCF ₃	0.54
5	C ₃ ClF ₆ HO	CF ₂ HOCClFCF ₃	0.49
6	C ₃ ClF ₆ HO	CF ₂ HO CF ₂ CClF ₂	0.22
7	C ₃ ClF ₅ H ₂ O	CClF ₂ OCH ₂ CF ₃	0.54
8	C ₃ ClF ₅ H ₂ O	CF ₂ HOCClHCF ₃	1.84
9	C ₃ ClF ₅ H ₂ O	CF ₂ HO CF ₂ CClFH	1.66
10	C ₃ Cl ₂ F ₅ HO	CClF ₂ OCClHCF ₃	1.31
11	C ₃ Cl ₂ F ₅ HO	CF ₂ HOCCl ₂ CF ₃	1.01
12	C ₃ Cl ₂ F ₅ HO	CClF ₂ OCF ₂ CClFH	1.52
13	C ₃ Cl ₂ F ₅ HO	CF ₂ HO CF ₂ CFCl ₂	1.04
14	C ₃ Cl ₂ F ₂ H ₄ O	CH ₃ OCF ₂ CHCl ₂	2.57
15	C ₃ F ₇ HO	CF ₂ HO CF ₂ CF ₃	-0.75
16	C ₃ F ₇ HO	CF ₃ OCFHCF ₃	-0.29
17	C ₃ F ₆ H ₂ O	CF ₂ HO CFHCF ₃	1.11
18	C ₃ F ₅ H ₃ O	CF ₂ HOCH ₂ CF ₃	0.96
19	C ₃ F ₅ H ₃ O	CFH ₂ OCFHCF ₃	1.33
20	C ₃ F ₅ H ₃ O	CFH ₂ OCF ₂ CF ₂ H	1.37
21	C ₃ ClF ₃ H ₄ O	CH ₃ OCF ₂ CClFH	1.80
22	C ₃ BrClF ₅ HO	CF ₂ HO CBrClCF ₃	1.82
23	C ₃ BrClF ₅ HO	CF ₂ HO CF ₂ CBrClF	1.82
24	C ₃ BrF ₅ H ₂ O	CF ₂ HO CBrHCF ₃	2.28
25	C ₃ BrF ₃ H ₄ O	CH ₃ OCF ₂ CBrFH	2.16

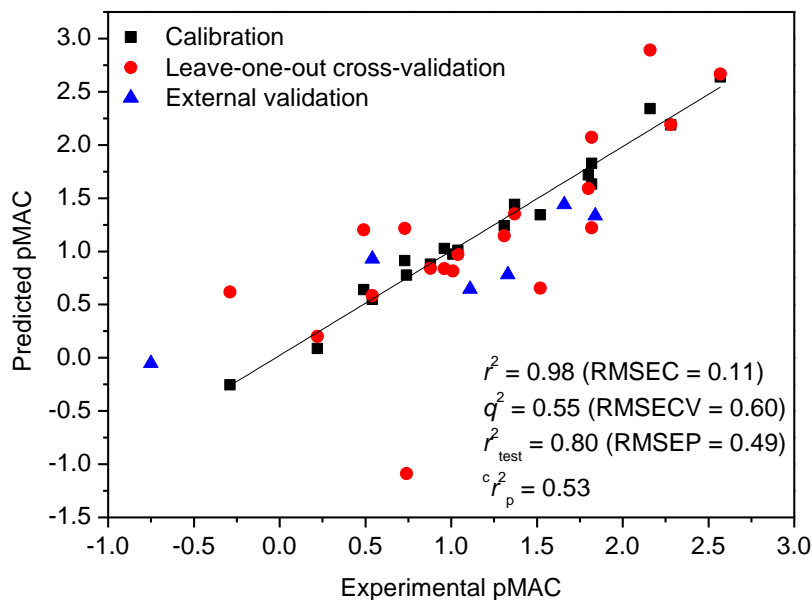
Table 5 - Correspondence of Pauling's electronegativity with atomic colors and respective pixel values (as a combination of RGB components) used in the MIA-QSAR model.

Atom	E	Color	Pixel value
H	2.1	Charcoal	210
C	2.5	Teal	250
O	3.5	Red	350
F	4.0	Turquoise	400
Cl	3.0	Green	300
Br	2.8	Maroon	280
Chemical bond	–	Grey	612
Blank space	–	White	765

A good correlation between MIA descriptors and the pMAC values for the series of anesthetic haloethers indicates that 2D chemical structures encode biological activities. A previous study using logP (the octanol/water partition coefficient) as descriptor showed a good correlation with pMAC [18], but the resulting model was not validated. Thus, the present study provided an internal (through leave-one-out cross-validation) and external validation to attest the reliability of the MIA-QSAR model. This was checked through the respective root mean square errors (RMSE) and the determination coefficients for the plot of actual vs. predicted pMAC (q^2 and $r^2_{\text{test}} > 0.5$ are considered acceptable). Also, since the MIA-QSAR model was obtained from PLS regression, a robustness test (y-randomization test) was performed to guarantee that calibration was not overfitted nor obtained by chance correlation; the y block was randomized and subsequently regressed against the intact matrix (ten times). Reliable models are achieved when $r^2_{\text{y-rand}} \ll r^2$, which is evaluated by ${}^c r^2_{\text{p}}$, defined as ${}^c r^2_{\text{p}} = r \times (r^2 - r^2_{\text{y-rand}})^{1/2}$ [19]. Values above 0.5 for ${}^c r^2_{\text{p}}$ are considered acceptable. The statistical results for 9 latent variables (PLS components) illustrated in Figure 3 attest the predicatibility and reliability of the MIA-QSAR, thus suggesting that 2D chemical structure indeed encodes biochemical

properties and that MIA descriptors can be useful to anticipating pMAC of prospective congeneric drug-like candidates.

Figure 3 - Plot of actual vs. predicted pMAC obtained from the MIA-QSAR model.



A major goal in QSAR is to determine and interpret the chemical motifs/properties responsible for the observed biological effects. Because of the numerous MIA descriptors generated to build the model, a straightforward analysis would not be an easy task using the raw data. Thus, pre-filtration procedures were performed in order to reduce the number of variables. In this sense, the first approach was a measure of Shannon's entropy (SE), corresponding to an unsupervised classification variables filtering [20, 21] applied to a 25 discrete intervals scheme. Variables with less than 10% of the maximum SE ($SE_{\text{MAX}} = \log 25$) were discarded. The variables were filtered again through the correlation coefficient (x/x), with variables removed for each set of variables with a $x/x = 0.98$ [22]. As the data set under study is highly

correlated, only two variables (**X1876** and **X5979**) were identified as containing further information regarding the anesthetic activity (pMAC). A careful analysis of the reduced matrix (Table 6) reveals the approximate positions of pixels **X1876** and **X5979** in the structural scaffold. Therefore, principal component analysis (PCA) was applied to obtain information on how these coordinates affect the bioactivity in terms of the two principal components PC1 (56.59%) and PC2 (43.41%) (Figure 4). Since PCA is a pattern recognition tool, the set of compounds were classified in three levels of anesthetic intensity: low (pMAC \leq 1.00), medium ($1.00 < \text{pMAC} < 2.00$), and high (pMAC ≥ 2.00) levels.

Figure 4 - PCA plots for the mean-centered data of the 25 anesthetic haloethers.

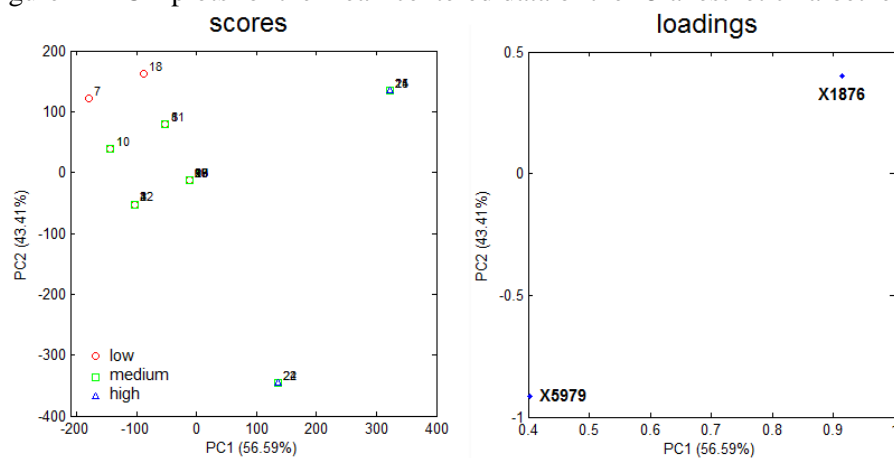


Table 6 - Reduced matrix with the selected descriptors and their approximate coordinates in the images.



Cpd	X1876	X5979
1	300	300
2	300	400
3	300	400
4	300	400
5	400	300
6	400	400
7	300	210
8	400	300
9	400	400
10	300	300
11	400	300
12	300	400
13	400	400
14	765	400
15	400	400
16	400	400
17	400	400
18	400	210
19	400	400
20	400	400
21	765	400
22	400	765
23	400	400
24	400	765
25	765	400

From the PCA scores plot of Figure 4, three clusters are observed: one at positive scores in PC1 (compounds with moderate-high activities), another with negative scores in PC1 and around null scores in PC2 (compounds with

moderate-low activities), and the third one at negative scores in PC1 and very positive scores in PC2 (compounds with low activities). Variables **X1876** and **X5979** in the loadings plot are responsible for clustering these compounds according to the scores plot. **X5979** clearly separates compounds **22** (moderate activity) and **24** (high activity) from the remaining ones in PC2, due to their high pixel value at this position (765, white - blank space in the image). Bulk and hydrophobic substituents (R) at position R4 (surrounding **X5979**, see substituent numbering in Table 4) have long C-R bonds and, therefore, the pixel variable is located where small substituents (such as hydrogen, fluorine and even chlorine) appears, while the bromine atom does not occupy this coordinate because of its long bond distance with carbon. From this, since compounds **22** and **24** contain a bromine substituent at position R4 and pertain to the moderate-high activity class, such a halogen tends to favor an improved anesthetic activity of congeneric haloethers in this position.

PC1 explains most of the data variance and, therefore, variable **X1876** should play a significant for the bioactivity pattern of the data set compounds, because of its high loading in PC1. In addition to **22** and **24**, compounds **14**, **21** and **25** (also pertaining to the moderate-high activity class) lie in the region with positive scores in PC1, while *e.g.* compounds **7** and **18** (low activity compounds) have very negative scores in PC1. In this case, the pixel value 765 at position **X1876** indicates absence of substituent (hydrogen at R1) and, consequently, atoms encoded by high pixel values (*e.g.* 400 for the small and electronegative fluorine atom, but mainly for the small hydrogen atom) at R1 tend to strongly favor the increase in pMAC of anesthetic derivatives.

Because the absence of intermolecular hydrogen bonds between enflurane (a similar anesthetic haloether) and the integrin LFA1 enzyme [13], the structural requirements mentioned above should be dictated by hydrophobic interactions of R4 with amino acid residues and by low steric effects

surrounding R1. Thus, 2D structural information provided by MIA descriptors (particularly related to the connectivity of different atoms in the present case) was capable of modeling and interpreting biochemical information of a series of anesthetic haloethers, without considering conformation.

Conclusion

It has been shown that a predictive and biochemically interpretable QSAR model can be obtained through bidimensional descriptors. Conformational insights could refine the analysis, but the risk of using wrong conformations could cause misinterpretation of the results. This eminent risk was assessed by investigating the conformational isomerism of isoflurane, whose bioconformation in at least one biological target is significantly different from the geometries optimized in a biological-free environment. A more secure option would be to obtain the ligand geometries inside an enzyme active site. Since conformational search inside a receptor normally gives the mode of interaction between substrate and enzyme, as well as the intermolecular interaction energy (related to the ligand-receptor affinity and, consequently, to the bioactivity), further QSAR analysis would be of limited utility.

Acknowledgements

Authors thank FAPEMIG (grant number APQ-00383-15), CAPES and CNPq for the financial support of this research, as well as by the scholarships (to M.C.G, M.H.D and J.M.S.) and fellowship (to M.P.F.).

References

- [1] Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* **1964**, *86*, 1616-1626.
- [2] Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.;

- Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. *J. Med. Chem.* **2014**, *57*, 4977-5010.
- [3] Cramer III, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959-5967.
- [4] Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130-4146.
- [5] Hopfinger, A.; Wang, S.; Tokarski, J. S.; Jin, B.; Albuquerque, M.; Madhav, P. J.; Duraiswami, C. *J. Am. Chem. Soc.* **1997**, *119*, 10509-10524.
- [6] Pan, D.; Tseng, Y.; Hopfinger, A. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1591-1607.
- [7] Freitas, M. P.; Brown, S. D.; Martins, J. A. *J. Mol. Struct.* **2005**, *738*, 149-154.
- [8] Barigye, S. J.; Freitas, M. P. *Int. J. Quant. Struct. Prop. Relat.* **2016**, *1*, 62-75.
- [9] Liu, R.; Loll, P.J.; Eckenhoff, R.G. *Faseb J.* **2005**, *19*, 567-576.
- [10] Zhang, H.; Astrof, N.S.; Liu, J.H.; Wang, J.H.; Shimaoka, M. *Faseb J.* **2009**, *23*, 2735-2740.
- [11] Hermann, A.; Mack, H. -G.; Oberhammer, H. *J. Fluor. Chem.* **2000**, *101*, 223-231.
- [12] Lesarri, A.; Vega-Toribio, A.; Suenram, R. D.; Brugh, D. J.; Nori-Shargh, D.; Boggs, J. E.; Grabow, J. -U. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6610-6618.
- [13] Andrade, L.; Silla, J. M.; Stephens, S. L.; Marat, K.; da Cunha, E. F. F.; Ramalho, T. C.; van Wijngaarden, J.; Freitas, M. P. *J. Phys. Chem. A* **2015**, *119*, 10735-10742.
- [14] Freitas, M. P.; Bühl, M.; O'Hagan, D.; Cormanich, R. A.; Tormena, C. F. *J. Phys. Chem. A* **2012**, *116*, 1677-1682.

- [15] Michalska, D.; Bienko, D. C.; Czarnik-Matusewicz, B.; Wierzejewska, M.; Sandorfy, C.; Zeegers-Huyskens, T. *J. Phys. Chem. B* **2007**, *111*, 12228–12238.
- [16] Koch, U.; Popelier, P. L. A. *J. Phys. Chem.* **1995**, *99*, 9747–9754.
- [17] Koblin, D. D.; Laster, M. J.; Ionescu, P.; Gong, D.; Eger II, E. I.; Halsey, M. J.; Hudlicky, T. *Anesth. Analg.* **1999**, *88*, 1161-1168.
- [18] Mehdipour, A. R.; Hemmateenejad, B.; Miri, R. *Chem. Biol. Drug Des.* **2007**, *69*, 362.
- [19] Mitra, I.; Saha, A.; Roy, K. *Mol. Simul.* **2010**, *36*, 1067-1079.
- [20] Godden, J. W.; Stahura, F. L.; Bajorath, J. *J. Chem. Inf. Comp. Sci.* **2000**, *40*, 796-800.
- [21] Barigye, S. J.; Marrero-Ponce, Y.; Martínez-López, Y.; Torrens, F.; Artiles-Martínez, L. M.; Pino-Urias, R. W.; Martínez-Santiago, O. *J. Comput. Chem.* **2013**, *34*, 259-274.
- [22] Duarte, M. H.; Barigye, S. J.; Freitas, M. P. *Comb. Chem. High Throughput Screen.* **2015**, *18*, 208-216.

ARTIGO 2**Is the bioconformation of 5-deoxy-5-fluoro-D-xylulose affected by intramolecular hydrogen bond?**

Maria C. Guimarães, Josué M. Silla, Elaine F. F. da Cunha, Teodorico C. Ramalho, Matheus P. Freitas*

Department of Chemistry, Federal University of Lavras, 37200-000, Lavras, MG, Brazil

* corresponding author: matheus@dqi.ufla.br

(artigo publicado no *RSC Advances*)

Abstract

5-Deoxy-5-fluoro-D-xylulose (DFX) binds to the xylulokinase enzyme and, as free ligand, has the preferential conformations governed by intramolecular interactions, such as hydrogen bond and hyperconjugative interactions. The role of intramolecular hydrogen bond on the bioconformation of DFX has not been studied yet, despite the relevance of this topic to explain the mode of interaction between ligand and enzyme and, therefore, the action mechanism of this molecule. DFX presents several conformations in the gas phase and implicit water, as determined by theoretical calculations, but the main optimized geometries do not match the bioactive conformation nor the most stable docked structure. This indicates that even expected strong interactions, such as hydrogen bond, are overcome by the enzyme induced fit of DFX. The natural consequence of this finding in three-dimensional quantitative structure-activity relationship (3D-QSAR) analysis is that the use of conformations obtained in a receptor-free environment can cause misinterpretation of the chemical and biological results.

Resumo

A 5-deóxi-5-flúor-D-xilulose (DFX) liga-se à enzima xiluloquinase e, como ligante livre, tem suas conformações preferenciais governadas por interações intramoleculares, tais como ligação de hidrogênio e interações hiperconjugativas. O papel da ligação de hidrogênio intramolecular na bioconformação da DFX ainda não foi estudado, apesar da relevância deste tópico para explicar o modo de interação entre o ligante e a enzima e, conseqüentemente, o mecanismo de ação dessa molécula. A DFX apresenta várias conformações em fase gasosa e em água implícita, conforme determinado por cálculos teóricos, mas as principais geometrias otimizadas não coincidem com a conformação bioativa, nem com a estrutura ancorada mais estável. Isso indica que mesmo as interações fortes, como a ligação de hidrogênio, são superadas pelo ajuste induzido na interação com a enzima. A consequência natural desses resultados sobre a análise quantitativa da relação estrutura-atividade 3D (QSAR-3D) é que o uso de conformações obtidas em um ambiente livre de receptor pode causar uma interpretação equivocada dos resultados químicos e biológicos.

1. Introduction

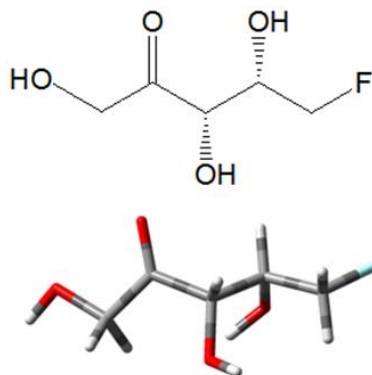
5-Deoxy-5-fluoro-D-xylulose (DFX) binds to the xylulokinase enzyme, which plays an important role in metabolic disease, since it catalyzes the ATP-dependent phosphorylation of D-xylulose to produce xylulose 5-phosphate, a key regulator of glucose metabolism and lipogenesis.¹ Molecular conformation is a property ruling the bioactivity of drug-like compounds, such as DFX, since the enzyme fitting depends on the three-dimensional arrangement of the small molecule. Nevertheless, the preferential conformations of an enzyme-free molecule are dictated by intramolecular interactions, such as steric and electrostatic effects, hyperconjugation (*e.g.* in the anomeric and gauche effects) and hydrogen bond.

It has been found that hydrophobic and hyperconjugative interactions are not strong enough to keep the bioconformation of organofluorine anesthetics as in the gas phase or solution.²⁻⁴ However, hydrogen bonds are known as being stronger forces than other noncovalent interactions,⁵ and usually ascribed as the responsible phenomenon for the architecture of supramolecular systems,^{6,7} ligand-enzyme interaction⁸ and conformation of small molecules.^{9,10} Despite the seeming lack of hydrogen bond involving fluorine as proton acceptor in solution,¹⁰⁻¹³ other nucleophilic sites in a molecule, such as oxygen and nitrogen atoms, can induce conformational changes due to intramolecular hydrogen bonds.¹⁴

This work reports the theoretical conformational analysis of DFX (Figure 1) in the gas phase, implicit water solution and docked inside the active site of xylulokinase (Protein Data Bank code: **4BC5**), in order to find out the role of intramolecular hydrogen bonds on the bioactive conformation of a small molecule containing fluorine and oxygen (as hydroxyl and carbonyl groups) as proton donors and acceptors. Ultimately, the outcomes of this study can aid the design of congeneric drugs. In addition, the suitability of three-dimensional

quantitative structure-activity relationship (3D-QSAR) techniques in drug design is evaluated, since such molecular modeling approaches require geometry optimization and 3D-alignment to obtain molecular descriptors for further correlation with bioactivity values; if the optimized enzyme-free conformation of DFX diverges from its bioconformation, such as in enflurane,³ isoflurane⁴ and 2,4-dichlorophenoxyacetic acid,¹⁵ the use of 3D molecular descriptors can be of limited utility.

Figure 1 - Chemical structure of 5-deoxy-5-fluoro-D-xylulose (DFX) and its corresponding bioactive conformation (Protein Data Bank code: **5FX**).



2. Theoretical Calculations

5-Deoxy-5-fluoro-D-xylulose (DFX) has 7 torsional angles and, therefore, the statistical Monte Carlo sampling was used to identify the possible energy minima for this compound. A total of 83 conformers were found and, after optimization at the ω B97X-D/6-311++g(d,p) level^{16,17} that includes empirical dispersion corrections, 65 different geometries converged to energy minima (Supplementary Material). From these, 15 conformers with non-zero population in the gas phase and 20 conformers in implicit water as solvent

(using the polarizable continuum model¹⁸) were obtained. Frequency calculations at this DFT level were employed to guarantee that the minima obtained were not saddle points and also to obtain the standard free energy values. Natural bond orbital (NBO)¹⁹ analysis was carried out to obtain the energies from hyperconjugative interactions, while QTAIM²⁰ analysis was performed to characterize hydrogen bonds. The docking calculations were performed in order to understand the physical–chemical impact of the interaction between the molecular target and the studied conformations. The crystal structure of DFX inside the xylulokinase active site was obtained from the Protein Data Bank (PDB codes: **5FX** for the ligand and **4BC5** for the enzyme) and used for the docking procedure and alignment of the optimized structures. The calculation of the docking energies of the rigid conformation inside the xylulokinase active site was performed using the software Molegro Virtual Docker (MVD).²¹

3. Results and discussion

Intramolecular hydrogen bond can drive the conformational isomerism of molecules and, therefore, it has been analyzed against conformational entropy and enzyme-ligand interaction (mostly due to intermolecular hydrogen bond) in order to find whether the enzyme-free conformation of DFX persists in a biological environment. First, the conformational isomerism of DFX was computationally studied in the gas phase (**g**) and implicit water solution (**w**) and, subsequently, the factors governing the conformer stabilization were evaluated using NBO and QTAIM analyses. From a diverse set of conformers, 15 geometries in the gas phase and 20 in implicit water solution have non-zero Gibbs population ($\geq 1\%$).

According to NBO analysis (Table 1 and Figure 2), the most populated conformer in the gas phase **1g** (23%) exhibits a network of hydrogen bonds involving hydroxyl groups and the carbonyl oxygen of DFX ($n_{\text{O}} \rightarrow \sigma^*_{\text{OH}}$). According to QTAIM analysis, a single bond path between C=O and O—H was observed in **1g** and characterized as a weak hydrogen bond, since both $\nabla^2\rho(r)$ and $H(r)$ are positive²² (Table 2). Hydrogen bonds also appear in the other conformers and can include the fluorine atom. However, other interactions lead to the greater stabilization of **1g** relative to the remaining conformers; the full electronic energy of a system can be decomposed into non-Lewis (electron delocalization) and Lewis-type (steric and electrostatic effects) interactions using NBO analysis and, according to Table 1, conformer **1g** combines high stabilization due to hyperconjugation and low repulsive effects. Table 3 shows that, in addition to hydrogen bond ($n_{\text{O}} \rightarrow \sigma^*_{\text{OH}}$), the *gauche* effect arisen from the *gauche* arrangement between electronegative substituents allows important antiperiplanar hyperconjugative interactions (especially $\sigma_{\text{CH}} \rightarrow \sigma^*_{\text{CO}}$).

Table 1 - Relative energies (kcal mol⁻¹), conformer Gibbs populations (%) and dihedral angles (degrees) for the conformations of DFX in the gas phase and implicit water, optimized at the ω B97X-D/6-311++g(d,p) level.
(Continue)

Gas	G ⁰ _{rel}	%	E _{rel}	E _L	E _{NL}	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
1g	0.0	23	0.0	9.6	9.6	-16.8	12.2	9.4	-13.4	57.9	76.2	173.4
2g	0.2	16	1.3	10.2	8.9	9.6	-4.8	-6.3	2.1	-166.4	-55.3	62.2
3g	0.4	11	1.0	5.9	4.9	4.0	-1.7	174.7	84.0	52.3	-51.0	56.8
4g	0.5	10	0.1	7.5	7.4	69.8	159.2	-11.1	6.2	-171.3	-53.5	60.1
5g	0.8	6	1.6	1.6	0.0	-1.2	0.6	176.7	164.2	-48.8	-56.1	57.7
6g	1.0	4	1.5	12.3	10.8	-16.6	10.8	101.6	5.1	74.9	-46.9	60.0
7g	1.1	4	1.5	8.7	7.2	4.1	-1.7	177.6	-172.9	48.6	81.7	-73.5
8g	1.3	3	1.6	6.2	4.6	-2.9	1.5	172.6	153.8	-42.2	40.7	-54.2
9g	1.3	3	1.6	15.9	14.3	-18.0	12.1	10.1	-13.5	66.7	64.8	65.2
10g	1.3	3	2.4	3.9	1.5	0.9	-3.0	-159.3	83.1	51.9	-145.1	178.8
11g	1.5	2	2.4	3.1	0.7	1.7	-3.6	-151.5	81.9	53.4	-148.7	179.0
12g	1.6	2	0.5	9.9	8.4	-55.5	156.9	-110.7	48.1	70.9	-51.5	57.8
13g	1.7	1	3.3	4.6	1.3	5.4	-1.8	177.7	-133.4	50.9	73.6	175.7
14g	1.8	1	2.1	16.1	14.0	2.7	-2.9	-171.9	83.1	-163.5	76.0	-176.5
15g	1.8	1	2.2	3.0	0.8	-9.0	6.2	-105.9	82.2	72.6	-166.3	176.9
H₂O												
1w	0.0	34	0.1	9.6	9.5	-18.9	10.7	7.2	-10.5	64.1	67.6	63.2
2w	0.7	10	0.1	7.3	7.2	-19.1	11.0	5.8	-8.2	59.7	72.9	174.9

Table 1 - Relative energies (kcal mol⁻¹), conformer Gibbs populations (%) and dihedral angles (degrees) for the conformations of DFX in the gas phase and implicit water, optimized at the ωB97X-D/6-311++g(d,p) level.
(Conclusion)

Gas	G ⁰ _{rel}	%	E _{rel}	E _L	E _{NL}	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
3w	0.7	10	0.7	5.0	4.3	2.0	-0.7	176.9	84.5	54.4	-53.2	59.8
4w	0.9	7	1.7	9.2	7.5	6.9	-3.2	-7.0	-165.7	176.9	-59.2	65.7
5w	0.9	7	0.6	7.4	6.8	-18.3	9.7	0.9	4.5	71.1	-59.2	62.3
6w	1.2	4	0.8	6.2	5.4	-0.8	0.6	-176.3	-170.5	50.3	78.6	-69.1
7w	1.3	4	1.3	1.7	0.4	2.1	-1.0	-176.6	-125.7	75.7	-48.9	60.6
8w	1.4	3	1.4	6.5	5.1	-16.8	9.4	-0.5	7.6	68.1	-74.0	177.4
9w	1.5	3	1.5	4.2	2.7	1.2	-0.3	179.0	83.5	53.5	-71.0	178.3
10w	1.6	2	0.7	13.0	12.3	6.4	-3.7	-2.0	2.6	-170.3	-161.0	-62.6
11w	1.7	2	0.5	10.1	9.6	68.2	159.7	-11.4	10.5	-168.6	-57.7	64.8
12w	1.7	2	1.6	3.3	1.7	-80.6	159.0	-7.0	13.9	70.1	-60.2	62.3
13w	1.7	2	1.3	14.2	12.9	0.5	1.0	-177.9	-102.8	-75.8	83.2	-66.1
14w	1.8	2	1.3	4.4	3.1	2.0	-0.7	176.8	84.5	54.4	-53.2	59.8
15w	1.9	1	1.5	5.5	4.0	-1.1	1.6	165.8	156.6	-45.2	47.1	-59.3
16w	2.0	1	0.0	14.1	14.1	68.8	153.4	-5.6	6.9	169.9	-66.6	-59.1
17w	2.0	1	1.9	1.9	0.0	-1.2	1.8	165.8	157.9	-42.3	-79.8	64.8
18w	2.1	1	2.0	3.1	1.1	2.3	-0.4	171.8	-90.5	66.2	-62.2	63.6
19w	2.4	1	0.4	21.4	21.0	-89.7	178.2	-6.1	5.0	-156.0	-43.0	-63.0
20w	2.5	1	2.0	3.0	1.1	19.35	-11.80	-35.74	39.39	-55.17	172.70	-71.94
5FX						175.0	4.4	167.1	-149.1	52.2	120.1	65.8

Figure 2 - Crystal structure and the optimized geometries of DFX in the gas phase and implicit water. The corresponding hydrogen bonds obtained by NBO analysis ($n_{O/F} \rightarrow \sigma^*_{OH}$, in kcal mol⁻¹) are shown.

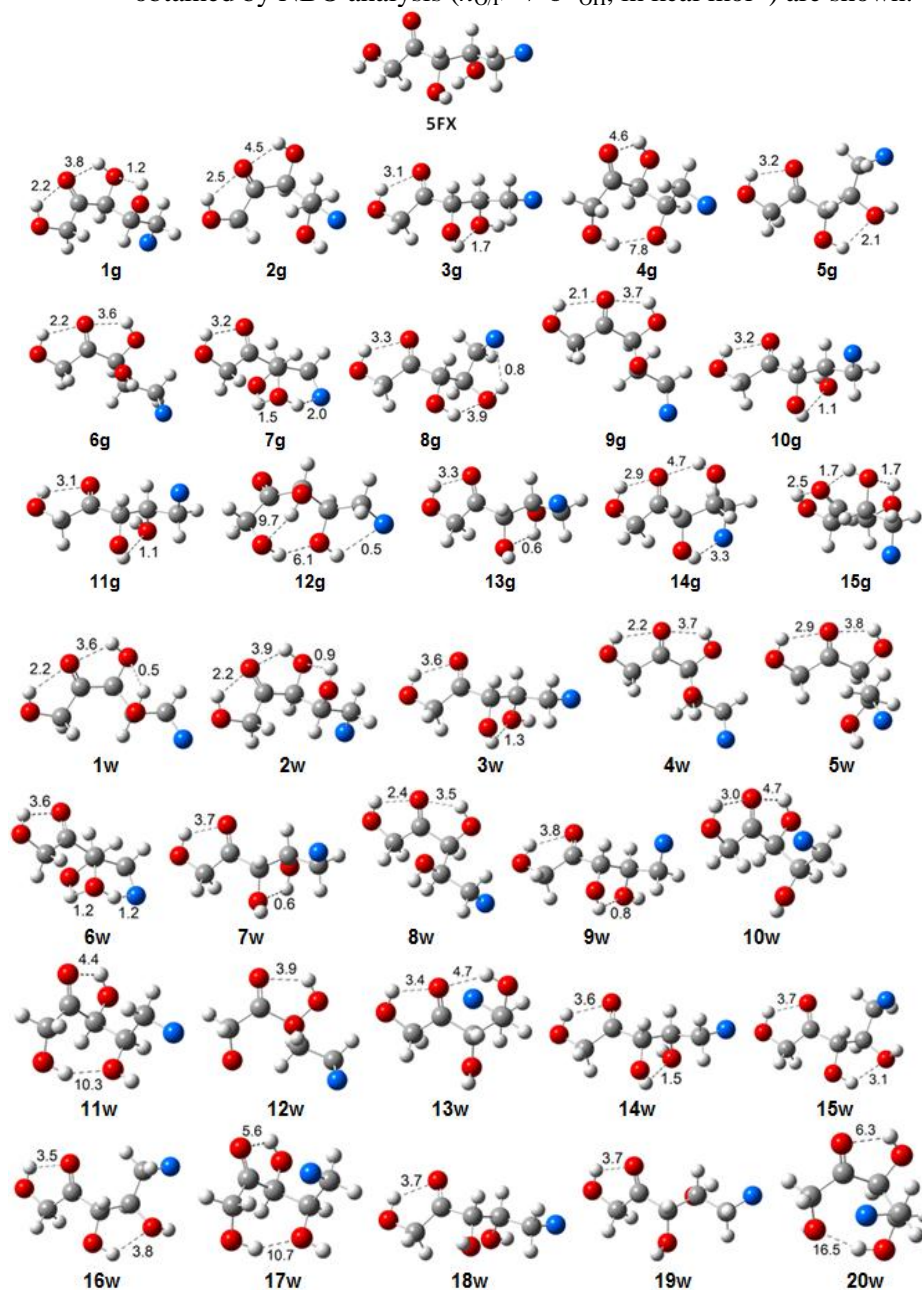
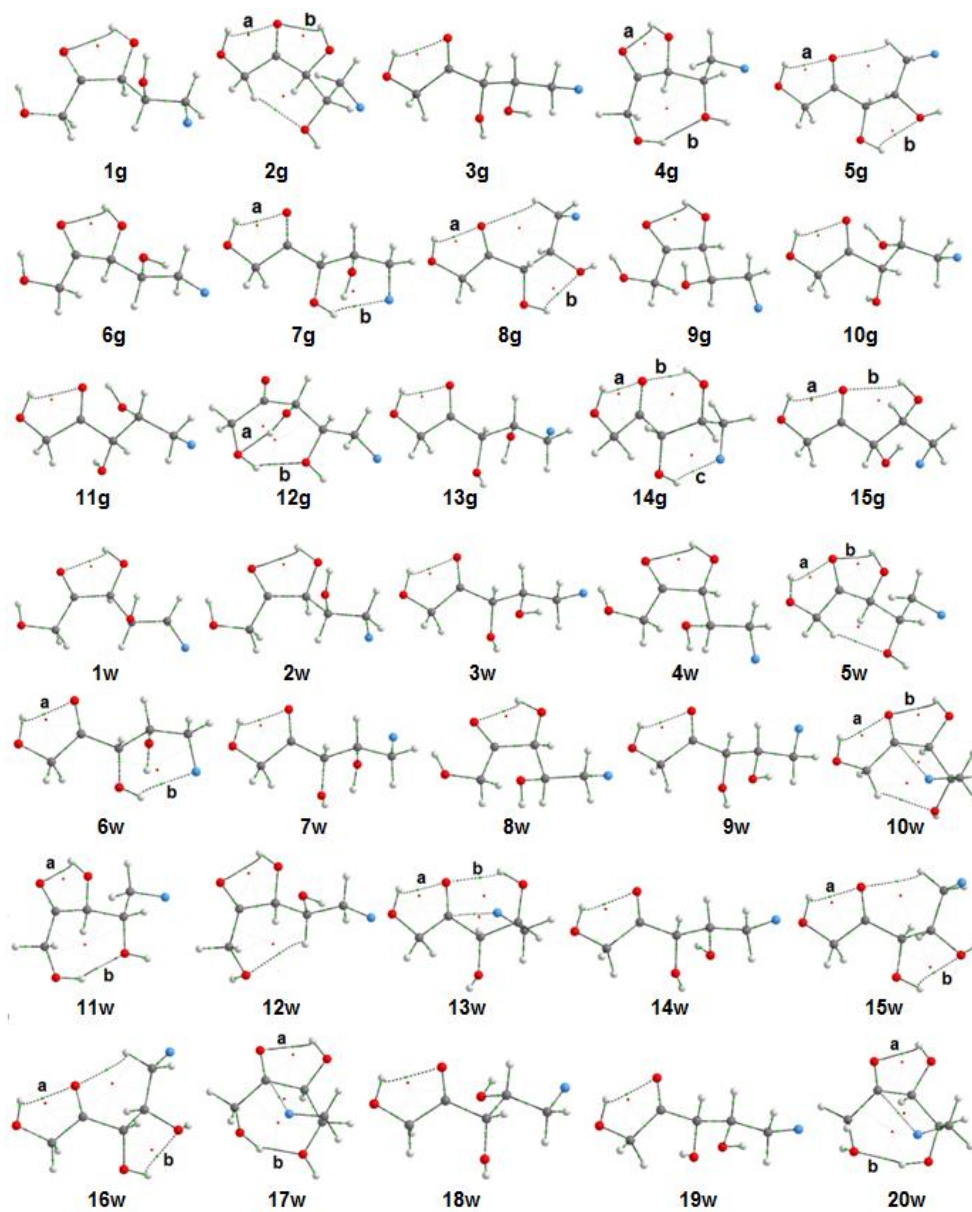


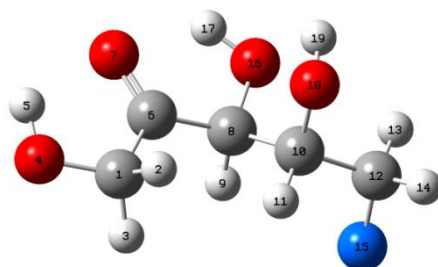
Table 2 - QTAIM parameters^a used to characterize hydrogen bonds in DFX.

Gas	$\rho(r)$	$\nabla^2\rho(r)$	$H(r)$
1g	+0.025374	+0.107403	+0.002732
2g a/b	+0.022415/+0.026278	+0.101106/+0.109487	+0.003063/+0.002569
3g	+0.023248	+0.101158	+0.002874
4g a/b	+0.026865/+0.027738	+0.110218/+0.107869	+0.002475/+0.001740
5g a/b	+0.023737/+0.020795	+0.101876/+0.095271	+0.002773/+0.003071
6g a/b	+0.025111	+0.106653	+0.002777
7g a/b	+0.023391/+0.017242	+0.101689/+0.066264	+0.002877/+0.000979
8g a/b	+0.023979/+0.023399	+0.102157/+0.093128	+0.002725/+0.002101
9g	+0.025217	+0.106525	+0.002722
10g	+0.023744	+0.101565	+0.002746
11g	+0.023546	+0.101048	+0.002759
12g a/b	+0.029000/+0.025460	+0.100256/+0.098074	+0.000656/+0.001851
13g	+0.023590	+0.102021	+0.002854
14g a/b/c	+0.023293/+0.021372/+0.019610	+0.100444/+0.074582/+0.075621	+0.002747/+0.001342/+0.001009
15g a/b	+0.021967/+0.016586	+0.098853/+0.058660	+0.002986/+0.001238
Water			
1w	+0.023990	+0.105896	+0.002748
2w	+0.025530	+0.107701	+0.002702
3w	+0.024078	+0.101946	+0.002712
4w	+0.025306	+0.106788	+0.002707
5w a/b	+0.023050/+0.025495	+0.100900/+0.107370	+0.002897/+0.002668
6w a/b	+0.024152/+0.014591	+0.102167/+0.056479	+0.002706/+0.000971
7w	+0.024449	+0.102938	+0.002677
8w	+0.024754	+0.105767	+0.002796
9w	+0.024544	+0.103050	+0.002656
10w a/b	+0.023425/+0.026389	+0.101164/+0.109546	+0.002807/+0.002544
11w a/b	+0.026365/+0.030194	+0.108854/+0.115734	+0.002536/+0.001356
12w	+0.025323	+0.106211	+0.002691
13w a/b	+0.024118/+0.022593	+0.101867/+0.081293	+0.002657/+0.001642

14w	+0.024192	+0.102243	+0.002696
15w a/b	+0.024560/+0.022415	+0.103014/+0.089973	+0.002641/+0.002163
16w a/b	+0.024189/+0.023866	+0.102220/+0.099989	+0.002695/+0.002483
17w a/b	+0.027762/+0.031069	+0.112484/+0.119232	+0.002350/+0.001245
18w	+0.024320	+0.102572	+0.002690
19w	+0.024322	+0.102578	+0.002690
20w	+0.029483/+0.035355	+0.119041/+0.124195	+0.002233/- 0.000144

^a Electron density - $\rho(r)$, Laplacian of the electron density - $\nabla^2\rho(r)$, and total energy density at the bond critical point - $H(r)$.

Table 3 - Hyperconjugative interactions between antiperiplanar donor and acceptor orbitals (kcal mol⁻¹) for the conformers of DFX.



Gas	1g	2g	3g	4g	5g	6g	7g	8g	9g	10g	11g	12g	13g	14g	15g	-	-	-	-	-
$\sigma_{C6C8} \rightarrow \sigma^*_{C10O18}$	-	-	-	-	2.2	0.7	-	2.5	-	-	-	0.6	0.6	-	-	-	-	-	-	-
$\sigma_{C8H9} \rightarrow \sigma^*_{C10O18}$	4.7	0.9	4.2	0.6	-	4.8	4.3	-	4.8	4.8	4.8	5.6	5.6	1.1	4.5	-	-	-	-	-
$\sigma_{C8O16} \rightarrow \sigma^*_{C10O18}$	-	1.7	-	2.0	-	-	-	-	-	-	-	-	-	1.4	-	-	-	-	-	-
$\sigma_{C10C12} \rightarrow \sigma^*_{C8O16}$	-	-	-	-	2.5	0.6	-	2.2	-	-	-	-	-	-	-	-	-	-	-	-
$\sigma_{C10H11} \rightarrow \sigma^*_{C8O16}$	4.5	1.1	4.6	0.9	-	5.0	4.8	-	4.8	4.8	4.9	5.4	5.4	1.0	4.9	-	-	-	-	-
$\sigma_{C10O18} \rightarrow \sigma^*_{C8O16}$	-	1.7	-	1.6	-	-	-	-	-	-	-	-	-	1.8	-	-	-	-	-	-
$\sigma_{C8C10} \rightarrow \sigma^*_{C12F}$	-	2.7	2.7	2.7	2.5	2.7	0.7	-	2.6	-	-	2.8	2.8	-	-	-	-	-	-	-
$\sigma_{C10H11} \rightarrow \sigma^*_{C12F}$	0.8	1.4	0.9	1.3	1.0	1.2	6.4	5.2	1.6	0.7	0.7	1.1	1.1	-	0.8	-	-	-	-	-
$\sigma_{C10O18} \rightarrow \sigma^*_{C12F}$	1.7	-	-	-	-	-	-	-	-	1.9	1.9	-	-	2.1	2.0	-	-	-	-	-
$\sigma_{C12H13} \rightarrow \sigma^*_{C10O18}$	-	5.2	5.3	5.3	5.0	5.3	1.2	0.6	5.6	-	-	5.5	5.5	0.7	-	-	-	-	-	-
$\sigma_{C12H14} \rightarrow \sigma^*_{C10O18}$	0.6	0.8	0.6	0.8	0.7	0.8	5.4	5.1	1.0	0.5	0.5	0.7	0.7	-	0.6	-	-	-	-	-
$\sigma_{C12F} \rightarrow \sigma^*_{C10O18}$	1.9	-	-	-	-	-	-	-	-	1.9	1.9	-	-	1.7	1.8	-	-	-	-	-
Water	1w	2w	3w	4w	5w	6w	7w	8w	9w	10w	11w	12w	13w	14w	15w	16w	17w	18w	19w	20w
$\sigma_{C6C8} \rightarrow \sigma^*_{C10O18}$	-	-	-	0.6	-	-	-	0.5	-	-	-	0.5	-	-	2.4	2.2	-	5.6	2.4	-
$\sigma_{C8H9} \rightarrow \sigma^*_{C10O18}$	4.9	4.8	4.2	4.8	0.9	4.5	4.4	4.8	4.3	4.8	0.8	4.8	-	4.4	-	-	0.7	4.9	-	0.9
$\sigma_{C8O16} \rightarrow \sigma^*_{C10O18}$	-	-	-	-	1.6	-	-	-	-	-	1.8	-	1.9	-	-	-	1.9	-	-	1.9
$\sigma_{C10C12} \rightarrow \sigma^*_{C8O16}$	-	-	-	0.5	-	-	-	-	-	-	-	-	-	-	2.2	2.4	-	-	2.2	-
$\sigma_{C10H11} \rightarrow \sigma^*_{C8O16}$	4.8	4.6	4.9	5.2	1.1	4.9	4.7	5.1	4.9	4.6	1	5.2	0.6	4.5	-	-	1.0	5.3	-	1.7

$\sigma_{C10O18} \rightarrow \sigma_{C8O16}^*$	-	-	-	-	1.7	-	-	-	-	-	1.7	-	1.6	-	-	-	1.6	-	-	1.1
$\sigma_{C8C10} \rightarrow \sigma_{C12F}^*$	2.7	-	2.8	2.8	2.9	0.5	-	-	-	-	2.8	2.9	-	2.9	-	2.7	-	2.7	-	-
$\sigma_{C10H11} \rightarrow \sigma_{C12F}^*$	1.5	0.7	1.1	1.3	1.6	6.5	0.7	0.7	0.7	0.7	1.6	1.3	6.5	1.4	5.8	1.4	6.3	1.4	5.8	6.6
$\sigma_{C10O18} \rightarrow \sigma_{C12F}^*$	-	1.8	-	-	-	-	1.9	1.8	1.8	1.8	-	-	-	-	-	-	-	-	-	-
$\sigma_{C12H13} \rightarrow \sigma_{C10O18}^*$	5.6	-	5.5	5.5	5.4	1.0	-	-	-	-	5.5	5.5	0.9	5.2	0.8	5.3	1.3	5.5	0.8	0.8
$\sigma_{C12H14} \rightarrow \sigma_{C10O18}^*$	0.9	0.6	0.7	0.8	0.9	5.8	0.6	0.6	0.6	0.6	1.0	0.8	5.6	1.1	5.4	1.0	1.4	0.9	5.4	5.4
$\sigma_{C12F} \rightarrow \sigma_{C10O18}^*$	-	1.9	-	-	-	-	1.8	1.9	1.9	1.9	-	-	-	-	-	-	-	-	-	-

A conformer with high energy in the gas phase became the most stable geometry in implicit water (**1w**) with 34% population, but it differs from **1g** only by the O—C—C—F dihedral angle. Indeed, the six most stable conformations in the gas phase appear within the top ten preferential conformations in solution, such as conformer **2w** (with a geometry similar to **1g**) as the second most stable (10%). Again, in addition to intramolecular hydrogen bond, a balance between steric effect and hyperconjugative interactions takes place and governs the conformational equilibrium of DFX in an enzyme-free environment (Tables 1-3 and Figure 2). Surprisingly, **1g** and **1w** do not match the crystal structure of DFX, which is the ligand bioconformation attached to the active site of xylulokinase.¹

In order to compare the optimized geometries with the experimental crystal structure, only the O—C—C=O, O=C—C—O(H), (H)O—C—C—O(H) and O—C—C—F dihedral angles were taken into account, since the coordinates of the hydroxyl hydrogens are not experimentally accurate. The bioconformation is reasonably consistent only with **3g** (11% in the gas) and **3w**, **7w** and **18w** (summing up 15% in implicit water). Consequently, the intramolecular forces ruling the optimized geometries are certainly not the same as those governing the bioactive conformation of DFX. In order to check the possible intermolecular interactions that overcome the intramolecular ones as driving forces of the bioconformation of DFX, docking studies were performed including all optimized geometries inside the bioactive site of xylulokinase.

The docking results were validated after finding the highest intermolecular (ligand-protein) interaction for the bioconformation of DFX (PDB code: **5FX**) in the active site of xylulokinase (PDB code: **4BC5**), which is $-78.8 \text{ kcal mol}^{-1}$. Conformers **3g** and **3w** were the most stable optimized structures in the biological target (Table 4), with intermolecular interaction energies of -76.6 and $-76.0 \text{ kcal mol}^{-1}$, respectively, more than 2 kcal mol^{-1} less

stable than **5FX**. Nevertheless, **3g** and **3w**, together with **7w** and **18w**, exhibited high stabilization in the active site of xylulokinase in comparison to most of the optimized geometries, since they have more structural similarity to **5FX**. The preferential enzyme-free conformations **1g** and **1w** interact attractively with the active site of xylulokinase (-31.1 and -39.8 kcal mol⁻¹, respectively), but they are far less stable than the bioconformation in the protein, revealing the important role of enzyme fit over intramolecular interactions for the conformer stabilization in a biological medium. Actually, many other enzyme-free conformations exhibit more favoring ligand-enzyme interaction than **1g** and **1w**, as a special consequence of more effective hydrogen bonds between ligand and amino acid residues in the enzyme active site. For instance, Figure 3 compares the main hydrogen bond interactions experienced by the bioconformation **5FX** and the optimized conformers **1g** and **1w**, in order to show the probable cause of conformer stabilization in the biological environment.

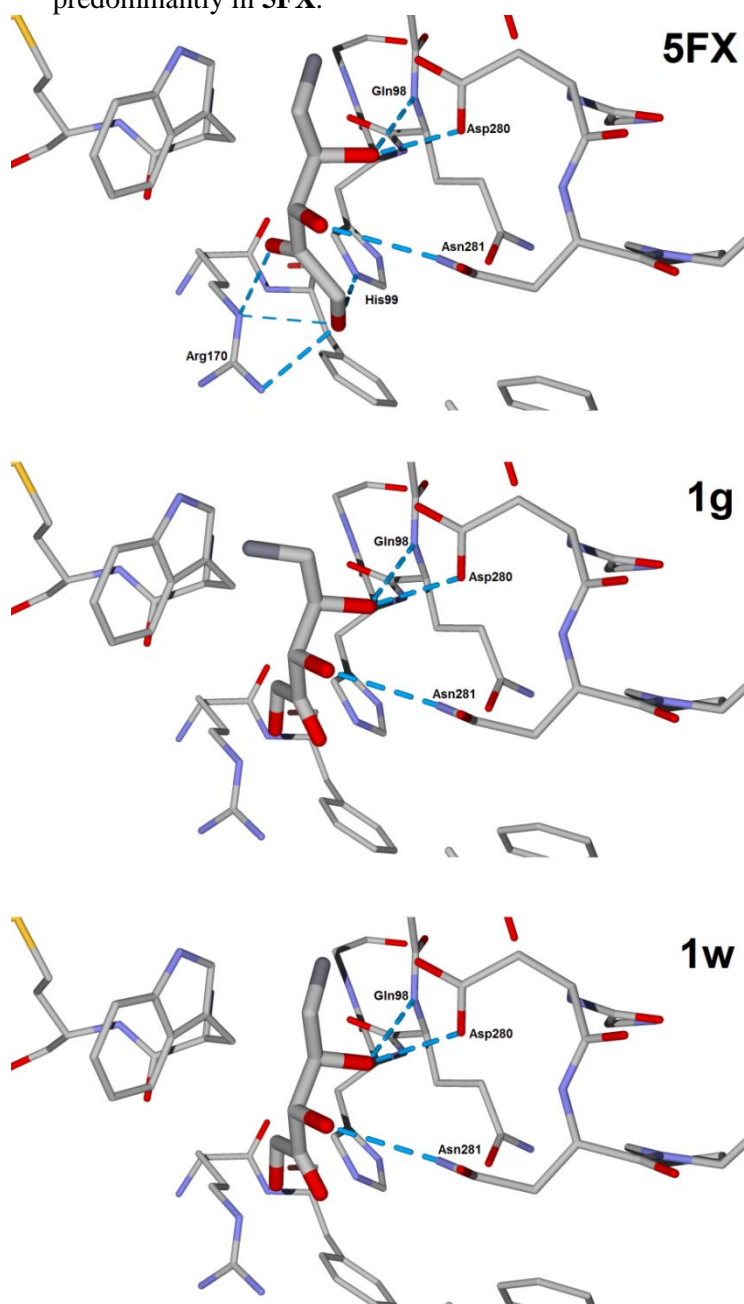
The stronger effect of intermolecular ligand-protein interactions over intramolecular hydrogen bonds and hyperconjugation as driving force of the conformer stabilization in a biological medium suggests that conformations obtained in an enzyme-free environment should be used with caution in conformation-dependent QSAR methods. Most 3D-QSAR methods use conformational screening and structure alignment rules that do not consider the biological target. Consequently, the chemical-biological interpretation obtained from molecular descriptors derived from optimized conformations (those not consistent with the bioconformation) can be inappropriate and inaccurate.

Table 4 - Intermolecular (ligand-protein) interactions (kcal mol^{-1}) for the main optimized conformers of DFX, obtained by docking studies.^a

Conformer	$E_{\text{ligand-protein}}$	Conformer	$E_{\text{ligand-protein}}$
1g	-31.1	1w	-39.8
2g	9.1	2w	-33.7
3g	-76.3	3w	-76.0
4g	8.2	4w	-35.6
5g	-48.0	5w	9.4
6g	-33.2	6w	-72.7
7g	-74.4	7w	-60.7
8g	-10.8	8w	-39.2
9g	-39.2	9w	-60.7
10g	-55.0	10w	-4.7
11g	-53.9	11w	9.3
12g	-64.0	12w	-33.6
13g	-67.8	13w	-42.0
14g	-31.1	14w	-75.4
15g	-45.9	15w	-10.0
		16w	-47.7
		17w	-2.8
		18w	-73.2
		19w	-73.2
		20w	0.6

^a The interaction energy for the remaining optimized conformers were all less stabilizing than $-78.8 \text{ kcal mol}^{-1}$ (the $E_{\text{ligand-protein}}$ for **5FX**).

Figure 3 - Conformers **5FX**, **1g** and **1w** docked inside the active site of xylulokinase. Hydrogen bonds with amino acid residues are shown, predominantly in **5FX**.



4. Conclusions

The main conformation of DFX in the gas phase and implicit water solution exhibits a network of cooperative intramolecular hydrogen bonds, in addition to hyperconjugative interactions and low steric/electrostatic effects that contribute to its stabilization. However, this conformation and many others optimized in an enzyme-free environment are not consistent with the bioconformation of DFX bound to the xylulokinase enzyme. Indeed, the enzyme-fit due to intermolecular hydrogen bonds with amino acid residues overcome intramolecular interactions in DFX, then strongly favoring the conformer stabilization in the biological medium.

Acknowledgements

Authors thank FAPEMIG (grant number APQ-00383-15), CAPES and CNPq for the financial support of this research, as well as by the scholarships (to M.C.G and J.M.S.) and fellowships (to E.F.F.C., T.C.R. and M.P.F.).

References

1. R.D. Bunker, E.M.M. Bulloch, J.M.J. Dickson, K.M. Loomes, E.N. Baker, Structure and Function of Human Xylulokinase, an Enzyme with Important Roles in Carbohydrate Metabolism, *J. Biol. Chem.* 288 (2013) 1643-1652.
2. M.P. Freitas, M. Bühl, D. O'Hagan, R.A. Cormanich, C.F. Tormena, Stereoelectronic Interactions and the One-bond C-F Coupling Constant in Sevoflurane, *J. Phys. Chem. A* 116 (2012) 1677-1682.
3. L.A.F. Andrade, J.M. Silla, S.L. Stephens, K. Marat, E.F.F. da Cunha, T.C. Ramalho, J. van Wijngaarden, M.P. Freitas, Conformational Exploration of

- Enflurane in Solution and in a Biological Environment, *J. Phys. Chem. A* 119 (2015) 10735-10742.
4. M.C. Guimarães, M.H. Duarte, J.M. Silla, M.P. Freitas, Is Conformation a Fundamental Descriptor in QSAR? A Case for Halogenated Anesthetics, *Beilstein J. Org. Chem.* 12 (2016) 760-768.
 5. G. Gilli, P. Gilli, *The Nature of the Hydrogen Bond*, Oxford University Press, Oxford, 2009.
 6. N.H. Evans, P.D. Beer, Advances in Anion Supramolecular Chemistry: From Recognition to Chemical Applications, *Angew. Chem. Int. Ed. Engl.* 53 (2014) 11716-54.
 7. N. Busschaert, C. Caltagirone, W. von Rossom, P.A. Gale, Applications of Supramolecular Anion Recognition, *Chem. Rev.* 115 (2015) 8038-8155.
 8. H. Kubinyi, Hydrogen Bonding: The Last Mystery in Drug Design? *In Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical, and Computational Strategies* (B. Testa, H. van de Waterbeemd, G. Folkers, R. Guy, Eds.), Verlag Helvetica Chimica Acta, Zürich, 2001.
 9. N.S. Nudelman, C.E.S. Alvaro, Inter- and Intramolecular Hydrogen Bonds in Polyamines: Variable-concentration $^1\text{H-NMR}$ Studies, *J. Phys. Org. Chem.* 24 (2011) 1067-1071.
 10. J.M. Silla, R.A. Cormanich, R. Rittner, M.P. Freitas, Conformational Analysis and Intramolecular Interactions in Aminofluorobenzoic Acids, *J. Phys. Chem. A* 117 (2013) 1659-1664.
 11. L.A.F. Andrade, J.M. Silla, C.J. Duarte, R. Rittner, M.P. Freitas, The Preferred All-gauche Conformations in 3-Fluoro-1,2-propanediol, *Org. Biomol. Chem.* 11 (2013) 6766-6771.
 12. R.A. Cormanich, M.P. Freitas, C.F. Tormena, R. Rittner, The F...HO Intramolecular Hydrogen Bond Forming Five-membered Rings Hardly

- Appear in Monocyclic Organofluorine Compounds, *RSC Adv.* 2 (2012) 4169-4174.
13. R.A. Cormanich, R. Rittner, M.P. Freitas, M. Bühl, The Seeming Lack of CF...HO Intramolecular Hydrogen Bonds in Linear Aliphatic Fluoroalcohols in Solution, *Phys. Chem. Chem. Phys.* 16 (2014) 19212-19217.
 14. A. Lämmermann, I. Szatmári, F. Fülöp, E. Kleinpeter, Inter- or Intramolecular N...H-O or N-H...O Hydrogen Bonding in 1,3-Amino- α/β -naphthols: An Experimental NMR and Computational Study, *J. Phys. Chem. A* 113 (2009) 6197-6205.
 15. M.P. Freitas, T.C. Ramalho, Employing Conformational Analysis in the Molecular Modeling of Agrochemicals: Insights on QSAR Parameters of 2,4-D, *Ciênc. Agrotec.* 37 (2013) 485-494.
 16. J.-D. Chai, M. Head-Gordon, Long-range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* 10 (2008) 6615-6620.
 17. K. Raghavachari, J.S. Binkley, R. Seeger, J.A. Pople, Selfconsistent Molecular Orbital Methods. XX. Basis Set for Correlated Wave-functions, *J. Chem. Phys.* 72 (1980) 650-654.
 18. C. Amovilli, V. Barone, R. Cammi, E. Cancès, M. Cossi, B. Menucci, C.S. Pomelli, J. Tomasi, Recent Advances in the Description of Solvent Effects with the Polarizable Continuum Model. *Adv. Quantum Chem.* 32 (1998) 227-261.
 19. F. Weinhold, C.R. Landis, *Discovering Chemistry with Natural Bond Orbitals*, Wiley, Hoboken, 2012.
 20. R.F.W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, Oxford, 1994.
 21. R. Thomsen, M.H. Christensen, MolDock: A New Technique for High-accuracy Molecular Docking, *J. Med. Chem.* 49 (2006) 3315-3321.

22. S.J. Grabowski, What is the Covalency of Hydrogen Bonding? *Chem. Rev.* 111 (2011) 2597-2625.

SUPPLEMENTARY MATERIAL**ARTIGO 2****Is the bioconformation of 5-deoxy-5-fluoro-D-xylulose affected by intramolecular hydrogen bond?**

Maria C. Guimarães, Josué M. Silla, Elaine F. F. da Cunha, Teodorico C. Ramalho, Matheus P. Freitas*

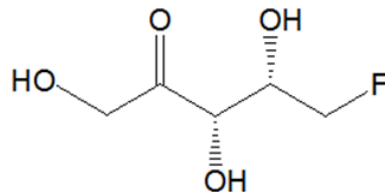
Department of Chemistry, Federal University of Lavras, 37200-000, Lavras, MG, Brazil

* corresponding author: matheus@dqi.ufla.br

Table S1. Optimized conformations for DFX in the gas phase at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol^{-1}) and key dihedral angles.

Table S2. Optimized conformations for DFX in implicit water at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol^{-1}) and key dihedral angles.

Table S1 - Optimized conformations for DFX in the gas phase at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol⁻¹) and key dihedral angles. (Continue)



Conf	ΔG_{rel}^0	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
1g	0.00	-16.8	12.2	9.4	-13.4	57.9	76.2	173.4
2g	0.22	9.6	-4.8	-6.3	2.1	-166.4	-55.3	62.2
3g	0.38	4.0	-1.7	174.7	84.0	52.3	-51.0	56.8
4g	0.47	69.8	159.2	-11.1	6.2	-171.3	-53.5	60.1
5g	0.78	-1.2	0.6	176.7	164.2	-48.8	-56.1	57.7
6g	1.00	-16.6	10.8	101.6	5.1	74.9	-46.9	60.0
7g	1.06	4.1	-1.7	177.6	-172.9	48.6	81.7	-73.5
8g	1.31	-2.9	1.5	172.6	153.8	-42.2	40.7	-54.2
9g	1.32	-18.0	12.1	10.1	-13.5	66.7	64.8	65.2
10g	1.34	0.9	-3.0	-159.3	83.1	51.9	-145.1	178.8
11g	1.51	1.7	-3.6	-151.5	81.9	53.4	-148.7	179.0
12g	1.58	-55.5	156.9	-110.7	48.1	70.9	-51.5	57.8
13g	1.71	5.4	-1.8	177.7	-133.4	50.9	73.6	175.7
14g	1.76	2.7	-2.9	-171.9	83.1	-163.5	76.0	-176.5
15g	1.80	-9.0	6.2	-105.9	82.2	72.6	-166.3	176.9
16g	2.07	-10.2	4.1	2.0	-3.0	-170.9	-159.7	-64.8
17g	2.26	19.1	-13.2	-32.6	31.0	-54.3	169.2	-75.9
18g	2.39	66.9	-142.9	1.9	-4.7	-156.4	76.4	-67.3
19g	2.57	-11.9	8.5	0.8	8.4	70.3	-60.9	172.7

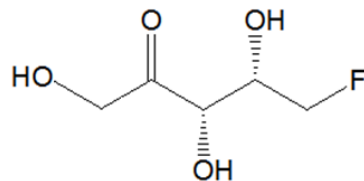
Table S1 - Optimized conformations for DFX in the gas phase at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol⁻¹) and key dihedral angles. (Continue)

Conf	ΔG_{rel}^0	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
20g	2.66	5.8	-7.5	-34.2	31.0	-53.9	168.4	165.9
21g	2.68	7.5	-1.5	-179.2	-154.8	-175.0	81.5	-68.5
22g	2.70	-56.1	-179.6	-115.9	81.4	50.4	-163.5	178.0
23g	2.83	86.8	-179.6	-26.7	24.4	-47.8	164.1	169.4
24g	2.83	65.0	168.5	-5.1	1.9	-161.6	-159.4	-60.9
25g	2.85	-71.9	-113.2	-0.1	-2.4	-164.9	-61.6	62.4
26g	2.88	-61.7	-157.1	-136.6	158.0	-43.8	42.1	-55.8
27g	2.97	65.8	139.3	157.8	166.3	-50.7	-59.6	59.7
28g	2.97	3.4	-3.9	-152.5	164.9	55.2	-165.4	-72.8
29g	3.13	-82.0	174.7	-3.3	1.0	-154.7	77.9	-63.7
30g	3.15	-64.3	-150.8	-142.3	164.0	-47.7	-56.2	56.3
31g	3.21	-82.7	152.9	-3.9	12.7	75.7	-48.9	60.6
32g	3.63	5.1	-5.0	-156.0	-157.6	74.8	-164.8	-169.3
33g	3.66	67.2	138.1	158.0	160.4	-45.5	40.0	-52.7
34g	3.71	7.2	-3.1	-174.9	63.5	-174.8	84.9	-83.0
35g	3.83	-65.2	139.3	168.9	38.8	-169.7	90.0	-69.5
36g	3.91	-53.7	153.5	152.6	47.0	68.8	-70.2	171.8
37g	4.11	7.4	-2.7	164.5	-77.1	70.0	-54.8	62.1
38g	4.24	0.3	-1.4	-163.4	-163.9	-167.3	68.4	70.9
39g	4.35	-66.3	-131.5	-166.7	-174.1	49.0	82.0	-73.9
40g	4.40	-66.4	-131.4	-166.7	-74.2	49.0	81.8	-73.9
41g	4.45	13.0	-7.9	-17.8	31.2	-161.2	-80.9	179.2
42g	4.48	75.3	157.3	-27.6	25.8	-55.7	171.2	-76.4
43g	4.49	13.3	-8.2	-17.2	29.4	-160.8	-79.5	178.9
44g	4.54	11.3	-7.3	-16.3	30.7	-160.0	-153.2	-178.7
45g	4.70	-1.1	2.5	161.3	-95.8	-52.4	179.6	72.0
46g	4.72	-1.3	2.4	161.4	-94.8	-52.0	177.8	71.8

Table S1 - Optimized conformations for DFX in the gas phase at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol⁻¹) and key dihedral angles. (Conclusion)

Conf	ΔG_{rel}^0	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
47g	4.85	17.5	-13.9	-38.7	35.8	-56.5	174.1	64.7
48g	4.85	79.2	-125.0	-177.5	-48.4	-173.9	82.9	-72.1
49g	5.18	67.8	163.0	-17.9	26.3	-166.7	-81.2	174.6
50g	5.37	-52.6	150.1	151.7	50.9	63.3	34.0	-45.0
51g	5.50	84.9	-163.5	-129.4	-72.3	-50.4	166.8	67.2
52g	5.64	65.5	166.3	-14.8	24.6	-159.6	-140.0	179.8
53g	6.21	-74.7	95.2	146.6	98.3	41.3	-167.3	67.1
54g	6.44	-79.3	102.9	-156.8	165.1	42.3	162.2	-72.7
55g	6.64	68.5	96.2	137.9	-161.3	41.4	80.3	-72.4
56g	6.65	-63.8	-141.6	-152.3	65.0	179.2	78.8	-94.1
57g	6.71	84.4	-153.4	-141.8	-65.7	-171.4	60.7	68.7
58g	6.86	-178.2	157.8	144.5	54.8	-50.4	163.6	71.8
59g	6.96	-1.2	1.9	164.5	65.3	-51.8	164.6	69.0
60g	7.26	177.2	-166.0	-12.1	-71.6	-48.8	169.5	70.4
61g	7.47	86.3	-154.9	139.0	-69.6	-47.8	169.7	69.6
62g	7.51	-69.9	-111.8	-10.2	18.5	-158.8	-78.9	174.7
63g	7.71	69.4	-149.1	-52.8	170.1	-52.0	-100.6	179.2
64g	8.45	-62.0	-163.8	-129.8	78.4	-165.2	-157.7	-172.0
65g	11.05	57.2	121.5	-139.3	79.5	-160.5	-84.3	-175.4

Table S2 - Optimized conformations for DFX in implicit water at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol⁻¹) and key dihedral angles. (Continue)



Conf.	ΔG_{rel}^0	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
1w	0.00	-18.9	10.7	7.2	-10.5	64.1	67.6	63.2
2w	0.66	-19.1	11.0	5.8	-8.2	59.7	72.9	174.9
3w	0.66	2.0	-0.7	176.9	84.5	54.4	-53.2	59.8
4w	0.86	6.9	-3.2	-7.0	-165.7	176.9	-59.2	65.7
5w	0.87	-18.3	9.7	0.9	4.5	71.1	-59.2	62.3
6w	1.22	-0.8	0.6	-176.3	-170.5	50.3	78.6	-69.1
7w	1.32	2.1	-1.0	-176.6	-125.7	75.7	-48.9	60.6
8w	1.40	-16.8	9.4	-0.5	7.6	68.1	-74.0	177.4
9w	1.47	1.2	-0.3	179.0	83.5	53.5	-71.0	178.3
10w	1.59	6.4	-3.7	-2.0	2.6	-170.3	-161.0	-62.6
11w	1.71	68.2	159.7	-11.4	10.5	-168.6	-57.7	64.8
12w	1.74	-80.6	159.0	-7.0	13.9	70.1	-60.2	62.3
13w	1.74	0.5	1.0	-177.9	-102.8	-75.8	83.2	-66.1
14w	1.82	2.0	-0.7	176.8	84.5	54.4	-53.2	59.8
15w	1.92	-1.1	1.6	165.8	156.6	-45.2	47.1	-59.3
16w	1.96	68.8	153.4	-5.6	6.9	169.9	-66.6	-59.1
17w	1.98	-1.2	1.8	165.8	157.9	-42.3	-79.8	64.8
18w	2.08	2.3	-0.4	171.8	-90.5	66.2	-62.2	63.6
19w	2.39	-89.7	178.2	-6.1	5.0	-156.0	-43.0	-63.0

Table S2 - Optimized conformations for DFX in implicit water at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol⁻¹) and key dihedral angles. (Continue)

Conf.	ΔG_{rel}^0	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
20w	2.46	19.4	-11.8	-35.7	39.4	-55.2	172.7	-71.9
21w	2.57	3.4	-2.6	-16.7	-12.5	-170.4	70.9	67.1
22w	2.61	0.6	0.8	173.7	-95.7	62.8	-79.0	178.3
23w	2.61	0.6	0.8	173.7	-95.7	84.1	-79.0	178.3
24w	2.63	-56.4	158.4	148.3	49.8	70.5	-60.6	61.3
25w	2.65	1.2	-0.9	-178.1	96.1	40.7	-147.3	-63.0
26w	2.85	-12.9	7.3	-109.0	79.4	72.4	-166.4	178.2
27w	2.93	-2.2	2.5	166.8	-94.3	-50.6	171.5	67.3
28w	3.26	11.0	-8.8	-34.2	34.0	-48.5	168.4	65.2
29w	3.30	17.9	-11.9	-37.5	39.2	-55.8	171.0	166.3
30w	3.32	-56.4	159.3	147.8	49.2	68.6	-75.5	176.5
31w	3.33	4.1	-2.7	-169.6	82.3	-164.7	74.7	-179.3
32w	3.48	75.7	-143.6	-1.0	-0.5	-156.4	76.7	-67.2
33w	3.66	62.5	146.1	151.8	161.6	-47.7	47.9	-59.1
34w	3.67	-55.2	-174.9	-123.3	77.9	52.0	-161.3	179.5
35w	3.68	12.1	-6.6	-16.6	32.6	-161.9	-157.6	-177.2
36w	3.71	13.5	-7.1	-17.7	32.9	-162.0	-77.9	-178.8
37w	3.72	13.9	-7.2	-17.7	32.7	-162.1	-78.1	-178.8
38w	3.74	-64.0	-136.9	159.5	-170.1	49.2	80.4	-70.7
39w	4.11	-79.6	108.5	141.2	98.1	39.9	-172.0	63.7
40w	4.19	-61.1	-156.3	-137.5	164.3	-46.1	-79.1	64.3
41w	4.31	67.6	161.6	-16.5	26.1	-166.2	-81.3	178.1
42w	4.33	62.5	143.7	154.2	163.5	-50.7	-79.6	64.1
43w	4.43	65.7	163.9	-16.1	28.2	-164.2	-124.6	-178.4
44w	4.52	87.8	-178.6	-29.8	30.7	-49.2	167.0	171.5
45w	4.62	87.6	-16.7	-128.9	-69.4	-49.0	168.9	-75.4
46w	4.70	172.7	161.3	142.4	52.9	-50.7	160.8	66.5

Table S2 - Optimized conformations for DFX in implicit water at the ω B97X-D/6-311++g(d,p) level and respective free energies (kcal mol⁻¹) and key dihedral angles. (Conclusion)

Conf.	ΔG_{rel}^0	HOC(H ₂)C(=O)	OCC=O	O=CCO(H)	C(=O)COH	(H)OCCO(H)	HOCC(F)	OCCF
47w	4.80	-72.8	-104.9	-3.7	2.6	-167.8	-69.5	65.7
48w	4.80	-1.9	2.3	167.0	-93.6	-50.0	169.8	67.2
49w	4.83	-174.4	-171.4	-124.3	-69.4	165.7	165.7	68.2
50w	4.85	76.7	159.2	-28.2	32.4	-55.5	174.3	-73.0
51w	4.86	-59.6	-157.4	-136.8	163.7	-43.9	45.1	-58.2
52w	4.91	-56.1	155.6	148.8	50.9	60.6	37.4	-50.7
53w	4.92	79.2	-16.3	-171.5	-47.2	174.6	83.9	-69.7
54w	5.02	-72.0	142.2	168.5	39.5	-168.7	87.3	-67.8
55w	5.11	87.9	-170.5	-27.6	-69.7	-47.5	165.6	67.8
56w	5.15	82.9	-156.7	-144.8	-59.3	-71.1	66.7	66.9
57w	5.15	-0.2	0.5	178.0	67.6	-51.2	168.2	63.3
58w	5.17	-173.4	-156.0	-144.5	-55.9	45.4	82.6	-63.8
59w	5.26	64.4	126.2	-178.5	-16.9	51.8	75.3	-69.1
60w	6.06	-83.7	114.8	143.0	108.0	35.3	-172.6	-66.0
61w	6.48	-71.6	-101.9	-12.1	20.4	-164.3	-78.8	176.2
62w	6.76	76.9	-160.6	-43.9	166.2	-53.8	-91.6	179.5
63w	7.09	-63.7	-133.2	-165.5	72.0	-173.9	80.9	-74.8
64w	8.18	-60.7	-154.5	-142.4	75.8	-162.3	-150.8	-169.5
65w	9.61	69.6	110.7	-148.4	87.2	-161.3	-94.9	-175.9

CONSIDERAÇÕES FINAIS

Em estudos QSAR, de modo geral, são utilizadas estruturas obtidas em um ambiente livre do receptor para a geração dos descritores moleculares que não é consistente com o que ocorre em meio biológico, podendo induzir à falha na interpretação do modelo construído. Por essa razão, descritores 3D não são indispensáveis na obtenção de modelos QSAR satisfatórios, sendo as metodologias 2D, como o MIA-QSAR, uma alternativa para a redução de esforços, de tempo e de custo computacional, pela busca racional de novos compostos bioativos e a melhoria de protótipos já existentes.

O modelo MIA-QSAR desenvolvido neste trabalho mostrou-se preditivo e bioquimicamente interpretável, demonstrando a utilidade dos descritores bidimensionais nesse tipo de análise. O isomerismo conformacional do isoflurano evidenciou significativa diferença entre suas conformações preferenciais obtidas em um ambiente livre da enzima e as estruturas bioativas disponíveis no PDB.

De modo similar, as conformações otimizadas da DFX não coincidiram com a estrutura bioativa (PDB: 5FX) ligada à enzima xiluloquinase. Tais resultados reforçam a necessidade de cautela no uso de conformações livres em metodologias QSAR, bem como evidenciam a importância da interação ligante-receptor nesses estudos. Como sugestão, algumas metodologias de QSAR bidimensional podem ser empregadas para a geração de modelos satisfatoriamente preditivos e interpretáveis, do ponto de vista bioquímico. Além disso, existem caminhos para a implementação e o aprimoramento de técnicas de QSAR 2D, visando à otimização do processo de desenvolvimento de novos fármacos e agroquímicos.