

## QSAR-3D e Docking Molecular de Derivados de Ácidos *N*-arilantranílicos com Atividade Inibitória na Enzima Catepsina L

Camila da Câmara Lopes<sup>a</sup>, Renato Queiroga Pordeus<sup>a</sup>, Arlene Gonçalves Corrêa<sup>b</sup>, Boaz Galdino Oliveira<sup>a</sup>, Mauro Alves Bueno<sup>a\*</sup>

<sup>a</sup>Universidade Federal da Bahia, Instituto de Ciências Ambientais e Desenvolvimento Sustentável, 47.808-021, Barreiras, BA, Brasil.

<sup>b</sup>Universidade Federal de São Carlos, Departamento de Química, 13565-905 São Carlos-SP, Brasil.

Article history: Received: 02 March 2015; revised: 21 September 2015; accepted: 09 December 2015. Available online: 08 February 2016. DOI: <http://dx.doi.org/10.17807/orbital.v1i1.714>

**Abstract:** In this work, a molecular modeling study of *N*-arylanthranilic acid derivatives with inhibitory action on the Cathepsin L enzyme is presented. Firstly, a 3DQSAR study was carried out in order to identify the most important molecular fields, in particular those formed by the C=O, C—O—C and —NO<sub>2</sub> groups, by which the experimental biological activity data were satisfactorily predicted through the correlation ( $R^2 = 0.99$ ) and determination ( $Q^2 = 0.66$ ) coefficient values. In agreement with the noncompetitive inhibition mechanism, the molecular docking calculations revealed that preferential interactions pose outside of the Cathepsin L active site. The correlation between the docking scores and biological activity data are fairly linear, indicating that the strongest and weakest bonded compounds are those with highest and lowest biological activities, respectively.

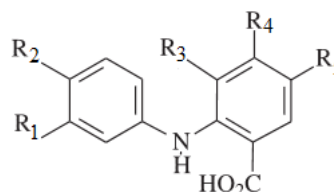
**Keywords:** 3DQSAR; docking; cathepsin L; *N*-arylanthranilic acids

### 1. INTRODUÇÃO

A aterosclerose é uma condição inflamatória crônica ocasionada pela formação de placas ou ateromas constituídos por lipídios ou tecidos fibrosos dentro dos vasos sanguíneos [1-3]. Uma das rotas bioquímicas preferenciais para o tratamento da aterosclerose é a inibição das proteases de cisteína ou catepsinas lisossomais [4]. Dentre a classe das catepsinas conhecidas, é bem estabelecido que algumas se originam nos lisossomos, mas também podem se acumular em outros tecidos, sendo exatamente através destes que a aterosclerose pode se desenvolver [5-6]. Segundo Brömme e coautores [7], a perda de elasticidade em tecidos vivos deve ser motivada pela Catepsina L, conforme também demonstrado nos trabalhos de Turk e colaboradores [8].

Na proposta de obter novos inibidores das Catepsina L e V, Vieira e Corrêa junto a seus colaboradores [9] investigaram a atividade (Tabela 1) de uma série de derivados de ácidos *N*-arilantranílicos (Figura 1). Os compostos foram sintetizados através de uma reação de Ullmann [10] sob irradiação de

micro-ondas. Os ensaios de inibição das enzimas Catepsina L e V mostraram um mecanismo não competitivo.



**Figura 1.** Scaffold de derivados dos ácidos *N*-arilantranílicos com a indicação dos grupos substituintes R<sub>1</sub>, R<sub>2</sub>, R<sub>3</sub>, R<sub>4</sub> e R<sub>5</sub> [9].

Com base nos resultados experimentais de cinética de reação, bem como pela necessidade de conhecer os potenciais sítios de ligação, no presente trabalho foi desenvolvido um estudo de modelagem molecular da série de 17 congêneres dos ácidos *N*-arilantranílicos (Figura 2) [11-15]. Independente do alvo biológico, os modelos de regressão conhecidos como QSAR (do inglês, *Quantitative Structure-Activity Relationship*) [16] bidimensionais apresentam uma vasta diversidade, embora metodologias multidimensionais tenham sido cada vez mais

\*Corresponding author. E-mail: [mauro.bueno@ufob.edu.br](mailto:mauro.bueno@ufob.edu.br)

utilizadas [17]. Dentre estas, o protocolo tridimensional, ou QSAR-3D, é um dos mais difundidos [18-19], cuja concepção está na utilização de descritores de natureza lipofílica, eletrônica e estérica originados do alinhamento das estruturas tridimensionais da série de congêneres. Com este alinhamento são gerados estatisticamente os campos moleculares [20], os quais mimetizam os grupos farmacofóricos mais importantes para a ação biológica de toda a série estudada. Concernente à modelagem dependente do alvo biológico, o *docking*

molecular [21] também tem sido muito utilizado na projeção de novos compostos, em particular aqueles com ação na protease Catepsina, conforme documentado por Brömme e colaboradores [22]. Na prática, a importância dos estudos de *docking* molecular está na identificação não apenas dos sítios ativos enzimáticos mais importantes, mas também das interações intermoleculares hidrofílicas e lipofílicas que norteiam o processo bioquímico de inibição da Catepsina L.

**Tabela 1.** Valores de IC<sub>50</sub> e pIC<sub>50</sub> para o conjunto de 17 congêneres do ácido *N*-arilantranílico.

N	Compostos	IC <sub>50</sub> ( $\mu$ M)	pIC <sub>50</sub>	N	Compostos	IC <sub>50</sub> ( $\mu$ M)	pIC <sub>50</sub>
1	<b>5a</b>	7,2	5,1	10	<b>5j</b>	2,1	5,7
2	<b>5b</b>	5,7	5,2	11	<b>5k</b>	50	4,3
3	<b>5c</b>	26	4,6	12	<b>5l</b>	50	4,3
4	<b>5d</b>	50	4,3	13	<b>5m</b>	50	4,3
5	<b>5e</b>	2,5	5,6	14	<b>5n</b>	18,2	4,7
6	<b>5f</b>	45	4,3	15	<b>5o</b>	7,2	5,1
7	<b>5g</b>	50	4,3	16	<b>5p</b>	11	4,9
8	<b>5h</b>	43	4,4	17	<b>5q</b>	3,2	5,5
9	<b>5i</b>	2,9	5,5	—	—	—	—

## 2. DETALHES COMPUTACIONAIS

### Geometrias otimizadas e alinhamento molecular

A otimização das 17 estruturas derivadas do ácido *N*-arilantranílico (Figura 2) foram teoricamente obtidas no nível da teoria B3LYP/6-31G(d,p) com todos os cálculos processados pelo *software* Gaussian 03W [23]. O alinhamento molecular foi realizado com os *softwares* OpenBabelGui [24] e Open3DAlign [25-26]. Para o alinhamento todas as estruturas foram consideradas rígidas [25, 33].

### Construção de modelos de QSAR

De acordo com os dados de atividade biológica em termos de IC<sub>50</sub> e pIC<sub>50</sub> listados na Tabela 1, os modelos de QSAR-3D foram construídos pelo *software* Open3DQSAR [27]. Dos 17 derivados do ácido *N*-arilantranílico, 14 foram distribuídos como pertencentes ao conjunto de treinamento enquanto 3 ao conjunto teste (Estruturas **5d**, **5j** e **5o** da Tabela 1), os quais apresentam valores baixos, moderados e altos de pIC<sub>50</sub>. As ilustrações dos mapas de campos moleculares foram geradas através do *software* Pymol 1.3 [28].

### Cálculos de *docking*

Para a simulação de *docking* molecular foi

utilizado o *software* AutoDock 1.5.6 [29], cuja estrutura cristalográfica da enzima Catepsina L foi obtida através do acesso ao *Protein Data Bank* [30] inserindo o código PDB = 3BC3 [31]. Na simulação de *docking*, considerando o mecanismo não competitivo previamente descrito, a enzima já sem o ligante foi colocada inteiramente dentro do *gridbox*, com espaçamento mínimo de 3,375 Å. A visualização dos resultados de *docking* foi feita com o *software* Discovery Studio Visualizer 4.0 [32].

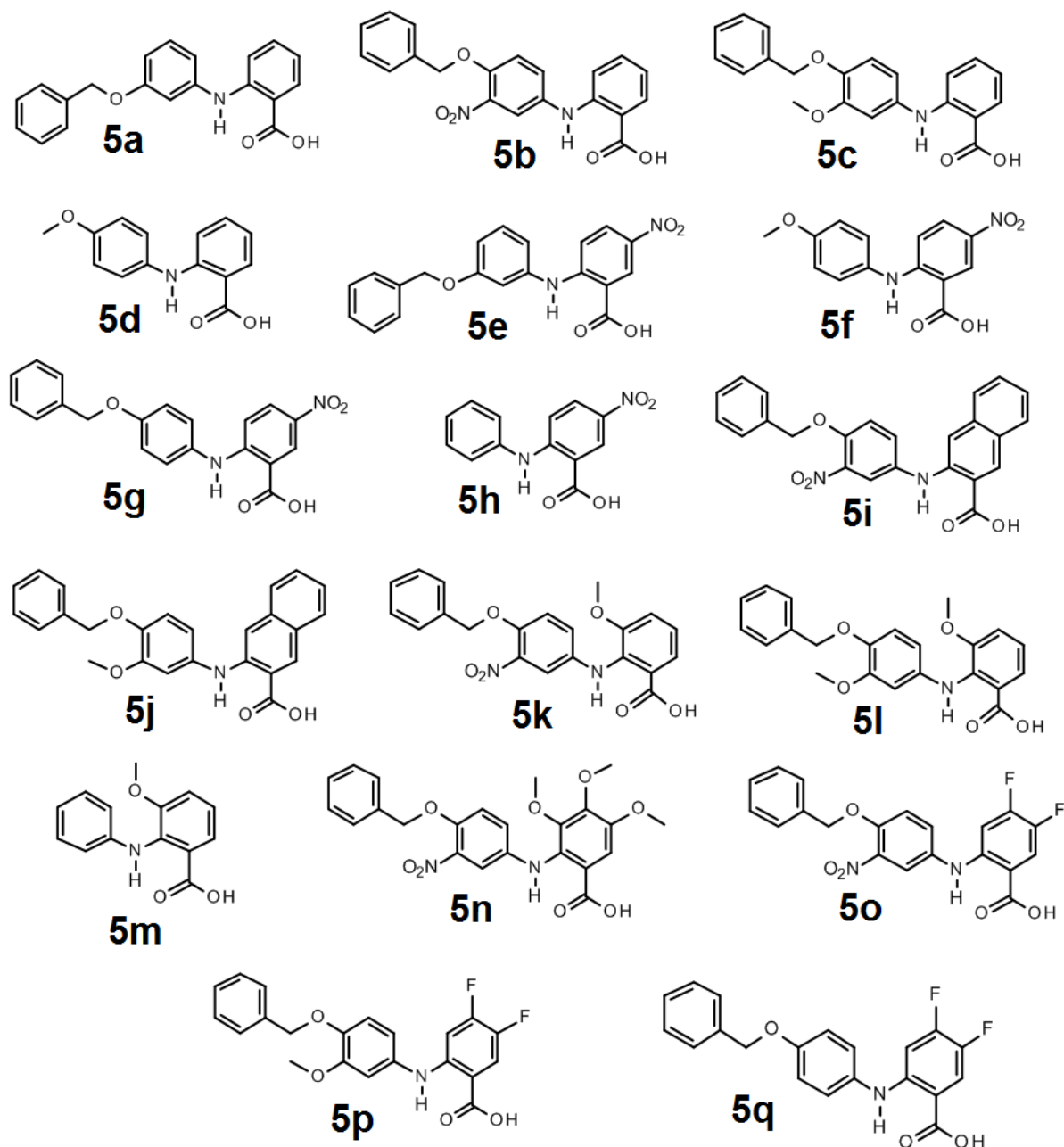
## 3. RESULTADOS E DISCUSSÃO

### 3.1 QSAR-3D

De acordo com o alinhamento apresentado na Figura 3, em geral observa-se que os anéis fenil-benzóicos, fenil-amino e benziloxi, como também os grupos —NH e —CO se mantiveram no mesmo plano. No procedimento estatístico, foram necessárias três componentes principais que alcançaram 99,46 % da variância explicada para o conjunto de dados. Para a construção do modelo de QSAR-3D, o algoritmo dos mínimos quadrados parciais ou PLS (do inglês, *Partial Least Squares*) [33-35] foi utilizado, gerando coeficientes de determinação R<sup>2</sup> = 0,99 e R<sup>2</sup><sub>(predição)</sub> = 0,83 para o conjunto de treinamento (14 estruturas) e teste (3 estruturas), respectivamente. Para a validação

cruzada, o coeficiente de determinação do conjunto de predição apresentou o seguinte valor  $Q^2 = 0,66$ . O valor do desvio padrão da validação cruzada do conjunto de teste ou *sPRESS* (do inglês, *Predicted Residual Sum of Squares*) apresentou um valor de 0,29 [36]. Conforme os valores dos resíduos menores que 0,07 e de acordo com Lima Neto *et al* [37], em

que a alta capacidade de predição de novos derivados biologicamente ainda mais potentes foi estabelecida pelos valores de  $Q^2 > 0,5$  e também  $R^2 > 0,9$  bem como  $R^2_{(pred)} > 0,7$ . Observa-se que os resultados obtidos neste trabalho estão em total conformidade com estes valores de referência.



**Figura 2.** Congêneres do ácido *N*-arilantranílico.

A Tabela 2 lista os valores experimentais ou observados (obs) e preditos (pred) de  $pIC_{50}$  para a série dos 17 derivados do ácido *N*-arilantranílicos,

cujas relações são graficamente ilustradas pela Figura 4 e matematicamente validadas pela equação (1), a qual gerou um coeficiente de correlação de 0,95:

$$pIC_{50(predito)} = 0,90.pIC_{50(observado)} + 0,41 \quad / \quad R^2 = 0,95 \quad (1)$$

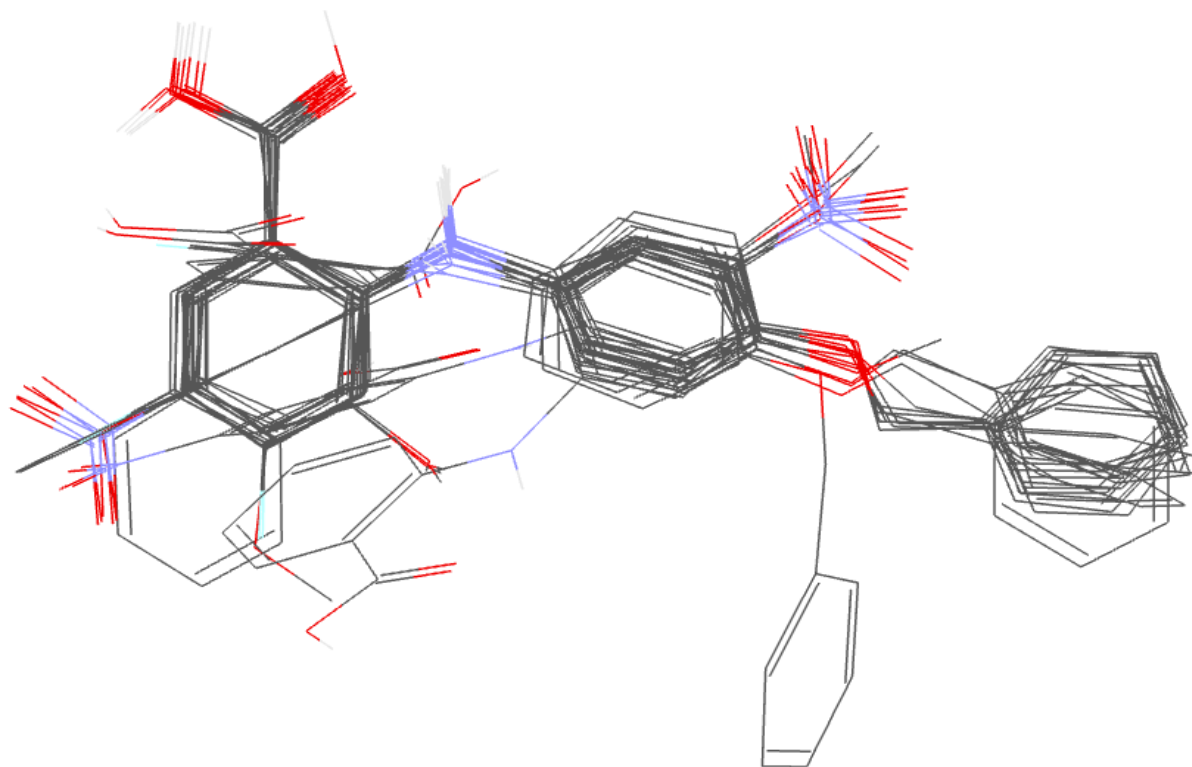


Figura 3. Alinhamento do conjunto de 17 derivados do ácido *N*-arilantranílico.

Tabela 2. Valores observados experimentalmente e preditos pelo modelo QSAR-3D.

N	Compostos	pIC <sub>50</sub> (pred)	pIC <sub>50</sub> (obs)	Resíduo	N	Compostos	pIC <sub>50</sub> (pred)	pIC <sub>50</sub> (obs)	Resíduo
1	5a	5,16	5,14	0,02	10	5j	5,27	5,68	-0,41
2	5b	5,20	5,24	-0,04	11	5k	4,34	4,30	0,04
3	5c	4,65	4,58	0,07	12	5l	4,30	4,30	0,00
4	5d	4,30	4,30	0,00	13	5m	4,29	4,30	-0,01
5	5e	5,57	5,60	-0,03	14	5n	4,76	4,74	0,02
6	5f	4,34	4,35	-0,01	15	5o	4,93	5,14	-0,21
7	5g	4,22	4,30	-0,08	16	5p	4,94	4,95	-0,01
8	5h	4,37	4,37	0,00	17	5q	5,49	5,50	-0,01
9	5i	5,56	5,54	0,02	—	—	—	—	—

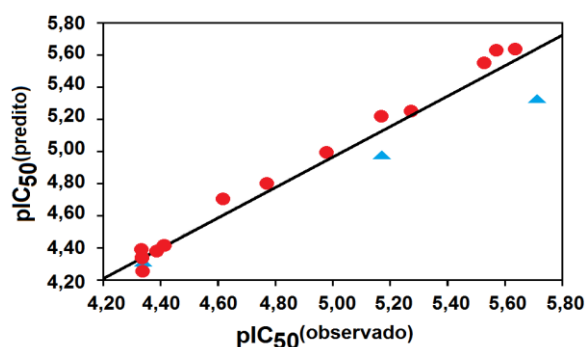
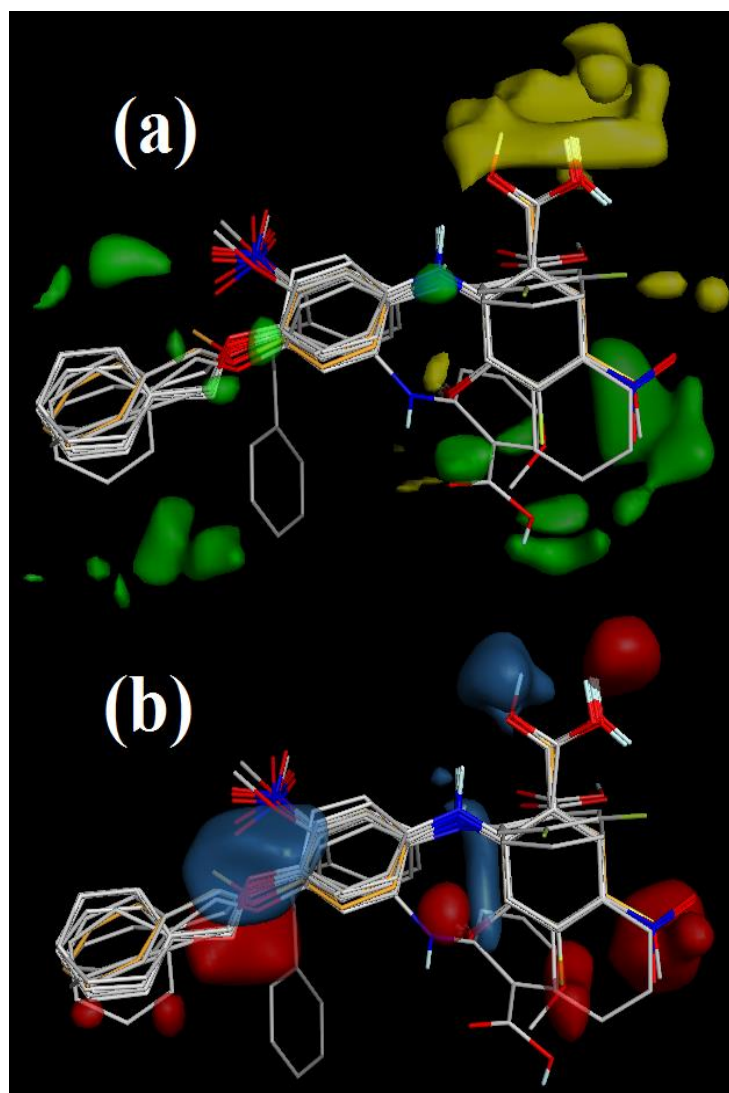


Figura 4. Correlação entre os valores de pIC<sub>50</sub> preditos e observados. Conjunto de treinamento (bolas vermelhas) e conjunto teste (triângulos azuis).

Uma das etapas da proposição do modelo de QSAR-3D concerne à localização de grupos farmacofóricos responsáveis pela atividade biológica da série de congêneres mediante a diagramação dos Campos de Interação Molecular ou MIFs (do inglês, *Molecular Interaction Fields*) [38]. Estes campos são resultantes das interações entre uma sonda padrão DRY e outra CH<sub>3</sub><sup>+</sup> [39] e o conjunto de moléculas alinhadas no *grid* tridimensional de 1,0 Å. Como resultado, na prática tem-se as contribuições das interações de natureza estérica e eletrostática representadas pelas energias de van der Waals e

coulômbica, respectivamente. Com o objetivo de reduzir a colinearidade entre os descritores bem como eliminar aqueles que não possuem informações relevantes para o modelo, utilizamos variância inferior a 1,0 e baseados no campo de força MMFF4

[40], as regiões que representam as contribuições estérica e eletrostática foram definidas entre 30 e -30 kcal.mol<sup>-1</sup>. Perante estes critérios, foram gerados os mapas de contorno apresentados nas Figuras 5(a) e 5(b).



**Figura 5.** Isosuperfícies dos campos moleculares estéricos (a) (positivo = verde e negativo = amarelo) e eletrostáticos (b) (positivo = azul e negativo = vermelho) do modelo QSAR-3D.

Na isosuperfície representada na Figura 5(a), as áreas amarelas, onde se localiza o grupo -COOH, são regiões em que a substituição por grupos volumosos contribuem negativamente para a atividade biológica. Em contra partida, nas regiões verdes, grupos volumosos são muito favoráveis para contribuir positivamente para a atividade biológica. Com relação à Figura 5(b), os contornos em vermelho representam regiões nas quais a presença de grupos doadores de densidade eletrônica aumenta a atividade, enquanto que, em geral, nas regiões destacadas em azul, a presença de grupos

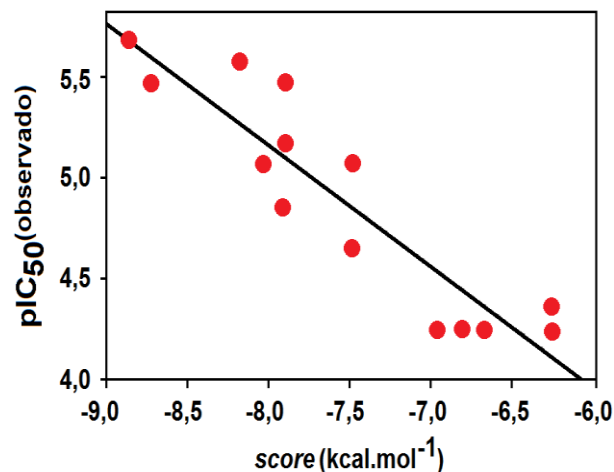
deficientes de elétrons aumenta a atividade biológica. As informações contidas na Figura 5 estão de acordo com o que se observa experimentalmente ao se comparar os valores de IC<sub>50</sub>, conforme documentado por Marques *et al* [9] que **5i** e **5j** são os compostos mais potentes. Nestes, a presença do grupo naftil cumpre os requisitos estérico e lipofílico; e também o requisito eletrônico, por ser um grupo doador de densidade eletrônica.

### 3.2 Docking molecular

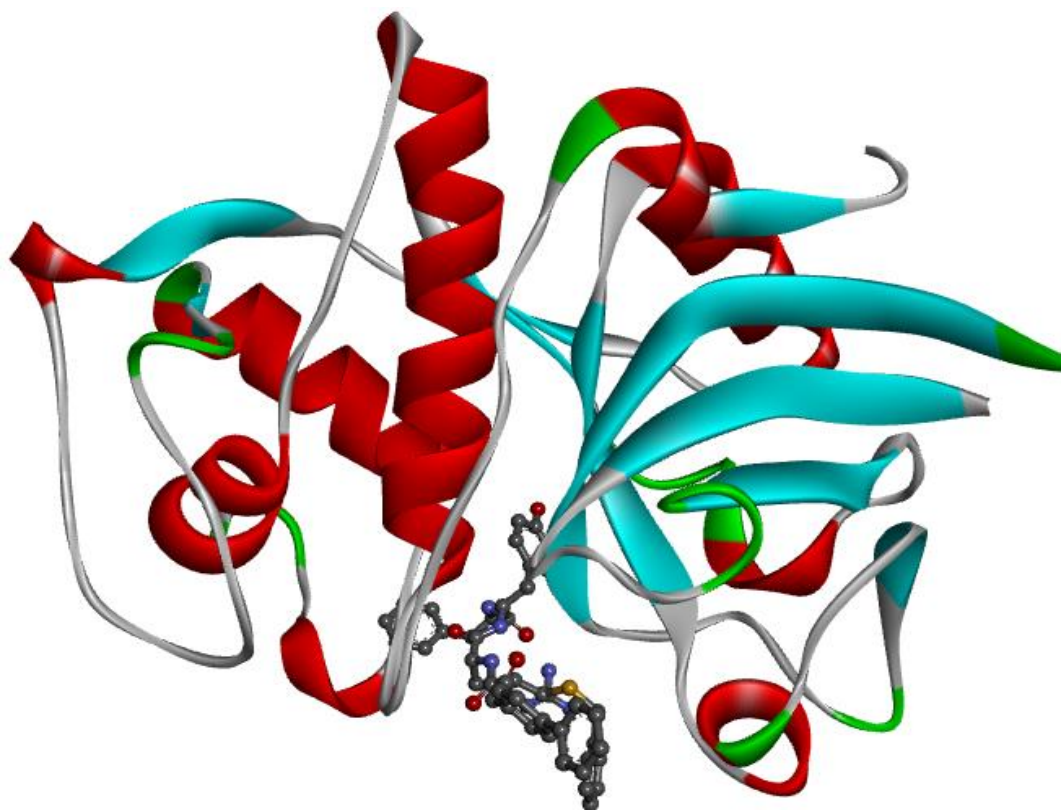


Os resultados das energias de *docking* para a série de congêneres do ácido *N*-arilantranílico, frente aos valores experimentais de atividade biológica  $pIC_{50}$  é ilustrado na Figura 6. Observa-se claramente uma tendência linear, em que os maiores valores de potência biológica apresentam *scores* mais negativos. Comparativo a estrutura cristalográfica da Catepsina L (Figura 7), as *poses* das conformações mais estáveis dos congêneres do ácido *N*-arilantranílico são apresentadas na Figura 8. Uma vez já sendo de conhecimento a não competitividade destes compostos no substrato da Catepsina L [9], os resultados de *docking* apresentam-se eficazes na reprodução desta propriedade experimental pelo fato de todas as estruturas interagirem fora da região do sítio ativo. Na estrutura da Catepsina L e o ligante cristalográfico (Figura 7), o qual interage com o sítio ativo através dos aminoácidos cisteína (Cys25) e histidina (His163) [31], as conformações mais estáveis dos ligantes (Figura 8) indicam que os compostos **5a** e **5e** alinham-se em uma região próxima ao sítio ativo, embora interajam com os aminoácidos triptofano (Trp169), glicina (Gln19) e

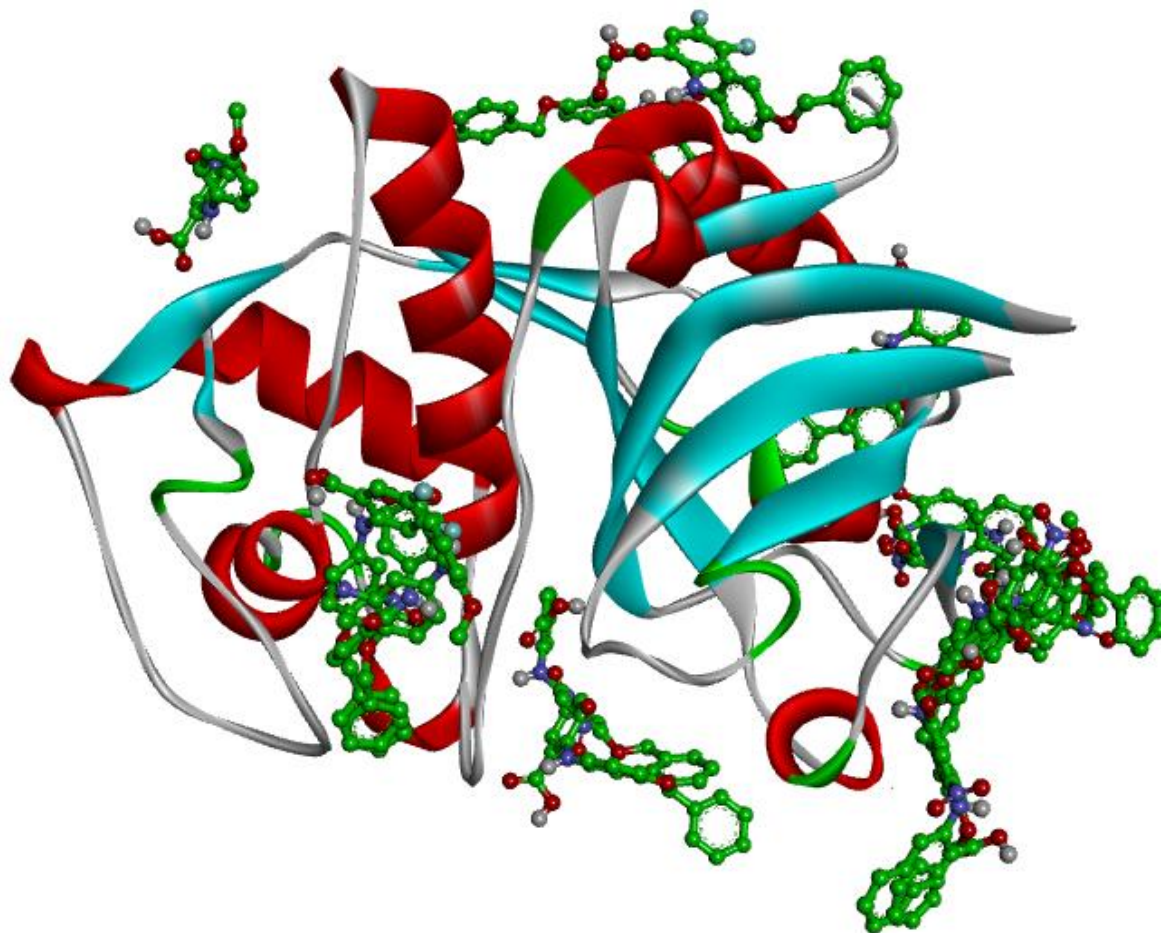
glutamina (Gly20). É digno de nota que a *pose* de *docking* mais estável do composto **5j** ilustrada na Figura 9 corrobora com o perfil experimental do mecanismo não competitivo, e que, diferencia-se totalmente das poses dos compostos **5a** e **5e** bem como do ligantes cristalográfico.



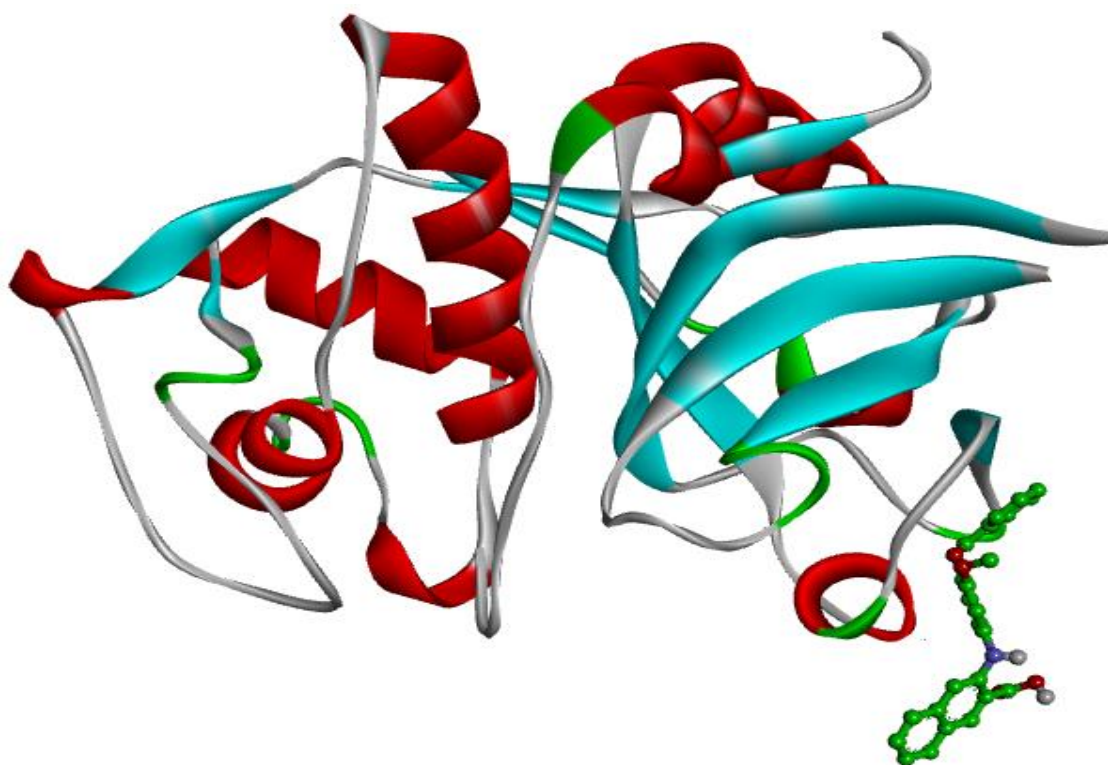
**Figura 6.** Gráfico de correlação entre os valores de *score* de *docking* e  $pIC_{50}$ .



**Figura 7.** Estrutura cristalográfica da enzima Catepsina L, contendo um ligante em seu sítio ativo [31] (código PDB = 3BC3).



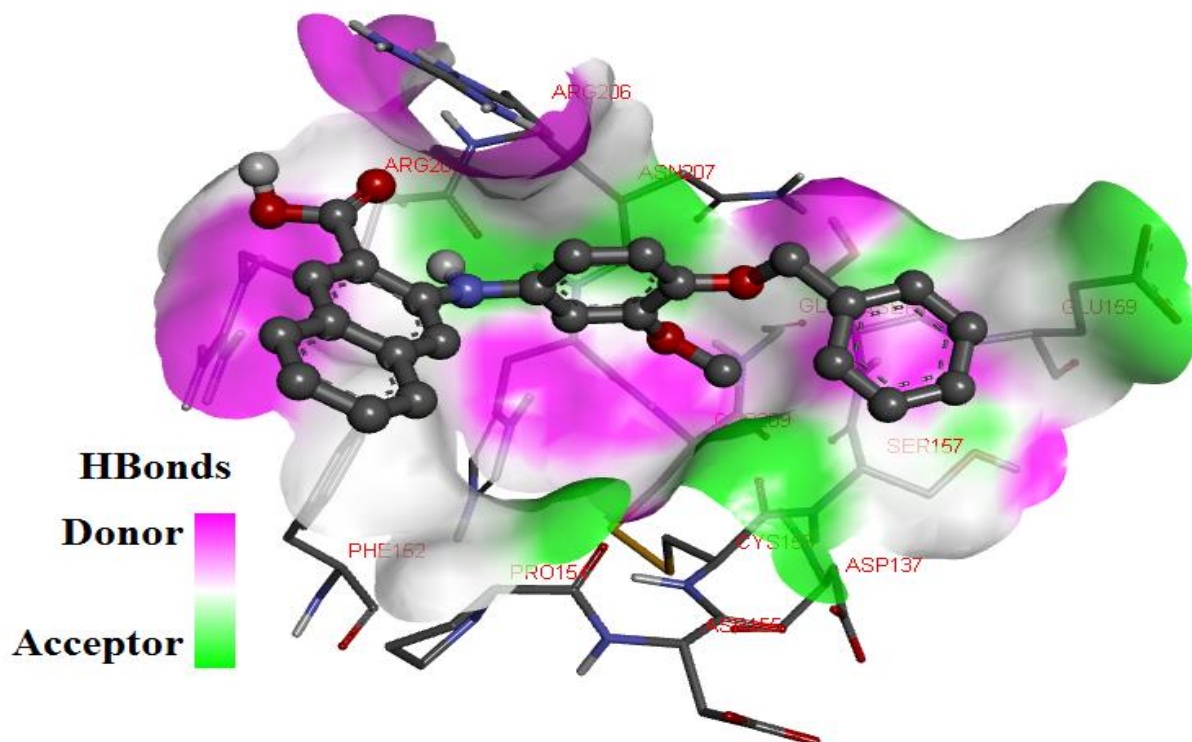
**Figura 8.** Superposição de *poses de docking* para toda a série de congêneres do ácido *N*-arilantranílico.



**Figura 9.** Composto 5j em cavidade fora do sitio catalítico da Catepsina L.

Analisando isoladamente o melhor resultado de *docking*, com *score* de  $-8,4 \text{ kcal.mol}^{-1}$  para o composto **5j**, a superfície que envolve o sítio de

ligação (Figura 10) indica as áreas doadoras e receptoras, nas quais as ligações de hidrogênio são favoráveis.



**Figura 10.** Regiões doadoras e receptoras de ligação de hidrogênio entre o composto **5j** e a enzima Catepsina L.

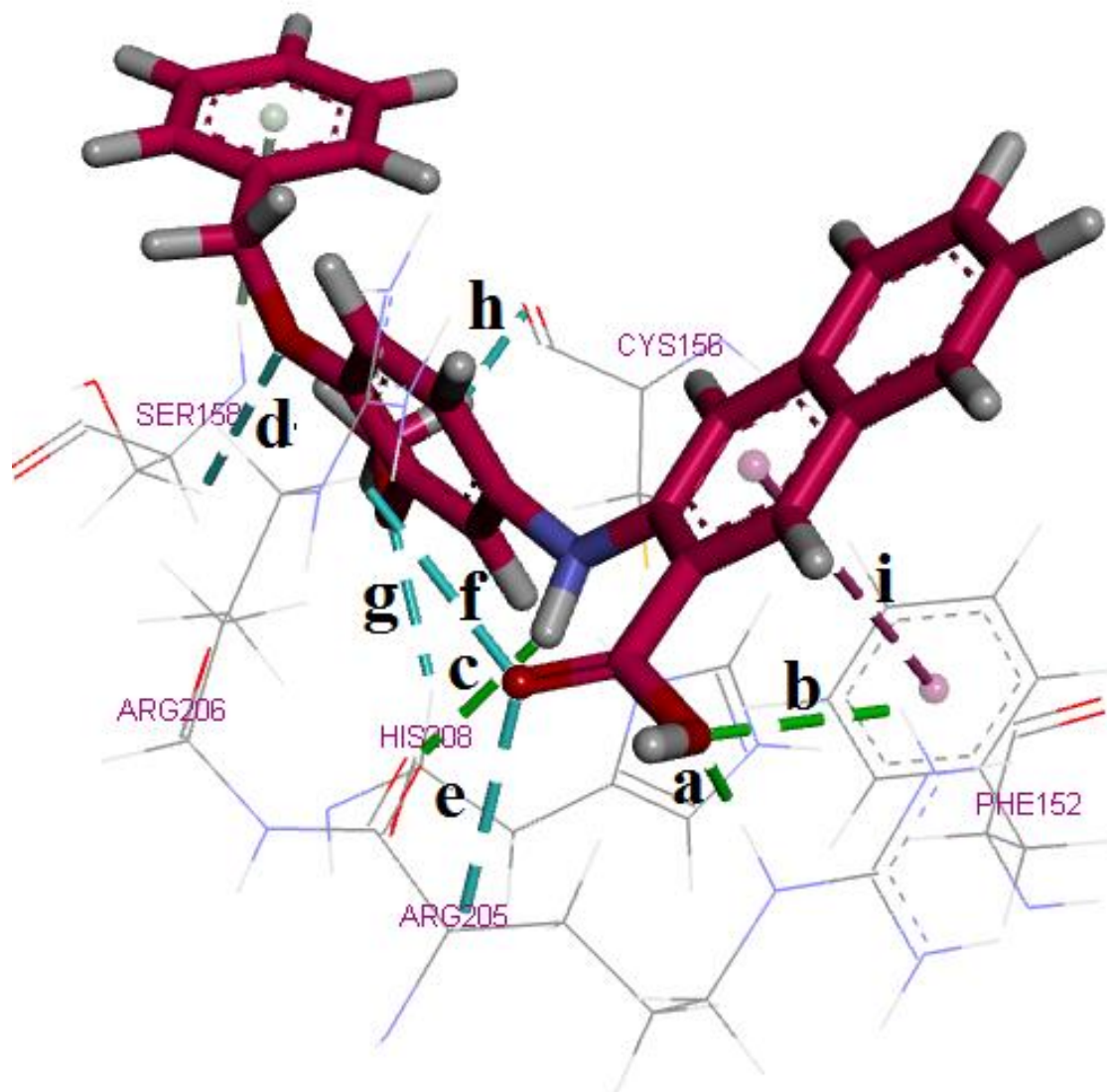
Analisando a Figura 10, entretanto, claramente os grupos C–O–C e C=O do composto **5j** estão aptos a interagirem com regiões eletrostaticamente positivas, ou seja, formarem ligações de hidrogênio com as estruturas dos aminoácidos da Catepsina L. Estas ligações de hidrogênio, em que, são concebidas como sendo hidrofóbica ou hidrofílica, são ilustradas na Figura 11, cujos valores de distância são organizados na Tabela 3. O grupo C–O–H destaca-se com duas ligações de hidrogênio (**a** e **b**) com a Arginina (Arg205), cujos valores são 1,60 e 2,81 Å. Estes valores são típicos de ligações de hidrogênio de força moderada e forte [41], enfatizando que a distância de 1,60 Å cabe na classificação de uma interação com energia muito estável. Por outro lado, o grupo C–O–C forma uma ligação de hidrogênio (**d**) com a Serina (Ser158), cuja distância de 3,07 Å pode ser considerada um valor muito alto típicos de sistemas intra- ou intermoleculares fracamente ligados [41]. Há também interações lipofílicas com distâncias de 4,95 Å entre o grupo naftil (**i**) e o aminoácido fanilalanina (Phe152). Este tipo de interação é considerado puramente lipofílico, sendo também reconhecido como *stacking* [42].

**Tabela 3.** Principais interações entre o composto **5j** e os resíduos da Catepsina L.

Grupos funcionais dos ligantes	Aminoácidos	Tipos de interações	Distâncias
O=C–O–H (a)	Arg205	Ligação de Hidrogênio	1,60
O=C–O–H (b)	Arg205	Ligação de Hidrogênio	2,81
N–H (c)	Arg205	Ligação de Hidrogênio	2,39
C–O–C (d)	Ser158	Ligação de Hidrogênio	3,07
H–O–C=O (e)	Arg205	Ligação de Hidrogênio	2,95
H–O–C=O (f)	Arg206	Ligação de Hidrogênio	2,99
CH <sub>3</sub> –O (g)	His208	Ligação de Hidrogênio	2,36
CH <sub>3</sub> –O (h)	Cys156	Ligação de Hidrogênio	2,09
Naftil (i)	Phe152	Interação lipofílica	4,95

Todos os valores são dados em Å.





**Figura 11.** Principais interações da molécula **5j** com os resíduos da Cathepsina L.

#### 4. CONCLUSÃO

O modelo de QSAR-3D para a série de congêneres do ácido *N*-arilantranílico apresentou uma ótima correlação entre as atividades biológicas observadas e previstas, além dos valores dos coeficientes de determinação  $R^2$  e  $Q^2$  acusarem que o modelo apresenta um ótimo perfil de predição. Também, os campos moleculares do modelo QSAR-3D apontam para importância da presença no anel do ácido benzoico de grupos doadores de elétrons (R3 e R4) como sendo aqueles que contribuem mais significativamente para o aumento da atividade biológica. Para a simulação de *docking*, as interações entre os grupos C–O–C e CH<sub>3</sub>–O com os aminoácidos Arginina e Cisteína para a estrutura mais estável (**5j**) são as mais prováveis, em que estes

grupos contribuem positivamente para o aumento da atividade biológica, assim como foi mostrado no QSAR-3D.

#### 5. AGRADECIMENTOS

Os autores agradecem ao CNPq.

#### 6. REFERÊNCIAS E NOTAS

- [1] Hopkins, P. N. *Physio. Rev.* **2013**, *93*, 1317. [[CrossRef](#)]
- [2] Gerber, Z. R. S.; Zielinsky, P. *Arq. Bras. Cardiol.* **1997**, *69*, 231. [[CrossRef](#)]
- [3] Santos, M. G.; Pegoraro, M.; Sandrini, F.; Macuco, E. C. *Arq. Bras. Cardiol.* **2008**, *90*, 301. [[CrossRef](#)]
- [4] Cheng, X. W.; Huang, Z.; Kuzuya, M.; Okumura, K.;

- Murohara, T. *Hypertension* **2011**, *58*, 978. [CrossRef]
- [5] Cheng, X. W.; Shi, G.; Kuzuya, M.; Sasaki, T.; Okumura, K.; Murohara, T. *Circulation* **2012**, *125*, 1551. [CrossRef]
- [6] Guha, S.; Padh, H. *Indian J. Biochem. Biophys.* **2008**, *45*, 75.
- [7] Yasuda, Y.; Li, Z.; Greenbaum, D.; Bogyo, M.; Weber, E.; Brömme, D. *J. Biol. Chem.* **2004**, *279*, 36761. [CrossRef]
- [8] Turk, V.; Stoka, V.; Vasiljeva, O.; Renko, M.; Sun, T.; Turk, B.; Turk, D. *Biochim. Biophys. Acta* **2012**, *1824*, 68. [CrossRef]
- [9] Marques, E. F.; Bueno, M. A.; Duarte, P. D.; Silva, L. R. S. P.; Martinelli, A. M.; dos Santos, C. Y.; Severino, R. P.; Brömme, D.; Vieira, P. C.; Corrêa, A. G. *Eur. J. Med. Chem.* **2012**, *54*, 10. [CrossRef]
- [10] Bueno, M. A.; Silva, L. R. S. P.; Corrêa, A. G. *J. Braz. Chem. Soc.* **2008**, *19*, 1269. [CrossRef]
- [11] Suleiman, M. M.; Isaev, S. G.; Klenina, O. V.; Ogurtsov, V. V. *J. Chem. Pharm. Res.* **2014**, *6*, 1219.
- [12] Puzyn, T.; Leszczynska, D.; Leszczynski, J. *Pract. Asp. Comput. Chem.* **2010**, 201.
- [13] Bala, S.; Kamboj, S.; Saini, V.; Prasad, D. N. *J. Chem.* **2013**, *1*, 1.
- [14] Yeates, T. O.; Kent, S. B. H. *Ann. Rev. Biophys.* **2012**, *41*, 41. [CrossRef]
- [15] Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. *FEBS J.* **2013**, *280*, 5705. [CrossRef]
- [16] Perkins, R.; Fang, H.; Tong, W.; Welsh, W. J. *Environ. Toxicol. Chem.* **2003**, *22*, 1666. [CrossRef]
- [17] Damale, M. G.; Harke, S. N.; Khan, F. A. K.; Shinde, D. B.; Sangshetti, J. N. *Min. Rev. Med. Chem.* **2014**, *14*, 35. [CrossRef]
- [18] Verma, J.; Khedkar, V. M.; Coutinho, E. C. *Curr. Top. Med. Chem.* **2010**, *10*, 95. [CrossRef]
- [19] Melo-Filho, C. C.; Braga, R. C.; Andrade, C. H. *Curr. Comput. Aided Drug. Des.* **2014**, *10*, 148. [CrossRef]
- [20] Cramer, R. D.; Wendt, B. J. *Chem. Inf. Model.* **2014**, *54*, 660. [CrossRef]
- [21] Meng, X. Y.; Zhang, H. X.; Mezei, M.; Cui, M. *Curr. Comput. Aided Drug. Des.* **2011**, *7*, 146. [CrossRef]
- [22] Du, X.; Chen, N. L. H.; Wong, A.; Craik, C. S.; Brömme, D.; *J. Biol. Chem.* **2013**, *288*, 34871. [CrossRef]
- [23] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; Gaussian, Inc., Pittsburgh PA, 2003.
- [24] O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. *J. Cheminform.* **2011**, *3*, 33. [CrossRef]
- [25] Tosco, P.; Balle, T. *J. Comput. Aid. Mol. Des.* **2011**, *25*, 777. [CrossRef]
- [26] Tosco, P.; Balle, T. *J. Cheminform.* **2011**, *3*, 39. [CrossRef]
- [27] Tosco, P.; Balle, T. *J. Mol. Model.* **2011**, *17*, 201. [CrossRef]
- [28] The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.
- [29] Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785. [CrossRef]
- [30] <http://www.rcsb.org/pdb/explore/explore.do?structureId=3BC3>, acessado em 10, janeiro, 2015.
- [31] Chowdhury, S. F.; Joseph, L.; Kumar, S.; Tulsidas, S. R.; Bhat, S.; Ziomek, E.; Menard, R.; Sivaraman, J.; Purisima, E. O. *J. Med. Chem.* **2008**, *51*, 1361. [CrossRef]
- [32] Discovery Studio Visualizer v4.0.100.13345, Copyright 2005-2013, Accelrys Software, Inc.
- [33] Tosco, P.; Balle, T. *J. Chem. Infor. Model.* **2012**, *52*, 302. [CrossRef]
- [34] Mehmood, T.; Liland, K.H.; Snipen, L.; Sæbø, S.; *Chemom. Int. Lab. Syst.* **2012**, *118*, 62. [CrossRef]
- [35] Vinzi, V. E.; Russolillo, G. *Partial least squares algorithms and methods. WIREs Comp. Stat.* **2013**, *5*, 1.
- [36] Ferreira, M. M. C. *J. Braz. Chem. Soc.* **2002**, *13*, 742. [CrossRef]
- [37] Lima Neto, Q. A.; Nogueira, M. A.; Reis, R. R.; Melo, E. B. *Rev. Bras. Canc.* **2006**, *52*, 345.
- [38] ElKerdawy, A.; Güssregen, S.; Matter, H.; Hennemann, M.; Clark, T. *J. Cheminform.* **2014**. [CrossRef]
- [39] Myint, K. Z.; Xie, X.-Q. *Int. J. Mol. Sci.* **2010**, *11*, 3846. [CrossRef]
- [40] Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490. [CrossRef]
- [41] Grabowski, S. J.; *Chem. Rev.* **2011**, *111*, 2597. [CrossRef]
- [42] Capim, S. L.; Santana, S. R.; Oliveira, B. G.; Rocha, G. B.; Vasconcellos, M. L. A. *J. Braz. Chem. Soc.* **2010**, *21*, 1718. [CrossRef]