# Internet Editonic Journal of Molecular Design

June 2003, Volume 2, Number 6, Pages 403–412

Editor: Ovidiu Ivanciuc

Special issue dedicated to Professor Haruo Hosoya on the occasion of the 65<sup>th</sup> birthday
Part 10

Guest Editor: Jun-ichi Aihara

# **Bone Resorption Inhibition Quantitative Structure—Activity Relationships for Aryl—Substituted Bisphosphonates**

Teodora Ivanciuc<sup>1</sup> and Ovidiu Ivanciuc<sup>2</sup>

<sup>1</sup> Texas A&M University at Galveston, Department of Oceanography and Marine Sciences, 5007 Avenue U, Galveston, Texas 77551

Received: June 9, 2003; Accepted: June 15, 2002; Published: June 30, 2003

# Citation of the article:

T. Ivanciuc and O. Ivanciuc, Bone Resorption Inhibition Quantitative Structure–Activity Relationships for Aryl–Substituted Bisphosphonates, *Internet Electron. J. Mol. Des.* **2003**, *2*, 403–412, http://www.biochempress.com.

<sup>&</sup>lt;sup>2</sup> Sealy Center for Structural Biology, Department of Human Biological Chemistry & Genetics, University of Texas Medical Branch, Galveston, Texas 77555–1157

Internet IFFONG Journal of Molecular Design

**BIOCHEM** Press http://www.biochempress.com

# Bone Resorption Inhibition Quantitative Structure–Activity Relationships for Aryl–Substituted Bisphosphonates<sup>#</sup>

Teodora Ivanciuc<sup>1</sup> and Ovidiu Ivanciuc<sup>2,\*</sup>

Received: June 9, 2003; Accepted: June 15, 2002; Published: June 30, 2003

Internet Electron. J. Mol. Des. 2003, 2 (6), 403-412

#### **Abstract**

Quantitative structure-activity relationships (QSAR) models for the bone resorption inhibition of 29 arylsubstituted bisphosphonates (ABP) were established with the CODESSA program. The QSAR models for the ABP bone resorption inhibition are obtained by selecting descriptors from a wide diversity of constitutional, topological, electrostatic and quantum structural indices. Standard quantum chemistry packages are used for optimizing the molecular geometry and for semi-empirical quantum computations at the AM1 level. A heuristic algorithm selects the best multiple linear regression equation according to the highest statistical indices; the predictive power of each QSAR model is estimated with the leave-one-out (LOO) cross-validation method. For the whole set of 29 compounds, the best QSAR model ( $r^2 = 0.8328$ ,  $r^2_{LOO} = 0.7479$ ,  $s^2 = 0.251$ , F = 29.88) is obtained with four quantum descriptors (minimum total interaction for a C-P bond, maximum valency of a N atom, minimum total interaction for a C-C bond, and hydrogen-acceptors surface area). A significant improvement of the statistical indices is obtained by deleting three outliers, when a fairly good QSAR is obtained ( $r^2 = 0.8827$ ,  $r^2_{LOO} = 0.8231$ ,  $s^2 = 0.118$ , F = 39.51) also with four quantum descriptors (minimum electron-electron repulsion for a H-N bond, minimum coulombic interaction for a C-P bond, maximum oneelectron reactivity index for a N atom, and minimum total interaction for a H-O bond). These results demonstrate the ability of AM1 quantum indices in modeling the bone resorption inhibition of aryl-substituted bisphosphonates.

**Keywords.** QSAR; quantitative structure–activity relationships; aryl–substituted bisphosphonates; bone resorption inhibition; CODESSA.

#### 1 INTRODUCTION

The bone resorption inhibition activity of geminal bisphosphonates makes them important compounds for the treatment of patients with postmenopausal osteoporosis, osteolytic bone metastases arising from breast cancer or multiple myeloma, tumor–induced hypercalcemia, and for Paget's disease of bone [1–5]. Bisphosphonates contribute up to \$5 billion to the global

<sup>&</sup>lt;sup>1</sup> Texas A&M University at Galveston, Department of Oceanography and Marine Sciences, 5007 Avenue U, Galveston, Texas 77551

<sup>&</sup>lt;sup>2</sup> Sealy Center for Structural Biology, Department of Human Biological Chemistry & Genetics, University of Texas Medical Branch, Galveston, Texas 77555–1157

<sup>&</sup>lt;sup>#</sup> Dedicated to Professor Haruo Hosoya on the occasion of the 65<sup>th</sup> birthday.

<sup>\*</sup> Correspondence author; E-mail: ovidiu\_ivanciuc@yahoo.com.

pharmaceutical market, and the design of new bone resorption inhibitors with improved potency is an active field of investigation.

In this paper we develop quantitative structure–activity relationships (QSAR) models for the bone resorption inhibition of aryl–substituted bisphosphonates (ABP). The relationships between ABP chemical structures and their activity in inhibiting bone resorption were established by selecting the structural descriptors for the multilinear regression equation from a wide range of topological, geometrical, electrostatic, and quantum indices.

## 2 MATERIALS AND METHODS

The success of the QSAR approach can be explained by the insight offered into the structural determination of chemical properties, and the possibility to estimate the properties of new chemical compounds without the need to synthesize and test them. The main QSAR hypothesis is that all properties (physical, chemical, and biological) of a chemical substance are statistically related to its molecular structure.

**Table 1.** Compound Number (See Figure 1 for Structure), Number from Ref. [4], Experimental  $ED_{50}$  ( $\mu g/kg$ ) [4], and  $pIC_{50}$  Experimental and Computed with Eq. (2)

1         5r         0.33         9.209         9.099           2         4j         0.4         8.979         8.422           3         5d         0.5         8.885         8.397           4         5f         0.6         8.835         8.348           5         4g         0.7         8.778         8.223           6         5p         0.7         8.741         8.417           7         4d         1         8.586         8.178           8         4i         1.0         8.565         7.964           9         5c         1.2         8.488         8.156           10         5h         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         <	No	Ref. [4]	$ED_{50}$	pIC <sub>50 exp</sub>	pIC <sub>50 Eq. (2)</sub>
3         5d         0.5         8.885         8.397           4         5f         0.6         8.835         8.348           5         4g         0.7         8.778         8.223           6         5p         0.7         8.741         8.417           7         4d         1         8.586         8.178           8         4i         1.0         8.565         7.964           9         5c         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.3869         8.307           17         5l         4         7.997         8.162           18         <	1	5r	0.33	9.209	9.099
4         5f         0.6         8.835         8.348           5         4g         0.7         8.778         8.223           6         5p         0.7         8.741         8.417           7         4d         1         8.586         8.178           8         4i         1.0         8.565         7.964           9         5c         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.3869         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20 <t< th=""><th>2</th><th>4j</th><th>0.4</th><th>8.979</th><th>8.422</th></t<>	2	4j	0.4	8.979	8.422
5         4g         0.7         8.778         8.223           6         5p         0.7         8.741         8.417           7         4d         1         8.586         8.178           8         4i         1.0         8.565         7.964           9         5c         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21 <th< th=""><th>3</th><th>5d</th><th>0.5</th><th>8.885</th><th>8.397</th></th<>	3	5d	0.5	8.885	8.397
6         5p         0.7         8.741         8.417           7         4d         1         8.586         8.178           8         4i         1.0         8.565         7.964           9         5c         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22 <td< th=""><th>4</th><th>5f</th><th>0.6</th><th>8.835</th><th>8.348</th></td<>	4	5f	0.6	8.835	8.348
7         4d         1         8.586         8.178           8         4i         1.0         8.565         7.964           9         5c         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23 <td< th=""><th>5</th><th><b>4g</b></th><th>0.7</th><th>8.778</th><th>8.223</th></td<>	5	<b>4g</b>	0.7	8.778	8.223
7         4d         1         8.586         8.178           8         4i         1.0         8.565         7.964           9         5c         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.306         8.136           24 <td< th=""><th>6</th><th><b>5</b>p</th><th>0.7</th><th>8.741</th><th>8.417</th></td<>	6	<b>5</b> p	0.7	8.741	8.417
9         5c         1.2         8.488         8.156           10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         <	7		1	8.586	8.178
10         5h         1.2         8.546         8.814           11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         5n         100         6.609         6.434           26	8	4i	1.0	8.565	7.964
11         5g         1.3         8.506         8.534           12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         5n         100         6.609         6.434           26         4a         300         6.053         6.168           27		5c	1.2	8.488	8.156
12         4b         1.4         8.407         8.259           13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         5n         100         6.609         6.434           26         4a         300         6.053         6.168           27         5k         500         5.930         5.203           28	10	5h	1.2	8.546	
13         4f         1.5         8.405         8.104           14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         5n         100         6.609         6.434           26         4a         300         6.053         6.168           27         5k         500         5.930         5.203           28         4l         1500         5.436         5.374	11	5g	1.3	8.506	
14         5a         1.5         8.402         8.605           15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         5n         100         6.609         6.434           26         4a         300         6.053         6.168           27         5k         500         5.930         5.203           28         4l         1500         5.436         5.374	12	4b	1.4	8.407	
15         5b         1.7         8.385         8.563           16         5e         1.7         8.369         8.307           17         5l         4         7.997         8.162           18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         5n         100         6.609         6.434           26         4a         300         6.053         6.168           27         5k         500         5.930         5.203           28         4l         1500         5.436         5.374		4f			
16       5e       1.7       8.369       8.307         17       5l       4       7.997       8.162         18       5q       7       7.741       7.903         19       5s       7.8       7.745       8.426         20       5j       10       7.624       7.697         21       4e       15       7.405       7.537         22       4c       20       7.253       8.023         23       4k       20       7.306       8.136         24       5i       20       7.142       8.070         25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374		5a	1.5	8.402	
17       5l       4       7.997       8.162         18       5q       7       7.741       7.903         19       5s       7.8       7.745       8.426         20       5j       10       7.624       7.697         21       4e       15       7.405       7.537         22       4c       20       7.253       8.023         23       4k       20       7.306       8.136         24       5i       20       7.142       8.070         25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374	15	<b>5</b> b	1.7	8.385	8.563
18         5q         7         7.741         7.903           19         5s         7.8         7.745         8.426           20         5j         10         7.624         7.697           21         4e         15         7.405         7.537           22         4c         20         7.253         8.023           23         4k         20         7.306         8.136           24         5i         20         7.142         8.070           25         5n         100         6.609         6.434           26         4a         300         6.053         6.168           27         5k         500         5.930         5.203           28         4l         1500         5.436         5.374			1.7	8.369	
19       5s       7.8       7.745       8.426         20       5j       10       7.624       7.697         21       4e       15       7.405       7.537         22       4c       20       7.253       8.023         23       4k       20       7.306       8.136         24       5i       20       7.142       8.070         25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374			4		
20       5j       10       7.624       7.697         21       4e       15       7.405       7.537         22       4c       20       7.253       8.023         23       4k       20       7.306       8.136         24       5i       20       7.142       8.070         25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374					
21       4e       15       7.405       7.537         22       4c       20       7.253       8.023         23       4k       20       7.306       8.136         24       5i       20       7.142       8.070         25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374					
22       4c       20       7.253       8.023         23       4k       20       7.306       8.136         24       5i       20       7.142       8.070         25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374	20				
23       4k       20       7.306       8.136         24       5i       20       7.142       8.070         25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374					
24     5i     20     7.142     8.070       25     5n     100     6.609     6.434       26     4a     300     6.053     6.168       27     5k     500     5.930     5.203       28     4l     1500     5.436     5.374					
25       5n       100       6.609       6.434         26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374		4k			
26       4a       300       6.053       6.168         27       5k       500       5.930       5.203         28       4l       1500       5.436       5.374	24				
<b>27 5k</b> 500 5.930 5.203 <b>28 4l</b> 1500 5.436 5.374					
<b>28 41</b> 1500 5.436 5.374		4a			
<b>29 5m</b> 7500 4.754 5.557					
	29	5m	7500	4.754	5.557

Figure 1. Structures of the 29 Aryl-Substituted Bisphosphonates [4].

The investigation of large and diverse molecular databases was made possible by the advent of general QSAR programs which integrate the computation of structural descriptors with the generation of structure–property models. CODESSA [6–8] is a widely used QSAR software [9–12] that describes numerically the chemical structure with more than one thousand structural descriptors from five classes: constitutional, graph theoretic and topological indices, geometrical, electrostatic,

and quantum-chemical descriptors. Using statistical methods, such as multiple linear regression (MLR) or PCA, the best descriptors are selected in the final structure-activity model. The bone resorption inhibition QSAR equations from this study were developed with CODESSA.

Figure 1. (Continued).

#### 2.1 Chemical Data

The experimental values for the bone resorption inhibition of 29 aryl–substituted bisphosphonates (Figure 1) taken from the literature [4] are presented in Table 1.

# 2.2 Previous QSAR Models

Using the Cerius<sup>2</sup> molecular field analysis QSAR, Kotsikorou and Oldfield [5] obtained a fairly good model for the set of 29 ABP:

$$pIC_{50} = 5.137 - 0.021 \text{ H}^{+}/326 + 0.052 \text{ CH}_{3}/335 - 0.073 \text{ H}^{+}/514 - 0.041 \text{ H}^{+}/564$$

$$n = 29 \quad r^{2} = 0.900 \quad F = 54.28 \quad r^{2}_{LOO} = 0.799$$
(1)

where  $H^+/i$  represents the interaction energy between a proton probe and the molecule at the grid point i, and  $CH_3/j$  represents the interaction energy between a methyl probe and the molecule at the grid point j.

# 2.3 Molecular Modeling

In the present investigation, the chemical structures were generated with HyperChem [13], the geometry optimization was performed with MOPAC [14] using the semiempirical quantum method AM1 [15] and the QSPR models were computed with CODESSA [16].

# 2.4 Structural Descriptors

The HyperChem structure files and the MOPAC output files were used by CODESSA to calculate 603 descriptors. CODESSA computes five classes of structural descriptors: constitutional (number of various types of atoms and bonds, number of rings, molecular weight); topological (Wiener index, Randić connectivity indices, Kier shape indices, information theory indices); geometrical (principal moments of inertia, shadow indices, molecular volume and surface area); electrostatic (when atomic charges are computed on the basis of atomic electronegativity: minimum and maximum partial charges, polarity parameter, charged partial surface area descriptors, hydrogen bond donor and acceptor surface indices); quantum (minimum and maximum partial charges, Fukui reactivity indices, dipole moment, HOMO and LUMO energies, molecular polarizability, minimum/maximum valency of an atom, minimum/maximum electron–electron repulsion for an atom, minimum/maximum exchange energy for a chemical bond, etc).

# 2.5 Multiple Linear Regression Model

From the whole set of 603 descriptors generated with CODESSA we have discarded descriptors with a constant value for all molecules in the data set. Descriptors for which values were not available for every molecule were assigned a zero value for the missing position. In the next step the number of descriptors was reduced by eliminating those with F-test values less than 1, t-test values less than 0.1 or correlation coefficients with pIC<sub>50</sub> less than 0.1; as a result of this descriptor

selection procedure, 348 descriptors remained for the 29 aryl–substituted bisphosphonates. CODESSA develops MLR models by a heuristic method that includes the following steps: (a) All quasi–orthogonal pairs of structural descriptors are selected from the initial set. Two descriptors are considered orthogonal if their intercorrelation coefficient  $r_{ij}$  is lower than 0.1. (b) CODESSA uses the pairs of orthogonal descriptors to compute the biparametric regression equations. The most significant 10 pairs of molecular descriptors are used in the third step. (c) To an MLR model containing n descriptors a new descriptor is added to generate a model with n+1 descriptors if the new descriptor is not significantly correlated with the previous n descriptors (intercorrelation coefficient lower than 0.8). Step (c) is repeated until MLR models with a given maximum number of descriptors are obtained.

#### 2.6 Model Validation

QSPR correlations can be observed not only because a causal relationship exists between a set of descriptors and a property, but also due to statistical bias resulting from errors in determining structural descriptors, experimental errors in measuring the property, or even due to chance alone. Model validation techniques are needed in order to distinguish between true and random correlations and to estimate the predictive power of the model. Although the QSAR equations developed with CODESSA are obtained by selection of descriptors from a large pool, several descriptor selection techniques are used in order to minimize the possibility of chance correlations. In a first step, from the initial pool of descriptors, CODESSA eliminates descriptors that do not correlate with pIC<sub>50</sub>, thus greatly reducing the dimensionality of the problem.

 Table 2. Notation of the CODESSA Descriptors Involved in the QSAR Models

_	Notation	Descriptor			
-	SD1	Maximum partial charge for a C atom (electrostatic)			
	SD2	Maximum partial charge for a O atom (electrostatic)			
	SD3	Maximum partial charge for a P atom (electrostatic)			
	SD4	Minimum total interaction for a C–P bond			
	SD5	Maximum valency of a N atom			
	SD6	Maximum electron-electron repulsion for a C atom			
	SD7	Maximum atomic state energy for a N atom			
	SD8	Minimum total interaction for a C–C bond			
	SD9	HASA H-acceptors surface area (quantum)			
	SD10	HBCA H-bonding charged surface area (quantum)			
	SD11	HBSA H-bonding surface area (quantum)			
	SD12	Minimum electron–electron repulsion for a H–N bond			
	SD13	Minimum coulombic interaction for a C–P bond			
	SD14	Minimum exchange energy for a C-N bond			
	SD15	Maximum one-electron reactivity index for a N atom			
	<b>SD16</b>	HA dependent HDSA-2/SQRT(TMSA) (electrostatic)			
	SD17	Minimum total interaction for a H–O bond			
.=	SD18	FPSA-3 Fractional PPSA (PPSA-3/TMSA) (electrostatic)			

Then, as described in the previous section, a heuristic algorithm selects only quasi-orthogonal groups of descriptors that are tested for correlation with  $pIC_{50}$  of aryl-substituted bisphosphonates. This selection algorithm ensures that the probability of obtaining a chance correlation is low, and maintains a reasonable searching time. Finally, the leave-one-out (LOO) cross-validation procedure is applied to each MLR equation in order to estimate the prediction power of bone resorption inhibition QSAR.

### **3 RESULTS AND DISCUSSION**

Table 2 presents the notation and a short description of the structural descriptors involved in the QSAR models reported in this investigation; more complete definitions of the descriptors can be found in the CODESSA manuals [16]. In Table 3 we present the descriptors and correlation coefficients for the best three QSAR models with one and two descriptors and with 29 and 26 compounds, respectively. For the whole set of 29 ABP, the best descriptors in monoparametric models are the electronegativity–based maximum partial charge for a C, O, and P atom, respectively. An improvement of r is obtained in QSAR models two descriptors, when  $r^2$  is between 0.7324 and 0.7431. All four descriptors involved in the biparametric QSAR are quantum–based, namely: minimum total interaction for a C–P bond, maximum valency of a N atom, maximum electron–electron repulsion for a C atom, and maximum atomic state energy for a N atom.

**Table 3.** Descriptors and Correlation Coefficients for the Best Three QSAR Models with One and Two Descriptors and with 29 and 26 Compounds, Respectively

n = 29		n = 26			
$x_1$	$x_2$	$r^2$	$x_1$	$x_2$	$r^2$
SD1		0.5899	SD1		0.4143
SD2		0.5899	SD2		0.4143
SD3		0.5899	SD3		0.4143
SD4	SD5	0.7431	SD6	<b>SD12</b>	0.7784
SD6	SD5	0.7327	<b>SD13</b>	<b>SD12</b>	0.7741
SD6	SD7	0.7324	SD13	SD14	0.7594

The QSAR statistics increase further in equations with three and four descriptors. Because adding more descriptors does not significantly improve the statistical indices, we have selected the QSAR with four descriptors as the best for the 29 ABP. The best three such QSAR models are:

$$pIC_{50} = -357.35 + 2.903 \text{ SD4} + 74.89 \text{ SD5} + 3.894 \text{ SD8} + 1.867 \cdot 10^{-2} \text{ SD9}$$

$$n = 29 \quad r^2 = 0.8328 \quad s^2 = 0.251 \quad F = 29.88 \quad r^2_{LOO} = 0.7479$$
(2)

$$pIC_{50} = -381.83 + 3.145 \text{ SD4} + 79.49 \text{ SD5} + 4.038 \text{ SD8} + 2.630 \cdot 10^{-2} \text{ SD10}$$

$$n = 29 \quad r^2 = 0.8313 \quad s^2 = 0.253 \quad F = 29.56 \quad r^2_{LOO} = 0.7374$$
(3)

$$pIC_{50} = -351.95 + 2.921 \text{ SD4} + 73.07 \text{ SD5} + 3.691 \text{ SD8} + 1.773 \cdot 10^{-2} \text{ SD11}$$

$$n = 29 \quad r^2 = 0.8296 \quad s^2 = 0.255 \quad F = 29.22 \quad r^2_{LOO} = 0.7428$$
(4)

Eqs. (1)–(3) shows that starting with a large selection of structural descriptors it is possible to find

several combinations of descriptors that provide models with similar good statistics. Moreover, owing to the errors in the experimental data, small statistical differences between QSRR equations are not significant, and it is almost impossible to select a "best" modeling equation in such cases. The plot of experimental vs. calculated pIC<sub>50</sub> for Eq. (2) is presented in Figure 2, and the pIC<sub>50</sub> calculated with Eq. (2) can be found in Table 1, column 5. Four quantum descriptors computed with the AM1 method are used in the QSAR model from Eq. (2), namely the minimum total interaction for a C–P bond, maximum valency of a N atom, minimum total interaction for a C–C bond, and hydrogen–acceptors surface area.

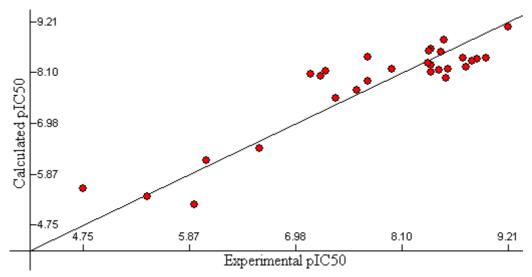


Figure 2. Experimental vs. calculated pIC<sub>50</sub> for Eq. (2).

Because the pIC<sub>50</sub> calculated with Eq. (2) for compounds **24**, **27**, and **29** have large deviations from the experimental values, we have removed these compounds and developed QSAR models with 26 ABP. In Table 3 we present the descriptors and correlation coefficients for the best three QSAR models with one and two descriptors, respectively. The best three QSAR models for the 26 ABP are:

$$pIC_{50} = -59.18 + 3.172 \text{ SD13} - 5.608 \cdot 10^{-2} \text{ SD12} - 1.081 \cdot 10^{3} \text{ SD15} + 4.985 \text{ SD16}$$

$$n = 26 \quad r^{2} = 0.8860 \quad s^{2} = 0.115 \quad F = 40.79 \quad r^{2}_{LOO} = 0.8055$$
(5)

$$pIC_{50} = 1.214 \cdot 10^{2} + 2.839 \text{ SD13} - 5.240 \cdot 10^{-2} \text{ SD12} - 9.812 \cdot 10^{2} \text{ SD15} - 13.07 \text{ SD17}$$

$$n = 26 \quad r^{2} = 0.8827 \quad s^{2} = 0.118 \quad F = 39.51 \quad r^{2}_{LOO} = 0.8231$$
(6)

$$pIC_{50} = -5.636 \cdot 10^{-1} + 3.099 \text{ SD13} - 5.500 \cdot 10^{-2} \text{ SD12} - 1.073 \cdot 10^{3} \text{ SD15} + 1.341 \cdot 10^{2} \text{ SD18}$$

$$n = 26 \quad r^{2} = 0.8816 \quad s^{2} = 0.120 \quad F = 39.09 \quad r^{2}_{LOO} = 0.8221$$
(7)

Compared with the QSAR obtained with 29 ABP, both calibration and LOO cross-validation r are significantly improved in Eqs. (5)–(7). Because  $r^2_{LOO}$  has a maximum value for Eq. (6), in Figure 3 we present the plot of experimental vs. calculated pIC<sub>50</sub> for Eq. (6). Four quantum descriptors computed with the AM1 method are used in the QSAR model from Eq. (6), namely the minimum electron-electron repulsion for a H–N bond, minimum coulombic interaction for a C–P

bond, maximum one–electron reactivity index for a N atom, and minimum total interaction for a H–O bond. Starting from a large collection of 603 constitutional, graph theoretic and topological indices, geometrical, electrostatic, and quantum–chemical descriptors, only electronic indices were selected in the best QSAR models, as can be seen from Table 2. These electronic indices, either based on the AM1 calculations or electronegativity–derived, offer good models for the ABP bone resorption inhibition. All these results suggest that electronic factors that control the molecular structure and inter–molecular interactions are the main factors that determine the bone resorption inhibition of aryl–substituted bisphosphonates.

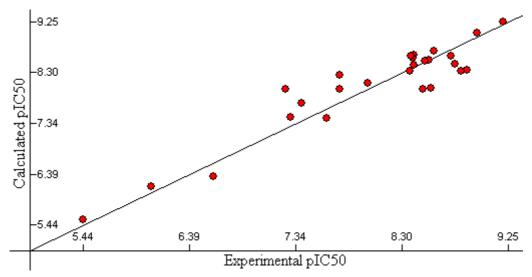


Figure 3. Experimental vs. calculated pIC<sub>50</sub> for Eq. (6).

#### 4 CONCLUSIONS

Quantitative structure–activity relationships (QSAR) models for the bone resorption inhibition of 29 aryl–substituted bisphosphonates were established with the CODESSA program. The QSAR models for the ABP bone resorption inhibition are obtained by selecting descriptors from a wide diversity of constitutional, topological, electrostatic and quantum structural indices. Standard quantum chemistry packages are used for optimizing the molecular geometry and for semi–empirical quantum computations at the AM1 level. A heuristic algorithm selects the best multiple linear regression equation according to the highest statistical indices; the predictive power of each QSAR model is estimated with the leave–one–out (LOO) cross–validation method. For the whole set of 29 compounds, the best QSAR model ( $r^2 = 0.8328$ ,  $r^2_{LOO} = 0.7479$ ,  $s^2 = 0.251$ , F = 29.88) is obtained with four quantum descriptors (minimum total interaction for a C–P bond, maximum valency of a N atom, minimum total interaction for a C–C bond, and hydrogen–acceptors surface area). A significant improvement of the statistical indices is obtained by deleting three outliers, when a fairly good QSAR is obtained ( $r^2 = 0.8827$ ,  $r^2_{LOO} = 0.8231$ ,  $s^2 = 0.118$ , F = 39.51) also with four quantum descriptors (minimum electron–electron repulsion for a H–N bond, minimum

coulombic interaction for a C–P bond, maximum one–electron reactivity index for a N atom, and minimum total interaction for a H–O bond). These results demonstrate the ability of AM1 quantum indices in modeling the bone resorption inhibition of aryl–substituted bisphosphonates. Together with the QSAR models proposed in this study, these three descriptors could be used to estimate the bone resorption inhibition for not yet synthesized or laboratory tested aryl–substituted bisphosphonates.

#### **5 REFERENCES**

- [1] C. M. Szabo, Y. Matsumura, S. Fukura, M. B. Martin, J. M. Sanders, S. Sengupta, J. A. Cieslak, T. C. Loftus, C. R. Lea, H.–J. Lee, A. Koohang, R. M. Coates, H. Sagami, and E. Oldfield, Inhibition of Geranylgeranyl Diphosphate Synthase by Bisphosphonates and Diphosphates: A Potential Route to New Bone Antiresorption and Antiparasitic Agents, *J. Med. Chem.* **2002**, *45*, 2185–2196.
- [2] P. Herczegh, T. B. Buxton, J. C. McPherson, III, A. Kovacs–Kulyassa, P. D. Brewer, F. Sztaricskai, G. G. Stroebel, K. M. Plowman, D. Farcasiu, and J. F. Hartmann, Osteoadsorptive Bisphosphonate Derivatives of Fluoroquinolone Antibacterials, *J. Med. Chem.* **2002**, *45*, 2338–2341
- [3] C. M. Szabo, M. B. Martin, and E. Oldfield, An Investigation of Bone Resorption and *Dictyostelium discoideum* Growth Inhibition by Bisphosphonate Drugs, *J. Med. Chem.* **2002**, *45*, 2894–2903.
- [4] L. Widler, K. A. Jaeggi, M. Glatt, K. Muller, R. Bachmann, M. A. R. Bisping, Born, R. Cortesi, G. Guiglia, H. Jeker, R. Klein, U. Ramseier, J. Schmid, G. Schreiber, Y. Seltenmeyer, and J. R. Green, Highly Potent Geminal Bisphosphonates. From Pamidronate Disodium (Aredia) to Zoledronic Acid (Zometa), *J. Med. Chem.* **2002**, *45*, 3721–3738.
- [5] E. Kotsikorou and E. Oldfield, A Quantitative Structure–Activity Relationship and Pharmacophore Modeling Investigation of Aryl–X and Heterocyclic Bisphosphonates as Bone Resorption Agents, *J. Med. Chem.* **2003**, *46*, 2932–2944.
- [6] A. R. Katritzky, U. Maran, V. S. Lobanov, and M. Karelson, Structurally Diverse Quantitative Structure–Property Relationship Correlations of Technologically Relevant Physical Properties, J. Chem. Inf. Comput. Sci. 2000, 40, 1–18.
- [7] A. R. Katritzky, L. Mu, V. S. Lobanov, and M. Karelson, Correlation of Boiling Points with Molecular Structure. 1. A Training Set of 298 Diverse Organics and a Test Set of 9 Simple Inorganics, *J. Phys. Chem.* **1996**, *100*, 10400–10407.
- [8] A. R. Katritzky, V. S. Lobanov, and M. Karelson, Normal Boiling Points for Organic Compounds: Correlation and Prediction by a Quantitative Structure–Property Relationship, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 28–41.
- [9] T. Ivanciuc and O. Ivanciuc, Quantitative Structure–Retention Relationship Study of Gas Chromatographic Retention Indices for Halogenated Compounds, *Internet Electron. J. Mol. Des.* **2002**, *1*, 94–107, http://www.biochempress.com.
- [10] R. Hiob and M. Karelson, Quantitative Relationship between Rate Constants of the Gas-Phase Homolysis of N-N, O-O and N-O Bonds and Molecular Descriptors, *Internet Electron. J. Mol. Des.* **2002**, *1*, 193–202, <a href="http://www.biochempress.com">http://www.biochempress.com</a>.
- [11] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, Quantitative Structure–Property Relationships for the Normal Boiling Temperatures of Acyclic Carbonyl Compounds, *Internet Electron. J. Mol. Des.* **2002**, *1*, 252–268, http://www.biochempress.com.
- [12] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, QSAR Models for the Dermal Penetration of Polycyclic Aromatic Hydrocarbons, *Internet Electron. J. Mol. Des.* **2002**, *1*, 559–571, <a href="http://www.biochempress.com">http://www.biochempress.com</a>.
- [13] HyperChem 5.1, Hypercube, Inc., Florida Science and Technology Park, 1115 N.W. 4th Street Gainesville, Florida 32601, U.S.A., E-mail info@hyper.com, www http://www.hyper.com.
- [14] MOPAC 6, adapted for Windows by V. Lobanov, www http://www.ccl.net.
- [15] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, AM1: A New General Purpose Quantum Mechanical Molecular Model, *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- [16] CODESSA 2.13, Semichem, 7204 Mullen, Shawnee, KS 66216, U.S.A., E-mail andy@semichem.com, www http://www.semichem.com.