Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Pharmaceutics 269 (2004) 51–60
international journal of pharmaceutics
www.elsevier.com/locate/ijpharm

Similarity approach to QSAR Application to antimycobacterial benzoxazines

Ana Gallegos [a], Ramon Carbó-Dorca [a], Robert Ponec [b,*], Karel Waisser [c]

[a] *Institute of Computational Chemistry, University of Girona, Campus Montilivi 17071, Girona, Spain*
[b] *Institute of Chemical Process Fundamentals, Czech Academy of Sciences, Suchdol 165 02, Prague 6, Czech Republic*
[c] *Faculty of Pharmacy in Hradec Králové, Charles University in Prague, Hradec Králové, Heyrovský str. 1203, CZ 500 55, Czech Republic*

Received 24 March 2003; received in revised form 22 July 2003; accepted 30 August 2003

The paper is dedicated to Professor Dr. Václav Horák (Georgetown University) on the occasion of his 80th anniversary

Abstract

The antimycobacterial activity in six series of substituted 3-phenyl-2*H*-benzoxazine-2,4(3*H*)-dithiones and 3-(phenyl)-4-thioxo-2*H*-benzoxazine-2,4(3*H*)-diones has been studied using a quantum molecular similarity approach. The approach is based on the use of fragment self-similarity measures as new universal molecular descriptors applicable for the design of novel theoretical QSAR models. Using this approach it was possible to show that while traditional QSAR models were able to describe the activity only within each of the six sets of studied molecules individually, the proposed approach is much more general and a single universal QSAR model describing the activity of all the 39 studied molecules in all the studied series together was built. The replacement of the oxo group by the thioxo group in position 4 on the benzoxazine ring of the antitubercular 3-(phenyl)-2*H*-benzoxazine-2,4(3*H*)-diones increases the activity, as well as the similar replacement in position 2.

Keywords: Quantum molecular similarity; Theoretical QSAR models; Mycobacterium; Tuberculostatics; Benzoxazine

1. Introduction

The current search for new antimycobacterial agents is very urgent, as tuberculosis has become the major emerging opportunistic infection. Approximately one-third of the world's population harbors *Mycobacterium tuberculosis* and it is at risk for developing the disease (Rouhi, 1999). The developing resistance to conventional antituberculotics is a stimulating factor in the research of new compounds. In addition, also the infections caused by non-tuberculous mycobacteria, e.g. *Mycobacterium avium* complex, show a rising occurrence among children, elderly, and HIV-infected patients (Bermudez and Young, 1995). In particular, 3-aryl-2*H*-1,3-benzoxazine-2,4(3*H*)-diones seemed a very prospective group of new antimycobacterial compounds (Waisser et al., 2001a). So far, several hundreds of new compounds have been synthesized, with the replacement of the oxo groups by thioxo groups. The replacement of oxygen by sulphur gave rise to antimycobacterial activity. In particular, we have described the synthesis and antimycobacterial activity of two groups of new antimycobacterial compounds, 3-phenyl-4-thioxo-2*H*-1,3-benzoxazine-2(3*H*)-ones and 3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dithiones (Waisser et al., 2000, 2001a). The most active deriva-

* Corresponding author. Tel.: +420-2-24311498; fax: +420-2-342073.
E-mail address: rponec@icpf.cas.cz (R. Ponec).

0378-5173/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpharm.2003.08.013

tives are comparable to INH with regard to the antimycobacterial activity against *M. tuberculosis* and more active against occasionally pathogenic strains. The main goal of this paper is to provide a theoretical rationale for the observed increase of antimycobacterial activity induced by the replacement of the oxo group in 3-aryl-2*H*-1,3-benzoxazine-2,4(3*H*)-diones by sulphur. Our approach is based on the use of molecular quantum similarity measures, which have recently been proved to provide new universal molecular descriptors applicable, instead of the traditional ones, in QSAR studies.

2. Materials and methods

2.1. Molecular quantum similarity measures

Molecular quantum similarity measures (MQSM) are the quantities allowing to give strict quantitative meaning to the intuitive concept of molecular similarity. These measures, seminally introduced by Carbó et al. (1980) are generally defined by Eq. (1):

$$Z_{AB}(\Omega) = \iint \rho_A(r_1)\Omega(r_1, r_2)\rho_B(r_2)dr_1dr_2 \quad (1)$$

from which it is evident that the similarity of two molecules A and B is quantitatively characterized by the value of the overlap of their corresponding electron density functions, $\rho_A(r)$ and $\rho_B(r)$, weighted in general case by a two-electron operator $\Omega(r_1,r_2)$. According to the actual choice of this operator, several types of QSM can be obtained. One of the most straightforward choices of the operator Ω is the Dirac delta function, which leads to the so-called overlap-like similarity measure, displayed in Eq. (2):

$$Z_{AB} = \int \rho_A(r)\rho_B(r)dr \quad (2)$$

In addition to these measures characterizing the similarity of two different molecules A and B, it is also useful to introduce the so-called quantum self-similarity measures, derived from the general Eq. (2), by identifying A = B, as can be seen in Eq. (3):

$$Z_{AA} = \int \rho_A(r)\rho_A(r)dr \quad (3)$$

2.2. Fragment self-similarity measures

The concept of fragment quantum self-similarity measures represents an empirical generalization of the general formula in Eq. (3), where the electron density function $\rho_A(r)$ of the whole molecule is replaced by the density function of a certain molecular fragment X, as shown in Eq. (4):

$$Z^X_{AA} = \int \rho^X_A(r)\rho^X_A(r)dr \quad (4)$$

The introduction of this concept arises from the empirical observation that a wealth of molecular properties is usually closely associated with the presence of certain molecular fragments (functional groups), which can be regarded as pharmacophores. As a consequence, it can be assumed that focusing just on the "active" fragment rather than on the whole molecule can hopefully improve the applicability of measures as new universal molecular descriptors in theoretical QSAR models.

The practical calculation of fragment self-similarity measures requires, of course, to specify the fragment electron density associated in the molecule A with the fragment X ($\rho^X_A(r)$). There are several possibilities to define the fragment densities (Gironés et al., 2003) but in this study we used a simple approach in which the fragment electron density is obtained from the electron density function of the whole molecule (Eq. (5)),

$$\rho_A(r) = \sum_\mu \sum_\nu D_{\mu\nu}\chi_\mu(r)\chi_\nu(r) \quad (5)$$

by appropriately restricting the summation over the basis functions (Eq. (6)):

$$\rho^X_A(r) = \sum_\mu \sum_{\nu\in X} D_{\mu\nu}\chi_\mu(r)\chi_\nu(r) \quad (6)$$

2.3. Theoretical QSAR models

The basic idea underlying the construction of theoretical QSAR models for the correlation of biological activities straightforwardly arises directly from the traditional approach by Hansch and Fujita (1964), in which the experimental data are being described by multi-linear correlation equations using various empirical descriptors like σ-Hammett constant, $\log P$, and others, as represented in Eq. (7):

$$\text{activity} = f(\log P, \ldots, \sigma, E_S) \quad (7)$$

In our approach, this traditional QSAR protocol is modified in such a way that the original empirical parameters used as descriptors in QSAR equations are replaced by the corresponding theoretical counterparts based on appropriate similarity and/or self-similarity measures. Thus, for example, the self-similarity measures of the whole molecule have been shown to act as good theoretical descriptors for molecular hydrophobicity, empirically measured by $\log P$ (Amat et al., 1998). Similarly, the quantum self-similarity measures associated with the appropriate molecular fragments can advantageously be used as universal theoretical descriptors of the substituent effect (Amat et al., 1999; Ponec et al., 1999a,b). Based on the above parallel, the original empirical QSAR model (Eq. (7)) is transformed into alternative theoretical QSAR models (Eqs. (8a) and (8b)), which will systematically be used in this study.

$$\text{activity} = a_0 Z_{\text{AA}}^{(\text{full})} + \sum_{\text{R=fragments}} \alpha_{\text{R}} Z_{\text{AA}}^{\text{R}} + b \qquad (8a)$$

$$\text{activity} = \sum_{\text{R=fragments}} \alpha_{\text{R}} Z_{\text{AA}}^{\text{R}} + b \qquad (8b)$$

The above methodology has been applied to the study of the biological activity of the six molecular sets of antituberculotics, with the aim to rationalize the observed increase of antimycobacterial activity induced by the replacement of the oxo group in 3-aryl-2*H*-1,3-benzoxazine-2,4(3*H*)-diones by thioxo.

2.4. Molecular sets

- A set of derivatives of 6,8-dichloro-3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dione (m001–m009), extracted from Waisser et al. (1998).
- A group of derivatives of 6,8-dibromo-3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dione (n001–n008), collected from Waisser et al. (1998).
- A collection of derivatives of 3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dione (s001–s011), gathered from Waisser et al. (1999).
- A series of derivatives of 3-phenylquinazoline-2,4(1*H*,3*H*)-dione (u001–u008), taken from Waisser et al. (1999).
- A set of derivatives of 6,8-dichloro-3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dithione (y001–y005), extracted from Waisser et al. (2001b).
- A number of derivatives of 3-phenylquinazoline-2,4(1*H*,3*H*)-dithione (z001–z005), taken from Waisser et al. (2001b).

The corresponding parent structures and the number of molecules belonging to each set are summarized in Table 1.

2.5. Experimental activity

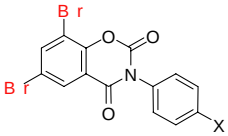
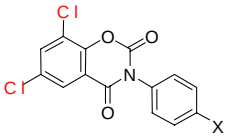
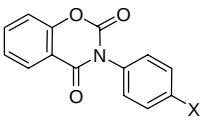
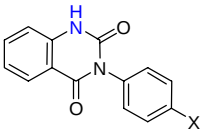
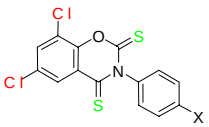
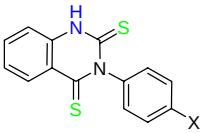
Antimycobacterial activities of the previous presented sets have been evaluated in vitro by Waisser et al. (1998, 1999, 2001b), using different standard strains of mycobacteria, obtained from the Czech National Collection of Type Cultures (CNCTC), of the Institute of Public Health of Prague. In this survey, only a single strain reported in the different studies in the same conditions has been considered: that is, *M. tuberculosis* CNCTC My 331/88. The activities of the compounds against this strain were measured in the Šula semisynthetic medium (SEVAC, Prague) and the compounds were added in different concentrations to the medium in dimethyl sulfoxide solutions. The minimum inhibitory concentration (MICs), i.e. the lowest concentration of a substance at which the inhibition of the growth occurs, was determined in all cases at 37 °C after incubation for 14 days and measured in μmol/l. For further information, additional details can be obtained from Waisser et al. (1999).

The logarithm of the activity values for the tuberculous strain, taken in order to have a narrower range of values, as well as the position and distinctive substitution for each compound are shown in Table 2. Those activities marked with an asterisk have not been precisely measured (Waisser et al., 1998, 1999, 2001b).

2.6. Definition of fragments

The main advantage of the above presented similarity approach to the construction of QSAR models is that it can be used even in the situation where, like in the present study, the fragment responsible for the observed activity is not known beforehand. In this case, namely, the systematic scrutiny of theoretical QSAR models based on similarity measures associated with different molecular fragments allows for detection and localization of the fragment most likely to be responsible for the observed activity. Here, in the case of

Table 1
Types of studied structures

Set	Number of molecules	Derivatives	Reference	Parent structure
M	9	6,8-Dichloro-3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dione	Waisser et al. (1998)	
N	8	6,8-Dibromo-3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dione	Waisser et al. (1998)	
S	11	3-Phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dione	Waisser et al. (1999)	
U	8	3-Phenylquinazoline-2,4(1*H*,3*H*)-dione	Waisser et al. (1999)	
Y	5	6,8-Dichloro-3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dithione	Waisser et al. (2001b)	
Z	5	3-Phenylquinazoline-2,4(1*H*,3*H*)-dithione	Waisser et al. (2001b)	

the series M–Z, the fragments considered as potential pharmacophores are summarized in the Table 3, and the numbering of atoms clarifying the definition of fragments is evident from the Fig. 1.

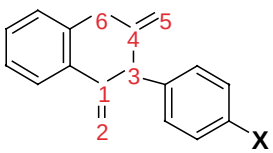


Fig. 1. Numbering of atoms for the definition of molecular fragments considered as potential pharmacophores in the studied series of compounds.

3. Results and discussion

3.1. Molecular modeling and geometry optimization

The calculations in this study were of two types. In the first one, the structures of all the studied molecules have been optimized at semi-empirical AM1 level of theory using the *PC Spartan Pro* software package (PC Spartan Pro, 1997). In the next step, the electron density functions for all the above determined structures have been processed by our own program MOLSIMIL (Amat et al., 1997), which used the densities as an input for the calculation of the similarity and self-

Table 2
Antituberculotic activity (My 3331/88) for the whole set

Compound	Substitution	Log(*M. tuberculosis*)
M01	4H	18
M02	$4CH_3$	1.5
M03	4Br	1.5
M04	$4OCH_3$	1.8
M05	4Cl	1.2*
M06	3Cl, 4Cl	0.9
M07	3Cl	0.9
M08	$3NO_2$	1.2
M09	$4N(CH_3)_2$	2.1*
N01	4H	1.8
N02	$4CH_3$	1.2
N03	4Br	1.2
N04	$4OCH_3$	2.1
N05	4Cl	12
N06	3Cl, 4Cl	0.9
N07	3Cl	0.9
N08	$4N(CH_3)_2$	1.8*
S01	4H	2.097
S02	$4CH_3$	1.792
S03	4Br	1.491
S04	$4OCH_3$	1.792
S05	4Cl	1.204
S06	3Cl, 4Cl	0.903
S07	3Cl	1.491
S08	4F	2.097
S09	$3NO_2$	1.204
S10	$4NO_2$	1.204
S11	$4N(CH_3)_2$	2.097*
U01	4H	2.699
U02	$4CH_3$	2.699
U03	4Br	2.398
U04	$4OCH_3$	2.097*
U05	3Cl, 4Cl	1.792*
U06	4F	2.097*
U07	$3NO_2$	2.398
U08	$3CH_3$, $4CH_3$	2.398
Y01	4H	0.903
Y02	4Br	1.204
Y03	$4CH_3$	0.602
Y04	4Cl	0.903
Y05	3Cl	0.602
Z01	4H	1.505
Z02	4Br	1.204
Z03	$4CH_3$	1.204
Z04	4Cl	1.505
Z05	3Cl, 4Cl	1.204

similarity measures considered in the construction of theoretical QSAR models.

3.2. QSAR models

In order to rationalize the observed biological activity in the studied series of compounds, the broad class of all possible single, two and three-parameter multilinear QSAR models (Eqs. (8a) and (8b)) based on previously defined molecular fragments was scrutinized. The set of all the considered QSAR models is summarized in the Table 4, where fr*i* represents the *i*th fragment and full stands for the similarity measure corresponding to the whole molecule.

The statistical importance of the aforementioned models was evaluated using a recently proposed analytical criterion (Pecka and Ponec, 2000), based on the calculation of the probability that QSAR models with the same correlation coefficient R as the actually observed one are obtained accidentally. This probability is given by the Eq. (9), where N and M denote the number of data points and the number of parameters, respectively.

$$P = \frac{\int_0^{\arccos(R)} \cos^{N-1}\theta \sin^{N-M-2}\theta \, d\theta}{\int_0^{\pi/2} \cos^{N-1}\theta \sin^{N-M-2}\theta \, d\theta} \quad (9)$$

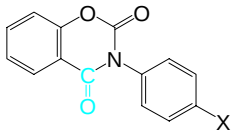
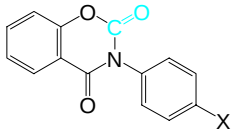
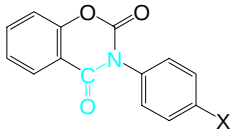
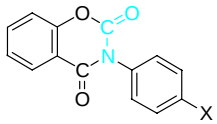
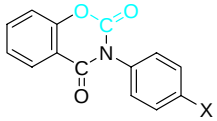
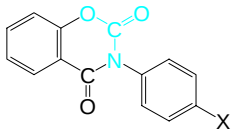
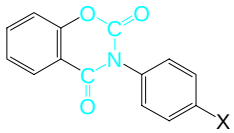
These probabilities are directly related to the so-called confidence level of a correlation, CL, defined in Eq. (10):

$$\text{Percent CL} = 100(1 - P) \quad (10)$$

On the basis of the above criterion it was possible to select a set of QSAR models that provide the "best" correlation of the observed data and the results of this search for the individual sets are summarized in the Table 5. The most important conclusion deduced from this table is that, from the whole set of analyzed QSAR models (Table 4), the "best" or the most statistically important correlations were in almost all cases obtained using single-parameter equations only.

In addition to the above primary results, there are also some other non-trivial deductions that can be extracted from the Table 5. Thus, for example, a more careful inspection reveals that descriptors the most often repeating in "successful" QSAR models are always associated with the fragments fr1 and fr5. This result is very interesting since these fragments just in-

Table 3
Definition of considered molecular fragments

Fragment	Fragment number	Number of atoms	Fragment
$C^1{=}O^2$	1	2	
$C^4{=}O^5$	2	2	
$N^3{-}C^1{=}O^2$	3	3	
$N^3{-}C^4{=}O^5$	4	3	
$O^6{-}C^4{=}O^5$	5	3	
$N^3{-}C^4{=}O^5{-}(O^6)$	6	4	
$C^1{=}O^2{-}N^3{-}C^4{=}O^5{-}(O^6)$	7	6	

O^2, O^5 and O^6 can be correspondingly replaced by S^2, S^5 and N^6.

Table 4
Full set of QSAR models considered

One-parameter model	Two-parameter model	Two-parameter model	Three-parameter model
Full			
fr1	fr1; full	fr1; fr2	fr1; fr2; full
fr2	fr2; full	fr1; fr4	fr1; fr4; full
fr3	fr3; full	fr1; fr5	fr1; fr5; full
fr4	fr4; full	fr1; fr6	fr1; fr6; full
fr5	fr5; full	fr2; fr3	fr2; fr3; full
fr6	fr6; full	fr3; fr5	fr3; fr5; full
fr7	fr7; full		

Table 5
QSAR equations and statistical significance of "best" QSAR models for individual sets of molecules

Set	N[a]	Fragment[b]	r^2	r^2 (cv)	P[c]	Percent CL[d]	Slope	Intercept
M	9	fr1	0.755	0.720	0.0023	99.77	−1069.578	119918.184
N	8	fr5	0.511	0.430	0.0463	95.37	−333.623	64292.314
N	8	fr1	0.510[e]	0.428	0.0467	95.33	−948.510	106344.478
S	11	fr5	0.558	0.509	0.0083	99.17	−232.364	44779.629
S	11	fr1	0.553[e]	0.504	0.0087	99.13	−732.153	82087.273
U	8	fr4	0.150	0.008	0.3438	65.62	260.272	−42728.963
Y	5	Full	0.634	0.512	0.1070	89.30	0.000	0.547
Z	5	fr3	0.277	0.036	0.362	63.80	−662.325	588694.097

[a] Number of molecules in the set.
[b] Label of the fragment.
[c] Previously defined probability.
[d] Confidence Level.
[e] Second best correlation model.

volve oxo and thioxo groups, whose role in determining the antimycobacterial activity of the studied series of molecules was the main goal of the scrutiny in this study.

To illustrate the previous comments, for example, for the series S, according to the literature (Waisser et al., 1999), the traditional approach yielded satisfactory correlations for the observed activity using two-parameter correlation equations (Eqs. (11) and (12)), while comparable or even better accuracy could be obtained using theoretical QSAR models based on single-parameter equations employing fragment similarity measures associated with fragments fr1 or fr5;

$$\log(\mathrm{MIC})^{-1} = a\sigma + b\pi + c$$
$$\log(\mathrm{MIC})^{-1} = 0.826\sigma + 0.273\pi - 1.873$$
$$r = 0.865, \quad s = 0.234, \quad F = 10.38, \quad n = 10(11),$$
$$P = 0.008, \quad \text{percent CL} = 99.20 \qquad (11)$$

$$\log(\mathrm{MIC})^{-1} = a\sigma + b\log P + c$$
$$\log(\mathrm{MIC})^{-1} = 0.826\sigma + 0.313\log P - 3.084$$
$$r = 0.844, \quad s = 0.25, \quad F = 8.64, \quad n = 10,$$
$$P = 0.013, \quad \text{percent CL} = 98.72 \qquad (12)$$

The sets U, Y and Z exhibit a slightly more complex situation. For these cases, the "best" QSAR models were obtained using the fragment similarity measure associated to other fragments but fr1 and fr5. However, as it is possible to notice in the same table the statistical importance of these "best" correlations is very low so that it is difficult to speak of reasonable correlation. This is partly due to the relatively small number of molecules in sets Z and Y, but the absence of a realistic correlation can be also attributed to the fact that the antimycobacterial activity of some of the studied molecules, especially of the derivatives of the substituted 3-phenyl quinazoline is so low that they are practically inactive (Waisser et al., 1999).

The universality and flexibility of the above similarity approach is demonstrated by the formulation of analogous theoretical QSAR models for wider data sets formed by joining several series of compounds. The results of such generalized models are summarized in Table 6.

In this connection it is fair to say that empirical QSAR models were also reported for some of the "extended" data sets (Waisser et al., 1998, 2001b). The fact that none of them is better than our theoretical models is particularly relevant. Thus, for example, the following empirical QSAR correlations were reported for the joined series M and N (Waisser et al., 1998), but in contrast to the three-parameter empirical model (Eq. (13)),

$$\log(\mathrm{MIC}) = a\sigma + b\pi + cI + d$$
$$\log(\mathrm{MIC}) = -0.675\sigma - 0.293\pi - 0.057I + 1.680$$
$$r = 0.856, \quad s = 0.234, \quad F = 9.14, \quad n = 14,$$
$$P = 0.003, \quad \text{percent CL} = 99.68 \qquad (13)$$

the proposed theoretical approach leads to a comparable statistical significance with a single-parameter Eq. (14), using the self-similarity measure associated

Table 6
Summary of "best" QSAR models for the data sets formed by joining several series of molecules

Set	N[a]	Fragment[b]	r^2	r^2 (cv)	P[c]	Percent CL[d]	Slope	Intercept
M/N	17	fr1	0.622	0.597	1.692E-04	99.98	−1007.783	112989.913
	14[e]	fr1	0.531	0.491	3.131E-03	99.69	−1194.523	133926.466
M/N/S	28	fr1	0.524	0.506	1.338E-05	100.00	−684.817	76780.271
	24[e]	fr1	0.422	0.395	5.967E-04	99.94	−651.294	73021.824
S/U	19	fr1	0.618	0.595	6,609E-05	99.99	−226.356	25379.508
	15[e]	fr1	0.768	0.750	1.834E-05	100.00	−288.370	32332.312
U/Z	13[f]	fr1	0.795	0.776	4.300E-05	100.00	−0.0014	2.4766
	13[f]	fr5	0.795	0.776	4.300E-05	100.00	−0.0014	2.5483
	10[e,f]	fr1	0.943	0.935	3.067E-06	100.00	−0.0016	2.7031
	10[e,f]	fr5	0.943	0.935	3.067E-06	100.00	−0.0016	2.7889
Y/Z	10	fr1	0.613	0.564	7.396E-03	99.26	−77.709	65023.499
		fr3					164.868	
S/Z	16	fr5	0.566	0.499	4.405E-03	99.56	−171.624	6004.685
M/N/S/U/Y/Z	39[e]	fr1	0.593	0.569	9.900E-08	100.00	0.0292	4.0303
		fr5					−0.0303	

[a] Number of molecules in the set.
[b] Label of the fragment.
[c] Previously defined probability.
[d] Confidence level.
[e] Compounds with not exactly measured activities omitted.
[f] Comparable models could be obtained using practically any of the fragments.

with the fragment fr1 as the corresponding descriptor.

$$\log(\mathrm{MIC})^{-1} = -1194.523 Z_{\mathrm{AA}}^{\mathrm{C^1{=}O^2}} + 133926.466$$
$$n = 14, \quad r = 0.728, \quad P = 0.003,$$
$$\text{percent CL} = 99.69 \qquad (14)$$

Similarly, the reported QSAR model for the joint series Y and Z (Waisser et al., 2001b) required the use of seven indicator variables, while only a single-parameter Eq. (15) based on the self-similarity measures associated with the fragment fr1 is again needed to obtain the same quality correlation using the similarity approach.

$$\log(\mathrm{MIC})^{-1} = -77.709 Z_{\mathrm{AA}}^{\mathrm{C^1{=}O^2}} + 65023.499$$
$$n = 10, \quad r = 0.783, \quad P = 0.007,$$
$$\text{percent CL} = 99.26 \qquad (15)$$

Another example of the broad applicability of similarity approach concerns the joint set of molecules involving the series S and U for which a single-parameter theoretical QSAR model (Eq. (16)) could again be formulated although no analogous QSAR model based on traditional approach was reported so far.

$$\log(\mathrm{MIC})^{-1} = -226.356 Z_{\mathrm{AA}}^{\mathrm{C_1{=}O_2}} + 25379.508$$
$$n = 19, \quad r = 0.786, \quad P = 0.000066,$$
$$\text{percent CL} = 99.99 \qquad (16)$$

This joined set is interesting because it combines the series U that involves the set of the practically inactive 3-phenylquinazolines with the set of active benzoxazines. The reasons for this successful correlation are clearly visible from the Fig. 2. While the scatter of the data within the inactive series U is relatively significant so that no reasonable correlation with fr1 exists, the inclusion of the active series S improves the situation dramatically. This is due to the big difference in the activity of the molecules in sets S and U, which dominates the successful correlation of the combined sets S/U.

A similar advantage of the theoretical approach is also evident from the fact that comparable theoretical QSAR models can be formulated for the more extensive series of molecules involving all the studied sets,

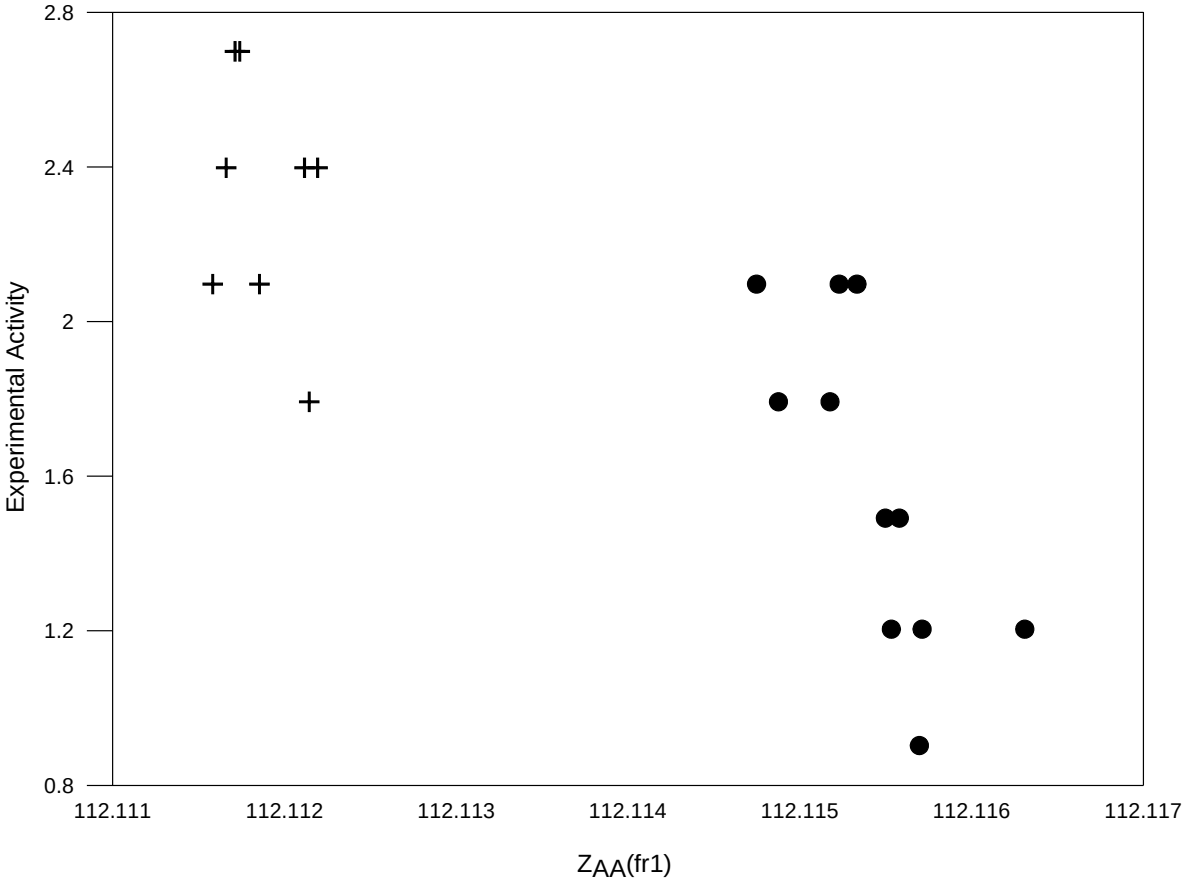


Fig. 2. Experimental activity vs. self-similarity measures for fragment 1, for the joint set S/U, where the circles symbolize the compounds of set S, and the crosses the ones of set U.

for which no analogous empirical equation has been reported using the traditional approach.

Thus, for example, the following two-parameter QSAR model (Eq. (17)) was found for the extensive set of molecules involving the joint series M, N, S, U, Y and Z.

$$\log(\mathrm{MIC})^{-1} = 0.0292 Z_{\mathrm{AA}}^{\mathrm{C_1{=}O_2}} - 0.0303 Z_{\mathrm{AA}}^{\mathrm{O_6{-}C_4{=}O_5}} + 4.0303$$
$$n = 39, \quad r = 0.770, \quad P = 0.00000009, \quad \text{percent CL} = 100 \qquad (17)$$

Summarizing the above results, it is possible to conclude that the reported similarity approach seem to be in complete harmony with previous experimental studies. Thus, for example, the replacement of the oxo group in the position 2 by the corresponding thioxo group was reported as the main factor responsible for the observed increase of antimycobacterial activity (Waisser et al., 2001b); hence, it is interesting that it is just the fragment involving these groups (fr1), which was found to yield the best theoretical QSAR correlations with the experimental activity. A similar albeit weaker activating effect was also reported to accompany the replacement of the oxo group by thioxo in the position 4 and the importance of this particular group (fr5) is again clearly revealed by the fact that the corresponding similarity measure was detected as the second most successful molecular descriptor in the reported theoretical QSAR models. On the other hand, the replacement of O by NH in the series of benzox-

azines and quinazolines is clearly accompanied by the drop of the activity which is also well reproduced by the calculated descriptors.

It can thus be concluded that the reported results clearly confirm the advantageous usefulness of fragment self-similarity measures as new efficient and general molecular descriptors. Hence, the application in the design of theoretical QSAR models and, specially, the rationalization of the effect of the systematic structural variation of substituents on the observed biological activity in the series of structurally related molecules has been widely proved.

Acknowledgements

This work was a part of the research project of MSM 111600001 of the Czech Ministry of Education and was supported by the Grant 234/2000/BCH of the Grant Agency of Charles University in Prague.

References

Amat, L., Constans, P., Besalú, E., Carbó-Dorca, R., 1997. MOLSIMIL 97. Institute of Computational Chemistry, University of Girona, Spain.

Amat, L., Carbó-Dorca, R., Ponec, R., 1998. Molecular quantum similarity measures as an alternative to log *P* values in QSAR studies. J. Comp. Chem. 19, 1575–1583.

Amat, L., Carbó-Dorca, R., Ponec, R., 1999. Simple linear QSAR models based on quantum similarity measures. J. Med. Chem. 42, 5169–5180.

Bermudez, L.E., Young, L.S., 1995. New drug for therapy of mycobacterial infection. Cur. Opin. Infec. Dis. 8, 428–437.

Carbó, R., Leyda, L., Arnau, M., 1980. How similar is a molecule to another? An electron density measure of similarity between two molecular structures. Int. J. Quantum Chem. 17, 1185–1189.

Gironés, X., Ponec, R., Carbó-Dorca, R., 2003. In: Sen, K. (Ed.), Nova Press.

Hansch, C., Fujita, T., 1964. ρ-σ-π Analysis. A method for the correlation of biological activity and chemical structure. J. Am. Chem. Soc. 86, 1616–1626.

PC Spartan Pro, version 1.4, 1997. Wavefunction Inc., Irvine.

Pecka, J., Ponec, R., 2000. Simple analytical method for evaluation of statistical importance of correlations in QSAR studies. J. Math. Chem. 27, 13–22.

Ponec, R., Amat, L., Carbó-Dorca, R., 1999a. Molecular basis of quantitative structure-properties relationship (QSPR): a quantum similarity approach. J. Comput. Aid. Mol. Des. 13, 259–270.

Ponec, R., Amat, L., Carbó-Dorca, R., 1999b. Quantum similarity approach to LFER: substituent and solvent effects on the acidities of carboxylic acids. J. Phys. Org. Chem. 12, 447–454.

Rouhi, A.M., 1999. Tuberculosis. A tough adversary. Chem. Ind. News 77, 52–69.

Waisser, K., Hladůvkova, J., Gregor, J., Rada, T., Kubicová, L., Klimešová, V., Kaustová, J., 1998. Relationships between the chemical structure of antimycobacterial substances and their activity against atypical strains. Part 14: 3-aryl-6,8-dihalogeno-2*H*-1,3-benzoxazine-2,4(3*H*)-diones. J. Arch. Pharm. Pharm. Med. Chem. 331, 3–6.

Waisser, K., Macháček, M., Dostál, H., Gregor, J., Kubicová, L., Klimešová, V., Kuneš, K., Palát, K., Hladůvkova, J., Kaustová, J., Möllmann, U., 1999. Relationships between the chemical structure of substances and their antimycobacterial activity against atypical strains. Part 18: 3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-diones and isosteric 3-phenylquinazoline-2,4(1*H*, 3*H*)-diones. Collect. Czech. Chem. Commun. 64, 1902–1924.

Waisser, K., Gregor, J., Kubicová, L., Klimešová, V., Kuneš, J., Macháček, M., Kaustová, J., 2000. New groups of antimycobacterial agents: 6-chloro-3-phenyl-4-thioxo-2*H*-1,3-benzoxazine-2(3*H*)-ones and 6-chloro-3-phenyl-2*H*-1,3-benzoxazine-2,4(3*H*)-dithiones. Eur. J. Med. Chem. 35, 733–741.

Waisser, K., Hladůvková, J., Holý, P., Macháček, M., Karajannis, P., Kubicová, L., Klimešová, V., Kuneš, J., Kaustová, J., 2001a. 2*H*-1,3-Benzoxazine-2,4(3*H*)-diones substituted in position 6 as antimycobacterial agents. Chem. Pap. 55, 323–334.

Waisser, K., Gregor, J., Dostál, H., Kuneš, J., Kubicová, L., Klimešová, V., Kaustová, J., 2001b. Influence of the replacement of the oxo function with the thioxo group on the antimycobacterial activity of 3-aryl-6,8-dichloro-2*H*-1,3-benzoxazine-2,4(3*H*)-diones and 3-arylquinazoline-2,4(1*H*,3*H*)-diones. Il Farmaco 56, 803–807.