Available online at www.sciencedirect.com
SCIENCE DIRECT®
Thermochimica Acta 417 (2004) 283–294
thermochimica acta
www.elsevier.com/locate/tca

Calorimetry and structure–activity relationships for a series of antimicrobial hydrazides

M.L.C. Montanari [a], A.D. Andricopulo [b], C.A. Montanari [c,*]

[a] *Departamento de Química, Universidade Federal de São Carlos, Rodovia, Washington Luiz, Km 235, P.O. Box 676, 13565-905, Brazil*
[b] *Centro de Biotecnologia Molecular Estrutural, CEPID/FAPESP, Instituto de Física de São Carlos, Universidade de São Paulo-USP, Av. Trabalhador São-carlense, 400, P.O. Box 369, 13560-970 São Carlos, SP, Brazil*
[c] *Núcleo de Estudos em Química Medicinal—NEQUIM, Departamento de Química, Universidade Federal de Minas Gerais, Campus da Pampulha, 31270-901 Belo Horizonte, MG, Brazil*

Received 14 June 2003; received in revised form 17 July 2003; accepted 18 July 2003

Available online 27 February 2004

Abstract

This paper presents some recent developments on the use of quantitative structure–activity relationships (QSAR) based on biological calorimetry. The calorimetric biological potency can be measured for structurally related compounds whose activity would not be easily determined with less accurate and precise methods. A series of antimicrobial hydrazides was assayed against two different cultured cell systems, *Escherichia coli* and *Saccharomyces cerevisiae*. The direct demonstration of a similar mode of action for the two biological systems was achieved with the use of calorimetry. The measured values were described in terms of 3D molecular interaction fields (MIF) by means of a recently developed GRID independent method (GRIND). The aim of this approach is to allow the analysis of a large number of quantitative descriptors by using chemometric tools such as partial least squares (PLS). The correlation between chemical structures and changes in bioactivity is described without the need for 3D molecular alignment according to a suitable conformational bioactive template. The proposed model for these molecular interaction fields has revealed the importance of the stereo-electronic properties on the cells metabolism. Throughout this paper, we describe the usefulness of the same cell systems in disclosing partitioning behaviour of study hydrazide antimicrobials employing the diffusion technique of Taylor–Aris. Since this variable may be of utility in pharmacokinetic studies, we have modelled and predicted it based on computed MIF and multivariate statistics by a procedure called GRID/VolSurf. This result was achieved with a small number of VolSurf descriptors encoding a balanced range of hydrophilic–lipophilic properties.

Keywords: Biological calorimetry; Taylor–Aris partitioning; QSAR; QSPR

1. Introduction

It has been written that "science moves forward according to what it can measure", and at present, there appear to be numerous promising advances among several analytical techniques that can be useful to describe drug–receptor interactions [1]. Calorimetric techniques are very useful in the field of medicinal chemistry for studying these interactions on very small quantities of biological molecules [2]. The robustness and sensitivity of thermal analysis methods [3] with automation, and the availability of reliable software tools are especially useful for the behavioural study of bioactive substances, excipients [4,5], and delivery systems [6]. Calorimetry is suitable for the investigation of the effect of drugs on microorganisms and animal cellular systems, and thus has been used as a method for the determination of bioactivity [7,8].

A number of previous studies have been performed associated with the applicability to derive structure–activity relationships (SAR), which can in turn help medicinal chemists gain insight about the key interactions between drug and its receptor, with the aim of producing new, more powerful antimicrobials. On the other hand, studies seldom showed the use of calorimetry in deriving quantitative structure–activity relationships (QSAR), a field where it is possible not only to set information on SAR, but also insight into modes of action can be envisaged [9]. We have already shown that QSAR based on biological calorimetry for a set

* Corresponding author. Tel.: +55-31-3499-5728; fax: +55-31-3499-5700.
E-mail address: montana@dedalus.lcc.ufmg.br (C.A. Montanari).

0040-6031/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.tca.2003.07.024

of antimicrobial hydrazides acting against *Saccharomyces cerevisiae* and *Escherichia coli* resulted in extrathermodynamic relationships (involve the correlation between variables which, from a strictly thermodynamic standpoint, are not related) between calorimetrically measured biopotencies and partitioning using the same cell systems [9,10]. To the best of our knowledge, the result reported was the first-ever demonstration of an extrathermodynamic relationship between calorimetric drug potencies (calorimetrically based dose–response curves) and two cell systems to behave in the same way with respect to the importance of partitioning and biopotencies. This meant that the same set of congeneric compounds experienced a similar environment in the two systems [10].

Biological calorimetry can also provide important possibilities for comparative QSAR. In order to gain a deeper understanding of the relationships and the meaning of parameters within the model it is necessary to establish some kind of lateral validation. The validation step is a major bottleneck, and can be accomplished by chemical procedures using physicochemical organic reactions and by means of biological systems [11]. Differences in binding affinities can be experimentally determined from calorimetric measurements. The isothermal calorimetric titration of dimeric 2-amino-1,8-naphthyridine selectively bond to the GG mismatch with 9-mer duplex d(CATCGGATG)$_2$ is enthalpy-controlled and supports the intercalation of both naphthyridine rings into DNA base pairs [12]. Induction of a remarkable conformational change in the human telomeric sequence by the binding of naphthyridine dimer has also been discussed in detail [13]. The relative binding energies between 2,6-disubstituted amidoanthracene-9,10-dione and 2,6-disubstituted acridine chromophores, and the folded human telomere DNA quadruplex were determined using molecular simulation methods. Results were in general agreement with binding enthalpies determined by isothermal calorimetry studies, thus supporting the hypothesis that guanine quadruplexes are the primary target for telomerase inhibitors with extended planar chromophores [14]. Nonetheless, Frisch et al. [15] have shown that the use of enthalpies for the analysis of SAR appears to be complicated by enthalpy–entropy compensation of weak intermolecular interactions, when studying the thermodynamics of the interaction of barnase and barstar by titration calorimetry. The changes in free energy, enthalpy, entropy, and heat capacity upon the binding of turkey ovomucoid third domain OMTKY3 to alpha-chymotrypsin was reported by Filfil and Chalikian [16]. They have observed that water molecules are released to the bulk state upon the binding of OMTKY3 to α-chymotrypsin, which is entropy driven with a large, unfavorable enthalpy contribution, thus highlighting the binding protein–protein pattern recognition process through calorimetric measurements.

In drug design, the study of protein–ligand interactions is one of the most important methods for the understanding of the SAR. The activity of N-2-phenylthioguanine HSV1 TK inhibitors was determined with kinetic measurements of thymidine phosphorylation by Folkers et al. [17]. They measured the interaction energies between thymidine and HSV1 TK via calorimetry, and reported binding to occur spontaneously since it was enthalpy driven. Raffa et al. have used isothermal calorimetry to characterize the reversible inhibition of cytidine 2′-monophosphate (2′-CMP) binding to RNase-A [18]. They have demonstrated that enthalpy-driven interactions may assist the design of small-molecule inhibitors. Calorimetry was also used to measure the change in enthalpy for the scission of calf thymus DNA (ct-DNA) induced by adriamycin® (ADM, doxorubicin). Liang et al. [19] have shown that calorimetry is a useful method to describe both the binding constant and the standard thermodynamic parameters for the binding of ADM, a potent anthracycline antibiotic active against many cancers and has been used in clinical medicine for many years for the scission of calf thymus DNA.

In QSAR studies and computer-aided molecular design, the information contents of descriptors increase in the following sequence: element-level, structural formulae, electronic structure, molecular shape, intermolecular, and interaction descriptors. Every subsequent class of descriptors normally covers information contained in the previous-level. It is, however, practically impossible to cover all the features of a molecular structure in terms of any single class of descriptors. Therefore, it is recommended to optimize the number of descriptors used by means of appropriate statistical procedures and characteristics of structure–property models based on these descriptors [20]. In the search for new methods of obtaining partition coefficients, we have modified the Taylor–Aris diffusion technique to allow the direct measurement of log *P* (octanol–water partition coefficient) within the cell system itself. This technique was successfully applied to two different cell systems, and an extrathermodynamic equation between log *P* for *E. coli* and *S. cerevisiae* seemed to be a promising way of obtaining this parameter directly from them [21]. Prior to the synthesis of a myriad of compounds with described biological activities, it might be interesting to create libraries of compounds with certain defined properties. When biological properties can also be included at least as a non-supervised descriptor, the definition of the chemical and biological space is of prime importance. Thus, this paper is intended to show that biological calorimetry, as a method for determining biopotencies of antimicrobials, can be the base input for obtaining physicochemical descriptors. This has been accomplished by using 3D QSAR methods, namely GRIND/GRID and GRID/VolSurf.

2. Experimental

2.1. Calorimetric measurements

A flow calorimeter (LKB type 10700-1; LKB Produkter AB, S-161 25 Bromma 1, Sweden) fitted with a flow-through

calorimetric vessel (0.5 mL working volume) was used throughout. The thermostatic air bath was maintained at 37 °C in a room kept at 25 ± 0.2 °C. The voltage was amplified with a Keithley 150B microvoltammeter (nominal setting of 3 μV equivalent to 43.2 μV full scale deflection). The power-time curves were recorded on a potentiometric recorder. The glucose buffer solution used for the calorimetric medium was phosphate buffer, pH 7.0 containing 172 g L^{-1}: D-glucose, 1.80; K_2HPO_4, 3.68; KH_2PO_4, 1.32; made to volume with distilled water. 45 mL of glucose buffer was added to 3 ml DMSO in a three-necked vessel, thermostated and stirred at 37 °C in a bath external to the calorimeter. This volume of DMSO was necessitated by the working solubility limits of the higher homologues. A pump rate of 0.78 mL min^{-1} was used (calculated from an independent volume of reaction medium expelled over a time of five minutes). The ampoule of cells was removed from a liquid nitrogen cryostat, thawed in a water bath at 37 °C for three minutes and shaken for one min. Five minutes after commencement of thawing, 0.5 mL of the cell suspension was pipetted into the reaction vessel. One min later, 1.5 mL of DMSO (carrying the appropriate amount of study compound to give the final concentration as required in the reaction vessel or control blank), was added to the mixture, with continuous stirring to promote homogeneity of the reaction mixture. Thus, the total volume of the incubation medium was 50 mL. This medium flowed through the calorimeter and the heat change associated with the metabolism of glucose, under the particular conditions of each assay, was registered. The flow calorimeter was washed out after each incubation with water followed by thermostatted glucose buffer prior to the establishment of each reaction loop. The calorimeter was left overnight in 10% RBS solution—a commercial surfactant.

2.2. The role of calorimetric output in drug design

One of the major problems of quantifying the relationship between chemical structure and activity is that the measurements of biological properties (i.e., IC_{50}, K_i, etc.) have to be accurate and reproducible within the experimental error. In fact, this is critical because many data sets consist of data collected from different studies (several sources), and therefore, there is no simple way of determining the quantitative relationship between the results obtained from different sources. Experiments must be meticulously designed to yield statistically valid data, producing sets of compounds that cover diverse chemical space, mapped onto biological space. QSAR modelling is the most successful method for analyzing experimental data for biological applications. However, it must be emphasized that the use of inappropriate biological activity data will not afford good-quality models. Experimental protocols involve the preparation of cultured cell systems, their frozen maintenance and recovery, were errors and all sorts of problems in incorporating high-level chemical and biological knowledge can possibly occur. There are two possible (out of many) errors that deserve to be noted: the quantitative nature of the measurement and the recognition that all compounds share a similar mode of action. The reliability of experimental data must be ascertained because such errors may not be easily detected after a model is conceived, mainly because they represent the sum of individual measurement errors. Biological calorimetry is efficient (least prone to errors), fast, and reproducible to better than 3%. In vitro screening can be carried out in complex and defined medium using frozen cells. Calorimetric output can reveal both biocide or biostatic properties of compounds from calorimetric output curves, and this is very important in order to control therapeutic drug dosage. Consistence of the data can be readily checked from derived experimental curves. An advantage of the use of biological calorimetry was observed in the antimicrobial hydrazide series (Fig. 1). The compounds were ranked in ascending order of potency according to their structural and chemical characteristics (substitution at the 4-position of the aryl ring in *p*-X-C_6H_4-C(O)NHNHC(O)Ph, Table 1), as follows: C_6H_5 *t*-Bu, C_5H_{11}, Br, OMe, Me, NO_2, H.

The results shown in Fig. 1 suggest that calorimetric biopotencies can be quantitatively measured, and thus seem to be an excellent choice for use in QSAR studies. Fig. 1a and c clearly shows that the interaction of the study hydrazide antimicrobials can be ranked in potency order according to substitution at the 4-position of the aryl ring. The Fig. 1b and d also demonstrates that these interactions resulted in a graded response over a limited range of drug concentrations. Though this is not surprising, it does indicate the concentration time-related response for such small subunit moiety allocations. The antimicrobial endpoints of interest can be readily seen from the calorimetric outputs shown in Fig. 1a and c. Each graded curve identifies the lowest concentration of a single compound at which response for a specific organism is inhibited (zero-order kinetics), thus reflecting its "biostatic activity". Results from this screening revealed the antimicrobial activity of these compounds against the two cell systems. Moreover, it is also feasible to evaluate the inner relationship between the modes of action for the same series of compounds against the two cell systems. Fig. 2 shows that the straight lines are quasi-parallel with slopes of 0.342–0.598 for *S. cerevisiae* and 0.276–0.520 for *E. coli*.

3. Theory and calculations

3.1. The Almond method

Almond [23] is a computational program that was developed for the generation, analysis, and interpretation of GRIND (GRid-INdependent Descriptors), a relatively new generation of alignment independent 3D-molecular descriptors [24]. These find application in 3D QSAR, QSAR, virtual screening, design of combinatorial libraries, binding site and

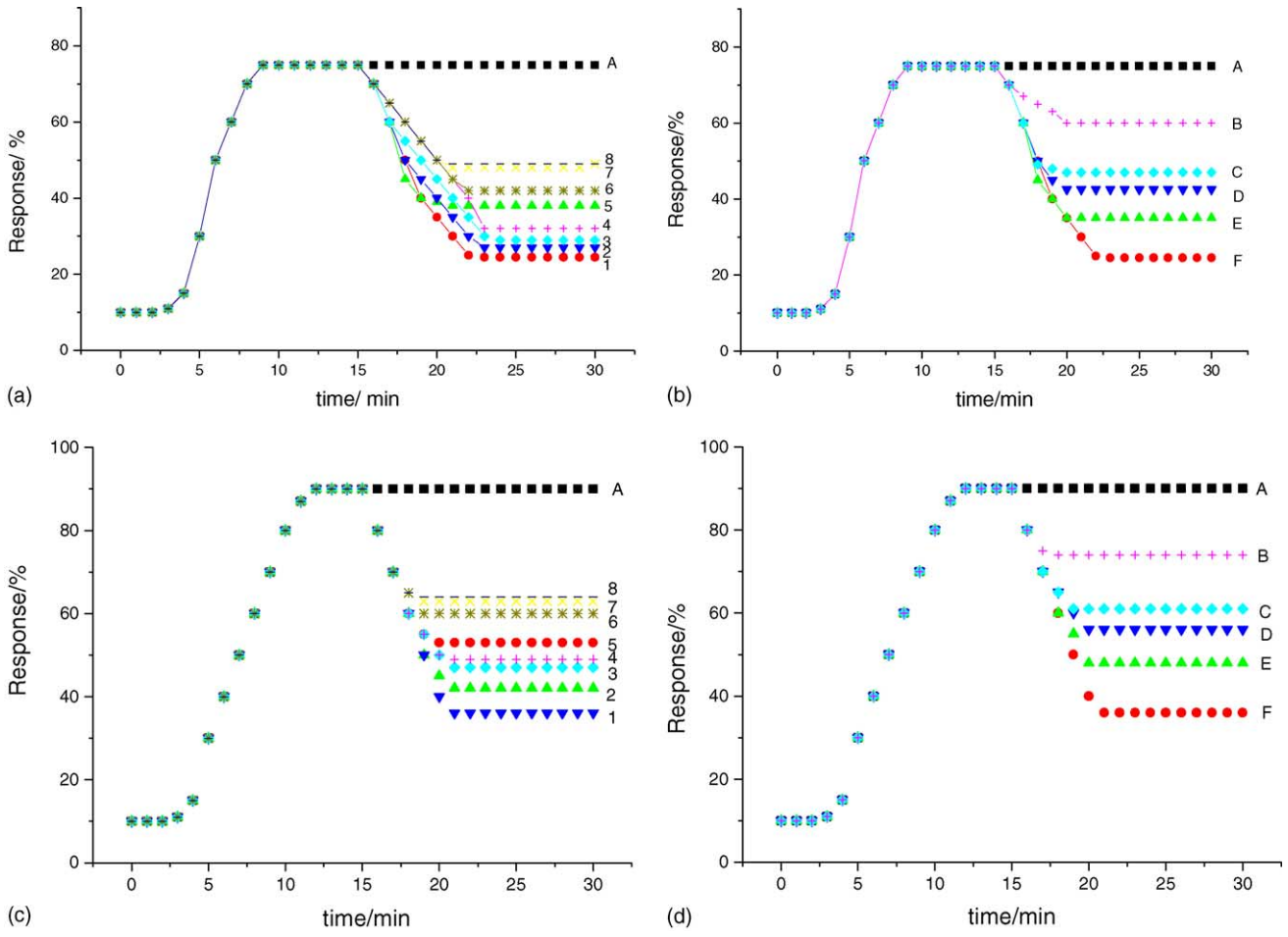
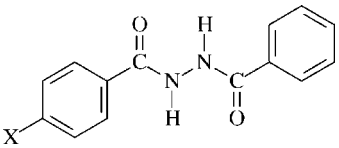


Fig. 1. Calorimetric outputs observed within the series of hydrazides (8) C_6H_5, (7) *t*-Bu, (6) C_5H_{11}, (5) Br, (4) OMe, (3) Me, (2) NO_2, (1) H (a and c, Table 1); compound 1 with A, control; concentration (μmol L^{-1}) B, 1.716; C, 2.017; D, 2.145; E, 2.318; F, 2.619 with *S. cerevisiae* and *E. coli* (b and d). Adapted from [9] and [22]: see text for explanation.

Table 1
Values of calorimetric potencies and Taylor–Aris partitioning of hydrazides against *E. coli* and *S. cerevisiae*[a]



X	log $IC_{50(Ec)}$[b]	log $IC_{50(Sc)}$	log $P_{TA(Ec)}$[c]	log $P_{TA(Sc)}$	$\pi_{TA(Ec)}$[d]	$\pi_{TA(Sc)}$
H	2.280	2.140	−1.094	−0.894	0	0
NO_2	2.38	2.216	−1.140	−0.944	−0.05	−0.05
Me	2.540	2.260	−0.956	−0.782	0.138	0.112
OMe	2.600	2.326	−1.001	−0.833	0.093	0.061
Br	2.730	2.473	−0.879	−0.693	0.215	0.207
C_5H_{11}	3.050	2.620	−0.652	−0.439	0.442	0.455
t-Bu	3.214	2.753	−0.799	−0.516	0.295	0.378
C_6H_5	3.240	2.900	−0.787	−0.532	0.307	0.362

[a] All values have been previously determined [9,10,21].
[b] The concentration of drug required to diminish 50% of cell metabolism measured as the heat flow rate by calorimetry.
[c] log P_{TA} is the Taylor–Aris diffusion technique of partitioning.
[d] Π values calculated from: $\pi = \log P_{TA(X)} - \log P_{TA(H)}$.

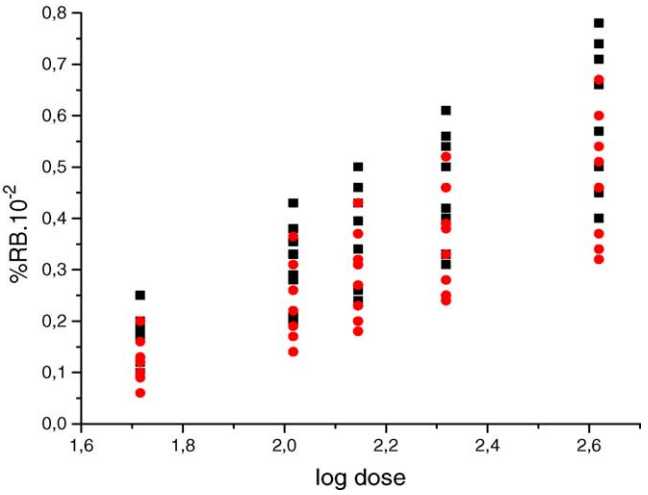


Fig. 2. Dose–response curves for *S. cerevisiae* (■) and *E. coli* (●) (BR means biological response).

selectivity studies, and in a variety of fields that require fast and accurate description for molecular structure. GRIND is insensitive to the position and orientation of the molecular structures in the space. From this perspective, there is a clear advantage of eliminating the need for determination of 3D molecular alignment according to a suitable conformational template (assumed to be the bioactive conformation), one of the most difficult and time consuming steps of the traditional 3D QSAR methods, such as comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) [25]. GRIND is based on molecular interactions fields (MIF) [26], which are generated using the program GRID, and represent the ability of a ligand to interact with a receptor. In ALMOND, the standard procedure involves three GRID [27] probes: the hydrophobic probe (DRY), the carbonyl oxygen (O) and the amide nitrogen (N1). These probes represent strong non-covalent interactions found in the biological targets (e.g., enzymes, receptors). When the three probes are used, the metric describing the series includes two distinct correlograms (auto- and cross-correlograms) of up to six blocks of *X* variables. The three auto-correlograms are DRY–DRY (hydrophobic); O–O (hydrogen bond donor); and N1–N1 (hydrogen bond acceptor); while the three cross-correlograms are DRY–O (hydrophobic and hydrogen bond donor); DRY–N1 (hydrophobic and hydrogen bond acceptor); and O–N1 (hydrogen bond donor and hydrogen bond acceptor). The assembly of these variables, usually less than a hundred per compound, constitutes the standard molecular description that can be used directly for the chemometric analysis. The *X*-matrix generated can be scaled using raw, remove baseline, or normalized block-wise methods. Some of the most common methods within chemometrics are principal component analysis (PCA) and partial least squares (PLS). PCA is most suitable for data overview while PLS for quantitative modeling, analysis and prediction. PCA [28] is a technique used to summarize and describe the information contained in the *X*-matrix. The PCA decomposes the *X*-matrix as the product of two similar matrices. These are the loading matrix (*P*) and the score matrix (*T*), which contains information about the variables and the objects, respectively. PCA is a very useful technique for understanding the distribution and differences of the objects, as well as the variables that contain similar or completely independent information. The PCA information is best represented by the 2D and 3D scores and loading plots of the matrices obtained, which represent the relative position of the objects and the original variables, respectively, in the space. As described above, PCA is a technique of multiple variables that deals only with *X*-variables. PLS, on the other hand, is a regression technique that generalizes and combines features from PCA and multiple regression [29]. It is particularly useful when one needs to predict a set of dependent variables from a large set of independent variables (i.e., predictors). In PLS regression, the *X* block of independent variables (descriptors) is correlated with the *Y* vector (activity) in such a way that the projected coordinates *T* are good predictors of *Y*. In this case, the biological activities are included in the decomposition procedure. The goal of PLS regression is to build a model that will be able to predict *Y* from *X* and to describe their common structure [30]. The best model obtained should be useful to calculate reliable predictions of *Y* values for new molecules, not included in the model. It is important to note that the PLS (not PCA) regression method can deal with the kind *X*-matrices used in 3D QSAR studies. In the present case, the *X*-matrix contains much less objects (molecules) than variables, and the use of multiple linear regression (MLR) is not appropriate. PLS decomposes the *X*-matrix as the product of two smaller matrices, *P* and *T*, in a similar way to that of PCA. The main difference is that PCA obtains PCs that represent at best the structure of the *X*-matrix, while PLS obtains the latent variables (LVs) that represent the structure of the *X* and *Y* matrices, maximizing the fitting between the *X*s and *Y*s [31]. Selecting the correct number of LVs (dimensionality) and testing the predictive power of the model is of critical importance in PLS. The predictive ability of a regression model is evaluated usually through an internal cross-validation (CV). This procedure creates reduced models (where some of the objects are removed), and are used to predict the *Y* variables of the objects held out. The regression model is usually validated through an internal leave-one-out (ideally, leave-two-out or leave-some-out) cross-validation procedure, a useful tool to verify its ability for future predictions. In general, the parameters used to assess the statistical quality of the model are the correlation coefficient r^2 and the cross-validated correlation coefficient q^2 (r^2_{cv}). The CV is a valuable technique because it performs an internal validation of the model and obtains an estimation of the predictive ability without the help of external data sets. As for PCA, the best way to examine the information from PLS is to plot the matrices obtained. The 2D *T*–*U* score plots represent objects in the space of *X*-scores (*T*) against the *Y*-scores (*U*). It gives an idea of the correlation between the

*X*s and the *Y*s obtained in the model for each one of the LVs. The 2D and 3D loading plots represent the original variables in the space of latent variables (*P*), while the 2D and 3D weight plots represent original variables in the space of the weights (*W*), which in turn, represent the coefficients that multiply the *X*s to best fit the *Y*s. Therefore, the loading best represents the first constraint used to build the PLS model (*X*-matrix), while the weights best represents the second constraint used to build the PLS model (the fitting of the *Y*s).

3.2. The GRID/VolSurf method

This powerful computer-automated approach has been used to correlate 3D MIF with physicochemical and pharmacokinetic properties [32–34]. Firstly, it generates MIF by using GRID program [27,35–37], then it treats the fields accordingly by producing descriptors that encode the information content from the chosen water and hydrophobic probes. VolSurf has the advantage of producing descriptors (Table 2) using the 3D information embedded in any map. VolSurf is also alignment independent and conformation insensitive. The VolSurf transformation is fast and its results are easy to interpret. The descriptors have a clear chemical meaning and are lattice-independent. Work reported herein demonstrates the usefulness of the method in describing the partition coefficients obtained from Taylor–Aris diffusion technique [21].

There is a chemical interpretation of VolSurf descriptors, which is outlined here. However, readers are referred to the specialized literature on this subject for a more detailed description [38]. The interaction of molecules with biological membranes is mediated by surface properties. These properties are determined from the size, shape, electrostatics and hydrophobicity obtained from calculations. Size and shape descriptors encode molecular volume, surface, globularity and the ratio volume/surface, and they are explained in Table 2. Descriptors of hydrophilic regions include a molecular envelope that is accessible to and attracts water molecules, and capacity factors that are represented by the hydrophilic surface per total molecular surface unit. Capacity factors are proportional to the concentration of exposed polar groups compared to the total surface area and are often relevant in membrane partitioning in which solvation–desolvation processes are of critical importance. The interaction energy (integy) moments express, like dipole moments, the unbalance between the centre of mass of a molecular and the barycentre of its hydrophilic regions [39]. The integy moment is calculated for both hydrophilic and hydrophobic regions. For the first, they are vectors pointing from the centre of mass to the centre of the hydrophilic regions, whereas for the latter they measure the unbalance between the centre of mass of a molecule and the barycentre of the hydrophobic regions. The high integy moments depicted in this study (Fig. 3) suggest a concentration of hydrated region in one part of the molecule, that is, the *N*,*N*-diacyl hydrazine moiety (–C(O)–NHNH–C(O)–). The hydrophilic–lipophilic balance is the ratio between the hydrophilic and the hydrophobic regions. In the study, molecules possess rather a hydrophilic predominance (Fig. 3). The descriptors of hydrophobic regions are molecular envelopes generating attractive hydrophobic interactions. All calculations were performed on a R10000 O_2 Silicon Graphics workstation.

4. Results and discussion

Table 1 shows the values of both calorimetric potencies and Taylor–Aris partitioning for the hydrazides studied. We have analyzed the possibility of quantifying the relationship between the antimicrobial hydrazides and the cell systems.

Table 2
Description of VolSurf Descriptors [38]

VolSurf number code	Definition
1 V	Volume: total volume (computed at 0.25 kcal mol^{-1})
2 S	Surface: total surface (computed at 0.25 kcal mol^{-1})
3 R	Rugosity: total volume/total surface
4 G	Globularity: surface of the compound divided by the surface of a sphere with the same volume
5–12 W1-W8	Volume of interaction with the H_2O probe at −0.2, −0.5, −1.0, −2. 0, −3.0, −4.0, −5.0, and −6.0 kcal mol^{-1} levels
13–20 IW1-IW8	Integy moment: proportional to the distance between the barycentre of the surface and the volume of interactions with the H_2O probe at the above energy levels
21–28 CW1-CW8	Capacity factor: volume of interaction with the H_2O probe divided by the surface
29–31 Min1-Min3	Energy minima: the first three energy minima interactions
32–34 D12, D13, D23	Distance: the distances between the energy minima
35–42 D1-D8	Volume of interaction with the DRY probe at −0.2, −0.4, −0.6, −0.8, −1.0, −1.2, −1.4, and −1.6 kcal mol^{-1} levels
43–50 ID1-ID8	Integy moment: proportional to the distance between the barycentre of the surface and the volume of interactions with the DRY probe at the different energy levels
51–52HL1, HL2	Balances of the hydrophilic-hydrophobic interactions, measured at −4 and −0.8 kcal mol^{-1}
53 A	Amphiphilic moment
54 CP	Critical packing
55 POL	Molecular polarizability
56 MW	Molecular weight

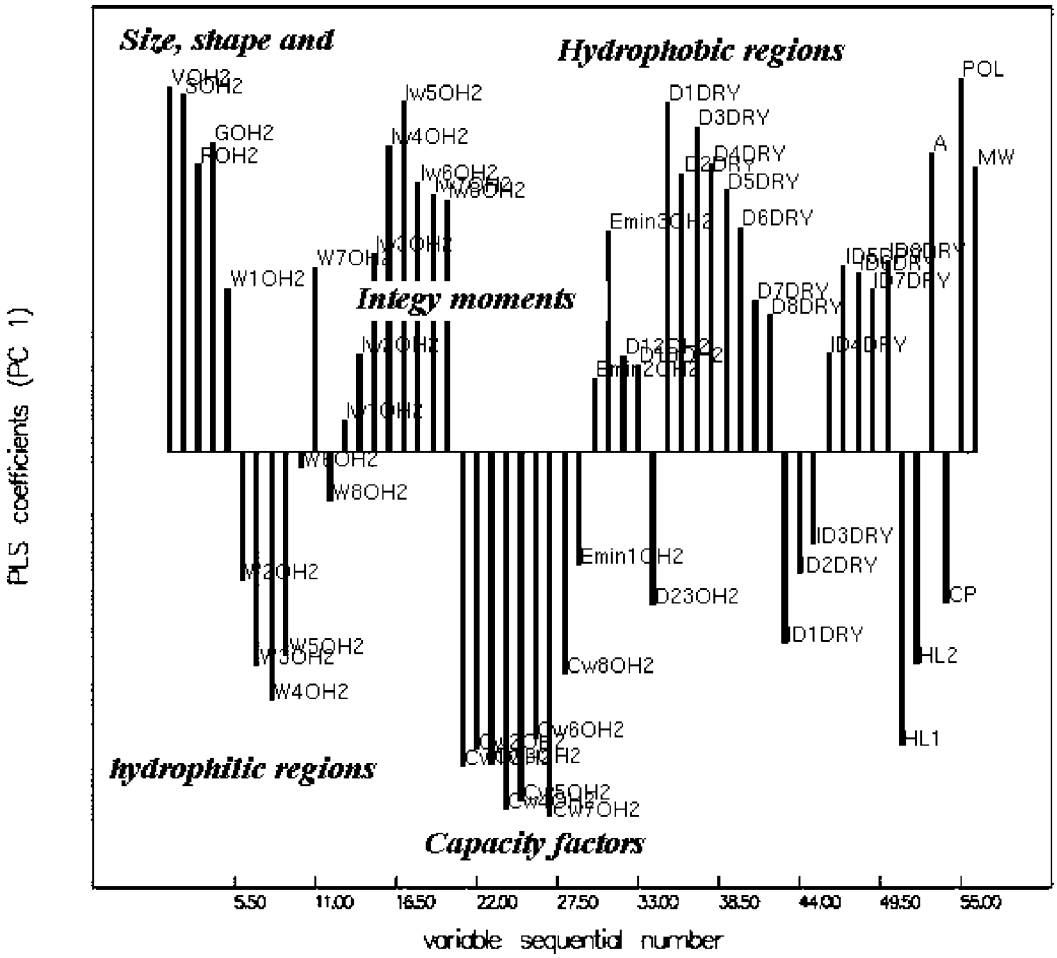


Fig. 3. Log P_{TA} partitioning of hydrazide antimicrobials on *S. cerevisiae*. The PLS coefficients plots show that there must be a hydrophobic–hydrophilic compensation that allows hydrazide antimicrobials to partitioning from solution to within cell system.

Eq. (1) shows that the log $1/IC_{50}$ (*S. cerevisiae*) values [9] are linearly correlated with those of log $1/IC_{50}$ (*E. coli*) [10].

Relationship between *E. coli* and *S. cerevisiae*

$$\log\frac{1}{IC_{50}}(Ec) = 1.368(\pm0.13)\log\frac{1}{IC_{50}}(Sc) - 1.605(\pm0.48)$$
$$(n = 8,\ r^2 = 0.978,\ s = 0.06,\ F = 269.69,\ r^2_{cv} = 0.938) \quad (1)$$

Eq. (1) indicates the existence of an extrathermodynamic relationship between the antimicrobial activity and the two cellular systems. The magnitude of the slope might be a consequence of similar modes of action (Fig. 2), including, for example, transport processes or binding interactions with receptors. Following this, we have tested the hypothesis that these hydrazides possess a similar transport process for both *E. coli* and *S. cerevisiae*, as shown in Eq. (2).

Relationship between log $P_{(TA)Ec}$ and log $P_{(TA)Sc}$

$$\log P_{(TA)E.c.} = 0.833(\pm0.05)\log P_{(TA)S.c.} - 0.323(\pm0.05)$$
$$(n = 8,\ r^2 = 0.955,\ s = 0.038,\ F = 128.40,\ r^2_{cv} = 0.931) \quad (2)$$

Hydrophobic properties do not differ much in the cell systems, since the environment in which transport takes place is similar. Consequently, a similar physicochemical environment should influence the binding modes of these hydrazides to the two cell systems. From the viewpoint of the Hansch–Fujita approach [40], it is well known in QSAR studies that the biological responses elicited by active compounds are correlated with a combination of hydrophobic, steric and electronic properties. However, a number of models do not consider the 3D alignment of the molecular structures. Moreover, chemical descriptors applied to biological calorimetric measurements and Taylor–Aris partitioning are rarely seen or used. One way of dealing with this is by means of ligand pharmacophore generation when calculating alignment-independent 3D molecular descriptors [41], such as those implemented in the Almond (GRIND) method. Therefore, we have used this approach for the description of measured calorimetric biopotencies. This was accomplished by carrying out the 3D generation of three MIF using the hydrophobic (DRY), N1 (amide nitrogen as hydrogen-bond donor) and O (carbonyl oxygen as hydrogen-bond acceptor) GRID probes (for all hydrazide

Table 3
Almond-PLS results of the model developed for the series of antimicrobial hydrazides

Cell system	Cross-validated		Components[c]	Non-cross-validated	
	r^2_{cv} [a]	SDEP[b]		r^{2}[d]	SDEC[e]
S. cerevisiae	0.699[f]	0.024	1	0.862	0.016
E. coli	0.702[f]	0.030	1	0.853	0.210

[a] Cross-validation term for estimated prediction.
[b] SDEP = standard deviation of estimated prediction.
[c] All values computed with the first principal component only.
[d] Goodness-of-fit.
[e] SDEC = standard deviation of estimated calculation.
[f] Up to five random groups left out in order to derive final model.

derivatives). The 3D structures of the hydrazides used in this study were generated with the Sybyl program [42] via a systematic force field based conformational analysis. The most stable conformation was taken from each set and subjected to a full minimization. The optimization process was performed using the Gasteiger-Hückel charges [43]. The study was carried out for the biological responses elicited by hydrazides against *S. cerevisiae* and *E. coli*.

The results in Table 3 show that the MIF analyses explain the antimicrobial activities of the hydrazides, as well as correlate them well (85%), with good predictive power (about 70%), for both *S. cerevisiae* and *E. coli*. The relationships between biopotencies and the three different sets of calculated descriptors were analyzed by PLS projection to latent structures (Fig. 4). The goodness-of-fit shown in Fig. 4 was achieved with no conformational sampling of the hydrazides. Thus, predictions of activity for this series of analogues based on the molecular structure, MIF and biological data, can be put forward to illustrate the proposed approach. The calculated Almond parameters are suitable for describing the physicochemical requirements for this congeneric series in order for them to elicit biological activity according with their 3D proposed pharmacophore models.

The most potent compounds have high positive values of T1. The biopotencies are explained by all three sets of descriptors calculated for the different interaction fields (Fig. 5). The 3-Latent Variables model shown in Fig. 5 is: first block for DRY auto-correlogram, second for O and third for N1. The PLS coefficient histograms for both cell systems show that all hydrophobic (DRY–DRY, DRY–O, DRY–N1) interaction energies correlate negatively with activity, whereas the N1–N1 interaction energies are positively correlated with it. The N1 auto-correlogram has only one peak, which corresponds to the interactions of the polar carbonyl groups (–C(O)–) as hydrogen-acceptors, being of similar importance for both cell systems. This suggests why more polar antimicrobial hydrazides have higher potencies as compared with those possessing lipophilicity moieties.

In drug design and development, properties such as ADME/Tox (absorption, distribution, metabolism, excretion and toxicity) have to be considered at an early step of the process to avoid loss of promising compounds at an advanced stage of the research. Computational approaches based on physicochemical parameters, and characterization of the partitioning behaviour of substances in biphasic systems (e.g., *n*-octanol/water, liposome/buffer, $\Delta \log P$) can be used for modelling transport studies. Standardization of cell cultures to be used for transport studies in the search of in vitro models represents a great potential on passive transport [44]. We have previously reported that partitioning via the modified diffusion technique of Taylor–Aris ($\log P_{TA}$) [21] is an important way of disclosing partitioning behaviour of drugs in the same cell system used for measuring biopotencies and transport properties. However, no physicochemical parameter was ever depicted as a descriptor of $\log P_{TA}$. Also, we envisaged a "hydrophobic" interaction for the antimicrobial hydrazides in order to explore the usefulness of such 3D fields in describing the elicited biological responses. Therefore, the interaction energies for hydrazides were calculated with hydrophobic (DRY) and hydrophilic (H_2O) probes using the GRID program [27,35–37]. The GRID program allows one to sample the potential energy for putative interactions with various probes around a given molecule. The $\log P$ can be described in terms of molecular surface area or volume for instance. Eqs. (3) and (4) established a relationship between $\log P_{TA}$, obtained for the *E. coli* and *S. cerevisiae* cell systems, and the GRID "hydrophobic" interaction energies, then mimicking the role of $\log P$ in 3D QSPR (quantitative structure–property relationships). All calculations used the most negative interaction potential, which represents the best interaction energy between the probe and the hydrazide molecules.

Linear dependence of $\log P_{TA(E.c.)}$ over 3D "hydrophobic" interaction energies (kcal mol^{-1}).

$$\log P_{TA(E.c.)} = -0.638(\pm 0.075)\mathrm{DRY} - 2.069(\pm 0.118)$$
$$(n = 7,\ r^2 = 0.921,\ s = 0.053,\ F = 58.04,\ r^2_{cv} = 0.866) \quad (3)$$

Linear dependence of $\log P_{TA(S.c.)}$ over 3D "hydrophobic" interaction energies (kcal mol^{-1}).

$$\log P_{TA(S.c.)} = -0.724(\pm 0.006)\mathrm{DRY} - 2.029(\pm 0.011)$$
$$(n = 7,\ r^2 = 0.948,\ s = 0.048,\ F = 91.28,\ r^2_{cv} = 0.920) \quad (4)$$

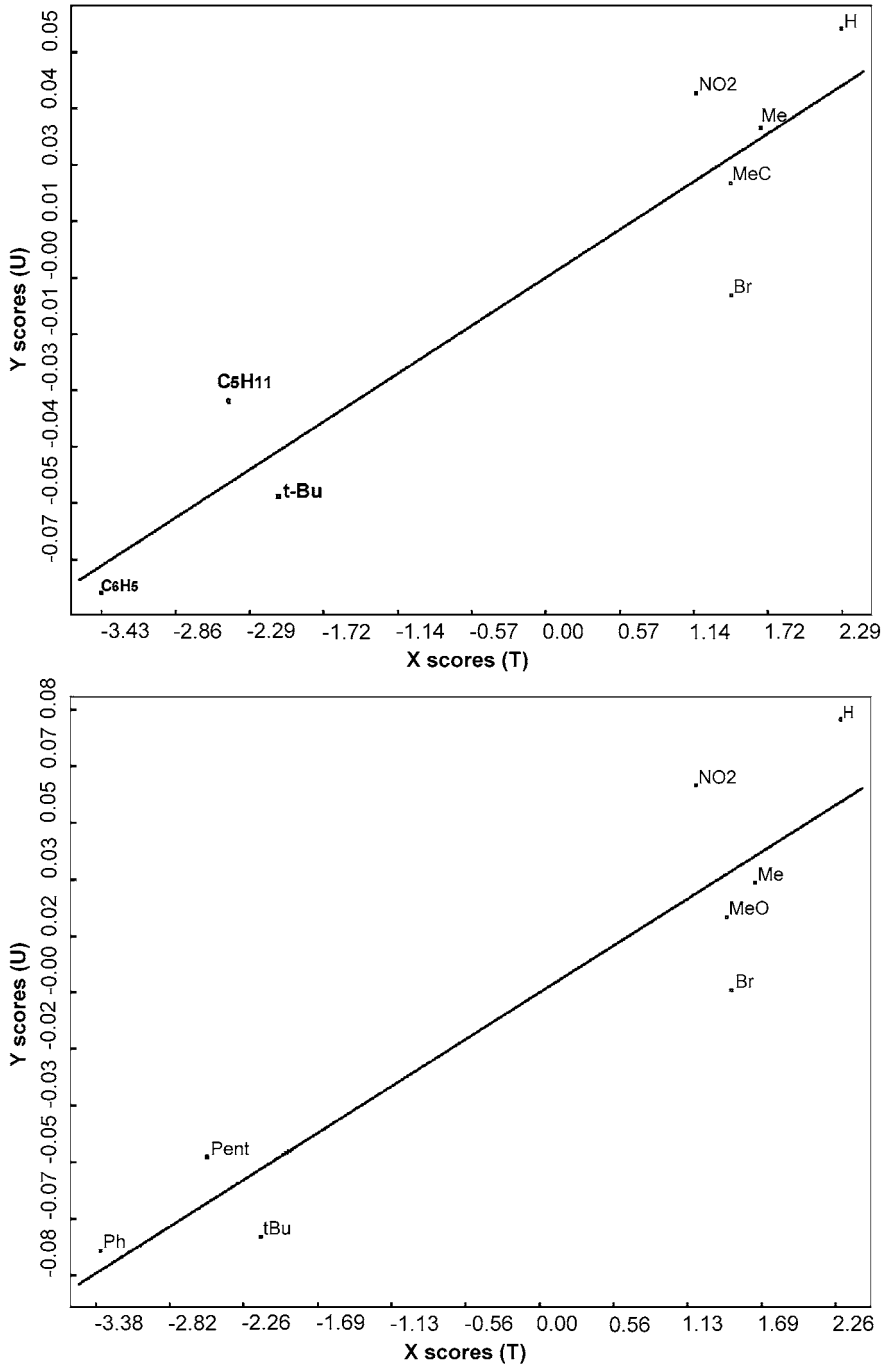


Fig. 4. PLS plots for the correlation between the Almond descriptors (T1) and the measured biological potencies for the hydrazide series (U1). (A) *S. cerevisiae*. (B) *E. coli*.

PLS analysis for all eight hydrazides resulted in quite reasonable statistical parameters ($r^2 = 0.672$ and $r^2_{cv} = 0.665$), with similar coefficients for both models, as described by Eqs. (3) and (4). However, compound 7 (*t*-Bu substituent) was left out (outlier), so improving the predictive power of the models. These results led us to use a new method for the modelling and prediction of pharmacokinetic properties based on computed MIF. VolSurf descriptors provide directly interpretable maps for the H_2O and DRY probes, LVs can be extracted by applying PLS to the VolSurf descriptors set, and the results are shown in Table 4. Hence, calculated molecular properties from GRID 3D molecular fields of interaction energies to correlate 3D hydrazide structures with physicochemical and Taylor–Aris partitioning properties are to be found in Table 4. The PLS plot for the first component, based on the two sets of descriptors (H_2O and DRY) for *S. cerevisiae* is shown in Fig. 6. The first component (29.9% of total variance) discriminates bulky moieties (C_5H_{11}, C_6H_5,

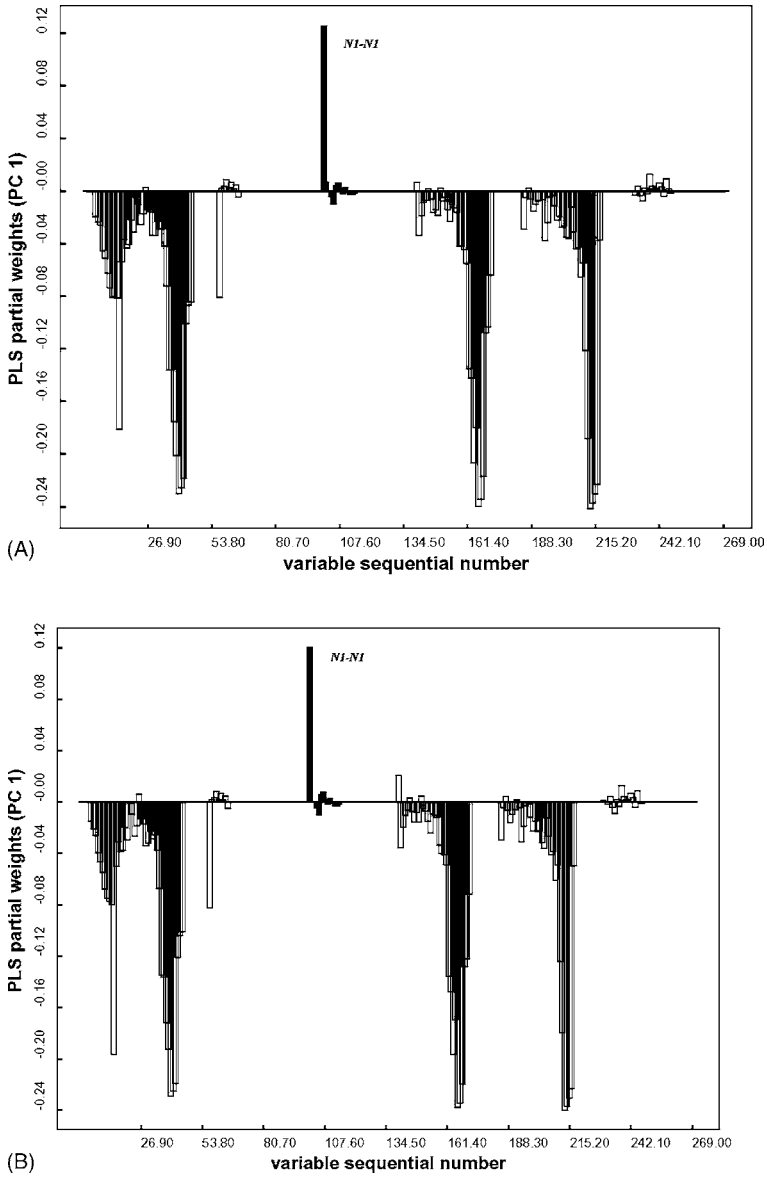


Fig. 5. PLS partial weights of PC1 for the hydrazide calorimetric biopotencies against (A) *S. cerevisiae* and (B) *E. coli*.

and *t*-Bu) from others. The bulk moieties have positive values of PC1, and higher π-lipophilic Taylor–Aris constant (Table 1). These can be seen from PLS coefficient plots shown in Fig. 3. The longer vertical bars indicate the contribution of each single descriptor, while the shorter ones are not important. Capacity factors, and to lesser extent low energy DRY integy moments and hydrophilic–lipophilic balance are inversely correlated with partitioning. This means that when less polar groups are attached to the 4-position of the hydrazide aryl ring (Fig. 2), $\log P_{TA}$ values increases,

Table 4
VolSurf PLS results of the model developed for the eight antimicrobial hydrazides

System	Crossvalidated		Components	Non-cross-validated	
	r^2_{cv}	SDEP		r^2	SDEC
$\log P_{S.cerevisiae}$	0.614	0.110	1	0.889	0.059
$\log P_{E.coli}$	0.512	0.106	2	0.932	0.040
$\pi_{S.cerevisiae}$	0.608	0.111	1	0.887	0.059
$\pi_{E.coli}$	0.524	0.107	2	0.931	0.041

Statistical parameters were defined in Table 3.

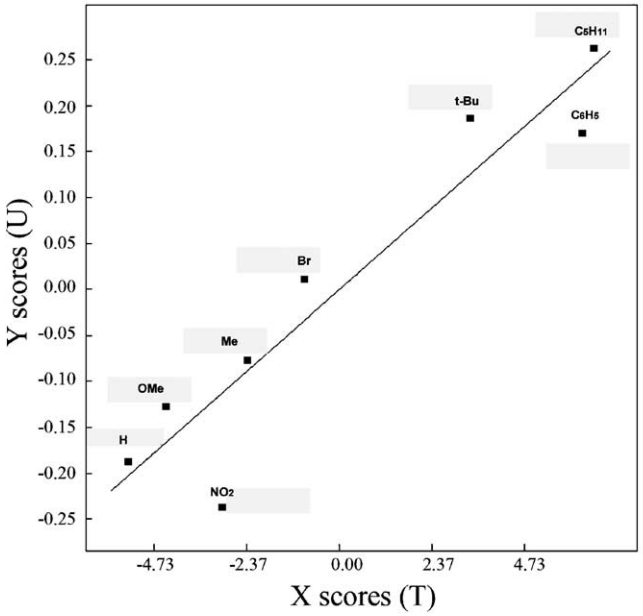


Fig. 6. PLS plot derived from VolSurf descriptors for partitioning data.

hence they tend to be more lipophilic. The most hydrophilic compound 2 (p-NO_2) has a high positive capacity factor, whereas the least one (compound 7, p-C_5H_{11}) has a high negative capacity factor. Size and molecular shape descriptors, water integy moments along with DRY hydrophobic regions and high-energy DRY integy moments are positively correlated with $\log P_{TA}$. It can be seen that compound 2 has a lower positive integy moment (0.00 at -6.0 kcal mol^{-1}) than compound 7, which is 1.51 at the same interaction energy level. Overall, the highly hydrated cell surface (Taylor–Aris partitioning) may account for the balance of all molecular properties responsible for hydrazide antimicrobials partitioning to within the cell system. Similar results were found when $\log P_{E.coli}$ values were used as dependent variable and the same analyses could be drawn from them all.

5. Conclusions

Statistically significant 3D QSAR models were developed based on biological calorimetry. The GRID/Almond and GRID/VolSurf proved to be useful methods leading to insights into the antimicrobial hydrazide properties, and Taylor–Aris partitioning through 3D molecular structures, respectively. $\log P_{TA}$ and biological calorimetric potencies can thus be used in QSAR and QSPR studies. In the dataset employed, it is worth to note that the bulk of the structural diversity lies in the nature of the substituent linked to the 4-position of the aryl ring. A more nearly suitable data set for statistical modeling would require more structural diversity and a larger number of compounds.

Biological calorimetry can be applied to the screening of a variety of chemicals, and for sensitive differentiation between related compounds, thereby generating a good and reliable quantitative biological descriptor. Finally, it maybe envisioned that many cell systems could be automated in order to allow an array of "well-measured" and reliable biological data that should be very useful in drug design, high-throughput screening (HTS), and QSAR studies.

Acknowledgements

We are grateful to Prof. Anthony Beezer (Tony) for his valuable intellectual contribution to the field of calorimetry, and his steadfast guidance and tireless dedication over the years. We are pleased to thank the following Brazilian Agencies for financial support: CNPq, FAPEMIG and FAPESP.

References

[1] P.W. Erhardt, Pure Appl. Chem. 74 (2002) 703–785.
[2] K.C. Thompson, Thermochim. Acta 355 (2000) 83–87.
[3] S. Vyazovkin, Anal. Chem. 74 (2002) 2749–2762.
[4] T.D. Morgan, A.E. Beezer, J.C. Mitchell, A.W. Bunch, J. Appl. Microbiol. 90 (2001) 53–58.
[5] S. Gaisford, G. Buckton, Thermochim. Acta 380 (2001) 185–198.
[6] A.E. Beezer, J.C. Mitchell, R.M. Colegate, D.J. Scally, L.J. Twyman, R.J. Willson, Thermochim. Acta 250 (1995) 277–283.
[7] A.E. Beezer, Thermochim. Acta 349 (2000) 1–7.
[8] P.J. Gittins, L.J. Twyman, Supramol. Chem. 15 (2003) 5–23.
[9] M.L.C. Montanari, A.E. Beezer, C.A. Montanari, Thermochim. Acta 328 (1999) 91–97.
[10] M.L.C. Montanari, A.E. Beezer, C.A. Montanari, D. Pilo-Veloso, J. Med. Chem. 43 (2000) 3448–3452.
[11] M.L.C. Montanari, A.C. Gaudio, C.A. Montanari, Quim. Nova 25 (2002) 231–240.
[12] K. Nakatani, S. Sando, H. Kumasawa, J. Kikuchi, I. Saito, J. Am. Chem. Soc. 123 (2001) 12650–12657.
[13] K. Nakatani, S. Hagihara, S. Sando, S. Sakamoto, K. Yamaguchi, C. Maesawa, I. Saito, J. Am. Chem. Soc. 125 (2003) 662–666.
[14] M.A. Read, A.A. Wood, J.R. Harrison, S.M. Gowan, L.R. Kelland, H.S. Dosanjh, S. Neidle, J. Med. Chem. 42 (1999) 4538–4546.
[15] C. Frisch, G. Schreiber, C.M. Johnson, A.R. Fersht, J. Mol. Biol. 267 (1997) 696–706.
[16] R. Filfil, T.V. Chalikian, J. Mol. Biol. 326 (2003) 1271–1288.
[17] G. Folkers, A. Prota, A. Merz, Farmaco 50 (1995) 449–454.
[18] R.B. Raffa, S.D. Spencer, R.J. Schulingkamp, Biochem. Pharmacol. 63 (2002) 1937–1939.
[19] T. Liang, C.X. Wang, G.L. Zou, Z.Y. Wang, Y.W. Liu, S.S. Ou, Thermochim. Acta 351 (2000) 21–27.
[20] O.A. Raevsky, Uspekhi Khimii 68 (1999) 555–576.
[21] M.L.C. Montanari, C.A. Montanari, D. Pilo-Veloso, A.E. Beezer, J.C. Mitchell, P.L.O. Volpe, Quant. Struct.-Act. Rel. 17 (1998) 102–108.
[22] M.L.C. Montanari, A.E. Beezer, C.A. Montanari, QSAR based on biological microcalorimetry. computational approaches to SAR and lipophilicity, in: B. Testa, H. van de Waterbeemd, G. Folkers, R. Guy, (Eds.), Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical, and Computational Strategies, Wiley-VCH, Zurich, 2001, CD-ROM Lecture 2.
[23] ALMOND version 3.2.0 Molecular Discovery Ltd., London, United Kingdom, 2003.
[24] M. Pastor, G. Cruciani, M. McLay, S. Pickett, S. Clementi, J. Med. Chem. 43 (2000) 3332–3343.
[25] J.K. Buolamwini, H. Assefa, J. Med. Chem. 45 (2002) 841–852.
[26] R.C. Wade, P.J. Goodford, J. Med. Chem. 36 (1993) 148–156.

[27] GRID version 20 Molecular Discovery Ltd., London, United Kingdom, 2002.
[28] S. Wold, K. Esbensen, P. Geladi, Chemom. Intell. Lab. Syst. 2 (1987) 37–52.
[29] (a) H. Martens, T. Naes, Trend Anal. Chem. 3 (1984) 204–210;
(b) H. Martens, T. Naes, Trend Anal. Chem. 3 (1984) 266–271.
[30] S. Wold, Partial least squares, in: S. Kotz, N.L. Johnson (Eds.), Encyclopedia of Statistical Sciences, vol. 6, Wiley, New York, 1985, pp. 581–591.
[31] N. Selitn, J.P. Keeves, Path analysis with latent variables, in: T. Husen, T.N. Postlethwaite (Eds.), International Encyclopedia of Education, 2nd ed., Elsevier, London, 1994, pp. 4352–4359.
[32] G. Cruciani, P. Crivori, P.-A. Carrupt, B. Testa, J. Mol. Struct. Theochem. 503 (2000) 17–30.
[33] G. Cruciani, M. Pastor, R. Mannhold, J. Med. Chem. 45 (2002) 2685–2694.
[34] G. Cruciani, M. Pastor, W. Guba, Eur. J. Pharm. Sci. 11 (2000) S29–S39.
[35] P.J. Goodford, J. Med. Chem. 28 (1985) 849–857.
[36] D.N.A. Boobbyer, P.J. Goodford, P.M. McWhinnie, R.C. Wade, J. Med. Chem. 32 (1989) 1083–1094.
[37] R.C. Wade, K.J. Clark, P.J. Goodford, J. Med. Chem. 36 (1993) 140–147.
[38] C. Cruciani, S. Clementi, P. Crivori, P.-A. Carrupt, B. Testa, VolSurf and its application in structure-disposition relationships, in: B. Testa, H. van de Waterbeemd, G. Folkers, R. Guy (Eds.), Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical, and Computational Strategies, Wiley-VCH, Zurich, 2001, pp. 539–550.
[39] T.I. Oprea, I. Zamora, A.-L. Ungell, J. Comb. Chem. 4 (2002) 258–266.
[40] C. Hansch, A. Leo, Exploring QSAR: Fundamentals and Applications in Chemistry and Biology, ACS, Washington, 1995.
[41] M. Pastor, I. McLay, S. Pickett, C. Clementi, S. J. Med. Chem. 43 (2000) 3233–3243.
[42] Sybyl, Version 6.5, Tripos, Inc. 1998.
[43] G. Folkers, A. Merz, D. Rognan, CoMFA: Scope and limitations, in: H. Kubinyi (Ed.), 3D QSAR in Drug Design, Kluwer/Escom, London, 1998, pp. 583–618.
[44] A. Braun, S. Hammerle, K. Suda, B. Rothen-Rutishauser, M. Gunthert, S.D. Kramer, H. Wunderli-Allenspach, Eur. J. Pharm. Sci. 11 (2000) S51–S60.