Can 'Bacterial-Metabolite-Likeness' Model Improve Odds of 'in Silico' Antibiotic Discovery?

Artem Cherkasov*

Division of Infectious Diseases, Faculty of Medicine, University of British Columbia, 2733 Heather Street, Vancouver, British Columbia V5Z 3J5, Canada

Received November 1, 2005

'Inductive' QSAR descriptors have been used to develop the series of QSAR models enabling 'in silico' distinguishing between antimicrobial compounds, conventional drugs, and druglike substances. The constructed neural network-based models operating by 30 'inductive' parameters have been validated on an extensive set of 2686 chemical structures and resulted in up to 97% accurate separation of the three types of molecular activities. The demonstrated ability of 'inductive' parameters to adequately capture molecular features determining 'antibiotic-like' and 'druglike' potentials have been further utilized to construct a model of 'Bacterial-Metabolite-Likeness' (BML). The same 'inductive' descriptors have been used to train a neural network that could very accurately recognize substances involved into bacterial metabolism (that have been experimentally identified). When the developed model has been applied to the mixed set of antimicrobials, drugs, and druglike chemicals (not used for training the BML model), it exhibited a 2−5-fold recognition preference toward antimicrobial compounds compared to general drugs and an 18- to 45-fold preference when compared to a druglike substance (depending on the model stringency). These results illustrate immanent similarity between conventional antimicrobials and native bacterial metabolites and suggest that the developed BML model can be an effective classification tool for 'in silico' antibiotic studies.

INTRODUCTION

In the series of our previous works we reported the development of 3D-sensitive QSAR descriptors called 'inductive' and demonstrated their successful application in a number of molecular modeling studies including quantification of antibacterial activity of organic compounds[1] and cationic peptides,[1,2] computation of partial charges in small molecules[3] and proteins,[4] and in comparative docking analysis[4,5] as well as in 'in silico' lead discovery.[4,5] The detailed description of 'inductive' QSAR descriptors and their rationale can be found in the recent review.[1]

In summary, all 'inductive' QSAR parameters are related to atomic electronegativity (χ), covalent radii (R), and intramolecular distances (r) and can be derived from the formulas for steric Rs and inductive σ^* parameters (eqs 1 and 2), 'inductive' electronegativity χ (eq 3), 'inductive' partial charge ΔN (eq 4), and 'inductive' analogues of chemical hardness η and softness s (eqs 5 and 6 respectively)

$$Rs_{j\rightarrow N-1} = \alpha\sum_{i\neq j}^{N-1}\frac{R_j^2}{r_{j-i}^2} = \alpha R_j^2\sum_{i\neq j}^{N-1}\frac{1}{r_{j-i}^2} \quad Rs_{G\rightarrow j=\alpha}\sum_{i\subset G, i\neq j}^{n}\frac{R_i^2}{r_{i-j}^2} \tag{1}$$

$$\sigma^*_{j\rightarrow N-1} = \beta\sum_{i\neq j}^{N-1}\frac{(\chi_j^0-\chi_i^0)R_j^2}{r_{j-i}^2}$$

$$\sigma^*_{G\rightarrow j} = \beta\sum_{i\subset G, i\neq j}^{n}\frac{(\chi_i^0-\chi_j^0)R_i^2}{r_{i-j}^2} \tag{2}$$

$$\chi_{N-1\rightarrow j}^0 = \frac{\sum_{i\neq j}^{N-1}\frac{\chi_i^0(R_i^2+R_j^2)}{r_{i-j}^2}}{\sum_{i\neq j}^{N-1}\frac{R_i^2+R_j^2}{r_{i-j}^2}} \quad \chi_{N-1\rightarrow j}^0 = \frac{\sum_{i\neq j}^{N-1}\frac{\chi_i^0(R_i^2+R_j^2)}{r_{i-j}^2}}{\sum_{i\neq j}^{N-1}\frac{R_i^2+R_j^2}{r_{i-j}^2}} \tag{3}$$

$$\Delta N_j = Q_j + \gamma\sum_{i\neq j}^{N-1}\frac{(\chi_j-\chi_i)(R_j^2+R_i^2)}{r_{j-i}^2} \tag{4}$$

(Q_j - formal charge of j)

$$\eta_i = \frac{1}{2\sum_{j\neq i}^{N-1}\frac{R_j^2+R_i^2}{r_{j-i}^2}} \quad \eta_{\mathrm{MOL}} = \frac{1}{s_{\mathrm{MOL}}} = \frac{1}{2\sum_{j\neq i}^{N-1}\frac{R_j^2+R_i^2}{r_{j-i}^2}} \tag{5}$$

$$s_i = 2\sum_{j\neq i}^{N-1}\frac{R_j^2+R_i^2}{r_{j-i}^2} \quad s_{\mathrm{MOL}} = \sum_{j\neq i}^{N}\sum_{j\neq i}^{N}\frac{R_j^2+R_i^2}{r_{j-i}^2} \tag{6}$$

* Corresponding author phone: (604)875-4588; fax: (604)875-4013; e-mail: artc@interchange.ubc.ca.

10.1021/ci050480j CCC: $33.50
Published on Web 01/20/2006

where the variables indexed with the *j* subscript describe the influence of a single atom onto a group of atoms *G* (typically the rest of the *N*-atomic molecule) while *G* indices designate group (molecular) quantities. The linear character of eqs 1−6 makes 'inductive' descriptors readily computable and suitable for sizable databases and position them as appropriate parameters for large-scale models of 'druglikeness', 'antibiotic-likeness', etc.[1] These binary QSAR classifiers represent an emerging topic of 'in silico' research that assist virtual screening studies, combinatorial library design, and large-scale data mining[6−13] etc.

A variety of QSAR descriptors as well as statistical and machine-learning techniques have been previously used for solving the drug/nondrug separation problem and for creating 'druglike', 'agrochemical-like', 'leadlike', and 'natural productlike' binary classifiers.[6−13] A number of QSAR models have been reported that distinguish between 249 antibacterials and general drugs presented in the Tomas-Vert data set.[14−16] Several related 'antibiotic-likeness' modeling studies have also been conducted on smaller sets of antimicrobials.[18−20]

For the purpose of the current study we assembled an extended molecular set consisting of 525 antimicrobials, 959 general drugs, and 1202 druglike substances and used them to evaluate the applicability of 'inductive' descriptors for large-scale modeling of 'antibiotic-likeness' properties and related QSAR categories.

MATERIALS AND METHODS

Training Set. The antimicrobial compounds and conventional drugs used in the current study have been identified from several public resources including the ChemIDPlus database[21] (for antibiotics), an antibiotic database hosted by *The Journal of Antibiotics*[22] (for antibiotics), and from the Merck Index Database[23] (for conventional drugs). Structures of general chemicals justifying druglikeness conditions have been randomly selected from the Assinex Gold collection.[24] The following druglike criteria have been imposed: number of H-bond acceptors between 1 and 10; number of H-bond donors between 1 and 5, molecular weight between 200 and 500 Daltons; number of rotating bonds below 12; hydrophobicity in the range 1−7; and the total polar surface area below 140 A^2. The redundancy of the resulting data set containing antimicrobials, drugs, and druglike substances has been ensured by checking the corresponding SMILES records and by descriptors-based clustering (using 30 'inductive' parameters). All duplicate entries have been removed; all organometallic structures as well as inorganic components (halogen hydrides, metal ions, etc.) have also been eliminated for the set. All molecules containing basic and/or acidic groups have been converted into an un-ionized form.

The resulting final data set included 525 antimicrobial compounds, 959 general drugs, and 1202 druglike chemicals that can be found in the Supporting Information file. All molecular structures have been further optimized with the MMFF94 force-field[25] as it is implemented within the MOE modeling package.[26]

Descriptors. The optimized structures of 2686 compounds have been used for the calculating of 50 'inductive' QSAR descriptors according to formulas 1−6 as it is described elsewhere.[1,2,5] The inductive QSAR descriptors have been calculated by SVL scripts;[26] the corresponding SVL programs for calculating 'inductive' QSAR descriptors can be freely downloaded through the SVL exchange.[27] To eliminate possible cross-correlation between independent variables we removed those descriptors that formed any cross-correlations with correlation coefficients *R* exceeding 0.9. Thus, 30 'inductive' parameters have been selected for further QSAR modeling (more detailed information on the selected QSAR parameters can be found in Table 1). The final set of 30 'inductive' descriptors calculated for 2686 compounds under investigation has been used to generate four neural network-based models distinguishing antimicrobial substances, general drugs, and druglike chemicals.

Neural Networks. To relate 'inductive' descriptors to the Boolean (0|1) indicators of antibiotic-, drug-, and druglike properties of the studied compounds we employed the Artificial Neural Networks (ANN) approach—one of the most effective pattern recognition methods broadly used in the current QSAR research.[28] On the basis of the extended set of 2686 compounds we have compiled three molecular data sets that have been used to train ANNs distinguishing the following bioactivity categories: (a) 'Antimicrobials versus Drugs'—in this case 525 antimicrobial compounds have been assigned a 1.0 activity value and 959 general drugs have been assigned to a null dependent variable; (b) 'Antimicrobials versus Druglikes'—525 antimicrobials assigned to 1.0 and 1202 druglike chemicals assigned to 0.0; (c) 'Antimicrobials versus Others'—525 antimicrobials assigned to 1.0 and 2161 drugs and druglikes have been considered inactive. We have also created a neural network that can simultaneously recognize all three types of bioactivity by classifying 'Antimicrobials versus Drugs versus Druglikes'. For that purpose, we assigned 2 dependent parameters to every compound in the data set: the antimicrobials corresponded to [1.0; 1.0] output (as antimicrobials can also be viewed as drugs); [0.0; 1.0] outputs have been assigned to general drugs; and [0.0; 0.0] − to druglike chemicals.

The values of all 30 'inductive' descriptors have been normalized into the range [0.0 ÷ 1.0], and nonoverlapping training and testing sets have been randomly drawn in 70%−30% proportion by customized Java scripts (and keeping the ratio of active−nonactive substances constant). The random separation of the input patterns into the training and testing groups has been done for effective training of the networks and to avoid their overfitting. All four neural networks have been trained with the standard back-propagation algorithm, input shuffling, weight decay, and using learning rate and threshold values set to 0.80 and 0.10, respectively. The initial network weights have been randomly assigned in a range of [−1.0 ÷ 1.0]. The number of ANN hidden nodes has been optimized on the system 'Antimicrobials versus Drugs' and has been set to 10. Thus, the 30−10−1 ANN configuration has been used in all the following studies, except for the 31−10−2 network that has been developed to separate all 3 groups of compounds simultaneously.

The training and testing of the neural networks have been conducted using the Stuttgart Neural Network Simulator.[29] Typically, 20 independent training and testing runs have been conducted for each ANN, and the resulting training/testing statistics have been reported for the averaged network outputs. More details on networks training, validation, and performance will be discussed in the following sections.

Table 1. 30 Inductive QSAR Descriptors Used in the Study

descriptor	characterization	parental formulas
	χ (Electronegativity)-Based	
EO_Equalized	iteratively equalized electronegativity of a molecule	(3)
Average_EO_Pos	arithmetic mean of electronegativities of atoms with positive partial charge	(3)
Average_EO_Neg	arithmetic mean of electronegativities of atoms with negative partial charge	(3)
	η (Hardness)-Based	
Sum_Hardness	sum of hardnesses of atoms of a molecule	(5)
Sum_Neg_Hardness	sum of hardnesses of atoms with negative partial charge	(5)
Average_Hardness	arithmetic mean of hardnesses of all atoms of a molecule	(5)
Average_Pos_Hardness	arithmetic mean of hardnesses of atoms with positive partial charge	(5)
Average_Neg_Hardness	arithmetic mean of hardnesses of atoms with negative partial charge	(5)
Largest_Pos_Hardness	largest atomic hardness among values for positively charged atoms	(5)
Largest_Neg_Hardness	largest atomic hardness among values for negatively charged atoms	(5)
Hardness_of_Most_Pos	atomic hardness of an atom with the most positive charge	(5)
Hardness_of_Most_Neg	atomic hardness of an atom with the most negative charge	(5)
	s (Softness)-Based	
Global_Softness	molecular softness − sum of constituent atomic softnesses	(6)
Total_Neg_Softness	sum of softnesses of atoms with negative partial charge	(6)
Average_Softness	arithmetic mean of softnesses of all atoms of a molecule	(6)
Largest_Neg_Softness	largest atomic softness among values for positively charged atoms	(6)
Softness_of_Most_Pos	atomic softness of an atom with the most positive charge	(6)
	q (Charge)-Based	
Total_Charge_Formal	sum of charges on all atoms of a molecule (formal charge of a molecule)	(4)
Average_Pos_Charge	arithmetic mean of positive partial charges on atoms of a molecule	(4)
Average_Neg_Charge	arithmetic mean of negative partial charges on atoms of a molecule	(4)
Most_Pos_Charge	largest partial charge among values for positively charged atoms	(4)
Most_Neg_Charge	largest partial charge among values for negatively charged atoms	(4)
	σ^* (Inductive Parameter)-Based	
Most_Pos_Sigma_mol_i	largest positive group inductive parameter *σ*(molecule→atom)* for atoms in a molecule	(2)
Most_Neg_Sigma_mol_i	largest (by absolute value) negative group inductive parameter *σ*(molecule→atom)* for atoms in a molecule	(2)
Most_Pos_Sigma_i_mol	largest positive atomic inductive parameter *σ*(atom→molecule)* for atoms in a molecule	(2)
Most_Neg_Sigma_i_mol	largest negative atomic inductive parameter *σ*(atom→molecule)* for atoms in a molecule	(2)
Sum_Neg_Sigma_mol_i	sum of all negative group inductive parameters *σ*(molecule →atom)* within a molecule	(2)
	Rs (Steric Parameter)−Based	
Largest_Rs_i_mol	largest value of atomic steric influence *rs(atom→molecule)* in a molecule	(1)
Most_Neg_Rs_mol_i	steric influence *Rs(molecule→atom)* ON the most negatively charged atom in a molecule	(1)
Most_Pos_Rs_i_mol	steric influence *Rs(atom→molecule) by* the most positively charged atom ONTO the rest of the molecule	(1)

RESULTS AND DISCUSSION

Model (a): Distinguishing Antimicrobials from Drugs. 1484 input patterns corresponding to 30 'inductive' descriptors for the studied antimicrobials and general therapeutics have been used to train 31−10−1 configured ANN (the structure is featured in Figure 1). As is has been previously described, the network outputs have been assigned to 1.0 for antimicrobial compounds and to 0.0 for all drugs in the training set (containing 70% of all patterns), and 20 independent training runs have been carried out. After each run, the produced network weights have been used to assess the ANN performance on the testing set consisting of the remaining 30% of the input patterns. The counts of the false/true positive and negative predictions have been estimated using 0.5 cutoff for ANN outputs, and the resulting values of Specificity, Sensitivity, Accuracy, and the Positive Prediction Values (PPV) have been collected into Table 2. The averaged 30−10−1 ANN outputs for antimicrobials and drugs can also be found in the Supporting Information.

Table 2 features averaged statistics for training (70% of the inputs used) and testing (30% of the inputs used) network runs as well as accuracy parameters for ANN classification of the entire (100%) set of antimicrobials and drug substances. The estimated values demonstrate that the use of 'inductive' QSAR descriptors results in up to a 92% accurate prediction of antimicrobial activity among the studied compounds. Such accuracy is similar or superior to the results of several similar 'antibiotic-likeness' studies conducted on smaller sets of molecular structures where the overall accuracy ranged from 78%[16] to 98%[18] depending on the QSAR methodology, size of molecular set, and validation techniques used.

Model (b): Distinguishing Antimicrobials from Druglike Substances. When the 30−10−1 neural network has been trained on 525 antimicrobials and 1202 general drugs (using the same set of 'inductive' descriptors as in the previous example)—the resulting separation accuracy achieved 96% during the training phase and 92% for the testing runs. The averaged Specificity, Sensitivity, Accuracy, and PPV values established with a 0.5 cutoff have also been collected into Table 2. The estimated parameters illustrate that the ANN can separate antimicrobials from druglike chemicals more effectively than from actual drugs. Less accurate separation of antimicrobials from general drugs may be explained by the fact that some conventional nonantibiotic therapeutics can possess side antibacterial activity. Thus, there have been reports on the antimicrobial potential of numerous types of conventional drugs,[30−39] and, therefore,

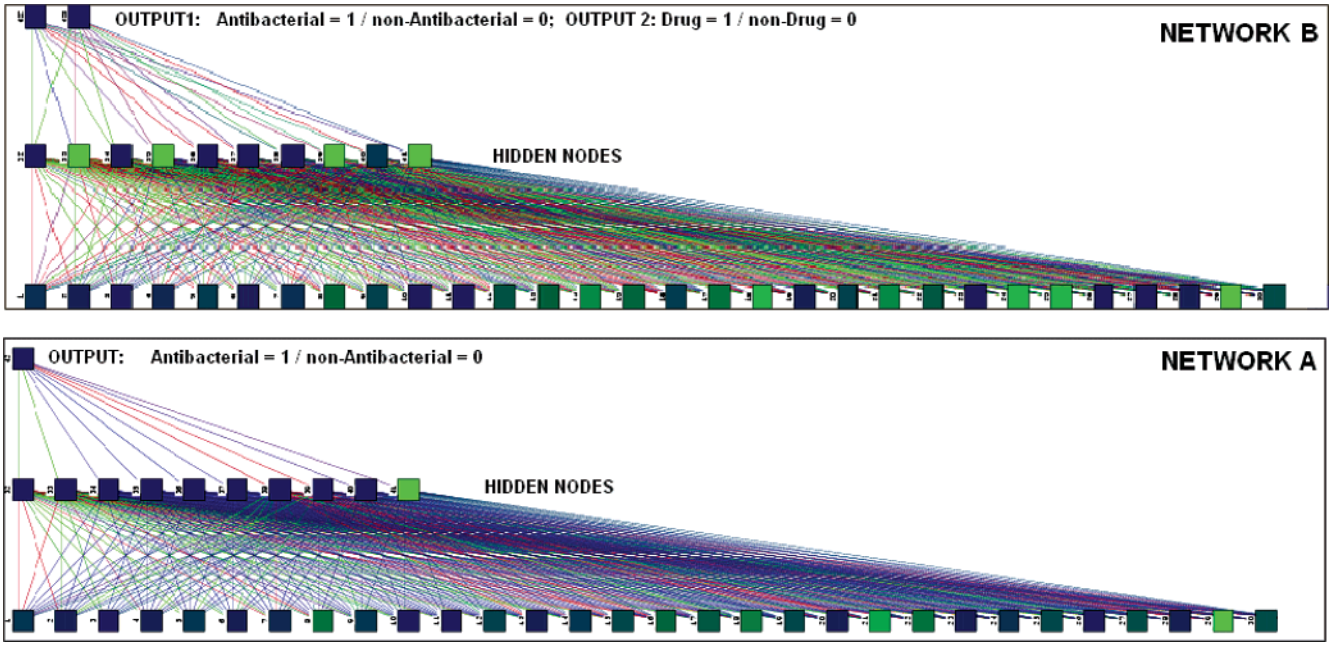


Figure 1. Neural networks used for distinguishing antibacterial compounds from a combined set of drugs and non-drugs (network A) and for classification of all 3 groups of compounds (network B). The edge color coding corresponds to the assigned weights (ranged from red: −1 to green +1), the nodes colors designate the last passed value calibrated from 0 to 1 (from blue to bright green).

Table 2. Specificity, Sensitivity, Accuracy, and Positive Predictive Values of Various QSAR Models

	antimicrobials from drugs		antimicrobials from druglikes		distinguishing antimicrobials from all others		distinguishing antimicrobials versus drugs versus druglikes		QSAR model for bacterial metabolites	
	train	test	train	test	train	test	train	test	train	test
T_P	327	130	332	140	294	124	270	89	360	139
T_N	631	248	841	342	1490	621	1486	644	792	347
F_P	49	33	7	14	32	20	17	14	39	26
F_N	33	35	30	23	66	41	108	58	48	19
SPEC	0.93	0.88	0.99	0.96	0.98	0.97	0.99	0.98	0.95	0.93
SENS	0.91	0.79	0.92	0.86	0.82	0.75	0.71	0.61	0.88	0.88
ACCUR	0.92	0.85	0.97	0.93	0.95	0.92	0.93	0.91	0.93	0.92
PPV	0.87	0.80	0.98	0.91	0.90	0.86	0.94	0.86	0.90	0.84
NPV	0.95	0.88	0.97	0.94	0.96	0.94	0.93	0.92	0.94	0.95

it is possible that some compounds form the 'negative' control group use in the example (a) can act as antimicrobials.

Model (c): Distinguishing Antimicrobials from All Others. The 30−10−1 network trained and tested on 525 antimicrobials and 2161 nonantimicrobial compounds of the 'negative control' group (corresponding to the combined set of general drugs and druglikes) allowed 92−94% accurate separation of the two activity classes. The corresponding values of average accuracy can be found in Table 2 along with all other statistics characterizing the performance of the developed QSAR model on the entire set of 2686 substances. These statistical parameters demonstrate that the QSAR model (c) represents a somewhat average version of models (a) and (b) confirming the overall good ability of inductive QSAR parameters to separate antimicrobials, drugs, and druglike compounds.

(d) Distinguishing Antimicrobials versus Drugs versus Druglikes. Considering that 'inductive' descriptors could confidently distinguish antimicrobials from drugs and from druglike chemicals, we attempted creating a neural network with two output nodes enabling simultaneous recognition of all three activity types. The expectation was that such an ANN configuration will produce more robust predictions and will result in reduction of false positives.

As it has been described in the 'Materials and Methods' section, we assigned two dependent variables to 2686 molecules under study. Namely, 525 substances have been assigned activity values [1.0; 1.0] corresponding to their classification as antimicrobials and as drugs; 959 molecules have been associated with [0.0; 1.0] outputs reflecting the absence of antimicrobial- and presence of drug activity; and the remaining 1202 druglike chemicals were considered as 'double inactive' and have been assigned [0.0; 0.0] values.

We have adopted 30−10−2 configuration for the ANN (more details can be found in Figure 1 and legends) and trained it in the same manner as described for previous models (a)−(c). The network outputs have also been interpreted with 0.5 threshold, and only when both output values were correct the corresponding prediction by the ANN has been considered true.

The performance by the 30−10−2 network has also been assessed with Specificity, Sensitivity, Accuracy, and PPV values. As it can be seen from the data in Table 2 the developed model resulted in ∼98% specific classification

of the three types of activity as the number of false positive predictions has been considerably reduced compared to the previously discussed models that utilized the 30−10−1 ANN configuration. The overall prediction accuracy of the network with two output nodes also appeared very high (93% and 91%, respectively, for the training and testing phases).

Thus, the results of classification of the three classes of compounds by the developed binary QSAR models have readily demonstrated that a limited set of 'inductive' descriptors can adequately capture those structural features of the studied chemicals that are relevant for their antimicrobial- and drug-behaviors. The rationale for this adequacy can be possibly explained by the fact that 'inductive' QSAR parameters cover a broad range of proprieties of bound atoms and molecules related to their size, polarizability, electronegativity, electronic, and steric interactions. Thus, despite certain drawbacks, such as a 'black-box' nature of ANN solutions, the developed QSAR models demonstrate a very high prediction accuracy and can be suggested as effective tools for large-scale QSAR studies/discovery of antimicrobial compounds. On the other hand, most of the effective antimicrobial leads identified to date are either derived from or similar to the substances naturally involved in bacterial metabolism. Thus, to enable more effective virtual screening for antibiotic candidates we attempted to develop a QSAR model for 'Bacterial-Metabolite-Likeness'.

QSAR Model for Bacterial Metabolites. According to the formally effective Section 507(a) of the Federal Food Drug and Cosmetic Act,[40,41] antibiotics have been defined as "... and drug intended for use by man containing any quantity of any chemical substance which is produced by a microorganism and has the capacity to inhibit or destroy microorganisms in dilute solution ...". This definition reflected the fact that, historically, bacterial metabolites (and chemically synthesized equivalents) served as the main source of antibiotic leads. Thus, the development of effective 'in silico' tool for assessing the lead resemblance to natural bacterial metabolites represents an important task.

We employed QSAR methodology described in the previous examples (a)−(d) to develop a binary classifier operating on bacterial metabolites. First of all, we assembled a set of 565 compounds recently isolated from bacteria and characterized by the Analyticon Discovery Company.[42] The names and the SMILES records for most of the studied bacterial metabolites can be found in the Supporting Information (though some metabolites structures have not been disclosed as they can only be obtained from Analyticon under the nondisclosure agreement). The previously described 1202 druglike substances from the Assinex Gold collection have been selected as the negative control for building the 'Bacterial-Metabolite-Like' (BML) model. Using SMILES-based comparison and 'inductive' descriptor-based clustering we ensured that the bacterial metabolites data set does not contain duplicates to the above-described 2686 chemical structures.

Similarly to the previous cases, we calculated 30 'inductive' descriptors featured in Table 1 for all metabolites and assigned 1.0 dependent value to all active entries and 0.0 to the negative control (druglike substances). The training of the 30−10−1 network has been similarly carried out using 70% to 30% split for the training and testing sets. Twenty independent runs have been conducted, and the separation of active and nonactive molecules by the resulting ANN solutions has been assessed using 0.5 output threshold.

The estimated average Accuracy, Specificity, Sensitivity, and PPV reflecting the ability of the developed model to distinguish bacterial metabolites from the druglikes (derived from the Assinex Gold collection[24]) have been colleted into Table 2. The prediction accuracy for the training and the testing runs appeared in the range of 92−93% which demonstrates a very good discriminative ability of 'inductive' descriptors and characterizes the resulting ANN-based model of 'Bacterial-Metabolite-Likeness' (BML) as very adequate.

Then, the developed BML model has then been applied to the validation set of 2686 molecular structures consisting of the conventional antimicrobials, drugs, and druglike compounds. In this case, however, we *did not* use the original set of 1202 druglike substances that have previously been used for training the BML model. Instead, we assembled the 'external' set of 1202 druglike structures randomly derived from the National Cancer Institute (NCI) data set.[43] To derived druglike substances from the NCI data set, we employed the same criteria as described in the previous sections. The nonredundancy of the selected 1202 NCI structures has been ensured through the SMILES records; the structures of the compounds have been optimized with the MMFF94 force field,[25] and all 30 inductive parameters previously used in the BML model training have been computed. Thus, the resulting set to be used for assessing performance of the BML model, included 525 antimicrobial compounds, 959 general drugs, and 1202 NCI druglike chemicals that can also be found in the Supporting Information file.

The objective for applying the BML model to such data set was to investigate the overall similarity of the three classes of chemicals (antimicrobials, drugs and druglikes) to native bacterial metabolites. The patterns of 30 'inductive' descriptors for each of 2686 molecular structures have been passed through the pretrained BML model to produce a single network output that could be considered as likelihood of the corresponding compound to be a metabolite. The generated ANN predictions for 525 antimicrobials, 959 drugs, and 1202 NCI druglike chemicals (can be found in the Supporting Information) have then been classified with 0.5 cutoff and transformed into the corresponding confusions matrices (false/true positive/negatives) as well as percent yield (%Y), percent accurate (%A), enrichment factor (E), and goodness of hit list (GH) parameters custom for in silico screening studies:

$$\%\mathrm{Y} = H_a/H_t$$

$$\%\mathrm{A} = H_a/A$$

$$E = (H_a/H_t)/(A/D)$$

$$\mathrm{GH} = \left(\frac{H_a(3A + H_t)}{4H_tA}\right) \times \left(1 - \frac{H_t - H_a}{D - A}\right)$$

where H_t is the total number of compounds in the hit list (in our case − ANN outputs exceeding the threshold), H_a is the number of known actives in the hit list (true positives), A is the number of the active compounds in the database, and D is the number of compounds in the database.

Table 3. Ability of the 'Bacterial-Metabolite-Likeness' Model To Recognize Antimicrobials, Drugs, and Druglike Compounds

type	threshold	TP	TN	FP	FN	accuracy	actives	active hits	total hits	% yield	% actives	enrichment factor	goodness of hits
antimicrobials recall	0.5	337	1895	266	188	0.83	525	337	603	55.89	64.19	2.86	0.51
	0.6	300	1972	189	225	0.85	525	300	489	61.35	57.14	3.14	0.55
	0.7	198	2078	83	327	0.85	525	198	281	70.46	37.71	3.61	0.60
	0.8	56	2135	26	469	0.82	525	56	82	68.29	10.67	3.49	0.53
drugs recall	0.5	232	1356	371	727	0.59	959	232	603	38.47	24.19	1.08	0.27
	0.6	167	1405	322	792	0.59	959	167	489	34.15	17.41	0.96	0.24
	0.7	73	1519	208	886	0.59	959	73	281	25.98	7.61	0.73	0.19
	0.8	19	1664	63	940	0.63	959	19	82	23.17	1.98	0.65	0.17
druglikes recall	0.5	34	915	569	1168	0.35	1202	34	603	5.64	2.83	0.13	0.03
	0.6	22	1017	467	1180	0.39	1202	22	489	4.50	1.83	0.10	0.03
	0.7	10	1213	271	1192	0.46	1202	10	281	3.56	0.83	0.08	0.02
	0.8	7	1409	75	1195	0.53	1202	7	82	8.54	0.58	0.19	0.06

These parameters have been computed using different ANN threshold values and have been collected into Table 3 for those cases when antimicrobials, drugs, and druglikes have been considered as active substances. By doing so we could evaluate the ability of the BML model to recognize these three groups of compounds from a mixed pool of chemicals. Figure 2 features compositions of the resulting BML hit lists produced with 0.5, 0.6, and 0.7 output thresholds; where numbers of identified antimicrobials, drugs, and druglikes are placed into separate circles color-coded by the number of constituent hits. Panels (b)−(d) in Figure 2 illustrate that the produced BML hit lists contain up to 56%, 61%, and 70% of antimicrobial compounds for 0.5, 0.6, and 0.7 thresholds, respectively. These data also graphically illustrate the fact that 30−10−1 neural network pretrained on bacterial metabolites can distinguish antibacterials from other substances with 83% accuracy, while general drugs could only be recognized by the BML model with 59% accuracy (see Table 3 for more details). The application of the 'Bacterial-Metabolite-Like' criteria to druglike substances produced only a few hits resulting in 35% prediction accuracy. Thus, only 34, 22, 10, and 7 compounds have been sufficiently recognized by the BML model from the set of 1202 NCI druglike substances, when 0.5, 0.6, 0.7, and 0.8 ANN threshold have been respectively applied.

To summarize this section, it is necessary to stipulate that the developed BML model is 2.6−5 times more likely to recognize antimicrobial compounds compared to general drugs and 18−45 times more likely to recognize antimicrobial compounds than just druglike substances (when judged by the corresponding enrichment factors). On one hand, these findings clearly demonstrate that there exists a definite similarity between conventional antimicrobial therapeutics and native bacterial metabolites. On the other hand, the results characterize the developed "Bacterial-Metabolite-Likeness" model as a potentially useful additional QSAR tool for in silico antibiotic discovery.

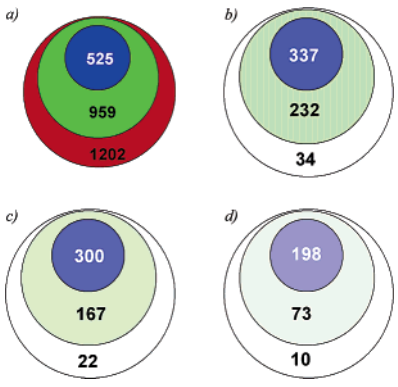
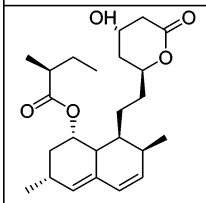
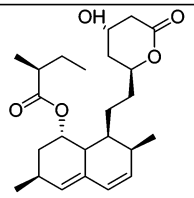
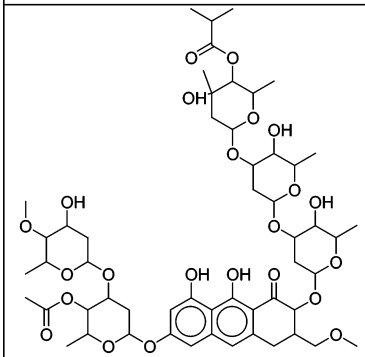
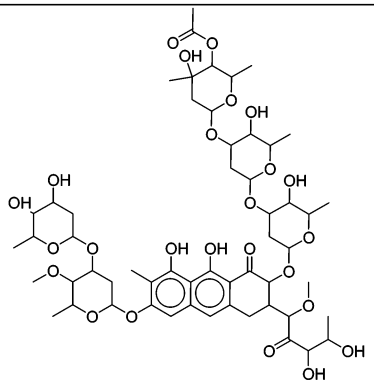
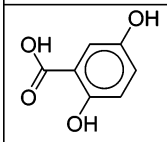
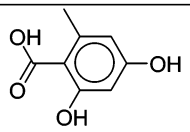


Figure 2. Composition of the studied molecular data set (a) and the distribution of active hits produced by the developed 'BML model' at various ANN thresholds (b−d): (b) ANN output threshold = 0.5, (c) ANN output threshold = 0.6, and ANN output threshold = 0.7.

Unfortunately, the ANN nature of the developed solutions and utilized 'inductive' QSAR descriptors do not allow direct identification of easily interpretable factors of intra- and/or intermolecular interactions that distinguish bacterial metabolites from drugs and druglike substances. We can only speculate that certain molecular features are likely to involve chemicals into bacterial uptake or can enhance their ability to penetrate bacterial cell walls or allow the compounds to fit the specific physical-chemical environment of bacterial cells. What is certain is, however, that it is possible to construct effective QSAR models distinguishing bacterial metabolites from other chemical substances, and therefore the current study may lay the foundations for further broad investigation of the BML systems. Thus, thorough investigation of bacterial metabolites is required in order to derive dependences similar to Lipinski's rule or to define special patterns of structural/fragmental features (such as ratio of heteroatoms, number of rings, number of chiral centers among others) as it has been previously done to distinguish natural produces, drugs, and combinatorial chemicals.[44] It is anticipated that identification of such distinguished factors of 'bacterial-metabolite-likeness' will help developing new antibiotics that more closely resemble native bacterial compounds and therefore may have enhanced potency or cause less pathogen resistance. Such investigations are currently under way.

Interpretation of Some False Positive Predictions. As it has been previously mentioned, some conventional drugs and/or untested chemicals may possess profound but unrecognized antimicrobial activity. On one hand, such a possibility may lower the accuracy of 'antibiotic-likeness' models (as our own results indicate, even a highly specific 30−10−2 neural network has produced significant 'antibiotic-like' predictions for a number of conventional drugs that do not have any antimicrobial annotation). On the other hand, some of these 'false positive' predictions may possess nonappreciated antibiotic potential and, thus, may be considered as interesting antimicrobial leads. Moreover, some of the studied conventional therapeutics produced significant

Table 4. Some Nonantimicrobial Compounds that Produced Significant Antibiotic-likeness Scores and Possess Close Bacterial Metabolite Analogues

Lead compound	Merck annotation	BML score	Antibiotic-likeness (30-10-2 ANN)	Metabolite analogue
 Lovastatin	Antilipemic	0.7091	0.5669	 NP-007587
 Olivomycin A	Antineoplastic Cytostatic agent	0.72190	0.997	 **NP-009248**
 Gentisic acid	Analgesic; Antiinflamatory	0.7456	0.6605	 NP-001423

scores by both 'antibiotic-likeness' and 'bacterial-metabolite-likeness' models, which makes these substances even more attractive for further testing.

Thus, Table 4 features three conventional therapeutics—lovastatin, gentisic acid, and olivomycin A, that all exhibited significant 'antibiotic-likeness' and 'bacterial-metabolite-likeness' potentials when processed by the developed QSAR models. Moreover, all these three compounds have structurally similar bacterial metabolites present in the BML training set (also featured in Table 4). The structural similarity to bacterial metabolites has been established by the descriptor-based clustering using 30 'inductive' parameters and Tanimoto similarity criteria.

In particular, the bacterial metabolite NP-007587 appeared to be lovastatin's steric isomer. This observation is not too surprising considering that lovastatin has originally been discovered as a fungal metabolite isolated from a strain of *Aspergillus terreus*.[45] Another compound featured in Table 4—the gentisic acid (scored 0.7456 as a bacterial metabolite-like and 0.6605 as an antimicrobial-like) closely resembles a bacterial substance NP-001423 that differs only by one methyl group and a substitution position of hydroxyl radical. Interestingly, olivomycin A, annotated as an 'antineoplastic, cytostatic agent' by the Merck Index Database[23] produced a very significant 0.997 ANN output as a potential antimicrobial and yielded a significant BML score of 0.7219. This large compound appeared to have a close bacterial metabolite analogue—a compound NP-009248 that is also featured in Table 4.

Naturally, we conducted a literature search for any evidences of antimicrobial activities of lovastatin, gentisic acid, and olivomycin A. It turned out that the antilipemic drug lovastatin (scoring 0.7091 out of 1.0 as a bacterial metabolite and 0.5669 as an antimicrobial) does, in fact, exhibit antibiotic potentials against a number of bacteria including *Escherichia coli*,[46] *Halobacterium holobiuim*,[47] and *Halobacterium volcanii*.[48] Antifungal[49−52] and antiparasite[53−58] activities of lovastatin have also been well documented with some studies reporting the minimal inhibitory concentration of lovastatin being as low as 0.25 μM.[59]

The analgesic, antiinflamatory agent gentisic acid and its close derivatives demonstrated profound antimicrobial activity against a range of human pathogens including *Campylobacter jejuni*, *Escherichia coli*, *Listeria monocytogenes*, and *Salmonella enterica*.[60−61]

Antimicrobial activity of olivomycin A—a compound isolated from the fungus *Actinomyces olivoreticuli*—has also been well documented in early antibiotic papers.[62−64] Olivomycins have even been annotated as antibiotics in several databases different from the Merck Index, such as ChemIDPlus[21] and the NIAID database.[65]

Thus, all three conventional therapeutics having very similar bacterial metabolite analogues do demonstrate a certain antimicrobial potential and, in principle, can serve as antimicrobial leads for future drug discovery attempts. A small number of other conventional drugs have also produced significant 'antibiotic-likeness' and 'bacterial-metabolite-likeness' predictions, but those molecules did not have close

structural analogues among experimentally determined bacterial metabolites. Possible antimicrobial activity of those substances will also be further investigated by means of extensive literature and a database search and by experimental verification. What is possible to state at the moment however is that the examples of lovastatin, onlivomycin, and gentisic acid illustrate a possible application of the developed 'antibiotic-likeness', and 'bacterial-metabolite-likeness' QSAR models for large-scale screening for novel antimicrobial candidates.

CONCLUSIONS AND FUTURE DIRECTIONS

The results of the present work demonstrate that a range of atomic, substituent, and molecular properties that are represented by the 'inductive' QSAR descriptors allow adequate separation of the four groups of chemicals: antimicrobial substances, general drugs, druglike compounds, and bacterial metabolites. Using only 30 'inductive' descriptors with no additional parameters we were able to achieve up to 97% correct separation of these activities using the Artificial Neural Network approach. Thus, the developed QSAR models can be suggested as useful 'in silico′ tool for distinguishing and ranking potential antibiotic leads.

It should also be mentioned, that our choice of the ANN approach for the current 'pilot-study' has not been dictated by any particular preferences rather than the factor that we have previously been using ANN for modeling 'antibiotic-likeness' potentials.[1] Thus, the accuracy of the developed approach operating by the inductive descriptors can presumably be further improved by applying other advanced classification techniques such as Support Vector Machines or Bayesian Neural Networks. Use of merely statistical techniques in conjunction with the 'inductive' QSAR descriptors will also be pursued in the following studies and the role of specific molecular features in antibiotic-, drug-, and bacterial metabolite-likeness will be investigated.

Supporting Information Available: Resulting set for assessing performance of the BML model included 525 antimicrobial compounds, 959 general drugs, and 1202 NCI druglike chemicals. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Cherkasov, A. 'Inductive' Descriptors. 10 Successful Years in QSAR. *Curr. Comput.-Aided Drug Des.* **2005**, *1*, 21−42.

(2) Cherkasov, A.; Jankovic, B. Application of 'Inductive' QSAR Descriptors for Quantification of Antibacterial Activity of Cationic Polypeptides. *Molecules* **2004**, *9*, 1034−1052.

(3) Cherkasov, A. Inductive Electronegativity Scale. Iterative Calculation of Inductive Partial Charges. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 2039−2047,

(4) Cherkasov, A. Electronegativity of Atoms in Proteins. *J. Chem. Inf. Model.* **2005**, *45*, submitted for publication.

(5) Cherkasov, A.; Shi, Z.; Fallahi, M.; Hammond, G. L. Successful in Silico Discovery of Novel Non-Steroidal Ligands for Human Sex Hormone Binding Globulin. *J. Med. Chem.* **2005**, *48*, 3203−3213.

(6) Anzali, S.; Barenickel, G.; Cezanne, B.; Krug, M.; Filimonov, D.; Poroikob, V. Discriminating between Drugs and Nondrugs by Prediction of Activity Spectra for Substances (PASS). *J. Med. Chem.* **2001**, *44*, 2432−2437.

(7) Byvalov, E.; Fechner, U.; Sadowski, J.; Schneider, G. Comparison of Support VEctr Machines and Artificial Neural Network Systems for Drug/Nondrug classification. *J. Chem. Inf. Comput. Sci*. **2003**, *43*, 1882−1889.

(8) Zernov, V.; Balakin, K. V.; Ivaschenko, A. A.; Savchuk, N. P.; Pletnev, I. V. Drug Discovery Using Support Vector Machines. The Case Studies of Drug-likeness, Agrochemical-likeness, and Enzyme Inhibition Predictions. *J. Chem. Inf. Comput. Sci*. **2003**, *43*, 2048−2056.

(9) Murcia-Soler, M.; Perez-Gimenez, F.; Garcia-March, F. J.; Salabert-Salvador, M. T.; Diaz-Villanueva, W.; Castro-Bleda, M. J. Drugs and Nondrugs: An Effective Discrimination with Topological Methods and Artificial Neural Networks. *J. Chem. Inf. Comput. Sci*. **2003**, *43*, 1688−1702.

(10) Frimurer, T. M.; Bywater, R.; Naerum, L.; Lauritsen, L. N.; Brunak, S. Improving the Odds in Discriminating 'Drug-like' from 'Non Drug-like' Compounds. *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 1315−1324.

(11) Galvez, J.; de Julian-Ortiz, J. V.; Garcia-Demenech, R. General Topological Patterns of Known Drugs. *J. Mol. Graphics Modell.* **2001**, *20*, 84−94.

(12) Sadowski, J.; Kubinyi, H. A scoring Scheme for Discriminating between Drugs and Nondrugs. *J. Med. Chem.* **1998**, *41*, 3325−3329.

(13) Ajay, Walters, W. P.; Murcko M. A. Can we: earn to Distinguish between "Drug-like" and "Nondrug-like" Molecules? *J. Med. Chem.* **1998**, *41*, 3314−3324.

(14) Tomas-Vert, F.; Perez-Gimenez, F.; Salabert-Salvador, M. T.; Garcia-March, F. J.; Jaen-Oltra, J. Artificial Neural Networks Applied to the Discrimination of Antibacterial Activity by Topological Methods. *J. Mol. Struct. (THEOCHEM)*. **2000**, *504*, 249−259.

(15) Cronin, M. T. D.; Aptula, A. O.; Dearden, J. C.; Duffy, J. C.; Netzeva, T. I.; Patel, H.; Rowe, P. H.; Schultz T. W.; Worth, A. P.; Voutzoulidis, K.; Schuurmann, G. *J. Chem. Inf. Comput. Sci*. **2002**, *42*, 869.

(16) Murcia-Soler, M,; Perez-Gimenez, F.; Garcia-March, F. J.; Salabert-Salvador, M. T.; Diaz-Villanueva, W.; Castro-Bleda, M. J.; Villanueva-Pareja, A. Artificial neural networks and linear discriminant analysis: a valuable combination in the selection of new antibacterial compounds. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1031−1041.

(17) Molina, E.; Diaz, H. G.; Gonzalez M. P.; Rodriguez, E.; Uriarte, E. Distinguishing Antibacterial Compounds through a Topological Substructural Approach. *J. Chem. Inf. Comput. Sci*. **2004**, *44*, 515−521.

(18) Garcia-Domenech, R.; de Julian-Ortiz, J. V. Antimicrobial Activity Characterization in a Heterogeneous Group of Compounds. *J. Chem. Inf. Comput. Sci*. **1998**, *38*, 445−449.

(19) Mishra, R. K.; Garcia-Demenech, R.; Galvez, J. Getting Discriminant Functions of Antibacterial Activity from Physicochemical and Topological Parameters. *J. Chem. Inf. Comput. Sci*. **2001**, *41*, 387−393.

(20) Jaen-Oltra, J.; Salabert-Sakvador, M. T.; Garcia-March, F. J.; Perez-Giminez, F. Tomas-Vert, Artificial neural network applied to prediction of fluorquinolone antibacterial activity by topological methods. F. *J. Med. Chem.* **2000**, *43*, 1143−1148.

(21) ChemIDPlus database: http://chem.sis.nlm.nih.gov/chemidplus/.

(22) The Journal of Antibiotics database: http://www.nih.go.jp/∼jun/NADB/byname.html.

(23) *The Merck Index 13.4 CD-ROM Edition*; CambridgeSoft: Cambridge, MA, 2004.

(24) Assinex Gold Collection; Assinex Ltd.: Moscow, 2004.

(25) Halgren, T. A. Merck molecular force field 0.1. Basis, form, scope, parametrization, and performance of MMFF94. *J. Comput. Chem*. **1996**, *17*, 490−519.

(26) Molecular Operational Environment; Chemical Computation Group Inc.: Montreal, Canada, 2004.

(27) MOE SVL exchange community: http://svl.chemcomp.com/index.php.

(28) Zupan, J.; Gasteiger, J. *Neural Networks in Chemistry and Drug Design*, 2nd ed.; Wiley: New York, 1999.

(29) *SNNS: Stuttgart Neural Network Simulator*, *Version 4.0*; University of Stuttgart: 1995.

(30) Annadurai, S.; Basu, S.; Ray, S.; Dastidar, S. G.; Chakrabarty, A. N. Antibacterial activity of the antiinflammatory agent diclofenac sodium. *Indian J. Exp. Biol.* **1998**, *36*, 86−90.

(31) Dash, S. K.; Dastidar, S. G.; Chakrabarty, A. N. Antibacterial property of promazine hydrochloride. *Indian J. Exp. Biol.* **1977**, *15*, 324−326.

(32) Dastidar, S. G.; Chaudhury, A.; Annadurai, S.; Roy, S.; Mookerjee, M.; Chakrabarty, A. N. In vitro and in vivo antimicrobial action of fluphenazine. *J. Chemother*. **1995**, *7*, 201−206.

(33) Dastidar, S. G.; Das, S.; Mookerjee, M.; Chattopadhyay, D.; Ray, S.; Chakrabarty, A. N. Antibacterial activity of local anaesthetics procaine & lignocaine. *Indian J. Med. Res.* **1988**, *87*, 506−508.

(34) Dastidar, S. G.; Jairaj, J.; Mookerjee, M.; Chakrabarty, A. N. Studies on antimicrobial effect of the antihistaminic phenothiazine trimeprazine tartrate. *Acta Microbiol. Immunol. Hung.* **1997**, *44*, 241−247.

(35) Dastidar, S. G.; Mondal, U.; Niyogi, S.; Chakrabarty, A. N. Antibacterial property of methyl-DOPA & development of antibiotic cross-resistances in m-DOPA mutants. *Indian J. Med. Res.* **1986**, *84*, 142−147.

(36) Kristiansen, J. E. Experiments to illustrate the effect of chlorpromazine on the permeability of the bacterial cell wall. *Acta Pathol. Microbiol. Scand. [B]* **1979**, *87*, 317−319.

(37) Kristiansen, J. E. The antimicrobial activity of nonantibiotics. Report from a congress on the antimicrobial effect of drugs other than

antibiotics on bacteria, viruses, protozoa, and other organisms. *APMIS Suppl.* **1992**, *30*, 7−14.
(38) Munoz-Bellido, J. L.; Munoz-Criado; Garcia-Rodrigez, J. A. Antimicrobial Activity of Psychotropic Drugs. Selective Serotonin Uptake Inhibitors. *Int. J. Antimicrob. Agents* **2000**, *14*, 177−180.
(39) Lind, K.; Kristiansen, J. E. Effect of some psychotropic Drugs and a Barbiturate on Mycoplasmas. *Int. J. Antimicrob. Agents* **2000**, *14*, 235−238.
(40) Conforming regulations regarding removal of section 507 of the Federal Food, Drug, and Cosmetic Act; confirmation of effective date. Food and Drug Administration, HHS. Direct final rule; confirmation of effective date. *Fed Regist*. **1999**, 17; 26657 [PMID: 10558515].
(41) The Federal Food Drug and Cosmetic Act: http://www.fda.gov/opacom/laws/fdcact/fdctoc.htm.
(42) Analyticon Discovery; Company: www.ac-discovery.com.
(43) NCI Open Database Compounds: http://cactus.nci.nih.gov/ncidb2/download.html.
(44) Feher, M.; Schmidt, J. M. Property Distributions: Differences between Drugs, Natural Products and Molecules from Combinatorial Chemistry. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 218−227.
(45) Alberts, A. W. Lovastatin and simvastatin—inhibitors of HMG CoA reductase and cholesterol biosynthesis. *Cardiology* **1990**, *77* (Suppl 4), 14−21.
(46) Zhou, D.; White, R. H. Early steps of isoprenoid biosynthesis in *Escherichia coli*. *Biochem. J.* **1991**, *273*, 627−634.
(47) Cabrera, J. A.; Bolds, J.; Shields, P. E.; Havel, C. M.; Watson, J. A. Isoprenoid synthesis in Halobacterium halobium. Modulation of 3-hydroxy-3-methylglutaryl coenzyme a concentration in response to mevalonate availability. *J. Biol. Chem.* **1986**, *261*, 3578−3583.
(48) Lam, W. L.; Doolittle, W. F. Shuttle vectors for the archaebacterium Halobacterium volcanii. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5478−5482.
(49) Lorenz, R. T.; Parks, L. W. Effects of lovastatin (mevinolin) on sterol levels and on activity of azoles in *Saccharomyces cerevisiae*. *Antimicrob. Agents Chemother*. **1990**, *34*, 1660−1665.
(50) Sud, I. J.; Feingold, D. S. Effect of ketoconazole in combination with other inhibitors of sterol synthesis on fungal growth. *Antimicrob. Agents Chemother*. **1985**, *28*, 532−534.
(51) Bejarano, E. R.; Cerda-Olmedo, E. Independence of the carotene and sterol pathways of Phycomyces. *FEBS Lett.* **1992**, *306*, 209−212.
(52) Engstrom, W.; Larsson, O.; Sachsenmaier, W. The effects of tunicamycin, mevinolin and mevalonic acid on HMG-CoA reductase activity and nuclear division in the myxomycete Physarum polycephalum. *J. Cell Sci*. **1989**, *92*, 341−344.
(53) Florin-Christensen, M.; Florin-Christensen, J.; Garin, C.; Isola, E.; Brenner, R. R.; Rasmussen, L. Inhibition of *Trypanosoma cruzi* growth and sterol biosynthesis by lovastatin. *Biochem. Biophys. Res. Commun*. **1990**, *166*, 1441−1445.
(54) Urbina, J. A.; Lazardi, K.; Marchan, E.; Visbal, G.; Aguirre, T.; Piras, M. M.; Piras, R.; Maldonado, R. A.; Payares, G.; de Souza, W. Mevinolin (lovastatin) potentiates the antiproliferative effects of ketoconazole and terbinafine against Trypanosoma (Schizotrypanum) cruzi: in vitro and in vivo studies. *Antimicrob. Agents Chemother*. **1993**, *37*, 580−591.
(55) Vandewaa, E. A.; Mills, G.; Chen, G. Z.; Foster, L. A.; Bennett, J. L. Physiological role of HMG-CoA reductase in regulating egg production by *Schistosoma mansoni*. *Am. J. Physiol.* **1989**, *257*, R618−625.
(56) Chen, G. Z.; Foster, L.; Bennett, J. L. Antischistosomal action of mevinolin: evidence that 3-hydroxy-methylglutaryl-coenzyme a reductase activity in *Schistosoma mansoni* is vital for parasite survival. *Naunyn-Schmiedeberg's Arch. Pharmacol*. **1990**, *342*, 477−482.
(57) Grellier, P.; Valentin, A.; Millerioux, V.; Schrevel, J.; Rigomier, D. 3-Hydroxy-3-methylglutaryl coenzyme A reductase inhibitors lovastatin and simvastatin inhibit in vitro development of Plasmodium falciparum and Babesia divergens in human erythrocytes. *Antimicrob. Agents Chemother*. **1994**, *38*, 1144−1148.
(58) Coppens, I.; Courtoy, P. J. Exogenous and endogenous sources of sterols in the culture-adapted procyclic trypomastigotes of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **1995**, *73*, 179−188.
(59) Ikeura, R.; Murakawa, S.; Endo, A. Growth inhibition of yeast by compactin (ML-236B) analogues. *J. Antibiot. (Tokyo)* **1988**, *41*, 1148−1150.
(60) Fernandez, M. A.; Garcia, M. D.; Saenz, M. T. Antibacterial activity of the phenolic acids fractions of Scrophularia frutescens and Scrophularia sambucifolia. *J. Ethnopharmacol.* **1996**, *26*, 11−4.
(61) Friedman, M.; Henika, P. R.; Mandrell, R. E. Antibacterial Activities of Phenolic Benzaldehydes and Benzoic Acids against Campylobacter jejuni, *Escherichia coli*, Listeria monocytogenes, and Salmonella enterica. *J. Food Protect.* **2003**, *66*, 1811−1821.
(62) Terekhova, L. P.; Galatenko, O. A.; Preobrazhenskaia, T. P. Sensitivity of actinomycetes of the genus Actinomadura to antibiotics. *Antibiotiki* **1981**, *26*, 345−349.
(63) Gol'dberg, L. E. Antibacterial and antineoplastic antibiotics. *Sov. Med.* **1977**, *10*, 115−123.
(64) Berlin, Y. A.; Esipov, S. E.; Kiseleva, O. A.; Kolosov, M. N. Olivomycin and related antibiotics X. Isolation and acid degradation of olivomycins A, B, C, and D. *Chem. Nat. Compd.* **1968**, *3*, 280−285.
(65) The NIAID database: http://chemdb.niaid.nih.gov.

CI050480J