MDPI

Article

New Pharmacokinetic and Microbiological Prediction Equations to Be Used as Models for the Search of Antibacterial Drugs

Jose I. Bueso-Bordils [1],*, Gerardo M. Antón-Fos [1], Antonio Falcó [2,3], Maria J. Duart [1], Rafael Martín-Algarra [1] and Pedro A. Alemán-López [1]

1   Departamento de Farmacia, Universidad Cardenal Herrera-CEU, CEU Universities C/Ramón y Cajal s/n, Alfara del Patriarca, 46115 Valencia, Spain; ganton@uchceu.es (G.M.A.-F.); mduart@uchceu.es (M.J.D.); rmartin@uchceu.es (R.M.-A.); paleman@uchceu.es (P.A.A.-L.)
2   ESI International Chair@CEU-UCH, Departamento de Matemáticas, Física y Ciencias Tecnológicas, Universidad Cardenal Herrera-CEU, CEU Universities San Bartolomé 55, Alfara del Patriarca, 46115 Valencia, Spain; afalco@uchceu.es
3   Departamento de Matemáticas, Física y Ciencias Tecnológicas, Universidad CEU Cardenal Herrera, CEU Universities, San Bartolomé 55, 46115 Alfara del Patriarca, Spain
*   Correspondence: jose.bueso@uchceu.es; Tel.: +34-96-136-9000

Abstract: Currently, the development of resistance of *Enterobacteriaceae* bacteria is one of the most important health problems worldwide. Consequently, there is a growing urge for finding new compounds with antibacterial activity. Furthermore, it is very important to find antibacterial compounds with a good pharmacokinetic profile too, which will lead to more efficient and safer drugs. In this work, we have mathematically described a series of antibacterial quinolones by means of molecular topology. We have used molecular descriptors and related them to various pharmacological properties by using multilinear regression (MLR) analysis. The regression functions selected by presenting the best combination of a number of quality and validation metrics allowed for the reliable prediction of clearance (CL), and minimum inhibitory concentration 50 against *Enterobacter aerogenes* (MIC50Ea) and *Proteus mirabilis* (MIC50Pm). The obtained results clearly reveal that the combination of molecular topology methods and MLR provides an excellent tool for the prediction of pharmacokinetic properties and microbiological activities in both new and existing compounds with different pharmacological activities.

Keywords: pharmacokinetics; antibacterial; multilinear regression (MLR); computational chemistry; molecular topology; molecular connectivity; topological indices; quinolones; QSAR

1. Introduction

The increasing emergence of resistance to known antibiotics is one of the most important problems that have appeared in recent years in the treatment of infectious diseases nowadays [1]. The emergence of resistance is associated not only with complications from infections in terms of morbidity and mortality, but also represents an increase in healthcare costs [2]. In this context, rapid and inexpensive methods are necessary for the immediate expansion of the therapeutic arsenal that we currently have [3].

Traditionally, in order to develop a drug, it was necessary to synthesize and evaluate the activity of hundreds to thousands of compounds with the intention of proving their biological activity, selectivity, and bioavailability, as well as their low toxicity. On average, this process of discovering new drugs involves, in addition to a high economic cost, several years of duration before, hopefully, finding a drug with the right characteristics to boost its commercialization [4].

Virtual screening along with combinatorial chemistry appear as a solution to this problem by allowing researchers to identify molecules that are likely to be active from a virtual

chemical library [5]. Given that these virtual libraries usually have many components, it is common to apply filters to discard molecules with a low probability of becoming a drug. In this sense, Lipinski was the first to apply these filters when describing the "Rule of Five" [6]. This standard predicts that a compound is "non-pharmacological" if it has too many hydrogen bond donors or acceptors (5 and 10 respectively), its molecular mass is greater than 500 and its calculated partition coefficient (ClogP) is greater than 5. With these rules, Lipinski created the first pre-filter prior to drug screening [7].

This rule has been revised based on data obtained from studies in rats [8]. Since then, other rules have appeared and been modified, such as the Rule of Three [9], used in "fragment-based" discovery, which classifies fragments as "drug-like" those with an average molecular weight ≤ 300 Da, a calculated partition coefficient (ClogP) ≤ 3, ≤3 hydrogen bond donors, ≤3 hydrogen bond acceptors, and <3 rotatable bonds. Another similar example is the "Rule of 3/75", which was described after analyzing 245 preclinical Pfizer compounds. This rule states that a compound will be 2.5 times safer in in vivo assays when its ClogP is <3 and the topological polar surface area (TPSA) > 75 [10].

Within drug screening, it is also necessary to highlight the rise of QSAR (Quantitative Structure-Activity Relationship) methods, which have been the foundation of the design and development of new molecules with pharmacological activities [11]. QSAR modelling is a technique that allows the interdisciplinary exploration of knowledge in compounds covering chemical, physical, biological, and toxicological aspects. It also provides formalisms for the mathematical development of models based on chemical characteristics and the activity of structurally similar compounds. This context is defined by mathematical algorithms and provides a reasonable foundation for creating a prediction model [12].

The use of QSAR methods for obtaining and developing new drugs is a key tool since it provides a noteworthy saving of time and resources, considering that many pharmacological activities of compounds can be predicted without having been synthesized by virtue of its rational approach [13,14].

There are numerous examples of drug designed molecules with very different pharmacological activities which have followed the application of QSAR techniques in combination with other in-silico methods, such as zanamivir, tirofiban, imatinib, raltegravir, donepezil, boceprevir, norfloxacin, captopril and dorzolamide, all of them approved by the US Food and Drug Administration [15].

Our research group uses molecular connectivity or molecular topology (MT), developed by Kier and Hall in the mid-1970s [16], a method derived from the QSAR techniques, which is based on the calculation of molecular descriptors or topological indices (TIs) from the structure of a molecule. These methods play an important role, since they provide useful information for the rational design of new molecules with minimal cost and are an alternative and effective tool in view of the reduction of investments by the pharmaceutical industry in the area of antibiotics [17,18].

By using MT, we can classify a compound as active or not active against a certain pharmacological activity using pattern recognition techniques such as multilinear regression (MLR), discriminant linear analysis, neuronal networks, random forests, support vector machines or principal component analysis [19].

All Gram-negative bacilli of the *Enterobacteriaceae* family, are resistant to common antimicrobials, such as penicillin, oxacillin, methicillin, macrolides, glycopeptides, linezolid, and clindamycin, among others [20]. Moreover, the World Health Organization (WHO) on its "WHO PRIORITY PATHOGENS LIST FOR R&D OF NEW ANTIBIOTICS" considers this family a critical or first priority pathogen [3]. Both bacteria studied in this work, *Proteus mirabilis* and *Enterobacter aerogenes*, are Gram-negative bacilli that belong to the *Enterobacteriaceae* family, and are an important cause of community and nosocomial infections, representing a major clinical and public health challenge due to the limited therapeutic arsenal and well-established resistances [21,22].

In this study, by combining MT and MLR after extending a thorough bibliographic research on physicochemical and biological properties on antibacterial quinolones from a

previous study [23], we developed three linear regression functions capable of predicting clearance and minimum inhibitory concentration 50 against *Enterobacter aerogenes* and *Proteus mirabilis*. We selected quinolones to build these equations, which can be applied to other quinolones not used in the study, as well as to molecules that share no structural relationship, seeing as the selection of hits or leads considers mathematical-topological similarity patterns, despite the possible core-structure diversity [24]. Thus, the obtained functions can be used as filters in the search process of new antibacterials in large databases of molecules. Examples of some commonly used chemical libraries where new molecules could easily be found on which to apply these techniques are ZINC, Benchmark DUD, PubChem, ChemBank, ChEMBL, DrugBank and Inter-bioscreen [15].

The first quinolone to be used clinically was discovered in 1962 by G. Lesher et al., who, during the purification of the antimalarial chloroquine from mother liquor, found 7-chloroquinoline, a by-product with antimicrobial activity. Starting from this lead, nalidixic acid was obtained. Despite the good tolerance and ease of synthesis of this compound, its use has been restricted to the treatment of some urinary infections due to its tendency to select resistance and to concentrate mainly in urine. From this discovery, derivatives with similar structures were developed, with still moderate activity against Gram-negative bacteria, but which were very valuable in uncomplicated urinary infections [25].

Pipemidic acid appeared in 1975, the first quinolone with a piperazine ring, with favorable activity, spectrum, and pharmacokinetics, which allowed the administration of lower doses. The appearance of bacterial resistance during treatment and side effects were also lower [26]. A year later, flumequin was discovered, which, by incorporating a fluorine atom in position 6 of the core double ring, was able to amplify its spectrum within *Enterobacteriaceae*, including some resistant strains and slightly improve the activity against Gram-positive bacteria [27].

These facts motivated the beginning of an active chemical synthesis campaign to refine the structure-activity relationships. The aim was to improve activity while optimizing the pharmacokinetic properties and reducing the toxicity and interactions with other drugs. This led to the consolidation of the second-generation quinolones, which differ from the classical quinolones in two characteristics common to all of them: the presence of a fluorine atom in position 6 and a substituted piperazine or pyrrolidine ring at position 7 of their core structure [28].

The first second-generation quinolone to be patented, in 1978, was norfloxacin, whose improved potency against Gram-negative bacteria was already in the natural antibiotics' range, in addition to having a longer half-life (3–4 h) and shorter protein binding (50%) than its predecessors [29]. Shortly after, ciprofloxacin was introduced, being for many years the most potent fluoroquinolone against Gram-negative bacteria and the most widely used quinolone [30]. This generation of quinolones has a much higher activity against aerobic Gram negatives and some important Gram-positive pathogens, as well as more favourable toxicological profiles and good absorption through the gastrointestinal tract [28].

In the 1990s, several quinolones, mostly di- or tri-fluorinated, were approved. They are the third-generation quinolones, which have a greater structural complexity than their predecessors. One of the first modifications was the introduction of an amino group at position 5, which caused an increase in the general activity against Gram-positive bacteria [28]. All of them have advantages in their bioavailability compared to second-generation quinolones, since the substituted rings at position 7 give them greater lipophilicity, resulting in longer elimination half-lives and greater tissue penetration [31].

In recent years, a new class of quinolones has been synthesized, since the optimization of the substituents in different positions of the core structure has allowed the suppression of the fluorine atom in position 6, thus avoiding possible toxic effects, without necessarily reducing its activity [32]. Probably, the most prominent examples of these novel des-fluoro(6) quinolones are garenoxacin (initially known as BMS-284756) and DX-619, which due to their high affinity for type II topoisomerases, have shown high intrinsic activity

against most Gram-negative, Gram-positive and anaerobic microorganisms with a low resistance selection potential [33,34].

The good pharmacokinetic profile of quinolones, as well as their high antibacterial activity and broad spectrum of action, makes their therapeutic indications very varied. In fact, quinolones are used in the treatment of many infections, both in the hospital setting and in outpatient clinical practice. More than ten thousand quinolones have been registered. Therefore, it has probably been the most widely used thoroughly studied class of chemotherapeutic agents [35,36]. For all these reasons, we thought that quinolones were the most suitable antibacterial family to be the focus of our study.

QSAR models have been widely attempted, using homologous and congeneric series of active compounds, obtaining excellent results for the correlation of fluorescent [37], antitumor [38], anti-HIV [39], antifungal [40] and antibacterial [41,42] properties, amongst others.

2. Results

We performed Y-randomization, inter-correlation, and leave-one-out (LOO) cross-validation tests on all equations. On the one hand, a representation of correlation coefficient (r^2) versus cross-validated correlation coefficient (r^2_{cv}) could show the presence of chance correlations. On the other hand, compounds that do not fit the selected function could be easily spotted by a representation of the Y-randomization analysis.

Hence, we rejected those equations with highly dependent indices (correlation coefficient greater than 0.8 or smaller than -0.8 between indices), suspected randomness (r^2_{vc} smaller than 0.5 for the selected equation) or outliers in the compound set (dissimilar residual values between the selected and the cross-validated equation for one of the compounds, which is indicative of instability).

We also considered adjusted linear coefficient of determination (r^2_a), which is represented by:

$$r^2_a = r^2 - (p(1 - r^2)/(N - p'))$$

where N is the number of observations; and p is the number of dependent variables selected in the RF ($p' = p$ when the y-axis intersect is zero and $p' = p + 1$ when it is different from zero). The correlation will be better the closer this value gets to unity.

Therefore, we finally selected three descriptor-independent regression functions (RFs) with high predictive capacity and significant statistical parameters:

$RF_{CL} = 8981.08 + 1596.68^6\chi_{ch} - 109.481S_{>N\text{-}} - 2637.68NI_2$

$n = 12$	$r = 0.95939$	$F = 30.85$	$p = 0.0001$	(1)
$SEE = 65.42953$	$r^2 = 0.92043$	$r^2_a = 0.89059$	$r^2_{cv} = 0.79791$	

$RF_{MIC50Ea} = 0.410929 - 0.123037S_{\text{-}CH2\text{-}} + 0.0685818S_{aCa} + 0.0569394S_{\text{-}NH\text{-}}$

$n = 13$	$r = 0.95930$	$F = 34.62$	$p = 0.0000$	(2)
$SEE = 0.033071$	$r^2 = 0.92025$	$r^2_a = 0.89367$	$r^2_{cv} = 0.83270$	

$RF_{MIC50Pm} = 0.376001 + 27.6676^9\chi^V_{ch} + 0.0650158S_{\text{-}CH3} - 0.0889109S_{>N\text{-}}$

$n = 16$	$r = 0.91837$	$F = 21.54$	$p = 0.0000$	(3)
$SEE = 0.076808$	$r^2 = 0.84340$	$r^2_a = 0.80425$	$r^2_{cv} = 0.70282$	

In Tables 1–3 we can find all the important information regarding the RFs for each of the compounds: the value of each independent variable or TI; the dependent variable or experimental value for each property; the predicted value for each property; and the residual values (difference between experimental and predicted value).

We can note that one of the calculated values is negative (Clinafloxacin, Table 2) and, of course, this is an impossible value for a MIC value. This fact can happen in compounds with a dependent variable's value very close to zero, since the linear regression function may calculate a predicted value for that compound just below the x-axis, i.e., just below zero.

Table 1. Prediction results obtained for the CL property (mL/min).

Compound	$^6\chi_{ch}$	$S_{>N-}$	NI_2	ExpCL [a]	CalcCL [b]	Res [c]
Moxifloxacin	0.1741	3.803	3.319	14.8	88.2486	−73.4486
Amifloxacin	0.1842	5.5294	3.262	31.0	65.715	−34.715
Garenoxacin	0.169	1.6095	3.408	86.1	85.4983	0.6017
Pefloxacin	0.1842	5.8196	3.262	106.716	33.9436	72.7724
Trovafloxacin	0.2245	2.731	3.392	106.8132	93.5342	36.2154
Fleroxacin	0.1741	4.5829	3.265	127.126	145.2989	23.0078
Levofloxacin	0.2296	5.8165	3.219	198.0	220.1923	38.148
Temafloxacin	0.2523	2.9448	3.352	206.6	220.0219	38.5829
Sitafloxacin	0.0907	3.1603	3.261	263.0	178.4345	42.066
Clinafloxacin	0.0907	3.4961	3.146	368.857	445.0039	−76.1469
Enoxacin	0.203	3.3811	3.166	562.875	584.1476	−21.2727
Ciprofloxacin	0.203	3.7812	3.17	617.945	529.7932	88.1518

[a] Experimental value of the CL property for each compound. [b] Calculated value of the CL property for each compound. [c] Residual value of RF_{CL} for each compound.

Table 2. Prediction results obtained for the MIC50Ea property (µg/mL).

Compound	S_{-CH2}	S_{aCa}	S_{-NH-}	ExpMIC50Ea [a]	CalcMIC50Ea [b]	Res [c]
Clinafloxacin	3.5533	0.1483	0	0.019	−0.0161	0.0351
Ciprofloxacin	4.8443	0.7434	3.2252	0.0351	0.0495	−0.0145
Tosufloxacin	1.5749	−2.6064	0	0.045	0.0384	0.0066
Gemifloxacin	2.7854	−0.3502	0	0.0465	0.0442	0.0023
Trovafloxacin	1.0072	−3.2996	0	0.048	0.0607	−0.0127
Levofloxacin	3.2033	0.6381	0	0.06	0.0606	−0.0006
Sparfloxacin	2.2119	−3.1814	3.3	0.084	0.1085	−0.0245
Fleroxacin	0.7962	−2.8908	0	0.09	0.1147	−0.0247
Norfloxacin	3.4104	0.6027	3.2081	0.1633	0.2153	−0.052
Enoxacin	3.1429	−0.1195	3.1803	0.2083	0.1971	0.0112
Amifloxacin	3.0361	0.3878	2.8144	0.25	0.2242	0.0258
Lomefloxacin	1.7121	−2.1985	3.2023	0.29	0.2318	0.0582
WIN-57273	1.8278	1.6636	0	0.29	0.3001	−0.0101

[a] Experimental value of the MIC50Ea property for each compound. [b] Calculated value of the MIC50Ea property for each compound. [c] Residual value of $RF_{MIC50Ea}$ for each compound.

Table 3. Prediction results obtained for the MIC50Pm property (µg/mL).

Compound	$^9\chi^V_{ch}$	S_{-CH3}	$S_{>N-}$	ExpMIC50Pm [a]	CalcMIC50Pm [b]	Res [c]
Clinafloxacin	0	0	3.4961	0.0225	0.0652	−0.0427
Sitafloxacin	0	0	3.1603	0.03	0.095	−0.065
Ciprofloxacin	0	0	3.7812	0.0426	0.0398	0.0028
Levofloxacin	0	3.8877	5.8165	0.06	0.1116	−0.0516
Flerofloxacin	0	1.8996	4.5829	0.12	0.092	0.028
Tosufloxacin	0	0	2.8277	0.12	0.1246	−0.0046
Gemifloxacin	0	1.437	3.3757	0.165	0.1693	−0.0043
Gatifloxacin	0	3.4792	3.6897	0.198	0.2741	−0.0761
Trovafloxacin	0	0	2.731	0.199	0.1332	0.0658
Amifloxacin	0	3.6197	5.5294	0.25	0.1197	0.1303
Enoxacin	0	1.8106	3.3811	0.25	0.1931	0.0569
Sparfloxacin	0	3.8376	3.0307	0.25	0.356	−0.1060
Lomefloxacin	0	3.6201	2.9493	0.29	0.3491	−0.0591
Garenoxacin	0.0049	2.0568	1.6095	0.5	0.5022	−0.0022
Moxifloxacin	0.0132	1.4781	3.803	0.5	0.4992	0.0008
WIN-57273	0	3.684	1.7957	0.583	0.4559	0.1271

[a] Experimental value of the MIC50Pm property for each compound. [b] Calculated value of the MIC50Pm property for each compound. [c] Residual value of $RF_{MIC50Pm}$ for each compound.

3. Discussion

Only those functions with a significance greater than 99.99% ($p \leq 0.0001$) were selected. The incorporation of more than three independent variables was studied in order to improve the statistical parameters but this was not achieved with any of the properties studied. Finally, the best statistical data were obtained with the presence of three independent

variables in each function, which indicates that all the indices provide important structural information to quantify each of the properties studied.

In all cases, the residual value for each compound was less than twice the standard error of estimation ($\pm$2SEE), indicating that the selected equations have high predictive ability.

The cross-validated correlation coefficient (r^2_{cv}) was determined to study Y-randomization. A function is considered to be non-randomly obtained when its $r^2_{cv} > 0.5$ [15]. In the three selected equations, the value of this coefficient was higher than 0.7028, which explains more than 70.28% of the variance of the group. The probability of finding a random sequence that fits the prediction model is very low since in the three selected equations they had significantly higher r^2_{cv} values than the non-selected equations (Supplementary Section S1). All these data guarantee the robustness of the selected equations.

For the clearance (CL) property, a connectivity index [43], which considers sixth order chain type graphs ($^6\chi_{ch}$), increases its value. This means that the presence of 6-member cycles increases this property. One the other hand, an electro-topological index ($S_{>N-}$) [44] and an information-theory index (NI_2) [45] have a negative effect on the RF's value. The first one increases with a high number of tertiary amine groups, so the presence of this group decreases the CL value. The latter can be applied to a molecule to quantify its atomic diversity [46].

For the minimum inhibitory concentration 50 against *Enterobacter aerogenes* (MIC50Ea), two electro-topological indices increase its RF's value. The first one (S_{aCa}) increases with a high number of aromatic carbons, while the second one (S_{-NH-}) increases with a high number of secondary amine groups. Thus, it should be noted that the presence of aromatic carbons and/or secondary amine groups have a negative effect on the activity against *E. aerogenes*, since the property in this case is MIC. On the contrary, another electrotopological index (S_{-CH2-}), which increases with an increasing number of methylene groups, helps decreasing the RF's value and, hence, improves the activity against *E. aerogenes*.

For the minimum inhibitory concentration 50 against *Proteus mirabilis* (MIC50Pm), a connectivity index and an electro-topological index increase its RF's value. The valence-connectivity index [47] considers ninth order chain type graphs ($^9\chi^V_{ch}$), present in garenoxacin and moxifloxacin, which present weak activity against *P. mirabilis* compared to most of the compounds in that data set. This is consistent with the structure-activity relationships established by Domagala et al. [48], which state that a 9-member bicycle at the quinolone's R7 position enhances activity against Gram-positive bacteria but is detrimental for Gram-negative bacteria such as *P. mirabilis*. Since this bicycle is present in only two compounds, the possibility of developing the RF ignoring this index was studied, but in all the attempts, the statistical consistence significantly decreased. The connectivity index's value (S_{-CH3}) increases with an increasing number of methyl groups, so we could state that an excessive amount of methyl groups in the quinolones' structure decreases the activity against *P. mirabilis*. In fact, most of the most active compounds from this data set lack this functional group, i.e., have a zero value for this index (see Table 3 for the MIC, $^9\chi^V_{ch}$ and S_{-CH3} values and Supplementary Section S2 for the quinolones' chemical structures). Nevertheless, an electro-topological index decreases the RF's value, i.e., enhances activity against *P. mirabilis*. Such is the case of $S_{>N-}$, which increases with a high number of tertiary amine groups. Therefore, we could state that a good proportion of tertiary amine groups in the quinolones' structure enhances the activity against *P. mirabilis*.

4. Materials and Methods

4.1. Compound Selection

Data were collected for a total of 85 physicochemical, pharmacokinetic, pharmacodynamic and microbiological properties of antibacterial quinolones (see Supplementary Section S2).

Experimental procedures for the obtention of experimental values of each property were studied. To select a property, it had to have been obtained with comparable experimental procedures for at least 10 compounds. Sufficient data for a total of 27 microbiological

properties (activities against different microorganisms measured as different MICs) were obtained from in vitro assays performed according to CLSI protocols [49] and 9 pharmacokinetic properties obtained in assays performed in healthy subjects.

Due to the diversity of bibliographical data found regarding a given property for each of the selected molecules, we considered the average of those data whose values were of the same order of magnitude for the study and development of the prediction equations.

4.2. Topological Descriptors

A total of 41 different quinolones were topologically characterized using DESMOL13 [50] and MOLCONN-Z [51] programs. We removed indices with value 0 for every compound and with identical values for all compounds, leaving a total of 140 topological descriptors specific to each molecule (see Supplementary Section S3). These descriptors were computed from the adjacency topological matrix obtained from the hydrogen-depleted chemical pseudographs, drawn with the ChemBioDraw Ultra 12.0 drawing program of the ChemBioOffice 2010 program package. The molecular descriptors used are described in Supplementary Section S4 along with their definitions and references.

4.3. Multilinear Regression

We used the statistical package BMDP module 9R [52] to perform the MLR analysis. This program is based on the Furnival-Wilson algorithm [53], which uses Mallow's Cp parameter [54] to assess how well a set of observations fit a given RF. Its purpose is to find the best combination of indices from all the possible ones to generate this RF. This statistical parameter can be depicted as:

$$Cp = PRESS/s^2 - (N - 2p)$$

where *PRESS* is the squared sum of the residual values of the dependent variables; s is the residual mean obtained in the RF; p is the number of dependent variables selected in the RF; and N is the number of observations.

In principle, the equation with the lowest Mallow's Cp value will be the optimal one. However, the residual sum of squares always decreases as dependent variables are added to the equation. Consequently, we will always select the equation with the highest number of indices if we aim for the minimum residual sum values.

However, we must always bear in mind that the predictions made from a linear regression function are not reliable if there is not a high degree of linear correlation between the independent and dependent variables. In addition, its statistical quality must be analyzed and the higher it is, the more reliable and accurate the prediction will be.

It is also recommended that the number of independent variables chosen be well below the number of observations, to avoid possible fortuitous correlations, so that the optimal relationship between them is the one that achieves a better prediction with the minimum number of descriptors possible, in order to avoid random variation as much as possible [55]. Randomness tests were also carried out to identify possible fortuitous regressions. To do this, the property value of each compound is randomly permuted and linearly correlated with the selected independent variables. This process is repeated as many times as necessary. The result of the randomness test is expressed by graphically representing the correlation coefficient (r^2) versus the prediction coefficient (r^2_{cv}).

For all those reasons, the number of observations (size of the set), the effect of the different independent variables and the degree of correlation between them will influence the selection of RFs under the criterion of this algorithm.

5. Conclusions

The increase of resistant *Enterobacteriaceae* bacteria such as *E. aerogenes* and *P. mirabilis* is one of the major challenges that have appeared in recent years in infectious disease treatments. Therefore, the urge to find new compounds with antibacterial activity has not

ceased to grow. It must also be noted that the pharmacokinetic profile of a compound is essential in determining the clinical outcome of an antibiotic treatment.

We have designed and developed three statistically significant, descriptor-independent regression functions with a high predictive capacity, which will aid us obtain theoretical values for the pharmacokinetic and microbiological properties selected. The results of the predictions obtained are supported by internal validation studies performed on all selected properties, as well as by all the statistical parameters, which can be classified as very satisfactory in all cases.

These equations, used as predictive models or as filters, can be extremely useful to find new molecules with antibacterial activity and even for their use in drug repositioning, where we find thousands of candidate compounds, which will have most certainly endured toxicological assays, to be screened in a short time for the search of safe antibacterial compounds.

We can conclude that the prediction functions obtained confirm molecular topology as a useful and efficient tool in the prediction of pharmacokinetic and microbiological properties.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ph15020122/s1. Supplementary Section S1: Internal validation results for each of the selected equations; Supplementary Section S2: Compounds used in the realization of the equations: molecule name/code, IUPAC name & structure and bibliographic references about activity for each compound; Supplementary Section S3: Microsoft Excel file containing all the significant and non-redundant topological indices for the selected compounds; Supplementary Section S4: Symbols and Definitions of Topological Indices used with DESMOL13 and MOLCONN-Z programs.

References

1. Tacconelli, E.; Magrini, N. *Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics*; World Health Organization: Geneva, Switzerland, 2017.
2. Archibald, L.K.; Jarvis, W.R. Health care-associated infection outbreak investigations by the centers for disease control and prevention, 1946–2005. *Am. J. Epidemiol.* **2011**, *174*, S47–S64. [CrossRef] [PubMed]
3. World Health Organization. WHO Priority Pathogens List for R&D of New Antibiotics. Available online: http://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-ET_NM_WHO.pdf?ua=1 (accessed on 24 October 2021).
4. Garrett, M.D.; Workman, P. Discovering novel chemotherapeutic drugs for the third millennium. *Eur. J. Cancer* **1999**, *35*, 2010–2030. [CrossRef]
5. Morris, G.M.; Lim-Wilby, M. Molecular Docking. *Methods Mol. Biol.* **2008**, *443*, 365–382.
6. Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26. [CrossRef]
7. Lipinski, C.A. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov. Today Technol.* **2004**, *1*, 337–341. [CrossRef] [PubMed]
8. Ridder, L.; Wang, H.; de Vlieg, J.; Wagener, M. Revisiting the rule of five on the basis of pharmacokinetic data from rat. *ChemMedChem* **2011**, *6*, 1967–1970. [CrossRef]

9. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* **2003**, *8*, 876–877. [CrossRef]

10. Hughes, J.D.; Blagg, J.; Price, D.A.; Bailey, S.; Decrescenzo, G.A.; Devraj, R.V.; Ellsworth, E.; Fobian, Y.M.; Gibbs, M.E.; Gilles, R.W.; et al. Physiochemical drug properties associated with in vivo toxicological outcomes. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4872–4875. [CrossRef]

11. Karelson, M. *Molecular Descriptors in QSAR/QSPR*; Wiley: New York, NY, USA, 2000.

12. Todeschini, R.; Consonni, V. *Molecular Descriptors for Cheminformatics*; Wiley: Weinheim, Germany, 2009.

13. Batool, M.; Ahmad, B.; Choi, S. A structure-based drug discovery paradigm. *Int. J. Mol. Sci.* **2019**, *20*, 2783. [CrossRef]

14. Tute, M.S. History and objectives of quantitative drug design. In *Comprehensive Medicinal Chemistry*; Hansch, C., Sammes, P.G., Taylor, J.B., Eds.; Pergamon Press: Oxford, UK, 2015; Volume 4, pp. 1–31.

15. Roy, K.; Kar, S.; Das, R.N. *A Primer on QSAR/QSPR Modeling: Fundamental Concepts*; Springer: Berlin/Heidelberg, Germany, 2015.

16. Kier, L.B.; Hall, L.H. *Molecular Connectivity in Chemistry and Drug Research*, 1st ed.; Academic Press: New York, NY, USA, 1976; pp. 41–80.

17. Zheng, W.; Sun, W.; Simeonov, A. Drug repurposing screens and synergistic drug-combinations for infectious diseases. *Br. J. Pharmacol.* **2017**, *175*, 181–191. [CrossRef]

18. Bennett, J.W. What is an antibiotic? In *Antibiotics: Current Innovations and Future Trends*; Sánchez, S., Demain, A.L., Eds.; Caister Academic Press: Poole, UK, 2015; Chapter 1; pp. 1–18.

19. Suay-Garcia, B.; Bueso-Bordils, J.I.; Falcó, A.; Pérez-Gracia, M.T.; Antón-Fos, G.; Alemán-López, P. Quantitative structure–activity relationship methods in the discovery and development of antibacterials. *WIREs Comput. Mol. Sci.* **2020**, *10*, e1472. [CrossRef]

20. Von Graevenitz, A. The use of antimicrobial agents as tools in epidemiology, identification and selection of microorganisms. In *Antibiotics in Laboratory Medicine*; Victor Lorian, M.D., Ed.; Williams and Wilkins: Baltimore, MD, USA, 1991; pp. 810–824.

21. Chen, C.Y.; Chen, Y.H.; Lu, P.L.; Lin, W.R.; Chen, T.C.; Lin, C.Y. Proteus mirabilis urinary tract infection and bacteremia: Risk factors, clinical presentation, and outcomes. *J. Microbiol. Immunol. Infect.* **2012**, *45*, 228–236. [CrossRef]

22. Davin-Regli, A.; Lavigne, J.P.; Pagès, J.M. Enterobacter spp.: Update on taxonomy, clinical aspects, and emerging antimicrobial resistance. *Clin. Microbiol. Rev.* **2019**, *32*, e00002. [CrossRef]

23. Bueso-Bordils, J.I.; Aleman-López, P.A.; Costa-Piles, S.; Duart, M.J.; Lahuerta-Zamora, L.; Martin-Algarra, R.; Anton-Fos, G.M. Obtaining microbiological and pharmacokinetic highly predictive equations. *Curr. Top. Med. Chem.* **2018**, *18*, 908–916. [CrossRef]

24. Suay-Garcia, B.; Falcó, A.; Bueso-Bordils, J.I.; Anton-Fos, G.M.; Pérez-Gracia, M.T.; Alemán-López, P.A. Tree-based QSAR model for drug repurposing in the discovery of new antibacterial compounds against Escherichia coli. *Pharmaceuticals* **2020**, *13*, 431. [CrossRef]

25. Mitscher, L.A. Bacterial topoisomerase inhibitors: Quinolone and pyridone antibacterial agents. *Chem. Rev.* **2005**, *105*, 559–592. [CrossRef]

26. Shimizu, M.; Nakamura, S.; Takase, Y.; Kurobe, N. Pipemidic acid: Absorption, distribution, and excretion. *Antimicrob. Agents Chemother.* **1975**, *7*, 441–446. [CrossRef]

27. Rohlfing, S.R.; Gerster, J.F.; Kvam, D.C. Bioevaluation of the antibacterial flumequine for urinary tract use. *Antimicrob. Agents Chemother.* **1976**, *10*, 20–24. [CrossRef]

28. Emami, S.; Shafiee, A.; Foroumadi, A. Structural features of new quinolones and relationship to antibacterial activity against gram-positive bacteria. *Mini Rev. Med. Chem.* **2006**, *6*, 375–386. [CrossRef] [PubMed]

29. Emmerson, A.M.; Jones, A.M. The quinolones: Decades of development and use. *J. Antimicrob. Chemother.* **2003**, *51*, 13–20. [CrossRef] [PubMed]

30. Van Bambeke, F.; Michot, J.M.; Van Eldere, J.; Tulkens, P.M. Quinolones in 2005: An update. *Clin. Microbiol. Infect.* **2005**, *11*, 256–280. [CrossRef]

31. O'Donnell, J.A.; Gelone, S.P. The newer fluoroquinolones. *Infect. Dis. Clin. N. Am.* **2004**, *18*, 691–716. [CrossRef]

32. Boteva, A.A.; Krasnykh, O.P. The methods of synthesis, modification, and biological activity of 4-quinolones. *Chem. Heterocycl. Compd.* **2009**, *45*, 757–785. [CrossRef]

33. Fung-Tomc, J.C.; Minassian, B.; Kolek, B.; Huczko, E.; Aleksunes, L.; Stickle, T.; Washo, T.; Gradelski, E.; Valera, L.; Bonner, D.P. Antibacterial spectrum of a novel des-fluoro (6) quinolone, BMS-284756. *Antimicrob. Agents Chemother* **2000**, *44*, 3351–3356. [CrossRef]

34. Fujikawa, K.; Chiba, M.; Tanaka, M.; Sato, K. In vitro antibacterial activity of DX-619, a novel des-fluoro (6) quinolone. *Antimicrob. Agents Chemother.* **2005**, *49*, 3040–3045. [CrossRef] [PubMed]

35. Rotschafer, J.C.; Ullman, M.A.; Sullivan, C.J. Optimal use of fluoroquinolones in the intensive care unit setting. *Crit. Care Clin.* **2011**, *27*, 95–106. [CrossRef]

36. Andersson, M.I.; MacGowan, A.P. Development of the quinolones. *J. Antimicrob. Chemother.* **2003**, *51*, 1–11. [CrossRef]

37. López-Malo, D.; Bueso-Bordils, J.I.; Duart, M.J.; Alemán-López, P.A.; Martín-Algarra, R.V.; Antón-Fos, G.M.; Lahuerta-Zamora, L.; Martínez-Calatayud, J. QSPR studies on the photoinduced-fluorescence behavior of pharmaceuticals and pesticides. *SAR QSAR Environ. Res.* **2017**, *28*, 609–620. [CrossRef]

38. Zanni, R.; Galvez-Llompart, M.; Morell, C.; Rodríguez-Henche, N.; Díaz-Laviada, I.; Recio-Iglesias, M.C.; Garcia-Domenech, R.; Galvez, J. Novel cancer chemotherapy hits by molecular topology: Dual Akt and Beta-catenin inhibitors. *PLoS ONE* **2015**, *10*, e0124244. [CrossRef] [PubMed]

39. Hdoufane, I.; Bjij, I.; Soliman, M.; Tadjer, A.; Villemin, D.; Bogdanov, J.; Cherqaoui, D. In silico SAR studies of HIV-1 inhibitors. *Pharmaceuticals* **2018**, *11*, 69. [CrossRef]

40. Chang, C.W.; Fosso, M.; Kawasaki, Y.; Shrestha, S.; Bensaci, M.F.; Wang, J.; Evans, C.K.; Takemoto, J.Y. Antibacterial to antifungal conversion of neamine aminoglycosides through alkyl modification. Strategy for reviving old drugs into agrofungicides. *J. Antibiot.* **2010**, *63*, 667–672. [CrossRef] [PubMed]

41. Bueso-Bordils, J.I.; Perez-Gracia, M.T.; Suay-Garcia, B.; Duart, M.J.; Algarra, R.V.M.; Zamora, L.L.; Anton-Fos, G.M.; Lopez, P.A.A. Topological pattern for the search of new active drugs against methicillin resistant Staphylococcus aureus. *Eur. J. Med. Chem.* **2017**, *138*, 807–815. [CrossRef]

42. Mor, S.; Nagoria, S.; Kumar, A.; Kumar, A.; Kaushik, C.P. QSAR studies and design of some tetracyclic 1, 4-benzothiazines as antimicrobial agents. *Drug Res.* **2016**, *66*, 436–443. [CrossRef]

43. Randic, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615. [CrossRef]

44. Kier, L.B.; Hall, L.H. The E-state as an extended free valence. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 548–552. [CrossRef]

45. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1949.

46. Vogt, M.; Wassermann, A.M.; Bajorath, J. Application of information-theoretic concepts in chemoinformatics. *Information* **2010**, *1*, 60–73. [CrossRef]

47. Kier, L.B.; Hall, L.H. General definition of valence delta-values for molecular connectivity. *J. Pharm. Sci.* **1983**, *72*, 1170–1173. [CrossRef] [PubMed]

48. Domagala, J.M.; Hagen, S.E. Structure-activity relationships of the quinolone antibacterials in the new millenium: Some things change and some do not. In *Quinolone Antimicrobial Agents*; Hooper, D.C., Rubinstein, E., Eds.; ASM Press: Washington, WA, USA, 2003; pp. 3–18.

49. CLSI. *Performance Standards for Antimicrobial Susceptibility Testing*; Twenty-Second Informational Supplement; CLSI Document M100-S22; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2012; Volume 32.

50. DESMOL13 Software. *Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular*; Facultad de Farmacia, Universitat de València: Valencia, Spain, 2000.

51. Hall, L.H. *MOLCONN-Z Software*; Eastern Nazarene College: Quincy, MA, USA, 1995.

52. Dixon, W.J. *BMDP Statistical Software*; University of California: Berkeley, CA, USA, 1990.

53. Furnival, G.M.; Wilson, R.W. Regressions by leaps and bounds. *Technometrics* **1974**, *16*, 499–511. [CrossRef]

54. Hocking, R. Criteria for selection of a subset regression: Which one should be used? *Technometrics* **1972**, *14*, 967–970. [CrossRef]

55. Hawkins, D.M. The problem of overfitting. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1–12. [CrossRef]