



Article

Molecular Topology for the Search of New Anti-MRSA Compounds

Jose I. Bueso-Bordils ^{1,*}, Pedro A. Alemán-López ¹ , Rafael Martín-Algarra ¹, Maria J. Duarte ¹, Antonio Falcó ² and Gerardo M. Antón-Fos ¹

¹ Departamento de Farmacia, Universidad Cardenal Herrera-CEU, CEU Universities C/Ramón y Cajal s/n, 46115 Alfara del Patriarca, Valencia, Spain; paleman@uchceu.es (P.A.A.-L.); rmartin@uchceu.es (R.M.-A.); mduart@uchceu.es (M.J.D.); ganton@uchceu.es (G.M.A.-F.)

² ESI International Chair@CEU-UCH, Departamento de Matemáticas, Física y Ciencias Tecnológicas, Universidad Cardenal Herrera-CEU, CEU Universities San Bartolomé 55, 46115 Alfara del Patriarca, Valencia, Spain; afalco@uchceu.es

* Correspondence: jose.bueso@uchceu.es; Tel.: +34-96-1369000

Abstract: The variability of methicillin-resistant *Staphylococcus aureus* (MRSA), its rapid adaptive response against environmental changes, and its continued acquisition of antibiotic resistance determinants have made it commonplace in hospitals, where it causes the problem of multidrug resistance. In this study, we used molecular topology to develop several discriminant equations capable of classifying compounds according to their anti-MRSA activity. Topological indices were used as structural descriptors and their relationship with anti-MRSA activity was determined by applying linear discriminant analysis (LDA) on a group of quinolones and quinolone-like compounds. Four extra equations were constructed, named DF_{MRSA1} , DF_{MRSA2} , DF_{MRSA3} and DF_{MRSA4} (DF_{MRSA} was built in a previous study), all with good statistical parameters, such as Fisher–Snedecor F (>68 in all cases), Wilk's lambda (<0.13 in all cases), and percentage of correct classification ($>94\%$ in all cases), which allows a reliable extrapolation prediction of antibacterial activity in any organic compound. The results obtained clearly reveal the high efficiency of combining molecular topology with LDA for the prediction of anti-MRSA activity.

Keywords: antibacterial; antibiotics; anti-MRSA; computational chemistry; linear discriminant analysis; molecular topology; molecular connectivity; topological indices; quinolones; QSAR



Citation: Bueso-Bordils, J.I.; Alemán-López, P.A.; Martín-Algarra, R.; Duarte, M.J.; Falcó, A.; Antón-Fos, G.M. Molecular Topology for the Search of New Anti-MRSA Compounds. *Int. J. Mol. Sci.* **2021**, *22*, 5823. <https://doi.org/10.3390/ijms22115823>

Academic Editor: Jesus Vicente De Julián Ortiz

Received: 30 April 2021

Accepted: 26 May 2021

Published: 29 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Staphylococcus aureus is a Gram-positive bacterium that causes important infections in humans. The first strain of methicillin-resistant *Staphylococcus aureus* (MRSA) was identified in 1960, almost immediately after the introduction of methicillin in therapeutics [1,2].

Antibiotic resistance occurs when the bacteria that cause the infection survive after exposure to a drug that, under normal conditions, would kill or inhibit its growth [3]. As a result, these surviving strains multiply and spread, due to the lack of competition from other strains sensitive to the same drug. This has led to the emergence of what we call “superbugs”, such as MRSA, which are difficult to treat with available antibiotics [4].

MRSA is an important cause of infections, both of community and hospital origin, and represents a major clinical and public health problem due to the limited treatment options (partly due to its worldwide spread), well-established resistance to vancomycin (formerly the antibiotic of choice for MRSA infections), and high number of therapeutic failures [5,6]. In fact, the World Health Organization (WHO) has placed MRSA as a priority 2 (high-priority) pathogen on its list of priority pathogens for the development of new antibiotics [7]. Furthermore, according to the latest antibacterial resistance report developed by the European Center for Disease Prevention and Control (ECDC), MRSA

strains are present in virtually the entire European continent, being one of the most common causes of nosocomial infections [8].

There are several factors that contribute to the success of this bacterium as a pathogen, among which are its ability to persist as a commensal, its frequent resistance to multiple antibacterials, and its variety of virulence determinants [9]. These bacteria have a great capacity to survive in adverse environments, which, after entering the hospital environment through patients, visitors, and/or healthcare workers, spread to other patients, mainly through the hands of healthcare personnel [10]. However, although initially related to the hospital setting, MRSA infections are common at the community level as well [11].

Its morbidity is variable and depends on factors specific to the host, the type of infection, and the precociousness of treatment. Although most of these infections affect the skin and soft tissues, these organisms are also capable of causing devastating diseases in certain patients. These infections include necrotizing fasciitis, septic thrombophlebitis of the extremities, Waterhouse–Frederickson syndrome, and rapidly progressive pneumonia [12].

Hence, MRSA can be regarded as a serious health problem worldwide. The current interest in the study of this pathogen is due to its high frequency and because it represents one of the main causes of nosocomial infection outbreaks around the world. The growing increase in resistant strains, which is now spreading faster in comparison to the development of new molecules, makes it necessary to investigate new antibacterial agents to expand the current therapeutic arsenal. In addition, WHO has repeatedly noted that current investment in the development of new antimicrobial compounds is insufficient [13].

Empirically establishing rules or filters for drugs, such as Lipinski's rule of five, will form the knowledge base to produce libraries tailored to drug discovery, which will require high-throughput identification of novel compounds within a large background of known substances [14].

This could be achieved by means of a rational drug design approach such as Quantitative Structure–Activity Relationships (QSAR). Within this field is molecular connectivity or molecular topology (MT), an effective and low-cost method capable of predicting molecular properties in new compounds, without the need to obtain or synthesize them previously [15].

By combining MT with pattern recognition techniques such as linear discriminant analysis (LDA) [16], neural networks [17], multilinear regression [18], factor analysis [19], or principal component analysis [20], and appropriately selecting the molecular descriptors to use, we can build mathematical–topological equations able to identify almost any molecular property. This makes MT a powerful tool for the search and design of new compounds with antibacterial activities, but it is not the only one in the field. In fact, a variety of linear and nonlinear statistical methods are used to develop models based on 2D or 3D representations of molecules [21]; 1D–3D-QSAR methods pose a series of limitations that have led to the development of higher dimensional QSAR models (4D–7D). Multi-dimensional models, although technically more complex, have been developed with the objective of finding the true binding mode [22].

In this context, this study aims to obtain mathematical–topological equations capable of predicting anti-MRSA activity. By combining MT and LDA, we use topological indices (TI) [23] to classify a compound as anti-MRSA or non-anti-MRSA. To do this, we simply select a group of compounds with antibacterial activity and another one lacking it.

We have selected structures from quinolones and quinolone-like compounds to build these equations (see Figure 1). Quinolones are a well-known and extensive group that will allow us to collect numerous data, leading to greater precision of the predictive equations [24]. In addition, their pharmacokinetic profile, as well as their high antibacterial activity and broad spectrum of action, makes them very versatile [25]. Furthermore, compared with the resistance levels of other classes of antibacterial, those of quinolones are relatively low [26].

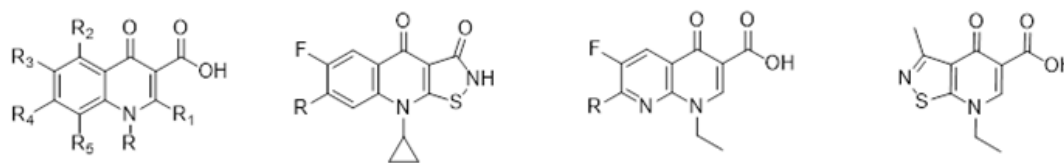


Figure 1. Core structures of the compounds used for the realization of the equations.

The discriminant functions obtained can be applied to quinolones and quinolone-like compounds not used in the study as well as in molecules that have no structural relationship, since they will select those that have a similar mathematical-topological relationship [27]. Since topology is the part of the mathematical analysis that studies the positions and interconnections of elements within a set, when applied to molecules, it gives rise to the discipline called MT, which analyzes the positions and interconnections of atoms within a molecule, giving structural information regarding length, ramifications, connection between atoms, shape, instaurations, etc.—in short, to the topological assembly or connectivity of the molecule [28]. Structurally related compounds will usually have similar values for their topological indices, but this can also occur in compounds with non-related structures.

The effectiveness of the topological method has also been proven, having been successfully applied to different therapeutic groups, such as anti-HIV [29], antihistaminic [30], antitumor [31], anti-Alzheimer [32], and antibacterial agents [33].

2. Results

To obtain the discriminant functions (DF), we used the data from a previous study [34], in which the statistical program BMDP randomly formed two training groups with 26 active and 30 inactive compounds and two test groups with seven active and six inactive compounds. These test groups allowed evaluation of the quality of the selected functions.

The DFs formed in this study along with their statistical parameters are shown in Equations (1)–(4), while their corresponding classification matrices are shown in Tables 1–4:

$$DF_{MRSA1} = 18.11717 - 3.16799S_{-CH3} + 8.24096S_{=C<} - 6.48843S_{aSa} + 1.91821S_{Cl} \quad (1)$$

N = 43 λ = 0.1223510 F = 68.145

$$DF_{MRSA2} = 62.90981 - 15.34013S_{>CH-} + 14.60322S_{=C<} + 3.06032S_{Cl} - 220.70956^3J^V \quad (2)$$

N = 43 λ = 0.0865409 F = 100.275

$$DF_{MRSA3} = 45.25057 - 3.70040S_{>N-} + 3.76891S_{Cl} - 400.80700^3J + 3.60644PR2 \quad (3)$$

N = 43 λ = 0.0894309 F = 96.727

$$DF_{MRSA4} = 70.32771 + 84.92148^3\chi_{ch} - 16.26568S_{-CH=} + 3.69496S_{Cl} - 394.47199^3J \quad (4)$$

N = 43 λ = 0.0628582 F = 141.634

Table 1. Classification matrix for DF_{MRSA1} .

Group	Active	Inactive	% Success
Training active	18	1	94.7
Training inactive	0	24	100
Test active	7	0	100
Test inactive	0	6	100
JK ¹ Training active	18	1	94.7
JK ¹ Training inactive	0	24	100

¹ Jack-Knife method.

Table 2. Classification matrix for DF_{MRSA2} .

Group	Active	Inactive	% Success
Training active	19	0	100
Training inactive	0	24	100
Test active	7	0	100
Test inactive	0	6	100
JK Training active	18	1	94.7
JK Training inactive	0	24	100

Table 3. Classification matrix for DF_{MRSA3} .

Group	Active	Inactive	% Success
Training active	19	0	100
Training inactive	0	24	100
Test active	7	0	100
Test inactive	0	6	100
JK Training active	19	0	100
JK Training inactive	0	24	100

Table 4. Classification matrix for DF_{MRSA4} .

Group	Active	Inactive	% Success
Training active	19	0	100
Training inactive	0	24	100
Test active	7	0	100
Test inactive	0	6	100
JK Training active	19	0	100
JK Training inactive	0	24	100

The classification criterion was determined by the value of the DF: if the value of the equation for a given compound was equal to or greater than 0, such a compound was classified as active, whereas if the value of the equation for a compound was smaller than 0, such a compound was classified as inactive.

3. Discussion

All equations have a low value of λ , indicating that there is a low linear dependence between independent variables. Furthermore, the high value of F in the equations indicates that the selected independent variables contribute largely to the separation of the active and inactive groups. Moreover, all equations correctly classify each compound with its corresponding group with very high success rates (100% in most cases).

We obtained the first three functions using combinations of different types of indices: DF_{MRSA1} used the electrotopological state (S_i) [35]; DF_{MRSA2} used electrotopological and charge indices [36]; while DF_{MRSA3} used all topological indices excluding the connectivity indices [37]. We obtained DF_{MRSA4} using all 136 topological indices.

DF_{MRSA1} involves four electrotopological indices (S_{-CH3} , $S_{=C<}$, S_{aSa} and S_{Cl}). Its value is influenced positively by the presence of sp^2 carbons ($S_{=C<}$) and chlorine atoms (S_{Cl}). However, the presence of methyl groups (S_{-CH3}) and sulfur atoms in aromatic rings (S_{aSa}) have a negative influence on the DF value. It should be noted that, in the latter case, some of the training and test inactive compounds possess it, while all active compounds lack this functional group in their structure (the structure of all compounds as well as bibliographic references about their activity can be found in Tables S1 and S2).

DF_{MRSA2} involves one valence charge index (${}^3J^V$) and three electrotopological indices ($S_{=C<}$, $S_{>CH-}$ and S_{Cl}). In this case, the presence of sp^2 carbons ($S_{=C<}$) and chlorine atoms (S_{Cl}) increases its value, while the presence of sp^3 carbons ($S_{>CH-}$), and sp^2 oxygens ($S_{=O}$) decreases it. Regarding the charge index, the value of the DF is influenced negatively by the topological charge present in third-order sub-pseudographs (${}^3J^V$). This type of index describes the distribution of the global charge in the molecule through the evaluation of charge transfer between pairs of atoms. Moreover, since it is also a valence index, it also considers heteroatoms and multiple bonds.

DF_{MRSA3} involves two electrotopological indices ($S_{>N-}$ and S_{Cl}), one charge index (3J) and one geometric index ($PR2$). In this case, the value of the equation increases due to the presence of chlorine atoms (S_{Cl}) and the geometric index $PR2$. Geometric indices are related to the shape and molecular surface [38]. The $PR2$ index in particular counts the number of pairs of branches (understanding as branches the points that contain three or more vertices) separated by two axes. This means that the presence of ramifications favors the anti-MRSA activity. On the other hand, the value of the equation is negatively influenced by the presence of tertiary amine groups ($S_{>N-}$) and very negatively by the topological charge in third-order sub-graphs (3J).

DF_{MRSA4} involves one connectivity index (${}^3\chi_{ch}$), two electrotopological indices (S_{Cl} and $S_{=CH-}$), and one charge index (3J). The equation also shows a clear dependence of the activity relative to the Kier–Hall chain-type third-order index (${}^3\chi_{ch}$), which implies that the presence of a cyclopropyl group greatly enhances the activity against MRSA. Most of the active compounds have this group. On the contrary, the inactive compounds generally lack this group (see Tables S1 and S2). This group has been proven to have a positive influence on the activity of a large number and variety of drugs [39]. The other index that has a positive influence on the value of the equation is S_{Cl} , meaning that the presence of chlorine atoms favors the anti-MRSA activity, while the presence of sp^2 carbons ($S_{=CH-}$) is detrimental for such activity. The topological charge in third-order sub-graphs (3J) has also a negative influence on the anti-MRSA activity.

After analyzing all the equations, we found that the electrotopological index S_{Cl} was common to all of them, with a positive influence in all cases. This indicates that the presence of this atom favors the anti-MRSA activity, as we concluded in one of our previous studies [34]. Foroumadi et al. [40] experimentally demonstrated the importance of this atom by substituting fluorine or hydrogen for chlorine in positions 6 and/or 8 from the double aromatic ring within the 4-quinolone structure. This change drastically increased the in vitro activity against two clinically isolated strains of MRSA and other bacterial species in numerous analog compounds.

We plotted the corresponding pharmacological distribution diagrams (PDD) for every function in order to visualize the values of the function in which the probability of classifying a compound as active or inactive is maximum—in other words, to find areas where the overlap between the two groups of compounds is minimal. The PDDs obtained for the DFs built along with the highest activity range for each function are shown in Figures 2–5. Compounds with values below the range were considered inactive while compounds with values over the range were considered unclassified. Thus, the value ranges derived from these PDDs establish the applicability domain for each of them [41].

DF_{MRSA1} Highest Activity Expectancy = 4–24

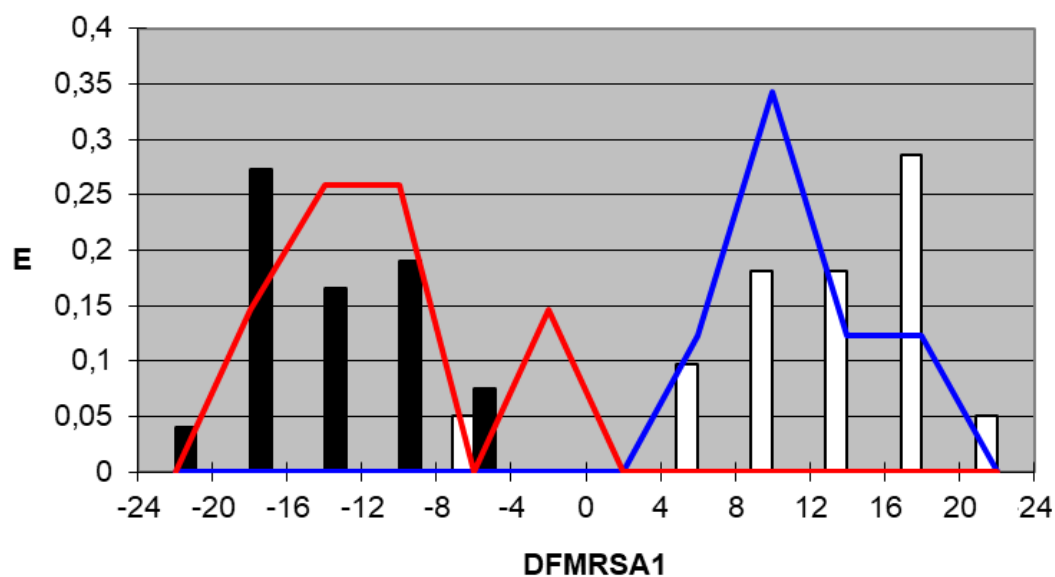


Figure 2. $DF_{M RSA1}$ pharmacological distribution diagram. Black bars: Training Inactives. White bars: Training Actives. Red curves: Test Inactives. Blue curves: Test Actives.

$DF_{M RSA2}$ Highest Activity Expectancy = 0–40

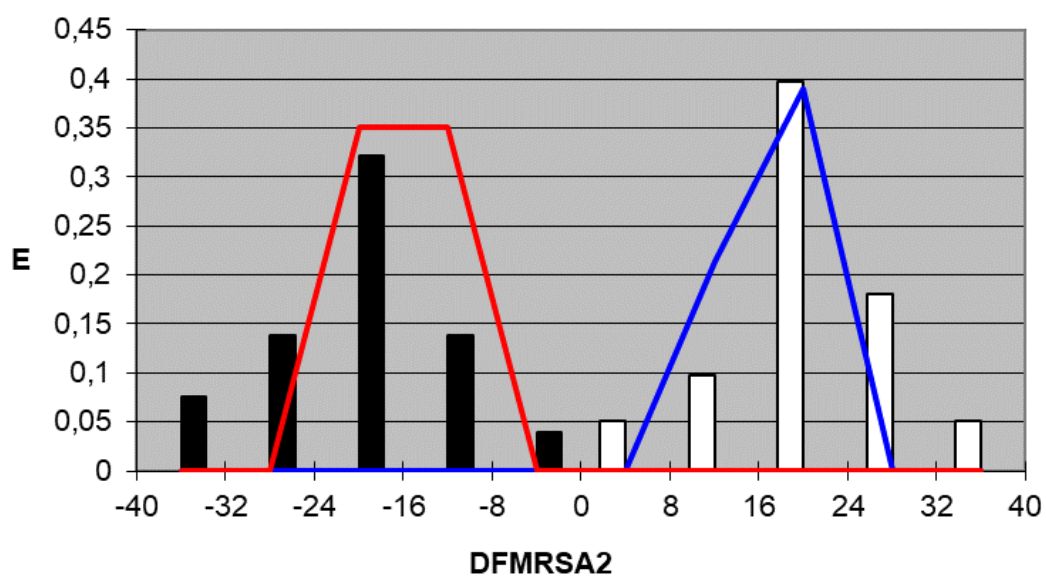


Figure 3. $DF_{M RSA2}$ pharmacological distribution diagram. Black bars: Training Inactives. White bars: Training Actives. Red curves: Test Inactives. Blue curves: Test Actives.

$DF_{M RSA3}$ Highest Activity Expectancy = 10–40

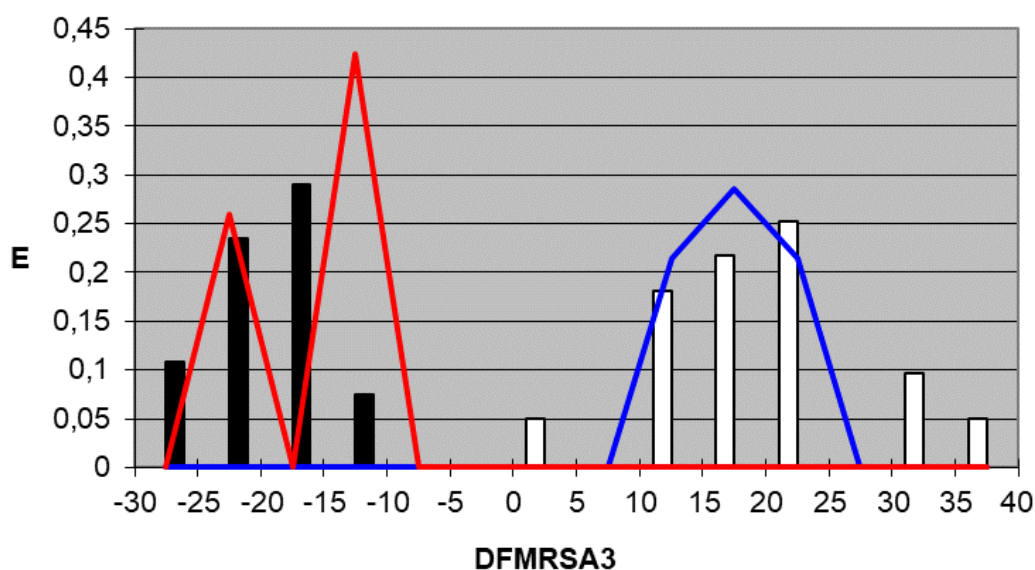


Figure 4. $DF_{M RSA3}$ pharmacological distribution diagram. Black bars: Training Inactives. White bars: Training Actives. Red curves: Test Inactives. Blue curves: Test Actives.

$DF_{M RSA4}$ Highest Activity Expectancy = 16–48

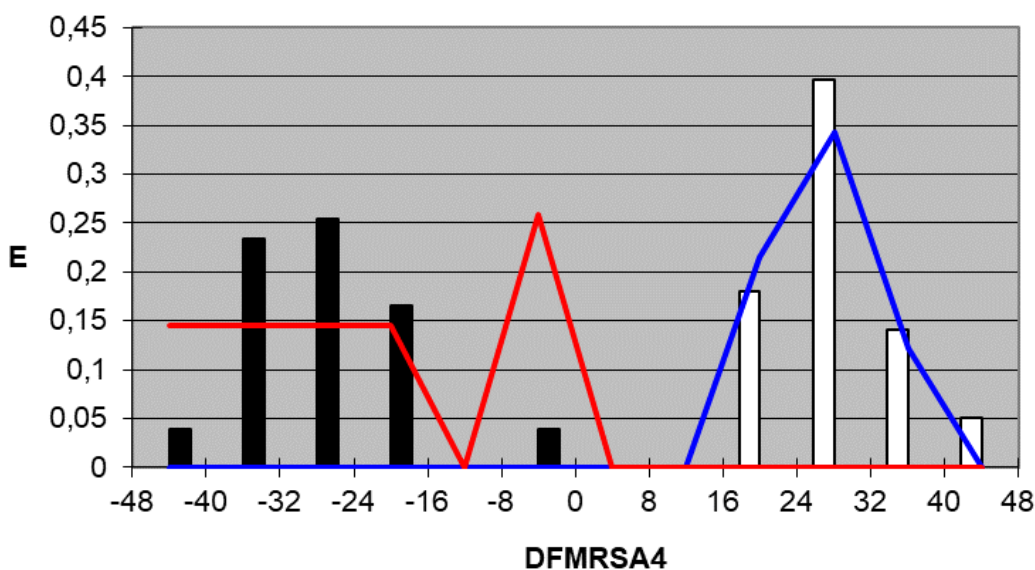


Figure 5. $DF_{M RSA4}$ pharmacological distribution diagram. Black bars: Training Inactives. White bars: Training Actives. Red curves: Test Inactives. Blue curves: Test Actives.

When PDDs are used, the accuracy for active compounds decreases, whereas that for inactive ones increases, so the probability of selecting a false active compound after applying the PDD filter decreases. The average percentage for the four equations of correctly classified inactive compounds is 100% for both the training and test group, and the average percentage of accurately classified active compounds is 97.4% for the training group and 100% for the test group. Tables 5 and 6 summarize the classification of the results obtained for all functions selected for both training groups, the active group and inactive group, respectively, and Table 7 summarizes the results for the test groups. As can be inferred from the tables, the training and test groups exhibit an average overall accuracy of 99.1%.

Table 5. Results obtained after combining LDA and PDD for the DFs. Training group: actives.

Compound	DF _{MRSA1} ¹	Clas ²	DF _{MRSA2} ¹	Clas ²	DF _{MRSA3} ¹	Clas ²	DF _{MRSA4} ¹	Clas ²
11act4	16.626	+	16.357	+	19.694	+	27.203	+
11act5	11.683	+	19.403	+	20.222	+	31.400	+
14act5	15.581	+	21.126	+	20.144	+	27.456	+
17act4	16.644	+	23.153	+	19.97	+	27.456	+
18act4	16.891	+	25.075	+	20.329	+	31.400	+
21act5	5.8776	+	16.021	+	18.056	+	18.668	+
22act4	11.789	+	20.928	+	13.000	+	20.577	+
28act4	17.471	+	26.089	+	20.238	+	31.400	+
30act4	15.632	+	22.201	+	16.962	+	24.358	+
34act5	17.661	+	24.684	+	32.461	+	39.959	+
35act5	17.614	+	23.094	+	32.469	+	39.959	+
4act5	17.676	+	23.776	+	23.800	+	34.920	+
5j-act1	21.524	+	36.244	+	36.171	+	42.102	+
5n-act1	11.424	+	18.699	+	11.661	+	25.305	+
5o-act1	9.171	+	8.2196	+	16.414	+	25.681	+
5q-act1	13.671	+	15.283	+	14.049	+	23.702	+
5r-act1	13.698	+	16.238	+	12.881	+	26.177	+
DX-619	-5.262	-	4.795	+	0.783	-	18.497	+
Sitafloxacin	6.404	+	26.299	+	24.666	+	31.369	+

¹ Value of the DF for each compound. ² The compounds are classified as active (+) if the DF value is within its highest expectancy range or inactive (-) for values outside of this range.

Table 6. Results obtained after combining LDA and PDD for the DFs. Training group: inactives.

Compound	DF _{MRSA1}	Clas	DF _{MRSA2}	Clas	DF _{MRSA3}	Clas	DF _{MRSA4}	Clas
5i-in1	-10.922	-	-21.009	-	-17.353	-	-31.118	-
in1-27j	-22.140	-	-32.034	-	-19.221	-	-19.988	-
in10-4Ib	-16.168	-	-22.602	-	-15.965	-	-26.738	-
in2-4b	-14.542	-	-18.473	-	-16.944	-	-23.370	-
in1-21a	-16.970	-	-14.755	-	-15.335	-	-35.360	-
in104IIa	-12.625	-	-17.332	-	-20.429	-	-34.317	-
in104IIb	-19.293	-	-21.486	-	-17.198	-	-34.565	-
in2-4e	-11.261	-	-19.599	-	-17.474	-	-30.255	-
in2-4f	-17.167	-	-18.987	-	-23.238	-	-28.553	-
in4-2a	-10.009	-	-18.479	-	-21.815	-	-18.213	-
in5-5d	-9.646	-	-15.358	-	-10.418	-	-4.134	-
in5-5e	-16.077	-	-18.117	-	-19.044	-	-32.963	-
in5-5f	-16.923	-	-27.682	-	-18.119	-	-31.946	-
in5-5i	-17.294	-	-26.764	-	-21.206	-	-31.331	-
in5-5k	-16.676	-	-16.353	-	-23.087	-	-29.665	-
in5-5l	-17.522	-	-25.918	-	-22.162	-	-28.649	-
in6-11a	-8.490	-	-13.852	-	-16.405	-	-28.032	-
in6-11c	-8.617	-	-21.238	-	-22.280	-	-37.466	-
in6-7c	-15.449	-	-32.295	-	-14.505	-	-19.063	-
in6-7e	-5.119	-	-6.883	-	-16.149	-	-19.063	-
in7pge61	-4.877	-	-13.953	-	-28.176	-	-34.778	-
in8-4b	-14.251	-	-20.785	-	-25.718	-	-36.492	-
in8-4c	-14.397	-	-19.425	-	-24.568	-	-35.308	-
in9-4	-16.367	-	-26.331	-	-25.571	-	-40.321	-

Table 7. Results obtained after combining LDA and PDD for DF1–4. Test group.

Compound	DF _{MRSA1}	Clas	Actives					
			DF _{MRSA2}	Clas	DF _{MRSA3}	Clas		
10act5	11.841	+	16.946	+	18.803	+	26.469	+
13act4	17.105	+	21.453	+	21.328	+	32.447	+
15act4	15.936	+	22.277	+	20.086	+	27.456	+
19act5	11.843	+	20.242	+	19.797	+	27.456	+
28act5	6.900	+	8.470	+	13.988	+	17.946	+
5m-act1	11.522	+	17.914	+	13.113	+	23.108	+
5p-act1	9.293	+	9.373	+	16.326	+	29.218	+
Compound	DF _{MRSA1}	Clas	Inactives					
			DF _{MRSA2}	Clas	DF _{MRSA3}	Clas		
25aArt13	−13.76	-	−22.394	-	−14.901	-	−47.111	-
in8-4a	−14.613	-	−19.761	-	−22.232	-	−29.320	-
in5-5g	−10.006	-	−14.539	-	−13.525	-	−3.536	-
in5-5h	−16.448	-	−17.199	-	−22.131	-	−32.348	-
in5-5j	−10.225	-	−13.780	-	−14.556	-	−0.929	-
in7pge52	−2.362	-	−12.941	-	−12.566	-	−20.154	-

4. Materials and Methods

4.1. Compound Selection

We collected in vitro activity information on quinolones and quinolone-like compounds using search engines such as ISI Web of Science, Medline, and SciFinder (Caplus). We only considered as valid those activity data from in vitro tests conducted according to the criteria of the CLSI [42].

We finally selected in vitro activity data of 56 quinolones and structurally related compounds (Figure 1) against MRSA, which we classified into two groups, active (26 compounds) and inactive (30 compounds), all of which had a 4-quinolone or a closely related structure (see Tables S1 and S2).

To consider a compound as active, it should have a minimum inhibitory concentration (MIC) equal to or under 1 mg/mL against MRSA, while those compounds considered inactive should have an MIC equal to or over 16 mg/mL against MRSA. Compounds with MICs between 1 and 16 mg/mL against MRSA were not included in the study. Regarding stereoisomers, if any of them were active, they were included in the active group. If all individual stereoisomers or the mixture of them in any ratio were inactive, they were included as a single graph in the inactive group.

4.2. Topological Descriptors

We calculated 301 topological indices for all 56 molecules. Those indices with values of 0 or with identical values for all compounds were removed. Finally, all compounds were characterized by a total of 136 non-redundant, significant descriptors specific to each molecule. These descriptors do not contain 3D parameters. The description of all the molecular descriptors can be found in Table S3 along with their definitions and references. Using MOLCONN-Z [43] and DESMOL13 [44] programs, we computed the adjacency topological matrix obtained from the hydrogen-depleted chemical pseudographs, previously drawn with the ChemBioDraw Ultra 12.0 molecule-editing program of the ChemBioOffice 2010 package.

4.3. Linear Discriminant Analysis (LDA)

LDA is a pattern recognition method that allows the classification of a compound in a given group or category (e.g., active and inactive) based on a combination of variables (e.g., topological indices). Based on the Fisher–Snedecor parameter F, which relates the variance explained by the equation to the residual variance, we chose the variables used to compute the linear classification functions in a stepwise manner. At each step, the

variable with the greater value of F (i.e., the variable that causes the larger contribution to the differentiation of groups in the discriminant function) is entered. On the other hand, selected variables with a small value of F (i.e., variables that lower the statistical significance of the classification function) are removed.

The percentage of correct classifications attained for each set assesses the discriminant ability of each DF. The classification criterion is the minimal Mahalanobis distance (distance of each case to the mean of all the cases in a category). The quality of the discriminant function was evaluated through Wilk's U -statistical parameter, λ , which was obtained by a multivariate analysis of variance that tests the equality of group means for the variable in the discriminant model.

The BMDP 7M Biomedical package [45] was the software used for the LDA study. The program randomly chooses compounds for the training and test groups. With the training group, a predictive mathematical model relating activity to structural descriptors is obtained. Internal validation is performed by the Jack-Knife (JK) method. Finally, the program performs external validation by applying the classification function to the test group and calculating the percentage of classification success.

4.4. Pharmacological Distribution Diagrams (PDD)

PDDs are histogram-like plots of connectivity functions used to determine the intervals of the discriminant function in which the expectancy, E , to find active compounds is maximum. In these plots, expectancies appear on the ordinate axis. For an arbitrary interval of values of a given function, we can define the expectancy of activity as $E_a = a / (i + 1)$, where " a " is the number of active compounds in the interval divided by the total number of active compounds, and " i " is the number of inactive compounds. The expectancy of inactivity is defined in a symmetrical way, as $E_i = i / (a + 1)$ [41]. This representation provides good visualization of the regions of minimum overlap and helps selecting intervals in the abscissa axis with maximum probability of finding active compounds.

PDDs allowed us to carry out the assignment of thresholds useful to discriminate active from inactive compounds with the highest probability of success.

5. Conclusions

Currently, the development of resistance of microorganisms such as *Staphylococcus aureus* is one of the most important problems that has appeared in recent years in the treatment of infectious diseases. Molecular topology has been demonstrated to be a useful methodology for identifying new compounds with antimicrobial activity against MRSA. By combining it with LDA, we developed four mathematical-topological equations with outstanding discriminant ability according to the success rates. These results are supported by internal and external validation performed on all functions, as well as by all the statistical parameters, which can be considered very satisfactory in all cases.

We can conclude that the equations obtained in this study confirm molecular topology as a powerful and efficient tool in the discrimination of anti-MRSA activity, offering new insights in the search for new compounds with this specific activity as opposed to classical structure-activity relationships.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms22115823/s1>, Table S1: Active group used in the discriminant functions: Paper name/code, IUPAC name & structure and bibliographic references about activity for each compound, Table S2: Inactive group used in the discriminant functions: Paper name/code, IUPAC name & structure and bibliographic references about activity for each compound, Table S3: Symbols and Definitions of Topological Indices used with DESMOL13 and MOLCONN-Z programs.

Author Contributions: Conceptualization, G.M.A.-F.; validation, P.A.A.-L. and G.M.A.-F.; formal analysis, G.M.A.-F. and A.F.; investigation, R.M.-A.; data curation, J.I.B.-B. and M.J.D.; writing—original draft preparation, J.I.B.-B.; writing—review and editing, P.A.A.-L.; supervision, G.M.A.-F.; project administration, A.F.; funding acquisition, A.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by ESI InternationalChair@CEU-UCH.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the Supplementary material.

Acknowledgments: J.I.B.-B. gratefully acknowledges his grants (BC/ICB-Santander 05/12 and BC/ICB-Santander 06/12) from the ICB (Biomedical Science Institute of the CEU-Cardenal Herrera University).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Uhlemann, A.C.; Otto, M.; Lowy, F.D.; DeLeo, F.R. Evolution of community- and healthcare-associated methicillin-resistant *Staphylococcus aureus*. *Infect. Genet. Evol.* **2014**, *21*, 563–574. [CrossRef]
2. Jevons, P.M. “Celbenin”-resistant staphylococci. *Br. Med. J.* **1961**, *1*, 124–125. [CrossRef]
3. Alvan, G.; Edlund, C.; Heddini, A. The global need for effective antibiotics—A summary of plenary presentations. *Drug Resist. Updates* **2011**, *14*, 70–76. [CrossRef] [PubMed]
4. Adegoke, A.A.; Faleye, A.C.; Singh, G.; Stenström, T.A. Antibiotic-Resistant Superbugs: Assessment of the Interrelationship of Occurrence in Clinical Settings and Environmental Niches. *Molecules* **2016**, *22*, 29. [CrossRef] [PubMed]
5. Bruniera, F.R.; Ferreira, F.M.; Saviolli, L.R.M.; Bacci, M.R.; Feder, D.; da Luz Gonçalves Pedreira, M.; Sorgini Peterlini, M.A.; Azzalis, L.A.; Campos Junqueira, V.B.; Fonseca, F.L.A. The use of vancomycin with its therapeutic and adverse effects: A review. *Eur. Rev. Med. Pharmacol. Sci.* **2015**, *19*, 694–700.
6. Boswihi, S.S.; Udo, E.E. Methicillin-resistant *Staphylococcus aureus*: An update on the epidemiology, treatment options and infection control. *Curr. Med. Res. Pract.* **2018**, *8*, 18–24. [CrossRef]
7. 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 16 March 2021).
8. European Center for Disease Prevention and Control. Antimicrobial resistance in the EU/EEA (EARS-Net)—Annual Epidemiological Report for 2019. Available online: <https://www.ecdc.europa.eu/sites/default/files/documents/surveillance-antimicrobial-resistance-Europe-2019.pdf> (accessed on 16 March 2021).
9. Foster, T.J. Immune evasion by staphylococci. *Nat. Rev. Microbiol.* **2005**, *3*, 948–958. [CrossRef]
10. Pujol, M.; Peña, C.; Pallares, R.; Ayats, J.; Ariza, J.; Gudiol, F. Risk factors for nosocomial bacteremia due to methicillin-resistant *Staphylococcus aureus*. *Eur. J. Clin. Microbiol. Infect. Dis.* **1994**, *13*, 96–102. [CrossRef]
11. DeLeo, F.R.; Otto, M.; Kreiswirth, B.N.; Chambers, H.F. Community-associated methicillin-resistant *Staphylococcus aureus*. *Lancet* **2010**, *375*, 1557–1568. [CrossRef]
12. Moellering, R.C. MRSA: The first half century. *J. Antimicrob. Chemother.* **2012**, *67*, 4–11. [CrossRef]
13. World Health Organization. 10 Facts on Antimicrobial Resistance. Available online: http://www.who.int/features/factfiles/antimicrobial_resistance/en/ (accessed on 16 March 2021).
14. Lipinski, C.A. Lead- and drug-like compounds: The rule-of-five revolution. *Drug Discov. Today Technol.* **2004**, *1*, 337–341. [CrossRef] [PubMed]
15. Khan, A.U. Descriptors and their selection methods in QSAR analysis: Paradigm for drug design. *Drug Discov. Today* **2016**, *21*, 1291–1302.
16. Maldonado, F.H. El análisis multivariante en la investigación con antimicrobianos. *Rev. Esp. Quimioterap.* **2007**, *20*, 300–309.
17. Baskin, I.I.; Palyulin, V.A.; Zefirov, N.S. Neural networks in building QSAR models. *Methods Mol. Biol.* **2008**, *458*, 137–158. [PubMed]
18. Furnival, G.M.; Wilson, R.W. Regressions by leaps and bounds. *Technometrics* **1974**, *16*, 499–511. [CrossRef]
19. Horn, J.L. A rationale and test for the number of factors in factor analysis. *Psychometrika* **1965**, *30*, 179–185. [CrossRef]
20. Meglen, R.R. Examining large databases: A chemometric approach using principal component analysis. *J. Chemom.* **1991**, *5*, 163–179. [CrossRef]
21. Suay-García, 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]
22. Wang, T.; Yuan, X.S.; Wu, M.B.; Lin, J.P.; Yang, L.R. The advancement of multidimensional QSAR for novel drug discovery—where are we headed? *Expert Opin. Drug Discov.* **2017**, *12*, 769–784. [CrossRef] [PubMed]
23. Ovidiu, I. Chemical graphs, molecular matrices and topological indices in chemoinformatics and quantitative structure-activity relationships. *Curr. Comput. Drug. Des.* **2013**, *9*, 153–163.
24. Hawkins, D.M. The problem of overfitting. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1–12. [CrossRef]
25. 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] [PubMed]

26. Mitscher, L.A. Bacterial topoisomerase inhibitors: Quinolone and pyridine antibacterial agents. *Chem. Rev.* **2005**, *105*, 559–592. [[CrossRef](#)] [[PubMed](#)]
27. 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](#)] [[PubMed](#)]
28. Rouvray, D.H. Predicting chemistry from topology. *Sci. Am.* **1986**, *255*, 40–47. [[CrossRef](#)]
29. Afantitis, A.; Melagraki, G.; Sarimveis, H.; Koutentis, P.A.; Markopoulous, J.; Igglessi-Mrkopoulou, O. A novel simple QSAR model for the prediction of anti-HIV activity using multiple regression analysis. *Mol. Divers.* **2006**, *10*, 405–414. [[CrossRef](#)]
30. Duart, M.J.; García-Doménech, R.; Gálvez, J.; Alemán, P.A.; Martín-Algarra, R.V.; Antón-Fos, G.M. Application of a mathematical topological pattern of antihistaminic activity for the selection of new drug candidates and pharmacology assays. *J. Med. Chem.* **2006**, *49*, 3667–3673. [[CrossRef](#)]
31. Zanni, R.; Galvez-Llompарт, 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](#)]
32. Wang, J.; Land, D.; Ono, K.; Galvez, J.; Zhao, W.; Vempati, P.; Steele, J.W.; Cheng, A.; Yamada, M.; Levine, S.; et al. Molecular topology as novel strategy for discovery of drugs with Ab lowering and anti-aggregation dual activities for Alzheimer’s disease. *PLoS ONE* **2014**, *9*, e92750.
33. Bueso-Bordils, J.I.; Alemán-López, P.A.; Suay-García, B.; Martín-Algarra, R.; Duart, M.J.; Falcó, A.; Antón-Fos, G.M. Molecular topology for the discovery of new broad-spectrum antibacterial drugs. *Biomolecules* **2020**, *10*, 1343. [[CrossRef](#)]
34. Bueso-Bordils, J.I.; Pérez-Gracia, M.T.; Suay-García, B.; Duart, M.J.; Martín-Algarra, R.V.; Lahuerta-Zamora, L.; Antón-Fos, G.M.; Alemán-López, P.A. Topological pattern for the search of new active drugs against methicillin resistant *Staphylococcus aureus*. *Eur. J. Med. Chem.* **2017**, *138*, 807–815. [[CrossRef](#)]
35. Kier, L.B.; Hall, L.H. The E-state as an extended free valence. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 548–552. [[CrossRef](#)]
36. Gálvez, J.; García-Domenech, R.; de Julián-Ortiz, J.V.; Soler, R. Topological approach to drug design. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 272–284. [[CrossRef](#)]
37. Randic, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615. [[CrossRef](#)]
38. Jain, D.V.S.; Singh, S.; Gombar, V. Correlations between topological features and physicochemical properties of molecules. *Proc. Indian Acad. Sci.* **1984**, *93*, 927–945.
39. Talele, T.T. The “cyclopropyl fragment” is a versatile player that frequently appears in preclinical/clinical drug molecules. *J. Med. Chem.* **2016**, *59*, 8712–8756. [[CrossRef](#)]
40. Foroumadi, A.; Ghodsi, S.; Emami, S.; Najjari, S.; Samadi, N.; Faramarzi, M.A.; Beikmohammadi, L.; Shirazi, F.H.; Shafiee, A. Synthesis and antibacterial activity of new fluoroquinolones containing a substituted N-(phenethyl)piperazine moiety. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3499–3503. [[CrossRef](#)]
41. Gálvez, J.; García-Domenech, R.; de Gregorio Alapont, C.; de Julián-Ortiz, J.V.; Popa, L. Pharmacological distribution diagrams: A tool for de novo drug design. *J. Mol. Graph.* **1996**, *14*, 272–276. [[CrossRef](#)]
42. 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.
43. *DESMOL13 Software*; Unidad de Investigación de Diseño de Fármacos y Conectividad Molecular, Facultad de Farmacia, Universitat de València: Valencia, Spain, 2000.
44. Hall, L.H. *MOLCONN-Z Software*; Eastern Nazarene College: Quincy, MA, USA, 1995.
45. Dixon, W.J. *BMDP Statistical Software*; University of California: Berkeley, CA, USA, 1990.