



Article

Application of Quantitative Structure-Activity Relationships in the Prediction of New Compounds with Anti-Leukemic Activity

Cristian Sandoval ^{1,2,3,*} , Francisco Torrens ⁴ , Karina Godoy ⁵, Camila Reyes ⁶ and Jorge Farías ^{2,*}

¹ Escuela de Tecnología Médica, Facultad de Salud, Universidad Santo Tomás, Los Carreras 753, Osorno 5310431, Chile

² Departamento de Ingeniería Química, Facultad de Ingeniería y Ciencias, Universidad de La Frontera, Temuco 4811230, Chile

³ Departamento de Ciencias Preclínicas, Facultad de Medicina, Universidad de La Frontera, Temuco 4811230, Chile

⁴ Institut Universitari de Ciència Molecular, Universitat de València, 46071 València, Spain; torrens@uv.es

⁵ Nucleo Científico y Tecnológico en Biorecursos (BIOREN), Universidad de La Frontera, Temuco 4811230, Chile; karina.godoy@ufrontera.cl

⁶ Carrera de Tecnología Médica, Facultad de Medicina, Universidad de La Frontera, Temuco 4811230, Chile; c.reyes24@ufromail.cl

* Correspondence: cristian.sandoval@ufrontera.cl (C.S.); jorge.farias@ufrontera.cl (J.F.)

Abstract: Leukemia invades the bone marrow progressively and, through unknown mechanisms, outcompetes healthy hematopoiesis. Protein arginine methyltransferases 1 (PRMT1) are found in prokaryotes and eukaryotes cells. They are necessary for a number of biological processes and have been linked to several human diseases, including cancer. Small compounds that target PRMT1 have a significant impact on both functional research and clinical disease treatment. In fact, numerous PRMT1 inhibitors targeting the S-adenosyl-L-methionine binding region have been studied. Through topographical descriptors, quantitative structure-activity relationships (QSAR) were developed in order to identify the most effective PRMT1 inhibitors among 17 compounds. The model built using linear discriminant analysis allows us to accurately classify over 90% of the investigated active substances. Antileukemic activity is predicted using a multilinear regression analysis, and it can account for more than 56% of the variation. Both analyses are validated using an internal “leave some out” test. The developed model could be utilized in future preclinical experiments with novel drugs.

Keywords: bone marrow microenvironment; cancer; leukemia; myeloproliferative disorders; model



Citation: Sandoval, C.; Torrens, F.; Godoy, K.; Reyes, C.; Farías, J. Application of Quantitative Structure-Activity Relationships in the Prediction of New Compounds with Anti-Leukemic Activity. *Int. J. Mol. Sci.* **2023**, *24*, 12258. <https://doi.org/10.3390/ijms241512258>

Academic Editors: Ioannis Kanakis and Claire Lucas

Received: 27 June 2023

Revised: 24 July 2023

Accepted: 25 July 2023

Published: 31 July 2023



Copyright: © 2023 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

Only 40% of <60 years old and 10% of >60 years old individuals have achieved long-term survival from acute myeloid leukemia (AML) [1]. The clones of leukemia that drive the disease’s persistence and recurrence are resistant to current chemotherapy and targeted treatments. Most AML patients have elevated expression of FLT3, a tyrosine kinase that may play a role in the etiology of the disease [2]. The continual ligand-independent activation of FLT3 kinase that this mutation causes makes it a promising therapeutic target [3]. When administered alone, tyrosine kinase inhibitors (TKIs) have extremely short-lived therapeutic effects and are only able to partially suppress AML cell proliferation [3,4]. Improved clinical response for FLT3-ITD+ AML patients requires the rapid development of effective combination treatments, such as TKI therapy.

To regulate signal transduction and protein–protein interactions, arginine residues can be modified into asymmetric dimethylarginine (ADMA) by adding two methyl groups to a single guanidino nitrogen [5,6]. About 85% of arginine methylation activities in human cells are carried out by PRMT1, which deposits an ADMA mark onto substrates [7]. PRMT1 methylates not just histones but also proteins, including RUNX1 and EGFR, and has been linked to processes as varied as cell proliferation, survival, and differentiation [5,8,9].

PRMT1 methylates AML1-ETO9, increasing its transcriptional activity in murine leukemia transformed by this fusion oncoprotein [9]. Recent research indicates that the oncogenic fusion protein MLL-GAS7 or MLL-EEN recruits PRMT1 to methylate H4R3, hence sustaining leukemic transcriptional pathways [10].

The PRMTs facilitate the transfer of methyl groups from the S-adenosylmethionine molecule to the guanidino nitrogen atoms of arginine. Based on how the arginine residues are methylated, methylarginines can be divided into three distinct forms: ω -NG,NG-asymmetric dimethylarginine (aDMA), ω -NG,N'-G-symmetric dimethylarginine (sDMA), and ω -NG-monomethylarginine (MMA). All PRMTs catalyze MMA, which is then used by Type I PRMTs (PRMT1, PRMT2, PRMT3, PRMT4, PRMT6, and PRMT8) to catalyze the production of aDMA or by Type II PRMTs (PRMT5 and PRMT9) to catalyze the formation of sDMA [5,6]. The PRMT7 is the only member of the Type III enzyme subclass that catalyzes MMA formation [5–7].

The PRMTs have been implicated in human tumorigenesis in numerous studies. PRMT enzymatic activity is necessary for many cellular processes in hematological malignancies, such as the activation of cell cycle and proliferation, inhibition of apoptosis, DNA repair processes, RNA splicing, and transcription by methylating histone tails' arginine [11–13]. In human malignancies, elevated PRMT expression is associated with aggressive clinical features and a poor prognosis [10,14]. PRMT1 overexpression promotes the survival and invasion of cancer cells, whereas PRMT1 suppression inhibits the proliferation of cancer cells [8,15–18]. In addition, growing evidence suggests that PRMT1 plays key roles in malignant hematopoiesis [19].

Nowadays, various *in silico* technologies are utilized to design and assess the efficacy of new medications, one of which is molecular topology, specifically molecular connectivity [20], which has proven to be useful in quantitative structure-activity relationship (QSAR) models. One of the most exciting features of molecular topology is the ease with which topological descriptors can be calculated. Each structure is described as a hydrogen-depleted network, with vertices representing atoms and edges representing bonds. Manipulation of such matrices gives multiple sets of integers known as topological indices [20], which have been shown to be capable of a simple and effective characterization of molecular structure. When these indices are chosen correctly, it is possible to obtain a highly specific characterization of each chemical compound, which can then be used in QSAR models [21–25].

In this way, the topological indices have shown their value for the selection and creation of novel pharmaceuticals [26,27], especially as antimalarial [28], antiviral [29], antihistaminic [30], hypoglycemic [31], analgesics [32,33], antituberculosis [34] and antileukemic drugs [35]. However, there are numerous limitations to consider when using QSAR modeling in the pharmaceutical business [36,37]. Due to the large number of variables involved in QSAR data—hundreds of thousands of chemicals, each represented by a unique set of descriptors, fingerprints that are typically very sparse, and some features that are highly correlated—it is anticipated that the dataset contains some mistakes since associations are examined by *in situ* studies.

As a result of these limitations, QSAR-based model prediction has been controversial. This has led to QSAR prediction being used along with machine-learning algorithms. For QSAR prediction, previous studies have turned to linear regression models [38], leave-one-out cross-validation [39], and multilinear regression analysis [40]. Our research examined the activities of a set of PRMT1 inhibitors of leukemia cell proliferation to develop QSAR models of prediction using molecular topology, linear discriminant analysis, and multilinear regression analysis. A screening method was also developed to identify novel compounds with the potential for greater bioactivity.

2. Results

2.1. Linear Discriminant Analysis

The search for an applicable mathematical-topological model to predict antileukemic activity was conducted in two phases. First, a discriminant function was chosen to differentiate between active and inactive antileukemic compounds. Second, the acquisition of a topological function capable of quantifying the efficacy of the activity in terms of IC_{50} . Both functions would constitute the framework of the mathematical model that permits the search for and selection of new potent antileukemic compounds.

To obtain the discriminant function, a linear discriminant analysis was applied to 17 compounds. The set consisted of both active and inactive compounds. The $IC_{50} \leq 1 \text{ M}/10^5$ concentration denotes the active compounds, whereas the $IC_{50} > 1 \text{ M}/10^5$ concentration denotes the inactive compounds. The selected discriminant function was:

$$\text{Discriminant function} = -14.123 + (0.848 \times \text{nDB}) + (1.680 \times \text{nN}) - (38.327 \times \text{GGI9}) + (392.190 \times \text{JGI4}), \quad (1)$$

$N = 17, \lambda \text{ (Wilk's Lambda)} = 0.172, F(4, 15) = 14.442, p < 0.001$

In Equation (1), there were topological descriptors that evaluated the constitutional characteristics of each compound (nDB), atom count (nN), and topological charge (GGI9 and JGI4).

Figure 1 depicts the antileukemic activity distribution diagram using the discriminant function (white bars represent inactive sets, and black bars represent active sets).

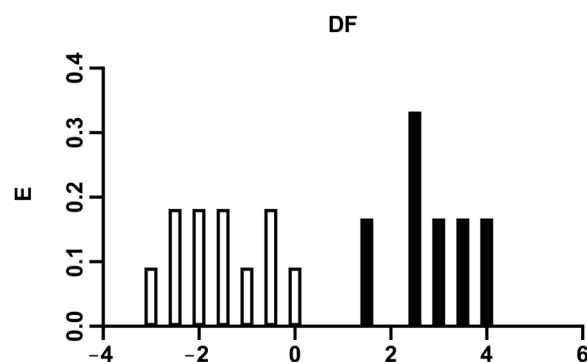


Figure 1. Pharmacological distribution diagram for antileukemic activity by plotting expectancy (E) vs. discriminant function (DF) (Equation (1)). The black bars represent the compounds with $IC_{50} \leq 1 \text{ M}/10^5$, and the white bars, the compounds with $IC_{50} > 1 \text{ M}/10^5$.

2.2. Multilinear Regression Analysis

Multiple linear regression allows one to generate a linear model in which the value of the dependent or response variable, pIC_{50} ($pIC_{50} = -\log[IC_{50}]$), is predicted from a set of independent variables called topological indices. To perform the multilinear regression analysis, compounds with quantitative IC_{50} values in $\text{M}/10^5$ units were used. In addition, the **17a** compound was removed as it is an outlier. The number of compounds needed to carry out the analysis was reduced to 16. The selected function was:

$$pIC_{50}^{\text{predicted}} = -42.089 + (314.536 \times \text{JGI4}) - (25.006 \times \text{GGI7}) + (59.732 \times \text{Mv}) + (0.450 \times \text{X0v}), \quad (2)$$

$N = 16, R^2 = 0.898, Q^2 = 0.863, \text{SEE} = 0.646, p = 0.009$

The predictive equation, Equation (2), exhibits an R^2 value above 0.80 (0.898), which explains over 89% of the variance. The descriptors selected are the mean topological charge index (JGI4), topological charge index (GGI7), mean atomic van der Waals volume (Mv), and the valence connectivity index (X0v). Particularly, the JGI4 index measures the charge transfer between pairs of atoms and, consequently, the global charge transfer in the molecule (e.g., the dipole moment) [32,33]; GGI7 index features the charge transfer between a pair of atoms, accounting for the overall charge transfer within the molecule [26], Mv

index shows the sum of the van der Waals volumes by the number of atoms, and the X0v index, the connectivity (Figure 2).

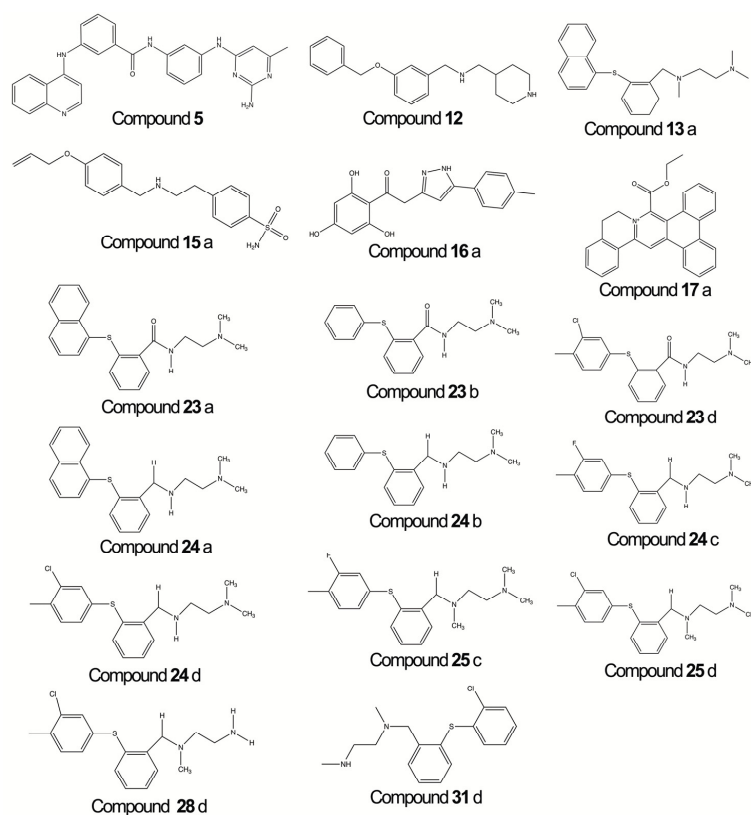


Figure 2. Chemical structures of the studied compounds.

Table 1 and Figure 3 summarize the predictions made for each compound in the training set using Equation (2). Overall, we can conclude that there is an acceptable level of efficacy because 56.3% of compounds exhibit residuals shorter than $\pm 1\text{SEE}$. The value of the greatest residual compound, compound 25d, is -1.271 . Linear discriminant analysis categorizes this compound as active, which implies that either the experimental IC_{50} is correct or the topological model used here is a valid classification for the compound's antileukemic activity.

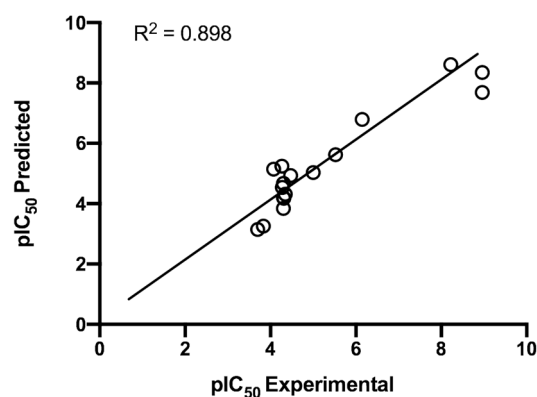


Figure 3. Prediction of growth inhibition of leukemia cell proliferation ($\text{M}/10^5$) for the PRMT1 derivatives. Graphic representation of experimental $\log \text{IC}_{50}$ against $\log \text{IC}_{50}$ calculated from Equation (2).

Table 1. Classification results were obtained from linear discriminant analysis and multilinear regression for each compound analyzed.

Comp.	SMILES	IC ₅₀ Exp. ^a (M/10 ⁵)	pIC ₅₀ Exp.	Class (Exp.)	Prob. Activ. ^b	DF ^c	Class (Pred.)	pIC ₅₀ Pred. ^d
5	NC1=NC(NC2=CC=CC(NC3=CC=CC(NC4=CC=NC5=CC=CC=C54)=C3)=O)=C2)=CC(C)=N1	1	5.000	A	1.000	3.743	A	5.033
12	C1(OCC2=CC=CC=C2)=CC=CC(CNCC3CCNCC3)=C1	14.7	3.833	I	<0.001	−2.443	I	3.258
13a	CN(CCN(O)C)CC1=C(C=CCC1)SC2=C3C(C=CC=C3)=CC=C2	0.3	5.523	A	0.971	1.586	A	5.620
15a	O=S(N)(C1=CC=C(CCNCC(C=C2)=CC=C2OCC=C)C=C1)=O	8.5	4.071	I	<0.001	−0.922	I	5.147
16a	OC1=CC(O)=C(C(C2=NNC(C3=CC=C(C)C=C3)=C2)=O)C(O)=C1	4.9	4.310	I	<0.001	−2.225	I	4.180
17a	O=C(OCC)C1=C(C2=CC=CC=C2C3=CC=CC=C43)C4=CC5=[N+]1CCC6=C5C=CC=C6	5.3	4.276	I	0.038	0.024	NC	4.536
23a	O=C(N([H])CCN(C)C)C1=C(C=CC=C1)SC2=CC=CC3=C2C=CC=C3	>20	<3.699	I	<0.001	−3.163	I	3.142
23b	O=C(N([H])CCN(C)C)C1=C(C=CC=C1)SC2=CC=CC=C2	>5	<4.301	I	<0.001	−2.123	I	4.199
23d	O=C(N([H])CCN(C)C)C1=C(C=CC=C1)SC2=CC=C(C)C(CI)=C2	>5	<4.301	I	<0.001	−1.255	I	3.843
24a	[H]C(N([H])CCN(C)C)C1=C(C=CC=C1)SC2=CC=CC3=C2C=CC=C3	3.4	4.469	I	<0.001	−2.431	I	4.932
24b	[H]C(N([H])CCN(C)C)C1=C(C=CC=C1)SC2=CC=CC=C2	>5	<4.301	I	<0.001	−1.527	I	4.673
24c	[H]C(N([H])CCN(C)C)C1=C(C=CC=C1)SC2=CC=C(C)C(F)=C2	4.5	4.347	I	0.008	−0.338	I	4.314
24d	[H]C(N([H])CCN(C)C)C1=C(C=CC=C1)SC2=CC=C(C)C(CI)=C2	5.4	4.268	I	0.008	−0.338	I	5.244
25c	[H]C(N(C)CCN(C)C)C1=C(C=CC=C1)SC2=CC=C(C)C(F)=C2	0.071	6.149	A	0.999	2.299	A	6.795
25d	[H]C(N(C)CCN(C)C)C1=C(C=CC=C1)SC2=CC=C(C)C(CI)=C2	0.039	8.959	A	0.999	2.299	A	7.688
28d	[H]C(N(C)CCN([H])[H])C1=C(C=CC=C1)SC2=CC=C(C)C(CI)=C2	0.00011	8.959	A	0.999	2.886	A	8.352
31d	CNCCN(CC1=CC=CC=C1SC2=CC=CC=C2C1)C	0.0006	8.222	A	1.000	3.931	A	8.612

^a Experimental IC₅₀ values taken from the work of Wang et al. [41], Valente et al. [42], Xie et al. [43]; ^b probability that the compound is active; ^c DF values obtained with Equation (1); ^d Activity values predicted with Equation (2). A: active; I: inactive; DF: discriminant function; NC: non-classified. The compound classified as an outlier has not been predicted.

Equation (2) was validated using leave-one-out cross-validation and an external test (Figure 4). The coefficient of prediction, $Q^2 = 0.863$, shows that Equation (2) has a high robustness. The test results are displayed in Table 1, where some anticipated $\log IC_{50}$ values are greater than 1.0, suggesting that the estimated IC_{50} value is less than $1 \text{ M}/10^5$.

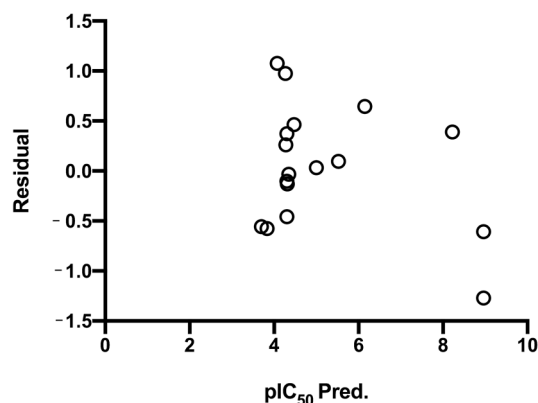


Figure 4. Graphic representation of the residual of pIC_{50} against predicted pIC_{50} obtained from the selected prediction function, Equation (2). The interval in which they are presented corresponds to $\pm 2SD$.

3. Discussion

One of the most common somatic mutations in AML is FLT3-ITD, which is seen in 25% of AML patients and is linked with a poor prognosis [3]. TKI therapy has shown relatively limited efficacy for individuals with AML, and recurrence is linked to the survival of FLT3-ITD+ AML clones [3,44]. Therefore, novel therapies are required to eradicate FLT3-ITD+ AML cells. In fact, PRMT1 specifically binds oncogenic FLT3, catalyzes its protein methylation, and hence promotes the survival and proliferation of FLT3-ITD+ AML cells [42]. In particular, FLT3 methylation levels are maintained in AML cells even after TKI treatment, and limiting this activity with a pharmacological inhibitor improved FLT3-ITD+ AML cell elimination by a TKI, suggesting that PRMT1 inhibition might function as a therapeutic for AML patients with FLT3-ITD [45].

The expression of PRMT1 is consistently high in malignant tumors [10,18,46–48]. However, the role it plays in these situations is likely to be determined by the role(s) of its substrate(s) [10,18,46–48]. Loss of PRMT1 activity inhibits MLL-GAS7- or MLL-EEN-driven leukemogenesis, according to recent research [10]. Specifically, PRMT1 methylates H4R3, which is critical for oncogenic transcriptional programs. There is growing evidence to show that, in contrast to the FLT3 WT receptor, the FLT3-ITD protein preferentially localizes intracellularly [49–51]. As a mostly nuclear and cytoplasmic protein, PRMT1 preferentially interacts with FLT3-ITD over FLT3 WT protein [52]. As a result, PRMT1 has significantly better access to FLT3-ITD protein than to FLT3 WT protein.

About 30% of individuals with AML have FLT3 changes, such as internal tandem duplication and point mutations within the tyrosine kinase domain (TKD) [53,54]. Inhibitors of FLT3 kinase-like sorafenib, quizartinib, and gilteritinib have been utilized in clinical practice [55]. However, clinical responses to these drugs are transient because of high rates of relapse and drug resistance after treatment, which contributes to disease progression and poor overall survival [4,56]. Therefore, finding effective compounds to overcome drug resistance is an urgent problem.

The steps involved in developing a QSAR model are as follows: (I) choosing a set of molecules that cover a wide range of chemical space and have verified bioactivity; (II) making 2D/3D structures of the molecules and optimizing them with the right molecular mechanics; (III) calculating molecular descriptors and pruning data with a good statistical method; and (IV) making a QSAR model with the right method. We used quantitative

structure-activity relationships to computationally screen 17 PRMT1 inhibitors of leukemia cell proliferation.

According to Mitteroecker and Bookstein [57], linear discriminant analysis is a statistical method for classifying data into two or more groups using a linear formalism [38]. Linear discriminant analysis, like multiple linear regression, creates a predicted association between categorical and continuous descriptor values. Another technique that looks for differentiation between groups is linear discriminant analysis. The equation below defines an arbitrary discriminant function: $DF = C_0 + C_1X_1 + C_2X_2 + \dots + C_kX_k$ where, X_1, X_2, \dots, X_k represents the predictor scores of the total k variables and C_1, C_2, \dots, C_k represents their respective weights. Thus, if the discriminant function was greater than zero, a given compound was selected as a potential antileukemic agent; otherwise, it was categorized as "inactive." The classification matrix was highly significant for the set (99.9% correct prediction for the active group, five out of five correctly classified, and 91.7% for the inactive group, 11 out of 12 correctly classified; Table 1).

It is evident that the regions with the least overlap for compounds with theoretical antileukemic activity occur when discriminant function >1 , indicating that the highest activity expectation occurs over this value. If discriminant function >0 and <1 , a compound will be classified as non-classified (NC). Based on the outcomes of linear discriminant analysis and multilinear regression (Equations (1) and (2)), a topological model for the search for novel antileukemic agents can be formulated. The search was made for PRMT1 inhibitor compounds according to the following requirements. If the discriminant function is >1 and <5 , and pIC_{50} is greater than 9, then the compound is labeled as potentially antileukemic. Otherwise, the compound would be considered inactive.

Antileukemic activity assays must support these suggestive findings in order to enable the validation or evaluation of the suggested model and to act as a useful tool in the search for novel compounds with higher activity against leukemia cell proliferation.

The QSAR method's main strength is that it can anticipate the characteristics of novel chemical compounds without first having to synthesize and test them. The chemical, industrial, medicinal, biological, and environmental fields all make use of this method for predicting physicochemical qualities [58]. In addition, QSAR techniques reduce costs and speed up the time it takes to create novel compounds for application as medications, materials, and additives, among others [59]. In contrast, molecular docking is a computer tool for assessing the affinity of active site residues for a given molecule or molecules [60]. The drug development industry makes use of molecular docking as a time-saving approach to examine ligand-target binding compatibility [61]. So, in future studies, molecular binding could help us to examine the ligand-target binding compatibility.

4. Materials and Methods

4.1. Analyzed Compounds and Tests Carried Out

In this QSAR study, we used the work of Wang et al. [41] to choose a group of 17 PRMT1 inhibitors of leukemia cell proliferation. As a reference, we used compound 5, which has a half-maximal inhibitory concentration (IC_{50}) against leukemia cell proliferation of $1 \text{ M}/10^5$. All these compounds have demonstrated inhibitory activity against leukemia cell proliferation, which has been previously tested and experimentally verified [41–43]. The respective IC_{50} ($\text{M}/10^5$) concentration values and their chemical structures are available (Table 1).

It is important to note that each molecule still has the same number as it did in the original work from which the compounds were taken [41–43]. To draw the chemical structure of the molecules, the ChemDraw® Professional 22.0.0.22 software was used (Figure 2).

4.2. Molecular Descriptors

Well-known topological descriptors have been used in this work: Subgraph Randić–Kier–Hall-like indices up to the fourth order (${}^m\chi_t$, ${}^m\chi_t^v$) [62,63], topological charge indices,

up to the fifth order (J_m, G_m, J^v_m, G^v_m) [31,32], quotients and differences between valence and non-valence connectivity indices (${}^mC_t = {}^m\chi_t / {}^m\chi^v_t$ and ${}^mD_t = {}^m\chi_t - {}^m\chi^v_t$) [64,65]. Each compound was characterized by a set of 100 descriptors. Table 2 shows the symbol, name, and definition of each descriptor. All descriptors used in this work were obtained with Alvascience[®] software, version 2 [66]. The descriptors are available on <https://doi.org/10.6084/m9.figshare.22784939.v1> (accessed on 23 June 2023)

Table 2. Symbol, name, and definition of each descriptor used in this study.

Symbol	Name	Definition
${}^k\chi_t, k = 0-4$ and $t = p, c, pc$	Randic-like indices of order k and type path (p), cluster (c), and path-cluster (pc)	${}^k\chi_t = \sum_{j=1}^{k_{nt}} \left(\prod_{i \in S_j} \delta_i \right)^{-1/2}$ where δ_i is the number of bonds, σ or π , of the atom i to non-hydrogen atoms. S_j is the j th sub-structure of order k and type t
${}^k\chi^v_t, k = 0-4$ and $t = p, c, pc$	Kier–Hall indices of order k and type path (p), cluster (c), and path-cluster (pc)	${}^k\chi^v_t = \sum_{j=1}^{k_{nt}} \left(\prod_{i \in S_j} \delta^v_i \right)^{-1/2}$ where δ^v_i is the Kier–Hall valence of the atom i . S_j is the j th sub-structure of order k and type t
$G_k, k = 1-5$	Topological charge indices of order k	$G_k = \sum_{i=1}^{N-1} \sum_{j=i+1}^N M_{ij} - M_{ji} \rightarrow \delta(k, D_{ij})$ where $M = AQ$ is the product of the adjacency and inverse square distance matrices for the hydrogen-depleted molecular graph. D is the distance matrix. δ is the Kronecker delta
$G^v_k, k = 1-5$	Valence topological charge indices of order k	$G^v_k = \sum_{i=1}^{N-1} \sum_{j=i+1}^N M^v_{ij} - M^v_{ji} \rightarrow \delta(k, D_{ij})$ where $M^v = A^vQ$ is the product of the electronegativity-modified adjacency and inverse square distance matrices for the hydrogen-depleted molecular graph. D is the distance matrix. δ is the Kronecker delta
$J_k, k = 1-5$	Normalized topological charge indices of order k	$J_k = \frac{G_k}{N-1}$
$J^v_k, k = 1-5$	Normalized valence topological charge indices of order k	$J^v_k = \frac{G^v_k}{N-1}$
${}^kD_t, k = 0-4$ and $t = p, c, pc$	Connectivity differences of order k and type path (p), cluster (c), and path-cluster (pc)	${}^kD_t = {}^k\chi_t - {}^k\chi^v_t$
${}^kC_t, k = 0-4$ and $t = p, c, pc$	Connectivity quotients of order k and type path (p), cluster (c), and path-cluster (pc)	${}^kC_t = \frac{{}^k\chi_t}{{}^k\chi^v_t}$

4.3. QSAR Algorithms

Linear Discriminant Analysis

Linear discriminant analysis is an algorithm that allows us to distinguish between two or more categories or objects by means of a linear function. In our case, it is about differentiating or discriminating between active and inactive compounds according to the values of the descriptors of their molecules [25].

Two sets of compounds: The first with proven activity against leukemia cell proliferation (in our case, all the compounds with $IC_{50} \leq 1 \text{ M}/10^5$), and the second comprised inactive compounds ($IC_{50} > 1 \text{ M}/10^5$) were considered for the analysis. The percentage of correct classifications tested the discriminant ability in each group. Linear discriminant analysis was performed using the SPSS[®] software, version 20 (IBM Corp., Armonk, NY, USA). The selection of the descriptors was based on the Fisher–Snedecor parameter, and the classification criteria were based on the shortest Mahalanobis distance (distance to the corresponding centroid) [67]. The statistical program selects the variables used for the calculation of the discriminant function in stepwise way. In fact, it reviews all the variables, and the one that contributes the most to the discrimination of the groups will be included in the model, while the variable that contributes the least to the prediction will be eliminated. The quality of the discriminant function was evaluated by Wilk’s lambda parameter, λ , which is a statistic of multivariate analysis of variance that tests the equality of group means for the variable(s) in the discriminant function [68].

From the selected discriminant function, the pharmacological activity distribution diagram was drawn. This diagram was pictured just to establish the intervals of the discriminant function in which the expectancy, E , of finding antileukemic compounds is maximum. Pharmacological activity distribution diagrams are histogram-like plots of connectivity functions in which expectancies appear on the ordinate axis. For each arbitrary interval of any function, the expectancy of activity Ea is defined as $Ea = a/(I + 1)$, where “ a ”

is the quotient between the number of active compounds in the interval divided by the total number of active compounds, and “i” is the number of inactive compounds in the interval divided by the total number of inactive compounds. The expectancy of inactivity, Ei , is defined in a symmetrical way as $Ei = i/(a + 1)$. This representation allows us to see the areas in which the overlap is minimal, as well as to determine the intervals of the discriminant function, where the probability of finding new active compounds is maximum in relation to the choice of a false active [69].

4.4. Multilinear Regression Analysis

The IC_{50} values have been softened as $pIC_{50} = -\log IC_{50}$, as is usual in QSAR studies. The regression equation was obtained by correlating the experimental pIC_{50} values with the topological index by multilinear regression analysis through the SPSS® software (IBM Corp., Armonk, NY, USA). The judgment for the selection of variables consisted of using the group with the least number of variables to avoid overfitting, and the value of the multiple correlation coefficient, R^2 , was high ($R^2 > 0.8$), and the standard error of the estimate was minimal. To accept the prediction function, an internal validation test was performed. The validation of the prediction function is performed through an internal cross-validation of the leave-one-out type; that is, determine the prediction coefficient (Q^2).

Leave-One-Out Cross-Validation

Each compound is removed from the model, and the activity value, pIC_{50} , is recalculated with the other compounds and descriptors from the selected equation. The process is repeated as many times as compounds are studied [70]. With the predicted values, the value of the prediction coefficient (Q^2) is determined and compared with the value of R^2 . The values of $Q^2 > 0.7$ show us that and, therefore, that the function obtained is robust and the selected model is of good quality.

4.5. Limitations

The purpose of this review was (1) to examine the activities of a set of PRMT1 inhibitors to develop QSAR models of prediction using molecular topology, linear discriminant analysis, and multilinear regression analysis and (2) to develop a screening method to identify novel compounds with greater antileukemic bioactivity. However, our *in silico* study should be replicated using *in vitro* and *in vivo* models with PRMT1 inhibitors must be made. In addition, future studies should perform molecular dynamics studies for the top three compounds and analyze the thermodynamic parameters, such as the free energy of binding using, for example, molecular mechanics/Poisson–Boltzmann surface area. Finally, molecular docking tools for this study of protein–ligand interactions and virtual screening should be performed.

5. Conclusions

Successfully using molecular topology to identify a QSAR model to predict the antileukemic activity of a group of PRMT1 inhibitor compounds. All employed molecular descriptors are graph-theoretic in nature. The mathematical model utilized in this study retains the primary structural features involving the correlated property, IC_{50} , and is therefore applicable to the virtual screening of databases for the discovery of new active compounds. In the next stage of research, multiple PRMT1 inhibitor compounds should be compiled into a virtual library for computationally pursuing and optimizing antileukemic activity against leukemia cell proliferation. Significant enhancements to the activity have been achieved.

Author Contributions: C.S.: conceptualization, methodology, software, validation, formal analysis, data curation, writing—original draft, writing—review and editing, funding acquisition; F.T.: conceptualization, methodology, formal analysis, supervision; K.G.: software, validation, formal analysis, writing—original draft; C.R.: formal analysis, writing—original draft; J.F.: methodology,

formal analysis, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universidad de La Frontera, DI22-0007 and FAPESP-UFRO, N°2020/06982-3.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available at “<https://doi.org/10.6084/m9.figshare.22784939.v1>” (accessed on 23 June 2023).

Acknowledgments: Authors thanks to SmartC-BIOREN (Service Management Analytical Research and Training Center), CCSS210005 Project, Agencia Nacional de Investigación y Desarrollo de Chile (ANID) for software acquisition, Programa de Formación de Investigadores Postdoctorales en la Universidad de La Frontera, PDT22-0001 Project, Universidad de La Frontera for design support; and Santander Iberoamérica Scholarship 2019, Banco Santander.

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

References

1. Döhner, H.; Weisdorf, D.J.; Bloomfield, C.D. Acute Myeloid Leukemia. *N. Engl. J. Med.* **2015**, *373*, 1136–1152. [\[CrossRef\]](#)
2. Ozeki, K.; Kiyoi, H.; Hirose, Y.; Iwai, M.; Ninomiya, M.; Koda, Y.; Miyawaki, S.; Kuriyama, K.; Shimazaki, C.; Akiyama, H.; et al. Biologic and clinical significance of the FLT3 transcript level in acute myeloid leukemia. *Blood* **2004**, *103*, 1901–1908. [\[CrossRef\]](#)
3. Kindler, T.; Lipka, D.B.; Fischer, T. FLT3 as a therapeutic target in AML: Still challenging after all these years. *Blood* **2010**, *116*, 5089–5102. [\[CrossRef\]](#)
4. Smith, C.C.; Wang, Q.; Chin, C.S.; Salerno, S.; Damon, L.E.; Levis, M.J.; Perl, A.E.; Travers, K.J.; Wang, S.; Hunt, J.P.; et al. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature* **2012**, *485*, 260–263. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Bedford, M.T.; Clarke, S.G. Protein Arginine Methylation in Mammals: Who, What, and Why. *Mol. Cell* **2009**, *33*, 1–13. [\[CrossRef\]](#)
6. Blanc, R.S.; Richard, S. Arginine Methylation: The Coming of Age. *Mol. Cell* **2017**, *65*, 8–24. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Tang, J.; Kao, P.N.; Herschman, H.R. Protein-arginine methyltransferase I, the predominant protein-arginine methyltransferase in cells, interacts with and is regulated by interleukin enhancer-binding factor 3. *J. Biol. Chem.* **2000**, *275*, 19866–19876. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Liao, H.W.; Hsu, J.M.; Xia, W.; Wang, H.L.; Wang, Y.N.; Chang, W.C.; Arold, S.T.; Chou, C.K.; Tsou, P.H.; Yamaguchi, H.; et al. PRMT1-Mediated Methylation of the EGF Receptor Regulates Signaling and Cetuximab Response. *J. Clin. Investig.* **2015**, *125*, 4529–4543. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Shia, W.J.; Okumura, A.J.; Yan, M.; Sarkeshik, A.; Lo, M.C.; Matsuura, S.; Komeno, Y.; Zhao, X.; Nimer, S.D.; Yates, J.R.; et al. PRMT1 interacts with AML1-ETO to promote its transcriptional activation and progenitor cell proliferative potential. *Blood* **2012**, *119*, 4953–4962. [\[CrossRef\]](#)
10. Cheung, N.; Fung, T.K.; Zeisig, B.B.; Holmes, K.; Rane, J.K.; Mowen, K.A.; Finn, M.G.; Lenhard, B.; Chan, L.C.; So, C.W. Targeting aberrant epigenetic networks mediated by PRMT1 and KDM4C in acute myeloid leukemia. *Cancer Cell* **2016**, *29*, 32–48. [\[CrossRef\]](#)
11. Sauter, C.; Simonet, J.; Guidez, F.; Dumétier, B.; Pernon, B.; Callanan, M.; Bastie, J.N.; Aucagne, R.; Delva, L. Protein Arginine Methyltransferases as Therapeutic Targets in Hematological Malignancies. *Cancers* **2022**, *14*, 5443. [\[CrossRef\]](#)
12. Sandoval, C.; Calle, Y.; Godoy, K.; Fariás, J. An Updated Overview of the Role of CYP450 during Xenobiotic Metabolization in Regulating the Acute Myeloid Leukemia Microenvironment. *Int. J. Mol. Sci.* **2023**, *24*, 6031. [\[CrossRef\]](#)
13. Nojszewska, N.; Idilli, O.; Sarkar, D.; Ahouiyek, Z.; Arroyo-Berdugo, Y.; Sandoval, C.; Amin-Anjum, M.S.; Bowers, S.; Greaves, D.; Saeed, L.; et al. Bone marrow mesenchymal/fibroblastic stromal cells induce a distinctive EMT-like phenotype in AML cells. *Eur. J. Cell Biol.* **2023**, *102*, 151334. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Yoshimatsu, M.; Toyokawa, G.; Hayami, S.; Unoki, M.; Tsunoda, T.; Field, H.I.; Kelly, J.D.; Neal, D.E.; Maehara, Y.; Ponder, B.A.; et al. Dysregulation of PRMT1 and PRMT6, Type I Arginine Methyltransferases, Is Involved in Various Types of Human Cancers. *Int. J. Cancer* **2011**, *128*, 562–573. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Wei, H.; Mundade, R.; Lange, K.C.; Lu, T. Protein Arginine Methylation of Non-Histone Proteins and Its Role in Diseases. *Cell Cycle* **2014**, *13*, 32–41. [\[CrossRef\]](#)
16. Baldwin, R.M.; Moretti, A.; Paris, G.; Goulet, I.; Cote, J. Alternatively Spliced Protein Arginine Methyltransferase 1 Isoform PRMT1v2 Promotes the Survival and Invasiveness of Breast Cancer Cells. *Cell Cycle* **2012**, *11*, 4597–4612. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Mathioudaki, K.; Papadokostopoulou, A.; Scorilas, A.; Xynopoulos, D.; Agnanti, N.; Talieri, M. The PRMT1 Gene Expression Pattern in Colon Cancer. *Br. J. Cancer* **2008**, *99*, 2094–2099. [\[CrossRef\]](#)
18. Papadokostopoulou, A.; Mathioudaki, K.; Scorilas, A.; Xynopoulos, D.; Ardavanis, A.; Kouroumalis, E.; Talieri, M. Colon Cancer and Protein Arginine Methyltransferase 1 Gene Expression. *Anticancer Res.* **2009**, *29*, 1361–1366.

19. Xu, W.S.; Parmigiani, R.B.; Marks, P.A. Histone Deacetylase Inhibitors: Molecular Mechanisms of Action. *Oncogene* **2007**, *26*, 5541–5552. [[CrossRef](#)]
20. Kier, L.B.; Murray, W.J.; Randic, M.; Hall, L.H. Molecular connectivity. V. Connectivity series concept applied to density. *J. Pharm. Sci.* **1976**, *65*, 1226–1230. [[CrossRef](#)]
21. Ivanciuc, O. Artificial Neural Networks Applications. Part 4. Quantitative structure-activity relationships for the estimation of relative toxicity of phenols for *Tetrahymena*. *Rev. Roum. Chim.* **1998**, *43*, 255–260.
22. Hosoya, H.; Gotoh, M.; Murakami, M.; Ikeda, S. Topological index and thermodynamic properties. 5. How can we explain the topological dependency of thermodynamic properties of alkanes with the topology of graphs? *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 192–196. [[CrossRef](#)]
23. De Gregorio, C.; Kier, L.B.; Hall, L.H. QSAR modeling with the electrotopological state indices: Corticosteroids. *J. Comput. Aided Mol. Des.* **1998**, *12*, 557–561. [[CrossRef](#)] [[PubMed](#)]
24. Duarte, M.J.; Anton-Fos, G.M.; De Julian-Ortiz, J.V.; Gonzalbes, R.; Gálvez, J.; García-Domenech, R. Use of molecular topology for the prediction of physico-chemical, pharmacokinetic and toxicological properties of a group of antihistaminic drugs. *Int. J. Pharm.* **2002**, *246*, 111–119. [[CrossRef](#)]
25. García-Domenech, R.; Espinoza, N.; Galarza, R.F.; Moreno-Padilla, M.J.; Rojas-Ruiz, B.; Roldán-Arroyo, L.L.; Sanchez-Lavado, M.I.; Gálvez, J. Application of molecular topology to the prediction of inhibition of *Trypanosoma cruzi* Hexokinase by bisphosphonates. *Ars. Pharm.* **2008**, *9*, 199–209.
26. Gálvez, J.; García-Domenech, R.; de Julian-Ortiz, J.V.; Soler, R. Topological approach to drug design. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 272–284. [[CrossRef](#)] [[PubMed](#)]
27. Mahmoudi, N.; de Julian-Ortiz, J.V.; Ciceron, L.; Gálvez, J.; Mazier, D.; Danis, M.; Derouin, F.; García-Domenech, R. Identification of new antimalarial drugs by linear discriminant analysis and topological virtual screening. *J. Antimicrob. Chemother.* **2006**, *57*, 489–497. [[CrossRef](#)]
28. Mahmoudi, N.; García-Domenech, R.; Gálvez, J.; Farhati, K.; Franetich, J.F.; Sauerwein, R.; Hannoun, L.; Derouin, F.; Danis, M.; Mazier, D. New active drugs against liver stages of *Plasmodium* predicted by molecular topology. *Antimicrob. Agents Chemother.* **2008**, *52*, 1215–1220. [[CrossRef](#)]
29. de Julian-Ortiz, J.V.; Gálvez, J.; Munoz-Collado, C.; García-Domenech, R.; Gimeno-Cardona, C. Virtual combinatorial syntheses and computational screening of new potential anti-herpes compounds. *J. Med. Chem.* **1999**, *42*, 3308–3314. [[CrossRef](#)]
30. Duarte, M.J.; Anton-Fos, G.M.; Aleman, P.A.; Gay-Roig, J.B.; Gonzalez-Rosende, M.E.; Gálvez, J.; García-Domenech, R. New potential antihistaminic compounds. Virtual combinatorial chemistry, computational screening, real synthesis, and pharmacological evaluation. *J. Med. Chem.* **2005**, *48*, 1260–1264. [[CrossRef](#)]
31. Calabuig, C.; Anton-Fos, G.M.; Gálvez, J.; García-Domenech, R. New hypo-glycemic agents selected by molecular topology. *Int. J. Pharm.* **2004**, *278*, 111–118. [[CrossRef](#)] [[PubMed](#)]
32. Gálvez, J.; García-Domenech, R.; De Julian-Ortiz, J.V.; Soler, R. Topological approach to analgesia. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1198–1203. [[CrossRef](#)] [[PubMed](#)]
33. Gálvez, J.; García, R.; Salabert, M.T.; Soler, R. Charge indexes. New topological descriptors. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 520–525. [[CrossRef](#)]
34. García-García, A.; Gálvez, J.; Vicente de Julian-Ortiz, J.; García-Domenech, R.; Munoz, C.; Guna, R.; Borrás, R. Search of chemical scaffolds for novel antituberculosis agents. *J. Biomol. Screen.* **2005**, *10*, 206–214. [[CrossRef](#)]
35. Shanmukha, M.C.; Basavarajappa, N.S.; Shilpa, K.C.; Usha, A. Degree-based topological indices on anticancer drugs with QSPR analysis. *Heliyon* **2020**, *6*, e04235. [[CrossRef](#)] [[PubMed](#)]
36. Ma, J.; Sheridan, R.P.; Liaw, A.; Dahl, G.E.; Svetnik, V. Deep neural nets as a method for quantitative structure–activity relationships. *J. Chem. Inf. Model.* **2015**, *55*, 263–274. [[CrossRef](#)]
37. Golbraikh, A.; Wang, X.S.; Zhu, H.; Tropsha, A. Predictive qsar modeling: Methods and applications in drug discovery and chemical risk assessment. In *Handbook of Computational Chemistry*; Leszczynski, J., Ed.; Springer International Publishing: New York, NY, USA, 2012.
38. Roy, J.; Kumar Ojha, P.; Carnesecchi, E.; Lombardo, A.; Roy, K.; Benfenati, E. First report on a classification-based QSAR model for chemical toxicity to earthworm. *J. Hazard. Mater.* **2020**, *386*, 121660. [[CrossRef](#)]
39. Shi, Y. Support vector regression-based QSAR models for prediction of antioxidant activity of phenolic compounds. *Sci. Rep.* **2021**, *11*, 8806. [[CrossRef](#)]
40. Hui, Z.H.; Aslam, A.; Kanwal, S.; Saeed, S.; Sarwar, K. Implementing QSPR modeling via multiple linear regression analysis to operations research: A study toward nanotubes. *Eur. Phys. J. Plus* **2023**, *138*, 200. [[CrossRef](#)]
41. Wang, C.; Jiang, H.; Jin, J.; Xie, Y.; Chen, Z.; Zhang, H.; Lian, F.; Liu, Y.C.; Zhang, C.; Ding, H.; et al. Development of Potent Type I Protein Arginine Methyltransferase (PRMT) Inhibitors of Leukemia Cell Proliferation. *J. Med. Chem.* **2017**, *60*, 8888–8905. [[CrossRef](#)]
42. Valente, S.; Liu, Y.; Schnakenburger, M.; Zwergel, C.; Cosconati, S.; Gros, C.; Tardugno, M.; Labella, D.; Florean, C.; Minden, S.; et al. Selective non-nucleoside inhibitors of human DNA methyltransferases active in cancer including in cancer stem cells. *J. Med. Chem.* **2014**, *57*, 701–713. [[CrossRef](#)]
43. Xie, Y.; Zhou, R.; Lian, F.; Liu, Y.; Chen, L.; Shi, Z.; Zhang, N.; Zheng, M.; Shen, B.; Jiang, H.; et al. Virtual Screening and Biological Evaluation of Novel Small Molecular Inhibitors against Protein Arginine Methyltransferase 1 (PRMT1). *Org. Biomol. Chem.* **2014**, *12*, 9665–9673. [[CrossRef](#)] [[PubMed](#)]

44. Levis, M. Midostaurin approved for FLT3-mutated AML. *Blood* **2017**, *129*, 3403–3406. [\[CrossRef\]](#)
45. Tecik, M.; Adan, A. Therapeutic Targeting of FLT3 in Acute Myeloid Leukemia: Current Status and Novel Approaches. *OncoTargets Ther.* **2022**, *15*, 1449–1478. [\[CrossRef\]](#)
46. Le Romancer, M.; Treilleux, I.; Boucheikioua-Bouzaghrou, K.; Sentis, S.; Corbo, L. Methylation, a key step for nongenomic estrogen signaling in breast tumors. *Steroids* **2010**, *75*, 560–564. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Wang, Y.; Hsu, J.M.; Kang, Y.; Wei, Y.; Lee, P.C.; Chang, S.J.; Hsu, Y.H.; Hsu, J.L.; Wang, H.L.; Chang, W.C.; et al. Oncogenic functions of Gli1 in pancreatic adenocarcinoma are supported by its PRMT1-mediated methylation. *Cancer Res.* **2016**, *76*, 7049–7058. [\[CrossRef\]](#)
48. Zou, L.; Zhang, H.; Du, C.; Liu, X.; Zhu, S.; Zhang, W.; Li, Z.; Gao, C.; Zhao, X.; Mei, M.; et al. Correlation of SRSF1 and PRMT1 expression with clinical status of pediatric acute lymphoblastic leukemia. *J. Hematol. Oncol.* **2012**, *5*, 42. [\[CrossRef\]](#)
49. Puissant, A.; Fenouille, N.; Alexe, G.; Pikman, Y.; Bassil, C.F.; Mehta, S.; Du, J.; Kazi, J.U.; Luciano, F.; Rönstrand, L.; et al. SYK is a critical regulator of FLT3 in acute myeloid leukemia. *Cancer Cell* **2014**, *25*, 226–242. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Reiter, K.; Polzer, H.; Krupka, C.; Maiser, A.; Vick, B.; Rothenberg-Thurley, M.; Metzeler, K.H.; Dörfel, D.; Salih, H.R.; Jung, G.; et al. Tyrosine kinase inhibition increases the cell surface localization of FLT3-ITD and enhances FLT3-directed immunotherapy of acute myeloid leukemia. *Leukemia* **2018**, *32*, 313–322. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Choudhary, C.; Schwäble, J.; Brandts, C.; Tickenbrock, L.; Sargin, B.; Kindler, T.; Fischer, T.; Berdel, W.E.; Müller-Tidow, C.; Serve, H. AML-associated Flt3 kinase domain mutations show signal transduction differences compared with Flt3 ITD mutations. *Blood* **2005**, *106*, 265–273. [\[CrossRef\]](#)
52. Herrmann, F.; Lee, J.; Bedford, M.T.; Fackelmayer, F.O. Dynamics of human protein arginine methyltransferase 1 (PRMT1) in vivo. *J. Biol. Chem.* **2005**, *280*, 38005–38010. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Lemonnier, F.; Inoue, S.; Mak, T.W. Genomic classification in acute myeloid leukemia. *N. Engl. J. Med.* **2016**, *375*, 900. [\[PubMed\]](#)
54. Cancer Genome Atlas Research Network; Ley, T.J.; Miller, C.; Ding, L.; Raphael, B.J.; Mungall, A.J.; Robertson, A.; Hoadley, K.; Triche, T.J., Jr.; Laird, P.W.; et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **2013**, *368*, 2059–2074.
55. Wu, M.; Li, C.; Zhu, X. FLT3 inhibitors in acute myeloid leukemia. *J. Hematol. Oncol.* **2018**, *11*, 133. [\[CrossRef\]](#)
56. Smith, C.C.; Paguirigan, A.; Jeschke, G.R.; Lin, K.C.; Massi, E.; Tarver, T.; Chin, C.S.; Asthana, S.; Olshen, A.; Travers, K.J.; et al. Heterogeneous resistance to quizartinib in acute myeloid leukemia revealed by single-cell analysis. *Blood* **2017**, *130*, 48–58. [\[CrossRef\]](#)
57. Mitteroecker, P.; Bookstein, F. Linear Discrimination, Ordination, and the Visualization of Selection Gradients in Modern Morphometrics. *Evol. Biol.* **2011**, *38*, 100–114. [\[CrossRef\]](#)
58. Wong, K.Y.; Mercader, A.G.; Saavedra, L.M.; Honarparvar, B.; Romanelli, G.P.; Duchowicz, P.R. QSAR analysis on tacrine-related acetylcholinesterase inhibitors. *J. Biomed. Sci.* **2014**, *21*, 84. [\[CrossRef\]](#)
59. Larif, M.; Chtita, S.; Adad, A.; Hmamouchi, R.; Bouachrine, M.; Lakhli, T. Predicting biological activity of Anticancer Molecules 3-ary 1-4-hydroxyquinolin-2-(1H)-one by DFT-QSAR models. *Int. J. Clin. Exp. Med.* **2013**, *3*, 32–42.
60. Lill, M.A.; Danielson, M.L. Computer-aided drug design platform using PyMOL. *J. Comput. Aided Mol. Des.* **2011**, *25*, 13–19. [\[CrossRef\]](#)
61. Hawkins, P.C.D.; Skillman, A.G.; Nicholls, A. Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* **2007**, *50*, 74–82. [\[CrossRef\]](#)
62. Kier, L.B.; Hall, L.H. Molecular connectivity. VII. Specific treatment of heteroatoms. *J. Pharm. Sci.* **1976**, *65*, 1806–1809.
63. Kier, L.B.; Hall, L.H. General definition of valence delta-values for molecular connectivity. *J. Pharm. Sci.* **1983**, *72*, 1170–1173. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Morgan, M.J.; Mukwambi, S.; Swart, H.C. On the eccentric connectivity index of a graph. *Discret. Math.* **2011**, *311*, 1229–1234. [\[CrossRef\]](#)
65. Das, K.C.; Jeon, H.; Trinajstić, N. Comparison between the Wiener index and the Zagreb indices and the eccentric connectivity index for trees. *Discret. Appl. Math.* **2014**, *171*, 35–41. [\[CrossRef\]](#)
66. Mauri, A.; Bertola, M. Alvascience: A New Software Suite for the QSAR Workflow Applied to the Blood–Brain Barrier Permeability. *Int. J. Mol. Sci.* **2022**, *23*, 12882. [\[CrossRef\]](#)
67. Gálvez, J.; Gálvez-Llompert, M.; García-Domenech, R. Introduction to molecular topology: Basic concepts and application to drug design. *Curr. Comput. Aided Drug Des.* **2012**, *8*, 196–223. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Gramatica, P.; Sangion, A. A historical excursus on the statistical validation parameters for QSAR models: A clarification concerning metrics and terminology. *J. Chem. Inf. Model.* **2016**, *56*, 1127–1131. [\[CrossRef\]](#) [\[PubMed\]](#)
69. 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. [\[CrossRef\]](#)
70. Besalu, E. Fast computation of cross-validated properties in full linear leave-many-out procedures. *J. Math. Chem.* **2001**, *29*, 191–204. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.