current issues in molecular biology

MDPI

Review

Perturbation-Theory Machine Learning for Multi-Target Drug Discovery in Modern Anticancer Research

Valeria V. Kleandrova, M. Natália D. S. Cordeiro and Alejandro Speck-Planche *

LAQV@REQUIMTE/Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal; valeria.kleandrova@gmail.com (V.V.K.); ncordeir@fc.up.pt (M.N.D.S.C.)
* Correspondence: alejspivanovich@gmail.com

Abstract: Cancers constitute a group of biological complex diseases, which are associated with great prevalence and mortality. These medical conditions are very difficult to tackle due to their multi-factorial nature, which includes their ability to evade the immune system and become resistant to current anticancer agents. There is a pressing need to search for novel anticancer agents with multi-target modes of action and/or multi-cell inhibition versatility, which can translate into more efficacious and safer chemotherapeutic treatments. Computational methods are of paramount importance to accelerate multi-target drug discovery in cancer research but most of them have several disadvantages such as the use of limited structural information through homogeneous datasets of chemicals, the prediction of activity against a single target, and/or lack of interpretability. This mini-review discusses the emergence, development, and application of perturbation-theory machine learning (PTML) as a cutting-edge approach capable of overcoming the aforementioned limitations in the context of multi-target small molecule anticancer discovery. Here, we analyze the most promising investigations on PTML modeling spanning over a decade to enable the discovery of versatile anticancer agents. We highlight the potential of the PTML approach for the modeling of multi-target anticancer activity while envisaging future applications of PTML modeling.

Keywords: PTML; fragment-based topological design; multi-target drug discovery; de novo drug design; topological indices; anticancer

check for updates

Academic Editor: Muhammad Khattab

Received: 26 March 2025
Revised: 23 April 2025
Accepted: 24 April 2025
Published: 25 April 2025

Citation: Kleandrova, V.V.; Cordeiro, M.N.D.S.; Speck-Planche, A. Perturbation-Theory Machine Learning for Multi-Target Drug Discovery in Modern Anticancer Research. *Curr. Issues Mol. Biol.* **2025**, *47*, 301. https://doi.org/10.3390/cimb47050301

Copyright: © 2025 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

Cancers constitute a wide and heterogeneous group of malignant neoplasms that represent a great threat to human life. In 2022, nearly 20 million new cases and 9.7 million deaths were caused by cancers [1]. Two factors of paramount importance prevent the discovery of more efficacious therapeutic solutions against cancers. On one side, many types of cancers are characterized by poor health outcomes, which are directly related to their grade and stage [2]. On the other hand, cancers have an intrinsic multi-factorial nature associated with multiple genetic mutations [3–7]; such biological adaptability provides cancers with the necessary means to evade the immune system, reach metastasized stages, and become resistant to current chemotherapeutic treatments [8]. As a consequence of these two factors, many of the current anticancer medications, which act through one mechanism of action, have become increasingly inefficient in tackling cancer progression [9,10]. All this points to the direction of a paradigm-shifting moment in anticancer research: from single-target modulation to multi-target drug discovery (MTDD) paradigm; the latter has emerged as a more promising option in the sense of providing more versatility, efficacy, and safety [11].

In modern drug development campaigns, in silico approaches have been demonstrated to be essential pillars, accelerating the design of (and/or the search for) novel and versatile molecular entities in multiple MTDD scenarios [12,13]. Particularly, within MTDD-based anticancer research, different in silico approaches have been employed to identify multi-target anticancer agents. These include ligand- and structure-based drug design methods [14–22], as well as complex network modeling [15–17,20,23]. In addition, predictive models derived from machine learning algorithms have also been applied in the context of either target-focused drug discovery or phenotypic drug search [24–28]. Although all these in silico approaches have been at the forefront of MTDD-based anticancer research, they present one or more of the following major disadvantages. First, some of them have relied on homogenous datasets of chemicals, impeding the search for wider regions of the chemical space. Second, they predict activity by considering only one measure of activity and one biomolecular (e.g., protein) or cellular (i.e., cancer cell line) target, thus neglecting the multi-factorial genetic nature of cancers; consequently, chemicals lacking multi-target activity are unlikely to become effective anticancer drugs. Third, usually, insufficient information is provided on the impact of the diverse assay protocols on the assessment of the different activity measurements; this can create a bias when performing activity prediction and subsequent biological experimentation. Last, in terms of physicochemical and structural interpretability, all the in silico approaches mentioned until now do not offer an accurate rationale on how to design new molecular entities with the desired multi-target anticancer activity.

The in silico approach known as perturbation-theory machine learning (PTML) has emerged to solve all the aforementioned disadvantages [29–34]. In this sense, PTML models can simultaneously predict multiple biological effects (activity, toxicity, pharmacokinetics) against dissimilar targets, which include proteins, microbial strains, cell lines, laboratory animals, and humans, among others [29–34]; information on different assay conditions have also been considered. Thus, PTML modeling has been applied to the discovery and/or design of molecules against microbial infections [35–44], neurological disorders [45–50], nano-systems for drug release or treatment [51–59], and different subfields at the interface of immunology and toxicology [60–64]. In this review, we focus on the emergence, development, and application of PTML modeling for MTDD-based anticancer research in the realm of small-molecule drug discovery. In this sense, we first briefly discuss key concepts and elements of PTML modeling. We then analyze the works that have applied the PTML approach to the modeling of the multi-target anticancer activity of chemicals at both the protein inhibition and phenotypic (cell-based) levels. We also discuss cutting-edge investigations that demonstrate that the physicochemical and structural interpretations of the PTML models can lead to the de novo design of molecules with the desired multi-target anticancer activity. Finally, we envisage some future perspectives on the use of PTML modeling in MTDD-based anticancer research.

2. An Overview of PTML Modeling

Currently, PTML models are viewed as advanced models for quantitative structure-activity relationships (QSAR) [29–64]. Because the creation and application of PTML models have been discussed in detail elsewhere (Figure 1), we will focus here only on the key aspects.

First, PTML models can simultaneously consider different biological effects (endpoints) at both in vitro and in vivo levels (b_m). Second, while predicting multiple endpoints, diverse targets (t_s) are considered. This means, that, in addition to biomolecules (i.e., cancer-related proteins), all biological systems (e.g., cancer cell lines, subcellular components, and

mammals including humans) on which a biological endpoint is determined are regarded as targets.

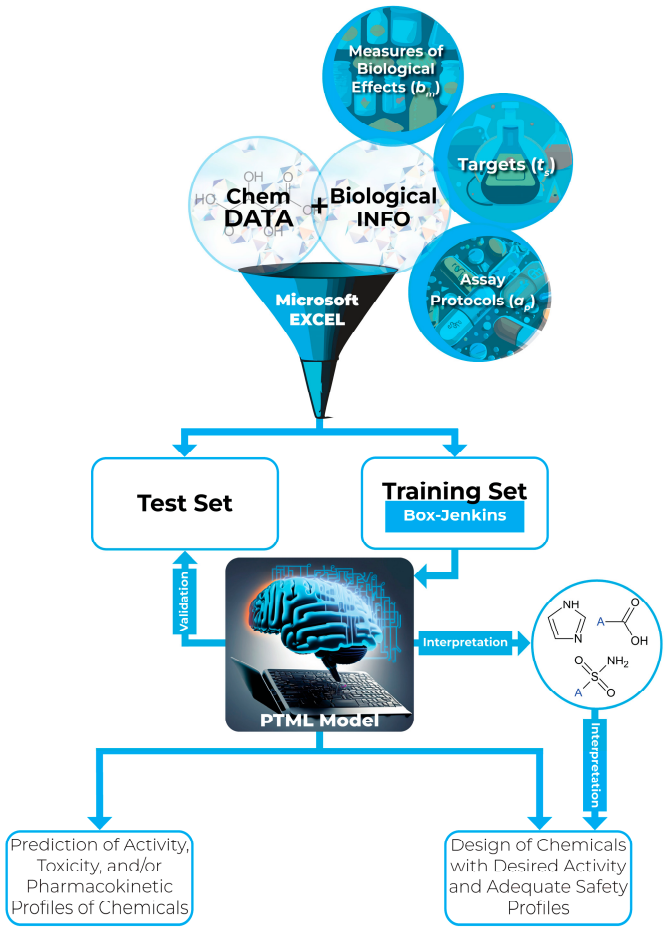


Figure 1. Generation and applications of a PTML model for small-molecule drug discovery. Chemical data related to the molecular structures are usually provided as simplified molecular-input line-entry system (SMILES) codes, which allow the calculation of the molecular descriptors (X). Interpretation is carried out by applying the fragment-based topological design (FBTD) methodology, where fragments are analyzed (with A being any atom or functional group).

Third, a wide variety of assay conditions (a_p) can be considered when developing PTML models. Fourth, the essential step before developing any PTML model is the application of the Box-Jenkins approach:

$$avg[X]e_j = \frac{1}{n(e_j)} \times \sum_{i=1}^{n(e_j)} X_i \quad (1)$$

$$D[X]e_j = \left(\frac{X - avg[X]e_j}{Num} \right) \times p(e_j)^Y \quad (2)$$

Notice that in Equations (1) and (2), X is a molecular descriptor; this can be an experimentally determined property or a theoretical value calculated from the 2D or the 3D structure of a molecule [65]. Also, e_j refers to an experimental condition, which is a combination of b_m, t_s, and a_p. On the other hand, in Equation (1), we have the average value $avg[X]e_j$ and $n(e_j)$; the latter indicates the number of chemicals that comply with a specific experimental aspect of e_j. Thus, for instance, if $e_j = b_m$, then, $n(e_j) = n(b_m)$, with $n(b_m)$ indicating the number of molecules experimentally tested by measuring the same biological effect/endpoint b_m. The same reasoning is applied to the elements t_s and a_p. This means

that Equation (1) is applied to b_m, t_s, and a_p, separately. It is important to highlight that because many chemicals/molecules are tested in diverse experimental conditions e_j, the values $avg[X]e_j$ and $n(e_j)$ will be different across the different biological endpoint measures (b_m), targets/biological systems (t_s), and assay conditions/protocols (a_p). In Equation (2), the term "*Num*" can be the range (difference between the maximum and minimum values of X), the standard deviation of X values, or a value of 1. At the same time, $p(e_j)$ is an *a priori* probability of finding a chemical tested by considering a specific experimental aspect of e_j (b_m, t_s, or a_p); the exponent Y usually is equal to -1, -0.5, 0, 0.5, or 1. It should be highlighted that in Equations (1) and (2), $avg[X]e_j$, $n(e_j)$, and $p(e_j)$ are calculated exclusively from the chemicals/molecules annotated to belong to the training set. The most important term of Equation (2) is $D[X]e_j$, which is known as a multi-label index (*MLI*). In this sense, the multi-label indices $D[X]e_j$ (*MLIs*) are used as inputs for the creation of the PTML models; they maintain the same physicochemical and structural meaning as X while also fusing the chemical information (contained within X) with the different biological aspects of e_j. The *MLIs* are the ones that allow PTML models to simultaneously predict activity, toxicity, and/or pharmacokinetic profiles of molecules under dissimilar experimental conditions e_j.

A fifth important aspect of PTML modeling is that the use of Equations (1) and (2) allows the classification of PTML models into different categories. If Equations (1) and (2) are employed only in the case of the experimental aspect t_s, then, we will be in the presence of a multi-target QSAR (mt-QSAR) model. If all the experimental aspects of e_j (b_m, t_s, and a_p) are used to simultaneously predict activity, toxicity, and pharmacokinetics, the PTML model will be classified as a multi-tasking model for quantitative structure-biological effect relationships (mtk-QSBER). Any PTML model will be classified as multi-condition QSAR (mtc-QSAR) when using b_m, t_s, and a_p to predict only activity endpoints; the term PTML can be used either instead of mtc-QSAR or any other case. Sixth, as commented in the upcoming section, PTML models are based on machine learning algorithms, particularly artificial neural networks (ANN) and linear discriminant analysis (LDA); however, random forests (RF), support vector machines (SVM), and other algorithms can also be easily implemented to generate PTML models. Seventh, the PTML models are created from the training set and validated by using the test set. Last, because of their ability to predict a multitude of endpoints, PTML models can be very useful as both virtual screening tools and guides for de novo molecular design.

3. PTML Models for MTDD-Based Anticancer Research

3.1. Key Aspects of the Analysis of PTML Modeling for MTDD-Based Anticancer Research

In the upcoming subsections, we have reported cutting-edge investigations devoted to the development of PTML models for MTDD-based anticancer research. In this sense, we would like to point out the following key aspects. In the discussion, when referring to a specific PTML modeling, we have used a notation that includes the type of PTML model and the machine learning algorithm used to create it. For instance, if a PTML model classified as mt-QSAR was based on the LDA algorithm, then, the notation for that model was "mt-QSAR-LDA". On the other hand, we have mentioned the type of molecular descriptors (e.g., topological indices, fragment-based descriptors, physicochemical properties, etc.) that were used to calculate the *MLIs* (see Equations (1) and (2)); the main software used to calculate such molecular descriptors were MODESLAB v1.5 [66], DRAGON v5.3 [67] or above, and QUBILs-MAS v1.0 [68–70]. During the analysis of the different PTML models, we have mentioned the numbers of the biological effect/endpoint measures and targets employed in the experiments; when considered relevant, assay protocols and other experimental aspects have also been mentioned. We have also reported the number of statistical cases or data points (e.g., molecule/target, molecule/endpoint measure/target,

or molecule/endpoint measure/target/assay protocol combinations) in each dataset used to generate the PTML models. It should be highlighted, that, before creating the PTML models, when splitting the datasets into training and test sets, 3:1 or 4:1 splitting ratios were used. This means that the training sets used to create the PTML models were formed by up to 80% of their corresponding datasets; test sets contained at least 20% of the data and were used to validate the PTML models. We have also reported the statistical performance of the PTML models. In doing so, because the statistical performance metrics (SPM) varied across different works, we applied the normalized terminology SPM > v_{sm}. Notice that SPM was *Sn* (sensitivity), *Sp* (specificity), or *Acc* (accuracy) while v_{sm} was the value of that particular SPM; the terminology SPM > v_{sm} (or a written similar phrase) meant that for a specific SPM, a value higher than v_{sm} was reported for both training and test sets. The search for PTML models mentioned in the upcoming subsections was performed by the STATISTICA software up to version 13.5.0.17 [71].

3.2. The PTML Approach for Modeling of Multi-Target Anticancer Activity

The first report using PTML modeling to predict multi-target anticancer activity dates back to 2011. In this work, the *MLIs* used to build several mt-QSAR-LDA models were derived from the topological indices known as bond spectral moments; the purpose of each model was to simultaneously predict inhibitory activity against eight tyrosine kinase proteins associated with the emergence and/or progression of different types of cancer [72]. The mt-QSAR-LDA models were developed from a dataset containing 1771 cases; five activity cutoff values based on the half-maximal inhibitory concentration (IC_{50}) were used, namely 0.05 μM, 0.1 μM, 0.5 μM, 1 μM, and 5 μM. All the mt-QSAR-LDA models exhibited *Sn* and *Sp* values higher than 75%. The best mt-QSAR model was the one based on the IC_{50} cutoff of 0.1 μM. This model was used to predict two small datasets, one formed by 13 molecules reported in the scientific literature and assayed against seven of the eight tyrosine kinases and the other containing three previously synthesized but newly tested quinazoline derivatives. The best mt-QSAR-LDA model satisfactorily predicted both datasets, demonstrating its capacity to accelerate the identification of multi-target tyrosine kinase inhibitors against cancers. Particularly, multi-target anticancer agents such as pelitinib and sorafenib (Figure 2) were predicted by the mt-QSAR-LDA model to exhibit versatile anti-kinase inhibitory activity, thus confirming previous experimental findings.

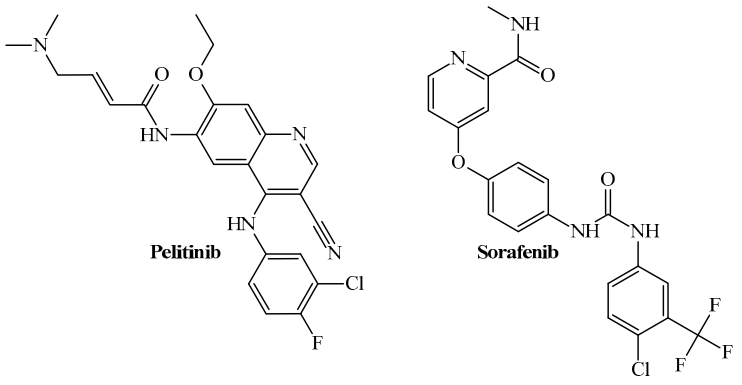


Figure 2. Chemical structures as the two most powerful versatile kinase inhibitors identified by the mt-QSAR-LDA model.

Another work reported the development of an mt-QSAR-LDA model that was devoted to predicting the multi-cell inhibitory potency of chemicals against 12 sarcoma cell lines [73]. This mt-QSAR-LDA model was generated from 3017 cases using *MLIs* based on bond spectral moments combined with fragment descriptors as inputs. The mt-QSAR-LDA

model achieved *Sn*, *Sp*, and *Acc* values above 90%. The mt-QSAR-LDA model enabled the calculation of quantitative contributions of different molecular fragments to the multi-cell anticancer activity against the 12 sarcoma cell lines. It is important to highlight that this work was the first report providing a rationale for the relation of the multi-target phenotypic (cell-based) activity and the chemical structure at the fragment/functional group level.

In a 2014 study, an mtk-QSBER-LDA model was obtained from *MLIs* based on the bond spectral moments to simultaneously predict multi-target activity against proteins associated with bladder cancer and multi-cell anticancer activity against several bladder cancer cell lines, as well as multiple in vitro and in vivo toxicity and pharmacokinetic endpoints against dissimilar targets [74]. The mtk-QSBER-LDA model was constructed from a dataset formed by 39,198 cases, which included 16 diverse endpoints b_m, 49 dissimilar targets t_s, and three different levels of curations/reliability of the assay protocols. The mtk-QSBER-LDA model displayed *Sn* and *Sp* above 95%. Through the calculation of the quantitative contributions, several molecular fragments were identified to simultaneously influence the increment of the multi-protein and multi-cell anticancer activities as well as the reduction of the toxicity and the enhancement of the pharmacokinetic profiles. This work, in addition to generalizing the calculation of quantitative fragment contributions beyond the activity domain (quantitative contribution of fragments were also performed for toxicity and pharmacokinetic profiles), is the only report that has successfully attempted to predict multi-protein/multi-cell anticancer activity while also considering toxic effects and pharmacokinetics.

Nowadays, it is recognized that the ubiquitin-proteasome pathway plays an important role in cancer [75–77]. Therefore, in 2015, research was conducted to predict multi-target inhibitors of the ubiquitin-proteasome pathway [78]. Here, an mtc-QSAR-LDA model was built from *MLIs* derived from linear indices and a dataset comprising 5602 cases distributed across 20 different b_m, at least 20 t_s, and 474 assay protocols a_p. The mtc-QSAR-LDA model exhibited *Sn* and *Sp* values higher than 70%. This work constitutes the first attempt to model the multi-target activity of chemicals in the context of MTDD-based anticancer research associated with the ubiquitin-proteasome pathway.

A 2018 study reported the development of two PTML-LDA models for the prediction of multi-protein and multi-cell activities of chemicals against diverse cancers [79], including (but not limited to) those of the breast, ovary, colon, lung, and prostate, as well as melanoma. In this research, the global physicochemical properties known as the logarithm of the n-octanol/water partition coefficient (logP) and polar surface area (PSA) were used to generate the *MLIs*, which subsequently served as inputs for the PTML-LDA models. By employing a dataset containing more than 100,000 cases (involving > 70 b_m, > 300 biomolecular and non-biomolecular t_s, and 4 labels of general assay protocols a_p, among other experimental considerations). Both PTML-LDA models achieved $Sn > 70\%$ and $Sp > 89\%$. The best of the two PTML models was used to carry out simulations on 115,000 data points, enabling the detection of tendencies on multi-target and multi-cell anticancer activity across many assay conditions and their corresponding associations with chemical structure variability.

Last, an investigation was carried out by creating several PTML-LDA and PTML-ANN models for the prediction of the multi-protein and multi-cell activity of chemicals against sarcoma [80]. Here, for the creation of these PTML models, *MLIs* derived from the global physicochemical properties logP and PSA were used as inputs. The dataset used by these PTML models consisted of >37,900 cases distributed across 155 endpoints b_m and 79 biomolecular and cellular targets t_s; in addition, 17 assay organisms were reported as relevant labels during the creation of the PTML models. In this work, the PTML models achieved $Sn > 79\%$ and $Sp > 95\%$. When compared to classical machine

learning models employing larger numbers of molecular descriptors, it was concluded that the developed PTML models were considerably simpler yet more informative than their classical counterparts. This work also offered a brief guideline on how to use the different PTML models to perform virtual screening of chemicals with potential multi-protein and/or multi-cell anti-sarcoma activity.

3.3. PTML Modeling for De Novo Drug Design in MTDD-Based Anticancer Research

The applications of the PTML approach are not limited only to the modeling of the multi-target anticancer activity; PTML modeling can also be applied to de novo drug design [81–84]. We would like to highlight that de novo drug design is devoted to generating novel molecules from scratch by using atoms/fragments as building blocks [81–84]; this means that no starting templates (e.g., known molecular scaffolds) are used.

When applying PTML modeling for de novo drug design, the methodology now formally known as fragment-based topological design (FBTD) should be used [32,33,85]. In this sense, as indicated in its name, FBTD enables a deep physicochemical and structural interpretation of the topological indices used as inputs in linear and non-linear machine learning models (including those based on the PTML approach) [32,33,85]. The reason topological indices are essential for the FBTD methodology comes from the fact that they are fairly easy to calculate, they provide considerable information on different important 3D structural aspects (e.g., dihedral angles, molecular accessibility, volume, etc.) [86–88] despite having a 2D nature, and can be expressed as linear combinations of different generic fragments (GF) in a molecule [89–95] (Figure 3); the latter aspect enables the calculation of the quantitative contribution of any fragment to the biological effect under study.

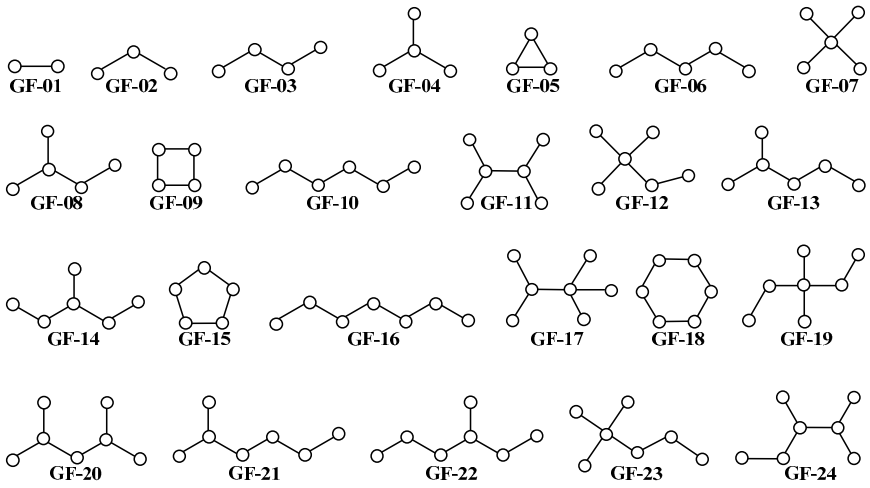


Figure 3. A non-exhaustive list of subgraphs/generic fragments (GF) used by the topological indices to characterize the chemical structure of the molecules. For instance, GF-05, GF-09, GF-15, and GF-18 characterize the presence of three-, four-, five-, and six-membered rings respectively; similarly, GF-04 (e.g., carbonyl, amide, urea, sulfoxide, tert-butyl, and other) and GF-07 (sulfone, sulfonamide, phosphorus-containing moieties, etc.) describe the presence of many important functional groups. The same reasoning can be applied to the other subgraphs/generic fragments present in this illustration. The joint use of generic fragments and atom- or bond-based physicochemical properties (atomic weight, bond dipole moment, atomic molar refractivity, bond distance, atomic hydrophobicity, molecular accessibility, etc.) allow topological indices to discriminate the many diverse functional groups/moieties present in the molecules. Therefore, topological indices can be successfully applied to the physicochemical and structural characterization of large and heterogeneous datasets of chemicals. Topological indices (and consequently, the *MLIs* derived from them) can be used as inputs for the creation of highly predictive machine learning models (including the ones based on the PTML approach).

Therefore, when using the FBTD methodology, the *MLIs* derived from topological indices in a PTML model are physicochemically and structurally interpreted; such interpretations lead to the chemistry-driven extraction and subsequent fusion/connection of suitable molecular fragments, ultimately yielding novel and rationally designed molecular entities with desired biological profiles (e.g., high activity, low toxicity, and/or enhanced pharmacokinetics) [29–34,96]. The upcoming ideas will focus on describing the cutting-edge works that have combined PTML modeling with FBTD for de novo drug design in MTDD-based anticancer research.

The first reports on the applications of PTML modeling for de novo design were a series of three works where the mt-QSAR-LDA models were created from fragment descriptors combined with *MLIs* derived from bond spectral moments. The first of these works aimed to accelerate the rational design of multi-cell inhibitors against four prostate cancer cell lines [97]. The mt-QSAR-LDA model was developed from a dataset containing 1668 cases, exhibiting $Sn > 88\%$ and $Sp > 92\%$. By using the mt-QSAR-LDA model, the calculation of the quantitative contributions revealed suitable fragments with positive contributions to the inhibitory activity against the four prostate cancer cell lines. After applying FBTD, some of these molecular fragments were merged, yielding six novel (structurally related) molecules with multi-cell inhibitory potency against the prostate cancer cell lines under study. In the second work, the mt-QSAR-LDA model was intended to rationalize the design of versatile inhibitors against 13 different breast cancer cell lines with varying degrees of sensitivity to anti-breast cancer drugs [98]. This mt-QSAR-LDA model was constructed from a dataset formed by 2272 cases and achieved $Acc > 90\%$. The use of the FBTD methodology and the posterior calculations of the quantitative activity contributions permitted the selection and subsequent connection/fusion of different molecular fragments; this led to the design of nine molecules, which were predicted by the mt-QSAR-LDA model as multi-cell inhibitors against the 13 breast cancer cell lines. The third work involved the generation of an mt-QSAR-LDA model as a tool to enable the design of anti-brain tumor agents against seven diverse brain tumor cell lines [99]. The mt-QSAR-LDA model built here from 1236 cases displayed $Acc > 88\%$. The joint combination of FBTD and the fragment quantitative contribution provided the necessary information for the design of six novel (but structurally related) chemicals that were predicted to exhibit multi-cell inhibitory potency against the seven brain tumor cell lines.

A second series of works on PTML models for de novo drug design followed, and, in this case, both linear and non-linear PTML models were generated. This series consisted of two works where the following procedure was applied. Two mt-QSAR models were built; the linear (mt-QSAR-LDA) model employed fragment descriptors and *MLIs* calculated from bond spectral moments while the non-linear (mt-QSAR-ANN) model relied on *MLIs* derived from different global topological indices. The mt-QSAR-LDA model was used to (a) extract molecular fragments, (b) calculate the quantitative contribution of those fragments to the multi-cellular anticancer activity, and (c) design new molecules as potential multi-cell inhibitors against multiple cancer cell lines. The mt-QSAR-ANN model was employed to theoretically validate the molecules designed by the mt-QSAR-LDA model. Thus, in the first work of this second series, two mt-QSAR models were built from 1651 data points to predict and design multi-cell inhibitors of 10 colorectal cancer cell lines [100]. The mt-QSAR-LDA model exhibited $Acc > 93\%$; for the mt-QSAR-ANN model, $Acc > 92\%$ was achieved. After applying the FBTD methodology and calculating the fragments' quantitative contributions, nine molecules were designed by using the mt-QSAR-LDA model; these molecules were predicted as multi-cell inhibitors against the 10 colorectal cancer cell lines. The mt-QSAR-ANN model predicted the nine designed molecules as multi-cell inhibitors in at least 8 of the 10 colorectal cancer cell lines. The second work of this series enabled the design and

prediction of chemicals with multi-cell anticancer activity against four bladder cancer cell lines [101]. The two mt-QSAR models reported in this work were developed from a dataset comprising 664 cases, exhibiting *Acc* > 92%. Here, the posterior physicochemical and structural interpretation of the descriptors (via the FBTD methodology) and subsequent computation of the quantitative contributions by the mt-QSAR-LDA model led to the merging of different suitable fragments. As a result, eight molecules were generated. The mt-QSAR-LDA model predicted the eight designed molecules to exhibit multi-cell inhibitory potency against the four bladder cancer cell lines. The same was confirmed by the mt-QSAR-ANN model, which predicted that 6 of the 8 designed molecules presented multi-cell inhibitory activity against the four bladder cancer cell lines while the other two design molecules were predicted as active against 3 of the four bladder cancer cell lines.

Two works were developed in the context of versatile inhibitors of proteins associated with breast cancer. In the first of them, research was conducted to create a PTML-LDA model for the fragment-based design and prediction of multi-target inhibitors of protein related to breast cancer [102]. In this sense, the PTML-LDA model was created from a dataset formed by 24,285 cases where chemicals were experimentally tested by considering at least 1 out of 2 b_m (IC_{50} or the inhibition constant K_i), at least 1 out of 19 breast cancer-related proteins (t_s), and at least 1 out of 2 labels of assay information; in addition, the reliability of the assay was taken into account. The PTML-LDA model used *MLIs* derived from atom-bases quadratic indices and had very good performance, with *Sn* and *Sp* values higher than 93%. In this work, eight new molecules were designed by a combination of the PTML-LDA model and the FBTD methodology; the designed molecules were predicted as multi-target inhibitors against the 19 breast cancer-related proteins under study. The second work focused on building an mt-QSAR-ANN model to predict dual-target inhibitors of the proteins named cyclin-dependent kinase 4 (CDK4) and human epidermal growth factor receptor 2 (HER2) [103]. Here, the mt-QSAR-ANN model was built from a dataset containing 2213 data points and *MLIs*-based topological indices (path-based atomic connectivity indices and 2D-autocorrelations). The mt-QSAR-ANN model displayed with *Sn* > 86% and *Sp* > 75%. The application of FBTD led to the design of six molecules with half of them being predicted as dual target inhibitors of CDK4 and HER2.

In a 2019 report, researchers focused their attention on the in silico design of multi-target inhibitors of the bromodomain-containing proteins 2, 3, and 4 (BRD2, BRD3, and BRD4, respectively) [104], which could lead to future versatile anticancer chemotherapies. In this report, two mt-QSAR models were developed from 1166 cases; the linear (mt-QSAR-LDA) used *MLIs* calculated from the total atom-based quadratic indices as the inputs while the non-linear model was an ensemble of neural networks (mt-QSAR-EL-ANN) using total atom-based quadratic indices as its inputs. The mt-QSAR-LDA model had *Sn* > 86% and *Sp* > 87%; for the mt-QSAR-EL-ANN model, *Sn* > 91% and *Sp* > 92% were obtained. Through the joint use of FBTD and the fragments' quantitative contributions, it was possible to design six molecules. These molecules were predicted by both mt-QSAR models as triple target inhibitors of BRD2, BRD3, and BRD4. To provide another in silico perspective, molecular docking calculations were performed to explore the potential binding mechanisms of the designed molecules. Molecular docking calculations confirmed the triple-target profile of the designed molecules; the interaction determined by molecular docking matched the physicochemical and structural insights provided by the FBTD methodology. The two most promising molecules were predicted by docking to exhibit more favorable binding energies than the reference ligands present protein-ligand complexes determined by x-ray (Figure 4).

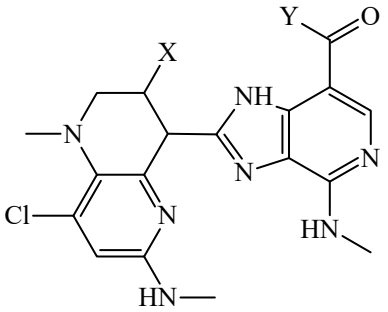


Figure 4. General structure of the two most promising multi-target inhibitors of BRD2, BRD3, and BRD4. The original codes of these molecules are ABD-001 and ABD-002. For the case of ABD-001, X = $–NH_2$ and Y = (pyridin-2-yl)oxidanyl; for the case of ABD-002, X = (pyridin-2-yl)oxidanyl and Y = $–NH_2$.

Three works were carried out on the applications of PTML modeling for de novo design in MTDD-based anticancer with emphasis on two of the cancers with the poorest prognoses: liver and pancreatic cancers. In this sense, these works mainly focused on designing chemicals with phenotypic multi-cell potency against different cancer cell lines. The first report aimed to discover multi-cell inhibitors of 17 liver cancer cell lines [105]. Here, an mt-QSAR-ANN model was built from 3079 data points and using *MLIs* derived from total and local atom-based quadratic indices as inputs. The mt-QSAR-ANN exhibited $Sn > 80\%$ and $Sp > 85\%$. The FBTD methodology permitted the physicochemical and structural interpretation of the mt-QSAR-ANN and the subsequent direct extraction of the molecular fragments responsible for the increase of the multi-cell anticancer activity. As a result, eight molecules were designed, with six of them being predicted as multi-cell inhibitors against the 17 liver cancer cell lines. The second and third work were focused on PTML modeling applied to the design of anti-pancreatic cancer agents. In the second work, an mt-QSAR-EL-ANN model was developed to speed up the design of multi-cell inhibitors against 31 pancreatic cancer cell lines [106]. To do so, the mt-QSAR-EL-ANN model used 5797 data points and *MLIs* obtained from the stochastic and non-stochastic (total and local) atom-based quadratic indices. The mt-QSAR-EL-ANN model achieved $Sn > 81\%$ and $Sp > 82\%$. Six novel molecules were designed through the application of the FBTD methodology; 4 out of six molecules were predicted by the mt-QSAR-EL-ANN model as multi-cell inhibitors against at least 28 out of 31 pancreatic cell lines. In the third work, two PTML models based on multilayer perceptron networks (PTML-MLP) were created by considering a dataset containing 9705 cases [107], which included two measures of activity endpoints b_m, 34 targets t_s (the 31 pancreatic cancer cell lines mentioned above and three pancreatic cancer-related proteins), and five labels of assay protocols a_p. For the creation of the first of these PTML-MLP models, the *MLIs* derived from bond spectral moments and atom- and bond-based connectivity indices were used as inputs; the second PTML-MLP model was based on *MLIs* calculated from the atom-based local stochastic quadratic indices. Both PTML-MLP models displayed $Sn > 78\%$ and $Sp > 84\%$. Although both PTML-MLP models were physicochemically and structurally interpreted, the FBTD methodology was applied only to the second of these PTML-MLP models. Consequently, the second PTML-MLP model was employed to design new molecules while the first PTML-MLP model was used as a filter to validate the design performed by the first model. Three molecules were designed and the three of them were predicted as multi-target inhibitors against pancreatic cancer (Figure 5); this means that both PTML-MLP models predicted the designed molecules as multi-protein inhibitors against the three pancreatic cancer-related proteins and multi-cell inhibitors against at least 30 of 31 pancreatic cancer cell lines.

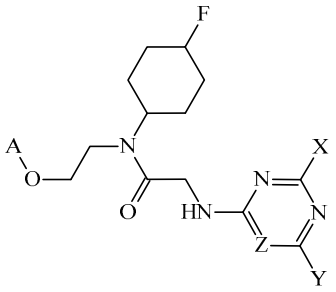


Figure 5. Chemical structure of three molecules designed as simultaneous multi-protein and multi-cell inhibitors against pancreatic cancer. The original codes of these molecules are MPMCI-001, MPMCI-002, and MPMCI-003. For the case of MPMCI-001, A = H, X = $–NH_2$, Y = H, and Z = C. For MPMCI-002, A = H, X = Y = $–NHCH_3$, and Z = C. For MPMCI-003, A = $–CH_3$, X = $–NHCH_3$, Y = H, and Z = N.

Last, the joint use of PTML modeling and FBTD was reported for the de novo generation of multi-cell inhibitors against different lung cancer cell lines [108]. In this work, the PTML-MLP, which was created from fifteen *MLIs* and achieved *Sn* > 77% and *Sp* in the interval 79–87.8%, was able to classify/predict the anti-lung cancer activity of a dataset formed by 7379 cases of molecules against nine different lung cancer cell lines (of varying degrees of drug sensitivity to current anticancer drugs) while also considering information regarding the ability of each lung cancer cell line to be sensitivity to immunotherapy [109]. The interpretation of the PTML-MLP model, which was carried out by utilizing the FBTD methodology, permitted the analysis of diverse molecular fragments responsible for the multi-cell inhibitory activity against the nine lung cancer cell lines, leading to the design of four new molecules belonging to two different chemical families (Figure 6). The designed molecules we confirmed as multi-cell anti-lung cancer agents by both the developed PTML-MLP model and CLC-Pred 2.0 [110]; the latter is a state-of-the-art webserver, which was built to predict anticancer activity against a panel of the 60 most known cancer cell lines.

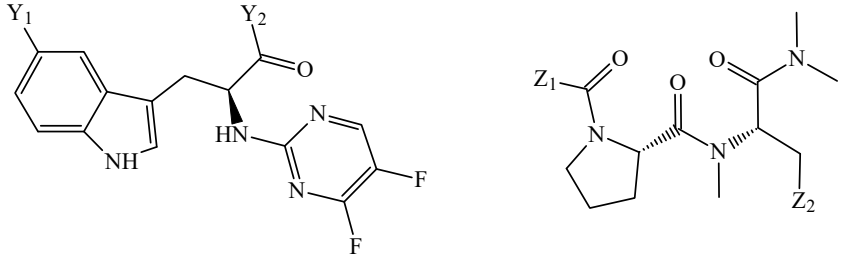


Figure 6. Two chemical families of molecules designed by combining a PTML-MLP model and FBTD. The original codes of these molecules were ASP-VALC-01, ASP-VALC-02, ASP-VALC-03, and ASP-VALC-04. The first chemical family (formed by ASP-VALC-01 and ASP-VALC-02) contained the substituents Y_1 and Y_2; Y_1 = H and Y_2 = 2-oxomorpholin-4-yl for the case of ASP-VALC-01 while Y_1 = OH and Y_2 = morpholin-4-yl for the case of ASP-VALC-02. The second chemical family (formed by ASP-VALC-03 and ASP-VALC-04) contained the substituents Z_1 and Z_2; Z_1 = 2,6-difluorophenyl and Z_2 = 4-hydroxyphenyl for the case of ASP-VALC-03 while Z_1 = pyridin-3-yl and Z_2 = 1H-indol-3-yl for the case of ASP-VALC-04.

3.4. Future Perspectives on PTML Modeling

As demonstrated in the previous subsections, PTML modeling is very useful in both target-based and phenotypic drug discoveries. In fact, in the context of de novo drug design, a major hurdle is that the designed molecules have not been able to dock well in the protein pockets [111]; however, PTML modeling, when combined with the FBTD methodology, has evolved to the point where the designed molecules have been confirmed by well-

established in silico approaches such as molecular docking [104,112], thus overcoming the aforementioned major hurdle.

In any case, there is always room for improvement and there are three main directions. One of them is, that, in the context of MTDD-based anticancer research, PTML modeling has focused mainly on the activity domain, providing deeper insights on the prediction and design of multi-protein and/or multi-cell inhibitors against many different types of cancer. Therefore, future efforts should be focused on developing mtk-QSBER models, which would enable the simultaneous prediction of multi-target anticancer activity while considering toxic effects and pharmacokinetic profiles at both in vitro and in vivo levels.

Second, PTML modeling combined with FBTD (see all the works reported in the previous subsection) has permitted to establish the theoretical foundations for the efficient computer-aided de novo design of chemicals with multi-protein and/or multi-cell inhibitory activity. To date, the majority of designed molecules have not yet reached experimental validation; thus, confirming the in vitro or in vivo anticancer efficacy of these novel molecules generated according to the de novo design paradigm remains a crucial next step for the field. Therefore, we recommend the synthesis and posterior biological evaluation to confirm the versatile anticancer activity of the designed molecules; this will consolidate PTML modeling (and subsequently, FBTD) as a powerful and innovative computational approach for antineoplastic discovery.

Last, the other main direction is that PTML modeling should also be devoted to exploring other chemical families beyond the realm of small organic molecules. This is the case of biosequence-based molecules such as peptides, micro-ribonucleic acids, and aptamers, which have emerged as potential therapeutic solutions in anticancer research [113–115].

4. Conclusions

Anticancer discovery continues to be an intensive research area, which requires the support of powerful in silico approaches to accelerate the discovery of versatile, effective, and safe MTDD-based chemotherapies. In this sense, how PTML models are mathematically conceived through the integration of chemical information with biological data, allows them to be used as tools for both virtual screening and de novo tasks across different biological endpoint measures, targets, and assay protocols. Consequently, we consider that PTML modeling considerably adds new and promising insights to the arsenal of in silico approaches for MTDD. We envisage that PTML modeling can be a great ally of modern drug development campaigns, rationalizing and prioritizing the identification/design of novel molecular entities in different MTDD scenarios.

Author Contributions: Conceptualization, A.S.-P.; methodology, A.S.-P. and V.V.K.; validation, A.S.-P.; formal analysis, A.S.-P. and V.V.K.; investigation, A.S.-P., V.V.K. and M.N.D.S.C.; resources, A.S.-P. and V.V.K.; writing—original draft preparation, A.S.-P., V.V.K. and M.N.D.S.C.; writing—review and editing, A.S.-P.; visualization, A.S.-P. and V.V.K.; supervision, A.S.-P.; funding acquisition, M.N.D.S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work received financial support from the PT national funds (FCT/MCTES, Fundação para a Ciência e Tecnologia and Ministério da Ciência, Tecnologia e Ensino Superior) through the project UID/50006—Laboratório Associado para a Química Verde—Tecnologias e Processos Limpos.

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

Abbreviations

The following abbreviations are used in this manuscript:

Acc	Accuracy
ANN	Artificial neural networks

a_p	Assay conditions or protocols
$avg[X]e_j$	Average value
b_m	Biological effects
BRD2	Bromodomain-containing protein 2
BRD3	Bromodomain-containing protein 3
BRD4	Bromodomain-containing protein 4
CDK4	Cyclin-dependent kinase 4
e_j	Experimental condition
FBTD	Fragment-based topological design
GF	Generic fragments
HER2	Human epidermal growth factor receptor 2
IC_{50}	Half-maximal inhibitory concentration
LDA	Linear discriminant analysis
logP	Logarithm of the n-octanol/water partition coefficient
MLIs	Multi-label indices; these are also denoted as $D[X]e_j$ in Equation (2) of this article
mt-QSAR	Multi-target QSAR model
mt-QSAR-ANN	Multi-target QSAR model based on an artificial neural network
mt-QSAR-EL-ANN	Multi-target QSAR model based on an ensemble of artificial neural networks
mt-QSAR-LDA	Multi-target QSAR model based on linear discriminant analysis
mtc-QSAR	Multi-condition QSAR
mtc-QSAR-LDA	Multi-condition QSAR based on linear discriminant analysis
MTDD	Multi-target drug discovery
mtk-QSBER	Multi-tasking model for quantitative structure-biological effect relationships
mtk-QSBER-LDA	Multi-tasking QSBER model based on linear discriminant analysis
$n(e_j)$	Number of chemicals that comply with a specific experimental aspect of e_j
Num	Numerator, which can be the range (difference between the maximum and minimum values of X), the standard deviation of X values, or a value of 1
$p(e_j)$	*A priori* probability of finding a chemical tested by considering a specific experimental aspect of e_j
PSA	Polar surface area
PTML	Perturbation-theory machine learning
PTML-ANN	PTML model based on an artificial neural network
PTML-LDA	PTML model based on linear discriminant analysis
PTML-MLP	PTML model based on a multilayer perceptron network
QSAR	Quantitative structure-activity relationships
t_s	Targets
RF	random forests
SMILES	Simplified molecular-input line-entry system
Sn	Sensitivity
Sp	Specificity
SPM	Statistical performance metrics
SVM	Support vector machines
X	Molecular descriptors
v_{sm}	Numerical value for a particular SPM
Y	Exponent, which can take the values of -1, -0.5, 0, 0.5, or 1

References

1. Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2024**, *74*, 229–263. [CrossRef] [PubMed]
2. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer Statistics, 2020. *CA Cancer J. Clin.* **2020**, *70*, 7–30. [CrossRef]

3. Dominguez-Valentin, M.; Nakken, S.; Tubeuf, H.; Vodak, D.; Ekstrom, P.O.; Nissen, A.M.; Morak, M.; Holinski-Feder, E.; Holth, A.; Capella, G.; et al. Results of multigene panel testing in familial cancer cases without genetic cause demonstrated by single gene testing. *Sci. Rep.* **2019**, *9*, 18555. [CrossRef] [PubMed]
4. Martin-Morales, L.; Rofes, P.; Diaz-Rubio, E.; Llovet, P.; Lorca, V.; Bando, I.; Perez-Segura, P.; de la Hoya, M.; Garre, P.; Garcia-Barberan, V.; et al. Novel genetic mutations detected by multigene panel are associated with hereditary colorectal cancer predisposition. *PLoS ONE* **2018**, *13*, e0203885. [CrossRef] [PubMed]
5. Mezina, A.; Philips, N.; Bogus, Z.; Erez, N.; Xiao, R.; Fan, R.; Olthoff, K.M.; Reddy, K.R.; Samadder, N.J.; Nielsen, S.M.; et al. Multigene Panel Testing in Individuals With Hepatocellular Carcinoma Identifies Pathogenic Germline Variants. *JCO Precis. Oncol.* **2021**, *5*, 988–1000. [CrossRef]
6. Hamilton, J.G.; Symecko, H.; Spielman, K.; Breen, K.; Mueller, R.; Catchings, A.; Trottier, M.; Salo-Mullen, E.E.; Shah, I.; Arutyunova, A.; et al. Uptake and acceptability of a mainstreaming model of hereditary cancer multigene panel testing among patients with ovarian, pancreatic, and prostate cancer. *Genet. Med.* **2021**, *23*, 2105–2113. [CrossRef]
7. Hu, C.; LaDuca, H.; Shimelis, H.; Polley, E.C.; Lilyquist, J.; Hart, S.N.; Na, J.; Thomas, A.; Lee, K.Y.; Davis, B.T.; et al. Multigene Hereditary Cancer Panels Reveal High-Risk Pancreatic Cancer Susceptibility Genes. *JCO Precis. Oncol.* **2018**, *2*, 1–28. [CrossRef]
8. Nussinov, R.; Tsai, C.J.; Jang, H. Anticancer drug resistance: An update and perspective. *Drug Resist. Updat.* **2021**, *59*, 100796. [CrossRef]
9. Zhong, L.; Li, Y.; Xiong, L.; Wang, W.; Wu, M.; Yuan, T.; Yang, W.; Tian, C.; Miao, Z.; Wang, T.; et al. Small molecules in targeted cancer therapy: Advances, challenges, and future perspectives. *Signal Transduct. Target Ther.* **2021**, *6*, 201. [CrossRef]
10. Hilal, T.; Gonzalez-Velez, M.; Prasad, V. Limitations in Clinical Trials Leading to Anticancer Drug Approvals by the US Food and Drug Administration. *JAMA Intern. Med.* **2020**, *180*, 1108–1115. [CrossRef]
11. Ramsay, R.R.; Popovic-Nikolic, M.R.; Nikolic, K.; Uliassi, E.; Bolognesi, M.L. A perspective on multi-target drug discovery and design for complex diseases. *Clin. Transl. Med.* **2018**, *7*, 3. [CrossRef] [PubMed]
12. Schaduangrat, N.; Lampa, S.; Simeon, S.; Gleeson, M.P.; Spjuth, O.; Nantasenamat, C. Towards reproducible computational drug discovery. *J. Cheminformatics* **2020**, *12*, 9. [CrossRef] [PubMed]
13. Brogi, S.; Ramalho, T.C.; Kuca, K.; Medina-Franco, J.L.; Valko, M. Editorial: In silico Methods for Drug Design and Discovery. *Front. Chem.* **2020**, *8*, 612. [CrossRef] [PubMed]
14. Pirhadi, S.; Damghani, T.; Avestan, M.S.; Sharifi, S. Dual potent c-Met and ALK inhibitors: From common feature pharmacophore modeling to structure based virtual screening. *J. Recept. Signal Transduct. Res.* **2020**, *40*, 357–364. [CrossRef]
15. He, Q.; Liu, C.; Wang, X.; Rong, K.; Zhu, M.; Duan, L.; Zheng, P.; Mi, Y. Exploring the mechanism of curcumin in the treatment of colon cancer based on network pharmacology and molecular docking. *Front. Pharmacol.* **2023**, *14*, 1102581. [CrossRef]
16. Khalid, H.R.; Aamir, M.; Tabassum, S.; Alghamdi, Y.S.; Alzamami, A.; Ashfaq, U.A. Integrated System Pharmacology Approaches to Elucidate Multi-Target Mechanism of Solanum surattense against Hepatocellular Carcinoma. *Molecules* **2022**, *27*, 6220. [CrossRef]
17. Batool, S.; Javed, M.R.; Aslam, S.; Noor, F.; Javed, H.M.F.; Seemab, R.; Rehman, A.; Aslam, M.F.; Paray, B.A.; Gulnaz, A. Network Pharmacology and Bioinformatics Approach Reveals the Multi-Target Pharmacological Mechanism of Fumaria indica in the Treatment of Liver Cancer. *Pharmaceuticals* **2022**, *15*, 654. [CrossRef]
18. Ahmed, B.; Khan, S.; Nouroz, F.; Farooq, U.; Khalid, S. Exploring multi-target inhibitors using in silico approach targeting cell cycle dysregulator-CDK proteins. *J. Biomol. Struct. Dyn.* **2022**, *40*, 8825–8839. [CrossRef]
19. Al-Khafaji, K.; Taskin Tok, T. Amygdalin as multi-target anticancer drug against targets of cell division cycle: Double docking and molecular dynamics simulation. *J. Biomol. Struct. Dyn.* **2021**, *39*, 1965–1974. [CrossRef]
20. Deng, Z.; Chen, G.; Shi, Y.; Lin, Y.; Ou, J.; Zhu, H.; Wu, J.; Li, G.; Lv, L. Curcumin and its nano-formulations: Defining triple-negative breast cancer targets through network pharmacology, molecular docking, and experimental verification. *Front. Pharmacol.* **2022**, *13*, 920514. [CrossRef]
21. Sharma, A.; Sinha, S.; Rathaur, P.; Vora, J.; Jha, P.C.; Johar, K.; Rawal, R.M.; Shrivastava, N. Reckoning apigenin and kaempferol as a potential multi-targeted inhibitor of EGFR/HER2-MEK pathway of metastatic colorectal cancer identified using rigorous computational workflow. *Mol. Divers.* **2022**, *26*, 3337–3356. [CrossRef] [PubMed]
22. Prabhavathi, H.; Dasegowda, K.R.; Renukananda, K.H.; Karunakar, P.; Lingaraju, K.; Raja Naika, H. Molecular docking and dynamic simulation to identify potential phytocompound inhibitors for EGFR and HER2 as anti-breast cancer agents. *J. Biomol. Struct. Dyn.* **2022**, *40*, 4713–4724. [CrossRef] [PubMed]
23. Elasbali, A.M.; Al-Soud, W.A.; Mousa Elayyan, A.E.; Al-Oanzi, Z.H.; Alhassan, H.H.; Mohamed, B.M.; Alanazi, H.H.; Ashraf, M.S.; Moiz, S.; Patel, M.; et al. Integrating network pharmacology approaches for the investigation of multi-target pharmacological mechanism of 6-shogaol against cervical cancer. *J. Biomol. Struct. Dyn.* **2023**, *41*, 14135–14151. [CrossRef] [PubMed]
24. De Simone, G.; Sardina, D.S.; Gulotta, M.R.; Perricone, U. KUALA: A machine learning-driven framework for kinase inhibitors repositioning. *Sci. Rep.* **2022**, *12*, 17877. [CrossRef]

25. Brindha, G.R.; Rishiikeshwer, B.S.; Santhi, B.; Nakendraprasath, K.; Manikandan, R.; Gandomi, A.H. Precise prediction of multiple anticancer drug efficacy using multi target regression and support vector regression analysis. *Comput. Methods Programs Biomed.* **2022**, *224*, 107027. [CrossRef]
26. Al Taweraqi, N.; King, R.D. Improved prediction of gene expression through integrating cell signalling models with machine learning. *BMC Bioinform.* **2022**, *23*, 323. [CrossRef]
27. Nguyen, L.C.; Naulaerts, S.; Bruna, A.; Ghislat, G.; Ballester, P.J. Predicting Cancer Drug Response In Vivo by Learning an Optimal Feature Selection of Tumour Molecular Profiles. *Biomedicines* **2021**, *9*, 1319. [CrossRef]
28. Simeon, S.; Ghislat, G.; Ballester, P. Characterizing the Relationship Between the Chemical Structures of Drugs and their Activities on Primary Cultures of Pediatric Solid Tumors. *Curr. Med. Chem.* **2021**, *28*, 7830–7839. [CrossRef]
29. Gonzalez-Diaz, H.; Arrasate, S.; Gomez-SanJuan, A.; Sotomayor, N.; Lete, E.; Besada-Porto, L.; Ruso, J.M. General theory for multiple input-output perturbations in complex molecular systems. 1. Linear QSPR electronegativity models in physical, organic, and medicinal chemistry. *Curr. Top. Med. Chem.* **2013**, *13*, 1713–1741. [CrossRef]
30. Speck-Planche, A.; Cordeiro, M.N.D.S. Multitasking models for quantitative structure-biological effect relationships: Current status and future perspectives to speed up drug discovery. *Expert Opin. Drug Discov.* **2015**, *10*, 245–256. [CrossRef]
31. Halder, A.K.; Moura, A.S.; Cordeiro, M.N.D.S. Moving Average-Based Multitasking In Silico Classification Modeling: Where Do We Stand and What Is Next? *Int. J. Mol. Sci.* **2022**, *23*, 4937. [CrossRef] [PubMed]
32. Kleandrova, V.V.; Cordeiro, M.N.D.S.; Speck-Planche, A. Optimizing drug discovery using multitasking models for quantitative structure-biological effect relationships: An update of the literature. *Expert Opin. Drug Discov.* **2023**, *18*, 1231–1243. [CrossRef] [PubMed]
33. Kleandrova, V.V.; Cordeiro, M.N.D.S.; Speck-Planche, A. Current in silico methods for multi-target drug discovery in early anticancer research: The rise of the perturbation-theory machine learning approach. *Future Med. Chem.* **2023**, *15*, 1647–1650. [CrossRef]
34. Kleandrova, V.V.; Cordeiro, M.N.D.S.; Speck-Planche, A. Perturbation-Theory Machine Learning for Multi-Objective Antibacterial Discovery: Current Status and Future Perspectives. *Appl. Sci.* **2025**, *15*, 1166. [CrossRef]
35. Velasquez-Lopez, Y.; Ruiz-Escudero, A.; Arrasate, S.; Gonzalez-Diaz, H. Implementation of IFPTML Computational Models in Drug Discovery Against Flaviviridae Family. *J. Chem. Inf. Model.* **2024**, *64*, 1841–1852. [CrossRef] [PubMed]
36. Dieguez-Santana, K.; Gonzalez-Diaz, H. Machine learning in antibacterial discovery and development: A bibliometric and network analysis of research hotspots and trends. *Comput. Biol. Med.* **2023**, *155*, 106638. [CrossRef]
37. Santiago, C.; Ortega-Tenezaca, B.; Barbolla, I.; Fundora-Ortiz, B.; Arrasate, S.; Dea-Ayuela, M.A.; Gonzalez-Diaz, H.; Sotomayor, N.; Lete, E. Prediction of Antileishmanial Compounds: General Model, Preparation, and Evaluation of 2-Acylpyrrole Derivatives. *J. Chem. Inf. Model.* **2022**, *62*, 3928–3940. [CrossRef]
38. Dieguez-Santana, K.; Casanola-Martin, G.M.; Torres, R.; Rasulev, B.; Green, J.R.; Gonzalez-Diaz, H. Machine Learning Study of Metabolic Networks vs ChEMBL Data of Antibacterial Compounds. *Mol. Pharm.* **2022**, *19*, 2151–2163. [CrossRef]
39. Vasquez-Dominguez, E.; Armijos-Jaramillo, V.D.; Tejera, E.; Gonzalez-Diaz, H. Multioutput Perturbation-Theory Machine Learning (PTML) Model of ChEMBL Data for Antiretroviral Compounds. *Mol. Pharm.* **2019**, *16*, 4200–4212. [CrossRef]
40. Nocedo-Mena, D.; Cornelio, C.; Camacho-Corona, M.D.R.; Garza-Gonzalez, E.; Waksman de Torres, N.; Arrasate, S.; Sotomayor, N.; Lete, E.; Gonzalez-Diaz, H. Modeling Antibacterial Activity with Machine Learning and Fusion of Chemical Structure Information with Microorganism Metabolic Networks. *J. Chem. Inf. Model.* **2019**, *59*, 1109–1120. [CrossRef]
41. Quevedo-Tumailli, V.; Ortega-Tenezaca, B.; Gonzalez-Diaz, H. IFPTML Mapping of Drug Graphs with Protein and Chromosome Structural Networks vs. Pre-Clinical Assay Information for Discovery of Antimalarial Compounds. *Int. J. Mol. Sci.* **2021**, *22*, 13066. [CrossRef]
42. Barbolla, I.; Hernandez-Suarez, L.; Quevedo-Tumailli, V.; Nocedo-Mena, D.; Arrasate, S.; Dea-Ayuela, M.A.; Gonzalez-Diaz, H.; Sotomayor, N.; Lete, E. Palladium-mediated synthesis and biological evaluation of C-10b substituted Dihydropyrrolo[1,2-b]isoquinolines as antileishmanial agents. *Eur. J. Med. Chem.* **2021**, *220*, 113458. [CrossRef]
43. Herrera-Ibata, D.M.; Pazos, A.; Orbegozo-Medina, R.A.; Romero-Duran, F.J.; Gonzalez-Diaz, H. Mapping chemical structure-activity information of HAART-drug cocktails over complex networks of AIDS epidemiology and socioeconomic data of U.S. counties. *Biosystems* **2015**, *132–133*, 20–34. [CrossRef] [PubMed]
44. Herrera-Ibata, D.M.; Orbegozo-Medina, R.A.; Gonzalez-Diaz, H. Multiscale mapping of AIDS in U.S. countries vs anti-HIV drugs activity with complex networks and information indices. *Curr. Bioinform.* **2015**, *10*, 639–657. [CrossRef]
45. Baltasar-Marchueta, M.; Llona, L.; M-Alicante, S.; Barbolla, I.; Ibarluzea, M.G.; Ramis, R.; Salomon, A.M.; Fundora, B.; Araujo, A.; Muguruza-Montero, A.; et al. Identification of Riluzole derivatives as novel calmodulin inhibitors with neuroprotective activity by a joint synthesis, biosensor, and computational guided strategy. *Biomed. Pharmacother.* **2024**, *174*, 116602. [CrossRef] [PubMed]
46. Sampaio-Dias, I.E.; Rodriguez-Borges, J.E.; Yanez-Perez, V.; Arrasate, S.; Llorente, J.; Brea, J.M.; Bediaga, H.; Vina, D.; Loza, M.I.; Caamano, O.; et al. Synthesis, Pharmacological, and Biological Evaluation of 2-Furoyl-Based MIF-1 Peptidomimetics and the Development of a General-Purpose Model for Allosteric Modulators (ALLOPTML). *ACS Chem. Neurosci.* **2021**, *12*, 203–215. [CrossRef] [PubMed]

47. Diez-Alarcia, R.; Yanez-Perez, V.; Muneta-Arrate, I.; Arrasate, S.; Lete, E.; Meana, J.J.; Gonzalez-Diaz, H. Big Data Challenges Targeting Proteins in GPCR Signaling Pathways; Combining PTML-ChEMBL Models and [(35)S]GTPgammaS Binding Assays. *ACS Chem. Neurosci.* **2019**, *10*, 4476–4491. [CrossRef]
48. Ferreira da Costa, J.; Silva, D.; Caamano, O.; Brea, J.M.; Loza, M.I.; Munteanu, C.R.; Pazos, A.; Garcia-Mera, X.; Gonzalez-Diaz, H. Perturbation Theory/Machine Learning Model of ChEMBL Data for Dopamine Targets: Docking, Synthesis, and Assay of New l-Prolyl-l-leucyl-glycinamide Peptidomimetics. *ACS Chem. Neurosci.* **2018**, *9*, 2572–2587. [CrossRef]
49. Abeijon, P.; Garcia-Mera, X.; Caamano, O.; Yanez, M.; Lopez-Castro, E.; Romero-Duran, F.J.; Gonzalez-Diaz, H. Multi-Target Mining of Alzheimer Disease Proteome with Hansch's QSBR-Perturbation Theory and Experimental-Theoretic Study of New Thiophene Isosters of Rasagiline. *Curr. Drug Targets* **2017**, *18*, 511–521. [CrossRef]
50. Romero-Duran, F.J.; Alonso, N.; Yanez, M.; Caamano, O.; Garcia-Mera, X.; Gonzalez-Diaz, H. Brain-inspired cheminformatics of drug-target brain interactome, synthesis, and assay of TVP1022 derivatives. *Neuropharmacology* **2016**, *103*, 270–278. [CrossRef]
51. He, S.; Segura Abarrategi, J.; Bediaga, H.; Arrasate, S.; Gonzalez-Diaz, H. On the additive artificial intelligence-based discovery of nanoparticle neurodegenerative disease drug delivery systems. *Beilstein J. Nanotechnol.* **2024**, *15*, 535–555. [CrossRef] [PubMed]
52. He, S.; Nader, K.; Abarrategi, J.S.; Bediaga, H.; Nocedo-Mena, D.; Ascencio, E.; Casanola-Martin, G.M.; Castellanos-Rubio, I.; Insausti, M.; Rasulev, B.; et al. NANO.PTML model for read-across prediction of nanosystems in neurosciences. computational model and experimental case of study. *J. Nanobiotechnol.* **2024**, *22*, 435. [CrossRef]
53. Diéguez-Santana, K.; Rasulev, B.; González-Díaz, H. Towards rational nanomaterial design by predicting drug–nanoparticle system interaction vs. bacterial metabolic networks. *Environ. Sci. Nano* **2022**, *9*, 1391–1413. [CrossRef]
54. Ortega-Tenezaca, B.; Gonzalez-Diaz, H. IFPTML mapping of nanoparticle antibacterial activity vs. pathogen metabolic networks. *Nanoscale* **2021**, *13*, 1318–1330. [CrossRef] [PubMed]
55. Munteanu, C.R.; Gutierrez-Asorey, P.; Blanes-Rodriguez, M.; Hidalgo-Delgado, I.; Blanco Liverio, M.J.; Castineiras Galdo, B.; Porto-Pazos, A.B.; Gestal, M.; Arrasate, S.; Gonzalez-Diaz, H. Prediction of Anti-Glioblastoma Drug-Decorated Nanoparticle Delivery Systems Using Molecular Descriptors and Machine Learning. *Int. J. Mol. Sci.* **2021**, *22*, 11519. [CrossRef]
56. Dieguez-Santana, K.; Gonzalez-Diaz, H. Towards machine learning discovery of dual antibacterial drug-nanoparticle systems. *Nanoscale* **2021**, *13*, 17854–17870. [CrossRef] [PubMed]
57. Urista, D.V.; Carrue, D.B.; Otero, I.; Arrasate, S.; Quevedo-Tumailli, V.F.; Gestal, M.; Gonzalez-Diaz, H.; Munteanu, C.R. Prediction of Antimalarial Drug-Decorated Nanoparticle Delivery Systems with Random Forest Models. *Biology* **2020**, *9*, 198. [CrossRef]
58. Santana, R.; Zuluaga, R.; Ganan, P.; Arrasate, S.; Onieva, E.; Montemore, M.M.; Gonzalez-Diaz, H. PTML Model for Selection of Nanoparticles, Anticancer Drugs, and Vitamins in the Design of Drug-Vitamin Nanoparticle Release Systems for Cancer Cotherapy. *Mol. Pharm.* **2020**, *17*, 2612–2627. [CrossRef]
59. Santana, R.; Zuluaga, R.; Ganan, P.; Arrasate, S.; Onieva, E.; Gonzalez-Diaz, H. Predicting coated-nanoparticle drug release systems with perturbation-theory machine learning (PTML) models. *Nanoscale* **2020**, *12*, 13471–13483. [CrossRef]
60. Tenorio-Borroto, E.; Castanedo, N.; Garcia-Mera, X.; Rivadeneira, K.; Vazquez Chagoyan, J.C.; Barbabosa Pliego, A.; Munteanu, C.R.; Gonzalez-Diaz, H. Perturbation Theory Machine Learning Modeling of Immunotoxicity for Drugs Targeting Inflammatory Cytokines and Study of the Antimicrobial G1 Using Cytometric Bead Arrays. *Chem. Res. Toxicol.* **2019**, *32*, 1811–1823. [CrossRef]
61. Vazquez-Prieto, S.; Paniagua, E.; Solana, H.; Ubeira, F.M.; Gonzalez-Diaz, H. A study of the Immune Epitope Database for some fungi species using network topological indices. *Mol. Divers.* **2017**, *21*, 713–718. [CrossRef] [PubMed]
62. Martinez-Arzate, S.G.; Tenorio-Borroto, E.; Barbabosa Pliego, A.; Diaz-Albiter, H.M.; Vazquez-Chagoyan, J.C.; Gonzalez-Diaz, H. PTML Model for Proteome Mining of B-Cell Epitopes and Theoretical-Experimental Study of Bm86 Protein Sequences from Colima, Mexico. *J. Proteome Res.* **2017**, *16*, 4093–4103. [CrossRef]
63. Tenorio-Borroto, E.; Penuelas-Rivas, C.G.; Vasquez-Chagoyan, J.C.; Castanedo, N.; Prado-Prado, F.J.; Garcia-Mera, X.; Gonzalez-Diaz, H. Model for high-throughput screening of drug immunotoxicity—Study of the anti-microbial G1 over peritoneal macrophages using flow cytometry. *Eur. J. Med. Chem.* **2014**, *72*, 206–220. [CrossRef] [PubMed]
64. Daghighi, A.; Casanola-Martin, G.M.; Iduoku, K.; Kusic, H.; Gonzalez-Diaz, H.; Rasulev, B. Multi-Endpoint Acute Toxicity Assessment of Organic Compounds Using Large-Scale Machine Learning Modeling. *Environ. Sci. Technol.* **2024**, *58*, 10116–10127. [CrossRef]
65. Todeschini, R.; Consonni, V. (Eds.) *Molecular Descriptors for Chemoinformatics*; WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2009; Volume I–II.
66. Estrada, E.; Gutiérrez, Y. *MODESLAB*, Santiago de Compostela, Spain, 2004; v1.5.
67. Todeschini, R.; Consonni, V.; Mauri, A.; Pavan, M. *DRAGON for Windows (Software for Molecular Descriptor Calculations)*, Milano Chemometrics and QSAR Research Group: Milano, Italy, 2005; v5.3.
68. Valdés-Martini, J.R.; García-Jacas, C.R.; Marrero-Ponce, Y.; Silveira Vaz 'd Almeida, Y.; Morell, C. QUBILs-MAS: Free Software for Molecular Descriptors Calculator from Quadratic, Bilinear and Linear Maps Based on Graph-Theoretic Electronic-Density Matrices and Atomic Weightings, v1.0. CAMD-BIR Unit, CENDA Registration Number: 2373-2012: Villa Clara, Cuba, 2012. Available online: https://tomocomd.com/ (accessed on 15 February 2025).

69. Medina Marrero, R.; Marrero-Ponce, Y.; Barigye, S.J.; Echeverria Diaz, Y.; Acevedo-Barrios, R.; Casanola-Martin, G.M.; Garcia Bernal, M.; Torrens, F.; Perez-Gimenez, F. QuBiLs-MAS method in early drug discovery and rational drug identification of antifungal agents. *SAR QSAR Environ. Res.* **2015**, *26*, 943–958. [CrossRef]
70. Valdes-Martini, J.R.; Marrero-Ponce, Y.; Garcia-Jacas, C.R.; Martinez-Mayorga, K.; Barigye, S.J.; Vaz d'Almeida, Y.S.; Pham-The, H.; Perez-Gimenez, F.; Morell, C.A. QuBiLS-MAS, open source multi-platform software for atom- and bond-based topological (2D) and chiral (2.5D) algebraic molecular descriptors computations. *J. Cheminformatics* **2017**, *9*, 35. [CrossRef]
71. TIBCO-Software-Inc. *STATISTICA (Data Analysis Software System)*, TIBCO-Software-Inc.: Palo Alto, CA, USA, 2018; v13.5.0.17.
72. Marzaro, G.; Chilin, A.; Guiotto, A.; Uriarte, E.; Brun, P.; Castagliuolo, I.; Tonus, F.; Gonzalez-Diaz, H. Using the TOPS-MODE approach to fit multi-target QSAR models for tyrosine kinases inhibitors. *Eur. J. Med. Chem.* **2011**, *46*, 2185–2192. [CrossRef] [PubMed]
73. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D.S. Fragment-based QSAR model toward the selection of versatile anti-sarcoma leads. *Eur. J. Med. Chem.* **2011**, *46*, 5910–5916. [CrossRef]
74. Speck-Planche, A.; Cordeiro, M.N.D.S. In *Bladder Cancer: Risk Factors, Emerging Treatment Strategies and Challenges*; Haggerty, S., Ed.; Multi-tasking chemoinformatic model for the efficient discovery of potent and safer anti-bladder cancer agents. Nova Science Publishers, Inc.: New York, NY, USA, 2014; pp. 71–94.
75. Pelon, M.; Krzeminski, P.; Tracz-Gaszewska, Z.; Misiewicz-Krzeminska, I. Factors determining the sensitivity to proteasome inhibitors of multiple myeloma cells. *Front. Pharmacol.* **2024**, *15*, 1351565. [CrossRef]
76. Velez, B.; Razi, A.; Hubbard, R.D.; Walsh, R.; Rawson, S.; Tian, G.; Finley, D.; Hanna, J. Rational design of proteasome inhibitors based on the structure of the endogenous inhibitor PI31/Fub1. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2308417120. [CrossRef]
77. Bennett, M.K.; Li, M.; Tea, M.N.; Pitman, M.R.; Toubia, J.; Wang, P.P.; Anderson, D.; Creek, D.J.; Orlowski, R.Z.; Gliddon, B.L.; et al. Resensitising proteasome inhibitor-resistant myeloma with sphingosine kinase 2 inhibition. *Neoplasia* **2022**, *24*, 1–11. [CrossRef] [PubMed]
78. Casanola-Martin, G.M.; Le-Thi-Thu, H.; Perez-Gimenez, F.; Marrero-Ponce, Y.; Merino-Sanjuan, M.; Abad, C.; Gonzalez-Diaz, H. Multi-output model with Box-Jenkins operators of linear indices to predict multi-target inhibitors of ubiquitin-proteasome pathway. *Mol. Divers.* **2015**, *19*, 347–356. [CrossRef] [PubMed]
79. Bediaga, H.; Arrasate, S.; Gonzalez-Diaz, H. PTML Combinatorial Model of ChEMBL Compounds Assays for Multiple Types of Cancer. *ACS Comb. Sci.* **2018**, *20*, 621–632. [CrossRef]
80. Cabrera-Andrade, A.; Lopez-Cortes, A.; Munteanu, C.R.; Pazos, A.; Perez-Castillo, Y.; Tejera, E.; Arrasate, S.; Gonzalez-Diaz, H. Perturbation-Theory Machine Learning (PTML) Multilabel Model of the ChEMBL Dataset of Preclinical Assays for Antisarcoma Compounds. *ACS Omega* **2020**, *5*, 27211–27220. [CrossRef] [PubMed]
81. Atz, K.; Cotos, L.; Isert, C.; Hakansson, M.; Focht, D.; Hilleke, M.; Nippa, D.F.; Iff, M.; Ledergerber, J.; Schiebroek, C.C.G.; et al. Prospective de novo drug design with deep interactome learning. *Nat. Commun.* **2024**, *15*, 3408. [CrossRef]
82. Wang, M.; Wang, Z.; Sun, H.; Wang, J.; Shen, C.; Weng, G.; Chai, X.; Li, H.; Cao, D.; Hou, T. Deep learning approaches for de novo drug design: An overview. *Curr. Opin. Struct. Biol.* **2022**, *72*, 135–144. [CrossRef]
83. Moret, M.; Pachon Angona, I.; Cotos, L.; Yan, S.; Atz, K.; Brunner, C.; Baumgartner, M.; Grisoni, F.; Schneider, G. Leveraging molecular structure and bioactivity with chemical language models for de novo drug design. *Nat. Commun.* **2023**, *14*, 114. [CrossRef]
84. Mouchlis, V.D.; Afantitis, A.; Serra, A.; Fratello, M.; Papadiamantis, A.G.; Aidinis, V.; Lynch, I.; Greco, D.; Melagraki, G. Advances in de Novo Drug Design: From Conventional to Machine Learning Methods. *Int. J. Mol. Sci.* **2021**, *22*, 1676. [CrossRef]
85. Kleandrova, V.V.; Speck-Planche, A. The QSAR Paradigm in Fragment-Based Drug Discovery: From the Virtual Generation of Target Inhibitors to Multi-Scale Modeling. *Mini Rev. Med. Chem.* **2020**, *20*, 1357–1374. [CrossRef]
86. Estrada, E.; Molina, E.; Perdomo-Lopez, I. Can 3D structural parameters be predicted from 2D (topological) molecular descriptors? *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1015–1021. [CrossRef]
87. Estrada, E. Physicochemical Interpretation of Molecular Connectivity Indices. *J. Phys. Chem. A* **2002**, *106*, 9085–9091. [CrossRef]
88. Estrada, E. Edge adjacency relationship and a novel topological index related to molecular volume. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 31–33. [CrossRef]
89. Estrada, E. Spectral moments of the edge adjacency matrix in molecular graphs. 1. Definition and applications for the prediction of physical properties of alkanes. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 844–849. [CrossRef]
90. Estrada, E. Spectral moments of the edge adjacency matrix in molecular graphs. 2. Molecules containing heteroatoms and QSAR applications. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 320–328. [CrossRef]
91. Estrada, E. Spectral moments of the edge adjacency matrix in molecular graphs. 3. Molecules containing cycles. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 23–27. [CrossRef]
92. Estrada, E.; Pena, A.; Garcia-Domenech, R. Designing sedative/hypnotic compounds from a novel substructural graph-theoretical approach. *J. Comput. Aided Mol. Des.* **1998**, *12*, 583–595. [CrossRef]
93. Kier, L.B.; Hall, L.H. *Molecular Connectivity in Structure-Activity Analysis*; John Wiley & Sons: New York, NY, USA, 1986.

94. Baskin, I.I.; Skvortsova, M.I.; Stankevich, I.V.; Zefirov, N.S. On the basis of invariants of labeled molecular graphs. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 527–531. [CrossRef]
95. Baskin, I.; Varnek, A. In *Chemoinformatics Approaches to Virtual Screening*; Varnek, A., Tropsha, A., Eds.; Fragment descriptors in SAR/QSAR/QSPR studies, molecular similarity analysis and in virtual screening. Royal Society of Chemistry: Cambridge, UK, 2008; pp. 1–43.
96. Kleandrova, V.V.; Cordeiro, M.N.D.S.; Speck-Planche, A. In Silico Approach for Antibacterial Discovery: PTML Modeling of Virtual Multi-Strain Inhibitors Against *Staphylococcus aureus*. *Pharmaceuticals* **2025**, *18*, 196. [CrossRef]
97. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D.S. Multi-target drug discovery in anti-cancer therapy: Fragment-based approach toward the design of potent and versatile anti-prostate cancer agents. *Bioorg. Med. Chem.* **2011**, *19*, 6239–6244. [CrossRef]
98. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D.S. Chemoinformatics in anti-cancer chemotherapy: Multi-target QSAR model for the in silico discovery of anti-breast cancer agents. *Eur. J. Pharm. Sci.* **2012**, *47*, 273–279. [CrossRef]
99. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D.S. Chemoinformatics in multi-target drug discovery for anti-cancer therapy: In silico design of potent and versatile anti-brain tumor agents. *Anticancer Agents Med. Chem.* **2012**, *12*, 678–685. [CrossRef] [PubMed]
100. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D.S. Rational drug design for anti-cancer chemotherapy: Multi-target QSAR models for the in silico discovery of anti-colorectal cancer agents. *Bioorg. Med. Chem.* **2012**, *20*, 4848–4855. [CrossRef] [PubMed]
101. Speck-Planche, A.; Kleandrova, V.V.; Luan, F.; Cordeiro, M.N.D.S. Unified multi-target approach for the rational in silico design of anti-bladder cancer agents. *Anticancer Agents Med. Chem.* **2013**, *13*, 791–800. [CrossRef]
102. Speck-Planche, A.; Cordeiro, M. Fragment-based in silico modeling of multi-target inhibitors against breast cancer-related proteins. *Mol. Divers.* **2017**, *21*, 511–523. [CrossRef] [PubMed]
103. Kleandrova, V.V.; Scotti, M.T.; Scotti, L.; Speck-Planche, A. Multi-Target Drug Discovery Via PTML Modeling: Applications to the Design of Virtual Dual Inhibitors of CDK4 and HER2. *Curr. Top. Med. Chem.* **2021**, *21*, 661–675. [CrossRef]
104. Speck-Planche, A.; Scotti, M.T. BET bromodomain inhibitors: Fragment-based in silico design using multi-target QSAR models. *Mol. Divers.* **2019**, *23*, 555–572. [CrossRef]
105. Kleandrova, V.V.; Scotti, M.T.; Scotti, L.; Nayarisseri, A.; Speck-Planche, A. Cell-based multi-target QSAR model for design of virtual versatile inhibitors of liver cancer cell lines. *SAR QSAR Environ. Res.* **2020**, *31*, 815–836. [CrossRef]
106. Speck-Planche, A. Multicellular Target QSAR Model for Simultaneous Prediction and Design of Anti-Pancreatic Cancer Agents. *ACS Omega* **2019**, *4*, 3122–3132. [CrossRef]
107. Kleandrova, V.V.; Speck-Planche, A. PTML Modeling for Pancreatic Cancer Research: In Silico Design of Simultaneous Multi-Protein and Multi-Cell Inhibitors. *Biomedicines* **2022**, *10*, 491. [CrossRef]
108. Kleandrova, V.V.; Cordeiro, M.N.D.S.; Speck-Planche, A. Perturbation Theory Machine Learning Model for Phenotypic Early Antineoplastic Drug Discovery: Design of Virtual Anti-Lung-Cancer Agents. *Appl. Sci.* **2024**, *14*, 9344. [CrossRef]
109. Budczies, J.; Kazdal, D.; Menzel, M.; Beck, S.; Kluck, K.; Altbürger, C.; Schwab, C.; Allgäuer, M.; Ahadova, A.; Kloor, M.; et al. Tumour mutational burden: Clinical utility, challenges and emerging improvements. *Nat. Rev. Clin. Oncol.* **2024**, *21*, 725–742. [CrossRef]
110. Lagunin, A.A.; Rudik, A.V.; Pogodin, P.V.; Savosina, P.I.; Tarasova, O.A.; Dmitriev, A.V.; Ivanov, S.M.; Biziukova, N.Y.; Druzhilovskiy, D.S.; Filimonov, D.A.; et al. CLC-Pred 2.0: A Freely Available Web Application for In Silico Prediction of Human Cell Line Cytotoxicity and Molecular Mechanisms of Action for Druglike Compounds. *Int. J. Mol. Sci.* **2023**, *24*, 1689. [CrossRef]
111. Cieplinski, T.; Danel, T.; Podlewska, S.; Jastrzebski, S. Generative Models Should at Least Be Able to Design Molecules That Dock Well: A New Benchmark. *J. Chem. Inf. Model.* **2023**, *63*, 3238–3247. [CrossRef] [PubMed]
112. Kleandrova, V.V.; Scotti, L.; Bezerra Mendonça Junior, F.J.; Muratov, E.; Scotti, M.T.; Speck-Planche, A. QSAR Modeling for Multi-Target Drug Discovery: Designing Simultaneous Inhibitors of Proteins in Diverse Pathogenic Parasites. *Front. Chem.* **2021**, *9*, 634663. [CrossRef] [PubMed]
113. Chinnadurai, R.K.; Khan, N.; Meghwanshi, G.K.; Ponne, S.; Althobiti, M.; Kumar, R. Current research status of anti-cancer peptides: Mechanism of action, production, and clinical applications. *Biomed. Pharmacother.* **2023**, *164*, 114996. [CrossRef] [PubMed]
114. Gurbuz, N.; Ozpolat, B. MicroRNA-based Targeted Therapeutics in Pancreatic Cancer. *Anticancer Res.* **2019**, *39*, 529–532. [CrossRef]
115. Venkatesan, S.; Chanda, K.; Balamurali, M.M. Recent Advancements of Aptamers in Cancer Therapy. *ACS Omega* **2023**, *8*, 32231–32243. [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.