UNIVERSITY OF MEDICINE AND PHARMACY
"CAROL DAVILA", BUCHAREST
DOCTORAL SCHOOL
PHARMACY FIELD

CHEMOINFORMATIC STUDIES AND MOLECULAR DESIGN STRATEGIES USED TO IDENTIFY NEW ANTITUMOR COMPOUNDS

SUMMARY OF THE DOCTORAL THESIS

PhD Supervisor:
PROF. UNIV. DR. NIȚULESCU GEORGE MIHAI

PhD student:
ION GEORGE NICOLAE DANIEL

2022

Cuprinsul tezei de doctorat

Introduction

According to the World Health Organization, oncological pathologies are currently one of the conditions with the greatest negative impact globally, both because of the diversity of cancer forms and because of the versatility of the ways of installing resistance to treatment, thus being a priority objective in the field of drug discovery.

The choice of doctoral topic was justified by the constant need to discover new antitumor inhibitors with high therapeutic efficiency and low systemic toxicity.

The importance of the chosen theme is given by the physiological impact of protein-kinases in cell signaling and functioning, these enzymes being involved in the regulation of crucial processes that control the cell cycle, proliferation, tumor survival but also the development of chemoesistance mechanisms. Protein-kinases have an important role in tumor development, in most oncological pathologies various quantitative or qualitative dysfunctions are reported.

An intensely studied therapeutic target at present is represented by Pim-1 kinase, a protein-kinase that phosphorylates a wide variety of protein substrates, involved both in tumorigenesis and metastasis and in the onset of chemoresthesinescence. Inhibition of this enzyme has been shown to be associated with tumor apoptosis, decreased cell proliferation and combating resistance to chemotherapy in various types of cancer.

The general objective of the thesis is represented by the investigation of the characteristics of the existing protein-kinase inhibitors, approved or under study, with the aim of optimizing the identification and development of new potent and selective inhibitors of Pim-1 kinase, applicable in the oncological treatment.

To this end, this research has focused on the development of statistical and chemoinformatic methods for analyzing the data made available online by the National Cancer Institute of the United States of America (NCI) about numerous compounds with antiproliferative activity on a collection of different cancer cell lines ("*NCI-60 panel"*).

The current study aims to find a use of the data provided by the NCI, since *in vitro* tests have already been carried out for these compounds, favoring the avoidance of quite high time and financial costs.

The development of methods for determining a correlation between the physico-chemical characteristics, the antiproliferative profile NCI-60 and the pharmacological mechanism of action of a compound is an important element of originality of the thesis. This

method allows the future identification of new inhibitors of Pim-1 kinases, with molecular skeletons different from those of the inhibitors described so far in the literature.

In order to achieve the proposed objective, the research methodology involves the following steps:

- identification of a specific antiproliferative profile for protein-kinase inhibitors, useful for differentiating other antitumor mechanisms of anticancer drugs,
- identification of structural and physico-chemical characteristics specific to molecules with increased antiproliferative activity,
- virtual screening of pim-1 kinase inhibitory compounds by molecular docking studies,
- development and optimization of a method of predicting *in vitro* activity based on *in silico* results from molecular dostation,
- identification of compounds with promising activity of inhibiting Pim-1 kinase following a therapeutic reorientation study based on developed methods.

The doctoral thesis is structured in two sections, the theroetic part, which presents the current state of knowledge in the field and the experimental part, which presents the personal researches carried out in order to identify potential inhibitors of Pim-1 kinase.

The first three chapters achieve a systematization of the literature, describing protein-kinases and their inhibitors approved in oncological treatment, as well as Pim-1 kinase as a particular case, along with recent discoveries made to inhibit its activity.

The experimental part begins in Chapter 4 with the development of a method of prediction and identification of the antiproliferative profile specific to protein-kinase inhibitors for compounds based on NCI, based on statistical methods, comparing 18 protein-kinase inhibitors and 80 oncological drugs with another known pharmacological mechanism.

Chapter 5 describes the exploration of the structural and physico-chemical characteristics of the compounds in the database and the development of a method of analyzing the basic structures for over 90000 compounds, with the obtaining of information about the ability of certain chemical structures to inhibit cell proliferation more strongly compared to others.

Chapter 6 describes a molecular docking study on a set of protein-kinase inhibitors on the chosen therapeutic target, Pim-1 kinase, optimized based on analysis of how the

algorithm interprets the interactions between compounds and protein by the doping algorithm. The developed method represents an important element of originality of the thesis by implementing a predictive score of *in vitro* activity based on *in silico* results that takes into account the interactions favorable to an increased inhibitory activity.

The last chapter brings together the three methods of analysis developed, through the therapeutic reorientation (*drug repurposing*) of 22 drugs from the DrugBank database with antiproliferative activity on cancer cell lines. The virtual screening was completed with the identification of potential inhibitors of Pim-1 kinase, phenothiazine and butyrophenonic nucleus derivatives, confirmed by recent results from the literature.

Further studies can be conducted for the thorough exploration of the Inhibitory Potential on Pim-1 Kinase of the compounds NSC177380, NSC186946 and NSC185040, predicted by the method developed as the most promising in terms of *in vitro* inhibition of the enzyme. The compounds NSC35949 and NSC1760, drugs already approved in the treatment of other pathologies, have also been identified as suitable for therapeutic reorientation towards the inhibition of Pim-1 kinase with antitumor effect.

The interdisciplinary character of the doctoral thesis of the research carried out results from the use of statistical methods for the identification of outlier values (extreme) of the antiproliferative effect, the use of chemoinformatic methods for the structural characterization of the studied compounds, the use of knowledge in the field of biochemistry and cell biology for the interpretation of data collected from the NCI base available online and the use of *in silico* screening methods for the assessment of the pharmacological potential of the studied inhibitors and the interpretation of the structure-action relationships reported in the literature or obtained as a result of molecular docking.

The developed methods, applicable to any therapeutic target involved in cell proliferation, can be improved to increase accuracy by incorporating more data on the compounds on the basis of which the mathematical formulas for predicting activity were generated, which can be achieved in subsequent studies to identify several potential anticancer drugs.

I. The general part

1. Protein-kinases as targets in antitumor therapy

Cancer is a widespread disease globally, being one of the first causes of death in the world, with an increased socio-economic impact [1].

The identification of oncogenic biomolecular disturbances and the development of personalized anticancer therapy is a topical topic in the medical field [2], and protein-kinases are one of the most studied therapeutic targets, being involved in the regulation of the processes that control the cell cycle, proliferation, tumor survival, but also the development of chemoresistance mechanisms [3].

The totality of protein kinases encoded in the human genome, known as the human kinomial, makes up about 2% of the sequences encoded in total at the genetic level and comprises 518 protein-kinases, of which 478 have similar genetic sequences, classical to eukaryotes, and another 40 are considered atypical protein-kinases [3], [4] . Depending on the nature of the phosphorylated protein substrate, enzymes are divided into tyrosine-kinases and serin-threonine kinases.

Various protein-kinase inhibitors approved for antitumor therapy have already proven their effectiveness, having better results compared to classical cytotoxic therapy, with an increase in survival in various malignancies [5], [6].

With the approval in 2001 of the first inhibitor that specifically targeted a protein-kinase, imatinib, for the treatment of chronic myeloid leukemia (CML), the way for the development of this class of drugs was opened, by the beginning of 2021 being approved over 70 protein-kinase inhibitors [6].

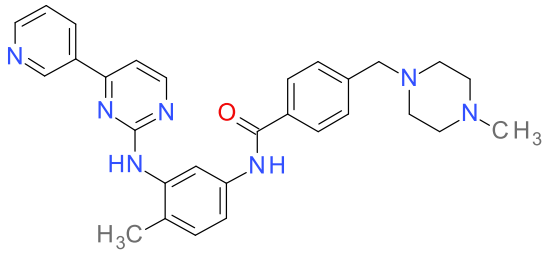

Fig. 1. 1 – Chemical structure of the first approved protein-kinase inhibitor, imatinib (CGP57148B)

Currently, the development of selective and potent inhibitors of protein-kinases has evolved significantly from the synthesis of staurosporin analogues (the first inhibitor investigated) to structurally based rational design methods using numerous modern

techniques, such as molecular docking studies, crystallography, nuclear magnetic resonance spectroscopy, and since the approval of imatinib in 2001 until now, more than 10000 applications have been submitted for the patenting of inhibitors of protein-kinases only in the United States of America [7].

Depending on the interaction with the target enzyme, protein-kinase inhibitors can be divided into classes, based on the type and position of macromolecule binding. Type I, $I^{1/2}$ and II inhibitors bind the pocket formed by the hydrogen bonds between adenine and the "hinge" region of protein kinase, while type III and IV inhibitors are allosteric inhibitors, and type V is represented by bivalent inhibitors, which bind to two different regions of kinase. Type VI compounds are irreversible inhibitors, binding covalently to the enzyme [5,6].

2. Pim-1 kinase as a therapeutic target in cancer

The family of protein-kinases Pim, named after the oncogenics of the same name (engl. "*proviral integration site for Moloney murine leukemia virus"*), belongs to the class of serine/threonine kinases (STPK) and comprises three main forms, Pim-1, Pim-2 and Pim-3. Representatives of this family are proteins with major involvement in tumor proliferation, playing important roles in the appearance of various types of cancer [8].

The three oncoproteins are part of the calcium/calmodulin-dependent protein-kinase family (engl). *Calcium/calmodulin-dependent protein kinase*, CAMK), with enzymatic framing EC (engl. *enzyme classification*) 2.7.11.1 [9]. The position of Pim-1 kinase in relation to the CAMK group in the human kinom is highlighted in Fig. 2. 1, with the help of the online platform KinMap (http://www.kinhub.org/kinmap/) [10].

Fig. 2. 1 – Classification of CAMK kinases encoded in the human kinomial. The pim-1 position of the human kinase is evidencedby the yellow and red marker. Classification illustration made using the online platform KinMap http://www.kinhub.org/kinmap/ [10].

The most relevant member of the Pim family for tumor pathologies is Pim-1, with high importance in signaling oncogenic processes, and is also the most widespread. The aberrant expression of this kinase has been reported to be closely related to numerous malignant pathologies, such as prostate cancer, pancreatic cancer, colon cancer, chronic lymphocytic leukemia, non-Hodgkin's lymphoma, as well as multiple myeloma [11].

The PIM-1 oncogene is located on the short arm of chromosome 6 and encodes 313 (the pim-1S isoform) and 404 amino acids (pim-1L isoform), the sequence being divided along 6 exons and 5 introns [12].

For the most part, the Pim-1protein has a typicalprotein-kinase structure, with the kinase domain consisting of two lobes, N-terminal and C-terminal, separated by a "hinge" region. *hinges*). Pim-1 possesses a structural characteristic that differentiates this protein from other protein-kinases by a different architecture of the "hinge" region that unitesthe two lobes andinfluencesthe binding of ATP, which allows obtaining inhibitors of Pim-1 kinase with high selectivity [11].

The three-dimensional structure of the Pim-1 protein crystallized with the staurosporin ligand is shown in Fig. 2. 2, together with the structural regions characteristic of the kinasic domain, as obtained and described by Arrouchi et al. [11].

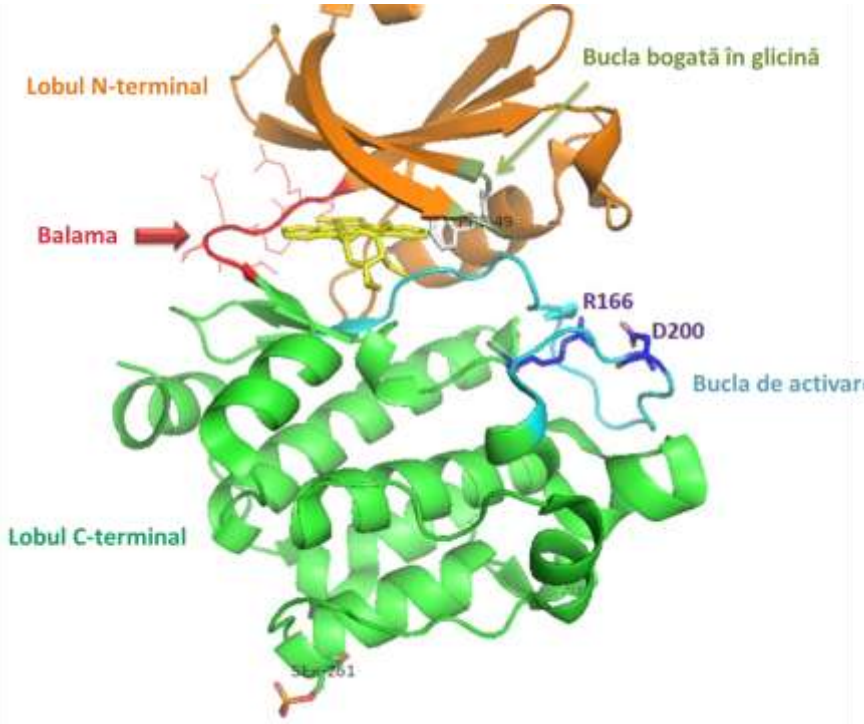

Fig. 2. 2 – Three-dimensional structure of Pim-1 kinase (PDB code: 1YHS) crystallized with the staurosporin ligand (highlighted in yellow). Image taken after Arrouchi et al. [11].

In Fig. 2. 3 The most important pim-mediated processes of Pim-1 kinase can be observed, along with the substrates or metabolic pathways with which it interacts and the involvement in all cellular functions [13].

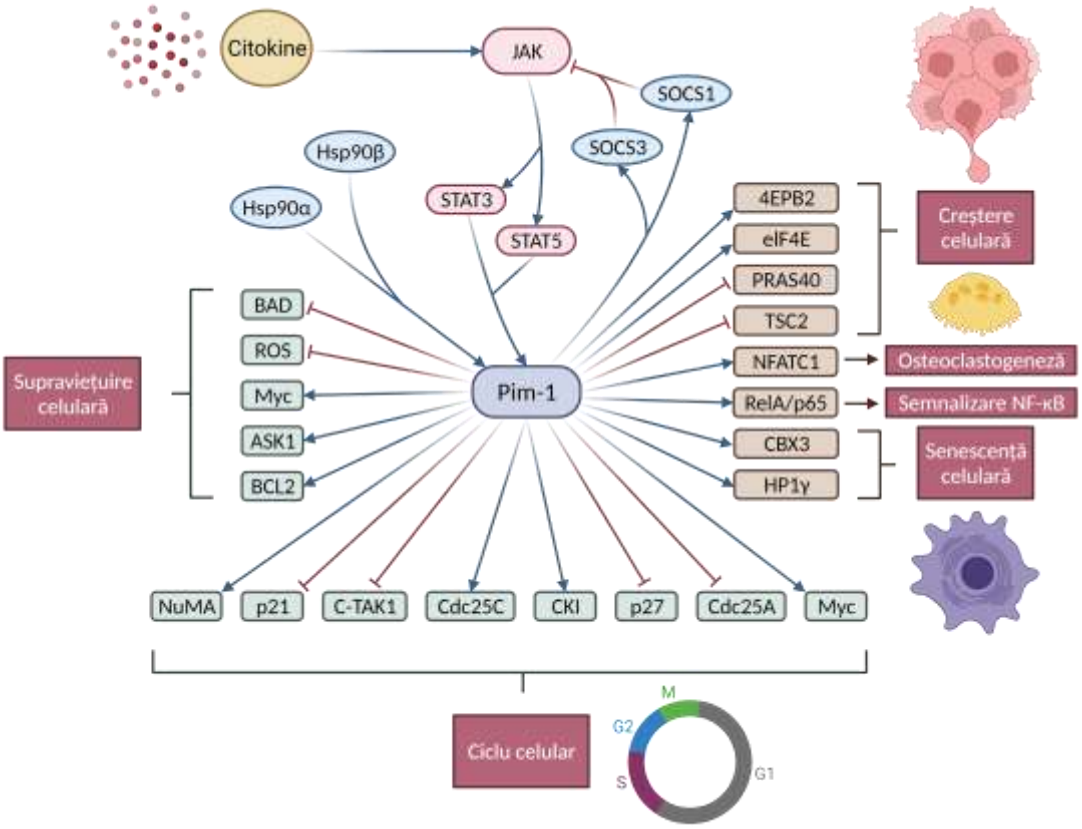

Fig. 2. 3 – Regulation and involvement of Pim-1 kinase in the cell signaling pathways, with the different metabolic pathways in which it is involved [13]. Image taken using BioRender

Of particular importance was attributed to Pim-1 kinase in the occurrence of resistance to chemotherapy in numerous types of cancer [14] with the main suspected mechanism in resistance to antitumor drugs represented by inhibition of apoptosis induced by proapoptotic factor p53.

3. Pim-1 kinase inhibitors

Hematological malignancies, but also prostate cancer and breast cancer, are the main pathologies in which inhibition of Pim-1 kinase exerts a significant impact on the evolution of the disease, so that numerous studies of the development and testing of some Pim-1 inhibitors are carried out on these types of cancers. The potential to reduce resistance to chemotherapy by joining some inhibitors of Pim-1 kinase to oncological treatment is promising, which further increases the value of inhibition of the enzyme in anticancer therapy.

Although to date of current research no inhibitor of Pim-1 kinase has been approved, numerous compounds with inhibitory activity on this therapeutic target are described in the literature. From the point of view of protein binding, Pim-1 kinase inhibitors are of the ATP-competitive type and are divided into ATP-mimetic inhibitors and non-ATP-mimetic inhibitors. A third category of inhibitors includes ligands acting on Pim-1 by binding mechanisms characteristic of both groups.

Given the high chemical diversity of Pim kinase inhibitors reported in the literature to date, a division can be made according to the basic structures. Among the chemical classes of Pim inhibitors are heterocycles such as indolocarbazoles, bisindolilmaleimids, naphthyridines, pyridazines, isoxazoles, thiazolidin-2-4-diones, tienopyimidinones, pyridones, isoxazoloquines, pyridine and pyrimidine derivatives, organometallic complexes, but also other types of structures [11], [15]

There are a limited number of clinical trials investigating the targeted inhibition of Pim kinases, most of which are on hematological malignancies. Some of the inhibitors arriving in phase I or II clinical trials with promising results are SGI-1776, SMI-4a, and TP-3654, non-selective Pim inhibitors with imidazopyridinic nucleus, studied in relapsed hematological malignancies or refractory to treatment, but also in her-2 positive breast cancer [16]. At the same time, the derivative of thiazolidin-2,4-dione, AZD1208, studied in

triple-negative breast cancer [17] and the aminopyridine derivatives PIM44 [18] and INCB053914 [19], with potential in the treatment of hematological malignancies, are noted.

II. Personal contributions

4. Investigations on the antiproliferative profile of antitumor compounds in the NCI-60 database

In the '90s, the U.S. National Cancer Institute (NCI) established a collection of 60 human cancer cell lines ("*NCI-60 panel*"), representing nine main types of cancerous tissues (brain, blood and bone marrow, breast, colon, kidney, lung, ovary, prostate, and skin). In recent decades, compounds discovered by various groups of researchers around the world have been tested using this NCI-60 collection to determine their proliferation inhibiting effect on different cell lines.

The objective of this research is the development of a simple but effective method of determining a correlation between the antiproliferative "imprint" NCI-60 and the pharmacological mechanism of action of a compound, with the aim of identifying potential protein-kinase inhibitors (PKI). For this purpose, a number of 9137 compounds were collected from the NCI base. After cleaning and filtering the data, a set of 18 protein-kinase inhibitors was formed, which constituted the PKI group, and an AOD (approved oncological drugs) set containing 80 drugs with a different mechanism of action than the inhibition of protein-kinases. The results expressed as logarithmic values from semi-maximal proliferation inhibitory concentrations, pGI_{50}, calculated from μM concentrations, have been used for the assessment of antiproliferative activity.

Descriptive analyses of the entire set were performed in order to identify the outlier values, which indicated the cell lilines that were more sensitive or resistant to the action of a compound. An example is the compound identified by the code NSC652680, with 4 outlier values for pIC50 whose distribution is shown in Fig. Fig. 4. 1.

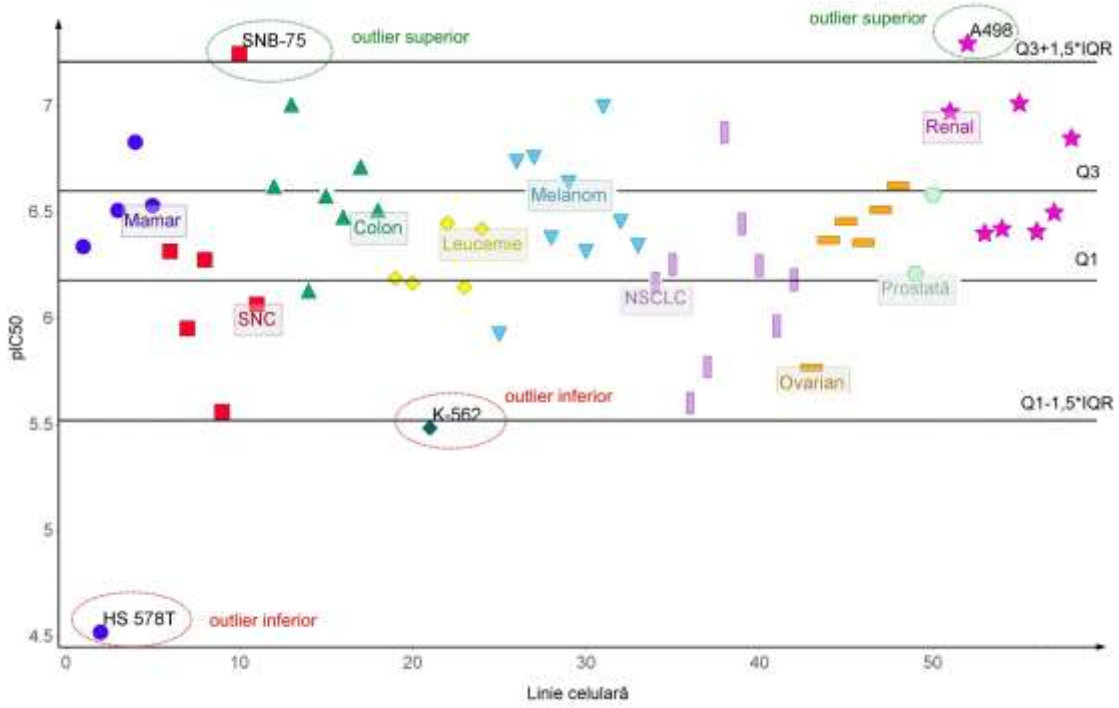

Fig. 4. 1 – Distribution of the outlier values of pIC_{50} on the 60 cell lines in the NCI-60 panel for the NSC652680 compound

The analysis of the studied sets from the point of view of the antiproliferative profiles revealed the existence of a growth inhibition pattern characteristic of the PKI set. The results showed that certain types of cells in the NCI-60 cell line panel are more sensitive than the rest to the action of protein-kinase inhibitors, in relation to the 80 antitumor drugs with which they were compared (AOD), which is expressed by the number of outlier values for those cell lines. Following the analysis of the number of upper outlier values for each of the 60 cell lines, a formula for predicting the mechanism of action of a compound was developed, based on the antiproliferative profile displayed under the NCI.

The frequency of the outlier values in each of the groups of compounds was used to calculate a weight factor (hereinafter also referred to as the 'cell line coefficient'). Thus, for each cell line, the frequencies of occurrence of the upper and lower type outlier values in the PKI and AOD sets, respectively, were calculated separately. A coefficient was calculated for each cell line using the difference between the frequency of occurrence in the PKI set and AOD, respectively, according to the mathematical formulas:

$$WUi = (\frac{uPKI_i}{nPKI} - \frac{uAOD_i}{nAOD}) * 100$$

$$WLi = (\frac{lPKI_i}{nPKI} - \frac{lAOD_i}{nAOD}) * 100$$

where WU_i represents the upper weight factor (engl. "*Weight Upper*"), and WL_i the lower weight factor (engl. "*Weight Lower*") for cell line *i*. The parameters $uPKI_i$ and $uAOD_i$ represent the number of values of the upper outlier type (engl. "*upper*") detected in the PKI set, respectively in the AOD set for cell line *i*. Analogously, $lPKI_i$ and $lAOD_i$ parameters can also be explained, with reference to the lower outlier values (engl. "*lower*"). nPKI and nAOD are the total number of compounds that make up each of the two sets.

Thus, the weight factor for each cell line is equivalent to the probability of a pGI50 result being an outlier value in the PKI set, rather than the AOD set. Based on the presence of an outlier value, coded by the values 1 (present) or 0 (absent) of the variables U_i ("*upper outlier*") and L_i ("*upper outlier*"), a predictive score of the potential to inhibit a protein-kinase was calculated, based on the formula:

$$S_c = \sum_{i=1}^{60} (WU_i * U_i + WL_i * L_i)$$

Values greater than 10.21 for the score calculated using the predictive formula indicated an increased likelihood of inhibiting a protein kinase, while lower values indicated a reduced probability. Scores for the 9137 compounds in the test set ranged from −74.58 to 118.33, and 409 compounds in the studied set were predicted to act by inhibiting protein-kinases, obtaining values greater than 10.21 for the predictive score. Following the analysis of the TBR for internal validation, the developed method was characterised as having a sensitivity value of 0,833 and a value for 1-specificity of 0,05.

The first 10 compounds in terms of calculated predictive scores are shown in Fig. 4. 3 depending on the antiproliferative profiles, compared to the set of 18 PKI used to establish the predictive model.

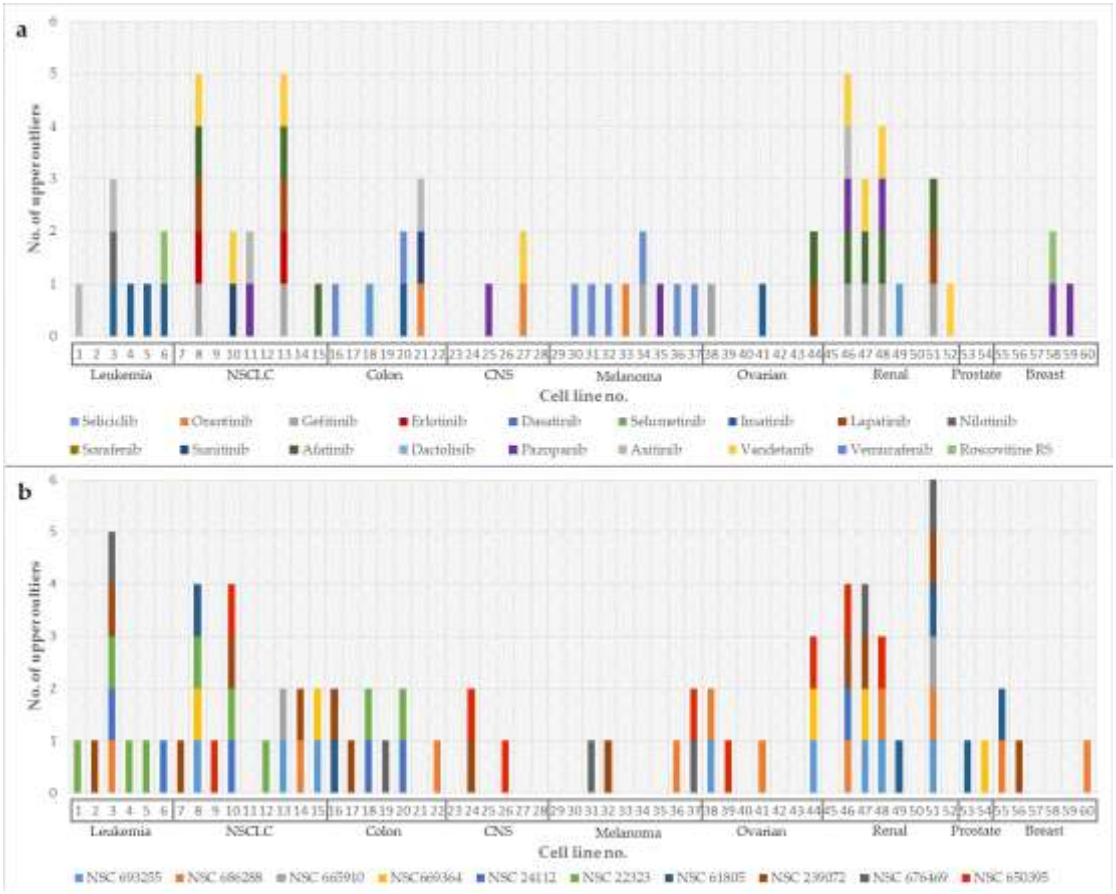

Fig. 4. 2 - (**a**) Distribution of upper outlier pGI50 values obtained for cell lines in the NCI-60 collection for the 18 PKI in the predictive set. **(b)** Distribution of the upper outlier pGI50 values obtained for NCI-60 cancer cell lines for compounds with the highest 10 scores, illustrated in Table 5. The graph shows colored bars corresponding to a particular compound, wherever the pGI50 value has been identified as a higher outlier for that compound. By comparison, some similarities can be observed between the profiles of some of the inhibitors in the predictive set and the first 10 potential PKI predicted.

The external validation of the method using the NCI COMPARE algorithm showed correlation values between 0.41 and 0.76 for 4 of the first 10 compounds with the highest predictive score values, displaying antiproliferative fingerprints similar to those of known imatinib protein kinase inhibitors, lapatinib, gefitinib, dasatinib and erlotinib, which confirmed the usefulness of the method developed [20].

The difference between the chemical structures of the identified compounds from the typical structures of known protein-kinase inhibitors indicates the usefulness of this method in orienting researchers in the field of medicinal chemistry to new chemical spaces for the identification of *lead-type* compounds for the *design* of PKI. As can be seen in Fig. 4. 3, most of the compounds discussed have in structure aromatic cycles or heterocycles different from the classical structures of known protein-kinase inhibitors.

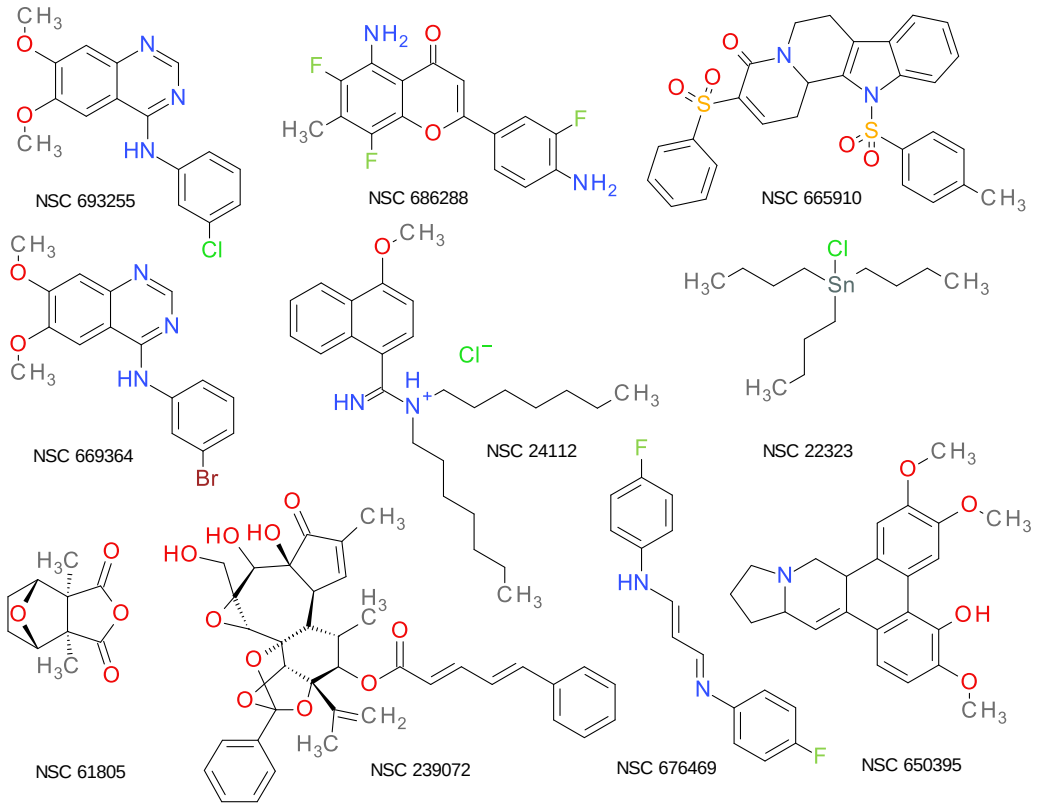

Fig. 4. 3 – Chemical structures of the first 10 potential protein-kinase inhibitors identified on the basis of the developed method

With the help of predictive score, 409 NCI-based compounds were identified that were predicted to act on cell proliferation by inhibiting protein kinases. Out of the first 10 potential PKI identified, the antiproliferative profiles correlated with those of some PKI already approved in the therapy of different types of cancer, such as hematological malignancies and lung, pancreatic and breast cancer, indicating the usefulness of the method for detecting compounds already studied for other effects or the therapeutic reorientation of some drugs.

5. Structural investigations on compounds with antitumor action

In order to increase selectivity, the next stage of the research involved an exhaustive analysis of the antiproliferative effect of the compounds from the NCI base, both from the point of view of the physico-chemical properties and from the point of view of the chemical structure, assessed on the basis of molecular skeletons that describe the topology of the analyzed compounds. Two different sets of compounds were collected from the NCI base, characterizing a total of 91438 compounds tested at one or five concentrations on cell lines (set G_{1D} and PGI set respectively). The process of generating the sets is schematized in Fig. 5.1.

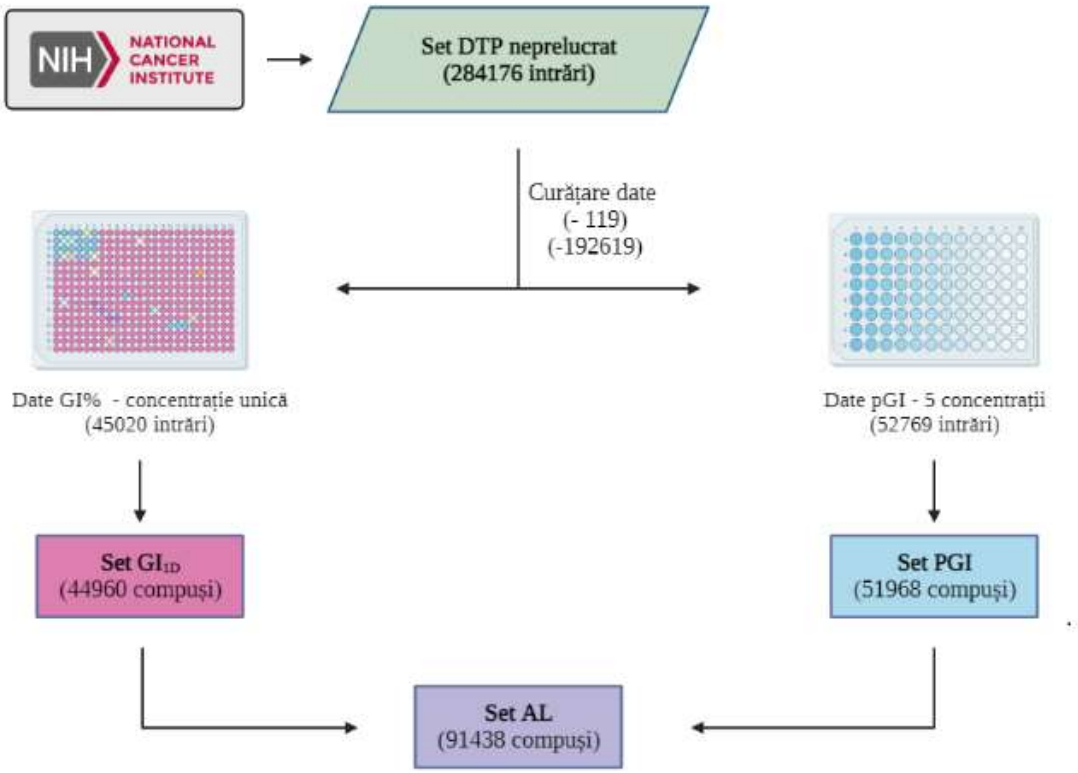

Fig. 5. 1 - Schematization of the process of generating the analyzed compound sets

Two types of characteristic molecular skeletons (Bemis-Murcko and *plain rings*) were generated based on the chemical structures of all compounds in the analyzed sets, using the methods illustrated in Fig. 5. 2and Fig. 5. 3, and then were calculated frequencies, selectivity scores (O_{1D}, O_{pGI}), performance scores (P_{1D}, P_{pGI}) and averages of biological results (A_{1D}, A_{pGI}) for each molecular skeleton.

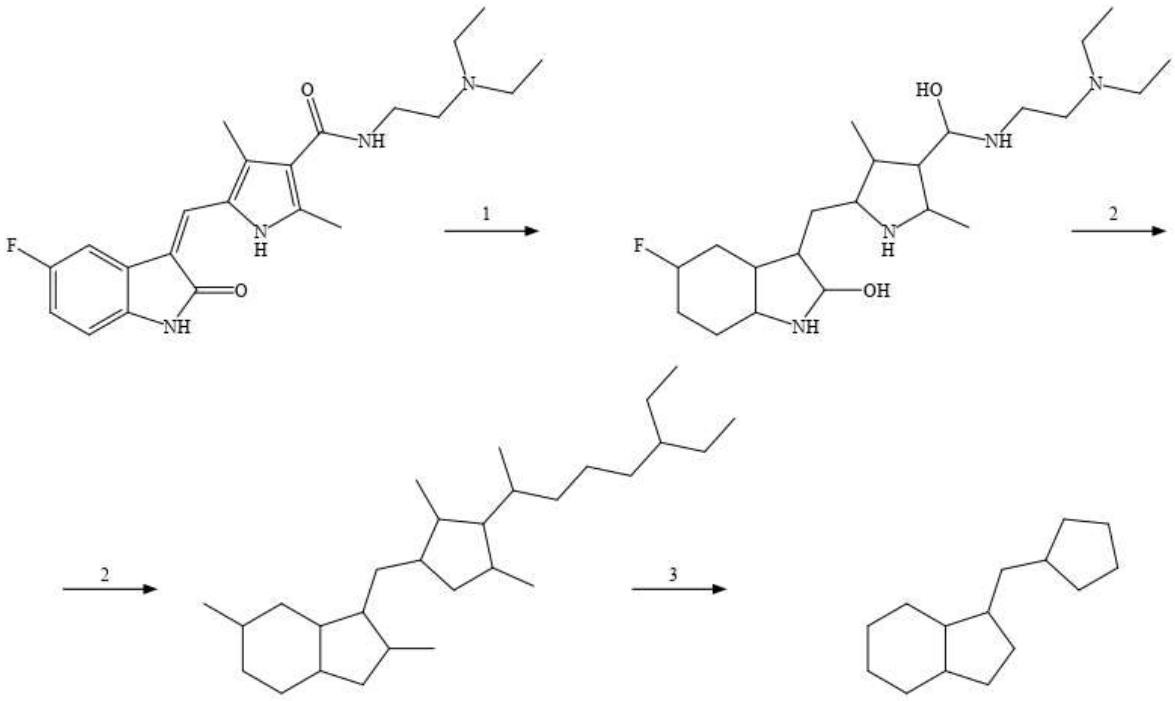

Fig. 5. 2 - Example of the transformation of a compound into a molecular skeleton of the Bemis-Murcko type, with the steps: 1 - the transformation of all double, triple or aromatic bonds into simple bonds; 2 - the transformation of all heteroatomas (except H) from the structure into carbon atoms; 3- the removal of all the lateral chains that do not unite two cycles.

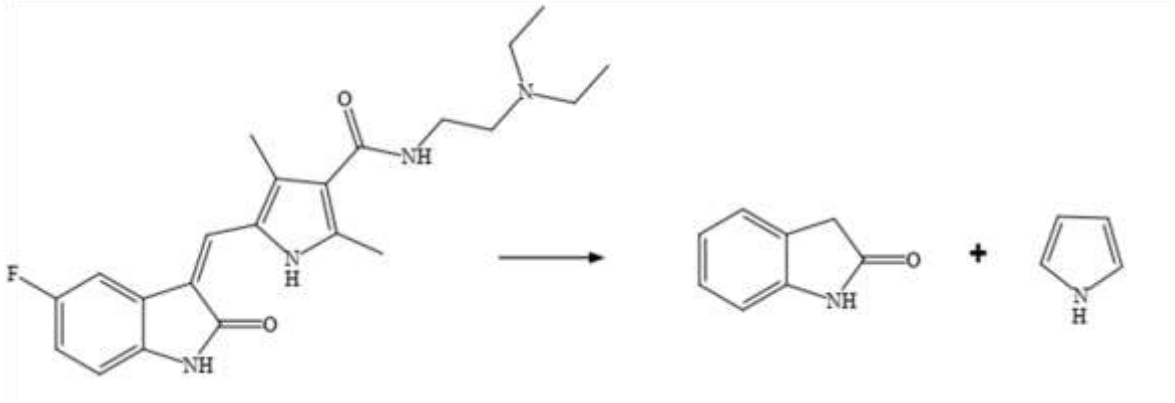

Fig. 5. 3 - Exemplification of the transformation of the compound NSC750690 (sunitinib) into the basic heterocyclic nuclei, indole-2-one and pyrol

The analysis of these scores for the 11763 Molecular Skeletons of the Bemis-Murcko type revealed that the 10 most popular structures are less promising as precursors of compounds with increased antiproliferative effects, but also the most effective, much more structurally complex, illustrated in Fig. 5. 4.

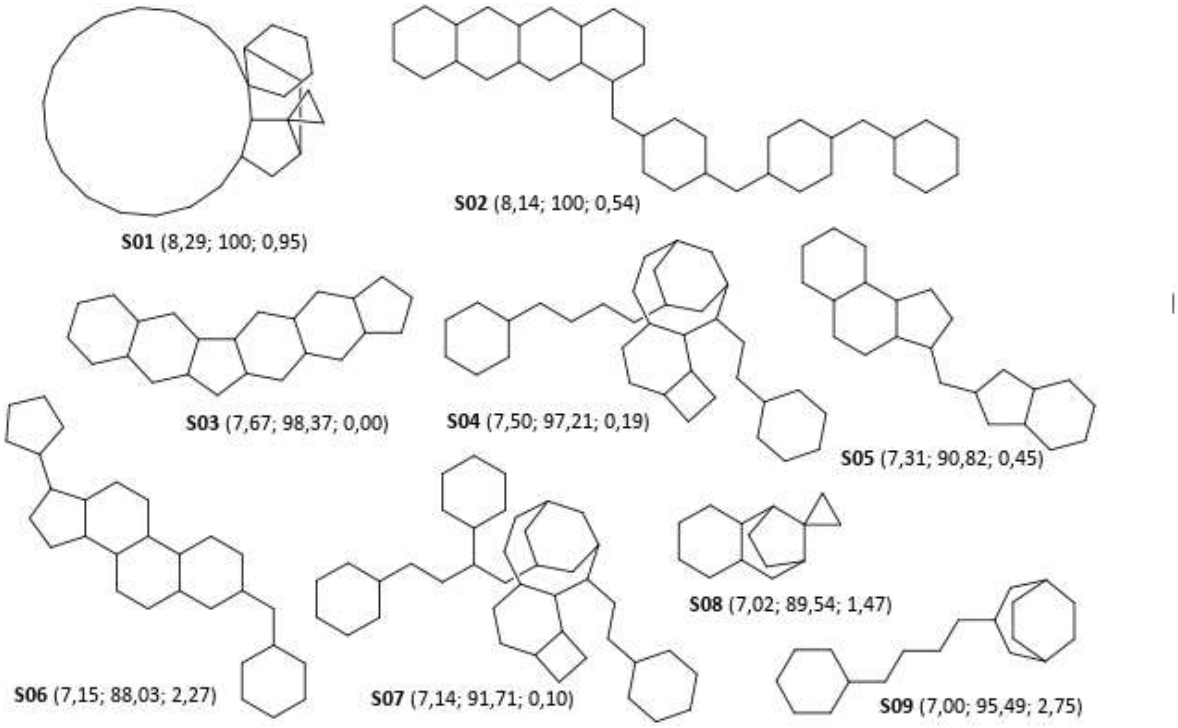

Fig. 5. 4 – Molecular skeletons in the PGI set with the highest performance scores, with A_{pGI}, PGI, and O_{pGI} scores displayed in parenthesis.

Similar analysis for *plain rings* molecular skeletons identified the nuclei of quinoline, tetrahydropyrtan, benzimidazole, and pyrazole with the highest performance scores. Also, various specific structures and patterns in the design of antitumor drugs were identified following this analysis. The results indicated that the nature of the molecular skeleton has a significant impact on the antiproliferative potential, and the acquired conclusions about effective molecular skeletons can be used for the design of antitumor drugs with increased potential for development. The most promising plain rings identified are shown in Fig. 5. 5 and Fig. 5. 6 [21].

The method thus developed in this study can be used for similar analyses based on data extracted from the NSC database of the National Cancer Institute. This method can be directed at a particular cyclical structure and used to identify other structures with similar inhibition profiles. Although the developed method ignores the effect of placeholders on cyclic structures, the results clearly indicate that the nature of the molecular skeleton has a significant impact on the antiproliferative potential. The method can also be used in further research to quantify the influence of the molecular skeleton in comparison with the position or nature of the placeholders in terms of antitumor effect.

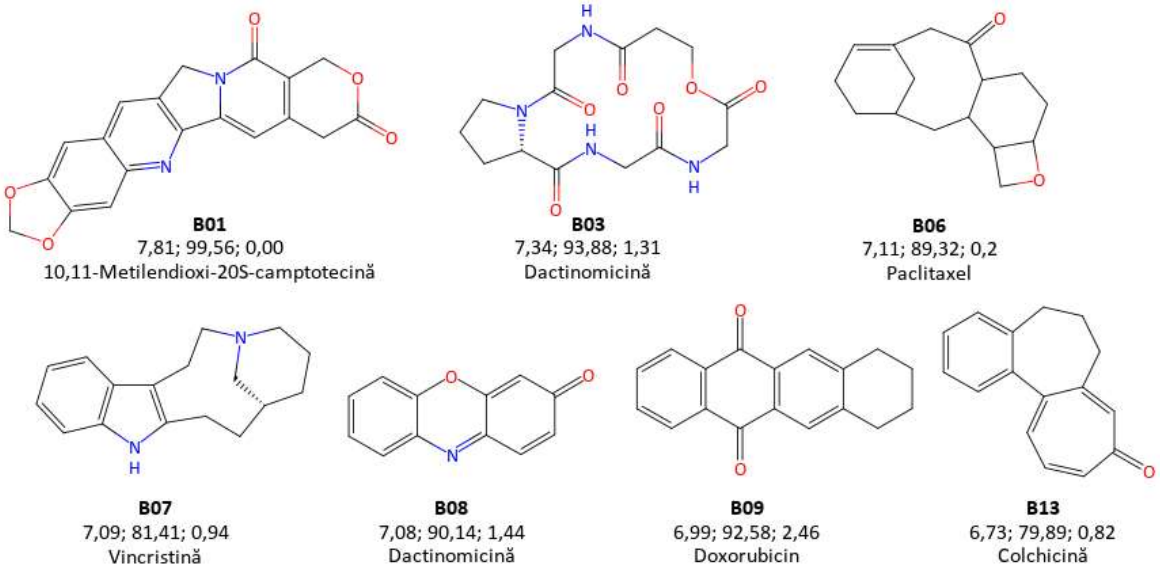

Fig. 5. 5 – The cores with the best performance scores, along with the A_{pGI}, PGI and O_{pGI} scores and drugs based on the respective structures.

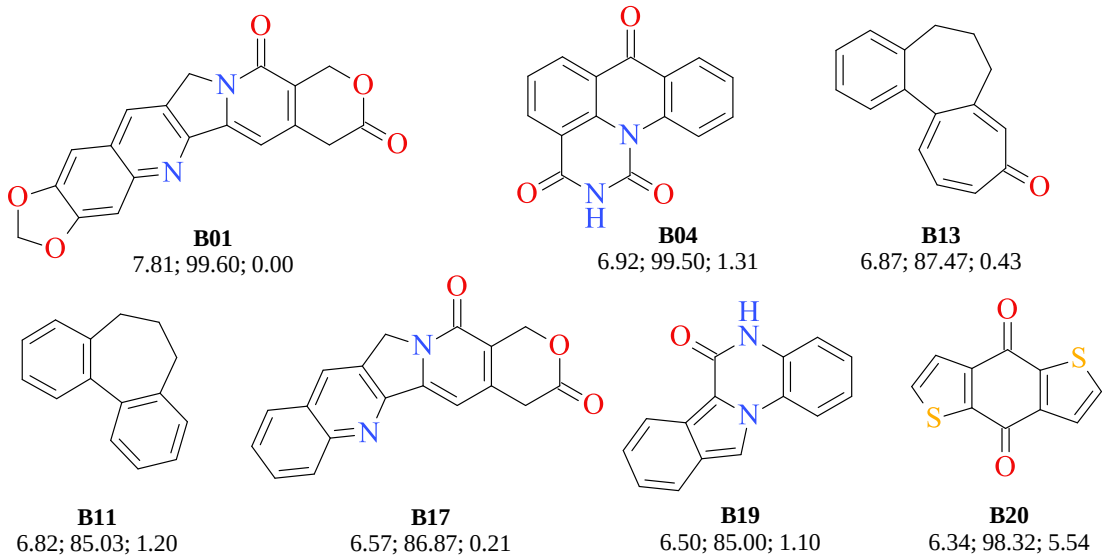

Fig. 5. 6 – The most promising 7 basic nuclei, according to the values of the A_{pGI}, PGI and O_{pGI} scores presented for each, calculated as independent effects on the molecule to which they belong.

6. Optimization of the virtual docking method of Pim-1 kinase inhibitors

Chapter 6 presents the molecular docking study, conducted to differentiate the structures specific to Pim-1 inhibition from the total of different protein-kinase inhibitors. Thus, data were collected on 3067 substances tested for inhibition of human Pim-1 kinase in various *in vitro* studies and available in the online database ChEMBL. Of these, 36 compounds were subjected to the molecular docking study to determine the correlation between the predicted activity *in silico and the* experimentally obtained IC_{50} value . The best binding energy was calculated for the compound CHEMBL2037200, shown in Fig. Fig. 6. 1.

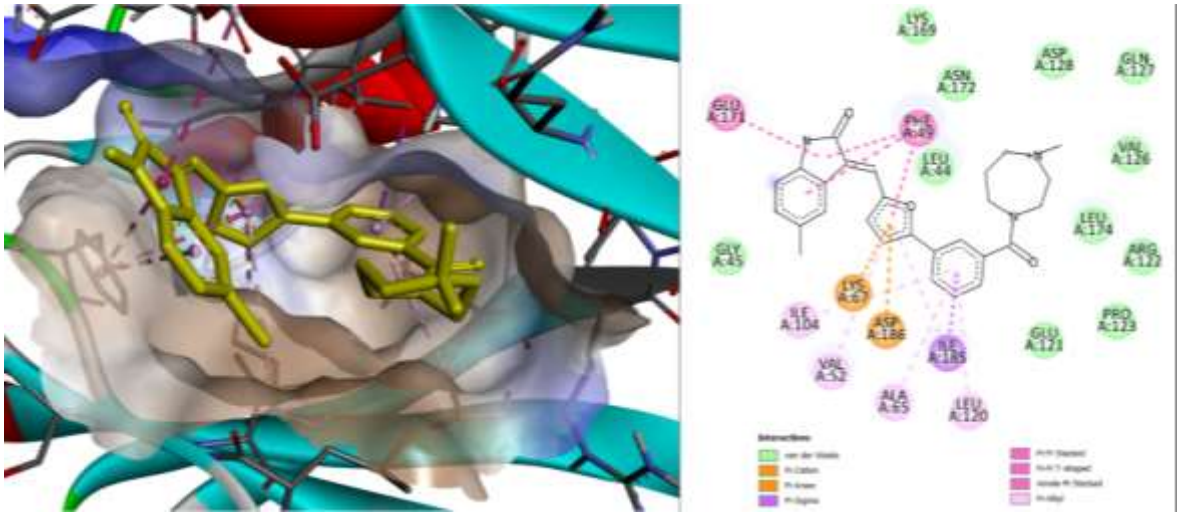

Fig. 6. 1 Docare results for compound 3 (CHEMBL2037200): (a) 3D view of the docated conformation (conformation with the lowest binding energy and without unfavorable interactions); (b) 2D diagram showing interactions with amino acid residues.

By comparing *the in vitro* results with the *in silico* results for already known protein-kinase inhibitors, the molecular docking study found that the docking algorithm used (Autodock Vina) overestimates or underestimates the binding energies, and therefore the calculated affinity to the Pim-1 kinase enzyme is different from the actual affinity, expressed by the $IC_{value\ 50}$. The results of dosing, compared with the results of in vitro testing, are illustrated in Fig. Fig. 6. 2.

To correct this aspect, a method of adjusting the affinity calculated on the basis of interactions with certain key amino acids has been developed. The study analysis process is shown in Fig. Fig. 6. 3.

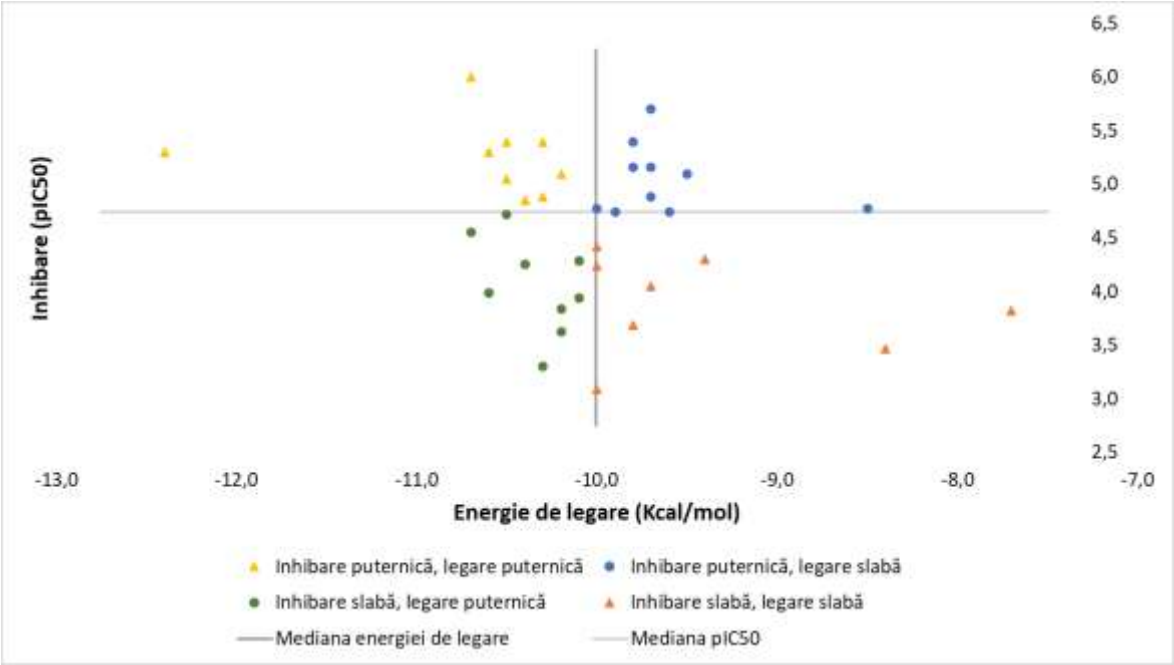

Fig. 6. 2 - Distribution of the analyzed compounds based on pIC_{50} and the binding energy resulting from the docking.

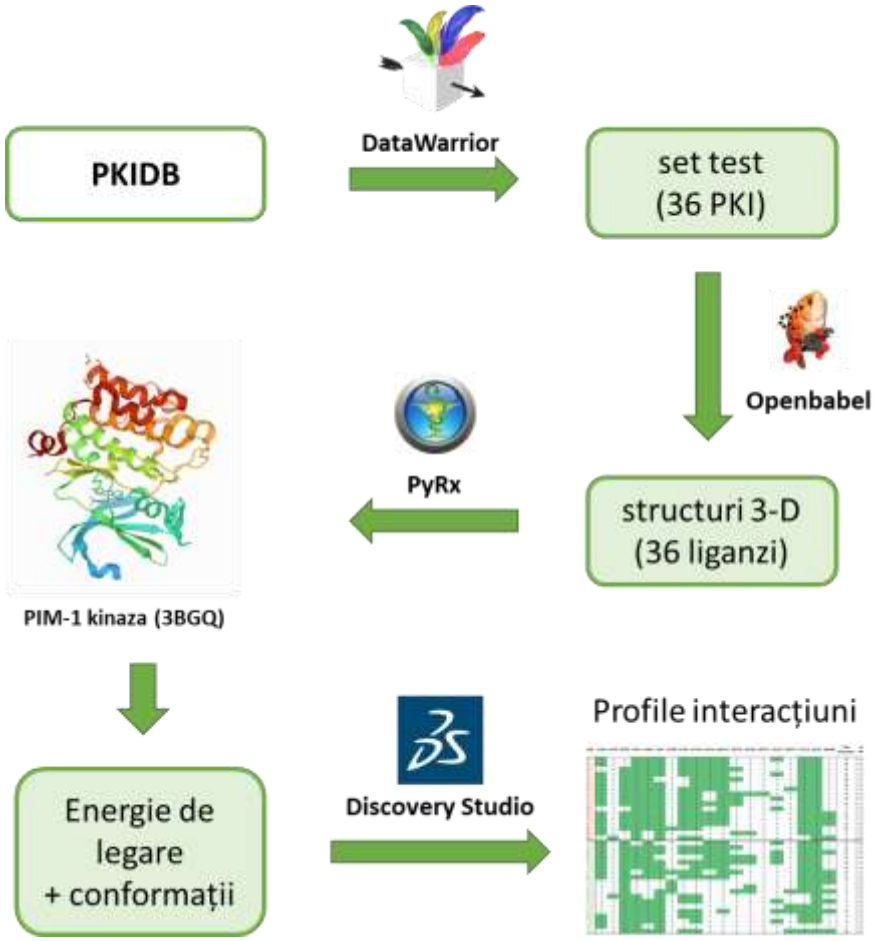

Fig. 6. 3 - General scheme of the stages of the study carried out on the interactions of the test compounds with Pim-1 kinase obtained from the doping process

Following the visualization and interpretation of the molecular dosing results, for each of the 36 ligands, the list of interactions specific to viable conformations was manually drawn. Interactions with each amino acid were encoded as a binary result for each of the amino acid residues participating in the interactions, and between the two sets the distribution of interactions was evaluated. Thus, some interactions have been noted that are supposed to favor more the binding of a ligand in the inhibition site than others. Based on this reasoning, interactions that are correlated with higher inhibitory activity ($lower_{IC\ values}$) were defined as "positive" interactions, being interactions mostly observed for the "strong" group, while "negative" interactions were considered those correlated with lower inhibitory activity. The amino acids Phe49, Val52, Glu89, Arg122, Glu171 and Ile185 in the structure of the Pim-1 enzyme have been identified as key remnants with which a compound should interact to be a ligand with high affinity for the target protein.

A coefficient was calculated for each remaining amino acid interacting with a ligand, then correlation and regression analyses were performed for each participating amino acid, as well as for each compound, also taking into account the subgroup of each ligand. Based on a multifactoral logistic regression analysis, a predictive score was developed that estimates a $pIC_{value\ of\ 50}$ based on the presence or absence of interaction of a compound with each remaining amino acid of the Pim-1 protein. The docking score correction equation was chosen as the model with the highest adjusted value for R^2 (0.525).

$$scor\ predictiv = -0{,}133 * Energie\ liberă + 0{,}903 * Phe49 - 1{,}087 * Val52 - 0{,}619 * Glu89 - 0{,}479 * Arg122 + 0{,}247 * Glu171 - 1{,}055 * Ile185 + 4{,}762\ (R^2 = 0{,}620)$$

The physico-chemical analysis of the cleaned set of 2551 Pim-1 inhibitors revealed specific characteristics of compounds with good activity on Pim-1 kinase, the description of the whole group being useful for identifying molecular descriptors that influence the exercise of a strong inhibitory effect on the enzyme. Compounds with low molecular weight below the threshold of 300 g/mol are distributed almost exclusively in the set of weak inhibitors, and high molecular weight compounds are potent inhibitors of Pim-1 kinase if the LogP value does not exceed the threshold of 5.5. The results indicate the importance of having a sufficiently voluminous molecular skeleton to interact with the enzyme. 87.55% of the compounds of the strong set showed a number of H acceptors between 5 and 8. Compounds with no more than one donor group of H were associated with belonging to the weak group of inhibitors. The results of the structural analysis also indicate the importance of having structures with a certain degree of flexibility and that the $pIC\ value_{50}$ increases on average

by the number of disjointed cyclical fragments. These aspects are illustrated in Fig. 6. 4.

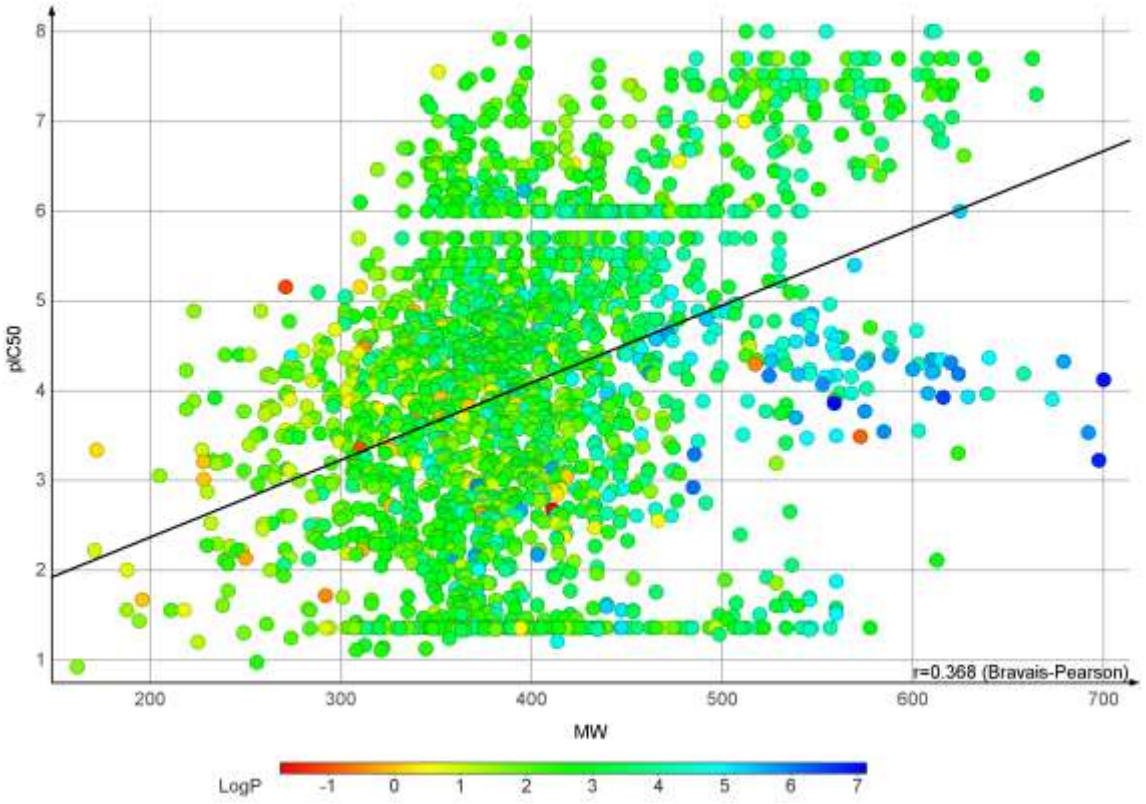

Fig. 6. 4 – Distribution of the pIC50 values of the 2551 compounds in the initial set according to the molecular weight and the n-octanol/water partition coefficient (LogP).

Although it highlights the errors in predicting a ligand's affinity to the target protein, the method developed in Chapter 6 is not validated per se, which can be solved in subsequent studies on a therapeutic target with inhibitors already known and well characterized, with IC50 values for a larger set of compounds. This could also help increase the predictive power of the method, improving accuracy.

7. Identification of potential inhibitors of Pim-1 kinase with antiproliferative effect based on therapeutic reorientation

Chapter 7 brought together the results gained in previous researches, being corroborated to increase the chances of identifying some Pim-1 kinase inhibitors among the structures located in the DrugBank drug base. The drug repurposing study thus conducted led to the selection of 22 compounds that meet the characteristics observed in the previous chapters as being associated with a higher affinity to Pim-1 and an increased antiproliferative effect, with a characteristic profile of protein-kinase inhibitors.

The molecules selected based on methods developed in Chapter 4 and then structurally studied using chapter 5 analysis were finally tested in silico for affinity with Pim-1 kinase by the optimized docking method in Chapter 6, with virtual screening ending with the identification of potential Pim-1 kinase inhibitors. Qn Fig. 7. 1 is illustrated the compound NSC35949, with the bestbinding energy predicted by the docking algorithm (AutoDock Vina), -11.1 Kcal/mol, and an $IC_{50\ predicted}$ score with an increased value of 7.3563.

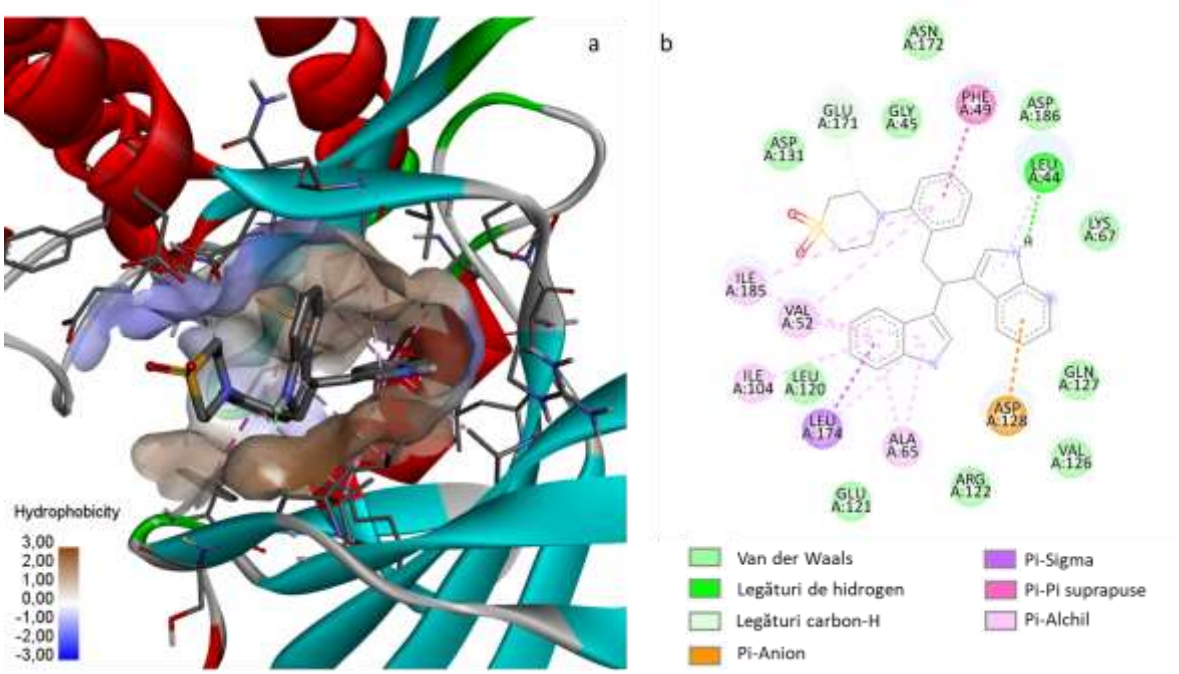

Fig. 7. 1 – a) three-dimensional conformation of the complex between the compound NSC35949 and the active site of the crystallized protein Pim-1 kinase (3BGQ); b) 2D diagram of the interactions between the ligand NSC35949 and Pim-1 kinase

According to the prediction method developed, some of the compounds in the set with high calculated affinities to the protein were overrated by the docking algorithm, with values

below average for the predictive score. Also, some of the compounds calculated to have a weak binding affinity showed above average predictive scores, indicating that they would be stronger inhibitors of the enzyme in reality. The compounds with the highest predictive scores were NSC177380, NSC186946 and NSC185040, which did not coincide with the results obtained from the docking, which identified the compounds NSC35949, NSC1760 and NSC178249 in the top three places. This underscores the predictive power of the developed correction factor, but raises the need to confirm the results by *in vitro* tests, which can be achievable in subsequent studies.

Table 7. 1 – Comparison between compounds with high affinity and those with low affinity, from the estimates of the doping algorithm, compared to the classification based on the predictive score. The orange colored values show the compounds overestimated by the docking algorithm, according to the calculated predictive score, and the values colored blue show the underestimation of the actual affinity

Inhibitors with increased affinity (Energy < -8.58 Kcal/mol)			**Inhibitors with poor affinity (Energy > -8.58 Kcal/mol)**		
NSC	Binding energy (Kcal/mol)	Predictive score	NSC	Binding energy (Kcal/mol)	Predictive score
35949	-11,1	7,3563	**177380**	-8,5	8,0975
1760	-10,5	6,5505	**186946**	-8,5	8,0975
178249	-9,8	5,5544	36398	-8,4	5,2282
174923	-9,3	5,9669	11897	-8,2	6,7236
176371	-9,3	6,2509	**185040**	-8,1	8,0443
186057	-9,3	6,8699	186063	-7,8	6,6704
49733	-9,1	6,8433	**325308**	-7,8	6,9174
174940	-9,1	6,6113	**183519**	-7,7	6,8411
84922	-8,9	6,3377	**5278**	-6,7	7,6111
181099	-8,9	7,9037	167897	-6,6	5,6398
36525	-8,8	7,2344	**23615**	-6,4	6,9522

Both the compounds identified by the docking algorithm and those identified by the predictive score, as potential inhibitors of Pim-1 kinase, are worth investigating further for the evaluation of the inhibitory effect in vitro, in enzymatic tests.

All the aspects studied in the research presented above, such as the physico-chemical properties, the number of outlier values, the Molecular Skeletons of the Bemis-Murcko type distributed according to the values of the binding energy in the docking and the calculated predictive score are corroborated in Fig. Fig. 7. 2.

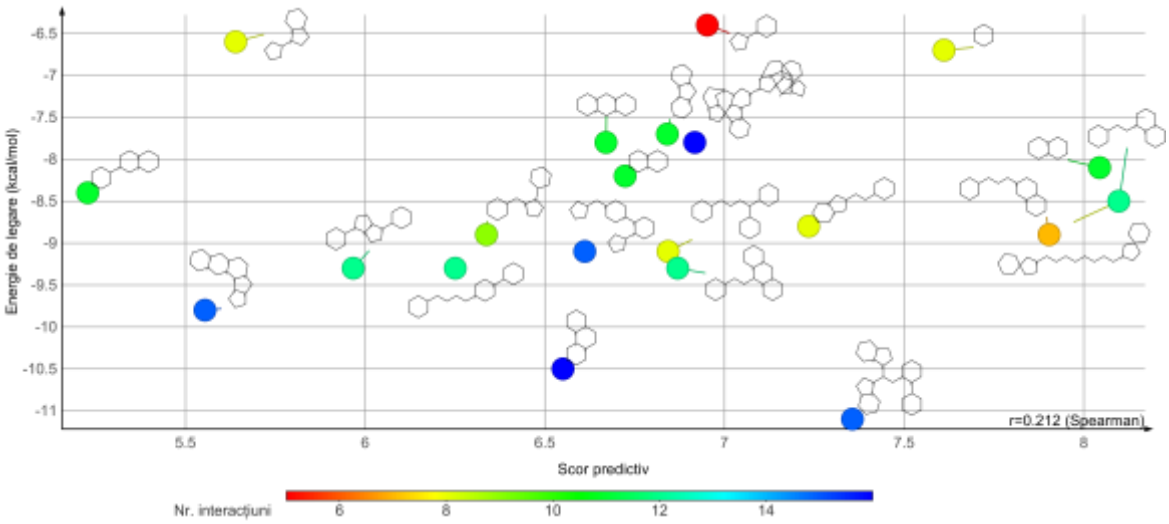

Fig. 7. 2 – Analysis of the correlation between the predictive score and the binding energy, according to the Bemis-Murcko molecular skeletons generated for the 22 studied compounds

Virtual screening identified potential inhibitors of Pim-1 kinase, some of which had antiproliferative effects described in the literature. Further studies can be conducted for the thorough exploration of the inhibitory potential on Pim-1 kinase of haloperidol compounds (NSC176371) and phenothiazine derivatives (NSC186067 and NSC 186063), drugs known as neuroleptic agents.

The method developed and validated throughout the 4 independent researches was confirmed by the identification of compounds, which proved *in vitro* confirmed activity on Pim-1 kinase in recent studies reported in the literature. The novelty and purpose of this research are not only represented by the identification in itself of potential therapeutic candidates, but also by the method developed to achieve this, a method that can be adapted for various therapeutic targets involved in the proliferation of tumor cells.

General conclusions

The general objective of the doctoral thesis was to identify new inhibitors of Pim-1 kinase, useful for the antiproliferative effect and to combat resistance to chemotherapy in various types of cancers, according to the literature.

An important stage for carrying out the researches was represented by the study of the current specialized literature, which describes the current state of knowledge in the field, laying the foundations of the research hypothesis and establishing the premises for the development of Pim-1 inhibitors, detailed aspects in the theoretical part.

In order to achieve the general objective, the experimental studies were based on the use of already existing data in well-characterized databases, made available online by the National Cancer Institute (NCI), about the testing of anticancer activity for more than 90,000 test compounds on the antiproliferative effect on different types of cancers, represented by 60 tumor cell lines. On the basis of these data, methods have been developed to identify patterns that characterize compounds with selectivity on protein-kinases, thus differentiating antitumor substances with another mechanism from protein-kinase inhibitors.

Following the *drug repurposing* study, phenothiazine and butyrophenonic derivatives were also identified, confirmed by recent results from the literature, such as haloperidol (NSC176371) phenothiazine derivatives (NSC186067 and NSC 186063), drugs known as neuroleptic agents.

Personal contributions from experimental research are represented by the methods of identification and prediction of the antiproliferative potential, the inhibition potential of a protein-kinase and the inhibition potential of Pim-1 kinase, respectively, as well as the identification of structural parameters characteristic of Pim-1 inhibitors, which can serve to further identify new antitumor candidates.

In conclusion, the objectives of the research were fully fulfilled, being identified characteristics of the antiproliferative profile NCI-60 specific to protein-kinase inhibitors, with the outlining of a structural profile characteristic of antitumor drugs and with the final identification of some potential inhibitors of Pim-1 kinase, among which haloperidol, oxyridazine and thioridazine are noted, which are suitable for conducting further studies to evaluate the activity *Vitro.*

Selective bibliography

[1]C. Mattiuzzi and G. Lippi, "Current Cancer Epidemiology.", *J. Epidemiol. Globe. Health*, vol. 9, no. 4, pp. 217-222, Dec. 2019, doi: 10.2991/jegh.k.191008.001.

[2]J. E. Debreczeni *et al.* , "Ruthenium half-sandwich complexes bound to protein kinase Pim-1.", *Angew. Chem. Int. Ed. Engl.* , vol. 45, no. 10, pp. 1580-1585, Feb. 2006, doi: 10.1002/anie.200503468.

[3]G. Manning, D.B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, "The protein kinase complement of the human genome.", *Science*, vol. 298, no. 5600, pp. 1912-1934, Dec. 2002, doi: 10.1126/science.1075762.

[4]K.C. Duong-Ly and J. R. Peterson, "The human kinome and kinase inhibition.", *Curr. Protoc. Pharmacol.* , vol. Chapter 2, p. Unit2.9, Mar. 2013, doi: 10.1002/0471141755.ph0209s60.

[5]S. P. Davies, H. Reddy, M. Caivano, and P. Cohen, "Specificity and mechanism of action of some commonly used protein kinase inhibitors.", *Biochem. J.* , vol. 351, no. Pt 1, pp. 95-105, Oct. 2000, doi: 10.1042/0264-6021:3510095.

[6]L. Zhong *et al.* , "Small molecules in targeted cancer therapy: advances, challenges, and future perspectives", *Signal Transduct. Target. Ther.* , vol. 6, no. 1, p. 201, 2021, doi: 10.1038/s41392-021-00572-w.

[7]K. S. Bhullar *et al.* , "Kinase-targeted cancer therapies: progress, challenges and future directions.", *Mol. Cancer*, vol. 17, no. 1, p. 48, Feb. 2018, doi: 10.1186/s12943-018-0804-2.

[8]Y. N. Zhukova, M. G. Alekseeva, N. V Zakharevich, A. A. Shtil, and V. N. Danilenko, "Pim family of protein kinases: Structure, functions, and roles in hematopoietic malignancies", *Mol. Biol.* , vol. 45, no. 5, p. 695, 2011, doi: 10.1134/S0026893311040170.

[9]J. A. Ubersax and J. E. J. Ferrell, "Mechanisms of specificity in protein phosphorylation.", *Nat. Rev. Mol. Cell Biol.* , vol. 8, no. 7, pp. 530-541, Jul. 2007, doi: 10.1038/nrm2203.

[10]S. Eid, S. Turk, A. Volkamer, F. Rippmann, and S. Fulle, "KinMap: a web-based tool for interactive navigation through human kinome data", *BMC Bioinformatics*, vol. 18, no. 1, p. 16, 2017, doi: 10.1186/s12859-016-1433-7.

[11]H. Arrouchi, W. Lakhlili, and A. Ibrahimi, "A review on PIM kinases in tumors", *Bioinformation*, vol. 15, no. 1, pp. 40-45, Feb. 2019, doi: 10.6026/97320630015040.

[12]J. Chen and G. Tang, "PIM-1 kinase: a potential biomarker of triple-negative breast cancer.", *Onco. Targets. Ther.* , vol. 12, pp. 6267-6273, 2019, doi: 10.2147/OTT. S212752.

[13]Y. Tursynbay *et al.* , "Pim-1 kinase as cancer drug target: An update", *Biomed. reports*, vol. 4, no. 2, pp. 140-146, Feb. 2016, doi: 10.3892/br.2015.561.

[14]M. Isaac, A. Siu, and J. Jongstra, "The oncogenic PIM kinase family regulates drug resistance through multiple mechanisms.", *Drug Resist. Updat. Rev. Comment. Antimicrobial. Anticancer Chemother.* , vol. 14, nr. 4-5, pp. 203-211, 2011, doi: 10.1016/j.drup.2011.04.002.

[15]A. L. Merkel, E. Meggers, and M. Ocker, "PIM1 kinase as a target for cancer therapy.", *Expert Opin. I'm investigating. Drugs*, vol. 21, no. 4, pp. 425-436, Apr. 2012, doi: 10.1517/13543784.2012.668527.

[16] B.-W. Wang *et al.* , "Pim1 Kinase Inhibitors Exert Anti-Cancer Activity Against HER2-Positive Breast Cancer Cells Through Downregulation of HER2", *Frontiers in Pharmacology*, vol. 12. 2021, [Online]. Valid at: https://www.frontiersin.org/article/10.3389/fphar.2021.614673.

[17]W. Zhao, R. Qiu, P. Li, and J. Yang, "PIM1: a promising target in patients with triple-negative breast cancer", *Med. Oncol.* , vol. 34, no. 8, p. 142, 2017, doi: 10.1007/s12032-017-0998-y.

[18]P. Mondello, S. Cuzzocrea, and M. Mian, "Pim kinases in hematological malignancies: where are we now and where are we going?", *J. Hematol. Oncol.* , vol. 7, no. 1, p. 95, 2014, doi: 10.1186/s13045-014-0095-z.

[19]H. Koblish *et al.* , "Preclinical characterization of INCB053914, a novel pan-PIM kinase inhibitor, alone and in combination with anticancer agents, in models of hematologic malignancies.", *PLoS One*, vol. 13, no. 6, p. e0199108, 2018, doi: 10.1371/journal.pone.0199108.

[20]G. N. D. Ion and G.M. Nitulescu, "In Search of Outliers. Mining for Protein Kinase Inhibitors Based on Their Anti-Proliferation NCI-60 Cell Lines Profile.", *Molecules*, vol. 25, no. 8, Apr.

2020, doi: 10.3390/molecules25081766.
[21]G. N. D. Ion *et al.* , "Improving the odds of success in antitumoral drug development using scoring approaches towards heterocyclic scaffolds", *Oncol Rep*, vol. 44, no. 2, pp. 589-598, 2020, doi: 10.3892/or.2020.7636.