Development of *A. solani β*-tubulin models and comparison of docking results for benzo[*d*]azoles derivatives as potential antifungal agents

Konstantin L. Obydennov * ⓘ, Tatiana A. Kalinina ⓘ, Tatiana V. Glukhareva ⓘ, Vasiliy A. Bakulev ⓘ

Institute of Chemical Engineering, Ural Federal University, Ekaterinburg 620009, Russia
* Corresponding author: k.l.obydennov@urfu.ru

This paper belongs to a Regular Issue.

Abstract

An approach to developing a docking protocol using free software such as KNIME, DataWarrior, AutoDock Vina, AutoDockTools, OpenBabel and the SWISS-MODEL service was described. In particular, the process of generating possible isomeric structures using KNIME chemoinformatics libraries is described. A library of benzo[*d*]azoles containing 145 compounds with homologous models of *Alternaria solani β*-tubulin obtained in two ways: using the commercial Prime program and the free SWISS-MODEL service, the comparison of results was shown. Despite the less preorganization of the binding cavity of the homologous model obtained using SWISS-MODEL, the correlation of the results between the two methods were observed. The correlation coefficients were as follows: Pearson was 0.65, Spearman was 0.62. According to the docking results, 99% of the studied 2-aminobenzo[*d*]azoles derivatives showed a docking score of no more than –7, which indicates that these compounds are promising for studying the fungicidal activity, in particular against *A. solani*. Without taking into account pharmacokinetic characteristics, benzo[*d*]imidazole derivatives containing a sulfanilamidine substituent at the 2-amino group and thioacetyl derivatives of 2-aminobenzo[*d*]imidazole are of particular interest in the search for new antitubulin fungicides.

Key findings

● A convenient approach was found to generate isomers of organic compounds using the KNIME and DataWarrior programs.

● The optimal parameters for flexible redocking of nocodazole were found in the AutoDock Vina program.

● The use of two homologous *β*-tubulin models in virtual screening results in the similar hit selection.

Supplementary materials

1. Introduction

The heterodimer tubulin is a structural unit of microtubules in eukaryotic cells. Disruption of the processes of formation and destruction of these microtubules, their dynamics, leads to a stop in cell division at the G2/M stage. This process is used in the development of antitumor [1], anthelmintic [2], and antifungal [3] compounds. An important group of antitubulin agents are benzimidazole fungicides. They bind to the *β*-subunit of tubulin in a cavity located near the site of interaction between the *α*- and *β*-subunits (the "colchicine binding cavity") [2, 4–6], disrupting the geometry of the heterodimer and thus disrupting its ability to participate in microtubule dynamics. Anticancer [7–9], antifungal [10–12] and anthelmintic [2, 13, 14] agents have been developed based on antitubulin benzimidazoles. However, the ability of organic compounds to bind in the "colchicine binding cavity" does not warrant antitubulin activity. This type of activity is primarily determined by changes in the conformation of the *β*-tubulin subunit due to the binding of small molecules. In various crystal structures of antitubulin agent complexes with β-tubulins, a similar

conformation of all peptide chains is observed. It was suggested that in this conformation β-tubulins lose their ability to participate in microtubule dynamics [15].

Well-known fungicides containing a benzo[*d*]imidazole moiety are carbendazim, benamyl, thiabendazole, fuberidazole. Benzo[*d*]imidazoles containing 2-acyl(thioacyl)amino group were shown to exhibit antifungal properties against phytopathogenic fungi. 2-Aminobenzo[*d*]oxazole and 2-aminobenzo[*d*]thiazole are structural analogs of 2-aminobenzo[*d*]imidazole (Scheme 1); therefore, it is of interest to study their interaction with *β*-tubulins of phytopathogenic fungi.

One of the modern approaches to the search for biologically active compounds is computer modeling of their interaction with a potential biological target. Computer docking makes it possible to estimate the binding energy of small molecules with macromolecules, but this requires the development of a docking protocol. The docking protocol consists of the following 4 required stages: preparation of the three-dimensional structure of the small molecules (ligand preparation), preparation of the three-dimensional structure of the biological target cavity for binding to the ligand (protein preparation), implementation of the docking and hits identification (Scheme 2).

Using the AutoDock Vina [16] program for docking is a widespread practice in implementing Target Based Drug Design. However, the AutoDock Vina program alone is not enough to implement the Computer Aided Drug Design concept, since programs are also needed to carry out the ligand and receptor preparation procedure.

The well-proven software package from Schrodinger [17] contains all the necessary tools for developing a docking protocol. Previously, using this software package, we performed Induced Fit Docking (IFD) between 2-acyl(thioacyl)aminobenzimidazoles and a number of fungal *β*-tubulins [18]. The purpose of this work was to perform a virtual screening of potentially antifungal benzazoles with a homologous model of *β*-tubulin from the microscopic fungus *Alternaria solani* with using freely distributed programs we selected (KNIME [19], DataWarrior [20], AutoDock Vina [16], AutoDockTools [21], OpenBabel [22] and SWISS-MODEL service [23, 24]). At the same time, it was important to compare the screening results obtained during docking with two homologous protein models: the one obtained in the SWISS-MODEL service and the model obtained using a commercial Prime software. SWISS-MODEL is a web-based integrated service dedicated to protein structure homology modelling. SWISS-MODEL allows you to find a template for homology modelling based on the requested amino acid sequence. In this case, the choice of the template is limited to the structures offered by the service. The Prime commercial program does not have this limitation. Both software products have been successfully used to construct three-dimensional homologous models of proteins or parts thereof [25–28].

2. Experimental

Computer simulation was carried out on a workstation with Intel Xeon Core-core 3.5 GHz processors and the Linux Ubuntu release 18.04 operating system. Graphic material was made in Maestro 12.8 [17].

2.1 Preparing of the small-molecule structures for docking

The same procedure was used to prepare the three-dimensional structures of the native ligand and the studied compounds. SMILE strings of possible tautomers and *E/Z* isomers for each compound were generated in the KNIME program using the KNIME Base Chemistry Types and Nodes, RDKit Nodes Feature, KNIME-CDK, Indigo KNIME integration libraries. 3D structures were obtained in DataWarrior [20] using MMFF94s+ force field minimization, with initial torsion angles taken from the open crystallographic database. Using AutoDockTools [21] and obabel [22], hydrogen atoms on heteroatoms and Gasteiger charges were added to each structure and pdbqt files were generated.

2.2 Preparing of the protein for docking

The creation of homologous models of *Alternaria solani β*-tubulin was carried out using the Prime program and the SWISS-MODEL service. To create a homologous model of *A. solani β*-tubulin in Prime, the GenBank amino acid sequence: HQ413317.1 and the crystal structure of the *Gallus Gallus* β-tubulin D chain (pdb id: 5CA1) as a template were selected. To create a homologous model of *A. solani β*-tubulin in SWISS-MODEL service, the amino acid sequence (GenBank number HQ413317) was used as a template search query. Furthermore, the first best homologous models were choosen (on base 295 unit PDB id: 7UNG). Using AutoDockTools, hydrogen atoms bound to a heteroatom were added for each protein model, and Gasteiger charges and flexible groups at a distance of 4 Å from the nocodazole atoms were determined.

2.3 Docking execution

Docking was carried out in the AutoDock Vina program into the nocodazole binding cavity, defined as a cube with dimensions 30x30x30. The exhaustiveness parameter was set to 16. The selection of the best pose for each compound occurred among all isomeric forms.

Scheme 1 General structures of 2-aminobenzo[*d*]imidazoles **1** (further benzimidazoles), 2-aminobenzo[*d*]oxazoles **2** (further benzoxazoles) and 2-aminobenzo[*d*]thiazoles **3** (further benzothiazoles).
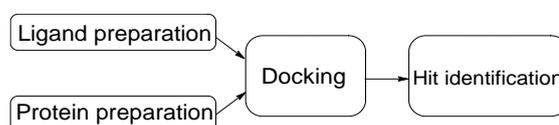
Scheme 2 Main stages of the docking protocol.

3. Result and Discussion

3.1 Development of redocking protocol

One of the ways to validate a docking protocol is its ability to correctly determine the spatial location of a molecule in the binding cavity of a native ligand. This procedure is called redocking. The redocking protocol depends both on the features of the software used and on the structural features of the ligand and receptor.

While the crystal structure of the tubulin complex with fungicidal benzimidazoles is unpublished, the structure of the mammalian tubulin complex with the anticancer drug nocodazole, which has an antitubulin mechanism of action, is known. This structure has pdb id number 5CA1 [15] in RCSB Protein Data Bank (RCSB PDB) [29] and can be used as a template for the creation of fungal β-tubulin. When developing a docking protocol, it is important to consider the possibility of cocrystallized nocodazole **4** to exist in two tautomeric forms (Scheme 3), which in solution are in equilibrium, and also an anomalous length of 1.42 Å of exocyclic C-N bond in the crystal structure. Typically, the length of this bond in the crystal structures of 2-aminobenzimidazoles is in the range of 1.34–1.39 Å [30–32].

To estimate the binding energy of native nocodazole in this complex by AutoDock programs, nocodazole structure was obtained in two ways: without adding hydrogen atoms to heteroatoms (structure **5**) and with adding hydrogen atoms to heteroatoms (structure **6**) (Scheme 4).

In order to ascertain the best structure, we compared the estimates of native ligand binding in these structures without changing the atomic coordinates using the Vina scoring function (equation 1). In this case, the energy of non-covalent interactions for two interacting atoms, i and j, can be decomposed into five energy contributions (equations 2, 3):

$$\Delta G = \frac{V}{1 + 0.058 \cdot N_{rot}}. \tag{1}$$

$$V = w_1 \cdot \Sigma \text{Hbonds}(d_{ij}) + w_2 \cdot \Sigma \text{Hydrophobic}(d_{ij}) + $$
$$+ w_3 \cdot \Sigma \text{Gauss1}(d_{ij}) + w_4 \cdot \Sigma \text{Gauss2}(d_{ij}) + \tag{2}$$
$$+ w_5 \cdot \Sigma \text{Repulsion}(d_{ij}),$$
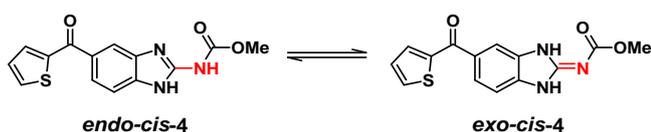
Scheme 3 Tautomeric forms of nocodazole
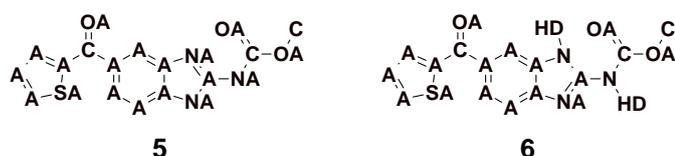
Scheme 4 Structure of nocodazole with AutoDock atom types: without the addition of **5** hydrogen atoms and with the addition of **6** hydrogen atoms.

where V is the conformation-dependent function of intermolecular contribution to scoring function (G); N_{rot} is the number of active rotatable bonds between heavy atoms in the ligand; HBonds is the function of hydrogen bonding term; Hydrophobic is the function of hydrophobic term; Gauss1, Gauss2 are the attractive terms of steric interaction represented as Gaussian functions; Repulsion is the repulsive terms of steric interaction represented as parabolic functions; $w_1 = -0.587$; $w_2 = -0.0351$; $w_3 = -0.0356$; $w_4 = -0.0516$; $w_5 = 0.840$;

$$d_{ij} = r - R_i - R_j \tag{3}$$

where r is the distance between atoms i and j; R_i and R_j are the van der Waals radius i and j atoms.

AutoDockTools was used to prepare the β-tubulin receptor by adding hydrogen atoms to the heteroatoms and partial charges on the atoms and calculated using the Gasteiger method. According to Table 1, the addition of polar protons slightly improved the estimate of binding energy due to hydrogen bond energies.

Furthermore, to make sure that docking structures **5** and **6** using AutoDock Vina with the exhaustiveness parameter equal to 32 will lead to finding the correct pose of the cocrystallized ligand, we performed rigid docking. As a result, form **6** also demonstrated the best performance: docking score –9.6 with an RMSD of 0.140. In this case, form **5** demonstrated a docking score of –8.8 at RMSD = 0.170, which is also acceptable for further modification of the docking protocol. Adding flexibility to the docking protocol in the form of the possibility of rotation around single bonds of the ligand (semi-rigid docking) had virtually no effect on the docking result. The ligand-receptor interaction in the ligand-receptor complex is mainly influenced by amino acid residues located in close proximity to the ligand. Usually, these closely located amino acid residues in the crystal structure are preorganized for better binding to the native ligand. Taking into account the mobility of these residues expands the capabilities of the docking protocol, as it allows one to change the preorganization of the binding cavity for each ligand specially. On the other hand, the appearance of additional degrees of freedom for atomic movement of amino acid residues in the cavity may lead to a less accurate pose prediction of the native ligand. In order to determine which amino acid residues can be set flexible, we also performed flexible docking. In the case of flexible amino acid residues at a distance of 2.5 Å from nocodazole, there is a slight effect on the docking result compared to rigid docking (see Table 2). The flexibility of amino acid residues located at a distance of 3 Å significantly improves the docking score and reduces the accuracy of predicting the pose of the native ligand. We assume that this is due to the "optimization geometry" of amino acid residues important for binding to nocodazole, such as GLU198, TYR200 and VAL236. A further increase in the distance from nocodazole, at which flexible amino acidic residues will be set, has virtually no effect on both the docking score and RMSD; however, an increase in the calculation time is observed.

Table 1 Comparison of energy contributions (Vina scores) to the binding of structural forms of nocodazole **5** or **6** to β-tubulin.

Structural forms	The terms of Vina score taking into account the weight coefficient					
	$w_1\Sigma$Hbonds	$w_2\Sigma$Hydrophobic	$w_3\Sigma$Gauss1	$w_4\Sigma$Gauss2	$w_5\Sigma$Repulsion	Vina score
5	–1.46	–1.12	–3.55	–6.79	4.52	–8.40
6	–2.36	–1.12	–3.55	–6.79	4.52	–9.30

Table 2 Results of redocking structures **5** and **6**.

Docking type	Exhaustiveness	5		6		Time, s [b]
		Vina Score	RMSD [a]	Vina Score	RMSD [a]	
Rigid	32	–8.8	0.171	–9.6	0.138	189
Semi-rigid	32	–8.8	0.171	–9.6	0.138	185
Flexible (2.5 Å)[c]	32	–8.8	0.175	-9.7	0.144	90
Flexible (3 Å)[d]	32	–10.4	0.391	–10.8	0.242	172
Flexible (3.5 Å)[e]	32	–10.8	0.398	–10.9	0.318	2201
Flexible (4 Å)[f]	32	–10.8	0.453	–11.0	0.386	8442
Flexible (4.5 Å)[g]	32	–10.1	0.587	–11.1	0.386	18147
Flexible (4 Å)[f]	16	–10.1	0.586	–11.2	0.434	4119

[a] RMSD of non-hydrogen atoms from the corresponding atoms of the native ligand.

[b] A total time spent docking structure **5** and **6**.

[c] The following amino acid residues set as flexible: ASN165 и TYR200.

[d] The following amino acid residues set as flexible: ASN165, GLU198, TYR200 и VAL236.

[e] The following amino acid residues set as flexible: VAL236, LEU240, ASN165, TYR200, GLU198, GLN134, CYS239, ILE316, ALA315, LYS350, THR351, ALA352.

[f] The following amino acid residues set as flexible: ASN165, TYR200, VAL236, THR237, LEU240, GLU198, LEU250, TYR50, GLN134, PHE167, LEU253, ILE368, CYS239, ALA314, MET257, ILE316, LEU246, ALA315, ALA352, LYS350, THR351.

[g] The following amino acid residues set as flexible: TYR50, ASN165, PHE167, GLU198, TYR200, VAL236, THR237, LEU240, LEU250, LEU253, GLN134, ILE4, ILE368, CYS239, ILE316, LEU246, ALA314, MET257, ALA315, ALA352, LYS350, THR351.

Therefore, we decided to leave the amino acid residues located at a distance of 4 Å as flexible. In order to test the possibility of reducing the time spent on calculations, we performed flexible redocking with the exhaustiveness parameter equal to 16. In this case, we obtained an acceptable result with an RMSD of 0.434, a docking score of –11.2 and a time spent half as much as in the case of exhaustiveness equal to 32. Also, since in all cases the best estimate of the binding energy was demonstrated by the form of nocodazole **6**, we decided to use only this form for further investigation.

Furthermore, we validated this docking protocol not only for nocodazole structures with bond lengths from the complex crystal structure, but also for nocodazole, which has the three-dimensional structure generated using molecular mechanics methods. For this, we generated nocodazole structures using the following sequence: in the KNIME program (see Figure S1 and Scheme S1) four forms of nocodazole were obtained as SMILES strings, and then these strings were converted into three-dimensional structures in the DataWarrior program and saved to an sdf file. After that, using the AutoDockTools tool, 4 corresponding pdbqt files were obtained.

Docking these structures with the receptor led to finding a pose with a binding energy of –9.6 and RMSD = 1.164, in which nocodazole is in the ***exo-cis-*4 form** (Figure 1A). When it is superimposed (Figure 1B) with the structure of the native ligand, it is clear that during docking an inversion of the thiophene ring occurs.
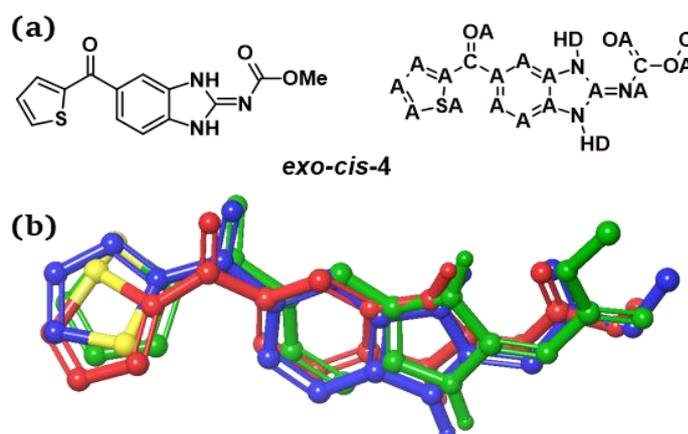
Figure 1 Structure of nocodazole showing the best estimate of binding energy with AutoDock atom types (a); overlay of best pose in this work (red), pose in our last work [18] (green) and native nocodazole structure (blue) (b). Non-polar hydrogens are omitted.

It should be noted that we obtained a similar result earlier when we carried out Induced Fit redocking of nocodazole into this cavity using Macromodel to optimize the geometry of the ligands [18] and Induced Fit Docking protocol. The best pose with a thiophene ring orientation close to that observed in the native ligand (RMSD = 0.839) showed a slightly worse binding score of –8.7. In order to prove whether the mobility of amino acid residues in the binding cavity will influence the orientation of the thiophene ring in the best pose, we performed semi-rigid

docking in Vina. As a result, the best position was also with inversion of the thiophene cycle. In this case, expectedly worse results were obtained compared to flexible docking: binding energy –8.5 and RMSD = 1.31. Manual removal of the thiophene ring from the best complex obtained during flexible docking leads to a deterioration in the binding score from –9.6 to –7.7 due to a decrease in the contribution of attractive steric and hydrophobic interactions (Table S1). Apparently, these structural differences from the crystal structure of the ligand are caused by a geometric difference in the structure of native nocodazole and its structures obtained using molecular mechanics methods. For example, an abnormally long exocyclic C-N bond (1.42 Å) is observed in the crystal. Thus, in this work, in the structure of *exo-cis-*4, the length of the C=N bond is 1.29 Å, and in our previous work the length of this bond was 1.28 Å. It should be noted that during induced docking not only the terminal atoms of the amino acid residues are flexible, but also the other carbon atoms of the amino acids. This leads to an increase in the variety of possible poses generated when searching for hit ones. Thus, in previous work, the best position found during redocking showed inversion of not only the thiophene ring, but also the methyl carboxyl group (Figure 1B). Despite this, the binding site for GLU198 and VAL236 was realized in both cases. In addition, in the case of AutoDock Vina, the pose

with the formation of a hydrogen bond with the amino acid residue ASN165 was better assessed. Based on the obtained RMSD value, as well as on the basis of the implementation of the expected binding site for GLU198 and VAL236, we adopted this docking protocol.

3.2 Ligands preparation

To search for potential antitubulin benzazoles, we selected 145 compounds containing a benzazole fragment with structural formulas **7** and **8**, presented in Scheme 5. This library contains 10 benzoxazole derivatives, 112 benzimidazole derivatives and 23 benzothiazole ones (Scheme 5).

Then the sequence of procedures that was used to generate the structures of nocodazole was used for benzazoles **5** and **6**. In the first step, possible tautomeric forms were generated from 145 SMILES strings of the compounds using the Tautomer Generator node of the KNIME-CDK library [33]. Duplicates and structures generated from incorrectly identified acidity H-atoms were removed. Possible *E/Z* isomers were generated in the Isomer Enumerator node of the Indigo KNIME library. All structures (1009 structures) were saved as isomeric SMILES. Furthermore, in the DataWarrior program, three-dimensional structures were obtained from each SMILES row by optimizing their geometry using the molecular mechanics method (MMFF94s+ force field).

$X = O, S, NH; Y = O, S;$
$R^1 = Me, CF_3, Ph, 3\text{-HOC}_6H_5, 3\text{-MeOC}_6H_5, 3\text{-ClC}_6H_5, 2\text{-thienyl};$
$R^2, R^3 = H, Cl, F, Me, OMe, CF_3$

$X = NH, O, S$
$R^1 = Me, 4\text{-FC}_6H_4, 4\text{-FC}_6H_4, 4\text{-MeC}_6H_4, 4\text{-MeOC}_6H_4,$ [thienyl-CO_2Me], [benzothiadiazole]
$R^2 = H, F, Me, CF_3, CO_2Et$
$R^3 = H, Me, CF_3, CO_2Et$
$R^4 = Me, CF_3, Ph$
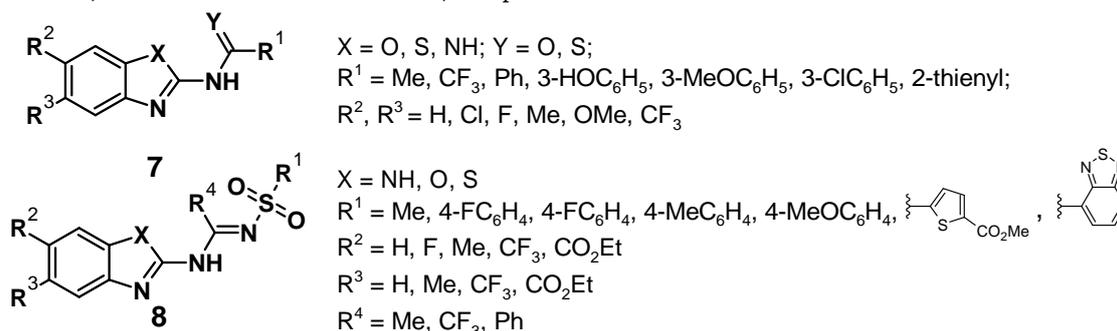
Scheme 5 General structures included in the library.
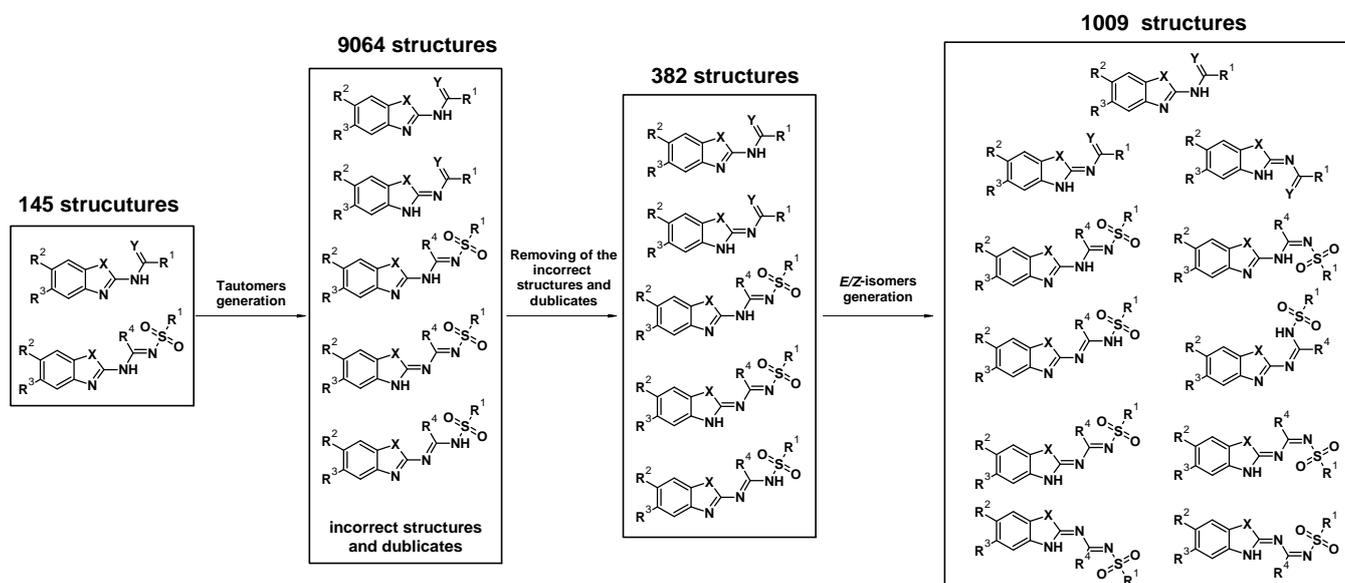
Scheme 6 Scheme for generating possible structures of benzazoles for docking.

In order to avoid mistakes when generating a three-dimensional structure, an option was selected in the Generate Conformers tool that allows specifying in the initial conformer the values of torsion angles that are most often found in the open crystallographic database. Then, for each structure from the sdf file, the corresponding pdbqt files were generated using AutoDockTools.

3.3 Receptor preparation

To create a homologous model of fungal β-tubulin, the amino acid sequence of *A. solani* (GenBank: HQ413317.1), obtained from the National Center for Biotechnology Information (NCBI) databases, was used [34]. Furthermore, we used two approaches to create the three-dimensional structure of β-tubulin: 1) the Prime program [17] and the structure of the D-chain pdb id:5CA1 as a template – the *A. solani* **I** model; 2) the SWISS-MODEL service [23, 24] and the structure 295 of a fragment of the crystal structure of human microtubules (pdb id: 7UNG [35]) – model of *A. solani* **II**. Despite the fact that the primary structures of proteins are identical, the spatial arrangement of atoms of amino acid residues are different. It is important to note that the location of amino acid residues in the potential binding cavity of benzazoles are different (Figure 2). This leads to the fact that different amino acid residues are located at a distance of 4 Å from nocodazole in the considered models: for model *A. solani* **I**: TYR50, GLN134, ALA165, TYR167, GLU198, VAL236, THR237, CYS239, LEU240, LEU246, LEU250, LEU253, MET257, SER314, ALA315, TYR316, GLN350, THR351, ALA352 (19 amino acid residues); for model *A. solani* **II**: GLN134, ALA165, TYR167, GLU198, PHE200, VAL236, THR237, SER248, LEU250, LEU253, MET257, SER314, ALA315, TYR316, GLN350, THR351, ALA352 (17 amino acid residues). Therefore, those amino acids that are contained in at least one of the lists given above were selected as flexible residues: TYR50, GLN134, ALA165, TYR167, GLU198, PHE200, VAL236, THR237, CYS239, LEU240, LEU246, SER248, LEU250, LEU253, MET257, SER314, TYR316, GLN350, THR351, ALA315, ALA352 (21 amino acid residues).

3.4 Docking execution and result analysis

Flexible docking of the benzazole library with tubulin models (*A. solani* model **I** and *A. solani* model **II**) was carried out according to the protocol developed for redocking nocodazole. In order to compare the results of two dockings, we determined the best scoring functions for each compound (among all found poses of each isomeric form). Figures 3 (A–C) show the Vina score scatter plots of the two models, as well as a histogram of the distribution of the differences in docking scores obtained for the two models. Most compounds have high binding scores (less than –7). It was found that for both models *N*-sulfonyl acetamidine binding had a better estimate compared to amides and thioamides (Figure 3A, Table S2, Figure S2). Thus, the median docking score estimates for *N*-sulfonyl acetamidines were –10.5 (model I) and –9.4 (model II), while for amides these

values were –9.4 and –8.3, and for thioamides –9.7 and –9.0, respectively. An analysis of the distribution of each of the contribution of the vina score revealed that the best binding of *N*-sulfonyl acetamidines occurred due to steric interactions and their ability to form hydrogen bonds (Figure S3–S8, Table S2). Thus, it was revealed that in some cases the sulfonyl fragment was involved in the formation of hydrogen bonds with the amino acid residue TYR316. It was also found that in the case of model **II**, benzimidazoles bind better compared to benzoxazoles and benzothiazoles (Figure 3A, Table S2, Figure S2). Moreover, in the case of model **I**, the difference in median binding estimates did not exceed 0.2. It is interesting to note that the median values of the contribution of hydrogen bonds were better for benzimidazoles compared to benzoxazoles and benzothiazoles, both in the case of model **I** and **II**. Amino acid residues that most often formed hydrogen bonds with benzimidazole derivatives were GLN134 and TYR316 in the case of amides and *N*-sulfonyl acetamidines, as well as TYR167 and THR237 in the case of thioamides. Combining benzimidazole moiety with a *N*-sulfonyl acetamidines group in one molecule led to excelent values of docking scoring, reaching values of –12, both due to the contribution of hydrogen bonds and steric interactions, represented by the sum of attractive and repulsive contributions (Table S2).

For some compounds, the best poses were found to be outside the binding cavity of the homologous *A. solani* **II** model and, therefore, have a larger difference in score compared to that obtained when docked with the *A. solani* **I** model (Figure 3B). Thus, these poses were excluded from further analysis. In general, a comparison of the results of the two docking protocols shows that there is a moderate correlation between the obtained values (Pearson's correlation coefficient was 0.65, Spearman's was 0.62).
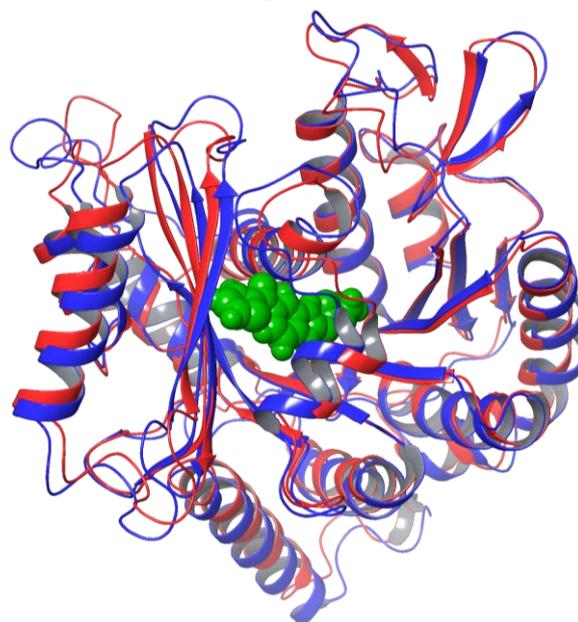
Figure 2 Overlay of structures of homologous models of β-tubulin: atoms of the *A. solani* **I** model are shown in red; atoms of the *A. solani* **II** model are shown in blue. Nocodazole is shown in green, showing a potential binding cavity for benzazoles.

The moderate value of the correlation coefficient is caused by the narrowing of the binding cavity when constructing the protein in the SWISS-MODEL service (Figure 2). This statement is also confirmed by the fact that for all classes of compounds the contribution of repulsive steric interactions is higher in the case of model **II** (Figure S4). Also, in the case of homologous model **II**, the appearance of poses outside the binding cavity is observed. Moreover, the presence of a correlation indicates the preservation of similar possibilities for the formation of non-covalent interactions in the cavities of both models **I** and **II**. Furthermore, for each compound, the differences between the best scoring functions were calculated for two dockings protocols: Vina Score I – Vina Score II. Then, histogram of the distribution of the obtained differences were plotted (Figure 3C). This histogram shows that for 67% of compounds, scoring functions were obtained that differed by no more than 1.0. From all the plots shown in Figure 3, it is clear that the use of model **II** generally underestimates the binding estimate compared to the *A. solani* **I** model. This fact indicates that the cavity in the *A. solani* model **I** is better organized for binding to the benzazoles compared to cavity in the *A. solani* **II** model. Of all the points in the scatter plot (Figure 3A, B), it is interesting to highlight the points corresponding to the following groups of compounds (Figure 3B and Table 3): those showing the lowest binding score (**9**, **10**, **11**, marked in blue), the compounds showing equal docking scores when using different protein models (**12**, **13**, **14**, **15**, marked with green crosses), the compound that previously showed fungicidal activity against *A. solani* [18] (**17**, marked with an orange cross), and the compounds that showed fungicidal activity against other strains of microscopic fungi (**18**, **19**, marked with yellow crosses).

Comparison of the best chosen poses of chosen compounds from different models shows that among these pairs there are both those that showed a head-to-head orientation in the binding cavity (**9**, **10**, **13**, **19**) and those that showed a head-to-tail orientation.

It should also be noted that for **12**, **16**, **18** and **19** the best poses are in different tautomeric forms. Overall, the differences in the best poses are caused by the different conformations of the binding cavities of *A. solani* models **I** and **II**.

4. Limitations

The prediction of the biological properties of the compounds under study cannot be based only on the results of docking. Compounds that show the best docking score values will not necessarily exhibit the greatest fungicidal activity, since docking results reflect only those properties of the compound that determine its pharmacodynamics, but not its pharmacokinetics. Thus, compounds that showed average results in docking may be more active, provided that they create high concentrations in fungal cells.
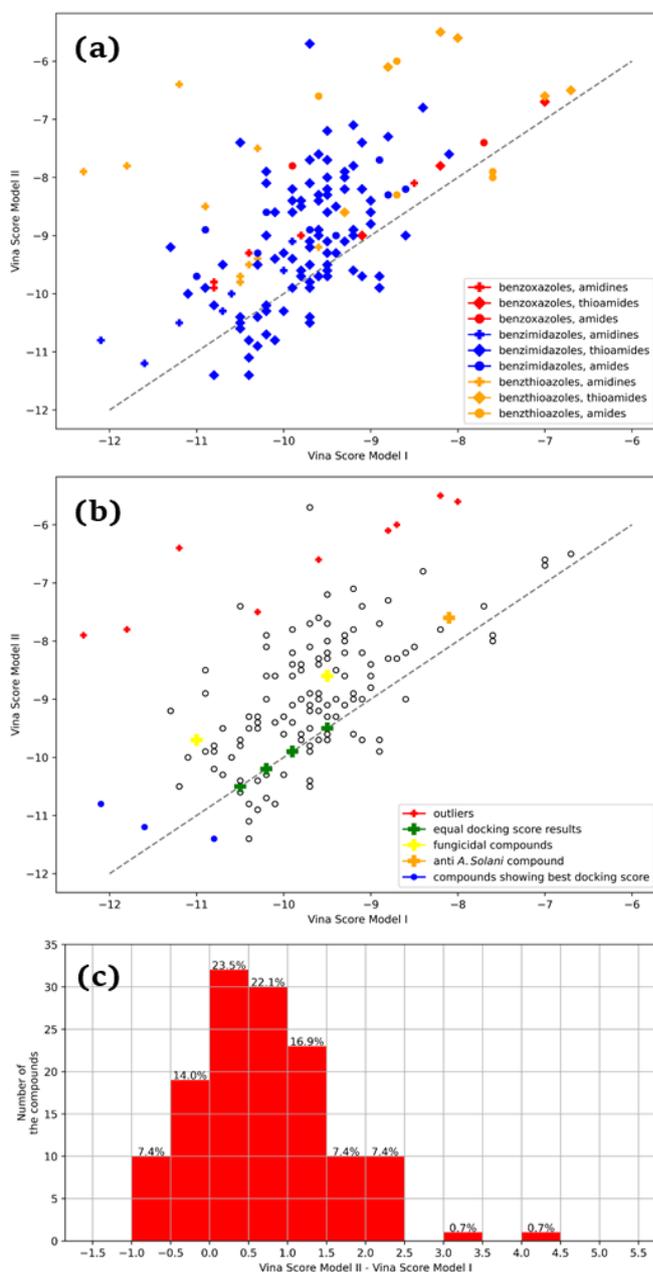
Figure 3 Vina Score scatterplots of two *A. solani* models (a and b). The dotted line is a line at an angle of 45° with respect to the coordinate axes. Distribution of differences between the best values of the scoring functions for compounds (all structural forms) (c).

5. Conclusions

Thus, we have described an approach to developing a docking protocol using only free software. In particular, the process of generating isomeric structures using KNIME libraries is described. Using the example of modeling the binding of benzazoles to homologous models of β-tubulin obtained using the Prime program and the free SWISS-MODEL service, the similarities and differences of the results are shown. Despite the worse preorganization of the binding cavity of the homologous model obtained using SWISS-MODEL, this did not affect the binding assessment of the best compounds compared to the results of docking using the Prime program.

Table 3 Comparison of the best poses for selected compounds after docking with *Alternaria solani* I and II models.

Compounds	Structure of best binding form	Model number	Docking score	Poses overlay
9		1	−12.1	
		2	−10.8	
10		1	−11.6	
		2	−11.2	
11		1	−10.8	
		2	−11.4	
12		1	−9.9	
		2	−9.9	
13		1	−9.5	
		2	−9.5	
14		1	−10.2	
		2	−10.2	
15		1	−10.5	
		2	−10.5	
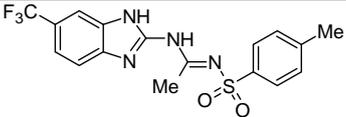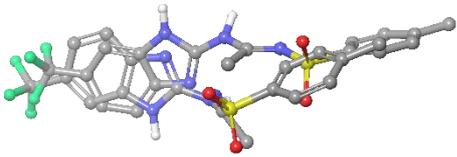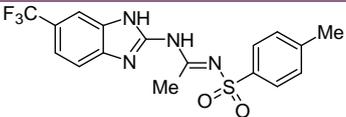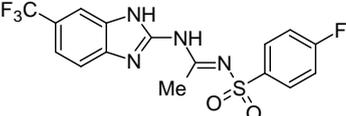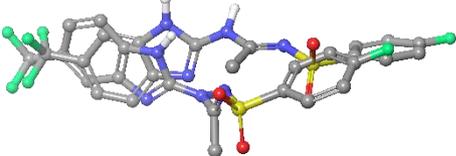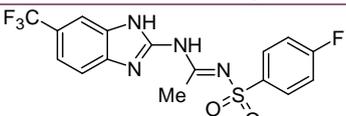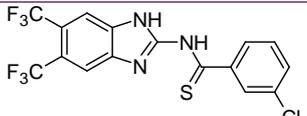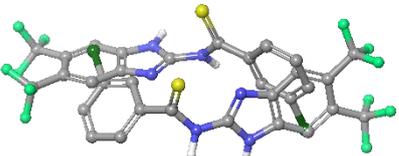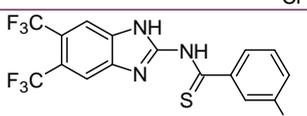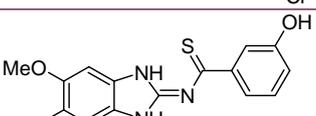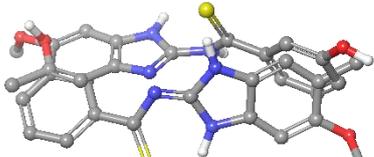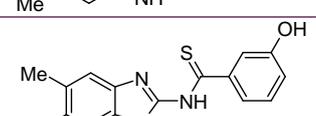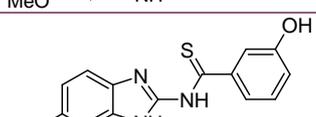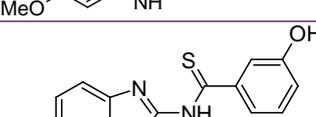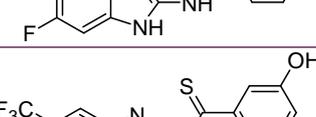
DOI:
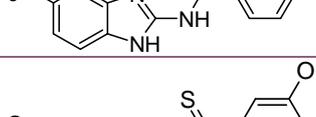
Compounds	Structure of best binding form	Model number	Docking score	Poses overlay
16		1	−10.2	
		2	−10.2	
17		1	−8.1	
		2	−7.6	
18		1	−11.0	
		2	−9.7	
19		1	−9.5	
		2	−8.6	
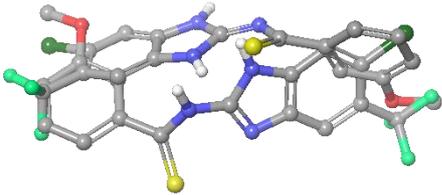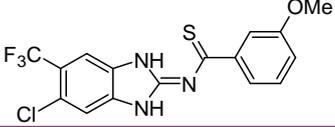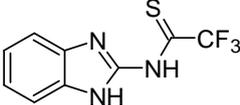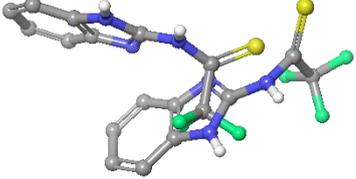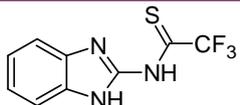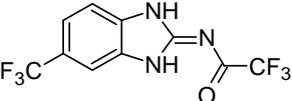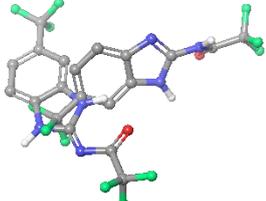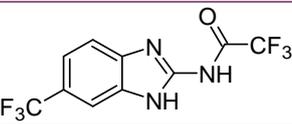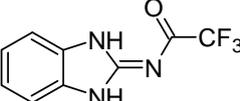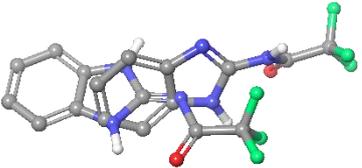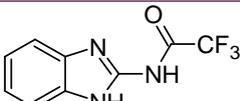
Docking results show that most of the studied 2-aminobenzazole derivatives are promising compounds for the search for fungicides, in particular against *A. solani*. Without taking into account the pharmacokinetic characteristics of the noted compounds (ADME), benzimidazole derivatives containing a sulfanilamidine substituent at the amino group at position 2 and thioacetyl derivatives of 2-aminobenzimidazole are of particular interest in the search for new antitubulin fungicides (Table 3).

● Supplementary materials

This manuscript contains supplementary materials, which are available on the corresponding online page

● Author contributions

Conceptualization: K.L.O., V.A.B.
Data curation: K.L.O., T.A.K.
Formal Analysis: T.A.K.
Funding acquisition: T.A.K.
Investigation: K.L.O., T.V.G.
Software: K.L.O.
Methodology: T.V.G., K.L.O.
Project administration: T.A.K.
Resources: T.A.K., T.V.G.
Supervision: T.V.G.
Validation: T.A.K., T.V.G.
Visualization: K.L.O.
Writing – original draft: K.L.O., T.A.K.
Writing – review & editing: K.L.O., T.V.G.

● Conflict of interest

The authors declare no conflict of interest.

● Additional information

Author IDs:

Konstantin L. Obydennov, Scopus ID 54684683000;
Tatiana A. Kalinina, Scopus ID 54684032600;
Tatiana V. Glukhareva, Scopus ID 6602582999;
Vasiliy A. Bakulev, Scopus ID 35516305000.

Website:

Ural Federal University, https://urfu.ru/en/.

References

1. Negi AS, Gautam Y, Alam S, Chanda D, Luqman S, Sarkar J, et al. Natural antitubulin agents: Importance of 3,4,5-trimethoxyphenyl fragment. Bioorg Med Chem. 2015;23:373–89. doi:10.1016/j.bmc.2014.12.027

2. Ranjan P, Kumar SP, Kari V, Jha PC. Exploration of interaction zones of β-tubulin colchicine binding domain of helminths and binding mechanism of anthelmintics. Comput Biol Chem. 2017;68:78–91. doi:10.1016/j.compbiolchem.2017.02.008

3. Davidse LC. Differential binding of methyl benzimidazol-2-yl carbamate to fungal tubulin as a mechanism of resistance to this antimitotic agent in mutant strains of Aspergillus nidulans. J Cell Biol. 1977;72:174–193. doi:10.1083/jcb.72.1.174

4. Li W, Sun H, Xu S, Zhu Z, Xu J. Tubulin inhibitors targeting the colchicine binding site: A perspective of privileged structures. Future Med Chem. 2017;9:1765–1794. doi:10.4155/fmc-2017-0100

5. Massarotti A, Coluccia A, Silvestri R, Sorba G, Brancale A. The Tubulin Colchicine Domain: a Molecular Modeling Perspective. ChemMedChem. 2012;7:33–42. doi:10.1002/cmdc.201100361

6. Wang Y, Zhang H, Gigant B, Yu Y, Wu Y, Chen X, et al. Structures of a diverse set of colchicine binding site inhibitors in complex with tubulin provide a rationale for drug discovery. FEBS J. 2016;283:102–111. doi:10.1111/febs.13555

7. Tuna BG, Atalay PB, Kuku G, Acar EE, Kara HK, Yilmaz MD, et al. Enhanced antitumor activity of carbendazim on HeLa cervical cancer cells by aptamer mediated controlled release. RSC Adv. 2019;9:36005–36010. doi:10.1111/febs.13555

8. Goyal K, Sharma A, Arya R, Sharma R, Gupta GK, Sharma AK. Double Edge Sword Behavior of Carbendazim: A Potent Fungicide With Anticancer Therapeutic Properties. Anticancer Agents Med Chem. 2018;18:38–45. doi:10.2174/1871520616666161221113623

9. Yenjerla M, Cox C, Wilson L, Jordan MA. Carbendazim Inhibits Cancer Cell Proliferation by Suppressing Microtubule Dynamics. J Pharmacol Exp Ther. 2009;328:390–398. doi:10.1124/jpet.108.143537

10. Vela-Corcía D, Romero D, de Vicente A, Pérez-García A. Analysis of β-tubulin-carbendazim interaction reveals that binding site for MBC fungicides does not include residues involved in fungicide resistance. Sci Rep. 2018;8:7161. doi:10.1038/s41598-018-25336-5

11. Xu S, Wang J, Wang H, Bao Y, Li Y, Govindaraju M, et al. Molecular characterization of carbendazim resistance of Fusarium species complex that causes sugarcane pokkah boeng disease. BMC Genomics. 2019;20:115. doi:10.1186/s12864-019-5479-6

12. Cai M, Lin D, Chen L, Bi Y, Xiao L, Liu X. M233I Mutation in the β-Tubulin of Botrytis cinerea Confers Resistance to Zoxamide. Sci Rep. 2015;5:16881. doi:10.1038/srep16881

13. Aguayo-Ortiz R, Méndez-Lucio O, Medina-Franco JL, Castillo R, Yépez-Mulia L, Hernández-Luis F, et al. Towards the identification of the binding site of benzimidazoles to β-tubulin of Trichinella spiralis: Insights from computational and experimental data. J Mol Graph Model. 2013;41:12–19. doi:10.1016/j.jmgm.2013.01.007

14. Aguayo-Ortiz R, Méndez-Lucio O, Romo-Mancillas A, Castillo R, Yépez-Mulia L, Medina-Franco JL, et al. Molecular basis for

15. benzimidazole resistance from a novel β-tubulin binding site model. J Mol Graph Model. 2013;45:26–37. doi:10.1016/j.jmgm.2013.07.008

15. Wang Y, Zhang H, Gigant B, Yu Y, Wu Y, Chen X, et al. Structures of a diverse set of colchicine binding site inhibitors in complex with tubulin provide a rationale for drug discovery. FEBS J. 2016;283:102–11. doi:10.1111/febs.13555

16. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455–461. doi:10.1002/jcc.21334

17. Schrödinger Release 2019-2: Maestro, Induced Fit Docking Protocol, MacroModel, Prime; Schrödinger, LLC: New York, NY, 2019.

18. Obydennov KL, Kalinina TA, Galieva NA, Beryozkina T V., Zhang Y, Fan Z, et al. Synthesis, Fungicidal Activity, and Molecular Docking of 2-Acylamino and 2-Thioacylamino Derivatives of 1H-benzo[d]imidazoles as Anti-Tubulin Agents. J Agric Food Chem. 2021;69:12048–12062. doi:10.1021/acs.jafc.1c03325

19. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, et al. KNIME: The Konstanz Information Miner. 2008. p. 319–326. doi:10.1007/978-3-540-78246-9_38

20. Sander T, Freyss J, von Korff M, Rufener C. DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis. J Chem Inf Model. 2015;55:460–473. doi:10.1021/ci500588j

21. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009;30:2785–2791. doi:10.1002/jcc.21256

22. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. J Cheminform. 2011;3:33. doi:10.1186/1758-2946-3-33

23. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018;46:296–303. doi:10.1093/nar/gky427

24. Voinkov EK, Drokin RA, Fedotov V V., Butorin II, Savateev K V., Lyapustin DN, et al. Azolo[5,1-c][1,2,4]triazines and Azoloazapurines: Synthesis, Antimicrobial activity and in silico Studies. ChemistrySelect. 2022;7. doi:10.1002/slct.202104253

25. Tavella D, Ouellette DR, Garofalo R, Zhu K, Xu J, Oloo EO, et al. A novel method for in silico assessment of Methionine oxidation risk in monoclonal antibodies: Improvement over the 2-shell model. PLoS One. 2022;17:e0279689. doi:10.1371/journal.pone.0279689

26. Lihan M, Lupyan D, Oehme D. Target-template relationships in protein structure prediction and their effect on the accuracy of thermostability calculations. Protein Sci. 2023;32. doi:10.1002/pro.4557

27. Rodriguez Moncivais OJ, Chavez SA, Estrada Jimenez VH, Sun S, Li L, Kirken RA, et al. Structural analysis of janus tyrosine kinase variants in hematological malignancies: implications for drug development and opportunities for novel therapeutic strategies. Int J Mol Sci. 2023;24:14573. doi:10.3390/ijms241914573

28. Champion C, Gall R, Ries B, Rieder SR, Barros EP, Riniker S. Accelerating Alchemical Free Energy Prediction Using a Multistate Method: Application to Multiple Kinases. J Chem Inf Model. 2023;63:7133–47. doi:10.1021/acs.jcim.3c01469

29. Berman HM. The Protein Data Bank. Nucleic Acids Res. 2000;28:235–242. doi:10.1093/nar/28.1.235

30. Alhalaweh A, Lou B, Bostrom D, Velaga SP. CCDC 668711: Experimental Crystal Structure Determination. CSD Commun. 2007;668711. doi:10.5517/ccqfvb9

31. Lin Y, Ong YC, Keller S, Karges J, Bouchene R, Manoury E, et al. Synthesis, characterization and antiparasitic activity of organometallic derivatives of the anthelmintic drug albendazole. Dalt Trans. 2020;49:6616–6626. doi:10.1039/D0DT01107J

32. Hiscock JR, Gale PA, Lalaoui N, Light ME, Wells NJ. Benzimidazole-based anion receptors exhibiting selectivity for lactate

over pyruvate. Org Biomol Chem. 2012;10:7780. doi:10.1039/c2ob26299a

33. Beisken S, Meinl T, Wiswedel B, de Figueiredo LF, Berthold M, Steinbeck C. KNIME-CDK: Workflow-driven cheminformatics. BMC Bioinformatics. 2013;14:257. doi:10.1039/c2ob26299a

34. National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), Na-tional Center for Biotechnology Infor-mation. 1988. https://www.ncbi.nlm.nih.gov/

35. Gui M, Croft JT, Zabeo D, Acharya V, Kollman JM, Burgoyne T, et al. SPACA9 is a lumenal protein of human ciliary singlet and doublet microtubules. Proc Natl Acad Sci. 2022;119. doi:10.1073/pnas.2207605119