



**HAL**  
open science

## Virtual screening of natural products to enhance melanogenesis

Colin Bournez, José-Manuel Gally, Samia Aci-Sèche, Philippe Bernard, Pascal Bonnet

► **To cite this version:**

Colin Bournez, José-Manuel Gally, Samia Aci-Sèche, Philippe Bernard, Pascal Bonnet. Virtual screening of natural products to enhance melanogenesis. *Molecular Informatics*, 2024, 43, pp.e202300335. 10.1002/minf.202300335 . hal-04662165

**HAL Id: hal-04662165**

**<https://hal.science/hal-04662165v1>**

Submitted on 25 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# Virtual screening of natural products to enhance melanogenesis

Colin Bournez <sup>[a]</sup>, José-Manuel Gally <sup>[a]</sup>, Samia Aci-Sèche <sup>[a]</sup>, Philippe Bernard <sup>[b]</sup> and Pascal Bonnet <sup>\* [a]</sup>

Natural products have long been an important source of inspiration for medicinal chemistry and drug discovery. In the cosmetic field, they remain the major elements of the composition and serve as marketing asset. Recent research showed the implication of salt-inducible kinases on the melanin production in skin via MITF regulation. Finding new potent modulators on such target could open the way to several cosmetic applications to attenuate visible signs of photoaging and improve the tan without sun.

**Keywords:** chemoinformatics; natural product; virtual screening; docking; cosmetic; kinase

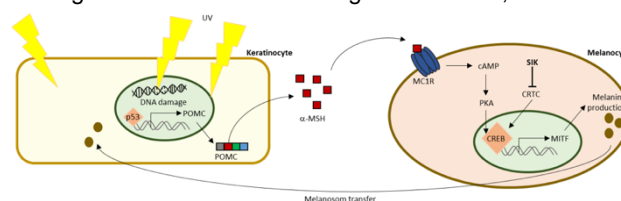
Since virtual screening can be a powerful tool for detecting hit compounds in the early stages of a drug discovery process, we applied this method on salt-inducible kinase 2 to discover potential interesting compounds. Here, we present the different steps from the construction of a database of natural products, to the validation of a docking protocol and the results of the virtual screening. Hits from the screening were tested *in vitro* to confirm their efficiency and results are discussed.

## 1 Introduction

Skin is the largest organ of human body and acts as a waterproof physical barrier against the external environment <sup>[1]</sup>. The products employed for skin care and protection belong to cosmetics; they provide their effect by topical administration directly on the skin. The global cosmetic market reached US\$500 billion in 2017 and is expecting to exceed US\$800 billion by 2023 <sup>[2]</sup>. Among this global market, skin care segment is the most important one, expected at \$181.2 billion by the same year <sup>[3]</sup>.

Skin alteration can result from multiple intrinsic factors including aging, diet and hormonal influence. However, due to its location at the body's surface, the normal evolution of the skin appearance may be influenced, by other external factors, mainly ultraviolet radiation (UVR). The environmental factors will thus play an important role in skin degradation. UVR effects on the skin, referred as photoaging, are important. They might induce a variety of mutagenic and cytotoxic DNA lesions. UVR chronic exposition traits includes oxidative stress, pigmented spots, loss of skin tone, skin wrinkles, increased risks for skin cancer, etc. <sup>[4]</sup>. As a defense, the organism secretes multiples pigments acting as absorbent filters to reduce the penetration through the epidermis of UV and thus reducing their damage <sup>[5]</sup>. In mammals, these pigments are known as melanin, a generic term employed to group the three different types existing: eumelanin, pheomelanin, and neuromelanin. However, only eumelanin and pheomelanin are found in human epidermis <sup>[6]</sup>. The production of melanin (melanogenesis) results from a complex cascade pathway (Figure 1) and rises after the exposure of keratinocytes, the epidermis cells, to UVR. Damages done by UVR on keratinocytes DNA trigger p53-

mediated transcription of the proopiomelanocortin (POMC) gene <sup>[7]</sup>. POMC peptide cleavage produces melanocyte-stimulating hormone ( $\alpha$ -MSH), which is afterward secreted from the keratinocytes. Once  $\alpha$ -MSH bound to the melanocortin 1 receptor (MC1R), located on the membrane of melanocytes, the level of cAMP increases via the activation of the adenylate cyclase. High level of cAMP activates protein kinase A (PKA), which phosphorylates the cAMP-responsive-element-binding protein (CREB). This enhances the transcription of the microphthalmia-associated transcription factor (MITF) gene <sup>[8]</sup>. This MITF induces the expression of tyrosinase catalysing the oxidation of L-tyrosine to L-DOPA and then to dopaquinone, serving as a common precursor to both pheomelanin and eumelanin <sup>[9]</sup>. Finally, once synthesized, the melanin is stored in melanosomes, which are transferred along melanocyte microtubules to basal keratinocytes <sup>[10]</sup>. The visible result of melanogenesis is the darkening of the skin, commonly



called tanning.

**Figure 1.** Melanin production induced by UV radiation and

[a] Institut de Chimie Organique et Analytique (ICOA), UMR CNRS-Université d'Orléans 7311 Université d'Orléans BP 6759, 45067, Orléans Cedex 2, France

[b] Greenpharma SAS 3, allée du Titane 45100 Orléans, France \*e-mail: pascal.bonnet@univ-orleans.fr, phone/fax: +33-238-417-254

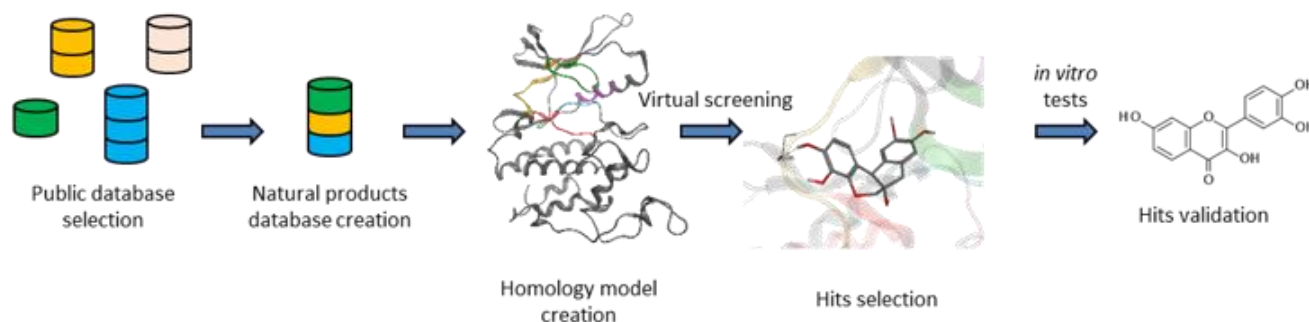
Supporting Information for this article is available on the WWW under [www.molinf.com](http://www.molinf.com)

Skin pigmentation deregulation might have a great impact on the wellness of individuals [11]. Furthermore, with a rising demand in western culture for golden and uniform tan [12], being able to control (positively or negatively) the melanogenesis remains a great opportunity for the dermatological and cosmetic fields. Recent studies report that salt-inducible kinases (SIK) plays an important role in melanin production by regulating the activity of the transcription factor CREB [13–15]. Thus, their inhibition can enhance melanogenesis and induce skin tanning without UVR damages. SIK are serine/threonine kinase belonging to AMP-activated protein kinase (AMPK) family with three distinct isoforms: SIK1, 2, and 3 [16]. The predominant isoform in melanocytes is SIK2 [13].

The trend observed nowadays is a rising demand for natural and organic products since consumers are more concerned about synthetic ingredients and chemical substances [17]. Although natural products (NPs) were traditionally used since ancient time, they were gradually replaced in the last one and half century by synthetic compounds, exhibiting similar properties. [18]. However, thanks to the shift and the increasing request for natural, such substances are now becoming prevalent in modern

cosmetic and cosmeceutical formulations [19]. Several NPs are already known to prevent UV damage with UV absorption property [20]. Some flavonoids are able to downregulate the melanogenesis or on the contrary to stimulate it [21,22]. In any event, these previous results prove that the melanogenesis can therefore be modulated with NPs. The goal of this work is to identify new promising NPs for cosmetic application. A topical application of such substances could be an interesting alternative to UV-tanning.

The application of structure-based virtual screening (VS) before experimental screening for discovering hit compounds on a target has been shown to be profitable [23]. To avoid high false-positive rate and thus testing compounds with low affinity, the methodology must be carefully validated before running on a whole library of compounds. In this study, a VS strategy applied on SIK2 kinase (Figure 2) was implemented. First, a group of several databases of NPs were selected and prepared to serve as a screening library. Second, a rigorous preparation of a 3D structure was made using homology modeling. Third, the docking strategy was validated by testing several software and methods. The docking was then applied to the curated NP dataset. Finally, after a careful selection, an experimental test was performed to confirm the hits and quantify the binding affinities.



**Figure 2.** Workflow of the different steps performed to find interesting molecules.

## 2 Methods

The public NP databases used to constitute our own natural database were all seven downloaded from their website or by request to their administrators (Table 1). Two standardization protocols were tested to prepare the molecules before docking. RDKit (version '2018-09-01') was used to standardize compounds and add hydrogens before generating 3D conformation. The standardization protocol of VSPrep notably includes the filtration of unwanted compounds, the removal of duplicates, the generation of most distributed tautomers at physiological pH (7.4) and the enumeration of stereoisomers depending on the unassigned chiral centers. SMILES formulas were calculated using RDKit (version '2018-09-01'). Molecule conformations were calculated using ETKDG method [24] followed by an optimization step using the MMFF94 forcefield [25].

All experiment and calculations on the molecules have been made with Python 3.6. We calculated the molecular descriptors using RDKit (version '2018-09-01').

The venn diagram was constructed using pyvenn (<https://github.com/tctianchi/pyvenn>).

The homology model was built using modeler version 9.16-1. It was prepared and refined for virtual screening with the "Structure Preparation" module from MOE (Chemical Computing Group, version 2018\_01).

The three docking software versions are GOLD 5.2.2 [26], Glide 7.5 [27] and rDock 2013.1 [28]. rDock is freely available at <http://rdock.sourceforge.net/>, the two others are commercial and need a license. The basic parameters of each program remained untouched. For consistency, the cavity was generated by the same way for all docking programs, by considering an area in a 6 Å spherical radius around the original co-crystallized ligand superposed on our model. Since no water molecule was involved in the binding mode, none was kept in the model. For rDock, we used the module "dock\_sol" instead of the default setting "dock". Once the cavity determined, we performed the redocking of the original ligand in the three software. The number of poses to return was ten in each case. To evaluate

performances two parameters were verified: the ability to predict poses close to the experimental one (RMSD < 2 Å) and their rank according to the native score function. If not ranked first, RMSD value of the first ranked pose were also verified. For Glide and rDock, the score function used, SP and  $S^{inter}$ , are empirical meaning they were determined via regression analysis of experimental dataset of protein-ligand complexes<sup>[27–29]</sup>. While for GOLD, the selected ASP fitness function belongs to knowledge-based ones and is derived from an atom-atom potential also calculated from a database of protein-ligand complexes<sup>[29]</sup>.

For the enrichment studies, the active molecules dataset came from ZINC15 database and decoys from DUD-E<sup>[30]</sup> and a random selection from dataset for JAK-2 kinase. For the virtual screening, all the molecules were fully prepared with VSPrep<sup>[31]</sup>. RDKit and Pandas were used to treat the results and select the best molecules. Poses visualisation and interactions calculation were realized using MOE software (Chemical Computing Group, version 2018\_01). The virtual screening of the entire database took about three days on a standard computer (Intel Core i7-950@3.07 GHz Quad-core, 6 Go RAM).

Experimental tests were conducted by DiscoverX KINOMEScan™ assay at 10 µM.

All the figures were made using matplotlib<sup>[32]</sup>, seaborn<sup>[33]</sup> or MOE. Molecules were drawn with Biovia Draw 2018.

### 3.1 Creation of a natural products database

The primary goal was to obtain a database of NPs containing the information on the precise origin of the compounds and the different organisms from which they can be extracted. Nowadays, many databases of NPs are available and a non-exhaustive list may be found on the ZINC15 database catalog (<http://zinc15.docking.org/catalogs/subsets/biogenic/>)<sup>[34]</sup>. However, they may not be free or may not contain all necessary information desired on the origin of the compounds. Therefore, seven public curated databases were manually selected: BioPhytMol<sup>[35]</sup>, KNApSack-3D<sup>[36]</sup>, NuBBE<sup>[37]</sup>, SANCDB<sup>[38]</sup>, StreptomeDB<sup>[39]</sup>, TCM<sup>[40]</sup> and TM-MC<sup>[41]</sup>. To appreciate the diversity of our dataset, a primary study was performed using several descriptors (see Table S1 and Figure S1 of the Supporting Information). First, the characteristics of these databases are recapitulated in the Table 1. Most of the compounds come from plants, except for StreptomeDB which is exclusively composed of metabolites from bacteria. Three databases contain compounds from around the world (BioPhytMol, StreptomeDB and KNApSack-3D) while the other ones cover a restraint geographical area (Brasil for NuBBE, South Africa for SANCDB and Asian countries for the others). The number of molecules per database varies from 633 to 60,556 but all databases exhibit a good diversity with molecular similarity mean between all compounds around 0.40. TM-MC is the database showing the best compound diversity with a molecular similarity mean of 0.31.

## 3 Results

**Table 1.** Characteristics of each database

| Database            | Localisation    | Source                 | Year | Molecules | Molecular similarity mean (SD) <sup>1</sup> | URL   |
|---------------------|-----------------|------------------------|------|-----------|---|---|
| BioPhytMol          | Worldwide       | Plant                  | 2014 | 633       | 0.39 (0.17)                                 | <a href="http://ab-openlab.csir.res.in/biophytmol/">http://ab-openlab.csir.res.in/biophytmol/</a>                 |
| NuBBE               | Brasil          | Plant                  | 2013 | 881       | 0.41 (0.19)                                 | <a href="http://nubbe.iq.unesp.br/portal/nubbe-search.html">http://nubbe.iq.unesp.br/portal/nubbe-search.html</a> |
| StreptomeDB         | Worldwide       | Bacteria               | 2013 | 4,040     | 0.4 (0.15)                                  | <a href="http://132.230.56.4/streptomedb2/">http://132.230.56.4/streptomedb2/</a>                                 |
| SANCDB              | South Africa    | Plant, marine life     | 2015 | 712       | 0.43 (0.18)                                 | <a href="https://sancdb.rubi.ru.ac.za/">https://sancdb.rubi.ru.ac.za/</a>   |
| KNApSack-3D         | Worldwide       | Plant                  | 2012 | 51,179    | 0.44 (0.16)                                 | <a href="http://knapsack3d.sakura.ne.jp/">http://knapsack3d.sakura.ne.jp/</a>                                     |
| TCM_Database@Taiwan | China           | Plant, mineral, animal | 2011 | 60,556    | 0.45 (0.16)                                 | <a href="http://tcm.cmu.edu.tw/">http://tcm.cmu.edu.tw/</a> <sup>2</sup>  |
| TM-MC               | Northeast Asian | Plant                  | 2015 | 25,518    | 0.31 (0.18)                                 | <a href="http://informatics.kiom.re.kr/compound/">http://informatics.kiom.re.kr/compound/</a>                     |

<sup>1</sup>. Calculated with MACCS keys (166 bits) and the Tanimoto coefficient.

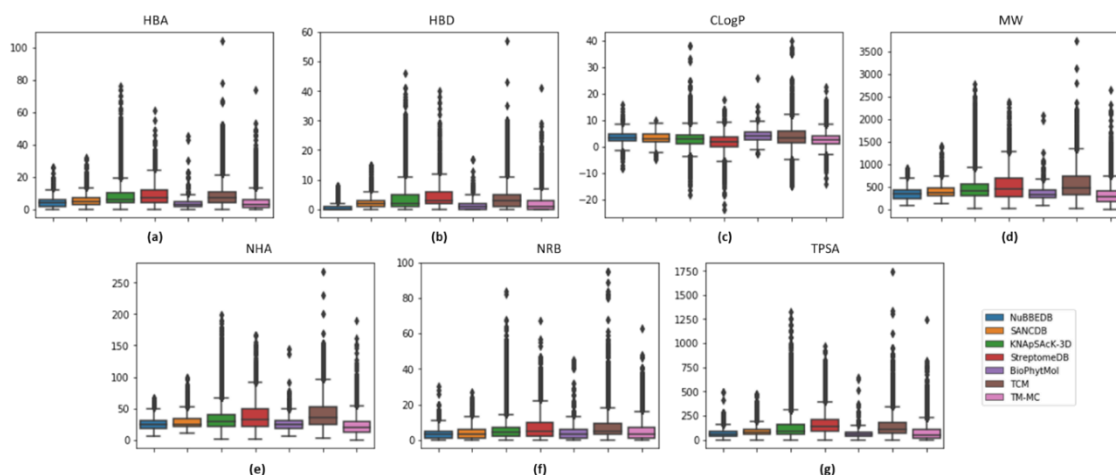
<sup>2</sup>. URL redirect to <https://tm-mc.kr/>, available via ZINC15: <http://zinc15.docking.org/catalogs/tcmnp/>

Second, each database's physico-chemical properties distribution is shown in Figure 3. According to all descriptors used, databases present all similar profiles of distribution. A few databases contain heavy

molecules (> 1,500 Da) but the average ranges from 327 Da for TM-MC to 556 Da for TCM. The extreme value of each descriptor corresponds to either compounds belonging to tannin family such as macabertin or ellagitannin, either complex glycosides as albizoside A

(saponin family), or peptides, as siamycin II, mainly in StreptomeDB. A representation of the heaviest molecule

from all database (MW: 3,738.30 Da) is provided Figure S2 of the Supporting Information.



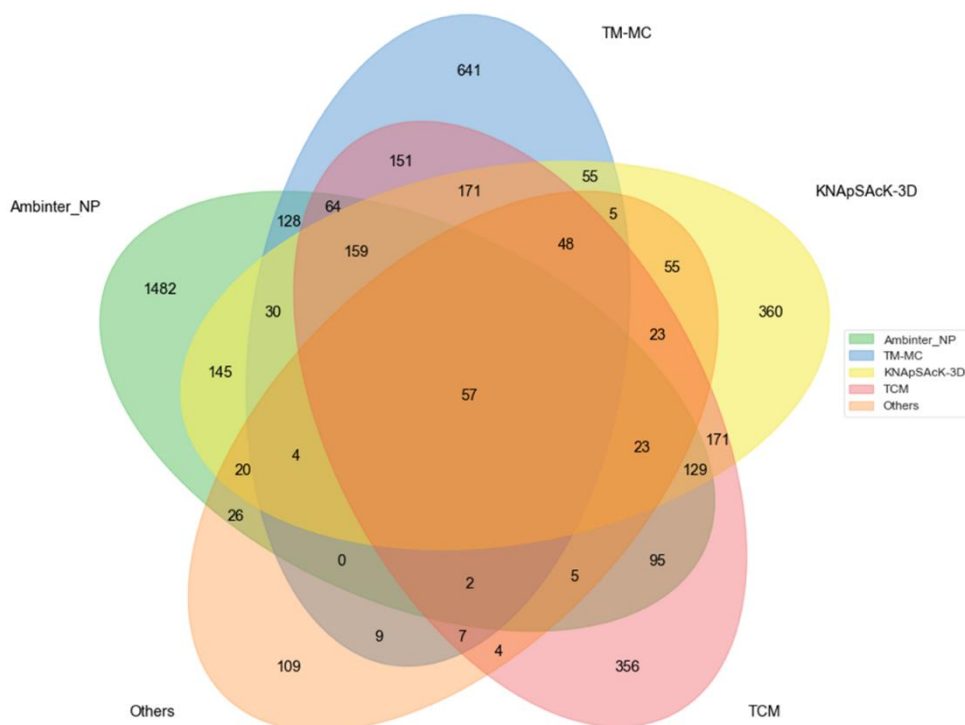
**Figure 3.** Distributions of physicochemical properties of gathered database: (a) Number of hydrogen bond acceptors (HBA); (b) Number of hydrogen bond donors (HBD); (c) Number of rotatable bonds (NRB); (d) Number of heavy atoms (NHA); (e) Molecular weight (MW); (f) ClogP (calculated with RDKit); (g) Topological polar surface area (TPSA).

The duplicates were eliminated based on the SMILES formula of molecules. Each database contained less than 5% of duplicates except TCM (16%). When compared two versus two (matrix depicted Figure S3 of the Supporting Information), all databases contain at least one common identical molecule. Not surprisingly, since they both contain the greatest number of molecules, Knapsack-3D and TCM possess the highest rate of similar molecules (2,394 in common, 2.4%). On another side, SANCDB and NuBBE only share five molecules, making sense since these two databases contain few molecules, which, moreover, come from geographically distant areas.

Before preparing the molecules for the virtual screening, the last step carried out was to check their

availability in Ambinter (<http://www.ambinter.com/>), a global chemical supplier, to ensure the possibility to purchase hits for further experimental testing. Ambinter also allowed us to incorporate its own natural product catalogue to our database (Ambinter\_NP). We ended with a database of 4,534 unique natural compounds. The Venn diagram on the Figure 4 represents the overlapping between the databases of origin of these natural compounds. Among the seven databases selected, four little ones gathered into the label “Others” in Figure 4 provide 109 molecules. The biggest part of original molecules come from Ambinter\_NP, followed by TM-MC and TCM databases. The database providing the fewest molecules is SANCDB, with only two compounds purchasable.

**Figure 4.** Venn diagram representing the compound overlap between the 8 NP databases forming the NP dataset.



## Running title

“Others” groups the databases containing less than 100 compounds (NuBBE, SANCDB, StreptomeDB and Biophytmol). “Ambinter\_NP” refers to Ambinter’s natural products.

A comparison based on the physicochemical properties between our database and protein kinase inhibitors (PKI) from PKIDB<sup>[42]</sup> showed that the chemical space covered by our database is different from the one from PKI (see Figures S4 and S5 in the Supporting Information).

In a VS project, two things are essential: a library of molecules and a 3D structure of the target. As no experimentally determined structure are available for SIK2 human kinase in the PDB<sup>[43]</sup>, we created one by homology modeling. Homology modeling is a technique allowing the prediction of a protein’s 3D structure from its sequence thanks to a template. Two prerequisites are important for a good model construction: a high sequence identity between the target and the template and a correct superposition between them. Since it has been shown that 3D structure of proteins in a family are more conserved than sequence itself, a good template needs to be a protein evolutionarily related to the target<sup>[44]</sup>.

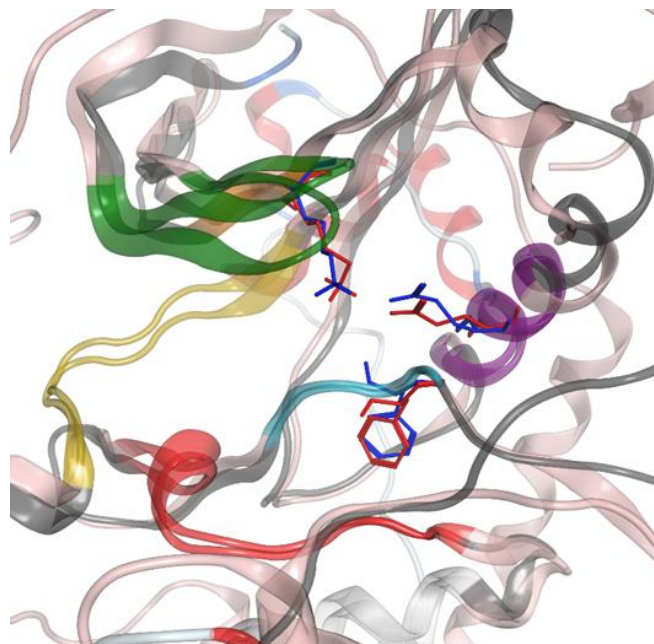
Nowadays, several program are available for homology modeling<sup>[45]</sup>. Our model was built using Modeller<sup>[46]</sup> and the SIK2 sequence retrieved from UniProt database (ID: Q9H0K1)<sup>[47]</sup>. The reference template is a MAP/microtubule affinity-regulating kinase 4 (PDB code: 5ES1, resolution of 2.8 Å)<sup>[48]</sup>, having a sequence similarity of 57% with SIK2. We independently evaluated our model with SwissModel “Structure Assessment” module. We thus obtained a MolProbity score of 2.97 and the Ramachandran plot returned 91.8% of favoured position of amino acids, with three outliers (Figure S6 of the Supporting Information)<sup>[49]</sup>. The global RMSD (root mean square deviation) of our model with the template calculated on backbone is 0.38 Å.

Protein kinases are flexible and may adopt several conformations. The consideration for kinase conformations is critical in the development of targeted kinase inhibitors, especially according to the aimed type<sup>[50]</sup>. Indeed, kinase inhibitors are classified into different types depending on the binding mode adopted and the conformation of the binding site<sup>[51]</sup>. The configurations of the  $\alpha$ C-helix and the conserved DFG motif into the ATP binding pocket were inspected to identify the conformation of our model. Depending on the positions of their amino acids, these motifs can either be in conformation “in” or in conformation “out”. When both are in conformation “in”, i.e.  $\alpha$ C-in/DFG-in, the kinase is in active state. All other possible configurations, i.e.  $\alpha$ C-in/DFG-out,  $\alpha$ C-out/DFG-in and  $\alpha$ C-out/DFG-out represent the inactive states<sup>[52]</sup>. Each conformation exhibits a unique shape allowing or not the binding of ATP and the access to a back pocket. The main characteristics of an active conformation are the opening of the binding site with the knockback of the activation loop, the orientation of the  $\alpha$ C helix toward the active site and a ion-pair interaction between its conserved Glu and the Lys of the  $\beta$ 3 strand. Another important feature is the orientation of the Phe from DFG motif inward toward an allosteric back pocket. In inactive conformation, the interaction between Lys and Glu disappears and Phe flips by  $\sim 180^\circ$  opening the way to the allosteric pocket<sup>[53]</sup>.

100 compounds (NuBBE, SANCDB, StreptomeDB and Biophytmol). “Ambinter\_NP” refers to Ambinter’s natural products.

## 3.2 Preparation of SIK2’s structure

The catalysis of ATP requires the precise positioning of highly conserved motifs. Therefore, the kinase active state is also highly conserved<sup>[54]</sup>. A comparison between our model and a reference structure (PDB code: 1ATP, chain E)<sup>[55]</sup> ensured that it was in active conformation and, as seen in Figure 5, it presents a  $\alpha$ C-in/DFG-in conformation. Kinase inhibitors targeting the active conformation of a kinase are categorized type I or I<sup>1/2</sup><sup>[56]</sup>.



**Figure 5.** Superposition of the model and a protein kinase in active conformation (PDB: 1ATP, chain E). The kinases family conserved motifs are highlighted: DFG motif in blue,  $\alpha$ C-helix in purple, hinge region in yellow, P loop in green and catalytic loop in red. The RMSD between the conserved motifs is 1.1 Å.  $\alpha$ C-helix’s Glu interacts with catalytic lysine and Phe from DFG is oriented toward (DFG-in,  $\alpha$ C-helix in).

## 3.3 Virtual screening

## 3.3.1 Validation of the method

The goal of this part was to find and select the best protocol to dock our natural product dataset. Three different docking software: Glide<sup>[27]</sup>, GOLD<sup>[26]</sup> and rDock<sup>[28]</sup> were compared using two preparation methods of compounds: RDKit and VSprep. To assess the enrichment in active molecules, a dataset composed of 33 active compounds and 200 decoys (ratio: 0.17) was used. The performance was judged based on the ability of the software to discriminate active compounds from inactive ones by using a score, illustrated by the value of area under the receiver operating characteristic curve (AUC). The higher the value of AUC, the better the enrichment of the true positives compared to a random method, where AUC would be equal to 0.5.

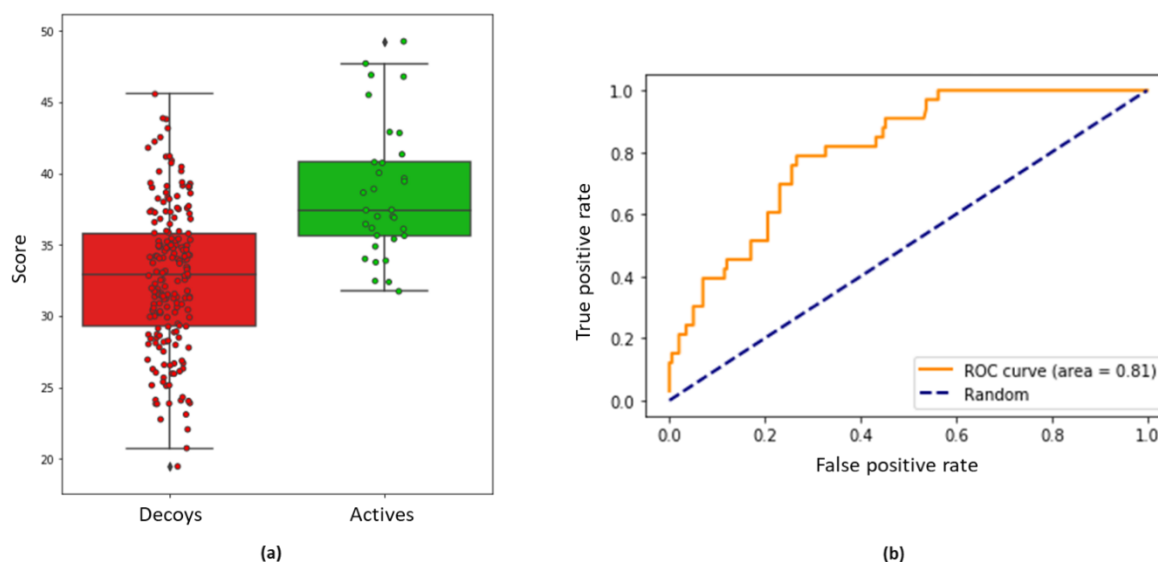
The results are summarized in the Table 2. We observed a higher AUC value for the three docking methods, meaning a better enrichment in true positives, when molecules are prepared with the VSprep protocol.

Furthermore, among the three docking methods and regardless the protocol used to prepare the compounds,

Gold is the only docking software showing significant score difference between actives and decoys.

**Table 2.** Median score and AUC values between the different software and preparation methods

| Software (score function)   | Preparation method of compounds | Median score for actives | Median score for decoys | AUC         |
|-----------------------------|---------------------------------|--------------------------|-------------------------|-------------|
| GLIDE (SP)                  | RDKit                           | -4.27                    | -4.75                   | 0.37        |
|                             | VSPrep                          | -5.22                    | -4.82                   | 0.66        |
| rDock (S <sup>inter</sup> ) | RDKit                           | -21.97                   | -21.51                  | 0.51        |
|                             | VSPrep                          | -23.13                   | -23.46                  | 0.55        |
| Gold (ASP)                  | RDKit                           | 35.97                    | 32.08                   | 0.74        |
|                             | <b>VSPrep</b>                   | <b>37.44</b>             | <b>32.98</b>            | <b>0.81</b> |



**Figure 6.** Gold docking's results between decoys (red) and actives (green) molecules prepared with VSPrep: (a) docking score dispersion; (b) ROC-curve plot.

To conclude, the protocol using VSPrep<sup>[31]</sup> to prepare the molecules for the virtual screening associated to GOLD software returned the best results with an AUC of 0.81 (Figure 7) and an enrichment factor at 1% of 7.03 (see Supporting Information for more details).

### 3.3.2 Docking of our natural product database

As the enrichment studies preconized, the molecules from the natural products database were prepared following the protocol with VSPrep. With the calculation of the different tautomeric forms, the total of molecules in our database rose to 10,939. At the rate of six poses returned per molecule, we ended with 65,586 poses meaning that GOLD was unable to perform calculation docking for eight compounds. The distribution range of score goes from -11.72 to 58.04 with an

average of 22.94. The average difference of score between the six poses of a same molecule is 1.69, showing a consistency in GOLD scoring results.

Three known synthetic inhibitors were randomly added as references: HG-9-91-01, YKL-06-061 and YKL-06-062<sup>[15]</sup>. Their best poses were all ranked in top 100 with a score of 42.49 (47<sup>th</sup>), 42.04 (54<sup>th</sup>) and 41.78 (56<sup>th</sup>) respectively, proving the robustness of the docking method. Surprisingly, while inspecting the top 100<sup>th</sup> compound, a compound focused our attention. A synthetic kinase inhibitor not placed on purpose, the gefitinib, was retrieved on 35<sup>th</sup> position with a score of 43.43. After further investigation, it appeared that it came from the database TM-MC native from A. Radix (<http://informatics.kiom.re.kr/compound/detail.do?id=gefitinib>). Even if this is not a natural product, it was also a proof that our method was able to score kinase inhibitors

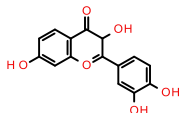
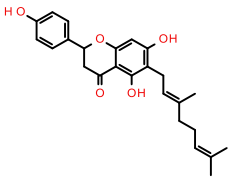
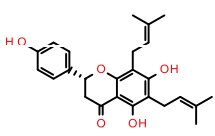
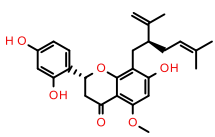
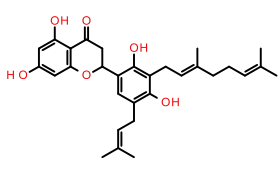
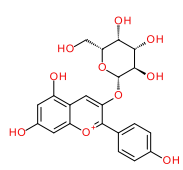
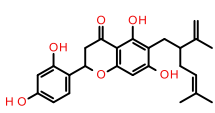
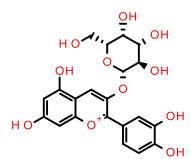
## Running title

at the top. In a way, this event could be compared to a successful double-blind test. However, this also serves as a reminder that even curated, a database can still contain mistakes and it remains important to check the data before and after their processing.

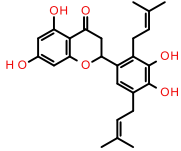
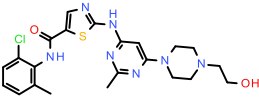
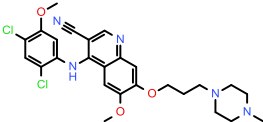
Only the top 1,000 poses were considered for further analysis. For the hit selections, multiple parameters were checked including physico-chemical properties, interactions with the receptor but also marketing issue

as the origin of the compound and the availability of intellectual property. Finally, after a rigorous visual inspection of remaining poses, 25 molecules were selected for experimental testing. Unfortunately, when purchasing the compounds, more than half of them were temporarily unavailable. All molecules sent for experimental test are recapitulated in Table 3. We also added two type I kinase inhibitors as positive test: bosutinib and dasatinib<sup>[42]</sup>.

**Table 3.** Compounds selected for experimental test. Score indicates the GOLD ASP fitness function.

| ID                       | Molecule  | CAS number   | MW     | CLogP | HBA | HBD | Score | Rank |
|--------------------------|---|--------------|--------|-------|-----|-----|-------|------|
| Fisetin                  |    | 528-48-3     | 286.05 | 2.28  | 6   | 4   | 44.25 | 23   |
| 6-geranylneringenin      |    | 97126-57-3   | 408.19 | 5.74  | 5   | 3   | 37.17 | 148  |
| 6,8-diprenylneringenin   |   | 68236-11-3   | 408.19 | 5.52  | 5   | 3   | 37.93 | 122  |
| Kuraninone               |  | 34981-26-5   | 438.20 | 5.61  | 6   | 3   | 37.29 | 140  |
| Sanggenol- P             |  | 1351931-30-0 | 492.25 | 6.96  | 6   | 4   | 41.33 | 62   |
| Fragarin                 |  | 574-57-2     | 468.08 | -2.32 | 9   | 7   | 40.61 | 75   |
| Kushenol F               |  | 34981-24-3   | 424.19 | 5.31  | 6   | 4   | 37.18 | 145  |
| Cyanidin-3-O-galactoside |  | 27661-36-5   | 484.08 | -2.61 | 10  | 8   | 40.14 | 80   |



|             |   |              |        |      |   |   |       |     |
|-------------|---|--------------|--------|------|---|---|-------|-----|
| Sigmoidin A |  | 176046-04-1  | 424.19 | 5.23 | 6 | 4 | 38.76 | 107 |
| Dasatinib   |  | 302962-49-8  | 487.16 | 3.31 | 9 | 3 | -     | -   |
| Bosutinib   |  | 1391063-17-4 | 529.17 | 5.19 | 8 | 1 | -     | -   |

**Table 4.** Biological data of the eleven compounds used in this study. Binding affinities (percent control %Ctrl and K<sub>d</sub>) were measured in the DiscoverX KINOME scan® platform. Compounds were first screened at 10 μM, and if % Ctrl < 50 %, K<sub>d</sub> determinations were performed in duplicate using 11-point dose response curves (top concentration 30μM, 3-fold dilutions).

| ID                       | SIK2<br>(% Ctrl)                                      | BRAF (%Ctrl)  | BRAF V600E (%Ctrl)                                    |
|--------------------------|---|---|---|
| Fisetin                  | 24 (pK <sub>d</sub> : 6.6, K <sub>d</sub> : 270 nM±1) | 9.6 (pK <sub>d</sub> : 6.8, K <sub>d</sub> : 150 nM ±5) | 11 (pK <sub>d</sub> : 7.0, K <sub>d</sub> : 110 nM±5) |
| 6-geranylningenin        | 100   | 97  | 95  |
| 6,8-diprenylningenin     | 100   | 75  | 91  |
| Kuraninone               | 100   | 88  | 86  |
| Sanggenol- P             | 100   | 89  | 87  |
| Fragarin                 | 100   | 82  | 93  |
| Kushenol F               | 100   | 88  | 86  |
| Cyanidin-3-O-galactoside | 100   | 90  | 97  |
| Sigmoidin A              | 100   | 79  | 91  |
| Dasatinib <sup>57</sup>  | 0 (pK <sub>d</sub> : 8.2, K <sub>d</sub> : 6.4 nM)    | 0 (pK <sub>d</sub> 6.3, K <sub>d</sub> 500 nM)          | 0.3 (pK <sub>d</sub> : 6.2, K <sub>d</sub> 570 nM)    |
| Bosutinib <sup>57</sup>  | 0 (pK <sub>d</sub> : 7.5, K <sub>d</sub> : 29 nM)     | 30 (pK <sub>d</sub> < 5.5, K <sub>d</sub> > 3000 nM)    | 35 (pK <sub>d</sub> < 5.5, K <sub>d</sub> > 3000 nM)  |

Binding affinities (K<sub>d</sub>) of the selected compounds were measured in the DiscoverX KINOME scan® platform. Among the nine compounds identified by virtual screening and available for purchase, only the fisetin presented an acceptable rate of inhibition of 76% (K<sub>d</sub>: 270 nM) against SIK2. No other compounds tested returned acceptable results (K<sub>d</sub> > 10,000 nM). The compounds used as positive tests, bosutinib and dasatinib, having a strong kinase inhibition on SIK2 fulfilled their function with both 100% inhibition (K<sub>d</sub>: 29 nM et 6.4 nM respectively).

The molecules were not only tested versus SIK2 kinase but also on a panel of other kinases including BRAF and BRAFV600E, since it was shown that Fisetin has activity against melanoma<sup>[58]</sup>. Against these targets, almost all molecules showed activity with around 10%

inhibition. Fisetin presents the highest inhibition rate with 90% on both kinases, followed with the 6,8-diprenylningenin inhibiting BRAF at 25% and Kushenol F inhibiting BRAF V600E at 15%.

## 4 Discussion and conclusion

UVR is responsible for various physiological effects on the skin, resulting in pigmentation alterations. In response, the melanin secreted by the melanocytes acts as a photoprotective shield inducing skin tanning. Indoor tanning is based on UVR and therefore does not allow the avoiding of DNA damage contrarily to sunless tanning. Topical application of natural products enhancing the melanogenesis remains an opportunity for cosmetic products to reduce

visible sign of photoaging and keep a golden and uniform tan. SIK kinase family are targets of importance to trigger melanogenesis, their inhibition can independently activate the tanning pathway without suffering from the damaging effects of UV.

The increase in the cost and time of the drug development process have led to greater usage of VS approaches. With the increase of the computing power and better consideration of receptor flexibility, these methods tend to become interesting alternatives to traditional screening methods for compounds prioritization. Before running a VS campaign, it is of major importance to prepare the ligand dataset. Molecular compounds should be cleaned from accompanying minor compounds or counter-ions. They should also be standardized with an appropriate protonation state at desired pH and the proper tautomeric forms accounted. Finally, a valid initial 3D conformation must be generated, especially if the chosen software for docking is not able to treat some conformational aspects as ring conformers. The target structure also needs preparation, including among others the assignment of hydrogen atoms, the determination of protonation states of amino acids and a step of minimization to prevent steric clashes. Here, several different protocols and multiple software were tested to retrieve the methodology returning best enrichment results. The best preparation step for molecular compounds was obtained with the VSPrep tool. Among the three docking software used, GOLD was the one giving the best results while Glide could not satisfy our validation request. The free software rDock showed interesting potential and proved it can be a promising tool for screening campaigns on protein kinase binding sites.

Despite a carefully validated strategy and a manual selection of hits after the VS, experimental results did not meet our expectations. Indeed, our methodology failed to find new potential SIK2 inhibitors, the fisetin being already known as SIK inhibitor<sup>[57]</sup>. Multiple factors can explain it, starting with the 3D structure of the target obtained by homology modeling and not experimentally, that might not be as accurate as an experimental structure. Moreover, since no structural information was available, the implication of water molecules in the binding site could not be checked and the VS campaign was performed without taking them into account. Another improvement could be to construct several models with different templates and compare them to get a consensus. Although our goal was to find type 1 inhibitors, realizing a VS on other conformations of the kinase might confirm or reveal other hits. The other important feature that might explain the lack of inhibition is the skeleton of selected compounds for experimental test. Most of the flavonoid chosen scaffolds are flavan based and not flavone based as fisetin or quercetin. However it seems that flavon scaffold gives better results on SIK kinase family<sup>[59]</sup>. Yet, the poor number of experimentally tested molecules (9) does not permit to conclude that our strategy is wrong. As soon as other molecules will be available for purchase, they will be tested to validate or not their inhibition. If proven, this strategy could be used on other kinases implicated in melanogenesis, as for example KIT, to discover natural product inhibitors<sup>[9]</sup>.

NPs remain an interesting class of compounds for marketing assets in the cosmetical field. The discovery of new NPs able to upstream regulate the melanogenesis would be valuable. Moreover, as seen in the study, they can cover a different chemical space than kinase inhibitors which is an important point since the intellectual property space of kinase inhibitors is crowded. However, the consideration for toxicity and safety for such compounds might raise interrogation. Previous results in mice over multiple months of treatment did not show any apparent associated toxicities<sup>[60]</sup>. Furthermore, the increase of melanin, especially for fair-skinned people, could act as an extra-protection against skin melanoma. Of note, even if some compounds exist for sunless tanning, they are not substitute to sunscreen which remains one of the most efficient tools for optimal skin protection against skin photoaging and damage from UVR.

## Author Contributions

The manuscript was written through the contribution of all authors and all of them approved the final version of the manuscript.

## Acknowledgements

Authors gratefully acknowledge the Fondation ARC pour la recherche sur le cancer (grant spring 2016) and the Région Centre Val de Loire (APR-IR isNatProd) which financially support this work. Authors thanks the projects CHemBio (FEDER-FSE 2014-2020-EX003677), Valbiocosm (FEDER-FSE 2014-2020-EX003202), Techsab (FEDER-FSE 2014-2020-EX011313), QUALICHIM (APR-IA-PF 2021-00149467), the RTR Motivhealth (2019-00131403) and the Labex programs SYNORG (ANR-11-LABX-0029) and IRON (ANR-11-LABX-0018-01) for their financial support of ICOA, UMR 7311, University of Orléans, CNRS.

## References

- [1] Y. Gilaberte, L. Prieto-Torres, I. Pastushenko, Á. Juarranz, in *Nanoscience in Dermatology* (Eds.: P. Avci, T.W. Prow), Boston, Hamblin, M.R, **2016**, pp. 1–14.
- [2] J. K. Liu, *Natural Products and Bioprospecting* **2022**, *12*, 40.
- [3] “Skin care industry: global skincare market size 2012-2024,” can be found under <https://www.statista.com/statistics/254612/global-skin-care-market-size/>, accessed on October 9, 2023.
- [4] D. J. Tobin, *Journal of Tissue Viability* **2017**, *26*, 37–46.
- [5] M. Brenner, V. J. Hearing, *Photochem Photobiol* **2008**, *84*, 539–549.
- [6] A. J. Thody, E. M. Higgins, K. Wakamatsu, S. Ito, S. A. Burchill, J. M. Marks, *J. Invest. Dermatol.* **1991**, *97*, 340–344.
- [7] R. Cui, H. R. Widlund, E. Feige, J. Y. Lin, D. L. Wilensky, V. E. Igras, J. D’Orazio, C. Y. Fung, C. F. Schanbacher, S. R. Granter, D. E. Fisher, *Cell* **2007**, *128*, 853–864.
- [8] C. Bertolotto, P. Abbe, T. J. Hemesath, K. Bille, D. E. Fisher, J. P. Ortonne, R. Ballotti, *J. Cell Biol.* **1998**, *142*, 827–835.
- [9] S. A. N. D’Mello, G. J. Finlay, B. C. Baguley, M. E. Askarian-Amiri, *Int J Mol Sci* **2016**, *17*.
- [10] H. Ando, Y. Niki, M. Ito, K. Akiyama, M. S. Matsui, D. B. Yarosh, M. Ichihashi, *J. Invest. Dermatol.* **2012**, *132*, 1222–1229.

- [11] S. Del Bino, C. Duval, F. Bernerd, *Int. J. Mol. Sci.* **2018**, *19*, 2668.
- [12] C. Chang, E. C. Murzaku, L. Penn, N. R. Abbasi, P. D. Davis, M. Berwick, D. Polsky, *Am J Public Health* **2014**, *104*, e92–e99.
- [13] N. Horike, A. Kumagai, Y. Shimono, T. Onishi, Y. Itoh, T. Sasaki, K. Kitagawa, O. Hatano, H. Takagi, T. Susumu, *Pigment Cell Melanoma Res* **2010**, *23*, 809–819.
- [14] A. Jagannath, L. Taylor, Y. Ru, Z. Wakaf, K. Akpobaro, S. Vasudevan, G. R. Foster, *Physiol. Rev* **2023**, *103*, 2231–2269.
- [15] N. Mujahid, Y. Liang, R. Murakami, H. G. Choi, A. S. Dobry, J. Wang, Y. Suita, Q. Y. Weng, J. Allouche, L. V. Kemeny, A. L. Hermann, E. M. Roider, N. S. Gray, D. E. Fisher, *Cell Reports* **2017**, *19*, 2177–2184.
- [16] M. N. Wein, M. Foretz, D. E. Fisher, R. J. Xavier, H. M. Kronenberg, *Trends Endocrinol. Metab.* **2018**, *29*, 723–735.
- [17] M. Emerald, A. Emerald, L. Emerald, V. Kumar, *Pharm Pharmacol Int J.* **2016**, *4*, 339–341.
- [18] A. Fatima, S. Alok, P. Agarwal, P. P. Singh, A. Verma, *Int J Pharm Sci Res* **2013**, *4*, 3746–3760.
- [19] J. F. Fowler, H. Woolery-Lloyd, H. Waldorf, R. Saini, *J Drugs Dermatol* **2010**, *9*, S72–81; quiz s82–83.
- [20] N. Saewan, A. Jimtaisong, *J. Cosmet. Dermatol.* **2015**, *14*, 47–63.
- [21] M.-H. Choi, H.-J. Shin, *Cosmetics* **2016**, *3*, 18.
- [22] H.-S. Yoon, S.-R. Lee, H.-C. Ko, S.-Y. Choi, J.-G. Park, J.-K. Kim, S.-J. Kim, *Biosci. Biotechnol. Biochem.* **2007**, *71*, 1781–1784.
- [23] E. Lionta, G. Spyrou, D. K. Vassilatis, Z. Cournia, *Curr Top Med Chem* **2014**, *14*, 1923–1938.
- [24] S. Riniker, G. A. Landrum, *J. Chem. Inf. Model.* **2015**, *55*, 2562–2574.
- [25] T. A. Halgren, *J Comp Chem* **1996**, *17*, 490–519.
- [26] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727–748.
- [27] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739–1749.
- [28] S. Ruiz-Carmona, D. Alvarez-Garcia, N. Foloppe, A. B. Garmendia-Doval, S. Juhos, P. Schmidtke, X. Barril, R. E. Hubbard, S. D. Morley, *PLoS Comput. Biol.* **2014**, *10*, 1003571.
- [29] W. T. M. Mooij, M. L. Verdonk, *Proteins: Structure, Function, and Bioinformatics* **2005**, *61*, 272–287.
- [30] M. M. Mysinger, M. Carchia, John. J. Irwin, B. K. Shoichet, *J. Med. Chem.* **2012**, *55*, 6582–6594.
- [31] Gally José- Manuel, Bourg Stéphane, Do Quoc- Tuan, Ací- Sèche Samia, Bonnet Pascal, *Molecular Informatics* **2017**, *36*, 1700023.
- [32] Thomas A Caswell, Michael Droettboom, John Hunter, Eric Firing, Antony Lee, David Stansby, Elliott Sales de Andrade, Jens Hedegaard Nielsen, Jody Klymak, Nelle Varoquaux, Benjamin Root, Phil Elson, Darren Dale, Ryan May, Jae-Joon Lee, Jouni K. Seppänen, Tim Hoffmann, Damon McDougall, Andrew Straw, Paul Hobson, cgohlke, Tony S Yu, Eric Ma, Adrien F. Vincent, Steven Silvester, Charlie Moad, Jan Katins, Nikita Kniazhev, Federico Ariza, Peter Würzt, *Matplotlib/Matplotlib v3.0.1*, Zenodo, **2018**.
- [33] Michael Waskom, Olga Botvinnik, Drew O’Kane, Paul Hobson, Joel Ostblom, Saulius Lukauskas, David C Gempferline, Tom Augspurger, Yaroslav Halchenko, John B. Cole, Jordi Warmenhoven, Julian de Ruiter, Cameron Pye, Stephan Hoyer, Jake Vanderplas, Santi Villalba, Gero Kunter, Eric Quintero, Pete Bachant, Marcel Martin, Kyle Meyer, Alistair Miles, Yoav Ram, Thomas Brunner, Tal Yarkoni, Mike Lee Williams, Constantine Evans, Clark Fitzgerald, Brian, Adel Qalieh, *Mwaskom/Seaborn: V0.9.0 (July 2018)*, Zenodo, **2018**.
- [34] T. Sterling, J. J. Irwin, *J. Chem. Inf. Model.* **2015**, *55*, 2324–2337.
- [35] A. Sharma, P. Dutta, M. Sharma, N. K. Rajput, B. Dodiya, J. J. George, T. Kholia, A. Bhardwaj, *J. Cheminformatics* **2014**, *6*, 46.
- [36] K. Nakamura, N. Shimura, Y. Otabe, A. Hirai-Morita, Y. Nakamura, N. Ono, M. A. Ul-Amin, S. Kanaya, *Plant Cell Physiol* **2013**, *54*, e4–e4.
- [37] M. Valli, R. N. dos Santos, L. D. Figueira, C. H. Nakajima, I. Castro-Gamboa, A. D. Andricopulo, V. S. Bolzani, *J. Nat. Prod.* **2013**, *76*, 439–444.
- [38] R. Hatherley, D. K. Brown, T. M. Musyoka, D. L. Penkler, N. Faya, K. A. Lobb, Ö. Tastan Bishop, *J. Cheminformatics* **2015**, *7*, 29.
- [39] D. Klementz, K. Döring, X. Lucas, K. K. Telukunta, A. Erxleben, D. Deubel, A. Erber, I. Santillana, O. S. Thomas, A. Bechthold, S. Günther, *Nucleic Acids Res* **2016**, *44*, D509–D514.
- [40] C. Y.-C. Chen, *PLOS ONE* **2011**, *6*, e15939.
- [41] S.-K. Kim, S. Nam, H. Jang, A. Kim, J.-J. Lee, *BMC Complement Altern Med* **2015**, *15*, DOI 10.1186/s12906-015-0758-5.
- [42] F. Carles, S. Bourg, C. Meyer, P. Bonnet, *Molecules* **2018**, *23*, 908.
- [43] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res* **2000**, *28*, 235–242.
- [44] S. Kaczanowski, P. Zielenkiewicz, *Theor Chem Acc* **2010**, *125*, 643–650.
- [45] J. A. R. Dalton, R. M. Jackson, *Bioinformatics* **2007**, *23*, 1901–1908.
- [46] B. Webb, A. Sali, *Curr Protoc Bioinformatics* **2016**, *54*, 5.6.1-5.6.37.
- [47] Uniprot Consortium, *Nucleic Acids Res* **2019**, *47*, 506–515.
- [48] J. S. Sack, M. Gao, S. E. Kiefer, J. E. Myers, J. A. Newitt, S. Wu, C. Yan, *Acta Cryst F* **2016**, *72*, 129–134.
- [49] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 12–21.
- [50] S. Müller, A. Chaikuad, N. S. Gray, S. Knapp, *Nat. Chem. Biol.* **2015**, *11*, 818–821.
- [51] R. Roskoski, *Pharmacol. Res.* **2016**, *103*, 26–48.
- [52] M. Huse, J. Kuriyan, *Cell* **2002**, *109*, 275–282.
- [53] R. S. K. Vijayan, P. He, V. Modi, K. C. Duong-Ly, H. Ma, J. R. Peterson, R. L. Dunbrack, R. M. Levy, *J Med Chem* **2015**, *58*, 466–479.
- [54] A. P. Kornev, N. M. Haste, S. S. Taylor, L. F. Ten Eyck, *Proc Natl Acad Sci U S A* **2006**, *103*, 17783–17788.
- [55] J. Zheng, E. A. Trafny, D. R. Knighton, N. Xuong, S. S. Taylor, L. F. Ten Eyck, J. M. Sowadski, *Acta Crystallogr. D* **1993**, *49*, 362–365.
- [56] R. Roskoski, *Pharmacological Research* **2016**, *103*, 26–48.
- [57] M.I. Davis, J.P. Hunt, S. Herrgard, P. Ciceri, L.M. Wodicka, G. Pallares, M. Hocker, D.K. Treiber, P.P. Zarrinkar, *Nat Biotechnol* **2011**, *29*, 1046–1051.
- [58] D.N. Syed, F. Afaq, N. Maddodi, J. J. Johnson, S. Sarfaraz, A. Ahmad, V. Setaluri, H. Mukhtar, *J Invest Dermatol.* **2011**, *131*, 1291-1299.
- [59] A. Kumagai, N. Horike, Y. Satoh, T. Uebi, T. Sasaki, Y. Itoh, Y. Hirata, K. Uchio-Yamada, K. Kitagawa, S. Uesato, H. Kawahara, H. Takemori, Y. Nagaoka, *PLoS One* **2011**, *6*, DOI 10.1371/journal.pone.0026148.
- [60] J. A. D’Orazio, T. Nobuhisa, R. Cui, M. Arya, M. Spry, K. Wakamatsu, V. Igras, T. Kunisada, S. R. Granter, E. K. Nishimura, S. Ito, D. E. Fisher, *Nature* **2006**, *443*, 340–344.

Received: ((will be filled in by the editorial staff))

Accepted: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

