Virtual screening of curcumin analogues as DYRK2 inhibitor: Pharmacophore analysis, molecular docking and dynamics, and ADME prediction [version 1; peer review: awaiting peer review]

La Ode Aman¹,², Rahmana Emran Kartasasmita¹, Daryono Hadi Tjahjono⁰¹

¹School of Pharmacy, Bandung Institute of Technology, Jalan Ganesha, 10 Bandung, 40132, Indonesia
²Department of Pharmacy, Universitas Negeri Gorontalo, Gorontalo, 96128, Indonesia

Abstract

Background: Curcumin reduces the proliferation of cancer cells through inhibition of the DYRK2 enzyme, which is a positive regulator of the 26S proteasome.

Methods: In the present work, curcumin analogues have been screened from the MolPort database using a pharmacophore model that comprised a ligand-based approach. The result of the screening was then evaluated by molecular docking and molecular dynamics based on binding the free energy of the interaction between each compound with the binding pocket of DYRK2. The hit compounds were then confirmed by absorption, distribution, metabolism, and excretion (ADME) prediction.

Results: Screening of 7.4 million molecules from the MolPort database afforded six selected hit compounds. By considering the ADME prediction, three prospective curcumin analogues have been selected. These are: 2-[2-(1-methylpyrazol-4-yl)ethyl]-1H,5H,6H,7H,8H-imidazo[4,5-c]azepin-4-one (Molport-035-369-361), methyl 4-(3-hydroxy-1,2-oxazol-5-yl)piperidine-1-carboxylate (Molport-000-004-273) and (1S)-1-[5-(furan-3-carbonyl)-4H,6H,7H-pyrazolo[1,5-a]pyrazin-2-yl]ethanol (MolPort-035-585-822).

Conclusion: Pharmacophore modelling, combined with molecular docking and molecular dynamics simulation, as well as ADME prediction were successfully applied to screen curcumin analogues from the MolPort database as DYRK2 inhibitors. All selected compounds that have better predicted pharmacokinetic properties than that of curcumin are considered for further study.

Keywords
Curcumin analogues, DYRK2, pharmacophore, docking, molecular dynamics simulation, ADME
Corresponding author: Daryono Hadi Tjahjono (daryonohadi@fa.itb.ac.id)

Author roles: Aman LO: Data Curation, Investigation, Methodology, Validation, Writing – Original Draft Preparation; Kartasasmita RE: Formal Analysis, Supervision, Validation, Writing – Review & Editing; Tjahjono DH: Conceptualization, Formal Analysis, Funding Acquisition, Methodology, Resources, Supervision, Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This research was supported partly by the Ministry of Research, Technology and Higher Education, Republic of Indonesia through HIBAH PUPT DIKTI 2017 (009/ADD/SP2H/LT/DRPM/VII/2017), 2018 (127/SP2H/PTNBH/DRPM/2018), 2019 (002/SP2H/PTNBH/DRPM/2019), LOA also received financial support from the Ministry of Research, Technology and Higher Education, Republic of Indonesia through a BPPDN Scholarship 2015-2019 (No. Scholarship Letter: 2279.3/E4.4/2015). The research was also partly funded by a Research and Innovation Grant from the Bandung Institute of Technology, Bandung 2017 (007/I1.B04/PL/2017) for DHT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Aman LO et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Aman LO, Kartasasmita RE and Tjahjono DH. Virtual screening of curcumin analogues as DYRK2 inhibitor: Pharmacophore analysis, molecular docking and dynamics, and ADME prediction [version 1; peer review: awaiting peer review]
F1000Research 2021, 10:394 https://doi.org/10.12688/f1000research.28040.1
First published: 17 May 2021, 10:394 https://doi.org/10.12688/f1000research.28040.1
Introduction
Curcumin is a compound derived from turmeric (Curcuma longa), which is responsible for the yellow rhizome extract coloration. Traditionally, a large number of people in India, China, Indonesia and other Asian countries have applied turmeric powder in therapeutic herbs and as a food additive. The curcumin (diferuloylmethane) constituent is a tautomeric compound known to exist as an enolic form in organic solvents, and in keto form in water. The wide range of biological activities are currently being tested in vivo and in vitro to develop the numerous potentials in treating various diseases. These include the application of curcumin as an antioxidant, antibacterial, antifungal, antiviral, anti-inflammatory, and anti-angiogenic agent. Furthermore, there are reports on the promising anti-Alzheimer, as well as anticancer properties of curcumin, and its antagonistic effects against other degenerative diseases. The curcumin component is also a non-toxic compound, as no toxicity has been reported following the administration of high doses to animals. Previous reports have shown biological activities related to cancer, including lymphomas, breast, prostate, cervical, lung and colorectal cancers, alongside leukemia. There are numerous pathways involved in regulation, including p53, BAX, cyclin D1, various BCL, p21, p27, AKT, COX-2, protein kinase, and others.

The facts and opinions on the biological effects of curcumin as a drug candidate indicate that it is a PAIN (pan-assay interference) compound or IMP (invalid metabolic panacea). Moreover, this unstable compound is known to easily degrade into others, and a total of eight have been reported to date. These include vanillin, ferulic acid, furfural, 2-hydroxy-6-(4-hydroxy-3-methoxyphenyl)-4-oxohexa-2,5-dienal, bicyclopentadiene, furfuraldehyde, vanillic acid, and 4-{1e}-3-{propan-2-yloxy}prop-1-en-1-yl]guaiacol. In addition, numerous pharmacokinetic evaluations have indicated the poor absorption, low solubility, rapid metabolism and elimination as well as poor bioavailability properties of curcumin.

There have been several suggestions of methods to solve these challenges, including co-administration with adjuvants. In addition, studies have shown the possible development into a nanoparticle form, complexations with metallic and radioactive elements, using the derivatives or analogue products, and application in the bioconjugation form. One of the current research strategies involves the screening of compounds from a large database to obtain analogues with the pharmacophore features of curcumin.

One of the purposes of studying curcumin and its analogues is to find analogues that are targeted to reduce cell proliferation by interacting with dual-specificity tyrosine-regulated kinase 2 (DYRK2). This is achieved through the positive regulation of the 26S proteasome, particularly in cancer cells. The inhibition property is observed in terms of cell proliferation with IC₅₀ 5 nM. DYRK2 is a family of protein kinases with members involved in cellular growth and development. In addition, there have also been reports on its function as a tumor suppressant by regulating cell survival, differentiation, proliferation and apoptosis. The mechanism adopted to control further involves the regulation of CDK14 expression. Furthermore, DYRK2 as an enzyme is capable of phosphorylating serine substrates and threonine residues. This action regulates apoptotic cell death in response to DNA damage by impacting the phosphorlation effect of Ser46 on p53. Reports have shown the negative regulatory impact on breast cancer formation through the transcriptional downregulation of Kruppel-like factor 4 (KLF4).

Therefore, the aim of this study is to explore the curcumin analogues for their potential application as a DYRK2 inhibitor through virtual screening by using pharmacophore molecular modelling as well as docking and molecular dynamics. The compounds of screening results are expected to be applied as lead compounds in discovering and developing a prospective anticancer molecule through DYRK2 inhibition.

Methods
Dataset preparation
Figure 1 shows the chemical structure of curcumin and the analogues used to model the dataset for ligand-based pharmacophore. The 2D chemical configuration was constructed with MarvinSketch 19.2, prior to an analysis with LigandScout 4.3 win64 evaluation version (the analysis can be replicated using PharmaGist Webserver). Subsequently, each structure’s geometry was optimized using the energy minimize module with MMFF94 force field set at a default setting.

Protein preparation
5ZTN was the Protein Data Bank accession number of the DYRK2 protein used in this study, and curcumin acted as the native ligand. In addition, the target for molecular docking and dynamics was prepared using Molecular Operating Environment (MOE) 2014.0901 software (this can be replicated using MGLTools 1.5.6 and UCSF Chimera 1.13.1) in order to correct the break residues, charging, and protonation of the protein structure. The protein molecule was opened in...
Figure 1. Chemical structure of dataset compounds.
AutoDockTool (ADT) 1.5.6, and the water molecule(s) were then removed. In grid menu, macromolecule was chosen, and the protein structure was then saved as a.pdbqt extension.

**Ligand-based pharmacophore modeling**

The ligand-based pharmacophore observed in this work was analyzed through multiple flexible alignment. Therefore, the model was generated from 24 dataset compounds using LigandScout 4.3 software (or PharmaGist Webserver). This was achieved through the 3D superposition of chemical features constructed by the flexible conformation alignments of all dataset compounds. In addition, the enrichment factor (EF) and receiver operating characteristics (ROC) analysis were used to validate the pharmacophore model using ROC Analysis: Web-based Calculator for ROC Curves. Active compounds were all of the 24 dataset compounds, and decoy compounds were obtained from zinc decoy database generated via DecoyFinder 2.0.

**Filtering the compounds database**

MolPort provided a large database with over 7.4 million catalogue compounds. The process of curcumin analogue filtration from the compounds database was performed on the Pharmit webserver. Filtering the compounds from database was then conducted using the pharmacophore query file as obtained from the above pharmacophore modeling.

**Molecular docking**

The goal of molecular docking was to assess the binding affinity of compound(s) upon interaction with the receptor. Therefore, all results obtained from the database filtering process were docked to the DYRK2 protein. In addition, the docking module of MOE was used for docking protocol detection and also for the docking score calculation of all hits (this can be replicated using AutoDock 4.2.6). Moreover, the molecular docking protocol was evaluated through virtual screening with the alpha triangle methods, London dG scoring and GridMin refinement.

**Molecular dynamics**

The aim of molecular dynamics was to evaluate the physical movement of molecules and atoms. This activity was intended to simulate the interaction stability between the ligand and DYRK2, and was further investigated in combination with protein-ligand complexes obtained from the docking score calculation and characterized by the highest binding affinity. In addition, the interaction dynamics between ligands and receptors was measured using Gromacs 2018.3.29 The stability of ligands in the binding pocket of DYRK2 protein were simulated by the molecular dynamic for 50 ns. In addition, Gromos96 54a7 force field was used to prepare the protein topology, while the PRODRG webserver was applied for ligand topology, using Gromos forcefield. The complex protein-ligand was solvated in a dodecahedron with 1 nm dimensions. Moreover, an aqueous environment was created in the system with the simple point charge (SPC) water model, and this was neutralized by adding Na+ and Cl⁻. The electrostatic interaction and periodic boundaries were calculated in all conditions using Particle-mesh Edward (PME) methods. Meanwhile, the cut-off radius for short-range van der Waals and Coulomb interactions was set to 0.9 nm. Furthermore, the linear constraint solver for molecular simulations (LINCS) method was used to constraint all bond lengths, while minimization, NPT and NVT equilibration as well as system production were performed at constant temperature (300K) and pressure (1 atm). The minimization process was conducted for 50 ps, NPT and NVT were collectively simulated for 100 ps, while the production process for 50 ns were saved at every 2 ps with coordinates of each simulation.

The interaction of ligand-receptor was visualized with LigandScout 4.3 (and can be replicated in USCF Chimera 1.13.1 and Discovery Studio Visualiser v20. The energy from ligand-receptor interactions were further estimated using the g-mmpbsa31 platform.

**ADME calculation**

The absorption, distribution, metabolism and excretion (ADME) prediction values for hit compounds were calculated using ADMET Prediction by ADMETLab Webserver. Furthermore, ADME properties applied in the current research include caco-2 permeability, bioavailability 30% (F30), plasma protein binding (PPB), blood-brain barrier (BBB), Cyp450 1A2 inhibitor, Cyp450 1A2 substrate, half-life (T1/2) and clearance (CL).

**Results**

**Ligand-based pharmacophore modeling**

The hypothetic pharmacophore was grouped based on the number of features, comprising 3 to 7, and each has 10 models, totaling to 50. Figure 1 shows the structure of the dataset molecules used to construct the pharmacophore model, where the validation process including ROC and EF analysis were implemented. These were performed on 24 active compounds and 717 decoy compounds obtained from zinc decoy database generated via DecoyFinder 2.0. Table 1 summarizes the values of area under the curve (AUC), ROC curve and EF for all models.
Table 1. Pharmacophore model candidates. Groups of model3, model4, model5, model6, and model7 have 3, 4, 5, 6 and 7 chemical features, respectively. Every group has 10 model candidates. AUC, area under the curve; EF, enrichment factor.

| Model   | AUC | EF | Numbers of hits |
|---------|-----|----|-----------------|
|         | 1%  | 5% | 10%  | 100% | 1%  | 5%  | 10%  | 100% |                  |
| Model3-1| 1   | 1  | 1    | 0.78  | 30.9 | 12.5 | 6.3  | 3.8  | 122               |
| Model3-2| 1   | 1  | 1    | 0.79  | 30.9 | 12.5 | 6.3  | 4.3  | 108               |
| Model3-3| 1   | 1  | 1    | 0.93  | 26.5 | 17.5 | 8.8  | 6.4  | 102               |
| Model3-4| 1   | 1  | 1    | 0.93  | 30.9 | 14.2 | 7.9  | 4.1  | 166               |
| Model3-5| 1   | 1  | 1    | 0.93  | 30.9 | 14.2 | 7.1  | 3.9  | 172               |
| Model3-6| 1   | 1  | 1    | 0.93  | 30.9 | 14.2 | 7.1  | 4.2  | 169               |
| Model3-7| 1   | 1  | 1    | 0.93  | 30.9 | 14.2 | 7.5  | 4.2  | 163               |
| Model3-8| 1   | 1  | 1    | 0.93  | 30.9 | 14.2 | 7.1  | 3.9  | 174               |
| Model3-9| 1   | 1  | 1    | 0.93  | 30.9 | 14.2 | 7.5  | 4.2  | 162               |
| Model3-10|1   |1  |1    |0.93  |30.9 |14.2 |7.5  |4.1  |164               |
| Model4-1| 1   | 1  | 1    | 0.86  | 30.9 | 12.5 | 9.7  | 9.7  | 57                |
| Model4-2| 1   | 1  | 1    | 0.86  | 30.9 | 12.5 | 9.4  | 9.4  | 59                |
| Model4-3| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.9  | 94                |
| Model4-4| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.8  | 95                |
| Model4-5| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.4  | 101               |
| Model4-6| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.8  | 95                |
| Model4-7| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 7    | 92                |
| Model4-8| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.8  | 96                |
| Model4-9| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.3  | 103               |
| Model4-10|1  |1  |1    |0.85  |30.9 |14.2 |11.7 |11.7 |45                |
| Model5-1| 1   | 1  | 1    | 0.89  | 30.9 | 15   | 7.9  | 7    | 322               |
| Model5-2| 1   | 1  | 1    | 0.96  | 30.9 | 16.7 | 9.6  | 2.6  | 269               |
| Model5-3| 1   | 1  | 1    | 0.92  | 30.9 | 15.9 | 8.8  | 7    | 93                |
| Model5-4| 1   | 1  | 1    | 0.92  | 26.5 | 15   | 8.8  | 3.6  | 181               |
| Model5-5| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.7  | 97                |
| Model5-6| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.5  | 99                |
| Model5-7| 1   | 1  | 1    | 0.92  | 30.9 | 14.2 | 8.8  | 6.9  | 94                |
| Model5-8| 1   | 1  | 1    | 0.92  | 26.5 | 15   | 8.3  | 5    | 129               |
| Model5-9| 1   | 1  | 1    | 0.95  | 30.9 | 14.2 | 8.8  | 2.8  | 256               |
| Model5-10|1  |1  |1    |0.85  |26.5 |14.2 |7.5  |4.9  |114               |
| Model6-1| 1   | 1  | 1    | 0.87  | 30.9 | 15   | 7.5  | 3.7  | 159               |
| Model6-2| 1   | 1  | 1    | 0.87  | 30.9 | 15   | 10.3 | 10.3 | 54                |
| Model6-3| 0.86|0.97 |0.99 |0.87 |26.5 |14.2 |7.9  |4    |148               |
| Model6-4| 1   | 1  | 1    | 0.96  | 30.9 | 16.7 | 9.2  | 2.8  | 256               |
| Model6-5| 1   | 1  | 1    | 0.92  | 30.9 | 16.7 | 8.3  | 6.7  | 97                |
| Model6-6| 1   | 1  | 1    | 0.92  | 30.9 | 15.9 | 8.3  | 6.4  | 101               |
| Model6-7| 1   | 1  | 1    | 0.96  | 30.9 | 15   | 9.6  | 2.8  | 257               |
| Model6-8| 1   | 1  | 1    | 0.97  | 30.9 | 16.7 | 9.6  | 2.8  | 254               |
| Model6-9| 1   | 1  | 1    | 0.96  | 30.9 | 15.9 | 9.2  | 2.8  | 257               |
| Model6-10|1   |1  |1    |0.91  |26.5 |15   |8.3  |3.6  |181               |
Table 1. Continued

| Model    | AUC | EF | Numbers of hits |
|----------|-----|----|-----------------|
|          | 1%  | 5% | 10% | 100% | 1%  | 5%  | 10% | 100% |          |
| Model7-1 | 1   | 1  | 1   | 0.87 | 26.5 | 15  | 7.5 | 4.2  | 139     |
| Model7-2 | 1   | 1  | 1   | 0.91 | 30.9 | 14.2| 7.9 | 3.1  | 211     |
| Model7-3 | 1   | 1  | 1   | 0.87 | 30.9 | 15  | 10.5| 10.5 | 53      |
| Model7-4 | 1   | 1  | 1   | 0.87 | 30.9 | 15  | 7.5 | 4.4  | 134     |
| Model7-5 | 1   | 1  | 1   | 0.9  | 26.5 | 14.2| 8.3 | 2.6  | 252     |
| Model7-6 | 1   | 1  | 1   | 0.9  | 26.5 | 14.2| 8.3 | 2.5  | 258     |
| Model7-7 | 1   | 1  | 1   | 0.87 | 30.9 | 14.2| 7.5 | 4.1  | 143     |
| Model7-8 | 1   | 1  | 1   | 0.87 | 26.5 | 14.2| 7.5 | 3.8  | 155     |
| Model7-9 | 1   | 1  | 1   | 0.87 | 26.5 | 15  | 7.5 | 4.1  | 142     |
| Model7-10| 1   | 1  | 1   | 0.91 | 30.9 | 15.9| 8.3 | 3.3  | 196     |

Figure 2. Receiver operating characteristics (ROC) curve with the area under the curve (AUC) and enrichment factor (EF) values in 1%, 5%, 10% and 100% of selected pharmacophore model.

Table 1 shows the adequacy of AUC values for all models, where the highest EF value of $EF^{100\%}$ was observed in model4-10 and was consequently selected as the best model. Figure 2 shows the ROC curve indicating the composition of 45 hit compounds, including 17 active and 28 decoy. Meanwhile, AUC values shown on 1%, 5%, 10% and 100% were 1.00, 1.00, 1.00 and 0.85, respectively, while the corresponding EF values were 30.8, 14.2, 11.7, and 11.7.

Filtering the curcumin analogues from database

The output file of the selected pharmacophore model generated from LigandScout 4.3 (or PharmaGist Webserver) was used to filter the compound database. Furthermore, the Pharmit interface was used for 3D visualization of the features, including details on the coordinate position and radius. The presence of hydrophobic, H-bond acceptor and donor features with radius 1.5 Å, and 0.9 Å for the aromatic variant were observed in the default setting. Moreover, filtering on Pharmit...
allowed the users to modify the feature’s radius, and consequently increase or decrease the amount of hit(s) as a result. However, manually changing this value is also possible by modifying the hit reduction and screening criteria. The Pharmit developers enter some criteria to set up the maximum hits per conformation, and molecule number, as well as the total limit for reduction, molar weight, rotated bond, logP, polar surface area (PSA), aromatic ring, H-bond donor and acceptor for screening.

This novel work involves database filtering with default settings for pharmacophore features and hit screening, with the exception of reduction. The model ultimately produced 1,130 hits, and a lot more compounds exist for continued screening. Therefore, the feature radius and screening procedure were modified. In particular, the hydrophobic, H-bond acceptor and donor radius were reduced to 1.4 Å for each, while the Rule of Three (RO3) was applied during hit screening. These rules include molar weight (300 Dalton), rotated bond (3), logP (3), aromatic ring (3), H-bond donor (3) and H-bond acceptor (3). Meanwhile, PSA was set to the maximum value according to another reference, at 90 Å, and these filtering protocols collectively produced 566 hits.

### Molecular docking

Binding affinity is an indicator of the connection strength between ligand and receptor. This was determined by a docking score calculated using the MOE software (and can be replicated using AutoDock 4.2.6). This procedure was then validated by redocking the native ligand of DYRK2 protein (5ZTN) present in curcumin. The lowest RMSD value for successful docking was 0.7788, indicating the propensity to apply this protocol to other ligands.

This protocol docking validation process involved calculating the binding affinity (docking score) for each curcumin analogue complex and DYSRK2 protein. Moreover, the value obtained for curcumin during redocking was −12.46 kcal/mol. This was used to filter the hits, as the specimens with greater and relatively close values between the native ligand and the binding pocket of target were selected. Figure 3 shows the chemical structure of ligand (hit compounds) and ligand-DYRK2 interaction with docking score lower than −10 kcal/mol, and Figure 4 indicates the overlay of each, with pose characterized by the highest value.

### Molecular dynamics

The complexes (ligand-protein) motion during simulation were expressed in Root Mean Square Deviation (RMSD) (Figure 4(a)), while the movements of protein backbone during simulation was estimated in a Root Mean Square Fluctuation (RMSF) curve (Figure 4(b)). Figure 4(a) showed that the interactions are stable after 30 ns of simulation.

### ADME Prediction

The ADME (absorption, distribution, metabolism, and excretion) provide a description for drug disposition within an organism. These pharmacokinetic parameters influence the overall level and kinetics in the tissues, and consequently influence the pharmacological effect of the active compounds. In addition, it is possible to represent the ADME of active compounds as a prediction value.

The results obtained with the selected compounds in this current investigation were caco-2 permeability and F30 for absorption; PB and BBB for distribution, Cyp450 1A2 inhibitor and Cyp450 1A2 substrate for metabolism; while T1/2 and CL represented excretion. Table 2 shows the summary of ADME prediction for selected compounds in contrast with curcumin.

### Discussion

#### Ligand-based pharmacophore modeling

The pharmacophore is the physicochemical feature of a molecule known to interact with a specific target receptor. This is modelled in a 3D pattern and the basic characteristics are also shared by a set of molecules. The ligand-based pharmacophore model was constructed without needing a protein target structure, through a 3D superposition of the ligand conformed physicochemical features.

The chemical characteristics of a selected pharmacophore model include three H-bond acceptors provided by the oxygen in keto, methoxy and hydroxyl groups, a H-bond donor from the hydroxyl, an aromatic feature and hydrophobicity contributed by aromatic group. Figure 5(a) shows the 2D visualization of pharmacophore model, while Figure 5(b) and 5(c) demonstrates it in 3D.

### Filtering and interaction study of curcumin analogue

There is rule of three (RO3) beside Lipinski’s rule of five (RO5), which applies in searches pertaining to drug-likeness. Particularly, RO5 stipulates the criteria to be satisfied, including the presence of no more than five hydrogen bond donors, 10 hydrogen bond acceptors, a molecular weight (MW) of less than 500 Daltons, and an octanol-water partition coefficient...
Figure 3. Chemical structure of hit compounds, its interaction with dual-specificity tyrosine-regulated kinase 2 (DYRK2), and their docking score $S (<-10 \text{ kcal/mol})$. 
Furthermore, all numbers are in the multiple of five, hence the RO3 criteria requires a multiple of three. These criteria include logP not greater than 3, MW less than 300 Daltons, the presence of no more than three hydrogen bond donors and acceptors, respectively, and no more than three rotatable bonds.36

The Pharmit platform allows users to customize the pharmacophore feature criteria based on RO5 or RO3. In addition, the value modification potentially increases or decreases the amount of hits. The application of RO3 in this research instigated a decline in the number of hit compounds from 1,130 (default criteria) to 566 (modified criteria). The latter was determined to be more rational during molecule screening.

Figure 4. Root Mean Square Deviation (RMSD) of ligand (a), and Root Mean Square Fluctuation (RMSF) of protein backbone (b) of all complexes in 50 ns simulation.
Molecular docking
Molecular docking is one of the in silico approaches of interaction studies between ligand(s) and receptor. The 566 hits are further filtered through this means, and 11 compounds were selected after using the specified protocol. In addition, the binding affinity (docking score) of samples in each hit were better than or close to curcumin (native ligand). Figure 6 shows the respective overlay with the binding pocket of DYRK2 protein. It was observed that all selected hit compounds occupy the DYRK2's binding site.

Molecular dynamics
Figure 4(a, b) shows the stability performance of RMSD for each ligand, as well as the RMSF for the respective protein backbone on each complex, after the initial 25 ns simulation. Therefore, an average of the system binding energy was calculated at the end. The van der Waals, electrostatic, polar solvation, non-polar (Solvent-Accessible Surface Area) and binding energy were calculated every 200 ps, in order to obtain an average value for each complex. Table 3 shows the summarized system energy report obtained from 200 snapshots.

The native ligand (curcumin) has a binding energy of $-53.058 \text{ kJ/mol}$ with the DYRK2 protein. Table 3 showed six selected hit compounds have more negative values compared to that of Curcumin, i.e. Molport-035-369-361 ($-71.35 \pm 24.85$), Molport-000-004-273 ($-83.56 \pm 18.24$), MolPort-029-697-986 ($-61.69 \pm 14.89$), MolPort-035-585-822 ($-74.24 \pm 15.55$), MolPort-000-156-336 ($-54.49 \pm 15.09$), and MolPort-002-747-457 ($-69.36 \pm 12.58$).

Table 2. The absorption, distribution, metabolism, and excretion (ADME) prediction of hit compounds and curcumin. F30, 30% bioavailability; PPB, plasma protein binding; BBB, blood-brain barrier, $T_{1/2}$, half-life CL, clearance.

| No. | Compound          | caco2 | F30 | PPB | BBB |
|-----|-------------------|-------|-----|-----|-----|
|     |                   |       | Category | Probability |       | Category | Probability |
| 1   | Curcumin          | -5.133| 0   | 0.452| 88.84| 1         | 0.579      |
| 2   | Molport-046-141-977| -4.969| 1   | 0.628| 20.79| 0         | 0.214      |
| 3   | Molport-046-149-133| -4.993| 1   | 0.628| 17.27| 0         | 0.277      |
| 4   | Molport-045-913-745| -4.914| 1   | 0.748| 34.89| 1         | 0.855      |
| 5   | Molport-035-369-361| -4.658| 1   | 0.692| 40.05| 1         | 0.890      |
| 6   | Molport-000-004-273| -4.869| 1   | 0.531| 36.79| 1         | 0.931      |
| 7   | MolPort-029-697-986| -5.029| 0   | 0.495| 44.41| 1         | 0.830      |
| 8   | Molport-046-067-592| -5.007| 1   | 0.638| 54.75| 1         | 0.789      |
| 9   | MolPort-028-957-184| -4.968| 1   | 0.652| 50.57| 1         | 0.986      |
| 10  | MolPort-035-585-822| -4.906| 1   | 0.672| 60.69| 1         | 0.845      |
| 11  | MolPort-000-156-336| -4.933| 1   | 0.536| 41.09| 0         | 0.183      |
| 12  | MolPort-002-747-457| -4.951| 1   | 0.652| 48.23| 1         | 0.726      |

| No. | Compound          | CYP1A2-inhibitor | CYP1A2-substrate | $T_{1/2}$ | CL |
|-----|-------------------|------------------|------------------|-----------|----|
|     |                   | Category | Probability | Category | Probability |       |       |
| 1   | Curcumin          | 0       | 0.449    | 0        | 0.456       | 1.653 | 1.560 |
| 2   | Molport-046-141-977| 1      | 0.586    | 0        | 0.354       | 0.906 | 1.421 |
| 3   | Molport-046-149-133| 1      | 0.567    | 0        | 0.362       | 1.069 | 1.291 |
| 4   | Molport-045-913-745| 0      | 0.037    | 0        | 0.460       | 1.260 | 1.910 |
| 5   | Molport-035-369-361| 0      | 0.158    | 1        | 0.678       | 1.310 | 1.991 |
| 6   | Molport-000-004-273| 0      | 0.103    | 0        | 0.462       | 0.910 | 1.756 |
| 7   | MolPort-029-697-986| 0      | 0.103    | 0        | 0.410       | 0.953 | 1.953 |
| 8   | Molport-046-067-592| 1      | 0.655    | 1        | 0.578       | 0.860 | 1.727 |
| 9   | MolPort-028-957-184| 1      | 0.910    | 1        | 0.542       | 0.847 | 1.572 |
| 10  | MolPort-035-585-822| 0      | 0.133    | 1        | 0.552       | 1.034 | 1.776 |
| 11  | MolPort-000-156-336| 0      | 0.034    | 1        | 0.608       | 0.972 | 1.801 |
| 12  | MolPort-002-747-457| 1      | 0.837    | 0        | 0.462       | 0.765 | 1.686 |

Molecular docking
Molecular docking is one of the in silico approaches of interaction studies between ligand(s) and receptor. The 566 hits are further filtered through this means, and 11 compounds were selected after using the specified protocol. In addition, the binding affinity (docking score) of samples in each hit were better than or close to curcumin (native ligand). Figure 6 shows the respective overlay with the binding pocket of DYRK2 protein. It was observed that all selected hit compounds occupy the DYRK2’s binding site.

Molecular dynamics
Figure 4(a, b) shows the stability performance of RMSD for each ligand, as well as the RMSF for the respective protein backbone on each complex, after the initial 25 ns simulation. Therefore, an average of the system binding energy was calculated at the end. The van der Waals, electrostatic, polar solvation, non-polar (Solvent-Accessible Surface Area) and binding energy were calculated every 200 ps, in order to obtain an average value for each complex. Table 3 shows the summarized system energy report obtained from 200 snapshots.

The native ligand (curcumin) has a binding energy of $-53.058 \text{ kJ/mol}$ with the DYRK2 protein. Table 3 showed six selected hit compounds have more negative values compared to that of Curcumin, i.e. Molport-035-369-361 ($-71.35 \pm 24.85$), Molport-000-004-273 ($-83.56 \pm 18.24$), MolPort-029-697-986 ($-61.69 \pm 14.89$), MolPort-035-585-822 ($-74.24 \pm 15.55$), MolPort-000-156-336 ($-54.49 \pm 15.09$), and MolPort-002-747-457 ($-69.36 \pm 12.58$).
This shows that van der Waals and polar solvation energy are the main impact on binding factors for curcumin and all selected compounds. The respective values were negative and positive. This indicates the tendency to generate more positive effect on binding energy at more negative van der Waals. The binding energy of all six selected compounds were larger (more negative) than that of curcumin. This discrepancy was attributed to the lesser polar solvation characteristics. Moreover, significantly higher values of positive polar solvation energy tend to decrease binding energy.

Figure 7 showed the position of curcumin and other selected hit compounds in the binding pocket of DYRK2 during the molecular dynamic simulation. These results confirm those of our docking study. The interaction of 5ZTN-curcumin (Figure 8(a)) and 5ZTN-Molport 000-004-273 (Figure 8(b)) show the binding mode which involves H-bonds and hydrophobic interactions.
Table 3. Binding energy of curcumin and hit compounds upon interaction with dual-specificity tyrosine-regulated kinase 2 (DYRK2) protein obtained from molecular dynamics simulation. SASA, Solvent-Accessible Surface Area.

| No. | Compound | van der Waals ($E_{VDW}$ (kJ/mol)) | Electrostatic ($E_{El}ec$ (kJ/mol)) | Polar solvation ($E_{polar}$ (kJ/mol)) | SASA ($E_{nonpolar}$ (kJ/mol)) | Binding energy ($E_{bind}$) (kJ/mol) |
|-----|----------|----------------------------------|------------------------------------|---------------------------------------|--------------------------------|-----------------------------------|
| 1   | Curcumin | -112.43 ± 18.69                  | -151.33 ± 22.12                   | 229.65 ± 25.56                       | -18.95 ± 1.38                  | -53.06 ± 20.80                    |
| 2   | Molport-046-141-977 | -64.33 ± 16.76                  | -97.88 ± 13.93                   | 141.26 ± 20.85                       | -10.81 ± 0.76                  | -31.75 ± 12.41                    |
| 3   | Molport-046-149-133 | -60.31 ± 12.01                  | -79.93 ± 14.52                   | 115.43 ± 19.96                       | -9.55 ± 0.58                   | -34.37 ± 13.61                    |
| 4   | Molport-045-913-745 | -141.79 ± 12.87                 | -71.03 ± 14.65                   | 180.21 ± 19.52                       | -16.46 ± 0.97                  | -49.07 ± 14.99                    |
| 5   | Molport-035-369-361 | -139.94 ± 13.10                 | -33.42 ± 30.50                   | 117.32 ± 42.56                       | -15.31 ± 1.19                  | -71.35 ± 24.85                    |
| 6   | Molport-000-004-273 | -128.37 ± 12.70                 | -69.26 ± 8.73                    | 128.63 ± 19.28                       | -14.56 ± 0.78                  | -83.56 ± 18.24                    |
| 7   | MolPort-029-697-986 | -104.59 ± 18.68                 | -116.96 ± 13.94                  | 174.61 ± 15.54                       | -14.75 ± 0.92                  | -61.69 ± 14.89                    |
| 8   | Molport-046-067-592 | -102.19 ± 15.47                 | -98.01 ± 22.78                   | 177.08 ± 31.05                       | -13.44 ± 0.98                  | -36.55 ± 18.86                    |
| 9   | MolPort-028-957-184 | -83.91 ± 11.48                  | -107.11 ± 28.64                  | 179.02 ± 46.31                       | -12.68 ± 1.10                  | -24.68 ± 25.28                    |
| 10  | MolPort-035-585-822 | -135.86 ± 11.74                 | -80.71 ± 15.28                   | 158.27 ± 18.93                       | -15.94 ± 0.82                  | -74.24 ± 15.55                    |
| 11  | MolPort-000-156-336 | -83.75 ± 10.68                  | -84.63 ± 16.67                   | 126.46 ± 19.53                       | -12.58 ± 0.66                  | -54.49 ± 15.09                    |
| 12  | MolPort-002-747-457 | -86.45 ± 13.99                  | -103.17 ± 26.38                  | 133.69 ± 15.44                       | -13.43 ± 1.06                  | -69.36 ± 12.58                    |
Table 4 summarizes the percentage of H-bond occupancy between the selected hit compounds and residues of DYRK2. It was observed that LYS178, GLU193, ASP295 have important roles in the ligand interaction in the DYRK2 binding pocket.

**ADME prediction**

The absorption prediction results recognized all 11 selected compounds to be better than curcumin in terms of Caco-2 permeability, where the optimal output for active compound is suggested to be more than $-5.15 \text{ log cm/s}$. $^{37}$ The calculated values were above $-5.0$, whereas a value of $-5.13$ was recorded for curcumin. In addition, the bioavailability (F30) was also better, as the curcumin was categorized 0 ($<30$), while the selected compounds were attributed as 1 ($>30\%$) except MolPort-029-697-986 or tert-butyl-3-hydroxy-2H,4H,5H,7H,8H-pyrazol[3,4-d] azepine-6-carboxylate.$^{38}$

The PPB was higher in curcumin, at 88.84% compared to the others. Hence, the native ligand is considered to have a more significant bond with plasma protein, although less than the suggested 90%, and consequently has a lower therapeutic

**Figure 7.** Overlay binding pose of hits in dual-specificity tyrosine-regulated kinase 2 (DYRK2) binding pocket at 0 ns (brown) and 50 ns (blue).
Conversely, the selected compounds have a greater propensity to attach to the target receptors and provide the desired therapeutic effect.

The BBB is a highly selective semipermeable border separating the circulating blood from the brain and other extracellular fluids in the central nervous system (CNS). This parameter is calculated as a ratio of compounds, and is improved by H-bond numbers as well as molecular weight. The ADMETLab webserver categorizes the BBB value of a drug as 0 (BBB negative) indicating the inability to penetrate, whereas 1 (BBB positive) demonstrates a barrier permeation potential. In addition, all compounds, including the curcumin were categorized with a BBB value of 1, and a positive BBB is derived in instances where the ratio probability was above 0.1.

The cytochrome P450 (CYP) gene family are responsible for drug metabolism. Here, the ADMETLab webserver was used to predict the interaction potentials of curcumin and all selected compounds as an inhibitor and substrate of the cytochrome P450 enzyme. However, the drugs were unable to act as inhibitors but could act as a substrate of cytochrome P450. The prediction results showed the tendency for 2-[2-(1-methylpyrazol-4-yl)ethyl]-1H,5H,6H,7H,8H-imidazo[4,5-c]azepin-4-one (Molport-035-369-361), (1S)-1-[5-(furan-3-carbonyl)-4H,6H,7H-pyrazolo[1,5-a]pyrazin-2-yl] ethanol (MolPort-035-585-822), and methoxytyramine (MolPort-000-156-336) to act as a substrate and not as CYP1A2 inhibitor. Furthermore, 4-hydroxy-1-[2-(imidazol-1-yl)ethyl]-6-methylpyridin-2-one (Molport-046-067-592), and 3-(1,3-oxazol-5-yl) phenylboronic acid (MolPort-028-957-184) potentially act as substrate and inhibitor. Curcumin or (1E,4Z,6Z)-5-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)hepta-1,4,6-trien-3-one, Methyl 4-(3-hydroxy-1,2-oxazol-5-yl)piperidine-1-carboxylate (Molport-000-004-273), and tert-butyl 3-hydroxy-2H,4H,5H,7H,8H-pyrazolo[3,4-d]azepine-6-carboxylate (MolPort-029-697-986)
### Table 4. Occupancy of hydrogen bond during 50 ns simulations.

| Acceptor-Donor | Occupancy (%) |
|----------------|---------------|
| Curcumin       |               |
| O16-SER232     | 66.64         |
| O41-ASP295     | 53.44         |
| O4-GLU237      | 75.10         |
| LYS178-O41     | 37.11         |
| LYS241-O4      | 18.67         |
| ASN234-O2      | 3.80          |
| Molport-046-141-977 | 13.11 |
| LYS178-O1      | 13.11         |
| O1-ASP295      | 86.55         |
| LYS178-N2      | 1.04          |
| Molport-046-149-133 | 88.90 |
| O1-GLU193      | 88.90         |
| LYS178-O1      | 2.32          |
| ASP295-O1      | 14.07         |
| PHE296-O1      | 1.52          |
| LYS178-N2      | 5.90          |
| Molport-045-913-745 | 29.65 |
| LYS178-O1      | 29.65         |
| N1-GLU193      | 3.32          |
| Molport-035-369-361 | 9.06 |
| N4-LEU230      | 9.06          |
| ASN234-O1      | 11.76         |
| LYS178-N2      | 80.34         |
| Molport-000-004-273 | 86.88 |
| O3-GLU193      | 86.88         |
| PHE296-O3      | 1.86          |
| LYS178-O3      | 1.78          |
| ASP295-O3      | 3.30          |
| MolPort-029-697-986 | 78.81 |
| O3-GLU193      | 78.81         |
| LYS178-O3      | 21.21         |
| N3-GLU193      | 11.24         |
| Molport-046-067-592 | 75.01 |
| O1-GLU193      | 75.01         |
| LYS178-O1      | 13.91         |
| LYS178-O2      | 17.21         |
| GLN189-O1      | 3.18          |
| MolPort-028-957-184 | 3.52 |
| ASP295-O1      | 3.52          |
| LYS178-O1      | 14.32         |
| MolPort-035-585-822 | 66.42 |
| O2-ASP295      | 66.42         |
| O1-ASP295      | 56.60         |
| O1-ASP275      | 9.86          |
| MolPort-000-156-336 | 52.82 |
| LYS178-O3      | 52.82         |
| LEU231-O1      | 21.69         |
| O3-GLU193      | 18.84         |
| MolPort-000-156-336 | 36.86 |
| O3-ASP295      | 36.86         |
| LYS178-O2      | 39.70         |
| O2-GLU193      | 56.06         |
| MolPort-002-747-457 | 47.72 |
| O3-GLU193      | 47.72         |
| O4-ASP295      | 67.69         |
were predicted as being unable to act as inhibitor and a substrate of cytochrome. The three other compounds can only act as a cytochrome inhibitor but cannot act as a substrate.

CL refers to the volume of plasma of the drug freed per unit time, where (T₁/₂) is half-life in plasma. The results indicate lower values with hits and curcumin at < 5 mL/min/kg and < 3 hours, respectively. Curcumin possessed the comparable CL and the highest T₁/₂. Thus, selected hit compounds are considered to have relatively better excretion properties.

Finally, the selected compounds demonstrated similar or better overall pharmacokinetic parameters than that of curcumin, as observed with the ADME discussed above.

Conclusions
A combined pharmacophore model and molecular docking for virtual screening has been conducted to find a potential DYRK2 inhibitor.

Based on a gradual virtual screening process using a ligand-based pharmacophore model and molecular docking, 11 hit compounds have been selected. Further detailed study using molecular dynamics simulation afforded six hit compounds with better binding interaction with DYRK2 compared to that of curcumin, i.e. Molport-035-369-361 (-71.35±24.85) kJ/mol, Molport-000-004-273 (-83.56±18.24) kJ/mol, MolPort-029-697-986 (-61.69 ±14.89) kJ/mol, MolPort-035-585-822 (-74.24±15.55) kJ/mol, MolPort-000-156-336 (-54.49±15.09) kJ/mol, and MolPort-002-747-457 (-69.36±12.58) kJ/mol.

The six compounds obtained after a gradual virtual screening process have similar pharmacophore characteristics. Considering the pharmacokinetic properties, Molport-035-369-361, MolPort-035-585-822 as well as Molport-000-004-273 are now under in vitro study for further investigation as lead compounds, and the results will be reported elsewhere.

Data availability
Source data
Protein Data Bank: DYRK2 data from Protein Data Bank. Accession number PBD5ZTN; https://identifiers.org/structure/5ztnt.

Dataset of compounds with biological activity for ligand-based pharmacophore modeling were obtained from21 and are shown in Figure 1.

Acknowledgements
LOA wishes to thank the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia for their financial support through a BPPDN Scholarship.

References
1. Aggarwal B, Sundaram C, Malani N: Curcumin: the Indian solid gold. The Molecular Targets and Therapeutic Uses of Curcumin in Health and Disease. 2007; 1–75. PubMed Abstract | Publisher Full Text
2. Akram M, Shahab-Uddin A, Usmanghani K: Curcuma longa and curcumin: A review article. Rom J Biol Plant. 2010; 55(2): 65–70.
3. Alamdari N, O’Neal P, Hasselgren PO: Curcumin and muscle wasting: A new role for an old drug? Nutrition. 2009; 25(2): 125–129. PubMed Abstract | Publisher Full Text | Free Full Text
4. Anand P, Thomas SG, Kunnunakkara AB, et al.: Biological activities of curcumin and its analogues (Congeners) made by man and Mother Nature. Biochem Pharmacol. 2008; 76(11): 1590–1611. PubMed Abstract | Publisher Full Text
5. Anand P, Kunnunakkara AB, Newman RA, et al.: Bioavailability of Curcumin: Problems and Promises reviews Bioavailability of Curcumin: Problems and Promises. 2007; 4(November): 807–818. PubMed Abstract | Publisher Full Text
6. Manolova Y, Deneva V, Antonov L, et al.: The effect of the water on the curcumin tautomerism: A quantitative approach.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com