

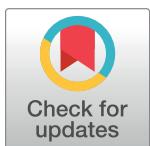
## RESEARCH ARTICLE

# Insights into the computer-aided drug design and discovery based on anthraquinone scaffold for cancer treatment: A systematic review

Hui Ming Chua<sup>1\*</sup>, Said Moshawih<sup>1</sup>, Nurolaini Kifli<sup>1\*</sup>, Hui Poh Goh<sup>1</sup>, Long Chiau Ming<sup>1,2</sup>

**1** PAP Rashidah Sa'adatul Bolkiah Institute of Health Sciences, Universiti Brunei Darussalam, Gadong, Brunei Darussalam, **2** School of Medical and Life Sciences, Sunway University, Bandar Sunway, Malaysia

\* [nurolaini.kifli@ubd.edu.bn](mailto:nurolaini.kifli@ubd.edu.bn) (NK); [hmchua.pharm@gmail.com](mailto:hmchua.pharm@gmail.com) (HMC)



## Abstract

### Background

In the search for better anticancer drugs, computer-aided drug design (CADD) techniques play an indispensable role in facilitating the lengthy and costly drug discovery process especially when natural products are involved. Anthraquinone is one of the most widely-recognized natural products with anticancer properties. This review aimed to systematically assess and synthesize evidence on the utilization of CADD techniques centered on the anthraquinone scaffold for cancer treatment.

### Methods

The conduct and reporting of this review were done in accordance to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) 2020 guideline. The protocol was registered in the “International prospective register of systematic reviews” database (PROSPERO: CRD42023432904) and also published recently. The search strategy was designed based on the combination of concept 1 “CADD or virtual screening”, concept 2 “anthraquinone” and concept 3 “cancer”. The search was executed in PubMed, Scopus, Web of Science and MedRxiv on 30 June 2023.

### Results

Databases searching retrieved a total of 317 records. After deduplication and applying the eligibility criteria, the final review ended up with 32 articles in which 3 articles were found by citation searching. The CADD methods used in the studies were either structure-based alone (69%) or combined with ligand-based methods via parallel (9%) or sequential (22%) approaches. Molecular docking was performed in all studies, with Glide and AutoDock being the most popular commercial and public software used respectively. Protein data bank was used in most studies to retrieve the crystal structure of the targets of interest while the main ligand databases were PubChem and Zinc. The utilization of in-silico techniques

### OPEN ACCESS

**Citation:** Chua HM, Moshawih S, Kifli N, Goh HP, Ming LC (2024) Insights into the computer-aided drug design and discovery based on anthraquinone scaffold for cancer treatment: A systematic review. PLoS ONE 19(5): e0301396. <https://doi.org/10.1371/journal.pone.0301396>

**Editor:** Daryono Hadi Tjahjono, School of Pharmacy - Bandung Institute of Technology, INDONESIA

**Received:** December 9, 2023

**Accepted:** March 14, 2024

**Published:** May 22, 2024

**Copyright:** © 2024 Chua 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 author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting Information](#) files.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

has enabled a deeper dive into the structural, biological and pharmacological properties of anthraquinone derivatives, revealing their remarkable anticancer properties in an all-rounded fashion.

## Conclusion

By harnessing the power of computational tools and leveraging the natural diversity of anthraquinone compounds, researchers can expedite the development of better drugs to address the unmet medical needs in cancer treatment by improving the treatment outcome for cancer patients.

## Introduction

The public health burden of cancer is rising rapidly. The American Cancer Society estimates that in year 2023 alone, nearly 2 million new cancer cases and half a million of cancer deaths will occur in the United States [1]. At the global level, 28.4 million of cancer cases is projected in year 2040 by the International Agency for Research on Cancer (IARC), translated to a 47% increment in 20 years of time [2]. Despite the remarkable achievements in oncology research for the past few decades that are extending the lives of many patients, there are still unmet medical needs due to resistance and relapse after a certain time of treatment [3]. On the other hand, almost all types of cancer treatment modalities cause different degrees of side effects, jeopardizing the patients' quality of life. For instance, conventional cancer therapies frequently result in organ toxicity leading to long term complications, and even the more advanced remedies such as immunotherapy may cause serious or deadly allergic reactions [4]. Therefore, the search for better anticancer drug with a good balance in between efficacy and safety continues to attract the attention of researchers.

It is an established fact that the journey to discover a novel drug is long, costly and fraught with challenges. A recent systematic review revealed that the research and development (R&D) cost of a new molecular entity could reach USD4.54 billion in estimation, with anticancer drug being the most expensive therapeutic to make [5]. Worst still, huge investment cost does not guarantee success in bringing one new medication from bench to bed due to the high attrition rate especially at the late stage (non-clinical and clinical trials) of the drug development process [6, 7]. It is crucial to improve the productivity of R&D and computer-aided drug design (CADD) comes right into the scene to facilitate this endeavour. CADD makes use of different software, mathematical models and algorithms to rationalize the drug design and speed up the drug discovery process especially at the early phases which include target identification, hit identification, hit-to-lead and lead optimization [8]. With the aid of in-silico tools, the number of chemical candidates to be tested in-vitro or in-vivo are greatly reduced, the success rate of clinical trials is also increased, leading to the optimization of resources and enhanced cost-effectiveness throughout the trajectory of drug discovery and development [9]. The advantages of CADD are also evident in discovering novel drug to tackle allosteric cancer targets or management of tumours that formed through complicated pathways [10].

In general, CADD can be divided into either structure-based or ligand-based approaches. As the name implies, structure-based methods rely on the availability of three-dimensional (3D) structure of the macromolecular target. Whereas ligand-based methods require the information of at least one chemical compound of interest or a set of known actives to begin with [9]. When the drug design and discovery process involved screening a set of compounds or

chemical databases to pinpoint promising hits by utilizing computer software and algorithm, the term is coined as ‘virtual screening’. Likewise, structure-based virtual screening (SBVS) and ligand-based virtual screening (LBVS) are the common categories to differentiate between the involvement of either target molecules or known active compounds as the starting point in the screening process. Both strategies can be used independently or in a hybrid manner for virtual screening [11].

The most established structure-based CADD tool is molecular docking which predicts the interactions and degree of complementary between the ligands and the target’s binding site [12]. Through docking-based virtual screening, potential hits can be shortlisted from the vast chemical space based on the scoring ranked by the docking software [7]. The more negative the scoring value, the tighter the binding of the ligand-target is deduced [13]. Structure-based pharmacophore modelling is another technique that utilizes the information gathered from the binding sites of target structures to generate a molecular framework that outlines the essential features required for binding, followed by virtual screening to map for potential binders from the chemical database [14].

Pharmacophore is defined as the ‘ensemble of steric and electronic features that is required to ensure optimal interactions with a target of interest or to exert its biological response (either by activating or inhibiting it). Pharmacophoric descriptors are used to define a pharmacophore, including hydrogen-bonding (acceptors or donors), hydrophobic groups, electrostatic interaction sites (positively or negatively ionizable groups), ring centres (aromatic groups) and virtual points (metal coordinating areas) [15]. When the structural information of the macromolecule target is unavailable, the physicochemical properties of a set of known actives are used to perform pharmacophore modelling and mapping instead [14]. This is one example of the ligand-based approach, in which molecular descriptors known to be essential for biological activities are gathered to retrieve other potential drug candidates based on the similarity principle that indicates similar molecules normally carry similar activities [16]. Other examples of ligand-based CADD methods include similarity search, scaffold hopping and quantitative-structure-activity relationship (QSAR).

After docking-based virtual screening, molecular dynamic simulation can be used to visualise the movement and interaction of ligand-target complex over time by simulating dynamical changes in the system. By analysing the snapshots taken throughout the simulation time, flexibility and stability of the ligand-target complex can be predicted, location of water molecules or change in entropy of special structures can also be observed. These hidden states of the system can by no means be tested by any wet-lab technique [7, 17]. The molecular mechanics energies combined with the Poisson-Boltzman (MM-PBSA) or generalized Born and surface area continuum solvation (MM-GBSA) are other commonly used tool for post-docking analysis to estimate the free binding energy of the ligand-target complex [18]. Both molecular dynamic simulation and MM-GBSA/PBSA have been shown to successfully improve the results of virtual screening and are particularly useful in the lead optimization stage [18, 19].

CADD tools are also useful to predict the pharmacokinetics properties and toxicology profile of potential drug candidates, in which various types of in-silico ADMET (adsorption-distribution-metabolism-excretion and toxicity) filters are available to remove compounds that carry undesired properties either before or after virtual screening. These web-based filters are being used extensively to finetune the virtual chemical database or combinatorial libraries, as well as during the lead optimization stage to enhance the pharmacological properties of the lead compounds and subsequently increase the success rate at the downstream stages of the drug discovery process [20].

Apart from that, making use of computational tools to build a combinatorial library for virtual screening is also getting popular in drug discovery. The combinatorial library refers to a

set of new compounds prepared by a single stepwise enumeration of existing ligands using different types of substitution [21]. With today's advancement in computational power, only a few seconds are needed to construct a virtual combinatorial library with millions of compounds [22]. Combinatorial library has been used in natural product research to create databases of natural product analogues with drug-like properties [23]. This strategy can uncover the potential of natural products with privileged scaffold for new drug design and discovery, for instance, the anthraquinone- and chalcone- derivatives that showed a wide spectrum of biological effects on many different macromolecular targets responsible for human diseases including cancer were used to construct virtual library as the starting point of new drug research [24].

Natural products and their derivatives have a long history in the pharmaceutical world owing to their rich bioactive constituents with remarkable therapeutic potential and contributed to the discovery of many new chemical entities especially in the early days [25]. Medicinal plants that worked in both minor ailments and severe illnesses including cancer are cheaper and cause lesser side effects as compared to pure chemical drugs [26]. Combining natural products with chemotherapeutics has been shown to provide a synergistic effect and overcome many of the chemo-resistance hurdles in cancer treatment [27].

Anthraquinone is one of the most widely-recognized natural products with great medicinal value especially in the oncology setting as evidenced in many of the published reviews [28–33]. There have been many pieces of research characterized and studied anti-cancer properties of naturally occurring anthraquinone derivatives in different cancer types, for example, emodin in leukaemia [34], colorectal cancer [35] and breast cancer [36]; aloe-emodin in oral cancer [37], lung cancer [38] and neuroectodermal cancer [39]; chrysophanol in liver cancer [40] and many others. Well-established cytotoxic drugs like doxorubicin and mitoxantrone also contain anthraquinone moiety [41]. This basic scaffold of anthraquinone is made up of two aromatic rings that are connected by two carbonyl groups (mostly at positions 9 and 10) to form a planar structure (S1 Fig). This unique scaffold has attracted intense interest in the research endeavours towards designing anthraquinone-derived medicines for many different medical conditions including malignancy. Typical molecular targets of anthraquinone derivatives in cancer therapeutics include enzymes that are involved in cellular signalling pathways such as topoisomerase and kinases, DNA intercalation, inflammatory processes and redox reactions [32]. Having a planar aromatic structure enables anthraquinone derivatives to intercalate with DNA or interact with DNA topoisomerases, leading to an interruption of the transcription and replication process. Likewise, interaction with kinases can also disrupt cellular differentiation and repair function, resulting in apoptosis of cancer cells [42]. On the other hand, the quinone moiety in anthraquinone exhibits antioxidant properties by acting as reactive oxygen species (ROS) regulators [43]. Although there have been a considerable amount of reviews done on the topic of the anticancer potential of anthraquinone, and there was also a recent publication discussed the past, present and future role of computer-aided drug discovery in cancer research [44], however none of these reviews linked together the role of virtual screening in facilitating the drug design and discovery in cancer treatment based on anthraquinone scaffold. More importantly, none of them are systematic reviews. To the best of our knowledge, this is the first systematic review synthesizing the evidence of computer-aided drug design and discovery based on an anthraquinone scaffold for the treatment of cancer.

The primary objective of this review was to systematically assess and synthesize evidence on the utilization of computer-aided drug design (CADD) techniques centered on the anthraquinone scaffold for cancer treatment. We highlight recent trends, popular computational methodologies, preferred software tools, and databases in discovering anticancer drugs that involved anthraquinone scaffolds. Additionally, the therapeutic potential of these

anthraquinone derivatives across various cancer types is summarized together with the identified target proteins that tackled multiple malignancy pathways. Our findings intend to provide a robust foundation for future research, offering evidence-based insights for successful drug design leveraging the CADD techniques and anthraquinone scaffold.

## Materials and methods

### Study protocol

This systematic review was conducted and reported based on the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) 2020 guideline [45, 46]. The protocol was published [47] and registered in the “International prospective register of systematic reviews” database (PROSPERO: CRD42023432904).

### Review question

The review was conducted to answer the main research questions as follows:

“What are the trends and types of computer-aided drug design and discovery tools used in virtual screening based on anthraquinone scaffold for cancer treatment”?

“What are the therapeutic potential and target protein of anthraquinone and derivatives elucidated by CADD to treat cancer?”

### Eligibility criteria

The review question and eligibility criteria were established according to the PECo strategy (P, problem; E, exposure; Co, context) for systematic review. Only original research studies that were published in English and utilized CADD tools as the primary method to discover or design anticancer drugs involving compounds with anthraquinone scaffold were included in the review. The details of the inclusion and exclusion criteria based on the PECo strategy are outlined in [Table 1](#).

**Table 1. Eligibility criteria based on the PECo strategy.**

Element	Inclusion Criteria	Exclusion Criteria
Problem (P)	Studies with clear descriptions of the CADD tools used in virtual screening involving compounds with anthraquinone scaffold were included.	Studies without details or clear descriptions of the CADD or virtual screening tools and studies not involving compounds with anthraquinone scaffold were excluded.
Exposure (E)	Studies investigating therapeutic potential and target protein of compounds involving anthraquinone scaffold for cancer treatment were included.	Studies investigating diseases other than cancer and not involving anthraquinone derivatives were excluded.
Context (Co)	Only original research studies published in English and utilized CADD techniques or virtual screening tools for either target protein prediction/ validation, hit identification, hit-to-lead and lead optimization were included.	Studies exclusively in-vitro, in-vivo or other types of in-silico tools that did not serve the purpose of target protein prediction/validation, hit identification, hit-to-lead or lead optimization were excluded. Network pharmacology was excluded as it did not align with the scope of this review. Review article, book chapter, letters, grey literatures (conference paper abstracts, theses/ dissertation, report) and articles published in languages other than English were excluded.

<https://doi.org/10.1371/journal.pone.0301396.t001>

## Information sources and search strategy

On 30 June 2023, the literature search was executed on four electronic databases which include PubMed, Scopus, Web of Science and MedRxiv. There was no restriction on the publication period, but only articles published in the English language were saved. The search strategy was designed based on the combination of three main concepts, namely CADD or virtual screening (Concept 1), anthraquinone (Concept 2) and cancer (Concept 3). The aim was to retrieve studies that used CADD tools such as molecular docking, molecular dynamic simulation or any other virtual screening method as the primary approach in the quest for anticancer drugs involving anthraquinone scaffold. The main search string was as follows: (“virtual screening” OR “computer aided drug design” OR “molecular docking” OR “molecular dynamics”) AND (“anthraquinone” OR “anthracenedione” OR “anthranoid” OR “anthradione” OR “dioxaanthracene” OR “anthracene-9,10-dione” OR “anthracene-9,10-quinone” OR “9,10-anthrachinon” OR “9,10-dihydro-9,10-dioxaanthracene”) AND (cancer OR tumour OR malignant OR neoplasm). The search was focused on the title, abstract and keywords of the articles and adjustments were made in each database based on their different characteristics. The [S1 Table](#) tabulates the search strategy executed in the PubMed database.

## Study selection

Results of the literature search from the databases were exported into the reference manager, Endnote X9.0 where duplicate publications were removed by following the steps as described by Bramer et al. [48]. After deduplication, two reviewers (HMC & MS) independently screened the title and abstract of the records to ascertain their relevance to the review questions. Afterward, the selected full-text articles were retrieved and read in detail by the same two reviewers according to the inclusion and exclusion criteria. The reasons for exclusion were recorded. Disagreements between the two reviewers were resolved through discussion with a third author (LM). The study selection process was recorded in the PRISMA flow diagram (refer to [Fig 1](#) under Result and Discussion).

## Risk of bias assessment

Due to the lack of a standardised tool for this type of study, the risk of bias of the selected papers was assessed using a checklist previously developed and applied by Taldaev and colleagues [49], with some modifications. The assessment was carried out separately by two independent reviewers (HMC & MS). Disagreement was resolved by discussion with another reviewer (LM). This tool was mainly focused on the reporting quality of the molecular docking study. The original checklist consists of 7 main bias domains and 12 sub-domains.

For this review, the bias domain of “Docking Software” was removed. The authors who developed this tool ranked Glide or GOLD docking software as “low risk of bias” [49] since the Monte-Carlo algorithm and empirical scoring functions used by these two software have been shown to perform better as confirmed by in-vitro validation [50]. However, this was not always the case. Cheng et al. concluded that ‘none of the scoring functions works best at all time’ after assessing the “docking power”, “ranking power” and “scoring power” of 16 scoring functions implemented in popular commercial and academic software [51]. Another study compared five commercial and five academic docking programs in which GOLD and LeDock (both are commercial software) had the best sampling power whereas the academic software, AutoDock Vina was superior in terms of scoring power, implying commercial programs did not outperform academic software as expected [52]. More recently, Reddy et al. demonstrated that Glide gave consistent results in terms of docking conformation, ranking and scoring accuracy, but AutoDock ranked as the best scoring accuracy among all other tested software [53]. Therefore

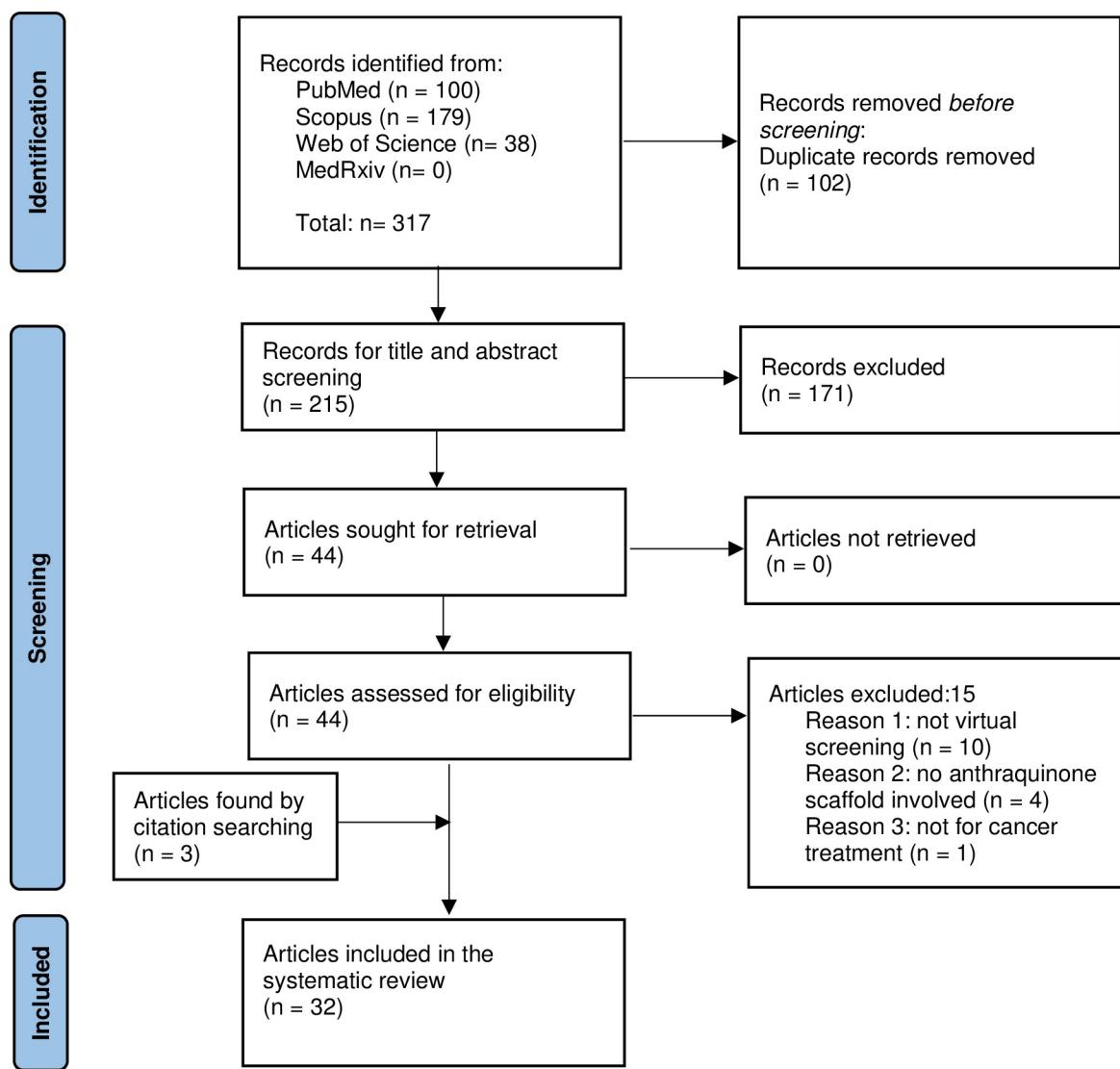


Fig 1. PRISMA flow diagram.

<https://doi.org/10.1371/journal.pone.0301396.g001>

there is no single docking method that gives the best outcome for all docking jobs and the quality of docking result is greatly influenced by the ligands and target of interest [54].

On the other hand, the sub-domains of “Ligand optimization” were merged under “Ligand Preparation”. Likewise, “Target Optimization” was modified to “Target Preparation” and all relevant sub-domains were merged. As long as the ligands and target structures were prepared by using a special tool before undergoing docking calculation, the studies were ranked as “Low risk of bias”. This made up a total of 5 main bias domains and 9 sub-domains (Table 2).

## Data extraction

Two independent reviewers extracted data from the eligible studies using a predefined data collection form (HMS & SM). The data extracted included the title of journal, authors, publication year, study context, CADD methods and the software/tools used, name and structure of the starting compounds or identified hits that contained anthraquinone scaffold, types of

**Table 2. Risk of bias assessment.**

Bias Domain	Sub-domain	Low Risk of Bias	High Risk of Bias	Unclear
Ligand selection	Ligand filtering	Performed	Not performed	No data
Ligand Preparation	Geometry-optimized and generation of energetically possible conformation	Performed by special tool	Not performed or performed without special tool	No data
Target Selection	Resolution of target structure	Not more than 2.5 Å	More than 2.5 Å	No data
	Method of obtaining target structure	NMR spectroscopy	X-ray crystallography or cryogenic electron microscopy	No data
Target Preparation	Protonation, addition of missing residue and side chain after X-ray crystallography or cryogenic microscopy	Performed with special tool	Not performed or performed without special tool	No data
	Control of histidine and addition of metals	Performed	Target structure did not reference to biological condition	No data
Results assessment	Visual control	Performed	Not performed or structure defects observed	No data
	Redocking/ Docking Validation	Performed	Not performed or the RMSD value is higher than 2 Å as compared to the initial structure	No data
	Verification of docking result by in-vitro study	Binding constant (eg: Ki, IC <sub>50</sub> ) was determined	No laboratory validation or no quantitative calculations	No data

<https://doi.org/10.1371/journal.pone.0301396.t002>

cancer involved, databases used to retrieve the structures of both target and ligand. (Refer [S2 Table](#). Data collection form).

## Data synthesis and analysis

All included papers in the final studies were used for data synthesis and analysis. The data were summarized via a narrative approach to address our review questions. Tables and figures were used to present the characteristics of the studies. The trends and types of different CADD approaches and tools used in virtual screening as the primary tools for designing or discovering anthraquinone analogues for cancer treatment were analysed and discussed. The respective macromolecular targets involved and their role in managing different types of cancer, together with the chemical structures of the identified hit compounds that contained anthraquinone scaffold were presented and reviewed.

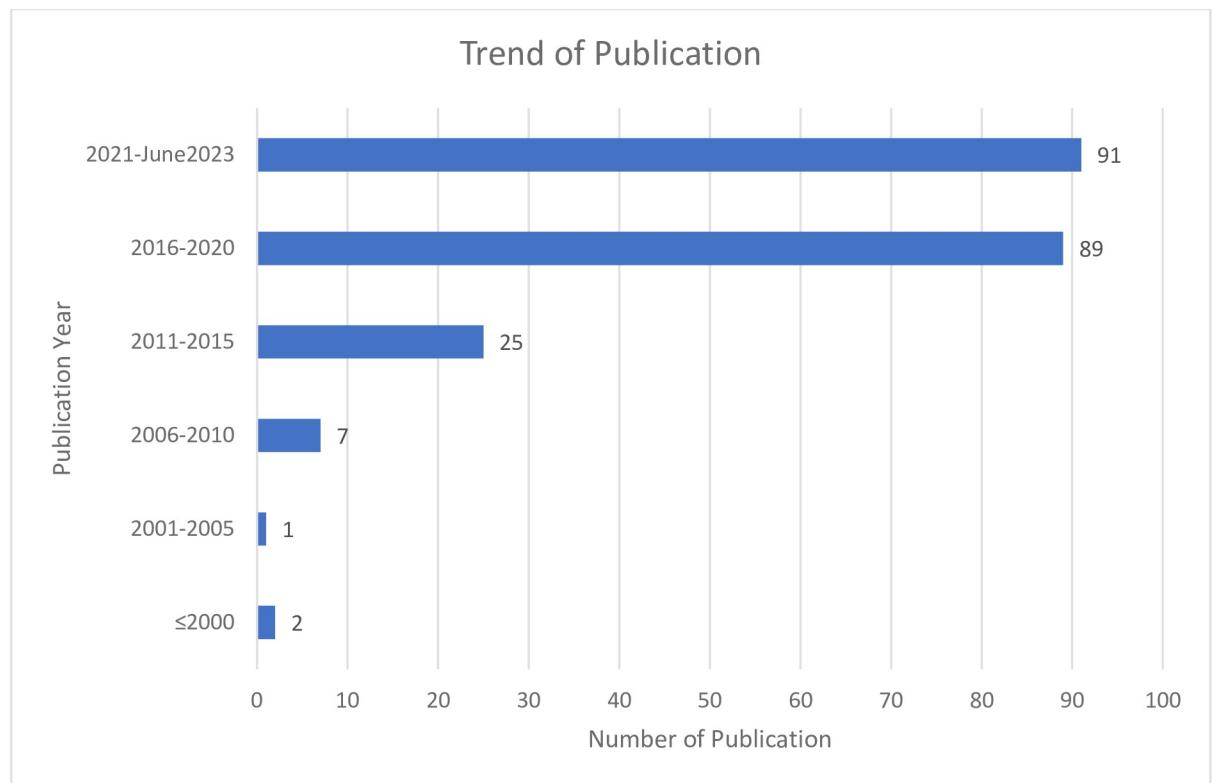
## Results and discussion

### Study selection

The literature search identified a total of 317 records. After importing the records into the reference manager (Endnote X9.0) where 102 records were deduplicated, 215 records were subjected to title and abstract screening by HMC and SM. The exclusion of 171 records that did not fulfil the inclusion criteria resulted in 44 records sought for full article retrieval. These articles were read in full and critically examined by the same two reviewers separately based on the eligibility criteria. The reasons for exclusion after full-text screening were recorded. Any disagreement was resolved by a third author (LM) through discussion and consensus. In addition, 3 more studies that fulfilled the inclusion criteria were found by citation searching, alongside the full screening process as they were somehow related to the included studies. The results of the study identification and selection process were recorded in the PRISMA flow diagram ([Fig 1](#)). A total of 32 articles were included in the final review.

### General view on the trend of publications

The number of publications in cancer research that involved CADD tools and anthraquinone was observed to be on the rising trend. From the total number of 215 articles retrieved from



**Fig 2. Trend of publications.**

<https://doi.org/10.1371/journal.pone.0301396.g002>

the search strategy after deduplication, only 2 articles were published before the twenty-first century. This number increased steadily after entering year 2006 (Fig 2). In fact, more than 90% of the identified articles were published in the last decade. From year 2021 up to June 2023 alone (less than 3 years), the highest number of publications (more than 40%) tagged with the searched keywords were recorded. This showed that the utilization of CADD methods has gained more popularity and anthraquinone is a compound of interest which attracted immense interest for biomedical research, especially in the oncology setting. One of the possible reasons may be also due to the worldwide lockdown caused by the Covid-19 pandemic in the past 3 years has switched many of the research focus from wet-lab to dry-lab (in-silico), hence more studies were performed virtually leading to more research publications on computer modelling. As virtual screening has been proven to offer an advantage in terms of resources, time and cost reduction in venturing novel drugs, it is expected that the number of research and publications in this field will continue to rise for the year 2023 and beyond.

### Types of CADD approaches used and the study context

Structure-based methods and ligand-based methods were used either independently or in combination. Table 3 illustrates the trends and different types of structure-based and ligand-based CADD methods used in the 32 included studies, together with their study context. Nearly one-third of the studies combined both structure-based and ligand-based methods in their drug design and discovery project (10 studies) while the rest used mainly the structure-based methods.

None of the studies used ligand-based method alone since the crystal structure of all identified targets in the studies were either available in the protein databases [55, 56, 58–60, 62–67, 69–75, 77–86] or successfully created by homology modelling [57, 61, 68, 76].

Table 3. Types of CADD methods used and the study context.

Authors (Year)	Structure-based methods	Ligand-based methods	Study Context	Reference
Ahire et al. (2016)	Docking	-	SBVS for hit identification	[55]
Arba et al. (2017)	Docking	-	SBVS for hit identification	[56]
Asnawi et al. (2022)	Docking	-	SBVS for hit identification	[57]
Bai et al. (2012)	Docking	Similarity search	SBVS + LBVS for hit identification	[58]
Ciaco et al. (2023)	Docking	Similarity search	LBVS + SBVS for hit identification	[59]
Choowong-komon et al. (2010)	Docking	-	SBVS for hit identification	[60]
Cozza et al. (2008)	Docking	-	SBVS for hit identification	[61]
Cozza et al. (2009)	Docking	-	SBVS for hit identification	[62]
Cozza et al. (2015)	Docking	-	SBVS for hit identification (re-evaluation)	[63]
Das et al. (2023)	Docking	-	SBVS for hit identification	[64]
DemiRezer et al. (2018)	Docking	-	SBVS for hit identification	[65]
Dong et al. (2019)	Docking	-	SBVS for drug repurposing	[66]
Gao et al. (2022)	Docking	-	SBVS for target verification and hit identification	[67]
Guan et al. (2020)	Docking	-	SBVS for drug repurposing	[68]
Golubovskaya et al. (2013)	Docking	-	SBVS for hit identification	[69]
Jordheim et al. (2013)	Docking	-	SBVS for hit identification	[70]
Khan et al. (2021)	Docking	Pharmacophore mapping	SBVS + LBVS for hit identification	[71]
Lankapalli et al. (2013)	Docking	-	SBVS for target fishing & hit identification	[72]
Leggett et al. (2022)	Docking	-	SBVS for hit identification	[73]
Liu et al. (2019)	Docking	-	SBVS for hit identification	[74]
Mhatre et al. (2017)	Docking	-	SBVS for hit identification	[75]
Nag et al. (2022)	Docking	Pharmacophore mapping	SBVS + LBVS for hit identification and poly-pharmacology study	[76]
Obounchoey et al. (2019)	Docking	Similarity search	LBVS + SBVS for hit identification	[77]
Rinne et al. (2020)	Docking, structure-based pharmacophore	Similarity search	SBVS + LBVS for hit identification	[78]
Roy et al. (2021)	Docking	-	SBVS for hit identification	[79]
Singh et al. (2021)	Docking	Scaffold hopping	LB-scaffold hopping SB-guided combinatorial library building & SBVS for hit identification	[80]
Song et al. (2019)	Docking	-	SBVS for hit identification & and hit-to-lead optimization	[81]
Taherkhani et al. (2021)	Docking	-	SBVS for hit identification	[82]
Wang et al. (2021)	Docking	3D-QSAR	LB + SB methods for novel drug design	[83]
Wu et al. (2018)	Docking	-	SBVS for Drug-repurposing	[84]
Wu et al. (2022)	Docking	Similarity search	LBVS+ SBVS for hit identification	[85]
Zyaater et al. (2019)	Docking	Similarity search	LBVS + SBVS for hit identification	[86]

<https://doi.org/10.1371/journal.pone.0301396.t003>

The use of both structure-based and ligand-based methods were combined in either a parallel manner [58, 78, 80] or performed one after another in sequence, termed as sequential screening [59, 71, 76, 77, 83, 85, 86]. The sequential approach was more popular as it enabled efficient computation in which the more straightforward and quicker ligand-based virtual screening served as the pre-docking filter to reduce the size of the screening library before the more computational-demanding docking job took place. This is evident in those studies that successfully reduced the size of the screening database from 100 million of compounds in PubChem to merely 200 compounds [77] or from the Key Organic database of 49415 compounds to less than 10 compounds (85) after similarity search. Another study utilized ligand-based

pharmacophore mapping from ZINC database with 20 million entries and identified only 12 phytochemicals analogues for docking [76]. Other studies that employed ligand-based pharmacophore or similarity screening prior to docking-based screening resulted in the shrinkage of library size that ranged from 30% to 85% [59, 86, 87].

On the other hand, running both approaches in parallel could enrich the hit discovery rate of virtual screening, and overcome the fundamental limitations of each approach [58] and this was particularly helpful when the sample size of experimental-validated known actives for use as references in virtual screen was limited [78]. For instance, Bai et al. designed a virtual screening protocol that started with a parallel screening of a selected database by molecular docking and shape-similarity search, followed by second stage docking in higher precision docking mode, then visually filtered out desired compounds for further bioassay, leading to a hit rate of 24.7% [58]. Rinne and colleagues ran five separate screening simultaneously (one by similarity search, two by pharmacophore mapping and two by molecular docking), each virtual screen shortlisted a hit list of 200 compounds for laboratory validation out of the database of 140000 compounds to tackle the limitation of virtual screen [78].

Majority of the studies aimed to discover novel hit or lead from the screened compounds or database, but three studies [66, 68, 84] looked into drug repurposing or repositioning from licensed medicines to explore additional indications in the oncology setting. The strategy of investigating new role of old drug offers benefits such as cheaper investment cost, shorter development time and reduced risk of failure since lesser efforts are required especially at the lead optimization stage [88].

There were three studies extended the exploration on the hit compounds identified from previous virtual screening studies. The starting compound, UM63 that contained anthraquinone scaffold used in one of the studies [59] to pre-filter the database based on similarity in SMARTS pattern was identified previously by the research group from the same university [86]. Similarly, the screening database of 200 compounds involved in another study [77] were defined based on Tanitomo coefficient of 0.95 from the eight hit compounds discovered by their colleagues previously. There was one study re-evaluated the selectivity of quinalizarin against CK2 by molecular docking and molecular dynamic simulation, and further measured the inhibitory effect of this compound on 14 kinases panels [63] compared to 7 kinases panels back in year 2009 when this hit was first shortlisted by the virtual screening study [62].

## Molecular docking

Molecular docking was the most popular structure-based method used in CADD. Notably, all of the 32 studies (100%) included in this review employed docking in their research (Table 3). This popular method is favoured due to its computational efficiency, in which the docking of one ligand on a single core typically requires only a few minutes of computing time [16]. Thus, docking-based virtual screening serves the purpose of filtering out good ideas from bad ones and subsequently eases the prioritization of more promising ligands out of the large size of the virtual library to be taken forward for subsequent studies.

Docking was also used in molecular modelling to investigate the interaction in between the shortlisted hits and the target protein in term of hydrogen bond, hydrophobicity, electron distribution and binding energy [55, 56, 63, 65, 75, 79, 82]. This was crucial to identify the structural determinants responsible for efficient binding with the amino acids or protein residues of the target. With this essential information in hands, promising drugs with more desired properties were designed [80, 81, 83].

One of the studies utilized molecular docking to verify the potential drug targets identified from the competitive endogenous RNA (ceRNA) network study by virtually screening out

small molecule binders against these targets from the selected database [67]. On the other hand, molecular docking was also used in target fishing and profiling, in which the ligands were docked into a variety of macromolecules to predict appropriate targets for further analysis [72]. When one particular compound regulates multiple targets at the same time in the same illness, it is referred as “Polypharmacology” and again, molecular docking was useful to provide insights in this case [76].

**Docking software.** Docking methods consist of both posing and scoring process. Small molecules are docked into the macromolecular target to generate different ligand-target conformation and the degree of complementarity for each binding was ranked by scoring function. There were thirteen different docking tools used in the included studies in this review (Table 4). These were either commercial software with subscription fees or freely available (mainly for academic researchers). Each docking software has a different searching algorithm to align the binding geometries of ligand-target complex to the preferred and stable pose. Different types of scoring functions are employed in these docking programs to evaluate the best binding pose of each bounded complex with minimum energy and rank the ligands accordingly from the most negative value to less negative value of the docking score.

The most popular commercial docking software used in this review was Glide provided by Schrodinger, followed by GOLD provided by the Cambridge Crystallographic Data Centre (CCDC). For the academic software, AutoDock was among the most sought-after freeware that used by seven of the included studies. This freeware was developed by the Scripps Research Institute and they also have another newer generation of docking software namely AutoDock Vina. It was used by a total of five studies, where two of them performed the docking study using AutoDock Vina via the commercial virtual screening platform, PyRx (Table 4).

The popular docking programs normally have user-friendly interfaces with readily available integrated tools required for smooth handling of the in-silico workflow. For example, structure-rendering, visualization, target preparation (via either interactive or automated protein preparation workflow), ligand preparation (via LigPrep) could be all managed by the Maestro graphical interface developed by Schrodinger before submitting the Glide docking calculation job [58, 61, 62, 64, 68, 78, 80, 83].

### Consensus scoring

Some of the studies used two or more scoring functions to evaluate the best hit. This technique is known as ‘consensus scoring’ whereby different types of scoring functions are combined with the hope to compensate deficiencies of each scoring function and to improve capability of the screening process in discriminating actives from decoys [60–62, 66, 75]. The docking protocols of two studies were made up of four different algorithms (MOE-Dock, Glide, FlexX and Gold) and five different scoring functions (MOE-Score, Glide-Score, Gold-Score, Chem-Score and X-score), each was performed independently. The top-ranked compounds taken forward for further analysis were prioritized from the ‘consensus scored list’ generated from combining these docking programs [61, 62]. Another study used an empirical score genetic algorithm from the GOLD program and force-field genetic algorithm from AutoDock program (combined with FRED calculation) to rank the binding of ligands against the target protein, in which the highest scored compounds were also consensus between these two programs [60].

Two other studies utilized different precision modes of the selected docking software prior to re-docking by second docking software with varied scoring functions, aimed to improve the efficiency and accuracy of the virtual screening. One of them performed the initial virtual screening by Surflex-Dock in screen mode, followed by redocking of shortlisted hits with Surflex-Dock GeomX mode which featured higher spin density and accuracy. Meanwhile, the

Table 4. Summary of different docking tools used in the studies.

Docking Tool	Features	Website	Studies
Glide (Grid-Based Ligand Docking with Energetics)	Commercial software. Complete systematic search of the orientation, conformational and positional space of the ligand in the target with the OPLS-AA force field (Optimized Potentials for Liquid Simulations). Available in HTVS mode, SP mode, XP mode and induced fit docking.	<a href="https://www.schrodinger.com/products/glide">https://www.schrodinger.com/products/glide</a>	[58, 61, 62, 64, 68, 78, 80, 83]
GOLD (Genetic Optimisation for Ligand Docking)	Commercial software. Uses empirical score genetic algorithm for exploration of ligand flexibility. Examples of scoring function are GoldScore, ChemScore, Kinase Scoring Function (KCS).	<a href="http://www.ccdc.cam.ac.uk/">http://www.ccdc.cam.ac.uk/</a>	[60–63, 70, 77, 85]
AutoDock	Freeware. Uses Lamarckian genetic algorithm (LGA) posing where the conformations changes of ligands after optimization are used as subsequent poses for the offspring. The force-field-based scoring function considers the intermolecular interaction energy, the sum of torsional free energy, total internal energy, and unbound system energy.	<a href="https://autodock.scripps.edu/">https://autodock.scripps.edu/</a>	[56, 57, 60, 66, 79, 81, 82]
AutoDock Vina	Freeware. A newer generation of AutoDock4. Uses knowledge-based scoring function and rapid gradient-optimization conformational search with Monte Carlo sampling technique.	<a href="https://vina.scripps.edu/">https://vina.scripps.edu/</a>	[67, 71, 75]
Surflex-Dock	Commercial software. A docking module in SYBYL software. Uses empirical scoring function by taking hydrophobic, polar, repulsive, entropic and solvated effects into consideration. The search engine is based on molecular similarity to dock ligands to target. Available in Normal, Screen, Geom and GenomX mode.	<a href="https://www.computabio.com/applications-of-surflex-dock-software.html">https://www.computabio.com/applications-of-surflex-dock-software.html</a>	[65, 66, 68, 73]
FlexX	Commercial software. An incremental fragment-based docking algorithm where the conformational space sampling is done using a tree search method. Provided by BioSolveIT in the LeadIT package.	<a href="https://www.biosolveit.de/products/#FlexX">https://www.biosolveit.de/products/#FlexX</a>	[55, 61, 62]
FRED	Free academic licensing program available. Uses exhaustive search algorithm that systematically searches conformers of each ligand within the active site at a specified resolution. Examples of scoring function are Chemgauss and Chemscore. Integrated in OpenEye Scientific Software.	<a href="https://www.eyesopen.com">https://www.eyesopen.com</a>	[59, 86]
MOE-DOCK (Molecular Operating Environment)	Commercial software. A docking software under MOE Suite, an Integrated Computer-Aided Molecular Design Platform.	<a href="http://www.chemcomp.com">http://www.chemcomp.com</a>	[61, 62]
UCSF DOCK	Free for non-commercial researchers only. Force-field based scoring.	<a href="https://dock.compbio.ucsf.edu/">https://dock.compbio.ucsf.edu/</a>	[69]
PatchDock	Free for non-commercial researchers only. Surface path matching and molecular shape complementary algorithms followed by filtering and scoring.	<a href="http://bioinfo3d.cs.tau.ac.il/PatchDock/patchdock.html">http://bioinfo3d.cs.tau.ac.il/PatchDock/patchdock.html</a>	[72]
iGEMDOCK (Generic Evolutionary Method for molecular DOCKing)	Free for non-commercial researchers only. An integrated graphical environment which utilizing post-screening analysis with pharmacological interactions for virtual screening.	<a href="http://gemdock.life.nctu.edu.tw/dock/igemdock.php">http://gemdock.life.nctu.edu.tw/dock/igemdock.php</a>	[75]
DockThor	Freeware. Uses MMFFLigand and PdbThorBox in-house tool for the docking algorithm along with MMFF94S53 force field.	<a href="https://www.dockthor.lncc.br/">https://www.dockthor.lncc.br/</a>	[76]
Internal Coordinate Mechanics (ICM)-Pro	Commercial software. The molecular system was described by using internal coordinates as variables. Energy calculations were based on the ECEPP/3 force field with a distance-dependent dielectric constant. The biased probability Monte Carlo (BPMC) minimization procedure was used for global energy optimization.	<a href="https://www.molsoft.com/technology.html">https://www.molsoft.com/technology.html</a>	[84]
PyRx virtual screen software	Commercial software. A virtual screening software that is using a variety of established open-source software including AutoDock4, AutoDock Vine, AutoDock Tools, Python, Visualization Toolkit, Open Babel etc.	<a href="https://pyrx.sourceforge.io/">https://pyrx.sourceforge.io/</a>	[74, 89]

<https://doi.org/10.1371/journal.pone.0301396.t004>

redocking was also conducted by AutoDock program that used a different scoring algorithm in parallel [66]. The other study combined Glide [Standard Precision (SP)], Extra Precision (XP), induced-fit docking and Surflex-Dock (Screen Mode, SurflexDock Geom and Surflex-Dock GeomX), in which each program was run by three different precision modes hierarchically [68]. Mhatre et al. also combined two different docking programs, namely iGEMDOCK and AutoDock Vina in their study but the operations were divided into two phases, iGEMDOCK was used in the first phase docking followed by redocking using AutoDock Vina in the second phase. The docking scores for each program were presented in a

separate table and comparison was performed to analyse the preferred binding modes of selected ligands against the target. The results revealed that the pharmacophoric and molecular space acquired by the selected ligands were similar to the known active, implying the therapeutic potential of the phytochemicals of interest [75].

Apart from combining different scoring functions of various docking programs, the concept of consensus scoring was also applied in virtual screening that combined structure-based and ligand-based methods. One of the recent studies integrated docking-based virtual screening with three other ligand-based screening tools (pharmacophore, shape similarity and QSAR) in which each method produced separate hit lists and consensus Z-score for each highest-ranked ligand was then calculated [90]. This paper was published after the article searching period of this review ended therefore it is not included in the final review. However, the proposed strategy is worth further exploration.

### Pharmacophore modelling and mapping

In ligand-based pharmacophore screening, the 3D structures of a set of known actives were retrieved to guide the development of the pharmacophore model. This model served as the tool for subsequent virtual screening to map a predefined database for best-fit compounds that presented shared common features responsible for binding and biological functions [71, 76]. LigandScout was one of the commercial computer software used to generate the pharmacophore model and to score the compounds from the large database based on the computed pharmacophore features so that potential hits could be identified from the top-ranked list of the pharmacophore-fit score [71]. There was also a free web server, ZINCPharmer (<http://zincpharmer.csb.pitt.edu/pharmer.html>) equipped with 'add feature' function to predict pharmacophore features from the uploaded ligands candidates. ZINC database was then screened virtually by ZINCPharmer to pinpoint compounds that demonstrated the highest complementary to the model in terms of chemical descriptors such as hydrogen bonds, ring groups, ionic groups, hydrophobic and lipophilic groups [76].

There was also structure-based pharmacophore screening performed by Rinne and colleagues. Discovery Studio was used to develop the pharmacophore model. Maestro and Pymol from Schrodinger were used to aid the manual selection of pharmacophore features based on key residues at two binding cavities of the 3D target structure obtained from X-ray crystallography. This resulted in two separate hits lists from the pharmacophore-based screening [78].

### Similarity searching

Ligand-based similarity searching utilized the two-dimensional (2D) or three-dimensional (3D) descriptors of the known actives to discover most-alike molecules from the screening library based on the concept of 'compounds with similar chemical structures tend to exhibit similar biological activities' [58, 59, 77, 78, 85, 86]. Examples of 2D descriptors included molecular fingerprints [77, 78], substructure-based descriptors [85] or SMART patterns of the molecules [59, 86]. SMART refers to 'SMILES arbitrary target specification' whereby the full name of SMILES is 'Simplified Molecular Input Line Entry System'. These are special languages developed by David Weininger and colleagues to describe structure or substructure patterns in molecules for cheminformatic purposes. FILTER application implemented in OMEGA from OpenEye Software was used to reduce the size of the screening database by performing SMART-based query prior to docking-based virtual screening [59, 86]. ChemBioFinder was another tool used to calculate the substructure search based on known active and to screen out potential candidates from the database with similar privileged motif as the query compound

for further analysis [85]. Tanimoto similarity coefficient was a common metric used to rank the magnitude of similarity especially for molecular fingerprint mapping [77, 78].

Apart from structural information, 3D descriptors also take structural alignment into account for predicting the similarity between two compounds. Pharmacophore modelling and shape similarity are among the popular 3D methods for ligand-based virtual screening [91]. Bai and colleagues analysed the 3D conformers of co-crystallized ligands for similarity search using SHAFTS program and aided by the in-house conformational generation tool Cyndi [58]. SHAFTS (SHApeFeaTure Similarity) was developed to merge the pharmacophore overlay and shape complementary approach in discovering drug candidates with desired properties. Hybrid similarities score was calculated to prioritize the best-matched compounds with query compounds in terms of molecular pose alignment and volumetric superposition [92]. Obviously, 2D similarity approach was more popular as it was simpler, quicker and more straightforward. Nevertheless, both 2D and 3D similarity approaches were proven to increase the efficiency of virtual screening especially by downsizing the large library to become more ‘target-focused library’ for subsequent in-silico research.

## QSAR

QSAR involves mathematics calculation and statistics to model the correlation in between the molecular descriptors and biological activities [93]. A group of active compounds with the corresponding binding constant or inhibitory concentration determined by in-vitro studies was gathered as the starting point of the process flow. The active compounds were randomly divided into training sets and test sets in a predefined ratio. SYBYL software was used for molecular alignment based on the most potent compound to fix a common substructure. The popular CoMFA (Comparative Molecular Field Analysis) and CoMSIA (Comparative Molecular Similarity Indices Analysis) methods were then used to construct the 3D-QSAR model [83]. CoMFA is a force field based method that involves linear function but CoMSIA uses an exponential function to compute ligand properties such as steric and electrostatic energies [94]. The 3D-QSAR model was useful to guide lead optimization and new drug design by modifying the structure of existing ligands [83]. Both QSAR model and pharmacophore model are considered essential features of the known actives but the focus of QSAR model was more on features that correlated closely to the biological effect.

## Scaffold hopping

Scaffold hopping can be categorized as one of the ligand-based virtual screening methods as it typically requires the core structures of known actives to be used as the template [95, 96], at the same time it is also one of the main aims of many drug discovery projects to identify novel chemotype [91, 92]. Singh et al. modified the base scaffold of anthrafuran that was predetermined to demonstrate desired biological properties, guided by the binding cavity of the target of interest to build a combinatorial library of over 2 million new compounds based on in-silico enumeration. This target-focused virtual library was then subjected to structure-based virtual screening in which novel compounds with better biological activities than the parent analogue were discovered [80].

## Drug-likeness, lead-likeness and ADMET filter

ADMET properties are important for the ultimate fate of a possible drug candidate. Unwanted effects in animal models or even human trials can be reduced by filtering drug candidates by their drug-likeness, lead-likeness and ADMET properties in early stages [55]. These filters can

be applied on the selected database prior to virtual screening [61, 69, 70, 76, 80] or afterward on the shortlisted hits [59, 64, 71, 75, 78, 81, 82], or even before and after [71].

The popular Lipinski's Rule of Five [97] was applied to filter out compounds that disobeyed the drug-likeness properties as predicted (more than 5 hydrogen donors, more than 10 hydrogen acceptors, molecular weight larger than 500 and CLogP, a measurement of lipophilicity greater than 5) [55, 69–71, 75, 76, 80–82]. The freely accessible SwissADME server (<http://www.swissadme.ch/>) was another useful tool to predict properties like bioavailability, lipophilicity, pKa, blood-brain barrier permeabilities etc [64, 76]. The toxicity prediction technique included TOPKAT (Toxicity Prediction by Komputer-Assisted Technology) that applied QSTR (Quantitative Structure-Toxicity Relationship) models to deduce toxicity profiles such as carcinogenicity, mutagenicity, skin irritation was executed via Discovery Studio software package by Accelrys [55]. PAINS (Pan-assay Interference Compounds) filter was also important to remove compounds with unwanted functional groups that might cause unexpected interactions with multiple targets leading to false positive results [59, 78].

Obviously, the studies involved drug repurposing could just skip the filtration step since the licensed small molecules are believed to possess desired physicochemical properties with acceptable bioavailability and safety aspects for oral consumption, as approved by the regulatory authority [66, 68, 84].

### Molecular dynamic simulation

The top-ranked protein-ligand complex shortlisted from docking studies were subjected to molecular dynamic simulations to observe how every atom in protein moved over time to assess the binding mode in depth, to confirm the stability of the docked pose and to gain insight on the protein flexibility of the ligand-protein complex [55–57, 63, 66, 76, 79, 80, 82, 83].

The most widely used software package to perform the simulation was AMBER [56, 66, 79, 83]. Throughout the years, different versions of the software package have been introduced, for instance AMBER12 [56], AMBER14 [66] and AMBER16 [79, 83]. Other molecular dynamic tools used were Discovery Studio Molecular Dynamic Protocol [55, 82], GROMAC [57], NAMD [63], DESMOND [80] and CABS-flex 2.0 server [76]. Among these tools, only DESMOND from Schrodinger required a license subscription fee whereas the rest were readily downloadable from the public domain for academic research purposes (Table 5).

All of the studies probed the dynamic of the protein-ligand complex in the time-scale of nanoseconds [55–57, 63, 66, 76, 79, 80, 82, 83]. Ideally, performing simulation in the order of microseconds or even milliseconds is better to reveal the biologically essential conformational evolution but unfortunately, this time-scale of simulation takes too long time to run and therefore rarely in use [54]. The stability of the ligand-protein complex was monitored by RMSD (Root Mean Square Deviation), a common metric used in evaluating the dynamic of

**Table 5. Molecular dynamic simulation software used.**

Molecular dynamic simulation software	Brief description	Website	Used by
AMBER	Freeware	<a href="https://ambermd.org/">https://ambermd.org/</a>	[56, 66, 79, 83]
Discovery Studio Molecular Dynamic Protocol	Freeware	<a href="https://discover.3ds.com">https://discover.3ds.com</a>	[55, 82]
GROMAC	Freeware	<a href="https://www.gromacs.org">https://www.gromacs.org</a>	[57]
NAMD	Freeware	<a href="https://www.ks.uiuc.edu/Research/namd/">https://www.ks.uiuc.edu/Research/namd/</a>	[63]
CABS-flex 2.0 server	Freeware	<a href="http://biocomp.chem.uw.edu.pl/CABSflex2">http://biocomp.chem.uw.edu.pl/CABSflex2</a>	[76]
DESMOND	Commercial software	<a href="https://www.schrodinger.com">https://www.schrodinger.com</a>	[80]

<https://doi.org/10.1371/journal.pone.0301396.t005>

macromolecules [55–57, 63, 66, 76, 79, 80, 82, 83]. The docked pose was considered stable if the docked ligand deviated from the macromolecules with an RMSD of less than 2 Angstrom over the simulated time range [57]. Some of the studies presented RMSF (Root Mean Square Fluctuation) to assess the protein flexibility of the ligand-protein complex [56, 76, 79].

### MM-GBSA/ MM-PBSA

There were seven studies predicted the binding free energy of the ligand-target complex using the molecular mechanics-Poisson-Boltzman solvent accessible surface area (MM-PBSA) [56, 57] or molecular mechanics-generalized Born surface area (MM-GBSA) methods [64, 66, 79, 80, 83]. These methods describe the binding free energy of a complex as the subtraction of the unbound macromolecule's and unbound ligand's free energy from the bounded macromolecule's free energy [56]. Binding free energies of a complex also known as molecular mechanics potential energy, which is composed of the energy of both bonded (bond energy, angle energy, torsion energy) and non-bonded interactions (electrostatic energy, van der Waals energy) [56, 57].

A variety of different software packages are used for this calculation including the Python modules and NMODE module in AMBER [56, 66, 79, 83], PRIME [64, 80] and GROMAC [57]. All of the studies calculated the free energy of each system based on the snapshots taken from the molecular dynamic simulation trajectories except one study that did not perform the simulation beforehand [64].

### Source of crystal structure for macromolecular target

The availability of 3D structures of macromolecular targets is the prerequisite for performing structure-based virtual screening. More than 80% of the macromolecular structures used in the studies were retrieved from RCSB PDB (Research Collaboratory for Structural Bioinformatics Protein Data Bank). The crystal structures of the protein were either determined by either X-ray crystallography, NMR spectroscopy or cryogenic microscopy method, and deposited in the Protein Data Bank for free download in 3D format (<https://www.rcsb.org/>) [55, 56, 58–60, 62–66, 69–75, 77–86]. There was only one study obtained the 3D structure of the protein from AlphaFold, an artificial-intelligence based database that also provides free downloads of protein structures (<https://alphafold.com>) [67].

There were four studies that used homology modelling to construct the protein structures since those 3D structures were not readily available in the existing database [57, 61, 68, 76]. To start with, template searching was performed by SWISSMODEL, the well-known protein modelling online server (<https://swissmodel.expasy.org/>). Next, the amino sequence of the templates was downloaded from NCBI protein resources (<https://www.ncbi.nlm.nih.gov/protein>) followed by alignment of the target sequence and template structure, and finally the models were built by SWISSMODEL [57, 76]. Another modelling tool available for constructing homology modelling was the Molecular Operating Environment (MOE) [61, 68].

Validation of the model was also performed to ensure the quality of the modeled protein. Various online tools were available to assess and validate the model, namely WHATCHECK, PROCHECK, ERRAT, Verify3D, PROVE, OMEAN and Ramachandran plot [57, 61, 68, 76].

### Ligand databases

The databases used in virtual screening were either a public database, corporate collection, vendor database, a set of potentially synthesizable compounds or a combinatorial library built by in-silico enumeration of pre-defined core structures. Table 6 briefly describes various online ligand databases used in the included studies.

Table 6. Online ligand databases used.

Database	Brief description	Website	Used by
PubChem	Freely accessible database contains data on chemical structures, identifiers, chemical and physical properties, biological activities, toxicity data and many others. Launched in year 2004 by National Institute of Health.	<a href="https://pubchem.ncbi.nlm.nih.gov/">https://pubchem.ncbi.nlm.nih.gov/</a>	[55, 57, 71, 72, 75–77, 82]
ZINC	Freely accessible database contains purchasable compounds in ready-to-dock format. Provided by the Irwin and Shoichet Laboratories in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF).	<a href="https://zinc.docking.org/">https://zinc.docking.org/</a>	[70, 71, 73, 76, 81, 84]
DrugBank	Freely accessible database contains FDA approved small molecule drugs, FDA approved biotech drugs, withdrawn drugs, and experimental drugs. Started in 2006 by Dr David Wishart's laboratory at the University of Alberta, Canada.	<a href="https://go.drugbank.com">https://go.drugbank.com</a>	[64, 66]
NCI/DTP database	Freely accessible database developed by National Cancer Institute (NCI), Developmental Therapeutic Program (DTP).	<a href="https://dtp.cancer.gov/">https://dtp.cancer.gov/</a>	[60, 69]
MMsINC	Freely accessible database contains around 4 million of synthetic and natural compounds. Maintained by University of Padova, Italy	<a href="http://mms.dsfarm.unipd.it/MMsINC/search">http://mms.dsfarm.unipd.it/MMsINC/search</a>	[61]
Chemical library by University of Helsinki	A library of about 240,000 compounds maintained by University of Helsinki, Finland. Includes a subset of approved drugs, natural products and diversified sets of drug-like compounds.	<a href="https://www.helsinki.fi/en/infrastructures/drug-discovery-chemical-biology-and-screening/infrastructures/high-throughput-biomedicine/chemical-compound-libraries">https://www.helsinki.fi/en/infrastructures/drug-discovery-chemical-biology-and-screening/infrastructures/high-throughput-biomedicine/chemical-compound-libraries</a>	[78]
FDA-approved library by Selleck	A unique collection of FDA-approved drugs and pharmacopoeia-included API.	<a href="https://www.selleckchem.com/screening/fda-approved-drug-library.html">https://www.selleckchem.com/screening/fda-approved-drug-library.html</a>	[68]
Maybridge	Commercial database contains ready-to-screen library, fragment library, drug-like compounds library and few others.	<a href="http://www.maybridge.com">http://www.maybridge.com</a> (old) <a href="https://www.thermofisher.com/bn/en/home/global/forms/lab-solutions/maybridge-library.html">https://www.thermofisher.com/bn/en/home/global/forms/lab-solutions/maybridge-library.html</a> (new)	[58]
SPECS database	Commercial database contains drug-like small molecules collected from academic sources worldwide.	<a href="https://www.specs.net/">https://www.specs.net/</a>	[58]
Key Organic Database	Commercial database contains fragments and screening libraries, building block libraries etc.	<a href="https://www.keyorganics.net/">https://www.keyorganics.net/</a>	[85]
Aldrich Market Select	Commercial database by Sigma Aldrich, USA.	<a href="https://www.aldrichmarketselect.com/">https://www.aldrichmarketselect.com/</a>	[59]
MolPort	Commercial database. Launched in Riga, Latvia.	<a href="https://www.molport.com/">https://www.molport.com/</a>	[59, 86]
Herb database	Traditional Chinese Medicine database guided by references and high-throughput experiments. Created by researchers in Beijing University of Chinese Medicine.	<a href="http://herb.ac.cn/">http://herb.ac.cn/</a>	[67]
TCM Database@Taiwan	Freely accessible small molecules database on Traditional Chinese Medicine. Constructed by Computational Systems Biology Laboratory at the China Medical University (Taiwan). <a href="http://researcher.nsc.gov.tw/ycc0929/">http://researcher.nsc.gov.tw/ycc0929/</a>	<a href="http://tcm.cmu.edu.tw">http://tcm.cmu.edu.tw</a>	[74]
Plant Database	Contains phytochemicals of medicinal plants from East Africa, North Africa, North East	Not found	[64]
Molecular Modelling Section (MMS) Database	In-house database by Molecular Modelling Section (MMS), Department of Pharmaceutical Sciences, University of Padova, Italy	Not found	[62, 63]

<https://doi.org/10.1371/journal.pone.0301396.t006>

PubChem and ZINC were among the most popular databases used. Both were public databases with more than millions of chemical structures readily downloadable in several common file formats. As of today, PubChem contains 110 million compounds whereas ZINC20 (the current version) contains 750 million molecules with 230 million purchasable compounds in ready-to-dock format (both websites accessed on 23 October 2023).

The size of these popular databases has grown in an exponential manner throughout the years. When Leggett et al. performed their virtual screening, '2007 ZINC drug-like library' was

Table 7. Target proteins and anticancer properties of anthraquinone derivatives elucidated by CADD methods.

Database/ Compounds Screened	Target(s) Involved: PDB ID	Cancer Type	Name & Structure of Identified Best Hit Derived from Anthraquinone	Observed Interactions/ Predicted Binding Mode by the Molecular Modelling	In-vitro Validation (Potency)	Study
Structural derivatives of 1-hydroxy-2-methyl-anthraquinone retrieved from PubChem	p53 Y220C mutant: 1TUP and 2XOW	General	Alizarin (Among the best six ligands)	Not reported	Not performed	[55]
Six cationic porphyrin-anthraquinone hybrids built with the help of GaussView	Cyclin-Dependent Kinases (CDKs): 1DI8	General	Porphyrin-anthraquinone hybrid with one substituent (mono-H2PyP-AQ)	Anthraquinone tail projected to the hydrophobic region while the porphyrin core interacted with the hinge residues Phe80, Glu81, and Leu83	Not performed	[56]
Sixty phytochemicals of <i>Acalypha indica</i> retrieved from PubChem	BRAF kinase: 6XFP (homology modelling template)	Melanoma	2-methyl anthraquinone (Among the top four phytochemicals)	Hydrophobic interactions were observed with Ile463, Val471, Ala481, Trp531 and Phe583.	Not performed	[57]
Vendor Database (Maybridge and SPECS)	Epidermal growth factor receptor (EGFR) wild type (WT) and EGFR T790M mutant: 2JIV, 2JIU and 3IKA	Lung Cancer	1,4-dihydroxy-5- methoxy-anthraquinone (Compound 4)	Hydrogen bonds observed with Met793, Gln791 and a conserved water molecule via two hydroxyl groups and a methoxyl group.	$IC_{50}$ EGFR-WT: 4.3 $\mu$ M; $IC_{50}$ EGFR-T790M/ L858R: 17.6 $\mu$ M	[58]
UM63 as starting compound for similarity search in MolPort (7,591,844 entries) and Aldrich Market Select (8,512,248 entries)	Ubiquitin-like containing PHD and RING fingers domains 1 (UHRF1): 3CLZ	General	Anthrarobin (AMSA2) -structurally related to UM63 with anthraquinone core	The isolated hydroxyl group H-bonded with the backbone of Val44 while the catechol moiety H-bonded with the backbone of Ala463 and Thr479, and the side chain of Asp469.	$IC_{50}$ UHRF1: 5.40 $\mu$ M	[59]
National Cancer Institute (NCI) diversity database (1990 compounds)	Epidermal growth factor receptor (EGFR): 1M17	Lung cancer	NSC125910 (Compound G)	Compound G formed hydrogen bond with Lys721 via O-atom of the dioxoanthracene group, while the O-atom and H-atom of the sulfophenyl group H-bonded to the Cys773 and Asp776.	Not performed	[60]
			NSC299137 (Compound J)	Compound J formed a weak hydrogen bond between the backbone-N of Met769 and the carbonyl oxygen of the dioxoanthracene-ring.		
MMsINC (in-house database contained around 4 million of synthetic and natural compounds)	Protein Kinase CK1 delta (CK1d): 2CSN and 1EH4 (homology modelling template)	General	1,4-diamino-anthraquinone (Compound 1)	Stabilizing interaction was made between one of the amino groups and Glu83 while another amino group H-bonded with Asp149. Another hydrogen bond was observed with Leu85. Hydrophobic bonds were also observed with Ile15, Ile23, Ala36, Leu135, Ile147.	$IC_{50}$ CK1d: 0.33 $\mu$ M	[61]
Molecular Modelling Section (MMS) database (more than 3000 of both synthetic and natural compounds)	Protein kinase CK2: 1JWH	General	Quinalizarin (1,2,5,8-tetrahydroxy-anthraquinone)	Not reported	$IC_{50}$ CK2: 0.15 $\mu$ M	[62]
Quinalizarin (Identified from previous SBVS study)	Protein kinase CK2: 4MD& (CK2 apo form) and 3QA0 (CK2 $\alpha$ apo form)	General	Quinalizarin (1,2,5,8-tetrahydroxy-anthraquinone)	Hydrogen bond formed between one of the hydroxyl groups with Val116 in the hinge region via a water molecule; another hydroxyl group H-bonded with His160 on one side, and with carbonyl of Arg47 on the other side.	$IC_{50}$ CK2 $\alpha$ $\beta$ : 0.15 $\mu$ M; $IC_{50}$ CK2 $\alpha$ : 1.35 $\mu$ M	[63]

(Continued)

Table 7. (Continued)

Database/ Compounds Screened	Target(s) Involved: PDB ID	Cancer Type	Name & Structure of Identified Best Hit Derived from Anthraquinone	Observed Interactions/ Predicted Binding Mode by the Molecular Modelling	In-vitro Validation (Potency)	Study
14000 phytochemicals from Plant Database and 14500 synthetic drugs from DrugBank	Maternal Embryonic Leucine Zipper Kinase (MELK): 5IH9	Breast cancer (especially Triple-Negative Breast Cancer)	Emodin	Hydrogen bond was formed with Cys89	IC <sub>50</sub> TNBC cell line: 30.30 μM	[64]
Five anthranoid skeletons and derivatives as hydroxy-anthraquinones constructed by the SYBYL sketcher module	Estrogen receptor alpha (ER $\alpha$ ): 1A52	Hormone-associated diseases including cancer	Glucofrangulin B	Hydrogen bond interactions were observed with residues Glu423, Lys520, His516, His516, Arg515, Arg515, Arg515 and Cys381.	Not performed	[65]
	Estrogen receptor beta (ER $\beta$ ): 1QKM		Emodin	Hydrogen bond interaction was observed with Glu305.		
DrugBank (contained 7097 compounds)	human O-GlcNAcase (hOGA): 5M7T	General	Mitoxantrone	The anthraquinone moiety pi-pi stacked with Phe223, Tyr286, and Trp645, and the hydroxyl groups on the anthraquinone ring H-bonded to Lys98 and Asp175. One of the hydroxyethyl side chains H-bonded with Lys98 and Asn313. The other side chain interacted with Lys648 outside the pocket.	IC <sub>50</sub> hOGA: 7.3 μM	[66]
HERB database for small molecules	Reticulo-calbin2 (RCN2): From AlphaFold database	Cholangiocarcinoma (CCA)	Emodin	Hydrogen bonds were observed with residues Gln261, Asn259, Tyr239 and Arg234.	Not performed	[67]
FDA-approved Drug Library (by Selleck, contained 1375 drugs)	eEF-2K (Eukaryo-tic elongation factor-2 kinase): 3LKM (homology modelling template)	Breast cancer	Mitoxantrone	Hydrogen bonds were observed with residues Arg140, Lys170, Ile232, Glu 233, and Gly234.	Kd to eEF-2K: 9.11 μM	[68]
NCI/DTP Small Molecules Database (contained more than 140,000 compounds)	Focal Adhesion Kinase (FAK): 2J0J and 2J0L	General	Mitoxantrone and derivative (A18 compound): 1,4-bis (diethylamino)-5-8-dihydroxy anthraquinone	A18 compounds docked into the K454 site of the FAK kinase domain	Not calculated	[69]
Acros Organics subset of the Zinc database (contained 13,754 compounds after filtering)	Human cytoplasmic nucleotidase (cN-II): 2JC9	General (especially acute myeloid leukemia)	Anthraquinone-2,6-disulfonic acid, disodium salt (AdiS)	Not reported	IC <sub>50</sub> RL (lymphoma) cell line: 750 μM	[70]
-166 known inhibitors retrieved from PubChem (filtered for pharmacophore modelling)-Zbc library (26432 natural compounds)	ATP-binding cassette Super-family G member 2 protein (ABCG2): 5NJ3	General	Rhein (ZINC4098704)	Twenty-seven hydrophobic bonds were observed with residues Tyr613(2), Ile423(4), Gln424(5), Ser420(4), Lys616(3), Tyr605(2), Ala606(2) and Thr607(5). Four hydrogen bonds were observed with Gln424, Thr607 and Lys616.	Not performed	[71]
9 compounds extracted from marine <i>Streptomyces</i> , retrieved from PubChem	human epidermal growth factor receptor 2 (HER2): 1N8Z	General	Marmycin A (marine-sourced anthraquinone-derivatives)	Interactions were observed with residues Cys312, Pro315, Cys316, Arg318, Val319, Cys320, Asn297, Gln298, Glu299, Cys312, Pro315, Cys316, Arg318, Glu326, Met324, Tyr321, and Phe349.	Not performed	[72]

(Continued)

Table 7. (Continued)

Database/ Compounds Screened	Target(s) Involved: PDB ID	Cancer Type	Name & Structure of Identified Best Hit Derived from Anthraquinone	Observed Interactions/ Predicted Binding Mode by the Molecular Modelling	In-vitro Validation (Potency)	Study
2007 Zinc drug like library (contained 2 million compounds)	Arylamine N-acetyl-transferase 1 (NAT1): 2PQT and 2PFR	General	9,10-dihydro-9,10-dioxo-1,2-anthracenediyl diethyl ester (Compound 10)	Not reported	IC <sub>50</sub> NAT1:0.75 μM	[73]
			Alizarin		IC <sub>50</sub> NAT1 0.89 μM	
TCM database@ Taiwan	G-quadruplex (G4s): 1KF1 (parallel) 143D (anti-parallel) 2JPZ (hybridized)	General	Emodin	Emodin docked at the binding site of 143D in a larger groove than in 2JPZ, and the hydroxyl of emodin H-bonded to the phosphoric acid oxygen atom of DT9. No ligands were observed for parallel G4 due to the lack of a pocket site on the surface.	Not calculated	[74]
Anthraquinones derivatives from <i>Morinda citrifolia</i> attained from PubChem	Dihydro-folate reductase (DHFR): 1DLS	General	Damnacanthal	Hydrogen bonds were observed with Asp21, Ser59, and Asn64. Van der waal interactions were observed with Tyr22, Phe31, Ile60.	Not performed	[75]
Four phytochemicals from <i>Amomum subulatum</i> , retrieved from PubChem for pharmacophore screening via ZINCPharmer	HPVE6:4GIZ; BCL2:2XA0; XIAP:1F9X; HPVE7 & LIVIN: homology modelling	Cervical cancer	Rhein (among the 4 phytochemicals identified)	Not reported	Not performed	[76]
200 compounds resulted from similarity search on PubChem based on eight starting compounds identified by previous study	Epidermal growth factor receptor (EGFR): 1M17	Lung cancer	Nogalamycin N-oxide (NSC116555),	Binding interaction was observed with Thr854 in the DFG-in conformation. Additional conserved hydrogen bonds were observed with Cys797, Ala743, Lys745, Asp855, Cys797, Leu844, and Phe856.	IC <sub>50</sub> EGFR: 31.56 nM	[77]
Chemical Library of about 140 000 compounds maintained by University of Helsinki	Toll-like receptor 4 (TLR4): 3FXI	General	Mitoxantrone (Structure as above)	Not reported	Not calculated	[78]
25 Derivatives of Monomeric anthraquinone based, bis-benzimidazole and hybrid carbazole-benzimidazole type ligands	G-quadruplex (G4) RNA: 2KBP	General	AOPH with the hydroxyethyl piperazine side chain (The best among nine anthraquinone-based ligands)	The anthraquinone moiety partially stacked on the G-tetrad surface interacted with G17 and G2, and the hanging piperazine part interacted with the groove containing U18 and A20 of strand B.	Not performed	[79]
Combinatorial library of over 2 million compounds created from in-silico modification of anthrafuran	human aurora kinase B (AurB): 4AF3	General	Naphthoisatin (Compound 2)	Hydrophobic interactions and hydrogen bonds were observed with residues Leu83, Gly84, Lys85, Phe88, Tyr156, Ala157, Arg159, Gly160, Glu161, Lys164, Glu 165, Glu204 and Leu207.	IC <sub>50</sub> AurB 7.4 μM	[80]

(Continued)

Table 7. (Continued)

Database/ Compounds Screened	Target(s) Involved: PDB ID	Cancer Type	Name & Structure of Identified Best Hit Derived from Anthraquinone	Observed Interactions/ Predicted Binding Mode by the Molecular Modelling	In-vitro Validation (Potency)	Study
235 compounds that have ring subsets from ZINC15 database after filtered off macrocyclic molecules	Histone deacetyl-lases isoform 6 (HDAC6): 5EF7 and 5EDU	General	Compound 3	Hydrogen bonds were observed with the side chain of Y745, backbone amide oxygen of G582, and imidazole side chain of H573. The middle quinone and phenyl rings were sandwiched between lipophilic side chains of F583 and F643, forming pi-pi interactions. Two carbonyl oxygens of the quinone ring formed hydrogen bond with H614 and S531.	IC <sub>50</sub> HDAC6: 56 nM	[81]
21 natural anthraquinone derivatives retrieved from PubChem	Matrix metallo-proteinase-13 (MMP-13): 5B5O	General	Pulmatin	Four hydrogens and four hydrophobic interactions formed with Gly183, Leu184, Ala186, Glu223, Ile243 and Tyr244.	Not performed	[82]
78 anthraquinone-based inhibitors retrieved from literatures used for 3D QSAR modelling	Phospho-glycerate mutase 1 (PGAM1): 5Y35	General	Compound Number 81 (among the seven new inhibitors designed based on the QSAR modelling)	Not reported	Not performed	[83]
3000 FDA-approved drugs from ZINC database	NEDD8-activating enzyme (NAE): 3GZN	General	Mitoxantrone	Three hydrogen bonds were observed with Thr103, Gln112, and Lys307 at the binding site. One of the alkyl chains extended out into the solvent region, while the other probed deeper into the binding pocket.	EC <sub>50</sub> Caco-2 cell: 1.3 μM	[84]
Emodin analogues queried by substructure search from Key Organic database (49,415 compounds)	Aurora Kinase A (AURKA): 5ORL	Ovarian cancer	8L-902 (among the 9 analogues identified)	One hydrogen bond interaction, two pi-pi interactions, one pi-sigma interaction, and four pi-alkyl interactions were observed.	Not calculated	[85]
MolPort (6,504,839 molecules)	Ubiquitin-like containing PHD & RING fingers domains 1 (UHRF1): 3CLZ	General	UM63	Interaction was stabilized by pi-pi stacking with the side chain of Tyr478 and several hydrogen bonds to key residues Asp469, Thr479, Gly448, Gly465, Ala463.	IC <sub>50</sub> to SRA-induced base flipping: 4.4 μM	[86]

<https://doi.org/10.1371/journal.pone.0301396.t007>

used whereby the number of molecules contained was reported at about 700 hundred thousand [73]. The reported number was increased to 120 million when another study screened the database (ZINC15) about a decade later (81). ZINC database also contains different subsets to tailor for different aims of screening. For example, the Acros Organic database used by Jordonheim et al. was one of the subsets after filtering with the Lipinski rules [70]. On the other hand, Khan and colleagues used the subset of synthesizable natural compounds (Zbc) loaded with biogenic lead-like compounds, primary and secondary metabolites of natural products [71]. Last but not least, the database of 'FDA-approved drugs (via DSSTOX)' used in another study was also obtained from ZINC (<http://zinc.docking.org/catalogs/fda>) [84].

It is inevitable that some of the compounds are overlapping in various libraries, but each library still carries unique features. In this review, some of the studies used more than one database to cater to their different objectives, these included a combination of two public databases [71, 76], two commercial databases [58, 59], or a plant database together with a synthetic

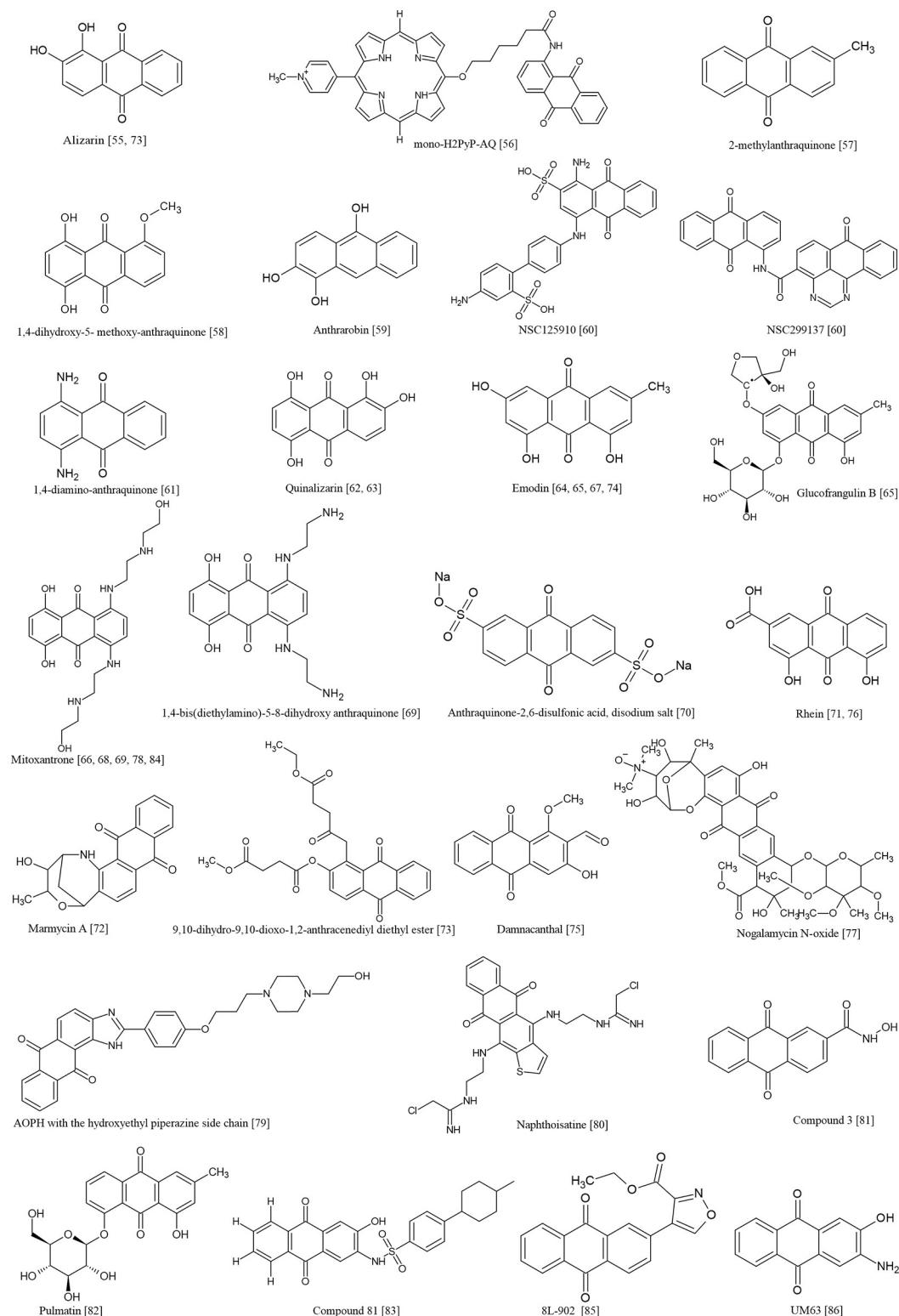


Fig 3. Chemical structure depiction of the hit compounds.

<https://doi.org/10.1371/journal.pone.0301396.g003>

database [64]. In order to optimize the screening outcome, another useful strategy was the construction of a virtual combinatorial library by enumerating the main scaffold of existing active compounds with a pre-selected drug block in which a target-focus library could be formed [80].

It is also worth highlighting the efforts of the research community in building natural product databases owing to the increasing momentum seen in herbal research. The HERB database used by Gao et al. [67] was built by linking targets and diseases to herbs, aiming to aid the modernization of Traditional Chinese Medicine especially in rational drug design [98]. Besides, another TCM database@Taiwan used in one of the studies [74] was designed in both English and Chinese languages, deemed as the largest collection of 3D structures for Traditional Chinese Medicine to date in a freely accessible ready-to-dock format to support the in-silico research [99].

Ligand constructors were used to draw the 3D structures of the ligands when the data required was not readily available. Among the tools used included Chemicalize [55], GaussView [56], SYBYL sketcher module [65], ACD/ChemSketch [75], Marvin Sketch [80], and ChemDraw [83].

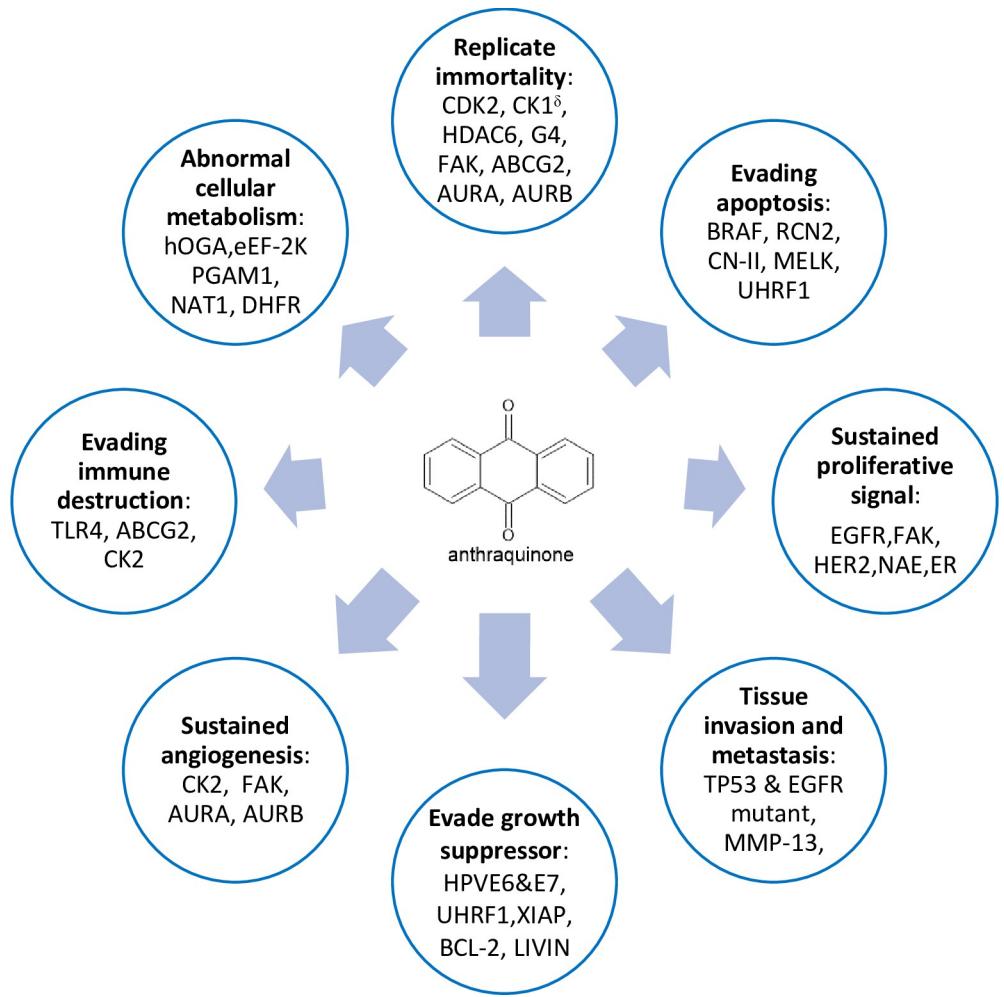
### Identified targets and anticancer properties of anthraquinone derivatives

The main target proteins investigated in the included studies for developing potential therapeutics against different types of cancer are presented. The best hit was either derived from a set of starting compounds that contained anthraquinone scaffold or discovered by screening varied sizes of databases. The identified hit compounds that contained anthraquinone scaffold, the observed interactions with the target proteins simulated by the molecular modelling and their corresponding potency (if tested experimentally) are also displayed (Table 7). (Refer to Fig 3 for the chemical structure depiction of the respective hit compounds).

Anthraquinone scaffold is a privileged scaffold that carries biological activities against a wide range of macromolecular targets. Privileged scaffold is defined as the core structure that can interact with more than one receptor with high affinity [100]. The prioritized anthraquinone derivatives from the included studies exhibit different types of interactions with the identified targets, these included hydrogen bonding, hydrophobic interactions, pi-pi stacking, pi-sigma and pi-alkyl interactions (Table 7). The calculated  $IC_{50}$  (half maximal inhibitory concentration) of these hit compounds determined from the in-vitro experiment ranged from micromolar [58, 59, 61–64, 66, 70, 73, 80, 86] to nanomolar [77, 81] scale, indicating the high affinity of these anthraquinone derivatives against the various macromolecule targets. There was one study reported the potency of the identified hit in the form of  $K_d$  [dissociation constant] [68] and another study calculated  $EC_{50}$  (half-maximal effective concentration) [84], both of these values were also in the micromolar scale, suggesting the potential of these hit compounds as the promising candidates for further development into new anticancer drugs.

The majority of the studies looked into the role of these targets in general cancer whereas some of them dived into specific cancer types. Among these were the top killer cancers in men and women namely lung cancer [58, 60, 77] and breast cancer [64, 68] respectively. Another two types of cancer that threaten the female population included cervical [76] and ovarian cancer [85]. Stubborn diseases with poor prognosis such as melanoma [57], cholangiocarcinoma [67] and acute myeloid leukaemia [70] were also investigated. The macromolecular targets involved are responsible for managing these different malignancies via various mechanisms, mainly to tackle the hallmarks of cancer [101] (Fig 4).

Many of the studies investigated the role of anthraquinone derivatives against the protein kinases family, the enzymes that carry multiple roles at once. The serine-threonine kinases like

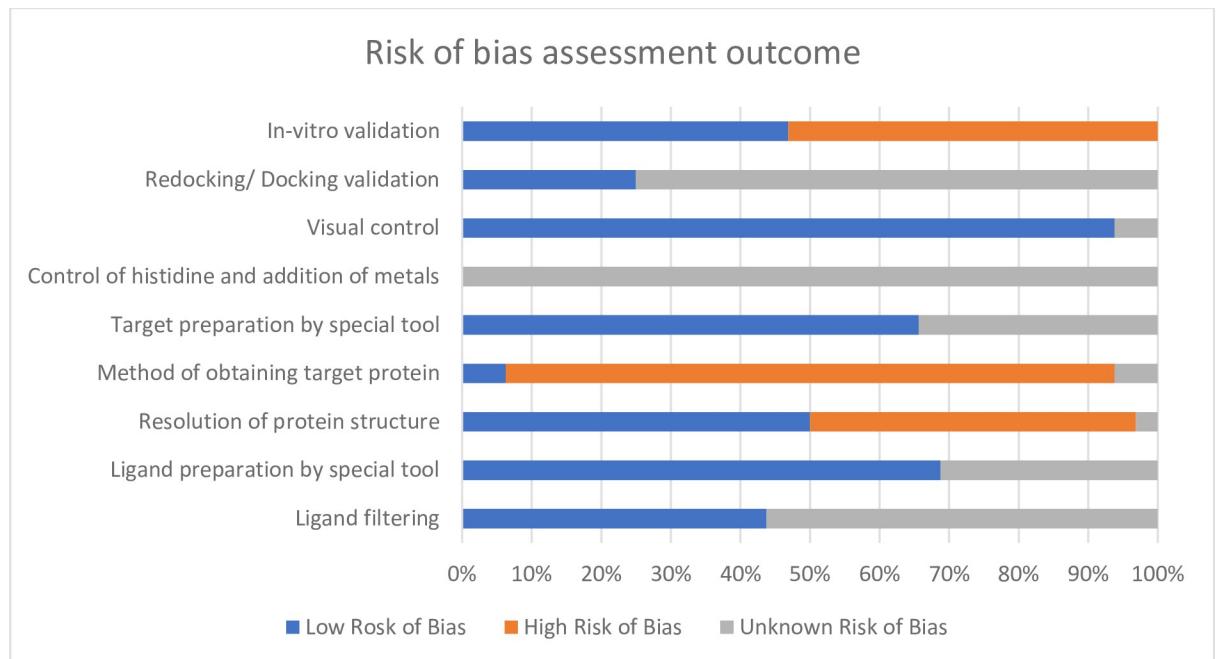


**Fig 4.** Targets involved that tackled the hallmarks of cancer.

<https://doi.org/10.1371/journal.pone.0301396.g004>

CDK2 [56], CK1 $\delta$  [61], CK2 [62, 63], MELK [64], AURA [85], AURB [80], BRAF [57] involved in many cellular processes including signal transduction and controlling cell cycle progression in various phases. Another tyrosine kinase family investigated were FAK [69], EGFR [58, 60, 77], and HER2 [72] that mediated cell proliferation and survival. Dysregulation of these kinases in malignant cells resulted in a sustained supply of blood to the tumours, uncontrolled growth and immortality of the cells, in which anthraquinone derivatives managed to serve as inhibitors against these targets to activate apoptosis. Inhibition of targets like RCN2 [67] and cN-II [70] also induced programmed cell death by activation of natural killer cells.

Another calcium/calmodulin-dependent kinase known as eEF-2K stimulated the cancer growth by mediating through autophagy and facilitating the switch from oxidative phosphorylation to glycolysis in response to metabolic stresses [68]. Other cancer-promoting enzymes that involved in carcinogen metabolism included hOGA [66], PGAM1 [83], NAT1 [73] and DHFR [75]. Anthraquinone derivatives such as mitoxantrone, alizarin and damnacanthal managed to turn off this special metabolic pathway and increase the susceptibility of cancer cells to apoptosis induced by chemotherapy.



**Fig 5. Risk of bias assessment outcome.**

<https://doi.org/10.1371/journal.pone.0301396.g005>

Apart from that, anthraquinone derivatives were also found to trigger telomere dysfunction by targeting G4 and suppressing tumour growth [74, 79]. Emodin, one of the naturally-occurring anthraquinones was reported to work synergistically with telomerase inhibitor in sustaining the telomere defect and enhancing cancer cell damage [74]. On the other hand, oncoproteins such as HPVE6, HPVE7 [76], UHRF1 [59, 86] and anti-apoptotic protein BCL-2, XIAP, LIVIN [76] are involved in protecting the cancer cells from growth suppressor and hence facilitating their growth. Anthrarobin and rhein were among the potential hits found to modulate these target proteins leading to cancer cells suppression.

Over-expression of proteins like mutated TP53 [55], mutated EGFR [58] and MMP-13 [82] enhanced the ability of cancer cells to evade adjacent tissues leading to metastasis. Alizarin, Pulmatin and 1,4-dihydroxy-5-methoxy-anthraquinone were identified as potential inhibitors for mutated TP53, MMP-13 and mutated EGFR respectively and help in preventing the spread of cancer cells. Another investigated target, TLR4 was found to serve as a mediator of innate immune system activation [78]. In addition, the elevated level of ABCG2 on the plasma membrane reduced the effect of anti-cancer drugs by enhancing the efflux process leading to multi-drug resistance (MDR) [71]. This phenomenon was also observed with overexpression of CK2 [62, 63]. Virtual screening discovered quinalizarin, rhein and mitoxantrone as modulators of these targets and exhibit therapeutic roles in managing multidrug resistance. Last but not least, one of the studies showed that phytochemicals like emodin and glucofrangulin B also carried phytoestrogenic activities in which their effect on estrogen receptors was investigated [65]. These anthraquinone derivatives may play an important role in hormone-related disorders including breast cancer that commonly associated with elevated levels of estrogen in blood.

Taken altogether, this review shows that anthraquinone was proven to be a valuable compound containing the most privileged scaffold with great therapeutic potential to be developed into wide-spectrum anticancer drugs.

## Risk of bias assessment

The risk of bias checklist was designed mainly to assess the reporting quality of docking-based studies (Table 2). Since molecular docking was involved as the main screening tool in all of the studies included in this review, the risk assessment managed to cover all papers. The results are illustrated in Fig 5.

For all 32 papers, none of them described on the control of histidine and addition of metal as part of their target preparation steps. It may be due to the lacking of clear guidelines to guide the proper reporting of in-silico research. Nevertheless, nearly 70% of the studies reported on the ligand optimization steps which include ionization assessment and generation of possible conformation, as well as the description on general target protein preparation mainly performed by the built-in tools associated with the docking software, whilst there were still some studies did not elaborate this domain.

For the method of target protein generation, further search from the RCSB Protein Data Bank revealed that the majority of the crystal structures used in the studies were obtained via X-ray diffraction method, where this method was more commonly in use compared to NMR spectroscopy. However, X-ray crystallography was ranked as “High Risk of Bias” in the checklist as it was opined that only NMR spectroscopy captures the three-dimensional structure details in a medium close to the actual biological environment [49]. Perhaps the ranking of this domain is worth revisiting since both NMR spectroscopy and X-ray crystallography offer different pros and cons, and both tools are highly complementary [102].

What is worth mentioning is almost all studies except two studies (data not found) performed visual control to inspect the docking pose to ensure there were no structural artifacts derived from the computer calculation. However, for docking validation by redocking, only a quarter of the studies elaborated on this crucial process. Since CADD methods are based on prediction by computational calculation and algorithm, experimental validation is still required to confirm the outcome of the study. Unfortunately, more than half of the studies did not pursue or yet report on the in-vitro validation.

In general, this assessment revealed that there is a gap in a proper or standard guide to enhance the reporting quality of studies involving molecular docking in particular. Nevertheless, Monks and colleagues have developed a 20-items checklist to improve the reporting of discrete-event simulation, system dynamics and agent-based simulation models within the field of Operational Research and Management Science, termed as “Strengthening The Reporting of Empirical Simulation Studies (STRESS)” [103]. However, the applicability of this checklist for this type of review is yet to be discovered.

## Limitation

One of the limitations of this review was that only articles written in English were included, resulting in missing important papers which were written in other languages. Besides, the included studies utilized different methods and approaches hence resulting in heterogeneity and difficulties in performing pooled analysis. On the other hand, there is no other systematic review that evaluated the anticancer drug design and discovery of anthraquinone derivatives based on CADD methods. Therefore, there is no comparison that could be done for assessing agreements or disagreements with other studies.

## Conclusion

The outcome of bias risk assessment implied the need for proper and more standardized guidelines in order to improve the reporting quality of in-silico studies. Nevertheless, the increasing number of publications retrieved throughout the years has proven the role of

CADD methods as an indispensable tool in the era of modern drug design and discovery. These methods were particularly useful in the early stages of the drug design and discovery trajectory to screen, identify and optimize the potential hits in a systematic manner for shortlisting only the most promising compounds for further analysis.

Structure-based and ligand-based methods were either used alone or in combination to obtain consensus prediction. Combining different tools offered the advantage of enrichment enhancement, enabling the synergizing of the strengths and complementing the weaknesses of each method. The choices of software used were mainly project specific although user-friendly interface could have also served as one of the driving factors.

The findings in this review also strengthened the fact that CADD methods enabled deeper exploration of the anticancer potential of anthraquinone-based compounds up to the molecular level. The utilization of in-silico techniques for the study of anthraquinone derivatives has made it possible to obtain further insights into their structural, biological and pharmacological properties. Notably, anthraquinone derivatives demonstrated remarkable anticancer properties by targeting a wide spectrum of biological targets that tackled the abnormalities of cancer cells in an all-rounded fashion. The synergy between computational and experimental approaches contributes to a more comprehensive understanding of anthraquinones' potential as anticancer therapeutics.

As cancer continues to pose a threat to the global healthcare system, the role of anthraquinones, coupled with CADD methods, offers a promising avenue for drug discovery. By harnessing the power of computational tools and leveraging the natural diversity of anthraquinone compounds, researchers can expedite the development of better drug to address the unmet medical needs in cancer treatment by improving the treatment outcome for cancer patients.

The insights gained from this review can serve as the scientific evidence-based guidance to improve the success rate of future cancer research. It is recommended for upcoming research to follow closely with the rapid advancement of CADD and make full use of integrated tools to facilitate the design and discovery of novel anticancer therapeutics expanded from the privileged anthraquinone scaffold.

## Supporting information

**S1 Checklist. PRISMA 2020 checklist.**  
(DOCX)

**S1 Fig. Chemical structure of 9,10-anthraquinone scaffold.**  
(TIF)

**S1 Table. Example of search strategy in PubMed.**  
(DOCX)

**S2 Table. Data extraction form.**  
(DOCX)

## Acknowledgments

We thank Universiti Brunei Darussalam for the University Graduate Scholarship awarded to HMC and SM.

## Author Contributions

**Conceptualization:** Hui Ming Chua, Nurolaini Kifli.

**Data curation:** Hui Ming Chua, Said Moshawih.

**Formal analysis:** Hui Ming Chua, Said Moshawih.

**Investigation:** Hui Ming Chua, Said Moshawih.

**Methodology:** Hui Ming Chua, Said Moshawih, Nurolaini Kifli, Hui Poh Goh, Long Chiau Ming.

**Project administration:** Nurolaini Kifli.

**Supervision:** Nurolaini Kifli.

**Validation:** Long Chiau Ming.

**Writing – original draft:** Hui Ming Chua.

**Writing – review & editing:** Hui Ming Chua, Said Moshawih, Nurolaini Kifli, Hui Poh Goh, Long Chiau Ming.

## References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA: A Cancer Journal for Clinicians. 2023; 73(1):17–48. <https://doi.org/10.3322/caac.21763> PMID: 36633525
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021; 71(3):209–49. <https://doi.org/10.3322/caac.21660> PMID: 33538338
3. Wang X, Zhang H, Chen X. Drug resistance and combating drug resistance in cancer. Cancer Drug Resist. 2019; 2(2):141–60. <https://doi.org/10.20517/cdr.2019.10> PMID: 34322663
4. Bidram E, Esmaeili Y, Ranji-Burachaloo H, Al-Zaubai N, Zarrabi A, Stewart A, et al. A concise review on cancer treatment methods and delivery systems. Journal of Drug Delivery Science and Technology. 2019; 54:101350.
5. Schlander M, Hernandez-Villafuerte K, Cheng C-Y, Mestre-Ferrandiz J, Baumann M. How Much Does It Cost to Research and Develop a New Drug? A Systematic Review and Assessment. PharmacoEconomics. 2021; 39(11):1243–69. <https://doi.org/10.1007/s40273-021-01065-y> PMID: 34368939
6. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nature Reviews Drug Discovery. 2010; 9(3):203–14.
7. Doytchinova I. Drug Design-Past, Present, Future. Molecules. 2022; 27(5). <https://doi.org/10.3390/molecules27051496> PMID: 35268598
8. Hughes JP, Rees S, Kalindjian SB, Philpott KL. Principles of early drug discovery. Br J Pharmacol. 2011; 162(6):1239–49. <https://doi.org/10.1111/j.1476-5381.2010.01127.x> PMID: 21091654
9. Chang Y, Hawkins BA, Du JJ, Groundwater PW, Hibbs DE, Lai F. A Guide to In Silico Drug Design. Pharmaceutics. 2023; 15(1):49.
10. Moshawih S, Lim AF, Ardianto C, Goh KW, Kifli N, Goh HP, et al. Target-Based Small Molecule Drug Discovery for Colorectal Cancer: A Review of Molecular Pathways and In Silico Studies. Biomolecules. 2022; 12(7):878. <https://doi.org/10.3390/biom12070878> PMID: 35883434
11. Vázquez J, López M, Gibert E, Herrero E, Luque FJ. Merging Ligand-Based and Structure-Based Methods in Drug Discovery: An Overview of Combined Virtual Screening Approaches. Molecules. 2020; 25(20). <https://doi.org/10.3390/molecules25204723> PMID: 33076254
12. Pinzi L, Rastelli G. Molecular Docking: Shifting Paradigms in Drug Discovery. Int J Mol Sci. 2019; 20(18). <https://doi.org/10.3390/ijms20184331> PMID: 31487867
13. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nature Reviews Drug Discovery. 2004; 3(11):935–49. <https://doi.org/10.1038/nrd1549> PMID: 15520816
14. Giordano D, Biancaniello C, Argenio MA, Facchiano A. Drug Design by Pharmacophore and Virtual Screening Approach. Pharmaceuticals. 2022; 15(5):646. <https://doi.org/10.3390/ph15050646> PMID: 35631472

15. Wermuth CG, Ganellin CR, Lindberg P, Mitscher LA. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure and Applied Chemistry*. 1998; 70(5):1129–43.
16. Lill M. Virtual screening in drug design. *Methods Mol Biol*. 2013; 993:1–12. [https://doi.org/10.1007/978-1-62703-342-8\\_1](https://doi.org/10.1007/978-1-62703-342-8_1) PMID: 23568460
17. Baltrukevich H, Podlewska S. From Data to Knowledge: Systematic Review of Tools for Automatic Analysis of Molecular Dynamics Output. *Front Pharmacol*. 2022; 13:844293. <https://doi.org/10.3389/fphar.2022.844293> PMID: 35359865
18. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov*. 2015; 10(5):449–61. <https://doi.org/10.1517/17460441.2015.1032936> PMID: 25835573
19. Salo-Ahen OMH, Alanko I, Bhadane R, Bonvin AMJJ, Honorato RV, Hossain S, et al. Molecular Dynamics Simulations in Drug Discovery and Pharmaceutical Development. *Processes*. 2021; 9(1):71.
20. Brogi S, Ramalho TC, Kuca K, Medina-Franco JL, Valko M. Editorial: In silico Methods for Drug Design and Discovery. *Frontiers in Chemistry*. 2020; 8. <https://doi.org/10.3389/fchem.2020.00612> PMID: 32850641
21. Podlewska S, Czarnecki WM, Kafel R, Bojarski AJ. Creating the New from the Old: Combinatorial Libraries Generation with Machine-Learning-Based Compound Structure Optimization. *Journal of chemical information and modeling*. 2017; 57(2):133–47. <https://doi.org/10.1021/acs.jcim.6b00426> PMID: 28158942
22. Suay-García B, Bueso-Bordils JI, Falcó A, Antón-Fos GM, Alemán-López PA. Virtual Combinatorial Chemistry and Pharmacological Screening: A Short Guide to Drug Design. *Int J Mol Sci*. 2022; 23(3).
23. Moshawih S, Goh HP, Kifli N, Idris AC, Yassin H, Kotra V, et al. Synergy between machine learning and natural products cheminformatics: Application to the lead discovery of anthraquinone derivatives. *Chemical Biology & Drug Design*. 2022; 100(2):185–217.
24. Moshawih S, Hadikhani P, Fatima A, Goh HP, Kifli N, Kotra V, et al. Comparative analysis of an anthraquinone and chalcone derivatives-based virtual combinatorial library. A cheminformatics “proof-of-concept” study. *Journal of Molecular Graphics and Modelling*. 2022; 117:108307. <https://doi.org/10.1016/j.jmgm.2022.108307> PMID: 36096064
25. Patridge E, Gareiss P, Kinch MS, Hoyer D. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov Today*. 2016; 21(2):204–7. <https://doi.org/10.1016/j.drudis.2015.01.009> PMID: 25617672
26. Khan AW, Farooq M, Haseeb M, Choi S. Role of Plant-Derived Active Constituents in Cancer Treatment and Their Mechanisms of Action. *Cells*. 2022; 11(8):1326. <https://doi.org/10.3390/cells11081326> PMID: 35456005
27. Naeem A, Hu P, Yang M, Zhang J, Liu Y, Zhu W, et al. Natural Products as Anticancer Agents: Current Status and Future Perspectives. *Molecules*. 2022; 27(23). <https://doi.org/10.3390/molecules27238367> PMID: 36500466
28. Lown JW. Anthracycline and anthraquinone anticancer agents: current status and recent developments. *Pharmacol Ther*. 1993; 60(2):185–214. [https://doi.org/10.1016/0163-7258\(93\)90006-y](https://doi.org/10.1016/0163-7258(93)90006-y) PMID: 8022857
29. Hussain H, Al-Harrasi A, Al-Rawahi A, Green IR, Csuk R, Ahmed I, et al. A fruitful decade from 2005 to 2014 for anthraquinone patents. *Expert Opin Ther Pat*. 2015; 25(9):1053–64. <https://doi.org/10.1517/13543776.2015.1050793> PMID: 26036306
30. Malik EM, Müller CE. Anthraquinones As Pharmacological Tools and Drugs. *Med Res Rev*. 2016; 36(4):705–48. <https://doi.org/10.1002/med.21391> PMID: 27111664
31. Tian W, Wang C, Li D, Hou H. Novel anthraquinone compounds as anticancer agents and their potential mechanism. *Future Medicinal Chemistry*. 2020; 12(7):627–44. <https://doi.org/10.4155/fmc-2019-0322> PMID: 32175770
32. Siddamurthi S, Gutti G, Jana S, Kumar A, Singh SK. Anthraquinone: a promising scaffold for the discovery and development of therapeutic agents in cancer therapy. *Future Medicinal Chemistry*. 2020; 12(11):1037–69. <https://doi.org/10.4155/fmc-2019-0198> PMID: 32349522
33. Tikhomirov AS, Shtil AA, Shchekotikhin AE. Advances in the Discovery of Anthraquinone-Based Anti-cancer Agents. *Recent Pat Anticancer Drug Discov*. 2018; 13(2):159–83. <https://doi.org/10.2174/1574892813666171206123114> PMID: 29210664
34. Chun-Guang W, Jun-Qing Y, Bei-Zhong L, Dan-Ting J, Chong W, Liang Z, et al. Anti-tumor activity of emodin against human chronic myelocytic leukemia K562 cell lines in vitro and in vivo. *European journal of pharmacology*. 2010; 627(1–3):33–41. <https://doi.org/10.1016/j.ejphar.2009.10.035> PMID: 19857484

35. Dai G, Ding K, Cao Q, Xu T, He F, Liu S, et al. Emodin suppresses growth and invasion of colorectal cancer cells by inhibiting VEGFR2. *European journal of pharmacology*. 2019; 859:172525. <https://doi.org/10.1016/j.ejphar.2019.172525> PMID: 31288005
36. Liu Q, Hodge J, Wang J, Wang Y, Wang L, Singh UP, et al. Emodin reduces breast cancer lung metastasis by suppressing macrophage-induced breast cancer cell epithelial-mesenchymal transition and cancer stem cell formation. *Theranostics*. 2020; 10(18):8365. <https://doi.org/10.7150/thno.45395> PMID: 32724475
37. Li Q, Wen J, Yu K, Shu Y, He W, Chu H, et al. Aloe-emodin induces apoptosis in human oral squamous cell carcinoma SCC15 cells. *BMC complementary and alternative medicine*. 2018; 18(1):1–5.
38. Lee H-Z, Hsu S-L, Liu M-C, Wu C-H. Effects and mechanisms of aloe-emodin on cell death in human lung squamous cell carcinoma. *European journal of pharmacology*. 2001; 431(3):287–95. [https://doi.org/10.1016/s0014-2999\(01\)01467-4](https://doi.org/10.1016/s0014-2999(01)01467-4) PMID: 11730720
39. Pecere T, Gazzola MV, Mucignat C, Parolin C, Vecchia FD, Cavaggioni A, et al. Aloe-emodin is a new type of anticancer agent with selective activity against neuroectodermal tumors. *Cancer research*. 2000; 60(11):2800–4. PMID: 10850417
40. Lu CC, Yang JS, Huang AC, Hsia TC, Chou ST, Kuo CL, et al. Chrysophanol induces necrosis through the production of ROS and alteration of ATP levels in J5 human liver cancer cells. *Molecular nutrition & food research*. 2010; 54(7):967–76. <https://doi.org/10.1002/mnfr.200900265> PMID: 20169580
41. Malik MS, Alsantali RI, Jassas RS, Alsimaree AA, Syed R, Alsharif MA, et al. Journey of anthraquinones as anticancer agents—a systematic review of recent literature. *RSC Adv*. 2021; 11(57):35806–27. <https://doi.org/10.1039/d1ra05686g> PMID: 35492773
42. Smith PJ, Blunt NJ, Desnoyers R, Giles Y, Patterson LH. DNA topoisomerase II-dependent cytotoxicity of alkylaminoanthraquinones and their N-oxides. *Cancer Chemother Pharmacol*. 1997; 39(5):455–61. <https://doi.org/10.1007/s002800050598> PMID: 9054961
43. Zhao L, Zheng L. A Review on Bioactive Anthraquinone and Derivatives as the Regulators for ROS. *Molecules*. 2023; 28(24):8139. <https://doi.org/10.3390/molecules28248139> PMID: 38138627
44. Iwaloye O, Ottu PO, Olawale F, Babalola OO, Elekofehinti OO, Kikiwo B, et al. Computer-aided drug design in anti-cancer drug discovery: What have we learnt and what is the way forward? *Informatics in Medicine Unlocked*. 2023; 41:101332.
45. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021; 372:n71. <https://doi.org/10.1136/bmj.n71> PMID: 33782057
46. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*. 2021; 372:n160. <https://doi.org/10.1136/bmj.n160> PMID: 33781993
47. Chua HM, Moshawih S, Goh HP, Ming LC, Kifli N. Insights into the computer-aided drug design and discovery based on anthraquinone scaffold for cancer treatment: A protocol for systematic review. *PLoS One*. 2023; 18(9):e0290948. <https://doi.org/10.1371/journal.pone.0290948> PMID: 37656730
48. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *Journal of the Medical Library Association: JMLA*. 2016; 104(3):240. <https://doi.org/10.3163/1536-5050.104.3.014> PMID: 27366130
49. Taldaev A, Terekhov R, Nikitin I, Zhevlavkova A, Selivanova I. Insights into the Pharmacological Effects of Flavonoids: The Systematic Review of Computer Modeling. *International Journal of Molecular Sciences*. 2022; 23(11):6023. <https://doi.org/10.3390/ijms23116023> PMID: 35682702
50. Kellenberger E, Rodrigo J, Muller P, Rognan D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins: Structure, Function, and Bioinformatics*. 2004; 57(2):225–42. <https://doi.org/10.1002/prot.20149> PMID: 15340911
51. Cheng T, Li X, Li Y, Liu Z, Wang R. Comparative assessment of scoring functions on a diverse test set. *J Chem Inf Model*. 2009; 49(4):1079–93. <https://doi.org/10.1021/ci9000053> PMID: 19358517
52. Wang Z, Sun H, Yao X, Li D, Xu L, Li Y, et al. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys Chem Chem Phys*. 2016; 18(18):12964–75. <https://doi.org/10.1039/c6cp01555g> PMID: 27108770
53. Reddy KK, Rathore RS, Srujana P, Burri RR, Reddy CR, Sumakanth M, et al. Performance Evaluation of Docking Programs- Glide, GOLD, AutoDock & SurflexDock, Using Free Energy Perturbation Reference Data: A Case Study of Fructose-1, 6-bisphosphatase-AMP Analogs. *Mini Rev Med Chem*. 2020; 20(12):1179–87. <https://doi.org/10.2174/1389557520666200526183353> PMID: 32459606
54. Leelananda SP, Lindert S. Computational methods in drug discovery. *Beilstein Journal of Organic Chemistry*. 2016; 12:2694–718. <https://doi.org/10.3762/bjoc.12.267> PMID: 28144341

55. Ahire V, Das D, Mishra KP, Kulkarni GR, Ackland L. Inhibition of the p53 Y220C Mutant by 1-Hydroxy-2-Methylanthraquinone Derivatives: A Novel Strategy for Cancer Therapy. *J Environ Pathol Toxicol Oncol.* 2016; 35(4):355–64. <https://doi.org/10.1615/JEnvironPatholToxicolOncol.2016012256> PMID: 27992315
56. Arba M, Ihsan S, Ramadhan OA, Tjahjono DH. In silico study of porphyrin-anthraquinone hybrids as CDK2 inhibitor. *Comput Biol Chem.* 2017; 67:9–14. <https://doi.org/10.1016/j.combiolchem.2016.12.005> PMID: 28024230
57. Asnawi A, Aman LO, Nursamsiar, Yuliantini A, Febrina E. MOLECULAR DOCKING AND MOLECULAR DYNAMIC STUDIES: SCREENING PHYTOCHEMICALS OF *Acalypha indica* AGAINST BRAF KINASE RECEPTORS FOR POTENTIAL USE IN MELANOCYTIC TUMOURS. *Rasayan Journal of Chemistry.* 2022; 15(2):1352–61.
58. Bai F, Liu H, Tong L, Zhou W, Liu L, Zhao Z, et al. Discovery of novel selective inhibitors for EGFR-T790M/L858R. *Bioorganic and Medicinal Chemistry Letters.* 2012; 22(3):1365–70. <https://doi.org/10.1016/j.bmcl.2011.12.067> PMID: 22227214
59. Ciacò S, Mazzoleni V, Javed A, Eiler S, Ruff M, Mousli M, et al. Inhibitors of UHRF1 base flipping activity showing cytotoxicity against cancer cells. *Bioorg Chem.* 2023; 137:106616. <https://doi.org/10.1016/j.bioorg.2023.106616> PMID: 37247564
60. Choo Wongkamon K, Sawatdichaikul O, Songtawee N, Limtrakul J. Receptor-based virtual screening of EGFR kinase inhibitors from the NCI diversity database. *Molecules.* 2010; 15(6):4041–54. <https://doi.org/10.3390/molecules15064041> PMID: 20657425
61. Cozza G, Gianoncelli A, Montopoli M, Caparrotta L, Venerando A, Meggio F, et al. Identification of novel protein kinase CK1 delta (CK1delta) inhibitors through structure-based virtual screening. *Bioorg Med Chem Lett.* 2008; 18(20):5672–5. <https://doi.org/10.1016/j.bmcl.2008.08.072> PMID: 18799313
62. Cozza G, Mazzorana M, Papinutto E, Bain J, Elliott M, di Maira G, et al. Quinalizarin as a potent, selective and cell-permeable inhibitor of protein kinase CK2. *Biochemical Journal.* 2009; 421(3):387–95. <https://doi.org/10.1042/BJ20090069> PMID: 19432557
63. Cozza G, Venerando A, Sarno S, Pinna LA. The Selectivity of CK2 Inhibitor Quinalizarin: A Reevaluation. *BioMed Research International.* 2015; 2015. <https://doi.org/10.1155/2015/734127> PMID: 26558278
64. Das A, Prajapati A, Karna A, Sharma HK, Uppal S, Lather V, et al. Structure-based virtual screening of chemical libraries as potential MELK inhibitors and their therapeutic evaluation against breast cancer. *Chem Biol Interact.* 2023; 376:110443. <https://doi.org/10.1016/j.cbi.2023.110443> PMID: 36893906
65. DemiRezer LO, Uzun M, Ozanver N, Guvenalp Z. Determination of phytoestrogenic potential of anthranoids by molecular docking studies. *Fabad Journal of Pharmaceutical Sciences.* 2018; 43(3):129–41.
66. Dong L, Shen S, Chen W, Xu D, Yang Q, Lu H, et al. Discovery of Novel Inhibitors Targeting Human O-GlcNAcase: Docking-Based Virtual Screening, Biological Evaluation, Structural Modification, and Molecular Dynamics Simulation. *J Chem Inf Model.* 2019; 59(10):4374–82. <https://doi.org/10.1021/acs.jcim.9b00479> PMID: 31487462
67. Gao X, Zhang W, Jia Y, Xu H, Zhu Y, Pei X. Identification of a prognosis-related ceRNA network in cholangiocarcinoma and potentially therapeutic molecules using a bioinformatic approach and molecular docking. *Sci Rep.* 2022; 12(1):16247. <https://doi.org/10.1038/s41598-022-20362-w> PMID: 36171401
68. Guan Y, Jiang S, Ye W, Ren X, Wang X, Zhang Y, et al. Combined treatment of mitoxantrone sensitizes breast cancer cells to rapalogs through blocking eEF-2K-mediated activation of Akt and autop-hagy. *Cell Death Dis.* 2020; 11(11):948. <https://doi.org/10.1038/s41419-020-03153-x> PMID: 33144562
69. Golubovskaya VM, Ho B, Zheng M, Magis A, Ostrov D, Cance WG. Mitoxantrone targets the ATP-binding site of FAK, binds the FAK kinase domain and decreases FAK, Pyk-2, c-Src, and IGF-1R in vitro kinase activities. *Anti-Cancer Agents in Medicinal Chemistry.* 2013; 13(4):546–54. <https://doi.org/10.2174/1871520611313040003> PMID: 22292772
70. Jordheim LP, Marton Z, Rhimi M, Cros-Perrial E, Lionne C, Peyrottes S, et al. Identification and characterization of inhibitors of cytoplasmic 5'-nucleotidase cN-II issued from virtual screening. *Biochem Pharmacol.* 2013; 85(4):497–506. <https://doi.org/10.1016/j.bcp.2012.11.024> PMID: 23220537
71. Khan MS, Mehmood B, Yousaf Q, Bibi S, Fazal S, Saleem S, et al. Molecular Docking Studies Reveal Rhein from rhubarb (*Rheum rhabarbarum*) as a Putative Inhibitor of ATP-binding Cassette Super-family G member 2. *Med Chem.* 2021; 17(3):273–88. <https://doi.org/10.2174/1573406416666191219143232> PMID: 31854281

72. Lankapalli AR, Kannabiran K. Interaction of marine *Streptomyces* compounds with selected cancer drug target proteins by *in silico* molecular docking studies. *Interdiscip Sci.* 2013; 5(1):37–44. <https://doi.org/10.1007/s12539-013-0146-0> PMID: 23605638

73. Leggett CS, Doll MA, Salazar-González RA, Habil MR, Trent JO, Hein DW. Identification and characterization of potent, selective, and efficacious inhibitors of human arylamine N-acetyltransferase 1. *Arch Toxicol.* 2022; 96(2):511–24. <https://doi.org/10.1007/s00204-021-03194-x> PMID: 34783865

74. Liu R, Liu J, Wang S, Wang Y, Zhang T, Liu Y, et al. Combined treatment with emodin and a telomerase inhibitor induces significant telomere damage/dysfunction and cell death. *Cell Death Dis.* 2019; 10(7):527. <https://doi.org/10.1038/s41419-019-1768-x> PMID: 31296842

75. Mhatre B, Gupta PP, Marar T. Evaluation of drug candidature of some anthraquinones from *Morinda citrifolia* L. as inhibitor of human dihydrofolate reductase enzyme: Molecular docking and *in silico* studies. *Computational Toxicology.* 2017; 1:33–8.

76. Nag A, Verma P, Paul S, Kundu R. *In Silico* Analysis of the Apoptotic and HPV Inhibitory Roles of Some Selected Phytochemicals Detected from the Rhizomes of Greater Cardamom. *Applied Biochemistry and Biotechnology.* 2022; 194(10):4867–91. <https://doi.org/10.1007/s12010-022-04006-3> PMID: 35670907

77. Obounchoey P, Tabtimmai L, Suphakun P, Thongkhao K, Eurtivong C, Gleeson MP, et al. *In silico* identification and *in vitro* validation of nogalamycin N-oxide (NSC116555) as a potent anticancer compound against non–small-cell lung cancer cells. *Journal of Cellular Biochemistry.* 2019; 120(3):3353–61. <https://doi.org/10.1002/jcb.27605> PMID: 30324706

78. Rinne M, Mätilik K, Ahonen T, Vedovi F, Zappia G, Moreira VM, et al. Mitoxantrone, pixantrone and mitoxantrone (2-hydroxyethyl)piperazine are toll-like receptor 4 antagonists, inhibit NF-κB activation, and decrease TNF-alpha secretion in primary microglia. *Eur J Pharm Sci.* 2020; 154:105493.

79. Roy S, Ali A, Bhattacharya S. Theoretical Insight into the Library Screening Approach for Binding of Intermolecular G-Quadruplex RNA and Small Molecules through Docking and Molecular Dynamics Simulation Studies. *Journal of Physical Chemistry B.* 2021; 125(21):5489–501. <https://doi.org/10.1021/acs.jpcb.0c10991> PMID: 34029082

80. Singh M, Malhotra L, Haque MA, Kumar M, Tikhomirov A, Litvinova V, et al. Heteroarene-fused anthraquinone derivatives as potential modulators for human aurora kinase B. *Biochimie.* 2021; 182:152–65. <https://doi.org/10.1016/j.biochi.2020.12.024> PMID: 33417980

81. Song Y, Lim J, Seo YH. A novel class of anthraquinone-based HDAC6 inhibitors. *Eur J Med Chem.* 2019; 164:263–72. <https://doi.org/10.1016/j.ejmech.2018.12.056> PMID: 30597327

82. Taherkhani A, Moradkhani S, Orangi A, Jalalvand A, Khamverdi Z. *In silico* study of some natural anthraquinones on matrix metalloproteinase inhibition. *Research Journal of Pharmacognosy.* 2021; 8(4):37–51.

83. Wang YW, Guo YF, Qiang SJ, Jin RY, Li Z, Tang YP, et al. 3D-QSAR, Molecular Docking, and MD Simulations of Anthraquinone Derivatives as PGAM1 Inhibitors. *Frontiers in Pharmacology.* 2021;12. <https://doi.org/10.3389/fphar.2021.764351> PMID: 34899321

84. Wu KJ, Zhong HJ, Li G, Liu C, Wang HD, Ma DL, et al. Structure-based identification of a NEDD8-activating enzyme inhibitor via drug repurposing. *Eur J Med Chem.* 2018; 143:1021–7. <https://doi.org/10.1016/j.ejmech.2017.11.101> PMID: 29232579

85. Wu FL, Chu PY, Chen GY, Wang K, Hsu WY, Ahmed A, et al. Natural anthraquinone compound emodin as a novel inhibitor of aurora A kinase: A pilot study. *Chem Biol Drug Des.* 2022; 99(1):126–35. <https://doi.org/10.1111/cbdd.13938> PMID: 34411446

86. Zaayter L, Mori M, Ahmad T, Ashraf W, Boudier C, Kilin V, et al. A Molecular Tool Targeting the Base-Flipping Activity of Human UHRF1. *Chemistry—A European Journal.* 2019; 25(58):13363–75. <https://doi.org/10.1002/chem.201902605> PMID: 31322780

87. Khan MS, Mehmood B, Yousafi Q, Bibi S, Fazal S, Saleem S, et al. Molecular docking studies reveal rhein from rhubarb (*Rheum rhabarba-rum*) as a putative inhibitor of atp-binding cassette super family g member 2. *Medicinal Chemistry.* 2021; 17(3):273–88. <https://doi.org/10.2174/1573406416666191219143232> PMID: 31854281

88. Pushpakom S, Iorio F, Evers PA, Escott KJ, Hopper S, Wells A, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov.* 2019; 18(1):41–58. <https://doi.org/10.1038/nrd.2018.168> PMID: 30310233

89. Song Y, Lim J, Seo YH. A novel class of anthraquinone-based HDAC6 inhibitors. *European Journal of Medicinal Chemistry.* 2019; 164:263–72. <https://doi.org/10.1016/j.ejmech.2018.12.056> PMID: 30597327

90. Moshawih S, Goh HP, Kifli N, Darwesh MAE, Ardianto C, Goh KW, et al. Identification and optimization of TDP1 inhibitors from anthraquinone and chalcone derivatives: consensus scoring virtual screening

and molecular simulations. *J Biomol Struct Dyn.* 2023;1–25. <https://doi.org/10.1080/07391102.2023.2256870> PMID: 37697727

91. Kumar A, Zhang KYJ. Advances in the Development of Shape Similarity Methods and Their Application in Drug Discovery. *Frontiers in Chemistry.* 2018; 6. <https://doi.org/10.3389/fchem.2018.00315> PMID: 30090808
92. Liu X, Jiang H, Li H. SHAFTS: a hybrid approach for 3D molecular similarity calculation. 1. Method and assessment of virtual screening. *J Chem Inf Model.* 2011; 51(9):2372–85. <https://doi.org/10.1021/ci200060s> PMID: 21819157
93. Lavecchia A, Di Giovanni C. Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem.* 2013; 20(23):2839–60. <https://doi.org/10.2174/0929867311320999001> PMID: 23651302
94. Shaker B, Ahmad S, Lee J, Jung C, Na D. In silico methods and tools for drug discovery. *Computers in Biology and Medicine.* 2021; 137:104851. <https://doi.org/10.1016/j.combiomed.2021.104851> PMID: 34520990
95. Böhm HJ, Flohr A, Stahl M. Scaffold hopping. *Drug Discov Today Technol.* 2004; 1(3):217–24. <https://doi.org/10.1016/j.ddtec.2004.10.009> PMID: 24981488
96. Sun H, Tawa G, Wallqvist A. Classification of scaffold-hopping approaches. *Drug Discov Today.* 2012; 17(7–8):310–24. <https://doi.org/10.1016/j.drudis.2011.10.024> PMID: 22056715
97. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2001; 46(1–3):3–26. [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0) PMID: 11259830
98. Fang S, Dong L, Liu L, Guo J, Zhao L, Zhang J, et al. HERB: a high-throughput experiment- and reference-guided database of traditional Chinese medicine. *Nucleic Acids Research.* 2020; 49(D1): D1197–D206.
99. Chen CY. TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening in silico. *PLoS One.* 2011; 6(1):e15939. <https://doi.org/10.1371/journal.pone.0015939> PMID: 21253603
100. Zhao H, Dietrich J. Privileged scaffolds in lead generation. *Expert Opinion on Drug Discovery.* 2015; 10(7):781–90. <https://doi.org/10.1517/17460441.2015.1041496> PMID: 25959748
101. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; 144(5):646–74. <https://doi.org/10.1016/j.cell.2011.02.013> PMID: 21376230
102. Yee AA, Savchenko A, Ignachenko A, Lukin J, Xu X, Skarina T, et al. NMR and X-ray Crystallography, Complementary Tools in Structural Proteomics of Small Proteins. *Journal of the American Chemical Society.* 2005; 127(47):16512–7. <https://doi.org/10.1021/ja053565+> PMID: 16305238
103. Monks T, Currie CSM, Onggo BS, Robinson S, Kunc M, Taylor SJE. Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. *Journal of Simulation.* 2019; 13(1):55–67.