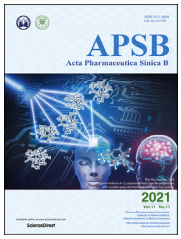
Acta Pharmaceutica Sinica B 2021;11(11):3417−3432

Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com



REVIEW

Gains from no real PAINS: Where 'Fair Trial Strategy' stands in the development of multi-target ligands

Jianbo Sun[a,†], **Hui Zhong**[b,†], **Kun Wang**[a], **Na Li**[a,*], **Li Chen**[a,*]

[a]*State Key Laboratory of Natural Medicines, Department of Natural Medicinal Chemistry, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing 210009, China*
[b]*Department of Pharmacology of Traditional Chinese Medicine, School of Traditional Chinese Pharmacy, China Pharmaceutical University, Nanjing 210009, China*

Received 24 December 2020; received in revised form 15 February 2021; accepted 25 February 2021

KEYWORDS

Multitarget-directed ligands;
PAINS suspects;
In silico filtering;
Biochemical experiment;
Fair trial strategy

Abstract Compounds that selectively modulate multiple targets can provide clinical benefits and are an alternative to traditional highly selective agents for unique targets. High-throughput screening (HTS) for multitarget-directed ligands (MTDLs) using approved drugs, and fragment-based drug design has become a regular strategy to achieve an ideal multitarget combination. However, the unexpected presence of pan-assay interference compounds (PAINS) suspects in the development of MTDLs frequently results in nonspecific interactions or other undesirable effects leading to artefacts or false-positive data of biological assays. Publicly available filters can help to identify PAINS suspects; however, these filters cannot comprehensively conclude whether these suspects are "bad" or innocent. Additionally, these *in silico* approaches may inappropriately label a ligand as PAINS. More than 80% of the initial hits can be identified as PAINS by the filters if appropriate biochemical tests are not used resulting in false positive data that are unacceptable for medicinal chemists in manuscript peer review and future studies. Therefore, extensive offline experiments should be used after online filtering to discriminate "bad" PAINS and avoid incorrect evaluation of good scaffolds. We suggest that the use of "Fair Trial Strategy" to identify interesting

Abbreviations: AD, Alzheimer disease; ALARM NMR, a La assay to detect reactive molecules by nuclear magnetic resonance; CADD, computer-aided drug design technology; CoA, coenzyme A; EGFR, epidermal growth factor receptor; GSH, glutathione; HER2, human epidermal growth factor receptor 2; HTS, high-throughput screening; LC−MS, liquid chromatography−mass spectrometry; MTDLs, multitarget-directed ligands; QSAR, quantitative structure−activity relationship; PAINS, pan-assay interference compounds; ROS, radicals and oxygen reactive species.

*Corresponding authors. Tel.: +86 25 83271447.
E-mail addresses: 409091029@qq.com (Na Li), chenli627@cpu.edu.cn (Li Chen).
†These authors made equal contributions to this work.

Peer review under responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

https://doi.org/10.1016/j.apsb.2021.02.023

2211-3835 © 2021 Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

molecules in PAINS suspects to provide certain structure–function insight in MTDL development.

© 2021 Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The advantages of multitarget-directed ligands (MTDLs) suggest that agents capable of modulating multiple targets in a selective manner can improve the balance between clinical and therapeutic benefit and safety compared to the characteristics of unique target-directed agents[1,2]. The types of scaffolds needed to design MTDLs require accurate analysis of target−disease associations, pathway−target−drug−disease relationships, and adverse event profiling[3]. Therefore, the application of computer-aided drug design technology (CADD), advanced systems biology, and chemical biology is shifting MTDL development paradigm from low affinity inhibition of multiple-targets to an approach involving interactions with mutually regulated targets to achieve synergistic and detoxifying effects[4]. However, meaningful interference of MTDLs with multiple targets has been flagged as undesirable for a long time due to questionable rationality and selectivity of MTDLs, which are considered "bad" scaffold suspects (known as pan-assay interference compounds, PAINS)[5]. The instances of "bad" scaffolds are often disguised as a drug combination and may be involved in nonspecific interactions, leading to artefacts in biological assays[6]. For example, arylpiperazine substructure has been considered a suitable scaffold for fine balancing of D2, 5-HT1A and 5-HT2A receptor activities to improve antipsychotic efficacy or mitigate adverse effects[7]. The prototype of the third-generation of antipsychotics, aripiprazole, launched into the market in 2015 was the first designed serotonin-dopamine activity modulator. However, the compound also displayed unwanted side effects probably due to sustained interaction as a false positive structure with post-synaptic D2 receptors[7,8].

Exemplary PAINS recognized as small molecule suspects include anilines, rhodanines, curcuminoids, Michael acceptors as irreversible inhibitors, and Mannich bases, which are typically present as substructures in other molecules[9−11]. PAINS are usually attributed to nonspecific binding or assay artefacts based on the structure or binding interactions or to indirect potential functionalization effects, such as sample fluorescence, which is the most difficult issue in biochemical assays[12,13]. To improve the affinity of PAINS toward targets, medicinal chemists spend considerable time and effort to produce various analogs and optimize the activity of PAINS[11]. In contrast, excessive screening and unjustified assessment may often result in abandonment of "good" scaffolds due to suspected PAINS. These considerations may confuse peer reviewers in their assessment of the submitted manuscripts and places many inexperienced chemists in a complex situation due to required activity evaluation and fragment selection in the developments of MTDLs[5]. Thus, these suspects compounds should be treated with rigorous and investigative "Fair Trail Strategy" to refrain from advancing a "bad" PAINS or discarding a "good" scaffold[11,14,15]. It is appropriate and relevant to screen selective scaffolds that combine specified multi-target properties according to the paradigm of targeted pharmacology and "network pharmacology"[3]. In case of a concern about PAINS suspects, detailed and reasonable follow-up experiments are essential to exculpate the innocent PAINS suspects to enable their development and to validate the expected functions of the "bad" suspects prior to discarding these compounds from further consideration[16]. Publications should allow authors to quantify the roles of each functional group of the PAINS suspects in their submitted manuscript rather than firmly reject the candidates due to biased perception.

2. Performance of PAINS suspects in MTDL development

The research and development strategies of MTDLs have been extensively focus on high-throughput screening (HTS), fragment based synthetic approaches, and physicochemical, pharmacodynamic, and network pharmacology aspects[17]. PAINS alerts are obviously more frequent for the combination of two or more specifically selected targets than those for a single target regardless of the approach used to develop MTDLs. The false-positive results caused by real PAINS are predominantly caused by a variety of mechanisms including the formation of colloidal aggregates (**1**, Fig. 1)[18], chelation (**2**, Fig. 1)[19], covalent protein reactivity (**3**, Fig. 1)[20], interference with assay spectroscopy (**4**, Fig. 1)[21,22], redox activity (**5**, Fig. 1)[23], membrane disruption[24], high molecular flexibility and hydrophobicity[25], decomposition in buffers[26], and photoactivation[27]. The first five mechanisms are considered to be the most important factors in the assignment of PAINS (Table 1[6,18−21,23]). Hence, there is no single diagnostic analysis system for the entire list of "bad" compounds[5].

Most of MTDLs obtained from HTS have relatively small molecular weight[28]. Introduction of appropriate fragments or removal of unnecessary groups from the favorable position is completely determined by the target and can be used to balance the effect of multi-target combination therapy. Therefore, general screening compound decks can be compiled based on the scaffolds with low structural complexity to increase the number of multi-target combinations and facilitate rapid hit-to-lead expansion[12]. The strategy of fragment recombination may be more popular and desirable for the customized design of MTDLs. Regardless of the methods used to obtain the ligands, the structural environment of potential scaffolds plays a critical role in the development of MTDLs due to their potential to be flagged as PAINS suspects[29]. Multiple evidence provided by the analysis of analog series containing suspected artefacts enables the detection of PAINS in various scaffolds[30]. In addition to PAINS, a number of other agents may have interference potential[31]. Similarly, in the case of certain highly promiscuous scaffolds, there is no direct evidence demonstrating possible assay artefacts[31]. Relationships between lack of activity, specific activity, and artificial performance of the agents with potential liabilities are highly confusing and difficult to distinguish[32]. Hence, comprehensive investigation of

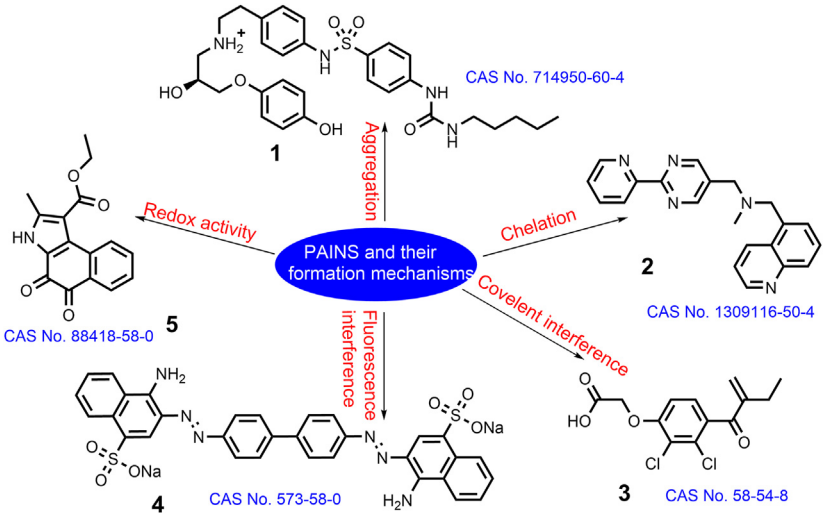


Figure 1 Examples of PAINS suspects and their activity modes and action profiles.

multitarget activities and assay interference patterns in HTS and assessment of fragment recombination are the major tasks in MTDL development (Fig. 2)[14,33].

2.1. *Performance of HTS in MTDL discovery and PAINS identification*

HTS has become a regular strategy in the discovery of screening hits as a good starting point for the growing list of validated multi-targets or as a useful tool for MTDL discovery. Artificial intelligence, such as neural networks, is critical for prediction of interactions between the screening hits and related targets to prioritize the ligands that can be further developed as MTDLs or to provide analysis of interactions with multiple targets[34]. The issues surrounding HTS received considerable attention because some of the screening hits are truly promiscuous ligands that potently, specifically, and reversibly bind to a multi-target combination related to a complex disease[35,36]. However, detection of desired and undesired ligands by HTS identifies certain scaffolds that do not necessarily reflect the binding promiscuity of the individual molecules. In contrast, new conceptual frameworks, such as systems biology and polypharmacology, may suggest that molecules, which often produce biological effects in multiple tests, do not interfere with bioassays; optimization of these compounds is challenging but can yield innovative and safe drugs[15].

The use of HTS for exclusion of PAINS suspects from chemical libraries is another controversial issue. The applications of public PAINS filters and other HTS methods are important for identification of undesired scaffolds; however, these methods are frequently insufficiently developed and are not completely reliable in identification of all possible relationships between the scaffolds and protein structures[5]. Up to 80%−100% of initial HTS hits in various screening models and various target activators/inhibitors can be labeled as artefacts if appropriate control experiments are not employed[37−40]. The activity of the majority of PAINS is positive only in a small number of assays[14] and depends on specific experimental conditions and modes of action[5]. Usually, PAINS are the compounds with a high hit rate, which interfere with the screening and detection methods, and poor drug properties of PAINS are an important issue in drug development[35]. The most frequent "bad" PAINS identified by HTS are rhodanine and its derivatives, which have been extensively reported to have a wide range of antibacterial activities at the cellular and compound level in more than 2000 publications[41,42]. However, due to extensive and complex mechanism of action, specific targets of

Table 1 Main chemotypes of PAINS and their mechanisms to cause promiscuity.

Interference	Principle	Chemotypes
Covalent interaction	Covalently bind to all sorts of macromolecules	Quinones, alkylidene barbiturates, rhodanines, omeprazole, carbidopa, ethacrynic acid, enones, related heterocycle [6,20]
Colloidal aggregation	Non-specifically bind to proteins, confounding and irrelevant enzymatic responses	Miconazole, nicardipine, trifluralin, cinnarizine, tetraiodophenolphthalein, staurosporine aglycone[18]
Redox cycling	Generate ROS and indirectly inhibit the catalytic activity of proteins	Phenol-sulphonamides, pyrimidotriazinediones, β-lapachone, arylsulfonamides, tolyl hydrazides, quinones and catechols [6,23]
Ion chelation	Can form chelate for a lot of potential proteins and functional systems	Hydroxyphenyl hydrazones, quinones and catechols, rhodanines, 2-hydroxybenzylamine [6,19]
Sample fluorescence	Fluorophoric properties can affect the biological evaluation results	Daunomycin, topotecan, and riboflavin, quinoxalin-imidazolium substructures[21]

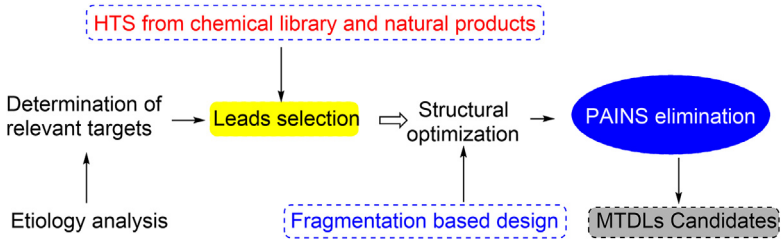


Figure 2 Rational design of multitarget-directed ligands.

these compounds cannot be determined, which becomes the main obstacle for further development[32].

2.2. *Performance of PAINS suspects in fragmentation-based MTDL design*

Fragmentation-based MTDL design generally requires a certain degree of synergy and balance between the optimized promiscuous ligands and the targets due to the mutual interactions of the targets in the clinical treatment of multiple cause diseases. The theoretical premise of fragmentation-based MTDL design is to determine the important role of related targets in the pathological process[43]. Then, rational structural optimization based on a pharmacophore skeleton can yield effective ligands for multiple targets. To achieve a better synergistic effect, the targets of ligand binding should be confirmed, and then suitable scaffolds need to be introduced to increase the affinity to the target combination[5]. A comprehensive analysis of ligand−target interactions indicates that multifamily ligands frequently result in multiple interactions in the binding sites even if the compounds have homologous bound conformations; these properties rationalize promiscuous binding events at the molecular level[44].

Generally, the mechanisms of action of drugs and targets involve regulation of the functions of the target proteins through noncovalent interactions. However, PAINS scaffolds have relatively high chemical activities and can be covalently combined with proteins, DNA, etc.[45]. Proteins are often very sensitive to the PAINS scaffolds containing electrophilic groups, and these scaffolds can influence enzyme activity through irreversible reactions with proteins, which in turn cause false positives screening results. Additionally, PAINS scaffolds containing electrophilic groups are easily hydrolyzed or decomposed by organic solvents to generate active fragments, which act on targets and change enzyme activity[35,46]. On the other hand, certain PAINS scaffolds, such as aromatic hydrocarbons, polyphenols, and hyperlipophilic and conjugated scaffolds, can form aggregates due to molecular interactions and can bind to the targets with high affinity in the reaction systems thus causing false positive results[35].

3. Exculpation of cunning PAINS suspects

As mentioned above, several mechanisms have been proposed to account for PAINS interference in MTDL development, and it is impossible to distinguish "bad" compounds only by an independent assessment system. The advanced chemical databases of PAINS filtering are currently available with annotated biological activities, and several successful applications of PAINS filtering were described. For example, the application of the Bioassay and Drug Screening Platform managed by LaBECFar-FIOCRUZ can routinely and rapidly identify PAINS with selectivity screens[31]. However, current public PAINS filters are not sufficient to identify all possible relationships between chemical and multiple targets as indicated by various types of false negatives detected in the analysis. Certain PAINS can be inadvertently or mechanically distinguished from nonspecific interactions or other undesirable effects leading to detection of irrelevant activities of the suspects (real PAINS)[6]. These "bad" PAINS frequently escape the judgment by peer reviewers and majestically appear in publications, since target reactivity or specificity can be higher than that of the other compounds identified by screening[5].

PAINS are defined in terms of a curated screening library, which is one of the reasons for exculpation of the suspects in filtering. These screening libraries reasonably represent chemical resources of subsequent modifications and were used to eliminate multiple PAINS suspects in many cases by preliminary filtering[11]. For example, dicyanoalkene **1** is a recognized "bad" PAINS; however, public filter ene_cyano_A did not identify **2** (substructure of **1**) as a PAINS because **2** did not include any scaffolds in the initial screening library used to define PAINS[11]. Other reactive scaffolds, such as isothiazolones **3** and pyrimidotriazinediones **5**, cannot be detected by PAINS filters because their "bad" characteristics were identified only after filter definition[11,47] (Fig. 3).

Additionally, compounds with PAINS substructures usually have variable performance in the screening assays, and many substructures are occasionally or consistently inactive[14]. Thus, vulnerability of PAINS filtering may be due to the presence of these inactive PAINS-containing substructures during the detection process (**1**−**4**, Fig. 4)[11]. An increase in the activity threshold of the filtering will result in a lower limit of the hit rates. Therefore, potential false positive rates of PAINS are generally higher than the corresponding anticipated values for the compounds that are expected to produce analytical artefacts[14]. In these cases, multiple assay activities should be monitored to identify PAINS interference.

Another reason for incorrect identification of "bad" PAINS is partially due to tautomerism, such as compounds **1**−**2** and **3**−**4** in Fig. 5; other compounds may cause assay artefacts that are not defined as PAINS[31]. However, the compounds initially defined as PAINS should be considered candidates for subsequent investigation of multitarget activities and molecular mechanisms of polypharmacology[31] to assess the details of "good" promiscuity or "bad" PAINS. Similarly, the suspects not identified by PAINS filters provide opportunities for subsequent investigation of assay interference.

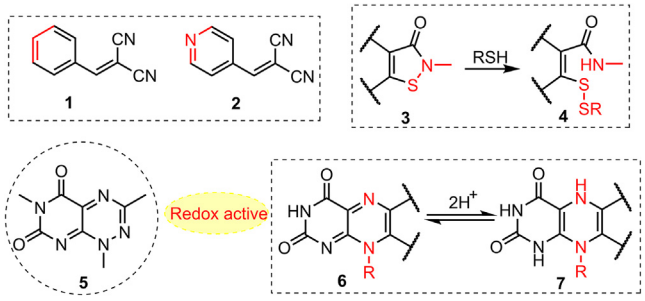


Figure 3 Structures defined in public filters that are not generally recognized as PAINS. Compounds such as **3** and **5** could not be considered useful or progressable and should be excluded from screening libraries. Compound **5** should be redox active and is not unexpected considering its similarity to the isoalloxazine ring of flavins (**6** and **7**)[48,49].

4. Injustice to innocent PAINS suspects

The application of public virtual filters for exclusion of PAINS suspects is controversial, and many drugs approved in the clinic (such as atovaquone **1**, mitomycin **2**, and mitoxantrone **3** in Fig. 6) would never reach the market if medicinal chemists have relied on these filters[10,15,37]. These tools do not comprehensively identify all suspected substructures or promiscuity as PAINS and may also inappropriately label a compound as an artefact[50]. Particularly, some suspects are truly undesired false positives; however, very simple structural modifications can change the potency and selectivity of these compounds. Recent publications suggested that chemical and biological methods identify many screening hits that contain PAINS alerts and directly define PAINS as "bad actors" that exhaust the energy and lower the expectations thus warning medicinal chemists[6]. Public reports of private trials make PAINS suspects to leave a subconscious bad impression on many editors and peer reviewers. Fortunately, this issue has attracted the attention of many editors-in-chief of ACS journals, including *Journal of Medicinal Chemistry*, *ACS Central Science*, and *ACS Chemical Biology*[5]. According to these fair and experienced editors, even if some screening hits can be defined as general PAINS alerts by public filters, the compounds cannot be considered useless in subsequent well-planned studies[5].

Naphthoquinone is one of the PAINS suspects that offers an easily accessible and extensively investigated scaffold for design of new drugs. Naphthoquinones and their derivatives undergo redox cycles *in vivo* generating semiquinone radicals and reactive oxygen species (ROS) that may interfere with multiple targets[51] resulting in their identification as PAINS by several filtering tools. In fact, some naphthoquinones are druggable agents with options for structure optimization by established synthetic pathways[52]. Many antitumor naphthoquinone derivatives are selective toward cancer versus normal cells due to specific metabolism[53]. For example, naphthalene-5,8-dione-1-sulfonamide-based STAT3 inhibitor **9** (Fig. 7) was designed by advanced multiple ligand simultaneous docking (AMLSD) and is characterized by superior druggability compared with other representative inhibitors. Compound **9** selectively binds to p-STAT3(Y705) versus STAT1/5, p-AKT, and p-ERK kinases in the similar pathways and versus other proximal sites. Analysis by the online public filters and biochemical assays identified compound **9** as a PAINS suspect[54]. Quinoline derivatives are also considered to cause PAINS interference in many cases[55]. Neratinib (Fig. 7) is a new irreversible dual-target inhibitor on the market that has antitumor activity and targets EGFR and HER2. Neratinib was designed to bear a Michael acceptor warhead and to take advantage of high sequence identity of the ATP-binding domains of EGFR and HER2 (82%)[56]. However, neratinib displays no activity against other serine−threonine kinases, such as AKT, cyclin D1/CDK4, cyclin E/CDK2, cyclin B1/CDK 1, IKK-2, MK-2, PDK1, c-RAF, and TPL-2 and tyrosine kinase c-MET[57].

5. "Fair Trial Strategy" for validation of PAINS suspects in MTDL development

The development of MTDLs is typically driven by the nature of the targets, availability of the starting points, and chemical tractability. The rational design of MTDLs is far from being an easy task and has to consider the crucial issues of interaction with specific targets[58], achieving a synergistic effect toward them, and "designing out" unwanted interactions with any undesired target while retaining drug-like properties[59]. Discrimination between multitarget activity and assay interference is a major problem in biological screening and medicinal chemistry[5]. However, rationalization and prediction of potential assay interference of PAINS suspects is a complex problem[14,15,31,60]. The importance of PAINS issues is clearly indicated by medicinal chemists; however, PAINS alerts do not necessarily disqualify the agents from further consideration and do not invalidate the available data[31]. The compounds can be characterized as PAINS only after *in silico* filtration is augmented by experimental follow-up; defining a compound as PAINS allows to stop their development and has a greater significance for future studies[61].

The "Fair Trial Strategy" (Fig. 8) encourages assessment that can eliminate unnecessary problems in MTDL developments by avoiding PAINS and should avoid elimination of good compounds and excellent resources for the development of new drugs. In particular, PAINS suspects that have frequent biological effects in diverse assays and do interfere with bioassays can be optimized *via* a challenging process into innovative and safe MTDLs using new conceptual frameworks, such as systems biology and polypharmacology[62]. The verification of molecule selectivity and specificity for disease related multitargets is the key strategy for excluding false positive hits. Covalent protein reactivity, colloidal aggregates, redox activity, and ion chelation are considered the main mechanisms accounting for PAINS. These types of compound interference can result from the compounds themselves[63,64], *e.g.*, in the case of fluorescent compounds. Thus, detailed description of effective computational and experimental validation methods is presented below with case analysis to illustrate the 'Fair Trial Strategy' of these mechanisms.

5.1. Verification of covalently interacting PAINS suspects

Traditionally, promiscuous scaffolds that rely on reactive electrophilic groups are difficult to optimize toward lead compounds because these scaffolds often interfere with biochemical analysis rather than actually alter the activity of the target[65]. Even if the target modulation is real, the undifferentiated target binding response is considered to trigger insurmountable toxic events[65]. These toxic events were increasingly believed to be associated with covalent on-target or off-target binding because irreversible covalent interactions may exacerbate possible adverse reactions[65].

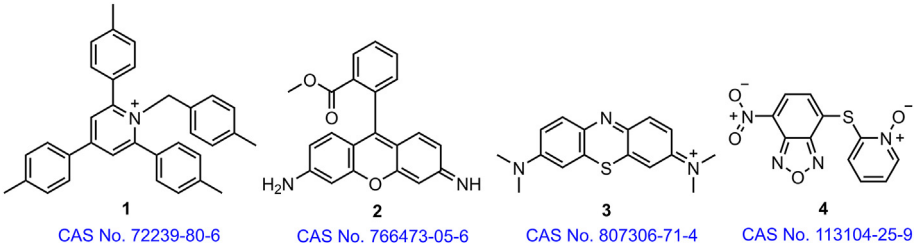


Figure 4 Representative PAINS structures of numerous consistently inactive compounds.

The main consequence of irreversible covalent binding is to covalently bind reactive metabolites to various proteins. In this case, the leads are metabolically bioactivated to form reactive species that potentially bind covalently to all types of macromolecules[65]. Hapten formation is another potential negative consequence that must be considered for covalent and irreversible kinase inhibitors because it can trigger an immune response to the protein adducts[65]. Nonspecific irreversible inhibitors are generally regarded as suspected PAINS by medicinal chemists and considered impurities and typical undesirable products unless special screening of selective covalent modification was originally planned[5]. Therefore, identification of irreversible covalent inhibitors by using specialized libraries and techniques is a necessary means to eliminate PAINS interference in MTDL developments.

5.1.1. Computational validation of non-specific irreversible covalent artefacts

Verification of a compound as a nonspecific irreversible covalent artefact requires complex and innovative methodology; the results have considerable variety in accuracy, precision, and acquired information. In contrast to single-target inhibitors, the development of MTDLs is more complex making traditional drug-like rules and empirical parameters unsuitable for assessment of safety[2]. MTDLs need to have high selectivity against unintended targets and good physical and chemical properties, including absorption, distribution, metabolism, excretion, and toxicity[2]. Rational structural optimization-based computational approaches, such as cheminformatics and virtual screening, pharmacophore development, molecular docking, and molecular dynamics, can yield effective ligands for multiple targets[66]. Efficient MTDL

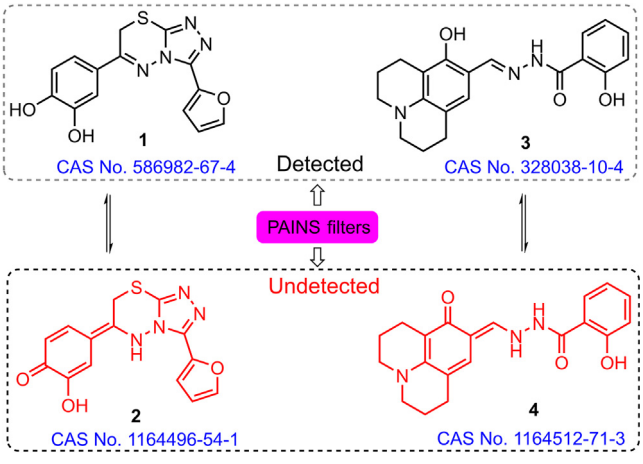


Figure 5 PAINS structures are not recognized by the public filters due to tautomerism.

design uses a number of computational approaches to scaffold optimization, including machine learning techniques of random forests, support vector machines, and Bayesian learning[67]. After acquisition of the feedback activity data from animal experimental models, multitarget QSAR analysis and skeleton transition can be used to optimize MTDLs to adjust their activity[68]. In addition to the developmental relevance of MTDLs, scaffolds with truly multitarget activities may be analyzed by computational methods to determine why and how these chemical entities specifically interact with multiple targets, especially if these targets are only distant relatives or are completely irrelevant and have different functions[69]. On the other hand, computational approaches have some defects and biases in assessment of the structure of PAINS; however, these methods are powerful for preliminary evaluation of the references and enable the use of more accurate and fair virtual approaches of reliable identification. When the designed MTDLs' information data are supplemented in time and more precise algorithms are developed, it is possible to make these calculation methods more precise and accurate for PAINS identification. Some computational approaches to MTDL development and PAINS recognition are discussed below.

5.1.1.1. Molecular docking. Molecular docking studies can be used to confirm nonspecific irreversible covalent interactions between the selected ligands and multiple targets before experimental synthesis[70]. Molecular docking is generally used to access possible conformations of the ligands, estimate the most stable binding modes of the ligands with the targets, and optimize the geometry of the docked ligand–target complexes. The most commonly used docking computational suites in MTDL design include Schrödinger[71], CDOCKER[72], DOCKovalent[73], GOLD[74], and DPubChem[75]. The docking simulations have been successfully used in MTDL development to design the treatments of many neurodegenerative diseases[76,77]. For example, the implementation of Autodock Vina in the design of MTDLs targeting AChE/BuChE/MAO-A/MAO-B for the treatment of AD led to the synthesis of DIH15[78] and DPH14[79], which have the best drug-like characteristics. However, due to rugged free energy landscape of these simulations and minor differences in the initial velocities, studies relying only on the outcome of a single molecular docking method instead of multiple simulations may lead to false positive conclusions[80]. Thus, multiple replicates and subsequent experimental validation are needed to draw conclusions based on molecular docking. Molecular docking is the critical step and a standard component that should be included in any ligand discovery process[80,81].

5.1.1.2. Machine learning techniques. Over the last decade, deep learning and machine learning techniques have been extremely successful and are widely used to develop artificial

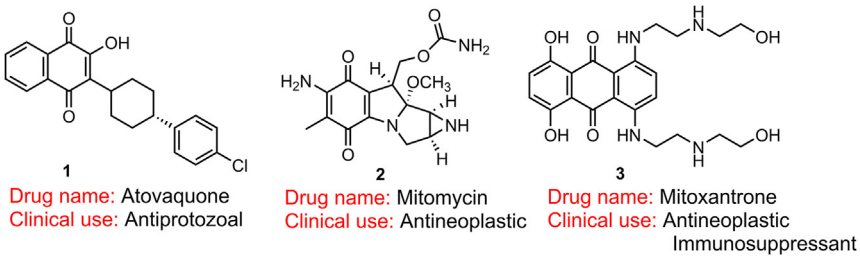


Figure 6 FDA- and worldwide-approved drugs contain PAINS chemotypes.

intelligence in almost every domain[82–86]. Numerous machines learning techniques have been successfully used in model optimization for efficient MTDL design and validation of nonspecific irreversible covalent interaction, including support vector machines, random forests, and Bayesian learning. For example, a screening platform for cloud computing proteomics called "Ligand Express" leverages the combined efficiency of biophysics, biological data, and artificial intelligence technologies[87]. These machine learning methods enable medicinal chemists to efficiently explore new pathways of drug discovery and PAINS identification.

Identification of unexpected drug−protein interactions is important for drug repurposing and PAINS identification. A comprehensive ligand homology modeling approach FINDSITE-comb is used for prediction of drug−protein interactions, human protein targets, and side effects of the drugs[88]. Construction of a library of target protein structures of a proteome enables expedient prediction of interactions of millions of molecules against a typical proteome using a medium-sized computing cluster. The application can be openly accessed on the DR. PRODIS (DRugome, PROteome and DISeasome) webserver at http://cssb.biology.gatech.edu/dr.prodis/ [88].

5.1.1.3. Public filtering tools. Generally, the complexity of the interactions between MTDLs and targets requires considerations of various readout methods used in various assays under various conditions to identify PAINS; thus, definition of PAINS suspects by virtual screening cannot be achieved[5]. However, public filtering tools can be implemented for approximate PAINS identification if all available chemical and biological information is taken into account[5]. Although virtual filtering cannot be 100% successful in ruling out PAINS, it can help medicinal chemists to exclude considerable interference in MTDL development. Unlike experimental biological assays, public filtering tools can analyze massive *in silico* libraries of molecules[59]. Application of filters to rule out PAINS requires an organic combination with subsequent experimental biological assays to improve the success of validated trial probability. Therefore, the improvement of filtering methods generating highly reliable activity data, such as characterization and identification of detailed thiol-reactive chemotypes, are needed for PAINS suspects to pass through the filters[33]. Moreover, experimental approaches should be supplemented with computational methods to accelerate triage of potential PAINS alert, guide screening library design, and prevent follow-up on undesirable chemical entities[5].

Several well-established public filtering tools are available online or as application software, including https://sandbox.ntp.niehs.nih.gov/interferences/, http://zinc15.docking.org/patterns/home, cbligand[5], RDKit[89], ZINC[90], ToxAlerts[48], and FAF-Drugs4 server[91]. However, due to insufficient analysis of PAINS definitions, the existing filtering tools always miss a substantial proportion of PAINS substructure-containing compounds in an independently curated molecular data-set. To address this shortcoming and allow visual exploration of the reasons for the prediction, Maksim Koptelov and coworkers[92] developed PrePeP. PrePeP uses benchmark datasets from the literature to compensate for a number of shortcomings of existing PAINS alerts that have been pointed out recently[92]. Another improved filter system, Vertex's REOS, can identify and remove toxic, reactive, or other undesirable molecules from the database by combining the characteristics of molecules with certain principles that are based on the requirements for the medicinal properties[93]. More than 90% of high hit rate molecules and 91% of nonhigh hit rate molecules could be distinguished by this filtering model[93].

5.1.1.4. BadApple. BadApple is a biological test data associative promiscuity pattern learning algorithm combining general and domain-specific features to assist with and accelerate identification of possible nonspecific irreversible covalent artefacts[13]. This engine generates a score associated with a pragmatic empirical definition of irreversible interferences with the overall goal to identify PAINS and streamline workflows[13]. Unlike methods that rely on experts to manage chemical substructure patterns, BadApple is a completely evidence-driven, automated, and self-improving method that can integrate additional data[13].

5.1.2. Experimental validation of non-specific irreversible covalent artefacts

As mentioned previously, virtual computer technologies can only assist medicinal chemists with the analysis of interaction modes of ligands or PAINS with proteins in targeted or untargeted MTDL design, and their target prediction and PAINS exclusion are not 100% correct. Therefore, the development of rapid and robust experimental methods to distinguish PAINS suspects in the validation of nonspecific irreversible covalent artefacts is particularly important. Irreversible covalent artefacts are usually manifested as confusing phenomena in biochemical assays, such as lack of correlation between structure and activity, time-dependent activity, and steep inhibition curve[5]. To determine whether these artefacts are "bad", conclusive experimental evidence from at least two different biological experimental tests should be provided. If both tests report that the compounds are specifically active and the apparent activity is not artificial, the suspect can be excluded[5]. In the "Fair Trial Strategy", full concentration−response curves are a

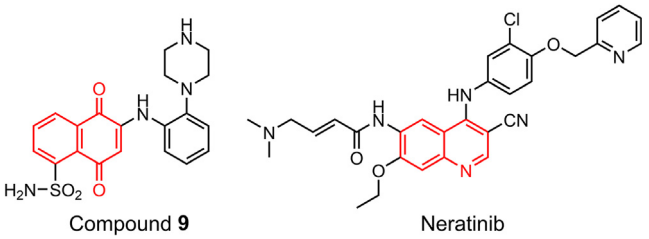


Figure 7 Structures of compound **9** and neratinib.

simple and critical evidence to rule out PAINS suspects because the investigator can obtain considerable important information based on the steepness of the curve and sampling quality[5].

5.1.2.1. Counter-screening validation by incubation time-dependent activity. Covalent inhibitors usually display concentration-dependent and incubation time-dependent activity in *in vitro* enzymatic assays[94]. Generally, a time-dependent increase in apparent inhibitory potency suggests irreversible covalent binding (Table 2)[4]. Counter-screening verification of an irreversible covalent inhibitor THZ531 (Fig. 9a) of CDK12/13 used changes in incubation time during radiometric kinase activity assay (Fig. 9)[95,96]. Counter-screening verification assessed the ability of recombinant CDK12 to phosphorylate the Pol-II-CTD peptide substrate in the presence of the cofactor cyclin K; the data were normalized to the relative [^{32}P] transfer, and DMSO was used as a control (Fig. 9b)[95]. The effect of incubation time on the activity may be caused by CDK12; however, the results still indicate covalent inhibition[95].

5.1.2.2. Washout experiments. In cell-based assays, cells are exposed to the solutions of PAINS suspects, washed, and allowed to grow in the medium without PAINS suspects. The washout experiment is an effective method to verify irreversible mechanism of action (Table 2)[95]. A continuous increase in inhibition in the washout experiments is attributed to irreversible target binding[95]. This method was also used in the verification of the effect of THZ531 in Jurkat T cells (acute lymphoblastic leukemia cells)[95]. In this experiment, the effects of THZ531 were maintained for 72 h after washout, while a negative control of a reversible compound THZ531R had no effect (Fig. 10)[95].

5.1.2.3. X-ray crystallography and in situ labeling followed by LC−MS/MS analysis. Due to the complexity of polypharmacology, the rational design of MTDLs is an equally challenging and attractive research area[97,98]. X-ray structures are used to analyze the covalent binding of MTDLs with target proteins, associate MTDLs with multitarget activities to protein binding site similarity or identify multifamily ligand PAINS suspects bound to targets (Table 2)[44]. Effective and direct methods for identification of covalent binding *in vitro* are X-ray crystallography and *in situ* labeling followed by LC−MS/MS analysis. These strategies were used to validate the covalent binding of THZ531 to CDK12/CDK13 *in vitro*[95]. Mass spectrometry analysis showed the formation of covalent adducts, and a peptide containing the exact site of modification was identified after proteolysis. Then, X-ray crystallography of CDK12-cyclin K bound to THZ531 confirmed irreversible mechanism of action[95].

5.1.2.4. Chemoproteomic approaches. A series of chemoproteomic approaches have been developed to identify specific targeted binding and to eliminate nonspecific covalent interactions. These methods include activity-based protein profiling (ABPP) in combination with MS-based proteomics (Table 2)[99] and stable isotope labeling with amino acids in cell culture or tandem mass tagging (TMT) click chemistry pull-down experiments[100]. Recently, a chemoproteomic approach called CITe-Id has been developed to capture, identify, and quantify dose-dependent covalently bound Cys sites in cell lysates[101]. CITe-Id verification demonstrated that THZ1, a covalent kinase inhibitor initially presumed to specifically bind to CDK7, forms covalent bonds with other proteins, including nonkinase targets[95].

5.1.2.5. ALARM NMR. The most common nonprotein thiol-based reporters may fail to mimic the local environment on the protein surface thus not reflecting the reactivity of the agents with physiologically relevant protein side chains[102]. Therefore, a fast and reliable NMR-based approach ALARM NMR (Table 2)[20] (a La assay to detect reactive molecules by nuclear magnetic resonance) has been developed to identify PAINS suspects, including the compounds that may target oxidized or alkylated protein targets[20]. The detection is based on monitoring of DTT-dependent ^{13}C chemical shift changes of the human La antigen in the presence of PAINS suspects. These chemical shifts can be attributed in an ALARM NMR experiment using a thiol in a potentially more biologically relevant environment that may be different from that used in the experiments with small molecule thiols (such as GSH)[99].

5.2. *Verification of PAINS suspects caused by colloidal aggregation*

Small colloidally aggregating molecules (SCAMs) are the largest causative source of PAINS in MTDL discovery and one of the most common mechanisms for false-positive readouts[5,83]. SCAMs are diverse and well represented in academic and commercial screening decks and even in approved drugs[103]. A mechanistic study demonstrated that SCAMs aggregate *via* phase separation and particle formation at concentrations above a compound-specific critical aggregation concentration (CAC)[5,18,103]. The consequences of colloidal aggregates include nonspecific binding to protein surfaces and partial denaturation of a protein to produce confounding and irrelevant enzymatic responses due to nonspecific molecular recognition mechanisms[103−105]. Tight binding of SCAMs produces partial and local protein denaturation resulting in time-dependent enzyme and protein inhibition[83,106]. The deceptive effects of the aggregates can be extended to cell-based detection systems, where SCAMs may lower anticipated

Figure 8 The principles of investigation of suspected PAINS.

Table 2 Comparison of experimental methods for the verification of irreversible covalent binding.

Assay	Principle	Feature	Readout
Incubation time-dependent activity	Time-dependent increase in inhibition indicates irreversible covalent binding	Easy to operate and low cost	Enzyme inhibition
Washout experiments	The continuous growth of covalent inhibitors is attributed to irreversible covalent binding	Can be used to verify irreversible MOR	Enzyme inhibition
X-Ray crystallography and *in situ* labeling	Can analyze the covalent binding situation with proteins	Provide more detailed binding information	Ligand–targets analysis
Chemoproteomic approaches	Can identify specific targeted binding and eliminate non-specific covalent interaction.	Provide more detailed binding information	Ligand–targets analysis
Alarm NMR	Monitor the DTT-dependent ^{13}C chemical shift changes of the human La antigen	Provide more detailed binding information	Ligand–targets analysis

activities due to reduced cell membrane diffusion thus reducing intracellular concentrations[107]. SCAMs can account for up to 95% of false positives in HTS campaigns, which makes SCAMs the most prevalent cause of erroneous ligand–target associations[83]. Therefore, the pernicious effects of SCAMs in MTDL discovery cripple the readout of target- and cell-based assays; hence, validation of these effects at the early stages of MTDL design is essential[83]. Considerable efforts have been invested in the development of methods for validation of potential SCAMs; however, only limited success has been achieved. Since the formation of the aggregates may be induced by minor changes in concentration, the formation of potential aggregates is difficult to predict based only on physical properties[83].

5.2.1. *Computational prediction of SCAMs*

Considering that SCAM and non-SCAM data span over the range of calculated log*P* and molecular weight values and have significant scaffold diversity, their biophysical properties may be reliably analyzed and recognized by computational approaches. Therefore, machine learning algorithms may play an important role in expedient identification of potentially liable chemical entities[83]. Publicly available aggregator advisors (*e.g.*, http://advisor.bkslab.org and http://advisor.docking.org)[6,108] can identify SCAMs based on the substructure fingerprints of a query entity. Verification results obtained by these public tools are similar to over 12,000 experimentally validated SCAMs[83]. The possibility of false positive query entities is calculated by triaging the cheminformatics-based similarity with known SCAMs, observed affinity range, and calculated log*P*[108]. Another effort to identify SCAMs based on physicochemical properties derived from chemical structure includes support vector machine-recursive feature elimination (SVM-RFE) developed as a combination of several independent studies using orthogonal algorithms[109]. Unlike aggregator advisors, SVM-RFE extends the learned data patterns beyond simple structure equivalence and relies on molecular similarity to assess aggregation propensity[83]. Additionally, the chemical space projection clearly emphasizes that the two computational methods share similar reference data but provide different solutions for identification of multiple SCAMs in various areas[83].

5.2.2. *Experimental detection methods for prediction of SCAMs*

The computational methods can preliminarily identify nuisance artefacts in pharmaceutical libraries; however, these methods regularly generate false positive and false negative results[110]. The reason for this inconsistency is due to fact that the formation of SCAMs is triggered by small changes in concentration; thus, it is difficult to predict potential aggregators strictly based on physical properties[25,110]. Moreover, various pH or buffer conditions may induce aggregation that is no longer related to the target compounds easily leading to false predictions[110]. Therefore, studies on the identification and elimination of SCAMs in biological experiments are necessary. The methods listed below are easy to

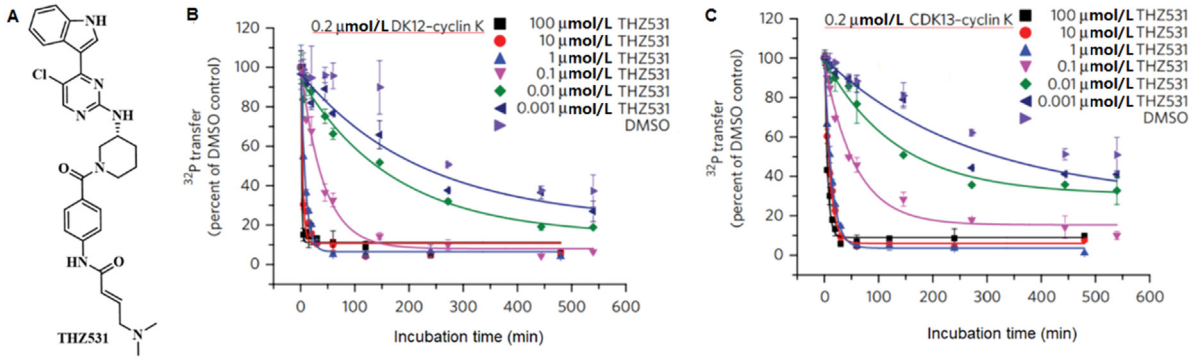


Figure 9 THZ531 (A) inhibition of CDK12/13 is time dependent. *In vitro* kinase activity assay of CDK12-cyclin K (B) and CDK13-cyclin K (C) with different concentrations of THZ531 and varying preincubation times. For all incubation time series, the counts per minute of the kinase activity measurements were normalized to the relative ^{32}P transfer. (B) and (C) Adapted with modification from Ref. 96. Copyright ©2016 Springer Nature.

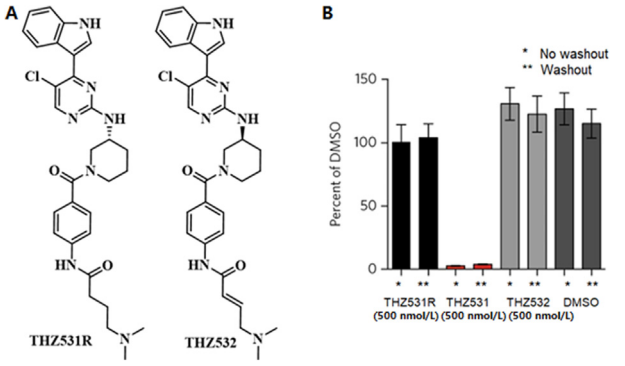


Figure 10 THZ531 retains activity in washout experiment compared with the effects of two negative control compounds. (A) THZ531R and THZ532. (B) Jurkat cells were treated with the indicated compounds for 6 h, inhibitor was washed out and cells were allowed to grow for the remainder of the 72 h. This growth was compared with the growth of cells treated with inhibitors for the full 72 h. (B) Adapted with modification from Ref. 96. Copyright ©2016 Springer Nature.

implement and can provide robust assessment of the aggregation propensity of the investigated molecules.

5.2.2.1. Parallel experiments with nonionic detergent washing. The addition of detergents facilitates the dissociation of the aggregates into monomers in the solution thus decreasing the apparent activity of SCAMs in the presence of detergents in the assay buffers (Table 3)[83]. If the activity of a compound can be attenuated by the addition of small amounts of a nonionic detergent, then the compound can be defined as SCAM[5]. A typical protocol involves the addition of 0.01% Triton X-100 in biochemical assays or 0.025% Tween-80 as a gentler detergent for cell-based assays[5]. Parallel screening with and without a nonionic detergent and reliance on model enzymes (such as β-lactamase) are more rigorous methods for identification of SCAMs[83]. Parallel experiments require preliminary confirmation that the detergent does not contribute to the readouts. Additionally, the changes in activity obtained by two assays should be statistically evaluated to ensure the results are not caused by experimental procedures[83].

5.2.2.2. Centrifugation before cell-based assay. A convenient procedure before cell-based assay involves centrifugation of the solutions of the compounds to precipitate colloidal aggregates. Parallel activity evaluation can be subsequently performed. If the activity is reduced after centrifugation, the compound is likely a SCAM (Table 3)[5]. The key point of this parallel experiment is to prove that the active concentration of a SCAM is lower than the concentration that produces cytotoxicity thus demonstrating that this apparent activity is not due to cytotoxicity. If the goal of the assay is assessment of cytotoxicity in tumor therapy, the selectivity of non-SCAMs in tumor cells should be higher than that in normal cells[5].

5.2.2.3. Competitive ligand–target binding experiment. The false positive SCAMs are often manifested by enzyme inhibition *via* nonspecific aggregation-type binding mechanisms[111]. Surface plasmon resonance (SPR) is a label-free technique that is often used to study the ligand−target binding kinetics and affinity constants of the active compounds. A powerful Biacore technology of SPR can contribute to the identification of potential SCAMs in the early drug discovery by observing and analyzing the ligand−target binding and interaction patterns (Table 3)[83,111]. The operation of this technology is relatively simple and requires only an injection of the dissolved compounds into the protein surface. The compounds are attached to the proteins immobilized on the surface of an optical sensor (association phase), and the binding data are recorded in real time based on the changes in the properties of the sensor surface[111]. This technology allows to evaluate the association and dissociation of SCAMs with proteins in real time to determine how SCAMs interact with proteins in detail[111]. SPR data can provide information about stoichiometry, reversibility, and changes in the properties of the compounds within a certain concentration range[111]. Based on this information, SPR can be used to analyze SCAMs in biochemical assays to expediently determine whether the interference is caused by specific or nonspecific interactions with the proteins. Thus, hits can be selected and prioritized for chemical processing by analytical and/or medicinal chemistry methods[111]. For specific experimental operation, please refer to Ref. 111.

The photonic crystal (PC) optical biosensor aggregation assay is another label-free biophysical technology that has been productively used to detect SCAMs by providing direct quantitative measurement of the mass density of the substances adsorbed on the transducer surface[107,112]. The PC sensors contain a subwavelength periodic surface that reflects a narrow wavelength band after illumination by a broadband collimated light source. Formation of the aggregates changes the sensor surface and adjusts the refractive index of the material to enable expedient identification of SCAMs[107,112]. For specific experimental operation, please refer to Ref. 112.

5.2.2.4. Printable hydrogel microarray. The inhibitory effects of SCAMs are usually due to the adsorption of a protein on the surface of the aggregates thus separating the enzyme from the substrate and resulting in partial denaturation of the protein[110]. The hydrogel-based enzyme immobilization platform provides a system for specific identification and elimination of SCAMs (Table 3)[110]. The adjustable porosity of the hydrogels allows selective transport of the substrates to and from the entrapped enzyme due to size selectivity potentially preventing drug aggregates from reaching the binding site on the enzyme and decreasing the effects associated with promiscuous inhibition[110]. The fabricated printable hydrogel microarray combines analytical functionality and easily generated multiple sample substrates to eliminate time-dependent SCAM interference, significantly improve the accuracy of lead compound discovery, and simplify the drug discovery process[110]. The hydrogel-based screening assays demonstrate minimal interference and can be applied to screening platforms based on high-throughput microarrays for rapid (<25 min) and low-cost elimination of interfering SCAMs to identify lead compounds with real inhibitory potential[110]. For specific experimental operation, please refer to the literature[113].

5.3. *Verification of redox cycling PAINS suspects*

Discussed mechanisms result in easily formed false positives; however, the compounds that manifest their apparent activity *via* targeted oxidation have not been considered. Certain redox cycling compounds (RCCs), such as naphthoquinones (NQ), undergo redox-dependent cycling in the presence of strong reducing

Table 3 Comparison of experimental methods for the verification of SCAMs.

Assay	Principle	Feature	Readout	Note
Nonionic detergent washing	Detergents will promote the formation of monomers, thus reducing the apparent activity	Easy to operate and low cost	Enzyme inhibition	Detergent should not contribute to the readouts
Centrifugation	Centrifugation will induce the formation of pellets originating, thus changing the apparent activity	Easy to operate and can be applied to HTS platforms	Enzyme inhibition	Selectivity of non-SCAMs to tumor cells should be higher than that of normal cells
Competitive ligand−target binding experiment	SPR can identify their nonspecific aggregation-type binding mechanisms	Easy to operate and provide more detailed mechanism information	Enzyme inhibition	Compounds should be injected into the protein surface and attached to the surface of the optical biosensor
Printable hydrogel microarray	The adjustable porosity of the hydrogels allows selective transport of substrates to and from the entrapped enzyme *via* size selectivity	Easy to operate	Enzyme inhibition	No special requirements

agents, such as dithiothreitol (DTT) and tris(2-carboxyethyl) phosphine (TCEP), resulting in the generation of hydrogen peroxide (H_2O_2)[22]. H_2O_2 can indirectly inhibit the catalytic activity of proteins by oxidizing accessible cysteine, tryptophan, methionine, histidine, or selenocysteine residues; thus, these compounds have pleiotropic effects manifested as nonspecific and promiscuous inhibition of protein activity (Fig. 11)[22,114]. RCC-based interference can lead to exaggerated values of apparent actives in HTS assays. In an HTS campaign to identify inhibitors of the dual specificity phosphatase cell division cycle 25B (Cdc25B), 55% of the concentration-dependent Cdc25B inhibitors were shown to be RCCs[115].

At present, several conventional processes can be used to identify and eliminate RCCs, including multiple counter-screening and secondary detection. This type of PAINS suspects cannot be completely validated; however, probability of their occurrence can be significantly reduced. Moreover, conventional validation methods enable significantly easier differentiation of 'off-target' activity from the targeted and pathway-specific activities[22]. These methods include: (a) use of LC/MS analysis to determine the oxidation of a cysteine in the active site of the target protein[116]; (b) measurement of the UV/Visible spectrum of RCCs with or without a reducing agent to determine whether RCCs are reduced in a time-dependent manner[22]; (c) investigation of the inhibitory effects of weak reducing agents, such as GSH, BME, or Cys on RCCs[22]; (d) verification that the inhibitory effect of RCCs on the target activity is time- and concentration-dependent in the presence of DTT or TCEP[22]; (e) inhibition of the target protein activity by RCCs can be abolished by the addition of catalase (CAT) to the assay to degrade any H_2O_2 produced[115]. CAT can abolish the target inhibition by RCCs in the presence of DTT or TCEP[22]. However, the verification process by these detailed analyses is material- and time-consuming, and the results are insufficient for complete and accurately characterization of PAINS. The following several methods are more feasible and can be used to mitigate the serious impact of RCCs on MTLDs.

5.3.1. *The phenol red-HRP assay*

Recently, a simple, rapid, sensitive, and economical method was used to investigate RCCs that indirectly modulate target activity and to identify promiscuous false positives. This assay is based on the H_2O_2-dependent horseradish peroxidase (HRP)-mediated oxidation of phenol red that produces a change in absorbance at 610 nm at alkaline pH and readily detects H_2O_2 generation (in the 1−100 μmol/L range, Table 4)[116]. The phenol red HRP assay can be performed in a 384-well format and was used to profile ~200,000 compounds available from the LOPAC and NIH MLSCN compound libraries for RCC activity[23] due to the ability of RCCs to generate H_2O_2[116]. Pyrimidotriazinedione is one of the active dominant scaffolds commonly used in drug design that was unfortunately identified as an RCC in the presence of 0.8 mmol/L DTT and has been promiscuously active against a number of target proteins[23]. More than 50% of RCCs identified in the RCC profiling screen are structurally similar to pyrimidotriazinedione and quinones; however, several other RCC pharmacophores were identified[114]. H_2O_2 is the common cellular messenger; hence, RCC interference is not limited to biochemical enzyme-based assays and can also produce promiscuous effects in cell-based analysis[23].

5.3.2. *A surrogate assay*

In addition to the mechanism described above, RCCS can cause inhibition by oxidizing susceptible enzyme targets, such as metalloenzymes and cysteine-containing enzymes. However, this redox phenomenon is rarely investigated, and the detection methods need optimization[114]. Interestingly, a surrogate assay using the conversion of resazurin to resorufin in a redox reaction in the presence of DTT and compounds was used to detect redox activity of small molecules (Table 4)[116]. The surrogate assay couples glucose-6-phosphate production to resorufin *via* glucose-6-phosphate dehydrogenase (G6PDH) and diaphorase[116]. Similar to other cases, 2203 out of 2262 compounds (considered nuisance hits) were able to produce resorufin from resazurin in the presence

Table 4 Comparison of experimental methods for the verification of redox cycling.

Assay	Principle	Feature	Readout
Phenol red-HRP assay	Based on the H_2O_2-dependent horseradish peroxidase mediated oxidation of phenol red that produces a change in its absorbance at 610 nm in alkaline pH, and readily detects H_2O_2 generation	Can produce promiscuous effects in enzymes and cell-based analysis	Enzyme and cell inhibition
Surrogate assay	Using the conversion of resazurin to resorufin redox reaction in the presence of DTT and compounds to detect small molecule redox activity	Easy to operate and can be applied to pharmaceutical targets that perform redox chemistry or targets with functional groups susceptible to redox modification	Enzyme inhibition

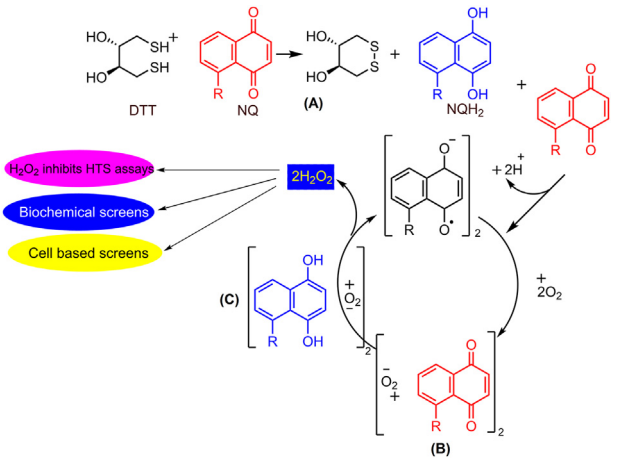


Figure 11 NQ generates H_2O_2 accounting for promiscuous bioactivity profiles in the HTS databases. (A) NQ can be reduced by DTT to form a hydroquinone (QH_2) capable of undergoing a comproportionation reaction with another NQ to form two identical radical anions. (B) In the presence of O_2 these radicals can form O_2^-. (C) The superoxide can then be reduced by QH_2 to form H_2O_2.

of only resazurin and DTT[116]. This surrogate assay was used to evaluate suspected redox compounds **4** and (±)-dunnione (Fig. 12) discovered by GSK as irreversible and noncompetitive inhibitors of caspase-8. A nonredox pancaspase peptide inhibitor Z-VAD (OMe)-FMK (Fig. 12) was used as a competitive control. The known inhibitor was not positive in this assay, whereas compounds **4** and (±)-dunnione were identified as oxidative nuisance compounds[117]. The reduction of resazurin to resorufin is a popular assay format used to measure cell viability; pharmaceutical targets that are involved in redox chemistry or targets with functional groups susceptible to redox modification are likely to benefit from analysis using this assay[117].

5.4. Verification of PAINS suspects involved in metal ion chelation

These false positive results are caused by the pan interference effects of the compounds or organic impurities; however, organic entities do not fully account for all false positive results. Conventional identification methods, such as MS and NMR, used during synthesis are suitable only for confirmation of the structure and purity of organic components, while potential inorganic impurities, such as transition metals that may be used in compound synthesis, are not identified[118]. These inorganic impurities can also cause positive reactions in low micromolar range in a number of potential protein and functional systems, including biochemical and biosensor assays[118]. These positive effects may result in the selection of organic entities thus making inorganic impurities the best candidates for PAINS suspects. For example, organic entities containing zinc impurities may form complexes with multivalent cations that persist through preliminary procedures, and nonfunctional mechanisms of inorganic interference on the readout systems may cause false positive signals[119]. Additionally, inorganic contaminations in organic entities during synthesis can cause inconclusive SAR in later lead optimization. Therefore, design and synthesis of MTDLs requires constant consideration of the effects of

inorganic impurities that can be ruled out in the presence of a nonselective chelator, such as EDTA, or a more selective chelator, such as TPEN (*N*,*N*,*N*′,*N*′, tetrakis (2-pyridylmethyl) ethylenediamine)[119].

5.5. *Verification of PAINS suspects causing sample fluorescence*

Interference mediated by sample fluorescence is highly prevalent in the biological evaluation of MTDLs due to the variability of detection strategies that use fluorescent labels and light detection. Fluorophoric properties usually accompany heterocyclic scaffolds and impurities and can affect the results of biological evaluation to complicate the analysis and cause false positive and negative results[22] (Fig. 13). Certain promiscuous natively fluorescent compounds may be disguised as valid regulators of target functions leading to futile pursuit of biologically inactive compounds in chemical genomics and drug development campaigns[21]. Therefore, these PAINS suspects should be eliminated from subsequent consideration.

Generally, the analytical artefacts arising from fluorescent compounds are reproducible, and the apparent activity usually shows a concentration-dependent response[21]. In this case, orthogonal assays are a conventional and effective method to eliminate false positives caused by fluorescence interference subsequent to the retesting of the primary screening hits[21]. This assay is generally used to assess compound interference by reading the fluorescence intensity of the analysis mixture after the sample is added prior to initiation of the detection leading to a change in the fluorophore[21]. Fluorescence intensity of each well and the calculated polarization value can rule out PAINS suspects with above average fluorescence in the analysis based on fluorescence polarization (FP)[21].

In addition to the method of calibration using fluorophore standards, an assay similar to quantitative HTS is used to profile each compound within a range of concentrations that can be used in verification[125]. The concentration response-based analysis measures fluorescence intensity or potency of each sample based on the known fluorophores and thus can be used to identify PAINS suspects[21].

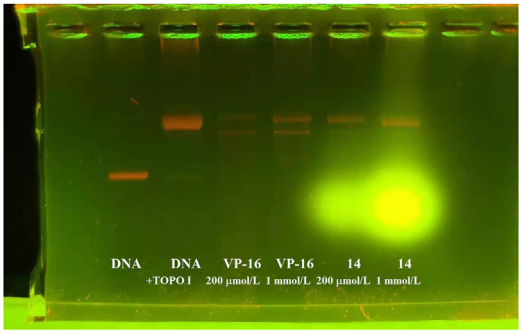


Figure 13 The effect of strong fluorescence interference on the results of topoisomerase II (Topo II) inhibition experiment. The image shows the results of Topo II inhibition assay of compounds currently investigated by us.

6. Summary

MTDLs can simultaneously regulate multiple associations in the disease network thus improving the efficacy and reducing the adverse reactions. Some MTDLs have been successfully used in many major diseases and have become very important therapeutics. Reasonable MTDL design may provide novel effective and promising products. PAINS has become an obstacle in the process of MTDL development, and medicinal chemists have a complex task to accurately distinguish "good" PAINS from "bad" PAINS based on currently available rules and screening tools. Up to 95% of "hits" identified from virtual screening can be considered PAINS; thus, this problem should be taken seriously to ruthlessly block some of the compounds from entering the next stage of investigation. Optimistically, use of the 'Fair Trail Strategy' and active investigations of the interference mechanism of PAINS by efficient and comprehensive approaches to exclude the interference suspects can double the output of MTDL research and development while halving the required effort.

Identification of the relevant targets for disease regulation and initiation of the search for lead ligands should be followed by the investigation of the ligand−target interactions using reasonable methods that can exclude the interference of PAINS. The complexity of biological systems and properties of ligand−target interactions may cause variability in the results of virtual screening. Thus, to avoid the inaccurate classification of compounds as PAINS, computational and experimental approaches should be combined in an ideal scenario. In the absence of accurate experimental evidence, false positive ligands should be only flagged as PAINS suspects that can be subsequently exculpated to avoid their disappearance from the chemical libraries or abandoned by medicinal chemists.

Acknowledgments

This work was financially supported by National Natural Science Foundation of China (31870325), Key Laboratory for Tumor Precision Medicine of Shanxi Province Research Fund (KLTPM-SX2022-B3, China), Innovation and Entrepreneurship Training Program for Undergraduate (202010316043S, China) and "Double First Class" Subject Innovation Team Construction Project of China Pharmaceutical University (CPU2018GY12, China) .

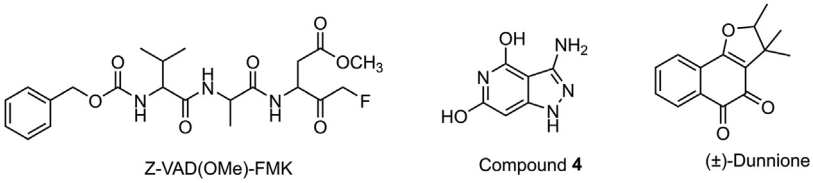


Figure 12 Structures of compounds studied in the redox assay.

Author contributions

Jianbo Sun and Hui Zhong wrote the manuscript; Kun Wang made the figures; Na Li reviewed and edited the manuscript; Li Chen supervised and edited the manuscript.

Conflicts of interest

The authors declare no conflict of interest for publishing this manuscript.

References

1. Ravikumar B, Aittokallio T. Improving the efficacy-safety balance of polypharmacology in multi-target drug discovery. *Expert Opin Drug Met* 2018;**13**:179−92.
2. Alcaro S, Bolognesi ML, García-Sosa AT, Rapposelli S. Multi-target-directed ligands (MTDL) as challenging research tools in drug discovery: from design to pharmacological evaluation. *Front Chem* 2019;**7**:71.
3. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* 2008;**4**:682.
4. Bottegoni G, Favia AD, Recanatini M, Cavalli A. The role of fragment-based and computational methods in polypharmacology. *Drug Discov Today* 2012;**17**:23−34.
5. Aldrich C, Bertozzi C, Georg GI, Kiessling L, Lindsley C, Liotta D, et al. The ecstasy and agony of assay interference compounds. *ACS Chem Neurosci* 2017;**8**:420−3.
6. Baell J, Walters MA. Chemistry: chemical con artists foil drug discovery. *Nature News* 2014;**513**:481−3.
7. Maeda K, Sugino H, Akazawa H, Amada N, Shimada J, Futamura T, et al. Brexpiprazole I: *in vitro* and *in vivo* characterization of a novel serotonin-dopamine activity modulator. *J Pharmacol Exp Therapeut* 2014;**350**:589−604.
8. Halberstadt AL, Vollenweider FX, Nichols DE. *Behavioral neurobiology of psychedelic drugs*. Berlin, Heidelberg: Springer; 2018. p. 36.
9. Nelson KM, Dahlin JL, Bisson J, Graham J, Pauli GF, Walters MAJ. The essential medicinal chemistry of curcumin: Miniperspective. *J Med Chem* 2017;**60**:1620−37.
10. Baell JB. Feeling nature's PAINS: natural products, natural product drugs, and pan assay interference compounds (PAINS). *J Nat Prod* 2016;**79**:616−28.
11. Baell JB, Nissink JWM. Seven year itch: pan-Assay interference compounds (PAINS) in 2017-utility and limitations. *ACS Chem Biol* 2018;**13**:36−44.
12. Schneider P, Schneider G. Privileged structures revisited. *Angew Chem Int Ed* 2017;**56**:7971−4.
13. Yang JJ, Ursu O, Lipinski CA, Sklar LA, Oprea TI, Bologa CG. Badapple: promiscuity patterns from noisy evidence. *J Cheminf* 2016;**8**:29.
14. Jasial S, Hu Y, Bajorath JR. How frequently are pan-assay interference compounds active? Large-scale analysis of screening data reveals diverse activity profiles, low global hit frequency, and many consistently inactive compounds. *J Med Chem* 2017;**60**:3879−86.
15. Capuzzi SJ, Muratov EN, Tropsha A. Phantom PAINS: problems with the utility of alerts for pan-assay in terference compound S. *J Chem Inf Model* 2017;**57**:417−27.
16. Dahlin JL, Walters MA. How to triage PAINS-full research. *Assay Drug Dev Technol* 2016;**14**:168−74.
17. Costantino L, Barlocco D. Designed multiple ligands: basic research *vs* clinical outcomes. *Curr Med Chem* 2012;**19**:3353−87.
18. Coan KE, Shoichet BK. Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors. *J Am Chem Soc* 2008;**130**:9606−12.
19. Schorpp K, Rothenaigner I, Salmina E, Reinshagen J, Low T, Brenke JK, et al. Identification of small-molecule frequent hitters from alphascreen high-throughput screens. *J Biomol Screen* 2014;**19**:715−26.
20. Huth JR, Mendoza R, Olejniczak ET, Johnson RW, Cothron DA, Liu Y, et al. Alarm NMR: a rapid and robust experimental method to detect reactive false positives in biochemical screens. *J Am Chem Soc* 2005;**127**:217−24.
21. Simeonov A, Jadhav A, Thomas CJ, Wang Y, Huang R, Southall NT, et al. Fluorescence spectroscopic profiling of compound libraries. *J Med Chem* 2008;**51**:2363−71.
22. Thorne N, Auld DS, Inglese J. Apparent activity in high-throughput screening: origins of compound-dependent assay interference. *Curr Opin Chem Biol* 2010;**14**:315−24.
23. Soares KM, Blackmon N, Shun TY, Shinde SN, Takyi HK, Wipf P, et al. Profiling the NIH small molecule repository for compounds that generate H_2O_2 by redox cycling in reducing environments. *Assay Drug Dev Technol* 2010;**8**:152−74.
24. Ingólfsson HI, Thakur P, Herold KF, Hobart EA, Ramsey NB, Periole X, et al. Phytochemicals perturb membranes and promiscuously alter protein function. *ACS Chem Biol* 2014;**9**:1788−98.
25. Feng BY, Shelat A, Doman TN, Guy RK, Shoichet BK. High-throughput assays for promiscuous inhibitors. *Nat Chem Biol* 2005;**1**:146−8.
26. Schneider C, Gordon ON, Edwards RL, Luis PB. Degradation of curcumin: from mechanism to biological implications. *J Agric Food Chem* 2015;**63**:7606−14.
27. Guo H, Eleftheriadis N, Rohr-Udilova N, Dömling A, Dekker FJ. Photoactivation provides a mechanistic explanation for pan-assay interference behaviour of 2-aminopyrroles in lipoxygenase inhibition. *Eur J Med Chem* 2017;**139**:633−43.
28. Zhou J, Jiang X, He S, Jiang H, Feng F, Liu W, et al. Rational design of multitarget-directed ligands: strategies and emerging paradigms. *J Med Chem* 2019;**62**:8881−914.
29. Jasial S, Gilberg E, Blaschke T, Bajorath JR. Machine learning distinguishes with high accuracy between pan-assay interference compounds that are promiscuous or represent dark chemical matter. *J Med Chem* 2018;**61**:10255−64.
30. Gilberg E, Stumpfe D, Bajorath J. Activity profiles of analog series containing pan assay interference compounds. *RSC Adv* 2017;**7**:35638−47.
31. Gilberg E, Jasial S, Stumpfe D, Dimova D, Bajorath J. Highly promiscuous small molecules from biological screening assays include many pan-assay interference compounds but also candidates for polypharmacology. *J Med Chem* 2016;**59**:10285−90.
32. Gilberg E. Gütschow M, Bajorath J. X-Ray structures of target−ligand complexes containing compounds with assay interference potential. *J Med Chem* 2018;**61**:1276−84.
33. Dahlin JL, Nissink JWM, Strasser JM, Francis S, Higgins L, Zhou H, et al. PAINS in the assay: chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *J Med Chem* 2015;**58**:2091−113.
34. Schneider G. Automating drug discovery. *Nat Rev Drug Discov* 2018;**17**:97.
35. Xie T, Du GH. Pain of high-throughput screening-pan assay interference compounds. *Acta Pharm Sin* 2015;**50**:925−30.
36. Baell JB. Observations on screening-based research and some concerning trends in the literature. *Future Med Chem* 2010;**2**:1529−46.
37. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* 2010;**53**:2719−40.
38. Arrowsmith CH, Audia JE, Austin C, Baell J, Bennett J, Blagg J, et al. The promise and peril of chemical probes. *Nat Chem Biol* 2015;**11**:536−41.
39. Shoichet BK. Screening in a spirit haunted world. *Drug Discov Today* 2006;**11**:607−15.
40. Walters WP, Namchuk M. Designing screens: how to make your hits a hit. *Nat Rev Drug Discov* 2003;**2**:259−66.

41. Tomašić T, Peterlin Mašič L. Rhodanine as a scaffold in drug discovery: a critical review of its biological activities and mechanisms of target modulation. *Expet Opin Drug Discov* 2012;**7**:549−60.
42. Mendgen T, Steuer C, Klein CD. Privileged scaffolds or promiscuous binders: a comparative study on rhodanines and related heterocycles in medicinal chemistry. *J Med Chem* 2012;**55**:743−53.
43. Zhou JT, Jiang XY, Feng F, Sun HP. Multi-target drug design strategy and its research progress. *Acta Pharm Sin* 2018;**53**:2012−25.
44. Gilberg E. Gütschow M, Bajorath J. Promiscuous ligands from experimentally determined structures, binding conformations, and protein family-dependent interaction hotspots. *ACS Omega* 2019;**4**: 1729−37.
45. Li S, Xu Y, Shen Q, Liu X, Lu J, Chen Y, et al. Non-covalent interactions with aromatic rings: current understanding and implications for rational drug design. *Curr Pharmaceut Des* 2013;**19**: 6522−33.
46. Babaoglu K, Simeonov A, Irwin JJ, Nelson ME, Feng B, Thomas CJ, et al. Comprehensive mechanistic analysis of hits from high-throughput and docking screens against β-lactamase. *J Med Chem* 2008;**51**:2502−11.
47. Baell JB, Ferrins L, Falk H, Nikolakopoulos G. PAINS: Relevance to tool compound discovery and fragment-based screening. *Aust J Chem* 2014;**66**:1483−94.
48. Sushko I, Salmina E, Potemkin VA, Poda G, Tetko IV. ToxAlerts: a web server of structural alerts for toxic chemicals and compounds with potential adverse reactions. *J Chem Inf Model* 2012; **52**:2310−6.
49. Zeller J, Turbiak AJ, Powelson IA, Lee S, Sun D, Showalter HH, et al. Investigation of 3-aryl-pyrimido[5,4-*e*][1,2,4]triazine-5,7-diones as small molecule antagonists of β-catenin/TCF transcription. *Bioorg Med Chem Lett* 2013;**23**:5814−20.
50. Baell JB. Screening-based translation of public research encounters painful problems. *ACS Med Chem Lett* 2015;**6**:229−34.
51. Neves AP, Pereira MX, Peterson EJ, Kipping R, Vargas MD, Silva Jr FP, et al. Exploring the DNA binding/cleavage, cellular accumulation and topoisomerase inhibition of 2-hydroxy-3-(amino-methyl)-1,4-naphthoquinone mannich bases and their platinum (II) complexes. *J Inorg Biochem* 2013;**119**:54−64.
52. Musiol R. An overview of quinoline as a privileged scaffold in cancer drug discovery. *Expert Opin Drug Met* 2017;**12**:583−97.
53. Cardoso MF, Rodrigues PC, Oliveira MEI, Gama IL, da Silva IM, Santos IO, et al. Synthesis and evaluation of the cytotoxic activity of 1,2-furanonaphthoquinones tethered to 1,2,3-1*H*-triazoles in myeloid and lymphoid leukemia cell lines. *Eur J Med Chem* 2014;**84**:708−17.
54. Yu W, Li C, Zhang W, Xia Y, Li S, Lin JY, et al. Discovery of an orally selective inhibitor of signal transducer and activator of transcription 3 using advanced multiple ligand simultaneous docking. *J Med Chem* 2017;**60**:2718−31.
55. Andrade MM, Martins LC, Marques GV, Silva CA, Faria G, Caldas S, et al. Synthesis of quinoline derivatives as potential cysteine protease inhibitors. *Future Med Chem* 2020;**12**:571−81.
56. Tsou HR, Overbeek-Klumpers EG, Hallett WA, Reich MF, Floyd MB, Johnson BD, et al. Optimization of 6,7-disubstituted-4-(arylamino) quinoline-3-carbonitriles as orally active, irreversible inhibitors of human epidermal growth factor receptor-2 kinase activity. *J Med Chem* 2005;**48**:1107−31.
57. Rabindran SK, Discafani CM, Rosfjord EC, Baxter M, Floyd MB, Golas J, et al. Antitumor activity of HKI-272, an orally active, irreversible inhibitor of the HER-2 tyrosine kinase. *Cancer Res* 2004; **64**:3958−65.
58. Decker M. *Design of hybrid molecules for drug development*. Würzburg: Elsevier; 2017.
59. Ramsay RR, Popovic-Nikolic MR, Nikolic K, Uliassi E, Bolognesi ML. A perspective on multi-target drug discovery and design for complex diseases. *Clin Transl Med* 2018;**7**:3.
60. Gilberg E, Stumpfe D, Bajorath J. Towards a systematic assessment of assay interference: identification of extensively tested compounds with high assay promiscuity. *F1000 Research* 2017;**6**. Chem Inf Sci-1505.
61. Wang Tao, Li Zhuyin, Cvijic Mary Ellen, Krause Carol, Zhang Litao, Sum Chi Shing. Measurement of β-arrestin recruitment for GPCR targets. In: Sittampalam GS, Grossman A, Brimacombe K, Arkin M, Auld D, Austin C, et al., editors. *Assay guidance manual.* Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2004.
62. Senger MR, Fraga CA, Dantas RF, Silva Jr FP. Filtering promiscuous compounds in early drug discovery: is it a good idea?. *Drug Discov Today* 2016;**21**:868−72.
63. Inglese J, Johnson RL, Simeonov A, Xia M, Zheng W, Austin CP, et al. High-throughput screening assays for the identification of chemical probes. *Nat Chem Biol* 2007;**3**:466−79.
64. Shapiro AB, Walkup GK, Keating TA. Correction for interference by test samples in high-throughput assays. *J Biomol Screen* 2009;**14**: 1008−16.
65. Barf T, Kaptein A. Irreversible protein kinase inhibitors: balancing the benefits and risks. *J Med Chem* 2012;**55**:6243−62.
66. Wilson GL, Lill MA. Integrating structure-based and ligand-based approaches for computational drug design. *Future Med Chem* 2011;**3**:735−50.
67. Carpenter KA, Huang XJ. Machine learning-based virtual screening and its applications to alzheimer's drug discovery: a review. *Curr Pharmaceut Des* 2018;**24**:3347−58.
68. Fu RG, Sun Y, Sheng WB, Liao DF. Designing multi-targeted agents: an emerging anticancer drug discovery paradigm. *Eur J Med Chem* 2017;**136**:195−211.
69. Gilberg E, Stumpfe D, Bajorath JR. X-ray-structure-based identification of compounds with activity against targets from different families and generation of templates for multitarget ligand design. *ACS omega* 2018;**3**:106−11.
70. Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput-Aid Drug* 2011;**7**:146−57.
71. Bhachoo J, Beuming T. Investigating protein−peptide interactions using the Schrödinger computational suite. *Methods Mol Biol* 2017; **1561**:235−54.
72. Wu G, Robertson DH, Brooks III CL, Vieth M. Detailed analysis of grid-based molecular docking: a case study of CDOCKER-A CHARMm-based MD docking algorithm. *J Comput Chem* 2003; **24**:1549−62.
73. London N, Miller RM, Krishnan S, Uchida K, Irwin JJ, Eidam O, et al. Covalent docking of large libraries for the discovery of chemical probes. *Nat Chem Biol* 2014;**10**:1066−72.
74. Jones G, Willett P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Biochem Mol Biol* 1995;**245**:43−53.
75. Soufan O, Ba-alawi W, Magana-Mora A, Essack M, Bajic VB. DPubChem: a web tool for QSAR modeling and high-throughput virtual screening. *Sci Rep-UK* 2018;**8**:1−10.
76. Cramer RD. The inevitable QSAR renaissance. *J Comput Aided Mol Des* 2012;**26**:35−8.
77. Ismaili L, Refouvelet B, Benchekroun M, Brogi S, Brindisi M, Gemma S, et al. Multitarget compounds bearing tacrine-and donepezil-like structural and functional motifs for the potential treatment of alzheimer's disease. *Prog Neurobiol* 2017;**151**:4−34.
78. Bautista-Aguilera OM, Samadi A, Chioua M, Nikolic K, Filipic S, Agbaba D, et al. *N*-Methyl-*N*-((1-methyl-5-(3-(1-(2-methylbenzyl) piperidin-4-yl) propoxy)-1*H*-indol-2-yl) methyl) prop-2-yn-1-amine, a new cholinesterase and monoamine oxidase dual inhibitor. *J Med Chem* 2014;**57**:10455−63.
79. Bautista-Aguilera OM, Esteban G, Chioua M, Nikolic K, Agbaba D, Moraleda I, et al. Multipotent cholinesterase/monoamine oxidase inhibitors for the treatment of alzheimer's disease: design, synthesis, biochemical evaluation, ADMET, molecular modeling, and QSAR analysis of novel donepezil−pyridyl hybrids. *Drug Des Dev Ther* 2014;**8**:1893−910.

80. Knapp B, Ospina L, Deane CM. Avoiding false positive conclusions in molecular simulation: the importance of replicas. *J Chem Theor Comput* 2018;**14**:6127−38.
81. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II , Cronin M, et al. QSAR modeling: where have you been? Where are you going to? *J Med Chem* 2014;**57**:4977−5010.
82. Jing Y, Bian Y, Hu Z, Wang L, Xie XQS. Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era. *AAPS J* 2018;**20**:58.
83. Reker D, Bernardes GJ, Rodrigues T. Computational advances in combating colloidal aggregation in drug discovery. *Nat Chem* 2019; **11**:402−18.
84. Stork C, Wagner J, Friedrich NO. de Bruyn Kops, CŠíchoM, Kirchmair J. Hit dexter: a Machine-learning model for the prediction of frequent hitters. *ChemMedChem* 2018;**13**:564−71.
85. David L, Walsh J, Sturm N, Feierberg I, Nissink JWM, Chen H, et al. Identification of compounds that interfere with high-throughput screening assay technologies. *ChemMedChem* 2019;**14**:1795−802.
86. Feldmann C, Bajorath J. Compounds with multitarget activity: structure-based analysis and machine learning. *Future Drug Discov* 2020;**2**:FDD44.
87. Molinski SV, Shahani VM, Subramanian AS, MacKinnon SS, Woollard G, Laforet M, et al. Comprehensive mapping of cystic fibrosis mutations to CFTR protein identifies mutation clusters and molecular docking predicts corrector binding site. *Proteins* 2018;**86**: 833−43.
88. Zhou H, Gao M, Skolnick J. Comprehensive prediction of drug–protein interactions and side effects for the human proteome. *SCI REP-UK* 2015;**5**:11090.
89. Landrum G. RDKit: open-source cheminformatics software. *GitHub and SourceForge* 2016;**10**:3592822.
90. Sterling T, Irwin JJ. ZINC 15-ligand discovery for everyone. *J Chem Inf Model* 2015;**55**:2324−37.
91. Jana S, Ganeshpurkar A, Singh SK. Multiple 3D-QSAR modeling, e-pharmacophore, molecular docking, and *in vitro* study to explore novel achE inhibitors. *RSC Adv* 2018;**8**:39477−95.
92. Koptelov M, Zimmermann A, Bonnet P, Bureau R, Crémilleux B. PrePeP: a tool for the identification and characterization of pan assay interference compounds. In: *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*; July 2018. p. 462−71.
93. Mignani S, Rodrigues J, Tomas H, Jalal R, Singh PP, Majoral JP, et al. Present drug-likeness filters in medicinal chemistry during the hit and lead optimization process: how far can they be simplified?. *Drug Discov Today* 2018;**23**:605−15.
94. Strelow JM. A perspective on the kinetics of covalent and irreversible inhibition. *SLAS Discov* 2017;**22**:3−20.
95. Zhang T, Hatcher JM, Teng M, Gray NS, Kostic M. Recent advances in selective and irreversible covalent ligand development and validation. *Cell Chem Biol* 2019;**26**:1486−500.
96. Zhang T, Kwiatkowski N, Olson CM, Dixon-Clarke SE, Abraham BJ, Greifenberg AK, et al. Covalent targeting of remote cysteine residues to develop CDK12 and CDK13 inhibitors. *Nat Chem Biol* 2016;**12**:876.
97. Anighoro A, Bajorath J, Rastelli GJ. Polypharmacology: challenges and opportunities in drug discovery: miniperspective. *J Med Chem* 2014;**57**:7874−87.
98. Hu Y, Bajorath J. Compound promiscuity: what can we learn from current data?. *Drug Discov Today* 2013;**18**:644−50.
99. Jackson PA, Widen JC, Harki DA, Brummond KM. Covalent modifiers: a chemical perspective on the reactivity of α,β-unsaturated carbonyls with thiols *via* hetero-michael addition reactions. *J Med Chem* 2017;**60**:839−85.
100. Drewes G, Knapp S. Chemoproteomics and chemical probes for target discovery. *Trends Biotechnol* 2018;**36**:1275−86.
101. Browne CM, Jiang B, Ficarro SB, Doctor ZM, Johnson JL, Card JD, et al. A chemoproteomic strategy for direct and proteome-wide covalent inhibitor target-site identification. *J Am Chem Soc* 2018;**141**: 191−203.
102. Dahlin JL, Nelson KM, Strasser JM, Barsyte-Lovejoy D, Szewczyk MM, Organ S, et al. Assay interference and off-target liabilities of reported histone acetyltransferase inhibitors. *Nat Commun* 2017;**8**:1−14.
103. Ganesh AN, Donders EN, Shoichet BK, Shoichet MS. Colloidal aggregation: from screening nuisance to formulation nuance. *Nano Today* 2018;**19**:188−200.
104. Mcgovern SL, Caselli E, Grigorieff N, Shoichet BK. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem* 2002;**45**:1712−22.
105. Mcgovern SL, Helfand BT, Feng B, Shoichet BK. A specific mechanism of nonspecific inhibition. *J Med Chem* 2003;**46**:4265−72.
106. Duan D, Torosyan H, Elnatan D, Mclaughlin CK, Logie J, Shoichet MS, Agard DA, Shoichet BK. Internal structure and preferential protein binding of colloidal aggregates. *ACS Chem Biol* 2017;**12**:282−90.
107. Owen SC, Doak AK, Ganesh AN, Nedyalkova L, McLaughlin CK, Shoichet BK, et al. Colloidal drug formulations can explain "bell-shaped" concentration−response curves. *ACS Chem Biol* 2014;**9**: 777−84.
108. Irwin JJ, Duan D, Torosyan H, Doak AK, Ziebart KT, Sterling T, et al. An aggregation advisor for ligand discovery. *J Med Chem* 2015; **58**:7076−87.
109. Rao H, Li Z, Li X, Ma X, Ung C, Li H, et al. Identification of small molecule aggregators from large compound libraries by support vector machines. *J Comput Chem* 2010;**31**:752−63.
110. Mateen R, Ali MM, Hoare T. A printable hydrogel microarray for drug screening avoids false positives associated with promiscuous aggregating inhibitors. *Nat Commun* 2018;**9**:1−9.
111. Giannetti AM, Koch BD, Browner MF. Surface plasmon resonance based assay for the detection and characterization of promiscuous inhibitors. *J Med Chem* 2008;**51**:574−80.
112. Chan LL, Lidstone EA, Finch KE, Heeres JT, Hergenrother PJ, Cunningham BT. A method for identifying small-molecule aggregators using photonic crystal biosensor microplates. *JALA Charlottesv Va* 2009;**14**:348−59.
113. Mateen R, Ali MM, Hoare T. A printable hydrogel microarray for drug screening avoids false positives associated with promiscuous aggregating inhibitors. *Nat Commun* 2018;**9**:602.
114. Johnston PA. Redox cycling compounds generate H_2O_2 in HTS buffers containing strong reducing reagents-real hits or promiscuous artifacts?. *CurrOpin Chem Biol* 2011;**15**:174−82.
115. Johnston PA, Foster CA, Tierno MB, Shun TY, Shinde SN, Paquette WD, et al. Cdc25B dual-specificity phosphatase inhibitors identified in a high-throughput screen of the NIH compound library. *Assay Drug Dev Technol* 2009;**7**:250−65.
116. Johnston PA, Soares KM, Shinde SN, Foster CA, Shun TY, Takyi HK, et al. Development of a 384-well colorimetric assay to quantify hydrogen peroxide generated by the redox cycling of compounds in the presence of reducing agents. *Assay Drug Dev Technol* 2008;**6**:505−18.
117. Lor LA, Schneck J, Mcnulty DE, Diaz E, Brandt M, Thrall SH, et al. A simple assay for detection of small-molecule redox activity. *J Biomol Screen* 2007;**12**:881−90.
118. Roughley SD, Jordan AM. The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. *J Med Chem* 2011;**54**:3451−79.
119. Hermann JC, Chen Y, Wartchow C, Menke J, Gao L, Gleason SK, et al. Metal impurities cause false positives in high-throughput screening campaigns. *ACS Med Chem Lett* 2013;**4**: 197−200.