



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Panel docking of small-molecule libraries - Prospects to improve efficiency of lead compound discovery

Citation for published version:

Sarnpitak, P, Mujumdar, P, Taylor, P, Cross, M, Coster, MJ, Gorse, A-D, Krasavin, M & Hofmann, A 2015, 'Panel docking of small-molecule libraries - Prospects to improve efficiency of lead compound discovery', *Biotechnology Advances*, vol. 33, no. 6 Pt 1, pp. 941-7. <https://doi.org/10.1016/j.biotechadv.2015.05.006>

Digital Object Identifier (DOI):

[10.1016/j.biotechadv.2015.05.006](https://doi.org/10.1016/j.biotechadv.2015.05.006)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Biotechnology Advances

Publisher Rights Statement:

<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Panel docking of small-molecule libraries – Prospects to improve efficiency of lead compound discovery

Pakornwit Sarnpitak^{1,†}, Prashant Mujumdar^{1,†}, Paul Taylor², Megan Cross¹, Mark Coster¹, Alain-Dominique Gorse³, Mikhail Krasavin^{1,*} & Andreas Hofmann^{1,4,5,*}

¹Eskitis Institute for Drug Discovery, Griffith University, Nathan, Qld, Australia

²School of Biological Sciences, The University of Edinburgh, Scotland, UK

³QFABBioinformatics, Institute for Molecular Bioscience, University of Queensland, St Lucia, Qld, Australia

⁴Faculty of Veterinary Science, The University of Melbourne, Parkville, Victoria, Australia

⁵Queensland Tropical Health Alliance, Smithfield, Australia

[†]These authors contributed equally

*Corresponding authors: Mikhail Krasavin
current address:
Department of Chemistry, Saint Petersburg State University
Peterhof, 198504 Saint Petersburg, Russian Federation
Tel: +7-931-361 7872. E-mail: m.krasavin@spbu.ru

Andreas Hofmann
Structural Chemistry Program, Eskitis Institute, Griffith University
N75 Don Young Road, Nathan, Qld 4111, Australia
Tel: +61-7-3735-4425, Email: a.hofmann@griffith.edu.au

Abstract

Computational docking as a means to prioritise small molecules in drug discovery projects remains a highly popular *in silico* screening approach. Contemporary docking approaches without experimental parametrisation can reliably differentiate active and inactive chemotypes in a protein binding site, but the absence of a correlation between the score of a predicted binding pose and the biological activity of the molecule presents a clear limitation.

Several novel or improved computational approaches have been developed in the recent past to aid in screening and profiling of small-molecule ligands for drug discovery, but also more broadly in developing conceptual relationships between different protein targets by chemical probing. Among those new methodologies is a strategy known as inverse virtual screening, which involves the docking of a compound into different protein structures. In the present article, we review the different computational screening methodologies that employ docking of atomic models, and, by means of a case study, present an approach that expands the inverse virtual screening concept.

By computationally screening a reasonably sized library of 1235 compounds against a panel of 48 mostly human kinases, we have been able to identify five groups of putative lead compounds with substantial diversity when compared to each other. One representative of each of the five groups was synthesised, and tested in kinase inhibition assays, yielding two compounds with micro-molar inhibition in five human kinases.

This highly economic and cost-effective methodology holds great promise for drug discovery projects, especially in cases where a group of target proteins share high structural similarity in their binding sites.

Contents

1. Introduction.....	2
2. Background.....	3
2.1 Kinase drug discovery.....	3
2.2 Ligand-protein docking and the virtual screening approach.....	3
2.3 The inverse virtual screening approach.....	4
2.4 Expanding inverse virtual screening to larger compound libraries.....	4
3. Towards a next-generation approach.....	5
3.1 Novel chemotypes for kinase drug discovery.....	5
3.2 Panel docking.....	6
3.3 In vitro kinase inhibition.....	6
4. Conclusion.....	7

1. Introduction

Because of the increasing importance placed on the pursuit of innovative chemical starting points (leads) in contemporary small-molecule drug discovery (Leeson & Springthorpe, 2007; Oprea, 2002), the challenge to identify non-promiscuous kinase inhibitors is made even more daunting on account of the limited number of known chemotypes. The search for new lead structures in contemporary drug discovery typically entails the high-throughput screening of megalibraries of small organic molecules (molecules with a molecular mass <500 Da) against modified cell lines or isolated recombinantly produced target proteins to identify novel modulators (Macarron et al., 2011). While in the field of protein kinases, for example, many ATP-competitive inhibitors were indeed discovered via this strategy, high-throughput screening has lost much of its initial momentum as most of the scaffolds present in existing libraries have now been identified (Li et al., 2004). Another limitation associated with the high-throughput screening of compound libraries (especially against a comprehensive panel of protein kinases) is that the process can be time and resource intensive (Camp et al., 2012). Because of these drawbacks, predictive methods have

increasingly been employed to identify new modulators that could potentially be developed into drugs.

In the past 15 years, a variety of different informatics-based and computational methods have been developed to aid in screening and profiling of small-molecule ligands for drug discovery (for a review see (Ekins et al., 2007)). A relatively new strategy known as inverse virtual screening (IVS) involves the docking of a compound into different protein structures described in protein databases (Chen & Zhi, 2001; Hui-fang et al., 2010; Lauro et al., 2011; Paul et al., 2004). Here, we present a brief overview of the different computational screening methodologies based on docking, and show that the resulting “panel docking” scores potentially facilitate the identification and analysis of promising modulators through statistical analysis of predicted binding to more than one target protein.

2. Background

2.1 Kinase drug discovery

Currently, just over 500 protein kinases have been identified in the human genome (Zhang et al., 2009). Kinases play crucial roles in the regulation of a wide range of cellular processes, including cell-cycle progression, transcription, DNA replication and metabolic functions by catalysing the transfer of phosphates to serine, threonine and tyrosine residues (Hanks et al., 1988). Most protein kinases possess a catalytic domain, which binds and phosphorylates target proteins, as well as a regulatory region, which, in turn, may be regulated by (auto)phosphorylation. The dysregulation of kinase function has established protein kinases as drug targets in cancer (e.g. Imatinib/Gleevec targeting BCR-Abl tyrosine kinase, c-Kit receptor tyrosine kinase, PDGF receptor) and, more recently, inflammatory diseases such as rheumatoid arthritis and myelofibrosis (e.g. Janus kinase inhibitors Tofacitinib and Ruxolitinib) (Cohen & Alessi, 2013). However, despite their attractiveness as drug targets, their relatively conserved ATP binding site presents a significant challenge in the discovery of non-promiscuous inhibitors. Structurally, the catalytic domains of eukaryotic protein kinases (Figure 1) share a highly conserved bilobal fold with 12 characteristic structural features (also known as subdomains and denoted with Roman numerals) (Hanks & Hunter, 1995). Given that the nucleotide binding sites of kinases are the most prominent, but also most highly conserved, druggable sites, two major challenges become crucial for drug discovery: (i) how can compounds be developed that selectively bind to a specific nucleotide binding site rather than any of the other >500 kinases, and; (ii) the newly developed compounds need to compete with ATP concentrations at the millimolar level and thus possess very high potency to be effective as a drug. The combination of selectivity and potency poses a serious challenge, but recent examples (see above) show that effective compounds can be developed. Intriguingly, in some therapeutic areas high specificity may not actually be required. One such example is Imatinib (Gleevec) where its broad range of efficacy is not undesired in the field of oncology (Cohen & Alessi, 2013).

2.2 Ligand-protein docking and the virtual screening approach

The concept of ligand-protein docking is widely used and has generally been proven to be a useful tool in aiding drug development. The fundamental process of docking comprises placement of a single small-molecule ligand in a pre-defined protein binding site. The binding pose of the placed ligand in a single, but more frequently in multiple conformations is evaluated by means of a scoring function. A variety of different algorithms have been established, and include shape matching (Kuntz et al., 1982; Miller et al., 1994), fragment-based docking (Böhm, 1992; Eisen et al., 1994; Ewing et al., 2001; Schellhammer & Rarey, 2004; Welch et al., 1996), simulated annealing (Morris et al., 2009), as well as genetic algorithms (Jones et al., 1997; Oshiro et al., 1995), evolutionary programming (Gehlhaar et al., 1995), and other approaches (Baxter et al., 1998; Blom & Sygusch, 1997; Lorber & Shoichet, 2005; McMartin & Bohacek, 1997; Wang et al., 1999). A detailed assessment of different docking programs and their scoring functions revealed that these approaches work well as an overall methodology, but the binding pose observed in the experimental crystal

structure is often not identified by the scoring functions, albeit the pose is reproduced as one of many by the docking algorithm (Warren et al., 2006). Accordingly, the scores achieved in ligand-protein docking are poor indicators of affinity.

The availability of established algorithms and the high degree of automation paired with the substantial computing power of modern processors has made automated ligand-protein docking a frequently used tool for virtual screening. In this process, the binding mode of a large number of small-molecule ligands with a single target protein can be evaluated in a fairly short time period. Two fundamental parameters characterise the performance of a virtual screening protocol (Warren et al., 2006): (i) the enrichment, which expresses the chance of identifying an active compound as compared to pure random choice, and (ii) the lead identification, which informs about how many compounds need to be screened until at least one representative of each active compound class has been identified.

2.3 The inverse virtual screening approach

In contrast to conventional virtual screening, which employs a single protein target and a library of small-molecule ligands, this situation is turned around in the inverse virtual screening approach. Conceptually, a single small-molecule ligand is docked into a variety of different protein targets. Inverse virtual screening has received some attention in the recent past due to its potential for identification of new target proteins for a given natural product. This methodology is based on the fact that proteins adopt a finite number of folds (~10,000) (Koonin et al., 2002) and that the recognition of any natural small-molecule ligand is coded into the particular protein fold by means of the three-dimensional conformations of the small-molecule as well as its possible interactions with protein residues. Because small-molecule natural products are made biosynthetically in nature by a particular enzyme, the imprint of interactions with proteins should be available from those synthetic enzymes, and, when deciphered, could be used to identify other target proteins for this specific natural product (Breinbauer et al., 2002; Kellenberger et al., 2011).

An important application for this approach lies in the prediction of off-targets for an individual ligand, as demonstrated by an inverse approach that identified the mycobacterial enoyl-acyl carrier protein reductase as a target for human catechol-O-methyl transferase inhibitors (Kinnings et al., 2009) and screened for staurosporine recognition in a library of some 6400 protein ligand binding sites (Defranchi et al., 2010). Those studies demonstrated that the concepts applied in inverse screening may not necessarily apply the same protocols as conventional screening. In both cases, the screening protocols did not employ an atomic approach, but represented the protein binding sites as abstract objects consisting of the 3D coordinates of the C α atoms and select physicochemical and geometrical properties. This concept has been embraced by the more recently proposed protein fold topology (McArdle et al., 2006), but to date, only few examples of such topologies have been established, mainly because the current computational methods do not adequately address the complex issue of fold similarity, as opposed to fold identity (for a review see (Kellenberger et al., 2011)).

At the same time, the methodology of inverse virtual screening based on conventional (all-atom) docking protocols was quickly expanded to test libraries of compounds against a panel of target proteins in order to evaluate the binding mode of several specific compounds (small compound libraries with 20-300 molecules) in order to achieve a reduction of cost and time requirement of a drug discovery project. However, while this approach is conceptually simple, the building of the panel of target proteins represents a critical step, because the targets must be carefully chosen if they do not belong to the same fold. Furthermore, analysis of the results is challenging and some elaborate post-processing strategies have been proposed (Lauro et al., 2011).

2.4 Expanding inverse virtual screening to larger compound libraries

In the present study, we sought to investigate the potential of inverse virtual screening in the identification of lead compounds from a moderately sized library (>1000 molecules) to reduce synthetic efforts and thus the costs of compound development and biological screening. Combined

into this effort, we also explored the feasibility of novel N-heteroaryl imidazoline scaffolds as protein kinase inhibitors. Protein kinases are ideal targets for the inverse virtual screening approach, since they comprise a panel of proteins that adopt a conserved general fold, and features such as the nucleotide binding site of individual kinases are structurally highly similar. Our belief that the identification of new chemotypes for protein kinases may be facilitated by IVS was bolstered by two recent reports in which this approach was employed to identify specific kinase targets from large panels using 3 and 10 ligands, respectively (Lauro et al., 2012; Zahler et al., 2007).

In particular, we explored whether it was possible to overcome the limitations of docking in predicting functional molecules by using a statistical approach that removes aspects of “individuality” from the predicted hits. Such a statistical approach would, in principle, allow the *in silico* screening and scoring of a focussed virtual library in excess of 1000 members, 1-2 orders of magnitude larger than any reports to date. By making use of the many experimental three-dimensional structures of kinases and docking the small-molecule library into multiple targets (i.e. panel docking), less emphasis is placed on individual docking hits, and the overall performance of a particular compound against kinases is evaluated. The hypothesis is that compounds which frequently score highly may be more likely to possess biological activity. Synthetic efforts would then be reduced to a small diverse set of the top scoring compounds in panel docking, and it was hoped that this would provide fast access to novel biologically active molecules.

3. Towards a next-generation approach

3.1 Novel chemotypes for kinase drug discovery

Addressing the need for novel chemical scaffolds in protein kinase drug discovery, we generated a reasonable sized virtual library of compounds, based on synthetically feasible combinations of (i) two imidazoline moieties attached to different azine cores, including symmetrical and non-symmetrical topologies, as well as (ii) combinations of an imidazoline moiety and a primary or secondary amine attached to varying central azine cores.

Considering the focus on 2-aminoazine and -azole scaffolds of current kinase inhibitor discovery, the structural core of the N-heterocycle allows variations such as 2-pyridyl, 2-pyrazinyl, 2-pyrimidinyl, and 4-pyrimidinyl and thus offers exploration of chemical scaffolds outside the previously explored chemical space. In particular, bis-imidazoliny azines such as **1** present an intriguing novel kinase inhibitor scaffold to increase selectivity, whereby one 2-imidazolinyazine moiety could serve as a transient Hinge segment binder while the other contributed additional interactions with the protein target (see Figure 2A).

A motivating factor in the design of this library was thus to interrogate chemistry space beyond established 2-aminoazine and -azole motifs. To this end, a non-classical motif based on an N-heteroaryl substituted 2-imidazoline, which we hypothesised may be particularly suited for binding to the protein kinase Hinge segments, was investigated. The 2-imidazoline moiety holds some promise as a novel scaffold for protein kinases as it possesses high solubility in aqueous environment (Remko et al., 2006) and, based on the reported broad range of biological activities against unrelated targets (Favaloro et al., 2003; Ferretti et al., 2002; Guo et al., 2008; Hong et al., 2005; Kahlon et al., 2009; Li & Zhang, 2011), can also be considered a privileged structure (Evans et al., 1988). Exploiting privileged structures with drug-like properties has been a popular approach to drug discovery, particularly within the pharmaceutical industry. The term was originally introduced (Evans et al., 1988) to describe motifs that could be elaborated into libraries of drug-like molecules (ligands) in which different members could selectively and potently interact with a variety of unrelated biological targets.

The design of this focussed library of 1235 small molecules (library MK1235; see Supplemental Procedures S1) was thus based on central di-nitrogen heterocyclic moieties, including pyrazine, pyrimidine and pyridine, and motivated by two main criteria. First, access to novel 2-imidazoline N-heteroarylation chemistry was recently secured (Krasavin, 2012) and, second, N-heteroaryl-2-imidazoline motifs such as **2** may become protonated in aqueous medium (see Figure 2B), and we

speculate that the resulting chemotype **3** may be binding to the Hinge segment of protein kinases in a non-classical fashion. Intriguingly, the binding pose obtained for one of the successful compounds in this case study (**9**) supports this notion (Figure 2D). The library was additionally populated with N-heteroaryl mono imidazolines **4** and **5** (see Figure 2C) where the azine periphery can be decorated with a simple secondary and primary amine (in lieu of the second imidazoline moiety in **1**). Such compounds, even in the absence of strong predicted kinase binding scores and experimentally measured kinase inhibition potency, could at least serve as a reference set and help reveal the unique value of the bis-imidazoline skeletons **1** as a hitherto undescribed scaffold for kinase inhibitor design. In consideration of decorating groups, care was taken to achieve sufficient diversity that enables conclusions as to structure-activity relationships.

3.2 Panel docking

Experimentally determined three-dimensional structures of kinase domains were retrieved from the PDB. The models were visually inspected for bound ligands (see Table 1), positioned and oriented with respect to a reference kinase domain, and all non-protein components were removed from the models. The library of 55 individual structures comprises 48 different kinases (see Supplemental Table S5).

In silico screening of the MK1235 library against the ATP-binding site of 55 structures from a kinase panel was conducted using AutoDock Vina (see Supplemental Procedures S2). The resulting scores were analysed by histogram analyses (Supplemental Figures S6 and S7) that considered the top five scored compounds for any kinase, based on the average score of each ligand for the predicted poses per individual kinase. From this histogram, 106 compounds were identified and subjected to cluster analysis to allow grouping based on Tanimoto similarity. Clustering was performed for both the “average score” and “highest score” docking selections (Supplemental Figure S8). Since the results based on the “average score” were found to generate more structurally dissimilar clusters on visual inspection, it was this clustering strategy that was used in the final nomination of the specific clusters for further study. The resulting 19 clusters with a population of >5 were appraised and the final five compounds (**6-10**) selected (Figure 3) that (i) represent their clusters, and (ii) provide a broad variety of different chemotypes (see Supplemental Figure S9).

3.3 *In vitro* kinase inhibition

Chemical synthesis of the selected compounds **6-10** required generation of imidazolines (Krasavin, 2012) that were coupled with aryl halides using the Buchwald-Hartwig reaction (Guram et al., 1995; Louie & Hartwig, 1995). For full details, as well as characterisation data, see Supplemental Procedures S3. The five compounds were then subjected to inhibition assays, using a panel of 48 kinases provided by Cerep (see Supplemental Procedures S4). Assays were also conducted in the presence of a specific inhibitor for each individual kinase, and the inhibition data of the five novel compounds were compared to the inhibition elicited by the control compounds. Final results were expressed as percentage inhibition relative to a control (see Figure 4). Instances where at least 30% inhibition compared to the control experiment were achieved, were considered as promising, and observed for six compound/kinase pairs: CDK1:**10**, CDK2:**10**, IRK:**6**, Src kinase:**9**, TAOK2:**10**, TRKA:**9**. Notably, two of the five compounds, **7** and **8**, could be excluded from the further work stream at this stage, as they only achieved a maximum of ~15% (**7**) and ~20% (**8**) inhibition as compared to control values.

For the six kinase-compound pairs identified above, dose-response data were then acquired (also at Cerep, France) in order to evaluate the IC₅₀ values of the novel molecules (see Figure 5). Two of the three compounds tested at this stage, showed reasonable efficacy with IC₅₀ values of around 14 μM (**9** and **10**). The third compound, **6**, possessed only a low effect (IC₅₀ = 0.1 mM).

4. Conclusion

Computational docking as a means to prioritise small molecules in drug discovery projects remains a highly popular *in silico* screening approach, despite its main shortcoming which is the general lack of a correlation between docking scores and biological activity.

Here, we investigated the possibility to improve the selection of small molecule compounds for a drug discovery project by computationally screening ligands from a virtual library comprised of N-heteroaryl-2-imidazoline derivatives against a panel of 48 mostly human kinases; the scheme shown in Figure 6 summarises this approach. In light of the exhaustion of the classical chemotypes of kinase inhibitors, the fundamentally new imidazolanyl azine chemotype is introduced here as promising starting point for kinase inhibitor design. Histogram analyses of the number of highest scoring compounds in the docking calculations were used to identify ~100 top scoring compounds. Subsequent grouping into clusters of maximal diversity enabled the selection of one representative each from five groups, which were synthesised and tested in kinase inhibition assays, with two of the five synthesised compounds achieving micro-molar inhibition in five human kinases.

In contrast to earlier inverse virtual screening approaches, where substantial numbers of compounds (26-265) needed to be synthesised (Muller et al., 2006; Urich et al., 2013), we present a case study of a highly economic and cost-effective pathway to the first lead molecules by putting less emphasis on the individual predicted docking poses but exploiting the statistical hit frequency. By synthesising just five compounds which were selected based on the results from panel docking, two compounds (**9** and **10**) with micro-molar IC₅₀ values were identified from the original library of 1235 members. These results also provide evidence that novel mono- and bis-imidazolanyl azines can serve as valid starting points for kinase inhibitor design.

Clearly, the panel docking approach as evaluated here does not abolish the need for a more rigorous prediction, or indeed experimental determination, of the binding pose of an individual ligand. The individual docking procedures in this survey have been carried out with conventional and widely used algorithms and scoring functions of AutoDock, and are thus subject to the inherent limitations of this methodology. Further more, in the analysis of the panel docking results, no attention is being paid to the individual binding pose of a ligand, but rather its ability to achieve many highly scored poses with different protein targets in the panel. We hypothesised that this might offer a new pathway to shortlist ligands from an initial library for synthesis in a drug discovery project. The results from this case study demonstrate that this methodology holds great promise for drug discovery projects in cases where a group of target proteins share high structural similarity in their binding sites. The methodology described here might provide a means of improving the efficiency and cost-effectiveness of target and lead identification. Since the concept of panel docking is reliant on a high degree of similarity in the structural fold of the target, the method should also be applicable to other target classes, such as serine/threonine phosphatases which have recently emerged as novel targets in the fight against parasitism (Campbell et al., 2011).

Acknowledgements

We thank David Camp for helpful discussions and critical reading of the manuscript. AH gratefully acknowledges funding by the National Health and Medical Research Council (APP1044022) and the Australian Research Council (DP140100599). AH and DG also gratefully acknowledge funding provided by ARC LIEF grant LE120100071.

Additional Information

Supplemental Procedures S1: Construction of compound library MK1235

Generation of the virtual library based on central di-nitrogen heterocyclic moieties.

Supplemental Procedures S2: Panel docking

Automated docking of library MK1235 to the panel of protein kinases.

Supplemental Procedures S3: Organic synthesis

Synthesis and characterisation of compounds **6-10**.

Supplemental Procedures S4: Protein kinase inhibition assays

Inhibition assays using the Cerep (France) panel of kinases.

Supplemental Table S5

The panel of three-dimensional structures of kinases used in the present study.

Supplemental Figure S6

Histograms of top five ligands based on average score of top poses per ligand using one model per fold.

Supplemental Figure S7

Histogram of top five ligands based on highest scoring pose per ligand using one model per fold.

Supplemental Figure S8

Groups of compounds obtained after panel docking and cluster analysis with description of the properties of the groups.

Supplemental Figure S9

Medicinal chemistry appraisal of and justification for the chosen five final clusters from which one compound each was synthesised.

References

- Andersen MD, Shaffer J, Jennings PA, Adams JA. Structural characterization of protein kinase A as a function of nucleotide binding. Hydrogen-deuterium exchange studies using matrix-assisted laser desorption ionization-time of flight mass spectrometry detection. *J. Biol. Chem.* 2001;276:14204-11.
- Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD. Flexible docking using Tabu search and an empirical estimate of binding affinity. *Proteins* 1998;33:367-82.
- Blom NS, Sygusch J. High resolution fast quantitative docking using Fourier domain correlation techniques. *Proteins* 1997;27:493-506.
- Böhm HJ. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J Comput Aided Mol Des* 1992;6:61-78.
- Breinbauer R, Vetter I, Waldmann H. From protein domains to drug candidates - Natural products as guiding principles in the design and synthesis of compound libraries. *Angew. Chem. Int. Ed. Engl.* 2002;41:2878-90.
- Camp D, Davis RA, Evans-Illidge EA, Quinn RJ. Guiding principles for natural product drug discovery. *Future Med. Chem.* 2012;4:1067-84.
- Campbell, B.E., McCluskey A, Hofmann A, Gasser R. Serine/threonine phosphatases in socioeconomically important parasitic nematodes – prospects as novel drug targets? *Biotechnol. Adv.* 2011;29:28-39.
- Chen YZ, Zhi DG. Ligand–protein inverse docking and its potential use in the computer search of protein targets of a small molecule. *Proteins* 2001;43:217-26.
- Cohen P, Alessi DR. Kinase drug discovery--what's next in the field? *ACS Chem. Biol.* 2013;8:96-104.
- Defranchi E, Schalon C, Messa M, Onofri F, Benfenati F, Rognan D. Binding of protein kinase inhibitors to synapsin I inferred from pair-wise binding site similarity measurements. *PLoS One* 2010;5:e12214.
- DeLano W. The PyMOL Molecular Graphics System. <http://www.pymol.org> 2002..
- Eisen MB, Wiley DC, Karplus M, Hubbard RE. HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding site. *Proteins* 1994;19:199-221.
- Ekins S, Mestres J, Testa B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br. J. Pharmacol.* 2007;152:9-20.
- Evans BE, Rittle KE, Bock MG, DiPardo RM, Freidinger RM, Whitter WL, et al. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med.*

Chem. 1988;31:2235-46.

Ewing TJ, Makino S, Skillman AG, Kuntz ID. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des* 2001;15:411-28.

Favaloro JL, Andrews KL, McPherson GA. Novel imidazoline compounds that inhibit Kir-mediated vasorelaxation in rat middle cerebral artery. *Naunyn Schmiedeberg Arch. Pharmacol.* 2003;367:397-405.

Ferretti G, Dukat M, Giannella M, Piergentili A, Pignini M, Quaglia W, et al. Homoazanicotine: a structure-affinity study for nicotinic acetylcholine (nACh) receptor binding. *J. Med. Chem.* 2002;45:4724-31.

Gehlhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DB, Fogel LJ, et al. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* 1995;2:317-24.

Guo X, Shi L, Wang R, Liu X, Li B, Lu X. Synthesis and biological activities of novel nonpeptide angiotensin II receptor antagonists based on benzimidazole derivatives bearing a heterocyclic ring. *Bioorg. Med. Chem.* 2008;16:10301-10.

Guram AS, Rennels RA, Buchwald SL. A simple catalytic method for the conversion of aryl bromides to arylamines. *Angew. Chem. Int. Ed. Engl.* 1995;34:1348-50.

Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* 1995;9:576-96.

Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 1988;241:42-52.

Hong S, Bavadekar SA, Lee S, Patil PN, Lalchandani SG, Feller DR, et al. Bioisosteric phentolamine analogs as potent alpha-adrenergic antagonists. *Bioorg. Med. Chem. Lett.* 2005;15:4691-5.

Hui-fang L, Qing S, Jian Z, Wei F. Evaluation of various inverse docking schemes in multiple targets identification. *J. Mol. Graphics Modell.* 2010;29:326-30.

Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 1997;267:727-48.

Kahlon DK, Lansdell TA, Fisk JS, Hupp CD, Friebe TL, Hovde S, et al. Nuclear factor-kappaB mediated inhibition of cytokine production by imidazoline scaffolds. *J. Med. Chem.* 2009;52:1302-9.

Kellenberger E, Hofmann A, Quinn RJ. Similar interactions of natural products with biosynthetic enzymes and therapeutic targets could explain why nature produces such a large proportion of existing drugs. *Nat. Prod. Rep.* 2011;28:1483-92.

Kinnings SL, Liu N, Buchmeier N, Tonge PJ, Xie L, Bourne PE. Drug discovery using chemical systems biology: repositioning the safe medicine Comtan to treat multi-drug and extensively drug resistant tuberculosis. *PLoS Comput. Biol.* 2009;5:e1000423.

Koonin EV, Wolf YI, Karev GP. The structure of the protein universe and genome evolution. *Nature* 2002;420:218-23.

Krasavin M. Novel diversely substituted 1-heteroaryl-2-imidazolines for fragment-based drug discovery. *Tetrahedr. Lett.* 2012;53:2876-80.

Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* 1982;161:269-88.

Lauro G, Masullo M, Piacente S, Riccio R, Bifulco G. Inverse Virtual Screening allows the discovery of the biological activity of natural compounds. *Bioorg. Med. Chem.* 2012;20:3596-602.

Lauro G, Romano A, Riccio R, Bifulco G. Inverse virtual screening of antitumor targets: Pilot study on a small database of natural bioactive compounds. *J. Nat. Prod.* 2011;74:1401-7.

Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* 2007;6:881-90.

Li B, Liu Y, Uno T, Gray N. Creating chemical diversity to target protein kinases. *Comb. Chem. High Throughput Screen.* 2004;7:453-72.

Li J, Zhang Y. Imidazoline I2 receptors: target for new analgesics? *Eur. J. Pharmacol.* 2011;658:49-

- Lorber DM, Shoichet BK. Hierarchical docking of databases of multiple ligand conformations. *Curr Top Med Chem* 2005;5:739-49.
- Louie J, Hartwig JF. Palladium-catalyzed synthesis of arylamines from aryl halides. Mechanistic studies lead to coupling in the absence of tin reagents. *Tetrahedron Lett.* 1995;36:3609-12.
- Macarron R, Banks MN, Bojanic D, Burns DJ, Cirovic DA, Garyantes T, et al. Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* 2011;10:188-95.
- McArdle B, Campitelli M, Quinn RJ. A Common Protein Fold Topology Shared by Flavonoid Biosynthetic Enzymes and Therapeutic Targets. *J. Nat. Prod.* 2006;69:14-7.
- McMartin C, Bohacek RS. QXP: powerful, rapid computer algorithms for structure-based drug design. *J. Comput. Aided Mol. Des.* 1997;11:333-44.
- Miller MD, Kearsley SK, Underwood DJ, Sheridan RP. FLOG: a system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. *J. Comput. Aided Mol. Des.* 1994;8:153-74.
- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* 2009;30:2785-91.
- Muller P, Lena G, Boilard E, Bezzine S, Lambeau G, Guichard G, et al. In silico-guided target identification of a scaffold-focused library: 1,3,5-triazepan-2,6-diones as novel phospholipase A2 inhibitors. *J. Med. Chem.* 2006;49:6768-78.
- Oprea TI. Current trends in lead discovery: are we looking for the appropriate properties? *Mol. Divers.* 2002;5:199-208.
- Oshiro CM, Kuntz ID, Dixon JS. Flexible ligand docking using a genetic algorithm. *J. Comput. Aided Mol. Des.* 1995;9:113-30.
- Paul N, Kellenberger E, Bret G, Muller P, Rognan D. Recovering the true targets of specific ligands by virtual screening of the protein data bank. *Proteins* 2004;54:671-80.
- Remko M, Swart M, Bickelhaupt FM. Theoretical study of structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives. *Bioorg. Med. Chem.* 2006;14:1715-28.
- Schellhammer I, Rarey M. FlexX-Scan: fast, structure-based virtual screening. *Proteins* 2004;57:504-17.
- Urich R, Wishart G, Kiczun M, Richters A, Tidten-Luksch N, Rauh D, et al. De novo design of protein kinase inhibitors by in silico identification of hinge region-binding fragments. *ACS Chem. Biol.* 2013;8:1044-52.
- Wang J, Kollman PA, Kuntz ID. Flexible ligand docking: a multistep strategy approach. *Proteins* 1999;36:1-19.
- Warren GL, Andrews CW, Capelli A, Clarke B, LaLonde J, Lambert MH, et al. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* 2006;49:5912-31.
- Welch W, Ruppert J, Jain AN. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. *Chem. Biol.* 1996;3:449-62.
- Zahler S, Tietze S, Totzke F, Kubbutat M, Meijer L, Vollmar AM, et al. Inverse in silico screening for identification of kinase inhibitor targets. *Chem. Biol.* 2007;14:1207-14.
- Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* 2009;9:28-39.

Table 1. Appraisal and preparation of kinase domains with known three-dimensional structures for panel docking in this study.

Site of ligand	No of structures	
	Total	Used
ATP-binding site	43	38
ATP-binding site and back pocket	6	6
ATP-binding site and other	1	1
ATP-binding site and substrate binding site	7	6
Substrate binding site	3	2
Other	1	1
None	8	4
Total	69	58

Figure legends

Figure 1

Illustration of fold and important structural elements of eukaryotic protein kinases, using protein kinase A as a representative example. The N-terminal lobe is coloured pale blue, and the C-terminal lobe is shown in pale brown. Residue side chains of the catalytic triad are rendered as stick models. The C-helix is coloured magenta, the activation loop is blue, the Hinge (Gly-rich loop) is orange, and the C-terminal tail is shown in pale red.

ATP binding in the nucleotide binding site between the Hinge and catalytic triad causes a clockwise rotation of the C-terminal tail, thus compacting the binding site and stabilising inter-lobe contacts (Andersen et al., 2001). Figure prepared with PyMOL (DeLano, 2002).

Figure 2

A Bis-imidazoliny azines may provide additional sites of interaction with target kinases.

B Protonation of N-heteroaryl-2-imidazoline motifs may enable binding to the Hinge segment of protein kinases in a non-classical fashion.

C N-heteroaryl mono imidazolines were included in the library in this study and, in the absence of inhibitory potency, may be used as probes.

D The binding poses of bis-imidazoliny azines and N-heteroaryl mono imidazolines obtained in the docking calculations support the hypothesised unique interaction features of these chemotypes. Shown is the binding pose of **9** obtained with TRKA (PDB accession code 4aoj); the dashed lines indicate polar interactions with distances $<3.8 \text{ \AA}$. The functionalisation of **1** with a benzyl moiety on **R** appears to be an added benefit of compound **9**, as this group is engaged in hydrophobic interactions with Phe251 of TRKA.

Figure 3

Compounds synthesised and tested, based on results from panel docking in this study. Compounds **6-10** represent the clusters of molecules shown in Supplemental Figures S8, S9.

Figure 4

Results of inhibition assays for each of the compounds **6-10** against 48 kinases. Every increment on the x-axis represents an individual kinase. For compounds **6**, **9** and **10**, inhibition of more than 30% was achieved for the kinases labelled.

Figure 5

Dose-response data for six kinase-compound pairs at five different compound concentrations (0.3 μM , 3 μM , 10 μM , 30 μM , 0.1 mM). Data points represent the average of two technical replica. The IC_{50} values for the pairings Src:**9**, TRKA:**9**, CDK1:**10**, CDK2:**10**, TAOK2:**10** are around 14 μM .

Figure 6

Schematic comparison of conventional single docking and the panel docking methodology applied in this study.

Schematic comparison of conventional single docking and the panel docking methodology applied in this study.

Figure 1

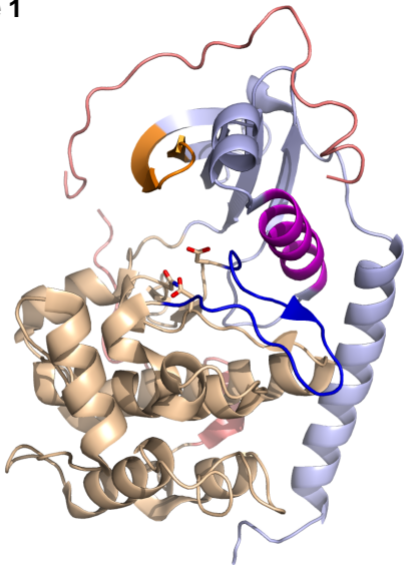
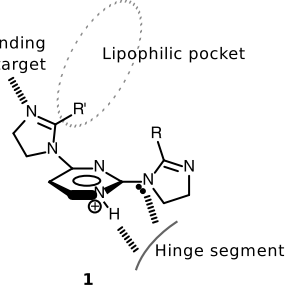
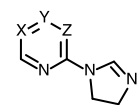
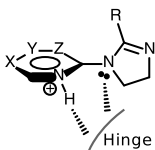
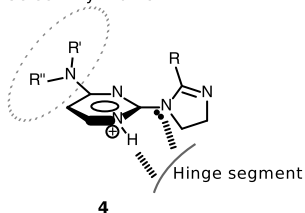


Figure 2additional H-bonding
with target

Lipophilic pocket

**B** $pK_a \approx 4.5$ (calc.)aqueous
medium**C**

Selectivity motifs

additional H-bonding
with target

Lipophilic pocket

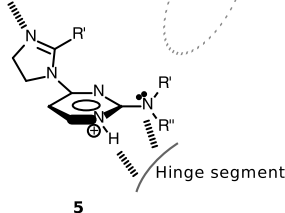
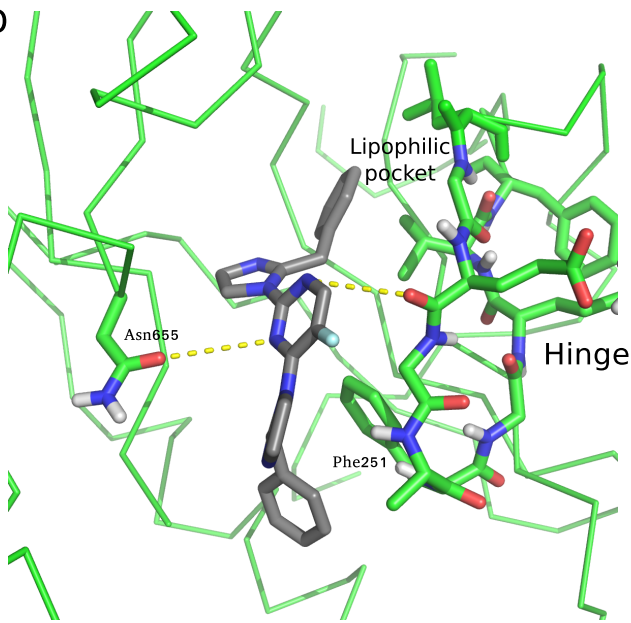
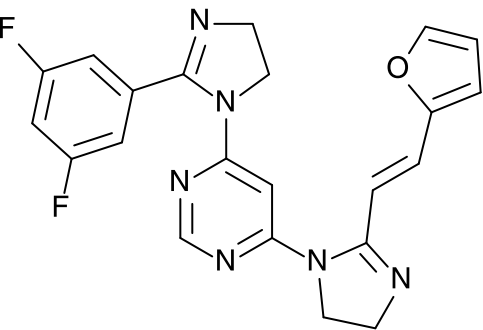
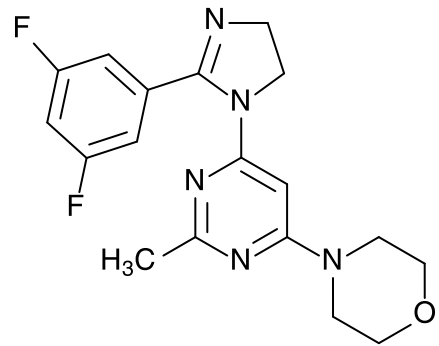
**D**

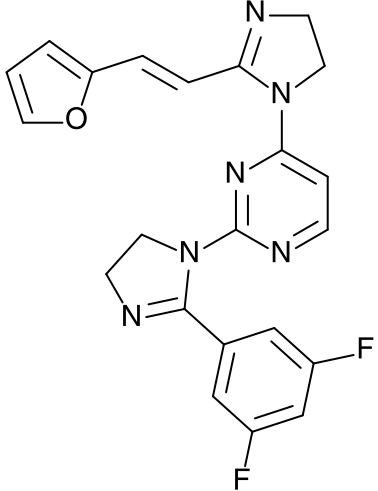
Figure 3



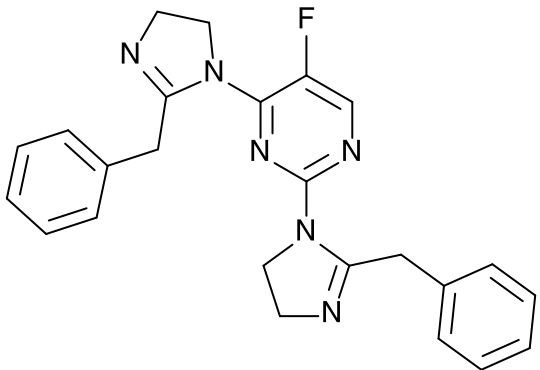
6



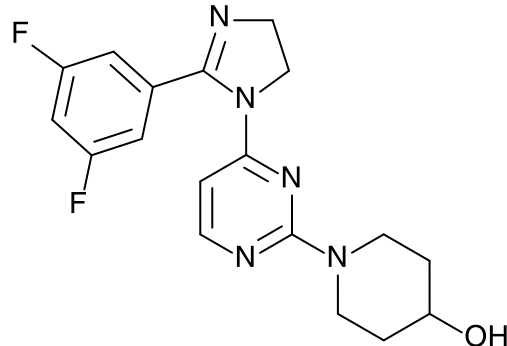
7



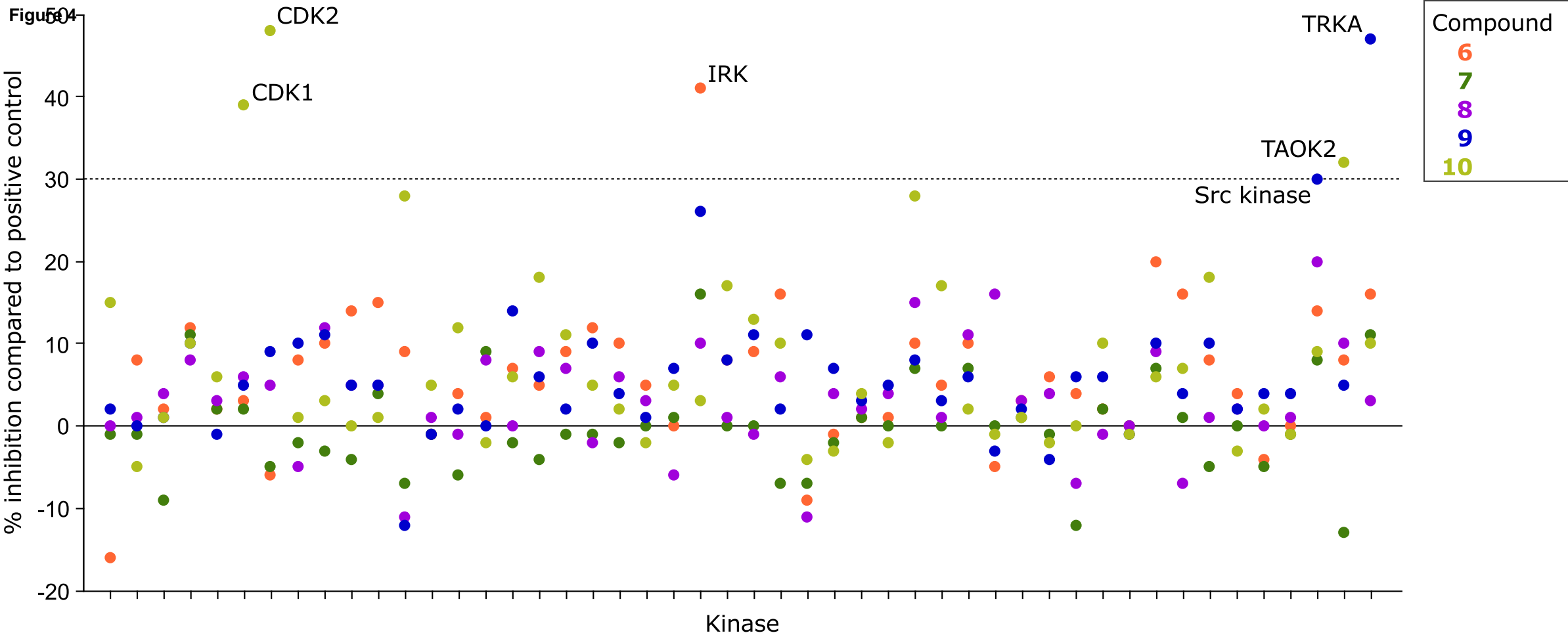
8



9



10



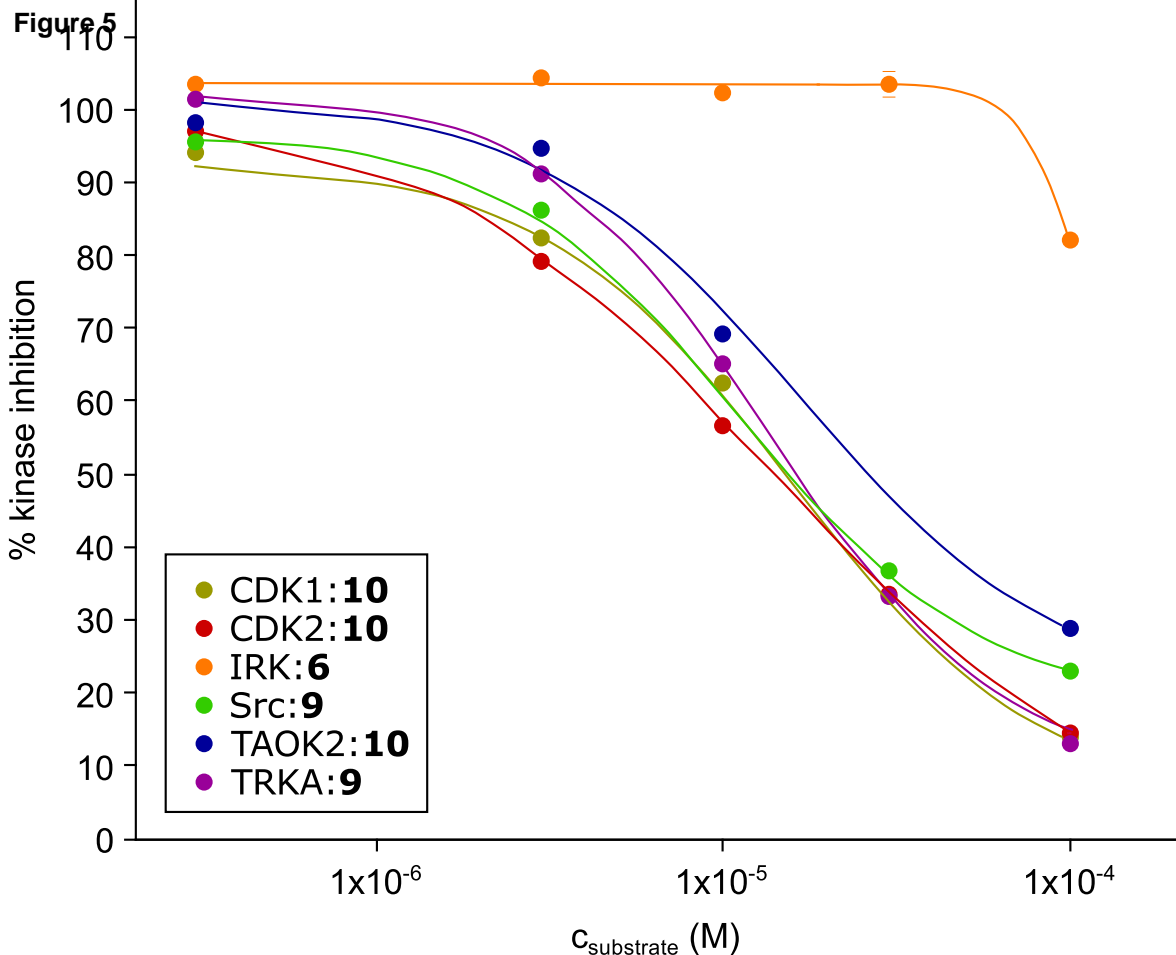
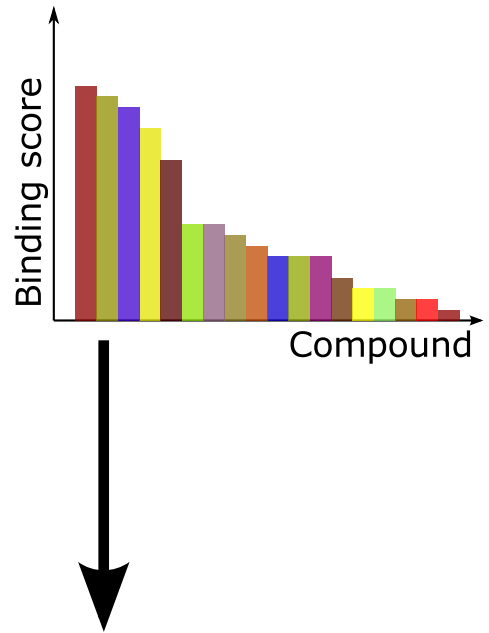
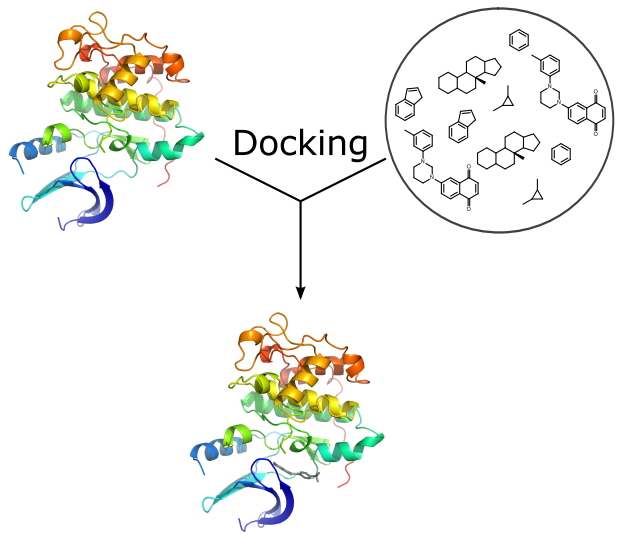
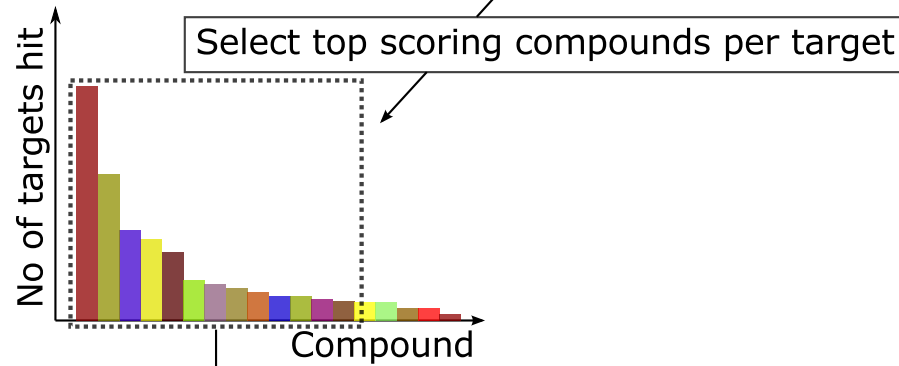
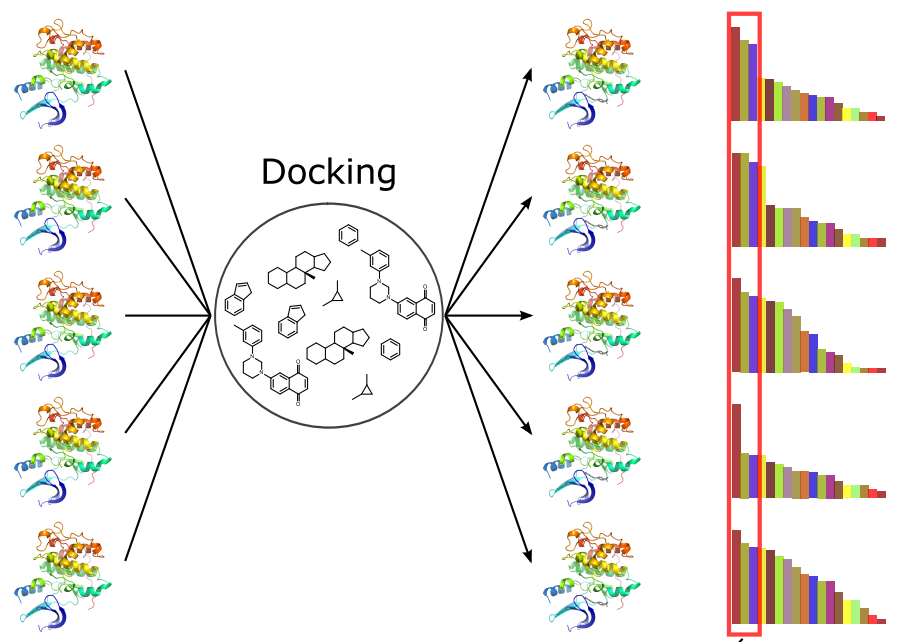


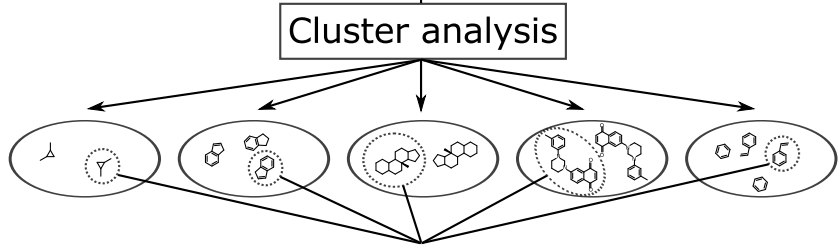
Figure 16
Single target docking



Panel docking



Select top scoring compounds per target



Synthesis

Kinase assays