HHS Public Access

Author manuscript

J Med Chem. Author manuscript; available in PMC 2020 February 14.

Published in final edited form as:

J Med Chem. 2019 July 25; 62(14): 6512–6524. doi:10.1021/acs.jmedchem.9b00089.

Kinase Atlas: Druggability Analysis of Potential Allosteric Sites in Kinases

Christine Yueh[1], **Justin Rettenmaier**[2], **Bing Xia**[1], **David R Hall**[3], **Krister Barkovich**[2], **Gyorgy Keseru**[4], **Adrian Whitty**[5], **James A. Wells**[2], **Sandor Vajda**[1,5,*], **Dima Kozakov**[6,7,*]

[1]Department of Biomedical Engineering, Boston University, Boston, MA 02215

[2]Departments of Pharmaceutical Chemistry and Cellular and Molecular Pharmacology, University of California, 1700 Fourth Street, San Francisco, CA 9415

[3]Acpharis Inc., Holliston, MA 01746

[4]Medicinal Chemistry Research Group, Research Center for Natural Sciences, Magyar tudósok krt. 2. H-1117 Budapest, Hungary

[5]Department of Chemistry, Boston University, Boston, MA 02215

[6]Department of Applied Mathematics and Statistics, Stony Brook University, NY 11794

[7]Laufer Center for Physical and Quantitative Biology, Stony Brook University, NY 11794

Abstract

The inhibition of kinases has been pursued by the pharmaceutical industry for over 20 years. While the locations of the sites that bind type II and III inhibitors at or near the ATP binding sites are well defined, the literature describes ten different regions that were reported as regulatory hot spots in some kinases and thus are potential target sites for type IV inhibitors. Kinase Atlas is a systematic collection of binding hot spots located at the above ten sites in 4910 structures of 376 distinct kinases available in the Protein Data Bank. The hot spots are identified by FTMap, a computational analogue of experimental fragment screening. Users of Kinase Atlas (https://kinase-atlas.bu.edu) may view summarized results for all structures of a particular kinase, such as which binding sites are present and how druggable they are, or they may view hot spot information for a particular kinase structure of interest.

Graphical Abstract

[*]**Corresponding Authors** For Dima Kozakov: Phone: 1-617-319-7952, midas@laufercenter.org, For Sandor Vajda: Phone: 1-617-353-4757, vajda@bu.edu.

Author Contributions

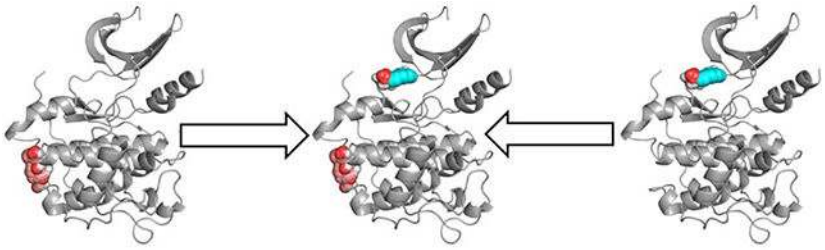
C.Y. and B.X. performed the analysis of kinases and built the Kinase Atlas database. J.R. and K.B. performed CDK2 tethering experiments. J.W. planned and supervised the experimental work. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acpharis Inc. offers commercial licenses to ATLAS, a software product in function similar to FTMap. Dima Kozakov and Sandor Vajda owns stock in the company. However, the FTMap program and the use of the FTMap server and the Kinase Atlas database are free for academic and governmental use.

ASSOCIATED CONTENT

Supporting Information

Figure S1: Diagram representing the definition of the 12 putative allosteric sites. Figure S2: Probe molecules used by FTMap; Figure S3: Flowchart of developing Kinase Atlas. Figure S4: Tethering, a disulfide-based fragment screening method.



Keywords

Kinase allosteric sites; type IV kinase inhibitors; binding hot spots; protein mapping; FTMap; PIF pocket of CDK2

INTRODUCTION

Members of the protein kinase family play vital roles in cellular physiology and are major drug targets,[1] as they have been implicated in many types of diseases such as cancer, diabetes, neurodegeneration, and inflammation.[2] Their association with a wide variety of ailments stems from their involvement in nearly all cellular processes, since they are responsible for regulating the activity of other proteins.[3] Although the development of new kinase inhibitors is a major focus in pharmaceutical research, a large number of kinases remained so far unexplored in drug discovery projects.[4] Currently, 37 small molecules have been approved by the FDA as kinase inhibitors,[5] and the vast majority of these target the active site, which is located between the N- and C-terminal domains and binds ATP.[6] Inhibitors that bind to the ATP pocket in the active form of the kinase are known as “type I” kinase inhibitors [1], and despite their popularity, the development process presents two major challenges: first, type I inhibitors must bind with enough potency to overcome high physiological concentrations of ATP, and second, the ATP site is highly conserved between all kinases, making it difficult to design inhibitors with enough selectivity to bind only to their intended targets.[7] Alternatively, kinase inhibitors can be classified as “type II” if they target the ATP site but bind to an inactive conformation known as “DFG-out”, in which a conserved DFG (Asp-Phe-Gly) motif partially blocks the ATP pocket.[1] Type II inhibitors, which make up about a quarter of approved kinase inhibitors,[5] often extend into the hydrophobic back pocket formed in the “DFG-out” state. As this pocket is less conserved than the active site, binding here offers the potential to achieve sufficient selectivity more easily.[8] Allosteric kinase inhibitors, which do not target the ATP site at all, are classified by where they bind: type III inhibitors bind to the active kinase conformation in a pocket adjacent to the ATP site, whereas type IV inhibitors bind away from the ATP site entirely.[1] Although type III and IV inhibitors would potentially face fewer issues with potency and/or selectivity than ATP-competitive inhibitors, they remain far less common: only two of the FDA-approved kinase inhibitors (trametinib and cobimetinib) would be considered type III,[9] and no inhibitors that bind to the kinase catalytic domain would be considered type IV.[1]

Here we have explored the potential of kinase allosteric sites as targets for inhibition, since they have been relatively underutilized for kinase drug development. To detect and assess the

druggability potential of these sites, we used FTMap, a computational analogue of experimental fragment screening,[10] to identify binding hot spots on all available kinase structures in the Protein Data Bank.[11] Hot spots are small regions within a binding site that contribute disproportionately to the binding free energy, and they can be detected in a protein structure even without ligand binding information or an obvious pocket, since they tend to bind a large number and variety of small molecules even in an unliganded state. FTMap finds consensus sites, which are regions on the protein surface that bind a large number of small molecule probe clusters; thus, we can use FTMap to identify binding hot spots even in unliganded structures, as well as to estimate the potency with which they could bind potential ligands, since consensus site strength relates to potential binding affinity.[12]

In addition to the pockets associated with type II and III inhibitors, we have identified ten sites on the kinase catalytic domain that have been described in the literature as being involved in either regulating the activity of a kinase, or the ability of a kinase to regulate the activity of its substrates. Seven of these sites are known to bind compounds that exhibit IC_{50} values in the micromolar or even nanomolar range, whereas the other three should be considered more speculative. Here we do not investigate the signal transduction mechanisms from the allosteric binding site to the kinase active site that may require extensive simulations,[13, 14] but focus on the detection of sites that can bind ligands with some measure of druggability.[15, 16] We consider a site druggable if it includes a primary hot spot with at least 16 probe clusters and at least one additional ho spot nearby. As shown in our earlier work,[12] these conditions guarantee that the site is potentially capable of binding a drug-size molecule with at least micromolar affinity, although this affinity may not be sufficient for developing a successful drug.[17] In addition, we emphasize that this notion of druggability does not imply that the binding of any ligand at the site will modulate the behavior of the target protein.[12] However, we assume that since the structure of the kinase catalytic domain is fairly conserved,[7] an allosteric site found on one kinase may be present in the same location on other kinases, although the structures and sequences of these analogous pockets would differ between kinases -- unlike the ATP site, which is relatively similar between all kinases. This is already known to be true of the DFG-out pocket, for example, which is found in many different kinases,[8] and many of the other potential allosteric sites we have identified are also associated with more than one kinase.

Our FTMap results for all kinase structures have been made available online as the Kinase Atlas (https://kinase-atlas.bu.edu), intended as a resource for researchers interested in kinases. Users may view summarized results for all structures of a particular kinase, such as which allosteric sites are present on the kinase and how druggable they are, or they may view or download FTMap results for a particular kinase structure of interest. The Kinase Atlas can thus be used either to discover where to target a kinase of interest, or whether a kinase is worth targeting at all, as some kinases do not have any structures with strong enough consensus sites to be considered druggable. In comparison to the active site, allosteric sites are rarely considered as targets for inhibitor development, but the Kinase Atlas shows that many of these pockets have the potential to be druggable, and may be worth pursuing. In particular, we present experimental data to show that the PIF pocket region, an allosteric site associated with PDK1,[18] but also found by FTMap in CDK2, is able to bind ligands leading to the discovery of novel micromolar type IV inhibitors. We also

demonstrate that our results can be used to find sites where binding has validated regulatory impact as reported in the literature.

RESULTS

The Kinase Atlas contains FTMap results for 3887 unique PDB IDs, corresponding to 4910 total kinase structures from 376 different kinases. The number of kinase structures is higher than the number of PDB IDs because some of the PDB files include multiple structures (multiple chains) in the unit cell. The consensus sites of probe clusters are determined by the protein mapping program FTMap and are assigned to 12 different putative binding sites that are potentially allosteric as defined on the basis of kinase literature (Figure 1). Each of these 12 sites was found in, or is most widely associated with, a particular kinase defined as the "source" kinase for the site (Table 1), although most sites can also be found in a variety of kinases. FTMap was able to identify hot spots located at each site in an unliganded structure of the "source" kinase for all sites. Further information on naming and selection of putative allosteric sites is given in the Discussion and in Figure S1.

As mentioned, the number of probe clusters in a ligand binding site predicts the druggability of the site. Not every binding site with druggable potential will have a liganded structure or binding data available, so being able to detect binding hot spots in unliganded structures is often useful in determining whether a protein is likely to be druggable as well as which regions to target.[12] We note that some of the strongest binding hot spots in all kinases are at the ATP binding site between the N and C lobes of the kinase domain. Since we are interested in allosteric sites, the two lobes are mapped separately, thereby decreasing the fraction of the FTMap probes that cluster in ATP binding pocket and increasing the relative importance of potential allosteric sites in other surface regions of the protein.

Each binding site and its FTMap results are described briefly in Table 1, shown in Figure 1, and in more detail in the following sections. Most of them were found to be druggable in at least one unliganded structure (see Figure 2); the structure with the strongest consensus site(s) for each allosteric site is listed in Table 1 and shown in Figure 3. The number of kinases that were found to be druggable at each allosteric site varied widely, with some sites being found in as few as 6 kinases, and others in well over 100 kinases, as seen in Figure 2 and Table 2. An allosteric site being "common" does not necessarily render it unsuitable as a target, however, as allosteric sites are not as conserved as the ATP site is between different kinases. Below we describe the 12 sites that are examined for the existence of hot spots. The first nine of these sites have known inhibitors in some kinases, whereas the last three sites have been reported to have some allosteric function.

DFG (DFG-out pocket)

The DFG-out pocket is a hydrophobic pocket that opens up when the conserved DFG motif changes conformation as a kinase switches from an active to an inactive state. In the active, DFG-in state, this site is occupied by the Phe residue in the DFG motif, and ATP can bind to the active site; in the inactive, DFG-out state, however, Phe instead partially occupies the ATP site, preventing ATP from binding and exposing the less conserved DFG-out pocket.[6] Type II inhibitors, which bind to the DFG-out conformation, are ATP-competitive, but they

frequently extend into the DFG-out pocket.[1] Since such inhibitors stabilize the DFG-out conformation of the DFG loop and lead to inhibition, the DFG-out pocket is sometimes considered allosteric,[8] in spite of the non-allosteric inhibition mechanism. As this site differs more between kinases than the ATP site, binding here may allow inhibitors to be more selective, although existing type II inhibitors do not necessarily have improved selectivity.[19] FTMap results for known targets of type II inhibitors suggest that the DFG-out pocket may not be particularly strong; many structures were shown to be only borderline druggable at this site, although some kinases had apo structures with strong enough consensus sites for this site to be considered druggable. The ability of type II inhibitors to bind with high affinity (for example, imatinib exhibits an IC_{50} value of 10.8 nM for c-Abl)[20] may be due mostly to the ATP pocket, with the DFG-out pocket playing a supporting role. In fact, since type II inhibitors bind to both the ATP and DFG-out pockets, it is difficult to separate the contributions of the corresponding inhibitor moieties to the binding free energy. However, in most kinases that bind both type I and type II inhibitors high binding affinity is already achieved by type I inhibitors that do not expand into the DFG-out pocket. Thus, while type II inhibitors generally do not further improve binding, they can provide improved selectivity

MT3 (MEK1/2 Type III inhibitor site)

This site is based on where type III kinase inhibitors bind; adjacent to the ATP and DFG-out sites and between the N- and C-terminal domains. Binding to this pocket in MEK1/2 disrupts the salt bridge between the conserved lysine residue K97/101 on the β3-strand and the conserved glutamate E114/118 on the αC-helix that is required for kinase activity.[6] Unlike the DFG-out pocket, this site is present in the active form of the kinase, and ligands can bind here even if the kinase is already bound to ATP.[1] FDA-approved type III inhibitors include trametinib and cobimetinib, both of which target MAPK/ERK kinase (MEK1/2) and have been used to treat melanoma.[9] Mapping results show that this site is highly druggable, with strong consensus sites present even in unliganded structures of MEK1/2, which is consistent with experiments showing that type III inhibitors can be highly potent (IC_{50} = 0.9 nM for cobimetinib/MEK1).[21] An analogous pocket appears to be present in EGFR, which binds the inhibitor EAI045 with IC_{50} values as low as 3 nM.[22].

PIF (PDK1 Interacting Fragment)

The PIF pocket is a hydrophobic groove found on the N-terminal domain of phosphoinositide-dependent kinase 1 (PDK1), which uses this pocket to recruit the C-terminal hydrophobic motif (HM) on other members of the AGC kinase family and thereby regulate their activity through phosphorylation.[23] PDK1 activity could thus be modulated through inhibitor binding at this site, as this would disrupt its interactions with its substrates and prevent PDK1 from either activating or inhibiting other kinases. Both activators [23] and inhibitors[24,18] have been developed for the PIF pocket, with many in the low affinity range (as low as Kd = 1.5 uM).[18] These compounds could likely be optimized for stronger binding, as mapping results indicate that this site is likely to be capable of binding compounds with higher affinity. Mapping showed that this was also true for other AGC kinases, many of which are known to possess a similar pocket.

CMP (c-Abl Myristoyl Pocket)

The myristoyl pocket in Abelson tyrosine-protein kinase 1 (c-Abl) binds the myristoyl group from the N-terminal cap of the kinase, allowing SH3 and SH2 domains to associate and induce an autoinhibited state. This N-terminal cap is not present, however, in the fusion BCR-Abl oncogene, which results from a chromosomal translocation that fuses the breakpoint cluster region (BCR) with c-Abl; BCR-Abl thus cannot be autoinhibited by myristoylation, and its elevated activity leads to disorders such as chronic myelogenous leukemia (CML).[25] CML has been treated with imatinib and other ATP-competitive inhibitors, but mutations near the ATP pocket in BCR-Abl are common and frequently lead to drug resistance.[26] As the myristoyl pocket is located in the C-terminal domain, far from the ATP site, these mutations would be less likely to affect binding there. GNF-2 has been found to bind at the myristoyl pocket in BCR-Abl, and similarly to the myristoyl group, it encourages SH2 and SH3 domain binding; it also induces conformational changes that promote binding of ATP-competitive inhibitors.[27] Alternatively, c-Abl can be activated by the binding of bulkier groups to the myristoyl pocket, such as DPH, which leads to conformational changes that prevent SH domains from associating with and inhibiting the kinase. [28] Although FTMap places a hot spot at the myristoyl pocket in most Abl kinase structures, the site is predicted to be at best borderline druggable. Nevertheless, a series of small molecules was shown to bind to the BCR–Abl myristoyl pocket and inhibit kinase activity via an allosteric mechanism,[29–31] in the case of ABL001 with high affinity.[30, 31] Even better results have been achieved by combining allosteric inhibitors with ATP-competitive inhibitors imatinib or nilotinib as such combinations can overcome resistance to either agent alone.[27, 30–32]

DRS (D-Recruitment Site)

The D-recruitment site is found in the C-terminal domain of all mitogen-activated protein (MAP) kinases, a family that includes extracellular signal-regulated kinases (ERKs), c-Jun N-terminal kinases (JNKs), and p38 MAPKs. Its name comes from the “D-motif” sequences found in MAPK binders, and it contains two subsites to which D-motifs bind, an acidic patch and a hydrophobic pocket.[33] MAP kinases control intracellular responses to extracellular stimuli, phosphorylating their substrates within the cell after being activated by their upstream regulators, MAPK kinases (MKKs).[34] Their interactions with other proteins in the MAPK signal transduction pathway are thus critical to the regulatory role MAPKs play in most cellular processes, and the D-recruitment site is a major docking site for these interactions.

Inhibitors have been developed to target the D-recruitment site, including peptides such as pepJIP1[35] and small molecules such as BI-78D3.[36] These inhibitors are mimetics of JIP1, a scaffold protein that enhances signaling in JNK, and compete with JIP1 to inhibit JNK activity. A small molecule natural product, rooperol, has also been shown to inhibit p38a MAP kinase.[37] None of these bind with the affinity that would be required for an appropriate drug candidate, however, with BI-78D3 being the most potent (IC50 = 280 nM).[36] Mapping results for MAPK structures suggest that developing compounds with greater affinity may not be likely, as the D-recruitment site was shown to be borderline druggable at best at the hydrophobic pocket, and not druggable at all at the acidic patch. Stronger binders that bind

at the D-recruitment site with high affinity (IC50 = 18 nM) have been reported, but these are long molecules that also extend into the ATP site, rather than targeting the D-recruitment site alone.[38] The D-recruitment site is thus not a promising pocket to be targeted on its own for MAP kinases, but this region was shown to have the potential to bind compounds with higher affinity in related kinases, such as members of the MKK family.

DEF (Docking site for ERK FXF)

The DEF site is found in the C-terminal domains of several members of the MAP kinase family, such as ERK1/2, p38a MAPK, and JNK1.[39, 40] It is also known as the FXF site, since it binds to the FXF motif found on several MAPK substrates, and it is located near the MAP kinase insert, where it appears after phosphorylation activates and induces conformational changes in ERK.[41] Compounds that bind to MAP kinases at the DEF site have been reported, such as biaryl tetrazoles identified by Comess et al that inhibit JNK1 activation by MKK7 (IC50 = 7.7 uM),[42] and mapping results suggest that the DEF site has the potential to be highly druggable, unlike the other major MAPK docking site, the D-recruitment site.

LBP (Lipid Binding Pocket)

The lipid binding pocket is located in the C-terminal domain of p38a MAP kinases, near the MAP kinase insert. The biological relevance of this pocket has not been verified conclusively, but it appears to be able to accommodate a variety of lipids, and it has been suggested that binding different lipids may affect p38a MAPK's catalytic activity and preference for specific substrates, particularly in cellular processes that involve lipids.[43] Several lipid-based molecules have been found to activate p38a MAPK upon binding to this site--such as phosphatidylinositol ether lipid analogues (PIAs) and perifosine, a phase II AKT/PKB inhibitor structurally similar to PIAs--by inducing conformational changes that lead to autophosphorylation.[44] These lipids bind with low affinity to p38a MAPK (IC50 = 1.2 uM), but mapping results indicate that this site would likely be able to bind ligands with higher affinity, since many structures of p38a MAPK had strong consensus sites at the lipid binding pocket.

PDIG (PDIG motif site)

This pocket is located near the PDIG motif in the C-terminal domain of Checkpoint kinase 1 (Chk1), an important regulator in the DNA damage response pathway. It appears to act as a substrate recognition site, and so Chk1 activity could potentially be inhibited by compounds that target this site and compete with substrate binding.[45] Several inhibitors have been developed that bind to this site with low micromolar affinity (as low as Ki = 146 nM), such as thioquinazolinones,[46] carbamates, and semicarbazides,[45] and mapping results for Chk1 structures suggest that the site has the potential to bind compounds with even greater affinity. A similar pocket appears to be present on PIM1 kinase, which was found to bind mitoxantrone, an FDA-approved chemotherapy drug that targets type II topoisomerase.[47] Mitoxantrone binds to two locations on PIM1, the substrate binding site (analogous to the one on Chk1) and the ATP site, with high affinity; this is in agreement with mapped structures of PIM1, which have strong consensus sites in this pocket.

EDI (EGFR Dimerization Interface)

Members of the epidermal growth receptor family (EGFR) are activated upon formation of an asymmetric dimer between two EGFR kinases, in which the C-terminal domain of one kinase interacts with the N-terminal domain of the second kinase. This dimerization is considered analogous to the interaction between cyclin and cyclin-dependent kinase (CDK), which activates CDK, and the inactive forms of CDK and EGFR are considered to be similar to each other.[48] Binding at the EGFR dimerization interface could thus be used to inhibit EGFR activity, since it would interfere with the ability of an EGFR monomer to be activated by another monomer. A peptide derived from the MIG6 (mitogen-inducible gene 6) protein was found to inhibit EGFR by binding at and blocking the dimerization interface, exhibiting a K_d value of 13 uM.[49] CDK2 appears to possess a similar pocket, as D-luciferin was found to bind in the same location and inhibit CDK2 (although it binds in two locations, with the other being the ATP site, so its contribution is less clear).[50] Mapping showed that this site appears to be druggable in both EGFR and CDK2, with strong consensus sites present in this pocket in structures of both kinases, although it appears to be stronger in CDK2.

PMP (PKA Myristoyl Pocket)

The myristoyl pocket in protein kinase A (PKA) is located in the C-terminal domain, in a different area from the c-Abl myristoyl pocket, and myristate binding here appears to activate membrane association with PKA.[51] This pocket has received less attention than the myristoyl pocket in c-Abl, and does not appear to be targeted by any known inhibitors, but based on the mapping results of PKA structures, the PKA myristoyl pocket appears to be highly druggable.

AAS (Aurora A Autophosphorylation Site)

This pocket is present in Aurora A kinases in the C-terminal domain, between the PDIG and DEF sites. Autophosphorylation of Aurora A occurs when the activation segment on one monomer binds to this region on a second monomer, activating the second monomer.[52] Inhibitors do not appear to have been developed for this site, but mapping indicates that it is likely to be highly druggable, with multiple strong consensus sites located there in some mapped structures of Aurora A.

MPP (MKK4 p38 Peptide site)

This site is located on the N-terminal domain of MAP kinase kinase 4 (MKK4), which phosphorylates and activates members of the MAP kinase family, such as JNKs and p38 kinases. A p38a peptide was found to inhibit MKK4 by binding at this site and inducing an auto-inhibition state.[53] Few structures of MKK4 are available, and none of them are unliganded at this site, but the available structures were found to have strong consensus sites there.

Experimental validation of CDK2 Inhibitors binding at the PIF pocket identified by FTMap

FTMap predicts a strong hot spot in the “PIF pocket” of CDK2 (Figure 4). To experimentally identify ligands that bind to that region, we used tethering, a disulfide-based fragment screening approach,[54] described in the Experimental Section (see Figure S4).

Briefly, a cysteine residue was introduced adjacent to the PIF pocket by mutation of Lys56 to Cys, and the mutant protein was used to screen a library of 1280 small-molecule disulfides for potential ligands. Hits were selected by measuring the formation of a protein-fragment mixed disulfide by intact protein mass spectrometry. The relative "affinities" of the top 12 hits were assessed by measuring the concentration of β-mercaptoethanol required to displace half of the fragment from the protein (βME_{50}). Higher βME_{50}values indicate a stronger noncovalent interaction between the fragment and the protein. Figure 5 shows the two most potent compounds bound to CDK2 as identified by tethering and the percent inhibition of disulfide labeling versus the βME_{50} concentration. We note that these results are fairly preliminary, as the concentration-response curves did not sufficiently flattened out to a maximum value, and there was only one replicate at each concentration. Nevertheless we further studied the 12 most promising compounds and used a radioactive peptide phosphorylation assay to measure whether any of the compounds could modulate CDK2 activity in the presence of Cyclin A2. Compounds 1 and 2 shown in Figure 5 were found to have the greatest inhibitory effect, reducing the phosphorylation of the Histone H1 peptide by CDK2-Cyclin by 42% and 81%, respectively (Figure 6). Additionally, at the reductant concentration used in the functional assay, there was no detectable labeling of CDK2 lacking the mutant K56C; this demonstrated that the modulation of CDK2 was due to ligand-binding in the PIF pocket region, likely through disruption of the CDK2-Cyclin interaction.

How frequently are druggable sites present in different structures of a kinase?

One strength of Kinase Atlas is that it can simultaneously display the binding hot spots predicted in all structures of a target kinase, and thus providing statistics to assess the robustness of a particular binding site. Here we use this option to explore 11 sites of the kinases we have studied in the paper, followed by the analysis of four kinases with validated allosteric sites, and one kinase that is the target of current drug discovery efforts. Results are presented in Table 3. We did not include the DFG-out site here, primarily because most structures, either unliganded or bound to a Type I inhibitor, are in the DFG-in conformation even for kinases that have a validated Type II inhibitor, and thus the statistics of site distribution is not very informative. In addition, as we already noted, it is questionable whether the DFG-out pocket is a genuine allosteric site.

Table 3 shows that the MT3 site is druggable (includes over 16 probe clusters) in 33 of the 39 MEK1/2 structures in the PDB. The number in parenthesis (36) shows the number of structures with more than 10 probe clusters at the site. The only other site that is predicted to be druggable in many MEK1/2 structures is the MPP pocket (note that the ATP site, while detected in all structures, is not shown in Table 3). The MT3 pocket is also druggable in 33 of the 110 EGFR structures in the PDB, and high affinity allosteric inhibitors targeting the site have been developed.[22] The PIF pocket is predicted to be druggable in a large fraction of the PDK1 structures, and we have just shown that this site is also important in CDK2. PIF is the most druggable site in MAPK8, although the MAPK8 structures in Table 1 were used to define the DRS and DEF pockets that are weaker sites than PIF. Based on the FTMap results, apart from the MT3 and PIF pockets druggable sires are predicted in smaller fractions of all structures, demonstrating that kinases are flexible molecules and druggability heavily depends on the conformation of the kinase. While sites that are present in many

structures are candidates for inhibitor development, a site can be important even when it is open only in specific structures. For example, the c-Abl myristoyl pocket is generally a weak site, but it is druggable in some structures, and it was recently shown to bind the inhibitor ABL001 with low nanomolar affinity in the interface of SH2 and SH3 binding domains.[30, 31] We also note that while some of the sites occur infrequently in the "source" kinases they are primarily associated with, they can be found in many other kinases (see Figure 2).

The last five proteins in Table 3 are not among the 12 "source" kinases considered in Table 1, and the results are shown to demonstrate further applications of Kinase Atlas using all structure of a target kinase. First, we already mentioned that the MT3 pocket is a promising site for development of allosteric EGFR inhibitors. The site, which is druggable in many EGFR structures (see Table 3), was reported to bind an allosteric inhibitor, EAI045, that targets selected drug-resistant EGFR mutants but spares the wild-type receptor.[22] The crystal structure shows that the allosteric site is created by the displacement of the regulatory C-helix in an inactive conformation of the kinase, thus it is clear why the pocket is druggable only in a fraction of structures. The application demonstrates the use of an allosteric site for improved selectivity, as the compound inhibits L858R/T790M-mutant EGFR with low-nanomolar potency in biochemical assays. As a single agent EAI045 is still not effective in blocking EGFR-driven proliferation in cells, but it works in synergy with cetuximab, an antibody therapeutic that blocks EGFR dimerization.[22]

The next row in Table 3 shows that the PIF pocket is druggable in many CDK2 structures, in agreement with the experimental validation presented in the previous section. However, the MT3 and CMP sites of CDK2 also show high levels of druggability. The PIF pocket is the most frequently occurring druggable site also in JNK3. However, here we focus on the PDIG pocket, which is druggable in some JNK structures, and was shown to bind irreversible inhibitors of JNK1/2/3. The co-crystal structure of one of the compounds in JNK3 (PDB ID 3V6R) shows a long molecule with one end in the PDIG pocket and the other end reaching the ATP site.[55] In addition the compound, JNK-IN-8, forms a covalent bond with a conserved cysteine residue, and it is a selective JNK inhibitor that inhibits phosphorylation of c-Jun, a direct substrate of JNK, in cells exposed to submicromolar drug.[55] The PDIG pocket is an allosteric substrate recognition site in DAPK,[56] and it consistently includes strong hot spots (Table 3). We also studied PKCθ, which was reported as a target for drug discovery by Hotspot Therapeutics. PKCθ is a well validated immunokinase in autoimmune and rare diseases, but its active-site inhibitors lack selectivity.[57] The protein has only five structures in the PDB, but the PDIG pocket is druggable in four of these structures, and there is no competing site that could offer comparable druggability (Table 3)

DISCUSSION

Existing kinase databases

The Kinase Atlas is the first database to focus on allosteric sites in kinases and their ability to bind potential ligands. The most similar existing database would likely be KLIFS, a structural kinase-ligand interaction database that covers the region of the kinase between the N- and C-terminal domains--which would include the DFG-out and MT3 sites--and provides detailed structural and ligand-binding information for all human and mouse kinase structures

available in the PDB.[58] KLIFS provides a consistent numbering scheme for kinase residues and descriptions of subpockets within the catalytic cleft, which makes it simpler to compare how known ligands bind to kinases and identify potential patterns in kinase-ligand interactions. The goal of the Kinase Atlas, on the other hand, is to identify which allosteric sites on each kinase might be suited for inhibitor development, even without ligand-binding data for that site--which is likely to be the case for many kinases of future interest.

How to use the Kinase Atlas

The Kinase Atlas is available at https://kinase-atlas.bu.edu and contains FTMap results for all kinase structures available in the PDB. Each structure has its own page, where users may access consensus site data (including whether consensus sites corresponded to any allosteric sites), or they may download or visualize mapping results for that structure; downloaded results are available as PyMOL session files containing the protein structure that was mapped and the resulting consensus sites, which indicate the regions on the surface that would be most likely to bind ligands. For users interested in a particular kinase, summarized mapping results for the structures associated with each kinase (based on their UniProt accessions) are also available, listing the strongest (if any) consensus sites associated with each allosteric site on each structure.

As an example, the serine/threonine kinase PKR (protein kinase R) has been linked to breast cancer,[59] hepatocellular carcinoma,[60] and Huntington's disease,[61] but it has just 3 structures of the full catalytic domain in the PDB (2A19, 2A1A, and 3UIU) and 271 bioactivities listed in ChEMBL (many other kinases have thousands of associated bioactivities). FTMap found that the AAS and PDIG sites were likely to be druggable in PKR, with both of them having strong consensus sites; several smaller consensus sites were also located near these pockets, fulfilling the criteria described by Kozakov et al. for being druggable by traditional druglike compounds.[12] Figure 7A shows the strong hot spots predicted by FTMap for the apo structure 3UIU_B of PKR. The strongest hot spot, 0(19), shown in magenta, is located at the AAS site. The second strongest hot spot, 1(18), shown in coral, is at the PDIG site. The latter also accommodates the secondary consensus sites 7(5) (in green) and 8(4) (in blue) indicating more minor hot spots. For orientation we superimposed the Chk1 inhibitor (compound 3, shown in yellow) from the Chk1 structure 3jvs. Figure 7B is a close-up of the PDIG site with PKR in surface representation to show that at this site PKR has a well-defined pocket that could accommodate small compounds.

Selection and naming of allosteric sites

The two most well known kinase allosteric sites would be the DFG-out and MT3 pockets, both of which have FDA-approved inhibitors, and a literature search turned up descriptions of six other sites (PIF, CMP, DRS, PDIG, DEF, and LBP). The remaining four sites (EDI, MPP, PMP, and AAS) were identified after mapping was performed on all kinase structures, and several regions were frequently found to have strong consensus sites that were not associated with any of the previously identified allosteric sites. Kinases that were found to have these hot spots then had their structures aligned with all other kinase structures to identify those that had either a protein-protein interaction site or a bound ligand/peptide in the same location as the consensus sites, and some of these structures were associated with

publications that described an allosteric site in the regions of interest (see Figure S1). Some kinase allosteric sites have established names; these tend to be based on either the "source" kinase (PIF, DEF), a motif located near the site (DFG), or a ligand/peptide that binds to the site (DRS, DEF). The remaining eight sites were thus named similarly, with all of them referencing the "source" kinase and ligand aside from PDIG (based on a motif) and LBP (based on the ligand only).

CONCLUSIONS

The existing FDA-approved kinase inhibitors are overwhelmingly ATP-competitive, and they are also disproportionate in which kinases and diseases they target. As of 2015, 18 out of the 27 protein kinase inhibitors targeted tyrosine kinases (that comprise only 90 out of the 518 human kinases),[3] and 26 out of 28 kinase inhibitors were intended to treat cancer, even though many other diseases are associated with kinases.[62] Similarly, the amount of structural and bioactivity data available for different kinases is also uneven, with over a quarter of the kinase bioactivity data from ChEMBL covering just 18 kinases, and the most popular 10 kinases accounting for over 40% of kinase structures.[6] However, a number of kinases that are well validated targets for the treatment of various diseases do not have efficient active-site modulators due to the limited selectivity of their ATP binding site, or due to the high mutation rate in the region leading to resistance. While developing allosteric modulators to avoid the problem could be a promising approach, only a few allosteric inhibitors have been approved by the FDA. Thus, for the majority of kinases without much ligand-binding data available, the Kinase Atlas could be a valuable resource for determining which regions are likely to be druggable, since FTMap can detect binding sites even in unliganded structures. We have shown examples of sites detected by Kinase Atlas that bind allosteric modulators reported in the literature. Although such modulators generally have moderate affinity, their combination with ATP competitive inhibitors can provide improved selectivity and resistance to mutations. Another avenue toward improved affinity can be the development of covalent inhibitors if the target includes a cystine residue at suitable location. Thus, we hope that Kinase Atlas will be used and results will be reported in the literature.

EXPERIMENTAL SECTION

In this section we describe both the computational tools used for developing the Kinase Atlas database and the experimental methods applied for the validation of the predicted PIF pocket in CDK2.

Computational solvent mapping by FTMap

FTMap identifies binding hot spots on the surface of a protein by finding the most favorable regions for 16 small molecular probes of various shapes and polarities (see Figure S2).[10] Given a protein structure, for each probe the algorithm places the tens of thousands of copies all over the surface on the basis of dense rotational and translational grids, retains the most favorable probe positions by energy and refines their orientations, then clusters the probe molecules by location and ranks them by their average energy. The lowest energy probe clusters of each probe type are retained, and clustering is performed once more on clusters of all probe types to form the consensus sites, which are ranked by their population of probe

clusters. Consensus sites identify the locations of binding hot spots on the protein surface, and their rank corresponds to the relative strength and importance of the associated hot spot.

Kinase structure selection

To obtain a list of kinase catalytic domain structures available for mapping, two resources were used: Pfam,[63] a database for protein families, which groups proteins by sequence and matches them to PDB structures through UniProt,[64] and the Gene Ontology (GO) project,[65] which describes gene products by their biological processes, molecular functions, and cellular components. A structure had to be classified as a "protein kinase domain" (Pkinase) by Pfam and having "protein kinase activity" as a molecular function by GO to be included for mapping. The final list contained 4910 structures of 376 distinct kinases available in the Protein Data Bank; 239 of these were human kinases.

Mapping preparation

After each kinase structure was downloaded from the Protein Data Bank,[66] it was split into its N- and C-terminal domains before mapping (see Figure S3). The active site in kinases is located between the domains and binds with high affinity to ATP, so separating the domains before mapping is required to break up the ATP site and allow potential allosteric sites to be detected. CATH, a database that classifies protein domains by secondary structure, was used to identify the domains in each structure.[67] For structures without an entry in CATH, the classification from a similar structure (identified using BLAST) was applied.

Assignment of mapping results to allosteric sites

Each allosteric site was assigned a representative structure from the kinase of origin with a ligand (small molecule/inhibitor or peptide) known to bind at the allosteric site.

After using PyMOL to align each representative structure to each mapped structure, consensus sites were assigned to allosteric sites based on whether they overlapped with a representative ligand. A consensus site that overlapped with multiple representative ligands was assigned to the allosteric site with which it had the greatest overlap. Overlap between a consensus site and representative ligand was found by using SciPy to calculate the convex hull of the ligand and determining whether any consensus site atoms were located within the convex hull.

We note that the splitting and separate mapping of N and C domains may result in a false positive hot spot for some kinases. This hot spot is on the C-lobe in a pocket normally occupied by the side chains of residues from the loop between the C-end of the αC helix and the next β-strand on the N-lobe. However, since there is no putative allosteric site in the vicinity, the probes binding to this region do not impact any of the 12 sites shown in Figure 1

Druggability assessment

The strength of a consensus site (based on its population) can be considered as a measure of the potential binding affinity at that site. A kinase was considered druggable at a particular allosteric site if at least one of its structures had a strong consensus site (at least 16 probe clusters) assigned to that allosteric site. For a site to be druggable using conventional small

molecule drugs, the positions of other nearby consensus sites would also need to be considered, but the main factor in determining whether a site would be an appropriate drug target is its potential to bind ligands with high affinity.[12] Pockets with a slightly weaker consensus site (at least 13 probe clusters) would be considered borderline druggable, but in Table 3 we have listed even the number of structures with 10 or more probe clusters at the considered site.

Experimental validation of the PIF pocket in CDK2

Full-length Human CDK2 was cloned into the pMCSG10 bacterial expression vector, which appends a 6xHis-GST purification tag to the N-terminus of CDK2, which can be later removed by TEV protease. To enable tethering at the PIF pocket of CDK2, Cys118 and Cys177 were mutated by alanine to prevent off-target labeling at these sites and Lys56 was mutated to Cys to enable targeting of fragments to the adjacent PIF pocket. CDK2 was expressed in the Rosetta *E. coli* strain by induction with IPTG at 16°C and growth overnight. Cells were collected by centrifugation, resuspended in lysis buffer [50 mM HEPES pH 7.5, 150 mM NaCl, 10 mM MgCl2, 2 mM DTT, 5% glycerol] with protease inhibitors and lysed using a microfluidizer. The clarified lysate was batch loaded onto Ni-NTA resin, washed, and eluted with imidazole. The sample was dialyzed overnight in lysis buffer with 10 mM imidazole and 1:20 w/w TEV protease to cleave the purification tag. The dialyzed mixture was passed over a Ni-NTA column and the flow-through was collected. Finally, the sample was purified by Size Exclusion Chromatography on a Superdex 200 16/60 column in lysis buffer. The protein was >95% pure by SDS-PAGE and its identity was confirmed by intact protein mass spectrometry. Although it was not validated that the mutations did affect structural integrity, the tethering technology has been successfully used for exploring suspected and orphan allosteric sites in a variety of proteins. In particular, the method was used to study ligand binding to the PIF pocket of PDK1 by introducing individual Cys mutation at six positions around the pocket.[68] It was reported that the mutants were somewhat less catalytically active that the WT PDK1. Nevertheless, the study discovered a variety of small molecule fragment disulfides that could either activate or inhibit PDK1 by conjugation to the PIF pocket.[68] Considering the success of this and a variety of other applications, the tethering method remains an important tool in the UCSF Small Molecule Discovery Center.[69]

Screening for fragment binding to CDK2 was performed by monitoring mixed disulfide formation between the protein and the fragment using intact protein mass spectrometry on a Waters LCT Premier (LC-TOF) with in-line C4 desalting column (Figure S4). Each tethering reaction was run in Tris-buffered saline (TBS) containing 2 μM CDK2, 100 μM of a small-molecule disulfide, and 1 mM βME as reductant. From the primary screen of 1280 fragments, the top 60 hits were cherry-picked and retested to confirm. The relative potency of the top 12 confirmed hits was measured by determining the concentration of β-mercaptoethanol required to displace half of the fragment from the protein ($βME_{50}$), with higher values indicative of a stronger non-covalent interaction between the fragment and the protein. The impact of the top 12 hits on CDK2 function was assessed by first labeling CDK2 with 100 μM of each fragment in the presence of 800 μM βME and then measuring

CDK2 catalytic activity in the absence or presence of its activator Cyclin A2 using a radioactive kinase assay with [γ–^{32}P]-ATP and Histone H1 peptide as the substrates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This investigation was supported by grant R35-GM118078 from the National Institute of General Medical Sciences.

ABBREVIATIONS USED

AAS	Aurora A activation segment
AGC	protein kinase AGC
BCR	breakpoint cluster region
CML	Chronic Myelogenous Leukemia
CMP	c-Abl Myristoyl Pocket
DEF	Docking site for ERK FXF
DFG	Asp-Phe-Glu
EDI	EGFR Dimerization Interface
HM	hydrophobic motif
JIP1	JNK-interacting Protein 1
JNK	c-Jun N-terminal Kinase
KLIFS	Kinase-Ligand Interaction Fingerprints and Structures database
LBP	Lipid Binding Pocket
MEK1/2	Dual Specificity Mitogen-Activated Protein Kinase Kinase ½
MKK	Mitogen-Activated Protein Kinase Kinase
MPP	MKK4 p38a Peptide
MT3	MEK1/2 Type III Inhibitor
PDIG	Pro-Asp-Ile-Gly motif
PDK1	Phosphoinositide-dependent Kinase 1
PIA	Phosphatidylinositol Ether Lipid Analogue
PIF	PDK1 Interacting Fragment

PKR	Protein Kinase R
PMP	PKA Myristoyl Pocket
SH2/3	Src Homology 2/3
TEV	Tobacco Etch Virus
TIE2	Tunica Interna Endothelial Cell Kinase

REFERENCES

(1). Roskoski R Jr. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. Pharmacol. Res. 2016, 103, 26–48. [PubMed: 26529477]

(2). Cell Signalling Technology, Kinase-disease associations. https://www.cellsignal.com/contents/resources-reference-tables/kinase-disease-associations/science-tables-kinase-disease

(3). Manning G; Whyte DB; Martinez R; Hunter T; Sudarsanam S The protein kinase complement of the human genome. Science 2002, 298, 1912–1934. [PubMed: 12471243]

(4). Brylinski M; Skolnick J Comprehensive structural and functional characterization of the human kinome by protein structure modeling and ligand virtual screening. J. Chem. Inf. Model 2010, 50, 1839–1854. [PubMed: 20853887]

(5). Blue Ridge Institute for Medical Research: FDA-approved protein kinase inhibitors. http://www.brimr.org/PKI/PKIs.htm

(6). Bajusz D; Ferenczy GG; Keseru GM Structure-based virtual screening approaches in kinase-directed drug discovery. Curr. Top. Med. Chem 2017, 17, 2235–2259. [PubMed: 28240180]

(7). Fang Z; Grutter C; Rauh D Strategies for the selective regulation of kinases with allosteric modulators: Exploiting exclusive structural features. ACS Chem. Biol 2013, 8, 58–70. [PubMed: 23249378]

(8). Vijayan RS; He P; Modi V; Duong-Ly KC; Ma H; Peterson JR; Dunbrack RL Jr.; Levy RM Conformational analysis of the DFG-out kinase motif and biochemical profiling of structurally validated type ii inhibitors. J. Med. Chem 2015, 58, 466–479. [PubMed: 25478866]

(9). Roskoski R Jr. Allosteric MEK1/2 inhibitors including cobimetanib and trametinib in the treatment of cutaneous melanomas. Pharmacol. Res 2017, 117, 20–31. [PubMed: 27956260]

(10). Kozakov D; Grove LE; Hall DR; Bohnuud T; Mottarella SE; Luo L; Xia B; Beglov D; Vajda S The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. Nat. Protoc 2015, 10, 733–755. [PubMed: 25855957]

(11). Berman HM; Westbrook J; Feng Z; Gilliland G; Bhat TN; Weissig H; Shindyalov IN; Bourne PE The Protein Data Bank. Nucleic Acids Res 2000, 28, 235–242. [PubMed: 10592235]

(12). Kozakov D; Hall DR; Napoleon RL; Yueh C; Whitty A; Vajda S New frontiers in druggability. J. Med. Chem 2015, 58, 9063–9088. [PubMed: 26230724]

(13). Foda ZH; Shan Y; Kim ET; Shaw DE; Seeliger MA A dynamically coupled allosteric network underlies binding cooperativity in src kinase. Nat. Commun 2015, 6, 5939. [PubMed: 25600932]

(14). Malmstrom RD; Kornev AP; Taylor SS; Amaro RE Allostery through the computational microscope: Camp activation of a canonical signalling domain. Nat. Commun 2015, 6, 7588. [PubMed: 26145448]

(15). Greener JG; Filippis I; Sternberg MJE Predicting protein dynamics and allostery using multi-protein atomic distance constraints. Structure 2017, 25, 546–558. [PubMed: 28190781]

(16). Caliman AD; Miao Y; McCammon JA Mapping the allosteric sites of the a2a adenosine receptor. Chem. Biol. Drug. Des 2018, 91, 5–16. [PubMed: 28639411]

(17). Hall DR; Kozakov D; Whitty A; Vajda S Lessons from hot spot analysis for fragment-based drug discovery. Trends Pharmacol. Sci 2015, 36, 724–736. [PubMed: 26538314]

(18). Rettenmaier TJ; Sadowsky JD; Thomsen ND; Chen SC; Doak AK; Arkin MR; Wells JA A small-molecule mimic of a peptide docking motif inhibits the protein kinase pdk1. Proc. Natl. Acad. Sci. U S A 2014, 111, 18590–18595. [PubMed: 25518860]

(19). Zhao Z; Wu H; Wang L; Liu Y; Knapp S; Liu Q; Gray NS Exploration of type II binding mode: A privileged approach for kinase inhibitor focused drug discovery? ACS Chem. Biol 2014, 9, 1230–1241. [PubMed: 24730530]

(20). Dietrich J; Hulme C; Hurley LH The design, synthesis, and evaluation of 8 hybrid DFG-out allosteric kinase inhibitors: A structural analysis of the binding interactions of Gleevec, Nexavar, and BIRB-796. Bioorg. Med. Chem 2010, 18, 5738–5748. [PubMed: 20621496]

(21). Rice KD; Aay N; Anand NK; Blazey CM; Bowles OJ; Bussenius J; Costanzo S; Curtis JK; Defina SC; Dubenko L; Engst S; Joshi AA; Kennedy AR; Kim AI; Koltun ES; Lougheed JC; Manalo JC; Martini JF; Nuss JM; Peto CJ; Tsang TH; Yu P; Johnston S Novel carboxamide-based allosteric MEK inhibitors: Discovery and optimization efforts toward XL518 (GDC-0973). ACS Med. Chem. Lett 2012, 3, 416–421. [PubMed: 24900486]

(22). Jia Y; Yun CH; Park E; Ercan D; Manuia M; Juarez J; Xu C; Rhee K; Chen T; Zhang H; Palakurthi S; Jang J; Lelais G; DiDonato M; Bursulaya B; Michellys PY; Epple R; Marsilje TH; McNeill M; Lu W; Harris J; Bender S; Wong KK; Janne PA; Eck MJ Overcoming EGFR(T790M) and EGFR(c797s) resistance with mutant-selective allosteric inhibitors. Nature 2016, 534, 129–132. [PubMed: 27251290]

(23). Hindie V; Stroba A; Zhang H; Lopez-Garcia LA; Idrissova L; Zeuzem S; Hirschberg D; Schaeffer F; Jorgensen TJ; Engel M; Alzari PM; Biondi RM Structure and allosteric effects of low-molecular-weight activators on the protein kinase PDK1. Nat. Chem. Biol 2009, 5, 758–764. [PubMed: 19718043]

(24). Bobkova EV; Weber MJ; Xu Z; Zhang YL; Jung J; Blume-Jensen P; Northrup A; Kunapuli P; Andersen JN; Kariv I Discovery of PDK1 kinase inhibitors with a novel mechanism of action by ultrahigh throughput screening. J. Biol. Chem 2010, 285, 18838–18846. [PubMed: 20385558]

(25). Nagar B; Hantschel O; Young MA; Scheffzek K; Veach D; Bornmann W; Clarkson B; Superti-Furga G; Kuriyan J Structural basis for the autoinhibition of c-ABL tyrosine kinase. Cell 2003, 112, 859–871. [PubMed: 12654251]

(26). Woessner DW; Lim CS; Deininger MW Development of an effective therapy for chronic Myelogenous Leukemia. Cancer J 2011, 17, 477–486. [PubMed: 22157291]

(27). Zhang J; Adrian FJ; Jahnke W; Cowan-Jacob SW; Li AG; Iacob RE; Sim T; Powers J; Dierks C; Sun F; Guo GR; Ding Q; Okram B; Choi Y; Wojciechowski A; Deng X; Liu G; Fendrich G; Strauss A; Vajpai N; Grzesiek S; Tuntland T; Liu Y; Bursulaya B; Azam M; Manley PW; Engen JR; Daley GQ; Warmuth M; Gray NS Targeting BCR-ABL by combining allosteric with ATP-binding-site inhibitors. Nature 2010, 463, 501–506. [PubMed: 20072125]

(28). Webb C; Upadhyay A; Giuntini F; Eggleston I; Furutani-Seiki M; Ishima R; Bagby S Structural features and ligand binding properties of tandem WW domains from YAP and TAZ, nuclear effectors of the Hippo pathway. Biochemistry 2011, 50, 3300–3309. [PubMed: 21417403]

(29). Adrian FJ; Ding Q; Sim T; Velentza A; Sloan C; Liu Y; Zhang G; Hur W; Ding S; Manley P; Mestan J; Fabbro D; Gray NS Allosteric inhibitors of BCR-ABL-dependent cell proliferation. Nat. Chem. Biol 2006, 2, 95–102. [PubMed: 16415863]

(30). Schoepfer J; Jahnke W; Berellini G; Buonamici S; Cotesta S; Cowan-Jacob SW; Dodd S; Drueckes P; Fabbro D; Gabriel T; Groell JM; Grotzfeld RM; Hassan AQ; Henry C; Iyer V; Jones D; Lombardo F; Loo A; Manley PW; Pelle X; Rummel G; Salem B; Warmuth M; Wylie AA; Zoller T; Marzinzik AL; Furet P Discovery of Asciminib (ABL001), an allosteric inhibitor of the tyrosine kinase activity of BCR-ABL1. J. Med. Chem 2018, 61, 8120–8135. [PubMed: 30137981]

(31). Wylie AA; Schoepfer J; Jahnke W; Cowan-Jacob SW; Loo A; Furet P; Marzinzik AL; Pelle X; Donovan J; Zhu W; Buonamici S; Hassan AQ; Lombardo F; Iyer V; Palmer M; Berellini G; Dodd S; Thohan S; Bitter H; Branford S; Ross DM; Hughes TP; Petruzzelli L; Vanasse KG; Warmuth M; Hofmann F; Keen NJ; Sellers WR The allosteric inhibitor ABL001 enables dual targeting of BCR-ABL1. Nature 2017, 543, 733–737. [PubMed: 28329763]

(32). Khateb M; Ruimi N; Khamisie H; Najajreh Y; Mian A; Metodieva A; Ruthardt M; Mahajna J Overcoming BCR-ABL T315I mutation by combination of GNF-2 and ATP competitors in an ABL-independent mechanism. BMC Cancer 2012, 12, 563. [PubMed: 23186157]

(33). Akella R; Moon TM; Goldsmith EJ Unique map kinase binding sites. Biochim. Biophys. Acta 2008, 1784, 48–55. [PubMed: 18068683]

(34). Johnson GL; Lapadat R Mitogen-activated protein kinase pathways mediated by ERK, JNK, AND P38 protein kinases. Science 2002, 298, 1911–1912. [PubMed: 12471242]

(35). Heo YS; Kim SK; Seo CI; Kim YK; Sung BJ; Lee HS; Lee JI; Park SY; Kim JH; Hwang KY; Hyun YL; Jeon YH; Ro S; Cho JM; Lee TG; Yang CH Structural basis for the selective inhibition of JNK1 by the scaffolding protein JIP1 and SP600125. EMBO J. 2004, 23, 2185–2195. [PubMed: 15141161]

(36). Stebbins JL; De SK; Machleidt T; Becattini B; Vazquez J; Kuntzen C; Chen LH; Cellitti JF; Riel-Mehan M; Emdadi A; Solinas G; Karin M; Pellecchia M Identification of a new JNK inhibitor targeting the JNK-JIP interaction site. Proc. Natl. Acad. Sci.U S A 2008, 105, 16809–16813. [PubMed: 18922779]

(37). Li J; Kaoud TS; LeVieux J; Gilbreath B; Moharana S; Dalby KN; Kerwin SM A fluorescence-based assay for p38α recruitment site binders: Identification of Rooperol as a novel p38α kinase inhibitor. Chembiochem 2013, 14, 66–71. [PubMed: 23225637]

(38). Stebbins JL; De SK; Pavlickova P; Chen V; Machleidt T; Chen LH; Kuntzen C; Kitada S; Karin M; Pellecchia M Design and characterization of a potent and selective dual ATP- and substrate-competitive subnanomolar bidentate C-JUN N-terminal kinase (JNK) inhibitor. J. Med. Chem 2011, 54, 6206–6214. [PubMed: 21815634]

(39). Tzarum N; Komornik N; Ben Chetrit D; Engelberg D; Livnah O DEF pocket in p38α facilitates substrate selectivity and mediates autophosphorylation. J. Biol. Chem 2013, 288, 19537–19547. [PubMed: 23671282]

(40). Liu X; Zhang CS; Lu C; Lin SC; Wu JW; Wang ZX A conserved motif in JNK/p38-specific mapk phosphatases as a determinant for JNK1 recognition and inactivation. Nat. Commun 2016, 7, 10879. [PubMed: 26988444]

(41). Lee T; Hoofnagle AN; Kabuyama Y; Stroud J; Min X; Goldsmith EJ; Chen L; Resing KA; Ahn NG Docking motif interactions in MAP kinases revealed by hydrogen exchange mass spectrometry. Mol. Cell 2004, 14, 43–55. [PubMed: 15068802]

(42). Comess KM; Sun C; Abad-Zapatero C; Goedken ER; Gum RJ; Borhani DW; Argiriadi M; Groebe DR; Jia Y; Clampit JE; Haasch DL; Smith HT; Wang S; Song D; Coen ML; Cloutier TE; Tang H; Cheng X; Quinn C; Liu B; Xin Z; Liu G; Fry EH; Stoll V; Ng TI; Banach D; Marcotte D; Burns DJ; Calderwood DJ; Hajduk PJ Discovery and characterization of non-ATP site inhibitors of the mitogen activated protein (MAP) kinases. ACS Chem. Biol 2011, 6, 234–244. [PubMed: 21090814]

(43). Diskin R; Engelberg D; Livnah O A novel lipid binding site formed by the map kinase insert in p38α. J. Mol. Biol 2008, 375, 70–79. [PubMed: 17999933]

(44). Tzarum N; Eisenberg-Domovich Y; Gills JJ; Dennis PA; Livnah O Lipid molecules induce p38α activation via a novel molecular switch. J. Mol. Biol 2012, 424, 339–353. [PubMed: 23079240]

(45). Vanderpool D; Johnson TO; Ping C; Bergqvist S; Alton G; Phonephaly S; Rui E; Luo C; Deng YL; Grant S; Quenzer T; Margosiak S; Register J; Brown E; Ermolieff J Characterization of the CHK1 allosteric inhibitor binding site. Biochemistry 2009, 48, 9823–9830. [PubMed: 19743875]

(46). Converso A; Hartingh T; Garbaccio RM; Tasber E; Rickert K; Fraley ME; Yan Y; Kreatsoulas C; Stirdivant S; Drakas B; Walsh ES; Hamilton K; Buser CA; Mao X; Abrams MT; Beck SC; Tao W; Lobell R; Sepp-Lorenzino L; Zugay-Murphy J; Sardana V; Munshi SK; Jezequel-Sur SM; Zuck PD; Hartman GD Development of thioquinazolinones, allosteric CHK1 kinase inhibitors. Bioorg. Med. Chem. Lett 2009, 19, 1240–1244. [PubMed: 19155174]

(47). Wan X; Zhang W; Li L; Xie Y; Li W; Huang N A new target for an old drug: Identifying mitoxantrone as a nanomolar inhibitor of PIM1 kinase via kinome-wide selectivity modeling. J. Med. Chem 2013, 56, 2619–2629. [PubMed: 23442188]

(48). Zhang X; Gureasko J; Shen K; Cole PA; Kuriyan J An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell 2006, 125, 1137–1149. [PubMed: 16777603]

(49). Zhang X; Pickin KA; Bose R; Jura N; Cole PA; Kuriyan J Inhibition of the egf receptor by binding of MIG6 to an activating kinase domain interface. Nature 2007, 450, 741–744. [PubMed: 18046415]

(50). Rothweiler U; Eriksson J; Stensen W; Leeson F; Engh RA; Svendsen JS Luciferin and derivatives as a DYRK selective scaffold for the design of protein kinase inhibitors. Eur. J. Med. Chem 2015, 94, 140–148. [PubMed: 25768698]

(51). Gaffarogullari EC; Masterson LR; Metcalfe EE; Traaseth NJ; Balatri E; Musa MM; Mullen D; Distefano MD; Veglia G A myristoyl/phosphoserine switch controls camp-dependent protein kinase association to membranes. J. Mol. Biol 2011, 411, 823–836. [PubMed: 21740913]

(52). Zorba A; Buosi V; Kutter S; Kern N; Pontiggia F; Cho YJ; Kern D Molecular mechanism of Aurora A kinase autophosphorylation and its allosteric activation by TPX2. Elife 2014, 3, e02667. [PubMed: 24867643]

(53). Matsumoto T; Kinoshita T; Kirii Y; Yokota K; Hamada K; Tada T Crystal structures of MKK4 kinase domain reveal that substrate peptide binds to an allosteric site and induces an auto-inhibition state. Biochem. Biophys. Res. Commun 2010, 400, 369–373. [PubMed: 20732303]

(54). Erlanson DA; Braisted AC; Raphael DR; Randal M; Stroud RM; Gordon EM; Wells JA Site-directed ligand discovery. Proc. Natl. Acad. Sci. U S A 2000, 97, 9367–9372. [PubMed: 10944209]

(55). Zhang T; Inesta-Vaquera F; Niepel M; Zhang J; Ficarro SB; Machleidt T; Xie T; Marto JA; Kim N; Sim T; Laughlin JD; Park H; LoGrasso PV; Patricelli M; Nomanbhoy TK; Sorger PK; Alessi DR; Gray NS Discovery of potent and selective covalent inhibitors of JNK. Chem. Biol 2012, 19, 140–154. [PubMed: 22284361]

(56). Temmerman K; de Diego I; Pogenberg V; Simon B; Jonko W; Li X; Wilmanns M A PEF/Y substrate recognition and signature motif plays a critical role in DAPK-related kinase activity. Chem. Biol 2014, 21, 264–273. [PubMed: 24440081]

(57). Chand S; Mehta N; Bahia MS; Dixit A; Silakari O Protein kinase c-theta inhibitors: A novel therapy for inflammatory disorders. Curr. Pharm. Des 2012, 18, 4725–4746. [PubMed: 22830352]

(58). Kooistra AJ; Kanev GK; van Linden OP; Leurs R; de Esch IJ; de Graaf C KLIFS: A structural kinase-ligand interaction database. Nucleic Acids Res. 2016, 44, D365–371. [PubMed: 26496949]

(59). Kim SH; Forman AP; Mathews MB; Gunnery S Human breast cancer cells contain elevated levels and activity of the protein kinase, PKR. Oncogene 2000, 19, 3086–3094. [PubMed: 10871861]

(60). Delhem N; Sabile A; Gajardo R; Podevin P; Abadie A; Blaton MA; Kremsdorf D; Beretta L; Brechot C Activation of the interferon-inducible protein kinase PKR by hepatocellular carcinoma derived-hepatitis C virus core protein. Oncogene 2001, 20, 5836–5845. [PubMed: 11593389]

(61). Peel AL; Rao RV; Cottrell BA; Hayden MR; Ellerby LM; Bredesen DE Double-stranded RNA-dependent protein kinase, PKR, binds preferentially to Huntington's disease (HD) transcripts and is activated in HD tissue. Hum. Mol. Genet 2001, 10, 1531–1538. [PubMed: 11468270]

(62). Wu P; Nielsen TE; Clausen MH FDA-approved small-molecule kinase inhibitors. Trends Pharmacol. Sci 2015, 36, 422–439. [PubMed: 25975227]

(63). Finn RD; Coggill P; Eberhardt RY; Eddy SR; Mistry J; Mitchell AL; Potter SC; Punta M; Qureshi M; Sangrador-Vegas A; Salazar GA; Tate J; Bateman A The PFAM protein families database: Towards a more sustainable future. Nucleic Acids Res 2016, 44, D279–285. [PubMed: 26673716]

(64). The UniProt C UniProt: The universal protein knowledgebase. Nucleic Acids Res. 2017, 45, D158–D169. [PubMed: 27899622]

(65). The Gene Ontology C Expansion of the gene ontology knowledgebase and resources. Nucleic Acids Res. 2017, 45, D331–D338. [PubMed: 27899567]

(66). Berman HM; Battistuz T; Bhat TN; Bluhm WF; Bourne PE; Burkhardt K; Feng Z; Gilliland GL; Iype L; Jain S; Fagan P; Marvin J; Padilla D; Ravichandran V; Schneider B; Thanki N; Weissig H; Westbrook JD; Zardecki C The Protein Data Bank. Acta Crystallogr. D Biol. Crystallogr 2002, 58, 899–907. [PubMed: 12037327]

(67). Dawson NL; Lewis TE; Das S; Lees JG; Lee D; Ashford P; Orengo CA; Sillitoe I CATH: An expanded resource to predict protein function through structure and sequence. Nucleic Acids Res. 2017, 45, D289–D295. [PubMed: 27899584]

(68). Sadowsky JD; Burlingame MA; Wolan DW; McClendon CL; Jacobson MP; Wells JA Turning a protein kinase on or off from a single allosteric site via disulfide trapping. Proc. Natl. Acad. Sci. U S A 2011, 108, 6056–6061. [PubMed: 21430264]

(69). Arkin MR; Ang KK; Chen S; Davies J; Merron C; Tang Y; Wilson CG; Renslo AR UCSF Small Molecule Discovery Center: Innovation, collaboration and chemical biology in the Bay area. Comb. Chem. High Throughput. Screen 2014, 17, 333–342. [PubMed: 24661212]

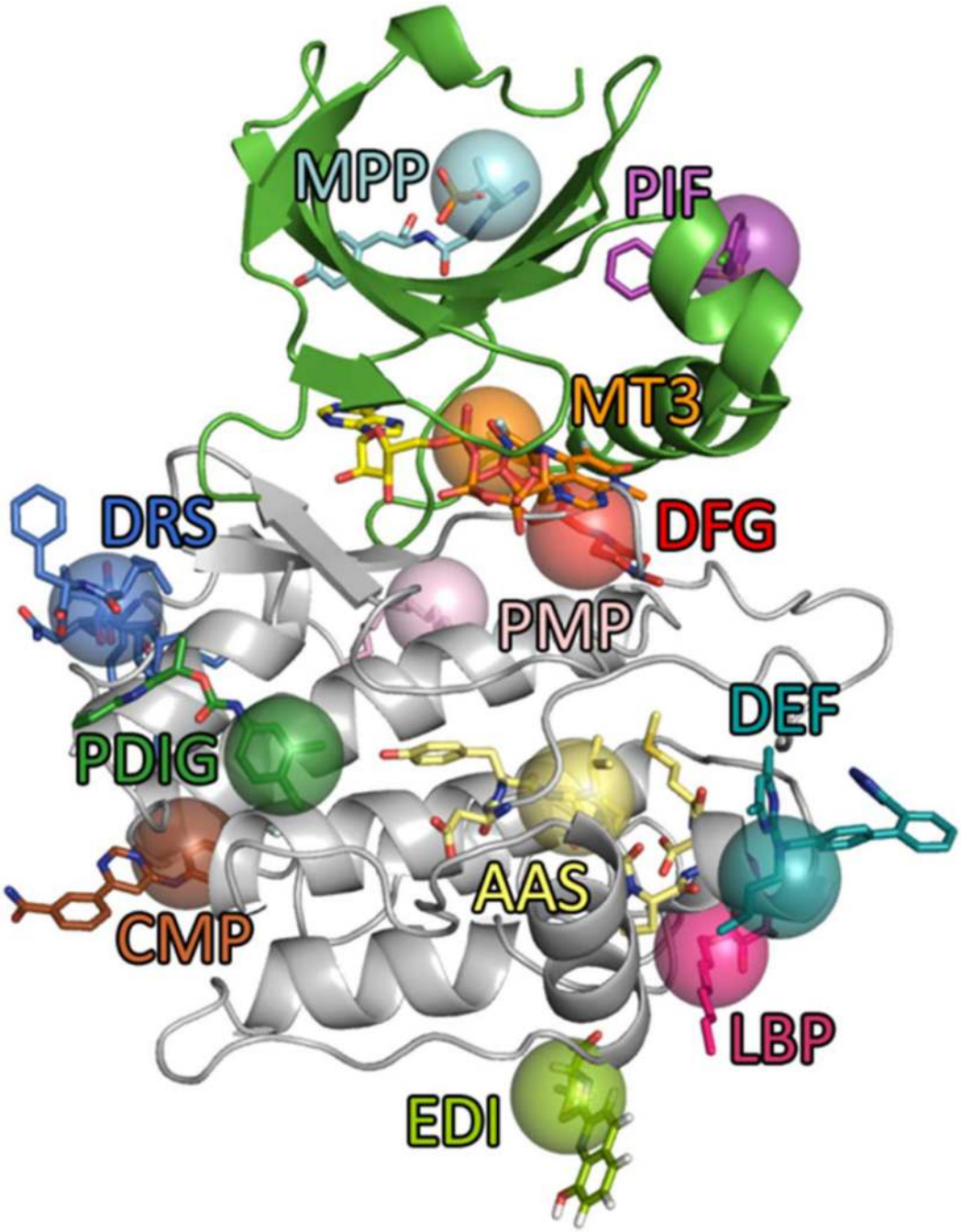


Figure 1.

Positions of all allosteric sites discussed in this paper. ATP is represented as yellow sticks, the N-terminal domain is shown in green, and the C-terminal domain is shown in gray. From top to bottom: MPP (cyan), PIF (purple), MT3 (orange), DFG (red), DRS (blue), PMP (pink), PDIG (dark green), AAS (light yellow), CMP (brown), DEF (teal), LBP (magenta), EDI (olive).

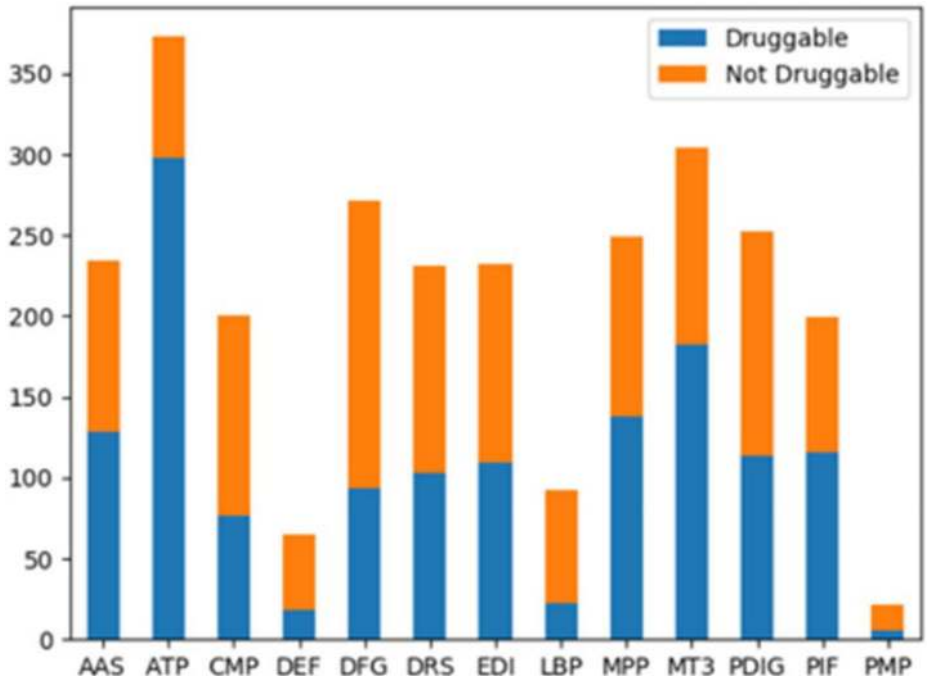


Figure 2.

Number of kinases found to have consensus sites at each putative allosteric site. ATP is not an allosteric site but is included here as a reference.

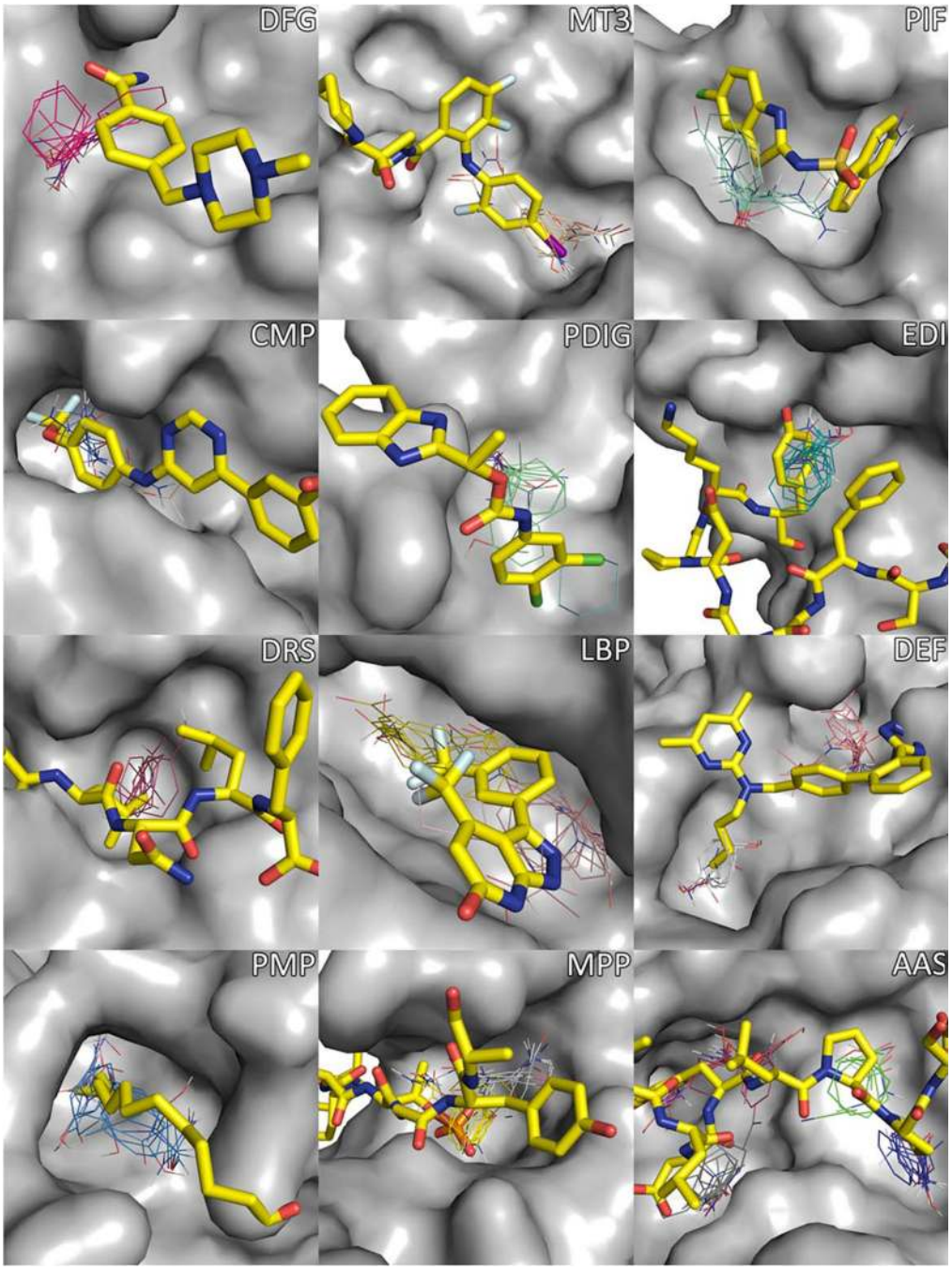


Figure 3.

Mapping results for unliganded structures (except MPP) of kinases associated with each allosteric site, with superimposed ligands from bound structures shown as yellow sticks. Each probe cluster in the consensus cluster is represented by a single probe shown in line representation. Structural, ligand, and mapping details for each site are given in Table 1.

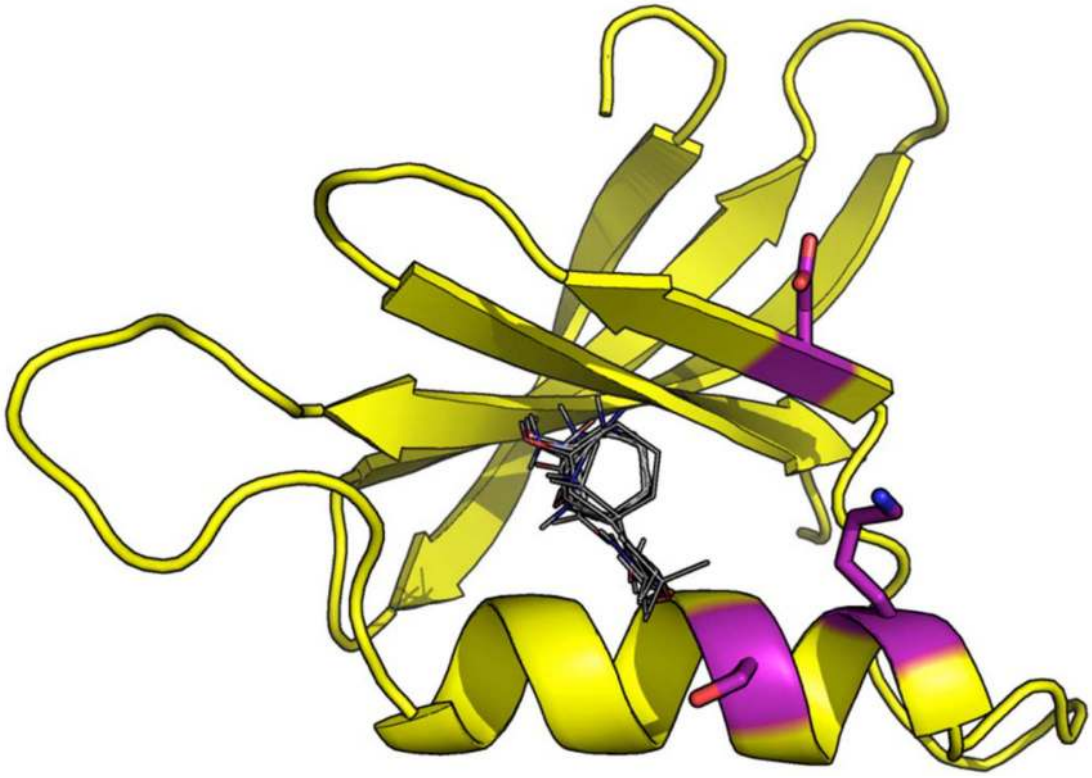


Figure 4.

FTMap finds a strong hot spot in the PIF pocket of CDK2. Only the N-lobe of CDK2 is shown, and each probe cluster placed by FTMap is represented by a single probe at its center (grey lines). Residues interacting with the probes are shown in magenta.

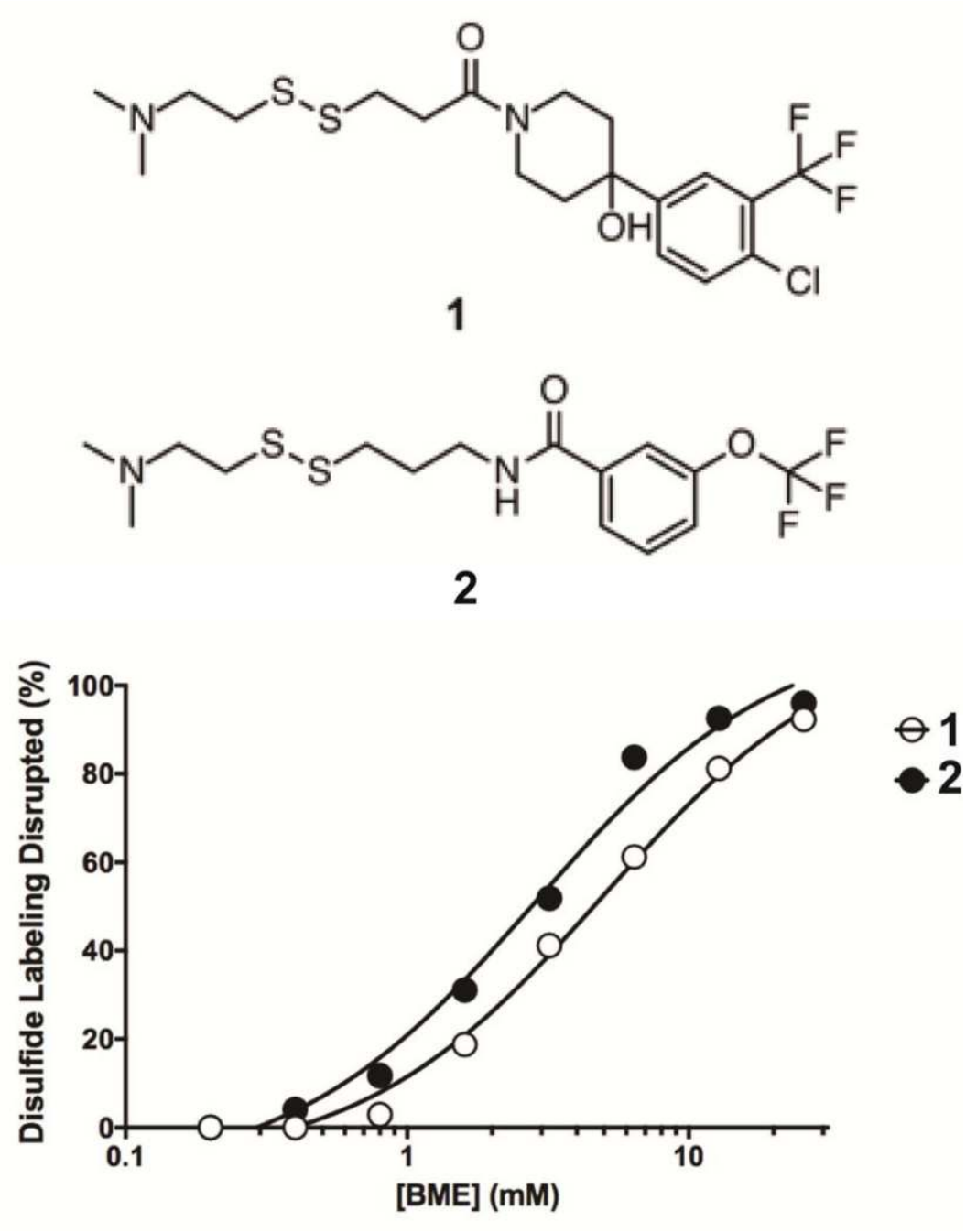


Figure 5.

Compounds bound to CDK2 as identified by tethering. **A.** The two most potent compounds. **B.** Preliminary potency measurements for compounds **1** and **2**.

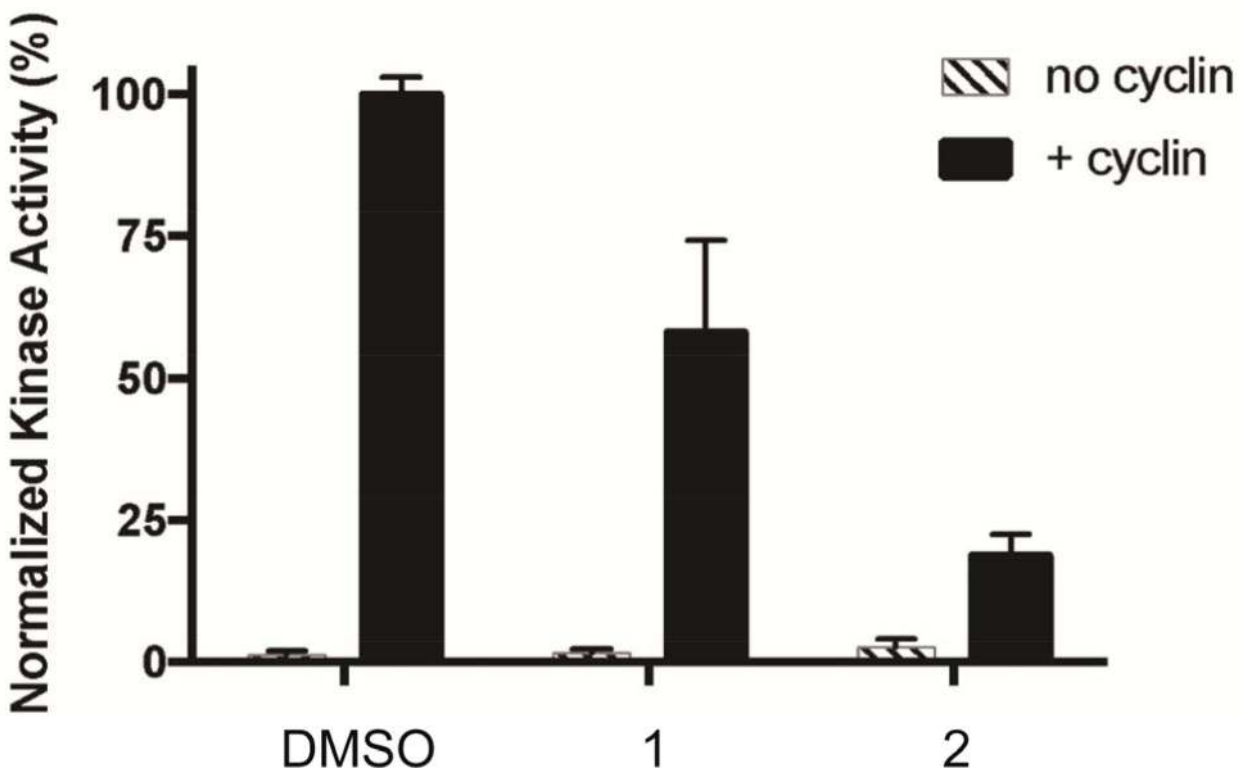


Figure 6.
Radioactive peptide phosphorylation assay results demonstrate the inhibitory effects of compounds 1 and 2 upon CDK2.

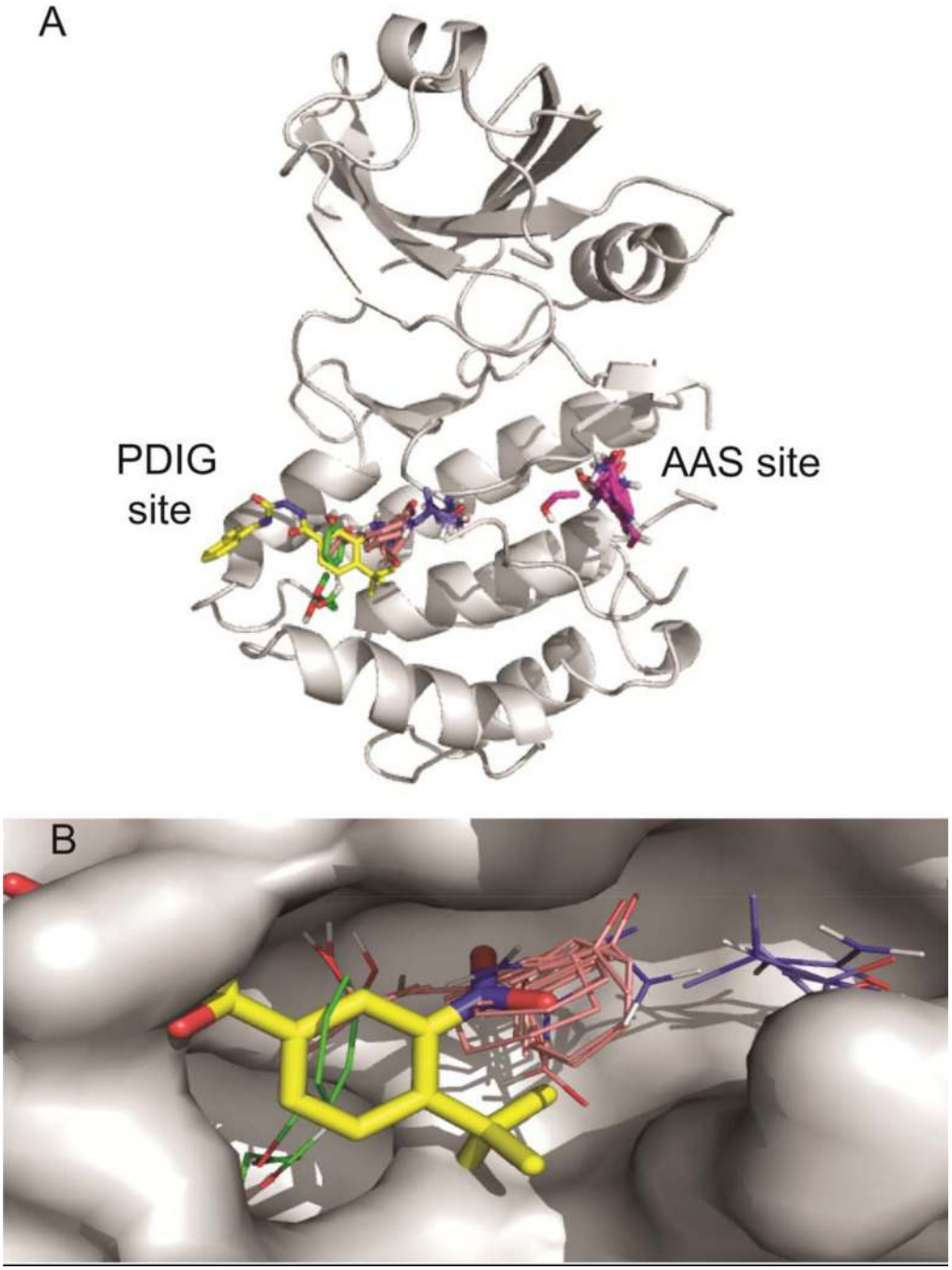


Figure 7.

FTMap results for apo apo structure 3UIU_B of PKR. **A.** The strongest hot spot, 0(19), shown in magenta, is located at the AAS site. The second strongest hot spot, 1(18), shown in coral, is at the PDIG site. The PDIG site also accommodates the secondary consensus sites 7(5) (in green) and 8(4) (in blue). For orientation we superimposed the Chk1 inhibitor (shown in yellow) from the Chk1 structure 3jvs. **B.** Close-up of the PDIG site with PKR in

surface representation to show that at this site PKR has a well-defined pocket that could accommodate small compounds.

J Med Chem. Author manuscript; available in PMC 2020 February 14.

Table 1.

Definitions and examples of allosteric sites.

Site	Site Name Origin	Inhibitor Type	Source Kinase	Example PDB	Ligand	Binding Data	Mapped Kinase	Mapped PDB	Consensus Sites	Pocket Description
DFG	DFG motif	II	many	1IEP	imatinib	IC50 = 10.8 nM	Tie2	1FVR_B	1(16)	Hydrophobic pocket that opens up when DFG motif switches to inactive "DFG-out" conformation; binding here may stabilize inactive kinase conformation
MT3	MEK1/2 type III inhibitor	III	MEK1/2	4AN2	cobimetinib	IC50 = 0.9 nM	MEK1	3EQF_A	0(20)	Adjacent to ATP and DFG-out pockets; binding disrupts salt bridge required for kinase activity
PIF	PDK1 interacting fragment	IV	PDK1	4RQK	RS1	Kd = 1.5 uM	PDK1	3IOP_A	0(21), 8(3)	PDK1 regulates other AGC kinases by recruiting them through this site
CMP	c-Abl myristoyl pocket	IV	c-Abl	3K5V	GNF-2	IC50 = 267 nM	c-Abl	3QRJ_B	3(12), 8(2)	Binding here leads to active or inactive state in c-Abl (depending on ligand size) by affecting SH domain binding
DRS	D-recruitment site	IV	MAPK8	1UKI	pepJIP1	Kd = 0.42 uM	JNK3	4Z9L_A	2(13)	Substrate docking site present in all MAP kinases
DEF	docking site for ERK FXF	IV	MAPK8	3O2M	A-82118	IC50 = 7.7 uM	JNK1	3V3V_A	0(22), 1(14)	Substrate docking site present in some MAP kinases; located near MAPK insert
LBP	lipid binding pocket	IV	p38a MAPK	3NEW	compound 10	IC50 = 1.2 uM	p38a MAPK	3S4Q_A	0(21), 3(15)	Binding of different lipids here affects p38a MAPK's preference and activity for different substrates
PDIG	PDIG motif	IV	Chk1	3JVS	compound 3	Ki = 146 nM	Chk1	4RVK_A	0(17), 9(1)	Substrate recognition site located near PDIG motif in Chk1
EDI	EGFR dimerization interface	IV	EGFR	2RFE	Mig6	Kd = 13 uM	EGFR	4RJ5_A	1(16)	An EGFR monomer activates another by binding at this interface on the C-terminal domain
PMP	PKA myristoyl pocket	IV	PKA	1CMK	myristoyl	n/a	PKA	4AE9_A	1(24), 4(08)	Myristoyl binding here activates membrane binding in PKA
AAS	Aurora A activation segment	IV	Aurora A	4C3P	Aurora A	Kd > 300 uM	Aurora A	3O51_A	0(22), 1(17)	An Aurora A monomer activates another through binding of its activation segment to this site
MPP	MKK4 p38a peptide	IV	MKK4	3ALO	p38a peptide	n/a	MKK4	3ALN_A*	1(20), 3(11)	p38a peptide binding inhibits MKK4 by inducing conformational changes that lead to auto-inhibition

*The PDB structure 3ALN_A does have a ligand at the MPP site, but MKK4 does not have any unliganded structures that contain a complete N-terminal domain.

Table 2.

Number of kinases found to be druggable at each allosteric site (out of total number of kinases containing each site). For human kinases, these numbers are further divided by family. ATP is not an allosteric site but is included here as a reference.

	All vs. Human Kinases		Human Kinases by Family							
Site	**All (376)**	**Human (239)**	**AGC**	**CAMK**	**CK1**	**CMGC**	**Other**	**STE**	**TKL**	**Tyr**
ATP	298/373	193/238	21/22	23/27	8/10	27/31	26/27	22/27	14/15	44/49
DFG	93/272	71/190	5/13	8/18	2/7	10/28	6/20	7/21	10/13	22/46
MT3	182/304	125/200	7/16	17/26	2/6	20/27	14/22	19/26	5/14	38/46
PIF	116/199	87/141	10/11	10/15	3/6	16/20	12/15	10/17	3/7	21/35
CMP	76/200	61/145	0/12	13/17	2/7	8/17	7/13	11/20	1/9	15/36
DRS	103/231	70/151	15/20	7/13	5/10	2/16	5/15	12/19	0/6	24/33
DEF	18/65	16/48	0/2	0/1	2/2	7/12	1/2	0/6	1/2	5/16
LBP	22/92	16/69	1/8	0/5	3/6	10/17	2/8	0/4	0/3	0/13
PDIG	114/252	80/166	10/20	15/21	2/5	13/24	11/17	9/21	0/2	17/36
EDI	109/232	69/148	6/12	12/19	2/7	26/29	6/13	0/4	6/11	9/32
PMP	6/21	2/13	2/5	0/4	0/0	0/1	0/1	0/0	0/0	0/2
AAS	128/234	86/158	8/17	16/22	2/7	5/16	12/16	13/17	7/11	20/33
MPP	128/249	102/171	6/9	7/17	2/4	8/21	9/17	21/25	6/10	40/47

J Med Chem. Author manuscript; available in PMC 2020 February 14.

Table 3.

Statistics of druggable (and weaker but still observable) allosteric sites in various structures of specific kinases

Site	Site Name	Inhibitor Type	Kinase	Example PDB	Number of structures	Site 1		Site 2		Site 3		Site 4	
						Type	Number	Type	Number	Type	Number	Type	Number
MT3	MEK1/2 type III inhibitor	III	MEK1/2	4AN2	39	**MT3**	**33 (36)**	MPP	22 (25)	CMP	5 (19)	DRS	2 (16)
PIF	PDK1 interacting fragment	IV	PDK1	4RQK	58	**PIF**	**43 (53)**	DRS	23 (47)	MT3	10 (14)	AAS	5 (5)
CMP	c-Abl myristoyl pocket	IV	c-Abl	2FO0	26	MPP	22 (23)	AAS	6 (12)	MT3	3 (6)	**CMP**	**1 (6)**
DRS	D-recruitment site	IV	MAPK8	1UKI	30	PIF	28 (30)	EDI	23 (28)	DEF	10 (22)	**DRS**	**0 (6)**
DEF	docking site for ERK FXF	IV	MAPK8	3O2M	30	PIF	28 (30)	EDI	23 (28)	**DEF**	**10 (22)**	DRS	0 (6)
LBP	lipid binding pocket	IV	p38a MAPK	3NEW	203	EDI	181 (202)	MT3	62 (109)	**LBP**	**35 (70)**	PIF	14 (72)
PDIG	PDIG motif	IV	Chk1	3JVS	113	MPP	80 (111)	EDI	79 (108)	AAS	28 (47)	**PDIG**	**19 (30)**
EDI	EGFR dimerization interface	IV	EGFR	2RFE	110	MT3	33 (35)	MPP	30 (46)	CMP	3 (20)	**EDI**	**2 (54)**
PMP	PKA myristoyl pocket	IV	PKA	1CMK	33	DRS	16 (25)	**PMP**	**14 (28)**	PIF	8 (31)	EDI	1 (1)
AAS	Aurora A activation segment	IV	Aurora A	4C3P	111	MPP	101 (108)	MT3	52 (67)	PIF	49 (72)	**AAS**	**12 (27)**
MPP	MKK4 p38a peptide	IV	MKK4	3ALO	3	AAS	3 (3)	**MPP**	**2 (2)**	CMP	2 (2)	DRS	1(2)
MT3	MT3 in EGFR	IV	EGFR	3W2R	110	**MT3**	**33 (35)**	MPP	30 (46)	CMP	3 (20)	EDI	2 (54)
PIF	PIF in CDK2	IV	CDK2	3QTU	371	MT3	185 (259)	CMP	130 (201)	**PIF**	**96 (109)**	EDI	80 (122)
PDIG	PDIG in JNK3	III/IV	JNK3	3V6R	48	PIF	41 (45)	EDI	40 (47)	**PDIG**	**19 (27)**	DEF	14 (42)
PDIG	PDIG in DAPK	IV	DAPK	5AV4	49	**PDIG**	**25 (43)**	PIF	10 (18)	CMP	7 (9)	MT3	5 (12)
PDIG	PDIG in PKCθ	?	PDIG	4RA5	5	**PDIG**	**(5)**	MPP	2 (4)	PIF	2 (3)	DRS	0 (2)