Bioorganic & Medicinal Chemistry Letters 25 (2015) 2372–2376

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Current kinase inhibitors cover a tiny fraction of fragment space

Hongtao Zhao, Amedeo Caflisch

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

ARTICLE INFO

Article history:
Received 15 March 2015
Revised 2 April 2015
Accepted 4 April 2015
Available online 10 April 2015

Keywords:
Kinase
Chemical space
Fragment-based drug discovery
Fragmentation
Tricyclic ring systems

ABSTRACT

We analyze the chemical space coverage of kinase inhibitors in the public domain from a fragment point of view. A set of 26,668 kinase inhibitors from the ChEMBL database of bioactive molecules were decomposed automatically by fragmentation at rotatable bonds. Remarkably, about half of the resulting 10,302 fragments originate from inaccessible libraries, as they are not present in commercially available compounds. By mapping to the established kinase pharmacophore models, privileged fragments in sub-pockets are identified, for example, the 5681 ring-containing fragments capable of forming bi-dentate hydrogen bonds with the hinge region in the ATP binding site. Surprisingly, hinge-binding fragments in current kinase inhibitors cover only 1% of the potential hinge-binders obtained by decomposing a library of nearly 7.5 million commercially available compounds, which indicates that a large fraction of chemical space is unexplored.

Kinases have been among the most intensively pursued classes of drug targets, mainly for the treatment of cancer.[1] In addition, deregulation of kinase activity has been implicated in a variety of disorders including immunological, neurological, metabolic, and infectious diseases. More than 20 years of intensive research has already generated more than two dozens of kinase inhibitors approved as drugs. The majority of kinase inhibitors target the ATP binding site with the so-called activation loop in the active conformation (type I) while a minority bind to an inactive conformation (type II).[2] In addition, type $I_{1/2}$ inhibitors bind to the DFG-in conformation but extend to the so-called back pocket establishing some characteristic type II interactions.[3] Type III and IV inhibitors bind to regions distal to the ATP site, and act through more or less pronounced allosteric effects.[4] With nearly 30,000 kinase inhibitors accumulated in the public domain, however, little is known about the coverage of chemical space. It is crucial to evaluate such coverage and find unexplored area for highly pursued targets in drug discovery projects.

Fragment-based drug discovery has progressed over years as a popular method in both pharmaceutical industry and academia.[5] Fragment-based high-throughput virtual screening approaches have yielded novel kinase inhibitors[6,7] as well as ligands for emerging therapeutic targets, for example, bromodomains (Table S1).[8,9] Verification of binding modes by X-ray and lead elaboration further emphasize the unique merits of the in silico fragment-based approaches.[6,8,10,11] Here we provide an analysis of the chemical space coverage of kinase inhibitors from a fragment perspective. Furthermore, we analyze the influence of phenol on cellular activity of three distinct inhibitors.

From the latest version of ChEMBL (Kinase SARfari, version 6.0), compounds with K_i, K_d, or IC_{50} value below 10 μM and molecular weight below 600 Da were extracted. This gives 26,668 kinase inhibitors with 70,273 activity data. About three quarters of the kinase inhibitors have inhibitory activity below 1 μM and half of them below 0.2 μM. In the ChEMBL database, 88 kinases out of the 367 kinases used for screening have more than 100 inhibitors with VEGFR2 (3,347 compounds), p38α (2,817), EGFR (2,534), CDK2 (2,023), SRC (1,848), GSK3β (1,273), LCK (1,271), CDK1 (1,205), ChK1 (1,130), PDGFRβ (1,114), HER2 (1,051), and CDK4 (1,014) being the most pursued targets. The properties of the inhibitors are analyzed (Fig. 1, black histograms) with particular interest on ring systems,[12] which constitute the essential core of a drug.[13] Most inhibitors have three to eight rotatable bonds, and two to four ring systems. Interestingly, 75% of the kinase inhibitors contain fused rings, which can offer a larger number of favorable interactions (van der Waals contacts and hydrogen bonds) with the binding site of the target than multiple non-fused rings of similar molecular weight.

The 26,668 kinase inhibitors are decomposed by an automatic fragmentation algorithm that was developed for fragment-based design by high-throughput docking.[6] Firstly, all rotatable bonds of a molecule are broken to obtain initial ring systems. Secondly, each ring system is extended at the cut(s) until it reaches another

E-mail addresses: h.zhao@bioc.uzh.ch (H. Zhao), caflisch@bioc.uzh.ch (A. Caflisch)

http://dx.doi.org/10.1016/j.bmcl.2015.04.005
0960-894X/

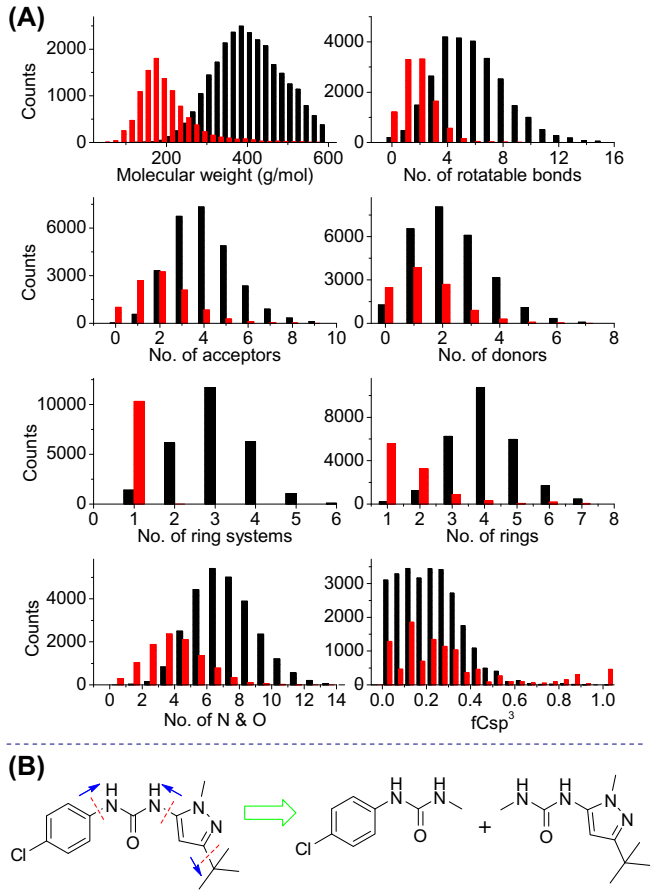

Figure 1. (A) Distributions of molecular properties of 26,668 kinase inhibitors (black) and their 10,302 unique ring-containing fragments (red) obtained by automatic decomposition. Distributions of molecular properties of the ZINC Drugs-Now 7.5 million compounds and their ring-containing fragments are shown in Figure S1. (B) Illustration of the fragmentation algorithm. (Left) The directions of elongation at the cuts (red dashed segments) are emphasized (blue arrows). (Right) Fragments obtained by automatic decomposition using the program described in Ref. 6. The cuts at rotatable bonds generate fragments with at least one ring and a single Murcko ring system.

ring or the length of the extension, that is, number of non-hydrogen atoms, reaches two. In case the atom on the second level is not an sp^3 carbon, the elongation goes further until it reaches an sp^3 carbon or a ring (Fig. 1B). In this way, 10,302 unique ring-containing fragments (i.e., Murcko ring systems[12] with rich chemical decorations) are extracted from the 26,668 kinase inhibitors. The molecular properties of the fragments are shown in Figure 1A (red histograms). Most of them consist of a single ring or fused two rings, and have one or two rotatable bonds in the functional group(s) attached to the ring(s). Remarkably, 53% of these fragments are not present in the 477,617 unique fragments obtained by automatic decomposition (using the same fragmentation algorithm as above) of the nearly 7.5 million compounds in the ZINC Drugs-Now library (Table 1).[14] To check for robustness, this analysis was repeated with Murcko ring systems,[12] that is, upon removal of ring decorations. About 40% of 2394 Murcko ring systems from kinase inhibitors are not present in the 81,715 Murcko ring systems from the ZINC Drugs-Now library. Both analyses of decorated fragments and Murcko ring systems show that a large portion of kinase fragments originates from compounds in proprietary libraries.

The most frequent fragments are six-member rings and pyrazole (Fig. 2A). They belong to four categories: (1) (substituted) benzene ring; (2) (substituted) pyridine or pyrimidine; (3) piperazine or morpholine; (4) benzamide or phenylurea. Fragments of top occurrence can be individually mapped to the binding site according to the pharmacophores[2] and the large amount of available crystal structures of kinase/inhibitor complexes (Fig. 2B). For example, pyridine and pyrimidine are known hinge-binding motifs, benzene is a preferred hydrophobic motif in the back pocket, piperazine and morpholine are solvent exposed motifs attached either to the hinge-binding motif (as in Gefitinib; PDB code 2ITY) or to the hydrophobic motif in the allosteric site (Gleevec; 1OPJ) to address solubility issue, and both benzamide and phenylurea are linkers to connect the hinge-binding motif and the hydrophobic motif in the allosteric site. Specifically, phenol is a privileged motif known to generate type $I_{1/2}$ inhibitors (see below)[3,6,10,11,15], whose binding modes are exemplified in Figure S2.

It is interesting to focus on the hinge-binding fragment space since such fragments are anchoring blocks suitable for lead elaboration. As the majority of kinase inhibitors form two to three hydrogen bonds with the hinge region, all fragments were filtered by the following two criteria: (1) at least one hydrogen bond acceptor, that is, sp^2 nitrogen or oxygen, and one donor including acidic CH;[16] and (2) geometry of acceptor and donor groups allows for bi-dentate hydrogen bonds with the hinge region.[17] Halogen bonds[18,19] are not considered because of their strict dependence on directionality which makes their analysis very difficult. More than half (precisely 5681) of the ring-containing fragments in kinase inhibitors satisfy both criteria. By applying the same criteria to the 477,617 fragments obtained from the decomposition of the ZINC Drugs-Now library,[14] a total of 196,904 potential hinge-binders are retrieved. Surprisingly, only 1% (precisely 1954) of the putative hinge-binding fragments in commercially available drug-like compounds are covered by those in kinase inhibitors (Table 1), which indicates that a large part of publically accessible chemical space remains unexplored.

Table 1
Coverage of chemical space by fragments of kinase inhibitors

	All fragments	Putative hinge-binding fragments		
		Monocyclic	Bicyclic	Multicyclic
In ZINC compounds[a]	477,617 (81,715)	102,097 (6902)	69,856 (17,694)	24,951 (12,669)
In kinase inhibitors[b]	10,302 (2394)	2031 (330)	2353 (678)	1297 (631)
ZINC coverage[c]	1.0% (1.8%)	1.0% (4.2%)	1.2% (2.7%)	0.4% (1.0%)
Proprietary[d]	53% (40%)	50% (12%)	64% (30%)	92% (80%)

[a] Number of unique ring-containing fragments obtained by automatic decomposition of the nearly 7.5 million compounds in the ZINC Drugs-Now library. Values in parentheses correspond to Murcko ring systems throughout this table.
[b] Same as above for the 26,668 current kinase inhibitors.
[c] Fraction of putative hinge-binding fragments in the ZINC Drugs-Now library that are present in kinase inhibitors. As an example, 36% of the 2353 bicyclic putative hinge-binding fragments in known kinase inhibitors are present in the 69,856 bicyclic putative hinge-binding fragments in ZINC, which gives a coverage of 1.2% (0.36*2353/69,856).
[d] Fraction of kinase inhibitor fragments that are not found in the ZINC Drugs-Now library.

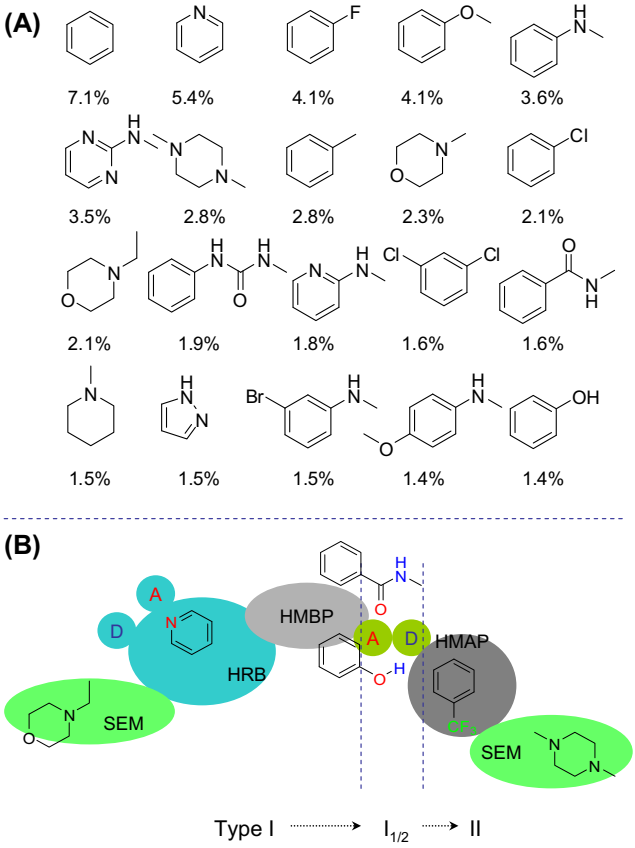

Figure 2. (A) Ring-containing fragments of top occurrence in the 26,668 kinase inhibitors. Benzene is present in 7.1% of the kinase inhibitors. (B) Illustration of type I, $I_{1/2}$, and II kinase inhibitors from a fragment point of view. A, H-bond acceptor; D, H-bond donor; SEM, solvent exposed motif; HRB, hinge-region binder; HMBP, hydrophobic motif in back pocket; HMAP, hydrophobic motif in allosteric pocket.

The 5681 potential hinge-binding fragments in kinase inhibitors were further classified into monocyclic (2031), bicyclic (2353), and multicyclic (1297). Similarly, the putative hinge-binding fragments of the ZINC Drugs-Now library can be clustered into 102,097 monocyclic, 69,856 bicyclic, and 24,951 multicyclic fragments. Interestingly, 92% multicyclic, 64% bicyclic, and 50% monocyclic hinge-binding fragments of kinase inhibitors are not present in the ZINC Drugs-Now library (Table 1). This analysis indicates that multicyclic fragments predominately originate from proprietary libraries. In an inverse sense, only 0.4% multicyclic putative hinge-binding fragments from commercial libraries are covered by kinase inhibitors, in comparison with 1.0% and 1.2% for monocyclic and bicyclic, respectively (Table 1). The same trend is observed by looking at Murcko ring systems. To rule out a potential size-dependence (i.e., a decreasing likelihood of finding a matching fragment by increasing the size), the ZINC coverage was evaluated using only the number of heavy atoms as a criterion for a match. This control shows the opposite trend, that is, a higher coverage for multicyclic ring systems followed by bicyclic and monocyclic (Fig. S3). These results provide strong evidence that multicyclic space has not been extensively explored.

Fragments of top occurrence are shown in Figure 3, Table S2 and Figure S5. Docking poses of the top 9 multicyclic fragments in the EphB4 receptor tyrosine kinase are shown in Figure S4, and all the 9 fragments are able to establish one to four hydrogen bonds (excluding acidic H-bonds) with the hinge region. The occurrence of top hinge-binding fragments decreases gradually from

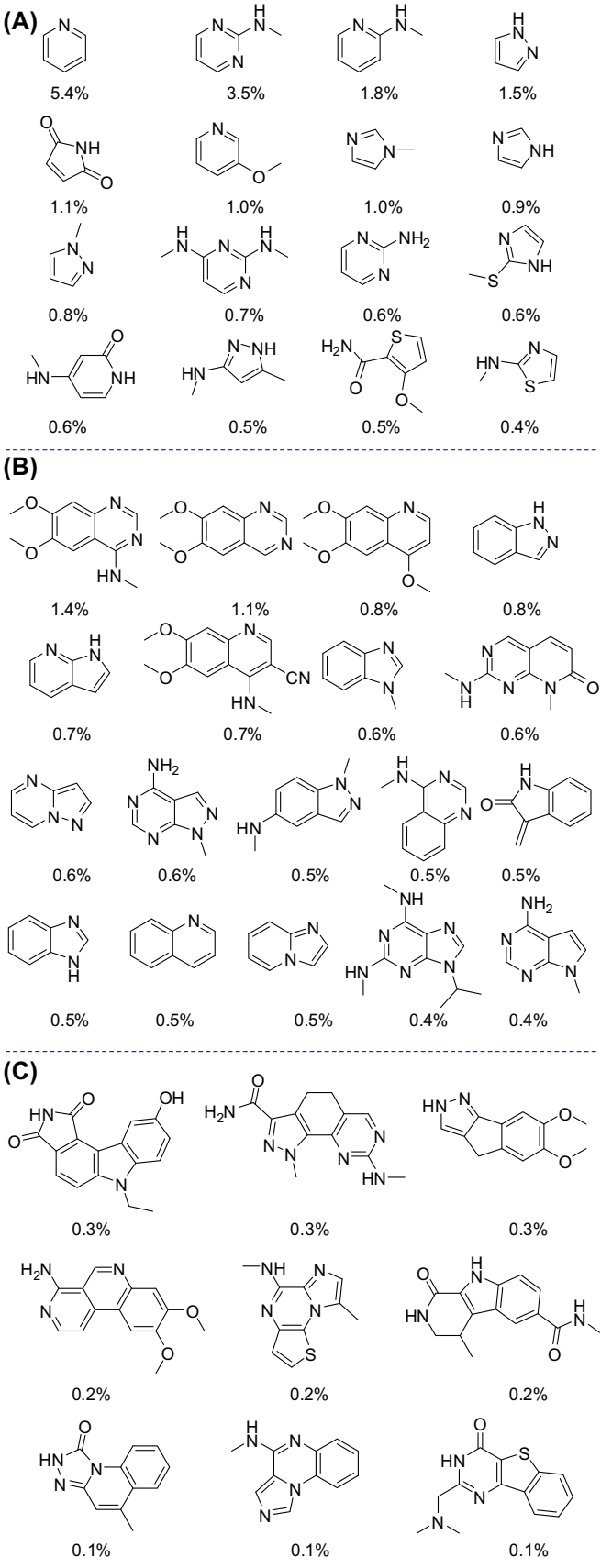

Figure 3. (A) Monocyclic, (B) bicyclic, and (C) tri/tetracyclic potential hinge-binding fragments of top occurrence in the 26,668 kinase inhibitors.

monocyclic to tricyclic, which is a consequence of two factors. Firstly, relatively large ring systems and their rigidity limit the design space of medicinal chemistry.[20] Secondly, multicyclic rings, particular with planar structure, are thought to be cytotoxic as they may interact with DNA by intercalation or have detrimental ADME effects.[21] On the other hand, DNA damage is tolerated for anti-cancer drugs. Notably, 55% of the more than 20,000 compounds in the Traditional Chinese Medicine Database[22] are multicyclic with 12,780 unique multicyclic fragments.

It is interesting to compare hinge-binding fragments of top occurrence in our study with those derived from the Pfizer database of crystal structures which includes nearly 4000 kinase-inhibitor complexes.[23] Five of the top 10 hinge-binding fragments of the Pfizer database[23] appear among the top 20 of our 5681 potential hinge binding fragments. These five fragments are pyridine, pyrimidin-2-amine, 1*H*-pyrrolo[2,3-b]pyridine, 1*H*-indazole, and 1*H*-pyrazolo[3,4-d]pyrimidin-4-amine. As in the Pfizer database, putative hinge-binding fragments in our study are mainly flat with a very small amount of sp^3 carbon atoms (Fig. S3), mainly because of the synthetic accessibility of heteroaromatic rings and in part as a consequence of the cleft-like shape of the ATP binding site. Importantly, our analysis of a comprehensive collection of kinase inhibitors reveals hinge-binding fragments not present in the previous study,[23] such as 1*H*-pyrrole-2,5-dione,[24] 3-methyleneindolin-2-one (as in Sunitinib), 6,7-dimethoxy-4-(methylamino)quinoline-3-carbonitrile (Bosutinib), quinoline and quinazoline.[25] Thus, putative hinge-binding fragments obtained by automatic fragmentation of all current kinase inhibitors could serve as a pool for rational design of kinase inhibitors via a combinatorial fragment assembly approach.[25]

Previously, we identified by high-throughput docking three tricyclic scaffolds as low-micromolar to high-nanomolar inhibitors of the Eph tyrosine kinase subfamily.[6,7,26] All three hits were successfully optimized into low-nanomolar leads (Fig. 4).[6,10,11,15] Particularly, pyrrolo[3,2-*b*]quinoxaline has evolved into a series of potent type $I_{1/2}$ and II tyrosine kinase inhibitors.[15] Moreover, compound 11d of Ref. 15 was shown to block tumor growth in a human breast cancer xenograft in nude mice. The recent discovery of three novel tricyclic scaffolds (after two decades of intensive pursuit of kinase inhibitors by a large number of research groups in academia and industries) suggests that tricyclic space remains largely unexplored, consistent with the above analysis that multicyclic fragment space in the public domain is least explored (0.4%). In addition, the commercial libraries contain only about 180,000 multicyclic compounds. Taken together, exploration of multicyclic space may open a new avenue for design of novel kinase inhibitors, with large available intellectual property space.

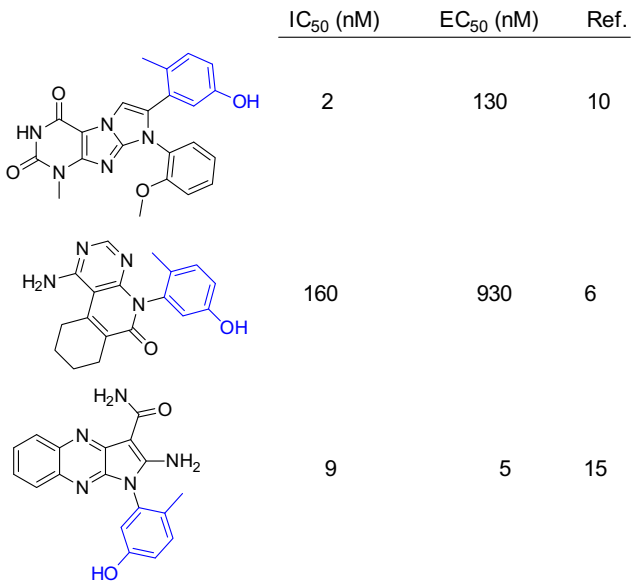

Figure 4. Type $I_{1/2}$ inhibitors with different hinge-binding scaffolds (black) and same privileged fragment in the back pocket (blue). Note the different ratios of enzymatic activity (IC_{50}, measured by radiolabeled ATP) vs. cellular activity (EC_{50}, measured in a murine fibroblast cell line overexpressing human EphB4).

In the efforts to pursue type $I_{1/2}$ inhibitors of Eph tyrosine kinases, we have added a hydroxyl group to the phenyl ring located in the back pocket (Fig. 4, see also PDB codes 4GK2 and 4G2F).[6,10,15] For three tricyclic scaffolds, the addition of a single hydroxyl group is able to improve the potency in vitro by a factor of 50–100 because the hydroxyl group forms two buried hydrogen bonds. However, the ratio of cellular activity (EC_{50}) to enzymatic activity (IC_{50}) ranges from about 1 to 65, reflecting differences in cell permeability and/or (active) efflux (Fig. 4). This indicates that (1) privileged fragments are able to boost potency in a biochemical assay equivalently on diversified scaffolds; (2) the influence of privileged fragments in a cellular assay may vary from scaffold to scaffold; and as a consequence (3) unfavorable cell-permeability and/or (active) efflux are not necessarily inherent to a fragment but rather the whole inhibitor.

In summary, the kinase inhibitors in the public domain have been decomposed by automatic fragmentation for the analysis of the coverage of chemical space. Focusing on the fragments in contact with the hinge region, bicyclic and multicyclic ring systems show more combinatorial diversity than monocyclic. Two main conclusions emerge from our analysis: (1) about half of the ring-containing fragments in known kinase inhibitors originate from proprietary libraries, and (2) the vast majority (99%) of potential hinge-binding scaffolds in the ZINC Drugs-Now library have not been explored. Based on the large unexplored chemical space (and our own discovery experience)[6,10,11,15] we suggest multicyclic systems for the design of novel kinase inhibitors.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmcl.2015.04.005.

References and notes

1. Zhang, J.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, *9*, 28.
2. Liu, Y.; Gray, N. S. *Nat. Chem. Biol.* **2006**, *2*, 358.
3. Zuccotto, F.; Ardini, E.; Casale, E.; Angiolini, M. *J. Med. Chem.* **2010**, *53*, 2681.
4. Gavrin, L. K.; Saiah, E. *Medchemcomm* **2013**, *4*, 41.
5. Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. *Biochemistry* **2012**, *51*, 4990.
6. Zhao, H. T.; Dong, J.; Lafleur, K.; Nevado, C.; Caflisch, A. *ACS Med. Chem. Lett.* **2012**, *3*, 834.
7. Kolb, P.; Kipouros, C. B.; Huang, D.; Caflisch, A. *Proteins* **2008**, *73*, 11.
8. Zhao, H.; Gartenmann, L.; Dong, J.; Spiliotopoulos, D.; Caflisch, A. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2493.
9. Zhao, H.; Caflisch, A. *Eur. J. Med. Chem.* **2015**, *91*, 4.
10. Lafleur, K.; Huang, D.; Zhou, T.; Caflisch, A.; Nevado, C. *J. Med. Chem.* **2009**, *52*, 6433.
11. Lafleur, K.; Dong, J.; Huang, D.; Caflisch, A.; Nevado, C. *J. Med. Chem.* **2013**, *56*, 84.
12. Bemis, G. W.; Murcko, M. A. *J. Med. Chem.* **1996**, *39*, 2887.
13. Taylor, R. D.; MacCoss, M.; Lawson, A. D. *J. Med. Chem.* **2014**, *57*, 5845.
14. Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.* **2005**, *45*, 177.
15. Unzue, A.; Dong, J.; Lafleur, K.; Zhao, H.; Frugier, E.; Caflisch, A.; Nevado, C. *J. Med. Chem.* **2014**, *57*, 6834.
16. Bissantz, C.; Kuhn, B.; Stahl, M. *J. Med. Chem.* **2010**, *53*, 5061.
17. Zhao, H.; Huang, D.; Caflisch, A. *ChemMedChem* **2012**, *7*, 1983.
18. Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. *J. Med. Chem.* **2013**, *56*, 1363.

19. Fedorov, O.; Huber, K.; Eisenreich, A.; Filippakopoulos, P.; King, O.; Bullock, A. N.; Szklarczyk, D.; Jensen, L. J.; Fabbro, D.; Trappe, J.; Rauch, U.; Bracher, F.; Knapp, S. *Chem. Biol.* **2011**, *18*, 67.
20. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3.
21. Ritchie, T. J.; Macdonald, S. J.; Young, R. J.; Pickett, S. D. *Drug Discovery Today* **2011**, *16*, 164.
22. Chen, C. Y. *PLoS ONE* **2011**, *6*, e15939.
23. Xing, L.; Rai, B.; Lunney, E. A. *J. Comput. Aided Mol. Des.* **2014**, *28*, 13.
24. Wagner, J.; von Matt, P.; Sedrani, R.; Albert, R.; Cooke, N.; Ehrhardt, C.; Geiser, M.; Rummel, G.; Stark, W.; Strauss, A.; Cowan-Jacob, S. W.; Beerli, C.; Weckbecker, G.; Evenou, J. P.; Zenke, G.; Cottens, S. *J. Med. Chem.* **2009**, *52*, 6193.
25. Zhao, Z.; Wu, H.; Wang, L.; Liu, Y.; Knapp, S.; Liu, Q.; Gray, N. S. *ACS Chem. Biol.* **2014**, *9*, 1230.
26. Zhao, H.; Huang, D. *PLoS ONE* **2011**, *6*, e19923.