Identification of allosteric binding sites for PI3K α oncogenic mutant specific inhibitor design

Michelle S. Miller^a, Sweta Maheshwari^b, Fiona M. McRobb^c, Kenneth W. Kinzler^d, L. Mario Amzel^b, Bert Vogelstein^d, Sandra B. Gabelli^{a,b,e*}

^aDepartment of Oncology

^bDepartment of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205

^cSchrödinger, Inc., 120 West 45th Street, New York, NY 10036

^dLudwig Center and Howard Hughes Medical Institutions, Johns Hopkins University School of Medicine, Baltimore, MD 21287

^eDepartment of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205

*To whom correspondence should be addressed: gabelli@jhmi.edu (S.B.G)

Keywords: *PIK3CA*, PI3K, PIP₂, PIP₃, fragment-based drug discovery

Abstract

PIK3CA, the gene that encodes the catalytic subunit of phosphatidylinositol 3-kinase α (PI3K α), is frequently mutated in breast and other types of cancer. A specific inhibitor that targets the mutant forms of PI3K α could maximize treatment efficiency while minimizing side-effects. Herein we describe the identification of novel binding pockets that may provide an opportunity for the design of mutant selective inhibitors. Using a fragment-based approach, we screened a library of 352 fragments (MW <300 Da) for binding to PI3K α by X-ray crystallography. Five novel binding pockets were identified, each providing potential opportunities for inhibitor design. Of particular interest was a binding pocket near Glu542, which is located in one of the two most frequently mutated domains.

1. Introduction

Phosphatidylinositol 3-kinase (PI3K) inhibitors have been gaining momentum as a viable cancer treatment strategy, as documented by the recent FDA approval of the first PI3K inhibitor, idelalisib (**1**, Fig. 1a). It was approved for use in combination with rituximab in relapsed chronic lymphoid leukemia, follicular non-Hodgkin's B-cell lymphoma and small lymphocytic lymphoma.¹ Idelalisib selectively targets the PI3K δ isoform, which is overexpressed in these hematological malignancies.²⁻⁴ During a Phase III clinical trial, the addition of idelalisib to rituximab produced a rapid and relatively long-lasting response in 81% of patients compared with a 13% response-rate in the group receiving rituximab alone.⁵

While the PI3K δ isoform is overexpressed in hematological malignancies, the PI3K α isoform is frequently mutated in solid tumors, particularly those of the breast, ovary, colorectum, and head and neck.⁶⁻¹⁰ As many as 80% of these mutations occur in three hotspot positions within the catalytic subunit, p110 α (encoded by the *PIK3CA* gene) of the heterodimeric enzyme (Fig. 1b). Of these three hotspots, the two helical domain mutants, E542K and E545K, are located in a loop of the helical domain of p110 α that binds in a groove of the nSH2 domain of the regulatory subunit, p85 α . This interaction is thought to result in the auto-inhibition of the enzyme. Both these mutations disrupt this interaction and relieve the auto-inhibition of PI3K α .^{11,12} In contrast, the H1047R mutation is located in the kinase domain and induces conformational changes that increase PI3K α membrane association and lipid binding.¹¹⁻¹³

Figure 1: PI3K inhibitors and domain organization a) Chemical structures of PI3K inhibitors approved or in clinical trials. **1.** idelalisib, **2.** buparlisib, **3.** alpelisib. b) Domain organization of PI3K α heterodimer. Colored domains are included in the construct used for crystallization.

The most advanced inhibitors in clinical trials for solid tumors harboring *PIK3CA* mutations include buparlisib (BKM-120, **2**), a pan-PI3K inhibitor and alpelisib (BYL-719, **3**), a PI3K α -selective inhibitor (Fig. 1a).^{14–16} Buparlisib increased progression-free survival by 4 months in a Phase III trial in aromatase inhibitor resistant breast cancer, and responses correlated with the presence of mutations in *PIK3CA*.¹⁷ However, 25% of patients experienced serious side effects, including hyperglycemia and liver toxicity, which limited treatment. Concerns about the toxicity of pan-PI3K inhibitors led to the investigation of PI3K α isoform selective agents such as alpelisib. Although not expected to completely abrogate adverse effects, treatment with PI3K α selective inhibitors may increase the therapeutic window, potentially leading to more effective treatments.¹⁸

All PI3K isoforms are involved in vital cellular processes such as metabolism, growth, proliferation, and migration. PI3K α , in particular, plays a central role in the regulation of glucose metabolism and growth.^{19–21} Many of the observed adverse effects stem from on-target toxicity, so selective targeting of the oncogenic mutant enzymes may be one approach to increase the therapeutic window. This should lead to more tolerable and efficacious treatments. One of the major limitations is the structural similarity of the ATP binding sites between the wild-type and mutant PI3K α . Comparison of the ATP-binding sites of the wild-type and the H1047R oncogenic mutant reveals no major differences.^{12,22,23} There is no structural information available for the helical domain mutants; however, given the inability thus far to develop an ATP-competitive inhibitor that is selective for these mutants, it is unlikely that there are any major differences in the ATP-binding site itself.

In the parallel field of protein kinase inhibitor development, selectivity between highly related kinases has posed a significant challenge. The development of Akt inhibitors provides a paradigmatic example. Early on, significant efforts were invested in the search for ATP-competitive Akt inhibitors. However, many of these inhibitors that advanced to the clinic were ultimately discontinued due to a lack of selectivity and the resultant severe toxicity issues.²⁴ The subsequent discovery of MK-2206, an allosteric inhibitor, revolutionized this field.^{25,26} Adopting a novel binding mechanism, this drug binds to the interface between the kinase and Plekstrin homology domains and stabilizes the inactive conformation of the enzyme. In contrast to the ATP-competitive inhibitors, MK-2206 has been well-tolerated in the clinic, and showed good efficacy.²⁷ Although some ATP-competitive Akt inhibitors with better selectivity are now under clinical evaluation, in

general, the allosteric inhibitors show greater selectivity due to the less conserved binding site.²⁸ The auto-inhibited conformation of PI3K α could potentially be targeted in a similar fashion.

Allosteric inhibitors have a number of potential benefits over the canonical ATP competitive inhibitors.^{29–31} Allosteric sites tend to be less homologous than ATP-binding sites among related kinases, potentially offering improved selectivity. Furthermore, as opposed to the highly competitive intellectual property space surrounding ATP-like inhibitors, allosteric binding sites allow for the exploration of novel chemical space. Binding sites with different topographies also have the potential to yield inhibitors with significantly improved pharmacokinetic properties. A final limitation of ATP-competitive inhibitors is the need to compete with millimolar concentrations of ATP within the cell, necessitating extremely high potency. Allosteric binding sites avoid this requirement and permit the identification of lower affinity binders as potential drug candidates.^{29–31}

Significant challenges exist, however, in the search for allosteric inhibitors. Due to the nature of their target binding site, typical kinase screening libraries predominantly consist of ATP analogs, biasing the search towards ATP-competitive inhibitors. However, even with a generic, non-target focused library of compounds, it is difficult to specifically screen for allosteric binders, since the enzymatic assays typically used either only detect binding at the ATP-binding site, or do not provide a simple way to discriminate between allosteric and orthosteric binders.^{29,31} An ATP-consumption based assay, for example, is inherently biased towards the discovery of ATP-competitive inhibitors.

Fragment screening by X-ray crystallography provides a way to identify novel binding sites.^{31,32} Fragments are small molecular weight compounds (MW < 300 Da) that are frequently used for the discovery of novel drugs. They generally possess good drug-like properties, high ligand efficiency (a measure of affinity corrected by size), and good coverage of available chemical space. Each of these properties makes fragments well-suited to the search for novel binding sites. Though laborious, crystallography is a uniquely valuable approach to fragment screening. It is sensitive, able to detect binders with affinities as low as 10 mM, provides structural information about the binding site and, unlike NMR, is suitable for large molecular weight proteins, such as PI3K α .^{31,33,34}

The pursuit of allosteric inhibitors for PI3K is an underexplored area, and could hold significant advantages in the search for mutant selective inhibitors. By screening a fragment library of 352 compounds, we have identified 10 novel binding sites on PI3K α , at least one of which holds promise for the development of oncogenic mutant selective inhibitors.

2. Materials and Methods

2.1 Molecular Modeling

2.1.1 Protein Preparation

The crystal structure of PI3K α (PDB ID: 4OVU) was prepared using the Protein Preparation Wizard in Maestro (Schrödinger Release 2016-2: Maestro, version 10.6, Schrödinger, LLC, New York, NY, 2016).³⁵ This involves the assignment of bond orders and formal charges and the addition of hydrogens. The hydrogen bonding network within the protein is optimized (including the reorientation of thiol and hydroxyl groups, sampling Asn, Gln and His side

chains, and the prediction of the protonation states of His, Asp and Glu), followed by a brief minimization.

2.1.2 Binding site prediction

SiteMap (Schrödinger Release 2016-2: SiteMap, version 3.9, Schrödinger, LLC, New York, NY, 2016) was used to identify potentially druggable binding sites, using default settings.^{36,37}

2.2 Protein expression

SF9 cells were grown in suspension culture in SF-900 III Serum Free Media (Invitrogen) supplemented with 0.5% penicillin-streptomycin at 27°C. At a density of 4×10^6 cells/mL, cells were infected with WT p110 α and p85 α (or niSH2) viruses at a multiplicity of infection ratio of 3:2. Media was supplemented with J-32, a PI3K inhibitor, as described in Mandelker *et al.*¹² Cells were harvested 48 hours after infection and the cell pellet collected through centrifugation at 500 x *g*. Protein purification was performed as previously described.^{12,22,23}

2.3 Crystallization and data collection

Crystallization was performed as previously described and improved with successive rounds of macroseeding.^{12,23} Crystals of p110 α /niSH2 were soaked for 30 minutes in mother liquor supplemented with 2-4 fragments each at 2 mM concentration. Fragments were obtained from Zenobia Therapeutics. X-ray diffraction data were collected at Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source or at the Berkeley Center for Structural Biology Beamline 502 at Advanced Light Source. Data from crystals that diffracted to a resolution of 3.5 Å or better were collected and processed with HKL2000 (Supp. Table 1).³⁸

2.4 Structure determination and analysis

The structures were determined by direct refinement using the coordinates of the previously determined WT p110 α /niSH2 (PDB ID 4OVU) as a model.²³ After rigid body and

positional refinement, the fragments were built in the difference electron density maps using the program Coot for model building.³⁹ Iterative rounds of refinement were performed with REFMAC 5.0.^{40–42} Model quality was assessed using Coot. Visualization, analysis and figure preparation were carried out with PyMOL (The PyMOL Molecular Graphics System, Version 1.6 Schrödinger, LLC).

2.5 Fluorescence Polarization Assay

The fluorescence polarization assay kit was purchased from Echelon Biosciences and adapted to a 96-well format. The fragments were dissolved at 200 mM in dimethyl sulfoxide (DMSO). Fragment dilutions at appropriate concentrations were prepared in 10% (v/v) DMSO and added to the reaction. PI3K α enzymatic activity was determined in the presence of 10 μ M PIP₂ and 25 μ M ATP and incubated at room temperature for 45 minutes. Increasing concentrations of fragments were used to generate a dose-response curve which was analyzed using GraphPad Prism v6 for OSX to calculate the IC₅₀.

3. Results and Discussion

To identify novel binding sites, we used X-ray crystallography as our primary screening method. We have previously reported the structure of wild-type PI3K α alone and in complex with its lipid substrate.^{22,23} The construct used for crystallization contains the full-length p110 α , which has five domains: the adaptor-binding domain (ABD), Ras-binding domain (RBD), C2 domain, helical domain and kinase domain; in addition to the truncated regulatory subunit, p85 α , which contains just the N-terminal SH2 domain and the inter-SH2 domain (hereafter termed niSH2) (Fig. 1b). This is the minimal portion of p85 α required for auto-inhibition and stable protein expression.^{22,43} We soaked wild-type PI3K α crystals with fragment cocktails containing 2-4 fragments from the Zenobia Fragment Library (352

compounds, Zenobia Therapeutics Inc.) at 2 mM individual concentrations. A total of 91 crystals were soaked, and we collected 24 datasets that diffracted better than 3.5 Å. In the structures determined with these datasets, we found ligand electron density in 10 distinct regions throughout the protein (Fig. 2a). Interestingly, most of the identified binding sites were located at the interfaces between multiple domains. To reduce the number of binding sites for further exploration, we examined these ten regions in more detail to decide whether they could represent druggable, novel allosteric binding sites. We determined that five of these regions were not in well-defined pockets within the protein; instead, in these cases, the fragment interacted with surface residues of the protein or made isolated interactions with a flexible loop. These five regions were not considered of interest for future drug design and will not be discussed further.

Figure 2: Identification of allosteric binding sites. a) Location of allosteric binding sites. Fragment binding sites are shown in blue; the substrate (ATP and PIP₂) binding sites are shown in gray for reference. PI3K domains are colored as in Fig. 1b. b) Binding site 1. Three fragments, **4-6**, were identified in this binding site at the interface between the RBD, kinase and helical domains. The surface representation of the protein is shown in gray. c) Binding site 2. Four fragments, **7-10**, were identified in this binding site between the RBD and helical domains. d) Binding site 3. Two fragments, **7** and **11**, were found at this binding site located between the ABD and kinase domain. e) Binding site 4. A single fragment, **12**, binds at this site located between the C2 and helical domains. f) Phosphopeptide binding site. A total of 12 fragments, **13-24**, were bound to this site which overlaps the physiological phosphopeptide site at the interface between the nSH2 and helical domains. This binding site is in close proximity to E542, which is mutated to a lysine residue in many cancers.

Of the five sites, the first binding site was a deep pocket located at the interface of the RBD, kinase and helical domains (Fig. 2b). Three different fragments bound in this pocket: an indole **4**, a benzimidazole **5**, and a piperidine **6**. The fragments were sandwiched between two residues, F666 of the helical domain and R818 of the kinase domain. These two residues

are both capable of π -stacking with the aromatic fragments, similar to the 'hydrophobic clamp' that has been observed with phosphodiesterase inhibitors.⁴⁴ The site is bordered by a number of residues that can participate in hydrogen bonding: N170, S629, R662, H670, Q815. The fragments overlap in their binding sites, but each one presents numerous distinct opportunities for further derivatization to increase binding affinity.

The second site is a shallow binding site on the surface of the helical domain, with some residues from the adjacent RBD contributing hydrogen bonding partners (Fig. 2c). The binding site is lined by polar and charged residues: K621, T624, D625 and K656 from the helical domain, with E263 and Y265 extending from a loop of the RBD. Four fragments (**7-10**), including monocyclic and bicyclic aromatic compounds, bound to this site.

Two other sites, which were very polar in nature, were also identified. Two fragments (**6** and **11**) bound in one of these sites located at the interface between the kinase domain and the ABD (Fig. 2d). This deep, narrow pocket has R115 and F119 at the base of the site and is lined with polar residues on either side: R704, E707, D746 from the kinase domain; and D84 and T86 from the ABD. A single fragment, **12**, was identified in the other binding site located at the interface between the C2 domain, helical domain, and ABD-RBD linker (Fig. 2e). This binding site is completely devoid of hydrophobic residues and is bordered by D133, E135, T462, S464, and Q682. Either of these two polar binding sites could be used as the basis for the development of compounds with good solubility and pharmacokinetic properties.

A total of 12 fragments (**13-24**) were modeled into a fifth binding site lying at the interface between the helical, kinase, C2 and nSH2 domains, also known as the phosphopeptide

binding site (Fig. 2f). Following activation of RTKs, the nSH2 domain of PI3K binds to the auto-phosphorylated tyrosine residues leading to the activation of kinase activity.¹² This fragment binding site overlaps with the binding site of such phosphorylated peptides. The sides of the binding site are lined by the residues E365, L540, E542, L570, C604 and N605 of p110 α and residues E341, N344 and N377 of p85 α . The top and bottom of the binding site are defined by residues I543 and F1016 of p110 α . The structures of the twelve fragments show some commonalities: six of the fragments are amino-substituted aromatic rings and four of the fragments contain a hydroxyl group (Fig. 2f). The identification of this binding site is particularly interesting because one of the residues lining it, E542, is one of the three most commonly mutated residues in p110 α ; the most frequent mutation is a substitution to lysine. It is conceivable that an inhibitor could be developed that makes an ionic interaction with the lysine present in this oncogenic mutant; such an inhibitor would bind to the mutant but would not interact with the glutamate present in the wild-type.

To determine if this binding site would be useful for the design of mutant selective inhibitors, we next evaluated the effects of these fragments on PI3K activity. Binding at this site could stabilize the auto-inhibited conformation, thereby inactivating the enzyme, or conversely activate the enzyme in the absence of a phosphorylated peptide. To distinguish between these possibilities, we used a fluorescence polarization assay (Echelon Biosciences), which detects the formation of PIP₃.⁴⁵ The fragments from each of the phosphopeptide binding site cocktails were screened for wild-type PI3K α inhibition against both p110 α /niSH2 (construct used for crystallography) and full-length enzymes at a concentration of 100 μ M. At higher concentrations than this, interference with the assay was observed, possibly due to aggregation of the fragments. The results for the niSH2 and

full-length constructs were similar. Most fragments tested had less than 25% inhibition at 100 μ M, but two compounds, **18** and **21**, showed, respectively, 92% and 87% inhibition of full-length PI3K α at 100 μ M (Fig. 2f). These high levels of inhibition are particularly surprising given the small size of the fragments and their intended use only as one component of a larger compound. Furthermore, the IC₅₀ values for **18** and **21** were determined to be 20 ± 4 μ M and 33 ± 4 μ M, respectively. Although rare, inhibitory activities like these by low molecular weight fragments are not without precedent; La *et al.* identified fragments that inhibited HIV reverse transcriptase with IC₅₀ values of 20-100 μ M.⁴⁶ The low IC₅₀ values could also reflect potential assay interference. Fragment **21** has previously been identified as a frequent hitter, and as such is unlikely to prove a good lead for the development of inhibitors.⁴⁷ However, to the best of our knowledge, **18** has not previously been identified as a frequent hitter, and may provide a more useful starting point.

To further investigate these novel binding sites on PI3K α , we used SiteMap, a program that predicts binding sites on a given protein using a grid-based approach.^{36,37} The number of predicted binding sites was limited to five. The two top ranked binding sites coincide with the phosphopeptide binding site, also identified through our fragment screen, at the interface between the nSH2 and helical domains. The binding sites extend to encompass the entire interface between p110 α and the nSH2 domain of p85 α (Fig. 2a). One of the predicted sites extends the binding site along the edge of the helical domain to a pocket between the nSH2 and C2 domains (Fig. 3a). Most of the residues contributed by the nSH2 domain are charged, lysines and aspartates, creating a “charged” base. One side of the site is lined by residues of the helical domain, extending towards E545 of the helical domain, suggesting inhibitors binding at this site could also potentially incorporate some selectivity

towards the other helical domain mutant, E545K. The β -loop of the C2 domain from 361-370 then forms the roof of the site. Fragments were identified at one extreme of this site, right up against the helical domain. This SiteMap predicted site defines a larger site that could incorporate drug-like molecules and potentially be used in virtual screening.

The other binding site adjacent to the phosphopeptide site extends instead along the boundary of the kinase domain towards the activation loop and into a large pocket formed between the iSH2 domain and p110 α . This predicted site is extremely large, and most likely represents several smaller potential binding sites. Of interest for the further development of these fragments into inhibitors is the pocket formed between the nSH2, C2 and kinase domains, with E542 from the helical domain forming the base of the site. The two sides of the site are lined by residues from nSH2 and C2, and the roof of the site is formed by the activation loop of the kinase domain. These two predicted sites adjacent to our experimental fragment sites suggest significant opportunities for fragment growth in either direction to design a potent inhibitor that selectively targets the oncogenic mutant, E542K PI3K α .

Figure 3: SiteMap predicted binding sites. The locations and details of the binding sites predicted by SiteMap are shown. PI3K domains are colored according to Fig. 1b. The location of the binding sites identified by SiteMap are highlighted with green spheres. a) Sites 1 and 2 cover most of the interface between p110 α and p85 α , including the phosphopeptide binding site. b) Site 3 is located at the interface between the ABD and kinase domains. c) Site 4 is located close to the ATP-binding site in the kinase domain. It is a large binding site that has many potential sub-sites for inhibitor binding. d) Site 5 is a smaller binding site located at the interface between the nSH2, C2 and helical domains.

The next highest ranked binding site was located at the interface between the ABD and the kinase domain. The binding site predicted by SiteMap extends to cover the interface between the ABD and kinase domain (Fig. 3b). Molecules binding at this site could potentially disrupt heterodimer formation or stability, however are unlikely to yield mutant selective inhibitors. The next predicted binding site sits behind the ATP-binding site between the kinase, RBD and helical domains (Fig. 3c). This binding site encompasses our experimentally determined binding site 1 (Fig. 2b). The last predicted binding site is between the nSH2 and helical domains, near the linker with the C2 domain (Fig. 3d).

In sum, we have identified a novel binding site containing one of the most frequently mutated amino acid residues in p110 α , E542K, found in multiple cancer types. SiteMap predicted binding sites extend this experimental site in two different directions, one encompassing another frequently mutated residue, E545K, and the other the activation loop. These binding sites hold significant promise for the design of novel, allosteric, selective inhibitors of the E542K or E545K oncogenic mutant enzymes.

Acknowledgements

We thank Jesse Yoder for helpful discussions and critical reading of the manuscript. This work was supported by NIH Grants CA062924 and CA043460 and by The Virginia and D.K. Ludwig Fund for Cancer Research. SBG is an Alexander and Margaret Stewart Trust Cancer fellow. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source was provided by Eli Lilly Company, which operates the facility. This research used resources from the Berkeley Center for Structural Biology supported in part by the National Institutes of Health, National Institute of General Medical Sciences, and the Howard Hughes Medical Institute. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We acknowledge the use and service of the JHU SOM Mass Spectrometry and Proteomics Core supported by the Sidney Kimmel Comprehensive

Cancer Center (NCI center grant 2P30 CA006973) and the use of the JHU SOM Eukaryotic Tissue Culture Facility.

Protein structures have been deposited in the Protein Data Bank with the following accession codes: 5SW8, 5SWG, 5SWO, 5SWP, 5SWR, 5SWT, 5SX8, 5SX9, 5SXA, 5SXB, 5SXC, 5SXD, 5SXE, 5SXF, 5SXI, 5SXJ, 5SXX.

Figure Legends

Fig. 1 PI3K inhibitors and domain organization a) Chemical structures of PI3K inhibitors approved or in clinical trials. **1.** idelalisib, **2.** buparlisib, **3.** alpelisib. b) Domain organization of PI3K α heterodimer. Colored domains are included in the construct used for crystallization.

Fig. 2 Identification of allosteric binding sites. a) Location of allosteric binding sites. Fragment binding sites are shown in blue; the substrate (ATP and PIP₂) binding sites are shown in gray for reference. PI3K domains are colored as in Fig. 1b. b) Binding site 1. Three fragments, **4-6**, were identified in this binding site at the interface between the RBD, kinase and helical domains. The surface representation of the protein is shown in gray. c) Binding site 2. Four fragments, **7-10**, were identified in this binding site between the RBD and helical domains. d) Binding site 3. Two fragments, **7** and **11**, were found at this binding site located between the ABD and kinase domain. e) Binding site 4. A single fragment, **12**, binds at this site located between the C2 and helical domains. f) Phosphopeptide binding site. A total of 12 fragments, **13-24**, were bound to this site which overlaps the physiological phosphopeptide site at the interface between the nSH2 and helical domains. This binding site is in close proximity to E542, which is mutated to a lysine residue in many cancers.

Fig. 3 SiteMap predicted binding sites. The locations and details of the binding sites predicted by SiteMap are shown. PI3K domains are colored according to Fig. 1b. The

location of the binding sites identified by SiteMap are highlighted with green spheres. a) Sites 1 and 2 cover most of the interface between p110 α and p85 α , including the phosphopeptide binding site. b) Site 3 is located at the interface between the ABD and kinase domains. c) Site 4 is located close to the ATP-binding site in the kinase domain. It is a large binding site that has many potential sub-sites for inhibitor binding. d) Site 5 is a smaller binding site located at the interface between the nSH2, C2 and helical domains.

References

1. Miller, B. W.; Przepiorka, D.; de Claro, R. A.; Lee, K.; Nie, L.; Simpson, N.; Gudi, R.; Saber, H.; Shord, S.; Bullock, J.; Marathe, D.; Mehrotra, N.; Hsieh, L. S.; Ghosh, D.; Brown, J.; Kane, R. C.; Justice, R.; Kaminskis, E.; Farrell, A. T.; Pazdur, R. *Clin. Cancer Res.* **2015**, *21*, 1525–1529.
2. Castillo, J. J.; Furman, M.; Winer, E. S. *Expert Opin. Invest. Drugs* **2012**, *21*, 15–22.
3. Herman, S. E. M.; Johnson, A. J. *Clin. Cancer Res.* **2012**, *18*, 4013–4018.
4. Wei, M.; Wang, X.; Song, Z.; Jiao, M.; Ding, J.; Meng, L.-H.; Zhang, A. *Med. Res. Rev.* **2015**, *35*, 720–752.
5. Furman, R. R.; Sharman, J. P.; Coutre, S. E.; Cheson, B. D.; Pagel, J. M.; Hillmen, P.; Barrientos, J. C.; Zelenetz, A. D.; Kipps, T. J.; Flinn, I.; Ghia, P.; Eradat, H.; Ervin, T.; Lamanna, N.; Coiffier, B.; Pettitt, A. R.; Ma, S.; Stilgenbauer, S.; Cramer, P.; Aiello, M.; Johnson, D. M.; Miller, L. L.; Li, D.; Jahn, T. M.; Dansey, R. D.; Hallek, M.; O'Brien, S. M. *N. Engl. J. Med.* **2014**, *370*, 997–1007.
6. Dirican, E.; Akkiprik, M.; Özer, A. *Tumour Biol.* **2016**, *37*, 7033–7045.
7. Du, L.; Chen, X.; Cao, Y.; Lu, L.; Zhang, F.; Bornstein, S.; Li, Y.; Owens, P.; Malkoski, S.; Said, S.; Jin, F.; Kulesz-Martin, M.; Gross, N.; Wang, X.-J.; Lu, S.-L. *Oncogene* **2016**, *35*, 4641–4652.
8. Wang, D.; Wang, M.; Jiang, N.; Zhang, Y.; Bian, X.; Wang, X.; Roberts, T. M.; Zhao, J. J.; Liu, P.; Cheng, H. *Oncotarget* **2016**, *7*, 13153–13166.
9. Hao, Y.; Samuels, Y.; Li, Q.; Krokowski, D.; Guan, B.-J.; Wang, C.; Jin, Z.; Dong, B.; Cao, B.; Feng, X.; Xiang, M.; Xu, C.; Fink, S.; Meropol, N. J.; Xu, Y.; Conlon, R. A.; Markowitz, S.; Kinzler, K. W.; Velculescu, V. E.; Brunengraber, H.; Willis, J. E.; LaFramboise, T.; Hatzoglou, M.; Zhang, G.-F.; Vogelstein, B.; Wang, Z. *Nat. Commun.* **2016**, *7*, 11971.
10. Samuels, Y.; Wang, Z.; Bardelli, A.; Silliman, N.; Ptak, J.; Szabo, S.; Yan, H.; Gazdar, A.; Powell, S. M.; Riggins, G. J.; Willson, J. K. V.; Markowitz, S.; Kinzler, K. W.; Vogelstein, B.; Velculescu, V. E. *Science* **2004**, *304*, 554.
11. Carson, J. D.; Van Aller, G.; Lehr, R.; Sinnamon, R. H.; Kirkpatrick, R. B.; Auger, K. R.; Dhanak, D.; Copeland, R. A.; Gontarek, R. R.; Tummino, P. J.; Luo, L. *Biochem. J.* **2008**, *409*, 519–524.
12. Mandelker, D.; Gabelli, S. B.; Schmidt-Kittler, O.; Zhu, J.; Cheong, I.; Huang, C. H.; Kinzler, K. W.; Vogelstein, B.; Amzel, L. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16996–17001.
13. Hon, W.-C.; Berndt, A.; Williams, R. L. *Oncogene* **2012**, *31*, 3655–3666.
14. Burger, M. T.; Pecchi, S.; Wagman, A.; Ni, Z.-J.; Knapp, M.; Hendrickson, T.; Atallah, G.; Pfister, K.; Zhang, Y.; Bartulis, S.; Frazier, K.; Ng, S.; Smith, A.; Verhagen, J.; Haznedar, J.; Huh, K.; Iwanowicz, E.; Xin, X.; Menezes, D.; Merritt, H.; Lee, I.; Wiesmann, M.; Kaufman, S.; Crawford, K.; Chin, M.; Bussiere, D.; Shoemaker, K.; Zaror, I.; Maira, S.-M.; Voliva, C. F. *ACS Med. Chem. Lett.* **2011**, *2*, 774–779.
15. Maira, S.-M.; Pecchi, S.; Huang, A.; Burger, M.; Knapp, M.; Sterker, D.; Schnell, C.; Guthy, D.; Nagel, T.; Wiesmann, M.; Brachmann, S.; Fritsch, C.; Dorsch, M.; Chène, P.; Shoemaker, K.; Pover, A. D.; Menezes, D.; Martiny-Baron, G.; Fabbro, D.; Wilson, C. J.; Schlegel, R.; Hofmann, F.; García-Echeverría, C.; Sellers, W. R.; Voliva, C. F. *Mol. Cancer Ther.* **2012**, *11*, 317–328.
16. Furet, P.; Guagnano, V.; Fairhurst, R. A.; Imbach-Weese, P.; Bruce, I.; Knapp, M.; Fritsch, C.; Blasco, F.; Blanz, J.; Aichholz, R.; Hamon, J.; Fabbro, D.; Caravatti, G. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3741–3748.
17. Baselga, J.; Im, S.-A.; Iwata, H.; Clemons, M.; Ito, Y.; Awada, A.; Chia, S.; Jagiello-Grusfeld, A.; Pistilli, B.; Tseng, L.-M.; Hurvitz, S.; Masuda, N.; Cortes, J.; De Laurentiis, M.; Arteaga, C.;

- Jiang, Z.; Jonat, W.; Hachemi, S.; Le Mouhaer, S.; Di Tomaso, E.; Urban, P.; Massacesi, C.; Campone, M. *SABCS Abstracts* **2015**, S6-1.
18. Mayer, I. A.; Abramson, V.; Balko, J.; Sanders, M.; Juric, D.; Solit, D.; Li, Y.; Cantley, L.; Winer, E.; Arteaga, C. *Cancer Res.* **2015**, *75*, CT232.
 19. Zhao, J. J.; Cheng, H.; Jia, S.; Wang, L.; Gjoerup, O. V.; Mikami, A.; Roberts, T. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16296–16300.
 20. Foukas, L. C.; Claret, M.; Pearce, W.; Okkenhaug, K.; Meek, S.; Peskett, E.; Sancho, S.; Smith, A. J. H.; Withers, D. J.; Vanhaesebroeck, B. *Nature* **2006**, *441*, 366–370.
 21. Knight, Z. A.; Gonzalez, B.; Feldman, M. E.; Zunder, E. R.; Goldenberg, D. D.; Williams, O.; Loewith, R.; Stokoe, D.; Balla, A.; Toth, B.; Balla, T.; Weiss, W. A.; Williams, R. L.; Shokat, K. M. *Cell* **2006**, *125*, 733–747.
 22. Huang, C. H.; Mandelker, D.; Schmidt-Kittler, O.; Samuels, Y.; Velculescu, V. E.; Kinzler, K. W.; Vogelstein, B.; Gabelli, S. B.; Amzel, L. M. *Science* **2007**, *318*, 1744–1748.
 23. Miller, M. S.; Schmidt-Kittler, O.; Bolduc, D. M.; Brower, E. T.; Chaves-Moreira, D.; Allaire, M.; Kinzler, K. W.; Jennings, I. G.; Thompson, P. E.; Amzel, L. M.; Vogelstein, B.; Gabelli, S. B. *Oncotarget* **2014**, *5*, 5198–5208.
 24. Lindsley, C. W. *Curr. Top. Med. Chem.* **2010**, *10*, 458–477.
 25. Yan, L. *Cancer Res.* **2009**, *69*, DDT01-1.
 26. Lindsley, C. W.; Zhao, Z.; Leister, W. H.; Robinson, R. G.; Barnett, S. F.; Defeo-Jones, D.; Jones, R. E.; Hartman, G. D.; Huff, J. R.; Huber, H. E.; Duggan, M. E. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 761–764.
 27. Tripathy, D.; Chien, A. J.; Hylton, N.; Buxton, M. B.; Ewing, C. A.; Wallace, A. M.; Forero, A.; Kaplan, H. G.; Nanda, R.; Albain, K. S.; Moulder, S. L.; Haley, B. B.; DeMichele, A.; Symmans, W. F.; Veer, L. van 't; Paoloni, M.; Esserman, L.; Berry, D. A.; Yee, D. J. *J. Clin. Oncol.* **2015**, *33*, (suppl; abstr 524).
 28. Nitulescu, G. M.; Margina, D.; Juzenas, P.; Peng, Q.; Olaru, O. T.; Saloustros, E.; Fenga, C.; Spandidos, D. A.; Libra, M.; Tsatsakis, A. M. *Int. J. Oncol.* **2016**, *48*, 869–885.
 29. Cowan-Jacob, S. W.; Jahnke, W.; Knapp, S. *Future Med. Chem.* **2014**, *6*, 541–561.
 30. Wu, P.; Clausen, M. H.; Nielsen, T. E. *Pharmacol. Ther.* **2015**, *156*, 59–68.
 31. Ludlow, R. F.; Verdonk, M. L.; Saini, H. K.; Tickle, I. J.; Jhoti, H. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 15910–15915.
 32. Noble, C. G.; Lim, S. P.; Arora, R.; Yokokawa, F.; Nilar, S.; Seh, C. C.; Wright, S. K.; Benson, T. E.; Smith, P. W.; Shi, P.-Y. *J. Biol. Chem.* **2016**, *291*, 8541–8548.
 33. Caliandro, R.; Belviso, D. B.; Aresta, B. M.; de Candia, M.; Altomare, C. D. *Future Med. Chem.* **2013**, *5*, 1121–1140.
 34. Chilingaryan, Z.; Yin, Z.; Oakley, A. J. *Int. J. Mol. Sci.* **2012**, *13*, 12857–12879.
 35. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. J. *Comput. Aided Mol. Des.* **2013**, *27*, 221–234.
 36. Halgren, T. A. *J. Chem. Inf. Model.* **2009**, *49*, 377–389.
 37. Halgren, T. *Chem. Biol. Drug Des.* **2007**, *69*, 146–148.
 38. Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307–326.
 39. Emsley, P.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
 40. Vagin, A. A.; Steiner, R. A.; Lebedev, A. A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. N. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2184–2195.
 41. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr. D Biol. Crystallogr.* **1997**, *53*, 240–255.

42. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242.
43. Yu, J.; Zhang, Y.; McIlroy, J.; Rordorf-Nikolic, T.; Orr, G. A.; Backer, J. M. *Mol. Cell. Biol.* **1998**, *18*, 1379–1387.
44. Manallack, D. T.; Hughes, R. A.; Thompson, P. E. *J. Med. Chem.* **2005**, *48*, 3449–3462.
45. Drees, B. E.; Weipert, A.; Hudson, H.; Ferguson, C. G.; Chakravarty, L.; Prestwich, G. D. *Comb. Chem. High Throughput Screen.* **2003**, *6*, 321–330.
46. La, J.; Latham, C. F.; Tinetti, R. N.; Johnson, A.; Tyssen, D.; Huber, K. D.; Sluis-Cremer, N.; Simpson, J. S.; Headey, S. J.; Chalmers, D. K.; Tachedjian, G. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 6979–6984.
47. Baell, J. B.; Holloway, G. A. *J. Med. Chem.* **2010**, *53*, 2719–2740.

Table 1 Data collection and refinement statistics

	5SW8	5SWG	5SWO	5SWP	5SWR	5SWT	5SX8
Fragments	7, 12	5, 21	4, 16	6, 24	20, 26	17, 27	13, 15
Data collection							
Space group	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁
Cell dimensions							
<i>a</i> , <i>b</i> , <i>c</i> (Å)	115.6, 117.8, 152.9	114.5, 116.7, 149.5	114.8, 116.4, 149.4	114.1, 116.1, 148.3	114.2, 116.3, 149.3	114.5, 115.9, 149.3	113.9, 116.3, 148.0
Resolution (Å)	50-3.30 (3.37-3.30)	50-3.12 (3.23-3.12)	50-3.50 (3.63-3.50)	50-3.40 (3.52-3.40)	50.00-3.30 (3.42-3.30)	50.00- 3.50(3.56- 3.50)	50-3.45 (3.57-3.45)
<i>R</i> _{sym}	0.101 (0.734)	0.085 (0.709)	0.130 (0.601)	0.121 (0.875)	0.097 (0.639)	0.123 (0.973)	0.098 (0.925)
<i>I</i> / σ <i>I</i>	26.5 (2.4)	27.1 (2.1)	15.2 (2.0)	17.3 (2.0)	27.5 (2.5)	19.2 (1.8)	24.2 (1.8)
Completeness (%)	99.9 (99.7)	99.9 (98.8)	99.7 (98.1)	99.9 (99.9)	99.7 (98.8)	100 (100)	99.9 (100)
Redundancy	7.1 (6.7)	7.2 (6.1)	6.8 (5.0)	6.7 (6.3)	7.2 (6.5)	7.3 (7.5)	7.3 (7.4)
Unique reflections	31,971	36,555	25,700	27,455	30,235	25,886	25,993
Total reflections	226,811	263,120	174,299	184,879	216,686	188,483	188,731
X-ray Source	1.0	1.0	1.0	1.0	1.0	1.0	1.0
wavelength							
Refinement							
Resolution (Å)	93.24-3.30	49.14-3.11	48.73-3.50	48.91-3.41	48.54-3.31	92.57-3.50	91.49-3.47
No. reflections	30,096	34,630	24,324	25,928	28,638	24,456	24,618
<i>R</i> _{work} / <i>R</i> _{free}	0.225/0.298	0.206/0.261	0.188/0.274	0.215/0.296	0.197/0.272	0.208/0.284	0.207/0.275
No. atoms							
Protein	10,166	10,695	10,864	10,852	10,748	10,781	10,657
Ligand	57	28	35	67	44	37	38
<i>B</i> -factors							
Protein	124	101	117	118	121	142	149
Ligand	187	116	142	142	153	155	180
R.m.s. deviations							
Bond lengths (Å)	0.011	0.009	0.011	0.009	0.010	0.009	0.009
Bond angles (°)	1.61	1.46	1.54	1.36	1.45	1.34	1.31

Fragment	5SX9 14	5SXA 10	5SXB 23	5SXC 8	5SXD 22	5SXE 19, 28	5SXF 9
Data collection							
Space group	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	113.4, 116.1, 148.0	115.2, 117.4, 152.1	114.6, 116.8, 150.1	114.7, 117.3, 150.3	114.9, 116.4, 150.0	113.8, 116.3, 149.0	112.3, 115.9, 147.2
Resolution (Å)	50-3.50 (3.63-3.50)	50-3.35 (3.42-3.30)	50-3.30 (3.42-3.30)	50.00-3.55 (3.68-3.55)	50-3.50 (3.63-3.50)	50-3.50 (3.63-3.50)	50-3.45 (3.51-3.45)
<i>R</i> _{sym}	0.108 (0.874)	0.104 (0.862)	0.107 (0.988)	0.113 (0.848)	0.111 (0.883)	0.112 (0.952)	0.112 (0.938)
<i>I</i> / σ <i>I</i>	19.2 (2.3)	24.4 (2.3)	23.7 (2.1)	19.9 (2.3)	18.2 (1.8)	15.7 (1.5)	21.1 (2.0)
Completeness (%)	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	99.9 (100)	100 (100)
Redundancy	7.2 (7.3)	7.3 (7.4)	7.3 (7.5)	7.3 (7.5)	7.2 (7.2)	6.5 (6.6)	7.3 (7.4)
Unique reflections	24,477	30,528	31,023	25,308	25,936	25,370	25,565
Total reflections	177,395	221,875	226,528	184,302	186,212	165,094	186,277
X-ray Source wavelength	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Refinement							
Resolution (Å)	91.34-3.52	92.91-3.35	92.16-3.30	92.46-3.54	91.97-3.5	91.65-3.51	91.07-3.46
No. reflections	23,166	28,743	29,298	23,959	24,555	24,018	24,200
<i>R</i> _{work} / <i>R</i> _{free}	0.192/0.277	0.221/0.274	0.204/0.278	0.193/0.267	0.194/0.277	0.191/0.271	0.194/0.266
No. atoms							
Protein	10,583	10,476	10,475	10,305	10,659	10,483	9,960
Ligand	19	44	29	29	31	35	22
<i>B</i> -factors							
Protein	134	136	121	138	138	139	129
Ligand	145	188	146	168	159	181	157
R.m.s. deviations							
Bond lengths (Å)	0.011	0.009	0.010	0.011	0.010	0.011	0.010
Bond angles (°)	1.56	1.33	1.50	1.53	1.52	1.54	1.48

	5SXI 25	5SXJ 29	5S XK 18
Data collection			
Space group	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁	p2 ₁ 2 ₁ 2 ₁
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	114.2, 116.8, 150.5	114.8, 117.9, 151.0	113.3, 115.8, 149.0
Resolution (Å)	50-3.40 (3.52-3.40)	50-3.45 (3.57-3.45)	50-3.55 (3.61-3.55)
<i>R</i> _{sym}	0.133 (0.907)	0.129 (0.955)	0.137 (0.934)
<i>I</i> / σ <i>I</i>	18.8 (1.7)	25.8 (2.5)	18.7 (1.7)
Completeness (%)	98.6 (96.8)	99.7 (99.8)	99.7 (99.6)
Redundancy	5.1 (3.6)	7.0 (5.7)	5.0 (6.4)
Unique reflections	28,264	28,041	24,161
Total reflections	144,489	197,575	155,578
X-ray Source wavelength	1.0	1.0	1.0
Refinement			
Resolution (Å)	92.27-3.38	92.94-3.42	91.45-3.55
No. reflections	26,461	26,321	22,868
<i>R</i> _{work} / <i>R</i> _{free}	0.201/0.278	0.212/0.279	0.193/0.277
No. atoms			
Protein	10,332	10,370	10,791
Ligand	28	50	19
<i>B</i> -factors			
Protein	132	138	142
Ligand	161	188	147
R.m.s. deviations			
Bond lengths (Å)	0.010	0.010	0.011
Bond angles (°)	1.50	1.43	1.54