



# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

## Structural basis of nSH2 regulation and lipid binding in PI3K

The Harvard community has made this article openly available.  
[Please share](#) how this access benefits you. Your story matters.

<b>Citation</b>	Miller, M. S., O. Schmidt-Kittler, D. M. Bolduc, E. T. Brower, D. Chaves-Moreira, M. Allaire, K. W. Kinzler, et al. 2014. "Structural basis of nSH2 regulation and lipid binding in PI3K ." <i>Oncotarget</i> 5 (14): 5198-5208.
<b>Accessed</b>	February 16, 2015 11:12:27 PM EST
<b>Citable Link</b>	<a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:12987346">http://nrs.harvard.edu/urn-3:HUL.InstRepos:12987346</a>
<b>Terms of Use</b>	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a>

*(Article begins on next page)*

## Structural basis of nSH2 regulation and lipid binding in PI3K $\alpha$

Michelle S. Miller<sup>1,8</sup>, Oleg Schmidt-Kittler<sup>2,9,12</sup>, David M. Bolduc<sup>3,10</sup>, Evan T. Brower<sup>2,11</sup>, Daniele Chaves-Moreira<sup>4</sup>, Marc Allaire<sup>7</sup>, Kenneth W. Kinzler<sup>2</sup>, Ian G. Jennings<sup>1</sup>, Philip E. Thompson<sup>1</sup>, Philip A. Cole<sup>3</sup>, L. Mario Amzel<sup>4</sup>, Bert Vogelstein<sup>2</sup> and Sandra B. Gabelli<sup>4,5,6</sup>

<sup>1</sup> Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Parkville, Victoria, Australia.

<sup>2</sup> Ludwig Center for Cancer Genetics and Therapeutics and Howard Hughes Medical Institutions, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

<sup>3</sup> Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

<sup>4</sup> Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

<sup>5</sup> Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

<sup>6</sup> Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

<sup>7</sup> Photon Sciences, Brookhaven National Laboratory, Upton, New York, USA.

<sup>8</sup> Present Address: Department of Oncology, Johns Hopkins University School of Medicine, Baltimore Maryland, USA.

<sup>9</sup> Present Address: Sanofi, Cambridge, Massachusetts.

<sup>10</sup> Present Address: Center for Neurologic Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

<sup>11</sup> Present Address: Paragon Bioservices, Baltimore, Maryland.

<sup>12</sup> Present Address: Berkeley Center for Structural Biology, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California

**Correspondence to:** Sandra B. Gabelli, *email:* gabelli@jhmi.edu

**Keywords:** PIK3R1, p85, PIK3CA, PI3K, PIP<sub>2</sub>, PIP<sub>3</sub>

**Received:** June 23, 2014

**Accepted:** July 23, 2014

**Published:** July 25, 2014

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### ABSTRACT

**We report two crystal structures of the wild-type phosphatidylinositol 3-kinase  $\alpha$  (PI3K $\alpha$ ) heterodimer refined to 2.9 Å and 3.4 Å resolution: the first as the free enzyme, the second in complex with the lipid substrate, diC4-PIP<sub>2</sub>, respectively. The first structure shows key interactions of the N-terminal SH2 domain (nSH2) and iSH2 with the activation loop that suggest a mechanism by which the enzyme is inhibited in its basal state. In the second structure, the lipid substrate binds in a positively charged pocket adjacent to the ATP-binding site, bordered by the P-loop, the activation loop and the iSH2 domain. An additional lipid-binding site was identified at the interface of the ABD, iSH2 and kinase domains. The ability of PI3K $\alpha$  to bind an additional PIP<sub>2</sub> molecule was confirmed in vitro by fluorescence quenching experiments. The crystal structures reveal key differences in the way the nSH2 domain interacts with wild-type p110 $\alpha$  and with the oncogenic mutant p110 $\alpha$ H1047R. Increased buried surface area and two unique salt-bridges observed only in the wild-type structure suggest tighter inhibition in the wild-type PI3K $\alpha$  than in the oncogenic mutant. These differences may be partially responsible for the increased basal lipid kinase activity and increased membrane binding of the oncogenic mutant.**

## INTRODUCTION

Dysregulation of the phosphatidylinositol 3-kinase (PI3K) pathway plays a critical role in tumor pathogenesis, with up to 50% of human cancers displaying aberrations in signaling [1]. A key modulator of the pathway, PI3K $\alpha$ , is mutated in a high fraction of breast, colon, brain, head and neck, gastric, and endometrial cancers [2–6]. About 80% of these mutations are somatic, missense mutations concentrated in three hotspots in the p110 $\alpha$  catalytic subunit of the protein (Fig. 1A) [7]. Two helical domain mutants, p110 $\alpha$ E542K and p110 $\alpha$ E545K, become constitutively active through the loss of auto-inhibition by the N-terminal SH2 domain (nSH2) [8,9]. As such, they are insensitive to further activation by phosphotyrosine peptide binding [8]. In contrast, the kinase domain oncogenic mutant, p110 $\alpha$ H1047R, remains sensitive to activation by receptor tyrosine kinases. The mutation enhances membrane binding, resulting in increased substrate accessibility and elevation of the basal lipid kinase activity [8,10–12].

Since the publication of the first p110 $\gamma$  crystal structures in 1999, more than 80 crystal structures of the four Class I PI3K isoforms (PI3K $\alpha$ , PI3K $\beta$ , PI3K $\gamma$  and PI3K $\delta$ ) have been deposited in the Protein Data Bank (PDB) [13]. The structures of PI3K in complex with ATP and with a wide range of pan-PI3K and isoform selective inhibitors, spanning multiple chemical classes, have been determined [14]. Despite this large amount of structural information, the lipid substrate binding site has yet to be identified crystallographically. Biochemical data has established an important role for the positively charged residues of the activation loop in determining substrate specificity and lipid kinase activity [15,16]. Replacement of the activation loop in PI3K $\alpha$  with that of Class II or III PI3K renders PI3K $\alpha$  unable to phosphorylate PIP<sub>2</sub>. Lipid binding appears to be unaffected, however, as pre-incubation of these hybrids with PIP<sub>2</sub> prevented binding of the covalent inhibitor, wortmannin [15]. Modeling studies of p110 $\alpha$  and p110 $\gamma$  support the biochemical data, placing the phospholipid head group between the positively charged residues of the P-loop equivalent of PI3K (hereafter P-loop; p110 $\alpha$  residues 772–778) and the activation loop (p110 $\alpha$  residues 935–958), with the 3'-hydroxyl group positioned for phosphate transfer near the  $\gamma$ -phosphate of ATP [13]. Computational studies have postulated a role for p110 $\alpha$  K776 within the P-loop in determining phosphoinositide substrate specificity [15]. Class II and III PI3Ks, which cannot phosphorylate phosphatidylinositol-4-5-bisphosphate (PIP<sub>2</sub>), do not have an analogous positively charged residue at this position.

The most common heterodimer construct used for PI3K $\alpha$  structure determination is full length p110 $\alpha$  with a truncated p85 $\alpha$  consisting of the nSH2 and the iSH2 domains (p85 $\alpha$  residues 322–600), hereafter referred to as p110 $\alpha$ /niSH2 (Fig. 1A) [11,17]. The first insights into

the mechanism of nSH2 domain-mediated auto-inhibition were gained from the structures of the p110 oncogenic mutant, H1047R [10]. These structures revealed that phosphotyrosine peptides bind at the interface of the helical and nSH2 domains, competing with the interaction of the nSH2 with the kinase domain and leading to release of its inhibition [10,18]. However, precisely what effect this has on the conformation of the kinase domain, if any, and how this release results in activation is still not known. Recently, two groups have reported structures that include density for the nSH2 domain, the first within a p110 $\alpha$ -p85niSH2 fusion construct (PDB IDs 4L1B, 4L23, 4L2Y), and the other within a double Ras-binding domain (RBD) mutant (M232K, L233K) (PDB ID 4JPS)[19,20]. However, the structure of the nSH2 domain in complex with wild-type p110 $\alpha$  has not yet been determined.

Herein we report the structure of wild-type p110 $\alpha$ /niSH2, free and in complex with the truncated lipid substrate mimetic, di-C<sub>4</sub>-phosphatidylinositol-4,5-bisphosphate (diC4-PIP<sub>2</sub>), refined to 2.9 and 3.4 Å, respectively (Table 1). The structures provide insights into lipid-binding and catalysis by PI3K $\alpha$ . In addition, differences in the interaction of the nSH2 with the kinase domains between the wild-type enzyme and oncogenic mutant H1047R suggest a possible mechanism for the inactivation of the enzyme by the nSH2 domain and for its release.

## RESULTS

### The lipid-substrate binding site of PI3K $\alpha$

The substrate mimetic diC4-PIP<sub>2</sub> sits in a positively-charged crevice bordering the ATP binding site, located in a groove delimited by the P-loop (p110 $\alpha$  residues 772–778) and the activation loop of the kinase domain (p110 $\alpha$  residues 935–958), and the iSH2 helix,  $\alpha$ 1 (p85 $\alpha$  residues 457–465) (Fig. 1B, 1C). The 3' hydroxyl group is oriented towards the ATP-binding site, with no direct interactions with the protein. While the electron density for diC4-PIP<sub>2</sub> is clear (Fig. S1A), only limited density is present for the side chains of the activation loop residues. The 4-phosphate group extends towards the activation loop, forming either a direct salt-bridge or water-mediated hydrogen bond with K941. One oxygen atom from the 5-phosphate group interacts with the iSH2 domain, forming a salt bridge with R461. The 1-phosphate group faces the P-loop and is located 4.8 Å from K776. Although this distance is too long for a direct interaction, the two groups probably form a water-mediated hydrogen bond. Both short hydrophobic tails of the diC4-PIP<sub>2</sub> are positioned in a way that would allow membrane insertion of the endogenous PIP<sub>2</sub> substrate's long hydrophobic tails. Although neither K776 or K941 are present at the

**Table 1: Data collection and refinement statistics.**

	p110 $\alpha$ /niSH2	p110 $\alpha$ /niSH2 + diC4-PIP <sub>2</sub>
Data collection		
Space group	p2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	p2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions		
a, b, c (Å)	114.7, 116.2, 149.1	114.3, 116.1, 148.7
Resolution (Å)	50.00-2.96(3.01-2.96)	50.00-3.36(3.42-3.36)
R <sub>sym</sub>	0.068 (0.69)	0.103 (0.71)
I / $\sigma$ I	3.6 (3.2)	3.5 (4.4)
Completeness (%)	99.6 (100)	99.9 (100)
Redundancy	7.2 (7.1)	9.0 (9.2)
Unique reflections	41,914	28,636
Total reflections	300,309	256,465
X-ray source		
wavelength	0.9788 Å	0.9788 Å
Refinement		
Resolution (Å)	37.79-2.96	37.45-3.5
No. reflections	39,719	27,082
R <sub>work</sub> / R <sub>free</sub>	0.19/0.27	0.24/0.33
No. atoms		
Protein	10,830	10,584
Ligand	-	83
Water	18	5
B-factors		
Protein	53.82	65.81
Ligand	-	90.00
Water	21.22	78.00
R.m.s. deviations		
Bond lengths (Å)	0.011	0.015
Bond angles (°)	1.7	1.9

equivalent positions in other Class I isoforms, there are lysine residues at adjacent positions in the other isoforms. This suggests that all Class I isoforms could interact with PIP<sub>2</sub> in a similar way (Fig. 1D). The binding orientation of diC4-PIP<sub>2</sub> in the crystal structure is compatible with biochemical data implicating the activation loop in determining substrate specificity [15]. The positively charged residues present in the activation loops of the Class I isoforms are missing in the Class II and III PI3K isoforms. This explains why substrates such as PIP<sub>2</sub>, containing 4- and 5-phosphates, would bind with a much lower affinity to the latter isoforms.

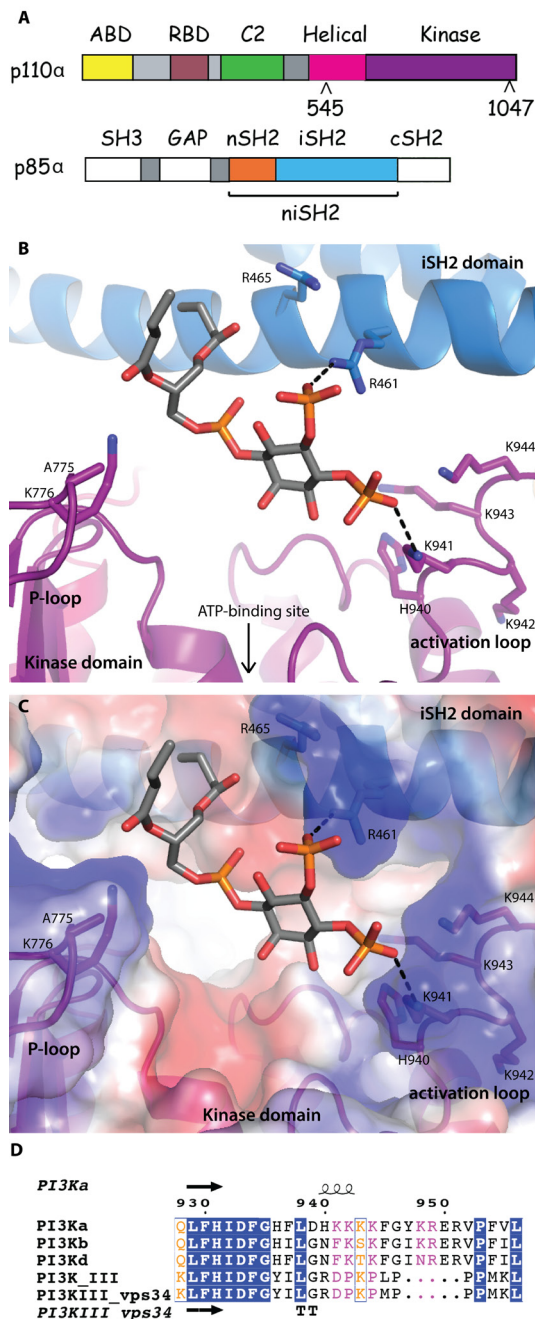
Clear electron density for the activation loop was present in the crystal structure of PI3K $\beta$  (PDB ID 2Y3A) and very recently, in some structures of PI3K $\alpha$  (PDB IDs 4A55, 4JPS, 4L1B, 4L23, 4L2Y) [11,19–21]. In the structure of wild-type PI3K $\alpha$  in complex with PIK-108 (PDB ID 4A55), the conformation of the activation loop was influenced by the inhibitor binding to a non-ATP binding site [11]. The conformations of the activation loop in the other PI3K $\alpha$  structures, including the two reported

here and in ref. 19 and 20, are very similar (Fig. S2A). It is nestled at the interface of the kinase and C2 domains of p110 $\alpha$ , with the iSH2 and nSH2 domains of p85 $\alpha$ . In protein kinases, activation loop phosphorylation results in activation of the kinase activity of the enzyme [22,23]. In PI3Ks, however, a similar phosphorylation event has not been shown to occur.

### Catalysis and nSH2 domain inhibition

The position of the bound ATP was modeled using a structural alignment of the PIP<sub>2</sub> bound structure (PDB ID 4OVV) with the p110 $\gamma$  crystal structure in complex with ATP (PDB ID 1E8X) (Fig. 2A) [13]. In this model, the 3'-hydroxyl group of PIP<sub>2</sub> is oriented toward the ATP  $\gamma$ -phosphate, in a position ideally suited for phosphate transfer. As in the p110 $\gamma$ -ATP complex structure, D915, the residue analogous to the catalytic base in protein kinases, is too far away to deprotonate the 3'-hydroxyl (10.9 Å), suggesting that in PI3Ks another residue may function as the base [13]. In the PI3K $\alpha$  structure, there are two





**Fig 1: Lipid-binding in PI3K.** (A) Domain organization of the PI3K $\alpha$  subunits, p110 $\alpha$  and p85 $\alpha$ , colored by domain. The positions of p110 $\alpha$  hotspot oncogenic mutations are indicated. The niSH2 p85 $\alpha$  construct used for crystallography is highlighted. (B) The substrate mimetic diC4-PIP<sub>2</sub> binding site. This site is adjacent to the ATP binding site, between the activation loop and the P-loop of the kinase domain (shown in purple) and the iSH2 domain (shown in blue). (C) PI3K $\alpha$  surface colored according to electrostatic potential, highlighting the positively charged PIP<sub>2</sub> binding-site. (D) Activation loop sequence alignment, performed between the Class IA, II and III PI3Ks. Blue represents identical residues, orange represents similar residues, pink represents differences that have been identified as being important for PIP<sub>2</sub> recognition and binding.

histidine residues in proximity to the 3'-hydroxyl, H917 (9.6 Å from the 3'-hydroxyl; equivalent to H948 in p110 $\gamma$ , proposed as an alternative base [13]) and H936 (7.4 Å from the 3'-hydroxyl). Both histidines are conserved in all Class I PI3K isoforms. Neither of these residues are close enough to deprotonate the hydroxyl. H936, however, is located on the activation loop, which is known to be flexible, and may undergo a conformational change that brings H936 into position to deprotonate the 3'-hydroxyl.

Interestingly, a salt-bridge formed between K948 of the activation loop and E342 in the nSH2 domain, suggests a mechanism through which the nSH2 domain inhibits catalytic activity (Fig. 2B): it presumably keeps the activation loop in an inactive conformation. Binding of phosphotyrosine-containing activators at the helical-nSH2 interface dislodges the nSH2 domain from its association with p110, disrupting the contacts with the activation loop and allowing it to adopt an active conformation in closer proximity to the lipid substrate.

The third helix of the iSH2 domain (p85 $\alpha$  residues 587-598),  $\alpha$ 3, forms an additional interface with the activation loop (Fig. 2C). This interface is mediated by hydrophobic interactions between L598 (p85 $\alpha$ ) and F945 (p110 $\alpha$ ), and a hydrogen bond between Q591 (p85 $\alpha$ ) and K948 (p110 $\alpha$ ). Deletions ( $\Delta$ 583-602) or truncations (p85-572<sup>STOP</sup>) of this section of the iSH2 domain are known to be oncogenic [24]. Previous studies have suggested that these iSH2 mutations activate the enzyme via a disruption of the iSH2-C2 interface [24,25]. Hydrogen and deuterium exchange mass spectrometry experiments have suggested that the disruption of this interface is a normal step in the PI3K $\alpha$  catalytic cycle, and may occur upon membrane binding [12]. These oncogenic deletions appear to function by mimicking this activation step. Our structures (PDB IDs 4OVU, 4OVV) suggest an additional mechanism through which these oncogenic mutations might activate the enzyme: in addition to disrupting the iSH2-C2 interface, the deletion of the iSH2  $\alpha$ 3 helix would release an inhibitory interaction between p85 and the activation loop, weakening the nSH2 domain mediated inhibition of kinase activity (Fig. 2D).

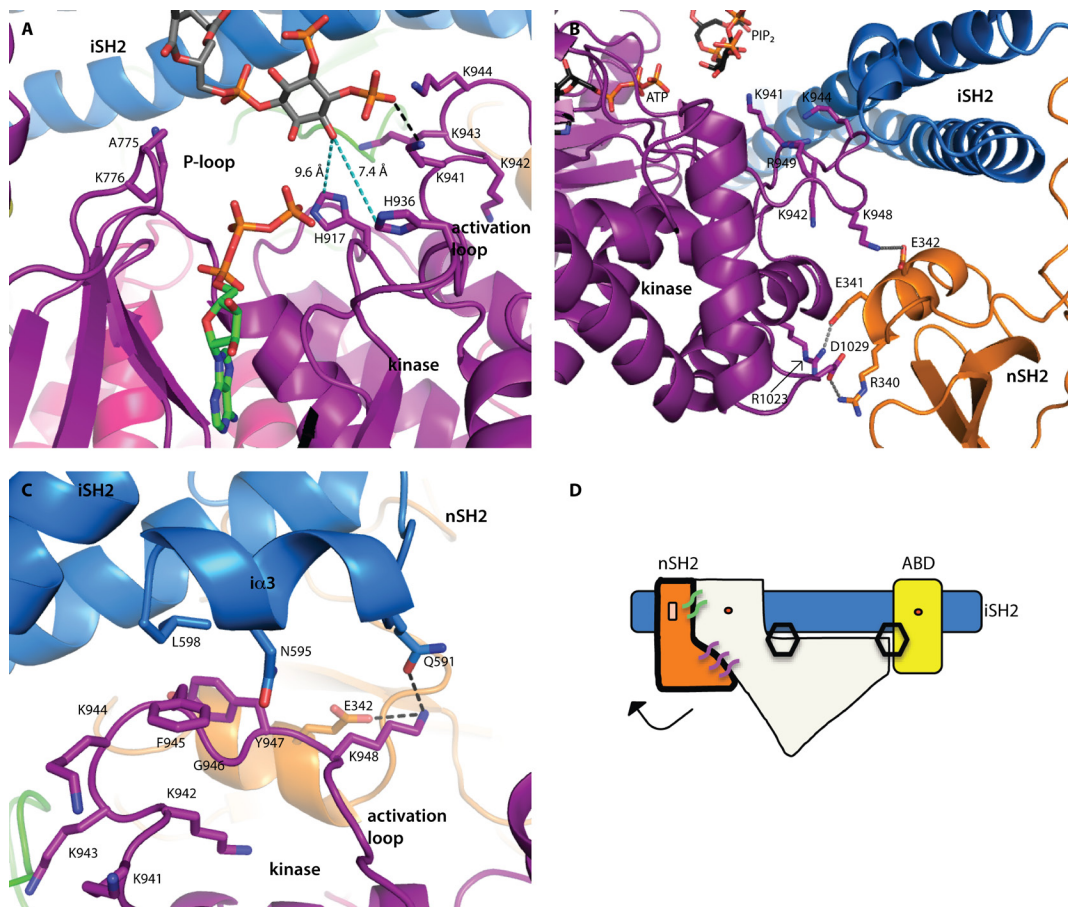
### Interactions of wild-type p110 $\alpha$ with the nSH2 domain

Comparison of the structure of the apo wild-type p110 $\alpha$  in complex with p85 $\alpha$ -niSH2 (reported here) and that of the H1047R mutant reveal key differences in the interactions between p110 $\alpha$  and nSH2 [10]. Many of these differences are similar to those between the wild-type and mutant reported by Mandelker *et al.* and will not be discussed further [10]. However, some important differences in the nSH2 and iSH2 domains and their interactions with p110 $\alpha$  were not evident in previous studies.

One such difference is a shift of up to 5.5 Å (measured at p85α 450, Fig. S3) in the position of the Ca atoms of the N-terminal helix at the end of the coiled coil iSH2 domain (p85α residues 443-475)(Fig. 3A). The change is less pronounced in the second helix, with an average movement of 2.5 Å (measured at p85α residues 565-579; maximum distance is 3.9 Å at p85α 577, Fig. S3). The positively charged residues on one face of the iSH2 domain, along with two key loops in the kinase domain (p110α residues 723-729 and 863-867) are thought to play a major role in mediating the interaction between PI3Kα and the cell membrane [17]. The iSH2 helices in the oncogenic mutant structure appear to be bent towards the membrane to a greater degree than those in the wild-type, which would be consistent with the increased

membrane binding of the oncogenic mutant [10,11]. These changes may act synergistically with the changes in the loops identified previously to further enhance membrane binding.

Striking differences are also observed in the interaction of the nSH2 domain with p110 (Fig. 3B). In the oncogenic mutant, the nSH2 domain interacts with the C2, helical and kinase domains of p110α, influencing the conformation of these adjacent domains [10,26]. In the wild-type structure, the nSH2 is rotated 14° towards the kinase and C2 domains. This change results in an increased buried surface area between the nSH2 domain and p110α in the wild-type protein compared to the H1047R mutant (calculated with PISA for the wild-type, 1083 Å<sup>2</sup>, and mutant, 820 Å<sup>2</sup>, considering only the



**Fig 2: Structural insights into catalysis.** All domains are colored according to the scheme in Fig. 1A. (A) The relationship between the two substrates was inferred by modeling a molecule of ATP into the binding site (from the alignment with the p110γ-ATP complex structure, PDB ID 1E8X). The 3'-hydroxyl is oriented toward the ATP γ-phosphate. There are two histidine residues in the binding site, which may deprotonate the 3'-hydroxyl for catalysis. Distances are shown in cyan colored dashed lines. (B) The nSH2 domain locks the activation loop in an inactive conformation via a salt-bridge between K948 (p110) and E342 (p85α). (C) The C-terminal residues of the iSH2 (p85α residues 587-602) form a short helix (iα3) which forms an interface with the activation loop. A hydrophobic stacking interaction is made between F945 (p110α) and L598 (p85α). Two key interactions between p85α (E342 and N591) are made with K948 of the activation loop, locking it in an inactive conformation. (D) Schematic representation of the p110α/nSH2 heterodimer showing the general position of the hydrogen bond network that locks p110α in an inactive conformation. In this scheme, p110α is represented in white. The two hexagons represent the PIP2 binding sites. The binding of phosphotyrosine residues at the helical-nSH2 interface causes a conformational change that breaks interactions with the activation loop, thereby activating the enzyme. In the basal state, this interface is maintained by key hydrogen bonds or salt-bridges between the subunits, represented by the purple and green lines.



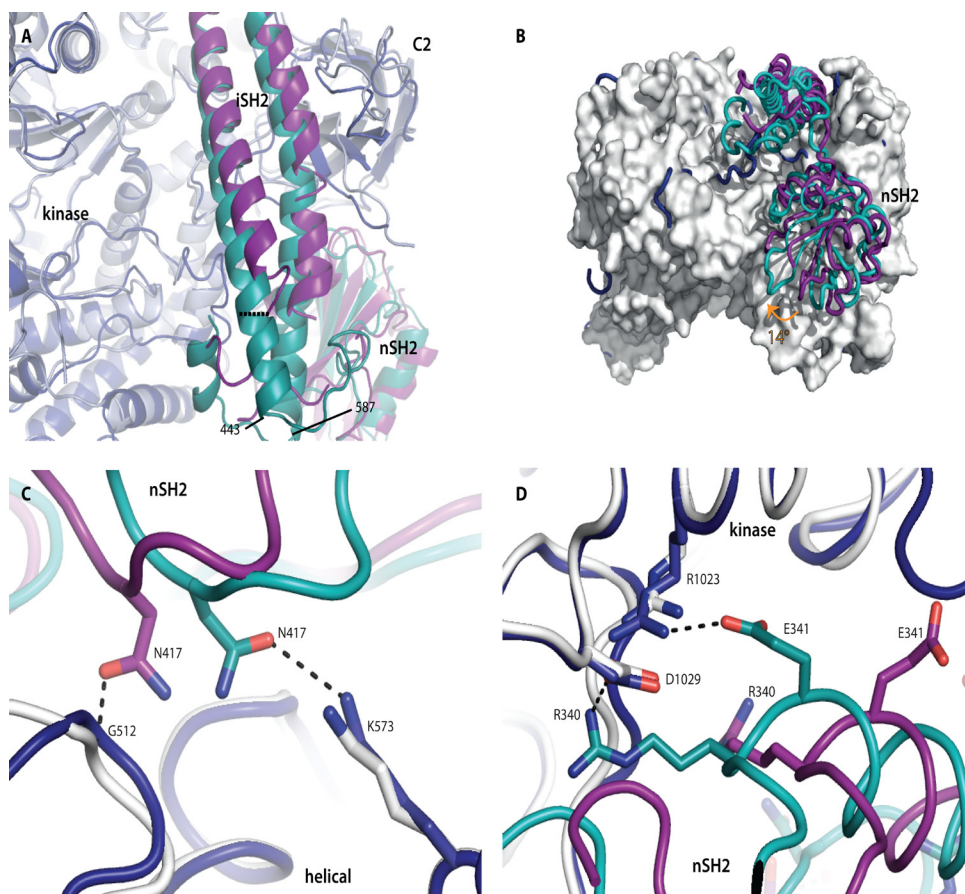
residues present in both structures) [27]. This suggests that the inhibition of the nSH2 domain in the oncogenic mutant may be less potent than in the wild-type protein. The best explanation of these observations can be summarized as follows: the p110 domains differ in their interactions with their nSH2 domains in such a way that the wild-type structure shows extra interactions that result in a tighter contact and consequently greater inhibition (Table S4).

In the helical domain of the wild-type structure, residues p110 $\alpha$  K573 and p85 $\alpha$  N417 form a hydrogen bond that is not present in the H1047 mutant, because p85 $\alpha$  N417 interacts instead with the backbone of p110 $\alpha$  G512 (Fig. 3C). A greater difference is observed at the interface between the C2 domain and nSH2, where a number of unique hydrogen bonds are made in each structure. The wild-type nSH2 domain makes three unique hydrogen bonds with the C2 domain, while the oncogenic

mutant makes five unique hydrogen bonds with these domains, reflecting the shift in nSH2 conformation (Table S4). Perhaps the most striking difference is the two salt bridges between the kinase domain (R1023 and D1029) and nSH2 (E341 and R340) that are only observed in the wild-type protein (Fig. 3D). These same kinase-nSH2 domain interactions are also present in the higher resolution structures published by Zhao *et al.* [20]. The lack of these key salt-bridges in the oncogenic mutant may reduce the nSH2 domain auto-inhibition in this mutant, resulting in higher lipid kinase activity.

### A second lipid binding site in PI3K $\alpha$

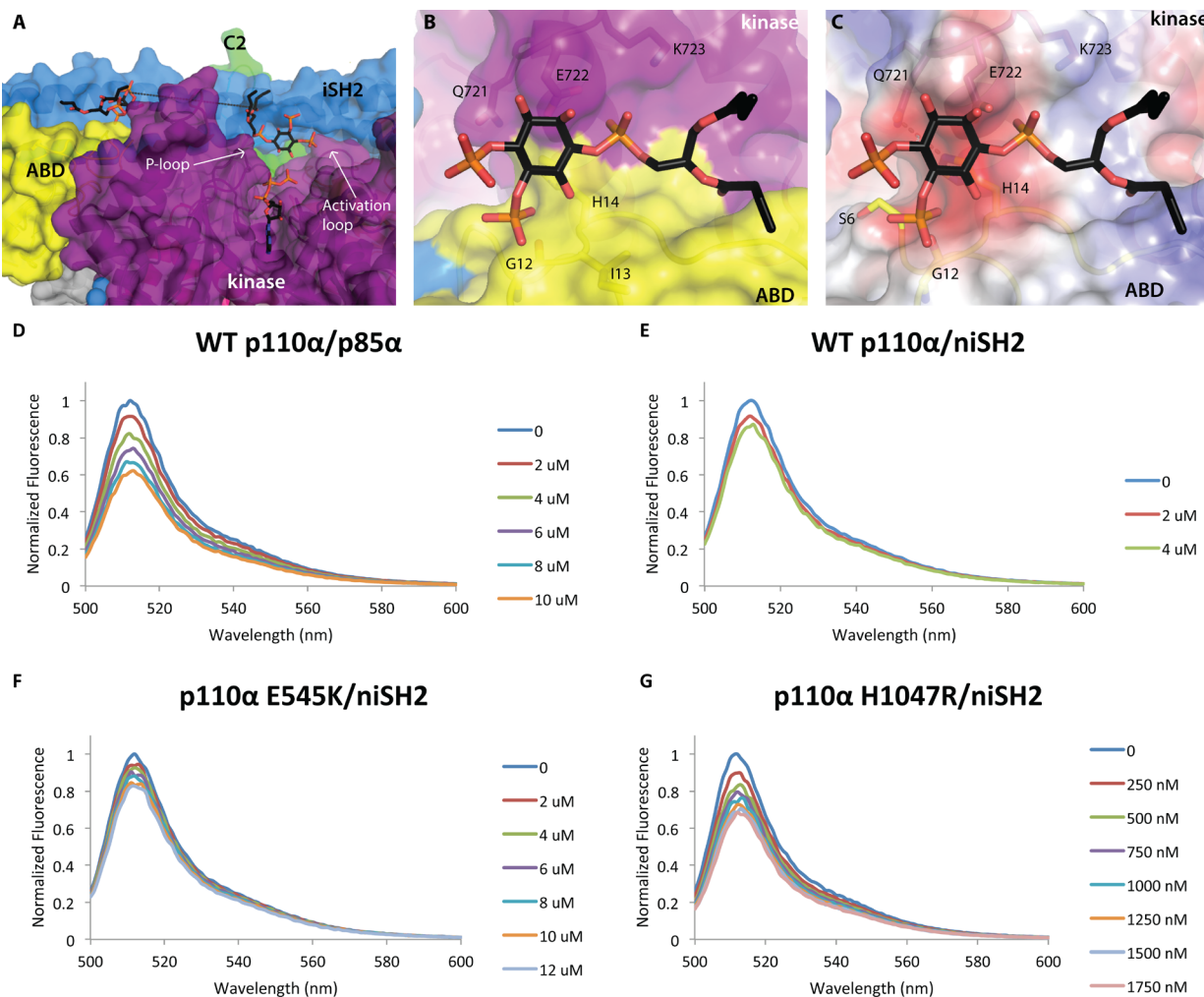
Unexpectedly, electron density for a second diC4-PIP<sub>2</sub> molecule was observed in the structure of the p110 $\alpha$ /niSH2 in complex with the lipid mimetic (Fig. S1B). This



**Fig 3: Wild-type p110 $\alpha$  has more interactions with the nSH2 domain than the oncogenic mutant H1047R.** Superposition of the wild-type crystal structure p110 $\alpha$ /niSH2 (PDB ID 4OVU) with the p110 $\alpha$ H1047R/niSH2 oncogenic mutant (PDB ID 3HHM), obtained by aligning the two p110 $\alpha$  molecules. The wild-type p110 and p85 $\alpha$  are shown in dark blue and teal, respectively, while the p110 $\alpha$ H1047R/niSH2 mutant structure is displayed as light grey (p110 $\alpha$ ) and purple (p85 $\alpha$ ). (A) The largest difference between the iSH2 domains is highlighted with a dashed line, measured between the C $\alpha$  atoms of p85 $\alpha$  450 in each structure. (B) The 14° rotation of the nSH2 domain is identified with an orange arrow. p110 $\alpha$ H1047R is shown as a surface representation. (C) Differences in the interactions between the helical and nSH2 domains of the wild-type and oncogenic mutant structures. (D) Two key salt-bridges between the kinase domain and nSH2 are present in the wild-type but absent in the oncogenic mutant structure.

second diC4-PIP<sub>2</sub> molecule binds in a groove between the ABD, kinase and iSH2 domains (Fig. 4A,B). One of the 4'-phosphate oxygens is within hydrogen bonding distance of the backbone of G12 and E722 is 3.6 Å from the 1-phosphate. As with the catalytic PIP<sub>2</sub> molecule, the truncated C4-hydrophobic tails are well positioned for membrane binding. The calculated electrostatic potential surface shows very few positive charges, possibly suggesting this may be a general lipid-binding site rather than a specific PIP<sub>2</sub> binding site (Fig. 4C). Its function may contribute to anchoring PI3K to the cell membrane, but further work is necessary to validate this hypothesis.

To determine whether PI3K $\alpha$  could indeed bind an additional PIP<sub>2</sub> molecule within the context of the lipid bilayer of a phospholipid membrane, we measured the ability of PI3K $\alpha$  to bind and cluster BODIPY® FL-PIP<sub>2</sub> embedded within phospholipid vesicles. Gambhir *et al.* reported that at a concentration of 0.1% PIP<sub>2</sub> in 100 nm diameter vesicles, the distance between PIP<sub>2</sub> molecules is ~300 Å [28]. Self-quenching of the BODIPY® fluorescence occurs when PIP<sub>2</sub> molecules are brought within 50-60 Å [29]. Therefore if PI3K binds and clusters two or more molecules of BODIPY®-FL-PIP<sub>2</sub>, fluorescence quenching should be observed. A similar



**Fig 4: Two PIP<sub>2</sub> molecules bound to p110 $\alpha$ /niSH2.** p110 $\alpha$ /niSH2 in complex with diC4-PIP<sub>2</sub> is shown as a molecular surface with the kinase domain colored in purple, ABD domain in yellow, helical domain in pink, C2 in green, iSH2 in blue and nSH2 in orange. PIP<sub>2</sub> molecules are shown as sticks with grey carbons. (A) Two molecules of PIP<sub>2</sub> bind at the interface between p110 $\alpha$  and iSH2 of p85 $\alpha$ . The distance between the two binding sites is ~21 Å. (B) A second molecule of PIP<sub>2</sub> binds at the interface between the ABD (yellow) and kinase (purple) domains. (C) The surface of the second PIP<sub>2</sub> binding site colored according to the electrostatic potential shows a very hydrophobic surface, suggesting possibly a general lipid-binding site rather than a specific PIP<sub>2</sub> binding site. (Fig. 4D-G) PI3K clusters PIP<sub>2</sub> in model membrane vesicles containing 50 nM of BODIPY®-FL-PIP<sub>2</sub>. The highest normalized emission intensity corresponds to the vesicles alone. Each subsequent spectrum represents an incremental addition of the corresponding protein. All experiments were performed with N=3. Graphs shown are representative and present the data from a single experiment. (D) Wild-type p110 $\alpha$ /p85 $\alpha$  quenches 20% of the signal at 4  $\mu$ M. (E) Wild-type p110 $\alpha$ /niSH2 quenches in a similar fashion to the full-length complex. (F) The displayed quenching by p110 $\alpha$ E545K/niSH2 is similar to wild-type p110 $\alpha$ /niSH2. (G) p110 $\alpha$ H1047R/niSH2 quenches the signal with a much higher potency than wild-type. Only 500 nM of protein is required to quench the signal by 20%.



approach has been used to demonstrate clustering of BODIPY®-FL-PIP<sub>2</sub> by MARCKS, NAP-22 and myelin basic protein (MBP) as well as clustering of BODIPY®-TMR-PIP<sub>2</sub> by dynamin [28,30–32].

BODIPY®-PIP<sub>2</sub> fluorescence quenching was measured using unilamellar vesicles with a diameter of 100 nm. The addition of wild-type, full-length PI3K (p110 $\alpha$ /p85 $\alpha$ ) to vesicles containing BODIPY®-FL-PIP<sub>2</sub> resulted in quenching of the fluorescence signal in a dose dependent manner suggesting PI3K $\alpha$  can bind and cluster two or more molecules of PIP<sub>2</sub> (Fig. 4D). Titration of vesicles with p110 $\alpha$ /niSH2 (Fig. 4E), a complex devoid of the RhoGap, SH3 and cSH2 domains of p85 $\alpha$ , demonstrated fluorescence quenching to the same degree as full length p110 $\alpha$ /p85 $\alpha$ . There was relatively little quenching by p85 $\alpha$  alone in the absence of p110 $\alpha$  (Fig. S5). Taken together, these data suggest that there is an additional binding site for PIP<sub>2</sub> on p110 $\alpha$  or at the interface of p110 $\alpha$  and p85 $\alpha$  subunits of PI3K $\alpha$ .

We also examined the effects of two oncogenic mutations on the ability of PI3K to bind two PIP<sub>2</sub> molecules. In similar experiments, p110 $\alpha$ E545K/niSH2 showed comparable fluorescence quenching to wild-type p110 $\alpha$ /niSH2, suggesting that this mutation does not affect PIP<sub>2</sub> binding (Fig. 4F). Interestingly, p110 $\alpha$ H1047R/niSH2, demonstrated greatly increased quenching compared to wild-type p110 $\alpha$ /niSH2 (Fig. 4G). Based on these results, along with previous data and inferences discussed in this work, it seems likely that this effect is due to increased membrane binding of the H1047R mutant.

## DISCUSSION

Binding of phosphotyrosine-containing effectors at the nSH2-helical interface activate PI3K $\alpha$  catalysis [8,10]. The structures of wild-type p110 $\alpha$ /niSH2 alone and in complex with diC4-PIP<sub>2</sub> reported here provide insight into how binding of effectors may be communicated to the kinase domain. We have identified key interactions between the nSH2 domain and the activation loop, along with an iSH2-activation loop interface through which the phosphotyrosine binding event may affect the activation state of the kinase domain and increase kinase activity.

In PI3K, lipid substrate specificity is determined by key positive residues on both the activation loop and the iSH2 domain which recognize the 4- and 5-phosphate groups of the lipid substrate, PIP<sub>2</sub>. The importance of these regions has previously been demonstrated biochemically [15,16]. H936, located on the activation loop, may act as a base to deprotonate the 3'-hydroxyl group as part of the phosphoryl transfer. Binding of a phosphotyrosine peptide at the helical-nSH2 interface would result in the dislocation of the nSH2 domain, releasing the activation loop to close in on the substrate and repositioning H936 to deprotonate the 3'-hydroxyl.

The structure also shows, surprisingly, that PI3K $\alpha$

binds an additional PIP<sub>2</sub> molecule. This observation was confirmed by fluorescence quenching experiments. Further studies are required to ascertain the functional relevance of this additional binding site. Regardless, the identification of multiple lipid binding sites provides additional targets that may enable more selective inhibition among the various isoforms, or even between mutants and wild-type PI3K $\alpha$ .

## METHODS

### Protein expression for crystallization

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 x 10<sup>6</sup> cells per milliliter, cells were infected with WT p110 $\alpha$  (or p110 $\alpha$ H1047R or p110 $\alpha$ E545K) and p85 $\alpha$ -niSH2 (or p85 $\alpha$ ) viruses at a multiplicity of infection ratio of 3:2. Media was supplemented with a PI3K inhibitor as described in Mandelker *et al.* [10]. Cells were harvested 72 hours after infection and the cell pellet collected through centrifugation at 900 x g. Protein purification was performed as previously described [10,17].

### p85 $\alpha$ Protein expression and purification

Full-length p85 $\alpha$  was expressed heterogeneously in *Escherichia coli* as described previously. Briefly, cells transformed with the pGEX 4T plasmid containing an N-terminal glutathione S-transferase (GST) fusion-p85 were grown at 37°C in LB medium. Expression was induced with 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG). Following 4 hours of induction at 37°C, cells were pelleted and stored at -80°C. Cell pellets were resuspended in PBS, 2 mM DTT, Roche complete EDTA-free protease inhibitor cocktail (Roche Diagnostics GmbH, Mannheim, Germany), pH 7.4 and lysed using a microfluidizer (Microfluidics, Newton, MA). Clarified lysate was incubated with glutathione sepharose HP resin (GE Healthcare) in binding buffer (PBS, 2 mM DTT, pH 7.4), at 4°C for 2 hours with gentle agitation. The GST-tagged p85 was eluted with 10 volumes of elution buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, and 10 mM reduced glutathione, pH 8.0). After thrombin cleavage, the p85 was purified by anion exchange Resource-Q anion exchange column (GE Healthcare). p85 was eluted using a linear gradient of 0 – 100% anion exchange buffer B (50 mM Tris, 500 mM NaCl, pH 8.0) over 60 column volumes. p85 of  $\geq$  90% homogeneity, was loaded onto a HiLoad 26/60 Superdex 200 prep grade gel filtration column (GE Healthcare) equilibrated in gel filtration buffer (50 mM Tris, 300 mM NaCl, pH 8.5); fractions containing p85 of  $\geq$  95% homogeneity, as determined by SDS-PAGE, were

pooled. p85 was concentrated to 15 mg/mL.

## Crystallization and data collection

Crystallization was performed as previously described and improved with successive rounds of macroseeding [17]. Crystals of p110 $\alpha$ /nSH2 were soaked for one hour with 1 mM diC4-phosphatidylinositol-4,5-bisphosphate (Echelon Biosciences). X-ray diffraction data were collected at beamlines X6A and X25 of NSLS at Brookhaven National Laboratory. The crystals diffracted to a resolution of 2.96 Å in the absence of diC4-PIP<sub>2</sub> and 3.37 Å in the presence of diC4-PIP<sub>2</sub>. Data were processed with HKL2000 (Table 1)[33].

## Structure determination and analysis

The free structure was determined by using the coordinates of the previously determined WT p110 $\alpha$ /nSH2 (PDB ID 2RD0) [17] as a model. After rigid body and positional refinement, the program Coot [34] was used for model building. Initial refinement revealed that the nSH2 domain of p85 $\alpha$  was present and ordered in the structure. Therefore, the nSH2 domain of p110 $\alpha$ H1047R/nSH2 structure (PDB ID 3HHM) [10] and the p85 $\alpha$  nSH2 crystal structure (PDB ID 2IUG) [35] were used as a guide to fit the additional electron density in this region. Iterative rounds of refinement using REFMAC 5.0 [36–38] yielded a final  $R_{work}$  of 0.191 and an  $R_{free}$  of 0.271 to 2.95 Å in the absence of diC4-PIP<sub>2</sub>. The refined coordinates of the free wild-type structure were used as an initial model for the determination of the structure in the presence of diC4-PIP<sub>2</sub>; refinement yielded an  $R_{cryst}$  of 0.23 and an  $R_{free}$  of 0.33. The overall quality of the final model was assessed by using the programs PROCHECK [39] and WHATIF [40]. Visualization, analysis and figure preparation were carried out with PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.1 Schrödinger, LLC). Sequence alignments were done using Clustal Omega and colored with EsPRIPT [41]. The calculation of the buried surface was done with the program PISA [27].

## Preparation of BODIPY-FL-PIP<sub>2</sub> vesicles

Unilamellar vesicles containing PC/PE/PS/BODIPY-FL-PIP<sub>2</sub>/cholesterol with a molar ratio of 54.9/25/5/0.1/15 were generated as follows: Lipids were dried under a stream of nitrogen in a glass test tube then held under vacuum for 24 hours to completely remove organic solvents. Lipids were then heated in the presence of assay buffer (50 mM Tris pH 7.6, 150 mM NaCl, 1 mM EDTA, 2 mM DTT) at 50°C for 10 mins. Lipids were resuspended by vortexing then extruded through a polycarbonate filter with 100 nm pore sizes to yield

unilamellar vesicles 100 nm in diameter. BODIPY®-FL-PIP<sub>2</sub> was purchased from Echelon Biosciences. All other lipids were purchased from Avanti Polar Lipids.

## Fluorescence quenching experiments

BODIPY®-FL-PIP<sub>2</sub> containing vesicles were diluted to a final concentration of 50  $\mu$ M lipid (50 nM BODIPY®-FL-PIP<sub>2</sub>) in assay buffer (50 mM Tris pH 7.6, 150 mM NaCl, 1 mM EDTA, 2 mM DTT) for all assays. Ovalbumin (0.5 mg/mL) was added to the assay solution to minimize signal loss due to vesicle binding to glass surfaces during mixing and binding measurements. Fluorescence measurements were taken on a FluoroLog fluorometer from HORIBA scientific. BODIPY®-FL-PIP<sub>2</sub> was excited at 490 nm and emission readings were recorded from 500 to 600 nm. The spectral bandwidths for excitation and emission were 2 and 5 nm respectively. All binding experiments took place at 25°C in a quartz cuvette. Quenching signals were normalized to the fluorescence signal from the vesicle solution alone and adjusted for dilution due to the addition of the corresponding protein.

## ACKNOWLEDGMENTS

The authors would like to thank A.M. Silva for helpful discussions. We acknowledge the use of the Johns Hopkins University School of Medicine Mass Spectrometry and Proteomics Core, Hopkins Digestive Diseases Basic and Translational Research Core Center (NIDDK center grant P30 DK089502). Funding: This work was supported by the Virginia and D.K. Ludwig Fund for Cancer Research and NIH grant CA 43460. Data collection was carried out at beamline X6A/X25, funded by the National Institute of General Medical Sciences, NIH under agreement GM-0080. The NSLS, Brookhaven National Laboratory is supported by the US Department of Energy under contract no. DE AC02-98CH10886. S.B.G. is a Stewart Trust Fellow. M.S.M. is a recipient of an Australian Postgraduate Award (APA), a Cooperative Research Centre for Cancer Therapeutics top-up scholarship and a Monash University Postgraduate Publications Award.

## Conflict of Interest Statement

Under agreements between The Johns Hopkins University and various commercial entities, K.W.K. and B.V. are entitled to a share of the royalties received by The Johns Hopkins University on sales of products related to the diagnosis of mutant PIK3CA genes. The terms of these arrangements are being managed by The Johns Hopkins University in accordance with its conflict of interest policies.

## Data deposition footnote

p110 $\alpha$ /nSH2 heterodimer structures, both free and in complex with diC4-PIP<sub>2</sub> have been deposited in the Protein Data Bank (PDB IDs 4OVU and 4OVV).

## REFERENCES

1. Fyffe C, Buus R, Falasca M. Genetic and Epigenetic Regulation of Phosphoinositide 3-kinase Isoforms. *Curr Pharm Des* 2013; 19: 680–6.
2. Broderick DK, Di C, Parrett TJ, Samuels YR, Cummins JM, McLendon RE, Fults DW, Velculescu VE, Bigner DD, Yan H. Mutations of PIK3CA in Anaplastic Oligodendrogliomas, High-Grade Astrocytomas, and Medulloblastomas. *Cancer Res* 2004; 64: 5048–50.
3. Campbell IG, Russell SE, Choong DYH, Montgomery KG, Ciavarella ML, Hooi CSF, Cristiano BE, Pearson RB, Phillips WA. Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res* 2004; 64: 7678–81.
4. Levine DA, Bogomolny F, Yee CJ, Lash A, Barakat RR, Borgen PI, Boyd J. Frequent mutation of the PIK3CA gene in ovarian and breast cancers. *Clin Cancer Res* 2005; 11: 2875–8.
5. Bachman KE, Argani P, Samuels Y, Silliman N, Ptak J, Szabo S, Konishi H, Karakas B, Blair BG, Lin C, Peters BA, Velculescu VE, Park BH. The PIK3CA gene is mutated with high frequency in human breast cancers. *Cancer Biol Ther* 2004; 3: 772–5.
6. Pedrero JMG, Carracedo DG, Pinto CM, Zapatero AH, Rodrigo JP, Nieto CS, Gonzalez MV. Frequent genetic and biochemical alterations of the PI 3-K/AKT/PTEN pathway in head and neck squamous cell carcinoma. *Int J Cancer* 2005; 114: 242–8.
7. Samuels Y, Waldman T. Oncogenic Mutations of PIK3CA in Human Cancers. *Curr Top Microbiol Immunol* 2011; 347: 21–41.
8. Carson JD, Van Aller G, Lehr R, Sinnamon RH, Kirkpatrick RB, Auger KR, Dhanak D, Copeland RA, Gontarek RR, Tummino PJ, Luo L. Effects of oncogenic p110 $\alpha$  subunit mutations on the lipid kinase activity of phosphoinositide 3-kinase. *Biochem J* 2008; 409: 519–24.
9. Chaussade C, Cho K, Mawson C, Rewcastle GW, Shepherd PR. Functional differences between two classes of oncogenic mutation in the PIK3CA gene. *Biochem Biophys Res Comm* 2009; 381: 577–81.
10. Mandelker D, Gabelli SB, Schmidt-Kittler O, Zhu J, Cheong I, Huang CH, Kinzler KW, Vogelstein B, Amzel LM. A frequent kinase domain mutation that changes the interaction between PI3K $\alpha$  and the membrane. *Proc Natl Acad Sci USA* 2009; 106: 16996–7001.
11. Hon W-C, Berndt A, Williams RL. Regulation of lipid binding underlies the activation mechanism of class IA PI3-kinases. *Oncogene* 2012; 31: 3655–66.
12. Burke JE, Perisic O, Masson GR, Vadas O, Williams RL. Oncogenic mutations mimic and enhance dynamic events in the natural activation of phosphoinositide 3-kinase p110 $\alpha$  (PIK3CA). *Proc Natl Acad Sci USA* 2012; 109: 15259–64.
13. Walker EH, Perisic O, Ried C, Stephens L, Williams RL. Structural insights into phosphoinositide 3-kinase catalysis and signalling. *Nature* 1999; 402: 313–20.
14. Walker E, Pacold ME, Perisic O, Stephens L, Hawkins PT, Wymann MP, Williams RL. Structural determinants of phosphoinositide 3-kinase inhibition by wortmannin, LY294002, Quercetin, Myricetin, and Staurosporine. *Mol Cell* 2000; 6: 909–19.
15. Pirola L, Zvelebil MJ, Bulgarelli-Leva G, Van Obberghen E, Waterfield MD, Wymann MP. Activation Loop Sequences Confer Substrate Specificity to Phosphoinositide 3-Kinase  $\alpha$  (PI3K $\alpha$ ): Functions of lipid kinase-deficient PI3K $\alpha$  in signaling. *J Biol Chem* 2001; 276: 21544–54.
16. Bondeva T, Pirola L, Bulgarelli-Leva G, Rubio I, Wetzker R, Wymann MP. Bifurcation of Lipid and Protein Kinase Signals of PI3K $\gamma$  to the Protein Kinases PKB and MAPK. *Science* 1998; 282: 293–6.
17. Huang CH, Mandelker D, Schmidt-Kittler O, Samuels Y, Velculescu VE, Kinzler KW, Vogelstein B, Gabelli SB, Amzel LM. The structure of a human p110 $\alpha$ /p85 $\alpha$  complex elucidates the effects of oncogenic PI3K $\alpha$  mutations. *Science* 2007; 318: 1744–8.
18. Nacht M, Qiao L, Sheets MP, St. Martin T, Labenski M, Mazdiyasi H, Karp R, Zhu Z, Chaturvedi P, Bhavsar D, Niu D, Westlin W, Petter RC, et al. Discovery of a Potent and Isoform-Selective Targeted Covalent Inhibitor of the Lipid Kinase PI3K $\alpha$ . *J Med Chem* 2013; 56: 712–21.
19. Furet P, Guagnano V, Fairhurst RA, Imbach-Weese P, Bruce I, Knapp M, Fritsch C, Blasco F, Blanz J, Aichholz R, Hamon J, Fabbro D, Caravatti G. Discovery of NVP-BYL719 a potent and selective phosphatidylinositol-3 kinase alpha inhibitor selected for clinical evaluation. *Bioorg Med Chem Lett* 2013; 23: 3741–8.
20. Zhao Y, Zhang X, Chen Y, Lu S, Peng Y, Wang X, Guo C, Zhou A, Zhang J, Luo Y, Shen Q, Ding J, Meng L, et al. Crystal Structures of PI3K $\alpha$  Complexed with PI103 and Its Derivatives: New Directions for Inhibitors Design. *ACS Med Chem Lett* 2014; 5: 138–42.
21. Zhang X, Vadas O, Perisic O, Anderson KE, Clark J, Hawkins PT, Stephens LR, Williams RL. Structure of Lipid Kinase p110 $\beta$ /p85 $\beta$  Elucidates an Unusual SH2-Domain-Mediated Inhibitory Mechanism. *Mol Cell* 2011; 41: 567–78.
22. Adams JA. Kinetic and Catalytic Mechanisms of Protein Kinases. *Chem Rev* 2001; 101: 2271–90.
23. Adams JA. Activation Loop Phosphorylation and Catalysis in Protein Kinases: Is There Functional Evidence for the Autoinhibitor Model?†. *Biochemistry* 2003; 42: 601–7.
24. Wu H, Shekar SC, Flinn RJ, El-Sibai M, Jaiswal BS, Sen KI, Janakiraman V, Seshagiri S, Gerfen GJ, Girvin

- ME, Backer JM. Regulation of Class IA PI 3-kinases: C2 domain-iSH2 domain contacts inhibit p85/p110 $\alpha$  and are disrupted in oncogenic p85 mutants. *Proc Natl Acad Sci USA* 2009; 106: 20258–63.
25. Jaiswal BS, Janakiraman V, Kljavin NM, Chaudhuri S, Stern HM, Wang W, Kan Z, Dbouk HA, Peters BA, Waring P, Dela Vega T, Kenski DM, Bowman KK, et al. Somatic Mutations in p85 $\alpha$  Promote Tumorigenesis through Class IA PI3K Activation. *Cancer Cell* 2009; 16: 463–74.
  26. Miled N, Yan Y, Hon W-C, Perisic O, Zvelebil MJ, Inbar Y, Schneidman-Duhovny D, Wolfson HJ, Backer JM, Williams RL. Mechanism of two classes of cancer mutations in the phosphoinositide 3-kinase catalytic subunit. *Science* 2007; 317: 239–42.
  27. Krissinel E, Henrick K. Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol* 2007; 372: 774–97.
  28. Gambhir A, Hangyás-Mihályiné G, Zaitseva I, Cafiso DS, Wang J, Murray D, Pentylala SN, Smith SO, McLaughlin S. Electrostatic sequestration of PIP2 on phospholipid membranes by basic/aromatic regions of proteins. *Biophys J* 2004; 86: 2188–207.
  29. Karolin J, Johansson LB-A, Strandberg L, Ny T. Fluorescence and Absorption Spectroscopic Properties of Dipyrrometheneboron Difluoride (BODIPY) Derivatives in Liquids, Lipid Membranes, and Proteins. *J Am Chem Soc* 1994; 116: 7801–6.
  30. Epand RM, Vuong P, Yip CM, Maekawa S, Epand RF. Cholesterol-dependent partitioning of PtdIns(4,5)P2 into membrane domains by the N-terminal fragment of NAP-22 (neuronal axonal myristoylated membrane protein of 22 kDa). *Biochem J* 2004; 379: 527–32.
  31. Musse AA, Gao W, Homchaudhuri L, Boggs JM, Harauz G. Myelin basic protein as a “PI(4,5)P2-modulin”: a new biological function for a major central nervous system protein. *Biochemistry* 2008; 47: 10372–82.
  32. Bethoney KA, King MC, Hinshaw JE, Ostap EM, Lemmon MA. A possible effector role for the pleckstrin homology (PH) domain of dynamin. *Proc Natl Acad Sci USA* 2009; 106: 13359–64.
  33. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997; 276: 307–26.
  34. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004; 60: 2126–32.
  35. Nolte RT, Eck MJ, Schlessinger J, Shoelson SE, Harrison SC. Crystal structure of the PI 3-kinase p85 amino-terminal SH2 domain and its phosphopeptide complexes. *Nat Struct Biol* 1996; 3: 364–74.
  36. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* 2004; 60: 2184–95.
  37. Murshudov GN, Vagin AA, Dodson EJ. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr D Biol Crystallogr* 1997; 53: 240–55.
  38. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 2011; 67: 235–42.
  39. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993; 26: 283–91.
  40. Hooft RW, Vriend G, Sander C, Abola EE. Errors in protein structures. *Nature* 1996; 381: 272.
  41. Gouet P, Courcelle E, Stuart DI, Métoz F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 1999; 15: 305–8.