6-12-2020

Phosphorylation-dependent substrate selectivity of protein kinase B (AKT1)

Nileeka Balasuriya
Western University

Norman E. Davey
The Institute of Cancer Research

Jared L. Johnson
Weill Cornell Medicine

Huadong Liu
Western University

Kyle K. Biggar
Western University

See next page for additional authors

JBC EDITORS' PICK

Phosphorylation-dependent substrate selectivity of protein kinase B (AKT1)

Received for publication, December 20, 2019, and in revised form, April 22, 2020 Published, Papers in Press, April 29, 2020, DOI 10.1074/jbc.RA119.012425

Nileeka Balasuriya[1], Norman E. Davey[3,‡], Jared L. Johnson[4,‡], Huadong Liu[1,5,‡], Kyle K. Biggar[1,6], Lewis C. Cantley[4], Shawn Shun-Cheng Li[1], and Patrick O'Donoghue[1,2,*]

From the Departments of [1]Biochemistry and [2]Chemistry, The University of Western Ontario, London, Ontario, Canada, [3]Division of Cancer Biology, The Institute of Cancer Research, London, United Kingdom, [4]Meyer Cancer Center, Department of Medicine, Weill Cornell Medical College, New York, New York, USA, [5]Center for Mitochondrial Biology and Medicine, Xi'an Jiaotong University, Xi'an, Shanxi, China, and [6]Department of Biology, Carleton University, Ottawa, Ontario, Canada

Edited by Alex Toker

Protein kinase B (AKT1) is a central node in a signaling pathway that regulates cell survival. The diverse pathways regulated by AKT1 are communicated in the cell via the phosphorylation of perhaps more than 100 cellular substrates. AKT1 is itself activated by phosphorylation at Thr-308 and Ser-473. Despite the fact that these phosphorylation sites are biomarkers for cancers and tumor biology, their individual roles in shaping AKT1 substrate selectivity are unknown. We recently developed a method to produce AKT1 with programmed phosphorylation at either or both of its key regulatory sites. Here, we used both defined and randomized peptide libraries to map the substrate selectivity of site-specific, singly and doubly phosphorylated AKT1 variants. To globally quantitate AKT1 substrate preferences, we synthesized three AKT1 substrate peptide libraries: one based on 84 "known" substrates and two independent and larger oriented peptide array libraries (OPALs) of ~10^{11} peptides each. We found that each phospho-form of AKT1 has common and distinct substrate requirements. Compared with pAKT1^{T308}, the addition of Ser-473 phosphorylation increased AKT1 activities on some, but not all of its substrates. This is the first report that Ser-473 phosphorylation can positively or negatively regulate kinase activity in a substrate-dependent fashion. Bioinformatics analysis indicated that the OPAL-activity data effectively discriminate known AKT1 substrates from closely related kinase substrates. Our results also enabled predictions of novel AKT1 substrates that suggest new and expanded roles for AKT1 signaling in regulating cellular processes.

Protein kinase B (PKB or AKT1) is a central node in the phosphoinositide 3-kinase (PI3K)/AKT signaling pathway that regulates cell survival (Fig. 1). There are three AKT isozymes (AKT1, AKT2, and AKT3) identified in mammalian cells with both distinct and overlapping roles (1). AKT1 is an oncogenic kinase that is overactive and hyper-phosphorylated in most cancers and a prime target in cancer therapy (2). Although direct AKT1 inhibitors have not been generally successful as monotherapies (3), AKT1 inhibitors do show promise in combination with, for example, mitogen-activated protein kinase inhibitors (4). Activated AKT1 phosphorylates downstream targets that regulate many cellular processes including cell survival and apoptosis (5). The diverse cellular functions of AKT1 are mediated by its ability to phosphorylate perhaps as many as hundreds of cellular substrates, although the precise number is not known. Based on a large volume of reports from the literature, a catalogue of ~200 known AKT substrates have been reported (6).

At present it is unclear if all of these substrates are truly phosphorylated by AKT1 or to what extent the known substrates are differentially phosphorylated by active AKT1 phosphovariants. Because the downstream phosphorylation profile of active AKT1 determines which cellular programs are turned on or off by the enzyme, the ability to identify novel AKT1 substrates will have significant impact on target selection and drug discovery in AKT1-mediated chronic diseases such as cancers (1) and diabetes (7).

We recently developed an approach to produce AKT1 variants with site-specific phosphorylation at either or both key regulatory sites (Thr-308 and Ser-473) (8, 9). Here we show the method provides an indispensable tool for identifying the role of AKT1-phosphorylation status in substrate selectivity. The substrate selectivity associated with specific AKT1 phosphoforms is challenging to resolve in the complex environment of cells (10). The phosphorylation status of AKT1 substrates depends on many factors, including the expression level and cellular localization of the substrate as well as the off-rate of phosphate on these substrates as a result of phosphatase activity (11). There remains a critical lack of understanding regarding how the two key activating phosphorylation sites on AKT1 lead to differences in substrate selection and downstream substrate phosphorylation (10) (Fig. 1).

The canonical AKT substrate recognition motif has been defined as $R_{-5}X_{-4}R_{-3}Z_{-2}\ Z_{-1}(S/T_0)\phi_{+1}$. X represents any amino acid; Z indicates preference small amino acids expect Gly; ϕ represents a bulky hydrophobic residue, and the phospho-accepting site at position 0 is a Ser or Thr residue (12). The AKT target consensus was initially determined using the amino acid sequences surrounding the phosphorylation site of the first reported AKT substrate, glycogen synthase kinase 3 (GSK-3) (12). Substrates with Arg residues at position −5 and −3 were found to be specific to AKT, whereas other AGC family kinases, such as the S6 kinase (S6K1), prefer Lys at −3 and −5 positions

This article contains supporting information.
‡ These authors contributed equally to this work.
* For correspondence: Patrick O'Donoghue, patrick.odonoghue@uwo.ca.

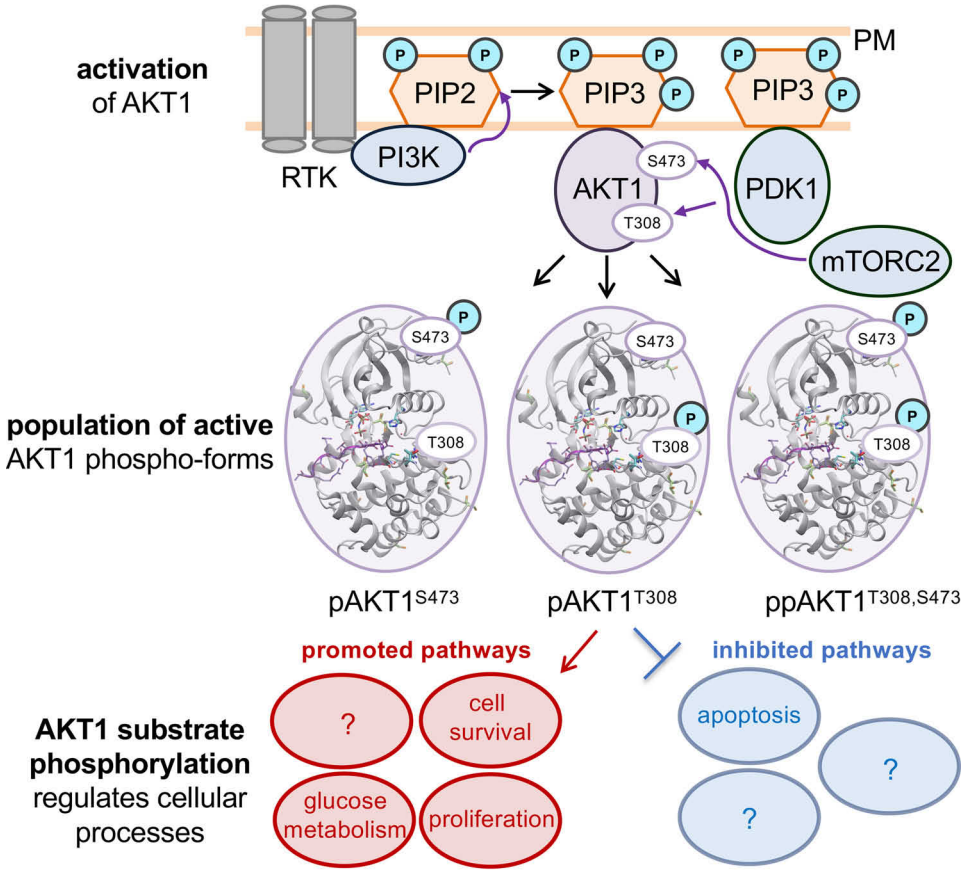

Figure 1. Schematic view of AKT1 activation and activity. AKT1 binds to phosphatidylinositol (3,4,5)-trisphosphate (PIP_3) molecules generated in the plasma membrane (*PM*) following receptor tyrosine kinase stimulation. AKT1 is activated by phosphorylation at key regulatory sites Thr-308 and Ser-473 by the upstream kinases PDK1 and mTOR complex 2 (*mTORC2*), respectively. The result of this activation process in combination with cellular phosphatase activities leads to a population of different phospho-forms of AKT1 in the cell (59). In cancer cells, these populations are known to change in response to altered metabolism or drug treatments (44). Active AKT1 phosphorylates many downstream substrates leading to inhibition or activation of cellular pathways linked to cell death, growth, and survival. These pathways only represent examples as AKT1 is thought to regulate a large number of proteins and specific pathways, possibly including some (?) that are not yet known. In addition, the role of each phospho-form of AKT1 in regulating these critical cellular processes is unknown.

(10). Using active preparations of AKT1 from Sf9 insect cells over-expressing PDK1 and human AKT1, the target motif was further investigated through oriented peptide array library (OPAL) screens. The results indicated selectivity for Arg residues at −3, −5, and −7 as well as Thr at −2. AKT1 also showed moderate preferences for aromatic residues at positions −1 and +1 and for small residues able to induce tight turns (Gly, Ser, Asn, or Thr) at the position +2 (13). Because Sf9 produced AKT1 normally contains a mixture of active AKT1 phospho-forms (8, 14), these data could not identify substrate preferences associated with differentially phosphorylated AKT1.

Peptide library screening is a gold standard method that has been widely used to determine specific amino acid preferences of the kinase substrate recognition motif. Degenerate peptide libraries or OPAL screens provide a systematic approach to identify the important residues surrounding the phosphorylation site on the substrate (13, 15). Here we designed and synthesized OPALs to define the substrate requirements for differentially phosphorylated AKT1 variants. We discovered that the phosphorylation status of AKT1 has a significant impact on substrate selectivity and on the specific residue preferences at each location in the consensus motif. Kinase activity data derived from our OPAL experiments was able to differentiate each AKT1 phospho-form from another. The OPAL data were also able to effectively discriminate known AKT1 substrates from closely related substrates in database searches, which also suggested high confidence putative AKT1 targets were enriched in pathways linked to RNA metabolism.

Results

For preparation of active and purified AKT1, the use of phosphoinositide-dependent kinase (PDK1) leads to efficient phosphorylation of AKT1 at Thr-308. Although in the cell mTORC2 is recognized as the major upstream kinase for Ser-473 (11, 16), incorporation of phosphate at Ser-473 is challenging in the test tube. Previous approaches relied on phosphorylation of Ser-473 by MAPKAPK-2, which was pre-activated by p38 in the presence of phosphatidylinositol (3,4,5)-trisphosphate and lipid vesicles (17). In contrast, our approach leads to recombinant and site-specifically phosphorylated AKT1 variants produced in *Escherichia coli* without the need to purify or activate additional kinases. Using a combination of genetic code expansion to encode phosphoserine at Ser-473 and co-expression of PDK1, we were able to produce active full-length and PH domain-truncated AKT1 variants in *E. coli* with either or both Thr-308 and Ser-473 sites phosphorylated. We previously characterized the activity and phosphorylation status of the enzyme preparations biochemically and with MS (8, 9). In that

work, we discovered that phosphomimetic substitutions are incapable of replacing or approximating the functionality of phosphorylated residues at the activation sites of AKT1. We further found that alanine substitutions at Thr-308 and Ser-473 inactivated the enzyme (8). One strength of our approach, therefore, is the ability to generate pAKT1 and ppAKT1 variants without the need to mutate the enzyme, which greatly alters enzyme function.

Our previous studies on pAKT1 and ppAKT1 variants focused on their differential activity with a single substrate peptide derived from GSK-3β (8). Here we quantified the substrate preference for phosphorylated AKT1 variants using a synthetic peptide library of known AKT1 substrates, and two independent oriented peptide arrays. We found that deletion of the PH domain abolishes its autoinhibitory activity and leads to a more soluble enzyme (9) appropriate for larger scale screening studies presented here.

AKT1 phosphorylation status alters activity among known substrates

A set of "known" AKT1 substrates have been identified and in some cases validated using multiple experimental approaches, including *in vitro* biochemical (18) and cell-based activity assays (19), proteomic analyses (20), as well as phospho-specific immunoblotting of AKT1 substrates (21) from cells (22) and animals (23). Our study allowed the evaluation of these separately reported substrates using a systematic approach to quantify AKT1 kinase activity. The doubly phosphorylated AKT1 was capable of phosphorylating most, but not all, of the substrate peptides in our peptide library representing known substrates (Fig. S1). The ppAKT1 enzyme showed significantly above background activity with 95% of the peptides tested (Fig. S1*C*). In our activity assays, the initial reaction velocity varied over an 80-fold range. We found that ppAKT1 produced between 0.1 and 8.0 pmol/min of phosphorylated peptide.

Overall, the level of phosphorylation observed was markedly lower with pAKT1^{S473} compared with pAKT1^{T308} and ppAKT1^{T308,S473} (Fig. S1). Compared with the singly phosphorylated variants (Fig. S1, *A* and *B*), ppAKT1^{T308,S473} (Fig. S1*C*) usually displayed the greatest activity over the widest range of substrate peptides. Indeed, we observed a clear trend in activity over the library with pAKT1^{S473} showing the most restricted substrate range. The pAKT1^{S473} enzyme showed low activity ($<$0.1 pmol/min) with more than 50% of the substrates. The variants phosphorylated at Thr-308 showed activity above this same relative level with $>$90% of the library.

Interestingly, some substrates representing interleukin-1 receptor-associated kinase 1 (IRAK1) and B-cell CLL/lymphoma protein 10 (BCL10), showed no significant activity with any of the AKT1 phospho-forms. Although BCL10 was found in association with AKT according to immunoprecipitation experiments (24), the ability of AKT1 to directly phosphorylate BCL10 was not demonstrated. Indeed, further studies in T cells found that AKT1 is not required to phosphorylate BCL10 (25). The suggestion of IRAK1 as an AKT1 substrate was based on work in human embryonic kidney cells showing that over-expression of calcium/calmodulin-dependent protein kinase kinase (CaMKKc) or AKT lead to increased phosphorylation at Thr-100, but a direct link between AKT and IRAK1 was not established (26). It is possible that AKT1 is not a direct kinase for these substrates or that the substrates require the complete protein structure or other co-incident modification to become active substrates for AKT1.

AKT1 phosphorylation status re-wires substrate selectivity

The three AKT1 phospho-forms have distinct preferences for particular peptides. The highest active substrates for the pAKT1^{S473} variant were completely orthogonal to or distinct from the most active substrates with pAKT1^{T308} or ppAKT1^{T308,S473} (Figs. S1 and S2). To visualize these distinct substrate preferences, we generated a heat map showing the relative activity of each AKT1 phospho-form over the library of 84 known substrates (Fig. 2). The substrates were categorized into groups based on shared biochemical function. Inside each functional group, the substrates were ranked according to decreasing ppAKT1 activity, and the functional groups were also listed in order of decreasing average activity with ppAKT1.

The ppAKT1 enzyme is most active on average with the Forkhead box O (FOXO) family of transcription factors, which are well-established substrates of AKT1 (21, 27). Peptides representing the FOXO family were among the most active substrates with ppAKT1 and AKT1 phosphorylated only at Thr-308 (Fig. S1, *B* and *C*). Certain substrates, including FOXOs and TSC2, have more than one AKT1-dependent phosphorylation site, which we represented with distinct synthetic peptides. Each AKT1 phospho-variant showed different preferences for the multiple sites in these substrates. For example, with pAKT1^{S473}, Ser-256 and Ser-319 of FOXO1 showed relatively high activity, but this enzyme showed low activity with a peptide derived from FOXO1 Thr-24. In contrast, for both pAKT1^{T308} and ppAKT1^{T308,S473} FOXO1 Ser-319 was the least preferred compared with Ser-256 and Thr-24 (Fig. 2).

Impact of Ser-473 phosphorylation on active AKT1

The accepted view that ppAKT1 is more active than singly phosphorylated AKT1 is not always true at the resolution of individual substrates. To better understand the impact of Ser-473 phosphorylation on the activity of the doubly phosphorylated AKT1 for our library of known substrates, we plotted kinase activity data of pAKT1^{T308} and ppAKT1^{T308,S473} together (Fig. 3). We analyzed these comparative data for statistically significant differences in activity. Interestingly, we observed that additional phosphorylation at Ser-473 significantly increased AKT1 activity for many (49%, 41 of 84) but not all AKT1 substrates (Fig. S1, and Fig. 3). In the context of AKT1 phosphorylated at Thr-308, for nearly half of the substrates (45%, 38 of 84), Ser-473 phosphorylation did not change activity, and for 6% (5 of 84) of the substrates, phosphorylation at Ser-473 actually decreased kinase activity (Table 1, Fig. 3). Phosphorylation at Ser-473 reduced activity for two peptides derived from distinct sites in the FOXO4 transcription factor (Thr-32, Ser-262), as well as sites in 14-3-3Z, Grb2-associated binding protein (Gab2), and TSC complex subunit 2 (TSC2). Together the data show that both regulatory phosphorylation sites in AKT1 have a significant and unexpectedly complex

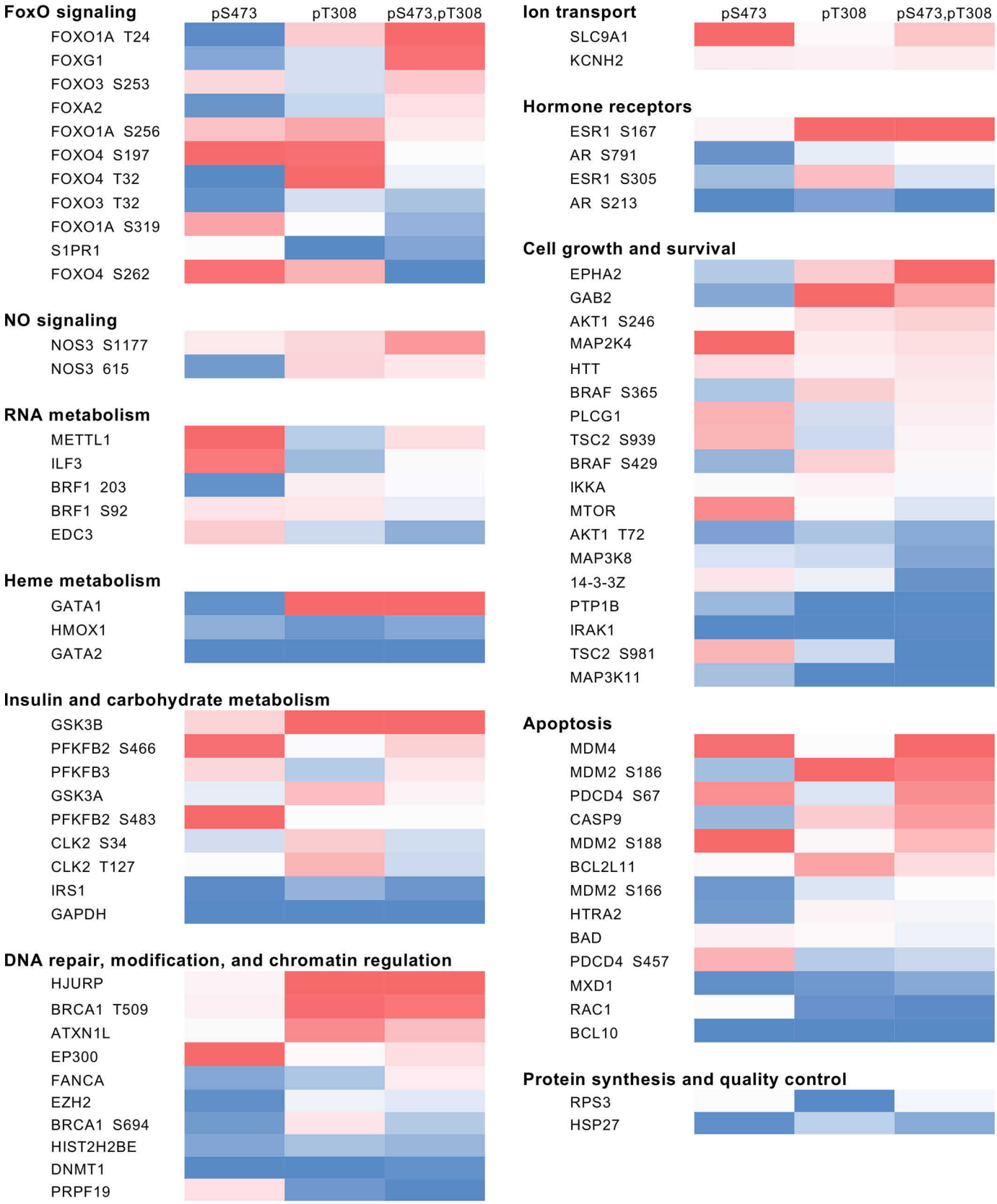

Figure 2. Relative activity of AKT1 variants across functional clusters of known substrates. The heatmap is colored according to high (*red*), medium (*white*), and low (*blue*) relative activity for pAKT1^{S473}, pAKT1^{T308}, and ppAKT1^{T308,S473}. The substrates were categorized into functional categories. Inside each group, the substrates are ranked according to decreasing ppAKT1 activity, and the functional groups are listed in order of decreasing average activity with ppAKT1.

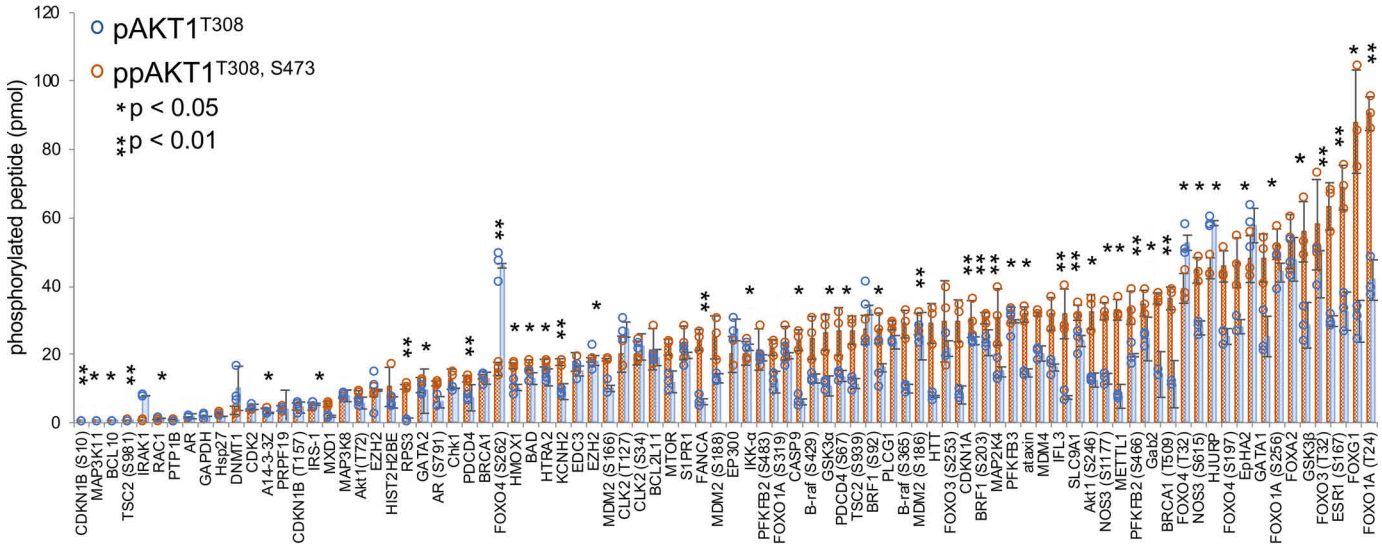

Figure 3. Impact of Ser-473 phosphorylation on active AKT1. Kinase activity of pAKT1^{T308} (*blue*) and ppAKT1 (*red*) were analyzed for statistically significant differences (*, $p < 0.05$; **, $p < 0.001$). *Error bars* represent ±1 S.D. based on 3 independent enzyme reactions.

Table 1
Activity of pAKT1^{T308} and ppAKT1 across known substrate peptides

AKT1 activity	Substrate peptides	
	Number/84	%
ppAKT1 > pAKT1^{T308}	41	49%
ppAKT1 = pAKT1^{T308}	38	45%
ppAKT1 < pAKT1^{T308}	5	6%

impact on substrate selectivity over a library of known AKT1 substrates.

Phosphorylation-dependent changes in the AKT1 target motif

Based on the activity data in Fig. S1, we computed sequence logos for each AKT1 phospho-form. The logos (Fig. S2) represent the relative importance of residues at each site for AKT1-dependent phosphorylation based on the activity level recorded for each substrate peptide (Fig. S1). The sequence alignments used to generate the logos were, therefore, populated with each of the 84 peptides such that the number of occurrences of each peptide in the alignment was linearly related to the enzyme activity (see "Materials and methods").

The preference for conserved consensus sequence residues Arg_{-5} and Arg_{-3} among known AKT1 substrates ($R_{-5}XR_{-3}XX$(S/T)) was observed with all three variants. As reported previously (12), the presence of a large hydrophobic residue following the phosphorylation site (+1 position) was observed among the peptides with all three AKT1 variants. Pro most frequently occupied the +2 position among highly active peptides for the pAKT1^{T308} and ppAKT1^{T308,S473} variants. These data are in agreement with the finding that the +2 position prefers amino acids that enable tight turns in protein structure (13). Interestingly, peptide substrates showing high activity with a pAKT1^{S473} variant were dominated by those peptides with a Ser rather than a Thr at the phosphorylation site. Conversely, for pAKT1^{T308} and ppAKT1^{T308,S473}, highly active peptides were approximately equally likely to have Ser or Thr at the phosphorylation site.

Determination of high-resolution AKT1 target specificity

Our analysis of known substrates revealed several new facets of AKT1 enzymology. These data, however, are biased by our current and perhaps limited knowledge of AKT1 substrates. We designed a far larger, unbiased oriented library of peptides (OPAL1) to test the most preferred amino acid residues at each site of the AKT1 substrate consensus motif. OPAL1 used 17 different amino acids at each of 8 variable positions in the synthetic peptides. We synthesized 136 different pools of peptide sublibraries to cover the complete library of ~10^{10} peptides (see "Materials and methods"). Using each phospho-form of AKT1, we conducted kinase assays with each of the substrate sublibraries (Fig. 4*A*, Figs. S4–S7 and Data File S1). In each sublibrary a single amino acid is held constant at a single position in the consensus motif (Fig. 4*A*), whereas all other variable positions are allowed to cover the complete sequence space with the exception of Cys, Thr, and Ser. Because Arg_{-3} is normally required for AKT1 activity (13), we held this position fixed in OPAL1.

The AKT1 activity for singly and doubly phosphorylated AKT1 variants over each of the peptide sublibrary pools are shown in terms of absolute activity (Fig. S7) and overlaid as a heat map indicating relative changes in activity (Fig. 4*A*). To visualize the amino acid preferences and anti-determinants we converted the activity data into position-specific score matrices (PSSMs, see "Materials and methods"). We used the PSSMs to generate sequence logos (28) representing the OPAL1 data (Fig. 5). The results demonstrate that the substrate selectivity of pAKT1^{S473} is distinct from the profile we observed for the more similar pAKT1^{T308} and ppAKT1^{T308,S473} enzymes. This observation is confirmed by principle component analysis (PCA) of the 3 OPAL1 matrices (Fig. 4*C*). PCA plots the variance between the OPAL matrices such that the first component explains most of the variance between the data sets. The principle components are linear combinations of the observed variables, which are the activity values at each location in the oriented array. The magnitude and direction of variance at each

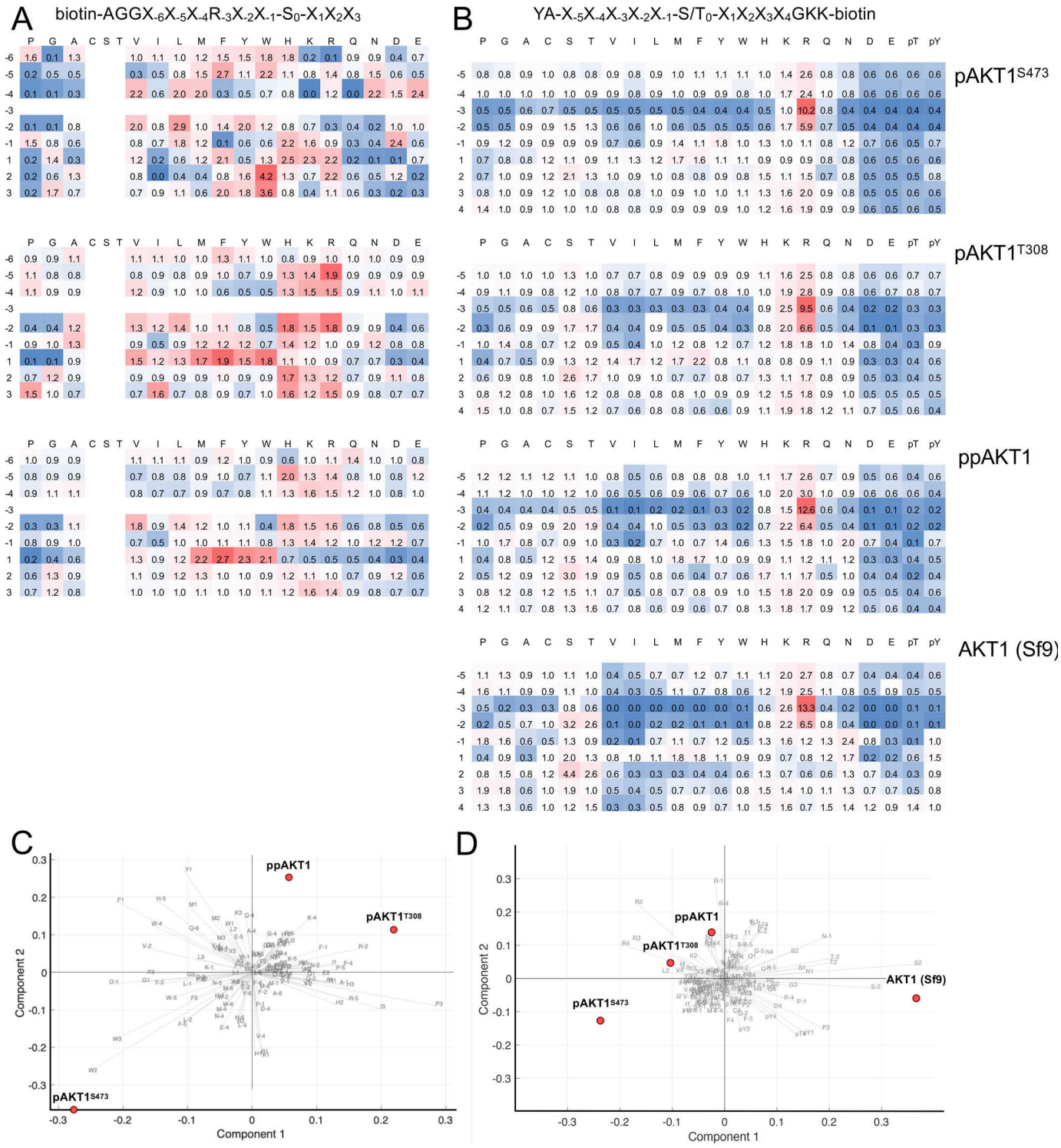

Figure 4. Oriented peptide array library 1 (*OPAL1*) and library 2 (*OPAL2*) activity assays. Row normalized activity data for OPAL1 (*A*) and OPAL2 (*B*). The color scale indicates high (*red*) and low (*blue*) relative activity. Each cell in the OPAL1 matrices (*A*) represents relative activity with the indicated sublibrary in reactions catalyzed by pAKT1^{S473}, pAKT1^{T308}, or ppAKT1^{T308,S473}. The data are based on 3 independent enzyme reactions (Figs. S4–S7). The 0 position is the phosphorylation site, which was fixed to a Ser; the −3 position was fixed to an Arg. *B,* the analysis of OPAL2 (Fig. S8) also includes data for AKT1 prepared from Sf9 cells (*B*, *bottom*). The OPAL1 (*C*) and OPAL2 (*D*) data were analyzed by principle component analysis. The first two principle components are plotted along with vectors representing the direction of greatest variation of the observed activities at each amino acid and position.

position in the array is shown as a vector. For example, the preference of pAKT1^{S473} for Trp at positions 2 (W_2) and 3 (W_3) most distinguishes this enzyme from the variants phosphorylated at Thr-308 (Fig. 4*C*).

Although consensus patterns are evident (Table 2), specific amino acids significantly preferred at each position show distinguishable profiles for each AKT1 phospho-form (Fig. 4, *A* and *C*) in concordance with the unique substrate selective pro-

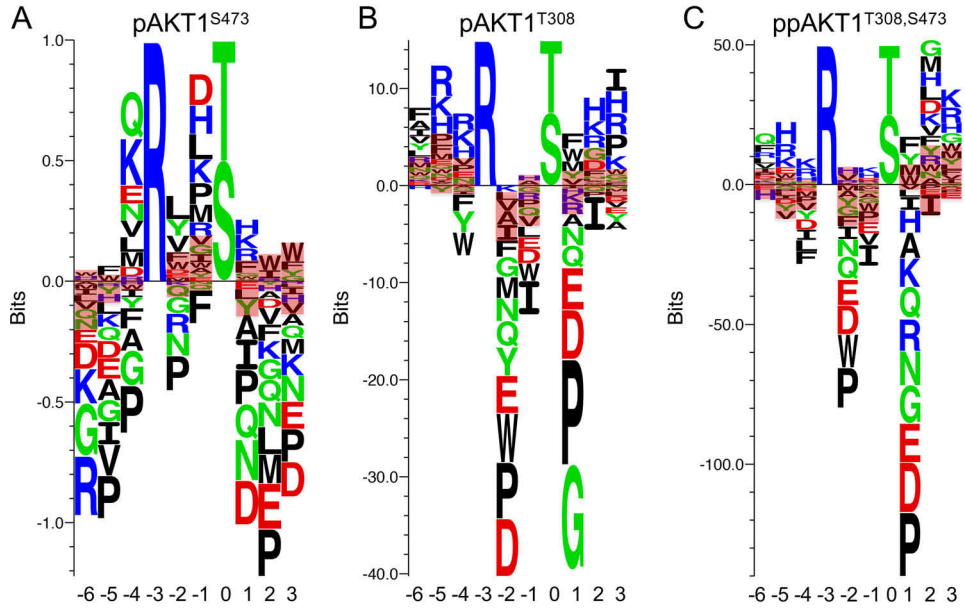

Figure 5. Sequence logos generated form the OPAL1 data. Sequence logos of (*A*) pAKT1^{S473}, (*B*) pAKT1^{T308}, and (*C*) ppAKT1^{T308,S473} are based on the conversion of activity values from Fig. S7 into a PSSM. Seq2Logo 2.0 (28) was used to convert the PSSMs to amino acid preferences as represented by sequence information bits. The average level of activity observed in the OPAL1 data (Fig. S7) was set to 0 bits; positive and negative values in bits refer to activity that is above or below the average activity over the array, respectively. The data were also statistically analyzed by comparing the overall average enzyme activity against the enzyme activities of all tested amino acids at each position (*i.e.* −6 to +3) using pairwise *t*-tests. Kinase activity that is not significantly different from the average activity is highlighted in *red*. The graph thus shows statistically significant amino acids preferences that are either more favored than average (above the *red bar*) or less favored than average (below).

files we recoded over the library of known substrates (Fig. S1, and Fig. 2). By varying the Arg$_{-5}$ position, we found that basic residues are highly favored and Arg$_{-5}$ is most preferred or nearly so for the enzymes phosphorylated at Thr-308. The pAKT1^{S473} enzyme is active with basic residues at the −5 position yet prefers aromatic residues at this site. Both ppAKT1^{T308,S473} and pAKT1^{T308} preferred large hydrophobic residues (Phe and Trp or Tyr) at the +1 position in the target peptide. In contrast, pAKT1^{S473} prefers basic residues (Lys, Arg, His) at +1. For the pThr-308 containing enzymes, basic amino acids (Arg, Lys, His) are among the most active residues at most of the variable positions (−5, −4, −2, −1, +2, and +3); this same preference is also seen at positions −4, −1, +1 for the pAKT1^{S473} enzyme.

In terms of anti-determinants to kinase activity, Ile at −1 and +2 positions were disfavored for both pAKT1^{T308} and ppAKT1^{T308,S473} variants; Phe and Pro are least preferred at −1 and +2 positions for pAKT1^{S473}. ppAKT1 was least selective at position +2, showing high activity with a large variety of amino acids. As observed before (12), we found Pro at +1 significantly inhibited activity in all of our OPAL matrices.

To confirm the generality of these findings, we conducted independent OPAL2 experiments based on a distinct peptide design (Fig. 4*B*, and Fig. S8). In contrast to OPAL1, this library probed the −3 position, which was fixed as Arg in OPAL1, included phosphorylated Thr and Tyr residues, and varied the +4 but not the −6 position. OPAL2 used 22 different amino acids at each of 9 positions leading to a library size of ~10^{12}. OPAL2 used a 15-residue peptide, rather than 13 in OPAL1, and the biotin tag was placed on the C rather than N terminus. The results confirm the dominant preference of Arg$_{-3}$ for all AKT1 phospho-forms, yet significant activity is observed with Lys$_{-3}$ and to a lesser extent His$_{-3}$. The OPAL2 data were also able to clearly discriminate each phospho-form of AKT1 from the other (Fig. 4*D*). We observed that the relative differences between the OPAL2 matrices were not as great as that observed in OPAL1 (Fig. 4), suggesting the OPAL1 peptide design may discriminate between the phospho-forms more effectively.

OPAL2 also confirmed that the pThr-308 containing enzymes, including AKT1 produced in insect cells (Sf9), were all far more active than pAKT1^{S473}. The finding is consistent with the data here (Figs. S1, S7, and S8) as well as our previous findings (8, 9). The Sf9 prepared AKT1 showed many commonalities with the AKT1 variants prepared in *E. coli*, however, the PCA indicates the Sf9 enzyme was distinguishable from the others with the majority of the variance explained by preferences for Ser/Thr$_{-2}$, Asn$_{-1}$, Ser/Thr$_{2}$, and Pro$_{3}$. Overall, the PCA indicates that although the Sf9 AKT1 is most similar to ppAKT1, it is a distinct AKT1 enzyme and not a linear combination of the pSer-473 or pThr-308 containing enzymes. Together the data further support our hypothesis that each AKT1 phospho-form has distinct substrate preferences.

Despite these differences, similarities among the OPAL data are also apparent. We summarized the consensus residues that are most or least favorable at each position across experiments and AKT1 variants (Table 2). His$_{-1}$ is among the most preferred for all phospho-forms, whereas Val and Ile are least preferred for all variants at −1. The addition of a basic residue, normally Arg or Lys but also His, at nearly any position will lead to increased activity with any AKT1 variant. The only clear exception to this rule is at +1 where Phe is commonly favored. Conversely, introduction of a negatively charged amino acid (Asp, Glu, as well as pTyr and pThr, Fig. 4*B*) at any position will lead to significantly reduced AKT1 activity (Fig. 5). The intro-

Table 2

Consensus of positive (+) and negative (−) AKT1 activity determinants from OPAL data

The table is based on sequence logo analysis in Fig. 5 (OPAL1)[a] and Fig. S9 (OPAL2)[b]. Bold residues indicate residue identities that are a consensus between AKT1 phospho-forms and experiments. Positive (black) and negative (red) activity determinants are listed in order of decreasing entropy. Data not determined (–).

position	pAKT1^{S473}		pAKT1^{T308}		ppAKT1		AKT1 Sf9		Consensus	
	+	-	+	-	+	-	+	-	+	-
-6[a]	W	KGR	FAV	H	QFR	H	--	--		
-6[b]	--	--	--	--	--	--	--	--		
-5[a]	F	IVP	**RKH**	Y	HRK	DYV	--	--	**RK**	**IDE**
-5[b]	RK	NED	**RKSH**	IDE	**RK**	EID	**RK**	DEV		
-4[a]	QK	AGP	**RKH**	FYW	**KRH**	DILF	--	--	**RK**	**D**
-4[b]	RK	IDE	**RK**	EID	**RK**	DFI	**RK**	DVI		
-3[a]	--	--	--	--	--	--	--	--		
-3[b]	**R**	NDE	**RK**	IVDE	**RK**	VFEDI	**RK**	DEVLI	**RK**	**DEV**
-2[a]	LYFVW	PN	KLV	WPDE	HV**R**	WPDE	--	--	**RK**	**PWDE**
-2[b]	RK	PWNED	RK	WPDE	**R**KST	PWDE	**R**STK	WEID		
-1[a]	DHLK	F	**H**	I	**KH**	IV	--	--	**KH**	**VIE**
-1[b]	YM**H**N	VIE	**K**RN	VEI	NRG**KY**H	IV	NP**K**	VI		
1[a]	HRK**F**	PQND	**FYMW**	EDPG	**FYMW**	GEDP	--	--	**FM**	**PDE**
1[b]	M**F**I	PDE	FMIV	PDE	**MF**	PDE	SFM	ED		
2[a]	WI	EP	**HKR**	I	GM**H** **KR**	EI	--	--	**RHK**	DP **E**
2[b]	SRH	DE	ST**RHK**	PDE	**RHK**	DEI	STG	MLI		
3[a]	W	EPD	I**HR**	DEYA	**KRH**	DPE	--	--	**RK**	EP **D**
3[b]	**RKH**	PDE	**R**SKH	LVED	**KR**	DE	SPG	LVI		
4[a]	--	--	--	--	--	--	--	--		
4[b]	RKP	VDE	**K**RP	FYE	**K**R	EFD	**K**H	LIV	**K**	**E**

duction of Pro or small hydrophobic residues (Ile, Val) is disfavored at the majority of positions for all AKT1 variants.

Identifying known and putative AKT1 substrates

We derived PSSMs from the OPAL1 data according to an approach established previously (29) and detailed under "Materials and methods." To begin to estimate the predictive value of these PSSMs, we plotted the activity data we generated for the 84 known substrates (Fig. S1) against the PSSM score that results from that particular peptide according to the OPAL1-derived PSSMs for each AKT1 phospho-form (Fig. S10). Interestingly, these data show statistically significant correlations ($p \sim 10^{-5}$) for the pAKT1^{T308} and ppAKT1^{T308,S473} enzymes, suggesting PSSMs derived from the OPAL1 data can be used to identify known and predict novel targets. The data, however, did not correlate for the pSer-473 enzyme (Fig. S10*A*). For this enzyme only, our data suggests that either the OPAL1-derived PSSM is not a good predictor, or the 84 known substrates are not true substrates for pAKT1^{S473}. The latter conclusion is supported by our data showing low activity of the pAKT1^{S473} enzyme (Figs. S1, S4, S7 and S8) as well as our previous work showing that AKT1 phosphorylated at Ser-473 alone was not

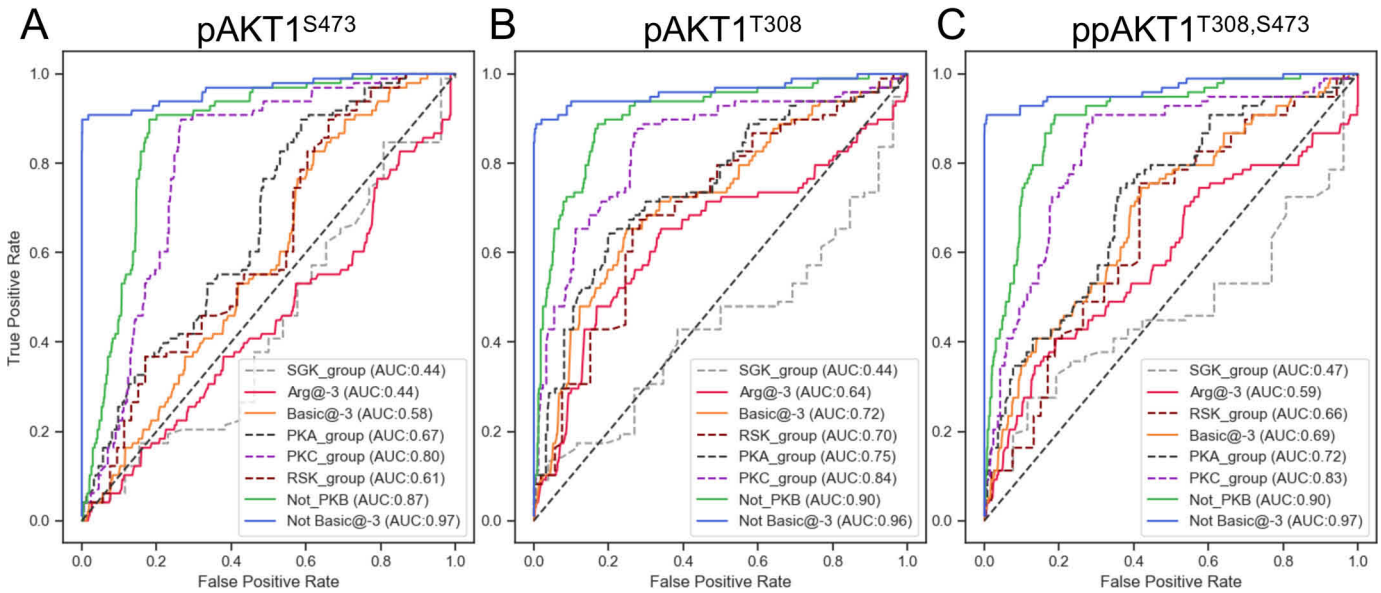

Figure 6. ROC plot of phospho-AKT1 PSSMs. The phosphoELM database was searched to determine the extent to which the OPAL1-derived PSSM could discriminate known AKT1 substrates (true positives) from the indicated groups. The PSSMs achieve discrimination against groups that include false-positives (*e.g.* peptides lacking a basic residue at −3 (*blue*); peptides that are not known as AKT1 (Not_PKB) substrates (*green*); substrates of PKC (*purple dashed line*)). As expected, if the indicated groups contained true positive AKT1 substrates (*e.g.* SGK group, R_{-3} group, basic at −3) the PSSM could not discriminate.

sufficient to propagate AKT1-dependent signaling in cells (8). The OPAL1 data (Fig. 4*A*) also suggests that the pAKT1^{S473} enzyme may prefer rather different substrates (Fig. 5*A*) than those included in the library of known substrates (Fig. S2).

We then used PSSMSearch (30) to rank potential AKT1 substrates across the entire human proteome with PSSMs derived from OPAL1 (Data File S2). We refined these data to include only hits with experimental evidence for phosphorylation at the predicted S_0/T_0 site (Data File S2). Collectively, the search identified both known and potentially novel AKT1 substrates. The known hits were significantly enriched in GO biological processes related to cell signaling and apoptosis, whereas the top novel hits among all AKT1 variants were enriched in processes linked to RNA processing and metabolism (Fig. S11).

To gauge the reliability of the PSSMs, we performed specificity and selectivity analysis (Table S1, Fig. 6, Fig. S12). The PSSMs were used with PSSMSearch over a database of known AKT1 substrates with a specified outgroup of kinase substrates as defined by PhosphoELM (31). Several outgroups were tested, including those lacking known AKT1 substrates (such as Not_PKB or Not $Basic_{-3}$) and other groups that do contain AKT1 (S6K, Arg_{-3}) or AKT1-like substrates (PKA, PKC). In all cases, the PSSMs scored known AKT1 substrates higher than other groups (Fig. S12), however, greater discrimination between known AKT1 substrates and other groups is evident in the enzymes phosphorylated at Thr-308.

Despite phospho-form differences in the OPAL data, noted above, activity data derived from each phospho-form performed well in identifying known AKT1 substrates (Fig. 6). We plotted the data in receiver operator characteristics (ROC) plots (Fig. 6). In a search including known AKT1 substrates along with all those lacking a basic residue at −3, nearly all of the true AKT substrates are found before the search hits "false-positives" that are from the Non-$Basic_{-3}$ group. The degree of discrimination between known AKT1 substrates and the specific outgroup is characterized by the area under the curve (AUC), which has an inverse correlation with the PSSM score of the outgroup (Table S1). The ROC plot shows the OPAL1 data are effective in discriminating known AKT substrates (true positives) from groups that do not contain AKT substrates. As expected, there is less discrimination and reduced AUC values for outgroups that contain known AKT1 substrates, *e.g.* Arg_{-3} group. We observe better discrimination using OPAL data for the pThr-308 containing enzymes compared with pAKT1^{S473}. In all cases, the OPAL data also provide discrimination against closely related protein kinase C (PKC) substrates and to a lesser extent protein kinase A (PKA) as well. The OPAL2 data performed similarly well in database searches (Fig. S13). Although according to the AUC values, we observed a marginally decreased ability of the OPAL2 data to discriminate known AKT substrates.

Next, we used an alternate and independent approach to score the known human phospho-proteome to identify potential new AKT1 substrates. We converted the OPAL1 AKT1 activity data into PSSMs based on calculating Z-scores (see "Materials and methods") as before (15). The known human phospho-proteome was scored based on the amino acid preferences at each position in the Z-score PSSM (Data File S3). The scoring of peptides was performed for data based on each phospho-form of AKT1. We then selected the top scoring substrates with a normalized sum of Z-scores >3. We included all of these high confidence hits in a Search Tool for Retrieval of Interacting Genes/Protein (STRING) (32) diagram to identify potential protein interactions and functional clusters among the known and predicted substrates (Fig. 7).

The searches identified both known and unknown AKT1 substrates among the high confidence hits. Interestingly, the prediction suggested the presence of substrates not previously

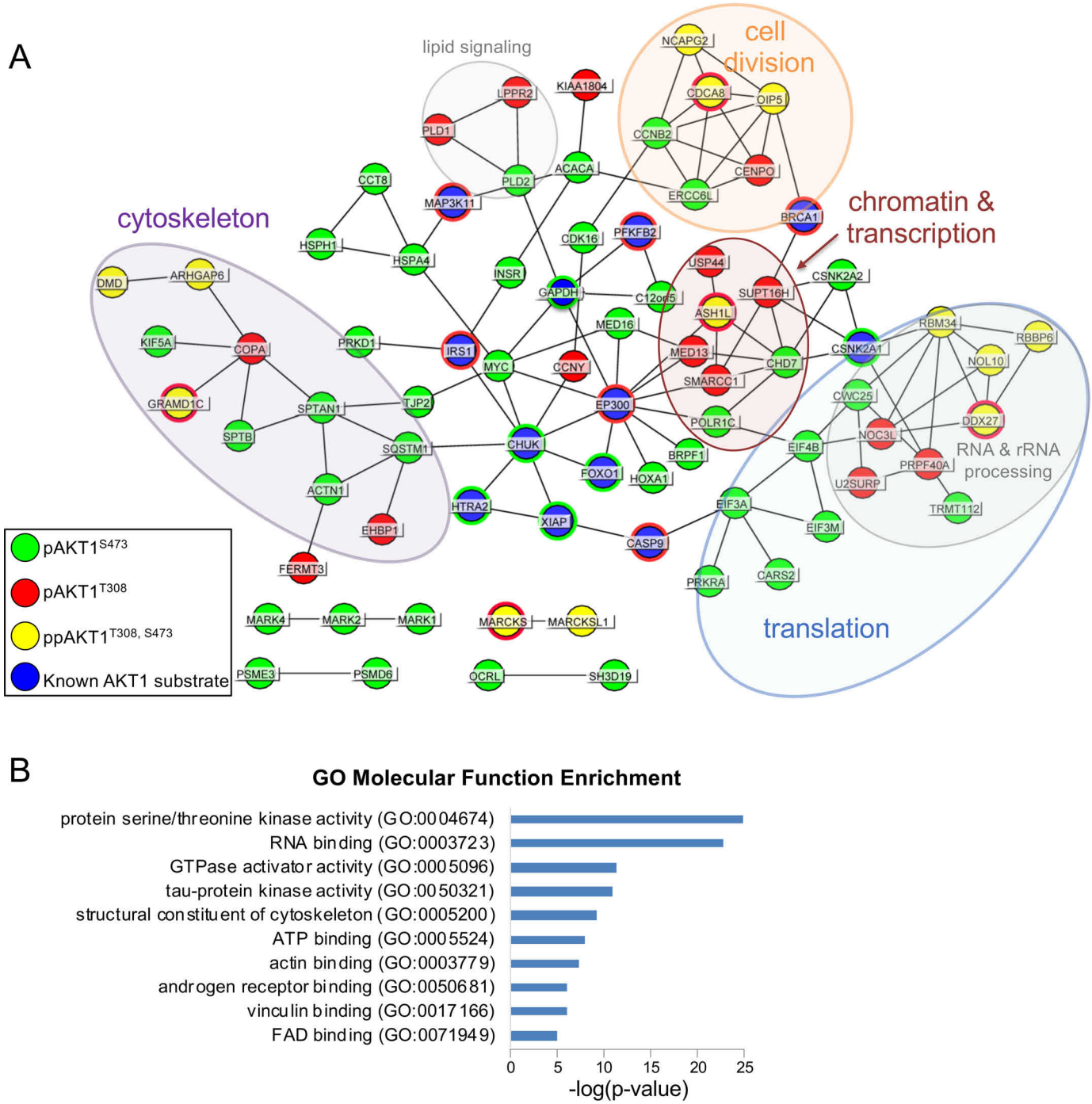

Figure 7. Identifying optimal phosphospecific AKT1 substrates. *A,* STRING diagram shows functional interactions of known (indicated by nodes with *blue inner filled circles*) and putative AKT1 substrates. The diagram includes high confidence hits to the known human phospho-proteome (Z = score >3) using OPAL1 data generated with pAKT1^{S473} (*green*), pAKT1^{T308} (*red*), and ppAKT1^{T308,S473} (*yellow*) variants. Novel clusters of substrates are involved in cell division (*orange*), transcription (*brown*), translation (*blue*), cytoskeleton (*purple*), RNA processing, and lipid signaling. *B,* GO enrichment analysis with the high confidence hits shown in *A*. The analysis suggests statistically significant enrichment in protein kinases and RNA-binding proteins.

known to be regulated by AKT1. These include clusters of proteins that participate in translation, transcription, cell division, cytoskeletal organization, and lipid signaling (Fig. 7*A*). The putative AKT1 substrates are significantly enriched in gene ontology (GO) (33) molecular functions related to protein kinases, cytoskeletal organization, and RNA binding (Fig. 7*B*). Putative AKT1 substrates include centromeric proteins involved in cellular mitosis. Chromosome-associated protein G (NCAPG2), borealin (CDCA8), opa protein 5 (OIP5), and centromere protein O (CENPO) were the identified strong candidates for activity with each phospho-form of AKT1. Another novel cluster of putative AKT1 targets includes nucleolar proteins involved in rRNA processing. RNA-binding motif protein 34 (RBM34), nucleolar protein 10 (NOL10), RB-binding protein 6 ubiquitin ligase (RBBP6), and a DEAD box RNA helicase (DDX27) are also included in this functional cluster. Similar to

the results from our independent PSSMSearch (Fig. S11*A*), the top putative AKT1 substrate hits are strongly enriched in biological processes associated with RNA metabolism.

Discussion

We recently developed a new method to produce AKT1 with programmed phosphorylation at either or both key regulatory sites (8). This breakthrough enabled resolution of the contribution of each phospho-site to the activation (8) or inhibition (9) of AKT1. Because this approach was only recently developed, there are still very few studies that report on the function or impact of each phosphorylation site on AKT1 substrates (34). This situation occurs despite the fact that both phospho-sites are current and routine clinical markers for multiple cancers, *e.g.* see Refs. 35–37.

Prior investigations into the substrate specificity of AKT have focused on characterizing the isozyme-specific substrate preferences that differ between AKT1, AKT2, and AKT3 (1, 38). Indeed, the previous lack of tools to isolate the function of each AKT1 phospho-form inhibited studies to characterize the impact of each phosphorylation site on substrate specificity.

Phosphorylation-dependent substrate specificity of AKT1

In our previous work (8), we found that phosphorylation at Thr-308 was necessary and sufficient for maximal AKT1 activity in mammalian cells. The addition of Ser-473 phosphorylation increased kinase in the test tube, but did not influence phosphorylation of AKT-specific live cell fluorescent reporter in COS-7 cells (8). These observations led to our hypothesis that phosphorylation of Ser-473 may function to tune rather than activate the AKT1. Using peptide library approaches here, we mapped the substrate preferences for each key phospho-form of AKT1. The data presented revealed that Ser-473 phosphorylation does indeed impact substrate specificity. More generally, we found that the phosphorylation status of AKT1 has a global impact on substrate preference as determined by kinase activity assays with multiple independent peptide libraries.

A multitude of studies (39, 40), including our own (Figs. S1, S4–S6 and S8) (8), have presented evidence indicating that AKT1 activity increases monotonically with $pAKT1^{S473} < pAKT1^{T308} <$ ppAKT1. Based on all of our assays, it remains clear that the $pAKT1^{S473}$ enzyme is far less active than the other two. Of the 84 known substrates we tested, this rule only holds completely true for half of the peptides (49%). When comparing $pAKT1^{T308}$ with $ppAKT1^{T308,S473}$, we found that many substrates (45%) show the same level of activity with $pAKT1^{T308}$ and ppAKT1, whereas a minority (6%) are actually more active with the singly phosphorylated $pAKT1^{T308}$ compared with the doubly phosphorylated enzyme.

Our current study further revealed that when only Ser-473 is phosphorylated, the substrate preference profile is distinct compared with that observed for AKT1 with Thr-308 phosphorylation. This observation was supported by data generated from known substrates as well as the much larger data set generated using the OPAL approach. Interestingly, several of the most active known substrates for $pAKT^{S473}$, including MDM2 (Ser-188), MDM4, EP300, and eNOS (Ser-1177), are members of the p53-mediated apoptotic pathway (41) (Fig. 2, Fig. S1*A*). In the cell, phosphorylation of MDM family members and EP300 by AKT1 leads to ubiquitination and degradation of p53, thus promoting cell survival (41).

The significance of Ser-473 phosphorylation in differential substrate phosphorylation had been highlighted anecdotally. Jacinto *et al.* (34) found disruption of Ser-473 phosphorylation resulted from genetic ablation of *sin1* in mice. SIN1 is a component of the TORC2 complex, which phosphorylates Ser-473. In mice and mammalian cells, downstream AKT1 substrate phosphorylation events were subsequently down-regulated for only a subset of AKT1 substrates, which included FOXO1 and FOXO3. In agreement with these findings, all of the FOXO peptides except the N-terminal phosphorylation sites Thr-24 and Thr-32 of FOXO1 and FOXO3, showed reduced activity in the absence of Ser-473 phosphorylation (Fig. 3).

Additional studies have revealed correlation between AKT1 phosphorylation status and substrate selectivity in conditions of endoplasmic reticulum (ER) stress. Upon ER stress, the ER chaperone protein GRP78 is down-regulated and this leads to increased phosphorylation of AKT1 at Ser-473 but not Thr-308 in human cells (JEG3). As a result, increased AKT1-dependent phosphorylation of MDM2 (at Ser-166), FOXO1 (at Ser-319), and GSK-3-β (at Ser-9) were observed by Western blotting (42). In our study, we found the additional of pSer-473 significantly increases activity with the GSK-3β peptide (Fig. 3). The FOXO1 319 site is one of the most active with $pAKT1^{S473}$ (Fig. S1).

Our data clearly indicate that ppAKT1 is normally (but not always) more active and has a broader substrate profile than either of the singly phosphorylated enzymes. In alignment with the essentiality of pThr-308 in the activation of AKT1 *in vitro* and in cells (8, 40), we found that the two variants phosphorylated at Thr-308 showed greatest similarities in substrate preferences compared with the $pAKT1^{S473}$ enzyme. Previous work (10) found that Ser-473 was not essential for phosphorylating the non-N-terminal phosphosites of the FOXO transcription factors, TSC2 or GSK-3β. These observations are in close agreement with and supportive of our findings. The $ppAKT1^{T308,S473}$ enzyme showed high activity with FOXO1 C-terminal phosphorylation sites (Ser-319, Ser-256), TSC2 and GSK-3β (Fig. S1, Fig. 3). For $pAKT1^{T308}$, all three FOXO1 phosphosites ranked among the highest active substrates, suggesting that Ser-473 is not a critical factor in FOXO1 phosphorylation. In the clinic, phosphorylation of Thr-308 was associated with poor cell survival with nonsmall cell lung cancer (43) and acute myeloid leukemia (44). According to our data here and previously (8), we suggest that phosphorylation of Thr-308 alone may be capable of switching-on downstream and disease-linked signaling to promote cell survival by inhibiting the FOXO group of transcription factors.

Phosphorylation dependent changes in the AKT1 substrate motif

As observed for AKT (13) and many other kinases (45–47), the molecular basis for kinase substrate selectivity can depend critically on the residues neighboring the phosphorylation site. Yet it is important to understand that not all substrates identified for AKT isozymes include the same exact consensus motif

(10). In the context of the cell, even with an optimal substrate according to the consensus motif requirements, phosphorylation by AKT1 will depend on multiple factors, including accessibility, cellular localization, substrate expression level, scaffolding interactions, and phosphatase activity (11). The binding affinity between AKT1 and full-length protein substrates may also contribute to the degree of substrate phosphorylation.

Our data obtained from the OPAL1 study for ppAKT1^{T308,S473} are in accord with but extended beyond previously reported results (12, 13). According to our results, all AKT1 phospho-forms showed a preference for Arg at position −5. ppAKT1^{T308,S473} and pAKT1^{T308} showed a preference for hydrophobic amino acids for position +1. These amino acid preferences are shared by other AGC family kinases including protein kinase C. Obata *et al.* (13) used degenerative peptide library studies to determine the most relevant substrate motif for active AKT1. In this previous work, active AKT1 was derived from Sf9 insect cells and the impact of AKT1 phosphorylation status on substrate selectivity could not be probed. The authors reported the amino acid preferences derived from a set of known substrates and a degenerative peptide library. Because the peptide design was identical between Obata *et al.* (13) and our OPAL2 library, the results are similar. We also observed a strong correlation with the previous results and the OPAL1 at the positions −5 and +1 with pAKT1^{T308} and ppAKT1^{T308,S473}. In Obata *et al.* (13), the most preferred amino acid at the −5 position was Arg and preferred amino acids at +1 were large hydrophobic residues. Also, both pAKT1^{T308} and ppAKT1^{T308,S473} enzyme variants showed preference for Phe at position −6 in agreement with the previous study (13). At the +2 position, as before (13), we also observed a preference for amino acids that enable tight turns, especially for pThr-308 containing enzymes: Gly_2 in OPAL1, and Ser_2 or Thr_2 in OPAL2. In addition, Ser and Thr at +2 were the favorable amino acids according to our known library data (Figs. S1 and S2). Finally, in previous work, Pro at +1 inhibited AKT1 activity similarly to p70 S6 kinase (12). We also observed significantly reduced activity in all OPAL experiments for Pro at +1. The +1 position provides some level of orthogonality between AKT1 and proline-directed kinases, such as MAP kinases and cyclin-dependent kinases that select substrates with Pro at +1 (48).

Expanding the AKT1-dependent phospho-proteome

Validation of predictions based on our biochemical data using cellular or ultimately animal models is warranted in future studies beyond the scope of this work. Our initial investigations, based on independent database searches, found that AKT1 may be involved in regulating RNA metabolism (Fig. 7*B*, Fig. S11). In a recent report published with our collaborators, we validated one of these predicted AKT1 targets, the terminal nucleotidyltransferase Gld2 (49). Our data suggests the pAKT1^{T308} enzyme would phosphorylate Gld2 at Ser-116 (p value $\sim 10^{-4}$) (Data File S2). Gld2 has been implicated in the stabilization and maturation of the tumor suppressor microRNAs miR-122 and let-7 (50, 51). We found Gld2 was indeed specifically phosphorylated by AKT1 at Ser-116, which abolished Gld2 activity in 3′-terminal addition of adenine to miRNA and mRNA substrates (49). As hyperactivity of AKT1 is common in many cancers (52), our results suggest that phosphorylation of Gld2 by AKT1 would be beneficial for cancer cells by decreasing the levels of miRNAs that act as tumor suppressors (49). Our data, thus, suggest a novel link between RNA regulation and oncogenic signaling by AKT1.

Conclusion

The phosphorylation status of AKT1 is a commonly used biomarker for human cancer (43). Here we found that AKT1 phosphorylation status has a global impact on substrate selectivity. Fascinatingly, ppAKT1^{T308,S473} showed higher activity with less than half of the known substrates we tested. We determined that the addition of Ser-473 phosphorylation did not lead to a significant change in activity for nearly half of the substrates tested, and a minority of substrates was significantly more active with pAKT1^{T308} than with ppAKT1. The pAKT1^{S473} variant had a distinct substrate preference profile compared with AKT1 variants phosphorylated at Thr-308. Because our data show the impact of Ser-473 phosphorylation on AKT1 activity is substrate-dependent, we conclude pSer-473 may tune substrate selectivity rather than truly activate the kinase. Indeed, whether AKT1 with only Ser-473 phosphorylated is active in the cell or what its true substrates might be remains unknown (8, 40). Our findings have major implications for the use of AKT1 phosphorylation status as a clinical diagnostic (43, 53) or a marker for cancer biology (52).

Materials and methods

Bacterial strains and plasmids

A codon optimized human *AKT1* gene lacking the PH domain (residues 109–480, 45 kDa) was synthesized at ATUM (Newark, CA, United States) previously (9) and subcloned, as follows: (NcoI/NotI) into an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible T7*lac* promoter driven pCDF-Duet vector with CloDF13-derived replicon and streptomycin/spectinomycin resistance (pCDF-Duet1-ΔPHAKT1). The human *PDPK1* gene, which was acquired from the Harvard PlasmidID repository service (HsCD00001584, Boston, MA, United States), was subcloned (KpnI/NdeI) into the second multiple cloning site of pCDFDuet-1. The codon for the residue in *AKT1* at position Ser-473 was mutated to the amber (TAG) codon in the synthetic gene from ATUM. Successful cloning was verified by DNA sequencing (London Regional Genomics Centre, London, ON, Canada, and Genewiz, Cambridge, MA, United States).

Protein and phosphoprotein production

Phosphorylated forms of AKT1 (pAKT1^{S473}, pAKT1^{T308}, and ppAKT1^{T308,S473}) were produced according to previously described methods (8, 9).

Library of known AKT1 substrate peptides

Based on reports from the literature collated at the Phosphosite database (6), we selected 84 different peptides to generate a library of known AKT1 substrates (Data File S1). The substrate motifs derived from these AKT1 substrates were synthesized in

the following form: biotin-A$X_{-6}X_{-5}X_{-4}X_{-3}X_{-2}X_{-1}$-$S_0$/$T_0$-$X_1X_2X_3X_4X_5X_6$YRR. For this library, X does not indicate randomization, rather the nature of the X residues is defined by the known sequence of AKT1 substrates. The phosphorylation site (Ser or Thr) is located at the 0 position. Each individual in-solution peptide was synthesized according to previously reported methods (54). For Fig. S1, multiple statistical tests were used to identify a group of peptide substrates that were not significantly different from the background measurements determined in experiments containing only kinase enzyme and buffer but lacking substrate. The statistical significance between enzyme activities was calculated using one-way analysis of variance in Minitab statistical analysis software. The level of significance was adjusted based on Bonferroni correction.

OPAL AKT1 activity screen

An oriented peptide array library (OPAL1) was designed as a series of soluble sublibrary pools according to previously reported methods (15). The known consensus sequence for AKT1 phosphorylation was used as the design for library synthesis: biotin-AGGG$X_{-6}X_{-5}X_{-4}R_{-3}X_{-2}X_{-1}S_0X_1X_2X_3$. In each of the 136 sublibraries, one of the X positions is held fixed to a specific amino acid. X represents any amino acid other than Ser, Thr, or Cys. The phosphorylatable residues Ser and Thr were excluded from the screen to avoid false-positive signals when probing position S_0. Each of the 17 other canonical amino acids were tested in the X positions. The −3 and 0 positions were fixed to Arg and Ser, respectively, where the 0 position is the phosphorylation site and Arg_{-3} is important for substrate phosphorylation by AKT1 (12). The total library contains $\sim 10^{10}$ different peptides. Accurate synthesis of the library was confirmed (Fig. S3) using a series of test peptides analyzed by MALDI-TOF MS as before (55).

The independent OPAL2 was based on a peptide design of YA$X_{-5}X_{-4}X_{-3}X_{-2}X_{-1}$-$S_0$/$T_0$-$X_1X_2X_3X_4$GKK-biotin and was synthesized and spotted exactly as described previously (56). AKT1 protein was produced from Sf9 cells and kinase activity assays were performed and processed as before (56); reaction products were spotted onto avidin-coated filter sheets (Promega SAM^2 biotin capture membrane) and visualized by phosphorimaging. PCA analysis was conducted to compare the OPAL matrices for each AKT1-phosphform. These were computed in Matlab (MathWorks, Natick, MA, United States) using custom scripts and default settings.

Kinase assays with peptides and libraries

Kinase assays with peptides or OPAL1 sublibraries were conducted using radiolabeled [γ-^{32}P]ATP according to previously reported methods (8) with modifications as following: AKT1 activity was determined using 400 μM substrate peptide for each peptide in the known substrate library. Peptides sublibraries were dissolved in 10 μl of dimethyl sulfoxide (DMSO) to a final total peptide concentration of 1 mM. Each reaction was conducted in kinase assay buffer (25 mM MOPS, pH 7.0), 12.5 mM β-glycerolphosphate, 25 mM $MgCl_2$, 5 mM EGTA, pH 8.0, 2 mM EDTA, 20 μM ATP, and 0.4 μCi of (33 nM) [γ-^{32}P]ATP). The indicated AKT1 variant was added to a final concentration of 6.5 nM to initiate the reaction. Each condition was assayed using 3 independent enzyme reaction replicates. Reactions were spotted and quantified as before (8).

Sequence logos and database searching

Sequence logos corresponding to substrate preferences for each AKT1 variant were generated using WebLogo version 2.8.2 (57) or Seq2Logo 2.0 (28) as indicated. Using the activity data over our known substrate library (Fig. S1), we generated an alignment of peptides in which the population of each peptide was linearly related to its activity level. The resulting peptide alignment was, thus, most populated by the highest active peptides and least populated by the low active peptides. This allowed us to convert the activity data over known peptides into a sequence logo using WebLogo (Fig. S2). The sequence logos based on the OPAL data were statistically analyzed by comparing the overall average enzyme activity against the enzyme activities of all tested amino acids at each position (*i.e.* −6 to +3) using pairwise *t* test analysis. Thus, activity that is not significantly different from the average activity is highlighted in *red* (Fig. 5).

Using the OPAL1 and OPAL2 data we generated PSSMs according to a similar method used in Scansite (29). For each data set, a selectivity matrix was calculated by dividing all activity values by the average value for the complete array, *i.e.* a selectivity factor. A scoring matrix was computed as the natural log of the selectivity matrix. Finally, the raw scoring matrices were scaled by a factor of 10 multiplied by the selectivity factor for each data set. The resulting PSSMs based on each OPAL data set were used to produce sequence logos with the program Seg2Logo 2.0 (28), where the PSSM-Logos were computed using default values. We then used the program PSSMSearch (30) to identify potential AKT1 substrates based on PSSMs created with the OPAL1 data.

As an alternate approach to generate a PSSM, we also converted the OPAL1 activity data to a set of Z-score matrices as before (58). Z-scores were calculated with the formula $Z_{i,j} = (x_{i,j} - \mu)/s$, *i* refers to the *i*th peptide sublibrary, and *j* refers to the *j*th randomized position in the target sequence. There is then a Z-score ($Z_{i,j}$) associated with each amino acid at each position in the consensus motif. The score is calculated by subtracting the average (μ) kinase activity over the library from the activity observed ($x_{i,j}$) normalized by the standard deviation (s) of activity values over the entire library. The preference rank of any potential substrate can then be calculated as a score based on the sum of Z-scores at each position in the test peptide. The known human phospho-proteome was scored based on the amino acid preferences at each position in the Z-score PSSM (Data File S3).

Data availability

All of the data described in Figs. 1–7 and Tables 1 and 2 as well as in supporting information Figs. S1–S13, Table S1, and Data Files S1–S3 are available on-line.

Acknowledgments—We are grateful to Ilka Heinemann for critical discussions and suggestions.

Author contributions—N. B., N. E. D., J. L. J., H. L., K. K. B., L. C. C., S. S.-C. L., and P. O. conceptualization; N. B., N. E. D., J. L. J., H. L., K. K. B., and P. O. data curation; N. B., N. E. D., J. L. J., K. K. B., and P. O. formal analysis; N. B. validation; N. B., N. E. D., J. L. J., H. L., K. K. B., L. C. C., S. S.-C. L., and P. O. investigation; N. B., N. E. D., and P. O. visualization; N. B., N. E. D., J. L. J., H. L., K. K. B., and P. O. methodology; N. B., H. L., and P. O. writing-original draft; N. B., L. C. C., S. S.-C. L., and P. O. project administration; N. B., N. E. D., J. L. J., K. K. B., L. C. C., S. S.-C. L., and P. O. writing-review and editing; N. E. D. and P. O. software; L. C. C., S. S.-C. L., and P. O. resources; L. C. C., S. S.-C. L., and P. O. supervision; L. C. C., S. S.-C. L., and P. O. funding acquisition.

Funding and additional information—This work was supported by the Canadian Institutes of Health Research Grant 165985 (to P. O.), Natural Sciences and Engineering Research Council of Canada Grant 04282 (to P. O.), Canadian Cancer Society Research Institute innovation Grant 704324 (to P. O. and S. S.-C. L.), Canada Foundation for Innovation Grant 229917 (to P. O.), Ontario Research Fund Grant 229917 (to P. O.), Canada Research Chairs 229917 and 232341 (to P. O.), and the Canadian Breast Cancer Foundation (to S. S.-C. L.).

Conflict of interest—With the following exception, there are no other conflicts to declare. L. C. C. is a founder and member of the BOD and SAB of Agios Pharmaceuticals; he is also a co-founder, member of the SAB, and shareholder of Petra Pharmaceuticals. These companies are developing novel therapies for cancer. L. C. C. laboratory receives some funding support from Petra Pharmaceuticals. J. L. J. reports consultant activities for Petra Pharmaceuticals.

Abbreviations—The abbreviations used are: PKB, protein kinase B; PI3K, phosphatidylinositol 3-kinase; GSK-3β, glycogen synthase kinase 3β; OPAL, oriented peptide array library; PH, pleckstrin homology; IRAK1, interleukin-1 receptor-associated kinase 1; TSC2, TSC complex subunit 2; PSSM, position-specific score matrices; PCA, principle component analysis; ROC, receiver operator characteristic; AUC, area under the curve; STRING, Search Tool for Retrieval of Interacting Genes/Protein; GO, gene ontology; ER, endoplasmic reticulum; PDK1, phosphoinositide-dependent kinase-1.

References

1. Gonzalez, E., and McGraw, T. E. (2009) The AKT kinases: isoform specificity in metabolism and cancer. *Cell Cycle* **8,** 2502–2508 CrossRef Medline
2. Martini, M., De Santis, M. C., Braccini, L., Gulluni, F., and Hirsch, E. (2014) PI3K/AKT signaling pathway and cancer: an updated review. *Ann. Med.* **46,** 372–383 CrossRef Medline
3. Ma, C. X., Sanchez, C., Gao, F., Crowder, R., Naughton, M., Pluard, T., Creekmore, A., Guo, Z., Hoog, J., Lockhart, A. C., Doyle, A., Erlichman, C., and Ellis, M. J. (2016) A phase I study of the AKT inhibitor MK-2206 in combination with hormonal therapy in postmenopausal women with estrogen receptor-positive metastatic breast cancer. *Clin. Cancer Res.* **22,** 2650–2658 CrossRef Medline
4. Dumble, M., Crouthamel, M. C., Zhang, S. Y., Schaber, M., Levy, D., Robell, K., Liu, Q., Figueroa, D. J., Minthorn, E. A., Seefeld, M. A., Rouse, M. B., Rabindran, S. K., Heerding, D. A., and Kumar, R. (2014) Discovery of novel AKT inhibitors with enhanced anti-tumor effects in combination with the MEK inhibitor. *PLoS ONE* **9,** e100880 CrossRef Medline
5. Zhang, X., Tang, N., Hadden, T. J., and Rishi, A. K. (2011) AKT, FoxO and regulation of apoptosis. *Biochim. Biophys. Acta* **1813,** 1978–1986 CrossRef Medline
6. Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43,** D512–520 CrossRef Medline
7. Hers, I., Vincent, E. E., and Tavaré, J. M. (2011) AKT signalling in health and disease. *Cell Signal.* **23,** 1515–1527 CrossRef Medline
8. Balasuriya, N., Kunkel, M. T., Liu, X., Biggar, K. K., Li, S. S., Newton, A. C., and O'Donoghue, P. (2018) Genetic code expansion and live cell imaging reveal that Thr-308 phosphorylation is irreplaceable and sufficient for AKT1 activity. *J. Biol. Chem.* **293,** 10744–10756 CrossRef Medline
9. Balasuriya, N., McKenna, M., Liu, X., Li, S. S. C., and O'Donoghue, P. (2018) Phosphorylation-dependent Inhibition of AKT1. *Genes (Basel)* **9,** 450 CrossRef
10. Manning, B. D., and Cantley, L. C. (2007) AKT/PKB signaling: navigating downstream. *Cell* **129,** 1261–1274 CrossRef Medline
11. Manning, B. D., and Toker, A. (2017) AKT/PKB signaling: navigating the network. *Cell* **169,** 381–405 CrossRef Medline
12. Alessi, D. R., Caudwell, F. B., Andjelkovic, M., Hemmings, B. A., and Cohen, P. (1996) Molecular basis for the substrate specificity of protein kinase B: comparison with MAPKAP kinase-1 and p70 S6 kinase. *FEBS Lett.* **399,** 333–338 CrossRef Medline
13. Obata, T., Yaffe, M. B., Leparc, G. G., Piro, E. T., Maegawa, H., Kashiwagi, A., Kikkawa, R., and Cantley, L. C. (2000) Peptide and protein library screening defines optimal substrate motifs for AKT/PKB. *J. Biol. Chem.* **275,** 36108–36115 CrossRef Medline
14. Fabbro, D., Batt, D., Rose, P., Schacher, B., Roberts, T. M., and Ferrari, S. (1999) Homogeneous purification of human recombinant GST-AKT/PKB from Sf9 cells. *Protein Expr. Purif.* **17,** 83–88 CrossRef Medline
15. Rodriguez, M., Li, S. S., Harper, J. W., and Songyang, Z. (2004) An oriented peptide array library (OPAL) strategy to study protein-protein interactions. *J. Biol. Chem.* **279,** 8802–8807 CrossRef Medline
16. Sarbassov, D. D., Guertin, D. A., Ali, S. M., and Sabatini, D. M. (2005) Phosphorylation and regulation of AKT/PKB by the rictor-mTOR complex. *Science* **307,** 1098–1101 CrossRef Medline
17. Zhang, X., Zhang, S., Yamane, H., Wahl, R., Ali, A., Lofgren, J. A., and Kendall, R. L. (2006) Kinetic mechanism of AKT/PKB enzyme family. *J. Biol. Chem.* **281,** 13949–13956 CrossRef Medline
18. Bonni, A., Brunet, A., West, A. E., Datta, S. R., Takasu, M. A., and Greenberg, M. E. (1999) Cell survival promoted by the Ras-MAPK signaling pathway by transcription-dependent and -independent mechanisms. *Science* **286,** 1358–1362 CrossRef Medline
19. Maiti, D., Bhattacharyya, A., and Basu, J. (2001) Lipoarabinomannan from *Mycobacterium tuberculosis* promotes macrophage survival by phosphorylating Bad through a phosphatidylinositol 3-kinase/AKT pathway. *J. Biol. Chem.* **276,** 329–333 CrossRef Medline
20. Michell, B. J., Harris, M. B., Chen, Z. P., Ju, H., Venema, V. J., Blackstone, M. A., Huang, W., Venema, R. C., and Kemp, B. E. (2002) Identification of regulatory sites of phosphorylation of the bovine endothelial nitric-oxide synthase at serine 617 and serine 635. *J. Biol. Chem.* **277,** 42344–42351 CrossRef Medline
21. Huang, H., Regan, K. M., Wang, F., Wang, D., Smith, D. I., van Deursen, J. M., and Tindall, D. J. (2005) Skp2 inhibits FOXO1 in tumor suppression through ubiquitin-mediated degradation. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 1649–1654 CrossRef Medline
22. Guan, K. L., Figueroa, C., Brtva, T. R., Zhu, T., Taylor, J., Barber, T. D., and Vojtek, A. B. (2000) Negative regulation of the serine/threonine kinase B-Raf by AKT. *J. Biol. Chem.* **275,** 27354–27359 Medline
23. Giamas, G., Hirner, H., Shoshiashvili, L., Grothey, A., Gessert, S., Kühl, M., Henne-Bruns, D., Vorgias, C. E., and Knippschild, U. (2007) Phosphorylation of CK1δ: identification of Ser370 as the major phosphorylation site targeted by PKA *in vitro* and *in vivo*. *Biochem. J.* **406,** 389–398 CrossRef Medline
24. Yeh, P. Y., Kuo, S. H., Yeh, K. H., Chuang, S. E., Hsu, C. H., Chang, W. C., Lin, H. I., Gao, M., and Cheng, A. L. (2006) A pathway for tumor necrosis factor-α-induced Bcl10 nuclear translocation. Bcl10 is up-regulated by NF-κB and phosphorylated by AKT1 and then complexes with Bcl3 to enter the nucleus. *J. Biol. Chem.* **281,** 167–175 CrossRef Medline
25. Cheng, J., Hamilton, K. S., and Kane, L. P. (2014) Phosphorylation of Carma1, but not Bcl10, by AKT regulates TCR/CD28-mediated NF-kap-

paB induction and cytokine production. *Mol. Immunol.* **59,** 110–116 CrossRef Medline
26. Chen, B. C., Wu, W. T., Ho, F. M., and Lin, W. W. (2002) Inhibition of interleukin-1beta-induced NF-kappa B activation by calcium/calmodulin-dependent protein kinase kinase occurs through AKT activation associated with interleukin-1 receptor-associated kinase phosphorylation and uncoupling of MyD88. *J. Biol. Chem.* **277,** 24169–24179 CrossRef Medline
27. Rena, G., Guo, S., Cichy, S. C., Unterman, T. G., and Cohen, P. (1999) Phosphorylation of the transcription factor Forkhead family member FKHR by protein kinase B. *J. Biol. Chem.* **274,** 17179–17183 CrossRef Medline
28. Thomsen, M. C., and Nielsen, M. (2012) Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Res.* **40,** W281–W287 CrossRef Medline
29. Obenauer, J. C., Cantley, L. C., and Yaffe, M. B. (2003) Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res.* **31,** 3635–3641 CrossRef Medline
30. Krystkowiak, I., Manguy, J., and Davey, N. E. (2018) PSSMSearch: a server for modeling, visualization, proteome-wide discovery and annotation of protein motif specificity determinants. *Nucleic Acids Res.* **46,** W235–W241 CrossRef Medline
31. Dinkel, H., Chica, C., Via, A., Gould, C. M., Jensen, L. J., Gibson, T. J., and Diella, F. (2011) Phospho.ELM: a database of phosphorylation sites: update 2011. *Nucleic Acids Res.* **39,** D261–D267 CrossRef Medline
32. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., and von Mering, C. (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43,** D447–D452 CrossRef Medline
33. Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., and Thomas, P. D. (2017) PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **45,** D183–D189 CrossRef Medline
34. Jacinto, E., Facchinetti, V., Liu, D., Soto, N., Wei, S., Jung, S. Y., Huang, Q., Qin, J., and Su, B. (2006) SIN1/MIP1 maintains rictor-mTOR complex integrity and regulates AKT phosphorylation and substrate specificity. *Cell* **127,** 125–137 CrossRef Medline
35. Soares, C. D., Borges, C. F., Sena-Filho, M., Almeida, O. P., Stelini, R. F., Cintra, M. L., Graner, E., Zecchin, K. G., and Jorge, J. (2017) Prognostic significance of cyclooxygenase 2 and phosphorylated AKT1 overexpression in primary nonmetastatic and metastatic cutaneous melanomas. *Melanoma Res.* **27,** 448–456 CrossRef Medline
36. Jin, Y., Yuan, Y., Yi, M., Han, H., Liu, B., and Li, Q. (2019) Phosphorylated-AKT overexpression is associated with a higher risk of brain metastasis in patients with non-small cell lung cancer. *Biochem. Biophys. Rep.* **18,** 100625 Medline
37. Kim, H.A., Kim, J. H., Choi, H. S., Kim, Y. G., Lee, S. J., Han, K. H., Seong, M. K., Seol, H. S., and Noh, W. C. (2017) Core needle biopsy specimens are more appropriate than surgical specimens for evaluating the expression of phosphoproteins as biological markers in invasive breast cancer. *Anticancer Res.* **37,** 1409–1412 CrossRef Medline
38. Lee, R. S., House, C. M., Cristiano, B. E., Hannan, R. D., Pearson, R. B., and Hannan, K. M. (2011) Relative expression levels rather than specific activity plays the major role in determining *in vivo* AKT isoform substrate specificity. *Enzyme Res.* **2011,** 720985 Medline
39. Alessi, D. R., Andjelkovic, M., Caudwell, B., Cron, P., Morrice, N., Cohen, P., and Hemmings, B. A. (1996) Mechanism of activation of protein kinase B by insulin and IGF-1. *EMBO J.* **15,** 6541–6551 CrossRef Medline
40. Hart, J. R., and Vogt, P. K. (2011) Phosphorylation of AKT: a mutational analysis. *Oncotarget* **2,** 467–476 Medline
41. Shi, D., and Gu, W. (2012) Dual roles of MDM2 in the regulation of p53: ubiquitination dependent and ubiquitination independent mechanisms of MDM2 repression of p53 activity. *Genes Cancer* **3,** 240–248 CrossRef Medline
42. Yung, H. W., Charnock-Jones, D. S., and Burton, G. J. (2011) Regulation of AKT phosphorylation at Ser-473 and Thr-308 by endoplasmic reticulum stress modulates substrate specificity in a severity dependent manner. *PLoS ONE* **6,** e17894 CrossRef Medline
43. Vincent, E. E., Elder, D. J., Thomas, E. C., Phillips, L., Morgan, C., Pawade, J., Sohail, M., May, M. T., Hetzel, M. R., and Tavaré, J. M. (2011) AKT phosphorylation on Thr-308 but not on Ser-473 correlates with AKT protein kinase activity in human non-small cell lung cancer. *Br. J. Cancer* **104,** 1755–1761 CrossRef Medline
44. Gallay, N., Dos Santos, C., Cuzin, L., Bousquet, M., Simmonet Gouy, V., Chaussade, C., Attal, M., Payrastre, B., Demur, C., and Récher, C. (2009) The level of AKT phosphorylation on threonine 308 but not on serine 473 is associated with high-risk cytogenetics and predicts poor overall survival in acute myeloid leukaemia. *Leukemia* **23,** 1029–1038 CrossRef Medline
45. Pearson, R. B., and Kemp, B. E. (1991) Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. *Methods Enzymol.* **200,** 62–81 CrossRef Medline
46. Ehrenberger, T., Cantley, L. C., and Yaffe, M. B. (2015) Computational prediction of protein-protein interactions. *Methods Mol. Biol.* **1278,** 57–75 CrossRef Medline
47. Bradley, D., and Beltrao, P. (2019) Evolution of protein kinase substrate recognition at the active site. *PLos Biol.* **17,** e3000341 CrossRef Medline
48. Vulliet, P. R., Hall, F. L., Mitchell, J. P., and Hardie, D. G. (1989) Identification of a novel proline-directed serine/threonine protein kinase in rat pheochromocytoma. *J. Biol. Chem.* **264,** 16292–16298 Medline
49. Chung, C. Z., Balasuriya, N., Manni, E., Liu, X., Li, S. S., O'Donoghue, P., and Heinemann, I. U. (2019) Gld2 activity is regulated by phosphorylation in the N-terminal domain. *RNA Biol.* **16,** 1022–1033 CrossRef Medline
50. Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., Baba, T., and Suzuki, T. (2009) Selective stabilization of mammalian microRNAs by 3′ adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev.* **23,** 433–438 CrossRef Medline
51. Rissland, O. S., Mikulasova, A., and Norbury, C. J. (2007) Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell Biol.* **27,** 3612–3624 CrossRef Medline
52. Bellacosa, A., Kumar, C. C., Di Cristofano, A., and Testa, J. R. (2005) Activation of AKT kinases in cancer: implications for therapeutic targeting. *Adv. Cancer Res.* **94,** 29–86 CrossRef Medline
53. Freudlsperger, C., Horn, D., Weißfuß, S., Weichert, W., Weber, K. J., Saure, D., Sharma, S., Dyckhoff, G., Grabe, N., Plinkert, P., Hoffmann, J., Freier, K., and Hess, J. (2015) Phosphorylation of AKT(Ser-473) serves as an independent prognostic marker for radiosensitivity in advanced head and neck squamous cell carcinoma. *Int. J. Cancer* **136,** 2775–2785 CrossRef Medline
54. Wei, R., Kaneko, T., Liu, X., Liu, H., Li, L., Voss, C., Liu, E., He, N., and Li, S. S. (2018) Interactome mapping uncovers a general role for Numb in protein kinase regulation. *Mol. Cell Proteomics* **17,** 2216–2228 Medline
55. Wang, R., Leung, P. Y. M., Huang, F., Tang, Q., Kaneko, T., Huang, M., Li, Z., Li, S. S. C., Wang, Y., and Xia, J. (2018) Reverse binding mode of phosphotyrosine peptides with SH2 protein. *Biochemistry* **57,** 5257–5269 CrossRef Medline
56. Hutti, J. E., Jarrell, E. T., Chang, J. D., Abbott, D. W., Storz, P., Toker, A., Cantley, L. C., and Turk, B. E. (2004) A rapid method for determining protein kinase phosphorylation specificity. *Nat. Methods* **1,** 27–29 CrossRef Medline
57. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14,** 1188–1190 CrossRef Medline
58. Huang, H., Li, L., Wu, C., Schibli, D., Colwill, K., Ma, S., Li, C., Roy, P., Ho, K., Songyang, Z., Pawson, T., Gao, Y., and Li, S. S. (2008) Defining the specificity space of the human SRC homology 2 domain. *Mol. Cell Proteomics* **7,** 768–784 CrossRef Medline
59. Iacovides, D. C., Johnson, A. B., Wang, N., Boddapati, S., Korkola, J., and Gray, J. W. (2013) Identification and quantification of AKT isoforms and phosphoforms in breast cancer using a novel nanofluidic immunoassay. *Mol. Cell Proteomics* **12,** 3210–3220 CrossRef Medline