Research Article

Computational Analysis of the Binding Specificities of PH Domains

Zhi Jiang,[1,2] Zhongjie Liang,[1] Bairong Shen,[1] and Guang Hu[1]

[1]*Center for Systems Biology, Soochow University, Suzhou 215006, China*
[2]*Department of Biochemistry and Molecular Biology, School of Medicine, Soochow University, Suzhou 215123, China*

Correspondence should be addressed to Guang Hu; huguang@suda.edu.cn

Pleckstrin homology (PH) domains share low sequence identities but extremely conserved structures. They have been found in many proteins for cellular signal-dependent membrane targeting by binding inositol phosphates to perform different physiological functions. In order to understand the sequence-structure relationship and binding specificities of PH domains, quantum mechanical (QM) calculations and sequence-based combined with structure-based binding analysis were employed in our research. In the structural aspect, the binding specificities were shown to correlate with the hydropathy characteristics of PH domains and electrostatic properties of the bound inositol phosphates. By comparing these structure properties with sequence-based profiles of physicochemical properties, PH domains can be classified into four functional subgroups according to their binding specificities and affinities to inositol phosphates. The method not only provides a simple and practical paradigm to predict binding specificities for functional genomic research but also gives new insight into the understanding of the basis of diseases with respect to PH domain structures.

1. Introduction

PH (pleckstrin homology) domains, consisting of about 100–120 amino acid residues, are found in a wide range of proteins involved in intracellular signalling or as constituents of the cytoskeleton [1–4]. Although their sequences bear very low similarity, all the determined three-dimensional structures have seven β-strands forming two perpendicular anti-parallel-sheets and one C-terminal α-helix [5]. PH domains bind either plasma-membrane phosphoinositides or cytosolic inositol phosphates with few exceptions. The specific binding of PH domains to different inositol phosphates is important for the signal-dependent membrane targeting [6–8]. To date, the most studied inositol phosphates include inositol 1,3,4-triphosphate (Ins(1,3,4)-P_3), inositol (1,4,5)-trisphosphate (Ins(1,4,5)-P_3), and inositol 1,3,4,5-tetrakisphosphate (Ins(1,3,4,5)P_4). Some PH domains have been shown to interact also with protein kinase C or heterotrimeric G proteins [9, 10] in signalling pathways. In addition, PH domains are increasingly found to be connected to human disorders. For examples, Bruton tyrosine kinase

(Btk) is involved in X-linked agammaglobulinemia (XLA) [11, 12], an immunodeficiency. FGD1 protein and AKT1 are connected to Aarskog-Scott syndrome (ASS) [13] cancer [14, 15], respectively.

In the present work, we will focus on eleven well-known PH domains whose three-dimensional structures have been determined. Five out of these PH domains bind inositol phosphates with different specificities [16–21]. In addition to evolution, PH domains have been divided into four functional subclasses according to binding affinities and specificities [2] (Table 1). In summary, PH domains in Group 3 bind their preferred ligands with similar affinities to those in Group 1, whereas PH domains in Group 2 have 4–8-fold weaker binding affinities. In addition, PH domains in Group 4 have low affinity and less specificity. Concerning the three inositol phosphates, the affinity of Ins(1,4,5)P_3 to its preferred PH domains is generally lower than those of Ins(1,3,4,5)P_4 and Ins(1,3,4)P_3 [22, 23]. To understand the sequence-structure relationship of the four groups of PH domains, we have investigated a number of sequence profiles to characterize their physicochemical properties [24]. More

TABLE 1: Binding specificities of representative PH domains.

	Group 1	Group 2	Group 3	Group 4
$PI(3,4,5)P_3/Ins(1,3,4,5)P_4$	$+^a$	$+$	$+$	$-$
$PI(4,5)P_2/Ins(1,4,5)P_3$	$-$	$+$	$-$	$-$
$PI(3,4)P_2/Ins(1,3,4)P_3$	$-$	$-$	$+$	$-$
PH domain	Btk, Grp1, Gap1^m Gap1^{IP4BP}, Vav, cytohesin-1, Sos, ARNO, TIAM1-N	PLCδ1, βArk, β-spectrin, DAGKδ, RasGAp, OSBP, IRS-1, Plec-N	Dapp1, Akt, PDK1	Dynamin, TIAM1-C

^{a}+, − represent specific binding or nonspecific binding. Gap1^{IP4BP} represents one of the members of Ras GTPase-activating proteins and Gap1^m represents the mammalian counterpart of the *Drosophila* Gap1 gene.

recently, molecular dynamics (MD) simulations were performed to study structural and binding affinity of functional mutations of Btk PH domains [25] and the role of membrane penetration and electrostatics in the interaction between GRP1 PH domain and $PI(3,4,5)P_3$ [26, 27].

Despite the large body of PH domain literature, some problems are urgently to be resolved, such as the following: (1) Why the binding affinity of $Ins(1,4,5)P_3$ to its preferred PH domains is weaker than those of the other two inositol phosphates? (2) What factors determine the binding specificities of PH domains for different inositol phosphates? (3) Considering the fact that PH domains have low sequence identities but highly conserved structures, is there any intrinsic relationship between the binding specificities and sequence profiles of physicochemical properties? To these aims, a systematic comparison of the sequence-structure-function relationship is needed for further research of the functions of PH domains. In this paper, we first calculated the properties of both the inositol phosphates and PH domains, and then the binding specificities and affinities of PH domains were analyzed from both the structural and sequence aspects.

2. Materials and Methods

2.1. QM Calculations. The geometries of the three inositol phosphates, $Ins(1,3,4)P_3$, $Ins(1,4,5)P_3$, and $Ins(1,3,4,5)P_4$, were optimized at the density functional theory (DFT) of B3LYP/6-31+G(d,p) level [28] and the equilibrium structures were verified with calculations of frequencies. The single point energies and the electronic properties of the optimized inositol phosphates were calculated at the B3LYP level with the same basis sets. All the calculations were performed with Gaussian03 [29]. The initial conformations of $Ins(1,4,5)P_3$ and $Ins(1,3,4,5)P_4$ were extracted from PDB entries 1B55 and 1MAI, respectively. The initial conformation of $Ins(1,3,4)P_3$ was modified from $Ins(1,3,4,5)P_4$.

2.2. Sequence Analysis. Because the sequence similarities of PH domains are limited (average identity is 16%), the multiple sequence alignment could not be obtained with the general alignment programs. The sequence alignment of the selected PH domains was retrieved from the protein family (Pfam) database, which is created based on hidden Markov model [30]. The sequence alignment was further modified based on structure-based alignment and shown

in Figure 1. The PH domains were selected to represent the different functional subgroups including Btk (PDB code: 1BTK, 1B55), Grp1 (1FGZ, 1FGY, 1FHW, 1FHX), Plc-δ (1MAI), spectrin (1BTN, 1MPH, 1DRO), pleckstrin (1PLS), β-Ark (1BAK), Dapp1 (1FB8, 1FAO), dynamin (1DYN), and UNC-89 (1FHO). MEGA6 was used to construct phylogenetic tree of PH domains based on maximum composite likelihood method [31]. The profiles of physiochemical properties of PH domains, including flexibility, hydropathy, isotropic surface area, and electronic charge concentration, were calculated as previously described [24].

2.3. Structure and Binding Analysis. The electrostatic potentials of PH domains were calculated using a finite different solution to the nonlinear Poisson-Boltzmann equation [32]. The grid was 20 Å larger than the PH molecule containing 123 grid points in the longest dimension. The solute dielectric was set to 2. Solvent accessible surface areas (SASA) were calculated according to the algorithm of Lee and Richards [33], and a solvent radius of 1.4 Å was used for water. The fractions of residues exposed to solvent were calculated directly from the experimental structures and were subsequently used to generate profiles with the sliding window averaging technique to facilitate comparison to the predicted properties. In the case of NMR structures missing the average structure (1MPH, 1PLS, 1BAK, and 1FHO), the first structure in the entry was used. The detailed interatomic contacts for inositol phosphate to PH domains were investigated with the LIGIN program [34]. All the other structural analyses were performed with InsightII software of Accelrys, Inc.

3. Results and Discussion

Binding PH domains have been identified in various species [35, 36], and a few reports have discussed the evolution of TFKs including PH domains [37]. Evolutionary relationship among different groups of PH domains is of interest to be compared with binding specificities. The phylogenetic tree for 12 PH domains (Figure 2) was constructed based on the sequence alignment. Apparently, four groups are not clearly classified in the phylogenetic tree. For example, Btk, Plc-δ, and PDK1, belonging to three different groups, have nearly phylogenetic relationship. This result may be attributed to the low sequence identities but extremely conserved structures in PH domains. Therefore, the detailed analysis of their

```
              β1                    β2              β3                  β4        β5
Btk          2 AAVI..LESIFLKRSQQKKK...TSPLNFKKRLFLLT..VHKLSYYEYDFE..RGRRGSKKGSIDVEKITCVETVVPEKN
Grp1       261 TFFNPDREGWLLKLGGRVKT........WKRRWFILT..DNCLYYFEYTTD......KEPRGIIPLENL.SIREVEDPRK
Cytohesin  256 TFFNPDREGWLLKLGGRVK.......TWKRRWFILT..DNCLYYFEYTTD......KEPRGIIPLENL.SIREVDDPRK
Plcδ        20 QALL..KGSQLLKVKSS.........SWRRERFYKLQEDCKTIWQESRKVMRS.....PESQLFSIEDIQEVRMGHRTEG
Spectrin  2198 Q.....MEGFLNRKHEWEAHNKKASSRSWHNVYCVIN..NQEMGFYKDAKSAASGIPYHSEVPVSLKEA.ICEVALDYKK
Pleckstrin   3 PKRI..REGYLVKKGSVF........NTWKPMWVVLL..EDGIEFYKKKSDNS......PKGMIPLKGS.TLTSPCQDFG
βArk1      557 KDCI..MHGYMSKMGNPFLTQ......WQRRYFYLF..PNRLEWRGEGEAPQ.......SLLTMEEIQSVEETNIKER
Dapp1      161 APSLGTKEGYLTKQGGLV........KTWKTRWFTLH..RNELKYFKDQMSPEPI......RILDLTEC.SAVQFDYSQE
PDK1       450 QFV...ENNLILKMGPVDK....RKGLFARRQLLLTE.GPHLYYVDPVNKVLK.......GEIPWSQE.LRPEAKNF..
Akt/PKB1     5 AIV...KEGWLHKRGEYI........KTWRPRYFLLKNDGTFIGYKERPQD........VDRESPLNNF.SVAQCQLMKT
Dynamin    518 ILVI..RKGWLTINNIGIM......KGGSKEYWFVLT..AENLSWYKDDEEKE......KKYMLSVDNL.KLRDVEKGFM
UNC89      348 RII...RHDAFQVWEGDE.........PPKLRYVFLF..RNKIMFTEQDASTS.PPSYTHYSSIRLDKY.NIRQHTTDE.

                                 β6                β7            α1
Btk         PPPERQIPRRGEESSEMEQISIIERFPYPFQVVYDE..........GPLYVFSPT...EELRKRWIHQLKNVI 132 (Q06187)
Grp1        PNCFELYN..............PSHKGQVIKACKTEADGRVVEGNHVVYRISAPS...PEEKEEWMKSIKASI 379 (O08967)
Cytohesin   PNCFELY..............IPNNKGQLIKACKTEADGRVVEGNHMVYRISAPT...QEEKDEWIKSIQAAV 375 (Q99418)
Plcδ        LEKFAR................DIPEDRCFSIVFKDQR........NTLDLIAPS...PADAQHWVQGLRKII 129 (P10688)
Spectrin    KK.....................HVFKLRLSDG.........NEYLFQAKD...DEEMNTWIQAISSAI 230 (Q62261)
Pleckstrin  KRM.....................FVFKITTTKQ.........QDHFFQAAF...LEERDAWVRDINKAI 100 (P08567)
βArk1       KC.....................LLLKI..RGGK.......QFILQC.DS...DPELVQWKKELRDAY 651 (P25098)
Dapp1       RV.....................NCFCLVFPF.........RTFYLCAKT...GVEADEWIKILRWKL 258 (Q9UN19)
PDK1        ......................KTFFVHTP...........NRTYYLMDP...SGNAHKWCRKIQEVW 543 (O15530)
Akt/PKB1    ERPRP...................NTFIIRCLQWTTV......IERTFHVET...PEEREEWTTAIQTVA 107 (P31749)
Dynamin     SSK.....................HIFALFNTEQRN..VYKDYRQLELACET...QEEVDSWKASFLRAG 124 (Q05193)
UNC89       ......................DTIVLQPQEPGLP.....SFRIKPKDFETSEYVRKAWLRDIAEEQ 450 (Q17362)
```

Figure 1: Three-dimensional structure-based sequence alignment of PH domains by the method of hidden Markov model. The SwissPort accession numbers are shown at the end of the sequences.
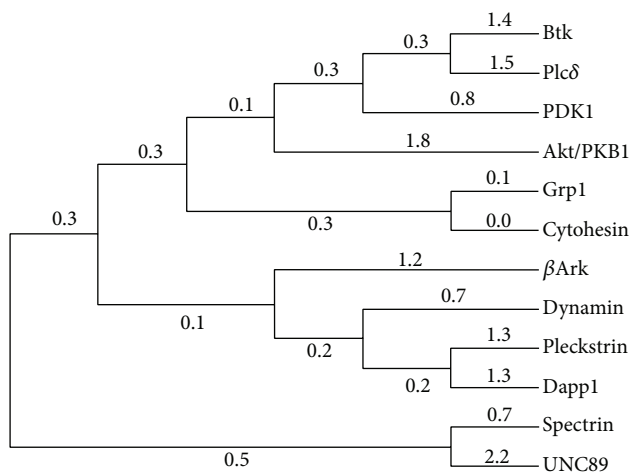

Figure 2: The phylogenetic tree for 12 PH domains. Bootstrap analysis was carried out using MCL approach, and bootstrap values are shown as scores of branches.

structural features is reasonably needed. We calculated and analyzed the binding specificity of PH domains (receptors) for inositol phosphates (ligands). First, the geometries and electronic properties of the three inositol phosphates were calculated. The electronic properties and hydropathy of the receptors were investigated with structure-based analysis. Then, based on the structural characteristics of PH domains and inositol phosphates, the binding affinities and specificities of PH domains to inositol phosphates were analyzed.

3.1. Geometries and Electronic Properties of Inositol Phosphates.

The electrostatic potentials of the three inositol phosphates were calculated by QM (Figure 3). The calculated single point energies, geometries, and electronic properties are listed in Table 2. In all the optimized structures, the *myo*-inositol ring adopted the conformation with 1-axial/5-equatorial oxygen positions (C2-hydroxyl in axial position and the other hydroxyls/phosphates in the equatorial orientation).

The electronic charge distribution is apparently different for $Ins(1,4,5)P_3$ and $Ins(1,3,4)P_3$ (or $Ins(1,3,4,5)P_4$) (Figure 3). In $Ins(1,4,5)P_3$, the negative charge (in red) is concentrated on one side of the molecule. The molecules were oriented based on the superimposition of their inositol carbon atoms. To compare the geometries and properties of $Ins(1,3,4)P_3$ and $Ins(1,4,5)P_3$, the $Ins(1,4,5)P_3$ was rotated by 180° to superimpose the phosphate groups of these two molecules. It has been reported that $Ins(1,3,4,5)P_4$ is bound to Btk in the opposite orientation compared to $Ins(1,4,5)P_3$ binding to Plcδ PH domain, although the interacting residues are in corresponding positions [3]. Figure 3 provides a qualitative explanation for the phenomenon, since the dipole moments of $Ins(1,4,5)P_3$ and $Ins(1,3,4,5)P_4$ point almost to the same direction when in inverted orientations.

The data in Table 2 confirmed the results of QM calculations. Although the chemical compositions of $Ins(1,3,4)P_3$ and $Ins(1,4,5)P_3$ are the same, the single point energies are different by 6.4 kcal/mol, indicating that $Ins(1,4,5)P_3$ conformation is more theoretically stable. The comparison of the geometries of the compounds provides an explanation
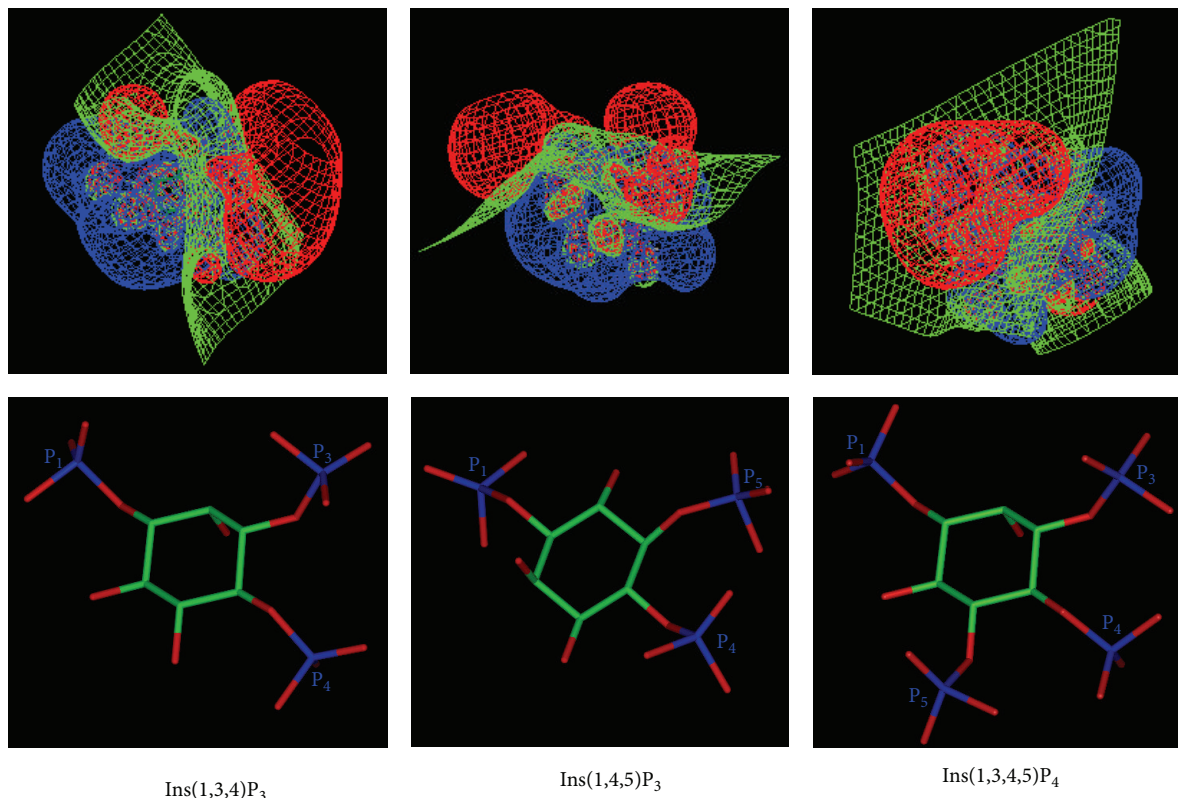
FIGURE 3: Comparison of the electrostatic potentials of three inositol phosphates. The contours of electrostatic potential at −5.0, 0.0, and 5.0 (kT/e) are coloured red, green, and blue, respectively.

TABLE 2: Comparison of the geometries and electronic properties of myo-inositol phosphates.

	$Ins(1,3,4)P_3$	$Ins(1,4,5)P_3$	$Ins(1,3,4,5)P_4$
Distance (O_2-O_{P_3}, Å)[a]	4.51; 3.89; 2.64	—	4.55; 3.40; 2.72
Distance (O_6-O_{P_5}, Å)[b]	—	4.57; 4.26; 2.66	4.62; 3.91; 2.84
Electronic spatial extent (Å^2)	10026	9868	13820
Dipole moment (Debye)	17.2	11.7	10.6
Energy (HF) (a.u.)	−2369.20	−2369.21	−2932.45
Energy (MP2) (a.u.)	−2371.76	−2371.77	−2951.77
ΔEnergy (MP2) (kcal/mol)[c]	6.4	0.0	—

[a] Distances between oxygen atom of 2-OH and three oxygen atoms of 3-PO_3 in $Ins(1,3,4)P_3$ or $Ins(1,3,4,5)P_4$.

[b] Distances between oxygen atom of 6-OH and three oxygen atoms of 5-PO_3 in $Ins(1,4,5)P_3$ or $Ins(1,3,4,5)P_4$.

There is no result for O_2-O_{P_3} distances in $Ins(1,4,5)P_3$ and O_6-O_{P_5} distance in $Ins(1,3,4)P_3$, because they do not contain 3-PO_3 and 5-PO_3 groups, respectively.

[c] The energy difference is calculated only for $Ins(1,3,4)P_3$ and $Ins(1,4,5)P_3$, because they belong to the same molecule but different conformations.

for the observation. The distances between the oxygen of 2-hydroxyl and the oxygen of 3-phosphate in $Ins(1,3,4)P_3$ are 4.51, 3.89, and 2.64 Å, respectively, while the corresponding distances in $Ins(1,4,5)P_3$ are 4.57, 4.26, and 2.66 Å, respectively. In $Ins(1,3,4)P_3$, the axial 2-hydroxyl is near negatively charged vicinal equatorial 3-phosphate and thus the repulsion between them is greater than in $Ins(1,4,5)P_3$. Both the dipole moment and the electronic spatial extent are greater in $Ins(1,3,4)P_3$ than in $Ins(1,4,5)P_3$.

The electronic spatial extent is a measure of molecular volume, whereas the dipole moment is an index of molecular polarizability. $Ins(1,3,4)P_3$ thus has wider electronic charge distribution and greater polarity than $Ins(1,4,5)P_3$. Since the electrostatic interaction is the main contributor for the interaction between PH domains and inositol phosphates. $Ins(1,4,5)P_3$, which has smaller electronic spatial extent, binding to its preferred PH domain is weaker compared with $Ins(1,3,4)P_3$. It is evident according to Figure 3 and Table 2 that the phosphate groups of $Ins(1,4,5)P_3$ are on one side of the molecule and the shape is much flatter than that for $Ins(1,3,4)P_3$. The different geometries and electronic properties between $Ins(1,3,4)P_3$ and $Ins(1,4,5)P_3$ contribute to the different specificities in PH domain interactions. The hydrophilic phosphate groups of inositol phosphates favour hydrophilic environment in the binding region of proteins. Among the three inositol phosphates, $Ins(1,3,4,5)P_4$ requires the most hydrophilic environment for binding, due to having four hydrophilic phosphate groups. For $Ins(1,4,5)P_3$, the binding environment in PH domain should be hydrophilic at one side of $Ins(1,4,5)P_3$. The binding environment for

TABLE 3: Comparison of ligand-protein contacts in 5 PH domain-inositol phosphate complexes.

| | | Contact surface area (Å^2) | | | Normalized complementarities |
		Legitimate contacts	Illegitimate contacts	Complementarities	
1B55	P_1-O_3	31.3	10.6	20.7	0.17
	P_3-O_3	85.0	16.0	69.0	0.65
	P_4-O_3	86.8	17.9	68.9	0.76
	P_5-O_3	81.6	19.7	61.9	0.57
1FGY	P_1-O_3	31.0	23.8	7.2	0.06
	P_3-O_3	89.6	21.8	67.8	0.61
	P_4-O_3	114.8	8.5	106.3	1.00
	P_5-O_3	104.5	11.1	93.4	0.80
1MAI	P_1-O_3	44.0	0.5	43.5	0.34
	P_4-O_3	83.0	17.8	65.2	0.55
	P_5-O_3	85.8	0.3	85.5	0.72
1BTN	P_1-O_3	33.6	0	33.6	0.27
	P_4-O_3	69.4	23.5	45.9	0.43
	P_5-O_3	72.6	0	72.6	0.68
1FAO	P_1-O_3	34.8	18.1	16.7	0.13
	P_3-O_3	82.9	23.1	59.8	0.55
	P_4-O_3	102.1	3.1	99.0	0.98
	P_5-O_3	45.0	20.9	24.1	0.20

$Ins(1,3,4)P_3$ is less hydrophilic than that for $Ins(1,4,5)P_3$, by virtue of the observation that the three phosphate groups in the molecule distribute separately and widely.

3.2. Electrostatic Properties and Hydropathy of PH Domains.

The calculated electrostatic potentials of PH domain structures (Figure 4) indicate that the binding sites for inositol phosphates are conserved and positively charged and thus electrostatic interactions play the key role in the binding of inositol phosphates. The localization and orientation of the inositol phosphate binding sites can be estimated by calculating the electrostatic properties of PH domains. On the other hand, hydrophobicity is also important for both the function and the stability of a protein. Hydropathy profiles may indicate functional sites [38]. The hydropathy analyses of PH domains are shown in Figure 4. Hydropathy environments of the PH domain binding regions are not as conserved as the electronic charge distributions. Hydropathy profiles can help to explain the different affinities and specificities in inositol phosphate binding. The binding environments are most hydrophilic for PH domains in group 1 (Figures 4(a) and 4(b)). The hydrophilic binding environments for PH domains in group 2 are on one side of the bound $Ins(1,4,5)P_3$ molecule (Figures 4(c) and 4(d)). The hydrophilic binding environments for PH domains in group 3 are less strict (Figure 4(e)). These observations are in agreement with the results obtained from the QM calculations of inositol phosphates.

We compared the structural profiles of electrostatic properties and hydropathy with results obtained by sequence profiles [24]. From both profiles, eight conserved extrema were found for structurally essential regions. For example, residues with smaller electronic charge correspond to the residues forming the hydrophobic core of PH domains, which may contribute to stabilizing the structure. According to the profiles of electronic charge concentration, the most charged segments are generally located in the $\beta 1/\beta 2$ and $\beta 7/\alpha 1$ loops. Indeed, the $\beta 1/\beta 2$ loop is positively charged and thus appears to be the most important segment for the binding of inositol phosphates.

3.3. Binding Specificities of PH Domains.

Table 3 lists the contact surface areas of phosphate groups in different PH domain-inositol phosphate complexes and the normalized complementarity (NC) function calculated by LIGIN program [34]. All the illegitimate contacts are of hydrophilic-hydrophobic type. In all cases, the C1-phosphate group has the lowest NC function indicating that the C1-phosphate group generally points outward and phosphoinositides can be replaced by inositol phosphates to study the binding specificities of PH domains. The binding affinities of inositol (tetra- and penta-) phosphates to Btk PH domain have the following order:

$$Ins(1,3,4,5)P_4 > Ins(1,3,4,5,6)P_5 \approx Ins(1,2,3,4,5,6)P_6 > Ins(3,4,5,6)P_4 > Ins(1,3,4,6)P_4 > Ins(1,2,5,6)P_4$$

By comparing with the experimental data of Btk PH domain [39], our computational results in Table 2 can be verified. In the structure of Btk PH domain-$Ins(1,3,4,5)P_4$ complex, C2 and C6 hydroxyls point to neutral and hydrophobic environment, respectively. Therefore, C2-phosphate has minor effect on the binding to the Btk PH domain, while the phosphate at position 6 inhibits the complex formation. The binding environment of C1-phosphate (NC = 0.17) is less hydrophilic than that for C3, C4, and C5-phosphates (NC = 0.65, 0.76, and 0.57, resp.), and consequently

(a) BtkPH-Ins(1,3,4,5)P$_4$

(b) Grp1PH-Ins(1,3,4,5)P$_4$

(c) PlcδPH-Ins(1,4,5)P$_4$

(d) SpecPH-Ins(1,4,5)P$_4$
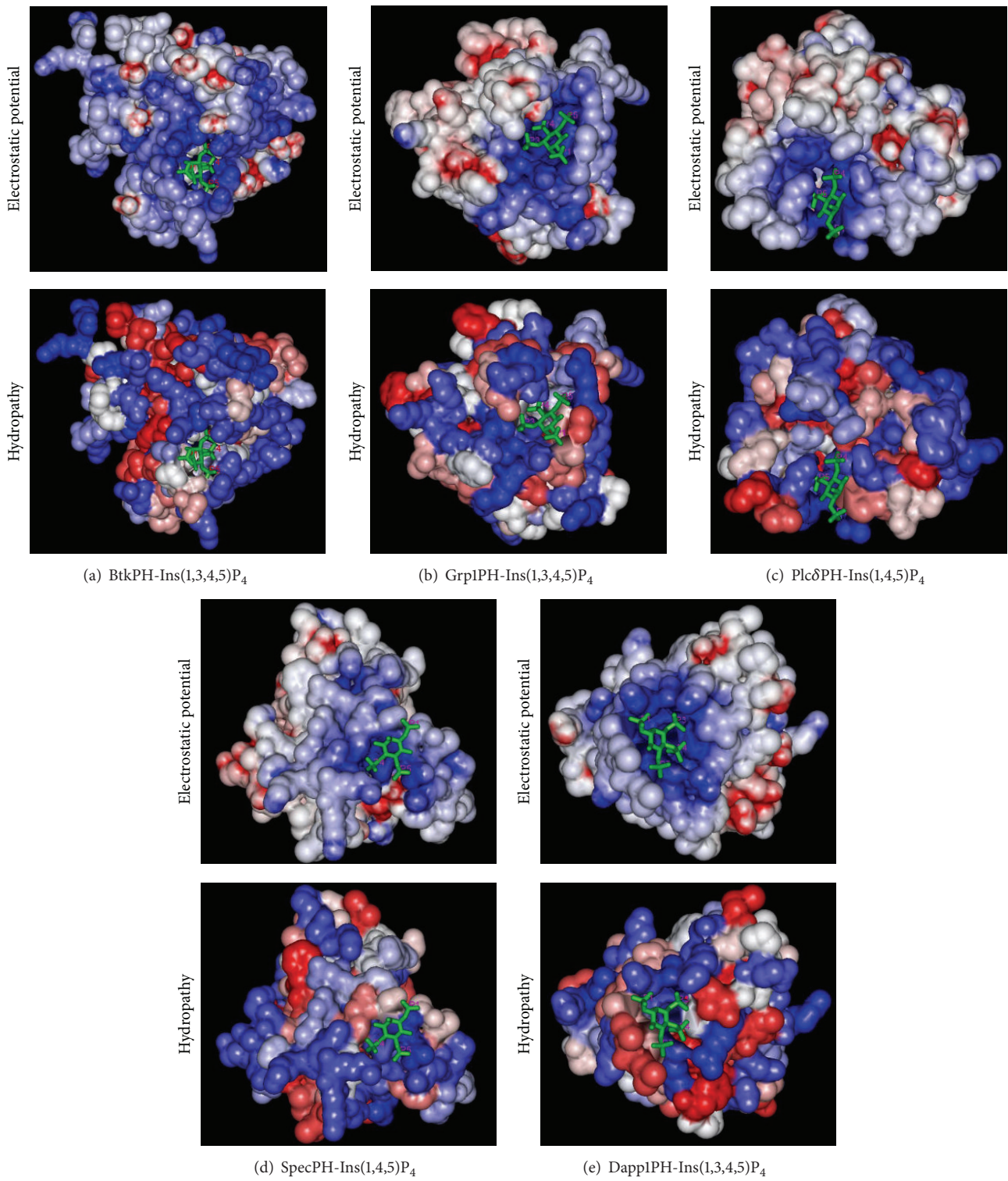
(e) Dapp1PH-Ins(1,3,4,5)P$_4$

FIGURE 4: Electrostatic and hydrophobic surface representations of PH domains. Top: molecular surfaces coloured by electrostatic potential, from red (−10 kT/e) to blue (+10 kT/e). Bottom: Hydropathic surface representations, hydrophobicity, and hydrophilicity are coloured from red to blue.

the contribution of C1-phosphate to the binding affinity is smaller. Hydropathy allows the prediction of the order of affinities of compounds. Here, the binding orientations of inositol (tetra- and penta-) phosphates binding to the Btk PH domain were presumed to be the same, analogous to the binding to the Grp1 PH domain. The recent crystal structure of Btk domain with phosphatidylinositol further identified a key residue located at $\beta1$-$\beta2$ loop for the binding [40]. The side chain of residues in $\beta1$-$\beta2$ loop form hydrogen bonds with multiple diacylglycerol groups of PtdIns(3,4,5)P_3. The binding specificities of $\beta1$-$\beta2$ loop are comparable with Ins(1,3,4,5)P_4, which will be discussed in the next section.

In the structure of the Dapp1 PH domain-Ins(1,3,4,5)P_4 complex the C6-hydroxyl is in a hydrophobic environment but it points outwards from the domain. It has only small effect on the affinity. The surroundings of C5-phosphate contain both hydrophobic and hydrophilic residues and thus C5-phosphate has minor effect on the interaction (NC = 0.20). In conclusion, the hydropathy analysis of the binding environment provides explanations for the experimentally obtained binding affinities as follows:

$$Ins(1,3,4,5,6)P_5 \cong Ins(1,3,4,5)P_4 > Ins(1,3,4,6)P_4 > Ins(1,4,5,6)P_4$$
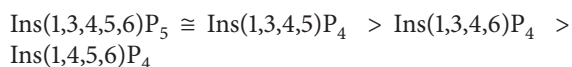
Contrary to the Plcδ domain (Figure 4(c)), the $\beta5$/$\beta6$ loop in the spectrin PH domain (Figure 4(d)) is more hydrophilic and more positively charged than the $\beta3$/$\beta4$ loop. Consequently, the inositol phosphate binds between $\beta1$/$\beta2$ and $\beta5$/$\beta6$ loops in the spectrin PH domain.

3.4. Analysis of Binding Affinity.

The positions of binding sites in SASA profiles of known PH domain structures are shown in Figure 5. The PH domain binding sites are generally hydrophilic, flexible, and charged. The charge concentration, hydrophilicity, and flexibility are the main factors, which determine the binding affinity and specificity. Since the $\beta1$/$\beta2$ loop is located between the $\beta3$/$\beta4$ and $\beta5$/$\beta6$ loops, the residues in this loop play crucial roles in inositol phosphate binding. In Figures 5(a) and 5(b), the binding sites in the $\beta1$/$\beta2$ loop of Group 1 PH domains (Btk and Grp1) are located in the region of high hydrophilicity, high flexibility, and high electronic charge concentration. Group 1 PH domains are specific and have high affinity for Ins(1,3,4,5)P_4. In addition, the Grp1 PH domain has high affinity to Ins(1,3,4,5)P_4 including the contribution of the $\beta6$/$\beta7$ loop. In the complex of Btk PH domain-Ins(1,3,4,5)P_4, the $\beta6$/$\beta7$ loop is not involved since it is hydrophobic.

The binding sites in the $\beta1$/$\beta2$ loop of Group 2 PH domains (Plc-δ and spectrin, Figures 5(c) and 5(d)) are also hydrophilic. The flexibility and electronic charge concentration are also high, but the electronic spatial extent and dipole moment of Ins(1,4,5)P_3 are relatively small. The hydrophilic phosphate groups are distributed on one side of the molecule. Therefore their affinities are reduced compared to Group 1 and Group 3. Compared to Group 2 PH domains, the binding environment of Group 3 PH domain is less hydrophilic (Figure 5(e)), because the phosphate groups of their binding ligands are distributed separately, even for Ins(1,3,4)P_3. The high binding affinity of the Dapp1 PH domain to Ins(1,3,4)P_3

could be related to the electronic properties of Ins(1,3,4)P_3. It can be seen that the Akt PH domain has similar profiles as Dapp1 [24, 41]. The Akt PH domain binds the Ins(1,3,4)P_3 to the loops $\beta1$/$\beta2$, $\beta3$/$\beta4$, and $\beta6$/$\beta7$. As for Group 4 PH domains, it may be identified by inspecting the profile of electronic charge concentration and hydropathy. The peaks in loop $\beta1$/$\beta2$ are generally low and the binding of Group 4 PH domains for inositol phosphates is less specific. Accordingly, the positively charged $\beta1$-$\beta2$ loop in all the structures of PH domains appears to be the most important segment for the binding of inositol phosphates. The inositol-binding affinities can thus been explained by the length of $\beta1$-$\beta2$ loops. The $\beta1$-$\beta2$ loops of the BTK PH domain (Figure 6(a)) and Plc-δ PH domain (Figure 6(c)) contain 11 and 9 residues, showing higher inositol-binding affinities. In comparison, the $\beta1$-$\beta2$ loop of Akt PH domain (Figure 6(b)) and dynamin PH domain (Figure 6(d)) are shorter, and thus they have lower inositol-binding affinities.

For PH domains without structure, the sequence profiles can pinpoint possible binding sites, guide experiments, and provide understanding of the sequence-function relationships. A signature motif for 3-phosphate binding has been suggested [23]; however it does not distinguish between Group 1 and Group 3 PH domains. With profile analysis, these two groups are distinguished, since the binding sites of Group 1 PH domains are generally more hydrophilic. Motif information should be combined with the profile analysis to predict binding specificities of PH domains. By mapping the signature motif for 3-phosphate binding, Group 1 and Group 3 PH domains can be distinguished from Groups 2 and 4. Then Group 1 and Group 3 PH domains are partitioned by hydropathy profile analysis. By analyzing the hydropathy and electronic charge concentration, it is possible to identify Group 4 PH domains. Since Group 4 PH domains have less specificity and lower binding affinity, the hydrophilicity of the sequence profile is weaker and the electronic charge concentration of the sequence profile is lower. Figure 7 gives an example of the prediction of the specificity of expressed sequence tag AA054961 PH domain, which bears the signature motif for 3-phosphate binding. Since the loop $\beta1$/$\beta2$ in this PH domain includes a notable hydrophobic peak, it is predicted to belong to Group 3.

4. Conclusions

The different binding affinities and specificities of PH domains to the three inositol phosphates of Ins(1,3,4)P_3, Ins(1,4,5)P_3, and Ins(1,3,4,5)P_4 were compared and explained from both the structural and sequence aspects. First, the electrostatics and geometric properties of the three inositol phosphates were calculated by a quantum mechanical method. Since the electronic charge distribution of the Ins(1,4,5)P_3 is smaller, its interaction with PH domains is generally weak. The phosphate groups in the Ins(1,4,5)P_3 are on one side of the molecule and the binding region is more hydrophilic on one side of the binding molecule than for Ins(1,3,4)P_3. Then, the structure-based electrostatic properties and hydropathy of PH domains profiles showed that hydrophobic environment is essential for the binding specificity. These structural results
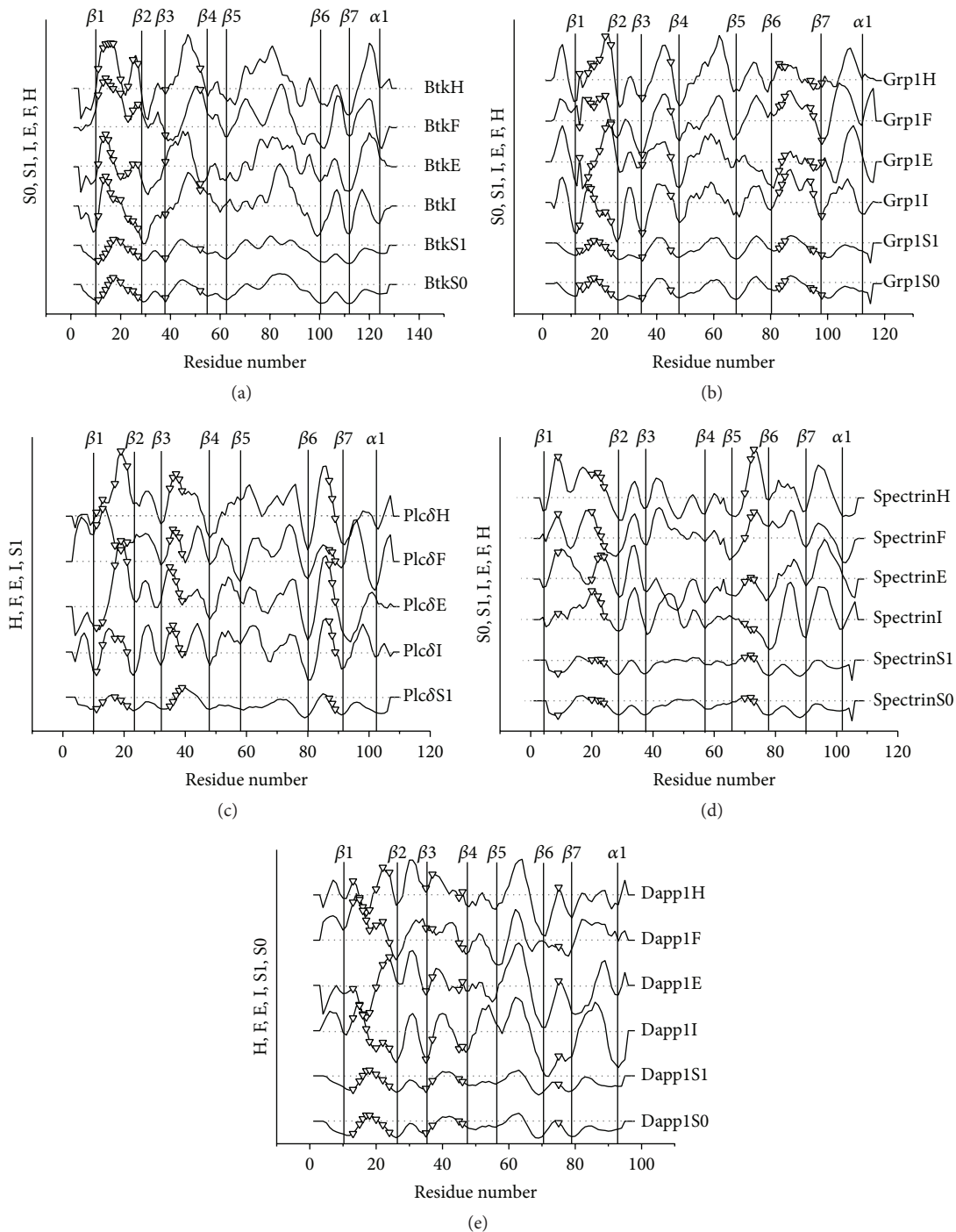
(a)

(b)

(c)

(d)

(e)

Figure 5: Ligand binding of the domains. The phosphatidyl inositol-binding residues are indicated by open triangles (∇). The insertions in the aligned sequences were deleted. The SASA for the PH domains and their complexes with inositol phosphates are marked with S0 and S1, respectively. H, F, E, and I represent hydropathy, flexibility, electronic charge concentration, and isotropic surface area, respectively.

are compared with sequence profiles for the analysis of binding specificity of PH domains, which also proved the essential role of hydrophobic environment for the binding specificity. The agreement of information from 1-dimensional sequence profiles and 3-dimensional structures provides a simple but practical method to investigate sequence-structure relationship of PH domains. The overall flowchart of our research

is summarized in Figure 8, which also contain two future directions.

PH domains can also be specifically identified and combined with signalling molecules, such as PTEN and PI3K [8]. It constitutes the basis for PH domains to participate in a variety of signalling pathways. Therefore, further understanding of the interaction between inositol phosphates
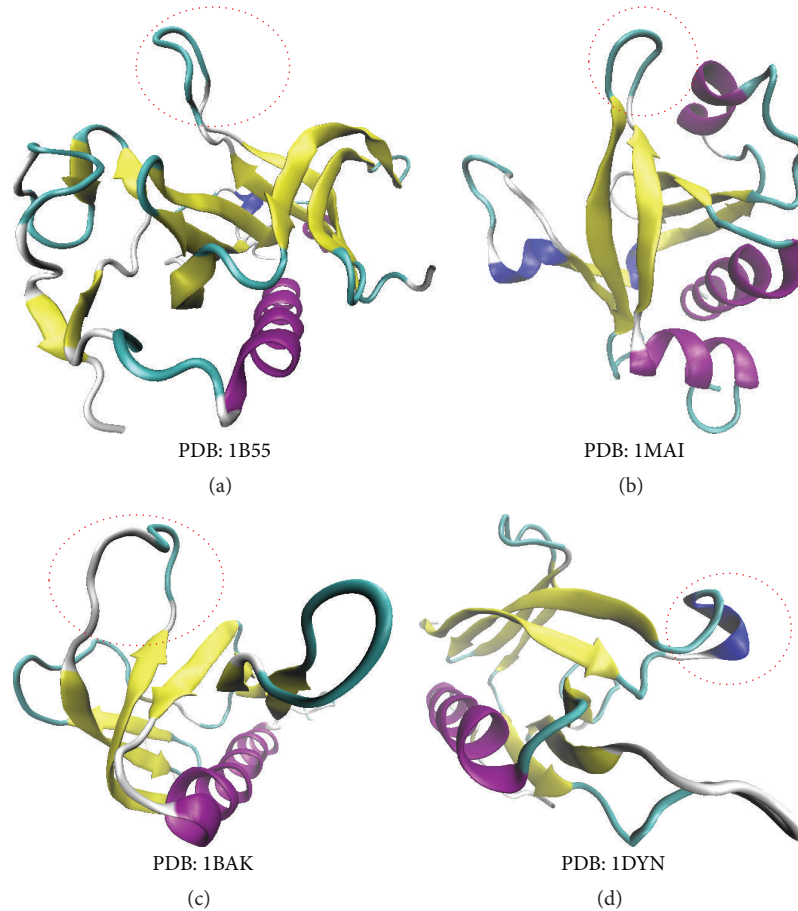
PDB: 1B55

(a)

PDB: 1MAI

(b)

PDB: 1BAK

(c)

PDB: 1DYN

(d)

FIGURE 6: The structures of $\beta 1$-$\beta 2$ loops (red circles) in four PH domains.
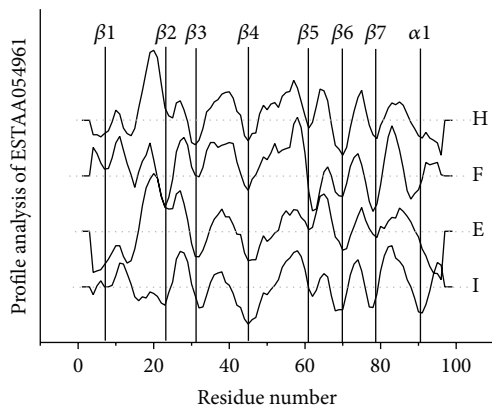
FIGURE 7: The profile analysis of ESTAA054961 PH domain.

FIGURE 8: The key mode for the classification of PH domain based on inositol-binding specificity, which is helpful to the analysis of PH domains mediated signalling pathways and disease-causing mutations.
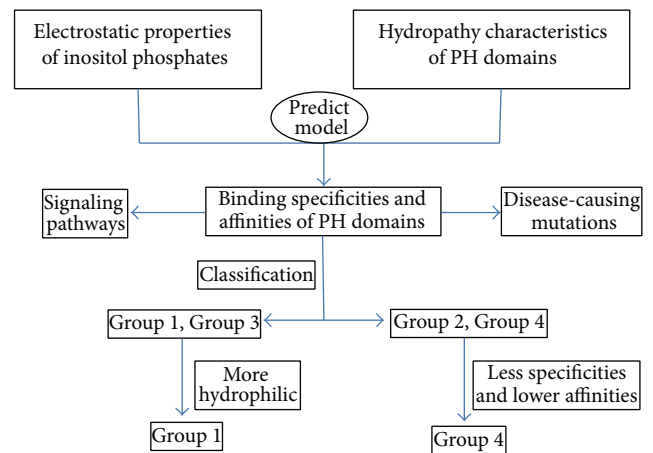
and its downstream molecules not only reveals a consistent picture of PH domain-mediated signal network system but also provides new insights into the mechanism of diseases with respect to these signalling pathways. Although MD simulation has been used to understand the interaction of PKB PH domain with inositol phosphates involved in the PI3K pathway [42], it remains a major challenge in the field.

In our previous work, the relationship between sequence profiles of binding sites and the effect of disease-causing PH domain mutation was analyzed, and MD method has been used to classify "folding mutation" and "disease-causing

mutation" [25]. However, the analysis and discussion of disease-causing variations affecting binding specificities and affinities pose another challenge. For example, Btk PH domain is the most studied PH domain which contains the highest number of unique disease-causing variations among the human protein kinases. The PON-BTK provides [43] a method for analyzing and classifying disease-causing mutations. With this mutation data, it is possible to reveal the basis of XLA by binding analysis. We hope that, in the new future, our method would be applied to PH domains to understand the basis of diseases with respect to inositol phosphates involving signalling pathways and harmful mutations for PH domains.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contribution

Zhi Jiang and Zhongjie Liang contributed equally to this paper.

Acknowledgments

References

[1] J. E. Harlan, P. J. Hajduk, H. S. Yoon, and S. W. Fesik, "Pleckstrin homology domains bind to phosphatidylinositol-4,5-bisphosphate," *Nature*, vol. 371, no. 6493, pp. 168–170, 1994.

[2] J. H. Hurley and S. Misra, "Signaling and subcellular targeting by membrane-binding domains," *Annual Review of Biophysics and Biomolecular Structure*, vol. 29, pp. 49–79, 2000.

[3] M. A. Lemmon and K. M. Ferguson, "Signal-dependent membrane targeting by pleckstrin homology (PH) domains," *Biochemical Journal*, vol. 350, no. 1, pp. 1–18, 2000.

[4] A. J. Mohamed, L. Yu, C.-M. Bäckesjö et al., "Bruton's tyrosine kinase (Btk): function, regulation, and transformation with special emphasis on the PH domain," *Immunological Reviews*, vol. 228, no. 1, pp. 58–73, 2009.

[5] J. Prashek, T. Truong, and X. Yao, "Crystal structure of the pleckstrin homology domain from the ceramide transfer protein: implications for conformational change upon ligand binding," *PLoS ONE*, vol. 8, no. 11, Article ID e79590, 2013.

[6] M. A. Lemmon, "Pleckstrin homology (PH) domains and phosphoinositides," *Biochemical Society Symposium*, vol. 74, pp. 81–93, 2007.

[7] Y. Yoon, "Small chemicals with inhibitory effects on PtdIns(3,4,5)P_3 binding of Btk PH domain," *Bioorganic & Medicinal Chemistry Letters*, vol. 24, no. 10, pp. 2334–2339, 2014.

[8] H. Li and A. J. Marshall, "Phosphatidylinositol (3,4) bisphosphate-specific phosphatases and effector proteins: a distinct branch of PI3K signaling," *Cellular Signalling*, vol. 27, no. 9, pp. 1789–1798, 2015.

[9] S. E. Boyken, D. B. Fulton, and A. H. Andreotti, "Rescue of the aggregation prone Itk Pleckstrin Homology domain by two mutations derived from the related kinases, Btk and Tec," *Protein Science*, vol. 21, no. 9, pp. 1288–1297, 2012.

[10] K. Scheffzek and S. Welti, "Pleckstrin homology (PH) like domains—versatile modules in protein-protein interaction platforms," *FEBS Letters*, vol. 586, no. 17, pp. 2662–2673, 2012.

[11] S. M. Tani, Y. Wang, H. Kanegane et al., "Identification of mutations of Bruton's tyrosine kinase gene (BTK) in Brazilian patients with X-linked agammaglobulinemia," *Human Mutation*, vol. 20, no. 3, pp. 235–236, 2002.

[12] S. Ponader and J. A. Burger, "Bruton's tyrosine kinase: from X-linked agammaglobulinemia toward targeted therapy for B-cell malignancies," *Journal of Clinical Oncology*, vol. 32, no. 17, pp. 1830–1839, 2014.

[13] T. Daubon, R. Buccione, and E. Génot, "The aarskog-scott syndrome protein Fgd1 regulates podosome formation and extracellular matrix remodeling in transforming growth factor β-stimulated aortic endothelial cells," *Molecular and Cellular Biology*, vol. 31, no. 22, pp. 4430–4441, 2011.

[14] J. D. Carpten, A. L. Faber, C. Horn et al., "A transforming mutation in the pleckstrin homology domain of AKT1 in cancer," *Nature*, vol. 448, no. 7152, pp. 439–444, 2007.

[15] C. Redaelli, F. Granucci, L. De Gioia, and L. Cipolla, "Synthesis and biological activity of Akt/PI3K inhibitors," *Mini-Reviews in Medicinal Chemistry*, vol. 6, no. 10, pp. 1127–1136, 2006.

[16] K. Saito, A. M. Scharenberg, and J.-P. Kinet, "Interaction between the Btk PH domain and phosphatidylinositol-3,4,5-trisphosphate directly regulates Btk," *The Journal of Biological Chemistry*, vol. 276, no. 19, pp. 16201–16206, 2001.

[17] K. A. Bethoney, M. C. King, J. E. Hinshaw, E. M. Ostap, and M. A. Lemmon, "A possible effector role for the pleckstrin homology (PH) domain of dynamin," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 32, pp. 13359–13364, 2009.

[18] H. Stocker, M. Andjelkovic, S. Oldham et al., "Living with lethal PIP3 levels: viability of flies lacking PTEN restored by a PH domain mutation in Akt/PKB," *Science*, vol. 295, no. 5562, pp. 2088–2091, 2002.

[19] D. B. van Rossum, R. L. Patterson, S. Sharma et al., "Phospholipase Cγ1 controls surface expression of TRPC3 through an intermolecular PH domain," *Nature*, vol. 434, no. 7029, pp. 99–104, 2005.

[20] E. J. McManus, B. J. Collins, P. R. Ashby et al., "The in vivo role of PtdIns(3,4,5)P-3 binding to PDK1 PH domain defined by knockin mutation," *The EMBO Journal*, vol. 23, no. 10, pp. 2071–2082, 2004.

[21] V. Calleja, M. Laguerre, P. J. Parker, and B. Larijani, "Role of a novel PH-kinase domain interface in PKB/Akt regulation: structural mechanism for allosteric inhibition," *PLoS Biology*, vol. 7, no. 1, Article ID e1000017, pp. 189–200, 2009.

[22] L. E. Rameh, A.-K. Arvidsson, K. L. Carraway III et al., "A comparative analysis of the phosphoinositide binding specificity of pleckstrin homology domains," *The Journal of Biological Chemistry*, vol. 272, no. 35, pp. 22059–22066, 1997.

[23] S. E. Lietzke, S. Bose, T. Cronin et al., "Structural basis of 3-phosphoinositide recognition by pleckstrin homology domains," *Molecular Cell*, vol. 6, no. 2, pp. 385–394, 2000.

[24] B. Shen and M. Vihinen, "Conservation and covariance in PH domain sequences: physicochemical profile and information theoretical analysis of XLA-causing mutations in the Btk PH domain," *Protein Engineering, Design & Selection*, vol. 17, no. 3, pp. 267–276, 2004.

[25] D. Lu, J. Jiang, Z. Liang et al., "Molecular dynamic simulation to explore the molecular basis of Btk-PH domain interaction with ins(1,3,4,5)P4," *The Scientific World Journal*, vol. 2013, Article ID 580456, 10 pages, 2013.

[26] C. N. Lumb and M. S. P. Sansom, "Finding a needle in a haystack: the role of electrostatics in target lipid recognition by PH domains," *PLoS Computational Biology*, vol. 8, no. 7, Article ID e1002617, 2012.

[27] C. N. Lumb, J. He, Y. Xue et al., "Biophysical and computational studies of membrane penetration by the GRP1 pleckstrin homology domain," *Structure*, vol. 19, no. 9, pp. 1338–1346, 2011.

[28] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic-behavior," *Physical Review A*, vol. 38, no. 6, pp. 3098–3100, 1988.

[29] M. J. Frisch, G. W. Trucks, H. B. Schlegel et al., *Gaussian 03, Revision C.02*, Gaussian Inc, Wallingford, Conn, USA, 2004.

[30] R. D. Finn, J. Clements, and S. R. Eddy, "HMMER web server: interactive sequence similarity searching," *Nucleic Acids Research*, vol. 39, supplement 2, pp. W29–W37, 2011.

[31] K. Tamura, G. Stecher, D. Peterson, A. Filipski, and S. Kumar, "MEGA6: molecular evolutionary genetics analysis version 6.0," *Molecular Biology and Evolution*, vol. 30, no. 12, pp. 2725–2729, 2013.

[32] B. Honig, K. Sharp, and A.-S. Yang, "Macroscopic models of aqueous solutions: biological and chemical applications," *Journal of Physical Chemistry*, vol. 97, no. 6, pp. 1101–1109, 1993.

[33] B. Lee and F. M. Richards, "The interpretation of protein structures: estimation of static accessibility," *Journal of Molecular Biology*, vol. 55, no. 3, pp. 379–400, 1971.

[34] V. Sobolev, A. Sorokine, J. Prilusky, E. E. Abola, and M. Edelman, "Automated analysis of interatomic contacts in proteins," *Bioinformatics*, vol. 15, no. 4, pp. 327–332, 1999.

[35] W. S. Park, W. Do Heo, J. H. Whalen et al., "Comprehensive identification of PIP3-regulated PH domains from *C elegans* to *H sapiens* by model prediction and live imaging," *Molecular Cell*, vol. 30, no. 3, pp. 381–392, 2008.

[36] Q. Xu, A. Bateman, R. D. Finn et al., "Bacterial pleckstrin homology domains: a prokaryotic origin for the PH domain," *Journal of Molecular Biology*, vol. 396, no. 1, pp. 31–46, 2010.

[37] C. Ortutay, B. F. Nore, M. Vihinen, and C. I. E. Smith, "Phylogeny of Tec family kinases: identification of a premetazoan origin of Btk, Bmx, Itk, Tec, Txk, and the Btk regulator SH3BP5," in *Advances in Genetics*, J. C. Hall, J. C. Dunlap, and T. Friedmann, Eds., vol. 64, pp. 51–80, Elsevier Academic Press, San Diego, Calif, USA, 2008.

[38] X. Gallet, B. Charloteaux, A. Thomas, and R. Brasseur, "A fast method to predict protein interaction sites from sequences," *Journal of Molecular Biology*, vol. 302, no. 4, pp. 917–926, 2000.

[39] S. Pfeiffer, D. Fushman, and D. Cowburn, "Impact of Cl$^-$ and Na$^+$ ions on simulated structure and dynamics of βARK1 PH domain," *Proteins: Structure, Function, and Bioinformatics*, vol. 35, no. 2, pp. 206–217, 1999.

[40] K. Murayama, M. Kato-Murayama, C. Mishima et al., "Crystal structure of the Bruton's tyrosine kinase PH domain with phosphatidylinositol," *Biochemical and Biophysical Research Communications*, vol. 377, no. 1, pp. 23–28, 2008.

[41] S.-B. Rong, Y. H. Hu, I. Enyedy et al., "Molecular modeling studies of the Akt PH domain and its interaction with phosphoinositides," *Journal of Medicinal Chemistry*, vol. 44, no. 6, pp. 898–908, 2001.

[42] S. A. J. Rosen, P. R. J. Gaffney, B. Spiess, and I. R. Gould, "Understanding the relative affinity and specificity of the pleckstrin homology domain of protein kinase B for inositol phosphates," *Physical Chemistry Chemical Physics*, vol. 14, no. 2, pp. 929–936, 2012.

[43] J. Väliaho, I. Faisal, C. Ortutay, C. I. Smith, and M. Vihinen, "Characterization of all possible single-nucleotide change caused amino acid substitutions in the kinase domain of Bruton tyrosine kinase," *Human Mutation*, vol. 36, no. 6, pp. 638–647, 2015.