Minireview

PTB or not PTB – that is the question

Kelley S. Yan[1], Miklos Kuti[1], Ming-Ming Zhou*

Structural Biology Program, Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York University, 1425 Madison Avenue, P.O. Box 1677, New York, NY 10029-6574, USA

Received 10 November 2001; revised 11 December 2001; accepted 17 December 2001

First published online 7 January 2002

Edited by Gianni Cesareni and Mario Gimona

Abstract Phosphotyrosine binding (PTB) domains are structurally conserved modules found in proteins involved in numerous biological processes including signaling through cell-surface receptors and protein trafficking. While their original discovery is attributed to the recognition of phosphotyrosine in the context of NPXpY sequences – a function distinct from that of the classical *src* homology 2 (SH2) domain – recent studies show that these protein modules have much broader ligand binding specificities. These studies highlight the functional diversity of the PTB domain family as generalized protein interaction domains, and reinforce the concept that evolutionary changes of structural elements around the ligand binding site on a conserved structural core may endow these protein modules with the structural plasticity necessary for functional versatility.

Key words: Phosphotyrosine binding domain; Three-dimensional structure; Fibroblast growth factor receptor substrate 2; Insulin receptor substrate-1; Numb; Shc; *Suc1*-associated neurotrophic factor target; X11

1. Introduction

The phosphotyrosine binding (PTB) domain, also known as the phosphotyrosine interaction domain, is a modular interface of ~100–170 amino acids important for protein–protein interaction [1–4]. First independently identified in the signaling proteins Shc [5–7] and insulin receptor substrate 1 (IRS-1) [8], this domain was discovered as an alternative to the *src* homology 2 (SH2) domain for phosphotyrosine recognition. Specifically, the PTB domains of Shc and IRS-1 preferentially bind to phosphorylated proteins containing an NPXpY motif (where pY is phosphotyrosine and X is any amino acid), with hydrophobic amino acids N-terminal to this sequence conferring additional specificity [8–13]. Recent studies show that PTB domain-like protein modules can also bind to proteins independent of tyrosine phosphorylation or even the canonical NPXY motif. For example, the PTB domains of the neuron-specific X11 protein and Fe65 both bind to a peptide derived from a region of β-amyloid precursor protein (β-APP) containing a non-phosphorylated NPTY sequence [14,15]. The *Drosophila* Numb PTB domain recognizes non-NPXY sequences, including GFSNMSFEDFP in a Ser/Thr protein kinase, Nak [16,17], and a GPY motif, identified through screening of a tyrosine-oriented synthetic peptide library [18,19]. In addition, the PTB domains of Shc [20] and mammalian protein Disabled (Dab) [21] can also interact with phospholipids. The enormous functional versatility of the PTB domain family is further highlighted by the PTB domain of membrane-anchored adapter proteins [22–24], SNT-1/2 (*suc1*-associated neurotrophic factor target; also known as fibroblast growth factor receptor substrate 2α/β) that has been shown to be capable of interacting with unrelated sequences from two different receptors, i.e. a tyrosine-phosphorylated NPXpY motif in tyrosine receptor kinases (TRKs) [25,26] and a non-phosphorylated segment in fibroblast growth factor receptor (FGFRs) that contains no tyrosine or asparagine residues [27,28].

Unlike SH2 domains, PTB domains share low sequence homology amongst themselves and exhibit extremely high ligand binding selectivity – the IRS-1 PTB domain does not bind to the NPXpY peptides derived from known Shc binding sites in the TRKs and epidermal growth factor receptors [12,29]. Despite their differences in amino acid sequence and ligand binding specificities, these PTB domains adopt a remarkably similar scaffold that is identical to that of the pleckstrin homology (PH) domain, which is known to bind to acidic phospholipids and localize proteins to the plasma membrane [30–32]. Recent structural studies on PTB domains reveal how a conserved protein modular domain has evolved to acquire the structural plasticity necessary to recognize a diverse set of biological ligands. In this minireview, we examine the unique structural features of PTB domains and their modes of ligand recognition.

2. A PTB/PH domain superfold

The three-dimensional structures of five PTB domains determined so far, in the proteins Shc [33], IRS-1 [34–36], X11 [37], Numb [17,19], and SNT-1 [38], show that all PTB domains adopt the structurally conserved PH domain-like fold [30,31], as highlighted in blue in the Shc PTB domain (Fig. 1).

*Corresponding author. Fax: (1)-212-849 2456.
E-mail address: zhoum@inka.mssm.edu (M.-M. Zhou).

[1] These authors contributed equally to this work.

Abbreviations: β-APP, β-amyloid precursor protein; FGFR1, fibroblast growth factor receptor 1; IR, insulin receptor; IRS, insulin receptor substrate; IL-4R, interleukin-4 receptor; PH, pleckstrin homology; PTB, phosphotyrosine binding; SH2, *src* homology 2; SNT, *suc1*-associated neurotrophic factor target; TRK, tyrosine receptor kinase

0014-5793/02/$22.00
PII: S0014-5793(01)03305-1

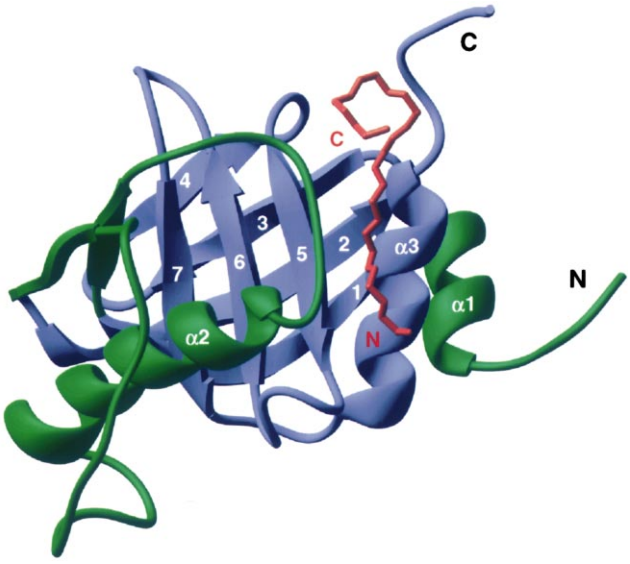

Fig. 1. Three-dimensional structure of the Shc PTB domain in complex with a tyrosine phosphorylated peptide derived from a sequence in TRKA containing pY490. The core PTB domain fold is shown in blue while the additional structural elements are in green. The phosphopeptide is depicted in red.

This PH domain fold consists of a β-sandwich containing two nearly orthogonal, antiparallel β-sheets capped at one end by a C-terminal α-helix (α3 in Shc). Despite the divergent sequences of the peptide ligands (see below), the major ligand binding site is located similarly in these different PTB domains, as illustrated by an elongated cleft formed by β5 and α3 in Shc. Although this highly conserved superfold is also seen in other functionally distinct protein interaction modules [39], which include the PH domain that commonly binds to phospholipids [31,32,40–43], the EVH1 domain (enabled/VASP homology 1) that binds proline-rich sequences [44,45], the Ran binding domain that interacts with Ran-GTPase [46], and lobe F3 of the FERM domain (band *f*our-point-one, *e*zrin, *r*adixin, *m*oesin homology) that functions in protein localization to the plasma membrane [47,48], the location of this ligand binding site between β5 and the C-terminal α helix (α3) is unique to the PTB domain family. Differences in ligand specificity among PTB domains are due to variations of structural features that lie outside the conserved secondary structure elements that comprise the basic scaffold.

3. NPXpY recognition

The structures of the Shc and IRS-1 PTB domains in complex with high-affinity peptides reveal the structural basis of NPXpY motif recognition. The structure of the Shc PTB domain in complex with a phosphopeptide derived from the nerve growth factor receptor TRKA (HIIENPQpYFSDA) was solved by nuclear magnetic resonance (NMR) [33], whereas the IRS-1 PTB domain structures were determined by NMR [34] and X-ray crystallography [35] when bound to an interleukin-4 receptor (IL-4R) peptide (LVIAGNPApYRS) or to a modified insulin receptor (IR) peptide (LYASSNPApY, where E(pY−1) in the wild type sequence was changed to an alanine for higher-affinity binding), respectively. Despite the lack of sequence homology between Shc and IRS-1, the PTB domains share a common fold, and the conformation and orientation of the phosphopeptides in the complexes are similar. The residues NPQpY in TRKA or NPApY in IR and IL-4R phosphopeptides form type I β-turns. Unlike SH2 domains [49,50], remarkably, arginine residues that interact directly with the phosphotyrosine on the protein surface are conserved only in three-dimensional space but not in the sequences of the Shc and IRS-1 PTB domains. The side chain atoms of N(pY−3) form specific intra- and intermolecular hydrogen bonds to the peptide residue at the (pY−1) position and to protein residues, explaining its crucial structural role in high-affinity binding to PTB domains [9–11]. The N-termini of both phosphopeptides adopt an extended conformation and meld into the second β-sheet of the protein through anti-parallel β-sheet interactions with β5.

Two divergent mechanisms are employed by the PTB domains to recognize hydrophobic residues N-terminal to the NPXpY motif. In Shc, L(pY−5) of TRKA peptide binds in a deep hydrophobic pocket formed between β5 and α3 [33], whereas in IRS-1, hydrophobic residues at the (pY−6) and (pY−8) positions in the IL-4R or IR phosphopeptides form extensive interactions with hydrophobic sites of the protein located on the surface of the second β-sheet [34,35]. The analogous site in Shc is covered by a loop and the N-terminal portion of an α-helix (α2), making it unavailable for ligand binding.

Additional differences in ligand selectivity between these two PTB domains are further attributed to the differences in ligand recognition. For example, the IRS-1 PTB domain favors a small hydrophobic amino acid such as alanine at the (pY−1) position for high-affinity binding – when A(pY−1) in IL-4R is substituted by a glutamate, a 30-fold loss in binding to IRS-1 is observed [12,29,35]. This result can be explained by the PTB domain structure, in which A(pY−1) interacts with a hydrophobic patch composed of three methionines, i.e. M209, M257 and M260 [34,35]. An alanine to glutamate mutation would result in an unfavorable interaction with the three methionines. It is interesting to note that IL-4R contains the optimal NPApY sequence (at pY497) for high-affinity ligand binding to the IRS-1 PTB domain, whereas the known IRS-1 binding site in IR at pY960 contains a non-optimal NPEpY motif. Indeed, it has been shown that while highly specific, the IRS-1 PTB domain binding to the phosphorylated NPEpY motif in IR ($K_D \approx 85$ μM) is significantly weaker than that to the NPApY sequence in IL-4R ($K_D = 1.8$ μM) [29]. These striking results emphasize the notion that biological interactions of high specificity but relatively low affinity may be just as important as those with high specificity and high affinity, allowing for additional regulation in signal transduction.

4. Non-NPXpY binding

Structural studies of the PTB domains from X11, Numb and SNT-1 have provided further insights into how the conserved PH domain-like fold can recognize non-phosphorylated NPXY sequences or even sequences lacking the NPXY motif. The PTB domain of X11 binds to a non-phosphorylated peptide (QNGYENPTYKFFEQ) derived from β-APP with high affinity ($K_D = 0.32$ μM) [14,37]. The crystal structure of the complex shows that the peptide, bound between β5 and the C-terminal helix (α2), forms an antiparallel β-strand with a β-turn (at the NPTY), followed by a 3^{10} helix at the C-ter-

minus (consisting of the KFFE residues) [37]. While Y(Y−5) does interact with the protein residues, most interestingly, the two aromatic residues F(Y+2) and F(Y+3) in the 3^{10} helix have extensive contacts with the protein residues in α2, which is longer by three turns as compared to the structurally analogous α3 in Shc. These interactions shown in the structure explain a 10-fold reduction in binding affinity when either of these phenylalanine residues is mutated to alanine [37]. This structure shows that the lack of coordination of phosphotyrosine by arginine residues, as seen in the Shc and IRS-1 PTB domains, can be compensated for by hydrophobic interactions with the surface on the X11 PTB domain for additional interactions to form a stable complex with the non-phosphorylated NPTY sequence of β-APP. Thus, tyrosine phosphorylation, in the case of β-APP, does not enhance its affinity for the X11 PTB domain [37].

The solution structure of the *Drosophila* Numb PTB domain solved in complex with a peptide (GFSNMSFEDFP) from the C-terminal tail of a serine/threonine protein kinase, Nak, has extended our understanding of how this PTB domain recognizes a ligand lacking the NPXY consensus sequence [16,17]. The structure reveals that the N-terminal residues (GFS) of the Nak peptide adopt an extended conformation to bind between β5 and the C-terminal α helix (α3), and the NMSFEDF sequence forms two consecutive β-turns [17]. While Numb recognition of Nak residues F(F−5) and those in the first β-turn (NMSF) is similar to that of the X11 PTB domain recognition of β-APP, binding of the C-terminal β-turn residues (FEDF) involves protein residues located in the β4/β5 and β6/β7 loops. This general mode of interaction may be used to extrapolate our understanding of the structural mechanism of the Numb PTB domain binding to an NPAY sequence of the protein Lnx, another reported biological target of Numb [51]. The structural versatility of the Numb PTB domain in its ligand binding cleft between β5 and C-terminal α helix is further illustrated by its high-affinity association with a helical turn-forming peptide (AYIGPpYL), identified through screening of a synthetic peptide library [18,19]. This unique mode of interaction with the GPpY sequence remains to be demonstrated as a biologically relevant function for the Numb PTB domain.

The SNT-1 PTB domain serves to further underscore the idea that the structural plasticity of the conserved PTB domain fold results in enhanced functional versatility [22–24]. This single PTB domain is capable of binding not only to the canonical NPXpY motif in TRKs [23–26], but also to a juxtamembrane region in FGFR containing no tyrosine or asparagine residues [27,28]. The recent solution structure of the SNT-1 PTB domain/human FGFR1 peptide (HSQMAVHKLAKSIPLRRQVTVS) complex shows that the peptide wraps around the protein with an unusual backbone conformation containing two nearly 90° turns that are oriented orthogonal to each other (Fig. 2) [38]. The C-terminal QVTVS segment of the peptide adopts an antiparallel β-strand sandwiched between β5 and β8, forming an intermolecular β-sheet. This unique β8, extended from the C-terminal α helix (α1), is not found in other PTB domains. On the other side of the β-sandwich, the peptide embeds its N-terminal MAVH segment into a large hydrophobic cavity bounded by the three loops connecting β1/β2, β3/β4 and β6/β7. Moreover, the LAKSIPL sequence located in the center of the peptide binds to otherwise solvent-exposed hydrophobic residues on the surface of the second β-sheet. In addition to hydrophobic interactions, complementary electrostatic interactions are observed, largely localized at the two turns in the peptide. This extensive network of intermolecular interactions confers the high specificity to the association between SNT-1 and FGFR1. It is not yet clear how SNT-1 interacts with the tyrosine-phosphorylated NPXpY sequence of TRKs (ΨXNPXpY, where Ψ is a bulky hydrophobic amino acid). Because the conserved arginine residues (R63 and R78) that are essential for the important phosphotyrosine binding activity are structurally analogous to those in the IRS-1 PTB domain, and the requirement for a large hydrophobic residue at the (pY−5) position for high-affinity association, binding of the NPXpY peptide likely involves regions including β5 and α1 in SNT-1 [38,52].

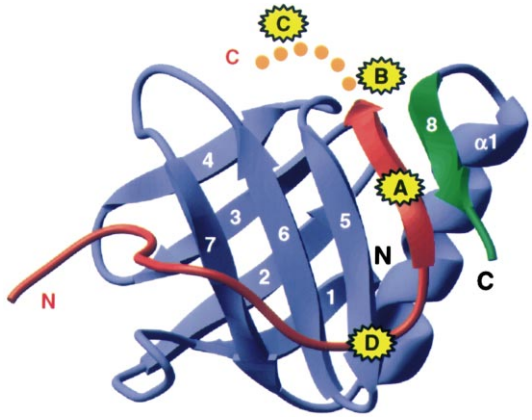

Fig. 2. Three-dimensional structure of the SNT-1 PTB domain in complex with a peptide derived from the juxtamembrane region of human FGFR1. The core PTB domain fold is illustrated in blue, additional structural elements in green and the peptide in red. Orange dots indicate the location of residues C-terminal to the β-strand of a putative PTB domain ligand. Highlighted with yellow stars are the general features of ligand recognition by the PTB domain fold: (A) the β-strand conformation of a ligand that melds into the second β-sheet of the PTB domain through antiparallel β-sheet interactions with β5; (B) an NPX(p)Y or related sequence with type I β-turn propensity; (C) the C-terminal specificity of a ligand achieved by phosphotyrosine-dependent or -independent interactions; and (D) the N-terminal specificity of a ligand conferred by hydrophobic residue(s) up to (pY−5) position. Note that the structure of the SNT-1 PTB domain/hFGFR1 peptide complex is used here to illustrate the general location of the major contact points for ligand recognition by PTB domains, and not all these contacts are necessary for hFGFR1 recognition by the SNT-1 PTB domain.

5. Unifying features?

The term 'PTB' is an inaccurate representation of the activities of this family of protein modules, but unifying features do exist. Despite the structural variations and the differences in ligand specificities, four major contact points appear to be crucial for PTB domain/ligand association. The location of these contact points is highlighted using the basic core elements of the SNT-1 PTB domain/FGFR1 peptide complex as an example (Fig. 2): (A) a ligand always adopts an antiparallel β-strand forming an extension of the second β-sheet of the protein. The length of the β-strand, bound between β5 and the C-terminal α helix, is influenced by its additional interactions with the protein, such as β8 in SNT-1; (B) a ligand often contains an NPX(p)Y motif or a related sequence

with a type I β-turn propensity. This asparagine residue at the (Y−3) position plays a key structural role in β-turn formation through intra- and intermolecular hydrogen bonds. This turn may be important for positioning a phosphotyrosine for coordination with solvent-exposed basic residues in the protein, or for anchoring C-terminal residues of the ligand for forming stabilizing interactions with the protein; (C) ligand residues C-terminal to the β-strand (either phosphotyrosine or hydrophobic in nature) form additional contacts with the protein; and (D) ligand residues N-terminal to the β-strand can form additional interactions with the PTB domain. These interactions could involve regions of the protein including (i) the hydrophobic core between the β5 and C-terminal α helix, (ii) the surface residues of the second β-sheet, and (iii) the hydrophobic pocket located on the other side of the β-sandwich opposite to β5. It is important to note that only the β-strand conformation appears absolutely required in biologically relevant ligands. The other major contacts can be used together or in combination to establish specific associations with the PTB domain. Together, these structural studies illustrate how functional diversity of PTB domains can be achieved by evolutionary changes of the structural features at the ligand binding sites outside the conserved scaffold.

Because of low protein sequence homology, secondary and tertiary structures, which contain more conserved, unifying features of PTB domains, may be better for identifying new family members and assigning their functions. It would not be surprising to see other unexpected modes of ligand recognition emerging from three-dimensional structures of new PTB domains in complex with their biological ligands.

Acknowledgements: The work was supported by grants from the American Cancer Society and the National Institutes of Health (to M.-M.Z.).

References

[1] Forman-Kay, J.D. and Pawson, T. (1999) Curr. Opin. Struct. Biol. 9, 690–695.
[2] Zhou, M.-M. and Fesik, S.W. (1995) Prog. Biophys. Mol. Biol. 64, 221–235.
[3] Shoelson, S.E. (1997) Curr. Opin. Chem. Biol. 2, 227–234.
[4] van der Geer, P. and Pawson, T. (1995) Trends Biochem. Sci. 20, 277–280.
[5] Kavanaugh, W.M. and Williams, L.T. (1994) Science 266, 1862–1865.
[6] Blaikie, P., Immanuel, D., Wu, J., Li, N., Yajnik, V. and Margolis, B. (1994) J. Biol. Chem. 269, 32031–32034.
[7] van der Geer, P. et al. (1995) Curr. Biol. 5, 404–412.
[8] Gustafson, T.A., He, W., Craparo, A., Schaub, C.D. and O'Meill, T.J. (1995) Mol. Cell. Biol. 15, 2500–2508.
[9] Kavanaugh, W.M., Turck, C.W. and Williams, L.T. (1995) Science 268, 1177–1179.
[10] Trüb, T., Cho, W.E., Wolf, G., Ottinger, E., Hen, Y.J., Weiss, M. and Shoelson, S.E. (1995) J. Biol. Chem. 270, 18205–18208.
[11] He, W., O'Neill, T.J. and Gustafson, T.A. (1995) J. Biol. Chem. 270, 23258–23262.
[12] Wolf, G. et al. (1995) J. Biol. Chem. 270, 27407–27410.
[13] Zhou, S., Margolis, B., Chaudhuri, M., Shoelson, S.E. and Cantley, L.C. (1995) J. Biol. Chem. 270, 14863–14866.
[14] Borg, J.-P., Ooi, J., Levy, E. and Margolis, B. (1996) Mol. Cell. Biol. 16, 6229–6241.
[15] Zambrano, N. et al. (1997) J. Biol. Chem. 272, 6399–6409.
[16] Chien, C.T., Wang, S., Rothenberg, M., Jan, L.Y. and Jan, Y.N. (1998) Mol. Cell. Biol. 18, 598–607.
[17] Zwahlen, C., Li, S.-C., Kay, L.E., Pawson, T. and Forman-Kay, J.D. (2000) EMBO J. 19, 1505–1515.
[18] Li, S.-C. et al. (1997) Proc. Natl. Acad. Sci. USA 94, 7204–7209.
[19] Li, S.-C., Zwahlen, C., Vincent, S.J., McGlade, C.J., Kay, L.E., Pawson, T. and Forman-Kay, J.D. (1998) Nature Struct. Biol. 5, 1075–1083.
[20] Ravichandran, K.S., Zhou, M.-M., Pratt, J.C., Harlan, J.E., Walk, S., Fesik, S.W. and Burakoff, S.J. (1997) Mol. Cell. Biol. 17, 5540–5549.
[21] Howell, B.W., Kanier, L.M., Frank, R., Gertler, F.B. and Cooper, J.A. (1999) Mol. Cell. Biol. 19, 5179–5188.
[22] Hadari, Y.R., Kouhara, H., Lax, I. and Schlessinger, J. (1998) Mol. Cell. Biol. 18, 3966–3973.
[23] Kouhara, H., Hadari, Y.R., Spivak-Kroizman, T., Schilling, J., Bar-Sagi, D., Lax, I. and Schlessinger, J. (1997) Cell 89, 693–702.
[24] Wang, J.-K., Xu, H., Li, H.-C. and Goldfarb, M. (1996) Oncogene 13, 721–729.
[25] Meakin, S.O., MacDonald, J.I.S., Gryz, E.A., Kubu, C.J. and Verdi, J.M. (1999) J. Biol. Chem. 274, 9861–9870.
[26] Peng, X., Greene, L.A., Kaplan, D.R. and Stephens, R.M. (1995) Neuron 15, 395–406.
[27] Ong, S.H., Guy, G.R., Hadari, Y.R., Laks, S., Gotoh, N., Schlessinger, J. and Lax, I. (2000) Mol. Cell. Biol. 20, 979–989.
[28] Xu, H., Lee, K.W. and Goldfarb, M. (1998) J. Biol. Chem. 273, 17987–17990.
[29] Farooq, A., Plotnikova, O., Zeng, L. and Zhou, M.-M. (1999) J. Biol. Chem. 274, 6114–6121.
[30] Blomberg, N., Baraldi, E., Nilges, M. and Saraste, M. (1999) Trends Biochem. Sci. 24, 441–445.
[31] Lemmon, M.A., Ferguson, K.M. and Schlessinger, J. (1996) Cell 85, 621–624.
[32] Ferguson, K.M., Lemmon, M.A., Sigler, P.B. and Schlessinger, J. (1995) Nature Struct. Biol. 2, 715–718.
[33] Zhou, M.-M. et al. (1995) Nature 378, 584–592.
[34] Zhou, M.-M. et al. (1996) Nature Struct. Biol. 3, 388–393.
[35] Eck, M.J., Dhe-Paganon, S., Trüb, T., Nolte, R. and Shoelson, S.E. (1996) Cell 85, 695–705.
[36] Dhe-Paganon, S., Ottinger, E.A., Nolte, R.T., Eck, M.J. and Shoelson, S.E. (1999) Proc. Natl. Acad. Sci. USA 96, 8378–8383.
[37] Zhang, Z., Lee, C.-H., Mandiyan, V., Borg, J.-P., Margolis, B., Schlessinger, J. and Kuriyan, J. (1997) EMBO J. 16, 6141–6150.
[38] Dhalluin, C. et al. (2000) Mol. Cell 6, 921–929.
[39] Murzin, A.G., Brenner, S.E., Hubbard, T. and Chothia, C. (1995) J. Mol. Biol. 247, 536–540.
[40] Yoon, H.S., Hajduk, P.J., Petros, A.M., Olejniczak, E.T., Meadows, R.P. and Fesik, S.W. (1994) Nature 369, 672–675.
[41] Harlan, J.E., Hajduk, P.J., Yoon, H.S. and Fesik, S.W. (1995) Nature 371, 167–170.
[42] Ferguson, K.M., Lemmon, M.A., Schlessinger, J. and Sigler, P.B. (1995) Cell 83, 1037–1046.
[43] Baraldi, E. et al. (1999) Struct. Fold. Des. 7, 449–460.
[44] Prehoda, K.E., Lee, D.J. and Lim, W.A. (1999) Cell 97, 471–480.
[45] Fedorov, A.A., Fedorov, E., Gertler, F. and Almo, S.C. (1999) Nature Struct. Biol. 6, 661–665.
[46] Vetter, I.R., Nowak, C., Nishimoto, T., Kuhlmann, J. and Wittinghofer, A. (1999) Nature 398, 39–46.
[47] Pearson, M.A., Reczek, D., Bretscher, A. and Karplus, P.A. (2000) Cell 101, 259–270.
[48] Hamada, K., Shimizu, T., Matsui, T., Tsukits, S. and Hakoshima, T. (2000) EMBO J. 19, 4449–4462.
[49] Pawson, T. and Nash, P. (2000) Genes Dev. 14, 1027–1047.
[50] Pawson, T. and Scott, J.D. (1997) Nature 278, 2075–2080.
[51] Dho, S.E., Jacob, S., Wolting, C.D., French, M.B., Rohrschneider, L.R. and McGlade, C.J. (1998) J. Biol. Chem. 273, 9179–9187.
[52] Yan, K.S., Kuti, M., Mujtaba, S., Farooq, A., Goldfarb, M.P. and Zhou, M.-M. (2001) submitted.