The EMBO Journal (2008) 27, 804–816 | © 2008 European Molecular Biology Organization | All Rights Reserved 0261-4189/08
www.embojournal.org

THE EMBO JOURNAL

Structural basis for a novel intrapeptidyl H-bond and reverse binding of c-Cbl-TKB domain substrates

Cherlyn Ng[1,3], Rebecca A Jackson[2,3], Jan P Buschdorf[2], Qingxiang Sun[1], Graeme R Guy[2,*] and J Sivaraman[1,*]

[1]Department of Biological Sciences, National University of Singapore, Singapore and [2]Institute of Molecular and Cell Biology, Proteos, Singapore

The c-Cbl tyrosine kinase binding domain (Cbl-TKB), essentially an 'embedded' SH2 domain, has a critical role in targeting proteins for ubiquitination. To address how this domain can bind to disparate recognition mofits and to determine whether this results in variations in substrate-binding affinity, we compared crystal structures of the Cbl-TKB domain complexed with phosphorylated peptides of Sprouty2, Sprouty4, epidermal growth factor receptor, Syk, and c-Met receptors and validated the binding with point-mutational analyses using full-length proteins. An obligatory, intrapeptidyl H-bond between the phosphotyrosine and the conserved asparagine or adjacent arginine is essential for binding and orientates the peptide into a positively charged pocket on c-Cbl. Surprisingly, c-Met bound to Cbl in the reverse direction, which is unprecedented for SH2 domain binding. The necessity of this intrapeptidyl H-bond was confirmed with isothermal titration calorimetry experiments that also showed Sprouty2 to have the highest binding affinity to c-Cbl; this may enable the selective sequestration of c-Cbl from other target proteins.

The EMBO Journal (2008) **27,** 804–816. doi:10.1038/emboj.2008.18; Published online 14 February 2008
Subject Categories: signal transduction; structural biology
Keywords: c-Cbl; EGFR; Met; reverse binding; Sprouty; TKB domain; X-ray crystallography

Introduction

Deregulation of growth factor receptors occurs in a large number of malignancies, predominantly resulting in enhanced or prolonged receptor signaling. Recent evidence implicates the loss of negative regulation, such as ubiquitination and endosome-directed degradation, as having a significant role in enhanced receptor activity (Peschard and Park, 2003). One of the key negative regulators of receptor tyrosine kinases (RTKs) and non-RTKs is the protein product of the c-Cbl proto-oncogene (Thien and Langdon, 2005). Originally assigned a role as an adapter/docker protein, c-Cbl was revealed to have E3-ubiquitin ligase activity via its RING domain, shedding light on the mechanism by which c-Cbl orchestrated the downregulation of tyrosine kinase-activated pathways (Joazeiro *et al*, 1999).

c-Cbl interacts with a diverse array of proteins via phosphorylatable tyrosine residues at its C terminus, via its proline-rich region, and through an N-terminal phosphotyrosine-like binding domain (Schmidt and Dikic, 2005) lying adjacent to the RING domain. Although this phosphotyrosine binding domain (PTB) appeared to have several hallmarks of a PTB consensus motif, the crystal structure demonstrated that it structurally resembled an SH2 domain but required a flanking four-helix bundle (4H) and an EF-hand subdomain to accomplish binding (Meng *et al*, 1999). Together, these three sub-domains make up the tyrosine kinase-binding (TKB) domain, which is unique to Cbl proteins and often referred to as a specialized or 'embedded' SH2 domain (Liu *et al*, 2006) (Figure 1A).

Binding to the TKB domain is required for the subsequent conjugation of ubiquitin through the RING domain during the ubiquitination process. The consensus binding sequence targeted by the TKB domain was originally identified by Lupher *et al* (1997) as D(N/D)XpY, but was later experimentally refined as (N/D)XpY(S/T)XXP. This motif is common to several protein tyrosine kinases (PTK), such as epidermal growth factor receptor (EGFR), colony-stimulating factor 1 receptor, ZAP-70 (ζ-associated protein of 70 kDa), Src and Syk, as well as members of the Sprouty (Spry) family of Ras/mitogen-activated protein kinase inhibitors (Figure 1B) (herein referred to as the PTK subgroup). Each of these targets binds to the TKB domain following phosphorylation of a central tyrosine residue in the consensus motif and, in most cases, their protein expression levels are subsequently downregulated by ubiquitination (Schmidt and Dikic, 2005).

In recent years, two additional consensus motifs have been identified as targeted by the c-Cbl-TKB domain. The first, RA(V/I)XNQpY(S/T), is a derivation of the original sequence, with conserved residues extending back to the arginine at the pY−6 position (Figure 1B). This sequence is conserved amongst a family of adaptor proteins based on the APS (adapter with a plekstrin homology and Src homology-2 domains) protein (Hu and Hubbard, 2005) and, unlike the PTK subgroup, they are not substrates for ubiquitination, but use c-Cbl as a docker protein in the insulin receptor complex (Chiang *et al*, 2001). The second consensus sequence was identified in a study investigating the binding of c-Cbl to the c-Met receptor (Met) but, unlike the APS sequence, it bears no resemblance to the previously characterized PTK-binding motif (Peschard *et al*, 2001). Alanine-scanning mutations

*Corresponding authors. J Sivaraman, Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore. Tel.: +65 6516 1163; Fax: +65 6779 5671; E-mail: dbsjayar@nus.edu.sg or GR Guy, Institute of Molecular and Cell Biology, Proteos, Singapore, Singapore. Tel.: +65 6586 9614; Fax: +65 6779 1117; E-mail: mcbgg@imcb.a-star.edu.sg
[3]These authors contributed equally to this work

Received: 20 November 2007; accepted: 17 January 2008; published online: 14 February 2008

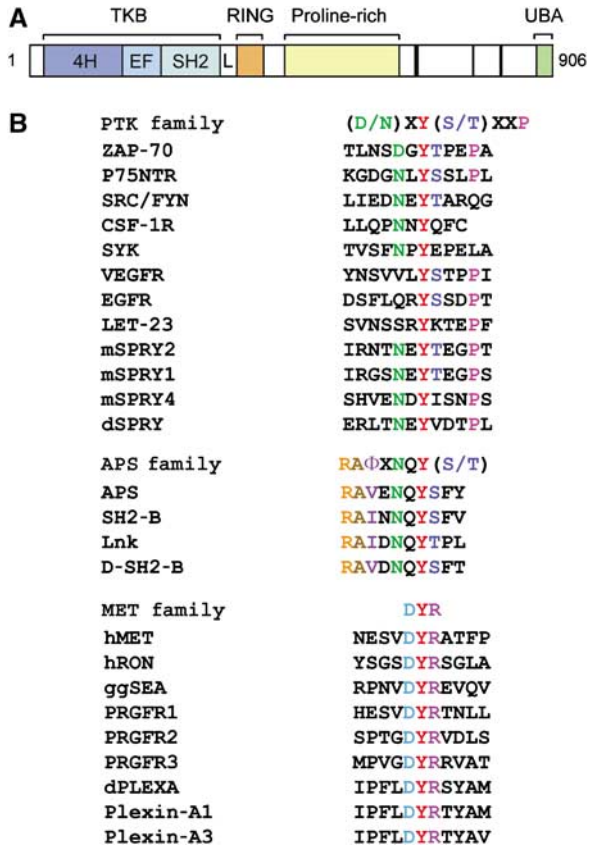

Figure 1 Schematic representation of c-Cbl domain architecture and targets of its TKB domain. (**A**) Structural organization of human c-Cbl. The N-terminal region contains the TKB comprising 4H, EF-hand (EF) and SH-2-like subdomains. The TKB domain is separated from the RING domain by a short linker (L) region. Cbl has a proline-rich stretch following the RING domain and a ubiquitin-associated (UBA) domain at its C terminus. (**B**) Sequence alignment of previously characterized c-Cbl-TKB binding sites for the PTK, APS and Met subgroups of proteins. Coloured residues refer to the conserved residues in the sequence: pY is coloured red, (pY−2)D/N is green, (pY+1)S/T is dark blue, (pY+4)P is fuschia. (pY−6)R in the APS family is orange, (pY−5)A is brown and the (pY−4)ϕ represents a hydrophobic residue, coloured violet. pY flanking residues of the Met family (pY−1)D are coloured light blue and (pY+1)R magenta.

identified that a DpYR motif was essential for the ubiquitination and downregulation of Met by c-Cbl (Peschard *et al*, 2004). This motif is conserved among the Met family members, Ron and Sea (Penengo *et al*, 2003), as well as in plexins—receptors for semaphorins that promote cell repulsion (Tamagnone *et al*, 1999; Figure 1B).

Of the two specialized domains that bind to phosphorylated tyrosines—SH2 and PTB—the SH2 domains evolved in parallel with tyrosine phosphorylation and are inextricably linked to the activity of PTKs, whereas only around 25% of characterized PTB domains have similar targets (Liu *et al*, 2006). PTB domains of proteins such as Shc and IRS-1/2 are reminiscent of the TKB domain and recognize residues N-terminal to the phosphotyrosine (pTyr), NPXpY, whereas the canonical SH2 domain primarily recognizes 3–5 residues C-terminal to the pTyr, and proteins are grouped based on their sequence and binding preferences (Machida and Mayer, 2005; Liu *et al*, 2006). Although the 120 currently documented SH2 domains use diverse strategies to aggregate signalling complexes, there is no precedent for a given SH2 or PTB domain to associate with such apparently unrelated consensus binding motifs as seen with the TKB domain.

Cbl proteins appear to have two major physiological functions—either in acting as scaffolds or in targeting proteins for ubiquitination (Thien and Langdon, 2005). These functions are reflected in the purpose of the TKB domain in (a) recognizing proteins and, in the case of ubiquitin substrates, (b) ensuring its appropriate orientation for ubiquitin conjugation. Cbl proteins have critical roles in the downregulation of both RTKs in growth and differentiation and the antigen receptors in the control of immune responses, and thus elucidating the TKB recognition sequence has been central to understanding the biochemistry of tyrosine kinase downregulation by Cbl proteins. The TKB domain is, in essence, an 'embedded' SH2 domain—a domain generally recognized to bind a 'tight' consensus motif (Liu *et al*, 2006). The abnormal range of the experimentally observed motifs that bind to the TKB domain invites the question as to whether this domain has an uncharacteristic flexibility of binding or whether there is a hitherto undiscovered common binding mechanism that encompasses the previously observed sequences. Furthermore, with such diverse sequence recognition strategies, binding affinity to c-Cbl is likely to be varied and may lead to the selective sequestration of c-Cbl by higher-affinity targets over lower-affinity ones. This model has been suggested as a mechanism whereby Spry2, the most well-characterized of the Spry proteins, recruits c-Cbl from EGFR and results in the ubiquitination of Spry2 and preservation of EGFR on the cell surface (Wong *et al*, 2002; Hall *et al*, 2003; Rubin *et al*, 2003).

We report here that TKB binding requires the formation of a unique intrapeptidyl hydrogen bond (H-bond) between the pTyr and the conserved asparagine (Asn) at the pY−2 position or an arginine (Arg) at the pY−1 position when the Asn is absent. We also demonstrate that Met binds to the TKB domain in the reverse direction, from N to C terminus with the formation of an intrapeptidyl H-bond between the pTyr and the Arg at the pY+1 position. To our knowledge, this is the first description of an SH2 domain having the capacity to bind to proteins in two orientations. This intrapeptidyl H-bond is critical for binding to c-Cbl, as exemplified by the absence of binding with vascular endothelial growth factor receptor 1 (VEGFR1) (lacking either an Asn or Arg residue), which can be recovered with the substitution of either of these amino acids. The other conserved residues of the TKB-binding motifs are not essential for binding but contribute to the binding affinity of the substrate, which may be physiologically relevant for the selective sequestration of c-Cbl by high-affinity substrates. Overall, these findings illustrate the unique binding strategy for the specialized and biologically vital function of c-Cbl.

Results

Complex crystal structures reveal structural conservation and novelty in binding to the Cbl-TKB domain

The ability of the TKB domain to bind to two different consensus motifs—(D/N)XpY(S/T)XXP or DpYR—but still appropriately align the protein for ubiquitination suggests that there would be a common binding mechanism. The motifs on Spry2, EGFR, Syk and Met that recognize c-Cbl-

TKB domain (herein referred to as cbl-TKB) have previously been identified and, as Spry4 has three of the four conserved residues, we also explored whether Spry4 binds to c-Cbl, as results for this are currently inconclusive. The structures of the complex crystals between Cbl-TKB and the phosphorylated 13-mer peptides of these five proteins were determined between 1.35–2.60 Å resolution and crystallized in two different space groups (Table I). Four of the five complexes resembled the binding pattern previously identified, with several exceptions (Meng *et al*, 1999; Hu and Hubbard, 2005), whereas the Met peptide bound to Cbl-TKB in the reverse orientation, with backbone directionality extending from C to N terminus. This reversed binding has previously not been seen for the TKB domain or for any other SH2 or PTB domain. The Cbl-TKB:Met complex also represents the highest-resolution structure currently available for c-Cbl. The structure of the Cbl-TKB:peptide complexes for Spry2$^{(49-61)}$, EGFR$^{(1063-1075)}$ and Met$^{(997-1009)}$ are shown as ribbon diagrams, respectively, in Figure 2A, C and E, to illustrate the position of the peptide relative to the subdomains of the TKB domain. The Cbl-TKB:peptide complexes for Spry4$^{(69-81)}$ and Syk$^{(317-329)}$ resemble the Cbl-TKB:Spry2 complex (Supplementary Figure 1A and C). Electrostatic surface representations for the Spry2, EGFR and Met complexes are shown in Figure 2B, D and F, respectively, and in the Supplementary Figure 1B and D for Spry4 and Syk, demonstrating peptide binding and the surface morphology of the TKB domain. Electron density maps for each of the peptides are provided in Supplementary Figure 2. In all of the complexes, extreme N- and C-terminal residues are absent in the electron density map and presumed to be disordered.

Conserved residues in the (D/N)XpY(S/T)XXP motif contribute to binding in varying degrees

The Cbl-TKB:peptide interactions from the PTK subgroup—Spry2, Spry4, EGFR and Syk—are primarily mediated by 9–13 H-bonds (<3.2 Å) (Figure 3; Table II). The essential pTyr residue occupies a positively charged binding pocket on Cbl-TKB and makes five to six H-bond contacts with several residues of the TKB domain (Arg294, Ser296, Cys297 and Thr298), supporting previous reports that the pTyr contributes almost 50% of the free energy of association to c-Cbl binding (Liu *et al*, 2006). A close-up view of this interaction is shown for Spry2 and EGFR in Figure 3A and C, with H-bonds represented by grey dotted lines, and an extended linear view of these interactions is shown (Figure 3B and D) with green dotted lines representing H-bonds.

The interactions mediated by the other conserved residues of the PTK subgroup were more varied. The (pY + 4)Pro of Spry2, Spry4 and EGFR peptides occupies a shallow hydrophobic cleft flanked by Tyr307, Thr317 and Phe336 on Cbl-TKB. The Syk peptide lacks this interaction and binds solely to the NXpY region, possibly further stabilized by additional H-bonds with Asn79 of Cbl-TKB through its nonconserved

Table I Data collection and refinement statistics

Data collection	Cbl-TKB Spry2	Cbl-TKB Spry4	Cbl-TKB EGFR[a]	Cbl-TKB Syk	Cbl-TKB Met
Resolution (Å)	50.00–1.98 (2.05–1.98)	50.00–2.00 (2.07–2.00)	50.00–2.60 (2.69–2.60)	50.00–1.45 (1.50–1.45)	50.00–1.35 (1.40–1.35)
Space group	P6	P6	$P2_1$	$P2_1$	$P2_1$
Unit cell (Å)	$a=122.26$, $b=122.26$, $c=54.71$	$a=122.85$, $b=122.85$, $c=55.54$	$a=63.86$, $b=110.17$, $c=55.82$, $\beta=89.94^\circ$	$a=63.79$, $b=104.81$, $c=52.78$, $\beta=89.83^\circ$	$a=63.18$, $b=104.80$, $c=52.60$, $\beta=90.41^\circ$
Molecules/a.s.u.	1	1	2	2	2
Reflections (total/unique)	528 974/32 750	566 291/31 352	59 344/21 945	181 855/36 701	930 080/142 157
R_{sym} (%)[b]	8.3 (34.6)	5.2 (15.1)	10.9 (33.5)	7.5 (25.4)	10.9 (23.6)
Completeness (%)[c]	99.9 (99.1)	96.3 (82.1)	91.9 (83.0)	92.7 (66.0)	95.0 (87.2)
Average redundancy	16.2 (8.2)	18.1 (13.7)	2.7 (2.0)	5.8 (3.3)	6.5 (4.9)
Refinement[d]					
Resolution (Å)	20.00–2.00	20.00–2.00	20.00–2.60	20.00–1.45	20.00–1.35
R_{work} (%)[e]	22.40 (29 469)	20.13 (29 166)	23.04 (15 455)	22.41 (95 507)	22.52 (127 952)
R_{free} (%)[f]	26.28 (2221)	24.24 (2128)	27.48 (2304)	23.84 (2176)	23.98 (2208)
Average B-factors (Å^2) (no. of atoms)					
Cbl-TKB	29.71 (2489)	40.11 (2489)	33.65 (4980)	21.79 (4994)	17.21 (4994)
Peptide	25.94 (74)	48.82 (63)	35.29 (190)	33.02 (162)	22.76 (140)
Water	42.86 (297)	53.97 (355)	32.18 (235)	30.71 (575)	26.66 (686)
r.m.s.d bond lengths (Å)	0.009	0.010	0.009	0.006	0.006
r.m.s.d bond angles (deg)	1.24	1.36	1.32	1.10	1.11
Ramachandran plot[g]					
Most favored regions (%)	89.8	90.5	81.3	89.3	89.1
Additionally allowed regions (%)	9.8	9.1	16.9	10.0	10.1
Generously allowed regions (%)	0.4	0.4	1.8	0.7	0.7
Disallowed regions (%)	0.0	0.0	0.0	0.0	0.0

Values in parentheses indicate statistics for the highest-resolution shells.
[a]NCS restraint was kept throughout the refinement of Cbl–TKB EGFR.
[b]$R_{sym} = |I_i - \langle I \rangle| / |I_i|$ where I_i is the intensity of the ith measurement, and $\langle I \rangle$ is the mean intensity for that reflection.
[c]Completeness = (number of independent reflections)/(total theoretical reflections).
[d]For all models, reflections with $I > \sigma I$ was used in the refinement.
[e]R factor $= 100 \times \sum |F_P - F_{P(calc)}| / \sum F_P$.
[f]R-free was calculated with approximately the same number of reflections in all the complexes' test set.
[g]Statistics for the Ramachandran plot from an analysis using PROCHECK (Laskowski *et al*, 1993).

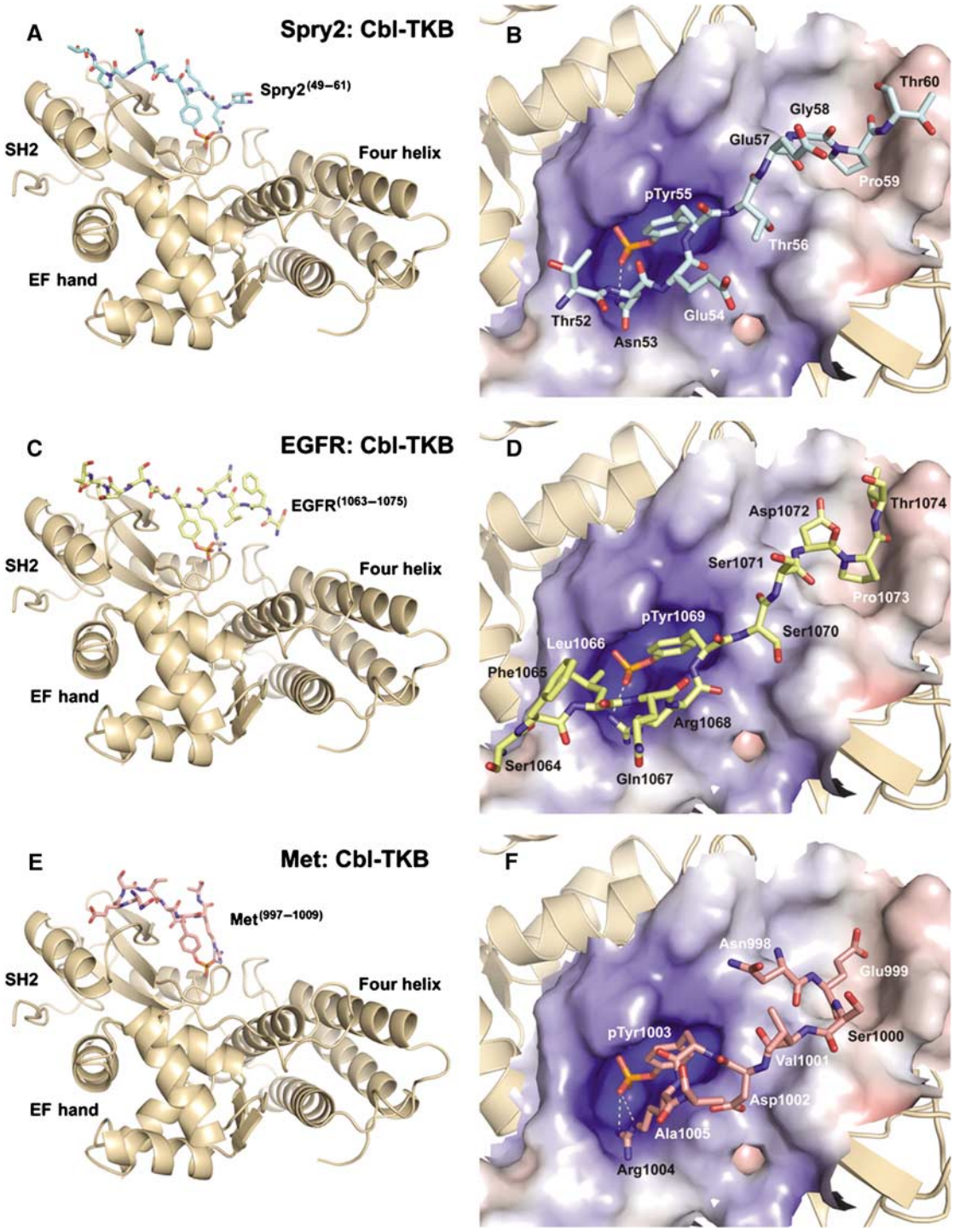

Figure 2 Crystal structure and electrostatic surface representations of TKB domain complexed with Spry2, EGFR and Met. Crystal structures of the Cbl-TKB complexed with each of the peptides (**A**) Spry2$^{49-61}$ (cyan), (**C**) EGFR$^{1063-1075}$ (yellow) and (**E**) Met$^{997-1009}$ (pink), and surface representations of the Cbl-TKB complexed with (**B**) Spry2$^{49-61}$, (**D**) EGFR$^{1063-1075}$ and (**F**) Met$^{997-1009}$. The c-Cbl-TKB domain is shown as a ribbon diagram with helices, β-strands and loops coloured light brown. Phosphopeptides are shown as stick models. Intrapeptidyl H-bonds are shown as grey dotted lines. Residues at the extreme N and C termini of the phosphopeptides are disordered and not included. The figure was prepared using Pymol (DeLano, 2002) and the APBS plugin (Baker *et al*, 2001).

(pY−4)Ser319, which is not observed with the other complexes. The conserved (pY + 1)Ser/Thr residue of Spry2 and EGFR interacts with Gln316 of Cbl-TKB through a backbone–backbone H-bonding contact. However, this pY + 1 residue was more tolerant of degeneracy than expected, as the amide N of both (pY + 1)Ile54 in Spry4 and (pY + 1)Glu324 in Syk also formed backbone H-bonding contacts with Gln316 of Cbl-TKB. This explains the lesser preference for conserved residues in the pY + 1 position in the degenerate peptide library study (Lupher *et al*, 1997), although residues with large side chains are likely to have steric clashes and may not be preferred. In all five complexes, the backbone carbonyl O of the pY−1 residue forms a H-bond with the OH of Tyr274 on Cbl-TKB, irrespective of the residue in that position.

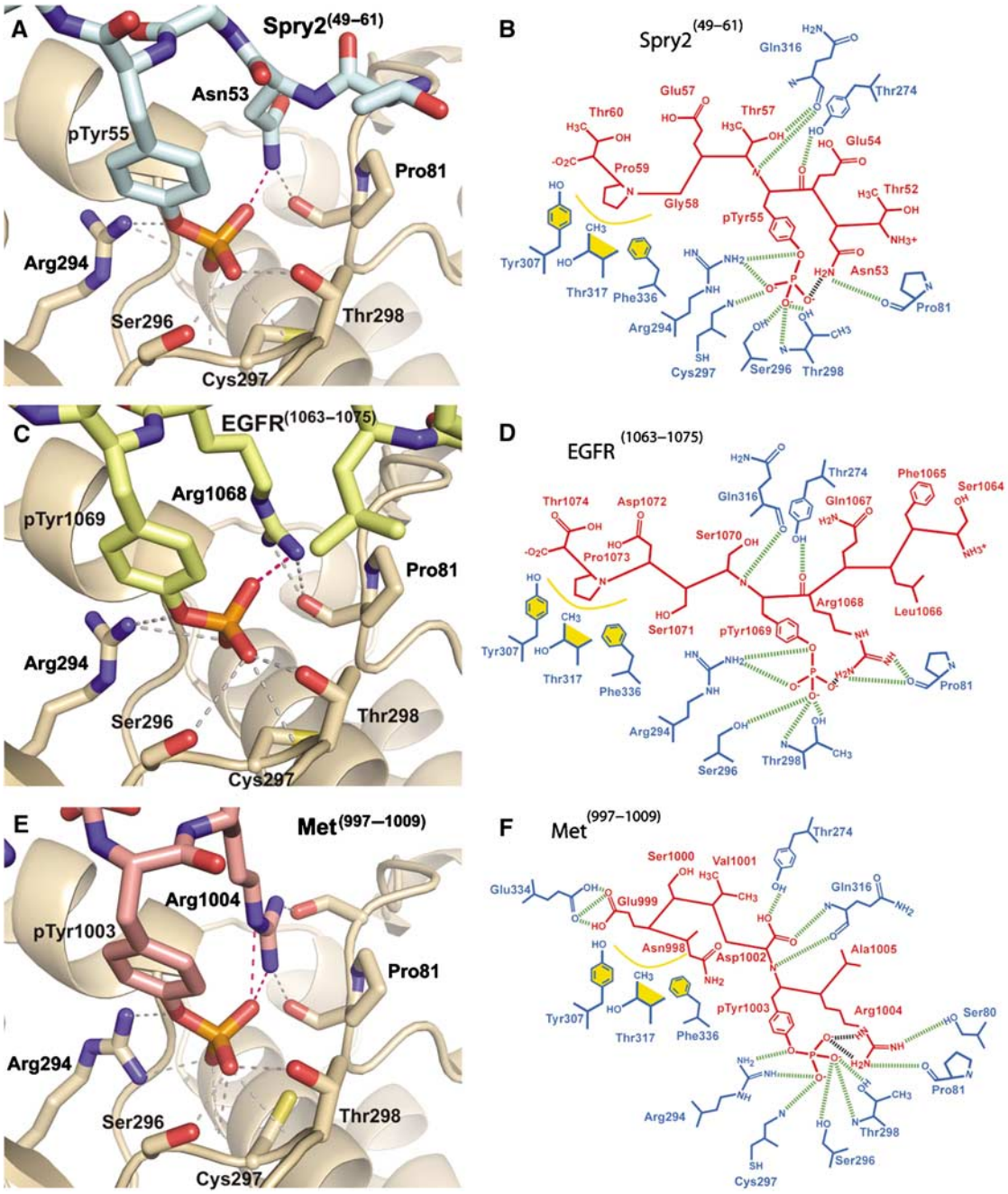

Figure 3 The conserved intrapeptidyl hydrogen (H) bond and neighbouring H-bonds for the conserved pTyr of Spry2, EGFR and Met. Close-up view of the tyrosine-binding pocket of the c-Cbl-TKB domain complexed with (**A**) Spry2$^{49–61}$ (cyan), (**C**) EGFR$^{1063–1075}$ (yellow) and (**E**) Met$^{997–1009}$ (pink) shows the conserved intrapeptidyl H-bonds between the pTyr and the (pY−2)Asn of Spry2 or the (pY−1)Arg of Met. Intrapeptidyl H-bonds are shown as dotted red lines and other H-bonds are shown as grey dotted lines. The figure was prepared using Pymol (DeLano, 2002). A schematic linear extended view of the c-Cbl-TKB domain with (**B**) Spry2, (**D**) EGFR and (**F**) Met. The phosphopeptides are coloured red and the Cbl residues are blue. Residues participating in van der Waals contacts are highlighted in gold with the interacting surfaces as gold curved lines. H-bonds are indicated by green dashed lines and the intrapeptidyl H-bond is depicted by a black dashed line between the phosphate group of the conserved tyrosine and the side chain of the Asn53 (Spry2), Arg1068 (EGFR) or the Arg1004 (Met).

The (pY−2)Asn in the peptides of the PTK subgroup formed a H-bonding contact through the N of the carboxamide side group with the carboxyl O of Pro81 on Cbl-TKB, as shown for Spry2 (Figure 3A and B). The (pY−2)Asn of Syk forms an additional H-bond with Ser80, which is not seen with the other peptides. Interestingly, for EGFR, which lacks the conserved (pY−2)Asn, the (pY−1)Arg appears to assume this role by forming a H-bond with the carboxyl O of Pro81 on Cbl-TKB in a manner similar to the other complexes (Figure 3C and D). This 'rescue'-binding mechanism of EGFR may enable it to acquire a higher binding affinity than would be expected.

Met binds to the TKB domain in the reverse orientation

One of the more unexpected findings was that the Met peptide bound to c-Cbl-TKB domain in the reverse orientation from C to N terminus (Figures 2E, F, 3E and F). This is contrary to previous modelling predictions (Peschard *et al*, 2004) and

Table II Hydrogen bonding interactions between Cbl-TKB and the various peptides

Cbl-TKB	Spry2$^{49-61}$	Spry4$^{47-59}$	EGFR$^{1063-1075}$	Syk$^{317-329}$	Met$^{997-1009}$
Asn79 Nδ2				Ser319 N, 3.04	
Ser80 Oγ				Asn321 Oδ1, 2.52	Arg1004 NH1, 2.62
Pro81 O	Asn53 Nδ2, 2.90	Asn51 Nδ2, 3.04	Arg1068 NH1, 2.81 Arg1068 NH2, 2.66	Asn321 Nδ2, 3.09	Arg1004 NH2, 2.85
Tyr274 OH	Glu54 O, 2.68	Asp52 O, 2.80	Arg1068 O, 2.81	Pro322 O, 2.71	Asp1002 Oδ2, 2.58
Arg294 NH1				pTyr323 PO2, 3.02	pTyr1003 PO2, 2.97
NH2	pTyr55 OH, 2.89 pTyr55 PO3, 2.91	pTyr53 OH, 2.89	pTyr1069 PO2, 3.03 pTyr1069 OH, 2.88	pTyr323 OH, 2.93	pTyr1003 OH, 2.93
Ser296 Oγ	pTyr55 PO3, 2.76	pTyr53 PO3, 2.71	pTyr1069 PO3, 3.13	pTyr323 PO3, 2.72	pTyr1003 PO3, 2.67
Cys297 N	pTyr55 PO3, 2.98	pTyr53 PO2, 2.95		pTyr323 PO2, 2.92	pTyr1003 PO2, 3.04
Thr298 Oγ1	pTyr55 PO3, 2.55	pTyr53 PO3, 2.70	pTyr1069 PO3, 2.94	pTyr323 PO3, 2.56	pTyr1003 PO3, 2.61
N	pTyr55 PO3, 2.89	pTyr53 PO3, 2.88	pTyr1069 PO3, 2.77	pTyr323 PO3, 3.01	pTyr1003 PO3, 2.99
Tyr307 OH					Glu999 O, 2.53
Gln 316 O	Thr56 N, 2.84 Thr56 Oγ1, 3.21	Ile54 N, 2.84	Ser1070 N, 2.75	Glu324 N, 2.85	pTyr1003 N, 2.86
N					Asp1002 Oδ1, 2.88
Thr317 Oγ1			Ser1070 O, 2.56		
His320 Nε		Asp55 Oδ2, 2.77			
Lys322 Nξ			Asp1072 Oδ2, 3.06		
Glu334 Oε1			Thr1074 Oγ1, 2.96		Glu999 Oε1, 3.03 Glu999 Oε2, 2.92
Oε2					Glu999 Oε1, 2.84

demonstrates that the conserved Asp1002 and Arg1004 residues have more than a supporting role for the pTyr in the binding of Met and c-Cbl. Similar to the other peptides of the PTK group, Met binding in the reverse direction forms six H-bond contacts from the essential pTyr residue with Arg294, Ser296, Cys297 and Thr298 on c-Cbl in a positively charged binding pocket (Figure 3E and F), and similar to the binding with EGFR, the Arg1004 in Met substitutes for the missing (pY−2)Asn and forms H-bond contacts with Pro81. Arg1004 also forms an additional H-bond with Ser80, similar to that observed with the Syk complex, but not with the other phosphopeptides.

The Asp1002 residue in Met assumes the role of the conserved (pY + 1)Ser/Thr in the PTK subgroup by interacting with Gln316 and Tyr274 on Cbl-TKB. In contrast to the modelling predictions by Peschard *et al* (2004), neither the backbone carbonyl nor amino groups of either residue interact with Cbl-TKB, and no salt bridge is formed between the side chains of Asp1002 and Arg1004 to orientate the phosphorylated tyrosine residue into its binding cleft on Cbl-TKB. Rather, the side chains of both residues are orientated into the Cbl-TKB:Met interface in the same direction as pTyr1003. The Glu999 residue of Met forms unique H-bonding contacts with Glu334 of the Cbl-TKB that is not observed in any other complex; its aliphatic part of the side chain partially occupies the hydrophobic cleft that is occupied by (pY + 4)Pro residues of the PTK subgroup.

An intrapeptidyl hydrogen bond is conserved across all TKB domain-binding proteins

Here, we report the formation of an intrapeptidyl H-bond between the phosphate oxygen of the pTyr and the (pY−2)Asn of the Spry2, Spry4 and Syk peptides (represented as a red dotted line in the Spry2 peptide in Figure 3A, or a black dotted line in Figure 3B) during binding to the TKB domain. This H-bond was reported in earlier studies (Meng *et al*, 1999; Hu and Hubbard, 2005); however, no significance was assigned to its existence. The presence of this H-bond across all peptides with an Asn at the pY−2 position strongly suggests that it is evolutionarily conserved. In addition, we also observed the formation of an intrapeptidyl H-bond between the phosphate oxygen of the pTyr and the guanidium group of the Arg at the pY−1 position in the EGFR peptide (Arg1068), which appears to be compensating for the loss of (pY−2)Asn (Figure 3C and D). Furthermore, this intrapeptidyl H-bond is formed between the (pY + 1)Arg1004 and the phosphate oxygen of the pTyr in the Met peptide in its reversed binding orientation (Figure 3E and F), in a similar manner to EGFR. Taken together, these findings demonstrate that an intrapeptidyl H-bond between the pTyr and the (pY−2)Asn, the (pY−1)Arg or the (pY + 1)Arg in the reverse orientation (hereafter also referred to as the adjacent Arg) is conserved across TKB domain-binding substrates, and we propose that this intrapeptidyl H-bond is essential for binding with the Cbl-TKB domain.

Full-length protein binding confirms that (pY−2)Asn and pTyr residues of Spry2 are indispensable for binding

To confirm and extend upon our structural studies using full-length proteins, we first tested the relative importance of each of the four conserved residues by creating alanine point mutations in the TKB-binding motif of Spry2 and examining

how this affects its interaction with endogenous c-Cbl under stimulated conditions (Figure 4A). When compared to WT Spry2 binding, we observed that binding to c-Cbl was disrupted with the Spry2 N53A and Spry2 Y55A mutants, reduced with the Spry2 T56A and P59A mutants and unaffected with the Spry2 E54A mutant. This supports our crystal structure observations that (pY−2)Asn and pTyr are essential for Cbl-TKB binding and suggests that Cbl's interactions with T56A and P59A are important, but not essential for binding, whereas the nonconserved Spry2 E54A mutant is not involved in binding.

Binding between full-length Cbl and its targets validates the peptide- and domain-derived structural studies

To further validate the crystal structures, key residues on c-Cbl that were shown to mediate binding in the crystal structures were mutated (Figure 4B) to test their interaction with WT proteins. In all cases, binding was observed for WT c-Cbl with Spry2 (Figure 4C), Spry4 (Figure 4D), EGFR (Figure 4E), Met (Figure 4F) and APS (Figure 4G). This binding was abrogated with the c-Cbl G306E mutant, a nonbinding mutation originally discovered in a genetic screen on *Sli-1*, the *Caenorhabditis elegans* homologue of c-Cbl (Levkowitz *et al*, 1999) that was shown to disrupt the binding with full-length ZAP-70 (Meng *et al*, 1999); included here as a negative control.

As with the G306E mutant, a phenylalaine to Tyr substitution (Y274F) in c-Cbl also caused a loss of binding to Spry2, Spry4, EGFR and Met. The OH of Tyr274 on Cbl-TKB binds to the invariant backbone carbonyl O of the pY−1 nonconserved residue of Spry2, Spry4 and EGFR or the (pY−1)Asp of Met, which is tested and confirmed by the E45A mutant that retains binding (Figure 4A). The side chains of residues Arg294, Ser296 and Thr298 on c-Cbl formed H-bond contacts with the essential pTyr on all of the peptides, and this is confirmed by a loss or reduction in binding with alanine substitutions of these residues. The S296 residue, however, appeared to be less important as some or all of the binding was retained. Phe336 participates in hydrophobic interactions with (pY+4)Pro, and an alanine substitution caused a reduction in binding between c-Cbl and all of the WT proteins, where present. (pY−4)Glu999 of Met occupies the same cleft, and an alanine mutation of its side chain causes a similar reduction in binding.

Alanine point mutations were also included for K322 and E334 as the crystal structures showed specific interactions for EGFR (Asp1072) and both EGFR (Thr1074) and Met (Glu999), respectively. However, in addition to EGFR, we observed a reduction in binding for Spry2 and Met with the c-Cbl K322A mutant. K322 is a solvent-exposed residue, and an alanine mutation may alter the surface conformation of c-Cbl and reduce its binding with proteins that are in close proximity to c-Cbl. This is possibly the case with the C termini of Spry2, Met and EGFR peptides when compared with C termini of Spry4 and Syk, which bend away from the c-Cbl's surface and are therefore unaffected. In contrast, binding was retained for all of the WT proteins with the c-Cbl E334A mutant, suggesting that the binding mediated by this residue is not essential.

Finally, we wanted to ensure that the structural changes created in the c-Cbl mutants did not influence c-Cbl's interaction with other proteins C-terminal to the TKB, such as Grb2, which binds via its SH3 domain to c-Cbl's proline-rich region. This was confirmed with Grb2 immunoprecipitations, with binding retained for all of the c-Cbl point mutants (Figure 4H).

Isothermal titration calorimetry reveals that Spry2 has the highest binding affinity to Cbl-TKB

To validate the necessity of the intrapeptidyl H-bond, we examined whether binding can still occur in proteins that lack either the (pY−2)Asn or the adjacent Arg, as well as the relative contributions to binding made by the conserved residues in the consensus motifs. The binding affinities of nine 13-mer phosphopeptides were determined by titration against the Cbl-TKB domain in isothermal titration calorimetry (ITC) experiments (Figure 5A): six sequences from the PTK subgroup (Spry2, Spry4, EGFR, Src, Syk and VEGFR), with the exception of Spry2, other sequences lacking at least one of the conserved residues, the APS sequence and two sequences belonging to the Met subgroup (Met and Ron). Importantly, although the binding affinity for APS differed from that previously reported by Hu and Hubbard (2005) who used a pH 8.0 for their measurements, when we repeated our measurements at pH 8.0 instead of pH 7.0, our findings replicated Hu and Hubbard's findings (Figure 5A; Supplementary Figure 3 for ITC profiles at pH 8.0).

Of the nine peptides examined, binding between the Cbl-TKB and the phosphorylated Spry2 peptide occurred with the highest affinity and the most negative Gibbs free energy change (Figure 5A), suggesting that this reaction is most favoured. In descending order of affinity, the dissociation constants (K_d) were as follows: $Spry2^{49-61}$ (0.32 μM) > $APS^{623-632}$ (0.39 μM) > $Spry4^{69-81}$ (0.61 μM) > $EGFR^{1063-1075}$ (0.88 μM) > $Src^{413-425}$ (1.32 μM), $Syk^{317-329}$ (1.49 μM), $Ron^{1011-1023}$ (1.55 μM) and $Met^{997-1009}$ (2.36 μM); the latter four were not significantly different from each other. Strikingly, the phosphorylated $VEGFR1^{1327-1338}$ did not bind, supporting our hypothesis that the intrapeptidyl H-bond is essential for binding to the Cbl-TKB domain. Representative ITC profiles for the Spry2, EGFR, Met and VEGFR peptides with Cbl-TKB are shown in Figure 5B–E, respectively, and the remaining ITC profiles are provided in Supplementary Figure 3.

When comparing between the (pY−2)Asn and the adjacent Arg, we observed that the Asn residue provides a greater contribution to binding affinity, as demonstrated by Spry2 versus EGFR, both of which contain the other three conserved residues, and by Src versus Ron/Met, which only have one other conserved residue. A degenerate peptide library characterization of the (D/N)XpY(S/T)XXP motif indicated that the most favourable residues for binding N-terminal of pTyr are: (pY−2)Asn, (pY−2)Asp or (pY−3)Asp (Liu *et al*, 2002). Surprisingly, an Arg residue N-terminal of pTyr was not significantly selected for binding. It is likely that there is no preference for Arg at the pY−1 position, because an Asn provides higher affinity binding and, as their roles overlap in this context, an Asn at the pY−2 position will be preferred.

Contributing next to the binding affinity is the (pY+4)Pro residue, as shown by the higher affinity of EGFR compared with Met and Ron, and the higher affinity of Spry2 compared with Src and Syk. APS, just below Spry2, also lacks the (pY+4)Pro, but appears to compensate for this by using other conserved residues N-terminal to the pTyr to enhance binding. Interestingly, the extent of the hydrophobic interac-

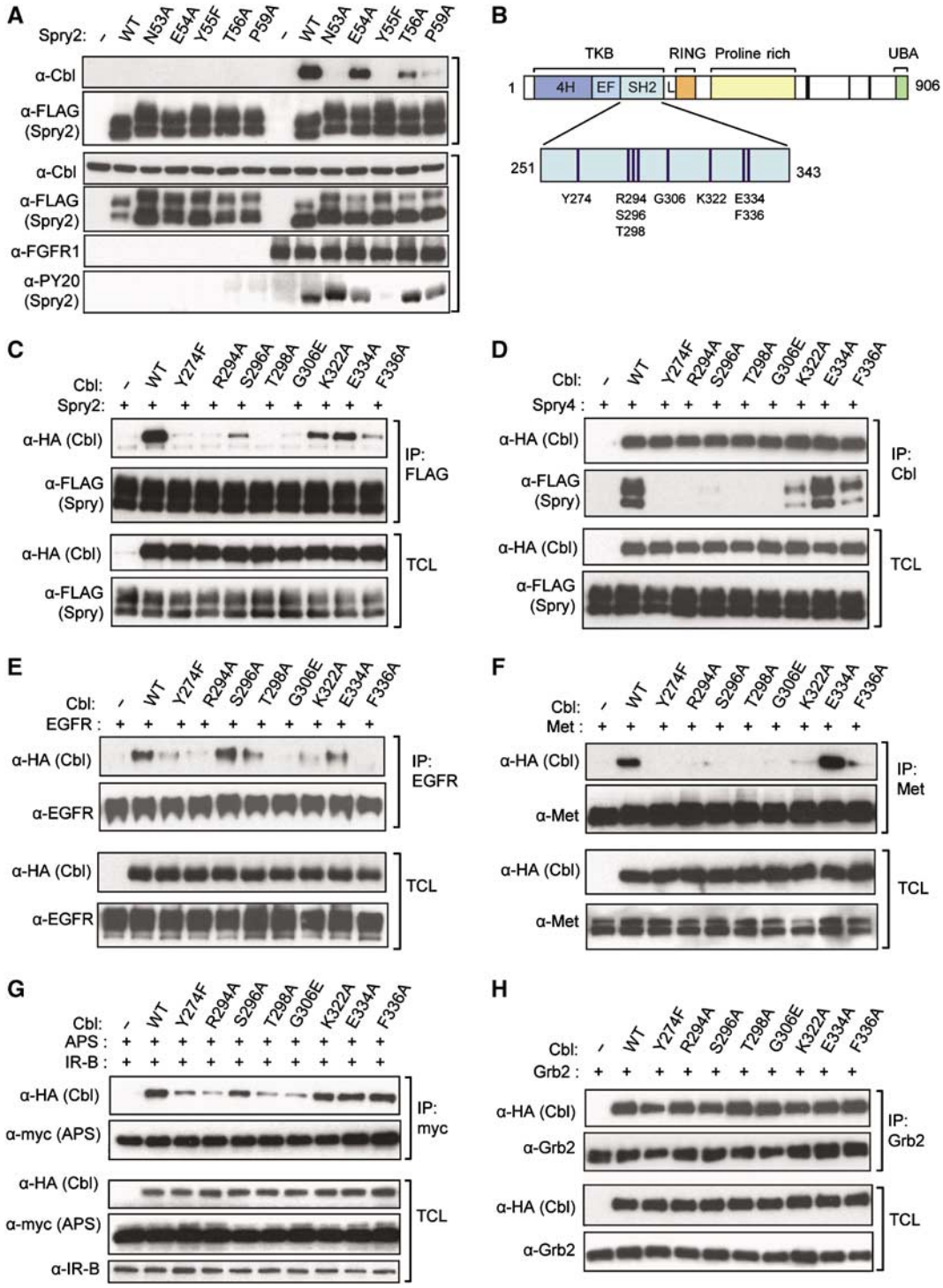

Figure 4 Site-directed mutagenesis to determine the importance of conserved residues. (**A**) 293T cells were transfected with FGFR1 and FLAG-tagged wild-type Spry2 or Spry2 bearing alanine point mutations as indicated. The cells were lysed 24 h post-transfection and FLAG–Spry2 proteins were immunoprecipitated (IP) using FLAG–M2 beads. IP and total cell lysates (TCL) were immunoblotted with anti-Cbl (C-15), anti-FLAG or anti-FGFR1 (C-15) as labelled. (**B**) Diagram of c-Cbl with an enlargement of the TKB domain. The point mutations created in the SH2 subdomain of Cbl are described. (**C**–**H**) 293T cells were transfected with HA-tagged WT c-Cbl and Cbl bearing various point mutations as described in (B), along with (C) FLAG–Spry2, (D) Spry4, (E) EGFR, (F) Met, (G) myc-APS in combination with insulin receptor-B (IR-B) or (H) alone for endogenous Grb2. Cells were stimulated as described in the Materials and methods and lysed 24 or 48 h post-transfection. In (C) and (D), anti-FLAG and anti-Cbl IPs (C-15) and TCLs were immunoblotted with anti-FLAG and anti-HA (Cbl). In (E), anti-EGFR IPs (528) and TCLs were immunoblotted with anti-HA (Cbl) and anti-EGFR (1005). In (F), anti-HGFR IPs and TCLs were immunoblotted with anti-HA and anti-Met (25H2). In (G), anti-myc IPs (A-14) and TCLs were immunoblotted with anti-HA, anti-c-myc (9E10) and IR-β (C-19). In (H), anti-Grb2 IPs (C-23) and TCLs were immunoblotted with anti-HA and anti-Grb2.

tion between the (pY + 4)Pro and its corresponding cleft on Cbl-TKB was directly correlated with the peptide's affinity for the latter: hydrophobic interaction was most extensive with Spry2 at this cleft, followed by Spry4 and, to a lesser extent, EGFR. The conserved (pY + 1)Ser/Thr appears to contribute least to the binding affinity, as suggested by the insignificant differences in affinity between Src and Syk. All of these results support the variations in binding we observed with the full-length proteins.

The (pY−1)Asp of Ron and Met does not appear to contribute substantially to binding affinity, as these peptides have a lower affinity than Syk, which has no additional residues other than the (pY−2)Asn. Interestingly, most of the interactions between Met and c-Cbl are mediated by the DpYR residues. The side chain of the (pY−1)Asp binds to the Tyr274 of Cbl-TKB, whereas this interaction is mediated by the backbone of the pY + 1 residue in the other complexes (Table II). As a D1002A mutation in Met resulted in the loss of binding to c-Cbl (Peschard *et al*, 2004), and as we showed here that a Y274F mutation in c-Cbl abolishes binding to Met, this may indicate why the Asp residue has been conserved and may be a unique binding feature for Met family members.

The intrapeptidyl H-bond is essential for binding to the TKB domain

To further validate the importance of the intrapeptidyl H-bond, two mutant phosphorylated peptides for VEGFR1 were included harbouring Asn or Arg substitutions that would permit formation of an intrapeptidyl H-bond and hence facilitate peptide binding if that was the key requirement for binding: (1) VEGFR1 with a Val1331Asn substitution at the pY−2 position, and (2) VEGFR1 with a Leu1332Arg substitution at the pY−1 position. This substitution would also endow VEGFR1 with four conserved residues, and it would therefore be expected to bind with similar affinities to Spry2 and EGFR, respectively. We show that both of these substitutions facilitate peptide binding to the Cbl-TKB domain (Figure 5F and G) and, as expected, with binding affinities similar to Spry2 and EGFR of 0.33 and 0.72 μM for VEGFR1 V1331N and L1332R, respectively (Figure 5A). These findings confirm the requirement for the formation of an intrapeptidyl H-bond between pTyr and (pY−2)Asn or an adjacent Arg at the pY−1 or pY + 1 position when reversed, and also substantiate the greater contribution of Asn towards a substrate's affinity for Cbl-TKB.

Discussion

We originally set out to investigate two major questions—why the TKB domain recognizes an unusually broad set of sequence-recognition motifs and whether variations in the binding affinity of the different TKB domain-binding proteins could result in the selective sequestration of c-Cbl by high-affinity binding proteins. Although we partially address the latter, demonstrating that differences in binding affinity do exist between different TKB-binding proteins, we cannot confirm whether this proposed sequestration of c-Cbl is of physiological importance. Given a favourable cellular localization, however, it is possible that Spry2 could influence the binding of c-Cbl to particular RTKs or other TKB-binding protein, provided it had a higher binding affinity. In addressing our first question, we have unexpectedly revealed two important elements of TKB binding: (1) the majority of binding is achieved through (D/N)XpY, RpY or pYR (in the reverse orientation) through the formation of an intrapeptidyl H-bond; (2) the Cbl-TKB domain can bind to substrates in two orientations.

The unique diversity of the TKB domain in its recognition of three consensus sequences is an unusual characteristic for

A

Sequence, affinity and favorability of the c-Cbl TKB-binding motifs

Peptide	Sequence	N	K_a ($\times 10^6 M^{-1}$)	K_d (μM)	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)
$Spry2^{49-61}$	IRNTNEpYTEGPTV	0.920 ± 0.007	3.11 ± 0.25	0.32	−13.44 ± 0.14	−4.663	−8.777
$APS^{623-632}$	RAVENQpYSFY	1.010 ± 0.007	2.54 ± 0.20	0.39	− 6.44 ± 0.07	2.217	−8.653
$Spry4^{47-59}$	SHVENDpYIDNPSL	1.060 ± 0.007	1.65 ± 0.09	0.61	− 5.11 ± 0.05	3.276	−8.386
$EGFR^{1063-1075}$	DSFLQRpYSSDPTG	0.993 ± 0.01	1.14 ± 0.11	0.88	− 9.11 ± 0.15	−0.871	−8.242
$Src^{413-425}$	LIEDNEpYTARQGA	0.969 ± 0.01	7.56 ± 0.55	1.32	− 6.33 ± 0.11	1.612	−7.939
$Syk^{317-329}$	TVSFNPpYEPELAP	0.969 ± 0.01	6.73 ± 0.66	1.49	− 3.07 ± 0.08	4.811	−7.882
$Ron^{1011-1023}$	LYSGSDpYRSGLA	1.000 ± 0.01	6.44 ± 0.56	1.55	− 1.58 ± 0.03	6.257	−7.835
$Met^{997-1009}$	SNESVDpYRATFPE	1.080 ± 0.03	4.24 ± 0.53	2.36	− 1.60 ± 0.06	5.991	−7.593
$APS^{623-632}$ pH 8.0	RAVENQpYSFY	0.911 ± 0.002	18.4 ± 1.1	0.054	−10.96 ± 0.04	−1.071	−9.889
$EGFR^{1063-1075}$ pH 8.0	DSFLQRpYSSDPTG	0.991 ± 0.004	2.97 ± 0.12	0.336	−6.178 ± 0.03	2.603	−8.781
$VEGFR1^{1327-1338}$	YNSVVLpYSTPPI	0	0	0	0	0	0
$VEGFR1^{1327-1338}$ V1331N	YNSVNLpYSTPPI	1.050 ± 0.002	3.01 ± 0.09	0.33	−5.003 ± 0.02	3.748	−8.751
$VEGFR1^{1327-1338}$ L1332R	YNSVVRpYSTPPI	0.997 ± 0.003	1.38 ± 0.04	0.72	−3.238 ± 0.02	5.047	−8.285

Figure 5 Binding affinity measurements with isothermal titration calorimetry. (**A**) The binding affinities of $Spry2^{49-61}$, $APS^{623-632}$, $Spry4^{69-81}$, $EGFR^{1063-1075}$, $Src^{413-425}$, $Syk^{317-329}$, $Met^{997-1009}$, $Ron^{1011-1023}$, $VEGFR^{1327-1338}$, $VEGFR^{1327-1338}$ V1331N and $VEGFR^{1327-1338}$ L1332R phosphopeptides to the c-Cbl-TKB domain were measured using ITC. ITC experiments for $APS^{623-632}$ and $EGFR^{1063-1075}$ were also repeated at pH 8.0 according to Hu and Hubbard (2005). The concentration of the TKB domain was 10.8 μM in all experiments. (**B**–**G**) Representative ITC profiles for Spry2, EGFR, Met, VEGFR, VEGFR V1331N and VEGFR L1332R are shown respectively. The figures show the injection profile after baseline correction and the bottom panels show the integration (heat release) for each injection (except the first one). The solid lines in the bottom panel show the fit of the data to a function based on a one-site binding model. The binding constants (K_a and K_d), number of binding sites (N), enthalpy (ΔH) and entropy (ΔS) changes of Cbl-N to the various peptides are provided in (A).

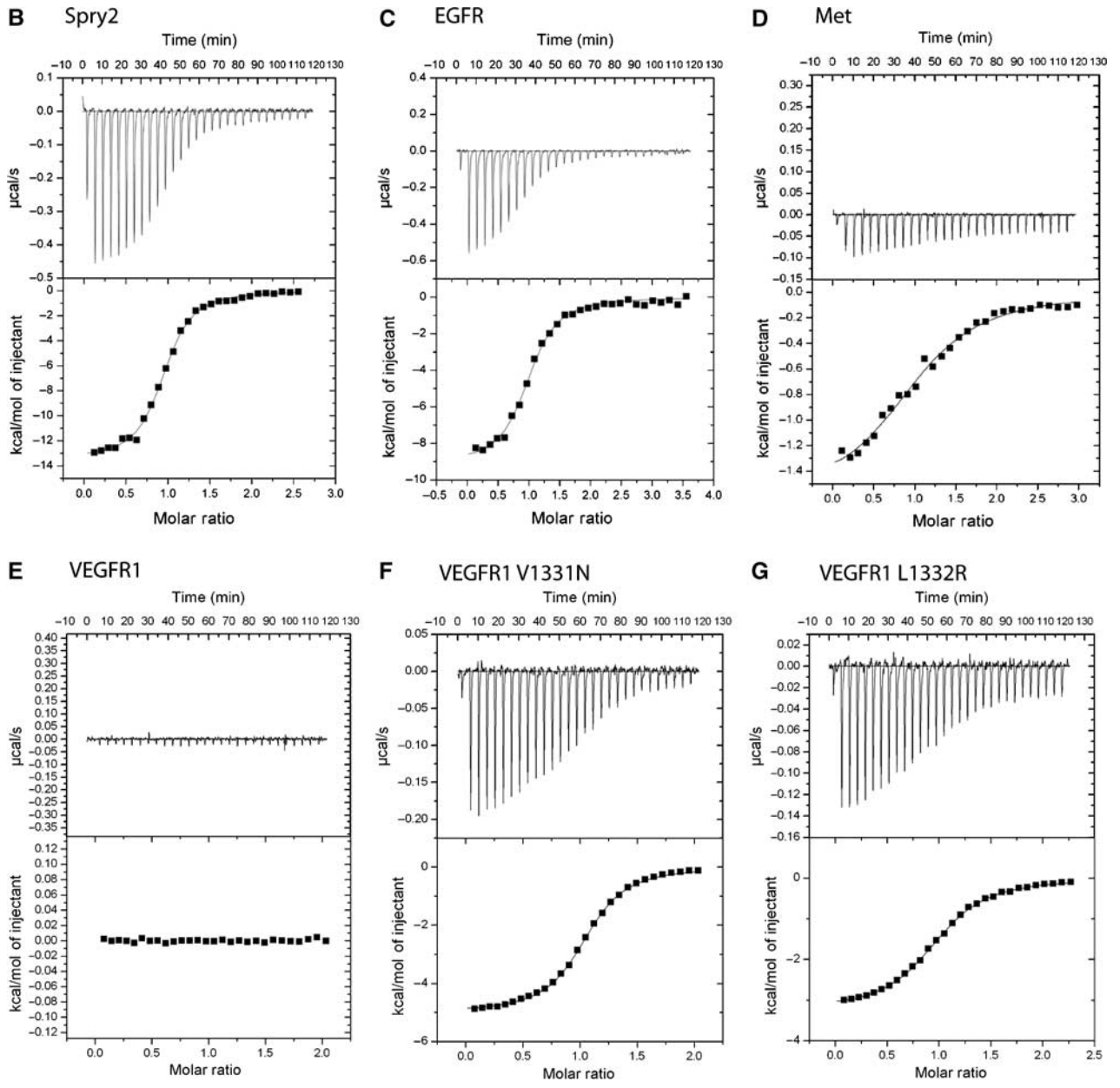

Figure 5 Continued.

an SH2 domain. These five complexes have revealed that there is some degree of consistency in binding between the Cbl-TKB targets and indicates that the TKB domain recognizes a less-radical set of motifs. The intrapeptidyl H-bond between pTyr and (pY−2)Asn, (pY−1)Arg or (pY+1)Arg in the reverse is necessary for the directed, 'hand-in-glove' fit of the peptide into a positively charged pocket of the SH2 region of the Cbl-TKB domain (depicted in Figure 6A–D for Spry2 and Met), whether before binding to c-Cbl or concomitantly. Although this intrapeptidyl H-bond was identified in both of the previously characterized c-Cbl structures (Meng *et al*, 1999; Hu and Hubbard, 2005), due consideration was not attributed to it due to a lack of supporting evidence to highlight its significance. When combined with these five structures, we are able to assign importance to the intrapeptidyl H-bond as a rule for binding and an event crucial for the directed positioning of the substrate on c-Cbl. Furthermore, with the ITC characterizations of the three VEFGR peptides, we show conclusively that binding to c-Cbl cannot occur without this intrapeptidyl H-bond, despite the presence of other conserved residues (depicted with a model diagram in Figure 6E–G). These other conserved residues—(pY+1)Ser/Thr, (pY+4)Pro and (pY−1)Asp—tweak the binding affinity, and such as in the case of APS, high-affinity binding can be compensated for by other residues. Substrate binding is further stabilized in concert with the 4H subdomain and EF-hand, particularly the 4H subdomain, which contributes to binding via the Pro81 residue to the conserved (pY−2)Asn or compensating Arg (Figure 3), and structural evidence suggests that these subdomains form an intimate, multifaceted association with the SH2 subdomain, linker sequence and RING domain in its interaction with the appropriate E2 enzyme, creating what appears to be a precise, rigid structure.

A comprehensive search of the known SH2 domain structures available in the protein data bank resulted in only four complex models where a pTyr formed a H-bond with the side

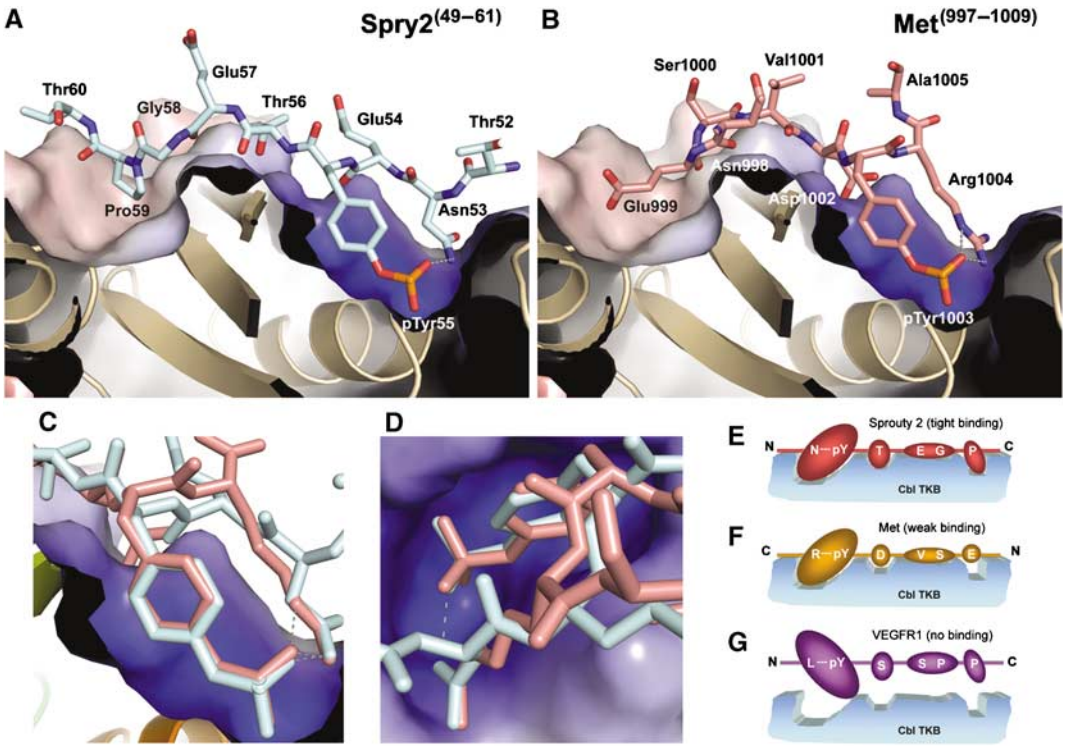

Figure 6 Orientation of the intrapeptidyl hydrogen (H) bond within the binding pocket of the Cbl-TKB. A cross-sectional perspective of the binding topograph in Figure 2B and F illustrates the fit of the phosphopeptides on the surface of the Cbl-TKB domain for (**A**) Spry2$^{49-61}$ (cyan) and (**B**) Met$^{997-1009}$ (pink). (**C**, **D**) Peptides superimposed from two different angles. (**E–G**) Schematic diagrams illustrating the importance of the intrapeptidyl H-bond between pTyr and (E) (pY−2)Asn for Spry2 or (F) (pY + 1)Arg for Met in the reverse orientation. In the absence of an intrapeptidyl H-bond (G), pTyr is disorientated and unable to dock into the pTyr-binding pocket on Cbl-TKB, as indicated by VEGFR.

chain of another residue in the peptide (Gunther *et al*, 2003; Cho *et al*, 2004; Frese *et al*, 2006). For the first case, the H-bond was not conserved among the different peptides complexed with the SH2 domain of Gads (Cho *et al*, 2004). In the Nck1-SH2:phospho-Tir and Nck2-SH2:phospho-Tir complexes, a degenerate peptide library scan showed a preference for hydrophobic residues over the interacting His at that position (Frese *et al*, 2006). The last complex model between the SH2 domain of PI3 kinase and the phospho-platelet-derived growth factor receptor peptide identified the preference for negatively charged residues N-terminal to the pTyr that were not essential for binding nor bound to its SH2 domain in the same manner as the conserved (pY−1)Asp binds to the TKB domain of c-Cbl in the Met family of proteins (Gunther *et al*, 2003).

Similarly, a search performed on known structures of PTB-binding peptides bearing the conserved D/NXpY and complexed with PTB domains also revealed that none of these binding conformations had the signature H-bond between the pTyr and the (pY−2)Asn. To the best of our knowledge, there are no conserved intrapeptidyl H-bonds between pTyr and its neighbouring Asn/Arg residues for SH2 and PTB domain-binding proteins. The formation of the intrapeptidyl H-bond in the D/NXpY, RpY or pYR (in the reverse) motifs is likely to be an essential specificity determinant for the recognition and docking of c-Cbl to its activated (phosphorylated) target proteins, whether before or during binding, and thus represents a mechanism that is unique to Cbl.

The most surprising finding from the crystal structure analyses is the reversed binding of the Met phosphopeptide. From these results, it would be presumed that other members of the Met subgroup also bind in the reverse direction. Interestingly, many of these family members have a proline residue at the pY−4 position, which would align with the (pY + 4)Pro residue in members of the PTK subgroup in the reverse orientation, and presumably increase their binding affinity to c-Cbl over Met.

There are several analogous situations where ligand binding can occur in two orientations of a specific domain. The SUMO-binding amino-acid sequence motif of RanBP2 and PIASX-P have been shown to bind to the surface of SUMO-1 in two different orientations (Reverter and Lima, 2005; Song *et al*, 2005) and similarly, the SH3 domain of Src can bind to proline-rich peptide ligands in two orientations (Feng *et al*, 1994). Furthermore, the PH domain of ArhGAP9 accommodates both PI(3,4)P2 and PI(4,5)P2 at 180° rotations via a shallow, noncanonical binding pocket (Ceccarelli *et al*, 2007). To our knowledge, however, there is no precedent for an SH2-domain-binding protein to bind to an SH2 domain in two orientations, and raises the question whether this may also occur. However, as the TKB domain accomplishes binding in conjunction with the 4H subdomain, its absence in other SH2 domains may restrict them from having such flexibility of binding.

The reverse binding of Met and the previously unidentified importance of the (pY−1)Arg in EGFR suggests that there is a larger cohort of Cbl-TKB domain targets, particularly those with an RpYD or RpYS/T motif. Searching the human proteome, we identified nearly 300 signal transduction proteins that contained the RYD sequence, and while it still remains to be determined whether these proteins are substrates for c-Cbl, this emphasizes the likelihood that other unrevealed

c-Cbl targets exist. In addition, the reversed binding of Met may reflect the need for Cbl to bind to targets in two orientations for steric reasons, as the DpYR residues in Met lie very close to the plasma membrane in the juxta-membrane domain. This reversed binding presumably allows the appropriate distance for the subsequent ligation of ubiquitin moieties. In contrast, the reversed binding may also allow c-Cbl to bind to proteins without resulting in their ubiquitination, expanding Cbl's role as an adapter protein in signalling complexes other than the APS family of proteins. Together, these findings provide not only further clues to Cbl's role in signalling but also the incentive to revisit previously identified binding motifs of other proteins, with the potential of redefining the apparently established paradigms of modular signalling.

Materials and methods

Plasmid constructs, cloning, expression and purification

The human c-Cbl-TKB domain (residues 25–351; Cbl-TKB) was cloned into pGEX4T-1 expression vector (GE Healthcare, Buckinghamshire, UK), overexpressed as a glutathione-*S*-transferase fusion protein and purified according to Meng *et al* (1999). The purified protein (1 mg/ml) was kept in a storage buffer (20 mM Na-HEPES pH 7.0, 0.2 M NaCl) and used for crystallization and ITC experiments. Wild-type full-length constructs of FLAG-tagged Spry2, Spry4, FGFR1, HA-tagged c-Cbl and EGFR have been described previously (Wong *et al*, 2001). pBABE-bleo human insulin receptor B (IR-B) plasmid 11211 was from Addgene (deposited by Dr C Ronald Kahn; Entingh *et al*, 2003). Rat myc-tagged APS was kindly provided by Dr W Langdon (University of Western Australia). c-Met was from Origene Technologies Inc. (Rockville, MD) and subcloned into PXJ40 vector (courtesy of Dr E Manser, Institute of Molecular and Cell Biology) using *Not*I/*Not*I restriction sites. Point mutations in FLAG–Spry2 and HA–Cbl were generated by site-directed mutagenesis using *Pfu* DNA polymerase from Promega (Madison, WI).

Complex formation and crystallization

Tyrosine phosphorylated peptides of Human Sprouty2 49–61 ($Spry2^{49-61}$), Sprouty4 69–81 ($Spry4^{69-81}$), EGFR 1063–1075 ($EGFR^{1063-1075}$), Syk tyrosine kinase 317–329 ($Syk^{317-329}$) and Met receptor 997–1009 ($Met^{997-1009}$), (Sigma Genosys) were reconstituted in Cbl-TKB storage buffer and incubated with the purified Cbl-TKB in two-fold molar excess at room temperature for 1 h for complex formation. The Cbl-TKB:peptide complexes were concentrated to 5 mg/ml in Amicon Ultra ultrafiltration devices (Millipore, Billerica, MA). Initial crystallization conditions were obtained using Hampton Research Screens (Hampton Research, CA). Each of the complexes crystallized in a different condition: Cbl-TKB:$Spry2^{49-61}$ complex crystals grew in 0.1 M Na-HEPES pH 7.5, 75 mM ammonium sulphate, 25% PEG3350 and 4% n-propanol; Cbl-TKB:$Spry4^{69-81}$ complex crystals grew in 0.2 Na/K tartrate and 20% PEG3350; Cbl-TKB:$EGFR^{1063-1075}$ crystals grew in 0.25 M sodium formate and 20% PEG3350; Cbl-TKB:$Syk^{317-329}$ crystals grew in 0.1 M Bis-Tris pH 6.5, 0.12 ammonium acetate and 18% PEG3350; Cbl-TKB:$Met^{997-1009}$ crystals grew in 0.15 M malic acid and 20% PEG3350.

Data collection, structure determination and refinement

The mother liquor was supplemented with 20% glycerol for cryo-protection. Data for the Cbl-TKB:$EGFR^{1063-1075}$ complex crystals were collected with an R-axis IV^{++} image plate detector mounted on a RU-H3RHB rotating anode generator (Rigaku Corp., Tokyo, Japan). Data for all other complex crystals were collected at X29 beamline in the National Synchrotron Light Source, Brookhaven National Laboratory (Upton, NY). Data were processed and scaled with HKL2000 (Otwinowski and Minor, 1997). The TKB domain of c-Cbl from the ZAP70:Cbl complex structure, PDB code 2CBL (Meng *et al*, 1999) was used as a starting model for molecular replacement solution using the programme Molrep (Vagin and Teplyakov, 1997). Subsequent rigid body refinement reduced the R-factors close to 30%. At this stage, the calculated difference electron density map clearly showed the presence of substrate peptides. Model building and refinement were carried out in O (Jones *et al*, 1991) and CNS (Brunger *et al*, 1998) programmes, respectively. After several cycles of map fitting and refinement the R-values were converged.

Protein Data Bank accession code

Coordinates and structure factors of all of the Cbl-TKB:phosphopeptide complexes have been deposited with RCSB Protein Data Bank with codes 3BUM (Spry2), 3BUN (Spry4), 3BUO (EGFR), 3BUW (Syk) and 3BUX (Met).

Isothermal titration calorimetry

Both phosphorylated and unphosphorylated peptides of $Spry2^{49-61}$, $Spry4^{69-81}$, $EGFR^{1063-1075}$, Src tyrosine kinase ($Src^{413-425}$), $Syk^{317-329}$, VEGFR ($VEGFR1^{1327-1338}$), $VEGFR1^{1327-1338}$ V1331N mutant, $VEGFR1^{1327-1338}$ L1332R mutant, c-$Met^{997-1009}$ and macrophage stimulating 1 receptor 1011–1023 ($Ron^{1011-1023}$) were titrated against 10.8 μM Cbl-TKB in a VP-ITC microcalorimeter (Microcal, Northampton, UK) performed under identical conditions. All peptides were maintained in storage buffer at 22°C and titrated in 10 μl injections into Cbl-TKB also in storage buffer. The heat of dilution, determined by titrating peptide into storage buffer, was subtracted from the raw data before curve fitting and refinement. The dissociation constants of Cbl-TKB to the various peptides were determined by least squares method and the binding isotherm was fitted using Origin v7.0 (Microcal) assuming a single-site binding model. All measurements were repeated at least twice for verification.

Antibodies and reagents

Rabbit anti-Cbl (C-15), anti-FGFR1 (C-15), anti-Myc (A-14), anti-EGFR (1005), anti-Grb2 (C-23) anti-insulin Rβ (C-19), and mouse anti-Myc (9E10) and anti-EGFR (528) were from Santa Cruz Biotechnology (Santa Cruz, CA). Mouse anti-c-Cbl and anti-Grb2 were from Transduction Laboratories (Lexington, KY). Mouse anti-Met (25H2) was from Cell Signaling Technology (Beverly, MA) and rat anti-HA was from Roche Diagnostics (Basel, Switzerland). Mouse anti-FLAG, rabbit anti-HA, goat anti-HGFR, agarose-conjugated anti-FLAG M2 beads, recombinant human (rh) FGF-2, and goat anti-rabbit-HRP and anti-mouse-HRP were from Sigma Aldrich (St Louis, MO). rhEGF was from Upstate Biotechnology Inc. (Lake Placid, NY) and rhHGF was from Calbiochem (EMD Chemicals, San Diego, CA).

Cell culture and transfection

Human embryonic kidney 293T cells (ATCC, Manassas, VA) were cultured, maintained and transfected as previously outlined (Wong *et al*, 2001; Yusoff *et al*, 2002). For binding between WT c-Cbl and Spry2 point mutants, cells were transfected with FLAG–Spry2 and FGFR1 for 16 h and precipitated for endogenous c-Cbl. Cells were transfected with WT c-Cbl or Cbl bearing various point mutations with WT Spry2, Spry4, Met, EGFR or APS with IR-B for 16 h or 40 h (in the case of Met). For Spry2, Spry4 or Grb2 experiments, cells were stimulated with 20 ng/ml FGF-2 for 2 h. Met-transfected cells were stimulated with 100 ng/ml of HGF 30 min. EGFR-transfected cells were stimulated with 100 ng/ml of EGF for 10 min. For APS experiments, cells were stimulated as described previously (Hu and Hubbard, 2005).

Cell lysis, immunoprecipitation and western blotting

Cells lysis and western blotting was performed as described previously (Wong *et al*, 2001). Immunoprecipitation was carried out using FLAG–M2 beads (Spry2), anti-HA (c-Cbl), anti-HGFR (Met), anti-myc (APS) or anti-EGFR (528).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank Dr Louis Payet for initial contributions and Dr Robert C Robinson, Dr Chee Wai Fong, Dr Permeen Yusoff, Dr Jormay Lim and Ms Sumana Chandramouli for critical review of the paper.

Data for this study were measured at beam line X29 of the National Synchrotron Light Source, Brookhaven National Laboratories (Upton, NY). We thank Dr Anand Saxena for assistance with data collection. We also thank Onmiway Solutions Pte Ltd for the Genbank sequence search. This work was supported by a BMRC grant (R154000273305), Singapore's Agency for Science Technology and Research (A-STAR) and the Institute of Molecular and Cell Biology (IMCB), Singapore. CN is research scholarship recipient from the National University of Singapore (Ministry of Education, Singapore).

References

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98:** 10037–10041

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* **54** (Part 5)**:** 905–921

Ceccarelli DF, Blasutig IM, Goudreault M, Li Z, Ruston J, Pawson T, Sicheri F (2007) Non-canonical interaction of phosphoinositides with pleckstrin homology domains of Tiam1 and ArhGAP9. *J Biol Chem* **282:** 13864–13874

Chiang SH, Baumann CA, Kanzaki M, Thurmond DC, Watson RT, Neudauer CL, Macara IG, Pessin JE, Saltiel AR (2001) Insulin-stimulated GLUT4 translocation requires the CAP-dependent activation of TC10. *Nature* **410:** 944–948

Cho S, Velikovsky CA, Swaminathan CP, Houtman JC, Samelson LE, Mariuzza RA (2004) Structural basis for differential recognition of tyrosine-phosphorylated sites in the linker for activation of T cells (LAT) by the adaptor Gads. *EMBO J* **23:** 1441–1451

DeLano WL (2002) *The PyMOL Molecular Graphics System*. San Carlos, CA, USA: DeLano Scientific

Entingh AJ, Taniguchi CM, Kahn CR (2003) Bi-directional regulation of brown fat adipogenesis by the insulin receptor. *J Biol Chem* **278:** 33377–33383

Feng S, Chen JK, Yu H, Simon JA, Schreiber SL (1994) Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science* **266:** 1241–1247

Frese S, Schubert WD, Findeis AC, Marquardt T, Roske YS, Stradal TE, Heinz DW (2006) The phosphotyrosine peptide binding specificity of Nck1 and Nck2 Src homology 2 domains. *J Biol Chem* **281:** 18236–18245

Gunther UL, Weyrauch B, Zhang X, Schaffhausen B (2003) Nuclear magnetic resonance structure of the P395S mutant of the N-SH2 domain of the p85 subunit of PI3 kinase: an SH2 domain with altered specificity. *Biochemistry* **42:** 11120–11127

Hall AB, Jura N, DaSilva J, Jang YJ, Gong D, Bar-Sagi D (2003) hSpry2 is targeted to the ubiquitin-dependent proteasome pathway by c-Cbl. *Curr Biol* **13:** 308–314

Hu J, Hubbard SR (2005) Structural characterization of a novel Cbl phosphotyrosine recognition motif in the APS family of adapter proteins. *J Biol Chem* **280:** 18943–18949

Joazeiro CA, Wing SS, Huang H, Leverson JD, Hunter T, Liu YC (1999) The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase. *Science* **286:** 309–312

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47** (Part 2)**:** 110–119

Laskowski RA, MacAuthur MW, Moss DS, Thornton JM (1993) Procheck: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26:** 283–291

Levkowitz G, Waterman H, Ettenberg SA, Katz M, Tsygankov AY, Alroy I, Lavi S, Iwai K, Reiss Y, Ciechanover A, Lipkowitz S, Yarden Y (1999) Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1. *Mol Cell* **4:** 1029–1040

Liu BA, Jablonowski K, Raina M, Arce M, Pawson T, Nash PD (2006) The human and mouse complement of SH2 domain proteins-establishing the boundaries of phosphotyrosine signaling. *Mol Cell* **22:** 851–868

Liu J, Kimura A, Baumann CA, Saltiel AR (2002) APS facilitates c-Cbl tyrosine phosphorylation and GLUT4 translocation in response to insulin in 3T3-L1 adipocytes. *Mol Cell Biol* **22:** 3599–3609

Lupher Jr ML, Songyang Z, Shoelson SE, Cantley LC, Band H (1997) The Cbl phosphotyrosine-binding domain selects a D(N/D)XpY motif and binds to the Tyr292 negative regulatory phosphorylation site of ZAP-70. *J Biol Chem* **272:** 33140–33144

Machida K, Mayer BJ (2005) The SH2 domain: versatile signaling module and pharmaceutical target. *Biochim Biophys Acta* **1747:** 1–25

Meng W, Sawasdikosol S, Burakoff SJ, Eck MJ (1999) Structure of the amino-terminal domain of Cbl complexed to its binding site on ZAP-70 kinase. *Nature* **398:** 84–90

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276** (Part A)**:** 307–326

Penengo L, Rubin C, Yarden Y, Gaudino G (2003) c-Cbl is a critical modulator of the Ron tyrosine kinase receptor. *Oncogene* **22:** 3669–3679

Peschard P, Fournier TM, Lamorte L, Naujokas MA, Band H, Langdon WY, Park M (2001) Mutation of the c-Cbl TKB domain binding site on the Met receptor tyrosine kinase converts it into a transforming protein. *Mol Cell* **8:** 995–1004

Peschard P, Ishiyama N, Lin T, Lipkowitz S, Park M (2004) A conserved DpYR motif in the juxtamembrane domain of the Met receptor family forms an atypical c-Cbl/Cbl-b tyrosine kinase binding domain binding site required for suppression of oncogenic activation. *J Biol Chem* **279:** 29565–29571

Peschard P, Park M (2003) Escape from Cbl-mediated downregulation: a recurrent theme for oncogenic deregulation of receptor tyrosine kinases. *Cancer Cell* **3:** 519–523

Reverter D, Lima CD (2005) Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. *Nature* **435:** 687–692

Rubin C, Litvak V, Medvedovsky H, Zwang Y, Lev S, Yarden Y (2003) Sprouty fine-tunes EGF signaling through interlinked positive and negative feedback loops. *Curr Biol* **13:** 297–307

Schmidt MH, Dikic I (2005) The Cbl interactome and its functions. *Nat Rev Mol Cell Biol* **6:** 907–919

Song J, Zhang Z, Hu W, Chen Y (2005) Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. *J Biol Chem* **280:** 40122–40129

Tamagnone L, Artigiani S, Chen H, He Z, Ming GI, Song H, Chedotal A, Winberg ML, Goodman CS, Poo M, Tessier-Lavigne M, Comoglio PM (1999) Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates. *Cell* **99:** 71–80

Thien CB, Langdon WY (2005) c-Cbl and Cbl-b ubiquitin ligases: substrate diversity and the negative regulation of signalling responses. *Biochem J* **391** (Part 2)**:** 153–166

Vagin A, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. *J Appl Cryst* **30:** 1022–1025

Wong ES, Fong CW, Lim J, Yusoff P, Low BC, Langdon WY, Guy GR (2002) Sprouty2 attenuates epidermal growth factor receptor ubiquitylation and endocytosis, and consequently enhances Ras/ERK signalling. *EMBO J* **21:** 4796–4808

Wong ES, Lim J, Low BC, Chen Q, Guy GR (2001) Evidence for direct interaction between Sprouty and Cbl. *J Biol Chem* **276:** 5866–5875

Yusoff P, Lao DH, Ong SH, Wong ES, Lim J, Lo TL, Leong HF, Fong CW, Guy GR (2002) Sprouty2 inhibits the Ras/MAP kinase pathway by inhibiting the activation of Raf. *J Biol Chem* **277:** 3195–3201

