Article

IRF2BP2 binds to a conserved RxSVI motif of protein partners and regulates megakaryocytic differentiation

Guanchao Wang [1], Tiantian Lu[1], Lei Zhang[1,2,3] & Jianping Ding [1,2] ✉

IRF2BP2 is a transcriptional coregulator that plays diverse regulatory roles in various cellular processes in either IRF2-dependent or IRF2-independent manner through interactions with protein partners via its RING domain; however, the underlying molecular mechanisms remain unclear. In this study, we conduct a motif discovery search on the sequences of interacting proteins IRF2 and VGLL4 of IRF2BP2 and identify a conserved RxSVI motif. Biochemical and structural data reveal that the RING domain binds to the motif-containing peptides of IRF2 and VGLL4 with comparable affinities and in a similar manner. The motif-containing peptides tend to form a short loop along with a short β-strand, which facilitates effective recognition and tight binding by the RING domain. Further exploration of this motif in the human proteome identifies the transcription factor ZBTB16 as an interacting protein of IRF2BP2. Biochemical, structural, and cell biological data demonstrate that the RING domain binds to the motif-containing peptide of ZBTB16 in a manner similar to that of IRF2 and VGLL4. Moreover, IRF2BP2 plays a crucial regulatory role in megakaryocytic differentiation through interaction with ZBTB16. These findings elucidate the molecular basis for how IRF2BP2 can engage with different protein partners, thereby exerting diverse regulatory functions in many cellular processes.

The interferon regulatory factor 2 binding protein (IRF2BP) family consists of a group of transcriptional regulators initially identified as IRF2 binding proteins, including IRF2BP1, IRF2BP2, and IRF2BPL[1]. IRF2BP2 was first recognized as a corepressor of the transcription factor IRF2, which negatively regulates the expression of interferon-responsive genes[1]. Subsequent research revealed that IRF2BP2 is widely expressed in various normal and tumor cells and tissues, some of which lack IRF2, suggesting that it functions as a general transcriptional cofactor in both IRF2-dependent and IRF2-independent manners[2]. More recent studies have shown that IRF2BP2 is involved in regulating numerous cellular processes, including proliferation, apoptosis, differentiation, migration, immune response, inflammation, and angiogenesis [see review by Pastor[3]]. However, the underlying molecular mechanisms by which IRF2BP2 exerts its diverse regulatory functions in these processes remain largely unknown.

IRF2BP2 consists of an N-terminal C4 type zinc finger (ZnF) domain (residues 1-72), a C-terminal C3HC4 type real interesting new gene (RING) domain (residues 498-578), and a long linker which contains a nuclear localization signal (Fig. 1A)[1,4]. The ZnF domain of IRF2BP2 can interact with IRF2BP1 and IRF2BPL to form a protein complex that mediates the transcriptional repression of FASTKD2[5]. The RING domain of IRF2BP2 is primarily responsible for its interactions with various protein partners, enabling IRF2BP2 to exert different regulatory roles in target gene expression. It interacts with the C-terminal domain of IRF2 to function as a transcriptional corepressor, thereby inhibiting both enhancer-activated and basal transcription[1].

[1]Key Laboratory of RNA Innovation, Science and Engineering, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai, China. [2]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [3]Sheng Yushou Center of Cell Biology and Immunology, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China. ✉e-mail: jpding@sibcb.ac.cn
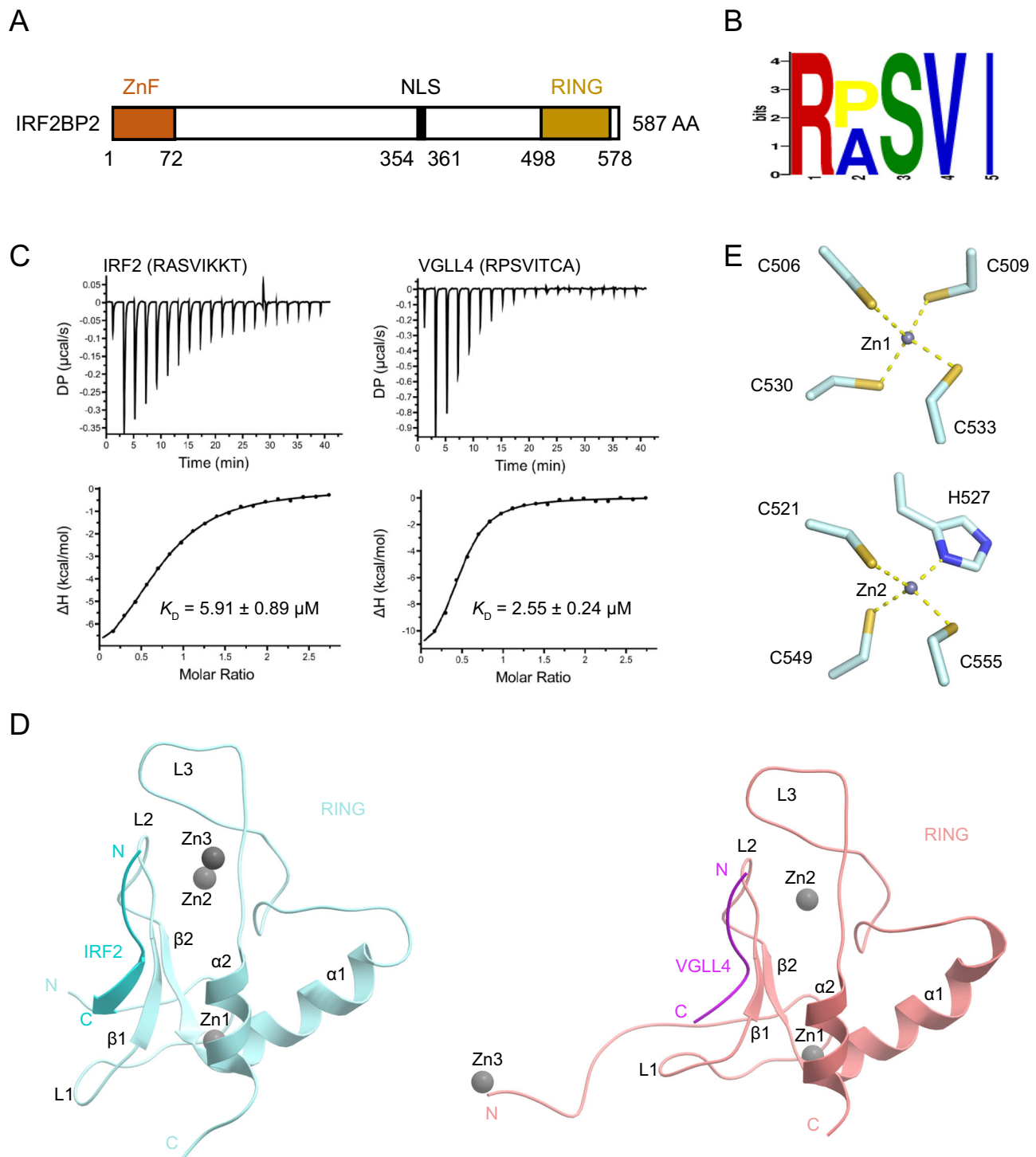
Fig. 1 | The RING domain of IRF2BP2 recognizes and binds to a conserved RxSVI motif of IRF2 and VGLL4. A Schematic domain diagram of IRF2BP2. **B** A conserved RxSVI sequence motif of IRF2 and VGLL4 was identified by GLAM2. **C** Binding affinities of the RING domain of IRF2BP2 with the RxSVI motif-containing peptides of IRF2 (left) and VGLL4 (right) as measured using isothermal titration calorimetry (ITC). Shown here is one representative result from three independent experiments for each peptide. **D** The overall structure of the RING-IRF2 (cyan/light-blue) and RING-VGLL4 (magenta/pink) complexes in ribbon representation. The three Zn^{2+} ions are shown as gray spheres. The secondary structure elements are indicated. **E** Structures of the two Zn^{2+}-binding sites of the C3HC4 RING motif in the RING-IRF2 complex. Up: Zn1-binding site. Down: Zn2-binding site.

Additionally, it interacts with the C-terminal domain of NFAT1 to repress NFAT1-mediated transcriptional activity[6], and interacts with ETO2 (also known as CBFA2T3) to recruit the NCOR1/SMRT complex, thereby inhibiting the expression of most erythroid genes[7]. IRF2BP2 also plays a role in regulating the Hippo pathway through its interaction with VGLL4[8–10]. Overexpression of IRF2BP2 in liver cancer cells

HepG2 and Huh7 inhibits YAP activity and consequently cell growth through a feedback loop. Mechanistically, IRF2BP2 interacts with VGLL4 via its RING domain, stabilizing VGLL4 by inhibiting its ubiquitination and further reducing YAP activity[8].

Clinically, several mutations in the RING domain of IRF2BP2 have been reported. The Q536delinsL* mutation, which results in a deletion

mutant lacking residues 536-587, was identified in patients with chronic diarrhea, severe eczema, and anemia[11]. The Q540* nonsense mutation, leading to a truncated protein, was found in a patient with recurrent respiratory infections during childhood, colitis, and rheumatoid arthritis[12]. The S551N mutation was identified in a family with common variable immunodeficiency disorder, and this mutation was shown to affect B cell differentiation[13]. Although the RING domain of IRF2BP2 is known to play crucial roles in its interaction with various partners and the regulation of diverse cellular processes, the molecular mechanisms by which the RING domain binds to different protein partners remain unexplored.

Zinc finger and BTB domain-containing protein 16 (ZBTB16), also known as promyelocytic leukemia zinc finger protein (PLZF), is a potent transcription factor that plays important roles in various cellular processes, including cell cycle, apoptosis, glucocorticoid response, immunity, hematopoiesis, and cancers [see reviews[14–18]]. Notably, ZBTB16 is highly expressed in hematopoietic progenitor cells and plays a crucial role in megakaryopoiesis. ZBTB16 binds to the promoter of miR-146a, inhibiting its expression and subsequently activating the translation of downstream CXCR4 and promoting megakaryocyte development[19]. Additionally, ZBTB16 may form a complex with GATA1, thereby playing a stimulating role in megakaryocyte development[20]. Despite advances in understanding the diverse regulatory roles of ZBTB16 in the development, proliferation, and differentiation of various cell types, little is known about the underlying molecular mechanisms that govern its regulation to achieve these functions.

In this study, we conducted a bioinformatic analysis on the sequences of two well-characterized interacting proteins of IRF2BP2, namely IRF2, and VGLL4, and identified a conserved RxSVI motif. Biochemical and structural data revealed that the RING domain of IRF2BP2 binds to the motif-containing peptides of IRF2 and VGLL4 with comparable affinities and in a similar manner. The RxSVI motif-containing peptides tend to form a short loop along with a short β-strand, which facilitates effective recognition and tight binding by the RING domain. Further exploration of this motif in the human proteome identified ZBTB16 as an interacting partner of IRF2BP2. Additional biochemical and structural data demonstrated that the RING domain of IRF2BP2 binds to the motif-containing peptide of ZBTB16 in a manner similar to that of IRF2 and VGLL4. Cell biological data indicated that the interaction between IRF2BP2 and ZBTB16 plays a crucial role in regulating megakaryocytic differentiation. In summary, we show that the RING domain of IRF2BP2 recognizes and binds to the conserved RxSVI motif of several interacting proteins in a similar manner, and that IRF2BP2 is involved in the regulatory process of megakaryocytic differentiation through its interaction with ZBTB16.

Results

The RING domain of IRF2BP2 binds to a conserved RxSVI motif of protein partners IRF2 and VGLL4

Previous studies have shown that IRF2BP2, as a transcriptional cofactor, can interact with various proteins via its RING domain (residues 498-578) (Fig. 1A) to exert distinct regulatory functions, with IRF2 and VGLL4 being two well-characterized interacting proteins[1,8]. To investigate whether the RING domain of IRF2BP2 interacts with these protein partners in a similar manner, we conducted a motif discovery search on the sequences of IRF2 and VGLL4 using the bioinformatics tool GLAM2[21]. This search identified a conserved RxSVI motif (x represents any amino acid residue) (Fig. 1B). To confirm this finding, we purified the RING domain of IRF2BP2 (Supplementary Fig. S1), and synthesized an IRF2 peptide (residues 331-338, RASVIKKT) and a VGLL4 peptide (residues 161-168, RPSVITCA) both containing the conserved motif. We then analyzed the interactions between the RING domain and these peptides using isothermal titration calorimetry (ITC). The results indicated that the RING domain binds to both

peptides with comparable binding affinities (the mean K_D values of three independent measurements are 6.45 μM and 2.90 μM, respectively) (Fig. 1C, Supplementary Table S1).

Given that the peptides used in the binding assay were relatively short (8 residues), we synthesized two longer peptides: IRF2-L (residues 325-343) and VGLL4-L (residues 156-171) based on the sequence conservations of these proteins in different species (human, mouse, chicken, and zebrafish) (Supplementary Fig. S2A, B). We then measured their binding affinities with the RING domain. The results indicated that these longer peptides exhibited slightly stronger binding affinities (K_D values) to the RING domain compared to the shorter peptides (increased by 3.9-fold and 2.1-fold, respectively) (Supplementary Table S1, Fig. S3A, C). This suggests that the N- and/or C-terminal ends of the longer peptides may also contribute to the interaction with the RING domain. Together, these findings indicate that the conserved RxSVI motif of IRF2 and VGLL4 plays an important role in their interactions with the RING domain of IRF2BP2.

Structures of the IRF2BP2 RING domain in complexes with the RxSVI motif-containing peptides of IRF2 and VGLL4

To investigate the molecular basis of this interaction, we co-crystallized the RING domain of IRF2BP2 in complexes with the shorter peptides of IRF2 and VGLL4, and determined the crystal structures of these complexes (RING-IRF2 and RING-VGLL4) at 1.45 Å resolution and 1.59 Å resolution, respectively (Table 1). Unfortunately, attempts to co-crystallize the RING domain with the longer peptides of IRF2 and VGLL4 did not yield any crystals. In both structures, the asymmetric unit (ASU) contains one complex molecule, with well-defined electron density for most residues of the RING domain and the first six residues of each peptide (Table 1, Fig. 1D, Supplementary Fig. S4A, B). In addition, there are three metal ions bound to the RING domain, which have been assigned as Zn^{2+} ions due to the presence of 0.1 mM $ZnCl_2$ in the purification and storage buffers of the protein. The refinement of these Zn^{2+} ions yielded reasonable B-factors (Table 1).

The RING domain of IRF2BP2 adopts a typical C3HC4 RING fold, consisting of two antiparallel β-strands (β1 and β2) followed by one α-helix (α1)[22] (Fig. 1D). Compared to most of other RING domains, the RING domain of IRF2BP2 contains an additional α-helix (α2) at the C-terminus and three long loops (L1-L3): the L1 loop at the N-terminus, the L2 loop connecting β1 and β2, and the L3 loop connecting α1 and α2. Each of the two Zn-binding sites of the C3HC4 RING fold is bound with a Zn^{2+} ion: one Zn^{2+} ion (Zn1) is coordinated by Cys506 (L1), Cys509 (L1), Cys530 (β2) and Cys533 (α1), while the other Zn^{2+} ion (Zn2) by Cys521 (L2), His527 (L2), Cys549 (L3) and Cys555 (L3) (Fig. 1E). The binding of these two Zn^{2+} ions appears to stabilize the overall structure of the RING domain. In the RING-IRF2 complex, a third Zn^{2+} ion (Zn3) is involved in crystal packing, which is coordinated by Ala565 of RING, Arg1 of the peptide, His517 of a symmetry-related RING, and a water molecule (Supplementary Fig. S5A). In the RING-VGLL4 complex, Zn3 is also involved in crystal packing but at a distinct site, which is coordinated by Ser497 (both mainchain amide and carbonyl) of RING, His510 of a symmetry-related RING, and three water molecules (Supplementary Fig. S5B). Since Zn3 binds to different sites at the inter-molecular interface in different structures, it is unlikely to have a functional role.

Comparison of the RING domain in the two structures revealed no substantial conformational differences (RMSD of 0.94 Å for 74 Cα atoms) except for the N-terminal L1 loop (Supplementary Fig. S6A). In the RING-VGLL4 complex, six additional residues of the L1 loop (Ser497-Ala502) could be modeled compared to the RING-IRF2 complex. Furthermore, part of the L1 loop (Leu513-His517) exhibited conformational changes due to differences in crystal packing and interactions with the peptide. In the RING-IRF2 complex, only the sidechain of His517 from the L1 loop of a symmetry-related RING molecule is coordinated with Zn3. In contrast, in the RING-VGLL4 complex, the sidechains of Glu514 and Asp515 of the L1 loop make

Table 1 | Summary of diffraction data and structure refinement statistics

PDB code	RING-IRF2 8YTG	RING-VGLL4 8YTF	RING-ZBTB16 8YTH
Diffraction data			
Wavelength (Å)	0.9792	0.9792	0.9792
Space group	$P2_1$	$P2_12_12_1$	$P6_4$
Cell parameters			
a, b, c (Å)	26.7, 39.9, 36.2	33.4, 43.2, 55.7	47.3, 47.3, 57.5
α, β, γ (°)	90, 108.2, 90	90, 90, 90	90, 90, 120
Resolution (Å)	26.06-1.45 (1.50-1.45)[a]	23.43-1.59 (1.65-1.59)[a]	33.37-2.40 (2.49-2.40)[a]
Total reflections	73,644 (6044)	140,905 (12,410)	58,126 (5917)
Unique reflections	12,800 (1,270)	11,128 (980)	2888 (285)
Average I/σ(I)	10.3 (1.8)	17.7 (5.6)	22.4 (7.4)
Average redundancy	5.8 (4.8)	12.7 (12.6)	20.1 (20.6)
Completeness (%)	98.8 (98.5)	98.1 (91.8)	99.6 (99.3)
R_{merge}[b]	0.163 (0.406)	0.085 (0.303)	0.161 (0.434)
$CC_{1/2}$ (%)	98.9 (76.3)	99.9 (97.9)	99.8 (99.1)
Refinement and structure model			
No. of reflections ($Fo > 0σ(Fo)$)	12,781	11,054	2879
Working set	12,156	10,535	2767
Test set	625	519	112
R_{work}/R_{free} (%)[c]	16.5/18.9	15.3/17.3	24.3/28.2
No. of atoms	782	777	561
Protein	621	646	557
Metal ion	3	3	3
Water	158	128	1
No. of protein residues	82	86	76
Observed residues	A: 503-578; B: 1-6	A: 497-576; B: 1-6	A: 503-542, 547-558, 560-576; B: 1-7
Wilson B-factor (Å²)	9.46	10.82	50.63
Average B-factor (Å²)	14.12	14.89	59.52
Protein	11.70	12.67	59.62
Metal ion	12.00	11.54	51.08
Zn1/Zn2/Zn3	10.63/7.00/18.38	10.00/9.10/15.53	61.38/48.27/43.59
Water	23.69	26.17	29.93
R.m.s. deviations			
Bond lengths (Å)	0.006	0.007	0.004
Bond angles (°)	0.890	1.060	0.760
Ramachandran plot (%)			
Favored	100.00	100.00	97.06
Allowed	0.00	0.00	2.94
Outliers	0.00	0.00	0.00

[a]Numbers in parentheses represent the highest resolution shell.
[b]$R_{merge} = \sum_{hkl}\sum_i |I_i (hkl) - \langle I (hkl) \rangle| / \sum_{hkl}\sum_i I_i (hkl)$.
[c]$R = \sum_{hkl}||Fo| - |Fc|| / \sum_{hkl}|Fo|$.

hydrophilic interactions with Arg1 of the peptide from a symmetry-related RING-VGLL4 complex, while the sidechain of His517 of the L1 loop engages in hydrophobic interactions with Val558 from a symmetry-related RING.

A structural homology search of the IRF2BP2 RING domain against the Protein Data Bank (PDB) using the DALI server[23] found that the most similar structures are the RING domains of E3 ubiquitin-protein ligases, such as the RING1 domain of PARKIN (PDB code 5N2W, Z-score of 5.3 and RMSD of 1.74 Å for 60 Cα atoms) (Supplementary Fig. S6B).
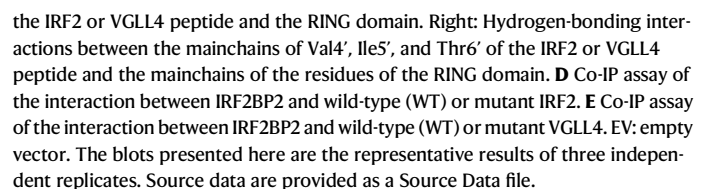
However, there are some notable structural differences. In the IRF2BP2 RING domain, the α1 and α2 helices are connected by a long loop (L3), whereas in the PARKIN RING1 domain, the linker between α1 and α2 forms two β-strands (β3 and β4) (Supplementary Fig. S6B). Additionally, the N-terminal region of the IRF2BP2 RING domain forms a long loop (L1), whereas the N-terminal region of the PARKIN RING1 domain forms an extra β-strand (β0) which folds antiparallel to the β1 strand (Supplementary Fig. S6C).

The RING domain of IRF2BP2 binds to the RxSVI motif-containing peptides of IRF2 and VGLL4 in a similar manner

In the RING-IRF2 complex, the IRF2 peptide forms a short loop (residues 1'–3') followed by a short β-strand (residues 4'−6') (residues of the peptide are designated with a prime to distinguish them from those of the RING domain) (Figs. 1D and 2A). The conserved Arg1', Ser3', Val4', and Ile5' of the peptide engage in numerous hydrophilic and hydrophobic interactions with several residues of the RING domain (Fig. 2B). Specifically, the sidechain of Arg1' forms a salt bridge with the sidechain of Glu570 and a hydrogen bond with the mainchain carbonyl of Ala565 of RING. Additionally, the sidechain of Arg1' forms a hydrogen bond with the mainchain carbonyl of Ala2' to stabilize the peptide's conformation. The sidechain of Ser3' forms a hydrogen bond with the sidechain of Glu570. The mainchain amide of Val4' forms a hydrogen bond with the sidechain of Glu570. Moreover, the hydrophobic sidechains of Val4' and Ile5' are embedded in a hydrophobic groove composed of Val519, Pro522, Phe531, Trp564, Phe566, Thr573, and Ile574 and make hydrophobic interactions with these residues (Fig. 2B). Furthermore, the mainchains of residues 4'−6' of the peptide participate in hydrogen-bonding interactions with the mainchains of residues His517-Gln520 of the β1 strand of RING. Specifically, the mainchain carbonyl of Val4' forms a hydrogen bond with the mainchain amide of Gln520; the mainchain amide of Lys6' forms a hydrogen bond with the mainchain carbonyl of Phe518; and the carboxyl of Lys6' forms two hydrogen bonds with the mainchain amide of Phe518 and the sidechain of His517 (Fig. 2B). Consequently, these residues of the peptide are assigned to form a short β-strand which forms a short antiparallel β-sheet with the β1 strand of RING (Fig. 2B).

In the RING-VGLL4 complex, the VGLL4 peptide adopts a loop conformation (Figs. 1D and 2A). Compared to the RING-IRF2 complex, the conserved Arg1', Ser3', Val4', and Ile5' of the VGLL4 peptide assume the same conformations as those of the IRF2 peptide and interact with the residues of the RING domain in a similar manner (Fig. 2C). However, while the mainchain carbonyl of Val4' and the mainchain amide of Thr6' form analogous hydrogen-bonding interactions with the RING domain as observed in the RING-IRF2 complex, the sidechain of Thr6' adopts a slightly different conformation and thus forms two hydrogen bonds: one with the mainchain amide of Phe518 and another with the mainchain carbonyl of Thr516. Additionally, the carboxyl group of Thr6' also adopts a slightly different conformation, which forms a hydrogen bond with the sidechain of His517 but not with the mainchain amide of Phe518 (Fig. 2C). Due to the conformational difference of Thr6', residues 4'−6' of the VGLL4 peptide are assigned to form a loop rather than a short β-strand. This conformational variation in Thr6' may stem from the conformational differences in part of the N-terminal loop of RING (Leu513-His517), which is involved in the intermolecular interactions. It is possible that, in the absence of crystal packing constraints, the VGLL4 peptide might adopt a conformation similar to that of the IRF2 peptide, featuring a short loop and a short β-strand.

In the structures of the RING-peptide complexes, the first half (residues 1'−3') of the peptide forms a short loop and the second half (residues 4'−6' or 7') forms or has the potential to form a short β-strand. This short β-strand folds along the β1 strand of the RING domain, creating an antiparallel β-sheet. This mode of intermolecular interaction is reminiscent of the intramolecular interactions observed

Fig. 2 | Interactions between the RxSVI motif-containing peptides of IRF2 and VGLL4 and the RING domain of IRF2BP2. A Electrostatic surface representation of the RING domain with the bound IRF2 (left) or VGLL4 (right) peptide shown as a stick model. The peptide binds to a surface groove of the RING domain. The unit of electrostatic potential is kT/e. **B** Interactions between the IRF2 peptide and the RING domain. **C** Interactions between the VGLL4 peptide and the RING domain. Left: Hydrophilic interactions between Arg1' and Ser3' of the IRF2 or VGLL4 peptide and the RING domain. Middle: Hydrophobic interactions between Val4' and Ile5' of the IRF2 or VGLL4 peptide and the RING domain. Right: Hydrogen-bonding interactions between the mainchains of Val4', Ile5', and Thr6' of the IRF2 or VGLL4 peptide and the mainchains of the residues of the RING domain. **D** Co-IP assay of the interaction between IRF2BP2 and wild-type (WT) or mutant IRF2. **E** Co-IP assay of the interaction between IRF2BP2 and wild-type (WT) or mutant VGLL4. EV: empty vector. The blots presented here are the representative results of three independent replicates. Source data are provided as a Source Data file.

in the RING1 domain of PARKIN. Notably, compared to the IRF2BP2 RING domain, the N-terminal region of the PARKIN RING1 domain forms an additional β-strand (β0) which forms an antiparallel β-sheet with β1 of the RING1 domain (Supplementary Fig. S6B). In particular, the sidechains of Leu228 and Ile229 in the β0 strand are embedded in a hydrophobic groove formed by β1 and α2 of the RING1 domain (Supplementary Fig. S6C). This similarity suggests that the ability of the RxSVI motif in interacting proteins to form a loop plus a short β-strand might be essential for ensuring that the RxSVI motif is effectively recognized and tightly bound by the RING domain.

In the RING-peptide complexes, the conserved Arg1', Val4' and Ile5' of the RxSVI motif are situated within a shallow groove formed by β1 and α2 of the RING domain. These residues engages in extensive hydrophilic and hydrophobic interactions with the RING domain through both their sidechains and mainchains. In contrast, the sidechain of Ser3' is oriented toward the solvent and forms only a hydrophilic interaction with the sidechain of Glu570 of the RING domain. This observation suggests that Arg1', Val4', and Ile5' of the motif are strictly conserved, while Ser3' may not be, indicating that the conserved RxSVI motif could extend to the RxxVI motif.

To validate the functional role of the key residues in the RxSVI motif of IRF2 and VGLL4 in their interactions with IRF2BP2, we first mutated the conserved residues of the motif in the IRF2 and VGLL4 peptides and measured their binding affinities with the RING domain of IRF2BP2 using ITC. The results indicated that mutations of either Arg1' or Val4' and Ile5' to Ala in both the IRF2 and VGLL4 peptides abolished their interactions with the RING domain, while mutation of Ser3' to Gly in both peptides led to a moderate decrease in binding affinities (Fig. S3B and D, Supplementary Table S1). Subsequently, we mutated the corresponding residues of the motif in full-length IRF2 and VGLL4 and conducted co-IP assays, and the results indicated that mutations of either Arg1' or Val4' and Ile5' to Ala in IRF2 and VGLL4 also disrupted their interactions with IRF2BP2 (Fig. 2D, E). Both in vitro and in vivo binding assay results confirmed the significance of the RxSVI motif in these proteins for their interactions with IRF2BP2. This is consistent with previous data showing that mutations of Ser163, Val164, and Ile165 in VGLL4 (corresponding to Ser3', Val4', and Ile5' of the RxSVI motif) to Ala disrupted the VGLL4-IRF2BP2 interaction, leading to reduced inhibition of VGLL4 on the YAP-induced TEAD4 reporter activity[8].

In summary, our structural and biochemical data demonstrate that the RING domain of IRF2BP2 binds to the conserved RxSVI motif of IRF2 and VGLL4 in a similar manner. The RxSVI motif-containing peptides tend to form a short loop plus a short β-strand, which facilitates effective recognition and tight binding by the RING domain. Additionally, residues Arg1', Val4', and Ile5' of the RxSVI motif play critical roles, while Ser3' a less critical role in their interactions with IRF2BP2, suggesting that the conserved RxSVI motif may be extended to the RxxVI motif.

ZBTB16 is an interacting protein of IRF2BP2

Based on these findings, we speculated that IRF2BP2 may interact with other proteins containing the RxSVI motif through its RING domain. To explore this possibility, we employed the bioinformatics tool FIMO to search for additional RxSVI motif-containing proteins in the human proteome[24]. This search identified a total of 1113 hits, of which 834 proteins are functionally annotated and involved in various biological processes (Fig. 3A). Given that previous studies have indicated that IRF2BP2 acts as a transcriptional coregulator, we focused on proteins within the gene-specific transcriptional regulator group. Among the 87 proteins in this category, we selected 6 representative proteins to conduct bidirectional co-IP assays in HEK 293 T cells, where both IRF2BP2 and the target protein were overexpressed. We used Flag-tagged IRF2BP2 to pull-down HA-tagged target proteins and Flag-tagged target proteins to pull-down Myc-

tagged IRF2BP2, respectively, to examine their interactions in vivo (Fig. 3B). The results revealed that among the examined proteins, ZBTB16 exhibited interactions with IRF2BP2 in both directions (Fig. 3C), ZNF774, ARID5B, USP36, and ZDBF2 showed interactions in one direction, and TGIF1 did not display detectable interactions in either direction (Supplementary Fig. S7). Additionally, we performed a co-IP assay in HEL cells without overexpressing either IRF2BP2 or ZBTB16, and the results clearly demonstrated that ZBTB16 can pull-down IRF2BP2, thereby validating the endogenous interaction between the two proteins (Fig. 3D). Moreover, we conducted a mammalian two-hybrid assay in HEK 293 T cells, and the results further confirmed the endogenous interaction of IRF2BP2 and ZBTB16 (Fig. 3E). Collectively, these findings demonstrate that IRF2BP2 can interact with ZBTB16 in vivo.

To investigate the molecular basis of this interaction, we first conducted an ITC analysis to determine whether the interaction occurs through the RxSVI motif of ZBTB16 and the RING domain of IRF2BP2. The results indicated that the RING domain of IRF2BP2 exhibited a reasonable binding affinity (K_D value of 12.60 μM) with the RxSVI motif-containing peptide of ZBTB16 (residues 285-292, RSSVITSA) (Fig. 3F, Supplementary Table S1). Similar to its interactions with IRF2 and VGLL4, the RING domain of IRF2BP2 displayed a slightly stronger binding affinity (K_D value of 7.21 μM) with a longer ZBTB16 peptide (residues 279-298, ZBTB16-L) (Fig. S3E, Supplementary Table S1). Sequence alignment of ZBTB16 across different species revealed that mouse ZBTB16 contains a glycine instead of a serine at position 2 of the RxSVI motif (Supplementary Fig. S2C). Our ITC analysis showed that the RING domain of IRF2BP2 exhibited a slightly weaker binding affinity with the mouse ZBTB16 peptide (residues 285-292, RGSVITSA) (K_D of 16.80 μM) compared to the human ZBTB16 peptide (Fig. 3F, Supplementary Table S1).

Next, we determined the crystal structure of the RING domain in complex with the ZBTB16 peptide (RING-ZBTB16) at 2.50 Å resolution (Table 1). In this structure, the asymmetric unit contains one complex molecule, with most residues of the RING domain and the first seven residues of the peptide well-defined in the electron density map (Table 1, Fig. 3G, Supplementary Fig. S4C). Similar to the RING-IRF2 and RING-VGLL4 complexes, there are three Zn^{2+} ions bound to each RING molecule: two Zn^{2+} ions (Zn1 and Zn2) occupy the Zn-binding sites of the C3HC4 RING motif, while the third Zn^{2+} ion (Zn3) binds to a distinct site of the inter-molecular interface which is coordinated by Cys505 and His510 (the L1 loop) of RING, His517 (the L1 loop) of a symmetry-related RING, and a water molecule (Supplementary Fig. S5C). Since the L1 loop is involved in the Zn3-mediated inter-molecular contacts, its conformation is slightly different from that in the other two complexes (Supplementary Fig. S6A).

Similar to the RING-IRF2 complex, the ZBTB16 peptide forms a short loop (residues 1'−3') followed by a short β-strand (residues 4'−7') which forms an antiparallel β-sheet with β1 of the RING domain. The ZBTB16 peptide binds to the RING domain in the same manner as observed in the other two complexes. The first five residues of the peptide adopt conformations similar to those of the IRF2 and VGLL4 peptides and make nearly identical interactions with the RING domain (Fig. 3H). However, unlike in the other two complexes, the mainchain carbonyl of Thr6' of the peptide does not form a hydrogen bond with the sidechain of His517 of the RING domain, as the sidechain of His517 is involved in the inter-molecular contacts and is coordinated with Zn3 from a symmetry-related RING (see discussion above) (Fig. 3H, Supplementary Fig. S5C).

Mutagenesis and ITC analysis demonstrated that mutations of either Arg1' or Val4' and Ile5' to Ala in the ZBTB16 peptide abolished its interaction with the RING domain in vitro (Supplementary Table S1, Fig. S3F). Additionally, mammalian two-hybrid and co-IP assays revealed that mutations of these residues in full-length ZBTB16 disrupted its interaction with IRF2BP2 in vivo (Fig. 3E and I). These results
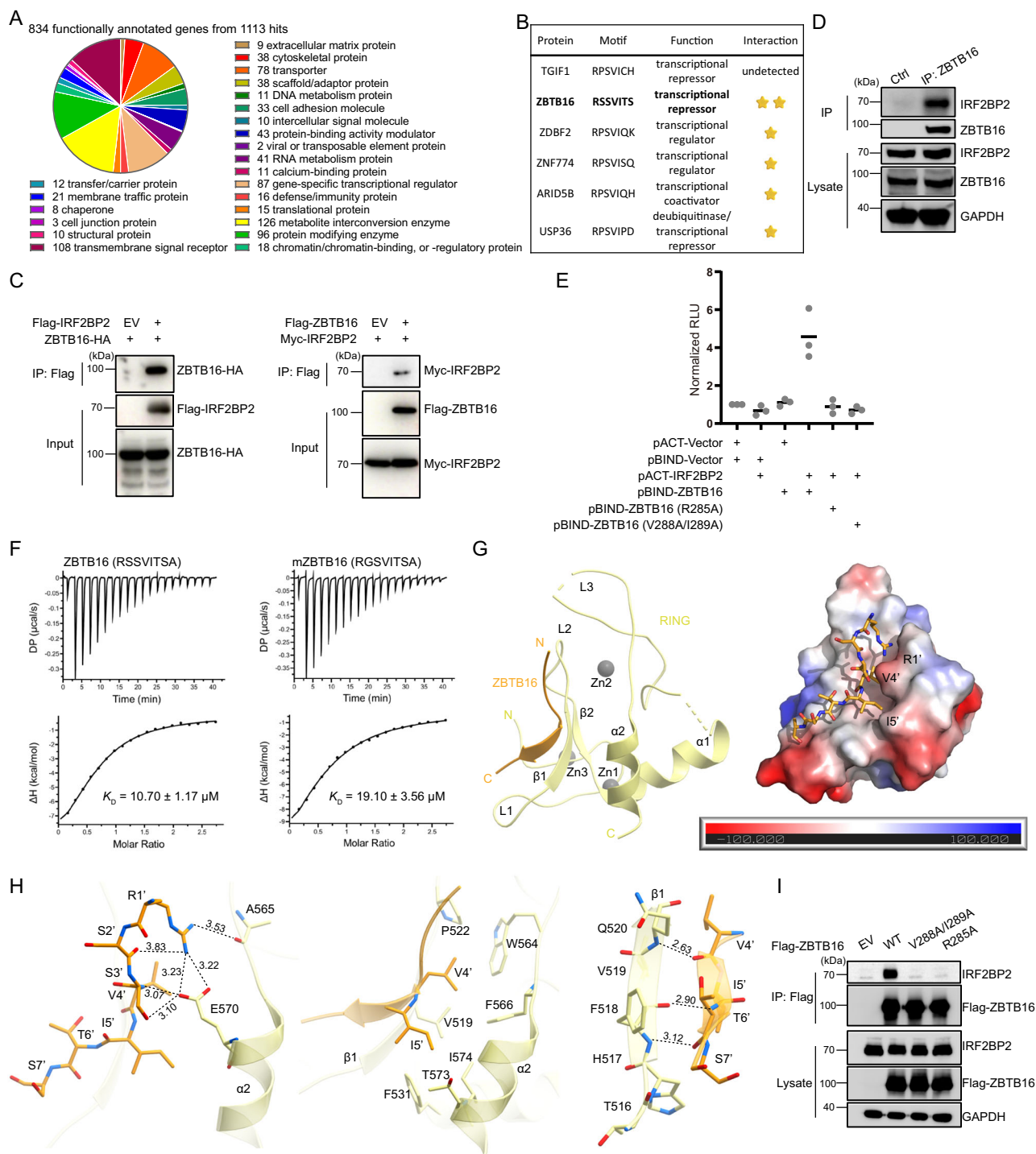
Fig. 3 | ZBTB16 is an interacting protein of IRF2BP2. A A motif search identified numerous potential interacting proteins containing the RxSVI motif in the human proteome. **B** Summary of the co-IP assay results to examine the interactions between IRF2BP2 and six representative potential interacting proteins in the gene-specific transcriptional regulator group. One star, interaction in one direction; Two stars, interactions in two directions. **C** Co-IP assay of the interaction between IRF2BP2 and ZBTB16 in HEK 293 T cells. Left: Flag-tagged IRF2BP2 to pull-down HA-tagged ZBTB16. Right: Flag-tagged ZBTB16 to pull-down Myc-tagged IRF2BP2. **D** Co-IP assay of the endogenous interaction between IRF2BP2 and ZBTB16 in HEL cells. **E** Mammalian two-hybrid assay of the interaction between IRF2BP2 and WT or mutant ZBTB16. The experiment was repeated three times independently. **F** Binding affinities of the RING domain of IRF2BP2 with the RxSVI motif-containing peptides of human ZBTB16 (left) and mouse ZBTB16 (right, mZBTB16) measured using the ITC method. Shown here is one representative result from three independent experiments for each peptide. **G** Overall structure of the RING-ZBTB16 complex in ribbon representation (left, orange/gold) and in electrostatic surface representation of the RING domain with the ZBTB16 peptide depicted as a stick model (right). The three Zn^{2+} ions are shown as gray spheres. The unit of electrostatic potential is kT/e. **H** Interactions between the ZBTB16 peptide and the RING domain. Left: Hydrophilic interactions between Arg1' and Ser3' of the ZBTB16 peptide and the RING domain. Middle: Hydrophobic interactions between Val4' and Ile5' of the ZBTB16 peptide and the RING domain. Right: Hydrogen-bonding interactions between the mainchains of Val4', Ile5', and Lys6' of the ZBTB16 peptide and the mainchains of the residues of the RING domain. **I** Co-IP assay of the interaction between IRF2BP2 and WT or mutant ZBTB16 in HEK 293 T cells. The blots presented here are the representative results of three independent replicates. Source data are provided as a Source Data file.

confirm the functional role of the RxSVI motif of ZBTB16 in its interaction with IRF2BP2 in vivo. Together, the structural and biochemical data demonstrate that ZBTB16 is an interacting protein of IRF2BP2, with the RING domain of IRF2BP2 binding to the conserved RxSVI motif of ZBTB16 in a manner similar to that of IRF2 and VGLL4.

The IRF2BP2-ZBTB16 interaction plays a crucial regulatory role in megakaryocytic differentiation

ZBTB16 is a potent transcription factor that plays a key regulatory role in hematopoietic cell proliferation and differentiation. Previous studies have shown that the pluripotent haemopoietic HEL cells can differentiate into megakaryocytic phenotype cells after treatment with phorbol 12,13-dibutyrate (PDBu) and ZBTB16 plays an important regulatory role in this process[19,25]. In PDBu-treated HEL cells, several markers associated with the megakaryocyte/platelet phenotype, including CD41 and CD61, were elevated.

To investigate whether the IRF2BP2-ZBTB16 interaction affects ZBTB16's regulatory role in megakaryocytic differentiation, we firstly knocked out either ZBTB16 or IRF2BP2 in HEL cells (Fig. 4A, C) and examined the expression levels of the megakaryocytic cell markers CD41 and CD61 as indicators of megakaryocytic cell populations before and after PDBu treatment. As expected, compared to wild-type (WT) cells, the populations of CD41+ and CD61+ cells were moderately decreased in ZBTB16-KO, PDBu-untreated (PDBu-) cells and substantially decreased in ZBTB16-KO, PDBu-treated (PDBu+) cells (Fig. 4B). Similarly, the populations of CD41+ and CD61+ cells were also substantially reduced to varying extents in IRF2BP2-KO cells both without and with PDBu treatment (Fig. 4D). These results indicate that, like ZBTB16, IRF2BP2 is also involved in the regulation of megakaryocytic differentiation.

A previous study demonstrated that the interaction between IRF2BP2 and VGLL4 enhances VGLL4 stability by inhibiting its ubiquitination, thereby protecting it from degradation[8]. To investigate whether IRF2BP2's regulatory role in megakaryocytic differentiation is mediated through modulation of ZBTB16 stability, we first analyzed the effects of IRF2BP2 overexpression on the protein expression and ubiquitination levels of ZBTB16. In both PDBu-untreated and PDBu-treated HEL cells, compared to the control (EV), IRF2BP2 overexpression (IRF2BP2-OE) did not significantly affect the protein expression level (Fig. 4E) or the ubiquitination level of ZBTB16 (Supplementary Fig. S8).

Next, we examined the impact of IRF2BP2 overexpression on the populations of CD41+ and CD61+ cells. Prior to PDBu treatment, the population of CD41+ cells was slightly decreased, while the population of CD61+ cells was slightly increased in the IRF2BP2-OE cells compared to the control (Fig. 4F), suggesting that IRF2BP2 may inhibit the CD41 expression through an unknown mechanism. However, after PDBu treatment, the population of CD41+ cells exhibited no significant difference, and the population of CD61+ cells was marginally increased in the IRF2BP2-OE cells. This suggests that IRF2BP2 overexpression had no notable effect on megakaryocytic differentiation, or that endogenous IRF2BP2 may have reached saturation level.

Subsequently, we examined whether disrupting the IRF2BP2-ZBTB16 interaction affects ZBTB16's regulatory role in megakaryocytic differentiation. Given that the V288A/I289A mutation shares more similar chemical properties and sizes with the wild-type residues compared to the R285A mutation, it is likely to primarily impact the interaction between ZBTB16 and IRF2BP2 while having a less effect on ZBTB16's structure and function. Therefore, we used the ZBTB16 mutant containing the V288A/I289A mutation for the following functional studies.

Western blot results indicated that overexpression of either WT or mutant ZBTB16 (ZBTB16-OE) did not significantly affect the protein expression level of IRF2BP2 in both PDBu- and PDBu+ cells compared to the control (EV) (Fig. 4G). When comparing cells overexpressing WT

and mutant ZBTB16 under PDBu- conditions, overexpression of WT ZBTB16 significantly increased the proportions of CD41+ and CD61+ cells compared to the control (EV), consistent with previous findings[19] (Fig. 4H). In contrast, overexpression of the ZBTB16 mutant substantially reduced the populations of CD41+ and CD61+ cells. This reduction may be due to the impaired regulatory role of the ZBTB16 mutant on the expression of CD41 and CD61 and/or its competition with endogenous ZBTB16 (Fig. 4H). Under PDBu+ conditions, overexpression of WT ZBTB16 significantly increased the populations of CD41+ and CD61+ cells, whereas overexpression of the ZBTB16 mutant had no noticeable effect on the populations of CD41+ and CD61+ cells (Fig. 4H). These results suggest that the ZBTB16 mutant has partially lost the regulatory function in megakaryocytic differentiation.

Moreover, as enlarged cell size is another characteristic of megakaryocytic differentiation, we conducted a statistical analysis of the cell size distribution in HEL cells after PDBu treatment. The results showed that overexpression of WT ZBTB16 significantly increased the proportion of large cells compared to the control (EV), while overexpression of the ZBTB16 mutant had a markedly weaker effect (Fig. 4I).

Furthermore, we analyzed the expression levels of several other megakaryocytic markers using RT-qPCR. Under both PDBu- and PDBu+ conditions, cells overexpressing WT ZBTB16 exhibited increased levels of megakaryocytic markers vwF, CD9, CD42a, CD49b, CD49f, and CD62p to varying extents compared to the control (Supplementary Fig. S9). In contrast, cells overexpressing the ZBTB16 mutant showed substantial decreases in these megakaryocytic markers to different degrees, which may be due to the impaired regulatory function of the ZBTB16 mutant on the expression of these marker genes and/or its competition with endogenous ZBTB16. Overall, our cell biological data suggest that the ZBTB16 mutant has an impaired regulatory function in promoting megakaryocytic differentiation, indicating that the regulatory role of ZBTB16 in megakaryocytic differentiation is dependent on its interaction with IRF2BP2.

Discussion

IRF2BP2 is a transcriptional coregulator that is widely expressed in various normal and tumor cells and tissues. It plays a crucial role in regulating multiple cellular processes, including transcription regulation, immunity, metabolism, and signaling pathways [see review by Pastor[3]]. IRF2BP2 performs diverse regulatory functions by interacting with different protein partners through its RING domain. So far, several interacting proteins of IRF2BP2 have been identified, including IRF2, NFAT1, ETO2 (or CBFA2T3), and VGLL4. However, the molecular mechanisms underlying how IRF2BP2 interacts with different protein partners remain unclear.

In this study, we employed a bioinformatics tool to search for conserved sequence motifs in the well-characterized IRF2BP2 interacting proteins IRF2 and VGLL4, and identified a conserved RxSVI motif. A search of this motif in the human proteome identified the transcription factor ZBTB16 as an interacting protein of IRF2BP2. Our structural and biochemical data indicated that the RING domain of IRF2BP2 binds to the conserved motif of IRF2, VGLL4, and ZBTB16 in a similar manner. In addition, we found that the conserved RxSVI motif could be extended to the RxxVI motif, and the RxSVI motif-containing peptides from the interacting proteins tend to form a short loop along with a short β-strand, facilitating effective recognition and tightly binding by the RING domain. These findings elucidate a molecular mechanism by which IRF2BP2 interacts with various protein partners, enabling it to exert different regulatory functions in diverse cellular processes.

In addition to IRF2 and VGLL4, the RING domain of IRF2BP2 has been shown to interact with the C-terminal TAD-C region (residues 781-867) of NFAT1 by yeast two-hybrid analysis[6], and the N-terminal US2 region (residues 74-101) of ETO2 by co-IP assay[7]. Sequence analysis reveals that the TAD-C region of NFAT1 contains an RSVLV sequence (residues 782-786), which shares some similarity to the RxSVI motif
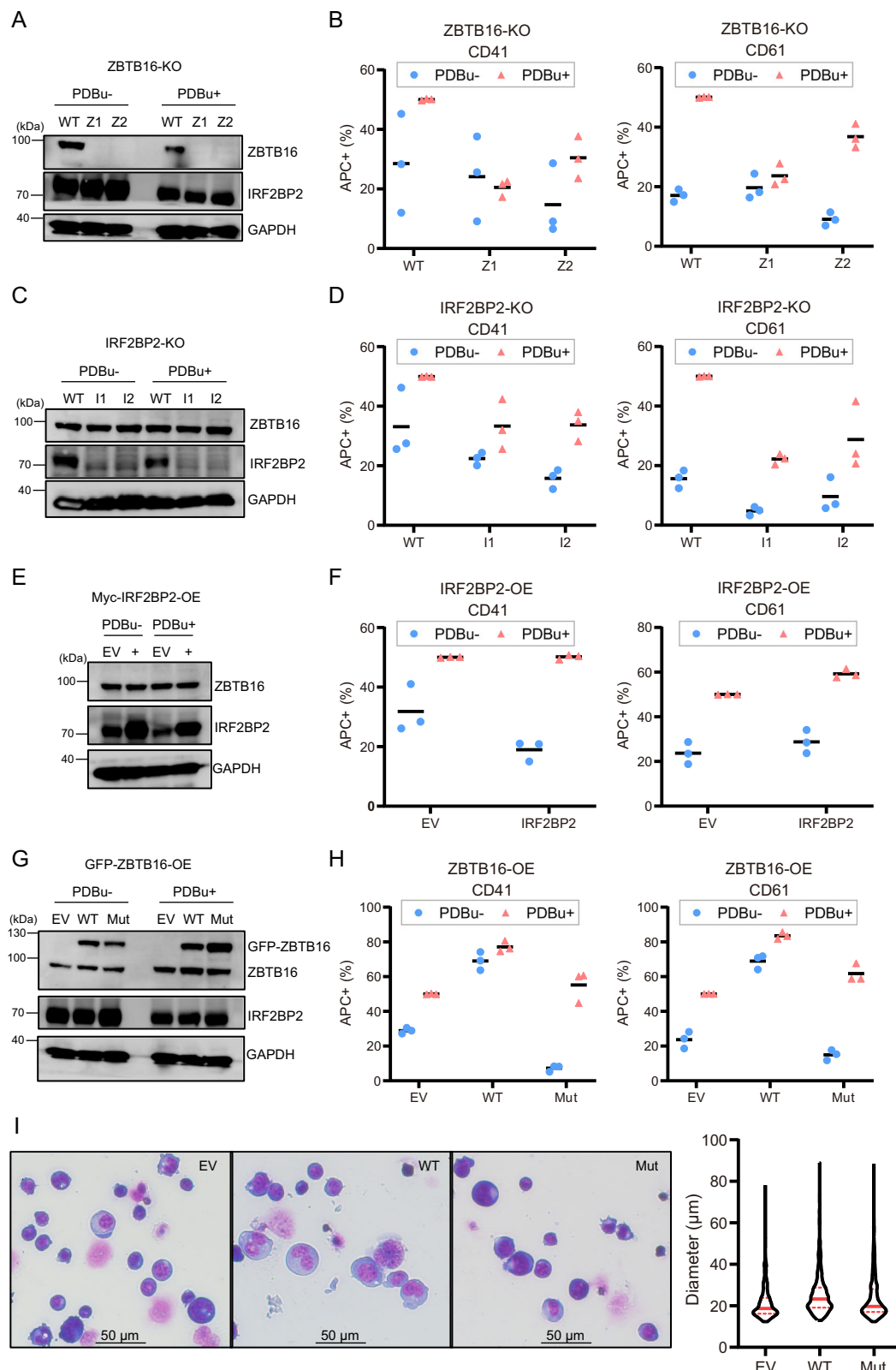
and is located in a flexible loop in the Alphafold2 predicted structure model of NFAT1[26]. These results suggest that the RING domain of IRF2BP2 may interact with the C-terminal TAD-C region of NFAT1 in a manner similar to that of IRF2, VGLL4, and ZBTB16. Indeed, our ITC analysis showed that the NFAT1 peptide (residues 782-789, RSVLVHAG) exhibited a weak binding affinity with the IRF2BP2 RING domain ($K_D = 514\,\mu M$) (Supplementary Fig. S10). Conversely, although ETO2 contains an RPFVI sequence (residues 229-233) which shares some similarity to the RxSVI motif, this sequence is located in a long α-helix in the Alphafold2 predicted structure model of ETO2, making it unlikely to form a loop or β-strand-like structure in order to be recognized and bound by the RING domain. These results suggest that the RING domain of IRF2BP2 may interact with ETO2 in a manner distinct from its interactions with IRF2, VGLL4, and ZBTB16.

Fig. 4 | IRF2BP2 plays a regulatory role in megakaryocytic differentiation through interaction with ZBTB16. A Western blot analysis of the expression levels of ZBTB16 in WT and ZBTB16-KO HEL cells (Z1 and Z2) before and after PDBu treatment (PDBu- and PDBu). **B** Flow cytometry analysis of the CD41+ and CD61+ cell populations as indicators of the megakaryocytic cell population in WT and ZBTB16-KO HEL cells (Z1 and Z2) before and after PDBu treatment. **C** Western blot analysis of the expression levels of IRF2BP2 in WT and IRF2BP2-KO HEL cells (I1 and I2) before and after PDBu treatment. **D** Flow cytometry analysis of the CD41+ and CD61+ cell populations in WT and IRF2BP2-KO HEL cells (I1 and I2) before and after PDBu treatment. **E** Effect of overexpression of IRF2BP2 (IRF2BP2-OE) on ZBTB16 expression levels before and after PDBu treatment. **F** Flow cytometry analysis of the CD41+ and CD61+ cell populations in IRF2BP2-overexpressing HEL cells before and

after PDBu treatment. **G** Effect of overexpression of WT or mutant ZBTB16 (ZBTB16-OE) on IRF2BP2 expression levels before and after PDBu treatment. **H** Flow cytometry analysis of the CD41+ and CD61+ cell populations in the WT or mutant ZBTB16-overexpressing HEL cells before and after PDBu treatment. **I** Cell size distribution of the HEL cells after PDBu treatment. Left panel: representative images of the cells with the scale bar indicated. Right panel: cell size distribution estimated from at least 328 cells. The western blots presented here are the representative results of three independent replicates. The flow cytometry data are the results of three independent assays, and the line indicates the average. The solid and dash lines in the violin graph of cell size distribution indicate the median and the quartile, respectively. EV: empty vector. WT: WT ZBTB16. Mut: V288A/I289A mutant ZBTB16. Source data are provided as a Source Data file.

In the BioGrid database (www.thebiogrid.org/131797/summary/homo-sapiens/irf2bp2.html), 113 proteins are annotated as potential IRF2BP2 interacting proteins. Among these proteins, only 7 proteins contain the RxxVI motif: CBFA2T3 (or ETO2), CEBPA, FOSL2, IRF2, SCRIB, UPRT, and VGLL4. Of these, IRF2, VGLL4, and ETO2 have been discussed above. The conserved motif in the remaining four proteins is located in flexible loops in the Alphafold2 predicted structure models[26], suggesting that these proteins may interact with the RING domain of IRF2BP2 in a manner similar to that of IRF2, VGLL4, and ZBTB16. Indeed, it was recently reported that IRF2BP2 interacts with FOSL2 and this interaction is crucial for maintaining the survival of tumor cells by regulating the chromatin accessibility of the neuroblastoma susceptibility gene ALK[27]. These findings further confirm that IRF2BP2 can bind to the conserved RxxVI motif of other interacting proteins via its RING domain.

Our cell biological data indicated that IRF2BP2 can modulate the regulatory function of ZBTB16 in megakaryocytic differentiation, and disruption of the IRF2BP2-ZBTB16 interaction impairs the regulatory role of ZBTB16 in promoting this differentiation. However, the underlying molecular mechanism remains unclear. ZBTB16 contains a BTB/POZ domain at the N-terminus[15,18], which is critical for ZBTB16 dimerization and transcriptional regulation, particularly mediating its interactions with its partners such as NCoR, SMRT, and Sin3A[16,18,28]. It also contains a nine Krüppel-like C2H2 zinc finger domain at the C-terminus[15,18], which confers a sequence-specific DNA binding to target gene regulatory elements, enabling ZBTB16 to exert transcriptional control. Additionally, the middle region includes a poorly characterized RD2 domain that may modulate its transcriptional activity[15,18]. The RxxVI motif of ZBTB16 is located within this RD2 domain. Previous studies have demonstrated that ZBTB16 interacts with the miR-146a promoter to suppress its transcription, which in turn activates CXCR4 translation and regulates megakaryocytic cell development[19]. ZBTB16 may also form a complex with GATA1 to play a stimulating role in megakaryocyte development[20]. The interaction of IRF2BP2 with the RxxVI motif of ZBTB16 may influence the interaction of ZBTB16 with the miR-146a promoter and/or GATA1, thereby suppressing the transcription of downstream target genes and subsequently impairing its regulatory role on megakaryocytic differentiation. Additionally, a previous study has shown that IRF2BP2 interacts with VGLL4 to stabilize VGLL4 by inhibiting its ubiquitination and further YAP activity[8]. However, our biological data indicated that IRF2BP2 overexpression does not significantly affect the protein expression and ubiquitination levels of ZBTB16, suggesting that the impact of the IRF2BP2-ZBTB16 interaction on megakaryocytic differentiation is unlikely mediated through the ubiquitination and stability of ZBTB16.

So far, it is unclear how the interaction of IRF2BP2 with its binding partners is regulated in cells. It is plausible that IRF2BP2 and its binding partners might form part of large protein complexes to exert their functions, and their interactions could be regulated by other proteins within these complexes. For instance, IRF2BP2 has been shown to recruit the NCoR/SMRT corepressor complex, thereby suppressing the expression of the archetypical erythroid genes[7], and ZBTB16 has been

shown to also interact with NCoR/SMRT[29]. This suggests that both IRF2BP2 and ZBTB16 may collaborate with NCoR/SMRT to exert their functions.

Our structural analysis revealed that the RING domain of IRF2BP2 exhibits a structure similar to the RING domains of E3 ubiquitin-protein ligases, particularly the RING1 domain of PARKIN, although there are some notable structural differences. To explore whether the RING domain of IRF2BP2 could function as an E3 ligase, we superimposed the IRF2BP2 RING-IRF2 complex onto the fruit fly PARKIN RING1-UBE2L3 complex (PDB: 6DJW) based on the RING domains, and analyzed whether IRF2BP2 could interact with UBE2L3 in a manner analogous to PARKIN. While the IRF2BP2 RING domain assumes a structural conformation similar to the PARKIN RING1 domain at the binding interface with UBE2L3, the residues of the PARKIN RING1 domain interacting with UBE2L3 are not conserved in the IRF2BP2 RING domain (Supplementary Fig. S11). This suggests that the IRF2BP2 RING domain may not interact with UBE2L3. In addition, we performed co-IP and MS analyses to determine whether IRF2BP2 could bind any E2 enzymes. These analyses identified 4 potential E2s in trace amounts, including UBE2NL, UBE2L3, UBE2M, and UBE2V1 (Supplementary Table S2). However, the co-IP assays indicated that none of these E2s interact with IRF2BP2 (Supplementary Fig. S12). Given that UBE2L6 has also been reported to interact with PARKIN, we performed a co-IP assay for UBE2L6 and found no interaction between UBE2L6 and IRF2BP2. These results suggest that the RING domain of IRF2BP2 is unlikely to function as an E3 ligase. Nevertheless, we cannot rule out the possibility that IRF2BP2 may interact with other unexamined or unidentified E2s and potentially function as an E3 ligase.

To date, three mutations in the RING domain of IRF2BP2 have been identified in patients with immune diseases. The Q536delinsL* and Q540* mutations result in partial truncations of the RING domain, which would disrupt its interactions with protein partners, leading to loss of the regulatory functions of IRF2BP2 in gene expression and cellular processes[11,12]. Ser551 is located in the L3 loop near the Zn2-binding site. The S551N mutation appears to alter the conformations of nearby residues, including Cys549 and Cys555 which are involved in the Zn2 binding, and thus would affect the structure and stability of the zinc finger motif and hence the RING domain-mediated interactions with protein partners[13]. Additionally, the ClinVar database (www.ncbi.nlm.nih.gov/clinvar/) lists several other mutations in the RING domain of IRF2BP2 which are identified in patients associated with immunodeficiency and inborn genetic diseases, such as A560V and V542L. These mutations may affect the stability of the RING domain of IRF2BP2 and thus impair its interaction with protein partners, subsequently disrupting the regulatory role of IRF2BP2 in immune response.

Methods
Cloning, expression, and purification of the RING domain of IRF2BP2
The gene fragment encoding the RING domain of IRF2BP2 (residues 498-578) was cloned into the pET22b vector (Novagen) and fused with

an N-terminal His$_6$-TEV tag. The protein was expressed in *E. coli* BL21-CodonPlus (DE3)-RIPL strain (WEIDI) at 16 °C in Luria-Bertani medium induced with 0.1 mM IPTG. The cells were harvested, resuspended, and then lysed using a sonicator in buffer A (20 mM HEPES, pH 8.0, 150 mM NaCl, and 0.1 mM ZnCl$_2$) supplemented with 1 mM PMSF. After centrifugation, Ni-NTA affinity chromatography was employed to purify the target protein, which was then eluted in buffer A supplemented with 300 mM imidazole. TEV protease was added to the elution, and the mixture was dialyzed into buffer A supplemented with 5 mM β-mercaptoethanol overnight. Subsequently, Ni-NTA affinity chromatography was used again to remove the cleaved tag. The flow-through was further purified by size exclusion chromatography (SEC) using a Superdex 200 10/300 column (Cytiva) in buffer A supplemented with 1 mM DTT. The purified protein exhibited high purity and homogeneity and existed as a monomer in solution as shown by gel filtration analysis. It was concentrated to 20 mg/mL and stored in the SEC buffer for structural and biochemical studies.

Isothermal titration calorimetry (ITC) analysis

The binding affinities of the RING domain with the motif-containing peptides from various protein partners (both wild-type and mutants) were measured by isothermal titration calorimetry (ITC) using a Micro Cal PEAQ-ITC system (Malvern). The RING domain protein (30 μM) was titrated with each peptide (400 μM). All protein samples were prepared in ITC buffer (20 mM HEPES, pH 8.0, 150 mM NaCl, 0.1 mM ZnCl$_2$, and 1 mM DTT). A total of 20 injections with a spacing of 120 s and a reference power of 5 μcal s^{-1} were conducted at 25 °C. The ITC curves were analyzed using Micro Cal PEAQ-ITC analysis software with a one-site fitting model. The small GTPase binding domain of WASL (which has a comparable molecular mass to the RING domain) was used as a negative control to confirm the absence of nonspecific interactions.

Crystallization, diffraction data collection, and structure determination

Crystallization was performed using the hanging-drop vapor diffusion method with a Mosquito instrument (TTP Labtech) at 20 °C. Prior to crystallization, the RING domain protein was incubated with the peptides from different protein partners at a molar ratio of 1:2 for 1 h at 4 °C. The protein-peptide complex was adjusted to about 15 mg/ml and then mixed with the Hampton Research crystallization screen kits as the reservoir solution with equal volume (100 nl). Crystals of the RING domain in complex with the IRF2 peptide (RING-IRF2) grew in drops containing 2.5 M ammonium sulfate and 0.1 M sodium acetate (pH 4.6), and belong to the space group of $P2_1$. Crystals of the RING domain in complex with the VGLL4 peptide (RING-VGLL4) grew in drops containing 20% 2-proponal, 0.1 M MES (pH 6.0), and 20% (w/v) PEG MME 2000, and belong to the space group of $P2_12_12_1$. Crystals of the RING domain in complex with the ZBTB16 peptide (RING-ZBTB16) grew in drops containing 0.1 M potassium thiocyanate and 30% (w/v) PEG MME 2000, and belong to the space group of $P6_4$. Prior to diffraction data collection, the crystals were cryoprotected using the reservoir solution supplemented with 20% ethylene glycol and then flash-cooled in liquid N$_2$. Diffraction data were collected at −173 °C at beamlines BL02U1 and BL18U1 of the Shanghai Synchrotron Radiation Facility. The diffraction data were processed using HKL2000[30] or XDS[31]. Statistics for the diffraction data are summarized in Table 1.

The crystal structure of the RING-VGLL4 complex was solved by the molecular replacement (MR) method implemented in Phenix[32] using the Alphafold2[26] predicted structure of the RING domain as the search model. The crystal structures of the RING-IRF2 and RING-ZBTB16 complexes were solved using the RING-VGLL4 complex as the search model. Structure refinement was carried out using Phenix[32] and manual model building was carried out using COOT[33]. The stereochemistry quality of the structure models was validated by the Validation and Deposition Services at the Research Collaboratory for

Structural Bioinformatics (RCSB) Protein Data Bank. Molecular graphics figures were generated using ICM-Browser (Molsoft) and Pymol (http://www.pymol.org)[34]. Statistics for the structure refinement and the quality of the structure models are also summarized in Table 1.

Cell cultures

HEK 293 T cells were maintained in DMEM supplemented with 10% FBS. HEL cells were maintained in RPMI-1640 supplemented with 10% FBS. For megakaryocytic differentiation, HEL cells were incubated for three days in the presence of 200 nM PDBu (MCE) and then analyzed. All cell lines were obtained from the Type Culture Collection Center of Chinese Academy of Sciences and validated through short tandem repeat (STR) testing.

Mammalian two-hybrid assay

The mammalian two-hybrid assay was conducted according to the technical manual of the Mammalian Two-Hybrid System from Promega. IRF2BP2 was cloned into the pACT vector (Promega) with an N-terminal VP16 fusion. Wild-type or mutant ZBTB16 was cloned into the pBIND vector (Promega) with an N-terminal Gal4 fusion. Then, the pACT-IRF2BP2, pBIND-ZBTB16 and pG5luc plasmids were co-transfected with a 1:1:1 ratio into HEK 293 T cells. After three days, dual-luciferase activity was measured using the Dual Luciferase Reporter Assay Kit (Vazyme) on a GloMax (Promega).

Plasmid constructions

The gene fragments encoding full-length IRF2BP2, IRF2, VGLL4, and ZBTB16 were cloned into the pcDNA3.1 (Invitrogen) vector with either an N-terminal 3×Flag tag or an N-terminal 3×Myc tag. The gene fragments encoding full-length ARID5B, ZNF774, USP36, ZDBF2, TGIF1, UBE2L3, UBE2L6, UBE2M, UBE2NL, and UBE2V1 were cloned into the pcDNA3.1 vector with either an N-terminal 3×Flag tag or a C-terminal HA tag.

For the construction of stable cell lines, the gene fragment encoding full-length ZBTB16 was cloned into the pCDH (System Biosciences) vector with an N-terminal GFP tag. To knock down IRF2BP2, the following oligonucleotide sequences were cloned into pLKO.1 (Sigma) at the AgeI/EcoRI site.

sh1: 5′-CCGGCCCAATGATTACCCTTAATTACTCGAGTAATTAAGG GTAATCATTGGGTTTTTGAATT-3′

sh2: 5′- CCGGGGGCCCTTCGAGAGCAAGTTTACTCGAGTAAACT TGCTCTCGAAGGGCCTTTTTGAATT-3′

Western blot assay

Protein expression was evaluated by western blotting using the following antibodies: anti-PLZF (ZBTB16) monoclonal antibody (sc-28319, Santa Cruz), anti-IRF2BP2 polyclonal antibody (18847-1-AP, Proteintech), anti-Flag monoclonal antibody (F1804, Sigma), anti-Myc monoclonal antibody (05-724, Millipore), anti-HA monoclonal antibody (3724, CST), and anti-GAPDH recombinant antibody (81640-5-RR, Proteintech). When necessary, the same membrane was stripped using Restore Western blot stripping buffer (21059, Thermo Scientific) between different antibody blots.

Immunoprecipitation (IP) and co-immunoprecipitation (co-IP) assays

For protein interaction detection, co-IP assays were performed in HEK 293 T cells. The cells were co-transfected with Flag-tagged IRF2BP2 and Myc/HA-tagged target proteins or Flag-tagged target proteins and Myc-tagged IRF2BP2 using Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions. After 48 hrs, cells were harvested and lysed in Western lysis buffer (Beyotime) supplemented with a protease inhibitor cocktail (APExBIO). After centrifugation, protein extracts were incubated with magnific beads covalently coated with anti-Flag antibody (MCE) overnight at 4 °C. The beads were then

isolated, washed with lysis buffer three times, and incubated in elution buffer (0.2 M glycine, pH 2.5) prior to western blot analysis.

To examine the endogenous interaction of IRF2BP2 and ZBTB16 without overexpressing the proteins, co-IP assays were conducted in HEL cells. Protein A/G magnetic beads (MCE) were preincubated with anti-PLZF (ZBTB16) monoclonal antibody (sc-28319, Santa Cruz) for 1 hr at 4 °C. The harvested HEL cells were lysed in Western lysis buffer (Beyotime) with a protease inhibitor cocktail (APExBIO). The protein extracts were then incubated with the antibody-bound protein A/G magnetic beads overnight at 4 °C. The beads were isolated, washed in lysis buffer three times, and incubated in elution buffer (0.2 M glycine, pH 2.5) prior to Western blot analysis.

For ubiquitination detection, HEK 293 T cells were co-transfected with Flag-tagged ZBTB16, HA-tagged Ub, and Myc-tagged IRF2BP2 or sh-IRF2BP2. After 48 hrs, cells were harvested and lysed in Western lysis buffer (Beyotime) supplemented with a protease inhibitor cocktail (APExBIO). The Flag-tagged ZBTB16 was then immunoprecipitated from the cell extracts using magnific beads covalently coated with anti-Flag antibody (MCE) for 4 hrs at 4 °C. The beads were isolated, washed in lysis buffer three times, and incubated in elution buffer (0.2 M glycine, pH 2.5) before Western blot analysis.

Flow cytometry analysis
500,000 cells were harvested and washed three times with PBS containing 2% FBS. Then, the cells were incubated for 30 min at 4 °C with either an APC-conjugated mouse anti-human CD41 antibody (20 μL per trial, 559777, BD Pharmingen), an APC-conjugated mouse anti-human CD61 antibody (5 μL per trial, 564174, BD Pharmingen), or an APC-conjugated mouse IgG1, κ isotype control antibody (555751, BD Pharmingen) of the same isotype (20 μL per trial). After washing with PBS containing 2% FBS, the cells were analyzed using a BD LSRFortessa flow cytometer (BD Biosciences), collecting a minimum of 10,000 events per sample. FACS sorting was performed with a Sony MA900 instrument (Sony) using the GFP tag.

Morphological analysis
Cells were collected on the third day after PDBu-induced megakaryocytic differentiation, transferred to glass slides using the smear method, and stained with May-Grünwald-Giemsa (Sigma). The slides were then observed and photographed under an Olympus BX53, and cell diameter was analyzed using Fiji[35].

RT-qPCR
For gene expression analysis, RNA was extracted using VeZol (Vazyme), and cDNA was trans-reversed by using HiScript III RT SuperMix (Vazyme). The cDNA served as the template for RT-qPCR, which was performed on a LightCycler 96 (Roche) using AceQ Universal SYBR qPCR Master Mix (Vazyme). Gene expression values were normalized to Rplp0 mRNA levels[36]. The fold change in gene expression was calculated using the ΔΔCt method. Primer sequences are listed in Supplementary Table S3.

Generation of knockout cell lines
Generation of knockout cell lines using the CRISPR/Cas9 technology was carried out with the pLenti-Cas9 vector and pGS vector (Genscript). Small guide RNA sequences for the targeted inactivation of Irf2bp2 and Zbtb16 in HEL cells were obtained from the Genscript library, which have been proven to be effective. Combinations of two different guide RNAs were used for each gene (Irf2bp2: 5′-CCGACCGCGTCGAGTTCGTCATC-3′ and 5′-TCTCGATGACGAACTC-GACG-3′; Zbtb16: 5′-AGCGGTTCCTGGATAGTTTG-3′ and 5′-TTCTCAGCCGCAAACTATCC-3′). HEL cells were transfected using NEPA21 (NEPA Gene), and single clones were isolated by FACS into 96 well plates. Targeted genomic deletions were verified by sequencing and Western blotting.

Statistics and reproducibility
Details on the statistical analyses and tests performed are provided in the figure legends. Standard error calculations were conducted using Graphpad Prism.

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
The crystal structures of the RING domain of IRF2BP2 in complexes with the IRF2, VGLL4, and ZBTB16 peptides generated in this study have been deposited in the Protein Data Bank with accession codes 8YTG, 8YTF, and 8YTH, respectively. Source data are provided with this paper.

References
1. Childs, K. S. & Goodbourn, S. Identification of novel co-repressor molecules for Interferon Regulatory Factor-2. *Nucleic Acids Res.* **31**, 3016–3026 (2003).
2. Fagerberg, L. et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell Proteom.* **13**, 397–406 (2014).
3. Pastor, T. P., Peixoto, B. C. & Viola, J. P. B. The transcriptional co-factor IRF2BP2: a new player in tumor development and micro-environment. *Front Cell Dev. Biol.* **9**, 655307 (2021).
4. Teng, A. C. et al. Identification of a phosphorylation-dependent nuclear localization motif in interferon regulatory factor 2 binding protein 2. *PLoS One* **6**, e24100 (2011).
5. Yeung, K. T. et al. A novel transcription complex that selectively modulates apoptosis of breast cancer cells through regulation of FASTKD2. *Mol. Cell Biol.* **31**, 2287–2298 (2011).
6. Carneiro, F. R., Ramalho-Oliveira, R., Mognol, G. P. & Viola, J. P. Interferon regulatory factor 2 binding protein 2 is a new NFAT1 partner and represses its transcriptional activity. *Mol. Cell Biol.* **31**, 2889–2901 (2011).
7. Stadhouders, R. et al. Control of developmentally primed erythroid genes by combinatorial co-repressor actions. *Nat. Commun.* **6**, 8893 (2015).
8. Feng, X. et al. The tumor suppressor interferon regulatory factor 2 binding protein 2 regulates Hippo pathway in liver cancer by a feedback loop in mice. *Hepatology* **71**, 1988–2004 (2020).
9. Wu, A. et al. Loss of VGLL4 suppresses tumor PD-L1 expression and immune evasion. *EMBO J.* **38**, https://doi.org/10.15252/embj.201899506 (2019).
10. Teng, A. C. et al. IRF2BP2 is a skeletal and cardiac muscle-enriched ischemia-inducible activator of VEGFA expression. *FASEB J.* **24**, 4825–4834 (2010).
11. Baxter, S. K. et al. Molecular diagnosis of childhood immune dys-regulation, polyendocrinopathy, and enteropathy, and implications for clinical management. *J. Allergy Clin. Immunol.* **149**, 327–339 (2022).
12. Korholz, J. et al. Novel mutation and expanding phenotype in IRF2BP2 deficiency. *Rheumatology* **62**, 1699–1705 (2023).
13. Keller, M. D. et al. Mutation in IRF2BP2 is responsible for a familial form of common variable immunodeficiency disorder. *J. Allergy Clin. Immunol.* **138**, 544–550.e544 (2016).
14. Cheng, Z. Y., He, T. T., Gao, X. M., Zhao, Y. & Wang, J. ZBTB transcription factors: key regulators of the development, differentiation and effector function of T cells. *Front Immunol.* **12**, 713294 (2021).
15. Suliman, B. A., Xu, D. & Williams, B. R. The promyelocytic leukemia zinc finger protein: two decades of molecular oncology. *Front. Oncol.* **2**, 74 (2012).
16. Jin, Y., Nenseth, H. Z. & Saatcioglu, F. Role of PLZF as a tumor suppressor in prostate cancer. *Oncotarget* **8**, 71317–71324 (2017).

17. Clotaire, D. Z. J., Wei, Y., Yu, X., Ousman, T. & Hua, J. Functions of promyelocytic leukaemia zinc finger (Plzf) in male germline stem cell development and differentiation. *Reprod. Fertil. Dev.* **31**, 1315–1320 (2019).

18. Sha, J. et al. Promyelocytic leukemia zinc finger controls type 2 immune responses in the lungs by regulating lineage commitment and the function of innate and adaptive immune cells. *Int. Immunopharmacol.* **130**, 111670 (2024).

19. Labbaye, C. et al. A three-step pathway comprising PLZF/miR-146a/CXCR4 controls megakaryopoiesis. *Nat. Cell Biol.* **10**, 788–801 (2008).

20. Labbaye, C. et al. PLZF induces megakaryocytic development, activates Tpo receptor expression and interacts with GATA1 protein. *Oncogene* **21**, 6669–6679 (2002).

21. Frith, M. C., Saunders, N. F., Kobe, B. & Bailey, T. L. Discovering sequence motifs with arbitrary insertions and deletions. *PLoS Comput. Biol.* **4**, e1000071 (2008).

22. Chasapis, C. T. & Spyroulias, G. A. RING finger E(3) ubiquitin ligases: structure and drug discovery. *Curr. Pharm. Des.* **15**, 3716–3731 (2009).

23. Holm, L., Laiho, A., Toronen, P. & Salgado, M. DALI shines a light on remote homologs: One hundred discoveries. *Protein Sci.* **32**, e4519 (2023).

24. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).

25. Zauli, G. et al. PMA-induced megakaryocytic differentiation of HEL cells is accompanied by striking modifications of protein kinase C catalytic activity and isoform composition at the nuclear level. *Br. J. Haematol.* **92**, 530–536 (1996).

26. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

27. Chen, Y. L. et al. Super-enhancer-driven IRF2BP2 enhances ALK activity and promotes neuroblastoma cell proliferation. *Neuro Oncol.*, https://doi.org/10.1093/neuonc/noae109 (2024).

28. Stogios, P. J., Downs, G. S., Jauhal, J. J., Nandra, S. K. & Prive, G. G. Sequence and structural analysis of BTB domain proteins. *Genome Biol.* **6**, R82 (2005).

29. Melnick, A. M. et al. The ETO protein disrupted in t(8;21)-associated acute myeloid leukemia is a corepressor for the promyelocytic leukemia zinc finger protein. *Mol. Cell Biol.* **20**, 2075–2086 (2000).

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

31. Kabsch, W. Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).

32. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).

33. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).

34. Schrodinger, L. L. C. *The PyMOL Molecular Graphics System, Version 2.4.1*, https://pymol.org (2015).

35. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

36. Hieronymus, K. et al. Validation of reference genes for whole blood gene expression analysis in cord blood of preterm and full-term neonates and peripheral blood of healthy adults. *BMC Genomics* **22**, 489 (2021).

Acknowledgements

Author contributions

G.W. conducted the biochemical, structural and functional studies, and participated in the data analyses and paper writing. T.L. and L.Z. participated in the discussion. J.D. conceived the study, participated in the experimental design and data analyses, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-54889-5.

Correspondence and requests for materials should be addressed to Jianping Ding.

Peer review information *Nature Communications* thanks Hsiao-Huei Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.