the FEBS Journal

Structure of human Sp140 PHD finger: an atypical fold interacting with Pin1

Chiara Zucchelli[1], Simone Tamburri[1,2,3], Giacomo Quilici[1], Eleonora Palagano[2], Andrea Berardi[1], Mario Saare[4], Pärt Peterson[4], Angela Bachi[2] and Giovanna Musco[1]

1 Dulbecco Telethon Institute, c/o Biomolecular NMR Laboratory, S. Raffaele Scientific Institute, Milano, Italy
2 IFOM-Firc Institute of Molecular Oncology, Milano, Italy
3 San Raffaele Vita-Salute University, Milano, Italy
4 Department of Molecular Pathology, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia

Keywords
histones; nmr; phd finger; pin1; sp140

Correspondence
G. Musco, Dulbecco Telethon Institute, c/o Biomolecular NMR Laboratory, S. Raffaele Scientific Institute, Via Olgettina 58, Milano 20132, Italy
Fax: +0039 02264 344153
Tel: +0039 02264 34824
E-mail: musco.giovanna@hsr.it

(Received 16 September 2013, revised 24 October 2013, accepted 25 October 2013)

doi:10.1111/febs.12588

Sp140 is a nuclear leukocyte-specific protein involved in primary biliary cirrhosis and a risk factor in chronic lymphocytic leukemia. The presence of several chromatin related modules such as plant homeodomain (PHD), bromodomain and SAND domain suggests a role in chromatin-mediated regulation of gene expression; however, its real function is still elusive. Herein we present the solution structure of Sp140-PHD finger and investigate its role as epigenetic reader *in vitro*. Sp140-PHD presents an atypical PHD finger fold which does not bind to histone H3 tails but is recognized by peptidylprolyl isomerase Pin1. Pin1 specifically binds to a phosphopeptide corresponding to the L3 loop of Sp140-PHD and catalyzes *cis–trans* isomerization of a $_{p}$Thr-Pro bond. Moreover co-immunoprecipitation experiments demonstrate FLAG-Sp140 interaction with endogenous Pin1 *in vivo*. Overall these data include Sp140 in the list of the increasing number of Pin1 binders and expand the regulatory potential of PHD fingers as versatile structural platforms for diversified interactions.

Introduction

Human Sp140 is an interferon inducible, 85.9 kDa nuclear leukocyte-specific protein expressed in mature B cells, plasma cell lines and in some T cells [1]. Originally identified as autoantigen in the serum of a patient affected by primary biliary cirrhosis [2], it is also implicated in innate immune response to HIV-1 by its interaction with the virus Vif protein [3]. Importantly, a genome-wide association study of 299 983 tagging single-nucleotide polymorphisms for chronic lymphocytic leukemia (CLL) showed that *Sp140* is a CLL risk locus [4]. In this case a significant dose relationship between genotype and Sp140 expression in lymphocytes was demonstrable, with risk alleles associated with reduced levels of mRNA [4]. In accordance with these data, 16 single-nucleotide polymorphisms were recently suggested to be involved in the etiology of CLL and linked to a decreased *Sp140* expression by means of expression quantitative trait loci analysis [5]. *Sp140* was also identified in a large set of new genes supposed to drive the development of CLL and displaying somatic mutations in CLL with relevant clinical correlates [6]. Sp140 localizes to LYSp100-associated nuclear dots (LANDs) in B-lymphocytic cell lines [7] and as a member of the Sp100 family of proteins it is also found in promyelocytic leukemia nuclear bodies (PML-NBs) of differentiated HL60 and NB4 cells and in adenovirus-Sp140-infected T24 and HeLa cells [1,8]. Sp140 localization to PML-NBs, which are subnuclear structures

Abbreviations
AIRE, autoimmune regulator; BRD, bromodomain; CLL, chronic lymphocytic leukemia; CSD, chemical shift difference; EBFP, enhanced blue fluorescent protein; HSQC, heteronuclear single quantum coherence; PHD, plant homeodomain; PPIase, peptidylprolyl isomerase; PTM, post-translational modification; TEV, tobacco etch virus.

involved in the regulation of gene transcription, cellular growth, apoptosis and maintenance of chromatin architecture [9], along with the presence of several chromatin related modules in its primary structure, suggest a role in chromatin-mediated regulation of gene expression. Indeed coactivator activity was inferred for Sp140 by virtue of its Gal4 DNA-binding domain fusion activity in transfected COS cells [10,11]; it has therefore been hypothesized that Sp140 might regulate the expression of genes involved in CLL development [5]. In line with its putative role in transcriptional regulation, Sp140 has strong sequence homology with autoimmune regulator (AIRE), a transcriptional activator governing the ectopic expression of peripheral tissue-specific antigens in the thymus [12]. Similarly to AIRE, Sp140 harbors a nuclear localization signal, a dimerization domain (HSR or CARD domain), a SAND domain, and a plant homeodomain (PHD) finger (Fig. 1A). At variance with AIRE, which contains a second PHD finger, Sp140 harbors a bromodomain (BRD). Both BRD and PHD fingers are evolutionarily conserved 'reader/effector' modules that bind to specific histone post-translational modifications (PTMs) promoting chromatin changes and/or protein recruitment [13]. The Zn^{2+} binding PHD finger (~ 60 amino acids) is the latest addition to the list of epigenetic readers. It is found in ~ 200 human proteins, many of which act as nucleosome interaction determinants playing a fundamental role in histone recognition and epigenetic mechanisms [14–16]. It can recognize the methylation status of histone lysines, such

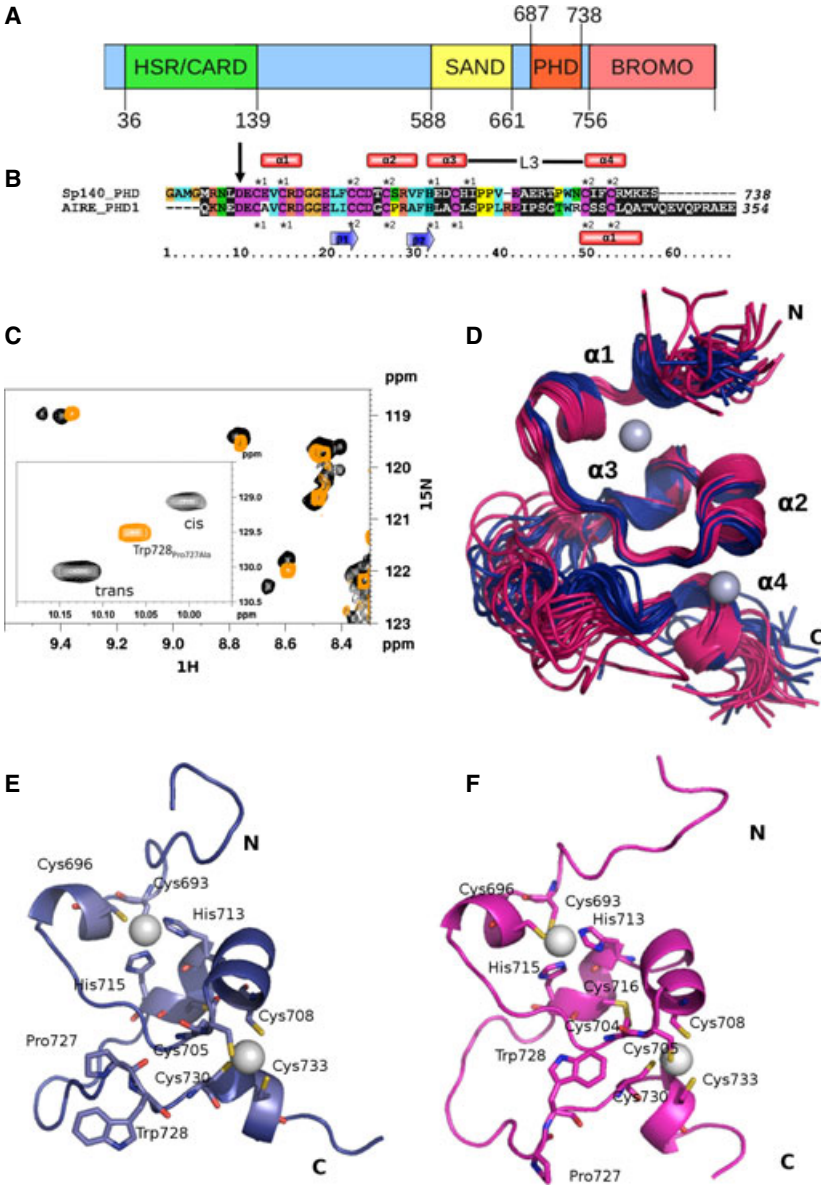

Fig. 1. Sp140-PHD solution structure. (A) Domain organization of human Sp140 full-length protein. (B) Sequence alignment of Sp140-PHD and AIRE-PHD1; their secondary structure elements are indicated at the top and the bottom of the alignment, respectively. The arrow indicates the conserved Asp in position 9. The label L3 indicates residues forming the variable L3 loop. The symbols *1 and *2 indicate the residues coordinating the first and the second Zn^{2+} ions, respectively. Italic numbering is for full-length Sp140 protein and AIRE sequences. Additional N-terminal residues (GAMG) in Sp140-PHD result from TEV protease cleavage of the recombinant protein. (C) Zoom of ^{1}H-^{15}N HSQC spectrum of 0.3 mM Sp140-PHD (black) and Sp140-PHD_{P727A} (orange), in 20 mM NaH_2PO_4/Na_2HPO_4 pH 6.3, 150 mM NaCl, 5 mM dithiothreitol, 50 μM $ZnCl_2$, T = 295 K. Mutation of Pro727 into Ala removes peak duplication due to Thr726-Pro727 *cis–trans* isomerization. In the inset are shown the NH_ε peaks corresponding to Trp726 in Sp140-PHD (black, *trans* and *cis* conformers) and in Sp140-PHD_{P727A} (orange), respectively. (D) Solution structure of Sp140-PHD; superposition of the best 20 structures of *trans* (blue) and *cis* (magenta) conformers. Grey spheres represent the Zn^{2+} ions. Cartoon representation of (E) the *trans* (blue) and (F) *cis* (magenta) conformers. Zn^{2+} binding residues, Cys704, Cys716, Pro727, Trp728, are represented with sticks.

as histone H3 lysine 4 (K4me0 versus K4me3/2) or H3 lysine 36 (H3K36), to smaller degree the methylation state of H3R2 (R2me0 versus R2me2), and the acetylation state of H3K14 [15,16]. Depending on the PTMs and the molecular context, decoding of histone H3 by PHD fingers can lead to gene activation or repression. Importantly, Sp140-PHD finger shares 52% sequence identity with AIRE-PHD1 and contains the typical N-terminal acidic hallmark usually suggestive of recognition of the unmodified histone H3 tail (H3K4me0; Fig. 1B). Notably, non-histone dependent functions have also been attributed to the PHD finger, which behaves as a versatile structural scaffold characterized by a wide functional diversity, ranging from protein–protein interaction hub to E3 ligase activity in sumoylation or ubiquitinylation reactions [14–17]. Despite several indications suggesting an implication of Sp140 in human malignancies, until now its function in both physiological and pathological conditions has remained extremely elusive and unexplored. As a first step towards the understanding of Sp140 function we have solved the solution structure of its PHD finger (Sp140-PHD) and investigated its possible role as epigenetic reader *in vitro*. Sp140-PHD presents an atypical PHD finger fold characterized by the presence of four short α helices and by *cis–trans* isomerization of a peptidyl–prolyl bond located in the variable L3 loop (according to the definition in [18]). Importantly, biochemical experiments and NMR titrations show that Sp140-PHD is not able to decode the histone H3 tail neither in its modified nor in its unmodified form. Conversely, NMR titrations provide evidence that the peptidylprolyl isomerase (PPIase) Pin1 binds directly to Sp140-PHD and is able to recognize a phosphorylated peptide corresponding to the Sp140-PHD finger L3 loop and to catalyze the rapid isomerization of its *cis–trans* peptidyl–prolyl bond. Importantly, co-immunoprecipitation experiments in cells transfected with FLAG-Sp140 demonstrate its interaction with endogenous Pin1 in a cellular context. Although Sp140 function needs further studies, data provided by this study include Sp140 in the list of the increasing number of Pin1 targets and suggest a Pin1-regulated modulation of the biological role of Sp140-PHD finger.

Results

Sp140-PHD solution structure shows prolyl isomerization in the L3 loop

We solved the solution structure of Sp140-PHD finger by multidimensional heteronuclear NMR spectroscopy. The recombinant protein (Met687-Ser738) behaves as a monomer in solution, as assessed by its rotational correlation time ($t_c \sim 4.7$ ns) determined from ^{15}N relaxation data. This is in agreement with the expected value for a folded 6 kDa protein. Importantly, the ^{1}H-^{15}N HSQC (heteronuclear single quantum coherence) spectrum presented peak duplication for 24 amide signals, compatible with the presence of two conformations in slow exchange. We hypothesized that the two sets of peaks might arise from propagation of structural changes due to *cis–trans* isomerization around the Thr726-Pro727 imide bond, a sequence which is known to favor this conformational rearrangement [19,20]. Indeed, mutation of Pro727 into Ala removed peak duplication in the ^{1}H-^{15}N HSQC spectrum, confirming our hypothesis (Fig. 1C). Peptidyl–prolyl bond conformations were assigned on the basis of the proline diagnostic chemical shift difference $\Delta = \delta^{13}C_\beta - \delta^{13}C_\gamma$, which showed Δ values of 4.43 and 9.67 ppm for the *trans* and *cis* configurations, respectively [21]. Further evidence was obtained from ^{13}C-edited nuclear Overhauser effect spectroscopy (NOESY), which showed two sets of NOE cross-peaks between H_αThr726-H_δPro727 and H_αThr726-H_αPro727, typical for the *trans* and *cis* conformations, respectively (Fig. S1H,I) [22]. Finally, volume integration of the duplicated amide cross-peaks in the ^{1}H-^{15}N HSQC spectrum indicated that at room temperature the two conformers were present with 66% in *trans* and 33% in *cis*. The exchange rate between the two conformations was too slow to be detected on the NMR timescale, as assessed by the absence of exchange peaks between *cis* and *trans* resonances. A single NMR data set contained the necessary information to simultaneously determine the structures of the two conformers [23] (Fig. 1D). In both families of structures the residues Leu690-Ile718, Cys730-Met735 adopt a well-defined tertiary structure with an rmsd of ~ 0.45 Å for backbone atoms and have all residues in the allowed regions of the Ramachandran plot (Table 1). Superposition of the backbone atoms of the structured regions of *cis* and *trans* structures indicates that the two conformers are virtually identical in these regions and that *cis–trans* isomerization increases structural heterogeneity in the L3 loop. Accordingly, the largest chemical shift differences between the two conformers were observed within this loop and in residues Phe703-Cys705 which are near in space to the L3 loop (Fig. S1A–G). In line with the paucity of the NOEs detected in this region, residues in L3 showed a reduction of the heteronuclear NOE intensities (Fig. S2).

Overall, the Sp140-PHD structure presents some peculiarities compared with the canonical PHD finger

Table 1. Structural statistics Sp140-PHD.

	<SA>*trans*[a]	<SA>-*cis*[a]
Restraints information		
Total number of experimental distance restraints[b]	885	753
Intra-residual	435	394
Sequential	233	208
Short–medium	141	87
Long	76	64
Zn^{2+} coordination restraints	8	8
Dihedral angle restraints (phi and psi)	52	48
Residual dipolar couplings	30	19
Deviation from idealized covalent geometry		
Bonds (Å)	0.003 ± 0.0001	0.1717 ± 0.0044
All dihedral angle restraints (°)	0.340 ± 0.003	0.50 ± 0.18
Coordinate rms deviation (Å)[b]		
Ordered backbone atoms (N, Cα, C′)	0.46 ± 0.10	0.48 ± 0.15
Ordered heavy atoms	0.95 ± 0.12	1.088 ± 0.15
Ramachandran quality parameters (%)[c]		
Residues in most favoured regions	88.3	87.2
Residues in allowed regions	11.7	12.8
Residues in additional allowed regions[c]	0	0
Residues in disallowed regions	0	0

[a] Simulated annealing. Statistics refer to the ensemble of 20 structures with the lowest energy calculated for the trans and cis conformers, respectively.

[b] Statistics are given for residues D9–I36 and C48–M53 with respect to the average structure.

[c] Statistics are given for residues D9–I36 and C48–M53.

fold. First, one Zn^{2+} binding site, usually formed by a CysCysHisCys motif, is replaced by a CysCysHisHis motif (Fig. 1E,F). This coordination pattern was confirmed by several NOEs involving the metal coordinating residues (H_βCys693 and $H_{\delta2}$His713, H_αCys696 and $H_{\varepsilon1}$His717). Both $N_{\delta1}$ of His713 and His717 are protonated and Zn^{2+} coordination occurs through the $N_{\varepsilon2}$ of the two imidazole rings, as judged from their chemical shifts in the 2D ^1H–^{15}N long-range HMQC spectrum (data not shown). Importantly, the involvement of both His713 and His717 in metal coordination excludes the conserved Cys716 from the Zn^{2+} binding site, thus allowing its side chain to be in close proximity to the Cys704 thiol group. Notably, NOEs between H_β atoms of Cys704 and Cys716 along with downshifts of their C_β resonances (50.9 and 42.8 ppm for Cys704 and Cys716, respectively) indicate the presence of a disulfide bond between these two cysteines (Fig. 1E,F). Another structural peculiarity of Sp140-PHD consists in the presence of two α helices involving residues Asp706-Val711 (α2) and His713-His717 (α3), respectively (Fig. 1D). Notably, in other PHD finger structures (e.g. AIRE-PHD1), residues corresponding to Val711-His713 usually form the second strand of a short antiparallel β-sheet, which is absent in Sp140-PHD (Fig. S4). A search for structural homologues using the DALI server [24] failed to identify any structural neighbor, indicating that Sp140-PHD belongs to a structurally different class of PHD fingers. Indeed, despite the high sequence identity, superposition of Sp140-PHD onto the AIRE-PHD1 structure shows a high rmsd (7.03 Å) on 50 equivalent residues (Fig. S4). Finally, a small hydrophobic cluster, composed of Phe703, Val721, Ile731, stabilizes the structure. Notably, the conserved Trp728, which is usually part of the hydrophobic core of the PHD finger fold, is partially accessible or totally exposed in the *cis* and *trans* conformer, respectively (Fig. 1E,F).

Sp140-PHD does not bind to histone H3 tail peptides

To investigate the possible role of Sp140-PHD in chromatin-regulating complexes, we examined its putative binding to histone tails. Prompted by the presence of a conserved N-terminal acidic hallmark, suggestive of a binding preference for the unmodified histone H3 tail (Fig. 1B), we analyzed binding of Sp140-PHD to a non-methylated peptide corresponding to the first 15 amino acids of histone H3 (H3K4me0) by using 2D ^1H-^{15}N NMR. Upon addition of a fivefold excess (1 mM) of H3K4me0 into ^{15}N-labeled Sp140-PHD we did not observe any interaction, as assessed by the absence of peak displacement in the ^1H-^{15}N HSQC spectrum (Fig. S5A). Further NMR titrations of ^{15}N Sp140-PHD with other H3 peptides bearing different epigenetic marks, such as H3K4me3, H3R2me2a (asymmetric di-methylation of R2) and H3K9ac, or with unmodified peptides corresponding to H3 (17–29) or H4 (1–10) did not show any binding (Fig. S6). Similar negative results were obtained with Sp140-

$PHD_{Pro45Ala}$ mutant (data not shown). To test whether other histone post-translational modifications and/or combinations thereof might be crucial for a possible interaction with Sp140-PHD, we performed binding assays using the MODified™ Histone Peptide Array (Active Motif, Carlsbad, CA, USA). The array contains 384 peptides (19 amino acids long) in various combinations of known and hypothetical modification states of the H3, H4, H2A and H2B histone tails. Despite the extensive coverage of histone modifications we did not observe any specific binding to GST-Sp140-PHD (Fig. S5B). We hypothesize that one of the reasons determining the lack of interaction might be related to some Sp140-PHD structural peculiarities. On the one hand we observed that the preformed anchoring pocket usually exploited by PHD fingers to anchor the positively charged N-terminus of histone H3 (e.g. AIRE-PHD1; Fig. S7A) [25] is absent or partially covered by Trp728 in Sp140-PHD *trans* and *cis* conformers, respectively (Fig. S7B,C). On the other hand, in both Sp140-PHD conformers the conserved aspartate in position 9 (Fig. 1B), which is usually a fundamental residue for the recognition of H3K4me0, is unfavorably oriented pointing in the opposite direction with respect to the canonical histone binding surface (Fig. S7).

Human Pin1 binds to the phosphorylated peptide corresponding to Sp140-PHD L3 loop and catalyzes the isomerization of the $_p$Thr-Pro bond *in vitro*

We next wondered whether the peptidyl–prolyl *cis–trans* isomerization observed in Sp140-PHD might have a functional relevance, as this conformational exchange process is emerging as a versatile regulatory strategy to modulate cell signaling, protein transcription, transport degradation and/or localization [26–29]. In this context, human PPIase Pin1 plays a fundamental role catalyzing the *cis–trans* isomerization of phosphorylated Ser/Thr-Pro peptide bonds in an increasing number of targets [30,31]. As Pin1 is emerging as a mediator of immune cell function [32,33], we asked whether the PHD of the leukocyte-specific protein Sp140 might be a substrate for human Pin1. With this aim we first tested *in vitro* Pin1 enzyme activity on two peptides, $EAER_pTPWN$ and EAERTPWN, corresponding to Sp140-PHD L3 loop (Glu722-Asn729) with or without threonine phosphorylation, respectively. Because of the slow exchange rate of the peptidyl–prolyl *cis–trans* isomerization, several residues in both the free peptides displayed two distinct sets of ^{1}H signals in 2D ROESY experiments (Fig. 2A,B). The *cis* and *trans* populations of the peptides were 15% and 85%, respectively, as estimated from 1D ^{1}H and 2D ^{1}H-^{13}C HSQC spectra at room temperature. Exchange cross-peaks were absent in the ROESY spectra of the free peptides, indicating that the exchange regime between the two conformations was too slow to be detected on the NMR timescale (Fig. 2A left, B left). Notably, addition of catalytic amounts of Pin1 to $EAER_pTPWN$ accelerated the isomerization rate of the phosphothreonine–prolyl bond, as shown by the appearance of exchange cross-peaks in the ROESY spectrum (Fig. 2A, right). As expected, in the presence of Pin1 no exchange peaks were observed for the non-phosphorylated control peptide (Fig. 2B, right).

We next determined the binding site of $EAER_pTPWN$ on Pin1, performing NMR based chemical shift mapping assays. To this end 2D ^{1}H-^{15}N HSQC spectra of full-length ^{15}N-labeled Pin1 were recorded to monitor possible changes in Pin1 ^{1}H-^{15}N chemical shifts upon successive additions of unlabeled peptides. A comparison of the spectra in the absence and presence of a fivefold excess (1 mM) of Sp140 peptides showed that only $EAER_pTPWN$ was able to bind Pin1, as revealed by the numerous peak displacements observed upon addition of the phosphopeptide (Fig. 2C, Fig. S8A). The unphosphorylated peptide did not show any binding evidence, confirming the phospho dependence of the Pin1–peptide interaction (Fig. S8B). The complex between Pin1 and $EAER_pTPWN$ was in the fast exchange regime on the NMR chemical shift timescale (Fig. 2C) with a dissociation constant of 138 ± 4 μM (Fig. 2D). Pin1 residues exhibiting significant amide chemical shift changes (Fig. 2E) where mapped on the Pin1 crystallographic structure (pdb code 1PIN). The binding surface mainly involved the β-sheet of the WW domain (Lys13-Ser16, Gly20, Val22-Asn26, Ala31, Gln33-Arg35; Fig. 2F). Chemical shift changes were observed also on the flexible linker (Ser41, Ser43, Lys46) and on the region of the PPIase domain facing (Lys97-Glu100) or nearby (Phe139-Arg142) the WW domain, probably induced by long-range conformational rearrangements upon complex formation. Overall, these data indicate that Pin1 binds to $EAER_pTPWN$ and catalyzes the *cis–trans* isomerization of its phosphothreonine–proline bond *in vitro*.

Pin1 recognizes the Sp140-PHD scaffold independently of phosphorylation

We next asked whether Pin1 was able to recognize the entire Sp140-PHD finger scaffold and we performed NMR binding assays titrating ^{15}N-labeled Pin1 with

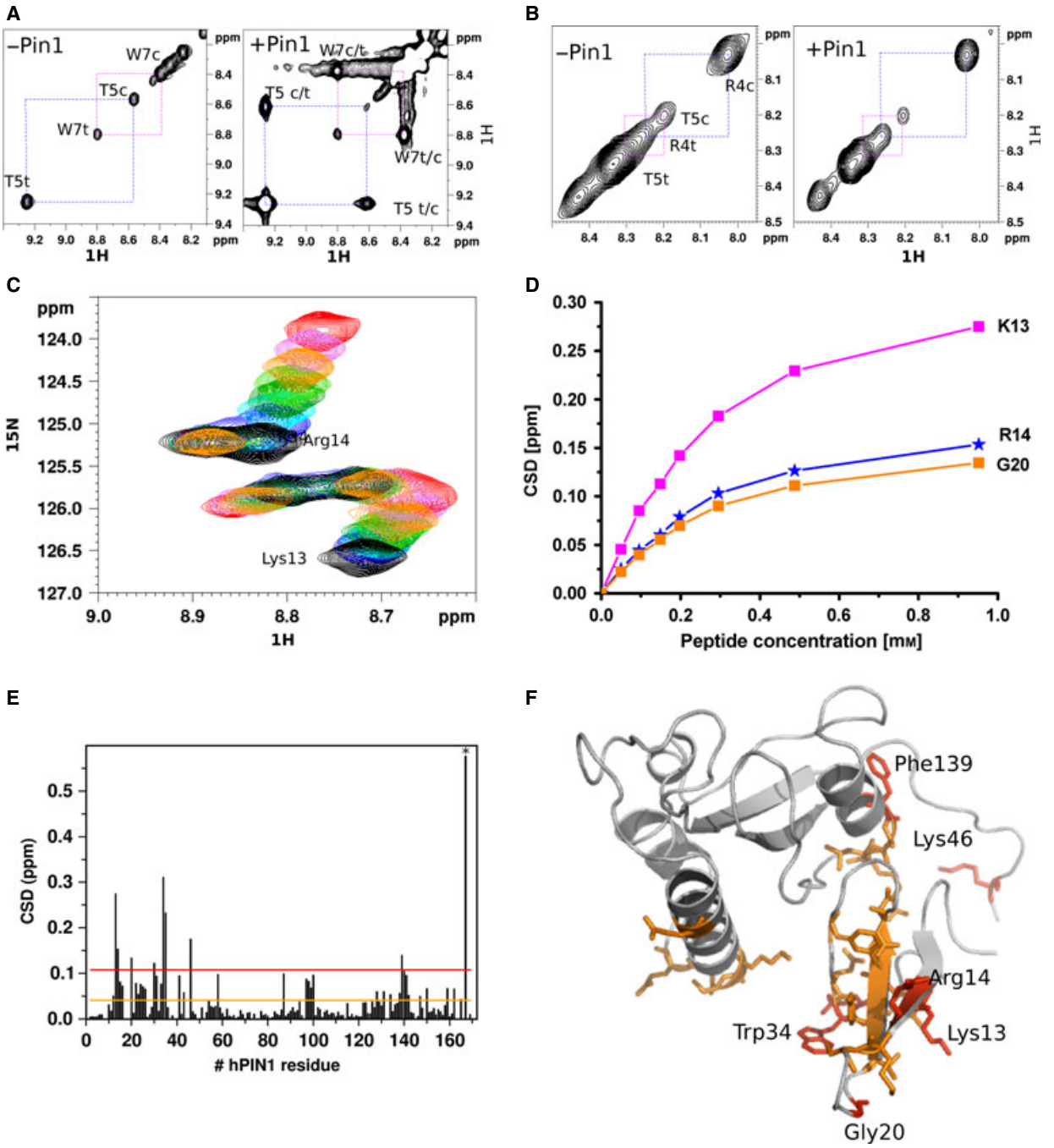

Fig. 2. Pin1 binds to EAER$_p$TPWN and catalyzes *cis–trans* isomerization of the $_p$Thr-Pro bond. Zoom of ^{1}H-^{1}H ROESY spectra of (A) EAER$_p$TPWN and (B) EAERTPWN peptides, free (left) or in the presence (right) of catalytic amounts of Pin1 (0.1 mM), $T = 301$ K, $t_{mix} = 300$ ms, 14 T. (C) Selected region of Pin1 (0.2 mM) ^{1}H-^{15}N HSQC spectra during the titration with increasing amounts of EAER$_p$TPWN peptide (0, 0.1, 0.2, 0.3, 0.4, 0.6, 1 mM). The starting and end titration points are represented in black and red, respectively. The observed chemical shift changes are a continuous and monotonic function of the amount of added peptide, indicating that the binding is in the fast exchange limit on the NMR timescale. (D) Representative isotherm binding curves derived from the analysis of the CSD of selected residues (Lys13, Arg14, Gly20) upon successive addition of EAER$_p$TPWN. (E) Histogram showing the values of Pin1 CSD induced upon addition of a fivefold excess of EAER$_p$TPWN (1 mM). Residue numbers are indicated on the *x* axis (residues for which the CSD is missing are either prolines or could not be detected because of exchange with the solvent). The star indicates the CSD of NH_ε of Trp34. (F) Cartoon representation of Pin1 (PDB code 1PIN); the residues showing CSDs larger than the mean value and the mean value plus one standard deviation are shown in orange and red, respectively.

unlabeled Sp140-PHD. Notably, upon addition of sub-stoichiometric amounts of Sp140-PHD a number of Pin1 resonances shifted in the ^{1}H-^{15}N HSQC spectrum. At equimolar ratio several Pin1 peaks disappeared broadening out from the spectrum, indicating binding in the intermediate exchange regime. Upon addition of a 1.5 excess (0.3 mM) of Sp140-PHD almost all Pin1 peaks disappeared, with the exception of residues from the flexible N-terminus (Fig. 3A–C). Interestingly, analogous line broadening effects have been observed in response to binding to Pin1 of the full-length substrate stem-loop binding protein [34]. Similarly to what was observed in the titration with the phosphopeptide, residues shifting upon Sp140-PHD addition involved the WW domain (Ser16, Gly20-Asn26, Thr29, Ser32, Gln33, Glu35) and the flexible linker (Ser41, Ser43, Gly45). Most importantly, spectral perturbation propagated throughout the protein involving additional residues in the PPIase domain around the catalytic pocket (Thr152, Asp153, Ser154, Gly155) and around the basic cluster (Lys63, His64, Arg69), suggesting that accommodation of the full-length substrate in the interdomain space induces further interactions and/or conformational rearrangements with respect to the phosphopeptide (Fig. 3D,E). Titrations with Sp140-PHD$_{Thr726Asp}$, a mutant mimicking the phosphorylation of Thr726, led essentially to similar results (Fig. S9), suggesting that Pin1 is able to recognize the PHD finger scaffold independently of phosphorylation. In this context we cannot exclude that the aspartate mimics the phosphorylation only partially, as it is smaller, unbranched and with a lower charge density with respect to a phosphorylated threonine. The reverse titration of ^{15}N Sp140-PHD with unlabeled Pin1 confirmed the interaction with both the *trans* (Fig. 4) and *cis* (Fig. S10) conformers, as

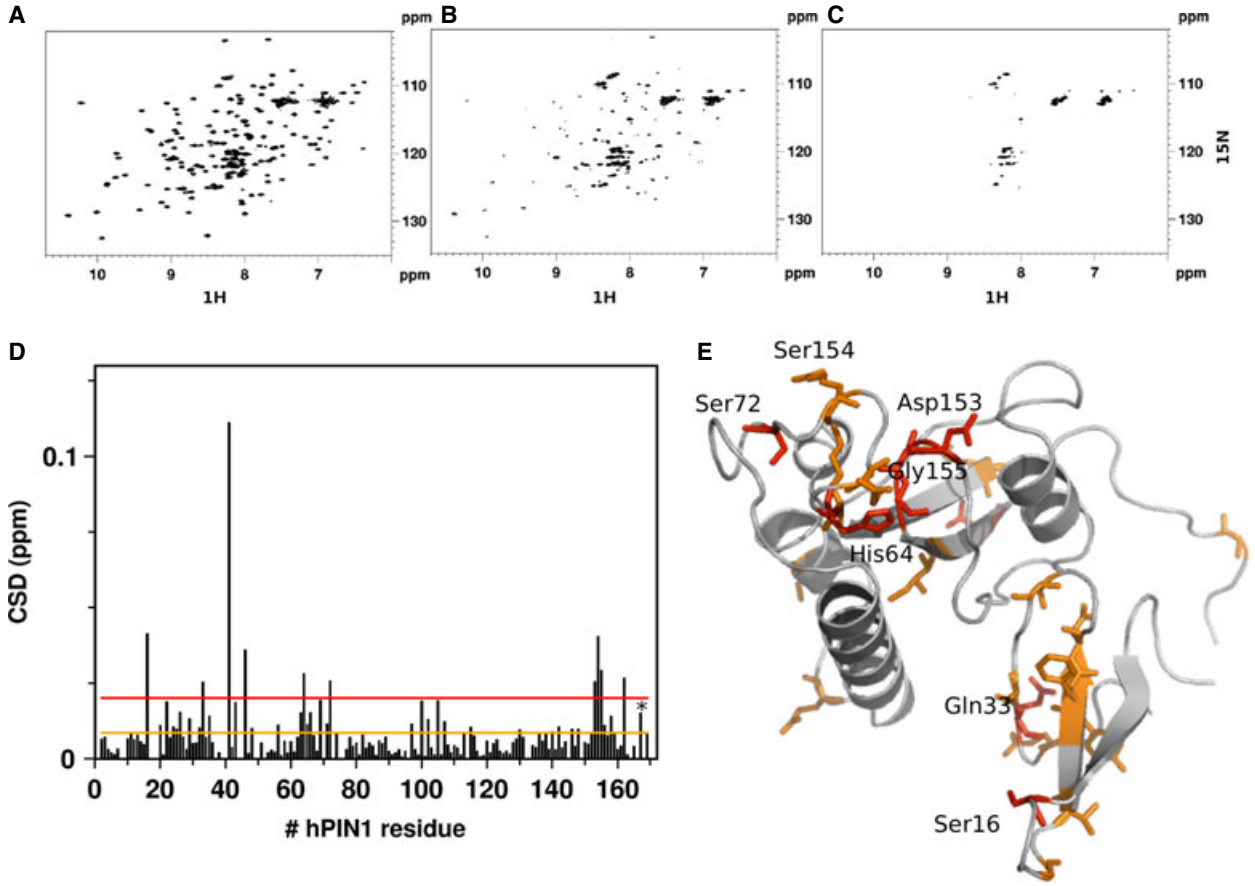

Fig. 3. Sp140-PHD binds to Pin1; mapping of the interaction. ^{1}H-^{15}N HSQC spectrum of ^{15}N Pin1 (0.2 mM) without (A) and with (B) 0.2 mM Sp140-PHD (1 : 1; 0.2 mM) and (C) with an excess (1 : 1.5) of Sp140-PHD (0.3 mM), in 150 mM NaCl, 20 mM Tris/HCl, pH 6.6, and 5 mM dithiothreitol, $T = 301$ K. (D) Histogram showing Pin1 CSD values upon binding to 0.1 mM Sp140-PHD (1 : 0.5). Residue numbers are indicated on the *x* axis (residues for which the CSD is missing are either prolines or could not be assigned because of exchange with the solvent). The star indicates the CSD of NH_{ε} of Trp34. (E) Cartoon representation of Pin1; the residues showing CSD values larger than the mean value and the mean value plus one standard deviation are shown in orange and red, respectively.

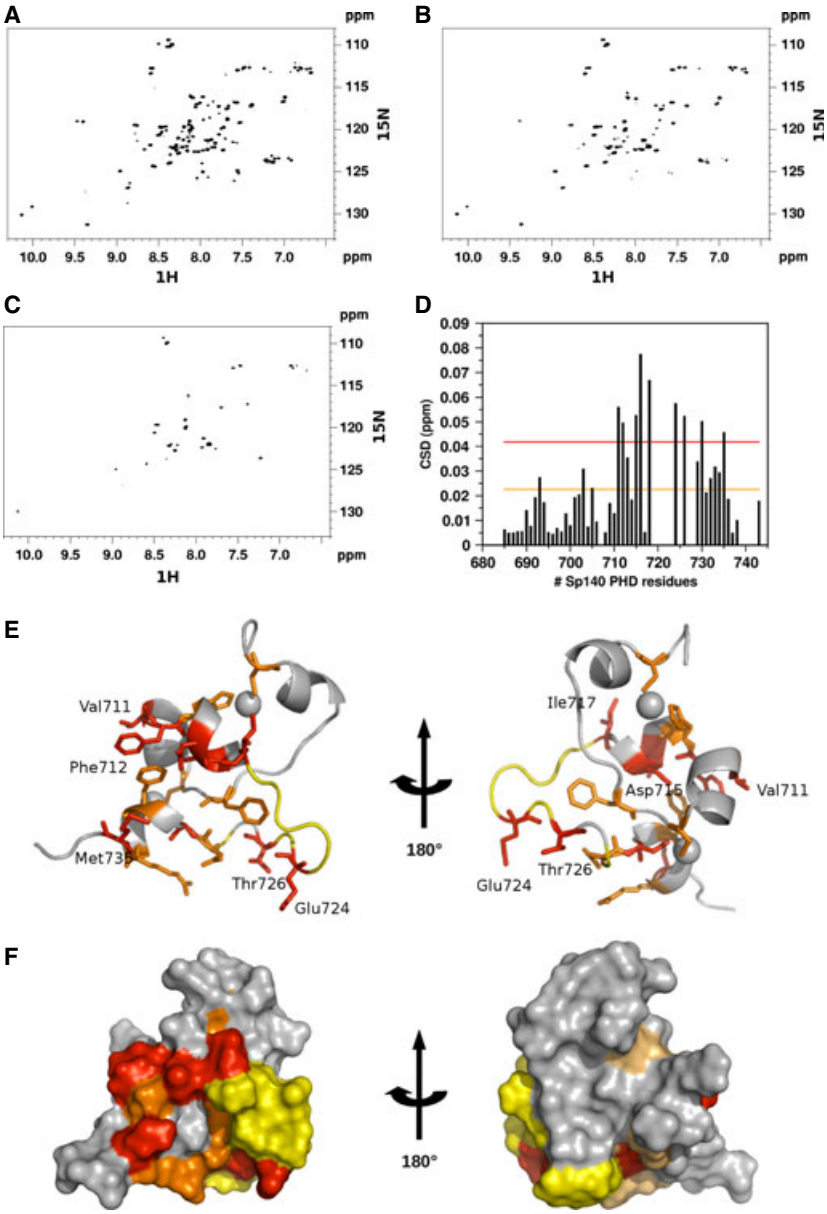

Fig. 4. Pin1 binds to Sp140-PHD; mapping of the interaction. ^{1}H-^{15}N HSQC spectrum of 0.2 mM ^{15}N Sp140-PHD without (A) and with (B) 0.1 mM Pin1 (molar ratio 1 : 0.5), and (C) with an excess (molar ratio 1 : 1.5) of Pin1 (0.3 mM). (D) Histogram showing Sp140-PHD (*trans* conformer) CSD values upon addition of 0.1 mM Pin1. Residue numbers are indicated on the *x* axis (residues for which the CSD is missing are either prolines or disappeared upon binding). (E) Cartoon and (F) surface representation of Sp140-PHD (*trans* conformation); the residues showing CSDs larger than the mean value and the mean value plus one standard deviation are shown in orange and red, respectively. Residues disappearing upon addition of Pin1 are colored in yellow. The CSD values and the mapping on the Sp140-PHD in *cis* conformation are reported in Fig. S10.

assessed by peak shifting and broadening upon addition of Pin1 (Fig. 4A–C). Sp140-PHD residues shifting in the presence of sub-stoichiometric amounts of Pin1 (0.1 mM; Sp140-PHD : Pin1 1 : 0.5) included not only the L3 loop (Ala41-Cys48) but also amino acids located on α2, α3 and α4 helices (Val711, Phe712, Asp715, Cys716, Ile717, Met735; Figs 4D–F, S10), suggesting either a direct contact or long-range conformational effects upon binding. As expected, despite Pin1 direct interaction with Sp104-PHD, it was not able to catalyze *cis–trans* isomerization, as assessed by the unaltered peak volume ratio (33% *cis* and 66% *trans*) observed in the ^{1}H-^{15}N HSQC spectra in the presence of a catalytic amount of Pin1. Taken together these data indicate that *in vitro* Pin1 is able to recognize the Sp140-PHD finger scaffold but does not catalyze *cis–trans* isomerization.

Sp140 interacts *in vivo* with Pin1

To verify whether the interaction between Sp140 and Pin1 occurs also *in vivo* we performed a co-immunoprecipitation assay in HEK293T cells, transiently transfected with FLAG-tagged Sp140 or with FLAG-tagged enhanced blue fluorescent protein (EBFP) as control. As reported in Fig. 5, anti-Pin1 western blot

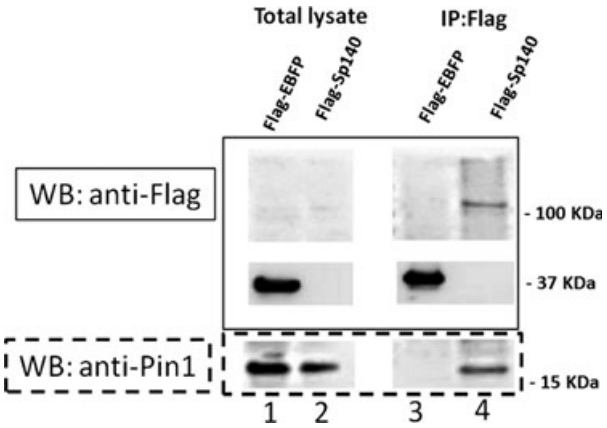

Fig. 5. Sp140 interacts with Pin1 *in vivo*. HEK293T cells were transiently transfected with FLAG-tagged Sp140 or with FLAG-tagged EBFP as control. In lanes 1 and 2 the levels of expression of FLAG-EBFP, FLAG-Sp140 and endogenous Pin1 are shown (50 μg of total lysate loaded). After anti-FLAG immunoprecipitation both EBFP and Sp140 are highly enriched. Notably, Pin1 is co-precipitated only with Sp140 (lane 4) and not with EBFP (lane 3) demonstrating that Pin1 and Sp140 can interact *in vivo*.

analysis of the FLAG immunoprecipitation shows that Pin1 is co-precipitated only with Sp140 and not in the control, thus demonstrating that the two proteins can interact also *in vivo*.

Discussion

In recent years the PHD finger domain, one of the most recurrent domains in nuclear proteins, has been extensively investigated from both the structural and functional point of view [14,15,35,36]. This small Zn^{2+} binding motif has emerged as a robust conserved scaffold with diversified activities: it can work not only as an epigenetic reader sensing the modification status of histone H3, but can also function as a general protein–protein interaction motif, thereby expanding its role in diverse cellular processes including transcriptional regulation and/or signal transduction [37]. Its high functional versatility relies on the low secondary structure content and on subtle but significant changes in amino acid compositions contributing to the domain functional and structural plasticity. In this context the structure of the PHD finger of the leukocyte-specific nuclear Sp140 protein represents a paradigmatic example for the structural and functional versatility attributed to this domain. In fact, structural comparison with AIRE-PHD1 reveals for Sp140-PHD an unexpected switch from an α/β to an all α-helical fold, conceivably imputable to few differences in the primary structure (Fig. 1B). For example, the presence in Sp140-PHD of a glutamate residue in position 12 (an alanine in AIRE-PHD1) favours the formation of a salt-bridge with Arg697, thus stabilizing a helical turn in this region. A second helix (α2), encompassing residues Asp706-Val711, is similarly stabilized by electrostatic interactions. In AIRE-PHD1 formation of this helix is probably hindered by the presence of a proline in position 27. In Sp140-PHD helix α2 is immediately followed by a third helix (α3) in which His713 and His717 form a helical zinc anchor site [38] that coordinates together with Cys693 and Cys696 the second Zn^{2+} ion, thus replacing the canonical CysCys-HisCys coordination scheme. The helix-turn-helix arrangement involving α2 and α3 impairs the formation of the short β-strands usually encompassing residues in positions 20–22 and 29–31. A further Sp140-PHD structural peculiarity consists in the unprecedented presence of *cis–trans* peptidyl–prolyl isomerization (Thr726-Pro7275 imide bond) in the variable L3 loop, conferring high structural heterogeneity to this region. Notably, in PHD fingers specialized in histone H3 tail recognition, this loop forms a narrow cavity to accommodate the positively charged N-terminus of histone H3 [14,15,35,36]. This preformed pocket is absent in the *trans* conformer and partially covered by Trp726 in the *cis* conformer (Fig. S7). Considering the importance of the H3A1 pocket as an anchoring element for histone H3 recognition, we hypothesize that the absence of an appropriate binding surface in this region might be one of the structural determinants hampering Sp140-PHD binding to histone H3 tail peptides *in vitro*. In this context, it should also be noted that the conserved aspartate in position 9 (Fig. 1B), which is considered the hallmark for unmethylated H3K4 recognition, points in the opposite direction with respect to the canonical histone binding surface (Fig. S7B,C). It is therefore conceivable that the combination of all these structural features strongly compromise the ability of Sp140-PHD to recognize the N-terminal part of the histone H3 tail, suggesting that the so-called acidic hallmark is not sufficient to predict recognition of unmodified H3K4. At this stage we cannot exclude that *in vivo* the Thr726-Pro727 bond isomerization, together with Thr726 phosphorylation, might play a role in chromatin–Sp140 interactions, as regulatory mechanisms via *cis–trans* peptidyl–prolyl isomerization are not unusual in the context of epigenetic readers. A remarkable example is offered by the MLL1 PHD-BRD cassette, where a *cis–trans* proline within the domain linker binds to the proline isomerase Cyp33, causing dramatic conformational changes in the domain orientation and preventing H3K4me3 interaction, ultimately resulting in HOX target gene

repression [39]. In the context of full-length Sp140 we also hypothesize that its putative involvement in chromatin interactions might be promoted and/or reinforced by other chromatin related domains, such as the SAND domain, a DNA binding module [40], and/or the BRD, another epigenetic reader specialized in the decoding of acetylated histones [41]. Notably, the L3 loop is characterized by high sequence and structural variability within the PHD family, contributing to the functional versatility of PHD fingers [37]. In fact, previous studies aimed at engineering the PHD finger scaffold with tailored functions have shown that grafting of CtBP2 binding motif into the L3 loop of Mi2β-PHD resulted in a functional domain switch [18]. Importantly, the L3 loop constitutes the structural determinant for the binding of several PHD fingers to non-histone proteins. This is the case for Pygo-PHD, where an α helix within the L3 loop constitutes the interaction surface with the homology domain 1 of BCL9 [42]. In line with this concept, we reasoned that the Sp140-PHD L3 loop, which is conserved among different mammalian species (Fig. S3), might play a relevant role in Sp140-PHD function. Prompted by the *cis–trans* isomerization of the Thr726-Pro727 bond, we asked whether Sp140-PHD could be a substrate for Pin1, a unique human PPIase which is able to catalyze the *cis–trans* isomerization of the phosphorylated Ser/Thr-Pro bond [43,44]. Pin1 is a component of the nuclear speckles macromolecular complex, including cell cycle proteins as well as elements of the splicing machinery [31]. Depending on specific target sites and local structural constraints, human Pin1 catalyzes *cis* to *trans* or *trans* to *cis* isomerization, thereby modifying the conformation, the stability and the activity of phosphorylated target proteins. In this respect the enzyme acts as an effective molecular timer playing a significant role in biological and pathological processes such as immune and cellular stress responses, microbial infections, cancer, cell cycle progression, growth-signal and gene regulation [30,31,45]. Herein we showed by co-immunoprecipitation experiments in HEK293T cells transfected with FLAG-Sp140 that Sp140 interacts *in vivo* with endogenous Pin1. Moreover, *in vitro* NMR binding experiments show that Pin1 specifically binds to a phosphopeptide corresponding to the L3 loop of Sp140-PHD, thus catalyzing *cis–trans* isomerization of the $_{p}$Thr-Pro bond. NMR-based chemical shift mapping indicates that the phosphopeptide mainly targets the WW domain. The chemical shift difference pattern is highly reminiscent of previously described interactions with other Pin1 phosphopeptide targets, which also bind to Pin1 with micromolar affinity and target mainly the WW domain [46–48]. Importantly, binding experiments performed with the entire Sp140-PHD domain showed that *in vitro* Pin1 recognizes the whole PHD scaffold, independently from threonine phosphorylation in L3, but does not catalyze *cis–trans* isomerization of the peptidyl–prolyl bond. In line with the NMR results obtained on phosphorylated and non-phosphorylated L3 loop peptides, the absence of isomerization catalysis of recombinant Sp140-PHD by Pin1 is probably due to a missing phosphorylation on Thr726. It is noteworthy that the residues whose backbone chemical shifts were affected by addition of Sp140-PHD are not only located in the WW domain but propagate throughout the domain involving the flexible interdomain linker and the PPIase domain, within and in spatial proximity to the basic cluster and the catalytic site. These results are in agreement with a recent study suggesting a non-catalytic participation of the PPIase domain in target binding [49]. However, the molecular mechanisms governing Pin1 interdomain rearrangement, substrate recognition and peptidyl–prolyl bond isomerization are still largely debated, as no 3D structure is available describing Pin1 interaction with a full-length substrate. Taken together, the data presented in this work provide an example of how malleable the PHD fold can be, thus expanding its regulatory potential as a versatile structural platform for diversified interactions. In this context Pin1 phosphorylation dependent *cis–trans* isomerization of the Thr726-Pro727 bond could act as a molecular switch to modulate Sp140 cellular fate and its interaction with chromatin. Herein this mechanism might then orchestrate the crosstalk between several Sp140 PTMs such as phosphorylation, acetylation, ubiquitylation and sumoylation, thus determining Sp140 turnover and/or cellular localization.

Experimental procedures

Sp140 PHD finger expression and purification

Human Sp140-PHD finger (residues Met687-Ser738, NM_007237) was cloned into *Nco*I/*Kpn*I sites of pETM11 expression vector (EMBL). The vector expresses the domain with N-terminal His6 tag, removable by cleavage with TEV (tobacco etch virus) protease. Site-directed mutations were made by standard overlap extension methods. BL21 (DE3) *Escherichia coli* cells were induced overnight at 30 °C with 1 mM isopropyl thio-β-D-galactoside (IPTG), in LB medium supplemented with 0.2 mM $ZnCl_2$. Cells were sonicated in buffer containing 20 $\mu g \cdot mL^{-1}$ RNase A, 2 $\mu g \cdot mL^{-1}$ DNase I, 150 mM NaCl, 20 mM Tris/HCl pH 8, 10 mM imidazole pH 8, 0.2% NP-40, 50 μM $ZnCl_2$, 0.4 mM dithiothreitol and

complete EDTA-free (Roche, Mannheim, Germany). The His6-tagged protein was purified on an Ni-nitrilotriacetic acid (NTA) column (GE Healthcare, Uppsala, Sweden) and eluted with 150 mM NaCl, 20 mM Tris/HCl pH 8, 50 μM $ZnCl_2$, 2 mM β-mercapto-EtOH and 300 mM imidazole pH 8. The His6 tag was cleaved off during overnight dialysis at 4 °C, by addition of His6-tagged TEV protease (home-made). The TEV protease was then removed by purification on an Ni-NTA column; Sp140-PHD finger was further purified by size exclusion chromatography (HiLoad 16/60 Superdex 30 pg column; GE Healthcare). The final buffer contained 20 mM Na_2HPO_4/NaH_2PO_4 pH 6.3, 150 mM NaCl, 5 mM dithiothreitol and 50 μM $ZnCl_2$. Protein identity was confirmed by mass spectroscopy. Uniformly ^{15}N- and ^{13}C-^{15}N-labeled Sp140-PHD finger was expressed by growing *E. coli* BL21 (DE3) cells in minimal bacterial medium containing $^{15}NH_4Cl$, with or without ^{13}C-D-glucose. For binding assays with histone peptide arrays the Sp140-PHD finger was cloned into *Nco*I/*Kpn*I sites of pETM30 expression vector (EMBL). Purification of the His6-GST-tagged protein was performed as described above, without cleavage of the GST-fusion protein and by size exclusion chromatography on a HiLoad 16/60 Superdex 75 pg column (GE Healthcare).

Pin1 expression and purification

His8-tagged human Pin1 (plasmid kindly provided by J. P. Noel, Salk Institute for Biological Studies, CA, USA) was expressed in *E. coli* BL21 (DE3) cells, induced for 4 h at 37 °C with 0.2 mM IPTG. Sonication buffer was 150 mM NaCl, 20 mM Tris/HCl pH 8, 10 mM imidazole pH 8, 0.2% NP-40, 4 mM β-mercapto-EtOH and complete EDTA-free (Roche). His8-tagged Pin1 was purified on an Ni-NTA column (GE Healthcare) and eluted with 150 mM NaCl, 20 mM Tris/HCl pH 8, 4 mM β-mercapto-EtOH, 300 mM imidazole pH 8. Protein was dialysed overnight against 20 mM Tris/HCl pH 7.2, 4 mM β-mercapto-EtOH, 150 mM NaCl and it was further purified by size exclusion chromatography (HiLoad 16/60 Superdex 75 pg column; GE Healthcare). Protein identity was confirmed by mass spectrometry. The final buffer contained 150 mM NaCl, 20 mM Tris/HCl pH 6.6 and 5 mM dithiothreitol. Uniformly ^{15}N-labeled Pin1 was expressed by growing *E. coli* BL21 (DE3) cells in minimal bacterial medium containing $^{15}NH_4Cl$. The Pin1 ^{1}H-^{15}N HSQC spectrum was assigned based on the human Pin1 ^{1}H and ^{15}N backbone chemical shifts deposited in the BMRB data bank (entry 5305) [50]. The deposited assignment is incomplete; the following residues were missing in the entry and were therefore not assigned in our spectra: Ser19, Gln75, Glu76 and Glu145 (numbering scheme according to Pin1 crystallographic structure, 1PIN.pdb). The peaks corresponding to Gly39, Asn40, Gly44, Gln49, Gly50, Ser114 and Lys132 were not detectable in our Pin1 ^{1}H-^{15}N HSQC spectra, probably because of solvent exchange phenomena.

NMR spectroscopy and resonance assignments

NMR experiments were performed at 295 K on Bruker Avance 600 and 900 MHz spectrometers equipped with inverse triple resonance cryoprobe and pulsed field gradients. Data were processed with NMRPIPE [51] or TOPSPIN 2.0 (Bruker, Karlsruhe, Germany) and analyzed using CCPNMR [52]. Sp140-PHD finger sample concentrations were 0.8–1.2 mM in 20 mM NaH_2PO_4/Na_2HPO_4 pH 6.3, 150 mM NaCl, 5 mM dithiothreitol, 50 μM $ZnCl_2$ and 10% or 100% (v/v) D_2O. ^{1}H, ^{15}N and ^{13}C backbone resonances were assigned through the following 2D and 3D experiments: ^{1}H-^{15}N HSQC, ^{1}H-^{13}C HSQC, HNCA, HNCO, CBCA (CO)NH, CBCANH. ^{1}H and ^{13}C side chain resonances were obtained by 2D and 3D experiments (^{1}H-^{1}H TOCSY, HCCH-TOCSY, CC(CO)NH and HCC(CO)NH). The tautomeric state of the histidine rings was determined by performing a long range ^{1}H-^{15}N HMQC, optimized to detect *J*-couplings in histidine side chains (J(HN) = 22 Hz) [53]. Proton–proton distance constraints were obtained from ^{15}N and ^{13}C separated 3D NOESY and from 2D ^{1}H-^{1}H NOESY spectra in H_2O and D_2O (120 ms mixing time). ^{3}J(HN, Ha) coupling constants were measured to derive restraints for Φ dihedral angles. Additional Φ/Ψ restraints were obtained from backbone chemical shifts using TALOS+ [54]. ^{1}H-^{15}N residual dipolar couplings were measured in isotropic and anisotropic phases created by the addition of 20 mg·mL^{-1} Pf1 phage (ASLA Biotech Ltd, Riga, Latvia). Heteronuclear ^{1}H-^{15}N NOEs as well as longitudinal and transversal ^{15}N relaxation rates were measured using standard 2D methods [55]. The relaxation delays were applied in an interleaved manner. The T1 and T2 decay curves were sampled at 14 (50–2600 ms) and 12 (14.4–244.8 ms) different time points, respectively. The heteronuclear NOE experiments were run twice in an interleaved fashion with and without (reference spectrum) proton saturation during proton recovery delay. The relaxation experiments were analyzed with the program NMRVIEW.5.03 [56].

Structure calculation

The structures of Sp140-PHD finger in *trans* and *cis* conformation were separately calculated using ARIA 2.3.1 software [57] in combination with CNS using the experimentally derived restraints (Table 1). NOESY spectra were manually assigned and calibrated by ARIA 2.3.2. A total of eight iterations was performed, 100 structures were computed in the last two iterations and ARIA default water refinement was performed on the 20 best structures of the final interaction. In ARIA 2.3.2, the geometry of the Zn^{2+} coordination is fixed through covalent bonds and angles in the CNS parameters; the tetrahedral angles and distances for Zn^{2+} coordinating residues were maintained also after water refinement. Several NOEs were observed

between Cys704 and Cys716 revealing spatial proximity between the two residues. Calculations with or without imposing the disulfide bond between Cys704 and Cys716 were virtually identical. Structural quality was assessed using PROCHECK-NMR [58] and molecular images were generated by PYMOL (http://pymol.org/). The family of the 20 lowest energy structures for Sp140-PHD finger in *trans* and *cis* conformations has been deposited in the Protein Data Bank with the accession codes 2md7 and 2md8 for the *cis* and *trans* conformation, respectively. Chemical shift and restraints lists that were used in the structure calculations have been deposited in BioMagResBank (19472 and 19473 for the *trans* and *cis* conformations, respectively).

NMR binding assays

For NMR binding assays the following synthetic peptides were purchased from Caslo Lyngby, Denmark: H3K4me0$_{1\text{-}15}$ ARTKQTARKSTGGKA; H3K4me3 (ARTKme3QTARKS, ARme2aTKQTARKS (asymmetric di-metylation of R2); H3K9ac (ARTKQTARK$_{ac}$S); H4 peptide (SGRGKGGKGL); phosphorylated EAER$_p$TPWN and non-phosphorylated EAERTPWN (both with N-acetylation and C-amidation). Peptide purity (> 98%) was confirmed by HPLC and mass spectrometry. Titration of ^{15}N Sp140-PHD with histone peptides was performed in 20 mM phosphate buffer, pH 6.8, 2 mM dithiothreitol, 150 mM NaCl. For NMR binding assays involving Pin1 all NMR titrations were performed in the absence of phosphate salts, in a buffer containing 20 mM Tris/HCl pH 6.6, 150 mM NaCl and 5 mM dithiothreitol, to avoid the presence of inorganic phosphate which inhibits binding between Pin1 and its targets [44]. Protein concentrations were determined by UV spectroscopy using the predicted extinction coefficients of ε_{280} 5599 $M^{-1}\cdot cm^{-1}$ and 20 970 $M^{-1}\cdot cm^{-1}$ for Sp140-PHD and Pin1, respectively. Peptide concentrations were estimated from their mass. In order to minimize dilution and NMR signal loss, titrations were carried out by adding to the protein samples small aliquots of concentrated (15 mM) peptide stock solutions. For each titration point (typically 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4, 5 equivalents of ligand) a 2D water-flip-back ^{15}N-edited HSQC spectrum was acquired with 512 (100) complex points, 55 ms (60 ms) acquisition times, apodized by 60° shifted squared (sine) window functions and zero filled to 1024 (512) points for ^{1}H (^{15}N), respectively. Assignment of the complex Pin1 : EAER$_p$TPWN was made by following individual cross-peaks through the titration series. For each residue the weighted average of the ^{1}H and ^{15}N chemical shift difference (CSD) was calculated as CSD = $[(\Delta^2HN + \Delta^2N/25)/2]^{1/2}$ [59]. The binding constant of EAER$_p$TPWN to Pin1 was estimated by monitoring the variation of CSD of individual peaks (nine peaks: Lys13, Arg14, Gly20, Asn30, Trp34, Glu35, Lys46, Phe139, Trp34ε). Assuming a simple binary reaction between protein and peptide, dissociation constants were obtained from least squares fitting of CSD as a function of total ligand concentration according to

$$\delta_i = \frac{b - \sqrt{b^2 - 4ac}}{2a} \quad (1)$$

with $a = (K_a/\delta_b)[P_t]$, $b = 1 + K_a([L_{ti}] + [P_t])$ and $c = \delta_b K_a[L_{ti}]$, where δ_i is the absolute change in chemical shift for each titration point, $[L_{ti}]$ is the total ligand concentration at each titration point, $[P_t]$ is the total protein concentration, $K_a = 1/K_d$ is the binding constant and δ_b is the chemical shift of the resonance in the complex. K_d and δ_b were used as fitting parameters using the ORIGIN program.

To monitor binding of Sp140-PHD to ^{15}N-Pin1, a concentrated solution of unlabeled Sp140-PHD (2.5 mM) was stepwise titrated into a 0.2 mM solution of ^{15}N-Pin1 up to a 1.5 molar excess (0.3 mM Sp140-PHD), and each titration step was monitored by recording a 2D ^{1}H-^{15}N HSQC spectrum. The reverse titration was also performed adding a concentrated solution of unlabelled Pin1 (1.8 mM) to a 0.2 mM solution of ^{1}H-^{15}N Sp140-PHD. Because of the large peak broadening effects in the HSQC spectra already at sub-stoichiometric protein : ligand ratio (1 : 0.25), it was not possible to determine a dissociation constant from CSD values and to proceed with the titration beyond a 1 : 1.5 molar ratio.

Enzyme activity analysis

NMR experiments were performed at $T = 301$ K on 6.5 mM solutions of EAER$_p$TPWN or EAERTPWN peptides, corresponding to phosphorylated and unphosphorylated Sp140-PHD finger L3 loop, in 20 mM Tris/HCl, 150 mM NaCl and 5 mM dithiothreitol, 90% H_2O and 10% D_2O (pH 6.6), with or without 0.1 mM Pin1 (molar ratio Pin1 : peptide 1 : 65). 2D ^{1}H-^{1}H TOCSY and ROESY spectra were recorded with spectral widths of 7183.91 and 6009.61 Hz in t1 and t2 dimensions, respectively. ROESY spectra were acquired at a mixing time of 300 ms with 32 scans, while TOCSY spectra were recorded at a mixing time of 60 ms with 32 scans.

Histone overlay assays

MODifiedTM Histone Peptide Arrays were purchased by Active Motif®. They enable screening in a single experiment 59 acetylation, methylation, phosphorylation and citrullination modifications on the entire N-terminal tails of histones H2A, H2B, H3 and H4. A series of synthetic 19mer histone H2A, H2B, H3 and H4 peptides, each of which may contain as many as four modifications, are spotted in duplicate onto a glass slide, generating a total of 384 unique histone modification combinations. Following overnight blocking at 4 °C with 5% milk in TTBS buffer (10 mM Tris/HCl pH 7.4, 150 mM NaCl, 0.05% Tween 20),

the array was washed twice with TTBS and once with binding buffer (50 mM Tris/HCl pH 7.5, 300 mM NaCl, 0.1% NP-40, proteinase inhibitors). The array was then incubated, for 2–4 h at room temperature, with 1 μM solution of GST-tagged recombinant protein in binding buffer. After three washes with binding buffer, the array was incubated with primary antibody anti-GST (1 : 1000) in 5% milk/TTBS for 1 h at room temperature. The array was then washed three times with TTBS and incubated for 1 h at room temperature with a secondary antibody HRP-conjugated (1 : 10 000) in 5% milk/TTBS. Three washes with TTBS followed, and ECLTM Western Blotting detection solution (GE Healthcare) was added and incubated on the array surface for 5 min at room temperature. The image was finally captured by the ImageQuant™ ECL image analysis system (GE Healthcare).

HEK293T transfection with FLAG-Sp140 and FLAG-EBFP

The FLAG-Sp140 was cloned in pFLAG-CMV-5a vector (Sigma, St. Louis, MO, USA). HEK293T cells were transiently transfected by using Turbofect reagent (Thermo).

Co-immunoprecipitation and western blot analysis

HEK293T cells transiently transfected with FLAG-Sp140 were cultured in DMEM with 10% fetal bovine serum and 5% glutamine whereas HEK293T cells transiently transfected with EBFP were used as control. Cells were harvested for 24 h and then lysed in JS buffer (75 mM NaCl, 50 mM HEPES pH 7.5, 1% glycerol, 1% Triton X-100, 1.5 mM $MgCl_2$, 5 mM EGTA) for 20 min. Phosphatase and protease inhibitors were added. Lysate was centrifuged at 16 100 ***g***. for 20 min at 4 °C. For co-immunoprecipitation experiments, lysates were precleaned and then incubated with anti-FLAG M2 affinity resin (Sigma) overnight at 4 °C. The day after, the sample was centrifuged at 1000 ***g*** for 1 min and the supernatant was discharged. Beads were washed three times with the JS buffer. The bound proteins were eluted by using the FLAG-peptide 1X from Sigma for 40 min at 4 °C by end-over mixing and run on a 4–12% SDS/PAGE gel (Invitrogen, Carlsbad, CA, USA). Western blot analysis was carried out using an antibody rabbit anti-FLAG (Antibodies-online) and an antibody rabbit anti-Pin1 (Abcam, Cambridge, UK).

Multiple sequence alignment

Sp140 protein sequence alignments and the phylogenetic tree were obtained from Ensembl gene tree ENSGT00510000046835 [60]. The alignments and phylogenetic tree were edited with JALVIEW [61] and TREEGRAPH 2 [62], respectively.

Acknowledgements

We thank Professor Joseph P. Noel (Salk Institute for Biological Studies, CA, USA) for PIN1 plasmid. GM thanks Dr Luca Mollica for acquiring NMR experiments in the first stage of the project and Telethon Foundation (TCP99035) and Associazione Italiana Ricerca sul cancro (Airc; grant no. 13159) for financial support. MS and PP were supported by the Estonian Research Agency grant IUT2-2, the Center of Excellence of Translational Medicine and Tartu University Development Fund (Center of Translational Genomics). ST conducted this study as partial fulfillment of his PhD in Molecular Medicine, Program in Cellular and Molecular Biology, San Raffaele University, Milan, Italy. CZ conducted this study as partial fulfillment of her PhD in biochemistry, University of Milan. We wish to thank CERM Infrastructure (Florence) for access to the 900 MHz spectrometer for NMR measurements.

References

1 Bloch DB, de la Monte SM, Guigaouri P, Filippov A & Bloch KD (1996) Identification and characterization of a leukocyte-specific component of the nuclear body. *J Biol Chem* **271**, 29198–29204.
2 Granito A, Yang WH, Muratori L, Lim MJ, Nakajima A, Ferri S, Pappas G, Quarneti C, Bianchi FB, Bloch DB *et al.* (2010) PML nuclear body component Sp140 is a novel autoantigen in primary biliary cirrhosis. *Am J Gastroenterol* **105**, 125–131.
3 Madani N, Millette R, Platt EJ, Marin M, Kozak SL, Bloch DB & Kabat D (2002) Implication of the lymphocyte-specific nuclear body protein Sp140 in an innate response to human immunodeficiency virus type 1. *J Virol* **76**, 11133–11138.
4 Di Bernardo MC, Crowther-Swanepoel D, Broderick P, Webb E, Sellick G, Wild R, Sullivan K, Vijayakrishnan J, Wang Y, Pittman AM *et al.* (2008) A genome-wide association study identifies six susceptibility loci for chronic lymphocytic leukemia. *Nat Genet* **40**, 1204–1210.
5 Sille FC, Thomas R, Smith MT, Conde L & Skibola CF (2012) Post-GWAS functional characterization of susceptibility variants for chronic lymphocytic leukemia. *PLoS ONE* **7**, e29632.
6 Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, Ramsay AJ, Bea S, Pinyol M, Martinez-Trillos A *et al.* (2011) Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* **44**, 47–52.
7 Dent AL, Yewdell J, Puvion-Dutilleul F, Koken MH, de The H & Staudt LM (1996) LYSP100-associated

nuclear domains (LANDs): description of a new class of subnuclear structures and their relationship to PML nuclear bodies. *Blood* **88**, 1423–1426.
8 Bloch DB, Chiche JD, Orth D, de la Monte SM, Rosenzweig A & Bloch KD (1999) Structural and functional heterogeneity of nuclear bodies. *Mol Cell Biol* **19**, 4423–4430.
9 Weidtkamp-Peters S, Lenser T, Negorev D, Gerstner N, Hofmann TG, Schwanitz G, Hoischen C, Maul G, Dittrich P & Hemmerich P (2008) Dynamics of component exchange at PML nuclear bodies. *J Cell Sci* **121**, 2731–2743.
10 Bloch DB, Nakajima A, Gulick T, Chiche JD, Orth D, de La Monte SM & Bloch KD (2000) Sp110 localizes to the PML-Sp100 nuclear body and may function as a nuclear hormone receptor transcriptional coactivator. *Mol Cell Biol* **20**, 6138–6146.
11 Zong RT, Das C & Tucker PW (2000) Regulation of matrix attachment region-dependent, lymphocyte-restricted transcription through differential localization within promyelocytic leukemia nuclear bodies. *EMBO J* **19**, 4123–4133.
12 Kisand K & Peterson P (2011) Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy: known and novel aspects of the syndrome. *Ann N Y Acad Sci* **1246**, 77–91.
13 Yap KL & Zhou MM (2010) Keeping it in the family: diverse histone recognition by conserved structural folds. *Crit Rev Biochem Mol Biol* **45**, 488–505.
14 Bienz M (2006) The PHD finger, a nuclear protein-interaction domain. *Trends Biochem Sci* **31**, 35–40.
15 Musselman CA & Kutateladze TG (2011) Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res* **219**, 9061–9071.
16 Sanchez R & Zhou MM (2011) The PHD finger: a versatile epigenome reader. *Trends Biochem Sci* **36**, 364–372.
17 Gaetani M, Matafora V, Saare M, Spiliotopoulos D, Mollica L, Quilici G, Chignola F, Mannella V, Zucchelli C, Peterson P *et al.* (2012) AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. *Nucleic Acids Res* **40**, 11756–11768.
18 Kwan AH, Gell DA, Verger A, Crossley M, Matthews JM & Mackay JP (2003) Engineering a protein scaffold from a PHD finger. *Structure (Camb)* **11**, 803–813.
19 Kay BK, Williamson MP & Sudol M (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J* **14**, 231–241.
20 Exarchos KP, Exarchos TP, Papaloukas C, Troganis AN & Fotiadis DI (2009) Detection of discriminative sequence patterns in the neighborhood of proline *cis* peptide bonds and their functional annotation. *BMC Bioinformatics* **10**, doi: 10.1186/1471-2105-10-113.
21 Shen Y & Bax A (2010) Prediction of Xaa-Pro peptide bond conformation from sequence and chemical shifts. *J Biomol NMR* **46**, 199–204.
22 Wüthrich K (1986) NMR of Proteins and Nucleic Acids. Wiley, New York.
23 Koo BK, Park CJ, Fernandez CF, Chim N, Ding Y, Chanfreau G & Feigon J (2011) Structure of H/ACA RNP protein Nhp2p reveals *cis*/*trans* isomerization of a conserved proline at the RNA and Nop10 binding interface. *J Mol Biol* **411**, 927–942.
24 Holm L, Kaariainen S, Rosenstrom P & Schenkel A (2008) Searching protein structure databases with DaliLite v. 3. *Bioinformatics* **24**, 2780–2781.
25 Chignola F, Gaetani M, Rebane A, Org T, Mollica L, Zucchelli C, Spitaleri A, Mannella V, Peterson P & Musco G (2009) The solution structure of the first PHD finger of autoimmune regulator in complex with non-modified histone H3 tail reveals the antagonistic role of H3R2 methylation. *Nucleic Acids Res* **37**, 2951–2961.
26 Mallis RJ, Brazin KN, Fulton DB & Andreotti AH (2002) Structural characterization of a proline-driven conformational switch within the Itk SH2 domain. *Nat Struct Biol* **9**, 900–905.
27 Min L, Fulton DB & Andreotti AH (2005) A case study of proline isomerization in cell signaling. *Front Biosci* **10**, 385–397.
28 Andreotti AH (2003) Native state proline isomerization: an intrinsic molecular switch. *Biochemistry* **42**, 9515–9524.
29 Wulf G, Finn G, Suizu F & Lu KP (2005) Phosphorylation-specific prolyl isomerization: is there an underlying theme? *Nat Cell Biol* **7**, 435–441.
30 Lu KP, Finn G, Lee TH & Nicholson LK (2007) Prolyl *cis–trans* isomerization as a molecular timer. *Nat Chem Biol* **3**, 619–629.
31 Liou YC, Zhou XZ & Lu KP (2011) Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. *Trends Biochem Sci* **36**, 501–514.
32 Esnault S, Shen ZJ & Malter JS (2008) Pinning down signaling in the immune system: the role of the peptidyl-prolyl isomerase Pin1 in immune cell function. *Crit Rev Immunol* **28**, 45–60.
33 Boussetta T, Gougerot-Pocidalo MA, Hayem G, Ciappelloni S, Raad H, Arabi Derkawi R, Bournier O, Kroviarski Y, Zhou XZ, Malter JS *et al.* (2010) The prolyl isomerase Pin1 acts as a novel molecular switch for TNF-alpha-induced priming of the NADPH oxidase in human neutrophils. *Blood* **116**, 5795–5802.
34 Krishnan N, Lam TT, Fritz A, Rempinski D, O'Loughlin K, Minderman H, Berezney R, Marzluff WF & Thapar R (2012) The prolyl isomerase Pin1 targets stem-loop binding protein (SLBP) to dissociate the SLBP-histone mRNA complex linking histone mRNA decay with SLBP ubiquitination. *Mol Cell Biol* **32**, 4306–4322.

35 Musselman CA & Kutateladze TG (2009) PHD fingers: epigenetic effectors and potential drug targets. *Mol Interv* **9**, 314–323.
36 Patel DJ & Wang Z (2013) Readout of epigenetic modifications. *Annu Rev Biochem* **82**, 81–118.
37 Li Y & Li H (2012) Many keys to push: diversifying the 'readership' of plant homeodomain fingers. *Acta Biochim Biophys Sin (Shanghai)* **44**, 28–39.
38 Andreini C, Bertini I & Cavallaro G (2011) Minimal functional sites allow a classification of zinc sites in proteins. *PLoS ONE* **6**, e26325.
39 Wang Z, Song J, Milne TA, Wang GG, Li H, Allis CD & Patel DJ (2010) Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. *Cell* **141**, 1183–1194.
40 Bottomley MJ, Collard MW, Huggenvik JI, Liu Z, Gibson TJ & Sattler M (2001) The SAND domain structure defines a novel DNA-binding fold in transcriptional regulation. *Nat Struct Biol* **8**, 626–633.
41 Filippakopoulos P & Knapp S (2012) The bromodomain interaction module. *FEBS Lett* **586**, 2692–2704.
42 Fiedler M, Sanchez-Barrena MJ, Nekrasov M, Mieszczanek J, Rybin V, Muller J, Evans P & Bienz M (2008) Decoding of methylated histone H3 tail by the Pygo-BCL9 Wnt signaling complex. *Mol Cell* **30**, 507–518.
43 Bayer E, Goettsch S, Mueller JW, Griewel B, Guiberman E, Mayr LM & Bayer P (2003) Structural analysis of the mitotic regulator hPin1 in solution: insights into domain architecture and substrate binding. *J Biol Chem* **278**, 26183–26193.
44 Ranganathan R, Lu KP, Hunter T & Noel JP (1997) Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* **89**, 875–886.
45 Dilworth D, Gudavicius G, Leung A & Nelson CJ (2012) The roles of peptidyl-proline isomerases in gene regulation. *Biochem Cell Biol* **90**, 55–69.
46 Namanja AT, Wang XJ, Xu B, Mercedes-Camacho AY, Wilson KA, Etzkorn FA & Peng JW (2011) Stereospecific gating of functional motions in Pin1. *Proc Natl Acad Sci USA* **108**, 12289–12294.
47 Jacobs DM, Saxena K, Vogtherr M, Bernado P, Pons M & Fiebig KM (2003) Peptide binding induces large scale changes in inter-domain mobility in human Pin1. *J Biol Chem* **278**, 26174–26182.
48 Wintjens R, Wieruszeski JM, Drobecq H, Rousselot-Pailley P, Buee L, Lippens G & Landrieu I (2001) 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J Biol Chem* **276**, 25150–25156.
49 Innes BT, Bailey ML, Brandl CJ, Shilton BH & Litchfield DW (2013) Non-catalytic participation of the Pin1 peptidyl-prolyl isomerase domain in target binding. *Front Physiol* **4**, 18.
50 Jacobs DM, Saxena K, Grimme S, Vogtherr M, Pescatore B, Langer T, Elshorst B & Fiebig KM (2002) 1H, 13C and 15N backbone resonance assignment of the peptidyl-prolyl *cis–trans* isomerase Pin1. *J Biomol NMR* **23**, 163–164.
51 Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293.
52 Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J & Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696.
53 Pelton JG, Torchia DA, Meadow ND & Roseman S (1993) Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Protein Sci* **2**, 543–558.
54 Shen Y, Delaglio F, Cornilescu G & Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213–223.
55 Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD & Kay LE (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33**, 5984–6003.
56 Johnson BA (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* **278**, 313–352.
57 Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE & Nilges M (2007) ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–382.
58 Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R & Thornton JM (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477–486.
59 Grzesiek S, Stahl SJ, Wingfield PT & Bax A (1996) The CD4 determinant for downregulation by HIV-1 Nef directly binds to Nef. Mapping of the Nef binding surface by NMR. *Biochemistry* **35**, 10256–10261.
60 Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R & Birney E (2009) EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res* **19**, 327–335.
61 Waterhouse AM, Procter JB, Martin DM, Clamp M & Barton GJ (2009) Jalview Version 2 – a multiple

sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191.
62 Stover BC & Muller KF (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics* **11**, doi: 10.1186/1471-2105-11-7.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Histograms of CSD values between *trans* and *cis* conformers.

Fig. S2. Heteronuclear values of Sp140-PHD.

Fig. S3. Alignment showing the conservation of Sp140-PHD among species.

Fig. S4. Superposition of AIRE-PHD1 and Sp140-PHD structures.

Fig. S5. NMR titrations and Modified Histone Peptide Arrays showing lack of interaction between unmodified histone H3 tail and Sp140-PHD.

Fig. S6. NMR titrations showing lack of interaction between modified histone H3 tail and Sp140-PHD.

Fig. S7. Surface representation of AIRE-PHD1 and Sp140-PHD.

Fig. S8. ^{15}N HSQC spectra of ^{15}N-Pin1 with EAER$_p$TPWN and EAERTPWN.

Fig. S9. ^{15}N HSQC spectra of ^{15}N-Pin1 with Sp140-PHD$_{T726D}$.

Fig. S10. Mapping of the interaction between ^{15}N Sp140-PHD in *cis* conformation and Pin1.