

# The MYC oncoprotein directly interacts with its chromatin cofactor PNUTS to recruit PP1 phosphatase

Yong Wei<sup>1,2,3,†</sup>, Cornelia Redel<sup>1,5,†</sup>, Alexandra Ahlner<sup>1,4,†</sup>, Alexander Lemak<sup>1,2,#</sup>, Isak Johansson-Åkhe<sup>1,4,#</sup>, Scott Houlston<sup>1,2,#</sup>, Tristan M.G. Kenney<sup>1,5</sup>, Aaliya Tamachi<sup>1</sup>, Vivian Morad<sup>1,4</sup>, Shili Duan<sup>1,4</sup>, David W. Andrews<sup>1,3,5</sup>, Björn Wallner<sup>1,4</sup>, Maria Sunnerhagen<sup>1,4,\*</sup>, Cheryl H. Arrowsmith<sup>1,2,5,\*</sup> and Linda Z. Penn<sup>1,5,\*</sup>

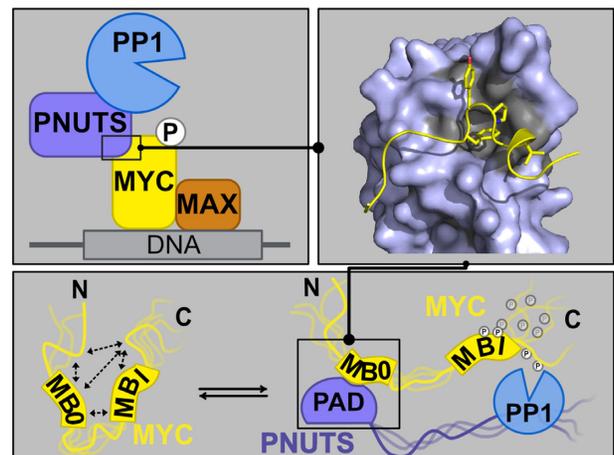
<sup>1</sup>Princess Margaret Cancer Centre, University Health Network, 101 College St, Toronto, ON M5G 0A3, Canada, <sup>2</sup>Structural Genomics Consortium (SGC), University of Toronto, 101 College St., Suite 700, Toronto, ON, M5G 1L7, Canada, <sup>3</sup>Sunnybrook Research Institute, 2075 Bayview Ave. Toronto, ON, M4N 3M5, Canada, <sup>4</sup>Department of Physics, Chemistry, and Biology, Linköping University, SE-581 83 Linköping, Sweden and <sup>5</sup>Department of Medical Biophysics, University of Toronto, 101 College St, Toronto, ON M5G 1L7, Canada

Received May 17, 2021; Revised January 11, 2022; Editorial Decision February 06, 2022; Accepted February 28, 2022

## ABSTRACT

Despite MYC dysregulation in most human cancers, strategies to target this potent oncogenic driver remain an urgent unmet need. Recent evidence shows the PP1 phosphatase and its regulatory subunit PNUTS control MYC phosphorylation, chromatin occupancy, and stability, however the molecular basis remains unclear. Here we demonstrate that MYC interacts directly with PNUTS through the MYC homology Box 0 (MB0), a highly conserved region recently shown to be important for MYC oncogenic activity. By NMR we identified a distinct peptide motif within MB0 that interacts with PNUTS residues 1–148, a functional unit, here termed PNUTS amino-terminal domain (PAD). Using NMR spectroscopy we determined the solution structure of PAD, and characterised its MYC-binding patch. Point mutations of residues at the MYC-PNUTS interface significantly weaken their interaction both *in vitro* and *in vivo*, leading to elevated MYC phosphorylation. These data demonstrate that the MB0 region of MYC directly interacts with the PAD of PNUTS, which provides new insight into the control mechanisms of MYC as a regulator of gene transcription and a pervasive cancer driver.

## GRAPHICAL ABSTRACT



## INTRODUCTION

Dysregulated MYC activity is a hallmark of >50% of human cancers and is often linked to aggressive disease and poor prognosis (1,2). MYC is a master-regulator that controls the transcription of ~15% of genes, thereby regulating numerous biological processes associated with cancer initiation and progression, including cell growth, metabolism and immune response (1,3–6). The core of MYC function is as a regulator of central control points of gene transcription, including epigenetic modifications as well as transcription initiation and elongation (7). MYC regulates these essential steps of chromatin control by directly interacting

\*To whom correspondence should be addressed. Tel: +1 416 634 8770; Email: Linda.Penn@uhnresearch.ca  
Correspondence may also be addressed to Cheryl H. Arrowsmith. Email: Cheryl.Arrowsmith@uhnresearch.ca  
Correspondence may also be addressed to Maria Sunnerhagen. Email: Maria.Sunnerhagen@liu.se

†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

#The authors wish it to be known that, in their opinion, the second set of three authors should be regarded as Joint Second Authors.

with specific proteins as part of the recruitment of particular chromatin complexes. These include TRRAP for recruitment of histone acetyl transferases, TBP a key component of TFIID, and TFIIF for RNA Polymerase II preinitiation complex formation (7). In non-transformed cells, the expression of this critical chromatin-regulator is tightly controlled. By contrast, in cancer MYC activity is dysregulated by a plethora of mechanisms, resulting in constitutive transcriptional activity that drives oncogenic growth. The MYC family of proteins also includes N-MYC and L-MYC, whose expression is normally restricted to fetal development, but can be reactivated and dysregulated in cancer (1). Thus, dysregulated MYC activity unleashes this powerful transcriptomic regulator and potent oncogenic driver of most human cancers.

Evidence from mouse models of cancer strongly suggest that inhibiting MYC oncogenic activity will dramatically improve cancer outcome (7–10), as treatment with cycles of systemic genetic suppression of MYC leads to tumour eradication, without adverse side-effects to normal cells (11–14). Despite these promising results, targeting MYC using traditional drug development approaches has failed (15–17), largely because MYC is intrinsically disordered in the absence of a binding partner (18). Structurally MYC contains a basic helix-loop-helix leucine zipper domain (bHLHLZ), common to many transcription factors, and six regions termed MYC boxes that are unique and highly conserved amongst the MYC family of proteins. MYC Box II (MBII), and more recently MYC Box 0 (MB0), have been shown to be functionally required for the full oncogenic activity of MYC, suggesting these MYC boxes are key regulatory regions (19). A promising strategy to develop MYC inhibitors is to identify MYC protein interactors that are essential for MYC oncogenic activity, and then disrupt these MYC-protein interactions by targeting the structured region of the partner proteins (20,21). The vision is that by identifying structurally unique binding modes of MYC to different protein interactors, an arsenal of inhibitors targeting functionally critical protein interactors will inhibit MYC activity with high specificity. However, only a handful of proteins that directly interact with MYC have been validated.

As a first step to filling this gap, we identified hundreds of novel MYC binding proteins using the BioID in-cell, proximity labelling technique followed by mass spectrometry, which includes both direct and indirect protein interactors (19,22). We then further characterized one of the hits, protein phosphatase 1 (PP1), a serine threonine phosphatase and its regulatory substrate-specifying subunit, PP1 nuclear targeting subunit (PNUTS), as we and others have shown that phosphorylation regulates MYC stability and/or activity (23–26). Indeed, inhibiting PP1 using RNAi or pharmacological inhibitors triggers MYC hyperphosphorylation, leading to chromatin eviction and MYC protein degradation (27). Exploiting this PP1:PNUTS-MYC regulatory axis to banish MYC from chromatin and target MYC destruction has enormous promise; however, inhibiting PP1 catalytic activity is not a viable approach as PP1 has several protein substrates (28–30). Thus, to advance our understanding of how MYC is regulated by this phosphatase complex and to evaluate the potential for pharmacological disruption of MYC interaction with PP1:PNUTS to pro-

mote MYC eviction from chromatin and degradation, we sought to delineate the molecular basis of this interaction.

Here, we report that MYC interacts directly with PNUTS. Using biolayer interferometry (BLI) analysis the interaction has been mapped to: (i) conserved residues 16–33 of MYC, recently termed MB0 and shown to be essential for MYC activity (19,31) and (ii) a functional unit of PNUTS consisting of the N-terminal 148 residues, which we have termed the PNUTS amino-terminal domain (PAD). Building on these results, the molecular interaction of PAD and MB0 was then resolved by Nuclear Magnetic Resonance (NMR) spectroscopy, analyzing each protein in the presence of the partner protein, and by analyzing a PAD-MB0-fusion protein, in conjunction with molecular modeling. These analyses revealed the critical residues and structural details essential for the PNUTS-MYC interaction. Validation of key residues important for the interaction within both proteins was achieved by demonstrating that point mutants of these residues disrupted the PNUTS-MYC interaction *in vitro* and *in vivo*. Moreover, elevated MYC phosphorylation was evident when interaction with PNUTS was abrogated in cells. Taken together, these results not only provide new insights into the molecular basis of MYC interaction with PNUTS, but also provide foundational data for the potential development of drugs targeting PNUTS to disrupt the PNUTS-MYC interaction and inhibit MYC oncogenic activity.

## MATERIALS AND METHODS

### Protein expression and purification

Several constructs within the 1–186 region of PNUTS were evaluated for expression by cloning into the pET28-MHL vector (RRID:Addgene.26096) containing an N-terminal His<sub>6</sub> tag and a TEV cleavage site. PNUTS(1–148) (PAD) was the shortest construct successfully expressed and stably purified in *Escherichia coli* BL21 (DE3) cells. These cells were lysed by sonication with a Misonix S-3000 sonicator for a total processing time of 10 min with cycles of 5 s sonication and 7 s rest in 1× PBS buffer (PBS415, Bioshop). PAD was purified by Ni-NTA agarose (QIAGEN) washed with 1× PBS buffer with 5% glycerol and 2 mM β-ME followed by gel filtration on Superdex 75/300 column (GE Healthcare), equilibrated with the NMR buffer containing 20 mM HEPES, pH 6.9, 200 mM NaCl, 2 mM DTT, 5% glycerol. Unlabeled proteins were grown in Terrific Broth medium and induced with 0.2 mM IPTG at 16 °C overnight. Double (<sup>13</sup>C, <sup>15</sup>N)-labeled PAD for NMR studies was grown in minimal M9 media supplemented with <sup>15</sup>NH<sub>4</sub>Cl and <sup>13</sup>C D-glucose.

For the PAD-MB0-fusion protein, cDNAs encoding PNUTS(1–148), a (GGGS)<sub>2</sub> linker and MYC(13–30) were cloned in frame and inserted into pET15-Trx-MHL (modified by SGC based on pET15-MHL (RRID:Addgene.26096)), containing an N-terminal His<sub>6</sub>-tagged thioredoxin (32) and a TEV cleavage site. Thioredoxin was cleaved using TEV protease in dialysis buffer (1× PBS, 5% glycerol, 2 mM β-ME) at 4 °C overnight, after which the fusion protein was purified as described above.

MYC(1–88) used for NMR experiments was expressed from a pNH-TrxT vector (RRID:Addgene\_26106) as a His<sub>6</sub>-tagged TEV cleavable fusion protein with Thioredoxin. Transformed BL21(DE3) cells (ROS-2, pRAR3 plasmid) were grown in LB medium and induced by 0.5 mM IPTG at 37°C for 3 h. When OD<sub>600</sub> reached a level of 0.6, the cells were harvested by centrifugation, resuspended in lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 10 mM Tris-HCl, 300 mM NaCl, 10% glycerol and 20 mM β-ME at pH 8.0), and sonicated. The supernatant containing the protein was purified under native conditions with Ni-NTA (Invitrogen), cleaved with TEV, and finally purified using reverse IMAC and gel filtration (Superdex 75, Cytiva), dialysed into NMR buffer (20 mM HEPES, pH 7.0, 100 mM NaCl, 5 mM DTT, 5% glycerol) and concentrated. For preparing the (<sup>13</sup>C, <sup>15</sup>N)-labeled MYC(1–88), cells were grown in minimal M9 media supplemented with <sup>15</sup>NH<sub>4</sub>Cl and <sup>13</sup>C D-glucose and purified as mentioned above. To the final NMR sample 1 mM TCEP, 100 μM NaN<sub>3</sub> and 10% D<sub>2</sub>O was added.

C-terminally biotinylated MYC(1–88) or MYC(1–88)ΔMB0, in which residues 16–33 were deleted, were expressed from a pET15-TrxT-MHL vector (RRID:Addgene\_26106) containing an AVI tag following the MYC(1–88) cDNA. The protein was expressed in the BirA-transformed Competent BL21 (DE3)-BirA strain (designed by SGC), which co-expresses BirA protein ligases after IPTG induction. Protein was purified as described above with an extra wash step using buffer (1 × PBS buffer with 5% glycerol and 2 mM β-ME) including 1 mM D-biotin. Mutations were introduced with the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene) and confirmed by DNA sequencing. Mutants were expressed and purified as the wild-type constructs above.

### Bio-layer interferometry interaction measurements

Bio-Layer Interferometry (BLI) assays were performed using an Octet RED384 instrument (ForteBio, USA). The following were immobilized onto SA biosensors (ForteBio, USA): (i) C-terminally biotinylated MYC(1–88) or MYC(1–88)ΔMB0, or (ii) N-terminally biotinylated peptides (Bachem Americas Inc., USA) consisting of biotinyl-glycine-lysine followed by MYC Box amino acids (aa) (MB0 aa 16–33, MBI aa 44–63, MBII aa 128–143, MBIIIa aa 188–199, MBIIIb aa 259–270, MBIV aa 304–324). These were dipped into serial dilutions of PAD (WT or mutant) in 384-well tilted-bottom microplates (ForteBio, USA). Assays were performed at room temperature in 1 × commercial HBS-EP buffer (GE Healthcare, 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% Surfactant P20) supplemented with 0.002 mg/ml BSA. *K<sub>D</sub>* values were calculated from concentration-dependent steady-state response curves, using the default values in the Octet software v10.0.

### NMR spectroscopy

NMR spectra were acquired on Bruker Avance spectrometers (operating at 600 and 800 MHz) and a Varian INOVA spectrometer (operating at 600 MHz) equipped with cryogenic probes. All spectra were processed with NMRPipe (33) and analyzed with SPARKY (34). Reconstruction

of non-uniformly sampled 3D spectra was performed with mddNMR (35). For PAD and PAD-MB0-fusion structure determinations, conventional 3D triple resonance backbone, and double resonance NOESY and TOCSY spectra were collected at 30°C as described previously (36). Proteins were buffered in 20 mM HEPES (pH 6.9, 200 mM NaCl, 2 mM DTT, 5% glycerol, 7% D<sub>2</sub>O). Assignments were performed with aid of the software FMCGUI (36). Assignments were 98.5% and 90.4% complete for backbone and sidechain resonances, respectively for PAD, and 96.4% and 90.2% complete for backbone and sidechain resonances, respectively for PAD-MB0-fusion. φ and ψ torsion angle restraints were derived from backbone chemical shifts using TALOS (37). Distance restraints were derived from cross peaks in NOESY spectra, and H-bond restraints, when applied, were for residues unambiguously determined to be in secondary structural elements based on NOE patterns, chemical shift assignments and backbone torsion angles. Automated NOE assignments and structure calculations were performed using CYANA 2.1 (38). The final 20 lowest energy structures were refined with CNSSOLVE by performing a short restrained molecular dynamics simulation in explicit solvent (29); the resulting 20 structures comprise the NMR ensemble.

Binding studies, where (<sup>13</sup>C, <sup>15</sup>N)-labeled PAD was titrated with unlabeled MYC(1–88) or biotinyl-MB0 peptide, with PAD to MYC molar ratios 1:1 and 1:2, were performed at 30°C, buffered in 20 mM HEPES (pH 6.9, 200 mM NaCl, 2 mM DTT, 5% glycerol, 7% D<sub>2</sub>O). The titrations were performed with biotinylated peptides to address solubility issues with simple peptides. CSPs were calculated with <sup>15</sup>N shifts weighted by a factor of 0.2 (for NH) and <sup>13</sup>C shifts weighted by a factor of 0.17 (for CHα). Binding studies, where (<sup>13</sup>C, <sup>15</sup>N)-labeled MYC(1–88) was titrated with unlabeled PAD, with PAD to MYC(1–88) molar ratios ranging from 0.03:1 to 1.9:1, were performed at 15°C and buffered in 20 mM HEPES (pH 7.0, 100 mM NaCl, 5 mM DTT, 1 mM TCEP, 5% glycerol, 100 μM NaN<sub>3</sub>, 10% D<sub>2</sub>O). Relative peak intensities (*I/I*<sub>0</sub>) were evaluated from HNC0 spectra as described previously (31) and were normalized to nine unaffected peaks to reduce the effect of dilution. NMR assignments for MYC(1–88) (untagged, natively expressed) used here were extrapolated from previous studies (31,39) and confirmed by conventional 3D triple resonance backbone spectra.

### Molecular docking of MB0 to the PAD

A model of the PAD-MB0 complex was constructed by docking the MB0 peptide comprising MYC residues 16–33 to PNUIS-PAD using the Rosetta FlexPepDock protocol (40). The docking was guided by pairwise ambiguous experimental constraints derived from all available chemical shift perturbations (CSPs) and intensity differences from NMR titration experiments, similarly to previous work (31). PAD CSPs were derived from <sup>1</sup>H–<sup>15</sup>N HSQC titration experiments with unlabeled biotinylated MB0, and from <sup>1</sup>H–<sup>15</sup>N HSQC and <sup>1</sup>H–<sup>13</sup>C HSQC experiments in titrations with unlabeled MYC(1–88). MB0 constraints were derived from <sup>1</sup>H–<sup>15</sup>N HNC0-derived CSPs as well as *I/I*<sub>0</sub> intensity changes when titrating unlabeled PAD into (<sup>15</sup>N, <sup>13</sup>C)-

labeled MYC(1–88) (Figure 2B). Only data from MYC residues 16–33 was used since CSPs in PAD from MYC(1–88) and biotinylated MB0 were very similar, suggesting MB0 as the interaction anchor point (Supplementary Figure S5a and b).

To efficiently steer the docking by per-residue ambiguous constraints representing all experimental data, we derived a combined chemical shift perturbation for each residue,  $CSP_i$ , and then merged this with intensity data where applicable ( $I/I_0$ ). In short, CSPs for assigned  $^1H$  (only amide),  $^{15}N$  and  $^{13}C$  nuclei in each residue were averaged by atom type and then combined to a single per-residue chemical shift perturbation,  $CSP_i$ , using a weighted average based on the magnetogyric ratio of the nucleus; 1.000, 0.102 and 0.251 for  $^1H$ ,  $^{15}N$  and  $^{13}C$ , respectively (41). The  $CSP_i$  standard deviation  $\sigma$  from 0 was calculated iteratively by omitting  $CSP_i$ s larger than  $3\sigma$  until  $\sigma$  converged (41).  $CSP_i$ s larger than  $2\sigma$  from 0 were considered significant indicators of binding and were included as constraints for the modeling after normalisation to a maximum of 1.0. For MB0,  $\sigma$  for the relative intensity change ( $I/I_0$ ) was obtained following a similar process, but using  $\sigma$  from 1 instead. Relative intensities larger than 1.0 were set to 1.0, and resulting signals were inversely scaled from 1.0 to the lowest observed value. The significant  $CSP_i$ s were averaged with the significant signals from the peak intensity (where applicable), in order to obtain residue-specific indicators of binding for any change detected by  $^1H$ ,  $^{15}N$  or  $^{13}C$  CSP or peak intensity.

Pairwise ambiguous constraints were constructed between every pair of PAD and MYC residues with significant indicators of binding, with a magnitude determined by the sum of the corresponding indicators. For PAD, only residues that had  $>5\%$  relative surface exposure as measured in comparison to an extended Gly-X-Gly chain were considered. The constraints were imposed during the docking using a square well scoring function that positively favour the atom pair when within 5.0 Å, and linearly decreasing the constraint influence in the distance range 5.0 Å to 10.0 Å, while not penalizing distances further than 10.0 Å. This corresponds to the FADE constraint function type in Rosetta with a lower bound -5.0, upper bound 10.0, and cubic splines of width 5.0. The overall constraint weight was set to contribute as much to the final scoring of the complex as all the unconstrained energy functions.

The PAD structure was energy-minimized using the Rosetta relax protocol. 50 000 decoys of PAD:MB0 interactions were generated using Rosetta Flex-PepDock ab-initio protocol guided by the constraints. Each decoy was generated starting from an extended MB0 peptide superpositioned on one of the constraint-pairs in a random orientation on the receptor surface. Rosetta FlexPepDock utilizes Monte-Carlo sampling to minimize the REF2015 energy function (42) in addition to any constraints, while sampling the translational, rotational, and torsional degrees of freedom of both MB0 and PAD. The translational and rotational degrees of freedom are sampled by randomly moving and rotating MB0 with respect to PAD. The backbone torsions of MB0 are sampled using fragment insertions, where the fragments are 3-, 5- or 9-residue backbone dihedrals with local sequence similarity to the MB0. After each structural change, the side chains are rebuilt using Dunbrack's

backbone-dependent rotamer library (43) followed by energy minimization before the trial energy is calculated. Following the standard Metropolis acceptance criterion, the structural change is accepted with a probability related to the difference in energy before and after the change.

To properly represent the plausible conformations of the MB0 peptide, a subset of representative decoys which best describe the observed CSP were selected. First, the 25 000 models (median filter) with best resulting energy scores were clustered by backbone position, with cluster radii ranging from 1.0 to 4.0 Å in increments of 0.25, ignoring the flexible loops on the receptor when superpositioning and calculating RMSD. Then, each model in a cluster was described as a binary vector denoting for each residue the involvement in inter-chain interaction, defined as an inter-chain distance between any pair of non-hydrogen atoms  $\leq 4.5$  Å. Each cluster was represented by the sum of all member vectors. The Lasso algorithm as implemented in scikit-learn (44), a linear model with L1 regularizer, was used to find the combination of clusters that had the best square-sum fit to the experimentally derived constraints. This was repeated for all cluster radii from above and the cluster radius 1.75 Å obtained the best fit to the constraints, with 9 clusters together describing the constraints with  $R^2$  of 0.71 (correlation  $R$  of 0.84).

All protein structure representations were made using PyMOL (<http://www.pymol.org/pymol>). The electrostatic surface analysis was generated by APBS electrostatics in PyMOL, with the potential range set from -6 to +6, all other parameters set as default. Multiple sequence alignment for the ConSurf analysis was obtained through two iterations of HHblits with a final  $E$ -value cutoff of 0.001, run on Uniref30 fetched June 2021 (45,46).

## Cell lines

Cell lines were grown as previously described (47). To evaluate the effect of MYC point mutations on the interaction with endogenous PNUTS, MCF10A cells were transduced with pLenti neo CMV V5-MYC or the respective mutants (P21A, Y22A, F23A, Y24A) and selected in 750  $\mu\text{g/ml}$  Neomycin (Gibco). To test the interaction of PNUTS(1–160) wild type or point mutants with MYC in cells, the MCF10A cell line was sequentially transduced with: pLenti neo CMV V5-MYC; pLenti blast CMV rtTA3 and pLenti hygro CMV/tight. The latter plasmid was also encoding a Flag-MCS-3xNLS cassette which contained either no insert (EV) or the PNUTS(1-160) wild-type or point mutant (A114K, M141W) coding sequence. After each transduction, cells were selected with the respective antibiotic (Neomycin/Geneticin: 750  $\mu\text{g/ml}$  (Gibco), Blasticidin: 5  $\mu\text{g/ml}$  (BioShop), Hygromycin B: 100  $\mu\text{g/ml}$  (BioShop)). All cell lines were validated by STR profiling and tested regularly for mycoplasma using the MycoAlert Kit (Lonza) and remained negative throughout.

## Protein extraction and immunoblotting

The protein expression of MYC and PNUTS(1–160) wild type and point mutants was assessed after selection and for each replicate of PLA. To this end, cells were seeded in

culture media and expression was induced with 0.5  $\mu\text{g/ml}$  doxycycline (Sigma-Aldrich) for 24 h prior to protein harvest, where applicable. For PNUTS(1–160) constructs, cells were treated with 10  $\mu\text{M}$  MG132 (Calbiochem) 4 h prior to harvest. Cells were washed twice with 1 $\times$  PBS (Wisent) and lysed on the plate using SDS lysis buffer (160 mM Tris–HCl pH 6.8, 2% SDS). Protein lysate was quantified using the Pierce 660 nm kit with added Ionic Detergent Compatibility Reagent (Pierce) and 10–15  $\mu\text{g}$  of total cell lysate were run on a 10% SDS-PAGE. Proteins were wet transferred onto 0.2  $\mu\text{m}$  Nitrocellulose membranes (Perkin Elmer) at 100 V for 1 h in chilled transfer buffer. For expression testing, membranes were blocked with 5% skim milk in 0.1% PBS-T and incubated overnight with mouse anti-Flag (B3111, Sigma-Aldrich, 1:500), mouse anti-V5 (ab27671, Abcam, 1:1000), or rabbit anti-Actin (Sigma-Aldrich, 1:3000) in 5% skim milk in PBS-T. To detect phosphorylated MYC species, membranes were blocked with 5% BSA in 0.1% TBS-T and incubated overnight with rabbit anti-MYC pT58 (Y011034, Applied Biological Materials, 1:1000) or mouse anti-MYC pS62 (ab78318, Abcam, 1:1000) in 5% BSA in TBS-T. Detection was performed using fluorescently labeled secondary antibodies against rabbit and mouse (LI-COR) on the LI-COR Odyssey imaging system.

### Proximity ligation assay

The expression of the PNUTS constructs was induced through the addition of 0.5  $\mu\text{g/ml}$  doxycycline (Sigma-Aldrich) for 24 h and cells were treated with 10  $\mu\text{M}$  MG132 (Calbiochem) 4 h prior to fixation. For all PLAs with PNUTS(1–160) and its mutant, cells were fixed 10 min in 4% paraformaldehyde (Sigma-Aldrich) in PBS at pH 7.2. PLAs for MYC and MYC point mutants were fixed with 2% paraformaldehyde (Sigma-Aldrich) in PBS at pH 7.2 for 20 min. Rabbit anti-MYC (Millipore, 06-340), mouse anti-Flag (Sigma, B3111), rabbit anti-V5 (Millipore, AB3792) and mouse anti-PNUTS (BD Bioscience, 611060) antibodies were optimized and used at a 1:200 (MYC, Flag), 1:500 (V5) or 1:50 (PNUTS) dilution, respectively. The PLA was performed as previously described (22). The fluorescent foci were quantified using the Blobfinder software (Carolina Wählby & Amin Allalou, CBA, Uppsala University), with a minimum of 200 nuclei per biological replicate quantified.

The data was plotted using R (version 4.0.2) using the function *ggviolin* from the package *ggpubr* (version 0.4.0). Statistical analysis was performed using the *ks.test* command in R.

## RESULTS

### MB0 interacts directly with PAD

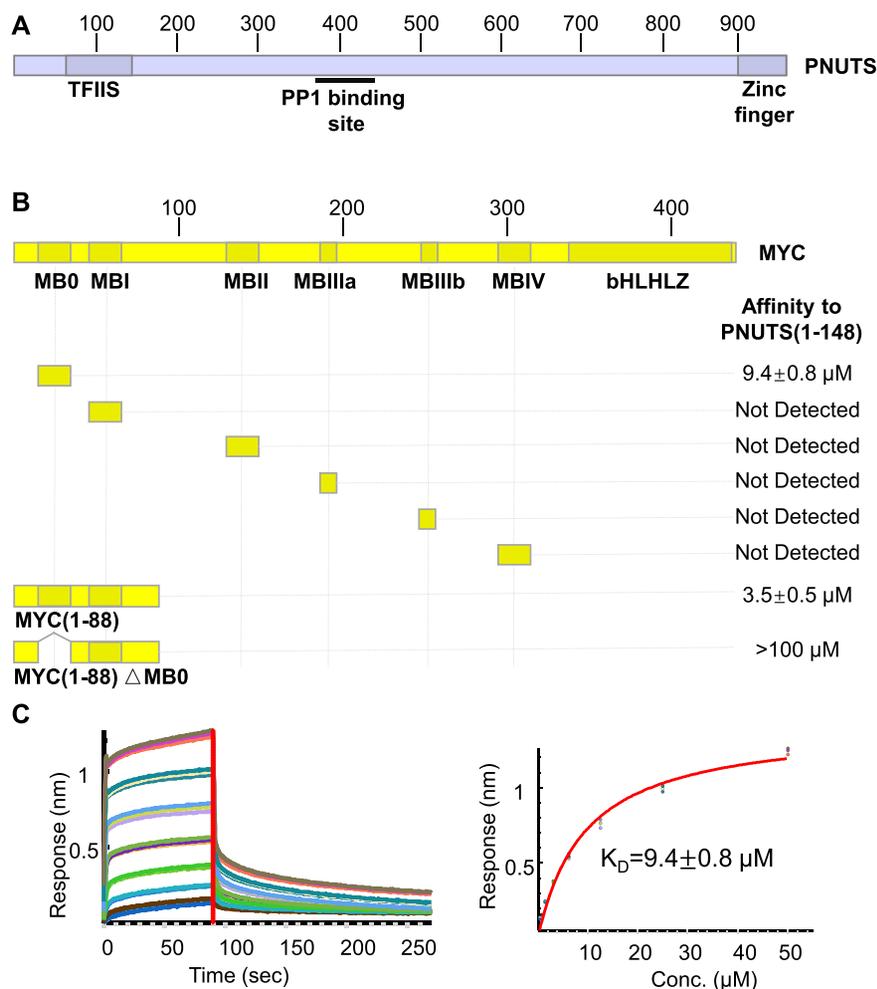
Our previous work identified MYC as a substrate of the PP1:PNUTS phosphatase complex, however the molecular basis of the interaction and whether it is direct remained unclear (27). As MYC does not contain a canonical PP1 recognition motif (RVxF), we hypothesized that MYC may interact with the non-catalytic, substrate-specifying subunit PNUTS. PNUTS contains two previously annotated domains: a TFIIS helical bundle-like domain (residues 73–147)

of the Med26 Pfam family (PF08711) with no functional annotation (48); and a zinc finger domain at the C-terminus (49,50) (Figure 1A). Since the former is a putative protein–protein interaction domain, we first interrogated the interaction potential of the N-terminal region by designing and testing several expression constructs of proteins within the first 186 amino acids of PNUTS. The shortest of these was successfully expressed and purified as a stable protein, encompassing residues 1–148 (PNUTS(1–148)). This suggests this region represents an independently folded functional domain, which was then further evaluated in binding assays with MYC peptides and protein.

To determine whether PNUTS(1–148) interacted directly with MYC, we initially focused our analysis on the six highly conserved and regulatory MYC Box (MB) regions. BLI was used to evaluate the binding of each MB (MB0, MBI, MBII, MBIIIa, MBIIIb, MBIV) peptide to PNUTS(1–148). Briefly, biotinylated MB peptides were immobilized onto streptavidin-conjugated biosensors and subsequently suspended in a serially diluted PNUTS(1–148) protein solution. The BLI signals were then used to calculate apparent binding affinities. MB0 interacted with PNUTS(1–148) with a  $K_D$  of  $9.4 \pm 0.8 \mu\text{M}$ , whereas the other MYC boxes failed to produce detectable signals (Figure 1B; Supplementary Figure S1a–e), suggesting that MB0 constitutes the primary interaction site of MYC with PNUTS(1–148). We next performed reciprocal BLI assays to test the interaction of recombinant biotinylated MYC N-terminal protein, comprising residues 1–88 (MYC(1–88)) that contains MB0 and MBI, and its deletion mutant lacking MB0 (MYC(1–88 $\Delta$ MB0)). MYC(1–88) directly interacted with PNUTS(1–148) with a  $K_D$  of  $3.5 \pm 0.5 \mu\text{M}$ , while MYC(1–88 $\Delta$ MB0) demonstrated a weak but still measurable interaction with PNUTS(1–148) ( $K_D > 100 \mu\text{M}$ ) (Figure 1B; Supplementary Figure S1f and g). These data suggest that while regions outside of MB0 can contribute to binding, MB0 is the dominant PNUTS anchor site. Having established the direct interaction of MB0 with PNUTS(1–148), and since the folding of this entity requires sequence components beyond the PNUTS TFIIS-like sequence, we termed this newly identified region the PNUTS amino-terminal domain (PAD).

### Specific residues within MB0 interact with the PAD

To identify MB0 residues critically involved in binding to PAD, we used NMR spectroscopy. We have previously shown that MB0 and MBI are transiently ordered regions, in an otherwise disordered His<sub>6</sub>-MYC(1–88) fragment, which was refolded from inclusion bodies (39). For this study, we expressed and purified MYC(1–88) under native conditions from a cleavable TRX-tagged construct, and found the NMR chemical shifts of (<sup>13</sup>C, <sup>15</sup>N)-labeled MYC(1–88) near-identical to those of His<sub>6</sub>-MYC(1–88). NMR titrations were then performed with unlabeled PAD at 15°C. As MYC(1–88) is predominantly intrinsically disordered, most of its amide resonances in the <sup>1</sup>H–<sup>15</sup>N HSQC spectra overlap heavily and are clustered between 7.7 and 8.6 ppm (39,51). Nevertheless, an <sup>1</sup>H–<sup>15</sup>N HSQC overlay at increasing PAD:MYC(1–88) ratios shows clear shifting and/or broadening of V19, Q20 and F23 peaks (Figure 2A).



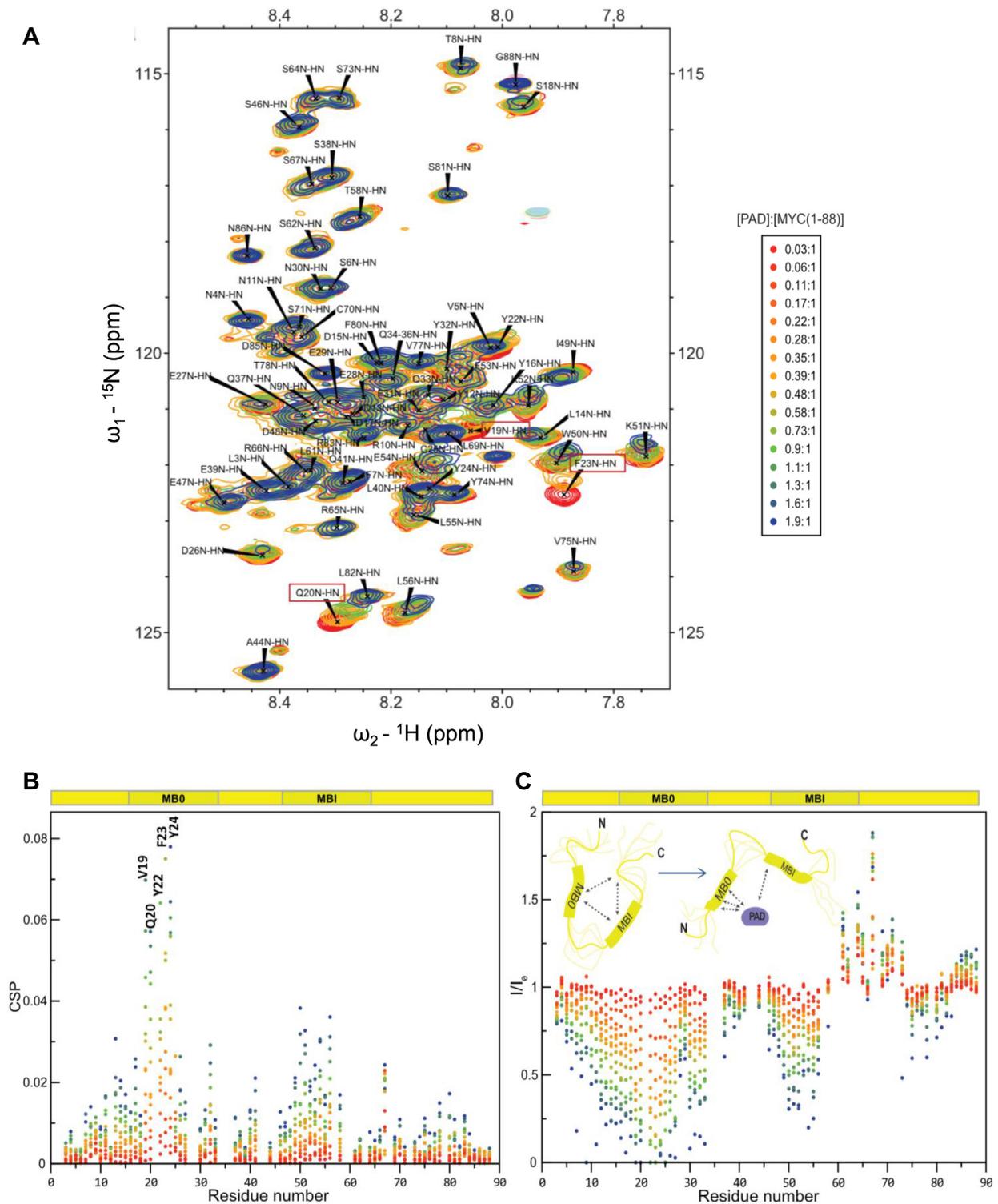
**Figure 1.** MB0 interacts directly with PAD. (A) Domain organization of PNUTS. Two previously annotated domains (TFIIS and Zinc finger) are highlighted with blue boxes and the PP1 binding region is indicated. (B) Domain organization of MYC with homology boxes (MB0, MBI, MBII, MBIIIa, MBIIIb, MBIV) as well as the basic helix-loop-helix leucine zipper domain (bHLHLZ) highlighted in dark yellow (top). The different MYC constructs that were tested by BLI for binding to PNUTS(1-148) are displayed below with the measured affinities denoted on the right. (C) Representative BLI sensorgrams (left) and steady-state curve fit (right) for MB0 binding to PNUTS(1-148). See also Supplementary Figure S1.

To more accurately determine chemical shift perturbations (CSPs) and broadening of MYC amide resonances due to PAD binding, it was necessary to acquire conventional 3D backbone spectra (HNCO and HNCA) to resolve the amide signals based on their coupling to neighbouring carbonyl groups or C $\alpha$  carbons. When MYC(1-88) was titrated with PAD, substantial CSPs and broadening of MYC resonances were observed (Figure 2B and C). A six-residue stretch of MB0 (V19 to Y24) exhibited the most substantial CSPs and broadening effects (Figure 2B), which is in agreement with MB0 being critical for PAD binding by BLI (Figure 1B). All assigned MYC residues between D13 and Q33 lose at least 75% of their peak intensity at a 1.9:1 PAD to MYC ratio with no signs of signal recovery in the bound state (Figure 2C, Supplementary Figure S2). This indicates continued chemical exchange between multiple bound states (31,52,53), which is characteristic for intrinsically disordered proteins in so-called fuzzy complexes (54). Consistent with our BLI data, other parts of MYC(1-88) also show resonance shifts and broadening, including

a sequence in MBI (residues W50-L55) with smaller CSPs and intensity changes compared to those of MB0. Interestingly, a region within and adjacent to MBI (L61 to S73) showed sharper amide signals in response to PAD binding, with S67 amide intensity increased 2-fold (Figure 2C, Supplementary Figure S2). This suggests that PAD interaction with MB0 leads to greater conformational mobility of this segment and more rapid interconversion between states than in free MYC(1-88), and would be in agreement with PAD-induced release of previously identified transient interactions between MB0 and MBI regions (39) (Figure 2C, inset).

#### PAD adopts a helical bundle with an Armadillo subdomain

To understand how PNUTS recognizes and binds to MYC, we first determined an NMR-derived solution structure of PAD (PDB ID: 6VTI, Table 1; Supplementary Figure S3). The topology of the overall structure consists of nine  $\alpha$ -helices ( $\alpha$ 1, aa8-18;  $\alpha$ 2, aa28-39;  $\alpha$ 3, aa44-56;  $\alpha$ 4, aa60-69;



**Figure 2.** Specific residues within MB0 interact with PAD. (A)  ${}^1\text{H}$ - ${}^{15}\text{N}$  HSQC spectral overlay of MYC(1–88) in the presence of increasing concentrations of PAD (left). The colour legend for panels A–C is displayed (right). (B) Composite CSPs of MYC amide resonances from HNCOSY spectra upon addition of PAD at PAD-to-MYC molar ratios ranging from 0.06 to 1.9. Residues with the highest detectable CSPs are labeled, with MB0 and MBI regions in MYC(1–88) highlighted by dark yellow boxes above the graph. (C) Changes in intensity ratios ( $I/I_0$ ) of MYC(1–88) amide resonances from the same HNCOSY spectra as in (B), and with MB0 and MBI regions similarly highlighted above the graph. Cartoon inset illustrates how increased  $I/I_0$  for residues in and adjacent to MBI, reflect more rapid interconversion between states in this region than in free MYC(1–88), agreeing with PAD-induced release of transient interactions between MB0 and MBI regions present in the unbound state. See also Supplementary Figure S2.

$\alpha$ 5, aa71–85;  $\alpha$ 6, aa88–100;  $\alpha$ 7, aa105–111;  $\alpha$ 8, aa113–123;  $\alpha$ 9, aa127–147; Figure 3A and B) that form a series of helical bundles. These can be grouped into two subdomains: (i) an N-terminal domain (NTD), consisting of  $\alpha$ 1–3 and (ii) an Armadillo (ARM) subdomain (aa60–148). This ARM subdomain is composed of two consecutive ARM repeats (ARM1 and ARM2), made up of  $\alpha$ 4–6 and  $\alpha$ 7–9, respectively (Figure 3B). ARM repeats, commonly associated with protein interactions (Pfam family PF00514), are made up of three  $\alpha$ -helices including a short  $\alpha$ -helix (e.g.  $\alpha$ 7 of ARM2) and two longer  $\alpha$ -helices ( $\alpha$ 8 and  $\alpha$ 9 of ARM2) (Figure 3B). A search of protein domains with similar structures to that of PNUTS-PAD using TMalign (55) yielded nearly 5000 proteins with similar helical bundle arrangements. Many of these contain a TFIIS helical bundle-like domain (Pfam family PF08711) with the closest human structural homologue being a region of transcription factor hMED26 (backbone RMSD of 2.6 Å). The PAD also structurally aligns well to the LEDGF integrase binding domain (backbone RMSD of 2.8 Å) and transcription factor IWS1 (backbone RMSD of 2.8 Å), both of which have previously been shown to bind intrinsically disordered binding motifs in a highly specific manner (56,57) (Supplementary Figure S4).

### MB0 interacts with the C-terminal facing surface of PAD ARM2 motif

To map the region of PAD that interacts with MB0, we performed NMR binding studies at 30°C using (<sup>13</sup>C, <sup>15</sup>N)-labeled PAD titrated with unlabeled biotinyl-MB0 peptide and MYC(1–88) (Supplementary Figure S5a and b). At a MB0 to PAD molar ratio of 2 to 1, there was a significant change in PAD amide resonances in <sup>1</sup>H–<sup>15</sup>N HSQC spectra with an average CSP of 0.012 ppm and a STDEV of 0.018. A highly similar perturbation pattern was observed, when MYC(1–88) was added to PAD (Supplementary Figure S5b), confirming that MB0 within MYC(1–88) is critical for the PNUTS-MYC interaction, consistent with our BLI data (Figure 1B and C) and NMR titrations of PAD with (<sup>13</sup>C, <sup>15</sup>N)-labeled MYC(1–88) (Figure 2). To visualise the MYC-MB0 imprint on PAD in titrations with MB0 and MYC(1–88), we analysed combined per-residue CSPs (CSP<sub>i</sub>s), including data from both <sup>1</sup>H–<sup>13</sup>C and <sup>1</sup>H–<sup>15</sup>N HSQC experiments (Figure 3C, see Materials and Methods). Eight PAD residues (K109, Q110, N111, A114, K115, Q119, M141 and Q147) had CSP<sub>i</sub>s  $\geq$  3 standard deviations from 0, and an additional 14 had values  $\geq$  2 standard deviations from 0, of which 7 have  $>$ 5% of their surfaces exposed according to the NMR-derived solution structure (V45, R58, T104, K118, W140, I144 and S146) (Figure 3C). A map of CSP<sub>i</sub>s on the PAD surface in response to MB0 interaction indicates that the highly perturbed residues form part of, or are close to the C-terminal facing surface of the ARM2 motif (Figure 3D, Supplementary Figure S5c). The methyl groups of A114 ( $\gamma$ ), I144 ( $\delta$ 1 and  $\gamma$ 2), T113 ( $\gamma$ 2) and M141 ( $\epsilon$ )—all in or near the C-terminal facing surface—were noticeably perturbed upon MB0 peptide binding suggesting hydrophobic contributions to binding (Supplementary Figure S5d). Lysines (K109, K115, K118) in this patch are also affected by MYC binding, and contribute positive charge to the MYC-binding patch (Figure 3E). Taken together, this shows that MYC interacts with PNUTS through a well-defined

**Table 1.** NMR restraints, structural statistics and quality scores of NMR-derived structures of PAD and PAD-MB0-fusion

	PAD	PAD-MB0-fusion
PDBID	6VTI	7LQT
BMRBID	30 722	30 861
<b>NMR distance &amp; dihedral restraints:</b>		
<i>Distance Restraints:</i>		
All	3365	3198
Intraresidue	843	912
Sequential ( i-j  = 1)	842	898
Medium range (2 ≤  i-j  ≤ 4)	888	694
Long range ( i-j  > 4)	792	694
Hydrogen Bonds	24 × 2	24 × 2
<i>Dihedral angle restraints:</i>		
All	224	224
$\phi$	112	112
$\psi$	112	112
<b>Structure Statistics:</b>		
<i>Number of violations in the NMR ensemble<sup>a</sup>:</i>		
Distance restraints (>0.5 Å)	0	0
Dihedral angle restraints (>5°)	2	3
<i>r.m.s.d. from experimental restraints:</i>		
Distance (Å)	0.016 ± 0.002	0.015 ± 0.001
Dihedral angle (°)	0.466 ± 0.071	0.482 ± 0.075
<i>r.m.s.d. from idealized covalent geometry:</i>		
bond (Å)	0.0143 ± 0.0002	0.0145 ± 0.0002
bond angles (°)	0.909 ± 0.016	0.923 ± 0.016
impropers (°)	1.74 ± 0.08	1.74 ± 0.11
<i>Average r.m.s.d. (Å):</i>		
Ordered regions <sup>b</sup>		
Backbone atoms	0.63 ± 0.09	0.69 ± 0.10
All heavy atoms	1.04 ± 0.10	1.13 ± 0.09
<b>Structural Quality Scores:</b>		
<i>Ramachandran plot (%)<sup>c</sup></i>		
Residues in most favoured regions	93.7	92.6
Residues in additional allowed regions	6.0	7.4
Residues in generously allowed regions	0.2	0.0
Residues in disallowed regions	0.0	0.0
<i>Global quality scores<sup>d</sup></i>		
Procheck (phi-psi)	0.22 1.18	0.14 0.87
Procheck (all)	-0.04 -0.24	-0.14 -0.83
MolProbity clash	6.04 0.49	15.3 -1.10

<sup>a</sup>NMR ensembles consist of the 20 lowest energy structures out of 100 calculated.

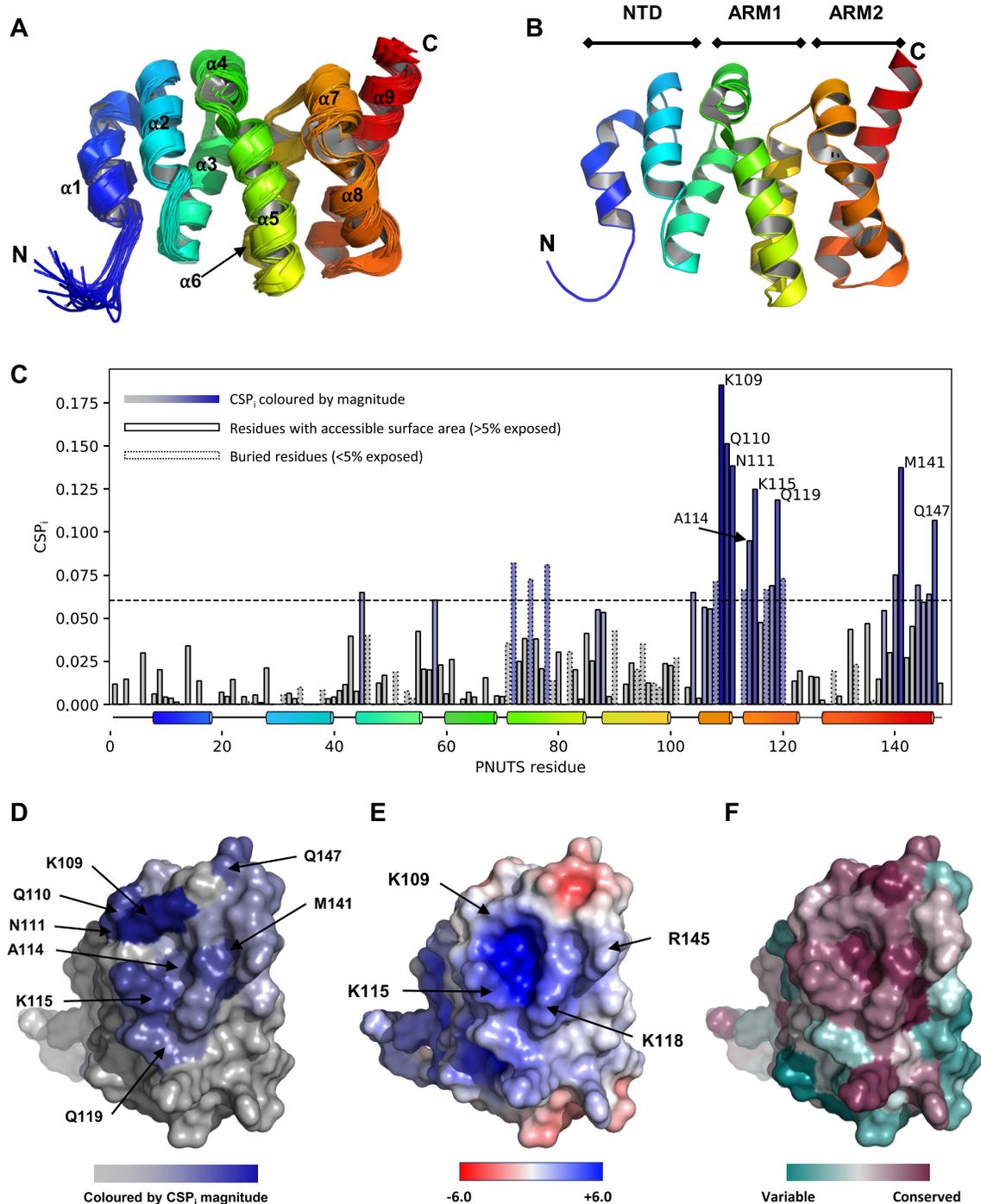
<sup>b</sup>Ordered regions, as reported by the PSVS server (77) for PAD are 5–22, 26–40 and 44–147; for PAD-MB0-fusion are 5–40, 43–120, 124–147 and 162–166.

<sup>c,d</sup>Calculated using the PSVS server - <http://psvs.nesg.org/> (77).

MYC-binding patch, with the potential to include both hydrophobic and electrostatic interactions. Importantly, a ConSurf analysis (58,59) indicates that this MYC-binding patch is conserved in PNUTS across species (Figure 3F), suggesting a potential conserved interaction surface.

### NMR guided model of PAD and MB0 interaction

To analyze structural features of the PAD-MB0 interaction, a computational modelling approach was taken to identify possible interaction modes of MYC with PNUTS,



**Figure 3.** MB0 interacts with the C-terminal facing surface of PAD ARM2 motif. (A) Backbone trace of the NMR ensemble of PAD (20 structures), as seen from the ‘front’ view. Helices are numbered with respect to their order from the N- to C-terminus coloured from blue to red. (B) Cartoon representation of the lowest energy PAD structure, which comprises an N-terminal terminal domain (NTD) consisting of a 3-helix bundle (helices 1 to 3), followed by two Armadillo (ARM) repeats (ARM1 and ARM2, helices 4 to 6 and 7 to 9, respectively). (C) Combined per-residue PAD CSPs ( $CSP_i$ s) upon complex formation with MB0, averaged between NMR titrations with MB0 peptide and MYC(1–88) protein (see Methods). Residue bars are coloured from grey to blue, with a more saturated blue hue for greater  $CSP_i$ , and with dashed contours for buried residues (<5% exposed). The horizontal dashed line indicates 2  $CSP_i$  standard deviations from 0. (D) Surface representation of apo-state PAD as seen from the view of a 70° rotation of panel (A) on the Y axis, indicating MYC-affected residues by colours and labels as in corresponding (C). (E) Representation of the PAD electrostatic surface potential using a colour gradient spanning red ( $kT/e = -6$ ) to blue ( $kT/e = 6$ ), in the same view as (D). (F) Analysis of PAD sequence conservation between species with a structure colour-coding bar spanning turquoise (for residues that are not conserved) to maroon (highly conserved residues), in the same view as (D). Conservation values were derived through ConSurf based on a multiple sequence alignment generated by HHblits (see Materials and Methods). See also Supplementary Figure S5.

guided by observed chemical shift perturbations and intensity changes. In short, significant CSPs and intensity changes for PAD and MB0 residues were used to construct combined modeling constraints (see Materials and Methods). A larger difference upon binding resulted in a proportionally larger contribution to the constraints, thereby focusing the modeling on key binding residues. We generated 50 000 models of the PAD-MB0 complex using the Rosetta FlexPepDock *ab-initio* protocol (40), with NMR-derived constraints accounting for circa 50% of the inter-chain interaction energy. These models were clustered with respect to the position of residues 19–26 within MB0 and, using a LASSO algorithm (60), the combination of the fewest possible clusters which together best satisfy the experimental constraints was determined. Nine clusters are required to describe 71% of the variance evident in the experimental data (Supplementary Figure S6a), with one dominant cluster (Cluster 0) representing 28.5% of all models (Supplementary Figure S6b and c).

The resulting NMR-driven ensemble of PAD-MB0 complexes (Figure 4A) has features of a ‘fuzzy complex’ (54) where MB0 adopts multiple conformations in the bound state, in full agreement with the nature of the NMR intensity changes (Figure 2). Analysis of the orientation of MB0 in the 25 000 models with the lowest energy shows a clear preference of direction of the peptide with the experimental-driven constraints compared to an analogous ensemble generated without constraints (Figure 4A; Supplementary Figure S6d). Thus, despite the fuzzy nature of the ensemble including several clusters of opposite orientation, there is a clear preference for MB0 binding along the conserved patch of PAD from the center of  $\alpha 9$  toward the center and C-terminus of  $\alpha 7$  of PAD.

### The solution structure of a PAD-MB0-fusion protein

To understand in greater detail how the PAD recognizes MB0, we engineered a fusion construct that could capture a predominant binding mode at the PAD C-terminus as indicated by the NMR-guided modeling (Figure 4A). This PAD-MB0-fusion protein (PAD-MB0-fusion) contains residues 1–148 of PNUTS and amino acids 13–30 of MYC connected by a flexible (GGGS)<sub>2</sub> linker. The PAD-MB0-fusion displays high quality, well dispersed NMR spectra (Supplementary Figure S7a) enabling the full assignment of backbone and side-chain resonances. In comparing <sup>1</sup>H–<sup>15</sup>N HSQC spectra of apo PAD vs. PAD-MB0-fusion, the CSP patterns for PAD amides are similar to those observed when apo PAD is titrated with MB0 peptide (Supplementary Figure S7b and c). For instance, several resonances in the ARM2 MYC-binding patch that shift upon PAD binding to MB0 (e.g. M141, V117 and A114) follow a similar trajectory to their position in the spectrum of PAD-MB0-fusion; however the magnitude of the associated CSPs are significantly higher (Supplementary Figures S6e and 7b). Differences in the perturbation patterns seen with PAD-MB0-fusion compared to those in the MB0 peptide binding assay are mainly localized to the PAD C-terminus (Figure 3C; Supplementary Figures S6e, 7b and c). The fusion linker displays a random behaviour, and some flexibility is also retained in the MB0 segment (Supplemen-

tary Figure S7d). Taken together, these data suggest that the fusion construct favours MB0-bound states that are also populated in the non-fused system, likely by increasing the local concentration of MB0 at the MYC-binding patch of PNUTS.

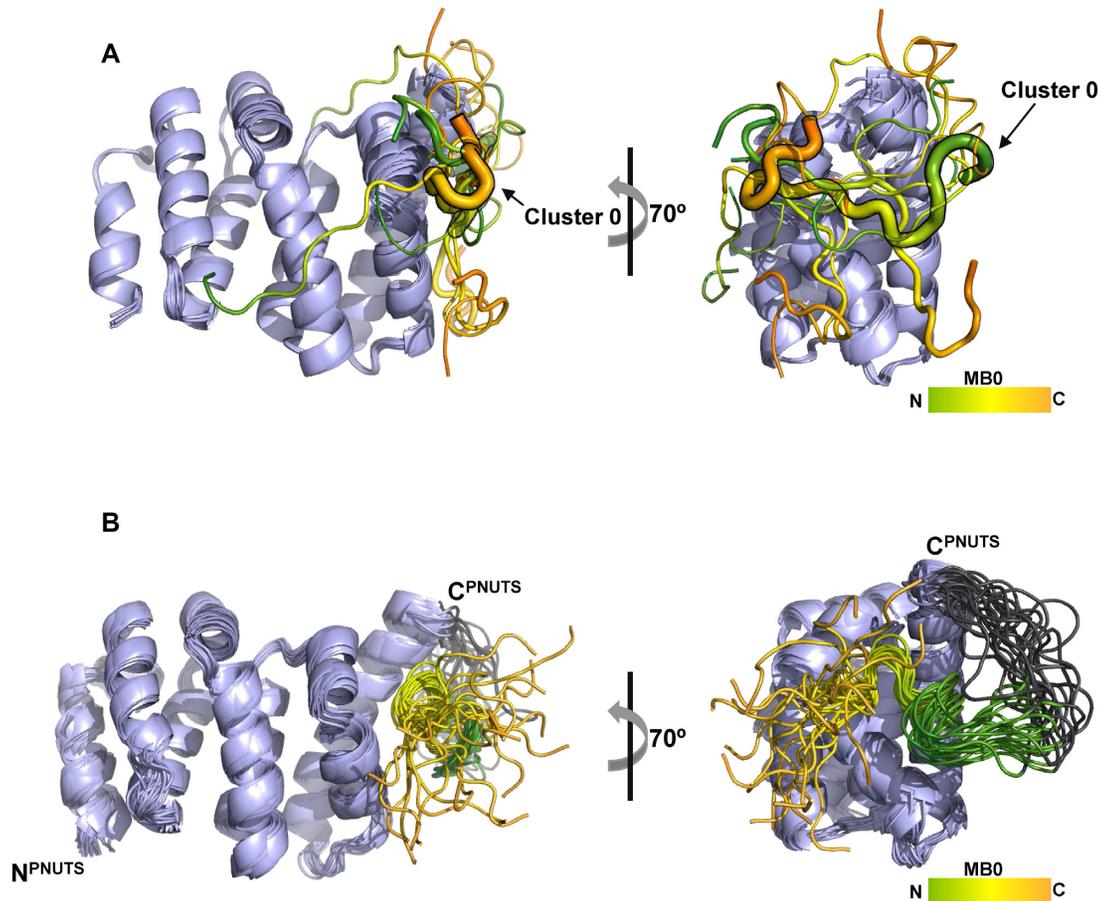
Having validated the PNUTS-MYC interaction within the fusion protein, we next performed a full, NOE-driven structure determination, including 43 experimental NOE restraints between the PAD and MB0 residues 13–30 (PDB ID: 7LQT, Figure 5A, B; Table 1). In this structure, MB0 is positioned across the conserved surface of ARM2 comprising helices  $\alpha 7$ ,  $\alpha 8$  and  $\alpha 9$  (Figures 4B and 5B) in good agreement with the NMR titration data (Figures 2B, C, 3C, D). There are small differences between the structure of apo and PAD-MB0-fusion which superimpose with a backbone RMSD of  $1.9 \pm 0.2 \text{ \AA}$  (Supplementary Figure S7e). The most important difference is confined to a small repositioning of the C-terminal helix ( $\alpha 9$ ) in the ARM2 motif, which we discuss in detail below.

There is excellent agreement between the solution structure of the PAD-MB0-fusion and the NMR-guided computational modeling of the interaction (Figure 4A, B; Supplementary Figure S6b and e). The position and direction of MB0 in all 20 structures of the fusion protein ensemble are consistent with the preferred binding ‘Cluster 0’ of the NMR-guided computational model (Supplementary Figure S6a and d). Indeed, the inter-chain contact frequency of residues over the total number of computational models are also similar to the fusion protein, especially the largest cluster of docked models with a correlation *R* of 0.82 (Supplementary Figure S6c).

### The specific interactions between PAD and MB0

Distinct hydrophobic and polar interactions are formed between the core motif of MB0-V<sub>19</sub>QPYFY<sub>24</sub>, and the conserved MYC-binding patch of PAD ARM2 (Figure 5A and B). This interaction is defined by NOEs from four MB0 residues (V19, P21, Y22, F23) and seven PNUTS residues (V105, L108, K109, A114, V137, M141, I144) (Figure 5B and C). This network of NOEs delineates a clear mode of interaction between PNUTS and MYC that is very similar to that of ‘Cluster 0’ in the NMR-guided computational model (Figure 4A). Notably, MYC(F23) sits in the MYC-binding patch formed by L108, A114, V137, W140, M141, I144 and the acyl chain of K118 at the bottom of the C-terminal facing surface (Figure 5C). A hydrogen bond between PNUTS(K109) and MYC(Y22), supported by NOEs, suggests additional binding specificity (Figure 5B and D). Two hydrogen bonds within MB0 are also supported by NOEs between residues MYC(F23)-MYC(Y24) and MYC(F23)-MYC(Y22), forming a  $\beta$ -turn motif (Figure 5D). The N- and C-termini of MB0 show no NOEs to any PAD residues (Figure 5B), and are more disordered than MYC-V<sub>19</sub>QPYFY<sub>24</sub> in all 20 structures of the ensemble (Supplementary Figure S7e and f).

Comparison of the MB0 interaction site of PAD in the absence (apo) and presence of MB0 reveals a subtle conformational adjustment of the three helices in the ARM2 motif upon MB0 binding in the fusion protein. This helix movement was also required in the NMR-guided computa-



**Figure 4.** NMR guided model of PAD and MB0 interaction. (A) CSP-driven computational models of PAD binding to MB0 based on CSPs of MB0 and PAD obtained from the experiments depicted in Figure 2B and C and Figure 3C. The geometric cluster centre of each of the nine major MB0 conformation clusters is shown. Sizes of the peptides are proportional to what percentage of states belong to the cluster represented by the peptide. The dominant cluster 0 is marked with a bold outline and highlighted by an arrow. The PAD is coloured light blue and MB0 is coloured green to orange from N-termini to C-termini. (B) Experimental backbone trace of the NMR structure of the PAD-MB0-fusion protein (20 structures) with PAD, linker, and MYC<sub>13–30</sub> coloured light blue, grey, and green to orange, respectively. See also Supplementary Figure S6.

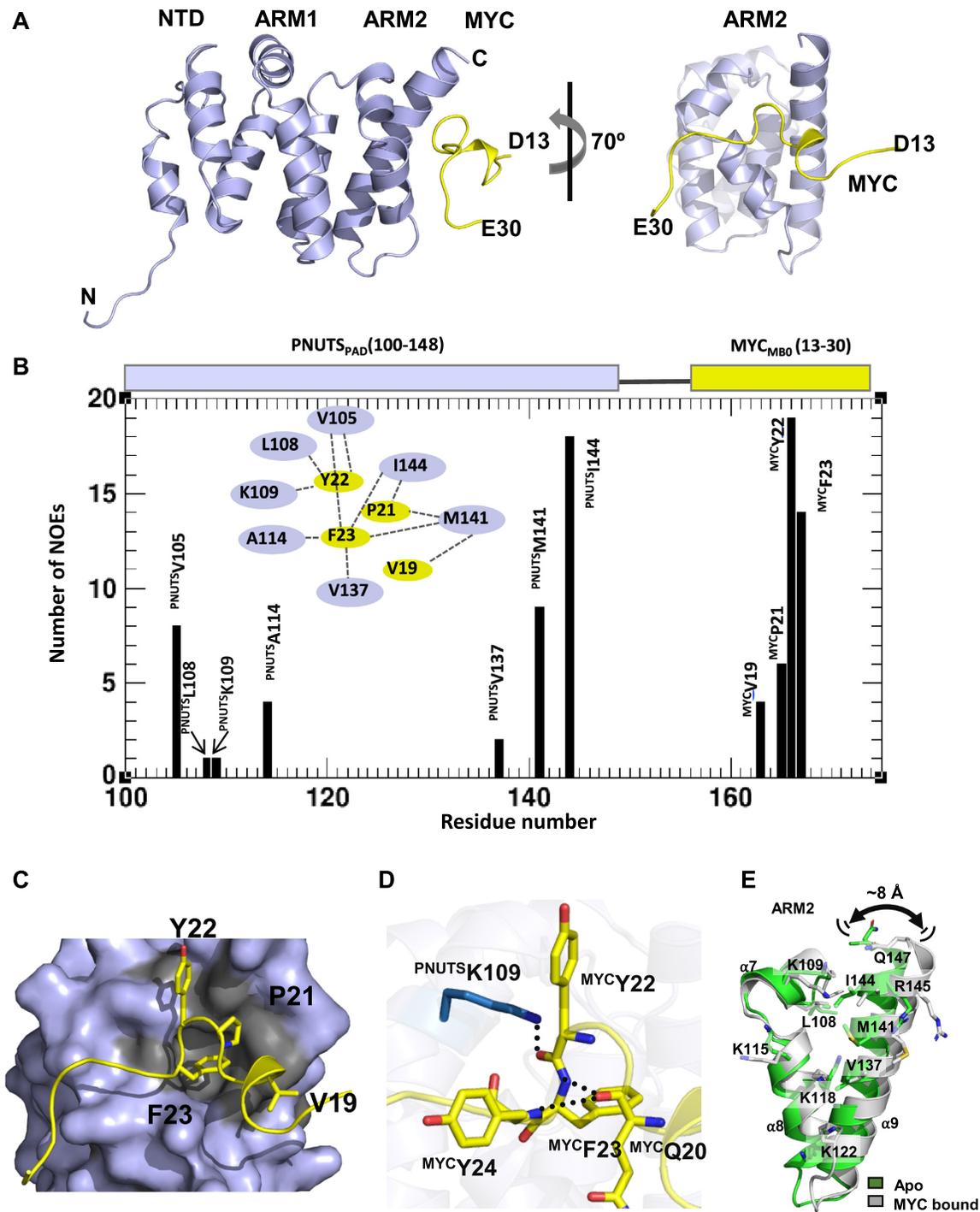
tional models for MB0 binding to PAD, further supporting this observation (Supplementary Figure S7g). In the MB0-bound state, helix  $\alpha 9$  is slightly tilted away from helices  $\alpha 7$  and  $\alpha 8$  translocating its C-terminus  $\sim 8$  Å in the presence of MB0 compared to the apo form of PAD (Figure 5E, Supplementary Figure S7h). As a result, key residues in  $\alpha 9$ , including binding patch residues I144, V137, M141, are slightly distanced from helices  $\alpha 7$  and  $\alpha 8$  in the presence of MB0 (Supplementary Figure S7i). Thus, PAD appears to accommodate MB0 binding by opening a binding patch between helices 7–9 in ARM2.

#### Mutations of key residues diminish PNUTS-MYC interaction *in vitro* and in human cells and increases MYC phosphorylation

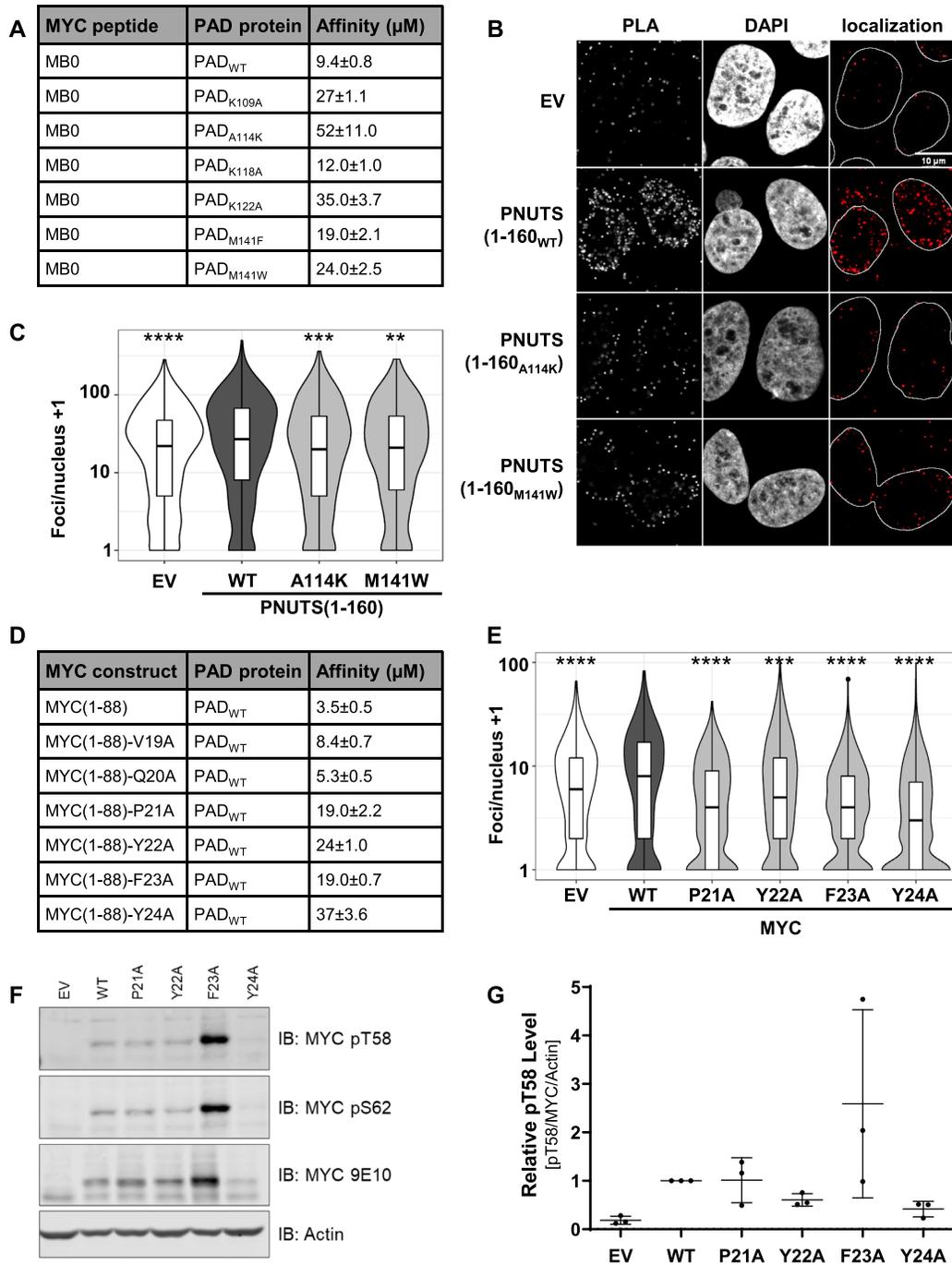
To further validate the mode of interaction, we evaluated the impact of single amino acid substitutions of PAD for their binding to MYC. Using FoldX (61), amino acid substitutions were chosen to preferentially alter PNUTS-MYC interaction without disrupting protein structure. All six of the PAD point mutants (K109A, A114K, K118A, K122A,

M141F and M141W) displayed weaker interaction with the MB0 peptide with five of these reducing the binding affinity by  $>2$ -fold (Figure 6A; Supplementary Figure S8a–f).

To assess whether these residues also contribute to the interaction of MYC and PNUTS in a cellular context, we evaluated the interaction using a proximity ligation assay (PLA). We generated a doxycycline-inducible, in-frame, fusion protein consisting of a Flag-tag, PNUTS amino acids 1–160 (PNUTS(1–160)) and three tandem SV40 nuclear localization signal (NLS) motifs. The longer PNUTS(1–160) construct was used to achieve sufficient expression of the PAD in mammalian cells and the NLS ensured nuclear localization. We expressed this allele in the MCF10A breast epithelial cell line, as this cell line was one of many that has been identified to have a functionally relevant PP1:PNUTS-MYC regulatory axis (27) and images well for PLA. Upon addition of doxycycline to the media, we observed the expected induction of PNUTS(1–160) expression as confirmed by immunoblot (Supplementary Figure S9a) and nuclear localization as determined by immunofluorescence (Supplementary Figure S9b). To evaluate binding of PNUTS(1–160) to full-length MYC in human cells,



**Figure 5.** The specific interactions between PAD and MB0. (A) Cartoon representation of the solution state NMR structure of the PAD-MB0 complex, as derived from the PAD-MB0-fusion protein. The PAD is depicted in light blue and MB0 is in yellow. (B) Histogram showing the number of assigned long-range NOEs between the PAD and MB0 residues, and vice-versa. The data was obtained from NOESY spectra of the PAD-MB0-fusion protein that is depicted above the diagram. Inset: Diagram of the distribution of inter-fragment NOEs - a dashed line connects MB0 and PAD residues with assigned NOEs. (C) Four residues within MB0 are crucial for the PAD-MB0 interaction within the fusion protein. MYC(13-30) is shown in cartoon (yellow) with the PAD ARM2 domain in surface representation (light blue). MYC residues V19, P21, Y22 and F23 are shown in stick format; the ring of F23 tucks into a MYC-binding patch on the surface of PNUTS, formed by V105, A114, V137, W140, M141 and I144 (grey). The MYC fragment adopts a  $\beta$ -turn (residues 20–24) and a helical turn (residues 15–19). (D) Hydrogen bonds (dotted lines) and polar interactions (stick representations) are indicated on the PAD-MB0 interaction surface. PNUTS and MYC residues are coloured in blue and yellow, respectively. One hydrogen bond is formed between PNUTS(K109)-MYC(Y22). Moreover, two hydrogen bonds are formed within the MYC chain between MYC(Q20)-MYC(F23) and MYC(Q20)-MYC(Y22). (E) Cartoon representation of PNUTS-ARM2 in apo (green) and MYC-bound (grey) states. Helices  $\alpha 7$  and  $\alpha 8$  superimpose well, while helix  $\alpha 9$  undergoes a conformational adjustment. Residues L108, K109, K115, K118, K122, V137, M141, I144, R145 and Q147 are shown in stick representation for reference. See also Supplementary Figure S7.



**Figure 6.** Mutations of key residues diminish PNUTS-MYC interaction *in vitro* and in human cells. (A) PAD residues that are important for MYC binding were mutated to amino acids that are able to disrupt the interaction while maintaining the folding of PAD. BLI was performed with MB0 peptide and measured affinities are displayed in the last column. (B) EV control, Flag-tagged PNUTS(1–160<sub>WT</sub>), PNUTS(1–160<sub>A114K</sub>), or PNUTS(1–160<sub>M141W</sub>) were expressed in MCF10A cells expressing V5-MYC. PLA was performed using FLAG and V5 antibodies. Representative images of magnified individual nuclei are shown. (C) The foci formed per nucleus were quantified and are displayed as a violin plot on a logarithmic scale. 1 was added to each measurement for visualization on the chosen scale. Statistical significance was tested using the Kolmogorov-Smirnov test (\*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ ; \*\*\*\* $P \leq 0.0001$ ;  $n = 3$ ). (D) Key residues in the V<sub>19</sub>QPYFY<sub>24</sub> stretch were individually mutated to alanine in the MYC(1–88) expression construct. The different MYC constructs were tested by BLI for binding to PAD and obtained affinities are displayed in the rightmost column. (E) Empty vector control (EV) or V5-tagged MYC constructs (Wild type (WT), P21A, Y22A, F23A or Y24A) were expressed in MCF10A cells and their proximity to endogenous full-length PNUTS assessed by PLA. Quantification is displayed on a logarithmic scale as foci/nucleus plus one. Representative images of magnified individual nuclei are displayed in Supplementary Figure S11a. Statistical significance was tested with the Kolmogorov-Smirnov test (\*\* $P \leq 0.01$ ; \*\*\*\* $P \leq 0.0001$ ;  $n = 3$ ). (F) Cells expressing MYC and MYC mutants were treated for 4h with 10  $\mu\text{M}$  MG132 prior to harvest. Lysates were immunoblotted to test the phosphorylation status of canonical phosphorylation sites on MYC (pS62, pT58). Actin was used as a loading control. Representative blot is shown. (G) Quantification of the pT58 levels of immunoblots shown in (F) and Supplementary Figure S11b. The phosphorylation level was normalized to MYC expression (MYC 9E10 signal) as well as loading (actin signal). See also Supplementary Figures S8 to S11.

we next performed PLA using Flag and MYC antibodies. With this assay co-localization of two proteins of interest within approximately 40 nm is scored as a fluorescent focus. Several fields of view were quantified and the foci enumerated on a per nucleus basis. Induction of PNUTS(1–160) dramatically increased the number of foci formed per nucleus compared to the EV control, giving us confidence that the PAD is also responsible for the interaction of MYC and PNUTS in human cells (Figure 6B). PLA analysis of point mutants of key PAD residues important for MYC interaction (Figure 6A), A114K and M141W, decreased the number of foci to levels similar to background empty vector (EV) control, which was significantly reduced compared to wild-type PNUTS(1–160) (Figure 6C). Taken together, these results validate our *in vitro* data, distinguishing important PNUTS residues for the PNUTS-MYC interaction.

Next, we aimed to assess point mutations of residues within MB0 that we identified as important for the PNUTS-MYC interaction (Figures 3C and 5B). Alanine point mutations of the MB0 residues within the binding stretch V<sub>19</sub>QPYFY<sub>24</sub> (Figure 2) were generated within MYC(1–88) and evaluated for interaction with PAD by BLI. Interaction of wild-type MYC(1–88) with PAD has a binding affinity of  $K_D \sim 3.5 \mu\text{M}$  (Figure 1B), whereas the MYC mutants MYC(1–88<sub>P21A</sub>), MYC(1–88<sub>Y22A</sub>), MYC(1–88<sub>F23A</sub>) and MYC(1–88<sub>Y24A</sub>) had reduced  $K_D$  values by a factor of 5 or more (Figure 6D; Supplementary Figure S8g–l). These reductions are similar in magnitude to what has been previously reported for point mutation of key residues of MYC that affect the interaction with other partner proteins such as TBP (47), Pin1 (31) and WDR5 (62).

To interrogate the role of key MYC residues for PNUTS-MYC interaction within cells, we chose to investigate the MYC point mutants that showed the largest reduction in  $K_D$  (Figure 6D), for their binding of endogenous PNUTS in human cell lines. To this end, we expressed V5-tagged MYC and MYC point mutants stably in the MCF10A cell line. Using a V5 antibody, we demonstrated similar levels of expression and nuclear localization of all mutants as well as wild-type MYC by immunoblot and immunofluorescence, respectively (Supplementary Figure S10a and b). We then performed PLA to measure the interaction of MYC(WT), MYC(P21A), MYC(Y22A), MYC(F23A) and MYC(Y24A) with endogenous full-length PNUTS. Indeed, each of the four MYC point mutants significantly reduced the number of fluorescent foci formed as compared to wild-type MYC (Figure 6E, Supplementary Figure S11a). Notably, all MYC point mutants reduced the interaction to background levels as determined by the EV control. Overall, our mutational analysis confirmed that the MYC residues that we identified as key determinants of PNUTS-MYC interaction *in vitro*, are also important within cells.

To assess whether the disruption of the MYC-PNUTS:PP1 interaction has a functional consequence, we used the MYC point mutants, which we determined to lose interaction with full-length endogenous PNUTS, and assessed MYC phosphorylation at two well characterised residues, T58 and S62. To prevent degradation of phosphorylated MYC species, we treated cells with the protease inhibitor MG132 prior to harvest. Immunoblotting for pT58 and pS62 revealed that mutation of MYC(F23A) in-

creased the level of phosphorylated MYC species, whereas the other mutations did not show a large effect (Figure 6F, G; Supplementary Figure S11b and c). Taken together, these findings indicate that the regions of MYC and PNUTS as well as specific residues within these regions are necessary to maintain MYC in an under-phosphorylated state.

## DISCUSSION

Distinguishing direct protein interactors of the MYC oncoprotein is essential, not only to understand the mechanism of MYC function, but also to potentially unveil novel strategies to inhibit MYC binding to key protein partners and thus disable MYC as a global transcriptional regulator and potent cancer driver. We have previously shown the PP1:PNUTS phosphatase complex regulates MYC phosphorylation, chromatin eviction and degradation (27), however the molecular basis of the interaction remained unclear. Here we show by NMR, molecular modeling, and both *in vitro* and *in vivo* validation assays, that PNUTS and MYC interact directly through the PAD and MB0 regions, respectively. Using NMR, we first determined the apo structure of the PAD (PDB ID: 6VTI), which was shown to consist of nine helices arranged as three helical bundles, with the two C-terminal bundles consisting of two ARM repeats; ARM1 and ARM2, consistent with the recently reported structure of PNUTS(1–148) (56). Using NMR and computational modeling, we then determined the molecular basis of the interaction by analyzing (i) PAD and MYC(1–88), and (ii) a PAD-MB0-fusion protein (PDB ID: 7LQT). Our high-resolution structural data support a model in which a core stretch of amino acids in MB0 binds into the MYC-binding patch on the C-terminal facing surface of the PAD ARM2 motif, primarily through hydrophobic and electrostatic interactions. Moreover, we show that this interaction ensemble is highly dynamic but comprises a narrow range of specific, bound conformations. Validation of the interaction was achieved by demonstrating that point mutants of key interacting residues in either MYC or PNUTS disrupted the interaction both *in vitro* and *in vivo*. Moreover, disruption of MYC-PNUTS interaction led to elevated MYC phosphorylation, as anticipated. Thus, we have determined the molecular basis of the MYC-PNUTS interaction.

The dynamic alterations in MYC in response to PAD binding, together with structural specifics of the bound state and cellular effects of its inhibition, suggest a model in which MB0 binding to PAD facilitates PP1:PNUTS access to phosphorylation sites within and adjacent to MBI for subsequent dephosphorylation (see Graphical Abstract). Our NMR analyses of MYC(1–88) showed that MB0 residues V19–Y24 comprise the primary anchor site of the PAD-MYC(1–88) interaction, with a weaker second touch-point within MBI (residues W50–L55). In contrast, MBI-neighbouring residues (L61 to S73) show dramatically sharper amide signals on PAD binding to MB0. This suggests a more rapidly interconverting ensemble state in this region compared to MYC(1–88) alone and a possible release of internal MYC interactions upon PAD binding, which is similar to previous observations for tumour suppressor Bin1 binding to the MBI region of MYC (39). In

the context of the MYC-PNUTS interaction, the region of MYC that appears to be released upon PAD binding comprises critical serine/threonine residues essential for functional phosphorylation (T58, S62). This agrees with the role of the PP1:PNUTS-MYC axis in the regulation of MYC phosphorylation, transcriptional activity and protein stability (27). Such an anchor-release mechanism of PAD-MB0 binding in the PNUTS-MYC complex resembles that of Pin1 binding to MYC, where its anchoring to MB0 enhances cis-trans isomerisation C-terminal of MBI, thereby enhancing phosphorylation of S62 (31). However, while Pin1 is a general cis-trans isomerase acting on many targets, PNUTS binding to MYC at the same site (MB0) holds a specific key regulatory role of MYC function and degradation. These observations raise the question, how can MYC interact and be regulated by multiple proteins through the same interaction domain? This question is still outstanding, however, a theoretical model, termed the ‘coalition model’, has been proposed by many of the authors of this manuscript in a recent perspective article (7), which is supported by previously published evidence (19). In brief, it is possible to conceptualize that each MYC molecule in a cell can interact with a specific protein complex at a given time. These complexes can be distinct, and it is the net effect of many MYC proteins interacting with a multitude of interactors, through which MYC function is realized. This means, that even though the specificity of the binding motif identified here is verified by single structure-based point-mutations that entirely abort the MYC-PNUTS interaction *in vitro* and in cells, it is important to consider that this and other regions of MYC can bind additional protein complexes. The regulatory conditions under which MYC interaction occurs with individual protein complexes, and the nature of their binding at a molecular level, remains to be determined. Advancing our understanding of the molecular basis of the MYC-PNUTS interaction, as shown here, provides new insight into MYC structure, function and MYC-protein interaction.

Our analyses revealed that the structure of the PAD is also affected by the PAD-MB0 interaction. Specifically, in the MB0-bound state, the  $\alpha 9$  helix of PNUTS ARM2 is tilted away from  $\alpha 8$  as compared to apo PAD. Both *in-silico* modeling and direct NMR studies suggest that this conformational change supports the formation of the MYC-binding patch on PNUTS required to accommodate the buried MYC(F23) residue. The PNUTS-MYC interaction may be further regulated by key PNUTS residues (I144, V137, M141) that could contribute to a more closed or open configuration of the ARM2 domain MYC-binding patch. Reciprocal structural changes in PNUTS that widen the binding patch upon MYC interaction suggest a mutual conformational adaptation, but whether this resembles induced fit or conformational selection models requires further investigation. The key structural and dynamic features of the PNUTS-MYC interaction revealed here may jointly contribute to the regulation of the phosphorylation state of MYC.

PNUTS is one of several PP1 phosphatase regulatory subunits that determines substrate specificity. In particular, PNUTS directs PP1 dephosphorylation activity to protein substrates localized in the nucleus, including TOX high-

mobility group box family member 4 (aka TOX4, LCP1) (63), phosphatase and tensin homolog (PTEN) (64), telomeric repeat-binding factor 2 (65), RNA polymerase II (66), retinoblastoma protein (67), and WDR82 (68). Of these only TOX4 and PTEN have been shown to bind to the PNUTS N-terminal region, which includes the PAD revealed in this work. The calponin-homology domain of TOX4 interacts with PNUTS residues 1-263 (63), and the C2 tensin-type domain of PTEN has been reported to interact with the first 146 residues of PNUTS (64). Here we show how the intrinsically disordered MYC(1–88) interacts with the PAD in a fuzzy complex, and identify and characterise a narrow ensemble of bound states by increasing the local concentration of MYC close to the binding site on PNUTS by means of a PAD-MB0-fusion protein. Taken together, these results suggest that the PAD recognizes and distinguishes specific PP1 nuclear substrates by a wide range of binding mechanisms, and thereby acts as a landing pad for protein interactions regulating the activity of the PP1:PNUTS complex. While our manuscript was under revision, a recent study showed how several proteins with TFIIIS N-terminal domains (TNDs), including PNUTS, interact with TND-interacting motifs (TIMs; (56)). The PNUTS-interacting motif in MYC is clearly distinct from these TIMs, and the regulatory role of this interaction in terms of MYC phosphorylation further supports the specificity of this interaction. The detailed structural and dynamic determinants of MYC binding to PNUTS that we describe here provide further understanding of how distinct peptides bind specifically to PAD-related motifs.

Targeting MYC activity by disrupting MYC-protein cofactor interactions essential for gene transcription is an area of intense interest as inhibiting MYC directly has not been fruitful to date. As the bHLHLZ interaction with MAX was the first binding partner of MYC (69), several groups endeavour to exploit this partnership for the development of inhibitors (15,21,70–72). More recently, the focus has been on targeting interactors of the unique MYC Box regions. The interaction of MBIV with WDR5 induces a specific subset of genes whose products regulate protein synthesis, and inhibitors to block this interaction are under development (73,74). Moreover, recent insight into the structural basis of MBII-TRRAP interaction has unveiled potential strategies to target this key interaction (75). Despite decades of MYC research, MB0 has only recently been recognized as a highly conserved MYC box that is functionally important for MYC oncogenic activity (19,31). Our finding here that PNUTS directly interacts with MB0 further reinforces the functional importance of this PP1:PNUTS-MYC regulatory axis, which we had previously shown controls MYC activity and stability (27). Indeed, PNUTS was recently shown by an independent group to regulate N-MYC stability (76), further emphasizing the critical nature of the PNUTS-MYC interaction to MYC family activity. This is consistent with MB0 being conserved amongst MYC family members, thereby enabling PP1:PNUTS to regulate the rapid turnover of these oncoproteins. Further research is needed to address the specifics of the PP1:PNUTS interaction with other MYC family members on a molecular basis. Moreover, it will be interesting to identify how MB0 is able to accommodate different interactors and which determi-

nants favour binding of one over the other. These insights will collectively inform how best to exploit MYC-partner protein interaction for the development of novel anti-MYC therapeutics.

## DATA AVAILABILITY

The NMR-derived structures and chemical shift assignments reported in this study have been deposited to the PDB and BMRB as 6VTI (30722) for PAD, and 7LQT (30861) for PAD-MB0-fusion.

The NMR-based model reported in this study has been deposited in Modelarchive with the following access: <https://modelarchive.org/doi/10.5452/ma-3ef73> code: LWSnMfLXwf.

## SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

## ACKNOWLEDGEMENTS

*Author Contributions:* Y.W., L.Z.P. conceived the project. Y.W. performed the bulk of protein purifications. V. M. cloned and purified MYC(1–88) for NMR. Y. W. performed BLI affinity measurements with the help from T.K. Y.W., S.H., S.D., A.L., and A.A. performed the NMR studies of the apo PAD and PAD-MB0-fusion. A.L., S.H. determined the NMR structures of apo PAD and PAD-MB0-fusion protein. A.A., I. J.-Å., M.S., B.W. performed and evaluated the NMR studies of MYC peptides with PAD. I. J.-Å., A. A. and B.W. carried out the NMR-guided structural modeling. C.R., A.T. performed all experiments in human cell lines. D.W.A., B.W., M.S., C.H.A., L.Z.P. supervised the work. Y.W., C.R., A.A., I. J.-Å., S.H., D.W.A., B.W., M.S., C.H.A., L.Z.P. wrote the manuscript. All authors discussed the results and contributed to the final manuscript.

## FUNDING

Canadian Institutes of Health Research [FRN156167 to L.Z.P., FDN154328 to C.H.A., FDN143312 to D.W.A.]; Swedish Cancer Society [20 1276 PjF 01 H to M.S.]; Swedish Childhood Cancer Fund [PR2019-0143 project grant to M.S., TJ2018-0103 postdoc award to A.A.]; Swedish Research Council [2018-04390 to M.S., 2016-05369 to B.W.]; Princess Margaret Cancer Centre; Princess Margaret Cancer Foundation; Ontario Ministry of Health; the Structural Genomics Consortium is a registered charity [1097737] that receives funds from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, Genome Canada through Ontario Genomics Institute [OGI-196]; EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking [EUbOPEN grant 875510]; Janssen, Merck KGaA (aka EMD in Canada and US); Pfizer; Takeda; NMR access at the ProLinC core facility was funded by Linköping University; the computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre (NSC) in Linköping; L.Z.P. and D.W.A. hold Tier 1 Canada Research Chairs in

Molecular Oncology and Membrane Biogenesis, respectively. Funding for open access charge: Canadian Institutes of Health Research.

*Conflict of interest statement.* None declared.

## REFERENCES

- Meyer, N. and Penn, L.Z. (2008) Reflecting on 25 years with MYC. *Nat. Rev. Cancer*, **8**, 976–990.
- Kalkat, M., De Melo, J., Hickman, K.A., Lourenco, C., Redel, C., Resetca, D., Tamachi, A., Tu, W.B. and Penn, L.Z. (2017) MYC deregulation in primary human cancers. *Genes (Basel)*, **8**, 151.
- Luscher, B. and Vervoorts, J. (2012) Regulation of gene transcription by the oncoprotein MYC. *Gene*, **494**, 145–160.
- Hsieh, A.L., Walton, Z.E., Altman, B.J., Stine, Z.E. and Dang, C.V. (2015) MYC and metabolism on the path to cancer. *Review Semin Cell Dev Biol*, **43**, 11–21.
- Poole, C.J. and van Riggelen, J. (2017) MYC-Master regulator of the cancer epigenome and transcriptome. *Genes (Basel)*, **8**, 142.
- Casey, S.C., Baylot, V. and Felsher, D.W. (2017) MYC: master regulator of immune privilege. *Trends Immunol.*, **38**, 298–305.
- Lourenco, C., Resetca, D., Redel, C., Lin, P., MacDonald, A.S., Ciaccio, R., Kenney, T.M.G., Wei, Y., Andrews, D.W., Sunnerhagen, M. et al. (2021) MYC protein interactors in gene transcription and cancer. *Nat. Rev. Cancer*, **21**, 579–591.
- Prochownik, E.V. and Vogt, P.K. (2010) Therapeutic targeting of Myc. *Genes Cancer*, **1**, 650–659.
- Whitfield, J.R., Beaulieu, M.E. and Soucek, L. (2017) Strategies to inhibit Myc and their clinical applicability. *Front. Cell Dev. Biol.*, **5**, 10.
- Ponzielli, R., Katz, S., Barsyte-Lovejoy, D. and Penn, L.Z. (2005) Cancer therapeutics: targeting the dark side of Myc. *Eur. J. Cancer*, **41**, 2485–2501.
- Soucek, L., Whitfield, J.R., Sodik, N.M., Masso-Valles, D., Serrano, E., Karnezis, A.N., Swigart, L.B. and Evan, G.I. (2013) Inhibition of myc family proteins eradicates KRas-driven lung cancer in mice. *Genes Dev.*, **27**, 504–513.
- Li, Y., Casey, S.C. and Felsher, D.W. (2014) Inactivation of MYC reverses tumorigenesis. *J. Intern. Med.*, **276**, 52–60.
- Bellovin, D.I., Das, B. and Felsher, D.W. (2013) Tumor dormancy, oncogene addiction, cellular senescence, and self-renewal programs. *Adv. Exp. Med. Biol.*, **734**, 91–107.
- Felsher, D.W. and Bishop, J.M. (1999) Reversible tumorigenesis by MYC in hematopoietic lineages. *Mol. Cell*, **4**, 199–207.
- Dang, C.V., Reddy, E.P., Shokat, K.M. and Soucek, L. (2017) Drugging the ‘undruggable’ cancer targets. *Nat. Rev. Cancer*, **17**, 502–508.
- McKeown, M.R. and Bradner, J.E. (2014) Therapeutic strategies to inhibit MYC. *Cold Spring Harb. Perspect. Med.*, **4**, a014266.
- Chen, H., Liu, H. and Qing, G. (2018) Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct. Target Ther.*, **3**, 5.
- Tu, W.B., Helander, S., Pilstal, R., Hickman, K.A., Lourenco, C., Jurisica, I., Raught, B., Wallner, B., Sunnerhagen, M. and Penn, L.Z. (2015) Myc and its interactors take shape. *Biochim. Biophys. Acta*, **1849**, 469–483.
- Kalkat, M., Resetca, D., Lourenco, C., Chan, P.K., Wei, Y., Shiah, Y.J., Vitkin, N., Tong, Y., Sunnerhagen, M., Done, S.J. et al. (2018) MYC protein interactome profiling reveals functionally distinct regions that cooperate to drive tumorigenesis. *Mol. Cell*, **72**, 836–848.
- Bugge, K., Brakti, I., Fernandes, C.B., Dreier, J.E., Lundsgaard, J.E., Olsen, J.G., Skriver, K. and Kragelund, B.B. (2020) Interactions by disorder – a matter of context. *Front. Mol. Biosci.*, **7**, 110.
- Lu, H., Zhou, Q., He, J., Jiang, Z., Peng, C., Tong, R. and Shi, J. (2020) Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduct. Target. Ther.*, **5**, 213.
- Dingar, D., Kalkat, M., Chan, P.K., Srikumar, T., Bailey, S.D., Tu, W.B., Coyaud, E., Ponzielli, R., Kolyar, M., Jurisica, I. et al. (2015) BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. *J. Proteomics*, **118**, 95–111.
- Wasylishen, A.R., Chan-Seng-Yue, M., Bros, C., Dingar, D., Tu, W.B., Kalkat, M., Chan, P.K., Mullen, P.J., Huang, L., Meyer, N. et al. (2013) MYC phosphorylation at novel regulatory regions suppresses transforming activity. *Cancer Res.*, **73**, 6504–6515.

24. Huang,Z., Traugh,J.A. and Bishop,J.M. (2004) Negative control of the Myc protein by the stress-responsive kinase Pak2. *Mol. Cell Biol.*, **24**, 1582–1594.
25. Chakraborty,A.A., Scuoppo,C., Dey,S., Thomas,L.R., Lorey,S.L., Lowe,S.W. and Tansey,W.P. (2015) A common functional consequence of tumor-derived mutations within c-MYC. *Oncogene*, **34**, 2406–2409.
26. Hemann,M.T., Bric,A., Teruya-Feldstein,J., Herbst,A., Nilsson,J.A., Cordon-Cardo,C., Cleveland,J.L., Tansey,W.P. and Lowe,S.W. (2005) Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. *Nature*, **436**, 807–811.
27. Dingar,D., Tu,W.B., Resetca,D., Lourenco,C., Tamachi,A., De Melo,J., Houlihan,K.E., Kalkat,M., Chan,P.K., Boutros,P.C. *et al.* (2018) MYC dephosphorylation by the PP1/PNUTS phosphatase complex regulates chromatin binding and protein stability. *Nat. Commun.*, **9**, 3502.
28. Casamayor,A. and Arino,J. (2020) Controlling Ser/Thr protein phosphatase PP1 activity and function through interaction with regulatory subunits. *Adv. Protein Chem. Struct. Biol.*, **122**, 231–288.
29. Bollen,M., Peti,W., Ragusa,M.J. and Beullens,M. (2010) The extended PP1 toolkit: designed to create specificity. *Trends Biochem. Sci.*, **35**, 450–458.
30. Cohen,P.T. (2002) Protein phosphatase 1–targeted in many directions. *J. Cell Sci.*, **115**, 241–256.
31. Helander,S., Montecchio,M., Pilstal,R., Su,Y., Kuruvilla,J., Elven,M., Ziauddin,J.M.E., Anandapadamanaban,M., Cristobal,S., Lundstrom,P. *et al.* (2015) Pre-Anchoring of pin1 to unphosphorylated c-Myc in a fuzzy complex regulates c-Myc activity. *Structure*, **23**, 2267–2279.
32. Savitsky,P., Bray,J., Cooper,C.D., Marsden,B.D., Mahajan,P., Burgess-Brown,N.A. and Gileadi,O. (2010) High-throughput production of human proteins for crystallization: the SGC experience. *J. Struct. Biol.*, **172**, 3–13.
33. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
34. Lee,W., Tonelli,M. and Markley,J.L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*, **31**, 1325–1327.
35. Kazimierzczuk,K. and Orekhov,V. (2015) Non-uniform sampling: post-Fourier era of NMR data collection and processing. *Magn. Reson. Chem.*, **53**, 921–926.
36. Lemak,A., Gutmanas,A., Chitayat,S., Karra,M., Fares,C., Sunnerhagen,M. and Arrowsmith,C.H. (2011) A novel strategy for NMR resonance assignment and protein structure determination. *J. Biomol. NMR*, **49**, 27–38.
37. Shen,Y., Delaglio,F., Cornilescu,G. and Bax,A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*, **44**, 213–223.
38. Guntert,P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.*, **278**, 353–378.
39. Andresen,C., Helander,S., Lemak,A., Fares,C., Csizmok,V., Carlsson,J., Penn,L.Z., Forman-Kay,J.D., Arrowsmith,C.H., Lundstrom,P. *et al.* (2012) Transient structure and dynamics in the disordered c-Myc transactivation domain affect Bin1 binding. *Nucleic Acids Res.*, **40**, 6353–6366.
40. Raveh,B., London,N., Zimmerman,L. and Schueler-Furman,O. (2011) Rosetta flexpepdock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. *PLoS One*, **6**, e18934.
41. Schumann,F.H., Riepl,H., Maurer,T., Gronwald,W., Neidig,K.P. and Kalbitzer,H.R. (2007) Combined chemical shift changes and amino acid specific chemical shift mapping of protein-protein interactions. *J. Biomol. NMR*, **39**, 275–289.
42. Raveh,B., London,N. and Schueler-Furman,O. (2010) Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins*, **78**, 2029–2040.
43. Shapovalov,M.V. and Dunbrack,R.L. Jr (2011) A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure*, **19**, 844–858.
44. Pedregosa,F., Varoquaux,G., Gramfort,A., Michel,V., Thirion,B., Grisel,O., Blondel,M., Prettenhofer,P., Weiss,R., Dubourg,V. *et al.* (2011) Scikit-learn: machine learning in python. *J. Mach. Learn. Res.*, **12**, 2825–2830.
45. Ashkenazy,H., Erez,E., Martz,E., Pupko,T. and Ben-Tal,N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.*, **38**, W529–W533.
46. Remmert,M., Biegert,A., Hauser,A. and Soding,J. (2011) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat. Methods*, **9**, 173–175.
47. Wei,Y., Resetca,D., Li,Z., Johansson-Akhe,I., Ahlner,A., Helander,S., Wallenhammar,A., Morad,V., Raught,B., Wallner,B. *et al.* (2019) Multiple direct interactions of TBP with the MYC oncoprotein. *Nat. Struct. Mol. Biol.*, **26**, 1035–1043.
48. Zacharchenko,T., Barsukov,I., Rigden,D.J., Bennett,D. and Mayans,O. (2016) Biophysical analysis of the N-terminal domain from the human protein phosphatase 1 nuclear targeting subunit PNUTS suggests an extended transcription factor TFIIIS-like fold. *Protein J.*, **35**, 340–345.
49. Kim,Y.M., Watanabe,T., Allen,P.B., Kim,Y.M., Lee,S.J., Greengard,P., Nairn,A.C. and Kwon,Y.G. (2003) PNUTS, a protein phosphatase 1 (PP1) nuclear targeting subunit. Characterization of its PP1- and RNA-binding domains and regulation by phosphorylation. *J. Biol. Chem.*, **278**, 13819–13828.
50. Allen,P.B., Kwon,Y.G., Nairn,A.C. and Greengard,P. (1998) Isolation and characterization of PNUTS, a putative protein phosphatase 1 nuclear targeting subunit. *J. Biol. Chem.*, **273**, 4089–4095.
51. Fladvad,M., Zhou,K., Moshref,A., Pursglove,S., Safsten,P. and Sunnerhagen,M. (2005) N and C-terminal sub-regions in the c-Myc transactivation region and their joint role in creating versatility in folding and binding. *J. Mol. Biol.*, **346**, 175–189.
52. Bozoky,Z., Krzeminski,M., Muhandiram,R., Birtley,J.R., Al-Zahrani,A., Thomas,P.J., Frizzell,R.A., Ford,R.C. and Forman-Kay,J.D. (2013) Regulatory R region of the CFTR chloride channel is a dynamic integrator of phospho-dependent intra- and intermolecular interactions. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E4427–E4436.
53. Forman-Kay,J.D. and Mittag,T. (2013) From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure*, **21**, 1492–1499.
54. Fuxreiter,M. (2018) Fuzziness in protein interactions - a historical perspective. *J. Mol. Biol.*, **430**, 2278–2287.
55. Zhang,Y. and Skolnick,J. (2005) TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.*, **33**, 2302–2309.
56. Cermakova,K., Demeulemeester,J., Lux,V., Nedomova,M., Goldman,S.R., Smith,E.A., Srb,P., Hexnerova,R., Fabry,M., Madlikova,M. *et al.* (2021) A ubiquitous disordered protein interaction module orchestrates transcription elongation. *Science*, **374**, 1113–1121.
57. Sharma,S., Cermakova,K., De Rijck,J., Demeulemeester,J., Fabry,M., El Ashkar,S., Van Belle,S., Lepsik,M., Tesina,P., Duchoslav,V. *et al.* (2018) Affinity switching of the LEDGF/p75 IBD interactome is governed by kinase-dependent phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E7053–E7062.
58. Ashkenazy,H., Abadi,S., Martz,E., Chay,O., Mayrose,I., Pupko,T. and Ben-Tal,N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.*, **44**, W344–W350.
59. Landau,M., Mayrose,I., Rosenberg,Y., Glaser,F., Martz,E., Pupko,T. and Ben-Tal,N. (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.*, **33**, W299–W302.
60. Tibshirani,R. (1996) Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B (Methodological)*, **58**, 167–288.
61. Schymkowitz,J., Borg,J., Stricher,F., Nys,R., Rousseau,F. and Serrano,L. (2005) The FoldX web server: an online force field. *Nucleic Acids Res.*, **33**, W382–W388.
62. Thomas,L.R., Wang,Q., Grieb,B.C., Phan,J., Foshage,A.M., Sun,Q., Olejniczak,E.T., Clark,T., Dey,S., Lorey,S. *et al.* (2015) Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC. *Mol. Cell*, **58**, 440–452.
63. Lee,S.J., Lee,J.K., Maeng,Y.S., Kim,Y.M. and Kwon,Y.G. (2009) Langerhans cell protein 1 (LCP1) binds to PNUTS in the nucleus: implications for this complex in transcriptional regulation. *Exp. Mol. Med.*, **41**, 189–200.

64. Kavela,S., Shinde,S.R., Ratheesh,R., Viswakalyan,K., Bashyam,M.D., Gowrishankar,S., Vamsy,M., Pattnaik,S., Rao,S., Sastry,R.A. *et al.* (2013) PNUITS functions as a proto-oncogene by sequestering PTEN. *Cancer Res.*, **73**, 205–214.
65. Kim,H., Lee,O.H., Xin,H., Chen,L.Y., Qin,J., Chae,H.K., Lin,S.Y., Safari,A., Liu,D. and Songyang,Z. (2009) TRF2 functions as a protein hub and regulates telomere maintenance by recognizing specific peptide motifs. *Nat. Struct. Mol. Biol.*, **16**, 372–379.
66. Booth,V., Koth,C.M., Edwards,A.M. and Arrowsmith,C.H. (2000) Structure of a conserved domain common to the transcription factors TFIIS, elongin A, and CRSP70. *J. Biol. Chem.*, **275**, 31266–31268.
67. De Leon,G., Sherry,T.C. and Krucher,N.A. (2008) Reduced expression of PNUITS leads to activation of Rb-phosphatase and caspase-mediated apoptosis. *Cancer Biol. Ther.*, **7**, 833–841.
68. Landsverk,H.B., Sandquist,L.E., Bay,L.T.E., Steurer,B., Campsteijn,C., Landsverk,O.J.B., Marteiijn,J.A., Petermann,E., Trinkle-Mulcahy,L. and Syljuasen,R.G. (2020) WDR82/PNUITS-PP1 prevents transcription-replication conflicts by promoting RNA polymerase II degradation on chromatin. *Cell Rep.*, **33**, 108469.
69. Blackwood,E.M. and Eisenman,R.N. (1991) Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science*, **251**, 1211–1217.
70. Han,H., Jain,A.D., Truica,M.I., Izquierdo-Ferrer,J., Anker,J.F., Lysy,B., Sagar,V., Luan,Y., Chalmers,Z.R., Unno,K. *et al.* (2019) Small-Molecule MYC inhibitors suppress tumor growth and enhance immunotherapy. *Cancer Cell*, **36**, 483–497.
71. Castell,A., Yan,Q., Fawcner,K., Hydbring,P., Zhang,F., Verschut,V., Franco,M., Zakaria,S.M., Bazzar,W., Goodwin,J. *et al.* (2018) A selective high affinity MYC-binding compound inhibits MYC:MAX interaction and MYC-dependent tumor cell proliferation. *Sci. Rep.*, **8**, 10064.
72. Carabet,L.A., Lallous,N., Leblanc,E., Ban,F., Morin,H., Lawn,S., Ghaidi,F., Lee,J., Mills,I.G., Gleave,M.E. *et al.* (2018) Computer-aided drug discovery of Myc-Max inhibitors as potential therapeutics for prostate cancer. *Eur. J. Med. Chem.*, **160**, 108–119.
73. Thomas,L.R., Adams,C.M., Wang,J., Weissmiller,A.M., Creighton,J., Lorey,S.L., Liu,Q., Fesik,S.W., Eischen,C.M. and Tansey,W.P. (2019) Interaction of the oncoprotein transcription factor MYC with its chromatin cofactor WDR5 is essential for tumor maintenance. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 25260–25268.
74. Chacon Simon,S., Wang,F., Thomas,L.R., Phan,J., Zhao,B., Olejniczak,E.T., Macdonald,J.D., Shaw,J.G., Schlund,C., Payne,W. *et al.* (2020) Discovery of WD repeat-containing protein 5 (WDR5)-MYC inhibitors using fragment-based methods and structure-based design. *J. Med. Chem.*, **63**, 4315–4333.
75. Feris,E.J., Hinds,J.W. and Cole,M.D. (2019) Formation of a structurally-stable conformation by the intrinsically disordered MYC:TRRAP complex. *PLoS One*, **14**, e0225784.
76. Tee,A.E., Ciampa,O.C., Wong,M., Fletcher,J.I., Kamili,A., Chen,J., Ho,N., Sun,Y., Carter,D.R., Cheung,B.B. *et al.* (2020) Combination therapy with the CDK7 inhibitor and the tyrosine kinase inhibitor exerts synergistic anticancer effects against MYCN-amplified neuroblastoma. *Int. J. Cancer*, **147**, 1928–1938.
77. Bhattacharya,A., Tejero,R. and Montelione,G.T. (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins*, **66**, 778–795.