

Structural Basis and Sequence Rules for Substrate Recognition by Tankyrase Explain the Basis for Cherubism Disease

Sebastian Guettler,^{1,2} Jose LaRose,³ Evangelia Petsalaki,^{1,2} Gerald Gish,¹ Andy Scotter,³ Tony Pawson,^{1,2,*} Robert Rottapel,^{3,4,*} and Frank Sicheri^{1,2,*}

¹Centre for Systems Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada

²Department of Molecular Genetics, University of Toronto, 1 Kings College Circle, Toronto, Ontario M5S 1A8, Canada

³Ontario Cancer Institute and the Campbell Family Cancer Research Institute, 101 College Street, Room 8-703, Toronto Medical Discovery Tower, University of Toronto, Toronto, Ontario M5G 1L7, Canada

⁴Division of Rheumatology, Department of Medicine, Saint Michael's Hospital, 30 Bond Street, Toronto, Ontario M5B 1W8, Canada

*Correspondence: pawson@lunenfeld.ca (T.P.), rottapel@uhnres.utoronto.ca (R.R.), sicheri@lunenfeld.ca (F.S.)

DOI 10.1016/j.cell.2011.10.046

SUMMARY

The poly(ADP-ribose)polymerases Tankyrase 1/2 (TNKS/TNKS2) catalyze the covalent linkage of ADP-ribose polymer chains onto target proteins, regulating their ubiquitylation, stability, and function. Dysregulation of substrate recognition by Tankyrases underlies the human disease cherubism. Tankyrases recruit specific motifs (often called RxxPDG “hexapeptides”) in their substrates via an N-terminal region of ankyrin repeats. These ankyrin repeats form five domains termed ankyrin repeat clusters (ARCs), each predicted to bind substrate. Here we report crystal structures of a representative ARC of TNKS2 bound to targeting peptides from six substrates. Using a solution-based peptide library screen, we derive a rule-based consensus for Tankyrase substrates common to four functionally conserved ARCs. This 8-residue consensus allows us to rationalize all known Tankyrase substrates and explains the basis for cherubism-causing mutations in the Tankyrase substrate 3BP2. Structural and sequence information allows us to also predict and validate other Tankyrase targets, including Disc1, Striatin, Fat4, RAD54, BCR, and MERIT40.

INTRODUCTION

ADP-ribosylation of proteins or other acceptors is catalyzed by a family of 22 known or putative human ADP-ribosyltransferases, which use nicotinamide adenine dinucleotide (NAD⁺) as a source for transferring ADP-ribose onto their substrates, either as monomers or by constructing poly(ADP-ribose) chains (reviewed in Amé et al., 2004; reviewed in Hassa and Hottiger, 2008; reviewed in Hottiger et al., 2010; Kleine et al., 2008). Protein ADP-ribosylation reportedly occurs on aspartate, glutamate,

asparagine, arginine, lysine, cysteine, phosphoserine, and diphthamide residues (reviewed in Hottiger et al., 2010). As a large posttranslational modification of substantial negative charge, protein poly(ADP-ribosylation) (PARsylation) can influence protein fate through several mechanisms, including a direct effect on protein activity, recruitment of binding partners that recognize poly(ADP-ribose), or by affecting protein turnover.

Tankyrase is a multidomain poly(ADP-ribose)polymerase (PARP) with an N-terminal region rich in ankyrin repeats, a sterile-alpha motif (SAM) domain that mediates Tankyrase oligomerization, and a C-terminal catalytic PARP domain (reviewed in Hsiao and Smith, 2008; Smith et al., 1998) (Figure 1A). The N-terminal ankyrin repeats cluster into five domains (“ankyrin repeat clusters,” ARCs), whose precise structures remain to be determined (Seimiya and Smith, 2002). The human genome encodes two similar Tankyrases, TNKS and TNKS2 (PARP5/ARTD5 and PARP6/ARTD6). Both recruit a variety of substrates involved in a broad range of biological functions (Table 1). Tankyrases recognize linear peptide motifs consisting minimally of six consecutive amino acids with high apparent degeneracy in sequence (the TNKS-binding motif, extended to 8 amino acids as shown below) (reviewed in Hsiao and Smith, 2008; Sbodio and Chi, 2002; Seimiya et al., 2004; Seimiya and Smith, 2002). This makes rationalization of known substrates and the prediction of additional substrates difficult. To date, TNKS-binding motifs have been validated or proposed in 17 proteins (Table 1). In many of the studied systems, binding of target proteins by Tankyrase results in their PARsylation. Tankyrase targets some of its substrates for ubiquitylation and proteasome-dependent degradation, as observed for TERF1/TRF1 (Chang et al., 2003; Smith et al., 1998), AXIN (Huang et al., 2009), MCL1 (Bae et al., 2003), and 3BP2 (see accompanying manuscript by Levaot et al., 2011 [this issue of *Cell*]). In the case of 3BP2, mutations in the TNKS-binding motif that abolish Tankyrase recognition underlie the human disease cherubism, a condition characterized by inflammatory lesions of the facial bone (see accompanying manuscript by Levaot et al., 2011). These findings highlight the essential role of substrate targeting in Tankyrase biological function.

Here we uncover the structural and functional basis of how Tankyrases recognize their substrates and derive a comprehensive set of rules that explain the basis for cherubism disease and enable the accurate prediction of Tankyrase substrates.

RESULTS

The TNKS2 N Terminus Contains Minimally Four Substrate-Binding Sites

Previous work indicated that the N-terminal ankyrin repeat region of Tankyrase has five ARCs (Seimiya et al., 2004; Seimiya and Smith, 2002), each predicted to bind substrate (Figure 1A). A bacterially expressed TNKS2 fragment containing the entire ankyrin repeat region (residues 20–800) bound a fluorescein-labeled 3BP2 substrate recruitment peptide (LPHLQRSPPDGQSFRRS) with an apparent affinity of $0.5 \pm 0.1 \mu\text{M}$ (using a one-site binding model; Figure 1B, left). TNKS2 fragments corresponding to ARCs 1, 4, and 5 and a double-ARC construct comprising both ARCs 2 and 3 (but not ARCs 2 or 3 individually) were also solubly expressed and purified. ARCs 1, 4, and 5 each bound the 3BP2 peptide (dissociation constant $[K_D] = 6.9 \pm 2.8 \mu\text{M}$, $6.3 \pm 1.2 \mu\text{M}$, and $1.3 \pm 0.2 \mu\text{M}$, respectively; Figure 1B, right). The ARC2-3 unit also bound the 3BP2 peptide ($K_D = 2.3 \pm 0.3 \mu\text{M}$), but we could not determine whether it contained one or two functional peptide-binding sites (Figure 1B, right). We concluded that there are four, or possibly five, functional substrate-binding sites in the TNKS2 N terminus, in agreement with previous work (Seimiya et al., 2004; Seimiya and Smith, 2002).

Structural Analysis of an ARC and Prediction of Its Peptide-Binding Pocket

A boundary-optimized form of ARC4 was crystallized in apo- and 3BP2 peptide-bound forms, and X-ray crystal structures were determined by molecular replacement (see Table S1 available online for data collection and refinement statistics and Extended Experimental Procedures for details). ARC4 consists of a stack of five ankyrin repeats (Figures 1A and 1C), the central three of which (repeats 2–4) possess a characteristic ankyrin repeat architecture (loop–helix1–loop–helix2–loop). Ankyrin repeat 1 is cryptic in nature, with an atypically long helix1, whereas ankyrin repeat 5 has an atypically short helix1. As in other ankyrin repeat proteins, the terminal loops of each ankyrin repeat extend from the pair of hydrophobically packed, antiparallel helices, forming a sheet of β hairpins (Figure 1C). The continuous succession of ankyrin repeats generates a concave surface at the “front” face of the ARC (according to the view shown in Figure 1C).

Projection of residue conservation for ARCs 1–5 onto the ARC4 surface identified a conserved concave surface in the central region of the ARC (Figure 1D, left). Omission of ARC3 from the conservation analysis increased the degree of sequence conservation on the mapped surface (Figure 1D, right). As demonstrated below, the conserved region corresponds to the peptide-binding pocket, and its lack of conservation in ARC3 reflects the fact that ARC3 does not bind substrates.

Overview of the ARC4:3BP2 TNKS-Binding Motif Complex

In the ARC4:3BP2 peptide complex, the 3BP2 peptide of sequence LPHLQRSPPDGQSFRRS (core binding motif as described below is underlined) adopts an extended conformation and binds to a pocket situated centrally to the ARC, perpendicular to its longitudinal axis (Figure 2A). The pocket is located within the peptide-binding site predicted by sequence conservation and is entirely formed by the central three ankyrin repeats (2–4). As such, the flanking ankyrin repeats, 1 and 5, likely serve as structural caps (Figure S1B), as noted for other ankyrin repeat proteins (reviewed in Forrer et al., 2003; reviewed in Li et al., 2006). The anticipated shortness of linkers (0 to 9 residues) between autonomously folding ARCs (Figure 1C) likely limits the flexibility of an otherwise beads-on-a-string architecture.

Description of the TNKS2 ARC4:3BP2 TNKS-Binding Motif Interface

The 8-residue 3BP2 core sequence RSPPDGQS is engaged by four groups of peptide-coordinating residues in ARC4: arginine at position 1 ($R415^{3BP2}$) of the 3BP2 peptide is engaged by an “arginine cradle,” glycine at position 6 ($G420^{3BP2}$) is engaged by an “aromatic glycine sandwich,” the central residues between the arginine and glycine (positions 2–5) are engaged by a “central patch,” and two C-terminal residues (positions 7 and 8) are engaged by two C-terminal contact residues (Figures 2B and S2A).

Four TNKS2 side chains contribute to the extensive arginine cradle (Figure 2C): $W591^{TNKS2}$ packs against the nonpolar portion of the $R415^{3BP2}$ side chain. $F593^{TNKS2}$ establishes a cation- π interaction, whereas $E598^{TNKS2}$ and $D589^{TNKS2}$ form salt bridges with the guanidinium group of $R415^{3BP2}$.

Two TNKS2 tyrosines, $Y536^{TNKS2}$ and $Y569^{TNKS2}$, form the aromatic glycine sandwich (Figure 2C). The absence of a side chain at $G420^{3BP2}$ allows for a close approach of the peptide main chain to the ARC such that a hydrogen bond is formed between the main-chain carbonyl of $G535^{TNKS2}$ and the main-chain amide of $Q421^{3BP2}$ at position 7.

Nine residues comprise the central patch of the peptide-binding pocket (Figures 2C and S2A). $R525^{TNKS2}$ forms a hydrogen bond with the side chain of $S416^{3BP2}$ at position 2, the main-chain carbonyl groups of $S416^{3BP2}$ and $P417^{3BP2}$ at positions 2 and 3, respectively, and, via a water molecule, the main-chain amino and carbonyl groups of $S416^{3BP2}$. TNKS2 makes no contact with the solvent-exposed side chain of $P417^{3BP2}$; however, $P417^{3BP2}$ directs the peptide backbone toward a subpocket into which the side chain of $P418^{3BP2}$ at position 4 inserts. Within this subpocket, $L560^{TNKS2}$ confers hydrophobic contact, whereas the main chains of $N565^{TNKS2}$ and $H564^{TNKS2}$ and the side chain of $S568^{TNKS2}$ confer Van der Waals contact. The side chain of $Y569^{TNKS2}$ additionally coordinates $P418^{3BP2}$ through a hydrogen bond to the main-chain carbonyl group.

A subpocket of the central patch defined by $F532^{TNKS2}$, $D521^{TNKS2}$, $S527^{TNKS2}$, and $R525^{TNKS2}$ accommodates the $D419^{3BP2}$ side chain at position 5. $S527^{TNKS2}$ forms a direct hydrogen bond with $D419^{3BP2}$, whereas $D521^{TNKS2}$ and the main-chain amino group of $R525^{TNKS2}$ and $S527^{TNKS2}$ form

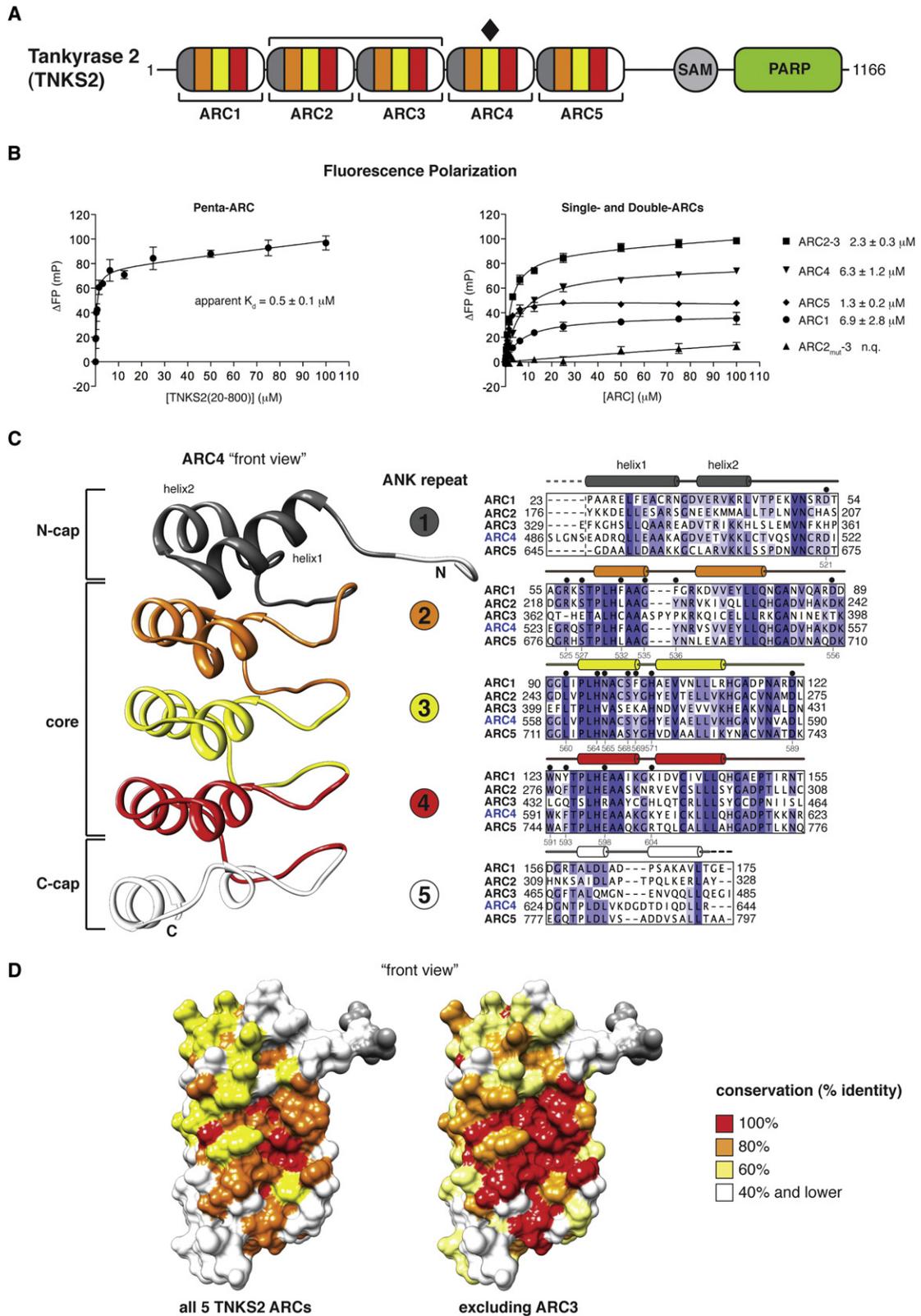


Figure 1. Functional Characterization of an Ankyrin Repeat Cluster

(A) Domain organization of human TNKS2. The ankyrin repeats group into five ARCs of five ankyrin repeats each. ARC4 was crystallized (indicated by \blacklozenge). SAM, sterile-alpha motif domain; PARP, poly(ADP-ribose)polymerase catalytic domain.

Table 1. Previously Identified and Proposed Tankyrase Binders

Protein Name	Function	Reference(s)
Proven Tankyrase Binders/Substrates with Identified TNKS-Binding Motifs		
TERF1/TRF1	telomere binder and negative regulator of telomere length	(Cook et al., 2002; Kaminker et al., 2001; Sbodio and Chi, 2002; Sbodio et al., 2002; Smith et al., 1998)
AXIN1/2	tumor suppressor in the Wnt signaling pathway	(Huang et al., 2009)
LNPEP/IRAP	integral-membrane aminopeptidase	(Chi and Lodish, 2000; Sbodio and Chi, 2002; Sbodio et al., 2002)
NUMA1	mitotic spindle regulator	(Chang et al., 2005a, 2005b; Sbodio and Chi, 2002)
MCL1	apoptosis regulator	(Bae et al., 2003)
TNKS1BP1/TAB182	largely uncharacterized Tankyrase binder	(Sbodio and Chi, 2002; Seimiya and Smith, 2002)
EBNA1	protein of Epstein Barr virus; functions in DNA replication and transcriptional regulation	(Deng et al., 2002, 2005)
FNBP1/FBP17	cytoskeletal regulator	(Fuchs et al., 2003)
BLZF1	regulator of protein transport; precise function unknown	(Zhang et al., 2011)
CASC3	regulator mRNA biogenesis	(Zhang et al., 2011)
3BP2	signaling adaptor protein; mutated in cherubism	(Levaot et al., 2011)
Proposed Tankyrase Interactors with Putative TNKS-Binding Motifs		
GRB14	signaling adaptor	(Lyons et al., 2001)
HOXB2	homeobox transcription factor	(Sbodio and Chi, 2002)
PPP1R12C/MBS85	polymorphic variant of the protein phosphatase 1 regulatory subunit 12C	(Sbodio and Chi, 2002)
TAX1BP1	apoptosis regulator	(Sbodio and Chi, 2002)
CACNA1S	alpha 1S subunit of the voltage-dependent L-type calcium channel	(Sbodio and Chi, 2002)
USP25	ubiquitin-specific protease	(Sbodio and Chi, 2002)

Proposed Tankyrase binders were identified by yeast-two-hybrid analysis; GRB14 was confirmed as an interactor by coimmunoprecipitation.

water-bridged hydrogen bonds to the D419^{3BP2} side chain. Similarly, S527^{TNKS2}, D556^{TNKS2}, and N565^{TNKS2} collaborate in a water-bridged hydrogen-bonding network to the main-chain amino group of D419^{3BP2}. The N565^{TNKS2} side chain also hydrogen bonds with the main-chain carbonyl group of D419^{3BP2}. Lastly, a water molecule bridges the Y536^{TNKS2} and D419^{3BP2} side chains.

Although the residues preceding the arginine at position 1 are poorly ordered and unlikely to contribute critical contacts, two C-terminally located amino acids of the peptide are directly coordinated by ARC4 side chains: the side chain of H571^{TNKS2} hydrogen bonds with the main-chain carbonyl group of Q421^{3BP2} at position 7, whereas the side chain of K604^{TNKS2} hydrogen bonds with the side chain of S422^{3BP2} at position 8 (Figure 2C). The structural order of the three C-terminal peptide residues is largely attributable to crystal contacts with no major direct contacts to the corresponding ARC.

The four ARC4:3BP2 complexes present in the asymmetric unit are highly similar (Figure S2B). A comparison of the apo- and peptide-bound forms of ARC4 demonstrates that the peptide-binding pocket is fully formed prior to peptide coordination (Figure S2C).

Validation of the ARC4-Peptide Interaction Surface by Mutagenesis

To biochemically validate the ARC4:3BP2 structure, we introduced mutations into the peptide-binding pocket predicted to impair binding while preserving the ankyrin repeat fold. We produced and purified 14 well-folded ARC4 mutant derivatives (Figures 3A and S3) and performed fluorescence polarization (FP) assays to test each for substrate (3BP2) binding (Figure 3B). As noted above, wild-type (WT) ARC4 bound the 3BP2 peptide with micromolar affinity ($K_D = 6.4 \pm 0.6 \mu\text{M}$) (Figure 3B).

(B) FP binding assay for the full ankyrin repeat region of TNKS2 (left) or the indicated single- and double-ARC constructs (right). ARCs 2 and 3 were analyzed as an ARC2-3 double-cluster. ARC2_{mut-3} specifies an L245W mutation that disrupts 3BP2 peptide binding to ARC2 (see Figure 3). K_D are indicated where derivable by nonlinear regression using a one-site total binding model; n.q., not quantifiable. n = 3 separate experiments performed in technical duplicate; error bars, SEM; K_D error values, standard error of the fit.

(C) Left, ribbon representation of apo-ARC4 with individually colored ankyrin repeats. Right, sequence alignment of all five TNKS2 ARCs colored by sequence identity with secondary structure elements for ARC4 above. ARC4 residues contacting the 3BP2 peptide (see Figure 2) are labeled with ● and numbered.

(D) Side-chain conservation across the TNKS2 ARCs mapped onto the ARC4 surface. The corresponding sequence alignment is shown in (C). Left, surface conservation across all five TNKS2 ARCs; right, surface conservation upon omission of ARC3.

See Figure S1 for domain architecture of substrates analyzed in this study and capping features of ARC4.

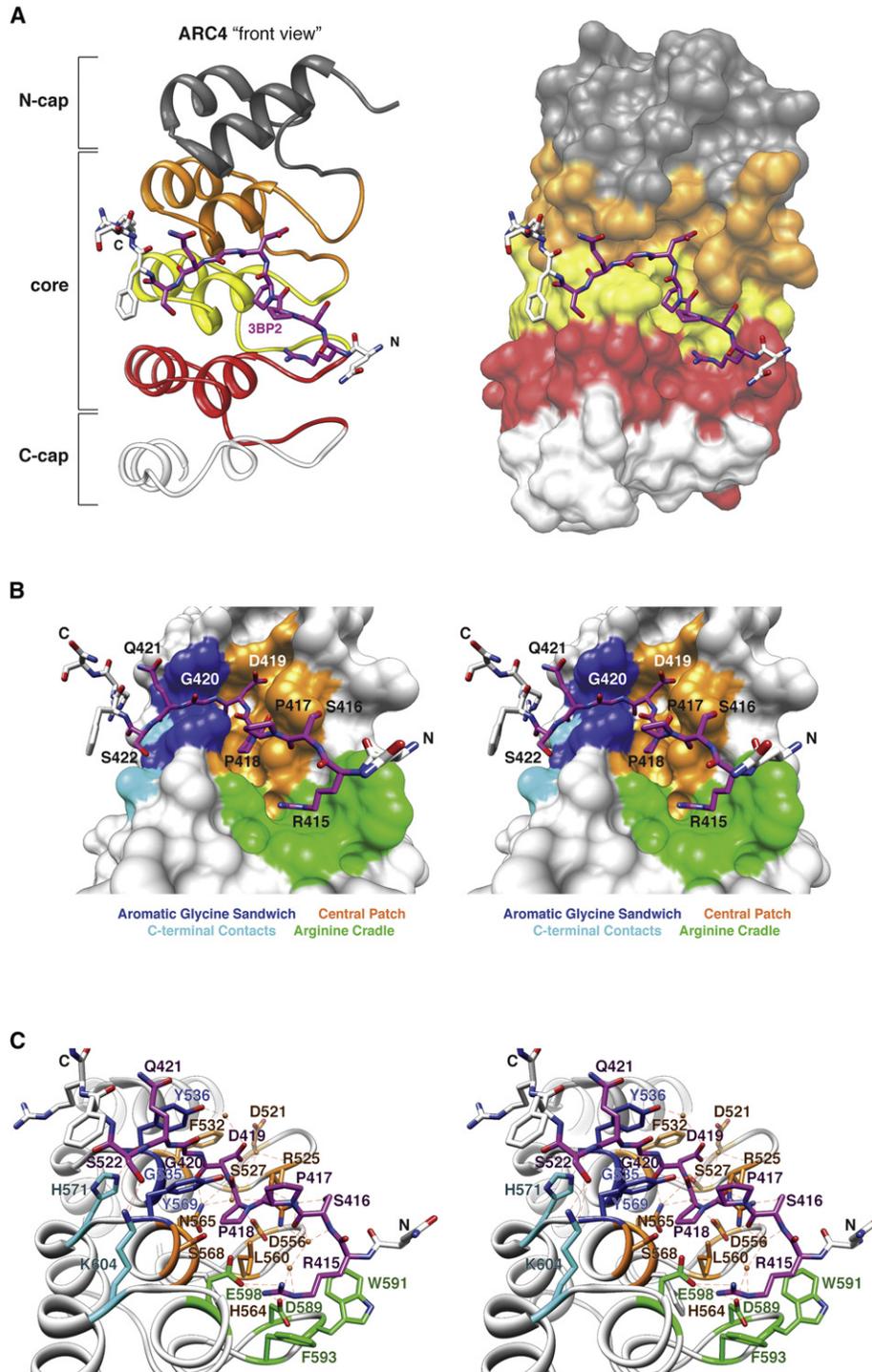


Figure 2. Structural Representation of the ARC4:3BP2 Targeting Peptide Complex

(A) Ribbon (left) and surface (right) representations of ARC4 colored as in Figure 1C with the 3BP2 peptide shown in stick representation. The 8 amino acids of the 3BP2 peptide that constitute the TNKS-binding motif (RSPDGQS) are shown in magenta with colored heteroatoms (oxygen, red; nitrogen, blue); remaining peptide residues are shown in white with colored heteroatoms.

(B) Stereo surface representation of the ARC4:3BP2 complex with the 3BP2 peptide shown as in (A). Key ARC4 residues participating in 3BP2 peptide coordination are color-coded; 3BP2 residues are labeled.

(C) Stereo ribbon representation of the peptide-binding pocket with the 3BP2 peptide and relevant contact residues in ARC4 shown as sticks. Coloring is as in (B). Dashed lines connect selected polar neighbors.

See Figure S2 for additional structural representations. See also Table S1.

Mutation of arginine cradle residues (W591A, F593A, E598A, WFE591/593/598AAA, D589T) strongly impaired peptide binding to ARC4 so that a K_D could not be determined (Figure 3B). Mutation of Y536, one of the glycine sandwich tyrosines (Y536A), had a weak effect on peptide binding ($K_D = 17.7 \pm 3.5 \mu\text{M}$), whereas mutation of Y569 (Y569A, YY536/569AA) abolished measurable peptide binding. TNKS2 residues coordinating the central positions 2–5 of the 3BP2 peptide also proved critical for peptide binding. An N565A mutation strongly reduced peptide binding ($K_D = 30.2 \pm 8.9 \mu\text{M}$), whereas the R525A, R525D, S527A, and a pocket-filling L560W mutation each reduced peptide binding below quantifiability. These results validate our structural observations and establish a set of residues critical for 3BP2 targeting peptide engagement by an ARC. Notably, despite the conservation of K604^{TNKS2} as a basic residue (see Figure 1C), its mutation to alanine had no measurable effect on 3BP2 peptide binding. As discussed below, K604 makes a more important contribution to binding of targeting peptides from other Tankyrase substrates.

Based on conservation of residues within the validated substrate-binding surface of ARC4, we can rationalize the peptide-binding function of the different ARCs: like ARC4, ARCs 1, 2, and 5 possess the required infrastructure for substrate binding, whereas ARC3 does not (see ARC alignment in Figure 1C). We therefore anticipated that the ARC2-3 construct discussed above contains only a single peptide-binding site formed by ARC2. Introduction of an L245W mutation into ARC2 (equivalent to the L560W mutation in ARC4) in the context of ARC2-3 (ARC2_{mut}-3) abolished ARC2-3 binding to the 3BP2 peptide, confirming this prediction (Figure 1B). The TNKS2 N terminus therefore contains four functional 3BP2 peptide-binding sites. If one considers the presence of four sites within the TNKS2 N terminus that contribute similarly to peptide binding, the overall affinity of the penta-ARC construct (Figure 1B, left) is in agreement with the affinities of the individual ARCs (Figure 1B, right).

ARCs 1, 2, 4, and 5 Display Redundancy for 3BP2 Binding

The binding behavior of ARCs 1, 2, 4, and 5 toward 3BP2 (Figure 1B) suggests that these ARCs may be functionally redundant for 3BP2 recognition. To test this prediction, we serially introduced inactivating L-to-W mutations (equivalent to the L560W mutation in ARC4) into each of the ARCs of TNKS2 in the context of the full ankyrin repeat region spanning all five ARCs (TNKS2(20–800); Figure 3C). Bacterially expressed His₆-GST-TNKS2(20–800) efficiently recruited purified, recombinant full-length His₆-3BP2 in a GST pull-down assay, whereas neither His₆-GST alone nor the corresponding TNKS2 mutant with five mutated ARCs (xxxxx) bound 3BP2 (Figure 3D). In agreement with our prediction, all five single-ARC mutants (x2345, 1x345, 12x45, 123x5, 1234x) showed robust binding to 3BP2, demonstrating that no single ARC is critical for 3BP2 binding by Tankyrase (Figure 3D).

To determine whether any of the five ARCs is sufficient for 3BP2 recruitment in the context of the full TNKS2 N terminus, we mutated four out of five ARCs, leaving only a single ARC intact (1xxxx, x2xxx, xx3xx, xxx4x, xxx5). As expected, ARC3 alone (xx3xx) did not bind 3BP2, whereas ARCs 1, 2, 4, and 5

alone did bind, but with reduced apparent affinities (Figure 3D). Reduced binding might reflect the simple reduction of the number of available binding sites and/or the loss of avidity effects.

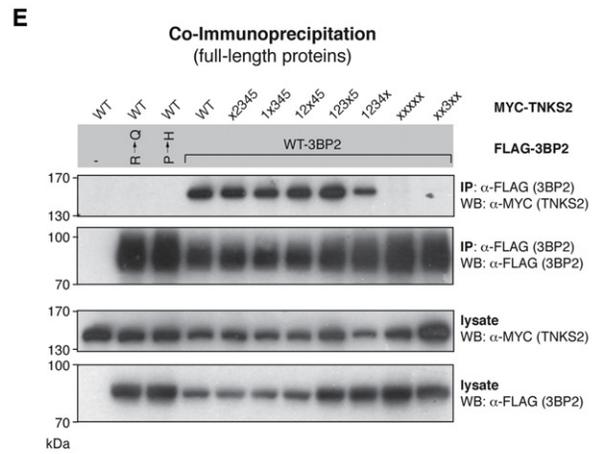
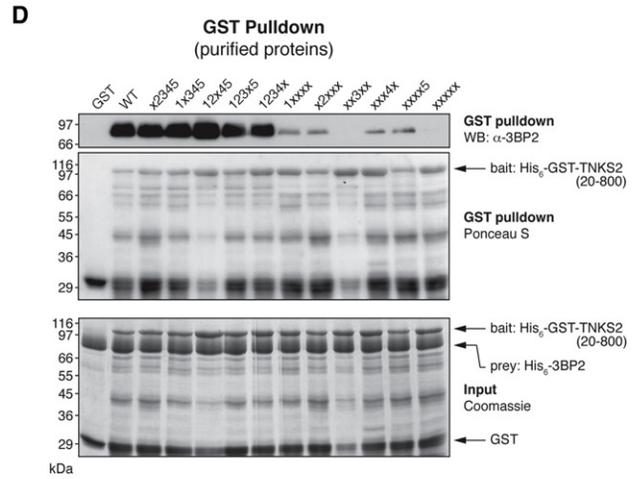
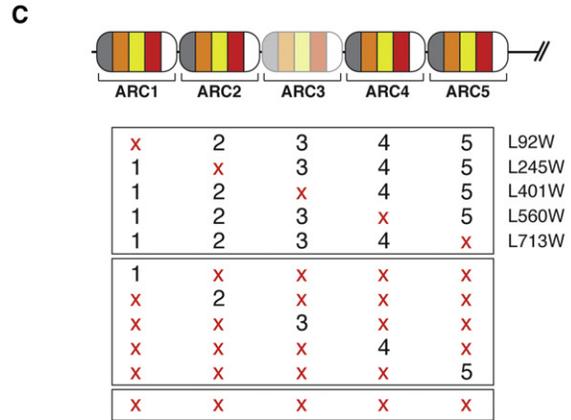
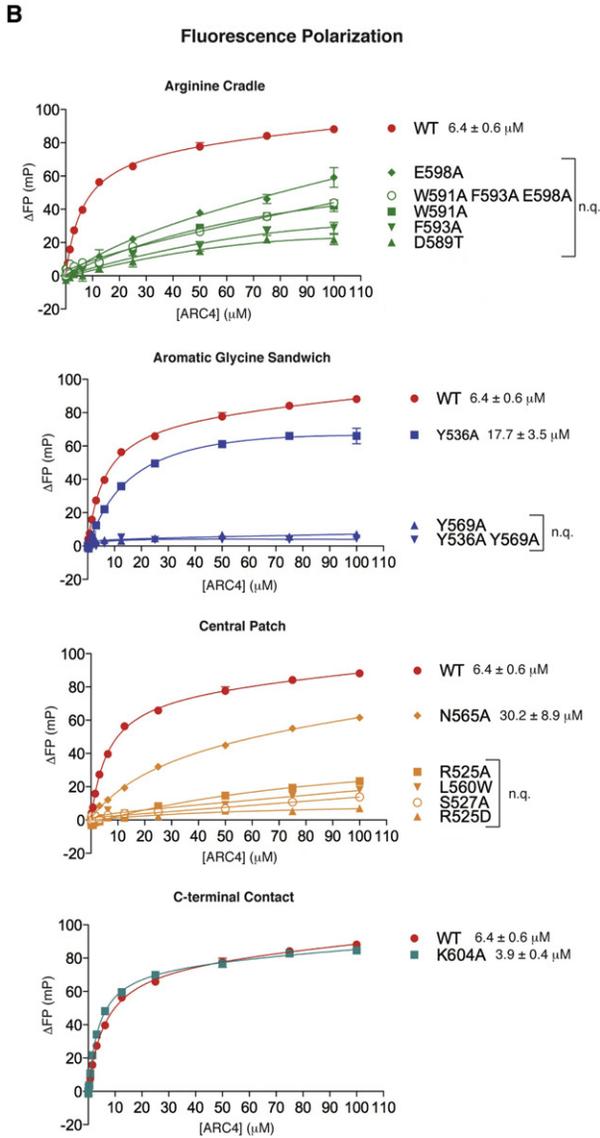
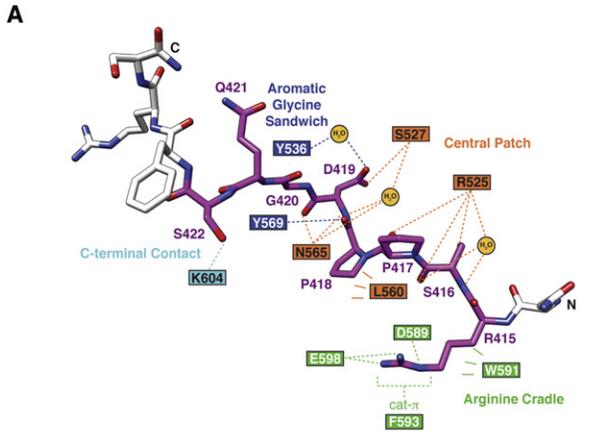
To extend our binding analysis to full-length Tankyrase, we coexpressed WT MYC-TNKS2 or a set of mutant derivatives with either WT FLAG-3BP2 or cherubism derivatives (murine R413Q and P416H) of FLAG-3BP2 in mammalian (HEK293T) cells (Figure 3E). Following immunoprecipitation of WT FLAG-3BP2, we detected robust binding of TNKS2 and any of its single-ARC mutants but not TNKS2 xxxxx or xx3xx (Figure 3E). The cherubic FLAG-3BP2 variants (murine R413Q or P416H) did not recruit Tankyrase detectably. These observations support a degree of functional redundancy across ARCs 1, 2, 4, and 5 for 3BP2 recruitment. A complete disabling of substrate recognition by Tankyrase in cells thus may require the inactivation of all four functional ARCs.

ARCs 1, 2, 4, and 5 Display Similar Substrate Recognition Abilities, and TNKS-Binding Motifs Are Octapeptides

We expanded our analysis from 3BP2 to other Tankyrase targets. Using the FP binding assay, we assessed ARC4 binding to proven or predicted targeting peptides from known or proposed Tankyrase substrates (3BP2, NUMA1, HOXB2, LNPEP, FNBP1, TNKS1BP1, a polymorphic variant of PPP1R12C, CACNA1S, MCL1, USP25, TERF1, AXIN1, TAX1BP1, EBNA1, and GRB14; see Table 1) (reviewed in Hsiao and Smith, 2008). All peptides except for those from TAX1BP1, EBNA1, GRB14, and nonpolymorphic PPP1R12C bound ARC4 with comparable (micromolar) affinities (Figures 4A and 4B).

Binding analysis of other ARCs to the same panel of peptides revealed similar peptide-discriminating abilities with only slight differences in overall binding affinities. On average, ARC2 (i.e., ARC2-3) and ARC5 bound TNKS-binding motif peptides more strongly than ARC1 or ARC4 (Figures S4A and S5 and Table S2). Among ARCs 1, 2, 4, and 5, ARC1 was the weakest peptide binder, which may be rationalized by two phenylalanines in place of the tyrosines in the aromatic glycine sandwich (see sequence alignment in Figure 1C). This would abolish a hydrogen bond between the residue equivalent to Y569 in ARC4 with the substrate main chain (see Figure 2C). Based on the hyperconservation of peptide-binding surface residues within each ARC across different species, we reason that the relative binding strength and function of each ARC will be conserved across life forms with Tankyrases that possess a penta-ARC organization (Figure S4B).

TNKS-binding motifs are typically referred to as “hexapeptides,” and the limited overall sequence conservation outside a 6-amino-acid stretch suggested that the hexapeptide contains the chief determinants for Tankyrase binding (Figure 4A). To revisit the actual extent of the TNKS-binding motif, we performed an alanine scan across a 9-amino-acid range for two prominent TNKS-binding motifs, 3BP2 and AXIN1. We then tested the effect of each mutation on ARC4 association in the FP assay (Figure 4C). The WT 3BP2 peptide bound ARC4 with a K_D of $3.8 \pm 0.4 \mu\text{M}$. Mutation of the residues preceding the invariant arginine at position 1, Q-1A, and mutation of residues at



positions 7 and 8 had little effect on ARC4 binding (K_D values of $4.6 \pm 1.1 \mu\text{M}$, $3.9 \pm 0.9 \mu\text{M}$, and $3.8 \pm 0.7 \mu\text{M}$, respectively). Strikingly, mutation of the invariantly conserved R1 and G6 positions to alanine (R415A and G420A) abolished quantifiable binding of ARC4 to the peptide, similar to the GST pull-down results obtained with the TNKS-binding motif peptide of LNPEP (Sbodio and Chi, 2002). Mutation of the less conserved peptide residues at positions 2–4 (S416A, P417A, P418) only had small effects (K_D values of $5.7 \pm 0.8 \mu\text{M}$, $7.3 \pm 1.1 \mu\text{M}$, $8.8 \pm 1.5 \mu\text{M}$, respectively). The D419A mutation at position 5 resulted in a larger decrease in peptide affinity ($K_D = 70.0 \pm 39.5 \mu\text{M}$), indicating a more critical contribution of aspartate at this site. The binding determinants for 3BP2 therefore map to a hexapeptide stretch. In the case of AXIN1, however, binding to ARC4 was sensitive to mutations across an octapeptide region (positions 1 to 8; Figure 4C). As observed for 3BP2, the arginine and glycine residues at positions 1 and 6 were absolutely required for binding. Mutation of residues at positions 2, 3, 4, 5, and 8 resulted in less dramatic decreases in affinity, whereas mutation of the residue at position 7 had no negative effect on binding. Unlike for 3BP2, mutation of residue 8 weakened the ARC4-AXIN1 interaction about 2.3-fold (from $6.5 \pm 1.4 \mu\text{M}$ to $15.1 \pm 4.0 \mu\text{M}$).

In summary, the four functional ARCs display similar binding properties, and ARC-binding determinants map to a TNKS-binding motif of 8 amino acids. The highly conserved arginine and glycine residues at positions 1 and 6 of the binding motif are critical anchor residues essential for peptide binding to ARC4. The limited sampling of the alanine scan, however, did not fully explain the remarkable diversity of residues at other positions of functional TNKS-binding motifs (explored below).

A Structural Comparison of Multiple TNKS-Binding Motif Peptides

The variability of the central peptide residues in various Tankyrase ligands is striking (see Figure 4A). For example, the TNKS-binding motif peptide of MCL1 contains three consecutive proline residues (RPPPIGAE), whereas that of TERF1 lacks prolines altogether (RGCADGRD). To elucidate whether such diverse ligands employ the same binding mode for ARC interaction, we crystallized ARC4 bound to peptides from TERF1, MCL1, LNPEP, NUMA1, and FNBP1 (Tables S1 and S6). All peptides bound in a similar configuration observed for the 3BP2-derived peptide (Figure 4D); notably, peptides without prolines adopted a geometry similar to proline-containing peptides. Thus, despite considerable variation in interaction motif sequence, all substrates investigated employ a common ARC-binding mode.

Unfortunately, we were unable to obtain crystals for ARC4 bound to AXIN1/2, one of the most prominent Tankyrase targets (Huang et al., 2009; Zhang et al., 2011). To probe whether AXIN1 employs an ARC-binding mode similar to that of the other targets analyzed, we investigated binding of the AXIN1-targeting peptide to the 14 ARC4 pocket mutants described above. Most of the mutations that impaired binding to the 3BP2 substrate peptide also perturbed binding to the AXIN1 peptide, arguing that a common binding mode is employed (compare Figures 3B and S4C). However, binding of the AXIN1 peptide was strikingly sensitive to mutation of the C-terminal contact residue K604^{TNKS2} to alanine (Figures 3B and S4C). Position 8 in the TNKS-binding motifs of AXIN1 and AXIN2 (and MCL1) is occupied by a conserved glutamate residue (versus a serine residue in 3BP2), which we predict forms a potent stabilizing salt bridge interaction (versus weaker hydrogen bond) with K604^{TNKS2}, as observed in the ARC4:MCL1 complex (Figures 4A and 4D). Furthermore, in comparison to 3BP2, the ARC4-AXIN1 interaction was less sensitive to the R525A^{TNKS2} mutation in the central patch (Figures 3B and S4C). As R525^{TNKS2} forms a stabilizing hydrogen bond to the polar side chain of substrate position 2 (S416 in the case of 3BP2), the presence of a nonpolar proline in position 2 of AXIN would render R525 more dispensable.

Thus, although the overall binding mode of substrates to an ARC of Tankyrase is highly conserved, the relative importance of component interactions can differ markedly for different substrates.

Defining Sequence Rules that Govern Substrate Recognition by Tankyrase

To fully understand the binding determinants within the TNKS-binding motifs in Tankyrase substrates, we designed a peptide library based on the TNKS-binding motif of 3BP2. We serially exchanged each residue of the 8-amino-acid stretch RSPPDGQS in the fluorescently labeled peptide QRSPPDGQS for each of the 20 standard amino acids, obtaining 153 unique peptides. We then used the FP binding assay to test their binding to ARC4. The results (Figures 5A and S6A and Table S3) can be rationalized by our ARC:peptide structures.

As expected, peptide recognition by ARC4 critically requires arginine and glycine residues at positions 1 and 6, respectively. The arginine cradle is specifically tuned for arginine with multiple residues collaborating in arginine coordination (see Figures 2B and 2C). The phi and psi peptide bond angles of glycine at position 6 are unique to glycine and cannot be adopted by any other

Figure 3. Probing the ARC4 Interaction Surface by Site-Directed Mutagenesis and Analyzing the Contribution of ARCs to 3BP2 Binding

- (A) Schematic representation of the ARC4:3BP2 peptide interaction. ARC4 residues targeted by site-directed mutagenesis are highlighted.
- (B) FP binding assays comparing WT ARC4 and mutants. K_D are indicated where derivable by nonlinear regression using a one-site total binding model; n.q., not quantifiable. n = 3 separate experiments performed in technical duplicate; error bars, SEM; K_D error values, standard error of the fit.
- (C) Outline of ARC mutations.
- (D) Equivalency of ARCs 1, 2, 4, and 5 for 3BP2 binding. GST-fusion proteins of the TNKS2 N terminus (as shown in C) were bacterially produced and probed for their ability to recruit purified full-length murine His₆-3BP2 in a GST pull-down assay. Bottom panel, Coomassie-stained gel of input samples; middle panel, PonceauS-stained membrane of bound material; top panel, western blot to detect bound 3BP2.
- (E) WT and the indicated mutant derivatives of full-length MYC-TNKS2 were coexpressed with either WT or cherubic (murine R413Q or P416H) FLAG-3BP2 in HEK293T cells. FLAG-3BP2 was immunoprecipitated. Bottom panels, lysates showing coexpression; top panels, immunoprecipitates probed with anti-FLAG (3BP2) or anti-MYC (TNKS2) antibodies. See Figure S3 for additional information.

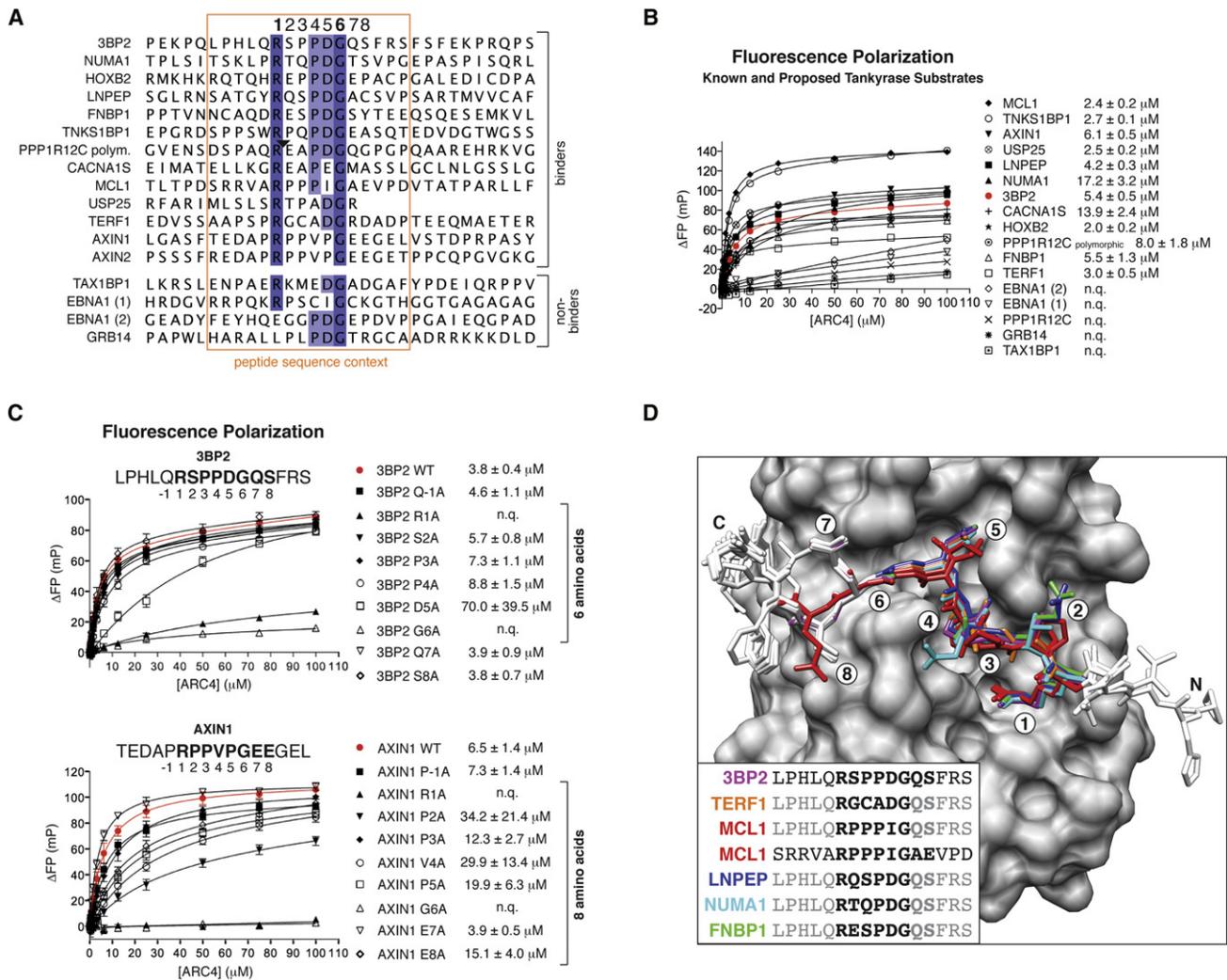


Figure 4. Determinants in TNKS-Binding Motifs for ARC4 Binding

(A) Sequence alignment of TNKS-binding motifs from previously known and proposed human Tankyrase binders; see [Extended Experimental Procedures](#) for accession numbers. Residue numbering within the TNKS-binding motif, as used throughout the analysis, is defined above the alignment. The arrow indicates the position of an additional alanine in the nonpolymorphic variant of PPP1R12C.

(B) FP binding analysis of peptides corresponding to the TNKS-binding motifs shown in (A). K_D are indicated where derivable by nonlinear regression using a one-site total binding model; n.q., not quantifiable. $n = 3$ separate experiments performed in technical duplicate; error bars, SEM; K_D error values, standard error of the fit.

(C) FP binding assay analyzing the contribution of the 3BP2 and AXIN1 targeting peptide residues to ARC4 binding. Each one of nine residues was exchanged for alanine. K_D are indicated where derivable by nonlinear regression using a one-site total binding model; n.q., not quantifiable. $n = 3$ separate experiments performed in technical duplicate; error bars, SEM; K_D error values, standard error of the fit.

(D) Diverse TNKS-binding motifs employ an identical ARC4:peptide binding mode. The sequences of the peptides and peptide chimeras are shown. See [Figures S4 and S5](#) for additional information. See also [Tables S1 and S2](#).

amino acid, explaining why even alanine cannot substitute at position 6 ([Figure S6B](#)).

Positions 2 and 3 of the TNKS-binding motif display a broad amino acid tolerance. However position 3 poorly tolerated phenylalanine, and the rank order of the individual residues was different between the two positions. Consistent with the broad tolerance, the side chains at positions 2 and 3 of all peptides in the crystal structures introduced above ([Figure 4D](#)) always face solvent. We cannot rationalize why phenylalanine is unfavorable at position 3.

The side chain at peptide position 4 is restricted to the small and/or hydrophobic amino acids proline, glycine, alanine, and cysteine. In the crystal structures, side chains at position 4 are always presented toward the peptide-binding pocket on the ARC. The space restraints imposed by the position-4 subpocket, and its hydrophobic nature, explain why only small and hydrophobic amino acids are tolerated at this position.

Aspartic acid is strongly preferred at position 5 but can be substituted to a far lesser degree by glutamic acid, valine, glutamine, tyrosine, isoleucine, and cysteine. In all but one crystal

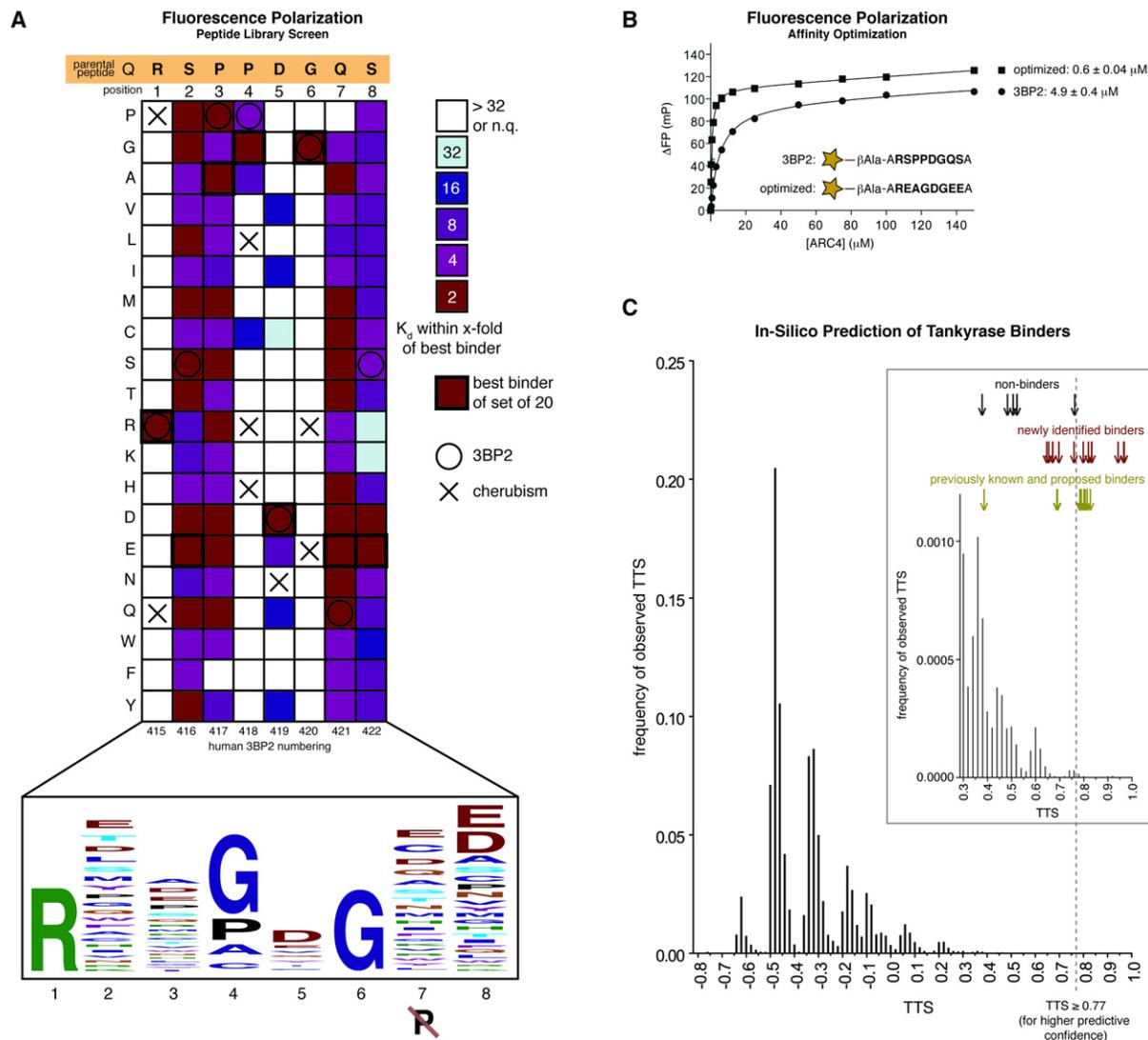


Figure 5. Defining a Rule-Based Consensus for TNKS-Binding Motifs to Predict Tankyrase Targets

(A) FP-based peptide library screen. Amino acids within the 3BP2-derived octapeptide RSPPDGQS were serially exchanged as indicated. The peptide library was subjected to an FP binding assay with ARC4. To obtain K_D , data ($n = 1$ technical duplicate) were analyzed by nonlinear regression using a one-site total binding model. Colors reflect K_D relative to the best binder in each group of 20 peptides per position scanned; n.q., not quantifiable. \circ indicates native 3BP2 residues; \times indicates cherubism mutations (Lietman et al., 2006; Ueki et al., 2001). See Figure S6 for FP titration graphs and Table S3 for K_D . Bottom, sequence logo representing the sequence rules, derived by dividing the inverse of the K_D values by the number of allowed amino acids per position and scaling the letters accordingly.

(B) Affinity optimization of the 3BP2 peptide based on the data shown in (A). The indicated peptides were analyzed by fluorescence polarization. K_D were derived by nonlinear regression using a one-site total binding model. $n = 3$ separate experiments performed in technical duplicate; error bars, SEM; K_D error values, standard error of the fit.

(C) In silico prediction of human Tankyrase binders based on a position-specific scoring matrix developed from the data shown in (A). Frequency of TTS is plotted against the corresponding TTS. The inset shows the high-TTS range of the same plot with a smaller TTS frequency scale. Arrows indicate TTS of all peptides tested by FP and present in the UniProt database for the human proteome.

See Figure S6 for additional information. See also Tables S3, S4, and S7.

structure, peptide position 5 is occupied by the highly preferred aspartate, as in 3BP2, establishing hydrogen bonds to S527^{TNKS2}. The crystal structures of the ARC4:MCL1 complex provide an example wherein position 5 is adopted by a hydrophobic amino acid (isoleucine). A phenylalanine, F532^{TNKS2}, and the hydrophobic portion of R525^{TNKS2} render the position-5 subpocket sufficiently hydrophobic to accommodate the less

optimal nonpolar residues identified in the peptide screen (see Figure 2C).

Position 7 of the TNKS-binding motif can be occupied by a wide range of amino acids. Surprisingly, proline was not allowed at this position, presenting a case for negative selection by the ARC against one particular amino acid.

Position 8 tolerates all 20 amino acids; however, it presents a substantial preference for acidic residues (glutamate, aspartate) and least prefers basic residues (arginine, lysine). This is in full agreement with acidic residues in position 8 forming a salt bridge with K604^{TNKS2}, as observed in the ARC4:MCL1 crystal structure (Figure 4D).

In summary, the findings of the positional scanning experiment support the binding mode we observe in the ARC4:peptide complex structures and can explain the ARC-binding behavior of most of the 16 confirmed or putative TNKS-binding motif peptides analyzed (see Figure 4A). The experiment also demonstrates that all characterized cherubism mutations (R415P, R415Q, P418L, P418R, P418H, D419N, G420E, G420R) act by abolishing binding of the 3BP2 TNKS-binding motif to ARCs (Figure 5) (Lietman et al., 2006; Ueki et al., 2001). Our peptide library screen suggests a “superconsensus” for the 8-amino-acid TNKS-binding motif. To test whether this is the case, we generated a pair of 10-amino-acid peptides, the first one containing the TNKS-binding motif of 3BP2 and the second one an optimized TNKS-binding motif with the preferred amino acid at each position (REAGDGEE). Whereas the peptide corresponding to 3BP2 bound with an affinity of $4.9 \pm 0.4 \mu\text{M}$, the sequence-optimized peptide bound almost an order more tightly, with an affinity of $0.6 \pm 0.04 \mu\text{M}$ (Figure 5B). Deviations from the superconsensus at any position, individually or jointly, may allow the affinity of substrate-targeting peptides to be differentially tuned, as appears to be the case. Isoleucine at position 5 conferred weak binding ($K_D = 72.4 \pm 31.6 \mu\text{M}$) in the screen, but the TNKS-binding motif peptide of MCL1, which contains isoleucine at position 5, bound the ARC robustly ($K_D = 2.4 \pm 0.2 \mu\text{M}$). Similarly, valine at position 4 or proline at position 5 were disallowed/unfavorable in our screen, but AXIN1, another robust ARC binder ($K_D = 6.1 \pm 0.5 \mu\text{M}$), contains exactly these residues at positions 4 and 5. Both of these TNKS-binding motifs contain the stabilizing glutamate residue at position 8, which can form a salt bridge interaction with K604^{TNKS2} that likely offsets the negative impact of suboptimal amino acids at positions 4 and 5. The targeting peptides of both MCL1 and AXIN1 are indeed highly sensitive to the K604A mutation in ARC4 (Figure S4C and Table S2). Furthermore, although introduction of the AXIN-like residues valine and proline at positions 4 and 5, respectively, into the TNKS-binding motif of 3BP2 abolished measurable binding, binding was partially restored when either glutamate or aspartate, but not glutamine or asparagine, was present at position 8 (Figure S6C).

An In Silico Prediction of Tankyrase Binders

To identify candidate Tankyrase binders, we performed an in silico search of the human proteome, using a position-specific scoring matrix (PSSM) approach based on the affinity data obtained in the peptide library screen (Yaffe et al., 2001; see Experimental Procedures for details). We calculated a Tankyrase-targeting score (TTS) of minimally -0.817 (no match) to maximally 1 (best match) for each peptide in the UniProt database (octapeptide scan with additional scoring of C-terminal hepta- and hexapeptides; see http://research.lunenfeld.ca/sicheri/files/file/PSSM_raw_list.txt.zip for a list of all peptides scored). The vast majority (10,420,456; 93.7%) of the

11,124,194 peptides scanned obtained a low TTS of <0 (Figure 5C). Eleven binders of the sixteen previously identified or predicted Tankyrase targets studied by FP (see Figure 4A) were enriched in the region of high TTS (Figure 5C). As points of reference, the prominent Tankyrase targets AXIN1 and AXIN2 displayed a low TTS of 0.385, due to suboptimal amino acids at positions 4 and 5 that are disallowed/unfavorable in the context of 3BP2 (see Figure 5A). All other previously identified and proposed Tankyrase binders that we observed to bind ARC4 in the FP assay scored above 0.68. An arbitrary cutoff at a TTS of 0.77, corresponding to a p value of ≤ 0.07 , contained only known binders and no proven nonbinders (Figure 5C).

To reduce the number of unlikely targets in our scoring list, we additionally filtered the ranked score list for topology (to remove proteins annotated as extracellular, secreted, luminal, or mitochondrial) and predicted disorder (to remove motifs that reside in structured regions) (Table S4). The resulting list contains 11,698; 447; and 257 peptide motifs with TTS cutoffs of ≥ 0.385 , ≥ 0.68 , and ≥ 0.77 , respectively. The apparent large number (257) and diversity of predicted Tankyrase binders at the highest stringency cutoff raise the possibility of a previously underappreciated pervasiveness of Tankyrase-dependent regulation of cellular processes.

Validation of Tankyrase Binders

We probed the validity of our Tankyrase substrate predictions by using FP to test whether peptides corresponding to a subset of predicted TNKS-binding motifs can bind ARC4. Out of 17 peptides tested, ranging in TTS from 0.483 to 0.967, 13 (from the 11 proteins DISC1, STRIATIN, FAT4, RAD54, BCR, MERIT40, CEP170, FASTK, TAOK2, MYO18B, and DIDO1) tested positive, whereas 4 (from the 4 proteins APC, TYK2, SMAD7, and LATS2) tested negative (Figure S7A; TTS highlighted in Figure 5C).

We next tested whether 8 of the 11 positive candidate substrate proteins (i.e., those for which we could generate or obtain epitope-tagged expression constructs) bound to Tankyrase as full-length proteins and whether the predicted TNKS-binding motifs were responsible for binding. We immunoprecipitated transiently expressed FLAG-tagged variants of Disc1, Striatin, Fat4 (lacking most of its extracellular portion), RAD54, BCR, MERIT40, FASTK, and Dido1. Coexpressed MYC-TNKS2 coimmunoprecipitated with all WT substrate candidates except for FASTK and Dido1, which expressed too weakly to allow for solid conclusions (Figure 6 and data not shown). Binding was abolished or reduced by mutation of the predicted TNKS-binding motifs (G at position 6 to R; equivalent to a cherubism mutation in 3BP2). In cases where two TNKS-binding motifs were predicted (Fat4 and MERIT40), both motifs contributed to Tankyrase binding. Using an anti-PAR antibody, we monitored the PARsylation status of these proteins to evaluate whether they were modified by TNKS2. Disc1, STRIATIN, Fat4, and BCR appeared strongly PARsylated by Tankyrase, whereas MERIT40 was weakly PARsylated and RAD54 was not detectably PARsylated (Figure 6). In all cases where PARsylation was observed, modification was sensitive to mutation of the TNKS-binding motifs (BCR to the least degree) and also dependent on the enzymatic integrity of the PARP domain of TNKS2.

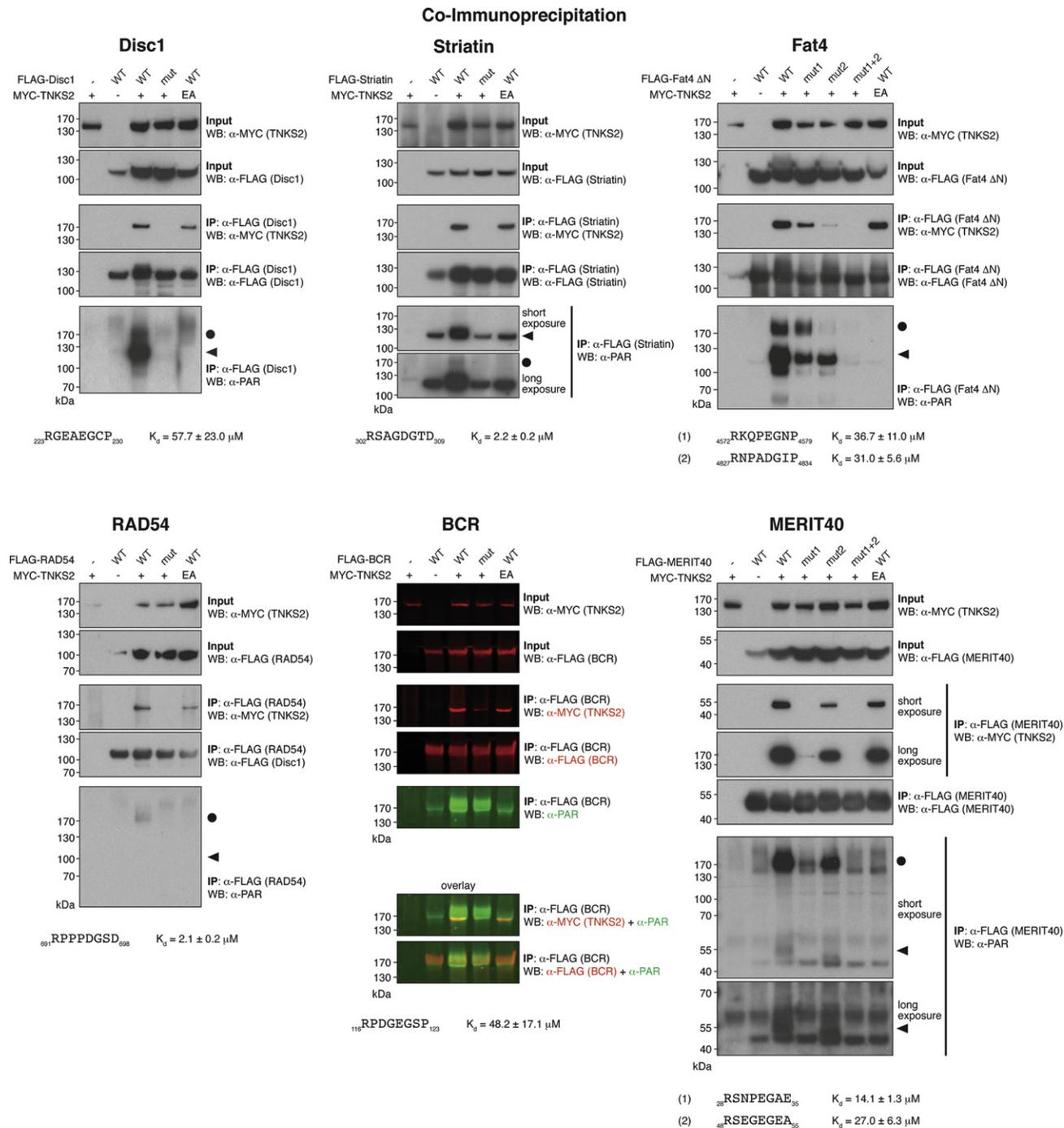


Figure 6. Validation of Predicted Tankyrase Binders and Substrates

The indicated proteins (FLAG-substrates, see Figure S1A for schematic representations; MYC-TNKS2) were transiently expressed in HEK293T cells (input). Labels “mut,” “mut1,” “mut2,” and “mut1+2” indicate mutation of the corresponding TNKS-binding motifs (G-to-R at position 6). MYC-TNKS2 “EA” corresponds to catalytically inactive TNKS2 E1138A. FLAG-substrates were immunoprecipitated (IP) and immunoprecipitates analyzed for coprecipitating MYC-TNKS2 and for the PARylation status of FLAG-substrates and MYC-TNKS2. ● and ▲ indicate expected approximate molecular masses of MYC-TNKS2 and FLAG-substrates, respectively. Sequences and positions of the TNKS-binding motifs (in the human orthologs, from which peptides were derived) are shown below with their affinities for ARC4, as determined by FP. See Figure S7 for additional information. For BCR, to assess whether the anti-PAR signal corresponds to MYC-TNKS2 or FLAG-BCR, samples were analyzed using a fluorescence imaging system, which allows for simultaneous probing of the same membrane. See also Figure S1.

To test whether binding of the Tankyrase targets was sensitive to mutation of the peptide-binding pocket in ARCs 1, 2, 4, and 5 and whether endogenous rather than overexpressed Tankyrase targets can bind the substrate recruitment module of Tankyrase, we performed GST pull-down experiments from a HEK293T cell lysate with a bacterially expressed TNKS2 N terminus encompassing all five ARCs (His₆-GST-TNKS2(20–800), WT, or xx3xx and xxxxx mutants; see above). Although we were unable to detect endogenous expression of DISC1 and FAT4 in HEK293T cells with the antibodies at hand (data not shown), we detected endogenous expression of RAD54, MERIT40, BCR, STRIATIN, and TERF1 as a positive control (Figure S7B). We observed ARC binding by each of the five proteins tested (Figure S7B). Binding of RAD54 and TERF1 was abolished by ARC mutations, as predicted, whereas binding of MERIT40, BCR, and STRIATIN was not greatly affected. These results suggest that Tankyrase substrates may be differentially sensitive to the specific L-to-W ARC mutation employed in this analysis. Further analysis bore this conclusion out (Figure S7C) and highlighted a requirement for compound binding pocket mutations to each functional ARC to ensure universal abrogation of substrate recruitment to Tankyrases.

DISCUSSION

In this study, we uncover the structural basis for substrate recognition by Tankyrase, mediated by its N-terminal ankyrin repeats. This provides a prototypic example for substrate recruitment by a PARP and also for the means by which an ankyrin repeat protein can recognize short, extended peptides rather than a globular protein domain, the more typical function described so far for ankyrin repeat proteins (reviewed in Li et al., 2006). Given the sequence similarity of both human Tankyrases, our findings likely apply to substrate recognition by both TNKS and TNKS2.

We demonstrate that the repeating unit of the Tankyrase N terminus, the ARC, consists of five stacked ankyrin repeats forming a modular unit and occurs five times within the Tankyrase N terminus. This clarifies previous uncertainties over the repetitive character of the Tankyrase N terminus.

The crystal structure of ARC4 bound to the 3BP2 peptide fully explains how mutations in this sequence lead to a loss of 3BP2 binding that gives rise to cherubism (Lietman et al., 2006; Ueki et al., 2001). The ARC4-3BP2 binding mechanism appears universal to ARCs 1, 2, 4, and 5 and is employed for numerous other Tankyrase substrates, as demonstrated by our crystal structures of the TNKS-binding motif peptides of TERF1, MCL1, LNPEP, NUMA1, and FNBP1 bound to ARC4.

Our peptide-binding studies reveal the sequence determinants for low-micromolar-affinity recognition of TNKS-binding motif peptides by Tankyrase: most importantly, positions 1 and 6 of a functional motif require arginine and glycine, respectively, whereas the amino acid at position 4 needs to be small and hydrophobic. The possibilities for position 5 are restricted: aspartate is strongly preferred, but glutamic acid, valine, glutamine, tyrosine, isoleucine, and cysteine are tolerated as well. Conversely, the side chains at positions 2 and 3 are solvent exposed and can thus adopt a wide range of identities. For such peptides to be recognized by Tankyrase, the amino acid

at position 7 must not be proline. Glutamate or aspartate at position 8 positively contribute to Tankyrase binding. Interestingly, most predicted TNKS-binding motifs deviate from the optimal “superconsensus,” suggesting that their affinity may be detuned to avoid potentially deleterious effects from too high affinity. The presence of multiple TNKS-binding motifs in a single substrate, as demonstrated for MERIT40 and Fat4, provides a means for increasing affinity beyond levels afforded by a single motif.

Our consensus-based *in silico* analysis predicts that a wide range, possibly hundreds, of proteins may be targeted by Tankyrase. Our analysis correctly predicts all known and proposed human Tankyrase binders in the UniProt database, including the recently identified Tankyrase substrates BLZF1 (UniProt ID GO45_HUMAN) and CASC3 (UniProt ID CASC3_HUMAN) (Zhang et al., 2011). We identified and validated six predicted Tankyrase binders, namely the neurogenesis regulator Disc1, the scaffold protein and PP2A phosphatase regulatory subunit Striatin, the atypical cadherin Fat4 involved in Hippo signaling and planar cell polarity, the DNA recombination protein RAD54, the GTPase-activating protein and Abl kinase fusion partner BCR, and the DNA repair protein MERIT40. Recruitment by Tankyrase did not strictly correlate with Tankyrase-mediated PARsylation, indicating that although Tankyrase binding via the TNKS-binding motif is required for modification, additional requirements for PARsylation are likely to exist.

Out of five ARCs present in TNKS2, only ARC3, whose surfaces are relatively poorly conserved across species (see Figure S4B), fails to bind TNKS-binding motif peptides. This agrees with an earlier study in which isolated ARC3, but not other ARCs, of TNKS failed to bind full-length TERF1, TNKS1BP1, LNPEP, and an uncharacterized protein termed TNK1BP2 (Seimiya et al., 2004). A function for ARC3 has yet to be determined. ARC3 may play a structural role in the context of the full-length enzyme or engage as-yet-undefined binding partners in a way that is difficult to predict from sequence conservation.

Future work is required to establish the contribution of individual ARCs to substrate PARsylation. In TNKS, ARC5 appears essential for TERF1 PARsylation, release of TERF1 from telomeres, and telomere elongation (Seimiya et al., 2004), suggesting that at the level of PARP activity, ARCs may not be functionally interchangeable. ARC5, however, was not sufficient for TNKS to induce telomere elongation (Seimiya et al., 2004), suggesting that the collaborative action of multiple ARCs is needed.

Much attention is currently devoted to the development of PARP inhibitors, which are of interest as potential anticancer therapeutics (Lord and Ashworth, 2008; Rouleau et al., 2010). Although TNKS and TNKS2, through their specific interactions with many substrates, are involved in a wide variety of biological processes, inhibition of Tankyrase may hold promise for treating BRCA1/2-deficient and Wnt-dependent cancers (Huang et al., 2009; Lord and Ashworth, 2008; McCabe et al., 2009) and a universal anticancer activity through the stabilization of TERF1 and a concomitant inhibition of telomerase function (Ancelin et al., 2002; Chong et al., 1995; Sbodio and Chi, 2002; van Steensel and de Lange, 1997). All PARP inhibitors characterized to date compete with NAD⁺. The universality of NAD⁺ as an

enzymatic cofactor and the likely emergence of drug resistance are two major challenges in pharmaceutical PARP inhibition (Edwards et al., 2008; Lord and Ashworth, 2008; Rouleau et al., 2010; Sakai et al., 2008). Our advancing understanding of substrate recognition by Tankyrases opens avenues for the development of Tankyrase substrate antagonists, which would address questions of specificity and possibly resistance. Given our evidence for functional redundancy of different ARCs, such antagonists may need to recognize all active ARCs.

EXPERIMENTAL PROCEDURES

Extended Experimental Procedures are available in the Supplemental Information online. Lists of plasmids and peptides used in this study are given in Tables S5 and S6.

Protein Expression, Crystallization, and Structure Determination

Purified human TNKS2 ARC constructs included residues 20–178 (ARC1), 173–487 (ARC2-3), 488–649 (ARC4) or 484–655 for apo-ARC4 structure, 641–800 (ARC5), and 20–800 (ARC1-5). Equivalent ARC mutations (denoted by “x”) were L92W (ARC1), L245W (ARC2), L401W (ARC3), L560W (ARC4), and L713W (ARC5). All crystals were grown at 20°C in hanging drops by mixing one volume of protein solution with one volume of precipitant solution. See Extended Experimental Procedures for details on expression plasmids, protein expression and purification, peptides, protein concentrations, precipitant solutions, cryoprotectants, and data collection and structure determination methodologies.

FP Binding Assays

Binding reactions were performed with 25 nM fluorophore-conjugated peptide and the indicated protein concentrations in 50 mM HEPES-NaOH (pH 7.5), 100 mM NaCl, 2 mM TCEP, 80 µg/ml BSA. For Figures 1B, S4A, and S5, the FP buffer contained 300 mM NaCl and 10% glycerol. Nonlinear regression with FP values was performed in GraphPad Prism 5 using a one-site total binding model.

GST Pull-Down Assay

For Figure 3D, His₆-GST-TNKS2(20–800) and mutants were immobilized on glutathione Sepharose and incubated with 48 µg purified murine His₆-3BP2 each in 50 mM HEPES-NaOH (pH 7.5), 300 mM NaCl, 2 mM DTT, 0.5 mM PMSF (total volume: 250 µl) for 90 min at 4°C. The resin was washed six times with binding buffer (without PMSF). Input and pull-down samples were analyzed by SDS-PAGE, Coomassie staining (input), or western blotting with a FLAG-tagged anti-3BP2 F_{ab} (1 ng/ml) and anti-FLAG M2-HRP.

Coimmunoprecipitation

For Figure 3E, HEK293T cells were transfected with expression plasmids for the indicated MYC-TNKS2 and FLAG-3BP2 derivatives or empty vector. After 24 hr, cells were lysed in 25 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1% Triton X-100, 1 mM Na₃VO₄, 5 mM EDTA, and protease inhibitors. Immunoprecipitation with soluble, precleared lysates was performed with anti-FLAG M2-agarose for 1 hr at 4°C. Immunoprecipitates were washed three times with lysis buffer.

For Figure 6, HEK293T cells were transfected with expression plasmids for MYC-TNKS2 (wild-type or PARP-dead E1138A), FLAG-substrate candidates (wild-type or G6R), or empty vector. Cells were cotransfected with dominant-negative PARG to limit PAR hydrolysis. After 20–24 hr, cells were lysed in 50 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100, 5 mM EDTA (pH 8.0), 2 µM ADP-HPD (PARG inhibitor), 5 mM DTT, and protease inhibitors. FLAG-substrate candidates were immunoprecipitated using anti-FLAG M2-agarose for 90 min at 4°C. Immunoprecipitates were washed five times with lysis buffer (without protease inhibitors and ADP-HPD). Lysates and immunoprecipitates were analyzed by SDS-PAGE and western blotting with anti-MYC 9E10, anti-FLAG M2, and polyclonal anti-PAR antibodies.

In Silico Tankyrase Substrate Prediction

An 8 × 20 PSSM (shown in Table S7) was generated using the ARC4 affinities for the peptides in the positional scanning peptide library (Table S3). To score heptapeptides or hexapeptides at the extreme C terminus of proteins, the scoring matrix was truncated after positions 7 or 6. All possible octapeptides, one heptapeptide, and one hexapeptide were scored according to this matrix for each human protein in the UniProt database (release v1.37). A TTS was calculated using the formula

$$TTS = \frac{\sum_{pos.=1}^n PSSM_{pos.}}{\max \left(\sum_{pos.=1}^n PSSM_{pos.} \right)}$$

with n = number of amino acids in peptide scored (6–8).

See Extended Experimental Procedures for additional details and filtering procedures.

ACCESSION NUMBERS

Structure coordinates and structure factors have been deposited in the Protein Data Bank under ID codes 3TWQ (apo-ARC4), 3TWR (ARC4:3BP2), 3TWS (ARC4:TERF1 chimera), 3TWT (ARC4:MCL1 chimera), 3TWU (ARC4:MCL1), 3TWW (ARC4:NUMA1 chimera), 3TWW (ARC4:LNPEP chimera), and 3TWX (ARC4:FNBP1 chimera).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and seven tables and can be found with this article online at doi:10.1016/j.cell.2011.10.046.

ACKNOWLEDGMENTS

We thank Mario Sanches, Elton Zeqiraj, Stephen Orlicky, and Hong Han for technical and computational help, Kay Perry, Kanagalaghatta Rajashankar, and Igor Kourinov at the Advanced Photon Source for assistance with data collection, Alessandro Datti, Frederick Vizeacoumar, and Thomas Sun (SLRI Robotics) for providing robotics, Rey Interior and Li Zhang (SickKids APTC) for N-terminal sequencing and mass spectrometry, Robert Weatheritt and Norman Davey for providing relative conservation scores for Table S4, Sachdev Sidhu for the recombinant antibody against 3BP2, Gregg Morin, Guy Poiret, John Roder, Anne-Claude Gingras, Marilyn Goudreau, Helen McNeill, Richelle Sopko, Rick Bagshaw, Roger Greenberg, Karen Colwill, Marina Olhovskoy, and Bernard Liu for cDNAs and plasmids, Thomas Güttler for helpful comments on the manuscript, and members of the Sicheri and Pawson laboratories for fruitful discussions. This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357. S.G. was supported by postdoctoral fellowships from the Human Frontier Science Program (HFSP) and the European Molecular Biology Organization (EMBO). This work was supported by grants from the Canadian Institutes of Health Research (MOP-36399 and MOP-6849, to F.S. and T.P.), the Ontario Research Fund (#GL2-01-025, to T.P.), and the Terry Fox Program (#20003, to R.R.). F.S. is supported by a Canada Research Chair in Structural Biology.

Received: March 5, 2011

Revised: August 9, 2011

Accepted: October 18, 2011

Published: December 8, 2011

REFERENCES

- Amé, J.C., Spenlehauer, C., and de Murcia, G. (2004). The PARP superfamily. *Bioessays* 26, 882–893.
- Ancelin, K., Brunori, M., Bauwens, S., Koering, C.E., Brun, C., Ricoul, M., Pommier, J.P., Sabatier, L., and Gilson, E. (2002). Targeting assay to study the cis functions of human telomeric proteins: evidence for inhibition of telomerase by TRF1 and for activation of telomere degradation by TRF2. *Mol. Cell Biol.* 22, 3474–3487.
- Bae, J., Donigian, J.R., and Hsueh, A.J. (2003). Tankyrase 1 interacts with Mcl-1 proteins and inhibits their regulation of apoptosis. *J. Biol. Chem.* 278, 5195–5204.
- Chang, P., Coughlin, M., and Mitchison, T.J. (2005a). Tankyrase-1 polymerization of poly(ADP-ribose) is required for spindle structure and function. *Nat. Cell Biol.* 7, 1133–1139.
- Chang, W., Dynek, J.N., and Smith, S. (2005b). NuMA is a major acceptor of poly(ADP-ribosylation) by tankyrase 1 in mitosis. *Biochem. J.* 391, 177–184.
- Chang, W., Dynek, J.N., and Smith, S. (2003). TRF1 is degraded by ubiquitin-mediated proteolysis after release from telomeres. *Genes Dev.* 17, 1328–1333.
- Chi, N.W., and Lodish, H.F. (2000). Tankyrase is a golgi-associated mitogen-activated protein kinase substrate that interacts with IRAP in GLUT4 vesicles. *J. Biol. Chem.* 275, 38437–38444.
- Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P., and de Lange, T. (1995). A human telomeric protein. *Science* 270, 1663–1667.
- Cook, B.D., Dynek, J.N., Chang, W., Shostak, G., and Smith, S. (2002). Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2 at human telomeres. *Mol. Cell Biol.* 22, 332–342.
- Deng, Z., Lezina, L., Chen, C.J., Shtivelband, S., So, W., and Lieberman, P.M. (2002). Telomeric proteins regulate episomal maintenance of Epstein-Barr virus origin of plasmid replication. *Mol. Cell* 9, 493–503.
- Deng, Z., Atanasiu, C., Zhao, K., Marmorstein, R., Sbdio, J.I., Chi, N.W., and Lieberman, P.M. (2005). Inhibition of Epstein-Barr virus OriP function by tankyrase, a telomere-associated poly-ADP ribose polymerase that binds and modifies EBNA1. *J. Virol.* 79, 4640–4650.
- Edwards, S.L., Brough, R., Lord, C.J., Natrajan, R., Vatcheva, R., Levine, D.A., Boyd, J., Reis-Filho, J.S., and Ashworth, A. (2008). Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* 451, 1111–1115.
- Forrer, P., Stumpp, M.T., Binz, H.K., and Plückthun, A. (2003). A novel strategy to design binding molecules harnessing the modular nature of repeat proteins. *FEBS Lett.* 539, 2–6.
- Fuchs, U., Rehkamp, G.F., Slany, R., Follo, M., and Borkhardt, A. (2003). The formin-binding protein 17, FBP17, binds via a TNKS binding motif to tankyrase, a protein involved in telomere maintenance. *FEBS Lett.* 554, 10–16.
- Hassa, P.O., and Hottiger, M.O. (2008). The diverse biological roles of mammalian PARPs, a small but powerful family of poly-ADP-ribose polymerases. *Front. Biosci.* 13, 3046–3082.
- Hottiger, M.O., Hassa, P.O., Lüscher, B., Schüle, H., and Koch-Nolte, F. (2010). Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem. Sci.* 35, 208–219.
- Hsiao, S.J., and Smith, S. (2008). Tankyrase function at telomeres, spindle poles, and beyond. *Biochimie* 90, 83–92.
- Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* 461, 614–620.
- Kaminker, P.G., Kim, S.H., Taylor, R.D., Zebardian, Y., Funk, W.D., Morin, G.B., Yaswen, P., and Campisi, J. (2001). TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes rapid induction of cell death upon over-expression. *J. Biol. Chem.* 276, 35891–35899.
- Kleine, H., Poreba, E., Lesniewicz, K., Hassa, P.O., Hottiger, M.O., Litchfield, D.W., Shilton, B.H., and Lüscher, B. (2008). Substrate-assisted catalysis by PARP10 limits its activity to mono-ADP-ribosylation. *Mol. Cell* 32, 57–69.
- Levaot, N., Voytyuk, O., Dimitriou, I., Sircoulomb, F., Chandrakumar, A., Deckert, M., Krzyzanowski, P.M., Scotter, A., Gu, S., Janmohamed, S., et al. (2011). Loss of Tankyrase-mediated destruction of 3BP2 is the underlying pathogenic mechanism of cherubism. *Cell* 147, this issue, 1324–1339.
- Li, J., Mahajan, A., and Tsai, M.D. (2006). Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry* 45, 15168–15178.
- Lietman, S.A., Kalinchinko, N., Deng, X., Kohanski, R., and Levine, M.A. (2006). Identification of a novel mutation of SH3BP2 in cherubism and demonstration that SH3BP2 mutations lead to increased NFAT activation. *Hum. Mutat.* 27, 717–718.
- Lord, C.J., and Ashworth, A. (2008). Targeted therapy for cancer using PARP inhibitors. *Curr. Opin. Pharmacol.* 8, 363–369.
- Lyons, R.J., Deane, R., Lynch, D.K., Ye, Z.S., Sanderson, G.M., Eyre, H.J., Sutherland, G.R., and Daly, R.J. (2001). Identification of a novel human tankyrase through its interaction with the adaptor protein Grb14. *J. Biol. Chem.* 276, 17172–17180.
- McCabe, N., Cerone, M.A., Ohishi, T., Seimiya, H., Lord, C.J., and Ashworth, A. (2009). Targeting Tankyrase 1 as a therapeutic strategy for BRCA-associated cancer. *Oncogene* 28, 1465–1470.
- Rouleau, M., Patel, A., Hendzel, M.J., Kaufmann, S.H., and Poirier, G.G. (2010). PARP inhibition: PARP1 and beyond. *Nat. Rev. Cancer* 10, 293–301.
- Sakai, W., Swisher, E.M., Karlan, B.Y., Agarwal, M.K., Higgins, J., Friedman, C., Villegas, E., Jacquemont, C., Farrugia, D.J., Couch, F.J., et al. (2008). Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature* 451, 1116–1120.
- Sbdio, J.I., and Chi, N.W. (2002). Identification of a tankyrase-binding motif shared by IRAP, TAB182, and human TRF1 but not mouse TRF1. NuMA contains this RXXPDG motif and is a novel tankyrase partner. *J. Biol. Chem.* 277, 31887–31892.
- Sbdio, J.I., Lodish, H.F., and Chi, N.W. (2002). Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1 (telomere-repeat-binding factor 1) and IRAP (insulin-responsive aminopeptidase). *Biochem. J.* 361, 451–459.
- Seimiya, H., and Smith, S. (2002). The telomeric poly(ADP-ribose) polymerase, tankyrase 1, contains multiple binding sites for telomeric repeat binding factor 1 (TRF1) and a novel acceptor, 182-kDa tankyrase-binding protein (TAB182). *J. Biol. Chem.* 277, 14116–14126.
- Seimiya, H., Muramatsu, Y., Smith, S., and Tsuruo, T. (2004). Functional subdomain in the ankyrin domain of tankyrase 1 required for poly(ADP-ribosylation) of TRF1 and telomere elongation. *Mol. Cell Biol.* 24, 1944–1955.
- Smith, S., Girit, I., Schmitt, A., and de Lange, T. (1998). Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science* 282, 1484–1487.
- Ueki, Y., Tiziani, V., Santanna, C., Fukai, N., Maulik, C., Garfinkle, J., Ninomiya, C., doAmaral, C., Peters, H., Habal, M., et al. (2001). Mutations in the gene encoding c-Abl-binding protein SH3BP2 cause cherubism. *Nat. Genet.* 28, 125–126.
- van Steensel, B., and de Lange, T. (1997). Control of telomere length by the human telomeric protein TRF1. *Nature* 385, 740–743.
- Yaffe, M.B., Leparo, G.G., Lai, J., Obata, T., Volinia, S., and Cantley, L.C. (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways. *Nat. Biotechnol.* 19, 348–353.
- Zhang, Y., Liu, S., Mickanin, C., Feng, Y., Charlat, O., Michaud, G.A., Schirle, M., Shi, X., Hild, M., Bauer, A., et al. (2011). RNF146 is a poly(ADP-ribose)-directed E3 ligase that regulates axin degradation and Wnt signalling. *Nat. Cell Biol.* 13, 623–629.