

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 (LPHLQRSPDGQSFRRS) 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 LPHLQRSPDGQSFRRS (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 RSPDGQS 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

Download English Version:

<https://daneshyari.com/en/article/2036058>

Download Persian Version:

<https://daneshyari.com/article/2036058>

[Daneshyari.com](https://daneshyari.com)