The ARTT motif and a unified structural understanding of substrate recognition in ADP-ribosylating bacterial toxins and eukaryotic ADP-ribosyltransferases

Seungil Han and John A. Tainer

Department of Molecular Biology
Skaggs Institute for Chemical Biology
The Scripps Research Institute, MB 4,
10550 North Torrey Pines Rd., La Jolla, California 92037, USA

*To whom correspondence should be addressed:
e-mail: jat@scripps.edu, telephone: 858-784-8119, fax:858-784-2289

Running title: The ARTT motif in ADP-ribosylation
Summary

ADP-ribosylation is a widely occurring and biologically critical covalent chemical modification process in pathogenic mechanisms, intracellular signaling systems, DNA repair, and cell division. The reaction is catalyzed by ADP-ribosyltransferases, which transfer the ADP-ribose moiety of NAD to a target protein with nicotinamide release. A family of bacterial toxins and eukaryotic enzymes has been termed the mono-ADP-ribosyltransferases, in distinction to the poly-ADP-ribosyltransferases, which catalyze the addition of multiple ADP-ribose groups to the carboxyl terminus of eukaryotic nucleoproteins. Despite the limited primary sequence homology among the different ADP-ribosyltransferases, a central cleft bearing NAD-binding pocket formed by the two perpendicular $\beta$-sheet core has been remarkably conserved between bacterial toxins and eukaryotic mono- and poly-ADP-ribosyltransferases. The majority of bacterial toxins and eukaryotic mono-ADP-ribosyltransferases are characterized by conserved His and catalytic Glu residues. In contrast, Diphtheria toxin, Pseudomonas exotoxin A, and eukaryotic poly-ADP-ribosyltransferases are characterized by conserved Arg and catalytic Glu residues. The NAD-binding core of a binary toxin and a C3-like toxin family identified an ARTT motif (ADP-ribosylating turn-turn motif) that is implicated in substrate specificity and recognition by structural and mutagenic studies. Here we apply structure-based sequence alignment and comparative structural analyses of all known structures of ADP-ribosyltransfeases to suggest that this ARTT motif is functionally
The ARTT motif in ADP-ribosylation

important in many ADP-ribosylating enzymes that bear a NAD binding cleft as characterized by conserved Arg and catalytic Glu residues. Overall, structure-based sequence analysis reveals common core structures and conserved active sites of ADP-ribosyltransferases to support similar NAD binding mechanisms but differing mechanisms of target protein binding via sequence variations within the ARTT motif structural framework. Thus, we propose here that the ARTT motif represents an experimentally testable general recognition motif region for many ADP-ribosyltransferases and thereby potentially provides a unified structural understanding of substrate recognition in ADP-ribosylation processes.

Keywords: ARTT motif/ADP-ribosylation/ADP-ribosyltransferase/NAD/catalytic mechanism
Introduction

ADP-ribosylation is a biologically important protein modification process that acts in pathogenic mechanisms, intracellular signaling systems, DNA repair, and cell division. ADP-ribosyltransferases catalyze this reaction by transferring the ADP-ribose moiety of NAD to a target protein and releasing nicotinamide. Many bacterial toxins are mono-ADP-ribosyltransferases specific for essential eukaryotic processes. Mono-ADP-ribosylation of diphthamide in elongation factor 2 by diphtheria toxin (DT) (Collier, 1990) and *Pseudomonas* exotoxin A (ETA) (Wick and Iglewski, 1990) inactivates protein synthesis. Cholera toxin (CT) (Fishman, 1990), pertussis toxin (PT) (Ui, 1990; Gierschik, 1992), and *Escherichia coli* heat-labile enterotoxins (LT) (Moss and Richardosn, 1978), interfere with signal transduction in human host cells by ADP-ribosylating regulatory G proteins on arginine or cysteine residues. A family of binary toxins, which includes *Bacillus cereus* vegetative insecticidal protein (VIP2) (Warren, 1997; Han et al., 1999), *Clostridium botulinum* C2 toxin (Aktories et al., 1986), *C. perfringens* iota toxin (Schering et al., 1988), *C. spiroforme* toxin (Popoff and Boquet, 1988), and *C. difficile* toxin (Popoff et al., 1988) modify an arginine residue of actin, inhibit actin polymerization, and block actin ATPase activity. Unlike the classical A-B toxins that assemble into A$_1$B$_3$ heterohexamers within the cell for activity, the binary toxins function through two independent proteins. The 'binding component' assembles into multimers at the cell membrane and is essential for the import of the 'catalytic component', the ADP-ribosyltransferase.
The Clostridium botulinum C3 exoenzyme family of bacterial ADP-ribosyltransferases selectively modifies low molecular weight GTP-binding proteins, RhoA, B, C (Aktories et al., 1987). All members of the C3-like family including C. limosum transferase (Just et al., 1992), Bacillus cereus transferase (Just et al., 1995), Staphylococcus aureus transferase (Wilde et al., 2001), and epidermal differentiation inhibitor (EDIN) from S. aureus (Sugai et al., 1992), have molecular masses of about 25-30 kDa and are very basic proteins (pI > 9). The C3-like toxins exhibit 35-70% amino acid identity and modify Rho proteins at an asparagine residue (Sekine, et al., 1989). This covalent modification inhibits effector binding, and/or nucleotide binding (Wei et al., 1997), to block Rho activity in regulating the actin cytoskeleton.

Mono-ADP-ribosylation in eukaryotes has been implicated in modulating cell-cell interactions, signal transduction, the architecture of the cellular cytoskeleton, and vesicular traffic (Koch-Nolte and Hagg, 1997). The primary amino acid sequences of eukaryotic mono-ADP-ribosyltransferases encode glycosylphosphatidylinositol (GPI)-anchored membrane proteins with entirely extracellular polypeptide chains. The majority of the eukaryotic enzymes are arginine-specific transferases. ADP-ribosylation of arginine appears to be a reversible process and free arginine can be regenerated by ADP-ribosylarginine hydrolases (Moss and Vaughan, 1990).

Distinct from mono-ADP-ribosylation, poly(ADP-ribose) polymerase (PARP) mono-ADP-ribosylates a glutamate of a target protein (initiation) like the other bacterial ADP-ribosyltransferases, but it then elongates this modification to a polymer. PARP is
located in the nucleus of most eukaryotes and helps to maintain genomic integrity in base excision repair, in DNA recombination, and in cellular differentiation (Ruf et al., 1998).

The recent crystal structures from VIP2 binary toxin and C3 families have provided insights into the mechanisms of toxicities and have suggested some unexpected relationships among different toxins (Han et al., 1999; Han et al., 2001). These structural results have enabled the classification of binary toxin and C3-like toxin classes into one family according to their structural features. Herein we define and extend our structure-based understanding of two toxin classes by comparative analyses of bacterial toxins, eukaryotic mono-ADP-ribosyltransferases and PARPs.

The conserved core structure of ADP-ribosyltransferases

All ADP-ribosylating toxins possess a NAD-binding catalytic domain of $\alpha/\beta$ fold with overall dimensions of approximately 35 x 40 x 55 Å (Figure 1(a)). Despite the limited primary sequence homology among the different ADP-ribosylating toxins, the crystallographic studies have identified a common core-fold, which is composed of 70-100 amino acids and includes the NAD binding active site. A central cleft bearing a NAD-binding pocket is formed by the two perpendicular $\beta$-sheet core and one or two $\alpha$-helices (one $\alpha$-helix in binary and C3 toxin classes) (Figure 1(a)). The cleft is significantly different from the Rossmann fold of several dehydrogenases (Rossmann et al., 1975) in which the NAD-binding site is at the carboxy terminus of a parallel $\beta$-sheet.
While the core-fold structures of ADP-ribosylating toxins are remarkably conserved, key differences do exist. Bacterial toxins DT, ETA, LT and PT have homologous active site loops consisting of ~15 residues which may act as arms to recognize ADP-ribose acceptor substrates such as EF-2, and G proteins (Bell and Eisenberg, 1996). In the crystal structures of a DT-NAD complex and an ETA-NAD analog complex, the active site loop becomes disordered upon NAD binding (DT) or undergoes significant conformational change upon NAD-analog binding (ETA) (Bell and Eisenberg, 1996; Li et al., 1996). Also, the difficulty of obtaining a stable PARP-NAD complex structure has been correlated with the presence of the equivalent active site loop blocking the active site cleft, suggesting that the additional α-helical domain of PARP participates in PARP activity modulation (Ruf et al., 1998). In contrast, the preformed NAD-binding pockets of both VIP2 and C3 are not obstructed by this loop and hence do not require significant conformational changes for activity. In fact, VIP2 and C3 replace the active site loop with a short two-residue loop connecting β2 and β3, and the corresponding space is filled with α3 (Figure 1(a)). Furthermore, VIP2 and C3 have four consecutive α-helices, α1-α4, which are mainly replaced with an extended antiparallel β-sheet in other structurally known toxins.

**NAD binding and catalysis**

The ADP-ribosyltransferases share three major structural features in the NAD binding pocket (Figure 1(b), 1(c)). Catalysis involves an absolutely conserved glutamic
acid residue, together with either an arginine or histidine on the opposite side of the pocket. The floor of the NAD-binding site consists of a β-strand (β3) followed by an α-helix (a loop in VIP2 and C3 toxins). Based on amino acid sequence variations in the active sites, the ADP-ribosyltransferases can be divided into two homology groups: the DT group, and the CT group (Domenighini and Rappuoli, 1996). The DT group contains DT, ETA and the family of eukaryotic PARPs. The characteristic features of this group are conserved Glu, the β/α structural element and a conserved His. In the crystal structures of DT-NAD, ETA-NAD analog and PARP-NAD analog complexes, two tyrosine residues appear to be important in NAD binding for anchoring the nicotinamide ring of NAD, creating a π pile of three aromatic rings (Bell and Eisenberg, 1996; Li et al., 1996; Ruf et al., 1998). The conserved His residue stabilizes NAD binding by forming hydrogen bond with the O2´ hydroxyl of the adenosine ribose (Figure 1(c)).

The CT group includes the majority of the mono-ADP-ribosyltransferases characterized by the catalytic Glu, a β/α structural element, and an Arg residue instead of His. In the recent crystal structures of C3 and the VIP2-NAD complex, Arg residue plays an important role in NAD binding by forming a hydrogen bond with a phosphate oxygen of the NAD molecule (Figure 1(b)). A conserved sequence (residues 386-388 in VIP2), called the ‘STS motif’, lies at the end of β3 and stabilizes the NAD-binding pocket by connecting two perpendicular β-sheets. Besides this structural role of the STS motif, the Ser386 side chain forms a hydrogen bond to the catalytic Glu428. In contrast to the two tyrosine residues in DT group, Phe397 and Ser386 within β3 and a loop in the VIP2-
The ARTT motif in ADP-ribosylation

NAD complex play similar role by stacking the nicotinamide portion of NAD into the cleft (Figure 1(b)). Many enzymes in the CT group have a Glu or Gln in the two residues upstream from the catalytic Glu residue, suggesting the deprotonation and/or stabilization of the substrate amino acid (Han et al., 1999, Han et al., 2001).

**The ARTT motif and substrate recognition in bacterial toxins**

The recently published structures of VIP2 binary toxin from *B. cereus* and exoenzyme C3 from *C. botulinum* shed light on substrate recognition of these two classes of toxins. Both the VIP2 and C3 structures implicate a bipartite recognition specificity motif for ADP-ribosylation, termed the ARTT motif, which consists of residues from two adjacent protruding turns, T1 and T2 joining β5 and β6 (Figure 1(b)). All of the residues in the ARTT motif of C3 toxin are well ordered with average B-factors of 18.4 Å² in each of the four molecules in the asymmetric unit (Han et al., 2001). The ARTT motif is positioned only 5 Å away from the N1N and C'1N atoms of NAD which form the scissile, N-glycosidic bond. The position of the solvent-exposed aromatic side chain of Phe in T1 suggests this residue is involved in the Rho-C3 and Actin-VIP2 interaction (Figure 1(b)). The significance of its placement and secondary structure was discovered by superimposing VIP2 with the new C3 structure, thereby supporting the value of comparative structural analyses to elucidate structure-function aspects of these enzymes.

The second turn in the ARTT motif (T2) contains solvent-exposed Gln/Gln, which is conserved among Rho- and Actin-ADP-ribosylating toxins. These structures
suggest that Gln may act in recognizing the Asn41 of Rho protein by forming a pair of hydrogen bonds between the carbonyl and amide group of Asn41 in Rho protein (Figure 5). Glu426 in VIP2, which is in the same position, may similarly recognize substrate by forming salt bridges with the Arg177 guanidinium group of actin, rather than facilitating a nucleophilic attack by the guanidinium as initially proposed (Figure 1(b)) (Han et al., 1999).

By using 15 amino acids of the β5-two turns-β6 motif as a template, we compared all of the known crystal structures of ADP-ribosyltransferases for structural and functional homologues. Interestingly, all members of the CT group contain both structural and functional Gln/Glu in the different sized loops connecting β5 and β6 (Figure 2). Both CT and LT have similar ARTT motifs with conserved aromatic Tyr and Glu. The presence of similar ARTT loops in CT and LT also supports its potential role as a substrate recognition motif since the Glu in the second turn can recognize the Arg residue of G protein for ADP-ribosylation. PT has a long insertion connecting β5 and β6, forming two antiparallel α-helices. These two α-helices interact with a B subunit pentamer. In contrast, DT and ETA of the DT group have a similar size loop connecting β5 and β6, but do not possess Glu/Gln in two residues upstream from the catalytic Glu. The catalytic domain of chicken PARP has a longest loop of 37 amino acid, and this loop appears rigid due to interactions with the catalytic core (Figure 2).

Although a common core structure forms the active site of ADP-ribosylating toxins, the limited sequence homology within the ARTT motif region suggests that the
The ARTT motif allows different mechanisms of substrate recognition to be used by the two different groups of toxins (the DT and CT groups) to perform their shared function.

**The ARTT motif in eukaryotic ADP-ribosyltransferases**

Structure-based amino acid sequence alignment and secondary structure prediction analyses of the mammalian ADP-ribosyltransferase family implicate the presence of ARTT motif and amino acid residues similar to those of bacterial mono-ADP-ribosyltransferases. Many eukaryotic enzymes have a conserved Arg observed in the CT group and contain Glu residue in two residues upstream of the crucial catalytic Glu. The presence of the Glu-x-Glu residue in the aligned ARTT motif region supports the recent hypothesis of the ARTT motif as a role of substrate recognition and specificity. These arginine-specific eukaryotic ADP-ribosyltransferases may recognize the substrate by a conserved Turn2 Glu for appropriate target side-chain hydrogen bonding recognition. Turn1 within the ARTT motif contains Phe or Tyr and is involved in a hydrophobic interaction with the substrate.

**Towards a unified structural understanding of substrate recognition in ADP-ribosylation**

Structure-based sequence alignments and comparative structural analyses prompt the proposal of the ARTT motif as a conserved structural framework for substrate recognition in ADP-ribosylation processes. Specifically, we show that this β5 to β6
The ARTT motif in ADP-ribosylation

region of the ADP-ribosyltransferase fold that was first noted in bacterial toxins is also evidently conserved in the sequences of mammalian ADP-ribosyltransferases. The ARTT motif evidently allows considerable sequence variation to control substrate specificity and thereby potentially provides a common structural framework for the widely variable substrate recognition specificities required for ADP-ribosylation processes. The ARTT motif thus suggests a unified understanding of the structural basis for the control of the ADP-ribosylation processes in the diverse biological functions of pathogenesis, intracellular signaling, DNA repair, and cell division.
Acknowledgments

We thank Novartis Agribusiness Biotechnology Research Inc. for initiation of the structural studies of C3 toxin and the Skaggs Institute for Research for fellowship support (SH).
The ARTT motif in ADP-ribosylation

References


Ruf, A., Murcia, G., Schulz, G. E. : Inhibitor and NAD* binding to Poly(ADP-ribose) polymerase as derived from crystal structures and homology modeling. Biochemistry, 37, 3893-3900 (1998)


The ARTT motif in ADP-ribosylation

Legends to Figures

**Figure 1.** ADP-ribosyltransferase folds and the ARTT motif (a) The overall structure of exoenzyme C3 from *C. botulinum*. Secondary structure elements are shown in blue (α-helices), orange (β-strand) and gray (loop). The conserved ARTT motif is shown in magenta. The Turn1 (T1) and Turn2 (T2) in the ARTT motif are indicated in green. The labels N and C indicate the locations of the termini. (b) The NAD binding site of VIP2 in CT group. NAD molecule is shown in black and the key residues are shown in green ball-and-stick representations. The ARTT motif is shown in magenta. (c) The NAD binding site of Diphtheria toxin with same color scheme as those of (b). Orientation matches (a). Figures were generated using programs MOLSCRIPT (Kraulis, 1991) and Raster 3D (Merrit and Murphy, 1994)

**Figure 2.** Sequence conservation and variation of the ARTT motif of ADP-ribosyltransferases. The ARTT recognition motif in the CT group has a conserved Gln/Glu (gray box) two residues upstream from the catalytic Glu (white). The Gln/Glu in Turn2 along with aromatic side chain (gray box) in Turn1 are implicated in substrate binding and recognition (Han *et al.*, 2001). In contrast, the DT group including PT, ETA and PARP evidently do not contain a substrate recognition motif in the region connecting β5 and β6. The structurally known ADP-ribosyltransferases are indicated by gray box. The numbers in the sequence indicate number of amino acid with big insertion. The
names of ADP-ribosyltransferases with known crystal structures are indicated by gray boxes.
Figure 1

(a) ARTT motif in ADP-ribosylation

(b) Specific amino acid residues

(c) Detailed view of active site loop
Figure 2

CT group

- E. carotae **VIP2**
- C. difficile **CTDb**
- C. sputiforms **Sa**
- C. perfringens **Ia**
- C. botulinum **C2I**
- C. botulinum **C3**
- C. limoen **C3**
- S. aureus **C3**
- S. aureus **EDIN**
- B. cereus **Partial seq.**
- E. faecalis **ARR1**
- R. norvegicus **ARR2**
- G. gallus **GMAT**
- M. musculus **ARR1**
- E. mars **ARR3**
- E. mars **ARR4**
- E. coli **LT**
- V. cholerae **CT**
- B. pertussis **PT**

DT group

- C. diptheriae **DT**
- F. arctigena **MTA**
- G. gallus **PARP**

The ARTT motif in ADP-ribosylation