Puzzling pyrophosphates: synthetic methodologies for ADP-ribosylated biopolymers

Wijngaarden, S.

Chapter 2

Synthetic dual cysteine-ADP ribosylated peptides from the androgen receptor are recognized by the DTX3L/PARP9 complex

This work was published:

S. Wijngaarden, C. Yang, C. Vela-Rodríguez, L. Lehtiö, H. S. Overkleeft, B. M. Paschal, D. V. Filippov, Synthetic dual Cys-ADP ribosylated Androgen Receptor derived peptides are recognized by the DTX3L/PARP9 complex. *ACS Chem. Biol.* **2023**, *17*, 2377-2384.

Introduction

Modification of proteins with adenosine diphosphate ribose (ADPr) is facilitated by ADPr-ribosyltransferase enzymes (for instance ARTs), that transfer ADPr from NAD^+ to nucleophilic acceptor amino acids Asp/Glu[1,2], Ser/Tyr[3,4], Arg[5,6], His[7] and Cys.[8-11] This process is implicated in a number of normal and pathophysiological pathways. ADP-ribosylation of cysteine in prostate cancer cells, where modification occurs with mono-ADP-ribose (MAR), was recently shown to be a critical feature of a gene expression mechanism involving the androgen receptor (AR), a ligand-regulated transcription factor.[12] Androgen binding to the ligand binding domain (LBD) of AR induces conformational change of the unstructured N-terminal domain (NTD), enabling its interaction with the LBD (**Figure 1A**). This process triggers AR transport into the nucleus, where it is ADP-ribosylated on several cysteine residues by the writer enzyme PARP7.[12,13] The agonist-bound, Cys-ADP-ribosylated AR is then recognized by the reader PARP9, which forms a heterodimer with the ubiquitin E3 ligase DTX3L via the Deltex binding domain (DeBD) on DTX3L.[12] Thus, PARP7-mediated ADP-ribosylation of AR (writing) results in highly selective recruitment (reading) of DTX3L/PARP9 and modulation of AR-dependent gene expression.[12] Point mutations in the PARP9 macrodomains (MDs) known to be critical for ADP-ribose binding abrogates DTX3L/PARP9 binding to AR.[12]

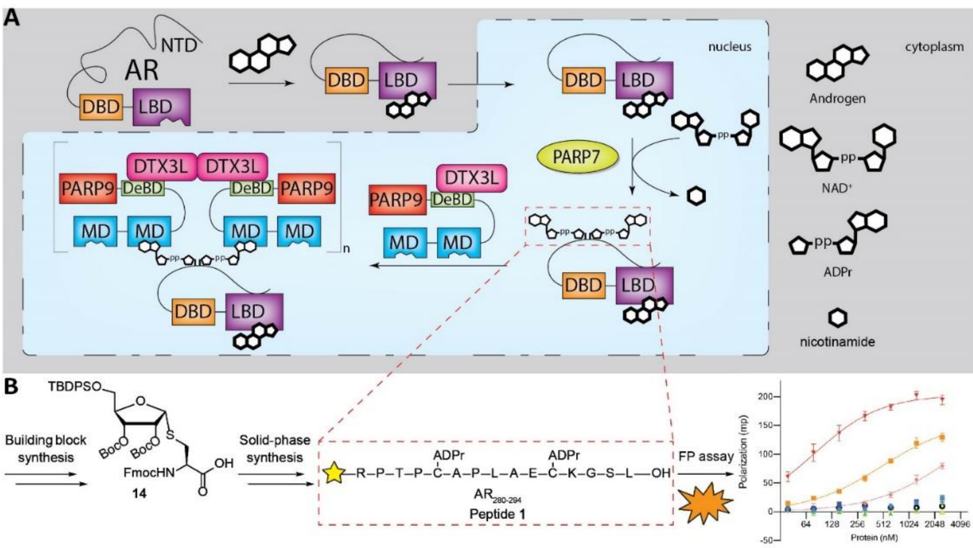


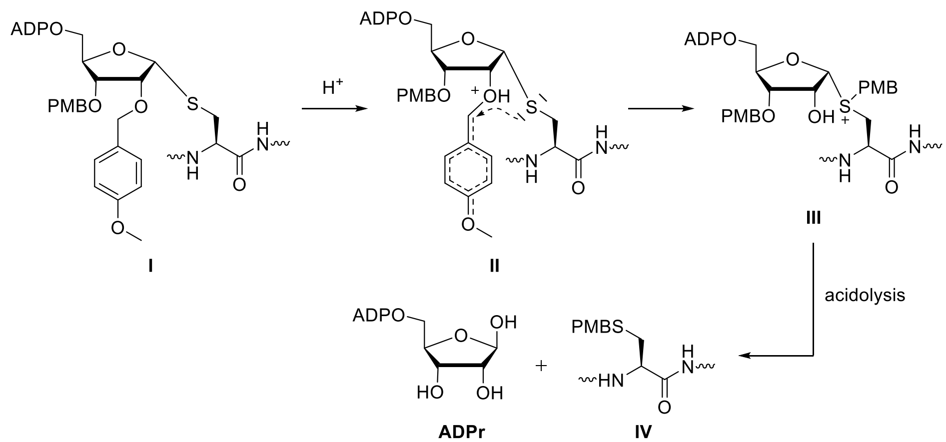
Figure 1A. Proposed mechanism of AR-DTX3L/PARP9 binding. Binding of androgen to the LBD of AR facilitates a conformational change enabling the unstructured NTD to interact with the LBD. Hereafter, the complex is transported to the nucleus, where it gets ADP-ribosylated by PARP7. After ADP-ribosylation, the agonist-bound AR is recognized by the MDs on PARP9 which exists as an oligomeric heterodimer with DTX3L, leading to modulation of AR-dependent gene expression. **B**. Workflow for obtaining binding affinity of model peptide **1**. Solid-phase synthesis of dual-ADPr containing peptide **1** based on building block **14** gave peptide **1** which could be evaluated for its affinity towards the oligomeric, heterodimerized reader complex DTX3L/PARP9.

Prior studies, which have shown that tandemly arranged macrodomains in PARP9 bind more efficiently to ADP-ribosylated AR than the individual macrodomains[12] provided the first clue that multivalent interactions might underpin DTX3L/PARP9 binding AR (**Figure 1A**). The observations that DTX3L/PARP9 assembles into an oligomer *in vitro*[14], and that the native complex from prostate cancer cells has an apparent size >500 kD[15], both indicated that the complex that binds ADP-ribosylated AR contains multiple macrodomains, which theoretically could engage multiple ADP-ribosyl Cys sites in AR. Such a mechanism would help explain the highly selective assembly of the AR-DTX3L/PARP9 complex, which is virtually undetectable prior to AR ADP-ribosylation by PARP7.[16]

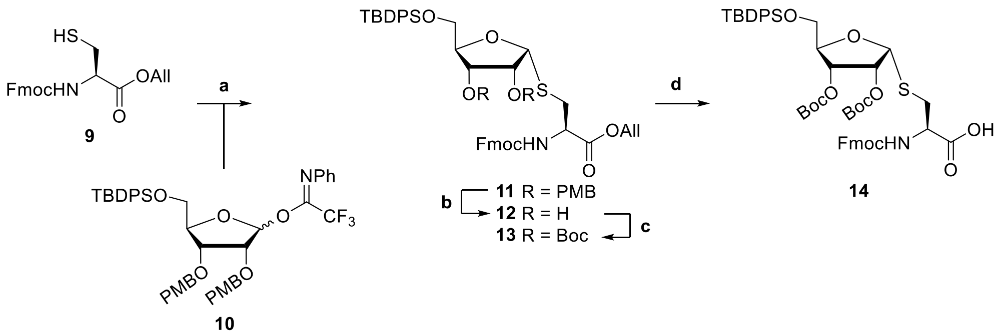
Within the unstructured NTD of AR, four ADP-ribosylation sites (Cys 125 and Cys 131; Cys 284 and Cys 290) appear as pairs, each separated by six amino acids. Given that PARP9 contains two macrodomains, it was hypothesized that pairs of ADP-ribosyl-Cys sites might be recognized by tandem macrodomains in PARP9. To help address this question, the site containing Cys 284 and Cys 290 was selected for focused analysis. To facilitate the preparation of bis-ADP-ribosylated $AR_{280\text{-}294}$ (**Figure 1B**, peptide **1**), key Boc-protected Cys-ribosyl building block **19** had to be developed, for which the solid-phase synthetic scheme was adjusted. To assess the importance of dual Cys-ADP-ribosylation in the context of binding the DTX3L/PARP9 oligomeric complex, a panel of mono-ADP-ribosylated peptides **2-5,** were prepared. The binding affinities of Cys-ADPr peptides for the DTX3L/PARP9 reader complex were then measured by fluorescence polarization (FP). The results indicated that dual ADP-ribosylation of AR is an essential determinant of ADP-ribose reading by the DTX3L/PARP9 complex.

Results and discussion

Recently, the first chemical synthesis of MAR-peptides modified on cysteine was developed by Voorneveld *et al*.[17] Target peptide **1** bearing two ADP-ribosyl-Cys residues, however, could not be prepared in this way (**Figure 3A**). This could possibly be explained by the formation of multiple side products as a consequence of the increased complexity (two ADP-ribosyl-Cys units instead of one) that characterizes peptide **1** in comparison to Cys-ADPr peptides prepared previously. A major side product observed was the migration of a *para*-methoxybenzyl (PMB) protective group installed in the S-ribosyl cysteine building block **I** (see **Scheme 1**) to the cysteine thiol. A mechanistic evaluation of this reaction led to the proposed mechanism of migration as depicted in **Scheme 1**. TFA-induced protonation of the PMB ether in **I** leads to cationic intermediate **II**, which can be intramolecularly intercepted by the sulfur via a 5- or 6-membered transition state. Acidolysis of anomeric sulfonium species **III** then gives free ADPr and the *para*-methoxybenzylated cysteine of type **IV**.



Scheme 1. Proposed mechanism of PMB migration during acidic deprotection. Protonation of the PMB-ether in **I** results in attack of the thioether on the benzylic position of **II**, instigating PMB migration. Hydrolysis of the sulfonium salt **III** then gives the para-methoxybenzylated thiol of cysteine **IV** and free ADPr.

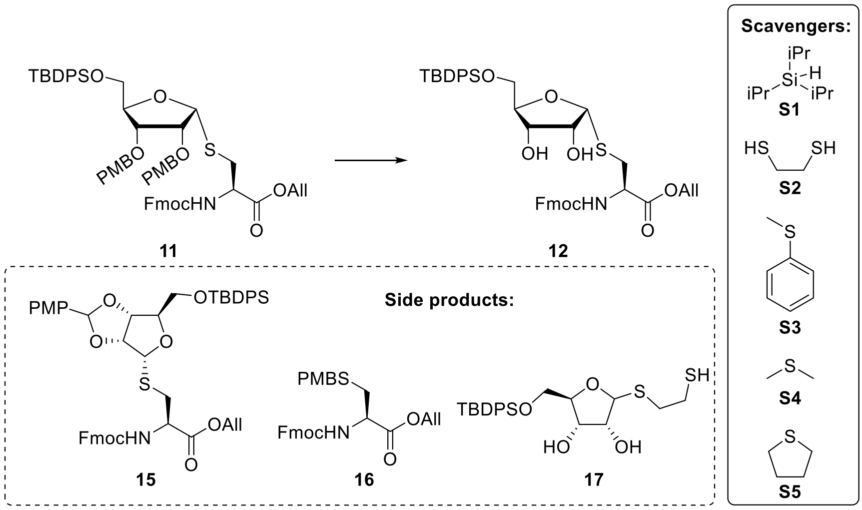


Scheme 2. Synthesis of suitably protected ribosylated cysteine building block **14**. Reagents and conditions: **a**) donor **10**, TBSOTf, DCM:dioxane 9:1, 1 h, -50 °C, 82%; **b**) TFA, TIS, THT, 1.5 h, RT, 66%; **c**) Boc_2O, DMAP, TEA, pyridine, 15 min, RT 77%; **d**) $Pd(PPH_3)_4$, DMBA, DCM, 1 h, RT quant.

To circumvent PMB migration, building block **14** featuring Boc-protection of the 2'- and 3'-hydroxyls was envisioned. (**Scheme 2**). The synthesis of **14** started with the fully α-selective glycosylation of the thiol in cysteine acceptor **9**, with trifluoroacetimidate ribosyl donor **10**. Compound **11** was thus obtained in 82% yield. TFA-mediated PMB removal in DCM as solvent proved abortive and resulted in a full migration of the PMB to the thiol to give **16** (**Table 1**). Oxidative deprotection of **11** with 2,3-dichloro-5,6-

dicyano-1,4-benzoquinone (DDQ) in contrast yielded **16** and the *para*-methoxybenzylidene derivative **15** (entry 1). Other acidic deprotection conditions, such

Table 1. Optimization op PMB deprotection in **11** towards compound **12**.

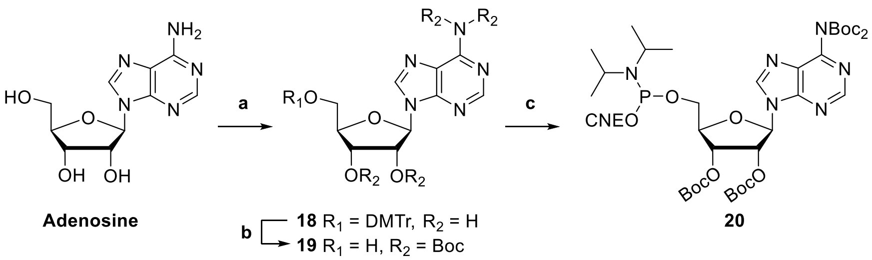


Entry	Reagent	Solvent	Scavenger[d]	Yield	Comments
1	3 eq. DDQ[a]	MeOH/DCM	-	N.A.	**15**+**16**
2	0.1 eq. HCl[b]	HFIP	-	N.A.	**16**
3	10% TFA[c]	DCM	**S1**, **S2**	54%	+side product **17**
4	10% TFA[c]	DCM	**S1**, **S3**, **S4**	61%	
5	10% TFA[c]	DCM	**S1, S3**, **S4**	27%	
6	10% TFA[c]	THT	**S1**	31%	
7	30% TFA[c]	THT	**S1**	56%	
8	30% TFA[c]	THT	**S1**	65%	1 mmol

[a]Reactions were conducted at a concentration of 0.1 M and a scale of 0.1 mmol unless stated otherwise. [b]Oxidative deprotection conditions (DDQ) resulted in only migration to PMB-cysteine **16** and *para*-methoxybenzylidene **15**. [c]Catalytic HCl/HFIP resulted in complete migration to compound **16**. [d]Acidic deprotection conditions using TFA and appropriate scavengers resulted in formation of **12**, as well as **17**. [e]Cation scavengers were used in a 2.5% v/v ratio. Triisopropylsilane (TIS, **S1**), ethane dithiol (EDT, **S2**), thioanisole (**S3**), dimethylsulfide (DMS, **S4**), and tetrahydrothiophene (THT, **S5**).

as catalytic HCl in HFIP, gave PMB migration to yield side product **16** as well (entry 2).[18] Deprotection of the PMB with 10% v/v TFA in DCM with the addition of ethane dithiol (EDT, **S2**) as a scavenger for the PMB cation, gave an inseparable mixture of the desired product **12** with ribosylated EDT (**17**) (entry 3). Thioanisole (**S3**) and dimethyl sulfide (DMS, **S4**) as thioether scavengers gave inconsistent results (entries 4-5). Next, PMB deprotection in tetrahydrothiophene (THT, **S5**) as thioether scavenger and as a solvent was examined (entry 6), and 30% TFA in THT gave the desired compound **12** in the highest yields (entries 7-8).

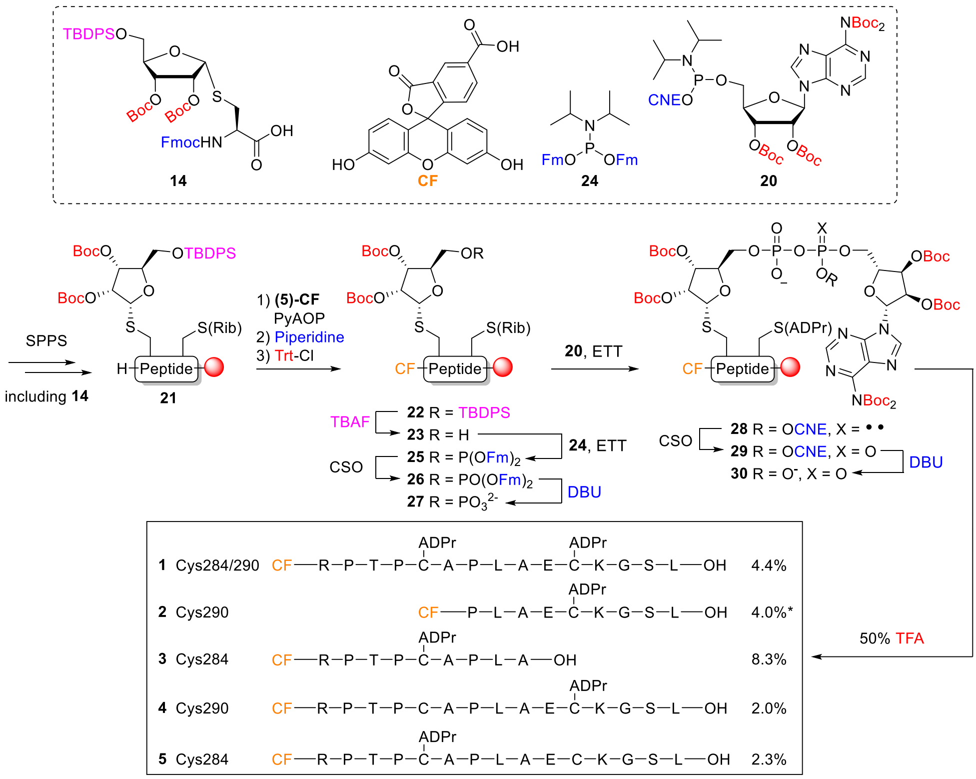
Next, both hydroxyls in diol **12** were Boc-protected, leading to the fully protected ribosylated cysteine **13**, which, after final deprotection of the allyl ester with $Pd(PPh_3)_4$, gave building block **14** ready for use in a solid-phase peptide synthesis.



Scheme 3. Synthesis of Boc-protected adenosine amidite **20**. Reagents and conditions: **a**) DMT-Cl, pyridine:DMF,8 h, 0 °C, 37%; **b**) (*i*) Boc_2O, DMAP, TEA, DCM, ON., RT; (*ii*) TFA, TIS, DCM, 30 min, RT, 68%; **c**) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM, 30 min, RT, 97%.

Boc-protected adenosine amidite **20**[19] was synthesized according to **Scheme 3**. First, adenosine was selectively protected with DMTr on the primary hydroxyl to give **18**. Next, compound **18** was exhaustively acylated with Boc_2O, after which the DMTr group was selectively cleaved using 1% TFA in DCM to give **19** in 68% yield. Lastly, phosphitylation with 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite gave amidite **20** in 97% yield.

With building blocks **14** and **20** in hand, the solid-phase synthesis of peptide **1** was undertaken (**Scheme 4**). Conventional Fmoc-SPPS chemistry was used for peptide elongation to obtain ribosylated and partially protected peptide **21** on the solid phase (**Figure 2B**). Next, N-terminal (5)-carboxy fluorescein (CF) was installed using PyAOP activation.[20] The self-acylation of CF on its phenolic alcohols was reversed by treatment with 20% v/v piperidine/DMF to cleave the unwanted esters.[21] The phenols of the fluorescein were then protected as the trityl (Trt) ether, enabling further on-resin modifications. Deprotection of the 5'-silyl ether of the ribose in **22** with tetra-butylammonium fluoride (TBAF) gave alcohol **23** ready for installation of the phosphate. Two key phosphorylation steps, utilizing phosphoramidite chemistry[22], were employed to yield the ADP-ribosylated peptide on-resin. First, the primary alcohol of ribose was phosphorylated using established phosphoramidite chemistry.[17,23] Coupling of the alcohol with bis-Fm phosphoramidite **24** under activation of 5-(ethylthio)-1H-tetrazole (ETT) yielded the phosphite triester P(III) intermediate **25**. Notably, due to the acidic nature of the activator ETT, the Trt protecting groups installed previously on the phenols were partially cleaved, resulting in over-phosphitylation. Usage of equimolar amounts of **24** and activator ETT prevented this undesired Trt cleavage and thus the unwanted phosphitylation of the phenolic groups at the fluoresceine. CSO oxidation of the P(III) intermediate to P(V) phosphotriester **26** and subsequent cleavage of the Fm groups using DBU yielded phosphate monoester **27** (**Figure 2C**). The pyrophosphate was then introduced through phosphitylation of **27** with adenosine amidite **20** to form P(V)-P(III) intermediate **28**. Oxidation of the P(III) to P(V) by CSO then gave semi-protected pyrophosphate intermediate **29**. Lastly, DBU-mediated cleavage of CNE gave immobilized partially protected ADPr peptide **30,** which was subjected to the acidic global deprotection, giving the final Cys-ADPr peptide **1** in good quality (**Figure 2D**).



Scheme 4. Synthetic strategy for the synthesis of N-terminally modified, tandem Cys-ADP-ribosylated peptides. (Rib) corresponds to the appropriately protected ribosyl moiety. *Peptide **2** was synthesized via the procedure described in Voorneveld *et al.*[17]

A critical feature of this route is the final acidic deprotection, enabling a more conventional cleavage of the product from the solid support, while earlier routes towards ADPr-peptides made use of an alkaline final deprotection[24-26] and mildly acidic conditions.[17] The relatively high stability of the thioglycoside linkage in Cys-ADPr permitted the use of a cleavage cocktail with 50% v/v TFA, resulting in shorter deprotection times while keeping the quality of the peptide the same as with 10% v/v TFA. Final deprotection towards the ADPr containing construct **1** showed no protecting group migration to the thiol and drastically improved the quality of the synthesis (**Figure 2A** vs **Figure 2D**). The tandem ADP-ribosylated peptide was isolated in 4.4% yield and in good purity (**Figure 2E**). Mono-ADPr-peptide **3** and extended mono-ADPr-peptides **4** - **5** were prepared as described for **1** and isolated in good yields. Note that the usage of 50% TFA facilitated the cleavage of the relatively acid-stable Pbf-protection from the arginine side chain. The other, simpler control mono-ADPr-peptide **2** was readily accessed using the method used by Voorneveld *et al*.[16]

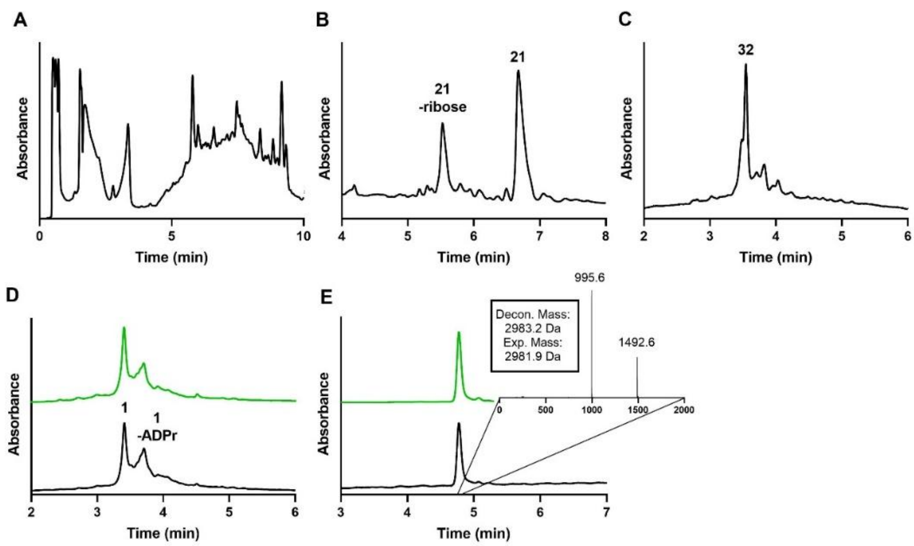


Figure 2. Solid-phase synthesis of intermediates *en route* to tandem-ADP-ribosylated peptide **1**. **A**. Analytical HPLC (linear gradient 10-50% ACN over 10 min.) of crude peptide **1**, synthesized via the procedure from Voorneveld *et al*.[17] **B**. Analytical HPLC (linear gradient 10-90% ACN over 10 min.) of intermediate **21**. **C**. Analytical HPLC (linear gradient 10-90% ACN over 10 min.) of intermediate **27**. **D**. Analytical HPLC (linear gradient 10-90% ACN over 10 min.) of crude peptide **1**. **E**. Analytical HPLC (linear gradient 10-50% ACN over 10 min.) of purified peptide **1**.

To determine whether synthetic AR peptides containing one or two ADP-ribosylated Cys residues are recognized by the PARP9 macrodomains, a series of binding reactions using fluorescent ADP-ribosylated peptides **1**-**5** and recombinant PARP9 were performed. The binding affinity of DTX3L/PARP9 with the fluorescently labeled, Cys-ADP-ribosylated peptides was determined using fluorescence polarization (FP).[27] Recombinant DTX3L was included in the analysis, which heterodimerizes with PARP9, but also promotes the formation of an oligomeric complex that, based on gel filtration and light scattering, contains at least four copies of the heterodimer.[14] Because the PARP9 monomer encodes two macrodomains, the fully assembled DTX3L/PARP9 complex is predicted to contain eight macrodomains. FP measurements revealed a low level of PARP9 binding to the dual ADP-ribosylated AR peptide **1**, and binding was too weak to infer a Kd (**Figure 3A**). In contrast, DTX3L/PARP9, preassembled as a complex prior to peptide addition, displayed robust binding to peptide **1** (Kd = 80.5 nM). This differs from the near-background level of DTX3L/PARP9 binding to AR peptides that contain a single ADP-ribose (peptides **2** and **3**). As expected, a control peptide **6** (**Table S1**) lacking ADP-ribose did not bind significantly to the DTX3L/PARP9 complex (**Figure 3A**). Thus, dual ADP-ribosylation of the AR peptide promotes binding to the DTX3L/PARP9 complex. To test if oligomerization of DTX3L/PARP9 affects its activity as an ADP-ribose reader, a mutant of DTX3L (DTX3LΔN) that heterodimerizes with PARP9 but fails to form oligomers was employed.[14] DTX3LΔN/PARP9 displayed low binding to the dual ADP-ribosylated peptide **1** that was not saturable and likely reflects a weak association (Kd> 500 nM).

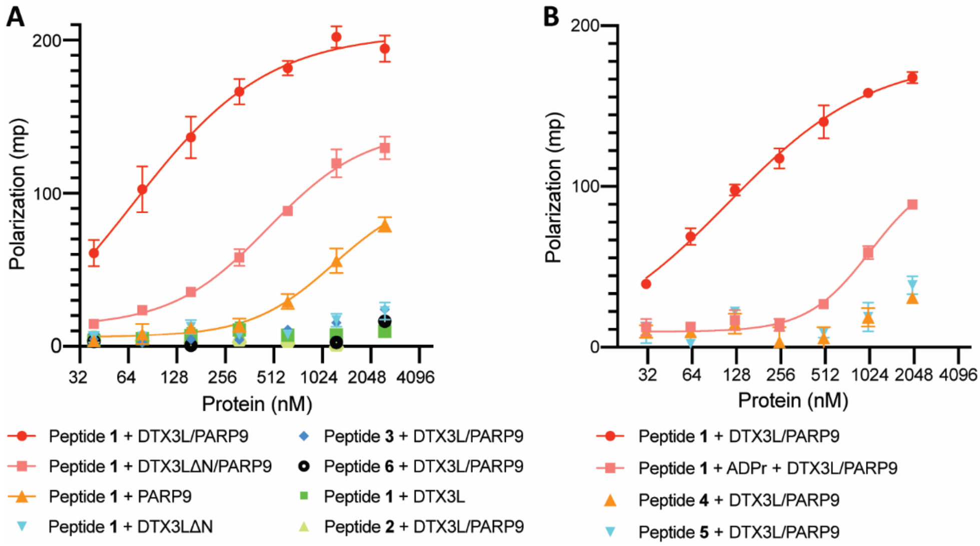


Figure 3. Interaction of DTX3L/PARP9 and ADP-ribosylated AR peptides determined by fluorescence polarization assays. **A**. Oligomerization of DTX3L/PARP9 is critical for recognition of dual ADP-ribosylated AR peptide. **B**. DTX3L/PARP9 binding is competed with ADP-ribose (50 μM) in solution. For all measurements: n=3, error bars = SD. CF = (5)-carboxy fluorescein. Peptides sequences are given in **Table S1**.

Hereafter, it was confirmed that the interaction between the dual ADPr peptide with DTX3L/PARP9 is outcompeted with free ADP-ribose (**Figure 3B**). Addition of excess ADPr (50 μM) to the FP assay reduced binding of DTX3L/PARP9 to dual ADP-ribosylated peptide **1** to ~17% (**Figure 3B**). In addition, ADP-ribosylated peptides (peptides **4** and **5**) that contain the same primary sequence as peptide **1**, but only a single ADP-ribose, were evaluated for their binding towards DTX3L/PARP9. This helped eliminate the concern that flanking sequences contributed to the superior binding observed with peptide **1** (**Figure 3A** and **B**). Overall, the ADP-ribosyl peptide binding results confirmed the prior analysis showing that loss-of-function point mutations in the two PARP9 MDs (G112, 311E) eliminates binding to AR, as do mutations in the Cys ADP-ribosylation sites in AR that that are modified by PARP7.[12]

The data in this chapter is consistent with a model of bivalent (and possibly multi-valent) ADPr interaction with PARP9 that are dependent on DTX3L/PARP9 oligomeric assembly. The spacing of the two Cys-ADPr moieties in the AR peptide may be too close to permit binding to the tandem macrodomains in a single PARP9 molecule. It seems plausible, however, that macrodomains juxtaposed in the DTX3L/PARP9 oligomeric complex may be sufficiently close to accommodate bivalent binding to the dual ADP-ribosylated -peptide. In any case, these experiments suggest that multi-site ADP-ribosylation is an important feature of AR binding DTX3L/PARP9, and that bivalent interactions may be fundamental. The other pair of ADP-ribosyl-Cys sites in the AR N-terminal domain (ADPr-C125, C131) which were not included in this chapter have the same spacing between the ADP-ribosylation sites as peptide **1** evaluated here by FP. These sites could potentially also be oriented in a manner that permits oligomeric DTX3L/PARP9 binding.

Conclusion

In conclusion, the work presented in this chapter reveals a new and effective strategy for the synthesis of fluorescently labeled Cys-ADPr-peptides, including a first preparation of a well-defined synthetic peptide containing a dual ADP-ribosylation motif. To facilitate the synthesis of tandem ADPr-peptide **1**, in which a fluorescent group was introduced on the N-terminus, Boc-protected Fmoc-cysteine-ribofuranosyl building block **14** was designed and synthesized. Sequential desilylation, phosphorylation, pyrophosphorylation and global acidolysis then yielded peptide **1** without migration of the protecting group on ribose to the cysteine thiol. The availability of fluorescent peptides **1**-**5** derived from the AR and ADP-ribosylated on Cys 284 and/or 290, enabled the assessment of whether tandem ADP-ribosylation is an essential factor in the recognition of the AR by the DTX3L/PARP9 heterodimeric complex. Binding affinities were determined via a fluorescence polarization assay. The binding of DTX3L/PARP9 to the dual ADP-ribosylated peptide **1** gave a saturable binding in the nanomolar range (Kd=80.5nM), while insignificant binding was observed of mono-ADPr-peptides **2**-**5** with DTX3L/PARP9 indicating the importance of the dual ADPr motif. The binding of peptide **1** to the mutant DTX3LΔN/PARP9 heterodimer, which does not oligomerize, or PARP9 without DTX3L, was not saturable. Together, these observations suggest that closely spaced ADP-ribosyl groups (for instance Cys 284 and Cys 290) might not be recognized by the tandem MDs in a single PARP9 polypeptide, but rather by the MDs positioning within the oligomer. It should be noted that the analysis presented in this chapter focused on a single pair of ADP-ribosylated Cys sites. Although the actual stoichiometry of ADP-ribosyl modified AR Cys sites has not been defined, it is conceivable that multiple sites are engaged by DTX3L/PARP9 at the same time.

Acknowledgements

Chunsong Yang from the University of Virginia School of Medicine, United States of America is kindly acknowledged for his work on the fluorescence polarization binding assays.

Appendix

Table S1. Overview of peptides synthesized for the fluorescence polarization binding assay.

Code	Name	Sequence
1	$AR_{280\text{-}294}$ADPr	CF-RPTP**C**(ADPr)APLAE**C**(ADPr)KGSL-OH
2	$AR_{286\text{-}294}$ADPr	CF-PLAE**C**(ADPr)KGSL-OH
3	$AR_{280\text{-}288}$ADPr	CF-RPTP**C**(ADPr)APLA-OH
4	$AR_{280\text{-}294}$ADPr	CF-RPTPCAPLAE**C**(ADPr)KGSL-OH
5	$AR_{280\text{-}294}$ADPr	CF-RPTP**C**(ADPr)APLAECKGSL-OH
6	$AR_{280\text{-}294}$	CF-RPTPCAPLAECKGSL-OH
7	$AR_{286\text{-}294}$	CF-PLAECKGSL-OH
8	$AR_{280\text{-}288}$	CF-RPTPCAPLA-OH

Amino acids are given in the one-letter code. Peptides are synthesized according to the general procedures for peptide synthesis and installation of CF. ADP-ribosylated amino acid residues are highlighted. Abbreviations CF = (5)-carboxy fluorescein (CF).

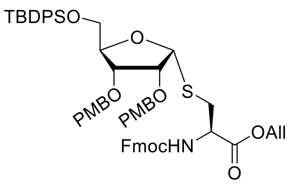
Experimental section

All reagents were of commercial grade and used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4 Å molecular sieves, except MeOH and MeCN which were stored over 3 Å molecular sieves. Column chromatography was performed on silica gel 60 Å (40-63 µm, Macherey-Nagel). TLC analysis was performed on Macherey-Nagel aluminium sheets (silica gel 60 F254). Compounds were visualized on TLC by irradiation with UV (254 nm) and by spraying with either cerium molybdate spray (25 g/L $(NH_4)_6Mo_7O_{24}$, 10 g/L $(NH_4)_4Ce(SO_4)_4 \cdot H_2O$ in 10% H_2SO_4 water solution), $KMnO_4$ spray (20 g/L $KMnO_4$ and 10 g/L K_2CO_3 in water) or H_2SO_4 (20% v/v in MeOH) followed by charring at about 250 °C. ^{1}H and ^{13}C NMR spectra were recorded on a Bruker AV-400, AV-500 or AV-600 NMR. Chemical shifts (δ) are given in ppm relative to tetramethylsilane as internal standard. Coupling constants (*J*) are given in Hz. For pyrophosphate-containing peptides, a small amount of EDTA was added to the NMR sample to sharpen the peaks for ^{31}P-NMR. All given ^{13}C-APT spectra are proton decoupled and are presented with even signals (Cq. and CH_2) pointing upwards and odd signals (CH and CH_3) pointing downwards.

LCMS analysis was performed on a Finnigan Surveyor HPLC system with a Nucleodur C18 Gravity 3 µm 50 x 4.60 mm column (detection at 200-600 nm) coupled to a Finnigan LCQ Advantage Max mass spectrometer with ESI or a Thermo Scientific Vanquish UHPLC coupled to a Thermo Scientific LCQ Fleet ion mass spectrometer with ESI. Buffers used were A= H_2O, B= MeCN and C= 1% TFA/H_2O. The methods used were 10→90% 13.5 min (0→0.5 min: 10% MeCN; 0.5→8.5 min: 10% to 90% MeCN; 8.5→11 min: 90% MeCN; 11→13.5 min: 10% MeCN) or 0→50% 13.5 min. HPLC purification was performed on a Gilson GX-281 preparative HPLC with a Gemini-NX 5u, C18, 110 Å, 250 x 10.0 mm column or on a Waters autopurifier HP-LC/MS system coupled to a Phenomenex Gemini 5 µm 150x21.2 mm column. HRMS was recorded on Thermo Scientific Q Exactive HF Orbitrap mass spectrometer equipped with an electrospray ion source.

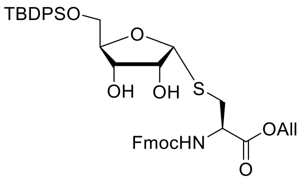
Building block synthesis

1-*S*-2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl cysteine allyl ester (11)



Acceptor **9** (1.70 g, 4.45 mmol) and imidate donor **10** (3.9 g, 4.90 mmol, 1.1 eq.) were co-evaporated with toluene (3x), dissolved in a mixture of dioxane and DCM (1:9 v/v, 45 mL, 0.1 M) and cooled to -50 °C. Activator TBSOTf (0.11 mL, 0.49 mmol, 0.11 eq.) was added and the mixture was stirred at -50 °C for 1 hour, whereupon, TLC analysis showed near full conversion of the starting material. The reaction was quenched with TEA and evaporated *in vacuo*. Purification by flash column chromatography (0% -> 0.5% -> 1.5% acetone/DCM) yielded the title compound as a clear oil (3.60 g, 3.66 mmol, 82%). **Rf** 0.47 (1.5% acetone/DCM). Spectra were in full accordance with the literature experimental data.[17]

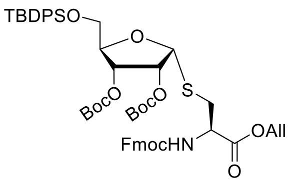
1-*S*-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl cysteine allyl ester (12)



Compound **11** (1.0 g, 1.0 mmol) was dissolved in tetra (7 mL) after which a mixture of 50% TFA, 4% TIS in THT (14 mL) was added to a final concentration of 0.05 M. The mixture was stirred for 1.5 hours, after which TLC analysis showed near-full conversion of the starting material. The reaction was quenched with sat. aq. $NaHCO_3$ and extracted with DCM (2x). The combined organics were dried over $MgSO_4$, filtered, concentrated *in vacuo* and co-evaporated with toluene (3x). Purification by flash column chromatography (0% -> 10%

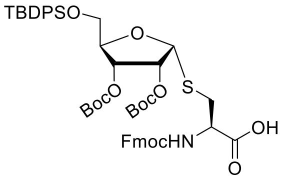
acetone/DCM) yielded the title compound as a white foam (0.50 g, 0.67 mmol, 66%). **Rf** 0.40 (5% Acetone/DCM). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.82 - 7.46 (m, 8H, Fmoc arom.+TBDPS arom.), 7.45 - 7.08 (m, 10H, Fmoc arom.+TBDPS arom.), 6.57 (d, *J* = 8.7 Hz, 1H, N*H*), 5.86 (ddt, *J* = 16.2, 10.7, 5.7 Hz, 1H, All), 5.41 (d, *J* = 5.4 Hz, 1H, *H*-1'), 5.33 - 5.12 (m, 2H. All), 4.73 (dt, *J* = 8.9, 4.7 Hz, 1H, $C_{\alpha}H$), 4.62 (d, *J* = 5.7 Hz, 2H, All), 4.47 - 4.31 (m, 2H, *H*-2'+ H-5'$_a$), 4.29 (t, *J* = 4.8 Hz, 1H, *H*-3'), 4.24 - 4.12 (m, 2H, *H*-4'+ C*H* Fmoc), 4.05 (t, *J* = 7.2 Hz, 1H, H-5'$_b$), 3.82 (d, *J* = 3.2 Hz, 2H, CH_2 Fmoc), 3.73, 3.32 (dd, *J* = 14.5, 5.3 Hz, 1H, $C_{\beta}H_a$), 2.99 (dd, *J* = 14.5, 3.8 Hz, 1H, $C_{\beta}H_b$), 1.01 (s, 9H, TBDPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.4 (C=O All), 156.4 (C=O Fmoc), 143.9, 143.7, 141.2 (*C*q. arom.), 135.5, 135.5 (*C*H arom.), 132.9, 132.9 (*C*q. arom.), 131.47 (All), 129.84, 129.78, 127.78, 127.74, 127.09, 127.03, 125.20, 125.08, 119.9 (*C*H arom.), 119.0 (All), 90.4 (*C*-1'), 83.9 (*C*-4'), 72.5 (*C*-2'), 71.5 (*C*-3'), 67.2 (*C*-5'), 66.3 (All), 63.7 (CH_2 Fmoc), 54.3 (C_{α}H), 47.0 (*C*H Fmoc), 34.5 ($C_{\beta}H_2$), 26.8 (*t*Bu TBDPS), 19.2 (*C*q. TBDPS) **HRMS** [$C_{42}H_{47}NO_8SSi + H^+$] found: 754.2862, calculated: 754.2864.

1-*S*-2,3-bis-*O*-(*tert*-butyloxycarbonyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl cysteine allyl ester (**13**)



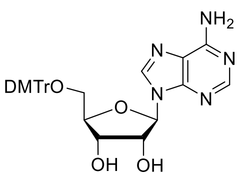
Compound **12** (506 mg, 0.67 mmol) was dissolved in pyridine and DCM (1:8 v/v, 6.7 mL, 0.1 M). TEA (0.33 mL, 2.35 mmol, 3.5 eq.), Boc_2O (0.47 mL, 2.02 mmol, 3.0 eq.) and DMAP (16 mg, 0.13 mmol, 0.2 eq.) were added and the solution was stirred for 15 minutes whereupon TLC analysis showed full conversion of the starting material. The reaction was quenched with 1 M aq. HCl and extracted with DCM (2x), dried over $MgSO_4$, filtered, concentrated *in vacuo* and loaded on celite. Purification by flash column chromatography (10% -> 20% -> 30% EtOAc/pentane) yielded the title compound as a white foam (495 mg, 0.52 mmol, 77%). **Rf** 0.65 (25% EtOAc/pentane). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.77 - 7.48 (m, 8H, Fmoc arom.+TBDPS arom.), 7.42 - 7.18 (m, 10H, Fmoc arom.+TBDPS arom.), 6.36 (d, *J* = 8.7 Hz, 1H, NH), 5.90 (ddt, *J* = 16.3, 10.7, 5.8 Hz, 1H, All), 5.56 (d, *J* = 5.1 Hz, 1H, *H*-1'), 5.37 - 5.18 (m, 4H, *H*-2'+ *H*-3'+All), 4.75 (dt, *J* = 8.6, 4.2 Hz, 1H, $C_{\alpha}H$), 4.65 (d, *J* = 5.7 Hz, 2H, All), 4.45 - 4.30 (m, 3H, H-5'$_a$+C*H* Fmoc), 4.20 - 4.10 (m, 1H, H-5'$_b$), 4.07 (q, *J* = 7.3, 6.4 Hz, 1H, *H*-4'), 3.84 (d, *J* = 2.8 Hz, 2H, CH_2 Fmoc), 3.40 (dd, *J* = 14.5, 4.9 Hz, 1H, $C_{\beta}H_{2a}$), 3.01 (dd, *J* = 14.5, 3.8 Hz, 1H, $C_{\beta}H_{2b}$), 1.50 (s, 18H, Boc), 1.03 (s, 9H, TBDPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.0 (C=O All), 156.1 (*C*=O Fmoc), 152.5, 152.2 (*C*=O Boc), 143.9, 143.7, 141.2, 141.2 (*C*q. arom.), 135.5 (*C*H arom.), 132.9, 132.8 (*C*q. arom.), 131.6 (All), 129.8, 129.7, 127.8, 127.6, 127.6, 127.1, 127.0, 125.2, 125.1, 119.9 (*C*H arom.), 118.9 (All), 88.0 (*C*-1'), 83.0, 82.8 (*C*q. Boc), 80.7 (*C*H Fmoc), 73.8, 72.9 (*C*-2'/3'), 67.0 (*C*-5'), 66.3 (All), 63.0 (CH_2 Fmoc), 54.2 (C_{α}H), 47.0 (*C*-4'), 34.9 ($C_{\beta}H_2$), 27.7 (*t*Bu Boc), 26.8 (*t*Bu TBDPS), 19.2 (*C*q. TBDPS). **HRMS** [$C_{52}H_{63}NO_{12}SSi + H^+$] found: 954.3909, calculated: 954.3913.

1-*S*-2,3-bis-*O*-(*tert*-butyloxycarbonyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl cysteine (**14**)



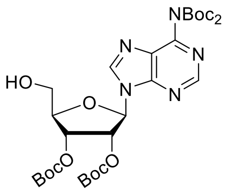
Fully protected thioglycoside **13** (509 mg, 0.53 mmol) was dissolved in DCM (5.3 mL, 0.1 M). DMBA (100 mg, 0.64 mmol, 1.2 eq.) and a catalytic amount of $Pd(PPH_3)_4$ were added and the mixture was stirred for 1 hour before TLC analysis showed full conversion of the starting material. The reaction was diluted with 10% w/w aq. citric acid and extracted with DCM (2x). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by flash column chromatography (0% -> 0.5% ->10% MeOH/DCM) yielded the title compound as a white foam (488 mg, 0.52 mmol, quant.). **Rf** 0.22 (5% MeOH/DCM). **^{1}H NMR** (500 MHz, $CDCl_3$ + MeOD) δ 7.77 - 7.47 (m, 10H, arom. Fmoc+arom. TBDPS), 7.42 - 7.20 (m, 8H, arom. Fmoc+arom. TBDPS), 6.59 (d, *J* = 7.9 Hz, 0.3H, N*H*), 5.60 (d, *J* = 4.6 Hz, 1H, *H*-1'), 5.39 - 5.13 (m, 2H, *H*-2'/3'), 4.48 - 4.42 (m, 1H, $C_\alpha H$), 4.34 - 4.30 (m, 1H, C*H* Fmoc), 4.29 - 4.22 (m, 3H, H-5'$_a$/5'$_b$+$C_\alpha H$+C*H* Fmoc), 4.17 (dd, J = 10.4, 7.4 Hz, 1H, H-5'$_b$+*H*-4'), 4.03 (t, J = 7.2 Hz, 1H, *H*-4'), 3.80 (d, *J* = 3.0 Hz, 2H, CH_2 Fmoc), 3.39 - 3.32 (m, 1H, $C_\beta H_{2a}$), 3.07 (dd, *J* = 14.0, 3.9 Hz, 1H, $C_\beta H_{2b}$), 1.51 - 1.41 (m, 18H, *t*Bu Boc), 1.02 (s, 9H, *t*Bu TBPDS).**^{13}C NMR** (126 MHz, $CDCl_3$ + MeOD) δ 175.04 (*C*=O COOH), 156.56 (*C*=O Fmoc), 152.43, 152.19 (*C*=O Boc), 143.69, 143.66, 141.05 (*C*q. arom.), 135.37 (*C*H arom.), 132.66, 132.61 (*C*q. arom.), 132.32, 132.30, 131.88, 131.79, 129.65, 129.61, 128.65, 128.55, 127.58, 127.50, 126.89, 124.96, 119.70 (*C*H arom.), 87.83 (*C*-1'), 82.97, 82.86 (*C*q. Boc), 80.15 (*C*H Fmoc), 73.96, 72.96 (*C*-2'/3'), 66.80 (*C*-5'), 62.71 (CH_2 Fmoc), 54.74 (C_αH), 46.83 (*C*-4'), 34.73 ($C_\beta H_2$), 27.41, 27.39 (*t*Bu Boc), 26.52 (*t*Bu TBPDS), 18.99 (*C*q. TBDPS). **HRMS** [$C_{49}H_{59}NO_{12}SSi + H^+$] found: 914.3596, calculated: 914.3600.

5-*O*-(4,4'-Dimethoxytrityl) adenosine (**18**)



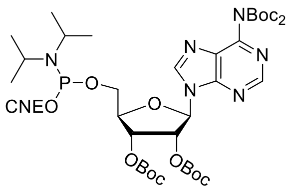
Adenosine (16.3 g, 40 mmol) was dissolved in pyridine and DMF (1:1 v/v, 80 mL, 0.5 M) and cooled to 0˚C. DMTr-Cl (16.3 g, 48 mmol. 1.2 eq.) was added in portions over 8 hours. Upon completion, the solution was poured in cold H_2O (800 mL). The resulting yellow precipitate was filtered off, washed with cold H_2O and dried on air. The crude product was recrystallized from hot xylene (1000 mL), giving a white precipitate after cooling to 0 °C which was filtered and washed with cold Et_2O. The filtrate was collected and dried in a high vacuum at 60˚C, giving the title compound as a white solid (8.46 g, 14.6 mmol, 37%). Spectra were in full accordance with the literature experimental data.[28]

2,3-bis-*O*-(*tert*-butyloxycarbonyl)-*N*^{6}-di(*tert*-butyloxycarbonyl) adenosine (**19**)



Adenosine derivative **18** (5.82 g, 10 mmol) was dissolved in DCM (100 mL, 0.1 M). TEA (11.4 mL, 82 mmol, 8 eq.), Boc_2O (14 mL, 60 mmol, 6 eq.) and DMAP (0.25 g, 2.0 mmol, 0.2 eq.) were added and the solution was allowed to stir over night. After TLC analysis showed full conversion of the starting material, the reaction was diluted with DCM, washed with sat. aq. $NaHCO_3$, dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude mixture was dissolved in DCM (146 mL, 0.07M). TIS (3.14 mL, 15 mmol, 1.5 eq.) and TFA (3.15 mL, 41 mmol, 4 eq.) were added and the mixture was stirred for 30 minutes. Upon completion as determined by TLC analysis, the reaction was quenched by addition of sat. aq. $NaHCO_3$ and extracted with DCM. The combined organic layers were washed with brine, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by flash column chromatography (20% -> 30% -> 50% EtOAc/pentane) yielded the title compound as a white foam (4.63 g, 6.9 mmol, 68%). **Rf** 0.49 (40% EtOAc/pentane). Spectra were in full accordance with the literature experimental data.[19]

5'-*O*-(2-cyanoethyl-*N*-*N*-diisopropylphosphoramidite)-2,3-bis-*O*-(*tert*-butyloxycarbonyl)-*N*^{6}-di(*tert*-butyloxycarbonyl) adenosine (**20**)



Compound **19** (4.63 g, 6.94 mmol) was dissolved in DCM (70 mL, 0.1 M). DIPEA (2.42 mL, 13.9 mmol, 2 eq.) and 2-cyanoethyl-*N*-*N*-diisopropylchlorophosphoramidite (1.70 mL, 7.63 mmol, 1.1 eq.) were added and the solution was stirred for 30 minutes until TLC analysis showed full conversion of the starting material. The reaction was quenched with MeOH and concentrated *in vacuo*. Purification by flash column chromatography (20% -> 30% EtOAc/pentane + 0.5% TEA) yielded the title compound as a white foam (5.84 g, 6.7 mmol, 97%). **Rf:** 0.49 (30% EtOAc/pentane). Spectra were in full accordance with literature experimental data.[19]

Solid-phase peptide synthesis

The amino acids (obtained from Novabiochem and Sigma Aldrich) applied in the synthesis were: Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Glu(O*t*Bu)-OH, Fmoc-Glu(*O*-2-Ph*i*Pr)-OH, Fmoc-Gly-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Lys(Mmt)-OH, Fmoc-Pro-OH, Fmoc-Ser(O*t*Bu)-OH, Fmoc-Ser(Trt)-OH, Fmoc-Thr(O*t*Bu)-OH, Fmoc-Thr(Trt)-OH, Fmoc-Cys(Trt)-OH. 5-carboxyfluorescein was bought from Bio-Connect B.V. TentaGel® S AC was bought from Rapp-Polymere GmbH and loaded by hand with the appropriate Fmoc-amino acid. ADPr peptides were synthesized with automated solid phase peptide synthesis on an CEM Liberty Blue Automated Microwave Peptide Synthesizer.

Automated synthesis of ADPr-peptides was performed on 25 or 50 µmol scale. The first amino acid was manually loaded on the resin using 2 eq. Fmoc-AA-OH, 2 eq. DIC and cat. DMAP while shaking overnight. Resin was first swollen for 5 minutes in DMF prior to amino acid coupling. Activation was achieved using DIC/Oxyma. Standard coupling was achieved using 5 eq. Fmoc-amino acid as a 0.2 M in DMF solution, 5 eq. DIC as a 0.5 M of DIC in DMF solution and 5 eq. Oxyma as a 1M solution in DMF which was buffered by DIPEA (0.1M) at room temperature for 60 minutes. Non-commercially available ribosylated Fmoc-amino acids were coupled using 3 eq. amino acid as a 0.1M solution in DMF, 3 eq. DIC as a 0.5M solution in DMF and 3 eq. Oxyma as a 1 M solution in DMF which was buffered by DIPEA (0.1M) at room temperature for 120 minutes. Standard Fmoc deprotection was achieved by 20% v/v piperidine/DMF at RT for 10 minutes (2 cycles). Synthesis quality could be monitored by UV absorption of dibenzofulvene released during Fmoc deprotection.

Coupling of (5)-carboxyfluorescein

After peptide elongation, (5)-carboxyfluorescein (2.5 eq.) and PyAOP (2.5 eq.) were suspended in DMF (0.25 M). NMP was added until the suspension was fully dissolved (0.1-0.15 M) and the mixture was added to the resin. DIPEA (5 eq.) was slowly added to the resin while shaking and the mixture was agitated over night after it was washed with DMF and DCM. To cleave over-acylated CF, the resin was treated with 20% v/v in DMF piperidine and washed with DMF. The treatment was repeated after which the resin was washed with DMF and DCM. CF was protected with Trt by addition of Trt-Cl (11 eq., 0.11 M) and DIPEA (11 eq. 0.11 M) in DCM to the resin, after which the mixture was shaken for 2 hours, followed by washing with DCM.

On-resin deprotection and phosphorylation

The resin was treated with TBAF in THF (1M, 10 mL /g resin) for 30 minutes while shaking. The resin was washed with DMF and the treatment was repeated once, after which it was extensively washed with DMF and DCM yielding the desilylated ribosyl peptide intermediate. Hereafter, the resin was washed extensively with MeCN and flushed with N_2 to remove all traces of water. A solution of **24** ($(FmO)_2PN(i\text{Pr})_2$) (5 eq., 0.25M in MeCN) was added, followed by ETT as the activator (5 eq., 0.25M in MeCN). The resin was agitated for 30 minutes, followed by washing with MeCN. A solution of CSO (8 mL /g resin, 0.5M in MeCN) was added and the mixture was shaken for 30 minutes, after which the resin was washed with MeCN and DCM. Hereafter, the Fm groups were deprotected by treatment of the resin with a 10% v/v solution of DBU in DMF for 15 minutes followed by washing with DMF. The treatment was repeated after which the resin was washed with DMF, yielding the phosphoribosylated intermediate.

On-resin pyrophosphate construction

The resin was extensively washed with MeCN and flushed with N_2 to remove traces of water. The resin was treated with a solution of adenosine amidite **20** (3 eq., 0.13 M in MeCN) and ETT (3 eq., 0.25 M in MeCN) for 30 minutes while being agitated. The resin was thoroughly washed with MeCN before a CSO solution (0.5 M in MeCN) was added after which it was shaken for 30 minutes. The resin was washed with MeCN and DMF and a solution of DBU (10% v/v in DMF) was added after which the resin was shaken for 10 minutes. The treatment was repeated once after which the resin was washed extensively with DMF and DCM, yielding the ADPr-peptide on-resin.

Global deprotection and purification

The resin was globally deprotected and cleaved from the resin with cleavage cocktail (16 mL/g resin, 50% v/v TFA, 2.5% v/v TIS in DCM) for 1 hour, after which it was filtered into cold Et_2O (5 times the volume of cleavage cocktail). The resin was washed with cleavage cocktail (2 mL/g resin), which was again filtered into the Et_2O. The precipitated peptide was centrifuged for 5 minutes after which the supernatant was discarded. The resulting pellet was resuspended in Et_2O and again centrifuged, followed by removal of the supernatant. The crude peptide was dissolved in a NH_4OAc solution of 1:1 MilliQ:MeCN (100 mM), lyophilized and purified by HPLC. Fractions containing the peptide were collected and lyophilized.

CF-Arg-Pro-Thr-Pro-Cys(5-O-adenosine-diphosphate-α-D-ribosyl)-Ala-Pro-Leu-Ala-Glu-Cys(5-O-adenosine-diphosphate-α-D-ribosyl)-Lys-Gly-Ser-Leu-OH (**1**)

The general procedures were followed as described to 50 µmol Tentagel® S AC resin unless stated otherwise. The amino acids used were Fmoc-Arg(Pbf)-OH, Fmoc-Pro-OH, Fmoc-Thr(O*t*Bu)-OH, Fmoc-Leu-OH, Fmoc-Ala-OH, Fmoc-Glu(O*t*Bu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Gly-OH, Fmoc-Ser(O*t*Bu)-OH and ribosyl building block **14**. For phosphorylation, a solution of **24** $(FmO)_2PN(iPr)_2$ (10 eq., 0.5M in MeCN) and ETT activator (10 eq., 0.5 M in MeCN) was used. For pyrophosphorylation, a solution of adenosine amidite **20** (6 eq., 0.25 M in MeCN) and ETT (6 eq., 0.5 M in MeCN) was used. After HPLC purification, compound **1** was obtained as an orange powder (6.72 mg, 2.25 µmol, 4.5%).**^{1}H NMR** (600 MHz, D_2O + CD_3CN) δ 8.62 (s, 2H, H-2), 8.47 - 8.44 (m, 1H, CF), 8.38 - 8.32 (m, 2H, H-8), 8.19 (t, *J* = 6.6 Hz, 1H, CF), 7.55 - 7.49 (m, 1H, CF), 7.22 (m, 2H, CF), 7.04 - 6.93 (m, 2H, CF), 6.92 - 6.83 (m, 2H, CF), 6.20 (m, 2H, H-1'), 5.72 - 5.63 (m, 0.2H, H-1"), 5.58 (d, *J* = 5.2 Hz, 1.3H, H-1"), 5.18 - 5.13 (m, 0.1H, H-1"), 5.11 (d, *J* = 5.4 Hz, 0.4H, H-1"). **^{31}P NMR** δ -10.3, -10.4, -10.4, -10.5, -10.6, -10.6, -10.6, -10.7. **LCMS** (10% -> 50% MeCN) Rt = 4.8 min. **HRMS** [$C_{116}H_{163}N_{29}O_{52}P_4S_2$ + $3H^+$] found: 995.3212, calculated: 995.3214.

CF-Pro-Leu-Ala-Glu-Cys(5-O-adenosine-diphosphate-α-D-ribosyl)-Lys-Gly-Ser-Leu-OH (**2**)

The general procedures described in Voorneveld *et al.*[17] were applied to 25 µmol Tentagel® S AC resin unless stated otherwise. The amino acids used were Fmoc-Pro-OH, Fmoc-Leu-OH, Fmoc-Ala-OH, Fmoc-Glu(O*t*Bu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Gly-OH, Fmoc-Ser(O*t*Bu)-OH and ribosyl building block from Voorneveld *et al.*[17] After HPLC purification, compound **2** was obtained as an orange powder (1.85 mg, 0.98 µmol, 4.0%). **^{1}H NMR** (600 MHz, D_2O + CD_3CN) δ 8.55 (d, *J* = 7.4 Hz, 1H, H-2), 8.28 (d, *J* = 10.6 Hz, 1H, H-8), 8.23 (d, *J* = 1.6 Hz, 0.5H, CF), 8.03 (d, *J* = 1.6 Hz, 0.5H, CF), 7.95 (dd, *J* = 7.7, 1.7 Hz, 0.5H, CF), 7.68 (dd, *J* = 7.8, 1.6 Hz, 0.5H, CF), 7.54 - 7.44 (m, 0.5H, CF) 7.36 (m, 0.5H, CF), 7.03 - 6.94 (m, 2H, CF), 6.91 - 6.83 (m, 2H, CF), 6.80 - 6.71 (m, 2H, CF), 6.12 (dd, *J* = 5.6, 2.4 Hz, 1H, H-1'), 5.49 (d, *J* = 4.6 Hz, 1H, H-1"). **^{31}P NMR** δ -10.29, -10.47, -10.59, -10.75. **LCMS** (10% -> 50% MeCN) Rt = 6.3 min. **HRMS** [$C_{75}H_{99}N_{15}O_{32}P_2S$ + $2H^+$] found: 908.7956, calculated: 908.7961.

CF-Arg-Pro-Thr-Pro-Cys(5-O-adenosine-diphosphate-α-D-ribosyl)-Ala-Pro-Leu-Ala-OH (**3**)

The general procedures were followed as described to 25 µmol Tentagel® S AC resin unless stated otherwise. The amino acids used were Fmoc-Arg(Pbf)-OH, Fmoc-Pro-OH, Fmoc-Thr(O*t*Bu)-OH, Fmoc-Leu-OH, Fmoc-Ala-OH and ribosyl building block **14**. After HPLC purification, compound **3** was obtained as an orange powder (3.86 mg, 2.07 µmol, 8.3%). **^{1}H NMR** δ 8.56 (m, 1H, H-2), 8.38 (s, 1H, CF), 8.13 (m, 1H, H-8), 7.45 (m, 1H, CF), 7.14 (m, 2H, CF), 6.92 (d, *J* = 11.7 Hz, 2H, CF), 6.82 (dd, *J* = 15.8, 9.2 Hz, 2H, CF), 6.11 (m, 1H, H-1'), 5.53 (d, *J* = 4.8 Hz, 1H, H-1"). **^{31}P NMR** (202 MHz, D_2O) δ -10.98, -11.08, -11.19, -11.22, -11.24, -11.32. δ **LCMS** (10% -> 50% MeCN) Rt = 5.02 min. **HRMS** [$C_{76}H_{99}N_{17}O_{30}P_2S$ + $2H^+$] found: 912.8037, calculated: 912.8043.

CF-Arg-Pro-Thr-Pro-Cys-Ala-Pro-Leu-Ala-Glu-Cys(5-O-adenosine-diphosphate-α-D-ribosyl)-Lys-Gly-Ser-Leu-OH (**4**)

The general procedures were followed as described to 50 µmol Tentagel® S AC resin unless stated otherwise. The amino acids used were Fmoc-Arg(Pbf)-OH, Fmoc-Pro-OH, Fmoc-Thr(O*t*Bu)-OH, Fmoc-Leu-OH, Fmoc-Ala-OH, Fmoc-Glu(O*t*Bu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Gly-OH, Fmoc-Ser(O*t*Bu)-OH, Fmoc-Cys(Trt)-OH and ribosyl building block **14**. After HPLC purification, compound **4** was obtained as an orange powder (2.4 mg, 0.98 µmol, 2.0%).**^{1}H NMR** (500 MHz, D_2O + CD_3CN) δ 8.62 (s, 2H, H-2), 8.47 - 8.44 (m, 1H, CF), 8.38 - 8.32 (m, 2H, H-8), 8.19 (t, *J* = 6.6 Hz, 1H, CF), 7.55 - 7.49 (m, 1H, CF), 7.22 (m, 2H, CF), 7.04 - 6.93 (m, 2H, CF), 6.92 - 6.83 (m, 2H, CF), 6.20 (m, 2H, H-1'), 5.72 - 5.63 (m, 0.2H, H-1"), 5.58 (d, *J* = 5.2 Hz, 1.3H, H-1"), 5.18 - 5.13 (m, 0.1H, H-1"), 5.11 (d, *J* = 5.4 Hz, 0.4H, H-1"). **^{31}P NMR** (202 MHz, CH_3CN+D_2O) δ 1.08, -4.03, -10.33, -10.42, -10.58, -10.69. **LCMS** (10% -> 90% MeCN) Rt = 3.7 min. **HRMS** [$C_{202}H_{282}N_{48}O_{78}P_4S_4$ + $4H^+$ (disulfide)] found: 1221.4456, calculated: 1221.4438.

CF-Arg-Pro-Thr-Pro-Cys(5-O-adenosine-diphosphate-α-D-ribosyl)-Ala-Pro-Leu-Ala-Glu-Cys-Lys-Gly-Ser-Leu-OH (**5**)

The general procedures were followed as described to 50 μmol Tentagel® S AC resin unless stated otherwise. The amino acids used were Fmoc-Arg(Pbf)-OH, Fmoc-Pro-OH, Fmoc-Thr(O*t*Bu)-OH, Fmoc-Leu-OH, Fmoc-Ala-OH, Fmoc-Glu(O*t*Bu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Gly-OH, Fmoc-Ser(O*t*Bu)-OH, Fmoc-Cys(Trt)-OH and ribosyl building block **14**. After HPLC purification, compound **4** was obtained as an orange powder (2.8 mg, 1.1 μmol, 2.3%). ^{1}**H NMR** (500 MHz, D_2O +$CD_{3C}N$) δ 9.03 (s, 1H ,H-2), 8.92 (s, 1H, CF), 8.77 (s, 1H, H-8), 8.69 (d, *J* = 8.0 Hz, 1H, CF), 7.82 (d, *J* = 8.2 Hz, 1H, CF), 7.27 (s, 2H, CF), 7.18 (d, *J* = 8.9 Hz, 2H, CF), 7.10 (d, *J* = 8.7 Hz, 2H, CF), 6.56 (d, *J* = 5.5 Hz, 1H, H-1'), 5.94 (m, 1H, H-1"). ^{31}**P NMR** (202 MHz, D_2O) δ 3.66, 3.33, -1.65, -7.68, -7,78, -7.94, -7.99, -8.05. **LCMS** (10% -> 90% MeCN) Rt = 3.7 min. **HRMS** [$C_{202}H_{282}N_{48}O_{78}P_4S_4$ + $4H^+$ (disulfide)] found: 1221.4457, calculated: 1221.4438.

Fluorescent Polarization (FP) Assays

Recombinant PARP9, DTX3L and DTX3LΔN (N-terminal deletion mutant) were expressed in Sf21 insect cells and purified as previously described.[14] Complex formation of DTX3L/PARP9 or DTX3LΔN/PARP9 was carried out in an equal-molar ratio by incubating the proteins at >1 mg/mL for each on ice for ~ 4 hours. Fluorescently labeled AR peptides (50 nM final concentration) **1**-**5** were mixed with various concentrations of the individual recombinant proteins or the protein complexes to a final concentration of 2,500 nM for **Figure 3A** or 2,000 nM for **Figure 3B**, and 2-fold serial dilution) in a reaction buffer containing 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.1mM EDTA and 2mM DTT, incubated at room temperature for 45 min, and then measured in a PHERAstar FSX Microplate Reader (BMG Labtech). The assays were done in triplicate in a 384-well plate (Corning 3575). Competition assays were performed with 50 μM ADP-ribose. After correction for signals due to fluorescent AR peptide alone, polarization values were plotted and analyzed with Prism (GraphPad v 9.3.1) using nonlinear fit for specific binding with Hill slope.

References

[1] Y. Zhang, J. Wang, M. Ding, Y. Yu. Site-specific characterization of the Asp-and Glu-ADP-ribosylated proteome. *Nat. Methods* **2013**, *10*, 981-984.

[2] R. Bredehorst, K. Wielckens, A. Gartemann, H. Lengyel, K. Klapproth, H. Hilz. Two different types of bonds linking single ADP-ribose residues covalently to proteins: quantification in eukaryotic cells. *Eur. J. Biochem*. **1978**, *92*, 129-135.

[3] L. Palazzo, O. Leidecker, E. Prokhorova, H. Dauben, I. Matic, I. Ahel. Serine is the major residue for ADP-ribosylation upon DNA damage. *Elife* **2018**, *7*, 1-12.

[4] D. M. Leslie Pedrioli, M. Leutert, V. Bilan, K. Nowak, K. Gunasekera, E. Ferrari, R. Imhof, L. Malmström, M. O. Hottiger. Comprehensive ADP-ribosylome analysis identifies tyrosine as an ADP-ribose acceptor site. *EMBO reports*, **2018**, *19*.

[5] S. Laing, F. Koch-Nolte, F. Haag, F. Buck. Strategies for the identification of arginine ADP-ribosylation sites. *J. Proteomics* **2011**, *75*, 169-176.

[6] J. Moss, M. Vaughan. Mechanism of action of choleragen. Evidence for ADP ribosyltransferase activity with arginine as an acceptor. *J. Biol. Chem*. **1977**, *252*, 2455-2457.

[7] S. C. Buch-Larsen, A. K. L. F. S. Rebak, I. A. Hendriks, M. L. Nielsen. Temporal and site-specific ADP-ribosylation dynamics upon different genotoxic stresses. *Cells* **2021**, *10*, 1-16.

[8] K. M. Rodriguez, S. C. Buch-Larsen, I. T. Kirby, I. R. Siordia, D. Hutin, M. Rasmussen, D. M. Grant, L. L. David, J. Matthews, M. L. Nielsen, M. S. Cohen. Chemical genetics and proteome-wide site mapping reveal cysteine MARylation by PARP-7 on immune-relevant protein targets. *Elife* **2021**, *10*, 1-94.

[9] S. C. Buch-Larsen, I. A. Hendriks, J. M. Lodge, M. Rykær, B. Furtwängler, E. Shishkova, M. S. Westphall, J. J. Coon, M. L. Nielsen. Mapping physiological ADP-ribosylation using activated ion electron transfer dissociation. *Cell Rep.* **2020**, *32*.

[10] S. Vyas, I. Matic, L. Uchima, J. Rood, R. Zaja, R. T. Hay, I. Ahel, P. Chang. Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nat. Commun.* **2014**, *5*.

[11] L. J. McDonald, J. Moss. Enzymatic and nonenzymatic ADP-ribosylation of cysteine. *Mol. Cell Biochem.* **1994**, *138*, 221-226.

[12] C. S. Yang, K. Jividen, T. Kamata, N. Dworak, L. Oostdyk, B. Remlein, Y. Pourfarjam, I. K. Kim, K. P. Du, T. Abbas, N. E. Sherman, D. Wotton, B. M. Paschal. Androgen signaling uses a writer and a reader of ADP-ribosylation to regulate protein complex assembly. *Nat. Commun.* **2021**, *12*, 1-18.

[13] T. Kamata, C. S. Yang, B. M. Paschal. PARP7 mono-ADP-ribosylates the agonist conformation of the androgen receptor in the nucleus. *Biochem. J.* **2021**, *478*, 2999-3014.

[14] Y. Ashok, C. Vela-Rodríguez, C. Yang, H. I. Alanen, F. Liu, B. M. Paschal, L. Lehtiö. Reconstitution of the DTX3L-PARP9 complex reveals determinants for high-affinity heterodimerization and multimeric assembly. *Biochem. J.* **2022**, *479*, 289-304.

[15] C. S. Yang, K. Jividen, A. Spencer, N. Dworak, L. Ni, L. T. Oostdyk, M. Chatterjee, B. Kuśmider, B. Reon, M. Parlak, V. Gorbunova, T. Abbas, E. Jeffery, N. E. Sherman, B. M. Paschal. Ubiquitin modification by the E3 ligase/ADP-ribosyltransferase Dtx3L/Parp9. *Mol. Cell* **2017**, *66*, 503-516.

[16] C. Yang, K. Wierbiłowicz, N. M. Dworak, S. Y. Bae, S. B. Tengse, N. Abianeh, J. M. Drake, T. Abbas, A. Ratan, D. Wotton, B. M. Paschal. Induction of PARP7 creates a vulnerability for

growth inhibition by RBN2397 in prostate cancer cells. *Cancer Res. Commun.* **2023**, *3*, 592-606.

[17] J. Voorneveld, J. G. M. Rack, L. van Gijlswijk, N. J. Meeuwenoord, Q. Liu, H. S. Overkleeft, G. A. van der Marel, I. Ahel, D. V. Filippov. Molecular tools for the study of ADP-ribosylation: A unified and versatile method to synthesise native mono-ADP-ribosylated peptides. *Chem. Eur. J.* **2021**, *27*, 10621-10627.

[18] A. G. Volbeda, H. A. V. Kistemaker, H. S. Overkleeft, G. A. Van Der Marel, D. V. Filippov, J. D. C. Codée. Chemoselective cleavage of p-methoxybenzyl and 2-naphthylmethyl ethers using a catalytic amount of HCl in hexafluoro-2-propanol. *J. Org. Chem.* **2015**, *80*, 8796-8806.

[19] N. Hananya, S. K. Daley, J. D. Bagert, T. W. Muir. Synthesis of ADP-ribosylated histones reveals site-specific impacts on chromatin structure and function. *J. Am. Chem. Soc.* **2021**, *143*, 10847-10852.

[20] P. J. Stahl, J. C. Cruz, Y. Li, S. M. Yu, K. Hristova. On-the-resin N-terminal modification of long synthetic peptides. *Anal. Biochem.* **2012**, *424*, 137-139.

[21] R. Fischer, O. Mader, G. Jung, R. Brock, Extending the applicability of carboxyfluorescein in solid-phase synthesis. *Bioconjug. Chem.* **2003**, *14*, 653-660.

[22] H. Gold, P. van Delft, N. Meeuwenoord, J. D. C. Codee, D. V. Filippov, G. Eggink, H. S. Overkleeft, G. A. van der Marel. Synthesis of sugar nucleotides by application of phosphoramidites. *J. Org. Chem.* **2008**, *73*, 9458-9460.

[23] J. Voorneveld, M. S. Kloet, S. Wijngaarden, R. Q. Kim, A. Moutsiopoulou, M. Verdegaal, M. Misra, I. Đikić, G. A. van der Marel, H. S. Overkleeft, D. V. Filippov, G. J. van der Heden van Noort. Arginine ADP-ribosylation: Chemical synthesis of post-translationally modified ubiquitin proteins. *J. Am. Chem. Soc.* **2022**, *144*, 20582-20589.

[24] G. J. Van Der Heden Van Noort, M. G. Van Der Horst, H. S. Overkleeft, G. A. Van Der Marel, D. V. Filippov. Synthesis of mono-ADP-ribosylated oligopeptides using ribosylated amino acid building blocks. *J. Am. Chem. Soc.* **2010**, *132*, 5236-5240.

[25] H. A. V. Kistemaker, A. P. Nardozza, H. S. Overkleeft, G. A. van der Marel, A. G. Ladurner, D. V. Filippov. Synthesis and macrodomain binding of mono-ADP-ribosylated peptides. *Angew. Chem.* **2016**, *55*, 10634-10638.

[26] J. Voorneveld, J. G. M. Rack, I. Ahel, H. S. Overkleeft, G. A. Van Der Marel, D. V. Filippov. Synthetic α- and β-Ser-ADP-ribosylated peptides reveal α-Ser-ADPr as the native epimer. *Org. Lett.* **2018**, *20*, 4140-4143.

[27] N. J. Moerke. Fluorescence polarization (FP) assays for monitoring peptide-protein or nucleic acid-protein binding. *Curr. Protoc. Chem. Biol.* **2009**, *1*, 1-15.

[28] G. H. Hakimelahi, Z. A. Proba, K. K. Ogilvie. New catalysts and procedures for the dimethoxytritylation and selective silylation of ribonucleosides. *Can. J. Chem.* **1982**, *60*, 1106-1113.