Synthetic peptides, nucleic acids and molecular probes to study ADP-Ribosylation

Voorneveld, J.

Synthetic Peptides, Nucleic Acids and Molecular Probes to Study ADP-Ribosylation

Jim Voorneveld

Cover design and layout by: © evelienjagtman.com
Printed by: Ridderprint | www.ridderprint.nl

ISBN: 978-94-6458-375-5

Synthetic Peptides, Nucleic Acids and Molecular Probes to Study ADP-Ribosylation

Proefschrift

Ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 8 september 2022
klokke 15:00 uur

door

Jim Voorneveld
Geboren te Haarlem
in 1991

Promotoren:
Dr. D.V. Filippov
Prof. dr. G.A. van der Marel

Promotiecommissie:
Prof. dr. Overkleeft, Leiden University
Prof. dr. J.D.C. Codée, Leiden University
Prof. dr. D. Fiedler, Leibniz-Institut für Molekulare Pharmakologie
Prof. dr. E. Tate, Imperial College London
Prof. dr. H.J. Jessen, Albert-Ludwigs-University of Freiburg

Table of Contents

Chapter 1

General introduction 7

Chapter 2

Synthesis of α- and β-Ser-ADP-Ribosylated Peptides 27

Chapter 3

A Unified and Versatile Method to Synthesise Native Mono-ADP-Ribosylated Peptides 51

Chapter 4

Solid-Phase Synthesis of Peptides with ADP-Ribosylated Tyrosine 85

Chapter 5

A Methodology for the Preparation of Peptides ADP-Ribosylated on Arginine and Arg-ADPr Modified Human Ubiquitin 103

Chapter 6

The Synthesis of ADP-Ribosylated Nucleic Acids 129

Chapter 7

Design and Synthesis of Potential Activity Based Probes for CD38 Based on Carba-Ribofuranose 145

Chapter 8

Olaparib-Based Photoaffinity Probes for PARP-1 Detection in Living Cells 169

Chapter 9

Summary and Future Prospects 185

Appendices

Samenvatting 207
List of Publications 213
Curriculum Vitae 219
Dankwoord 223

Chapter 1

General Introduction

Biological background

ADP-ribosylation

Post-translational modifications (PTMs) enable proteins to exist in multiple isoforms, adding another layer of complexity and functionality to the protein-encoding genes. With over 300 PTMs known e.g. phosphorylation, glycosylation and acetylation, the activity or functionality of a given protein that is modified after ribosomal synthesis can be widely varied.[1] One such PTM is adenosine diphosphate ribosylation (ADP-ribosylation) which occurs by the transfer of β-nicotinamide dinucleotide adenine (β-NAD^+) onto a nucleophilic amino acid side chain of the protein target. Nicotinamide (Nam) is then expulsed, resulting in a covalent glycosidic linkage with the ribosyl moiety by inversion of stereochemistry leading to an α-configured ADP-ribose (ADPr) unit to the protein (Figure 1). ADP-ribosylation is catalyzed by the ADP-ribosyl transferase family (ARTs, otherwise known as PARPs[2]). After the initial installation of one ADPr unit (*mono*-ADP-ribosylation, MARylation), some ART family members are able to continue their transferase activity (*poly*-ADP-ribosylation, PARylation[3]) adding multiple ADPr units to the 2'-OH[4] of adenosine. In doing so, long polymers of up to 200 ADPr units can be achieved.[5–7] If PARylation occurs, the chain can be branched at the 2'-OH position of the ribosyl moiety,[4,8] adding to a complex occurrence in heterogeneity of PARylated and MARylated protein populations in the cell.

ADP-ribosylation is a dynamic PTM which can be reversed by several hydrolases. If a protein is PARylated, the protein is first trimmed down to its MARylated isoform by *poly*-ADP-ribosyl glycohydrolase (PARG),[9] capable of cleaving the glycosidic linkages of the ADPr chain but showing no activity towards MARylated proteins.[10] Depending on the amino acid that is modified with ADPr, ADPr-hydrolases (ARHs, e.g. ARH1 for arginine[11] (Arg), ARH3 for serine[9,12,13] (Ser)) then remove the final ADPr unit, returning the protein to its unmodified form.

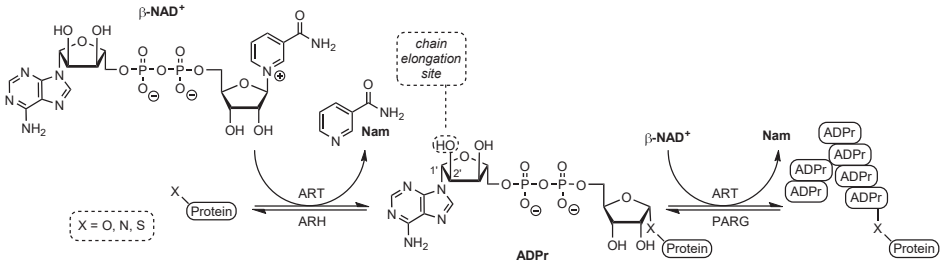


Figure 1. ADP-ribosylation of proteins on the side chains of nucleophilic amino acids by ARTs and their reverse reaction by PARG (in case of PAR) and ARHs (in case of MAR).

ADP-ribosyl transferases

ARTs are divided in subclasses based on two catalytic triad motifs[2,14]: H-Y-E and R-S-E containing transferases, both of which are structurally well conserved throughout viral, bacterial and mammalian ARTs. The H-Y-E motif was first encountered in *Corynebacterium diphtheriae* that secretes diphtheria toxin (DT).[15,16] DT is an ART that targets the diphthamide residue in elongation factor 2 with lethal effects[17] and carries this catalytic triad.[18] ARTs carrying the same H-Y-E motif are therefore named as diphtheria toxin-like ARTs (ARTDs). ARTs with the R-S-E motif are similarly named after cholera toxin (CT), an ART secreted by *Vibrio cholerae*, [19,20] and termed cholera toxin-like ARTs (ARTCs). In human cells, the ARTD subfamily is most abundant with seventeen family members, of which ARTD1 (commonly referred to as PARP1) is the most well studied.[5,21] Upon the detection of single stranded DNA breaks, PARP1 is activated and starts PARylating both itself and the core histones within seconds.[22] The resulting PAR chains are then involved in the recruitment of enzymes associated with DNA repair mechanisms.[23–26] Contrarily to ARTDs, which reside mostly in the nucleus but are found in several other compartments of the cell as well,[2,27] ARTCs are either membrane-bound (ARTC1, 3 and 4) or secreted to the exoplasm (ARTC5).[28–30] Despite these differences, mechanistically they work in a similar fashion (Figure 2). For the H-Y-E triad, encountered for example in human PARP1, the His862 residue provides hydrogen-bonding with the 2'-hydroxyl and exocyclic NH_2 on the adenosine part.[20,31,32] The Tyr896 engages in π-π stacking with the aromatic nicotinamide ring[32] and the Glu988 is involved in several processes i.e. polarization of the *N*-glycosidic bond of the distal ribose with Nam by hydrogen-bonding with its 2'-hydroxyl, stabilizing the hypothesized oxocarbenium ion[33–36] and proper positioning of the incoming nucleophile. This glutamic acid (Glu) residue is found in ARTCs carrying the R-S-E catalytic triad as well where it carries out the same function.[31,37–39] An Arg residue however provides electrostatic interactions with the pyrophosphate and the Ser helps the proper positioning of the nicotinamide ribosyl moiety by hydrogen bonds.[20,28–30] In both cases, the nucleophile is installed, forming an α-glycosidic bond with ADPr.

ADPr reversal

The importance of ADPr as a PTM is highlighted by *in vitro* studies of cells treated with ARTD inhibitors with lethal consequences backed up by *in vivo* studies with PARP1 knockout mice. [40] However, the reverse reaction can be considered equally important. The glycohydrolase PARG is solely active in the highly efficient turnover of PAR chains[9] and undegraded PAR chains, due to dysfunctional PARG, can drive neurodegeneration in Parkinson's disease[41] and embryonic lethality.[42] Another enzyme, called ARH3, was first identified in 2006 and also showed PAR degrading properties[43] albeit less efficient than PARG.[9,44] Its primary target later turned out to be Ser-linked mono-ADPr[12,13] but *O*-acetyl-ADP-ribose is efficiently hydrolyzed as well.[45] Not much is known about the substrate specificity

of ARH3 but crystal structures of this protein co-crystallized with different substrates were recently published[46] to provide insights into this process and the mechanism of hydrolysis. Unfortunately, progress in this area is hampered due to a lack of well-defined ADP-ribosylated material such as MARylated peptides. Synthetic, homogenous material can make for a great molecular tool to aid in this progress.

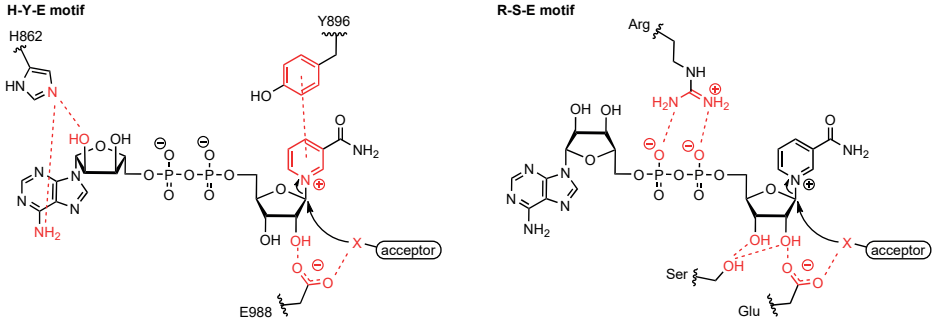


Figure 2. Mechanism of ADP-ribosylation by the ARTD class with the H-Y-E motif (left) of human PARP1 and the general ARTC class with the R-S-E motif (right). The interacting residues are colored red and the X signifies a generic nucleophile of the ADPr acceptor.

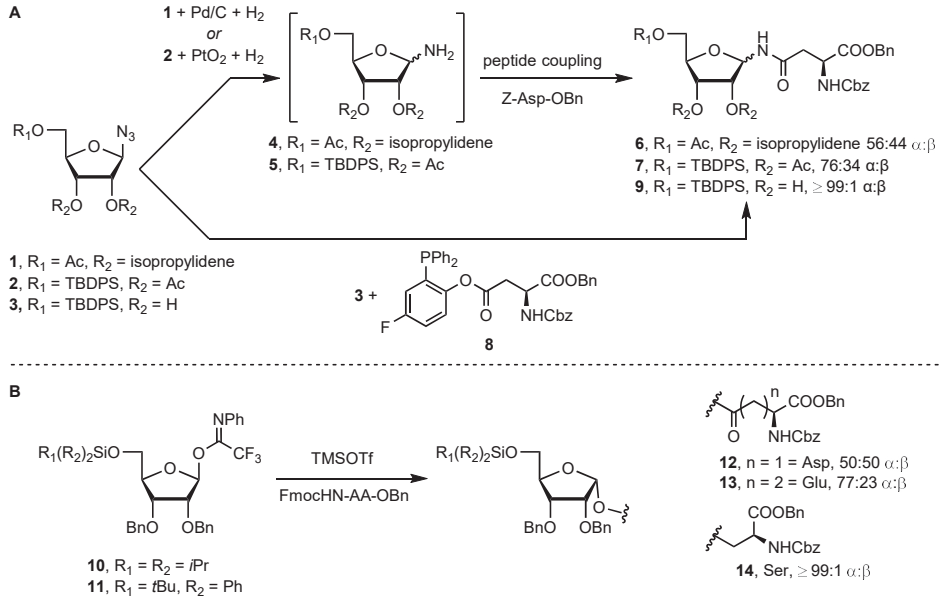
Synthetic studies on ADP-ribosylation

After its initial discovery, the PTM of ADP-ribosylation has gained increasing interest as area of biological research which led to a plethora of literature concerning its physiological role in the cell and mechanism of action.[5,20,40] In contrast, research on the acquirement of ADP-ribosylated biomolecules is still immature. Nevertheless, over the past decade, important steps have been taken to synthesize ADP-ribosylated molecules and their usefulness have been demonstrated in docking studies with macrodomain containing proteins, crystallization studies and MS-based proteomic studies as will be discussed in the next part.

α-Selective construction of ribosylated amino acids

ARTs are known to glycosylate their substrates with α-selectivity (see Figure 2). The first synthetic challenge towards ADP-ribosylated peptides is thus to construct a glycosidic bond that connects the (former) nicotinamidic ribosyl moiety with the amino acid side chain in an α-fashion. One of the earlier examples of ribosylated amino acids is found in the form of Asn-ribosyl **6** (Scheme 1A).[47] Therefore, the anomeric azide in **1** is reduced with Pd/C followed by a peptide coupling of the intermediate amine with a protected asparagine (Asn) giving a 56:44 α:β anomeric mixture. Although this procedure does yield

a decent amount of α-ribofuranosylated Asn **6**, van der Heden van Noort *et al*. successfully increased the α-selectivity of this reaction sequence by both careful kinetic control of the reduction step (10 °C and PtO_2 instead of Pd/C) and a change in the protecting group pattern on the 5' OH to a more bulky silylidene group.[48] These changes increased the α-selectivity, allowing the formation of **7** which was ultimately used for the construction of the first, synthetic MARylated peptides (discussed in more detail later). Although **7** can be obtained in sufficient amounts via this peptide coupling approach, Nisic *et al*. employed a different strategy to equip glycosyl azides with anomeric amides. In this method the glycosyl azide **3** is reduced with a phosphine, appropriately functionalized with an acyl group, via a Staudinger ligation.[49] By removing the β-steering acetyl groups and treating **3** with functionalized phosphine **8**, ribosylated Asn **9** was furnished in a fully α-selective fashion.[50]



Scheme 1. Stereoselective ribosylation of amino acids via **A:** reducing an anomeric azide followed by peptide coupling or Staudinger ligation or **B**: glycosylation of a ribofuranosyl donor with an amino acid acceptor.

Although the above-described method works well for the synthesis of *N*-ribosylated Asn and glutamine, another strategy has to be developed for the synthesis of *O*-ribosylated amino acids. This prompted Kistemaker *et al*. to develop a different approach relying on the glycosylation reaction, providing a more universal approach and an overall enhancement of stereoselectivity (Scheme 1B).[51] The steric bulk of a 5'-OH silyl protecting group in

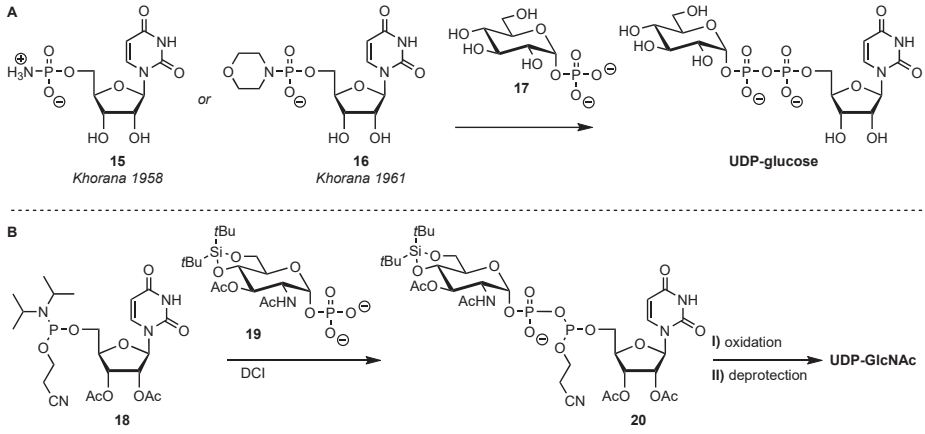
combination with benzyl ether protection of the 2'- and 3'-OH groups, promote α-selective glycosylations,[52,53] which has led to the choice of donors **10** and **11**. The glycosylation procedure was optimized by varying the protecting group patterns on the acceptors and donors as well as the use of different activators (TMSOTf, $HClO_4$-SiO_2), temperatures (-50 °C to room temperature) and solvent systems. Ultimately, these experiments led to an efficient glycosylation procedure for the α-selective formation not only of *N*-glycosylated ribofuranosides (Asn/Gln) but also of *O*-ribosylated **12** and **13** (Asp/Glu) and in particular **14** (Ser). Although α-selectivity is lost in the case of *O*-ribosylated **12** (Asp), column chromatography allowed for the isolation of the α-anomer and this is the first reported synthesis of ribosyl building blocks functionalized with Asp. Ser, although not officially confirmed as ADPr acceptor at that time, was also glycosylated with donors **10** and **11**, giving solely the α-anomer in both cases. Thus, with these results Kistemaker *et al.*[51] paved the way for the generation of a wide variety of both *N*- and *O*-ribofuranosylated amino acid building blocks, which can be used for synthesizing ADP-ribosylated peptides.

Synthesis of the pyrophosphate moiety

The second challenge in the synthesis of ADP-ribosylated molecules is the introduction of the non-symmetric pyrophosphate bridge, connecting the 5'-OH of adenosine with the 5-OH of ribosyl residue. Although a variety of methods are known to construct pyrophosphates, they can roughly be divided into two categories: i) coupling of a phosphate monoester to a different, activated phosphate ester which will henceforth be regarded as P^V – P^V coupling and ii) coupling of a phosphoramidite to a phosphomonoester (P^{III} – P^V coupling).

The P^V – P^V coupling method was pioneered by the group of Khorana who first used phosphoramidates, such as **15** to construct non-symmetric pyrophosphates in the form of UDP-glucose and FAD (Scheme 2A).[54] Not long after, they improved their method by the use of phosphormorpholidate (such as **16**) as activatable intermediate.[55] The employment of phosphormorpholidates has become the foundation of pyrophosphate chemistry ever since as is evidenced by the improvements and optimizations still being reported today (partly reviewed elsewhere[56]). The P^{III} – P^V coupling, is inspired by the extensive use of phosphoramidites in the synthesis of oligonucleotides. Early-stage development of oligonucleotide synthesis involved the condensation of deoxynucleosides with dichlorophosphines,[57,58] a procedure that was hampered by the need of temperatures of -78 °C and the formation of unwanted symmetrical and oxidized side products. The group of Caruthers came with the solution in the form of nucleoside phosphoramidites[59,60] as shelf stable synthons that provide fast reaction rates at room temperature. Shortly after their introduction, protocols with nucleoside phosphoramidites became the method of choice for automated oligonucleotide synthesis[61] and have been employed for solution-phase synthesis of nucleotide analogues as well.[62] Gold *et al.* were the first

who implemented phosphoramidites in the synthesis of pyrophosphates.[63] To efficiently obtain nucleotide sugars as substrates of glycosyltransferases, they synthesized protected 5'-phosphoramidite uridine nucleoside **18** and coupled it with the *tetra*-butylammonium salt of protected GlcNAc phosphate **19** (Scheme 2B). The resulting phosphate-phosphite intermediate **20** is then oxidized and deprotected resulting in the isolation of UDP-GlcNAc in 76% in a one-pot reaction sequence. This method of pyrophosphate construction results in a cleaner reaction profile as no unwanted symmetric pyrophosphates is formed.

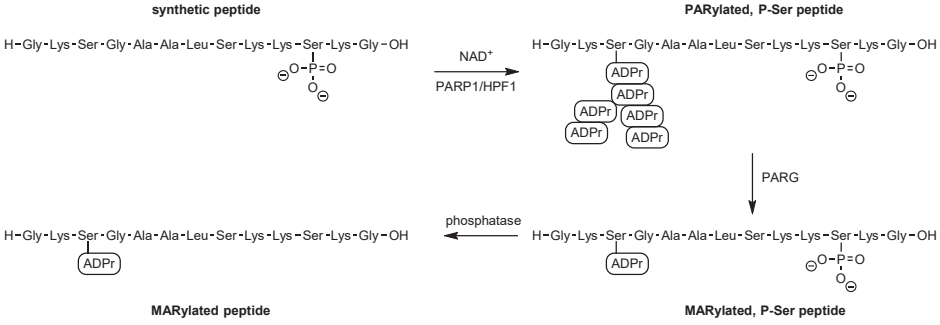


Scheme 2. Pyrophosphate construction in UDP derivatives via **A)** P^{V} – P^{V} coupling methods as pioneered by Khorana and co-workers[54,55] or **B)** P^{III} – P^{V} method as was developed by Gold *et al.*.[63]

Synthetic ADP-ribosylated peptides

As mentioned above, the availability of homogenous ADP-ribosylated biomolecules can greatly assist in studying ADP-ribosylation, either for crystallization studies, proteomics or modified ADPr analogues as molecular tools. Therefore, in the past decade a lot of research has been done towards the synthesis of fragments of biomolecules involved in ADP-ribosylation e.g., MARylated peptides, PAR chains with a well-defined chain length, trimeric branching point of ADPr chains or PARG substrates, all bearing their own synthetic challenges. In this framework, two main strategies can be discerned: i) the (chemo)enzymatic approach in which enzymes are utilized to generate ADP-ribosylated material *in vitro* and ii) a fully synthetic approach whereby protected building blocks are first synthesized and subsequently used to generate well-defined and homogenous ADP-ribosylated material. The next part will discuss the efforts made thus far to obtain synthetic ADP-ribosylated biomolecules with a focus on native MARylated peptides and analogues thereof.

When utilizing ARTs to generate ADP-ribosylated material *in vitro*, some inherent problems arise. PARP1 is a ubiquitous nuclear protein with a broad substrate selection that synthesizes PAR chains with varying polymer lengths and introducing seemingly random branching points. Therefore, a common practice in the field of studying ADP-ribosylation is the use of PARP1 followed by a second step where the proteins are treated with PARG to effectively trim down the chain length to furnish the MARylated proteins. However, PARP1 also modifies multiple amino acid side chains complicating the resulting proteomic mixture. The group of Matic elegantly addressed this problem by exploiting the recently found substrate selectivity of the PARP1:HPF1 complex towards Ser.[64,65] They used synthetic peptides found to be ADP-ribosylated with the aid of proteomics and incubated them with the PARP1:HPF1 complex and NAD^+ (Scheme 3).[66] The KS motif is specifically recognized by the PARP1:HPF1 complex[67,68] and ADP-ribosylates the Ser residue.[69] Incubation by PARG to trim the PAR chain furnished the chemoenzymatically generated MARylated peptide. Unwanted ADP-ribosylation of other Ser-residues are countered by protecting these Ser-residues with a phosphate group. In this case, a phosphatase treatment is incorporated to hydrolyze the phosphate monoesters in order to liberate Ser side chain residues.

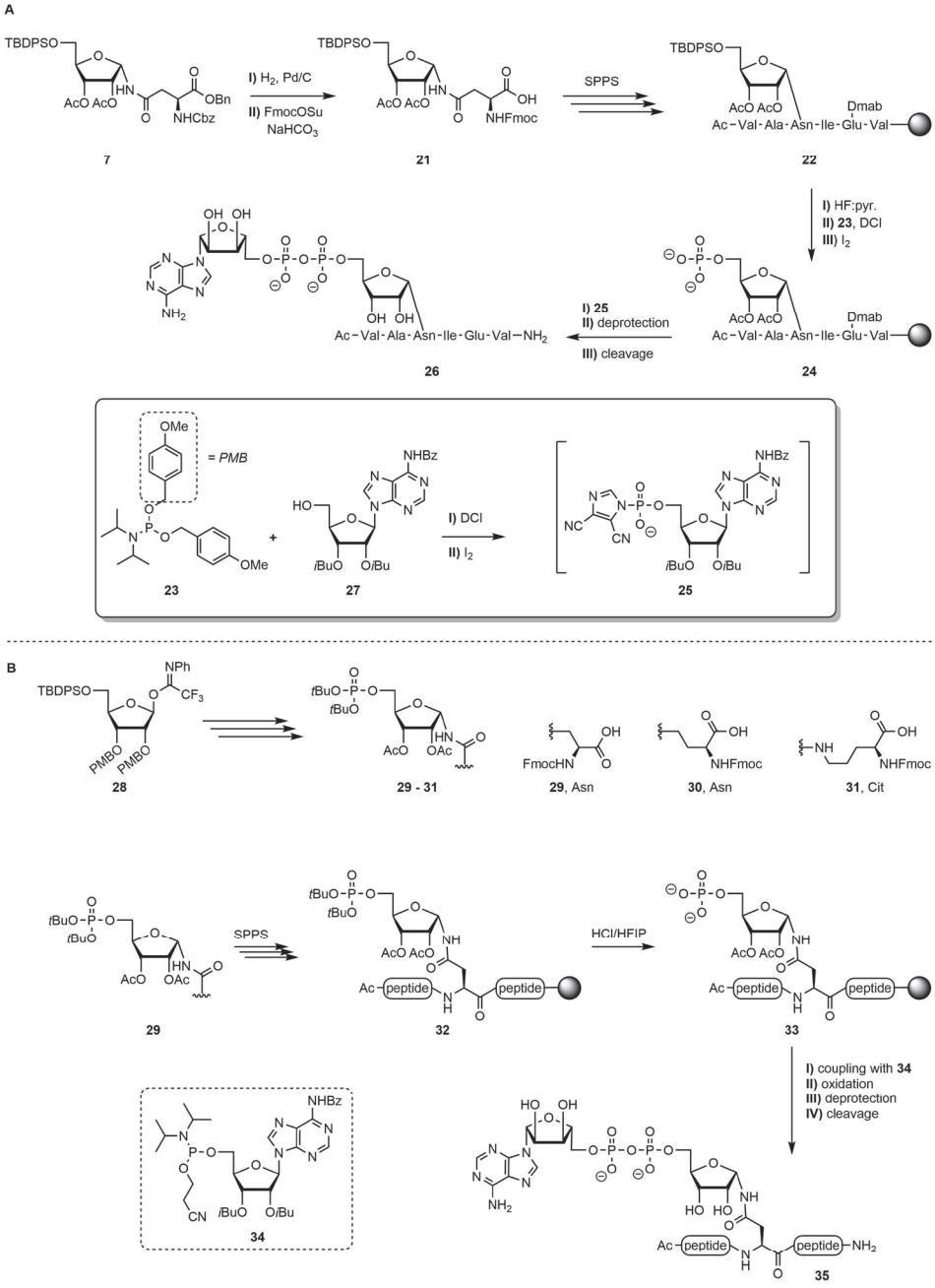


Scheme 3. Chemoenzymatic approach towards MARylated Ser-ADPr peptides developed in the group of Matic.[66]

Chemical synthesis can provide well-defined material and several studies have successfully employed fully synthetic MARylated peptides, PAR oligomers or branch points as molecular tools to study ADP-ribosylation.[70] The first report of fully synthetic MARylated peptide material originated from 2010 when van der Heden van Noort *et al.* developed ribosylated Asn building block **7** (described above, see Scheme 1A and a similar Gln analogue) to synthesize MARylated peptide **26**, Scheme 4A. To make **7** suitable for solid phase peptide synthesis (SPPS) the benzyl ether and Cbz-protecting group were removed by dehydrogenation and the released amino function was selectively masked with an Fmoc-group to furnish building block **21**. Implementation of ribosylated Asn **21** in

standard Fmoc-based SPPS led to immobilized hexamer **22**. Next, a phosphomonoester was introduced at the 5-OH of the ribose residue by desilylation, phosphitylation with phosphoramidite **23** and treatment with iodine which oxidized not only the intermediate phosphite but deprotected the resulting phosphotriester as well, to give immobilized phosphoribosylated **24**.[71] To install the ADP moiety, the phosphomonoester **24** was treated with *in situ* prepared phosphoramidate **25** and finally the resulting peptide was subjected to a set of deprotection conditions and finally cleavage from the resin, yielding the first fully synthetic ADP-ribosylated peptide **26**.

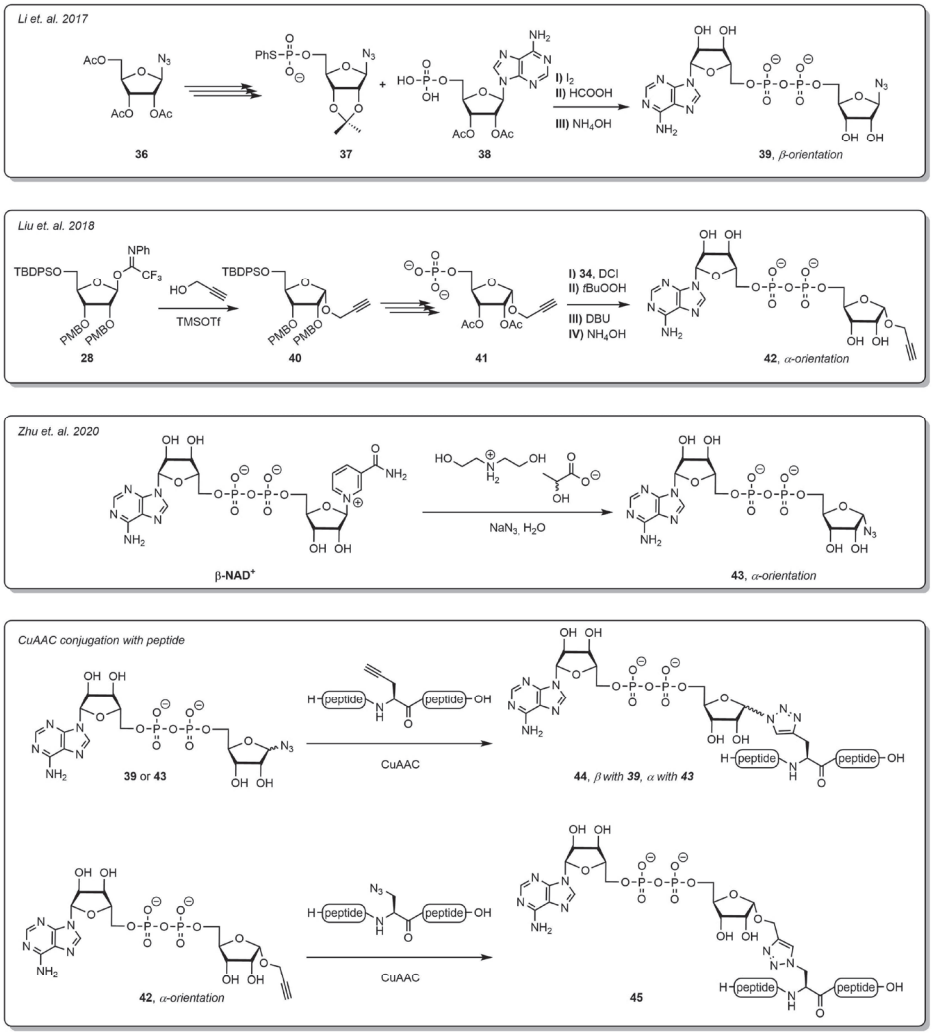
Kistemaker *et al.* furthered the field of synthetic ADPr peptides by providing the ribosylated amino acid building blocks at the 5'-OH with a tert-butyl protected phosphotriester, circumventing on-resin phosphorylation (**22** to **24**) and thereby avoiding the unwanted formation of an H-phosphonate as a result of hydrolysis during the iodine mediated oxidation.[72] To obtain these newly designed amino acid building blocks, ribose donor **28** (Scheme 4B) was coupled with the side chains of appropriately protected Asn, Gln and Cit, the latter of which can serve as an isostere for Arg. As described above, this glycosylation was also optimized to get sufficient α-selectivity and yield.[51] A multi-step synthesis route then furnished Asn, Gln and Cit building blocks **29** – **31** respectively, all of which prove to be compatible with standard Fmoc-based SPPS. Application of building blocks **29** – **31** in SPPS resulted in the formation of immobilized peptide **32**. The phosphotriester in **32** is converted to phosphomonoester **33** using HCl/HFIP, a process that conveniently was monitored by ^{31}P-NMR where phosphotriester **32** (δ≈10 ppm) shifts to phosphomonoester **33** (δ≈0 ppm). Next, the pyrophosphate moiety was introduced on-resin by adaptation of the P^{III} – P^{V} procedure of Gold *et al.* (see Scheme 2B). Accordingly, adenosine phosphoramidite **34** was coupled with immobilized phosphomonoester **33**, followed by oxidation of the phosphate-phosphite intermediate with CSO and finally removal of the cyanoethyl group with a DBU solution. To isolate the MARylated peptide the resin was then treated with a saturated NH_3 solution in 2,2,2-trifluoroethanol, resulting not only in removal of both the acetyl and iso-butyryl (*i*Bu) esters but also in cleavage from the solid support by the selective formation of the carboxamide at the *C*-terminus of the peptide. Addition of aqueous NH_4OH to the cleavage cocktail was needed to ensure complete removal of all protecting groups (i.e., Bz on the adenine ring and also possible side-chain protecting groups like trifluoroacetyl on lysine (Lys) residues) yielding generic peptide **35**. In this way, a variety of peptides was synthesized, MARylated on a Asn, Gln or Cit residue.



Scheme 4. Synthetic MARylated peptides on Asn, Gln or Cit residues via the **A)** van der Heden van Noort procedure[48] or **B)** the Kistemaker approach.[72]

Triazoles have been reported as isostere for the side chain functional group of Arg and as functionality by which generic ADP-ribosylated peptides can be obtained.[73–75] Triazoles are the products of copper catalyzed azide alkyne cyclo-additions (CuAAC), a bio-orthogonal reaction often employed in the field of chemical biology. In 2017, Li *et al.* used b-azide **36** (Scheme 5), a known ribose derivative (see also the synthesis of **1** – **3** in Scheme 1) to prepare phosphorothioate **37**.[74] Upon activation of phosphodiester **37** with I_2 in the presence of adenosine phosphate **38**, the pyrophosphate was formed and ensuing deprotection of the intermediate furnished β-azido-ADPr **39**. Guided by the fact that all known ADPr modifications are α-selective, Zhu *et al.* used an approach that deviates from conventional synthetic strategies (e.g. masking reactive groups with protecting groups) and they developed a reaction whereby a nucleophilic azide directly displaces the nicotinamide ring on β-NAD^+, mimicking the ART enzymatic reaction.[75] They found that the attack of a nucleophile could be steered towards the α-face of β-NAD^+ by the use of ionic liquids. After careful screening of 54 ionic liquids, they found that a mixture of 5:1 H_2O:[diethanolamine-lactate] resulted in the formation of a 24:1 α:β mixture which could be separated by prep-HPLC purification, yielding azide **43**. Next, oligopeptides provided with propargylglycine at the prospected modification site were synthesized and CuAAC reaction of these peptides with azides **39** or **43** in the presence of $CuSO_4$ yielded generic *N*-ADP-ribosylated peptide **44**.

Liu *et al.* reported an alternative approach in which an oligopeptide is provided with an azide at a predetermined position and the ADPr building block is functionalized with an alkyne.[73] In a high-yielding reaction, donor **28** was coupled with propargyl alcohol to give the α-configured alkyne **40**. A set of protecting group manipulations then gave phosphomonoester **41** which was used in a one-pot, three-step (coupling, oxidation and deprotection) P^{III} – P^{V} procedure with adenosine amidite **34** (see Scheme 4) to yield α-ADPr alkyne **42**. SPPS using azido-alanine furnished oligopeptides with an azide at the selected site of modification. CuAAC reaction of these peptides with ADPr building block **42** in the presence of $CuSO_4$ gave *O*-ADP-ribosylated peptide **45**. Liu *et al.* also synthesized a derivative of ubiquitin (Ub), [73] a 76 amino acid protein with azido-homoalanine on the position of Arg-42 (a known ADPr modification site) and upon CuAAC with **43**, homogenous triazole linked Ub-ADPr was isolated. This modified Ub-ADPr was recognized by ADPr-specific antibodies showing that although the triazole link is not native, it can be recognized as such in biological systems.

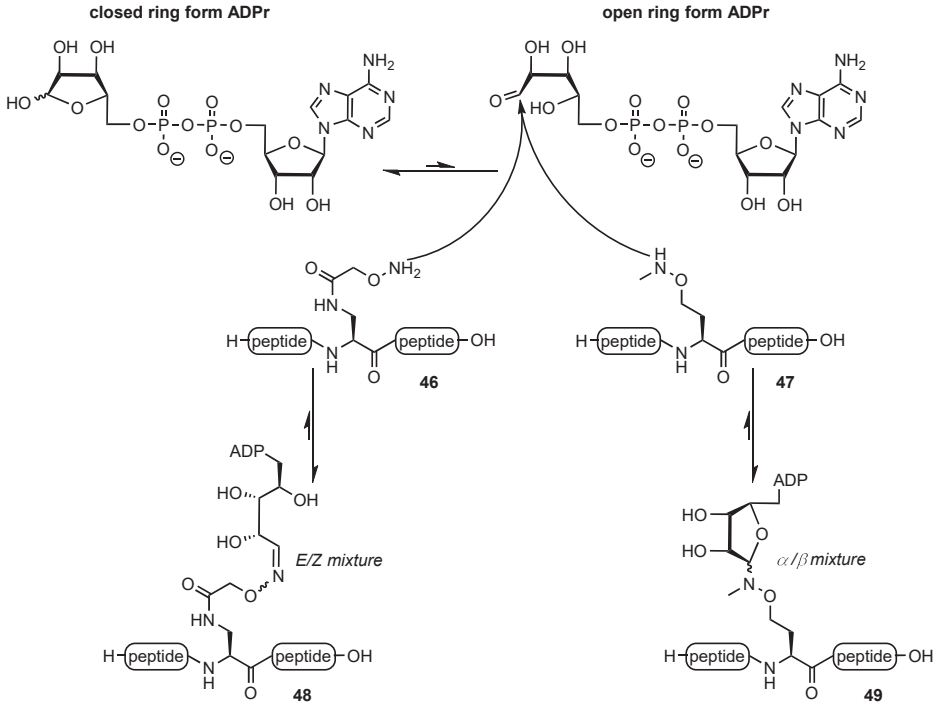


Scheme 5. Synthesis of ADPr building blocks **39**, **42** and **43** and their conjugation to synthetic peptides *via* click chemistry.

The ability to conjugate carbohydrates to peptides through oxime formation[76,77] was applied by the group of Muir to develop a chemoselective ligation method of aminooxy containing peptides with ADPr, producing another mimic of ADP-ribosylated peptides.[78] Oligopeptides corresponding to the *N*-terminus of human histone H2B were assembled with SPPS. The Glu residue in these peptides, a putative site of ADP-ribosylation, was replaced by either a protected aminooxy-containing amino acid or *N*-methyl aminooxy-containing amino acid to give peptides **46** and **47** respectively (Scheme 6). Reaction of ADPr with aminooxy **47** resulted in the formation of an E/Z mixture of the open ring isomer **48**.

On the contrary, reaction of the *N*-methyl aminooxy function in **47** with ADPr, resulted in an anomeric mixture of *N*-ADP-ribosylated peptide **49**. Chemoselectivity was achieved by performing the ligation reaction at slightly acidic pH values (around 4.5) as Lys and Arg are protonated at this pH value and therefore incapable of imine formation.

Some of the prepared ADPr peptides, equipped with a biotin tag, were applied in a pull-down assay with macrodomain containing proteins and it was found that peptide **48** efficiently pulled down mH2A1.1 (a macrodomain containing histone) and is thus able to engage in macrodomain binding. This encouraged a further investigation as to the ability of peptide **48** to bind macrodomains in complex biological systems. Nuclear HeLa cell lysates containing mH2A1.1 were treated with **48** equipped with biotin and a photo-crosslinker. After UV-irradiation to form a covalent bond, **48** was indeed able to pull down mH2A1.1 in a biological system demonstrating the usefulness of such material.



Scheme 6. Synthesis of ADP-ribosylated peptides via the oxime ligation approach. ADP = adenosine diphosphate.

Aim and outline of this Thesis

While in the past decade significant advances have been made towards synthesizing MARylated peptides, most of them however lack the native chemical linkage. Chapter 1 gives a concise summary on the developed methodologies thus far. For instance, MARylated peptides were prepared with Asn and Gln as relative stable Asp and Glu bioisosteres to prevent migration of the anomeric ester linkage to the 2'-OH position of the ribose residue.[48,72] In addition, artificial MARylated peptides, in which the ADPr residue is linked to the peptide via a triazole moiety were obtained via CuAAC, a reaction that is compatible with all natural amino acids.[73–75] Another type of artificial ADPr-peptides is the product of a chemoselective oxime ligation. MARylated peptides were prepared with either E/Z isomeric open chain ribose or an α/β-mixture of ribofuranose *N*-glycoside.

The recent discovery that Ser is the major target of PARP1 in response to DNA damage[79,80] was an incentive to develop an efficient synthetic route to peptides ADP-ribosylated on Ser. Chapter 2 describes the synthesis of phosphoribosylated Ser building blocks that are compatible with modified Fmoc-based SPPS, in which alkali sensitive semi-permanent protections are used to preserve the integrity of the ADPr moiety. With the objective to establish the native configuration of the anomeric centre of the ribose, both α- and β-configured building blocks were synthesized. With these building blocks, ADP-ribosylated peptides with either the α- or the b-epimer incorporated were prepared and used to determine that PARP1 synthesizes Ser-ADPr in an α-selective manner.

Because the alkali sensitive protecting groups are difficult to apply for the common amino acid side chain protections in SPPS, the approach from Chapter 2 is limited. Chapter 3 deals with a new synthetic methodology that addresses these limitations using acid sensitive semi-permanent protections, both for amino acid side chains as well as the 2- and 3-OH of the distal ribose. Moreover, this methodology provides less labor-intensive steps toward the ribosylated Fmoc-building blocks necessary for solid-phase synthesis and gave access not only to Ser-, but also to Thr- and Cys-ADP-ribosylated peptides. With the aid of these peptides the first Cys-ADPr specific hydrolase, the *Streptococcus pyogene* encoded protein *Spy*MacroD, was identified.

Unfortunately, the newly developed strategy proved not to be suitable for the synthesis of ADP-ribosylated peptides modified on Tyr-residues, a problem attributed to the heightened acid-sensitivity of the phenolic Tyr-ADPr glycosidic linkage. A solution was found by combining the strategies described in Chapter 2 and Chapter 3. The resulting

'hybrid' strategy uses alkali sensitive protections for the adenine base and proximal ribose residue from Chapter 2 whilst keeping the acid sensitive amino acid side chain protections from Chapter 3. The hybrid strategy described in Chapter 4 was highly efficient in synthesizing peptides, MARylated on a Tyr-residue.

Although Arg is a canonical target of ARTs, synthetic efforts have thus far been solely aimed at biomimetic Arg-ADPr linkages. Chapter 5 describes the first synthetic strategy capable of synthesizing the native Arg-ADPr linkage. With this newly developed methodology, not only oligopeptides MARylated on Arg but also Ub, ADP-ribosylated on Arg42 have been synthesized. This is the first description of a fully synthetic and functional ADP-ribosylated protein.

Recently, it has also been discovered that nucleic acids with a terminal phosphate monoester can be ADP-ribosylated *in vitro*.[81–86] Not much is known yet about the substrate specificity and which ARTs are responsible for this phenomenon. Chapter 6 is focused on the synthesis of ADP-ribosylated nucleotides in order to provide valuable chemical tools to further the studies on this newly discovered target of ARTs.

Chapter 7 deals with the design and synthesis of potential activity-based probes (ABPs) to target the membrane receptor CD38, which catalyzes the formation of cyclic ADPR from NAD^+. The structures of these ABPs are derived from NAD^+, and bear an epoxide or aziridine as a nucleophilic trap to covalently bind the active site which will allow for activity-based protein profiling (ABPP). In Chapter 8, the development of two photoaffinity probes for PARP1 is described. These probes are equipped with a diazirine containing photo affinity linker to perform photo cross linking of the probes with PARP1. The probes are used on living cells and after treatment of the cells with UV-light, PARP1 is efficiently visualized by in-gel fluorescence or pulled-down for MS-based proteomic studies. The final chapter summarizes all the research results obtained and outlines the directions of the future research.

References

[1] O. N. Jensen, *Curr. Opin. Chem. Biol.* **2004**, *8*, 33–41.

[2] M. O. Hottiger, P. O. Hassa, B. Lüscher, H. Schüler, F. Koch-Nolte, *Trends Biochem. Sci.* **2010**, *35*, 208–219.

[3] H. Juarez-Salinas, H. Mendoza-Alvarez, V. Levi, M. K. Jacobson, E. L. Jacobson, *Anal. Biochem.* **1983**, *131*, 410–418.

[4] M. Miwa, M. Ishihara, S. Takishima, N. Takasuka, M. Maeda, Z. Yamaizumi, T. Sugimura, S. Yokoyama, T. Miyazawa, *J. Biol. Chem.* **1981**, *256*, 2916–2921.

[5] B. A. Gibson, W. L. Kraus, *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 411–424.

[6] D. D'Amours, S. Desnoyers, I. D'Silva, G. G. Poirier, *Biochem. J.* **1999**, *342*, 249–268.

[7] P. Chambon, J. D. Weill, P. Mandel, *Biochem. Biophys. Res. Commun.* **1963**, *11*, 39–43.

[8] M. K. Jacobson, R. Alvarez-Gonzale, *Biochemistry* **1987**, *26*, 3218–3224.

[9] J. G. M. Rack, L. Palazzo, I. Ahel, *Genes Dev.* **2020**, *34*, 263–284.

[10] D. Slade, M. S. Dunstan, E. Barkauskaite, R. Weston, P. Lafite, N. Dixon, M. Ahel, D. Leys, I. Ahel, *Nature* **2011**, *477*, 616–622.

[11] J. Moss, S. chen Tsai, R. Adamik, H. chia Chen, S. J. Stanley, *Biochemistry* **1988**, *27*, 5819–5823.

[12] P. Fontana, J. J. Bonfiglio, L. Palazzo, E. Bartlett, I. Matic, I. Ahel, *Elife* **2017**, *6*, 1–20.

[13] J. Abplanalp, M. Leutert, E. Frugier, K. Nowak, R. Feurer, J. Kato, H. V. A. Kistemaker, D. V. Filippov, J. Moss, A. Caflisch, M. O. Hottiger, *Nat. Commun.* **2017**, *8*, DOI 10.1038/s41467-017-02253-1.

[14] H. Otto, P. A. Reche, F. Bazan, K. Dittmar, F. Haag, F. Koch-Nolte, *BMC Genomics* **2005**, *6*, DOI 10.1186/1471-2164-6-139.

[15] T. Honjo, Y. Nishizuka, O. Hayaishi, *J. Biol. Chem.* **1968**, *243*, 3553–3555.

[16] D. M. Gill, A. M. Pappenheimer, R. Brown, J. T. Kurnick, *J. Exp. Med.* **1969**, *129*, 1–21.

[17] R. J. Collier, *Toxicon* **2001**, *39*, 1793–1803.

[18] C. E. Bell, D. Eisenberg, *Adv. Exp. Med. Biol.* **1997**, *419*, 35–43.

[19] D. Cassel, T. Pfeuffer, *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 2669–2673.

[20] M. S. Cohen, P. Chang, *Nat. Chem. Biol.* **2018**, *14*, 236–243.

[21] P. O. Hassa, S. S. Haenni, M. Elser, M. O. Hottiger, *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 789–829.

[22] A. Buntz, S. Wallrodt, E. Gwosch, M. Schmalz, S. Beneke, E. Ferrando-May, A. Marx, A. Zumbusch, *Angew. Chemie Int. Ed.* **2016**, *55*, 11256–11260.

[23] A. J. Gottschalk, G. Timinszky, S. E. Kong, J. Jin, Y. Cai, S. K. Swanson, M. P. Washburn, L. Florens, A. G. Ladurner, J. W. Conaway, R. C. Conaway, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 13770–13774.

[24] G. Timinszky, S. Till, P. O. Hassa, M. Hothorn, G. Kustatscher, B. Nijmeijer, J. Colombelli, M. Altmeyer, E. H. K. Stelzer, K. Scheffzek, M. O. Hottiger, A. G. Ladurner, *Nat. Struct. Mol. Biol.* **2009**, *16*, 923–929.

[25] C. Remodeling, E. Alc, D. Ahel, Z. Ho, N. Wiechens, S. E. Polo, E. Garcia-wilson, I. Ahel, H. Flynn, M. Skehel, S. C. West, S. P. Jackson, T. Owen-hughes, S. J. Boulton, *Science* **2009**, *325*, 1240–1243.

[26] I. Ahel, D. Ahel, T. Matsusaka, A. J. Clark, J. Pines, S. J. Boulton, S. C. West, *Nature* **2008**, *451*, 81–85.

[27] S. Vyas, M. Chesarone-Cataldo, T. Todorova, Y. H. Huang, P. Chang, *Nat. Commun.* **2013**, *4*, DOI 10.1038/ncomms3240.

[28] M. Seman, S. Adriouch, F. Haag, F. Koch-nolte, *Curr. Med. Chem.* **2004**, *11*, 857–872.

[29] G. Glowacki, R. Braren, K. Firner, M. Nissen, M. Kühl, P. Reche, F. Bazan, M. Cetkovic-Cvrlje, E. Leiter, F. Haag, F. Koch-Nolte, *Protein Sci.* **2002**, *11*, 1657–1670.

[30] M. Di Girolamo, G. Fabrizio, *Biochem. Pharmacol.* **2019**, *167*, 86–96.

[31] E. E. Alemasova, O. I. Lavrik, *Nucleic Acids Res.* **2019**, *47*, 3811–3827.

[32] J. D. Steffen, J. R. Brody, R. S. Armen, J. M. Pascal, *Front. Oncol.* **2013**, *3*, 301.
[33] C. E. Bell, T. O. Yeates, D. Eisenberg, *Protein Sci.* **1997**, *6*, 2084–2096.
[34] G. C. Zhou, S. L. Parikh, P. C. Tyler, G. B. Evans, R. H. Furneaux, O. V. Zubkova, P. A. Benjes, V. L. Schramm, *J. Am. Chem. Soc.* **2004**, *126*, 5690–5698.
[35] S. L. Parikh, V. L. Schramm, *Biochemistry* **2004**, *43*, 1204–1212.
[36] P. J. Berti, S. R. Blanke, V. L. Schramm, *J. Am. Chem. Soc.* **1997**, *119*, 12079–12088.
[37] E. Barkauskaite, G. Jankevicius, I. Ahel, *Mol. Cell* **2015**, *58*, 935–946.
[38] A. Ruf, V. Rolli, G. De Murcia, G. E. Schulz, *J. Mol. Biol.* **1998**, *278*, 57–65.
[39] G. T. Marsischky, B. A. Wilson, R. J. Collier, *J. Biol. Chem.* **1995**, *270*, 3247–3254.
[40] B. Lüscher, M. Bütepage, L. Eckei, S. Krieg, P. Verheugd, B. H. Shilton, *Chem. Rev.* **2018**, *118*, 1092–1136.
[41] T. I. Kam, X. Mao, H. Park, S. C. Chou, S. S. Karuppagounder, G. E. Umanah, S. P. Yun, S. Brahmachari, N. Panicker, R. Chen, S. A. Andrabi, C. Qi, G. G. Poirier, O. Pletnikova, J. C. Troncoso, L. M. Bekris, J. B. Leverenz, A. Pantelyat, H. S. Ko, L. S. Rosenthal, T. M. Dawson, V. L. Dawson, *Science* **2018**, *362*, DOI 10.1126/science.aat8407.
[42] D. W. Koh, A. M. Lawler, M. F. Poitras, M. Sasaki, S. Wattler, M. C. Nehls, T. Stöger, G. G. Poirier, V. L. Dawson, T. M. Dawson, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17699–17704.
[43] S. Oka, J. Kato, J. Moss, *J. Biol. Chem.* **2006**, *281*, 705–713.
[44] M. Mashimo, J. Moss, *Curr. Protein Pept. Sci.* **2016**, *17*, 633–640.
[45] A. Kasamatsu, M. Nakao, B. C. Smith, L. R. Comstock, T. Ono, J. Kato, J. M. Denu, J. Moss, *J. Biol. Chem.* **2011**, *286*, 21110–21117.
[46] M. Wang, Z. Yuan, R. Xie, Y. Ma, X. Liu, X. Yu, *J. Biol. Chem.* **2018**, *293*, 14470–14480.
[47] M. A. Bonache, F. Nuti, A. Le Chevalier Isaad, F. Real-Fernández, M. Chelli, P. Rovero, A. M. Papini, *Tetrahedron Lett.* **2009**, *50*, 4151–4153.
[48] G. J. van der Heden van Noort, M. G. van der Horst, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *J. Am. Chem. Soc.* **2010**, *132*, 5236–5240.
[49] F. Nisic, G. Speciale, A. Bernardi, *Chem. - A Eur. J.* **2012**, *18*, 6895–6906.
[50] G. Speciale, A. Bernardi, F. Nisic, *Molecules* **2013**, *18*, 8779–8785.
[51] H. A. V Kistemaker, G. J. van der Heden van Noort, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2013**, *15*, 2306–2309.
[52] C. H. Larsen, B. H. Ridgway, J. T. Shaw, D. M. Smith, K. A. Woerpel, *J. Am. Chem. Soc.* **2005**, *127*, 10879–10884.
[53] G. J. van der Heden Van Noort, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2011**, *13*, 2920–2923.
[54] J. G. Moffatt, H. G. Khorana, *J. Am. Chem. Soc.* **1958**, *80*, 3756–3761.
[55] J. G. Moffatt, H. G. Khorana, *J. Am. Chem. Soc.* **1961**, *83*, 649–658.
[56] Z. Xu, *Bioorganic Med. Chem. Lett.* **2015**, *25*, 3777–3783.
[57] R. L. Letsinger, W. B. Lunsford, *J. Am. Chem. Soc.* **1976**, *98*, 3655–3661.
[58] R. L. Letsinger, V. Mahadevan, *J. Am. Chem. Soc.* **1965**, *87*, 3526–3527.
[59] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859–1862.
[60] L. J. Mcbride, M. H. Caruthers, *Tetrahedron Lett.* **1983**, *24*, 245–248.
[61] M. H. Caruthers, *Science* **1981**, *230*, 281–285.
[62] M. Kukhanova, X. Li, S.-H. Chen, I. King, T. Doyle, W. Prusoff, Y.-C. Cheng, *Mol. Pharmacol.* **1998**, *53*, 801–807.
[63] H. Gold, P. van Delft, N. Meeuwenoord, J. D. C. Codée, D. V. Filippov, G. Eggink, H. S. Overkleeft, G. A. van der Marel, *J. Org. Chem.* **2008**, *73*, 9458–9460.
[64] I. Gibbs-Seymour, P. Fontana, J. G. M. Rack, I. Ahel, *Mol. Cell* **2016**, *62*, 432–442.

[65] J. J. Bonfiglio, P. Fontana, Q. Zhang, T. Colby, I. Gibbs-Seymour, I. Atanassov, E. Bartlett, R. Zaja, I. Ahel, I. Matic, *Mol. Cell* **2017**, *0*, 1–9.
[66] J. J. Bonfiglio, O. Leidecker, H. Dauben, E. J. Longarini, T. Colby, P. San Segundo-Acosta, K. A. Perez, I. Matic, *Cell* **2020**, *183*, 1086–1102.
[67] V. Bilan, M. Leutert, P. Nanni, C. Panse, M. O. Hottiger, *Anal. Chem.* **2017**, *89*, 1523–1530.
[68] M. J. Suskiewicz, L. Palazzo, R. Hughes, I. Ahel, *FEBS J.* **2020**, 1–12.
[69] M. J. Suskiewicz, F. Zobel, T. E. H. Ogden, P. Fontana, A. Ariza, J. C. Yang, K. Zhu, L. Bracken, W. J. Hawthorne, D. Ahel, D. Neuhaus, I. Ahel, *Nature* **2020**, *579*, 598–602.
[70] G. J. van der Heden van Noort, *ACS Omega* **2020**, *5*, 1743–1751.
[71] G. J. van der Heden van Noort, C. P. Verhagen, M. G. van der Horst, H. S. Overkleeft, G. a van der Marel, D. V Filippov, *Org. Lett.* **2008**, *10*, 4461–4.
[72] H. A. V Kistemaker, A. P. Nardozza, H. S. Overkleeft, G. A. van der Marel, A. G. Ladurner, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10634–10638.
[73] Q. Liu, H. A. V. Kistemaker, S. Bhogaraju, I. Dikic, H. S. Overkleeft, G. A. van der Marel, H. Ovaa, G. J. van der Heden van Noort, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2018**, *57*, 1659–1662.
[74] L. Li, Q. Li, S. Ding, P. Xin, Y. Zhang, S. Huang, G. Zhang, *Molecules* **2017**, *22*, 1346.
[75] A. Zhu, X. Li, L. Bai, G. Zhu, Y. Guo, J. Lin, Y. Cui, G. Tian, L. Zhang, J. Wang, X. D. Li, L. Li, *Nat. Commun.* **2020**, *11*, 1–9.
[76] F. Peri, P. Dumy, M. Mutter, *Tetrahedron* **1998**, *54*, 12269–12278.
[77] S. E. Cervigni, P. Dumy, M. Mutter, *Angew. Chemie Int. Ed. Engl.* **1996**, *35*, 1230–1232.
[78] P. M. Moyle, T. W. Muir, *J. Am. Chem. Soc.* **2010**, *132*, 15878–15880.
[79] O. Leidecker, J. J. Bonfiglio, T. Colby, Q. Zhang, I. Atanassov, R. Zaja, L. Palazzo, A. Stockum, I. Ahel, I. Matic, *Nat. Chem. Biol.* **2016**, *12*, 998–1000.
[80] L. Palazzo, O. Leidecker, E. Prokhorova, H. Dauben, I. Matic, I. Ahel, *Elife* **2018**, *7*, 1–12.
[81] E. Matta, A. Kiribayeva, B. Khassenov, B. T. Matkarimov, A. A. Ishchenko, *Sci. Rep.* **2020**, *10*, 1–11.
[82] D. Munnur, I. Ahel, *FEBS J.* **2017**, *284*, 4002–4016.
[83] I. Talhaoui, N. A. Lebedeva, G. Zarkovic, C. Saint-Pierre, M. M. Kutuzov, M. V. Sukhanova, B. T. Matkarimov, D. Gasparutto, M. K. Saparbaev, O. I. Lavrik, A. A. Ishchenko, *Nucleic Acids Res.* **2016**, *44*, 9279–9295.
[84] E. A. Belousova, A. A. Ishchenko, O. I. Lavrik, *Sci. Rep.* **2018**, *8*, 1–12.
[85] G. Zarkovic, E. A. Belousova, I. Talhaoui, C. Saint-Pierre, M. M. Kutuzov, B. T. Matkarimov, D. Biard, D. Gasparutto, O. I. Lavrik, A. A. Ishchenko, *Nucleic Acids Res.* **2018**, *46*, 2417–2431.
[86] D. Munnur, E. Bartlett, P. Mikolčević, I. T. Kirby, J. G. M. Rack, A. Mikoč, M. S. Cohen, I. Ahel, *Nucleic Acids Res.* **2019**, *47*, 5658–5669.

Chapter 2

Synthesis of α- and β-Ser-ADP-Ribosylated Peptides

Introduction

Post-translational modifications (PTMs) of proteins are involved in many biological processes and entail covalent and reversible modification of specific amino acid side chains.[1] Adenosine diphosphate ribosylation (ADP-ribosylation) is a PTM that has been associated with DNA damage response pathways, transcriptional regulation, ageing and apoptosis.[2] ADP-ribosylation is catalyzed by a series of enzymes known as adenosine ribosyltransferases (ARTs).[3] In this process, ARTs consume β-NAD^+ and transfer ADP-ribose (ADPr) onto nucleophilic side chains of various amino acids, including arginine (Arg),[4] glutamate/aspartate (Glu, Asp),[5,6] and serine (Ser).[7,8] This mono-ADP-ribosylated (MARylated) site can then be modified further by some ARTs forming long and also branched poly-ADPr (PAR) chains.

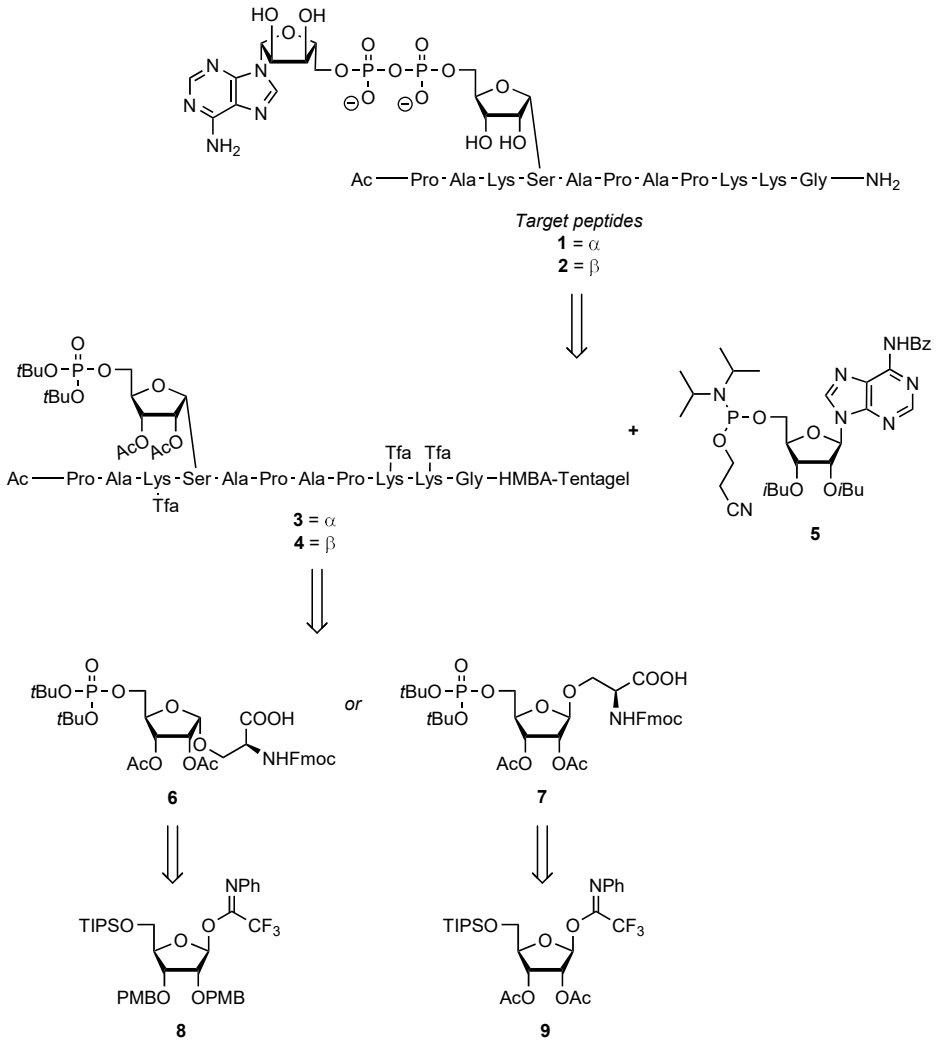
Recently, Ser residues were shown to be both acceptors of ADPr[8] and the primary target of ADP-ribosylation upon DNA damage.[9] Many of the proteins identified as Ser-ADP-ribosylated are critical for the maintenance of genome stability.[10,11] Moreover, amino acid selectivity in this pathway appears to be dependent on a switch of the catalytic preference of the ARTs PARP1 and PARP2 from Glu/Asp toward Ser residues upon the formation of a newly identified PARP1:HPF1 complex.[10,12] Transfer of ADPr by PARP1 onto Glu and Asp residues involves stereochemical inversion at the anomeric carbon and results in the formation of a-glycosidic linkages as the sole products of the reaction.[13] However, in case of Ser-ADPr and the previously unknown PARP1:HPF1 complex, this α-selectivity is not yet confirmed. Studies of this complex as well as other ADP-ribosylation systems has been restrained by the unavailability of sufficient quantities of well-defined fragments of Ser-ADPr containing peptides. This chapter describes the preparation of such Ser-ADPr peptides from phosphoribosylated Fmoc-Ser building blocks designed specifically for the incorporation in peptide chains via a modified Fmoc SPPS. These building blocks were applied in a solid phase approach similar to the one previously described for the synthesis of ADPr-peptides that contained Gln-, Asn-, or Cit-ADPr,[14] which are close isosteres of the native Glu-, Asp- and Arg-ADPr linkages. In this way, an *N*-terminal H2B peptide was prepared that contains Ser10, which is a confirmed *in vivo* ADPr acceptor site.[8,11] Preparation of the Ser-ADPr H2B peptide allowed for unambiguously determining the anomeric configuration of the ribose residue and thus to gain early mechanistic insights into the Ser-ADPr biochemistry.

Results and discussion

As shown in Scheme 1, retrosynthetic analysis of the target mono-ADP-ribosylated peptides **1** and **2** relies on the P^{III} – P^{V} method as described in Chapter 1. The intermediary phosphoribosylated peptides **3** and **4** (α- and β-configured respectively) can be coupled with adenosine 5'-phosphoramidite **5** as phosphorylating agent.[15] Adenosine amidite **5** as well as intermediates **3** and **4** bear base-labile protecting groups (e.g. acetyl/isobutyryl, Ac and *i*Bu, on the 2' and 3' hydroxyls and trifluoroacetyl, Tfa, on the side chain of lysines (Lys)) which are to be removed in the final stage of the synthesis. To furnish **3** and **4**, key phosphoribosylated Ser building blocks **6** (α) and **7** (β) were chosen to allow for the orthogonal cleavage of the benzyl ester prior to SPPS. Furthermore, the phosphates in **6** and **7** are protected with *tert*-butyl (*t*Bu) groups to enable the on-resin orthogonal deprotection and ensuing pyrophosphate formation via a procedure developed by Kistemaker *et al*.[14] In order to obtain the α-configured phosphoribosylated **6** in a stereoselective manner, donor **8** was synthesized for its known α-selectivity.[16] For a β-selective glycosylation to furnish **7**, it was expected that participation of the 2-O-acetyl group will steer the nucleophilic attack of the acceptor to the β-face leading to the design of donor **9**.

Building block synthesis

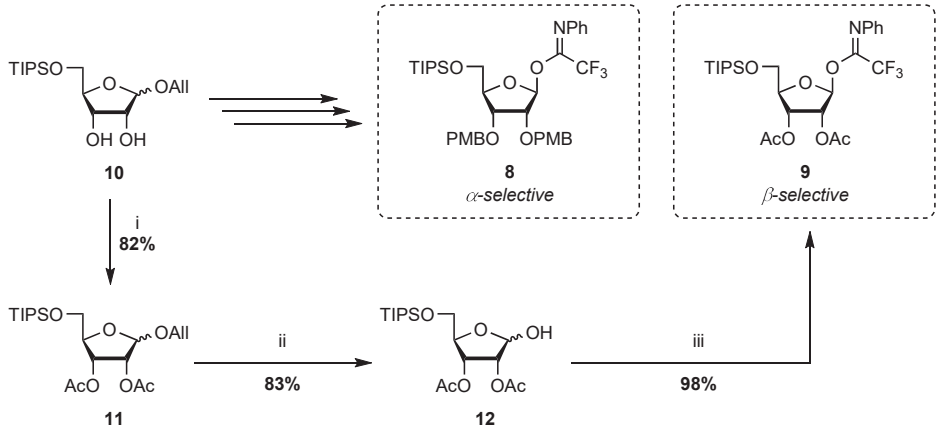
Known ribosyl donor **8** was prepared according to a reported procedure.[16] Synthesis of β-steering donor **9** could be accomplished via a three-step sequence from allyl 5-*O*-triisopropylsilyl-α,β-D-ribofuranoside **10** (Scheme 2), an intermediate towards donor **8**. Acetylation of the 2- and 3-OH groups followed by Pd mediated cleavage of the anomeric allyl group furnished alcohol **12** which was then converted to *N*-phenyltrifluoroacetimidate donor **9**.



Scheme 1. Retrosynthetic analysis for the synthesis of mono-ADP-ribosylated peptides **1** and **2.**

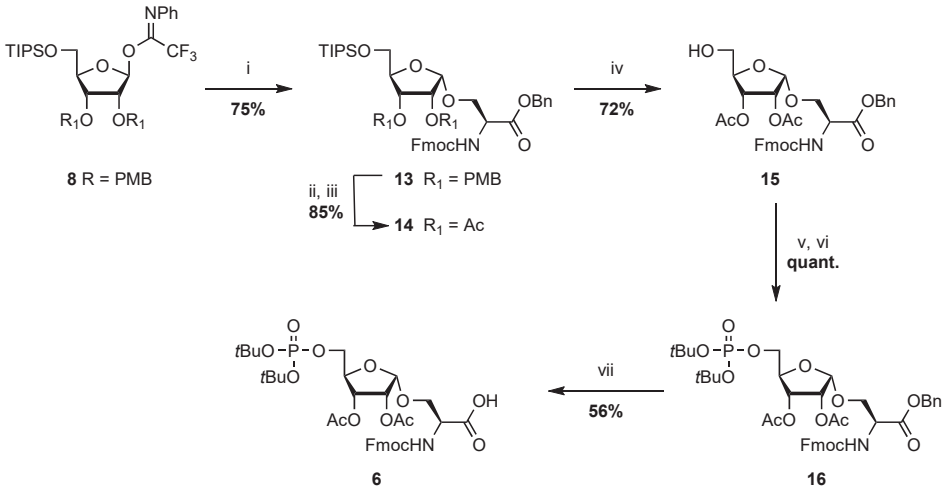
Having prepared the required ribofuranosyl donors, an α-selective glycosylation of Fmoc-Ser-OBn was performed with donor **8** (Scheme 3). When 0.2 eq of TMSOTf was employed at −50 °C, α-configured ribosylated Ser **13** was isolated in 75% yield, while the corresponding β-product could not be detected. Subsequent cleavage of the PMB protecting groups in **13** using a previously reported HCl/HFIP cocktail[17] (0.1 M) gave unsatisfactory yields due to acid-mediated hydrolysis of the Ser moiety. Although lowering the concentration of HCl (down to 0.01 M) in the mixture reduces the cleavage of the *O*-glycosidic bond, scaling up of the reaction resulted again in moderate yields. Fortunately, treatment of **13** with

TFA in DCM showed no sign of acidolysis of the *O*-glycosidic bond and upscaling of the reaction combined with the subsequent acetylation of the crude intermediate gave **14** in reproducible yields that exceeded 80% over two steps. Removal of the silyl ether in **14** using TEA·3HF proceeded uneventfully, while treatment with pyr·HF was accompanied by Fmoc cleavage. Phosphitylation of the primary hydroxyl in **15** using di-*tert*-butyl *N*,*N*-di-*iso*-propylphosphoramidite, and subsequent oxidation of the phosphite intermediate yielded phosphotriester **16**. Finally, hydrogenolysis of the benzyl ester gave orthogonally protected, α-phosphoribosylated Ser building block **6**.



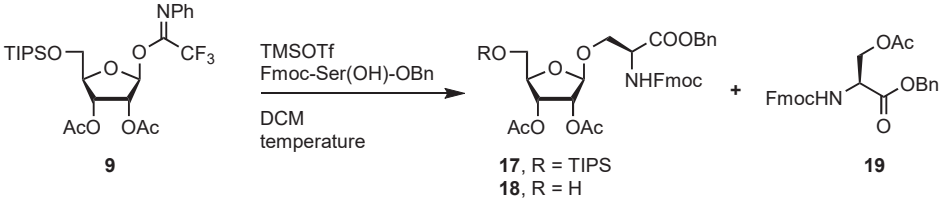
Scheme 2. Synthesis of the α- and β-selective ribosyl donors **8** and **9**. Reagents and conditions: i) Ac_2O, pyr., rt. ii) $PdCl_2$, O_2, $CHCl_3$, H_2O, 45 °C. iii) $Cl(C{=}NPh)CF_3$, Cs_2CO_3, acetone, H_2O, rt.

For the β-selective glycosylation of Fmoc-Ser-OBn with *N*-phenyltrifluoroacetimidate donor **9**, different reaction conditions proved to be necessary. The conditions that were used for α-selective glycosylation (–50 °C, 0.2 eq. of activator, entry 1 Table 1) gave a low yield and a significant amount of acetyl migration to the Ser aglycon giving **19** as a significant side product.[18] Therefore, an optimization of the β-selective glycosylation reaction with respect to the donor/activator ratio and the reaction temperature was performed (Table 1). Increasing the temperature to 0 °C (entry 2) not only improved the yield of **17** from 5% to 21% but also led to the formation of desilylated product **18** in 16% yield, giving a combined yield of 37%. Noteworthy, this acid-catalyzed desilylation was not observed in the α-directed glycosylation with donor **8**. Upon further increase of the temperature to room temperature (entry 3) **17** was not isolated but only the desilylated product **18** was obtained in a low yield. A reduced amount of activator (0.1 eq.) applied either at 0 °C or at room temperature (entries 4 and 5) increased the yields of both products significantly. As a result, the conditions in entry 4 (0 °C with 0.1 eq. TMSOTf) providing a cumulative yield of 55% were selected as the optimum conditions for β-selective glycosylation of Fmoc-Ser-OBn with donor **9**.



Scheme 3. Synthesis of orthogonally protected α-phosphoribosylated Ser building block **6**. Reagents and conditions: i) Fmoc-Ser(OH)-OBn, TMSOTf, DCM, -50 °C. ii) TFA, DCM, rt. iii) Ac_2O, pyr., rt. iv) TEA·3HF, THF, 0 °C to rt. v) $(tBuO)_2PN(iPr)_2$, 1-methyl imidazole, 1-methylimidazolium chloride, DMF, rt. vi) *t*BuOOH, nonane, rt. vii) H_2, 10 wt% Pd/C, *t*BuOH, 1,4-dioxane, H_2O.

Table 1. optimization of the glycosylation of β-selective donor **6** with Fmoc-Ser(OH)-OBn.

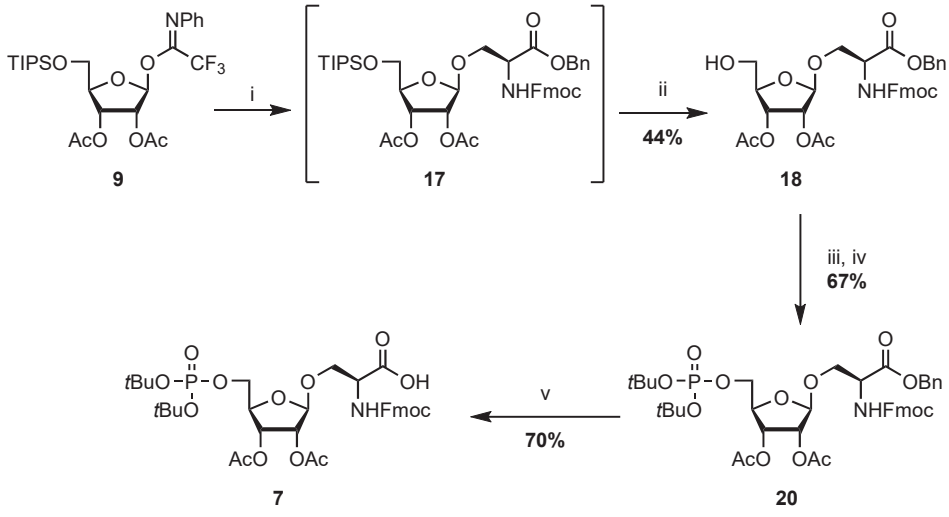


Entry	T (°C)	Activator (eq.)	Yield 17 (%)	Yield 18 (%)	Cumulative yield (%)
1	-50	0.2	5	n.d.	5
2	0	0.2	21	16	37
3	rt.	0.2	0	16	16
4	0	0.1	32	23	55
5	rt.	0.1	13	21	34

Thus, the route to β-phosphoribosylated Ser building block **7**, depicted in Scheme 4, started by condensation of donor **9** with Fmoc-Ser-OBn using the optimized conditions described above. Due to the partial desilylation observed, the crude mixture was immediately treated with TEA·3HF to produce **18** in 44% yield over two steps after silica gel chromatography. Analogous to the synthesis of α-phosphoribosylated Ser, standard phosphitylation of building block **18** ensuing oxidation gave **20**. Finally, cleavage of the benzyl ester furnished β-phosphoribosylated building block **7**.

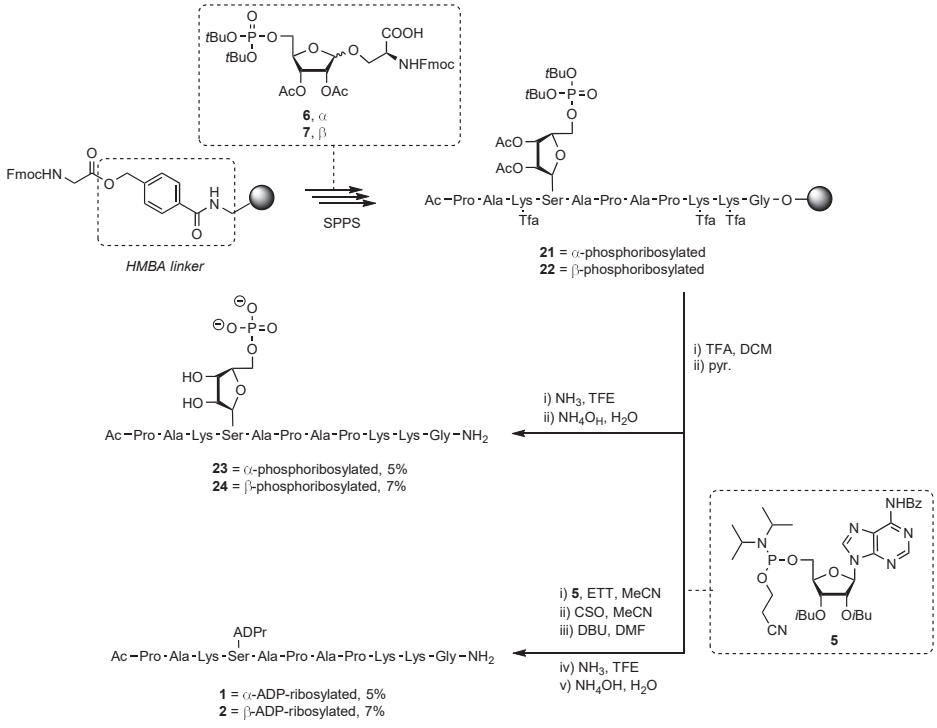
Peptide synthesis

Solid phase peptide synthesis (SPPS) of hendecapeptides **1** and **2** was performed using TentaGel® resin equipped with HMBA-linker as solid support (Scheme 5). Immobilized peptides **21** and **22** were assembled with standard Fmoc SPPS, utilizing trifluoroacetyl protected Lys (Fmoc-Lys(Tfa)-OH), Fmoc-Pro-OH, Fmoc-Ala-OH, Fmoc-Gly-OH and building blocks **6** or **7**. The *t*Bu groups of the phosphotriester were then removed by treatment of the immobilized peptides **21** and **22** with TFA in DCM. The conversion of the phosphotriester to the phosphate monoester could be monitored on-resin by ^{31}P-NMR (a shift of the phosphorus resonance from approximately -9 ppm for the phosphotriester to -4 ppm for the phosphodiester to 0 ppm for the unprotected phosphomonoester). After complete removal of the *t*Bu groups, the resin was washed with pyridine to obtain the more stable pyridinium salt of the phosphomonoester. To obtain phosphoribosylated peptides **23** and **24**, the protected phosphoribosylated intermediates were cleaved from the resin by treatment with a saturated NH_3 solution in 2,2,2-trifluoroethanol (TFE). The use of TFE ensures the formation of only carboxamide at the C-terminus of the target peptide chain and excludes formation of a carboxylic acid.[19] Finally, to effect complete removal of all protecting groups (i.e. the (trifluoroacetyl groups), NH_4OH was added, furnishing α- and β-phosphoribosylated peptides **23** and **24**.



Scheme 4. Synthesis of the β-phosphoribosylated Ser building block **20**. Reagents and conditions: i) Fmoc-Ser(OH)-OBn, TMSOTf, DCM, 0 °C. ii) TEA·3HF, THF, 0 °C to rt. iii) tBuO)$_2$PN(iPr)$_2$, 1-methyl imidazole, 1-methylimidazolium chloride, DMF, rt. iv) *t*BuOOH, nonane, rt. v) H_2, 10 wt% Pd/C, *t*BuOH, 1,4-dioxane, H_2O.

Next the synthesis of ADP-ribosylated peptides **1** and **2** was undertaken. Removal of the *t*Bu groups in the immobilized fully protected peptides **21** and **22** to the corresponding phosphomonoesters was followed by the installation of the pyrophosphate via a three step procedure developed by Kistemaker *et al.*[14] On-resin treatment of the obtained phosphate monoesters with activator ETT and phosphoramidite **5**, subsequent immediate oxidation of the P^{III} – P^{V} intermediate species with CSO and, finally, removal of the cyanoethyl group with DBU gave the immobilized and unprotected pyrophosphate. Similar conditions as described for phosphoribosyl peptides **23** and **24** were then used for the cleavage from the resin and deprotection of the ADP-ribosylated peptides to furnish, after purification, homogeneous α- and β-ADP-ribosylated peptides **1** and **2**, respectively.



Scheme 5. Solid phase peptide synthesis of the α- and β-ADP-ribosylated peptides **1** and **2** and their phosphoribosylated derivatives **23** and **24**. Tfa = trifluoroacetyl.

*Chemical and enzymatic degradation of peptides **1** and **2***

Having obtained Ser-ADPr peptides **1** and **2**, the native stereochemistry at the anomeric centre of the ribose residue attached to the side chain of Ser was investigated. Human ARH3 is the only enzyme known to hydrolyze the *O*-glycosidic Ser-ADPr linkage.[20] A homologue of ARH3, termed ARH1, is a stereoselective ADP-ribosylarginine hydrolase.[21,22]

It was hypothesized that ARH3 is likewise only capable of hydrolyzing glycosidic linkages with the native stereochemistry, leaving the non-natural epimer intact. Accordingly, a de-MARylation assay by treating the homogeneous α- and β-configured H2B peptides **1** and **2** with hARH3 was performed. In addition, the reaction mixture contained human NUDT5 (hNUDT5), which converts the released ADPr into AMP, and thus allows detection of the peptide turnover by a commercial assay. As expected, no turnover was observed either by hNUDT5 alone[23] or in the presence of the catalytically inactive hARH3 double mutant (D77N D78N) (Figure 1). On the other hand, wild-type hARH3 exclusively hydrolyzes the α anomeric Ser-ADPr linkage. Since no spontaneous epimerization in solution was observed previously, this data strongly suggest that the α-linked form is the endogenous epimer.

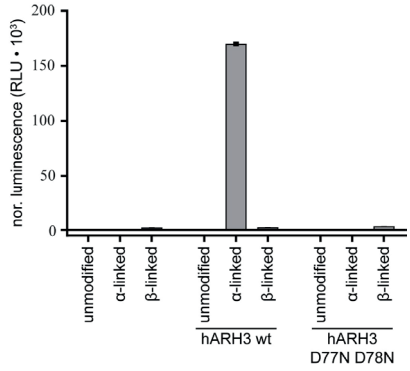
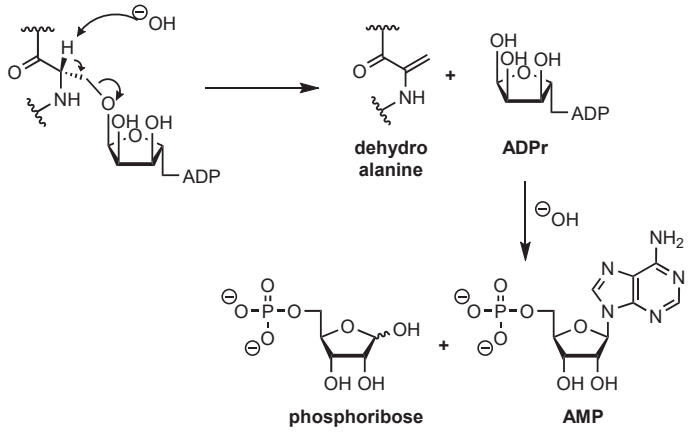


Figure 1. Analysis of the stereospecificity of hARH3. ADPr released in the ARH3 reaction was converted into AMP using hNUDT5 and subsequently measured using the AMP-Glo assay (Promega). Control reactions were carried out both in absence of hARH3 as well as using a catalytically inactive hARH3 mutant (D77N D788N). The data were normalized to reactions containing only hNUDT5 and represent triplicate measurement ± SD.

After confirming the stereochemistry of the naturally occurring Ser-ADPr modification, the chemical stability of the native α-Ser-ADPr was tested. Such knowledge about the stability of ADP-ribosylation sites is relevant for the future synthetic and proteomics studies and for a retrospective evaluation of already documented observations. A set of chemically divergent conditions were chosen (aqueous acid, diluted NaOH and 0.5 M NH_2OH) as these are broadly encountered in proteomic studies for identifying PAR- and MARylation sites (NH_2OH, NaOH)[22,24] or in the course of sample preparation (aqueous acid).[7] Peptide **1** was dissolved in aqueous solutions of either 0.1 M TFA, 0.1 M NaOH, or 0.5 M NH_2OH, and after various time points (15 min, 60 min, 2.5 h, 5 h, and 24 h), a sample was taken for LC-MS analysis of the quality of the peptide. For TFA and NH_2OH, no detectable degradation was observed after 24 hours. This leads to the conclusion that the isolation and purification of Ser-ADPr-peptides can safely be performed under acidic conditions and that the treatment

of ADPr-proteins with NH_2OH is indeed selective toward Asp and Glu residues since Ser-ADPr is unreactive toward this nucleophile.[24] However, treatment of our MARylated peptide **1** with 0.1 M aqueous NaOH (pH 13) showed a measurable elimination of ADP-ribose after 15 minutes (7%, Table 2). After 2.5 hours, the eliminated ADPr was partially hydrolyzed, giving phosphoribose and adenosine monophosphate (AMP). In a time span of 5 hours, more than half of the starting material was converted into a mixture of ADPr or AMP and complete degradation ensued after 24 hours. This elimination reaction leads to transformation of Ser-ADPr into dehydroalanine at the modification site, a reaction not possible with amino acids such as Asp, Glu, or Arg. This result is in line with the known behavior of Ser *O*-glycopeptides, in which the glycosyl linkage is also cleaved by β-elimination upon treatment with base.[25] The formation of dehydroalanine residue has been exploited in studies on phosphoproteomics[26,27] and glycomics,[28] and the present observation that Ser-ADPr sites behave similarly points to a possibility of extending such "foot printing" to the ADP-ribosylome as well.

Table 2. Elimination of ADPr from Ser and its subsequent hydrolysis of to phosphoribose and AMP.



t (h)	Intact ADPr peptide (%)	Free ADPr (%)	Free AMP (%)
0	100	0	0
0.25	93	7	0
1	85	15	0
2.5	66	26	8
5	45	32	23
24	0	50	50

Conclusion

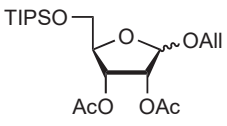
In conclusion, suitably protected phosphoribosylated Ser building blocks have been synthesized and used for the SPPS-mediated assembly of H2B-derived peptides, MARylated on their Ser residues both in α- and β-glycosidic form, via an SPPS based strategy. The work presented in this chapter showed the synthesis of the first Ser-ADP-ribosylated peptides via a modified literature procedure.[14] The thus prepared H2B-derived ADPr-peptides were then used to investigate fundamental, but as yet unknown, properties of the Ser-ADPr modification. The striking specificity of ARH3 for Ser-ADPr substrates was utilized to gain insight into its substrate preference and by extension the stereoselectivity by which the PARP1–HPF1 complex catalyzes MARylation of Ser residues. ARH3 catalysis resulted only in the hydrolysis of the α-linked peptide, thus strongly implying that this is the natural epimer. Also, since the Ser *O*-ribosidic bond is chemically distinct from the formerly studied Asp, Glu, and Arg modifications, the stability of this linkage was tested under conditions typically encountered in biochemical and proteomic studies. Ser-ADPr proved to be stable for a minimum of 24 hours when subjected to TFA (0.1 M) and NH_2OH (0.5 M) with no detectable degradation. In contrast, when Ser-ADPr was treated with NaOH (0.1 M) β-elimination occurred, in which the glycosidic bond was cleaved to give dehydroalanine residue and free ADPr.

Experimental section

General synthetic procedures

All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. 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 F_{254}). TLC was used to visualize compounds by UV at wavelength 254 nm and by spraying with either cerium molybdate spray (25 g/L $(NH_4)_6Mo_7O_{24}$, 10g/L $(NH_4)_4Ce(SO_4)_4{\cdot}H_2O$ in 10% H_2SO_4 water solution) or $KMnO_4$ spray (20 g/L $KMnO_4$ and 10g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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. High resolution mass spectra were recorded by direct injection (2 µL of a 2 µM solution in water:acetonitrile 50:50 (v/v) and 0.1% formic acid) on a mass spectrometer (Thermo Finnigan LTQ Orbitrap) equipped with an electrospray ion source in positive mode (mass range m/z = 150-2,000) and dioctylphthalate (m/z = 391.2843) as a "lock mass". The high-resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). HPLC purification was performed on an Agilent 1200 series using a 6130 quadrupole MS equipped with a Develosil C30 column, size 10.0/250. NMR spectra were recorded on a Bruker AV-400 or AV-600 NMR. Chemical shifts (δ) are given in ppm relative to tetramethyl silane as internal standard. Coupling constants (*J*) are given in Hz. For the 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. Optical rotation measurements were performed on an Anton Paar MCP100 modular circular polarimeter. FT-IR spectra were recorded on a Shimadzu IRAffinity-1 FT-IR spectrometer.

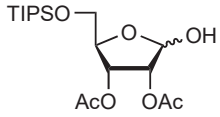
2,3-*O*-Acetyl-1-*O*-allyl-5-*O*-triisopropylsilyl-α,β-D-ribofuranoside (11)



Compound **10**[114] (1.04 g, 3.0 mmol) was dissolved in pyridine (6.0 mL, 0.5 M) . Ac_2O (1.13 mL, 12.0 mmol, 4 eq.) was added and the reaction was stirred overnight. The reaction was diluted with EtOAc and transferred into a separatory funnel. The reaction was subsequently washed with 1 M HCl, sat. aq. $NaHCO_3$ and brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (10% EtOAc in pentane) yielded the title compound as a clear oil (α and β combined 1.06 g, 2.47 mmol, 82%). *α-anomer* **^{1}H NMR** (400 MHz, $CDCl_3$) δ 5.90 (ddt, *J* = 16.4, 10.8, 5.4 Hz, 1H, OCH_2CHCH_2), 5.39 – 5.30 (m, 2H, H-3 + OCH_2CHCH_2), 5.28 (d, *J* = 4.1 Hz, 1H, H-1), 5.18 (dd, *J* = 10.4, 1.8 Hz, 1H, OCH_2CHCH_2), 5.03 (dd, *J* = 7.0, 4.8 Hz, 1H, H-2), 4.25 (dd, *J* = 13.5, 3.0 Hz, 1H, OCH_2CHCH_2), 4.17 (d, *J* = 3.2 Hz, 1H, H-4), 4.10 (dd, *J* = 13.7, 6.2 Hz, 1H, OCH_2CHCH_2), 3.95 (dd, *J* = 10.9, 2.8 Hz, 1H, H-5_a), 3.84 (dd, *J* = 11.0, 2.9 Hz, 1H, H-5_b), 2.18 – 2.02 (m, 6H, 2x Ac), 1.18 – 0.98 (m, 21H, TIPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.53, 169.85 (C=O Ac), 134.21 (OCH_2CHCH_2), 116.90 (OCH_2CHCH_2), 99.46 (C-1), 82.88 (C-4), 71.36 (C-2), 70.53 (C-3), 68.41 (OCH_2CHCH_2), 63.37 (C-5), 20.88, 20.57 (CH_3 Ac), 17.88 (CH TIPS), 11.87 (CH_3 TIPS). *β-anomer* **$[\alpha]_D^{20}$**: -22.2° (c = 1.0, $CHCl_3$). **IR** cm^{-1}: 2934, 2866, 1750, 1464, 1368, 1239, 1215, 1128, 1067, 881, 756. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 5.87 (ddt, *J* = 15.8, 10.8, 5.5 Hz, 1H, OCH_2CHCH_2), 5.41 (dt, *J* = 6.7, 3.6 Hz, 1H, H-3), 5.34 – 5.22 (m, 2H, H-2 + OCH_2CHCH_2), 5.18 (dt, *J* = 10.7, 2.0 Hz, 1H, OCH_2CHCH_2), 5.05 (s, 1H, H-1), 4.27 – 4.14 (m,

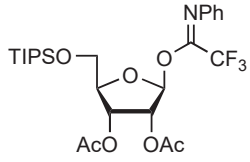
2H, H-4 + OC*H*$_2$CHCH$_2$), 4.04 – 3.91 (m, 1H, OC*H*$_2$CHCH$_2$), 3.91 – 3.73 (m, 2H, H-5), 2.10 (s, 3H, Ac), 2.04 (s, 3H, Ac), 1.14 – 1.00 (m, 21H, TIPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.63 (C=O Ac), 133.68 (OCH$_2$*C*HCH$_2$), 117.29 (OCH$_2$*C*HCH$_2$), 104.09 (C-1), 81.32 (C-4), 75.00 (C-2), 72.30 (C-3), 68.35 (OCH$_2$*C*HCH$_2$), 64.72 (C-5), 20.61 (CH$_3$ Ac), 20.54 (CH$_3$ Ac), 17.90 (CH TIPS), 11.88 (CH$_3$ TIPS). **HRMS:** [$C_{21}H_{38}O_7Si$ + Na]$^+$ found: 453.2277, calculated: 453.2279

2,3-*O*-Acetyl-1-5-*O*-triisopropylsilyl-α,β-D-ribofuranoside (12)



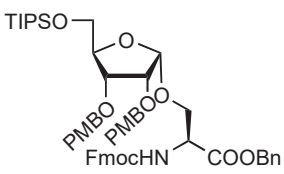
Compound **11** (995 mg, 2.31 mmol) was dissolved in a 3:1 (v/v) mixture of $CHCl_3$:H_2O (total volume 12 mL, 0.2 M) and $PdCl_2$ (62 mg, 0.35 mmol, 0.15 eq.) was added. Oxygen gas was bubbled through the solution for 15 minutes while the temperature was raised to 45°C. The reaction was stirred vigorously for 2 days under oxygen atmosphere. The reaction mixture was carefully concentrated *in vacuo* and the residue was dissolved in 12 mL EtOAc and 12 mL sat. aq. $NaHCO_3$. Iodine (1 eq.) was added and the reaction was transferred into a separating funnel and shaken. The organic layer was separated and washed with sat. aq. $Na_2S_2O_3$. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography yielded the title compound as a clear oil (750 mg, 1.92 mmol, 83%, α:β ratio: 0.6:1). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 5.56 (td, *J* = 5.3, 2.1 Hz, 1H, β H-3), 5.50 (d, *J* = 4.6 Hz, 1H, α H-1), 5.44 (dt, *J* = 6.0, 1.9 Hz, 1H, α H-3), 5.27 (s, 1H, β H-1), 5.22 (d, *J* = 5.1 Hz, 1H, β H-2), 5.20 – 5.14 (m, 1H, α H-2), 4.27 (t, *J* = 2.3 Hz, 1H, α H-4), 4.21 (dt, *J* = 5.2, 2.6 Hz, 1H, β H-4), 3.99 – 3.77 (m, 4H, α + β H-5), 2.15 – 2.04 (m, 12H, α + β Ac), 1.13 – 1.01 (m, 42H, α + β TIPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.02, 169.98, 169.88, 169.68 (α + β C=O Ac), 100.27 (β C-1), 95.61 (α C-1), 82.95 (α C-4), 82.61 (β C-4), 76.77 (β C-2), 71.86 (α C-3), 71.55 (α C-2), 71.22 (β C-3), 63.57, 63.37 (α + β C-5), 20.89, 20.65, 20.61, 20.59 (α + β CH$_3$ Ac), 17.95, 17.91, 17.88 (α + β CH TIPS), 11.89 (α + β CH$_3$ TIPS). **IR cm^{-1}:** 3456, 29492, 2866, 1746, 1464, 1369, 1220, 1090, 1066, 1013, 881, 754, 680. **HRMS:** [$C_{18}H_{34}O_7Si$ + Na]$^+$ found: 413.1968, calculated: 413.1966

2,3-*O*-Acetyl-1-((*N*-phenyl)-2,2,2-trifluoroacetimido)-5-*O*-triisopropylsilyl-α,β-D-ribofuranoside (9)



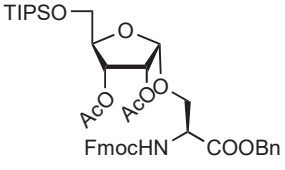
Compound **12** (623 mg, 1.60 mmol) was dissolved in acetone (8 mL, 0.2 M) and H_2O (0.16 mL) was added. Cs_2CO_3 (782, mg, 2.4 mmol, 1.5 eq.) and 2,2,2-trifluoro-*N*-phenylacetimidoyl chloride (0.36 mL, 2.4 mmol, 1.5 eq.) were added and the reaction was stirred overnight. The reaction was filtered and concentrated *in vacuo*. Flash column chromatography (5 -> 10% EtOAc in pentane) yielded the title compound as a pale oil (α and β combined 880 mg, 1.57 mmol, 98%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.32 – 7.27 (m, 3H Ph arom.), 7.10 (t, *J* = 7.5 Hz, 1H, Ph arom.), 6.82 (d, *J* = 7.1 Hz, 2H, Ph arom.), 6.32 (bs, 1H, H-1), 5.57 (d, *J* = 3.4 Hz, 1H, H-2), 5.50 (t, *J* = 5.5 Hz, 1H, H-3), 4.33 (dt, *J* = 6.2, 4.6 Hz, 1H, H-4), 3.90 (d, *J* = 4.6 Hz, 2H, H-5), 2.12 (s, 3H, Ac), 2.07 (s, 3H, Ac), 1.10 – 1.06 (m, 21H, TIPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.77, 169.53 (C=O Ac), 143.59 (Cq Ph), 128.82, 124.41, 119.54 (CH arom. Ph), 101.54 (C-1), 83.36 (C-2), 74.50 (C-3), 71.18 (C-4), 63.79 (C-5), 20.63 (CH$_3$ Ac), 17.97 (CH TIPS), 11.93 (CH$_3$ TIPS). **IR cm^{-1}:** 2943, 2867, 1753, 1713, 1599, 1451, 1322, 1237, 1205, 1153, 1088, 881, 776, 690.

***O*-(2,3-*O*-di-(4-methoxybenzyl)-5-*O*-triisopropylsilyl-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine benzyl ester (13)**



Fmoc-Ser(OH)-OBn (877 mg, 2.10 mmol, 1.0 eq.) together with donor **8** (1.65 g, 2.30 mmol, 1.1 eq.) was co-evaporated twice with 1,4-dioxane and once with DCE before the mixture was dissolved in DCM (23 mL, 0.09 M). The solution was cooled to -50 °C before TMSOTf (42 µL, 0.23 mmol, 0.1 eq. relative to the donor 5) was added. The reaction was stirred for 90 minutes at -50 °C before TLC showed full conversion of the acceptor and a higher running product. The reaction was quenched by the addition of TEA and concentrated *in vacuo.* Flash column chromatography (0.5 -> 3% acetone in DCM) yielded the title compound as a colorless oil (1.64 g, 1.73 mmol, 75%). **Rf:** 0.6 in 3% acetone in DCM. **$[\alpha]_D^{20}$** = +21° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.72 (dd, *J* = 7.6, 3.9 Hz, 2H, Fmoc arom.), 7.56 (d, *J* = 7.4 Hz, 2H, Fmoc arom.), 7.41 – 7.16 (m, 9H, Fmoc arom. + Bn arom. + $CHCl_3$), 6.90 – 6.70 (m, 4H, PMB arom.), 6.60 (d, *J* = 9.0 Hz, 1H, NH), 5.23 – 5.12 (m, 2H, CH_2 Bn.), 4.91 (d, *J* = 4.1 Hz, 1H, H-1), 4.69 – 4.56 (m, 2H, CH Ser + CH_2 PMB), 4.53 – 4.45 (m, 3H, 2 CH_2 PMB), 4.39 (dd, *J* = 10.3, 6.8 Hz, 1H, CH_2 Fmoc), 4.18 (m, 4H, CH_2 Fmoc + CH Fmoc + CH_2 Ser + H-4), 4.02 – 3.92 (m, 2H, CH_2 Ser + H-3), 3.76 (m, 4H, CH_3 PMB + H-2), 3.69 (s, 3H, CH_3 PMB), 3.66 – 3.55 (m, 2H, H-5), 0.98 (m, 21H, TIPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.56 (C=O COOBn), 159.40, 159.21 (C-OMe PMB), 156.57 (C=O Fmoc), 144.27, 143.87, 141.32, 141.23 (Cq arom. Fmoc), 135.69 (Cq arom. Bn), 130.43, 129.97 (Cq arom. PMB), 129.84, 129.52, 128.58, 128.22, 128.09, 127.68, 127.62, 127.12, 127.06 (CH arom. Fmoc + Bn + PMB), 125.48, 125.22, 119.96, 119.93 (CH Fmoc arom.), 113.84 113.72 (CH PMB arom.), 101.13 (C-1), 84.44 (C-4), 78.44 (C-2), 75.16 (C-3), 72.33, 72.03 (CH_2 PMB), 67.52 (CH_2 Ser), 67.19 (CH_2 Fmoc), 67.15 (CH_2 Bn), 63.74 (C-5), 55.31, 55.24 (CH_3 PMB), 54.70 (CH Ser), 47.17 (CH Fmoc), 18.03, 17.99 (CH TIPS), 11.94 (CH_3 TIPS). **IR cm^{-1}:** 2941, 2864, 1722, 1610, 1512, 1246, 1029, 908, 729. **HRMS** $[C_{55}H_{67}NO_{11}Si + Na]^+$: 968.4398 found, 968.4375 calculated.

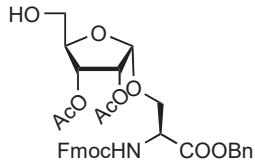
***O*-(2,3-*O*-diacetyl-5-*O*-triisopropylsilyl-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine benzyl ester (14).**



Compound **13** (1.64 g, 1.73 mmol) was dissolved in DCM (35 mL, 0.05 M) and TFA (1.0 mL, 13.5 mmol, 7.8 eq.) was added. Next, the reaction was stirred for 30 minutes. TLC showed full conversion of starting material and a product on the baseline (3:97 acetone:DCM). The reaction was diluted with toluene (30 mL) and the reaction was concentrated *in vacuo*. The diol was co-evaporated thrice with toluene to remove traces of TFA. Then residue was dissolved in pyridine (17.5 mL, 0.1 M) and Ac_2O (3.3 mL, 34.6 mmol, 20 eq.) was added. The reaction was stirred overnight after which TLC showed full conversion of the diol. The reaction was diluted with DCM and transferred into a separating funnel. The organic layer was washed with 1 M HCl, sat. aq. $NaHCO_3$, and brine. The organic layer was collected, dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (1 -> 3% acetone in DCM) yielded the title compound as a white foam (1163 mg, 1.47 mmol, 85%). **Rf:** 0.58 in 2% acetone in DCM. **$[\alpha]_D^{20}$** = +47.5° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.81 – 7.72 (m, 2H, Fmoc arom.), 7.58 (t, *J* = 6.9 Hz, 2H, Fmoc arom.), 7.44 – 7.23 (m, 9H, Fmoc arom. + Bn arom. + $CHCl_3$), 5.98 (d, *J* = 9.2 Hz, 1H, NH), 5.40 (dd, *J* = 6.4, 2.0 Hz, 1H, H-3), 5.21 (d, *J* = 4.4 Hz, 1H, H-1), 5.18 (m, 2H, CH_2Bn), 4.98 (dd, *J* = 6.4, 4.4 Hz, 1H, H-2), 4.57 (dt, *J* = 9.3, 2.8 Hz, 1H, CH Ser), 4.50 (dd, *J* = 10.7, 6.6 Hz, 1H, CH_{2a} Fmoc), 4.41 (dd, *J* = 10.6, 7.0 Hz, 1H, CH_{2b} Fmoc), 4.23 (t, *J* = 6.8 Hz, 1H, CH Fmoc), 4.08 (q, *J* = 2.5 Hz, 1H, H-4), 3.99 (qd, *J* = 10.1, 2.9 Hz, 2H, CH_2 Ser), 3.82 (ddd, *J* = 30.3, 11.0, 2.5 Hz, 2H, H-5), 2.00 (s, 3H, Ac), 1.96 (s, 3H, Ac), 1.08 – 1.03 (m, 21H, TIPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.40, 170.20, 169.96 (C=O COOBn/Ac), 156.20 (C=O Fmoc), 144.0, 143.79, 141.45, 141.40 (Cq arom.

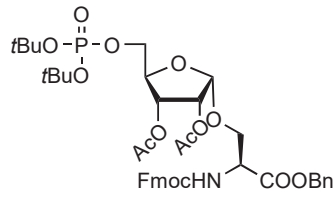
Fmoc), 135.43 (Cq arom. Bn), 128.75, 128.54, 128.12, 127.85, 127.82, 127.20, 127.13 (CH arom. Fmoc + Bn), 125.09, 124.97, 120.14, 120.10 (CH arom. Fmoc), 100.14 (C-1), 83.28 (C-4), 71.83 (C-2), 70.80 (C-3), 67.50 (CH_2 Ser), 67.32 (CH_2 Bn), 67.01 (CH_2 Fmoc), 63.47 (C-5), 54.34 (CH Ser), 47.25 (CH Fmoc), 20.68, 20.55 (CH_3 Ac), 18.05 (CH TIPS), 18.02 (CH 11.99 (CH_3 TIPS). **IR cm^{-1}:** 2943, 2866, 1732, 1506, 1221, 1057, 1030, 908, 729. **HRMS:** [$C_{43}H_{55}NO_{11}Si$ + Na]$^+$:812.3454 found, 812.3436 calculated.

O-(2,3-*O*-diacetyl-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine benzyl ester (15)



Compound **14** (1.16 g, 1.47 mmol, 1.0 eq.) was dissolved in THF (15 mL, 0.1 M) and cooled to 0 °C. TEA·3HF (4.8 mL, 29.4 mmol, 20 eq.) was added slowly and the reaction was stirred three days while slowly allowed to warm to room temperature. TLC showed full conversion of the starting material into a lower running product. The reaction was cooled to 0 °C and carefully quenched by the addition of sat. aq. $NaHCO_3$. The reaction was transferred into a separating funnel and extracted with EtOAc. The organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (3 -> 15% acetone in DCM) yielded the title compound as a white foam (673 mg, 1.06 mmol, 72%). **Rf:** 0.32 in 10% acetone in DCM. **[α]$_D^{20}$** = +69.9° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.79 (d, *J* = 7.5 Hz, 2H, Fmoc arom.), 7.65 – 7.59 (m, 2H, Fmoc arom.), 7.47 – 7.30 (m, 9H, Fmoc arom. + Bn arom.), 5.95 (d, *J* = 9.1 Hz, 1H, NH), 5.29 (dd, *J* = 6.9, 3.1 Hz, 1H, H-3), 5.24 (d, *J* = 4.3 Hz, 1H, H-1), 5.23 – 5.19 (m, 2H, CH_2 Bn), 4.93 (dd, *J* = 6.9, 4.3 Hz, 1H, H-2), 4.60 (dt, *J* = 9.3, 2.8 Hz, 1H, CH Ser), 4.54 (dd, *J* = 10.7, 6.6 Hz, 1H, CH_2 Fmoc), 4.46 (dd, *J* = 10.8, 6.9 Hz, 1H, CH_2 Fmoc), 4.26 (t, *J* = 6.7 Hz, 1H, CH Fmoc), 4.11 (q, *J* = 3.2 Hz, 1H, H-4), 4.08 – 3.95 (m, 2H, CH_2 Ser), 3.85 – 3.68 (m, 3H, H-5 + OH), 2.05 (s, 3H, Ac), 2.02 (s, 3H, Ac). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.47, 170.08, 170.01 (C=O COOBn/Ac), 156.14 (C=O Fmoc), 143.98, 143.75, 141.47, 141.42 (Cq arom. Fmoc), 135.34 (Cq arom. Bn), 128.77, 128.60, 128.20, 127.88, 127.86, 127.16 (CH arom. Fmoc + Bn), 125.07, 124.98, 120.15, 120.12 (CH arom. Fmoc), 100.39 (C-1), 82.45 (C-4), 71.49 (C-2), 70.20 (C-3), 67.99 (CH_2 Ser), 67.43 (CH_2 Bn), 66.99 (CH_2 Fmoc), 62.27 (C-5), 54.33 (CH Ser), 47.26 (CH Fmoc), 20.65, 20.52 (CH_3 Ac). **IR cm^{-1}:** 3431, 2939, 1738, 1511, 1451, 1369, 1219, 1055, 1027, 974, 739, 698. **HRMS:** [$C_{34}H_{35}NO_{11}$ + Na]$^+$: 656.2118 found, 656.2102 calculated.

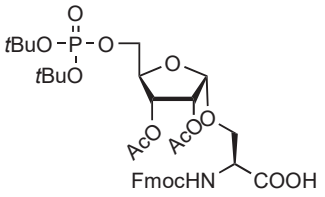
O-(2,3-*O*-diacetyl-5-*O*-(di-*tert*-butyl)-phosphoryl-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine benzyl ester (16)



Compound **15** (668 mg, 1.05 mmol, 1.0 eq.) was co-evaporated thrice with 1,4-dioxane and dissolved in a 0.2 M 1-methyl-imidazole and 0.3 M 1-methyl-imidazole·HCl solution in DMF (14.2 mL, 0.1 M). Di-*tert*-butyl-*N*,*N*-diisopropylphosphoramidite (167 µL, 2.12 mmol, 1.5 eq.) was added and the reaction was stirred for 20 minutes. TLC indicated full conversion (R_f = 1.0, 1:9 acetone:DCM) after which *t*BuOOH (5.5 M in nonane, 2.58 mL, 14.2 mmol, 10 eq.) was added. After 120 minutes, ^{31}P-NMR indicated full conversion of the amidite into the protected phosphate. The reaction was quenched by the addition of sat. aq. $NaHCO_3$ and transferred into a separating funnel. The water layer was extracted with Et_2O and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (30 -> 50% EtOAc in pentane) yielded the title compound as a white foam in quantitative yield. **[α]$_D^{20}$** = +53.6° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.6 Hz, 2H, Fmoc arom.), 7.59 (t, *J* = 6.6 Hz, 2H, Fmoc arom.), 7.45 – 7.29 (m, 9H, Fmoc arom. + Bn arom.), 5.87 (d, *J* = 9.2 Hz, 1H,NH), 5.34 (dd, *J* = 6.7, 2.5 Hz, 1H, H-3), 5.23 (d, *J* = 4.3 Hz, 1H, H-1), 5.17 (s, 2H, CH_2 Bn), 4.92 (dd, *J* = 6.7, 4.3 Hz, 1H, H-2), 4.57 (dt, *J* = 9.3, 2.9 Hz, 1H, CH Ser), 4.51 (dd, *J* = 10.7, 6.5 Hz, 1H, CH_2 Fmoc), 4.42 (dd, *J* = 10.7, 6.9 Hz, 1H, CH_2 Fmoc), 4.23 (t, *J* = 6.7 Hz, 1H,

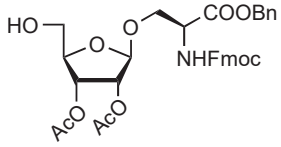
CH Fmoc), 4.15 (p, *J* = 4.0, 3.5 Hz, 1H, H-4), 4.14 – 4.03 (m, 2H, H-5), 3.98 (d, *J* = 2.9 Hz, 2H, CH_2 Ser), 2.00 (s, 3H, Ac), 1.97 (s, 3H, Ac), 1.52 – 1.47 (s, 18H, *t*Bu). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.09, 170.01, 169.72 (C=O COOBn/Ac), 156.04 (C=O Fmoc), 143.88, 143.65, 141.35, 141.30 (Cq arom. Fmoc), 135.26 (Cq arom. Bn), 128.67, 128.49, 128.08, 127.78, 127.76, 127.08 (CH arom. Fmoc + Bn), 124.97, 124.86, 120.06, 120.01 (CH arom. Fmoc), 100.02 (C-1), 82.93, 82.90, 82.86, 82.83 (Cq *t*Bu) 80.79, 80.71 (C-4), 71.31 (C-2), 70.21 (C-3), 67.49 (CH_2 Ser), 67.29 (CH_2 Bn), 66.89 (CH_2 Fmoc), 65.79, 65.73 (C-5), 54.18 (CH Ser), 47.14 (CH Fmoc), 29.87, 29.86, 29.83, 29.81 (CH_3 *t*Bu), 20.48, 20.38 (CH_3 Ac). **^{31}P NMR** (162 MHz, $CDCl_3$) δ -9.35. **IR cm^{-1}:** 2942, 1744, 1512, 1450, 1371, 1242, 1082, 1037, 996, 741. **HRMS:** $[C_{42}H_{52}NO_{14}P + Na]^+$: 848.3043 found, 848.3017 calculated.

O-(2,3-*O*-diacetyl-5-*O*-(di-*tert*-butyl)-phosphoryl-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine (6)



Compound **16** (1.15 g, 1.39 mmol) was dissolved in a 4:4:1 v/v mixture of *t*BuOH:1,4-dioxane:H_2O (total volume 27.8 mL, 0.05 M). The reaction was purged with nitrogen and Pd/C (10 wt.% 172 mg) was added. Hydrogen gas was then bubbled through the reaction mixture for 1 hour. TLC indicated full conversion of the starting material. The reaction was purged with nitrogen gas and filtered through a pad of Celite. The reaction was concentrated *in vacuo* and flash column chromatography (2 -> 10% MeOH in DCM, 1 v/v% AcOH) yielded the title compound as a white foam (568 mg 0.77 mmol, 56%). **Rf:** 0.47 in 10% MeOH in DCM + AcOH. **$[\alpha]_D^{20}$** = +66.0° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 10.55 (s, 1H COOH), 7.75 (d, *J* = 7.6 Hz, 2H, Fmoc. arom.), 7.59 (t, *J* = 7.6 Hz, 2H, Fmoc. arom.), 7.39 (t, *J* = 7.4 Hz, 2H, Fmoc. arom.), 7.33 – 7.25 (m, 2H, Fmoc. arom.), 5.86 (d, *J* = 8.9 Hz, 1H, NH), 5.31 (dd, *J* = 6.7, 2.6 Hz, 1H, H-3), 5.26 (d, *J* = 4.3 Hz, 1H, H-1), 4.93 (dd, *J* = 6.7, 4.3 Hz, 1H, H-2), 4.54 (dt, *J* = 9.2, 3.2 Hz, 1H, CH Ser), 4.48 (dd, *J* = 10.6, 6.7 Hz, 1H, CH_2 Fmoc), 4.40 (dd, *J* = 10.6, 7.0 Hz, 1H, CH_2 Fmoc), 4.27 – 4.15 (m, 2H, CH Fmoc + H-4), 4.15 – 4.03 (m, 2H, H-5), 4.03 – 3.90 (m, 2H, CH_2 Ser), 2.02 (s, 3H, Ac), 2.01 (s, 3H, Ac), 1.48 (s, 18H, *t*Bu). **^{31}P NMR** (162 MHz, $CDCl_3$) δ -10.17. **^{13}C NMR** (101 MHz, $CDCl_3$) δ 176.37 (C=O COOH), 170.33, 170.08 (C=O Ac), 156.16 (C=O Fmoc), 143.92, 143.73, 141.35 (Cq arom. Fmoc), 127.82, 127.13, 125.08, 124.98, 120.09, 120.06 (CH arom. Fmoc), 100.33 (C-1), 83.67, 83.64, 83.60, 83.57 (Cq *t*Bu), 80.82, 80.73 (C-4), 71.28 (C-2), 70.22 (C-3), 67.73 (CH_2 Ser), 67.02 (CH_2 Fmoc), 66.08, 66.02 (C-5), 54.01 (CH Ser), 47.18 (CH Fmoc), 29.89, 29.84 (CH_3 *t*Bu), 20.58 (CH_3 Ac), 20.36 (CH_3 Ac). **IR cm^{-1}:** 3421, 2981, 1739, 1512, 1373, 1238, 1041, 1002, 744. **HRMS:** $[C_{35}H_{46}NO_{14}P + Na]^+$: 758.2567 found, 758.2548 calculated.

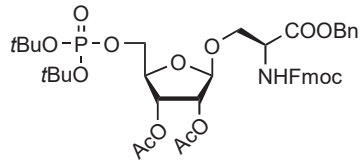
O-(2,3-*O*-diacetyl-β-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine benzyl ester (18)



Fmoc-Ser-OBn (584 mg, 1.40 mmol, 1.0 eq.) together with donor **9** (851 mg, 1.52 mmol, 1.1 eq.) was co-evaporated twice with 1,4-dioxane and once with DCE before the mixture was dissolved in DCM (15 mL, 0.1 M). The solution was cooled to 0 °C before TMSOTf (27 µL, 0.15 mmol, 0.1 eq. relative to donor 6) was added. The reaction was stirred for 1 hour at 0 °C before TLC (3% acetone in DCM) showed full conversion of the starting material into both the 5'-silylidene protected product (R_f = 0.58) and its deprotected equivalent. The reaction was quenched by adding three drops of TEA and concentrated *in vacuo*. Flash column chromatography (0 -> 10% acetone in DCM) obtained 87 mg of the title compound and 669 mg of a 1 : 2 mixture of Fmoc-Ser(Ac)-OBn and protected product (calculated by ^{1}H-NMR). This mixture was dissolved in THF (8.5 mL, 0.1 M) and the solution was cooled to 0 °C. TEA·3HF (2.77 mL, 17 mmol, 20 eq.) was added and the reaction was stirred for three days while the reaction temperature was slowly allowed to warm to room temperature. The reaction was cooled to 0 °C and carefully quenched by the addition of sat. aq.

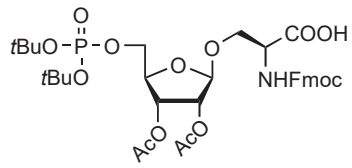
$NaHCO_3$. The mixture was transferred into a separating funnel and extracted with EtOAc. The organic layers were combined, were dried over $MgSO_4$, filtered and concentrated. Flash column chromatography (5 -> 10% acetone in DCM) yielded the title compound as a white foam (combined yield of 391 mg, 0.62 mmol, 44%). **Rf:** 0.16 in 3% acetone in DCM. **$[\alpha]_D^{20}$** = -1.40° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.5 Hz, 2H, Fmoc arom.), 7.62 (dd, *J* = 7.5, 3.6 Hz, 2H, Fmoc arom.), 7.44 – 7.28 (m, 9H Fmoc arom. + Bn arom.), 6.00 (d, *J* = 8.6 Hz, 1H, NH), 5.34 (dd, *J* = 6.9, 4.9 Hz, 1H, H-3), 5.27 (d, *J* = 5.0 Hz, 1H, H-2), 5.27 – 5.15 (m, 2H, CH_2 Bn), 4.98 (s, 1H, H-1), 4.58 (dt, *J* = 8.6, 3.3 Hz, 1H, CH Ser), 4.45 (dd, *J* = 10.5, 7.0 Hz, 1H, CH_2 Fmoc), 4.36 (dd, *J* = 10.5, 7.1 Hz, 1H, CH_2 Fmoc), 4.28 – 4.17 (m, 2H, CH Fmoc + CH_2 Ser), 4.14 (ddd, *J* = 6.7, 3.6, 2.7 Hz, 1H, H-4), 3.82 (dd, *J* = 10.4, 3.7 Hz, 1H, CH_2 Ser), 3.73 (dd, *J* = 12.6, 2.8 Hz, 1H, H-5_a), 3.51 (dd, *J* = 12.6, 3.7 Hz, 1H, H-5_b), 2.12 (s, 3H, Ac), 2.07 (s, 3H, Ac). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.47, 170.11, 169.74 (C=O COOBn/Ac), 156.30 (C=O Fmoc), 144.04, 143.87, 141.38 (Cq arom. Fmoc), 135.15 (Cq arom. Bn), 128.76, 128.64, 128.51, 127.83, 127.81, 127.26, 127.24 (CH arom. Fmoc + Bn), 125.33, 125.29, 120.07 (CH arom. Fmoc), 105.90 (C-1), 82.16 (C-4), 75.45 (C-2), 70.21 (C-3), 68.85 (CH_2 Ser), 67.88 (CH_2 Bn), 67.36 (CH_2 Fmoc), 61.68 (C-5), 54.56 (CH Ser), 47.22 (CH Fmoc), 20.77, 20.73 (CH_3 Ac). **IR cm^{-1}:** 3368, 2921, 1745, 1722, 1519, 1450, 1370, 1240,1215, 1045, 739. **HRMS:** $[C_{34}H_{35}NO_{11} + Na]^+$ found: 656.2108, calculated: 656.2102.

O-(2,3-*O*-diacetyl-5-*O*-(di-*tert*-butyl)-phosphoryl-β-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine benzyl ester (20)



Compound **18** (343 mg, 0.54 mmol, 1.0 eq.) was co-evaporated thrice with 1,4-dioxane before dissolved in a 0.2 M 1-methyl-imidazole and 0.3 M 1-methyl imidazole·HCl solution in DMF (5.4 mL, 0.1 M). Di-*tert*-butyl-*N,N*-diisopropylphosphoramidite (256 µL, 0.81 mmol, 1.5 eq.) was added and the reaction was stirred for 20 minutes. TLC indicated full conversion (R_f = 1.0, 1:9 acetone:DCM) and *t*BuOOH (5.5 M in nonane, 0.64 mL, 3.5 mmol, 10 eq.) was added. After 135 minutes, ^{31}P-NMR indicated full conversion of the amidite into the protected phosphate. The reaction was quenched by the addition of sat. aq. $NaHCO_3$ and transferred into a separating funnel. The water layer was extracted with Et_2O and the organic layers were combined, dried over $MgSO_4$, filtered and concentrated in vacuo. Flash column chromatography (30 -> 50% EtOAc in pentane) yielded the title compound as a white foam (298 mg, 0.36 mmol, 67%). **Rf:** 0.45 in 50% EtOAc in pentane. **$[\alpha]_D^{20}$** = -9.30° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.75 (d, *J* = 7.5 Hz, 2H, Fmoc arom.), 7.66 (t, *J* = 7.2 Hz, 2H, Fmoc arom.), 7.43 – 7.27 (m, 9H, Fmoc + Bn arom.), 6.51 (d, *J* = 8.7 Hz, 1H, NH), 5.28 (d, *J* = 5.0 Hz, 1H, H-2), 5.26 – 5.16 (m, 3H, CH_2 Bn + H-3), 5.05 (s, 1H, H-1), 4.65 (ddd, *J* = 8.7, 5.4, 3.3 Hz, 1H, CH Ser), 4.40 (dd, *J* = 10.4, 7.4 Hz, 1H, CH_{2a} Fmoc), 4.36 – 4.27 (m, 2H, CH_{2b} Fmoc + H-4), 4.27 – 4.19 (m, 2H, CH Fmoc + CH_{2a} Ser), 4.08 – 3.98 (m, 2H, H-5), 3.75 (dd, *J* = 10.0, 3.4 Hz, 1H, CH_{2b} Ser), 2.10 (s, 3H, Ac), 2.05 (s, 3H, Ac), 1.47 (m, *J* = 5.6 Hz, 18H, *t*Bu). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 169.71, 169.59, 169.55 (C=O COOBn/Ac), 156.40 (C=O Fmoc), 143.97, 143.88, 141.20 (Cq Fmoc), 135.29 (Cq Bn), 128.58, 128.35, 128.23, 127.62, 127.07 (CH arom. Fmoc + Bn), 125.35, 119.87 (CH Fmoc arom.), 105.51 (C-1), 82.87, 82.79, 82.76, 82.69 (Cq *t*Bu), 79.97, 79.89 (C-4), 74.63 (C-2), 71.58 (C-3), 67.47 (CH_2 Ser), 67.32, 67.30, 67.23, 67.17 (CH_2 Fmoc, CH_2 Bn, C-5), 54.47 (CH Ser), 47.06 (CH Fmoc), 29.83, 29.78 (CH_3 *t*Bu), 20.58, 20.51 (CH_3 Ac). **^{31}P NMR** (162 MHz, $CDCl_3$) δ -9.18. **IR cm^{-1}:** 2978, 1747, 1716, 1517, 1370, 1239, 1214, 1038, 989, 759, 739. **HRMS:** $[C_{42}H_{52}NO_{14}P + Na]^+$ found: 848.3040, calculated: 848.3017.

***O*-(2,3-*O*-diacetyl-5-*O*-(di-*tert*-butyl)-phosphoryl-β-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine (7)**



Compound **20** (298 mg, 0.36 mmol) was dissolved in a 4:4:1 v/v mixture of *t*BuOH:1,4-dioxane:H_2O (total volume 7.2 mL, 0.05 M). The reaction was purged with nitrogen and Pd/C (10 wt.%, 54 mg) was added. Hydrogen gas was then bubbled through the reaction mixture for 2 hours. TLC indicated full conversion of the starting material. The reaction was purged with nitrogen gas and filtered through a pad of Celite. The reaction was concentrated *in vacuo* and flash column chromatography (1 -> 5% MeOH in DCM, 1 v/v% AcOH) yielded the title compound as a white foam (185 mg 0.25 mmol, 70%). **Rf:** 0.45 in 5% MeOH in DCM + AcOH. **[α]** $_D^{20}$ = +10.40° (c = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.76 (dt, *J* = 7.6, 1.3 Hz, 2H, Fmoc arom.), 7.69 – 7.62 (m, 2H, Fmoc arom.), 7.44 – 7.37 (m, 2H, Fmoc arom.), 7.33 (tdd, *J* = 7.4, 2.7, 1.3 Hz, 2H, Fmoc arom.), 5.81 (d, *J* = 8.1 Hz, 1H, NH), 5.33 (d, *J* = 4.7 Hz, 1H, H-2), 5.18 (dd, *J* = 7.9, 4.7 Hz, 1H, H-3), 4.98 (s, 1H, H-1), 4.52 (dt, *J* = 8.2, 2.4 Hz, 1H, CH Ser), 4.40 – 4.34 (m, 2H, CH_2 Fmoc), 4.34 – 4.18 (m, 3H, H-4 + CH Fmoc + CH_{2a} Ser), 4.15 – 3.98 (m, 2H, H-5), 3.89 (dd, *J* = 9.0, 2.6 Hz, 1H, CH_{2b} Ser), 2.12 (s, 3H, Ac), 2.06 (s, 3H, Ac), 1.51 (s, 18H, *t*Bu). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.93 (COOH), 169.68, 169.63 (C=O Ac), 155.94 (C=O Fmoc), 144.10, 144.03, 141.38 (Cq Fmoc), 129.17, 128.36, 127.79, 127.26, 125.49, 125.46, 125.43, 120.04 (CH Fmoc arom.), 105.09 (C-1), 79.30, 79.19 (C-4), 74.59 (C-2), 71.03 (C-3), 68.73 (CH_2 Ser), 67.48, 67.42, 67.37 (C-5 + CH_2 Fmoc), 54.20 (CH Ser), 47.24 (CH Fmoc), 29.97, 29.93, 29.89, 29.85 (CH_3 *t*Bu), 20.75, 20.63 (CH_3 Ac). **^{31}P NMR** (162 MHz, $CDCl_3$) δ -8.66. **IR cm^{-1}:** 2981, 2935, 1750, 1717,1506, 1451, 1371, 1238, 1214, 1039, 999, 758, 738. **HRMS:** $[C_{35}H_{46}NO_{14}P + Na]^+$ found: 758.2565, calculated: 758.2548.

Solid phase synthesis

*General procedure A, synthesis of protected 5-phosphate-ribose intermediate peptides **21** and **22***

The first intermediate peptide was synthesized using standard automated SPPS protocols starting from preloaded TentaGel® HMBA resin. The following amino acids were used: Fmoc-Pro-OH, Fmoc-Lys(Tfa)-OH, Fmoc-Ala-OH, Fmoc-Gly-OH. For the manual couplings, the resin was transferred into a fritted syringe. From the α/β-phosphoribosylated serine building block 12 or 14 (2 eq.), was dissolved in 0.25 M HCTU solution in NMP (2 eq.). This solution was added to the resin together with a 1 M DIPEA solution in NMP (4 eq). The resin was shaken overnight before washed with DCM and NMP. Standard Fmoc-based SPPS protocols were used to couple the last three amino acids. Final Fmoc cleavage, capping and washing (DCM and NMP) obtained the protected 5-phosphate-ribose peptides **21** and **22**.

General procedure B, on-resin tBu removal

The resin was treated with 10 v/v% TFA in DCM for 30 minutes. The progress of the reaction was monitored by on-resin ^{31}P-NMR analysis by the use of an acetone capillary. If the *t*Bu removal was not fully complete the treatment was repeated until full deprotection was observed. After full deprotection of the phosphate the resin was washed twice with pyridine.

General procedure C, on-resin ADP-ribosylation

The resin was washed thoroughly with dry MeCN and purged with nitrogen gas. To the resin a solution of **5** in MeCN (0.3 M, 3 eq.) was added together with ETT solution in MeCN (0.25 M, 6 eq.). The resin was shaken for 30 minutes before the resin was washed with dry MeCN. CSO in MeCN (0.5 M, 10 eq.) was added to the resin and shaken for 10 minutes after which the resin was washed with MeCN, DCM and DMF. A solution of DBU in DMF (0.5 M, 10 eq.) was added to the resin and shaken for 10 minutes. The resin was washed extensively with DMF, MeCN and DCM.

General procedure D, cleavage of the crude peptides from the resin
The resin was treated with 5 mL of a saturated solution of NH_3 in TFE for 4 hours. To the cleavage cocktail 5 mL of sat. aq. NH_4OH was added and the resin was shaken overnight. The resin was filtered, washed with milli-Q water and the filtrate was concentrated *in vacuo*. The crude peptide was dissolved in a minimal amount of MeOH with three drops of AcOH and precipitated in cold Et_2O. The precipitate was centrifuged and the liquids were decanted to obtain the crude peptide.

Ac-Pro-Ala-Lys-Ser(5-*O*-phosphate-α-D-ribosyl)-Ala-Pro-Ala-Pro-Lys-Lys-Gly-NH_2 (23)
General procedure A was applied with **6** to 200 mg HMBA-TentaGel® S resin (50 µmol) followed by general procedure B and D. RP-HPLC with a gradient of 0 -> 20 % B in A (eluens A, 10 mM NH_4OAc in H_2O, eluens B MeCN) gave pure fractions which were concentrated, redissolved in H_2O and lyophilized to obtain the title compound as a white solid (3.20 mg, 2.45 µmol, 5%). **^{1}H NMR** (600 MHz, D_2O) δ 5.23 (d, *J* = 4.4 Hz, 1H, H1' ribose). **^{31}P NMR** (202 MHz, D_2O) δ 1.33. **LC-MS:** (0 -> 20% B in A): Rt = 7.56. **HRMS:** $[C_{54}H_{94}N_{15}O_{20}P + H]^+$ found: 1304.6610, calculated: 1304.6610. $[C_{54}H_{94}N_{15}O_{20}P + 2H]^{2+}$ found: 652.8356, calculated: 652.8341.

Ac-Pro-Ala-Lys-Ser(5-*O*-phosphate-β-D-ribosyl)-Ala-Pro-Ala-Pro-Lys-Lys-Gly-NH_2 (24)
General procedure A was applied with **7** to 200 mg HMBA-TentaGel® S resin (50 µmol) followed by general procedure B and D. RP-HPLC with a gradient of 0 -> 20 % B in A (eluens A, 10 mM NH_4OAc in H_2O, eluens B MeCN) gave pure fractions which were concentrated, redissolved in H_2O and lyophilized to obtain the title compound as a white solid (4.48 mg, 3.43 µmol, 7%). **^{1}H NMR** (600 MHz, D_2O) δ 5.12 (s, 1H, H1' ribose). **^{31}P NMR** (202 MHz, D_2O) δ 1.29. **LC-MS:** (0 -> 20% B in A): Rt = 7.11. **HRMS:** $[C_{54}H_{94}N_{15}O_{20}P + 2H]^{2+}$ found: 652.8345, calculated: 652.8341.

Ac-Pro-Ala-Lys-Ser(5-*O*-adenosine diphosphate-α-D-ribosyl)-Ala-Pro-Ala-Pro-Lys-Lys-Gly-NH_2 (1)
General procedure A was applied with **6** to 400 mg HMBA-TentaGel® S resin (100 µmol) followed by general procedure B, C and D. RP-HPLC with a gradient of 0 -> 20 % B in A (eluens A, 10 mM NH_4OAc in H_2O, eluens B MeCN) gave pure fractions which were concentrated, redissolved in H_2O and lyophilized to obtain the title compound as a white solid (7.58 mg, 4.65 µmol, 5%). **^{1}H NMR** (600 MHz, D_2O) δ 8.48 (s, 1H, H-2), 8.23 (s, 1H, H-8), 6.10 (d, *J* = 6.0 Hz, 1H H1' adenosine), 4.94 (d, *J* = 4.3 Hz, 1H, H1' ribose). **^{31}P NMR** (202 MHz, D_2O) δ -10.51, -10.61, -10.65, -10.76. **LC-MS:** (0 -> 20% B in A): Rt = 6.01. **HRMS:** $[C_{64}H_{106}N_{20}O_{26}P_2 + 2H]^{2+}$ found: 817.362, calculated: 817.3604.

Ac-Pro-Ala-Lys-Ser(5-*O*-adenosine diphosphate-β-D-ribosyl)-Ala-Pro-Ala-Pro-Lys-Lys-Gly-NH_2 (2)
General procedure A was applied with **7** to 200 mg HMBA-TentaGel® S resin (50 µmol) followed by general procedure B, C and D. RP-HPLC with a gradient of 0 -> 20 % B in A (eluens A, 10 mM NH_4OAc in H_2O, eluens B MeCN) gave pure fractions which were concentrated, redissolved in H_2O and lyophilized to obtain the title compound as a white solid (5.86 mg, 3.59 µmol, 7%). **^{1}H NMR** (600 MHz, D_2O) δ 8.40 (s, 1H, H-2), 6.26 (d, *J* = 5.5 Hz, 1H H1'adenosine), 4.87 (H1' ribose overlapped with D_2O). **^{31}P NMR** (202 MHz, D_2O) δ -10.46, -10.56, -10.63, -10.74. **LC-MS**: (0 -> 20% B in A): Rt = 6.77. **HRMS:** $[C_{64}H_{106}N_{20}O_{26}P_2 + 2H]^{2+}$ found: 817.3615, calculated: 817.3604.

Biochemical evaluation

Plasmids and Proteins

The expression construct for wildtype ARH3 was a gift from Prof Paul Hergenrother (University of Illinois) and mutation was introduced by site directed mutagenesis as described earlier.[29] For expression of recombinant hARH3 proteins plasmid were transformed in Rosetta (DE3) cells and grown to an OD_{600} of 0.6 in LB medium supplemented with 2 mM $MgCl_2$ and appropriate antibiotics. Expression was induced with 0.4 mM isopropyl β-D-1-thioglactopyranoside (IPTG), cells were grown overnight at 17 °C and harvested by centrifugation. Proteins were purified by Ni^{2+}-NTA chromatography (Jena Bioscience) according to the manufacturer's protocol using the following buffers: all buffers contained 50 mM TrisHCl (pH 8), 500 mM NaCl and 10 mM $MgCl_2$; additionally, the lysis buffer contained 25 mM, the washing buffer 40 mM, and the elution buffer 500 mM imidazole. Proteins were dialysed overnight against 50 mM TrisHCl (pH 8), 200 mM NaCl, 1 mM dithiothreitol and 5% (v/v) glycerol. Recombinant human NUDT5 was a kind gift from the Protein Science Facility (PSF) at Karolinska Institute (Sweden).

Peptide demodification assay

20 µM histone H2B peptides **1** and **2** (aa 4-14) were demodified by incubation with 1 µM hARH3 for 20 minutes at rt in assay buffer (50 mM TrisHCl [pH 8], 200 mM NaCl, 10 mM $MgCl_2$ and 1 mM dithiothreitol). The released ADPr was converted into AMP by co-incubation with 0.2 µM human NUDT5. Reactions were stopped and analyzed by performing the AMP-Glo™ assay (Promega) according to the manufacturer's protocol. Luminescence was recorded on a SpectraMax M5 plate reader (Molecular Devices) and data analyzed with GraphPad Prism 7. Control reactions were carried out in absence of peptide.

*Chemical degradation assay of peptide **1***

Peptide **1** (0.6 mg) was dissolved in 300 µL water to make a stock solution of peptide **1**. This stock solution was divided in three portions of 100 µL and each portion was mixed respectively with 100 µL 0.2 M TFA in water, 100 µL 1.0 M NH_2OH in water or 100 µL 0.2 M NaOH in water. Samples of 30 µL were taken for LC-MS analysis at various time points (15 minutes, 1 hour, 2.5 hours, 5 hours and 24 hours). The samples treated with TFA or NH_2OH were injected directly, the sample treated with NaOH was quenched with 2 µL AcOH prior to injection. Peptide degradation was monitored by UV-detection and mass fragments.

References

[1] B. Lüscher, M. Bütepage, L. Eckei, S. Krieg, P. Verheugd, B. H. Shilton, *Chem. Rev.* **2018**, *118*, 1092–1136.

[2] B. A. Gibson, W. L. Kraus, *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 411–424.

[3] M. O. Hottiger, P. O. Hassa, B. Lüscher, H. Schüler, F. Koch-Nolte, *Trends Biochem. Sci.* **2010**, *35*, 208–219.

[4] S. Laing, M. Unger, F. Koch-Nolte, F. Haag, *Amino Acids* **2011**, *41*, 257–269.

[5] N. Ogata, K. Ueda, H. Kagamiyama, O. Hayaishi, *J. Biol. Chem.* **1980**, *255*, 7616–7620.

[6] N. Ogata, K. Ueda, O. Hayaishi, *J. Biol. Chem.* **1980**, *255*, 7610–7615.

[7] C. M. Daniels, S. Ong, A. K. L. Leung, *Mol. Cell* **2015**, *58*, 911–924.

[8] O. Leidecker, J. J. Bonfiglio, T. Colby, Q. Zhang, I. Atanassov, R. Zaja, L. Palazzo, A. Stockum, I. Ahel, I. Matic, *Nat. Chem. Biol.* **2016**, *12*, 998–1000.

[9] L. Palazzo, O. Leidecker, E. Prokhorova, H. Dauben, I. Matic, I. Ahel, *Elife* **2018**, *7*, 1–12.

[10] J. J. Bonfiglio, P. Fontana, Q. Zhang, T. Colby, I. Gibbs-Seymour, I. Atanassov, E. Bartlett, R. Zaja, I. Ahel, I. Matic, *Mol. Cell* **2017**, *0*, 1–9.

[11] L. Palazzo, O. Leidecker, E. Prokhorova, H. Dauben, I. Matic, I. Ahel, *Elife* **2018**, 1–12.

[12] I. Gibbs-Seymour, P. Fontana, J. G. M. Rack, I. Ahel, *Mol. Cell* **2016**, *62*, 432–442.

[13] E. E. Alemasova, O. I. Lavrik, *Nucleic Acids Res.* **2019**, *47*, 3811–3827.

[14] H. A. V Kistemaker, A. P. Nardozza, H. S. Overkleeft, G. A. van der Marel, A. G. Ladurner, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10634–10638.

[15] G. J. van der Heden van Noort, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *J. Org. Chem.* **2010**, *75*, 5733–5736.

[16] H. A. V Kistemaker, G. J. van der Heden van Noort, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2013**, *15*, 2306–2309.

[17] A. G. Volbeda, H. A. V. Kistemaker, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, J. D. C. Codée, *J. Org. Chem.* **2015**, *80*, 8796–8806.

[18] A. Bérces, D. M. Whitfield, T. Nukada, I. S. Z, A. Obuchowska, J. J. Krepinsky, *Can. J. Chem.* **2004**, *82*, 1157–1171.

[19] G. J. van der Heden van Noort, P. van Delft, N. J. Meeuwenoord, H. S. Overkleeft, G. a van der Marel, D. V Filippov, *Chem. Commun.* **2012**, *48*, 8093–5.

[20] P. Fontana, J. J. Bonfiglio, L. Palazzo, E. Bartlett, I. Matic, I. Ahel, *Elife* **2017**, *6*, 1–20.

[21] J. Moss, S. chen Tsai, R. Adamik, H. chia Chen, S. J. Stanley, *Biochemistry* **1988**, *27*, 5819–5823.

[22] J. Moss, N. J. Oppenheimer, R. E. West, S. J. Stanley, *Biochemistry* **1986**, *25*, 5408–5414.

[23] L. Palazzo, B. Thomas, A. S. Jemth, T. Colby, O. Leidecker, K. L. H. Feijs, R. Zaja, O. Loseva, J. C. Vert, I. Matic, T. Helleday, I. Ahel, *Biochem. J.* **2015**, *468*, 293–301.

[24] Y. Zhang, J. Wang, M. Ding, Y. Yu, *Nat. Methods* **2013**, *10*, 981–984.

[25] K. Wakabayashi, W. Pigman, *Carbohydr. Res.* **1974**, *35*, 3–14.

[26] M. B. Goshe, T. P. Conrads, E. A. Panisko, N. H. Angell, T. D. Veenstra, R. D. Smith, *Anal. Chem.* **2001**, *73*, 2578–2586.

[27] Y. Oda, T. Nagasu, B. T. Chait, *Nat. Biotechnol.* **2001**, *19*, 379–382.

[28] L. Wells, K. Vosseller, R. N. Cole, J. M. Cronshaw, M. J. Matunis, G. W. Hart, *Mol. Cell. Proteomics* **2002**, *1*, 791–804.

[29] R. K. Morgan, M. S. Cohen, *ACS Chem. Biol.* **2015**, *10*, 1778–1784.

Chapter 3

A Unified and Versatile Method to Synthesise Native Mono-ADP-Ribosylated Peptides

Introduction

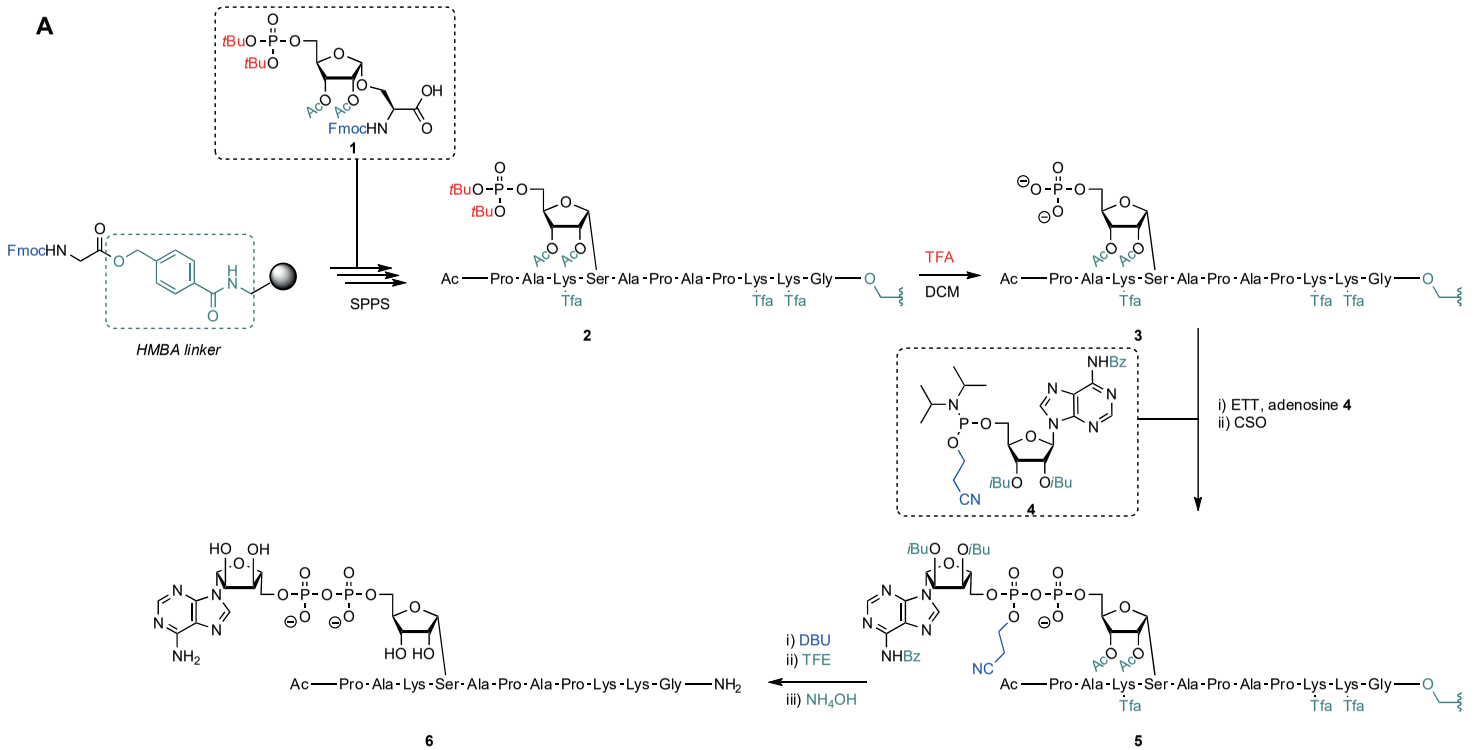
In the research on ADP-ribosylation, multiple types of amino acid side chains have been discovered to function as acceptors in this post-translational modification (PTM), that is introduced onto target proteins by (ADP-ribosyl)transferases (ARTs) from the PARP-family. Historically, arginine (Arg) was among the first to be identified as ADPr acceptor[1,2] and for decades acidic residues like glutamic acid (Glu) [3,4] and aspartic acid (Asp)[5,6] were thought to be the most common acceptors of ADPr.[5,7] Other residues such as lysine (Lys),[3] histidine or diphthamide[8,9] and cysteine (Cys)[10] were encountered as well but it was not until 2016 that serine (Ser) was added to this list.[11] After the initial discovery, it quickly became apparent that Ser in fact was the major acceptor of ADP-ribose upon DNA damage,[12] ushering in a new era for the field with a sharp increase of interest for Ser as a modification site. In the past few years, many insights have been acquired regarding the biological role of Ser-ADPr,[13] its chemical stability,[14] the distinctly different mechanism by which Ser-ADPr is introduced by a previously unknown protein complex of PARP1 with HPF1[15,16] and the reversal of this modification by the glycohydrolase ARH3.[17] Following Ser, tyrosine (Tyr) was discovered as a less frequently occurring acceptor of ADPr[18,19] and the synthetic efforts towards the synthesis of ADPr-Tyr peptides is the subject of Chapter 4. The present chapter describes the development of a novel strategy for the synthesis of ADP-ribosylated peptides, modified on Ser, Thr and Cys. Although Cys can function as acceptor of ADPr,[20] it can also serve as a chemically and enzymatically stable bioisostere of Ser, that is less susceptible to hydrolysis by ARH3. Thr-ADPr is a rare acceptor of ADPr and until now, no family member of mammalian PARP has been identified as being able to transfer ADPr onto threonine (Thr). Literature suggests however, that some bacterial enzymes are able to ADP-ribosylate Thr residues, for example, in human ubiquitin (Ub).[21] Thr-modified Ub-ADPr ablates the function of PolyUb on multiple levels, e.g. biosynthesis and Ub recognition and plays a crucial role in colonization of the bacteria, demonstrating the biological significance of Thr as acceptor of ADPr.

Until now, the synthesis of *mono*-ADP-ribosylated (MARylated) peptides relied on a solid phase strategy developed by Kistemaker *et al.*,[22] of which an application to Ser-ADPr peptides is described in Chapter 2. This strategy uses a phosphorylated building block **1** for the introduction of ribosylated amino acids in a peptide sequence to give immobilized phosphoribosylated peptide **2** (Scheme 1A). The pattern of the semipermanent protection in **2** is characterized by the protecting groups sensitive for alkali and nucleophilic bases, such as ammonia. The secondary hydroxyls of the ribose moiety are protected with acetyl groups, the side chains of the Lys residues in the peptide sequence bear trifluoroacetyl protection (Tfa) and a HMBA linker to the resin is used, which is cleaved by ammonolysis. Orthogonality for the installation of the ADPr moiety is achieved by

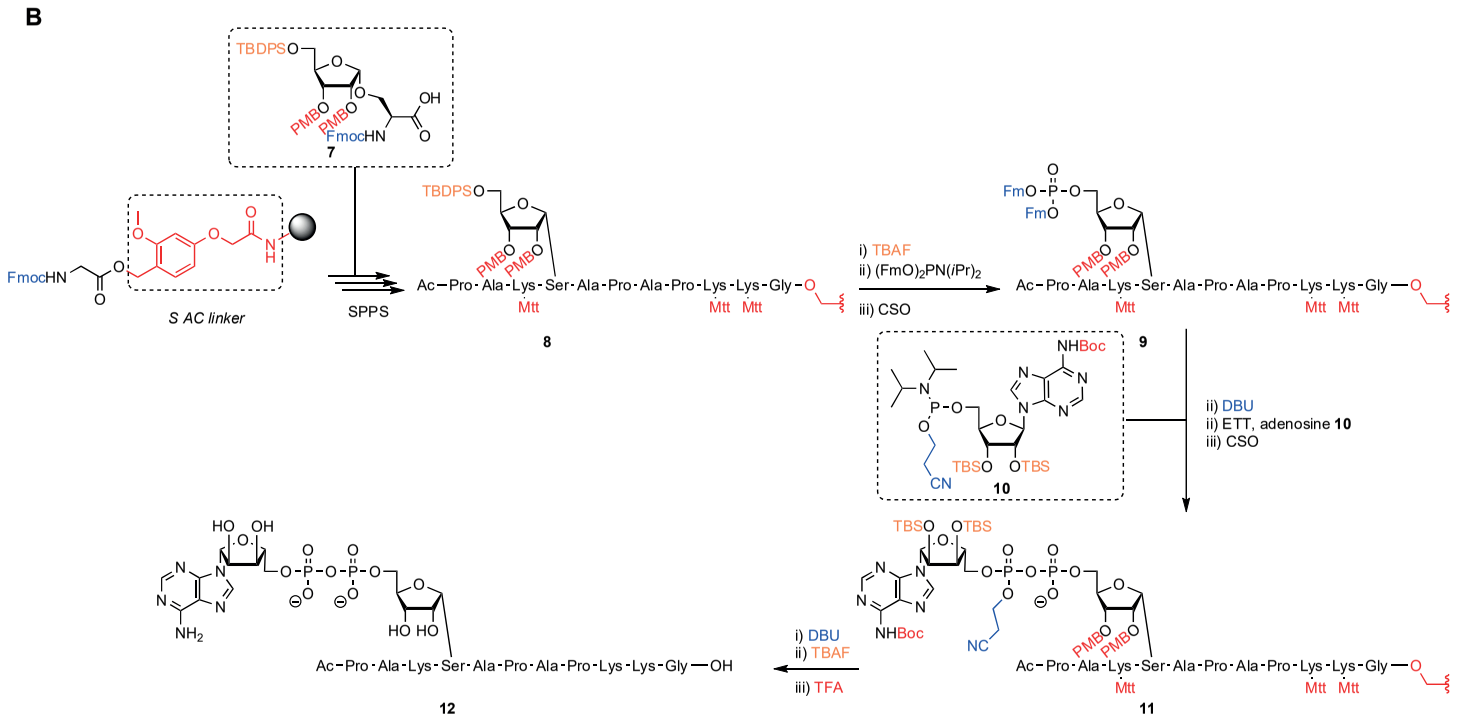
using a *t*Bu-protected phosphotriester that can be deprotected with TFA to give **3**. In the next step the pyrophosphate is constructed by coupling protected adenosine amidite **4** to phosphomonoester **3**, followed by oxidation of the resulting P^{III} – P^{V} intermediate to give protected ADPr-peptide resin **5**. Finally, a sequence of deprotection steps including DBU, saturated NH_3 in trifluoroethanol (TFE) and aqueous NH_4OH furnishes the desired MARylated peptide **6**. Although a variety of sequences, MARylated on the Gln, Asn or Cit site[22] as well as Ser[14] have been prepared using this method, some drawbacks remain such as the presence of a carboxamide at the *C*-terminus of the MARylated peptide. Also, not all peptide sequences can be synthesized equally efficiently. For instance, when the ADPr-modification site in a peptide is flanked by Ser or Thr, additional protecting group manipulations are needed.

For sequences with a Ser or Thr at the *C*-terminal flanking sequence to the modification site, the trityl (Trt) protecting group of the hydroxyls in these amino acids must be replaced by an acetyl to allow further processing of the incorporated *t*Bu-protected phosphotriester. For the same reason, sequences with Ser or Thr at the *N*-terminal flanking sequence are not amenable for the synthesis by the strategy outlined in the Scheme 1A. Based on these considerations a new strategy to MARylated peptides, that is outlined in Scheme 1B, was developed. Relying on standard Fmoc-based solid phase peptide synthesis (SPPS), the commercially available TentaGel® resin equipped with the highly acid sensitive S AC linker, cleavable by TFA at concentrations as low as 1% in DCM, was selected as solid support. For the side chain protecting groups on Ser-, Thr- or Lys-residues, the trityl (Trt) and 4-methyltrityl (Mtt) were chosen as they can be readily cleaved at low concentrations of TFA in DCM during global deprotection at the end of the synthesis. Ribosylated building block **7** (Scheme 1B) has been selected for the incorporation at the prospected ADPr site via SPPS. This newly designed building block is characterized by the following protecting groups: i) the PMB groups for the secondary hydroxyls not only allow α-selective glycosylation but are also made semipermanent to be cleaved in the final stage of the assembly of the MARylated peptides, ii) the bulky TBDPS group, aiding to attain α-selectivity during glycosylation allows for the orthogonal liberation of the 5' hydroxyl prior to the on-resin phosphorylation during the SPPS of the MARylated peptides. In this way the number of protecting group manipulations toward the solution phase synthesis of building block **7** is limited while the introduction of the phosphate ester is postponed to SPPS. Thus, after completion of the immobilized oligopeptide **8**, the TBDPS protecting group is removed using a fluoride anion. The released alcohol is phosphitylated using an Fm-protected phosphoramidite and the resulting phosphite triester is oxidized by CSO yielding protected phosphoribosylated peptide **9**. The ADP-ribose moiety is introduced, upon the deprotection of the phosphotriester by DBU, using the P^{III} – P^{V} methodology

from Chapter 2 by coupling with protected adenosine phosphoramidite **10** and CSO oxidation.[14,22] Finally, removal of the cyanoethyl on the pyrophosphate and the silyl groups on the 2' and 3' position of the adenosine, global deprotection of the peptide and cleavage from the resin will give mono-ADPr peptide **12**.



Scheme 1. Overview of the synthetic strategies for the synthesis of MARylated peptides with a model H2B sequence. **A)** Strategy based on orthogonality between acid and bases as utilized in Chapter 2 (nucleophilic and non-nucleophilic). **B)** Strategy as presented in this Chapter. Mtt = 4-methyl trityl Tfa = trifluoroacetyl. Tfe = 2,2,2-trifluoroethanol.



Scheme 1. Overview of the synthetic strategies for the synthesis of MARylated peptides with a model H2B sequence. **A)** Strategy based on orthogonality between acid and bases as utilized in Chapter 2 (nucleophilic and non-nucleophilic). **B)** Strategy as presented in this Chapter. Mtt = 4-methyl trityl Tfa = trifluoroacetyl. Tfe = 2,2,2-trifluoroethanol.

Results and discussion

Building block synthesis

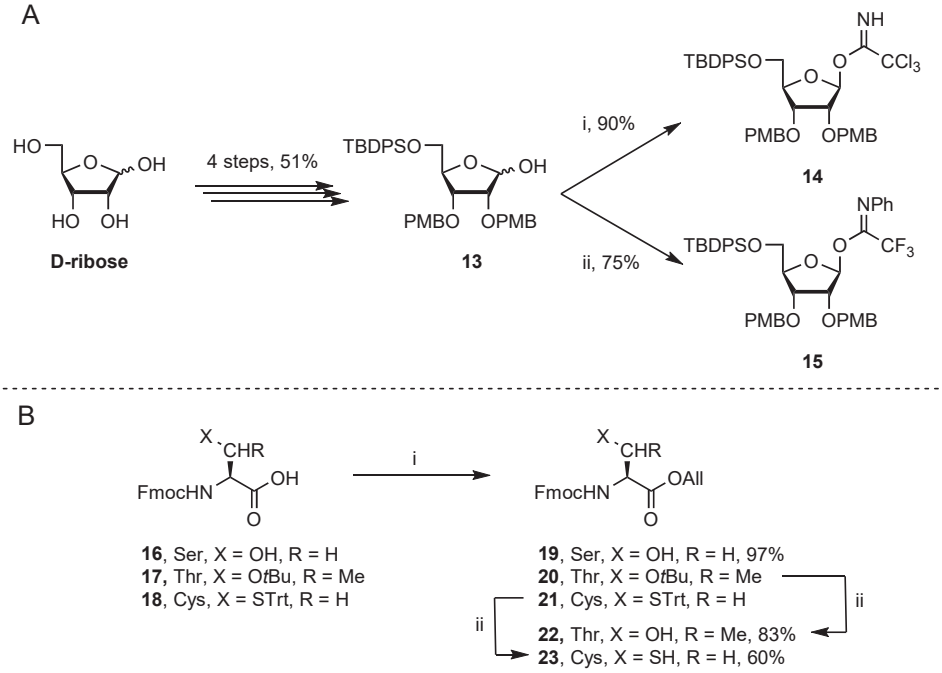
Ribofuranosylated Fmoc amino acids

Firstly, the glycosylation procedure towards the respective ribofuranosylated Fmoc amino acids was optimized by testing two ribosyl donors, the known *N*-(phenyl)trifluoroacetimidate donor **15**[23] and trichloroacetimidate donor **14** (Scheme 2A). The latter was chosen for its synthetic accessibility as donor **15** requires the labor-intensive synthesis of 2,2,2-trifluoro-*N*-phenylacetimidoyl chloride as a reagent. On the other hand, trichloroacetimidate donors can undergo a Chapman-like rearrangement to form the unwanted glycosylamides by-products under glycosylation conditions.[24,25] Trichloroacetamide donor **14** (Scheme 2A) was prepared in 90% yield by treating known partially protected ribose **13**[23] with trichloroacetonitrile in the presence of DBU whereas trifluoroacetamide donor **15** was obtained in 75% yield by reacting **13** with 2,2,2-trifluoro-*N*-phenyl-acetimidoyl chloride in the presence of Cs_2CO_3.[23]

Scheme 2B shows the synthesis of the appropriately protected Ser **19**, Thr **22** and Cys **23** to be used as the acceptors. The *C*-terminal benzyl ester, the use of which is described in Chapter 2, had to be replaced as the hydrogenation conditions of its removal are incompatible with the PMB protecting groups. The allyl ester was chosen as a suitable replacement since it can be selectively removed by treatment with catalytic $Pd(PPh_3)_4$ under neutral conditions. To prepare appropriately protected Ser acceptor **19**, commercially available Fmoc-Ser-OH (**16**) was treated with allyl bromide and DIPEA. For the synthesis of Thr and Cys acceptors **22** and **23** the corresponding side chain protected Fmoc amino acids **17** and **18** were converted in allyl esters **20** and **21** under the same conditions as used for **16**. Subsequent removal of the acid sensitive side chain protecting groups with TFA yielded acceptors **22** and **23**.

With all required building blocks in hand, the glycosylation reaction could be optimized. As a model reaction, the condensation of Ser acceptor **19** with ribosyl donors **14** and **15** was examined by varying the reaction conditions in terms of temperature, concentration and nature of the activator, the results of which are listed in Table 1. Coupling of donor **14** with acceptor **19** using the same conditions as described in Chapter 2 gave a low yield of the target ribofuranosylated Fmoc-Ser **25** together with a significant amount of side-product **24,** originating from the Chapman-type rearrangement (entry 1).[25,26] Changing the activator to TfOH (entry 2) led to acid-catalyzed cleavage of one or both PMB-protecting groups of which the corresponding cation was presumably scavenged by the acceptor, resulting in Fmoc-Ser(PMB)-OAll together with a complex mixture of ribose derived products. In an attempt to reduce the loss of the PMB group, TBSOTf was used as activator (entry 3). It

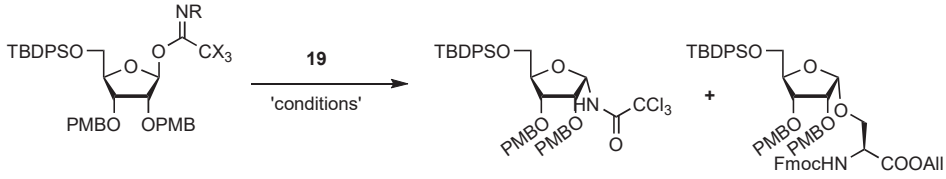
was reasoned that the softer character of the TBS cation would decrease the acidity of the glycosylation conditions and lead to less cleavage of the PMB groups. Although the use of TBSOTf significantly improved the yield of product **25**, side product **24** was still formed in 12%. A two-fold increase of the concentration of donor **14** augmented the formation of side product **24** to 48% (entry 4). Finally, increasing the temperature of the reaction to -40 °C did improve the yield of the glycosylation to 60% (entry 5).



Scheme 2. A) Synthesis of ribosyl donors **14** and **15**. Reagents and conditions i) Cl_3CCN, DBU, acetonitrile. ii) $Cl(C{=}NPh)CF_3$, Cs_2CO_3, acetone. B) Synthesis of amino acid acceptors **19**, **22** and **23**. Reagents and conditions: i) Allyl-Br, DIPEA, DMF. ii) TFA, DCM.

For donor **15**, the same set of reaction conditions were tested (entries 6 – 10). It is of interest to note that when using TfOH as activator with donor **15**, the amount of acid-catalyzed PMB ether cleavage has been significantly reduced in comparison with trichloroacetimidate counterpart **14**. However, the results of the glycosylations proved to be more reproducible with TBSOTf, particularly for scaling up the reaction.

Table 1. Optimization of the glycosylation conditions with donors **14** and **15** and acceptor **19**. The concentration (C) is the concentration of the donor in DCM as solvent. Activators were used in 0.1 equivalent relative to the donor and the reactions were carried out at a 0.2 mmol scale.



14 R = H, X = Cl
15 R = Ph, X = F

24

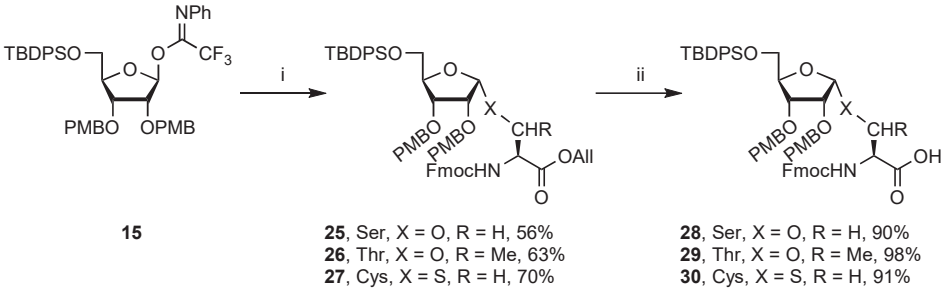
25

Entry	Donor	C (M)	T (°C)	Activator	Side product 24 (%)	Yield (%)
1	**14**	0.1	-50	TMSOTf	34	23
2	**14**	0.1	-50	TfOH	n.d.	0
3	**14**	0.1	-50	TBSOTf	12	39
4	**14**	0.2	-50	TBSOTf	48	17
5	**14**	0.1	-40	TBSOTf	12	60
6	**15**	0.1	-50	TMSOTf	-	32
7	**15**	0.1	-50	TfOH	-	53
8	**15**	0.1	-50	TBSOTf	-	56
9	**15**	0.2	-50	TBSOTf	-	23
10	**15**	0.1	-40	TBSOTf	-	48

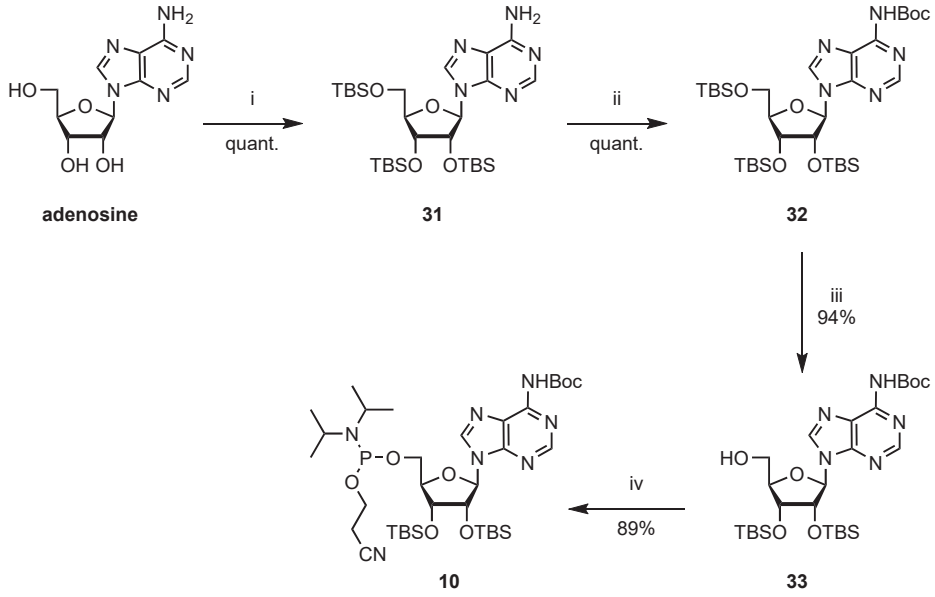
After the optimization of glycosylation conditions, the synthesis of ribosylated Fmoc amino acids **28**, **29** and **30** (Scheme 3) that are suitable building blocks for the solid phase synthesis of ADPr-peptides was undertaken. To this end, donor **15** was coupled to the appropriately protected amino acid acceptors **19** (Ser), **22** (Thr) and **23** (Cys) to give the desired products **25** – **27** in a high α-selectivity (no β-product was observed) and in good yields (56 – 70%). Removal of the allyl protecting group in **25** – **27** with catalytic amounts of $Pd(PPh_3)_4$ and 1,3-dimethylbarbituric acid (DMBA) as the allyl cation scavenger furnished the required building blocks **28** – **30** in excellent yields.

Adenosine building block

The final building block, needed for the assembly of MARylated peptides, is adenosine phosphoramidite **10** that enables the introduction of the ADPr moiety via the established P^{III} – P^{V} procedure (Scheme 4).[27] Silylation of the hydroxyl functions in adenosine with TBS-Cl was followed by the introduction of a Boc group to the exocyclic amine yielding **32** in a quantitative yield. An acid mediated regioselective desilylation[28] of the 5' position in **32** and subsequent phosphitylation of **33** under standard conditions gave phosphoramidite **10**.



Scheme 3. Synthesis of Fmoc-based SPPS building blocks **28** – **30**. Reactions and conditions: i) TBSOTf, acceptors **19**, **22** or **23**, DCM, -50 °C. ii) $Pd(PPh_3)_4$, DMBA, DCM.



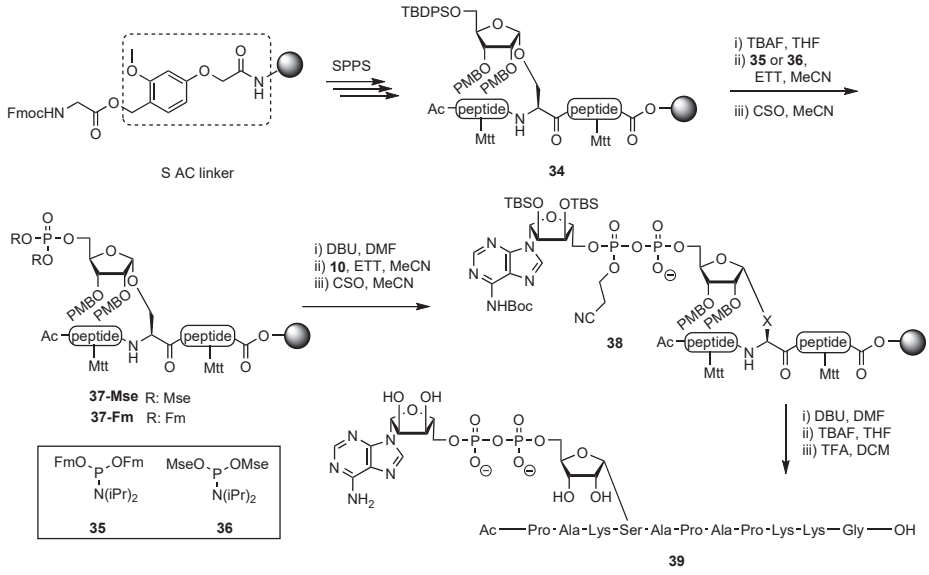
Scheme 4. Synthesis of adenosine amidite **10**. Reagents and conditions: i) TBS-Cl, imidazole, DMF, 50 °C. ii) Boc_2O, DMAP, THF, reflux. iii) TFA, H_2O, THF, 0 °C. iv) 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, DCM.

Solid phase synthesis of mono-ADPr-peptides

With all building blocks in hand the solid phase assembly of MARylated peptide **39**, derived from the *N*-terminus of histone H2B, was undertaken. The standard Fmoc-based methodology was combined with amino acid building blocks having highly acid sensitive protecting groups on the side chains (Mtt for Lys, Trt for Ser and Thr). As depicted in Scheme 5, TentaGel® S AC resin preloaded with glycine (Gly) was elongated using the selected protected amino acid building blocks, including ribofuranosylated Fmoc-Ser-OH

28, to give immobilized peptide **34** having the required sequence. The ensuing cleavage of the TBDPS-protecting group was tested using three different F^- anion sources: TEA·3HF, HF·pyridine and TBAF. Both TEA·3HF and HF·pyridine needed reaction times of up to 16 hours to fully remove the silyl protecting group whereas a 1 M TBAF solution in THF ensured full deprotection in 30 minutes. The TBAF treatment was not only superior with regard to reaction rate but also in the quality of the product according to LC-MS analysis of the peptides after desilylation and removal of all other protecting groups and cleavage from the solid support. After desilylation of the primary hydroxyl in the ribose moiety, two phosphoramidite reagents were investigated for obtaining the ribosyl-5-phosphomonoester. Two phosphoramidites, both bearing base sensitive protecting groups, were tested i.e., reagent **35** with fluorenylmethyl (Fm) groups and reagent **36** equipped with 2-methylsulfonylethyl (Mse) groups (Scheme 5). The phosphitylation of the immobilized peptide, having a ribose with a free 5-OH by either reagent **35** or **36**, followed by the CSO mediated oxidation of the formed phosphite to the phosphotriester intermediate provided the immobilized, fully protected phosphoribosyl peptides **37**-**Fm** and **37**-**Mse**, respectively. To convert the phosphotriester into a phosphomonoester, the Fm and Mse protecting groups were removed by treatment of the resins **37**-**Fm** and **37**-**Mse** with 10% DBU in DMF. Monitoring the reaction progress for 20 minutes by LC-MS showed that both Fm-protecting groups were completely eliminated whereas only one Mse-group had been removed, stalling this elimination at the phosphodiester stage. Although the crystalline Mse reagent **36** is easier to handle than the amorphous **35**, the Fm protecting group was chosen for its faster deprotection rate. Thus, the assembly of the MARylated peptide was continued with the deprotection of **37**-**Fm.** Coupling of the resulting phosphomonoester with adenosine phosphoramidite **10** and oxidation of the P^{III} – P^{V} intermediate gave immobilized peptide **38**, containing a partially protected pyrophosphate moiety.

Now the terminal stage of the SPPS has been reached. The cyanoethyl group was removed from the pyrophosphate in **38** by 10% DBU in DMF, the silyl ethers were deprotected with TBAF and finally, the remaining protecting groups were removed with concomitant cleavage of the target MARylated peptide from the resin by treatment with a 10% TFA solution in DCM containing 2.5% TIS as a scavenger for trityl and *para*-methoxybenzyl carbocations. Monitoring of the latter deprotection by LC-MS analysis revealed that the trityl and PMB-protecting groups were split off instantly while the more stable Boc protecting group on the exocyclic amine of adenosine needed at least 4 hours for its removal. Purification with RP-HPLC of the obtained crude product led to the isolation of MARylated peptide **39**, derived from the *N*-terminus histone H2B in a 3.6% overall yield. The synthesis of the same ADPr-peptide with a *C*-terminal carboxamide is described in Chapter 2.



Scheme 5. Synthesis of MARylated H2B peptide **39**.

The above described successful synthesis of Ser-ADPr peptide **39** by application of the newly developed SPPS methodology was an incentive to assemble Ser-ADPr oligopeptides **40** and **41.** The Ser residues in both sequences have been confirmed to be an ADPr acceptor by LC-MS/MS analysis of biological systems.[29,30] Synthesis of peptide **40** proceeded uneventful whereas *en route* to peptide **41**, the incorporation of Thr-7 proved to be difficult and the repetition of three coupling cycles was needed. Peptide **41** contains an Arg, a side chain of which no suitable acid labile protecting group is available, therefore, a bis-Alloc protection for the guanidine function was selected.[22,31] The Alloc groups required an additional treatment of the resin with $Pd(PPh_3)_4$ and DMBA as a scavenger prior to the final acidolysis that furnished **41** in 6.1% overall yield. After successfully demonstrating the robustness of the developed SPPS methodology sequences **40** and **41** were synthesized again but instead, one $^{13}C_6$ labelled Lys residue (**42**) or one $^{13}C_6$ labelled leucine residue (**43**) was incorporated. These isotopically enriched ADPr peptides **42** and **43** were obtained in a 4.6 and 1.2% yield, respectively and can be implemented as tools in proteomics for quantification studies.

Having effectively prepared the Ser-ADPr peptides **39** – **43**, attention was turned to the assembly of ADPr peptides with Thr or Cys as the ADP-ribosylation site. Thr-ADPr peptide **44**, which is sharing part of the sequence of peptide **41**, has been selected to help determine the exact site of modification as the analysis by LC-MS/MS of proteomic mixtures is not always irrefutable.[18] The aforementioned difficulties incorporating the Thr-7

residue in **41** were not encountered in coupling ribosylated Thr-building block **29** to obtain peptide **44**. Another relevant Thr-ADPr peptide is **45**, a motif around the ADP-ribosylation site from human ubiquitin that is modified on Thr-66 by the bacterial effector protein CteC as detected by LC-MS/MS analysis in proteomics studies.[21] It is of interest to note that ADPr peptide **45** includes the amino acids Glu and His. In the SPPS to **45**, the building blocks Fmoc-His(Trt)-OH and Fmoc-Glu(O-2-Ph*i*Pr)-OH were used as both protecting groups are removable by the levels of TFA present in the cleavage cocktail. This demonstrates that with the aid of the herein presented strategy, MARylated peptides with His incorporated in the sequence can be assembled for the first time.

Table 2. Overview of synthetic MARylated peptides **39** – **47**. The amino acids indicated by bold print are the modification sites. When applicable, isotopically labeled sites are indicated by brackets.

Number	Sequence	Yield (%)	Notes
39	Ac-PAK**S**APAPKKG-OH	3.6	n.a.
40	Ac-GK**S**GAALSKKG-OH	11	n.a.
41	Ac-GK**S**SGPTSLFAVTVAPPGARG-OH	6.1	Strenuous coupling of Thr-7
42	Ac-GK**S**GAALS(K)KG-OH	4.6	Indicated site was isotopically labelled
43	Ac-GK**S**SGPTS(L)FAVTVAPPGARG-OH	1.2	Indicated site was isotopically labelled
44	Ac-GKSSGP**T**SLF-OH	9.5	n.a.
45	Ac-KES**T**LHLVLRL-OH	0.94	n.a.
46	Ac-PAK**C**APAPKKG-OH	4.1	EDT added in cleavage cocktail
47	biotin-PAK**C**APAPKKG-OH	1.9	*t*BuOOH was used instead of CSO

As said, Cys in ADP-ribosylated peptides can be considered to be an isosteric Ser replacement that should be more stable to enzymatic hydrolysis. The SPPS of Cys-ADPr peptide **46** as an analogue of **39**, was performed using ribofuranosylated Cys building block **30**. Although the synthesis proceeded uneventfully, deprotection and cleavage of the immobilized Cys-ADPr peptide **46** from the resin led to the detection of peptide Ac-PAK**C**^{PMB}APAPKKG-OH, a side product originating from the migration of the PMB protecting group, to the thiol of the Cys side chain. Addition of ethane dithiol (EDT), a potent thiol-based cation scavenger, to the cleavage cocktail suppressed this side reaction. Lastly, to obtain a useful pull-down tag for biological experiments, the *N*-terminus of ADPr peptide **46** was acylated with biotin to give Cys-ADPr peptide **47**. After completion of the synthesis of **47**, LC-MS analysis of the crude product revealed a main product with a mass 16 Da higher than expected, presumably due to the oxidation of a dialkylsulfide. Since such overoxidation has not been detected during the synthesis of the similar Cys-ADPr peptide **46**, it is postulated that this unwanted reaction has occurred on the sulfur of the biotin tag. Oxidized biotin occurs as both α- and β-sulfoxide and

while the β-form nearly completely ablates its affinity towards avidin the α-biotin sulfoxide still has strong binding properties[32] and can be used without loss of streptavidin pull-down efficiency.[33] Still, to eliminate CSO-mediated overoxidation, the synthesis of ADPr-peptide **46** was repeated and using a milder oxidizing agent (0.5 M *t*BuOOH) chemoselective oxidation of the phosphite species was achieved, leaving biotin intact.

Biological evaluation

Having obtained MARylated peptides **39** – **47**, the differences in enzymatic turn-over of these modifications were investigated. As ARH3 is the only known hydrolase capable to deADP-ribosylate Ser residues,[17] its ability to hydrolytically remove the ADPr moiety from these peptides was first tested (Figure 1a, b). ARH3, but not its catalytic mutants D77N or D78N, was capable of hydrolyzing the glycosidic linkage in **39** (Ser) and **44** (Thr). The turnover of the latter proved not as efficient as the former (Figure 1a and b), which might be a result of steric clash of the additional methyl group of the Thr side chain within the enzyme active site. Note that an incubation time of 45 min was employed whereas Ser-MARylation turn-over is complete within 20 min.[14] In contrast, the ADPr-Cys glycosidic linkage in **46** was largely stable towards enzymatic hydrolysis under the conditions applied. Since ADP-ribosylation of Cys residues is a modification found in human cells, the investigation was expanded to all known human ADP-ribosyl hydrolases and confirmed that only ARH3 could remove the modification from Ser and by extension Thr. Surprisingly, none of the tested hydrolases were able to remove the modification from Cys-ADPr peptide **46** (Figure 1b). This suggests that either the modification is irreversible in human cells or it is reversed by a thus far unidentified enzyme.

To test whether the modification could in principle be reversed, several evolutionary divergent hydrolases of the macrodomain and (ADP-ribosyl)hydrolase family from various lower organisms were tested (Figure 1c). While none of the ARH-like enzymes were able to hydrolyze this particular linkage, *Streptococcus pyogenes* MacroD (*Spy*MacroD)[34] efficiently hydrolyzed the ADP-ribosyl-cysteinyl glycosidic bond. Earlier structural studies on SAV0325, the *Staphylococcus aureus* homologue of *Spy*MacroD, showed a Zn^{2+}-binding motif within the active site and the authors suggested that this zinc centre participates in substrate recognition and catalysis.[35] These results suggest that an increased Lewis acidity, due to the presence of the Zn^{2+} ion, relative to other macrodomains as well as production of Cys, a favorable zinc-coordination ligand, are supporting the reaction. This result clearly shows that hydrolases can readily evolve into efficient Cys deADP-ribosylating enzymes and that such activity may exist in humans. Together, this data provides new insights into the turnover of ADP-ribosylated substrate and highlight the suitability of the MARylated peptides **39** – **47** as tools for the study of ADP-ribosylation.

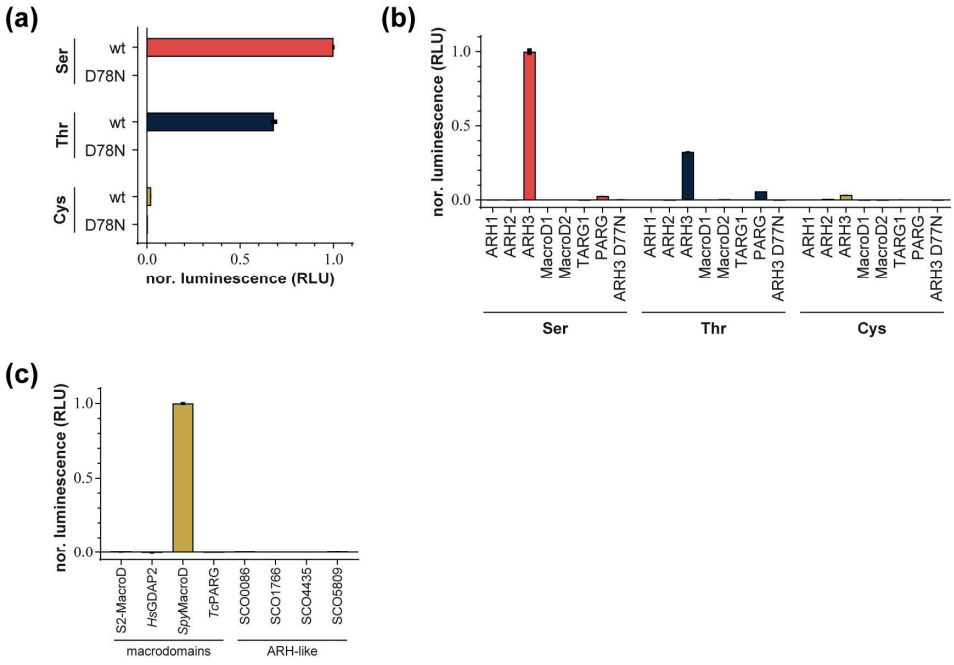


Figure 1. Enzymatic hydrolysis of glycosidic linkages in ADP-ribosylated Ser, Thr and Cys containing peptides **39**, **44** and **46**, respectively. (a-c) Measurements of hydrolase activity against the various ADPr-peptide linkages was assessed by converting released ADPr into AMP via NUDT5 and subsequently measured using the AMP-Glo™ assay (Promega). Samples are background corrected and normalized to ARH3 wt (a+b) or SpyMacroD (c). Data are represented as mean values ± SD measured in triplicates. (a) Hydrolysis of Ser-, Thr- and Cys-linked ADPr by ARH3 wt and catalytic mutant D77N. (b) Hydrolysis of Ser-, Thr- and Cys-linked ADPr by human (ADP-ribosyl) hydrolases. (c) Hydrolysis of Cys-linked ADPr by evolutionary diverged (ADP-ribosyl)hydrolases of the macrodomain and ARH-like families. Abbreviations: S2-MacroD, SARS-CoV-2 nsp3 macrodomain; HsGDAP2, Homo sapiens Ganglioside-induced differentiation-associated protein 2 (GDAP2); TcPARG, Thermomonospora curvata PARG, Streptomyces coelicolor (SCO) ARH-like hydrolases indicated by their respective gene identifiers

Conclusion

This chapter describes the development of a novel and robust synthetic strategy to obtain peptides modified with mono ADP-ribose at a predetermined Ser, Thr or Cys side chain. In doing so, novel ribofuranosylated Ser-, Thr- and Cys-building blocks **28**, **29** and **30** were developed and used in a solid phase methodology to obtain immobilized peptides, decorated with an orthogonally protected ribosyl unit at the prospected ADPr site. The 5-OH of this ribose residue was then functionalized with a phosphomonoester prior to construction of the pyrophosphate, yielding the immobilized and partially protected ADP-ribosylated peptides. The peptides were then subjected to a deprotection sequence and cleavage from the resin. RP-HPLC purification of the crude products furnished peptides **39** – **47**, MARylated at a predetermined Ser, Thr or Cys residue. The novel strategy presented here was used to incorporate isotopically enriched amino acids, providing well-defined homogenous 'heavy' peptides **42** and **43** that are crucial for quantification studies in proteomic mixtures. Furthermore, the set of peptides synthesized varies in functional side chains flanking the ADPr site, demonstrating that a broad set of amino acids can be incorporated in the peptide sequence. Also, functionalization of the peptide with a biotin tag is allowed, provided that the oxidation of the immobilized phosphite intermediate is performed with *t*BuOOH rather than CSO.

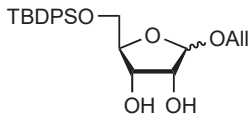
The availability of the Ser-ADPr, Thr-ADPr and Cys-ADPr peptides allowed the direct assessment and comparison of the sensitivity of these linkages towards enzymatic hydrolysis. It was found that the additional methyl group in Thr, as compared to Ser, leads to a pronounced reduction in turn-over by ARH3. This suggests that the additional methyl group hinders optimal substrate arrangement within the active site due to increased steric hindrance, but no other hydrolase was identified as being able to hydrolyze Thr-ADPr. In contrast, Cys-ADPr peptide **46**, having an *S*- instead of an *O*-glycosidic bond is almost completely resistant to ARH3-mediated hydrolysis. Given that Cys-modifications, which occur in human cells and are suggested to be involved in regulation of hypoxia, immunity, coronavirus response and nuclear receptors,[36–40] cannot be reversed by any of the known human hydrolases, it may be that the modification is either irreversible or is reversed by an as yet unidentified hydrolase or mechanism. The discovery of *Spy*MacroD as a Cys-(ADP-ribosyl)hydrolase shows that efficient and specific hydrolysis is possible, suggesting that a human enzyme harboring this activity may exist as well.

Experimental section

General synthetic procedures

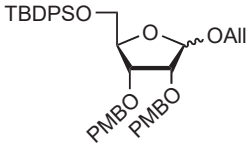
All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. Column chromatography was performed on silica gel 60 Å (40-63 µm, Macherey-Nagel). TLC analysis was performed on Macherey-Nagel aluminium sheets (silica gel 60 F_{254}). TLC was used to visualize compounds by UV at wavelength 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) or $KMnO_4$ spray (20g/L $KMnO_4$ and 10g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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. NMR spectra were recorded on a Bruker AV-400, AV-500 or AV-600 NMR. Chemical shifts (δ) are given in ppm relative to tetramethyl silane as internal standard. Coupling constants (*J*) are given in Hz. For compounds **39** – **47**, 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.

1-*O*-allyl-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside



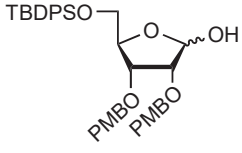
D-Ribose (11.3 g, 75 mmol) was suspended in allyl alcohol (190 mL, 0.4 M). Acetyl chloride (3.7 mL, 53 mmol, 0.7 eq.) was added and the suspension was stirred for 3 hours after which all starting material was dissolved. The reaction was quenched with pyridine and concentrated *in vacuo*. The residue was extensively co-evaporated with pyridine after which it was dissolved in pyridine (375 mL, 0.2 M) and TBDPS-Cl (21.5 mL, 82.5 mmol, 1.1 eq.) was added. The reaction was stirred overnight after which TLC indicated full conversion of starting material into a higher running product. The reaction was quenched with MeOH and concentrated *in vacuo*. The residue was taken up in DCM and washed with 1 M aq. HCl followed by sat. aq. $NaHCO_3$. The organic layer was dried over $MgSO_4$, filtrated and concentrated *in vacuo*. Purification by flash column chromatography yielded the title compound in a mixture of anomers (α:β ratio 1:3.8) as a clear oil (30.32 g, 70.74 mmol, 94%). Spectral data was in accordance with literary precedence.[23] **Rf**: 0.61 (β) and 0.71 (α) in 40% EtOAc in pentane **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.75 – 7.60 (m, 8H, TBDPS arom. α + β), 7.47 – 7.29 (m, 12H, TBDPS arom. α + β), 6.03 – 5.71 (m, 2H, CH_2CHCH_2 α + β), 5.40 – 5.06 (m, 2H, H-1 α, CH_2CHCH_2 α + β), 4.98 (s, 1H, H-1 β), 4.39 – 4.27 (m, 1H, H-3 β), 4.27 – 3.91 (m, 9H, H-2 α + β, H-3 α, H-4 α + β, CH_2CHCH_2 α + β), 3.86 – 3.62 (m, 4H, H-5 α + β), 2.86 (bs, 4H, OH α + β), 1.20 – 0.94 (m, 18H, *t*Bu TBDPS).**^{13}C NMR:** (101 MHz, $CDCl_3$) δ 135.7, 135.7, 135.7 (CH arom TBDPS. α/β), 134.1 (CH_2CHCH_2 β), 133.9 (CH_2CHCH_2 α), 133.4, 133.3 (Cq TBDPS β), 133.2, 133.1 (Cq TBDPS α) 129.9, 129.9, 129.9, 127.9, 127.9, 127.9, 127.9, 127.8 (CH arom. TBDPS α/β), 118.0 (CH_2CHCH_2 α), 117.4 (CH_2CHCH_2 β), 106.3 (C-1 β), 101.2 (C-1 α), 85.8 (C-4 α), 83.3 (C-4 β), 75.5 (C-2 β), 72.7 (C-3 β), 72.2 (C-2 α), 71.5 (C-3 α), 69.0 (CH_2CHCH_2 α), 68.5 (CH_2CHCH_2 β), 65.4 (C-5 β), 64.1 (C-5 α), 27.0, 26.9 (CH_3 *t*Bu α/β), 19.4, 19.3 (Cq *t*Bu α/β).

1-*O*-allyl-2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside



1-*O*-allyl-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside (30.3 g, 70.7 mmol) was co-evaporated with toluene, dissolved in DMF (350 mL, 0.2 M) and cooled to 0 °C. NaH (60 wt% in dispersion oil, 6.22 g, 155.6 mmol, 2.2 eq.) was added portion wise and stirred for 15 minutes after which TBABr (4,562 g, 14.15 mmol, 0.2 eq.) and PMB-Cl (21.2 mL, 155.6 mmol, 2.2 eq.) were added. The ice bath was removed and the reaction was stirred for 3 hours after which TLC indicated full conversion of the starting material into a higher running product (Rf 0.45 (β) and 0.75 (α) in 20% EtOAc in pentane). The reaction was cooled to 0 °C and carefully quenched with sat. aq. $NaHCO_3$. After the bubbling ceased, the reaction was transferred into a separatory funnel and partitioned with sat. aq. $NaHCO_3$ and Et_2O. The water layer was extracted twice with Et_2O and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography yielded the title compound (24.72 g for β, 6.21 g for α, 46.24 mmol combined yield, 65%). Spectral data was in accordance with literary precedence.[23] *α-anomer* **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.66 – 7.49 (m, 4H, TBDPS arom.), 7.46 – 7.31 (m, 6H, TBDPS arom.), 7.31 – 7.19 (m, 4H, PMB arom.), 6.88 – 6.77 (m, 4H, PMB arom.), 5.98 (dddd, *J* = 17.1, 10.4, 6.6, 4.8 Hz, 1H, CH_2CHCH_2), 5.34 (dd, *J* = 17.3, 1.7 Hz, 1H, CH_2CHCH_{2a}), 5.19 (dd, *J* = 10.4, 1.6 Hz, 1H, CH_2CHCH_{2b}), 5.03 (d, *J* = 4.3 Hz, 1H, H-1), 4.67 – 4.45 (m, 4H. 2x CH_2 PMB), 4.35 – 4.23 (m, 1H, $CH_{2a}CHCH_2$), 4.19 – 4.10 (m, 2H, H-4 + $CH_{2b}CHCH_2$), 3.95 (dd, *J* = 6.5, 3.0 Hz, 1H, H-3), 3.86 – 3.75 (m, 6H, 2x CH_3 PMB), 3.56 (ddd, *J* = 38.8, 11.2, 3.5 Hz, 2H, H-5), 0.95 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 159.4, 159.2 (Cq PMB), 135.7, 135.7 (CH arom. TBDPS), 135.0 (CH_2CHCH_2), 133.4, 133.2 (Cq TBDPS), 130.6, 130.1 (Cq PMB), 129.8, 129.8, 129.7, 129.7 (CH arom. TBDPS/PMB), 127.8, 127.8 (CH arom. PMB), 117.3 (CH_2CHCH_2), 113.9, 113.7 (CH arom. PMB), 100.2 (C-1), 83.5 (C-4), 77.7 (C-2), 74.8 (C-3), 72.1, 72.0 (CH_2 PMB), 68.7 (CH_2CHCH_2), 64.2 (C-5), 55.3 (2x CH_3 PMB), 26.9 (CH_3 *t*Bu), 19.3 (Cq *t*Bu). *β-anomer* **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.72 – 7.63 (m, 4H, TBDPS arom.), 7.46 – 7.32 (m, 6H, TBDPS arom.), 7.32 – 7.25 (m, 2H, PMB arom.), 7.21 – 7.17 (m, 2H, PMB arom.), 6.91 – 6.78 (m, 4H, PMB arom.), 5.81 (dddd, *J* = 16.8, 10.3, 6.2, 5.1 Hz, 1H, CH_2CHCH_2), 5.23 – 5.11 (m, 2H, CH_2CHCH_2), 5.06 (d, *J* = 1.2 Hz, 1H, H-1), 4.66 – 4.50 (m, 2H, CH_2 PMB), 4.51 – 4.34 (m, 2H, CH_2 PMB), 4.24 (dt, *J* = 6.9, 4.2 Hz, 1H, H-4), 4.20 – 4.10 (m, 2H, H-3 + $CH_{2a}CHCH_2$), 4.00 – 3.90 (m, 1H, $CH_{2b}CHCH_2$), 3.90 – 3.85 (m, 1H, H-2), 3.85 – 3.75 (m, 7H, 2x CH_3 PMB + H-5a), 3.70 (dd, *J* = 11.1, 4.6 Hz, 1H, H-5b), 1.03 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 159.4, 159.3 (Cq PMB), 135.7 (CH arom. TBDPS), 134.3 ($CHCHCH_2$), 133.6, 133.5 (Cq TBDPS), 130.2, 130.1 (Cq PMB), 129.8, 129.7, 129.5 (CH arom. TBDPS/PMB), 127.8, 127.8 (CH arom. PMB), 117.3 (CH_2CHCH_2), 113.9, 113.8 (CH arom. PMB), 104.5 (C-1), 82.2 (C-4), 79.7 (C-2), 77.5 (C-3), 72.1 (CH_2 PMB), 72.0 (CH_2 PMB), 68.5 (CH_2CHCH_2), 64.6 (C-5), 55.4, 55.4 (CH_3 PMB), 26.9 (CH_3 *t*Bu), 19.4 (Cq).

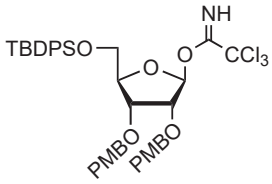
2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside (13)



1-*O*-allyl-2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside (20.5 g, 30.6 mmol) was dissolved in a 3:1 $CHCl_3$:H_2O mixture (167 mL, 0.18 M). $PdCl_2$ (852 mg, 4.81 mmol, 0.16 eq.) was added and the reaction was vigorously stirred at 45 °C under an O_2 atmosphere for 72 hours after which TLC indicated full conversion into the lactol (Rf = 0.23 in 20% EtOAc in pentane) plus the corresponding isomerized allyl ether. The reaction was filtered over a pad of Celite and the pad was thoroughly flushed with EtOAc. Iodine (8,122 mg, 32 mmol, 1.05 eq.) was added to the filtrate and the resulting mixture was stirred for 5 minutes, transferred into a separatory funnel and washed twice with sat. aq. $Na_2S_2O_3$ and once with brine. The organic layer was dried over $MgSO_4$, filtrated and concentrated *in vacuo*. Flash column chromatography (5 -> 20% EtOAc in pentane) yielded the title compound, as a pale yellow oil (16.2 g, 25.8 mmol, 84%). Spectral data was in accordance with literary precedence.[23] **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.67 – 7.56 (m, 8H, TBDPS

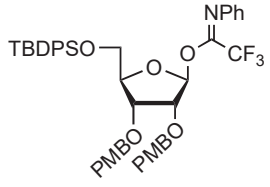
arom. α+β), 7.49 – 7.33 (m, 12H, TBDPS arom. α+β), 7.33 – 7.16 (m, 12H, PMB arom. α+β), 6.90 – 6.80 (m, 8H, PMB arom. α+β), 5.31 – 5.24 (m, 2H, H-1 α+β), 4.68 – 4.43 (m, 7H, 2x Ch_2 PMB α+β), 4.32 – 4.24 (m, 3H, H-3β + H-4α + CH_2 PMB), 4.21 (dt, *J* = 6.8, 2.6 Hz, 1H, H-4β), 4.05 (dd, *J* = 4.9, 1.8 Hz, 1H, H-3α), 4.00 – 3.93 (m, 1H, H-2α), 3.87 – 3.71 (m, 13H, H-2β + 2x CH_3 PMB α+β), 3.66 – 3.55 (m, 4H, H-5 α+β), 1.06 – 0.96 (m, 18H, *t*Bu TBDPS α+β). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 159.5, 159.5 (Cq PMB), 135.8, 135.7, 135.7, 135.7, 135.6 (CH arom. TBDPS), 133.2, 133.0, 132.7 (Cq TBDPS), 130.1, 130.0, 130.0, 129.9, 129.9 (CH arom. TBDPS/PMB), 129.8, 129.7 (Cq PMB), 129.6, 129.6, 129.6 (CH arom. TBDPS/PMB), 127.9, 127.9, 127.9, 127.8 (CH arom. TBDPS), 114.0, 113.9, 113.9, 113.7 (CH arom. PMB), 100.2 (C-β), 96.4 (C-1α), 82.6 (C-4α), 82.3 (C-4β), 80.2 (C-2β), 77.8 (C-2α), 77.5 (C-3α), 76.1 (C-3β), 72.6 (CH_2 PMB), 72.1 (CH_2 PMB), 72.1 (CH_2 PMB), 72.0 (CH_2 PMB), 64.1 (C-5α), 63.3 (C-5β), 55.4, 55.4 (CH_3 PMB), 26.9, 26.9 (CH_3 *t*Bu), 19.3 (Cq *t*Bu).

1-(2,2,2-trichlororoacetimido)-2,3-*O*-di-(4-methoxybenzyl)-5-O-*tert*-butyldiphenylsilyl -α,β-D-ribofuranoside (14)



Compound **13** (1.01 g, 1.6 mmol, 1 eq.) was co-evaporated thrice with toluene, dissolved in dry DCM (16 mL, 0.1 M) and the solution was cooled to 0 °C. DBU (0.11 mL, 0.74 mmol, 0.5 eq.) and trichloroacetonitrile (0.8 mL, 8.0 mmol, 5 eq.) were added and the solution was stirred at 0 °C for 1 hour after which the reaction was concentrated *in vacuo*. Purification by flash column chromatography in neutralized silica (5 -> 20% EtOAc in pentane with 1% TEA) yielded the title compound as a pale oil (1.11 g, 1.44 mmol, 90%). **Rf**: 0.85 (20% EtOAc in pentane). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 8.48 (s, 1H, NH acetimidate), 7.69 – 7.54 (m, 6H, H arom.), 7.46 – 7.33 (m, 6H, H arom.), 7.36 – 7.25 (m, 7H, H arom.), 7.18 (s, 1H, H arom.), 6.84 (dd, *J* = 25.9, 8.7 Hz, 3H, H arom.), 6.31 (s, 1H, H-1), 4.80 – 4.64 (m, 1H, CH_2 PMB), 4.60 (d, *J* = 11.7 Hz, 1H, CH_2 PMB), 4.46 (d, *J* = 11.3 Hz, 1H, CH_2 PMB), 4.42 – 4.34 (m, 2H, H-4 + CH_2 PMB), 4.17 – 4.11 (m, 1H, H-3), 4.06 (d, *J* = 4.8 Hz, 1H, H-2), 3.86 – 3.79 (m, 4H, CH_3 PMB + H-5), 3.79 – 3.76 (m, 4H, CH_3 PMB + H-5), 1.02 (s, 9H TBDPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 161.1, 159.4, 135.7, 133.4, 130.0, 129.6, 127.8, 113.9, 103.9 (C-1), 83.7 (C-4), 78.6 (C-2), 76.6 (C-3), 72.1, 71.8 (CH_2 PMB), 64.2 (H-5), 55.4, (CH_3 PMB), 27.0 (CH_3 TBDPS), 19.4 (Cq TBDPS).

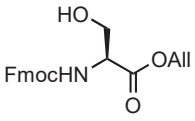
1-*O*-((*N*-phenyl)-2,2,2-trifluoroacetimido) 2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside (15)



Compound **13** (6.28 g, 10.0 mmol) was dissolved in acetone (50 mL, 0.2 M). Cs_2CO_3 (4.89 g, 15.0 mmol, 1.5 eq.) and 2,2,2-trifluoro-*N*-phenylacetimidoyl chloride (1.74 mL, 11.0 mmol, 1.1 eq.) were added. The suspension was stirred for 2 hours before TLC indicated full conversion into a higher running product. The reaction was filtered over a pad of Celite and the filtrate was concentrated *in vacuo*. Flash column chromatography in neutralized silica (10 -> 20% Et_2O in pentane) yielded the title compound as a pale oil (6.01 g, 7.51 mmol, 75%). Spectral data was in accordance with literary precedence.[23] **Rf:** 0.20 (10% Et_2O in pentane). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.74 – 7.66 (m, 4H, TBDPS arom. α+β), 7.66 – 7.50 (m, 4H, NPh arom. α+β), 7.46 – 7.32 (m, 12H, TBDPS arom. + NPh arom. α+β), 7.32 – 7.17 (m, 12H, PMB arom. α + β), 7.12 – 7.00 (m, 2H, PMB arom. α), 6.91 – 6.72 (m, 12H, PMB arom. α + β), 6.48 (bs, 1H, H-1α), 6.30 (bs, 1H, H-2β), 4.82 – 4.57 (m, 4H, 2x CH_2 PMB α/β), 4.57 – 4.40 (m, 4H, 2x CH_2 PMB α/β), 4.40 – 4.31 (m, 2H, H-4 α+β), 4.26 (t, *J* = 5.5 Hz, 1H, H-3β), 4.19 – 4.01 (m, 3H, H-2α + H-2β + H-3α), 4.00 – 3.86 (m, 1H, H-5_aβ), 3.84 – 3.56 (m, 15H, H-5_bβ + H-5α + 2x CH_3 PMB α+β), 1.07 (s, 9H, *t*Bu TBDPS β), 0.97 (s, 9H, *t*Bu TBDPS α).**^{13}C NMR:** (101 MHz, $CDCl_3$) δ 159.5, 159.4, 159.4, 159.2 (Cq PMB),

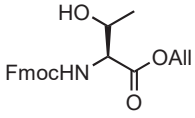
144.3, 143.9 (Cq NPh), 135.7, 135.6, 135.6, 135.5(CH arom. TBDPS/NPh), 133.3, 133.1, 133.0, 132.7, 130.4 (Cq TBDPS/PMB), 129.9, 129.9, 129.7, 129.7 (CH arom.), 129.6 (Cq), 129.5 (CH arom.), 129.4 (Cq), 129.4, 129.3, 129.2, 128.7, 128.7, 127.8, 127.8, 127.7 (CH arom.), 124.2, 120.6, 119.6, 114.3, 113.9, 113.8, 113.8, 113.7 (CH arom. PMB), 102.8 (C-1β), 85.7 (C-4α), 83.4 (C-4β), 78.7 (C-2α), 78.4 (C-2β), 76.0 (C-3β), 75.2 (C-3α), 72.9, 72.3, 72.2, 71.9 (CH_2PMB), 63.7 (C-5β), 63.3 (C-5α), 55.5, 55.2, 55.1, 55.1 (CH_3 PMB), 26.8, 26.7 (CH_3 *t*Bu), 19.2, 19.2 (Cq *t*Bu).

***N*-fluorenylmethoxycarbonyl-serine-*O*-allyl ester (19)**



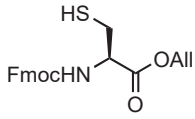
Fmoc-Ser-OH (3.45 g, 10 mmol, 1 eq.) was co-evaporated thrice with 1,4-dioxane and dissolved DMF (50 mL, 0.2 M). DIPEA (2.10 mL, 12 mmol, 1.2 eq.) was added followed by the addition of allyl-bromide (1.04 mL, 12 mmol, 1.2 eq.) and stirred overnight. The reaction was quenched with 100 mL water and transferred into a separatory funnel. The reaction mixture was extracted three times with Et_2O. The combined organic layers were washed three times with brine, dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (10 -> 40% EtOAc in pentane) afforded the title compound as a white solid (3.58 g, 9.74 mmol, 97%). **Rf:** 0.49 (40% EtOAc in pentane). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.5 Hz, 2H, Fmoc arom.), 7.61 (d, *J* = 6.8 Hz, 2H, Fmoc arom.), 7.41 (t, *J* = 7.4 Hz, 2H, Fmoc arom.), 7.32 (t, *J* = 7.5 Hz, 2H, Fmoc arom.), 5.99 – 5.84 (m, 1H, OCH_2CHCH_2), 5.73 (d, *J* = 7.2 Hz, 1H, NH), 5.35 (d, *J* = 17.1 Hz, 1H, CH_2 OCH_2CHCH_{2a}), 5.27 (d, *J* = 10.4 Hz, 1H, OCH_2CHCH_{2b}), 4.69 (d, *J* = 5.5 Hz, 2H, OCH_2CHCH_2), 4.51 – 4.40 (m, 3H, CH Ser + CH_2 Ser), 4.23 (t, *J* = 6.9 Hz, 1H, CH Fmoc), 4.04 (d, *J* = 10.7 Hz, 1H, CH_2 Fmoc), 3.95 (d, *J* = 10.6 Hz, 1H, CH_2 Fmoc), 2.15 (s, 1H, OH Ser). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 170.4 (C=O COOAll), 156.4 (C=O Fmoc), 143.9, 143.8, 141.4 (Cq Fmoc), 131.4 (OCH_2CHCH_2), 127.9, 127.2, 125.2, 120.1 (CH Fmoc arom.), 119.1 (OCH_2CHCH_2), 67.3 (CH_2 Fmoc), 66.5 (OCH_2CHCH_2), 63.3 (CH_2 Ser), 56.2 (CH Ser), 47.2 (CH Fmoc).

***N*-fluorenylmethoxycarbonyl-threonine-*O*-allyl ester (22)**



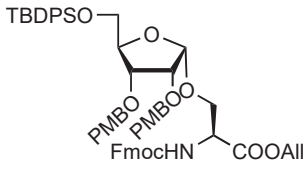
Fmoc-Cys(Trt)-OH (1.98 g, 5 mmol) was co-evaporated thrice with 1,4-dioxane and dissolved in DMF (25 mL, 0.2 M). DIPEA (1.04 mL, 6.0 mmol, 1.2 eq.) was added followed by allyl bromide (0.52 mL, 6.0 mmol, 1.2 eq.). The resulting solution was stirred overnight. The reaction was carefully quenched by the addition of water and the resulting slurry was transferred into a separatory funnel. The water layer was extracted thrice with Et_2O and the combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude allylated cysteine was dissolved in 25 mL of a 10% TFA solution in DCM and TIS (4.1 mL, 20 mmol, 4.0 eq.) was added. The resulting solution was stirred for 3 hours after which TLC showed full conversion of the starting material. The reaction mixture was diluted with toluene and concentrated *in vacuo*. The resulting oil was co-evaporated thrice with toluene. Flash column chromatography (10 -> 40% EtOAc in pentane) yielded the title compound as a white solid (1.58 g, 4.15 mmol, 83%). **Rf:** 0.65 (40% EtOAc in pentane). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.77 (dt, *J* = 7.6, 0.9 Hz, 2H, Fmoc arom.), 7.62 (dd, *J* = 7.7, 4.2 Hz, 2H, Fmoc arom.), 7.46 – 7.38 (m, 2H, Fmoc arom.), 7.33 (dt, *J* = 7.5, 1.5 Hz, 2H, Fmoc arom.), 5.92 (ddt, *J* = 16.4, 11.0, 5.8 Hz, 1H, OCH_2CHCH_2), 5.60 (d, *J* = 9.1 Hz, 1H, NH), 5.35 (dd, *J* = 16.2, 1.7 Hz, 1H, OCH_2CHCH_{2a}), 5.27 (dd, *J* = 1.1 Hz, 1H, OCH_2CHCH_{2b}), 4.69 (d, *J* = 5.7 Hz, 2H, OCH_2CHCH_2), 4.49 – 4.34 (m, 4H, CH_2 Fmoc + 2x CH Thr), 4.25 (t, *J* = 7.1 Hz, 1H, CH Fmoc), 1.77 (s, 1H, OH), 1.27 (d, *J* = 6.3 Hz, 3H, CH_3 Thr). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 144.0, 141. 5, 131.5 (OCH_2CHCH_2), 127.9, 127.2, 125.3, 120.1, 119.2 (OCH_2CHCH_2), 68.2 (CH Thr), 67.4 (CH_2 Fmoc), 66.4 (OCH_2CHCH_2), 59.2 (CH Thr), 47.3 (CH Fmoc), 20.1 (CH_3 Thr).

_N_-fluorenylmethoxycarbonyl-cysteine-_O_-allyl ester (23)



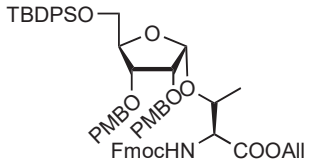
Fmoc-Cys(Trt)-OH (2.23 g, 3.81 mmol) was co-evaporated thrice with toluene and dissolved in DMF (19 mL, 0.2 M). DIPEA (0.80 mL, 4.6 mmol, 1.2 eq.) was added followed by allyl bromide (0.40 mL, 4.6 mmol, 1.2 eq.). The resulting solution was stirred overnight. The reaction was carefully quenched by the addition of water and the resulting slurry was transferred into a separatory funnel. The water layer was extracted thrice with Et_2O and the combined organic layers were washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. The crude allylated cysteine was dissolved in 19 mL of a 10% TFA solution in DCM and TIS (3.1 mL, 15.2 mmol, 4.0 eq.) was added. The resulting solution was stirred for 4 hours in a dark environment after which TLC showed full conversion of the starting material. The reaction was diluted with toluene and concentrated *in vacuo*. The resulting oil was co-evaporated thrice with toluene. Flash column chromatography (5 -> 10% EtOAc in toluene) yielded the title compound as a white solid (875 mg, 2.28 mmol, 60%). **Rf:** 0.53 (10% EtOAc in toluene). **^{1}H NMR:** (400 MHz, CDCl3) δ 7.76 (dd, J = 7.6, 1.1 Hz, 2H, Fmoc arom.), 7.60 (d, J = 7.5 Hz, 2H, Fmoc arom.), 7.40 (t, J = 8.3, 6.8 Hz, 2H, Fmoc arom.), 7.32 (tt, J = 7.4, 1.5 Hz, 2H, Fmoc arom.), 5.92 (ddt, J = 16.5, 10.4, 5.8 Hz, 1H, OCH_2CHCH_2), 5.72 (bs, 1H, NH), 5.43 – 5.18 (m, 2H, OCH_2CHCH_2), 4.76 – 4.61 (m, 3H, CH Cys + OCH_2CHCH_2), 4.51 – 4.32 (m, 2H, CH_2 Fmoc), 4.23 (t, J = 6.9 Hz, 1H, CH Fmoc), 3.01 (dqt, J = 14.0, 8.8, 4.2 Hz, 2H, CH_2 Cys), 1.37 (t, J = 9.0 Hz, 1H, SH). **^{13}C NMR:** (101 MHz, CDCl3) δ 169.8 (C=O COOAll), 155.8 (C=O Fmoc), 143.9, 143.7, 141.4, 141.4 (Cq Fmoc), 131.4 (OCH_2CHCH_2), 127.9, 127.2, 127.2, 125.2, 125.2, 120.2, 120.1 (CH Fmoc arom.), 119.5 (OCH_2CHCH_2), 67.2 (CH_2 Fmoc), 66.6 (OCH_2CHCH_2), 55.3 (CH Cys), 47.3 (CH Fmoc), 27.3 (CH_2 Cys).

1-_O_-(2,3-bis-_O_-(4-methoxybenzyl)-5-_O_-((_tert_-butyl)-diphenylsilyl)-α-D-ribosyl)-_N_-fluorenylmethoxycarbonyl serine allyl ester (25)



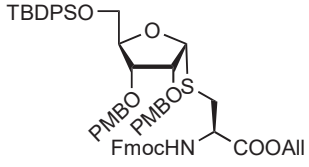
Compound **14** (1.60 g, 2.00 mmol, 1.1 eq.) and Fmoc-Ser-OAll **19** (669 mg, 1.82 mmol, 1.0 eq. relative to the donor) were co-evaporated thrice with toluene and dissolved in DCM (20 mL, 0.1 M relative to the donor). The reaction was cooled to -50 °C and TBSOTf (46 µL, 0.2 mmol, 0.1 eq. relative to the donor) was added. The reaction was stirred at -50 °C for 4.5 hours before TLC analysis showed near full conversion of the starting material into a higher running product. The reaction was quenched with TEA and concentrated *in vacuo*. Flash column chromatography (0.5 -> 3% acetone in DCM) yielded the title compound as a clear oil (994 mg, 1.02 mmol, 56%). **Rf:** 0.60 (3% acetone in DCM). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.72 – 7.65 (m, 2H, Fmoc arom.), 7.65 – 7.49 (m, 6H, Fmoc arom. + TBDPS arom.), 7.46 – 7.10 (m, 14H, Fmoc arom. + TBDPS arom. + PMB arom. + $CHCl_3$ overlap), 6.89 – 6.74 (m, 4H, PMB arom.), 6.61 (d, *J* = 9.0 Hz, 1H, NH), 5.88 (ddt, *J* = 16.1, 10.8, 5.5 Hz, 1H, $CHCH_2$), 5.31 (dd, *J* = 17.2, 1.6 Hz, 1H, CH_2CHCH_{2a}), 5.16 (dd, *J* = 10.5, 1.5 Hz, 1H, CH_2CHCH_{2b}), 4.98 (d, *J* = 4.1 Hz, 1H, H-1), 4.73 – 4.46 (m, 7H, CH Ser + CH_2CHCH_2 + 2x CH_2 PMB), 4.39 (dd, *J* = 10.5, 6.9 Hz, 1H, CH_{2a} Fmoc), 4.28 – 4.15 (m, 3H, H-4 + CH_{2b} Fmoc + CH_{2a} Ser), 4.12 (t, *J* = 7.3 Hz, 1H, CH Fmoc), 4.03 (dd, *J* = 6.1, 2.3 Hz, 1H, H-3), 3.97 (dd, *J* = 10.3, 3.3 Hz, 1H, CH_{2b} Ser), 3.89 (dd, *J* = 6.1, 4.1 Hz, 1H, H-2), 3.74 (s, 3H, CH_3 PMB), 3.69 (s, 3H, CH_3 PMB), 3.61 (ddd, *J* = 38.2, 11.3, 3.2 Hz, 2H, H-5), 0.97 (s, 9H, *t*Bu TBDPS).**^{13}C NMR:** (101 MHz, $CDCl_3$) δ 170.3 (C=O COOAll), 159.4, 159.1 (Cq PMB), 156.5 (C=O Fmoc), 144.1, 143.8, 141.2, 141.1 (Cq Fmoc), 135.6, 135.5 (CH arom. TBDPS), 133.1, 132.9 (Cq TBDPS), 131.8 (CH_2CHCH_2), 130.3 (Cq PMB), 129.9 (CH arom.), 129.8 (Cq PMB), 129.8, 129.7, 129.5, 127.8, 127.8, 127.6, 127.5, 127.0, 127.0, 125.4, 125.1 (CH arom.), 119.9, 119.8 (CH arom. Fmoc), 118.3 (CH_2CHCH_2), 113.8, 113.7 (CH arom. PMB), 101.0 (C-1), 84.1 (C-4), 78.6 (C-2), 75.2 (C-3), 72.3, 72.1 (CH_2 PMB), 67.3 (CH_2 Ser), 67.0 (CH_2 Fmoc), 66.0 (CH_2CHCH_2), 64.0 (C-5), 55.2, 55.1 (CH_3PMB), 54.5 (CH Ser), 47.1 (CH Fmoc), 26.8 (CH_3 *t*Bu), 19.2 (Cq *t*Bu). **HRMS:** $[C_{58}H_{63}NO_{11}Si + Na]^+$ found: 1000.4053, calculated: 1000.4063

1-*O*-(2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl threonine allyl ester (26)



Compound **14** (1.22 g, 1.53 mmol, 1.3 eq.) and Fmoc-Thr-OAll **22** (450 mg, 1.18 mmol, 1.0 eq. relative to the donor) were co-evaporated thrice with toluene and dissolved in DCM (15 mL, 0.1 M relative to the donor). The reaction was cooled to -50 °C and TBSOTf (35 µL, 0.15 mmol, 0.1 eq. relative to the donor) was added. The reaction was stirred at -50 °C for 2 hours before TLC analysis showed near full conversion of the starting material into a higher running product. The reaction was quenched with TEA and concentrated *in vacuo*. Flash column chromatography (0.5 -> 3% acetone in DCM) yielded the title compound as a clear oil (742 mg, 0.748 mmol, 63%). **Rf:** 0.57 (3% acetone in DCM). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.70 (d, *J* = 7.5 Hz, 2H, Fmoc arom.), 7.67 – 7.52 (m, 6H, Fmoc arom. + TBDPS arom.), 7.46 – 7.28 (m, 10H, Fmoc arom. + TBDPS arom.), 7.28 – 7.15 (m, 4H, PMB arom.), 6.91 – 6.75 (m, 4H, PMB arom.), 6.62 (d, *J* = 8.7 Hz, 1H, NH), 5.88 (ddt, *J* = 17.3, 10.5, 5.7 Hz, 1H, OCH_2CHCH_2), 5.37 – 5.26 (m, 1H, OCH_2CHCH_{2a}), 5.21 – 5.15 (m, 1H, OCH_2CHCH_{2b}), 5.13 (d, *J* = 4.2 Hz, 1H, H-1), 4.71 – 4.55 (m, 4H, OCH_2CHCH_2 + CH_2 PMB) , 4.55 – 4.36 (m, 5H CH_2 PMB + CH_{2a} Fmoc + 2x CH Thr), 4.30 (dd, *J* = 10.3, 7.6 Hz, 1H, CH_{2b} Fmoc), 4.23 (t, *J* = 7.3 Hz, 1H, CH Fmoc), 4.18 (q, *J* = 3.3 Hz, 1H, H-4), 3.95 (dd, *J* = 6.6, 3.0 Hz, 1H, H-3), 3.82 (dd, *J* = 6.6, 4.2 Hz, 1H, H-2), 3.76 – 3.61 (m, 7H, 2x CH_3 PMB + H-5_a), 3.55 (dd, *J* = 11.3, 3.4 Hz, 1H, H-5_b), 1.40 (d, *J* = 6.3 Hz, 3H, CH_3 Thr), 0.97 (s, 9H, *t*Bu TBDPS).**^{13}C NMR:** (101 MHz, $CDCl_3$) δ 170.7 (C=O, COOAll), 159.3, 159.1 (Cq PMB), 157.0 (C=O Fmoc), 144.0, 143.8, 141.2, 141.1 (Cq Fmoc), 135.5, 135.5 (CH arom. TBDPS), 133.1, 132.9 (Cq TBDPS), 131.6 (OCH_2CHCH_2), 130.1 (Cq PMB), 129.8, 129.7 (CH arom.), 129.6 (Cq PMB), 129.5, 129.4, 127.7, 127.7, 127.6, 127.5, 127.0, 127.0, 125.3, 125.2 (CH arom.), 119.8 (CH arom. Fmoc), 118.7 (OCH_2CHCH_2), 113.8, 113.7 (CH arom. PMB), 101.4 (C-1), 83.6 (C-4), 77.8 (C-2), 74.8 (C-3), 74.4 (RO-*C*H(CH_3)CH-R Thr), 72.0, 71.8 (CH_2 PMB), 67.1 (CH_2 Fmoc), 65.9 (OCH_2CHCH_2), 63.9 (C-5), 59.4 (RO-CH(CH_3)*C*H-R Thr), 55.1, 55.0 (CH_3 PMB), 47.1 (CH Fmoc), 26.7 (CH_3 *t*Bu), 19.1 (Cq *t*Bu), 19.1 (CH_3 Thr). **HRMS:** $[C_{59}H_{65}NO_{11}Si + H]^+$ found: 992.4402, calculated: 992.4400.

1-*O*-(2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl cysteine allyl ester (27)



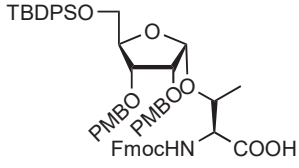
Compound **14** (1.60 g, 2.00 mmol, 1.3 eq.) and Fmoc-Cys-OAll **23** (595 mg, 1.55 mmol, 1.0 eq. relative to the donor) were co-evaporated thrice with toluene and dissolved in DCM (20 mL, 0.1 M relative to the donor). The reaction was cooled to -50 °C and TBSOTf (46 µL, 0.2 mmol, 0.1 eq. relative to the donor) was added. The reaction was stirred at -50 °C for 1.5 hours before TLC analysis showed near full conversion of the starting material into a higher running product. The reaction was quenched with TEA and concentrated *in vacuo*. Flash column chromatography (0.5 -> 1.5% acetone in DCM) yielded the title compound as a clear oil (1.08 g, 1.08 mmol, 70%). **Rf:** 0.65 (1.5% acetone in DCM). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.73 – 7.62 (m, 2H, Fmoc arom.), 7.62 – 7.43 (m, 6H, Fmoc arom. + TBDPS arom.), 7.42 – 7.06 (m, 14H, Fmoc arom. + TBDPS arom. + PMB arom.), 6.90 – 6.77 (m, 4H, PMB arom.), 6.67 (d, *J* = 8.8 Hz, 1H, NH), 5.87 (ddt, *J* = 16.3, 10.8, 5.7 Hz, 1H, OCH_2CHCH_2), 5.39 (d, *J* = 5.5 Hz, 1H, H-1), 5.28 (dq, *J* = 17.2, 1.5 Hz, 1H, OCH_2CHCH_{2a}), 5.15 (dd, *J* = 10.4, 1.4 Hz, 1H, OCH_2CHCH_{2b}), 4.77 (dt, *J* = 8.9, 4.4 Hz, 1H, CH Cys), 4.72 – 4.53 (m, 5H, OCH_2CHCH_2 + 1x CH_2 PMB + 1x CH_{2a} PMB), 4.47 (d, *J* = 11.8 Hz, 1H, CH_{2b} PMB), 4.42 – 4.26 (m, 2H, H-4 + CH_{2a} Fmoc), 4.21 – 3.97 (m, 4H, H-2 + H-3 + CH_{2b} Fmoc + CH fmoc), 3.77 (dd, *J* = 11.4, 3.3 Hz, 1H, H-5_a), 3.74 – 3.70 (m, 6H, 2x CH_3 PMB), 3.66 (dd, *J* = 11.3, 2.9 Hz, 1H, H-5_b), 3.44 (dd, *J* = 14.6, 5.0 Hz, 1H, CH_{2a} Cys), 2.95 (dd, *J* = 14.6, 3.9 Hz, 1H, CH_{2b} Cys), 0.96 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 170.2

(C=O COOAll, 159.3, 159.1 (Cq PMB), 156.1 (C=O Fmoc), 143.9, 143.7, 141.1, 141.0 (Cq Fmoc), 135.4, 135.4 (CH arom. TBDPS), 132.9, 132.8 (Cq TBDPS), 131.6 (OCH_2CHCH_2), 130.0 (Cq PMB), 129.7, 129.6, 129.5 (CH arom.), 129.4 (Cq PMB), 129.2, 127.7, 127.6, 127.6, 127.5, 127.5, 126.9, 126.9, 125.2, 125.0 (CH arom.), 119.7 (CH arom. Fmoc) 118.5 (OCH_2CHCH_2), 113.7, 113.6 (CH arom. PMB), 89.1 (C-1), 82.6 (C-4), 78.0 (C-2), 76.0 (C-3), 72.5, 72.0 (CH_2 PMB), 66.9 (CH_2 Fmoc), 65.9 (OCH_2CHCH_2), 63.3 (C-5), 55.1 (CH_3 PMB), 54.2 (CH Cys), 46.9 (CH Fmoc), 34.8 (CH_2 Cys), 26.7 (CH_3 *t*Bu), 19.1 (Cq *t*Bu). **HRMS:** $[C_{58}H_{63}NO_{10}SSi + H]^+$ found: 994.4017, calculated: 994.4015.

1-*O*-(2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine (28)

Compound **25** (836 mg, 0.855 mmol, 1.0 eq.) was dissolved in DCM (8.5 mL, 0.1 M). DMBA (265 mg, 1.70 mmol, 2.0 eq.) and $Pd(PPh_3)_4$ (10 mg, 8.5 µmol, 0.01 eq.) were added and the reaction was stirred for 1 hour before TLC showed full conversion of the starting material into a lower running product. The reaction was diluted with DCM, washed with 1 M HCl and the organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (4% MeOH in DCM + 0.1% AcOH) yielded the title compound as a white foam (724 mg, 0.772 mmol, 90%). **Rf:** 0.38 (5% MeOH in DCM + 0.1% AcOH). **$[\alpha]_D^{20}$:** +50.8° (c = 1.0, $CHCl_3$). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.70 (d, *J* = 7.6 Hz, 2H, Fmoc arom.), 7.61 – 7.49 (m, 6H, Fmoc arom. + TBDPS arom.), 7.49 – 7.13 (m, 14H, Fmoc arom. + TBDPS arom. + PMB arom.), 6.86 – 6.72 (m, 4H, PMB arom.), 6.27 – 6.18 (m, 1H, NH), 4.93 (d, *J* = 4.1 Hz, 1H, H-1), 4.62 – 4.44 (m, 5H, CH Ser + 2x CH_2 PMB), 4.42 – 4.26 (m, 2H, CH_2 Fmoc), 4.20 – 4.09 (m, 2H, H-4 + CH Fmoc), 4.06 – 3.99 (m, 1H, CH_{2a} Ser), 3.97 (dd, *J* = 6.1, 1.6 Hz, 1H, H-3), 3.94 – 3.85 (m, 2H, CH_{2b} Ser + H-2), 3.79 – 3.69 (m, 6H, 2x CH_3 PMB), 3.56 (dd, *J* = 11.3, 3.5 Hz, 1H, H-5_a), 3.44 (dd, *J* = 11.3, 2.8 Hz, 1H, H-5_b), 0.95 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 172.0 (C=O COOH), 159.6, 159.4 (Cq PMB), 156.4 (C=O Fmoc), 141.4, 141.3 (Cq Fmoc), 135.7, 135.6 (CH arom. TBDPS), 133.1, 132.8 (Cq TBDPS), 130.1, 130.0, 130.0, 129.9 (CH arom.), 129.7, 129.2 (Cq PMB), 127.9, 127.9, 127.8, 127.2, 125.3, 125.2 (CH arom.), 120.0 (CH arom. Fmoc), 114.1, 113.9 (CH arom. PMB), 101.5 (C-1), 84.6 (C-4), 78.2 (C-2), 74.9 (C-3), 72.3 (CH_2 PMB), 67.2 (CH_2 Fmoc), 66.7 (CH_2 Ser), 64.1 (C-5), 55.3 (CH_3 PMB), 53.6 (CH Ser), 47.2 (CH Fmoc), 26.9 (CH_3 *t*Bu), 19.3 (Cq *t*Bu). **HRMS:** $[C_{55}H_{59}NO_{11}Si + Na]^+$ found: 960.3756, calculated: 960.3750.

1-*O*-(2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl threonine (29)



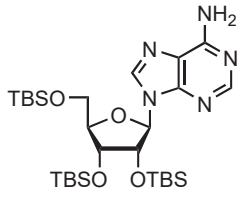
Compound **26** (525 mg, 0.529 mmol, 1.0 eq.) was dissolved in DCM (5.3 mL, 0.1 M). DMBA (372 mg, 2.38 mmol, 4.5 eq.) and a catalytic amount of $Pd(PPh_3)_4$ were added and the reaction was stirred for 15 minutes before TLC showed full conversion of the starting material into a lower running product. The reaction was diluted with DCM, washed with sat. aq. $NaHCO_3$. The water layer was extracted six times with DCM. The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. After purification by flash column chromatography (1 -> 10% MeOH in DCM) all fractions containing the title compound were combined and concentrated *in vacuo*. The residue was taken up in DCM and washed with 10% aq. citric acid. The water layer was extracted twice with DCM and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo* yielding the title compound as an orange foam (495 mg, 0.520 mmol, 98%). **Rf:** 0.05 (25% EtOAC in pentane + 0.1% AcOH). **$[\alpha]_D^{20}$:** +43.2° (c = 1.0, $CHCl_3$). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.74 (d, *J* = 7.2 Hz, 2H,

Fmoc arom.), 7.64 – 7.55 (m, 6H, Fmoc arom. + TBDPS arom.), 7.49 – 7.32 (m, 10H, Fmoc arom. + TBDPS arom.), 7.28 (tt, *J* = 7.5, 1.4 Hz, 4H, Fmoc arom. + TBDPS arom.), 7.22 (t, *J* = 8.3 Hz, 4H, PMB arom.), 6.86 – 6.74 (m, 4H, PMB arom.), 6.13 (d, *J* = 6.7 Hz, 1H, NH), 5.20 (d, *J* = 4.1 Hz, 1H, H-1), 4.58 – 4.47(m, 5H, 2x CH_2 PMB + CH Thr), 4.42 – 4.33 (m, 2H, CH_2 Fmoc), 4.26 (dd, *J* = 6.6, 4.4 Hz, 1H, CH Thr), 4.24 – 4.16 (m, 2H, CH Fmoc + H-4), 3.95 (dd, *J* = 6.4, 1.9 Hz, 1H, H-3), 3.89 (dd, *J* = 6.3, 4.2 Hz, 1H, H-2), 3.75 (s, 3H, CH_3 PMB), 3.71 (s, 3H, CH_3 PMB), 3.64 – 3.57 (m, 1H, H-5_a), 3.51 (dd, *J* = 11.3, 3.0 Hz, 1H, H-5_b), 1.27 (d, *J* = 6.5 Hz, 3H, CH_3 Thr), 0.97 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 171.4 (C=O COOH), 159.6, 159.3 (Cq PMB), 156.1 (C=O Fmoc), 143.9, 143.7, 141.3, 41.3 (Cq Fmoc), 135.6, 135.6 (CH arom. TBDPS), 133.1, 132.8 (Cq TBDPS), 129.9, 129.9, 129.8, 129.6 (CH arom.), 128.9 (Cq PMB), 127.8, 127.8, 127.8, 127.1, 125.2 (CH arom.), 120.0, 120.0 (CH arom. Fmoc), 114.0, 113.8 (CH arom. PMB), 103.1 (C-1), 84.7 (C-4), 78.0 (C-2), 74.9 (CH Thr), 74.5 (C-3), 72.2, 72.1 (CH_2 PMB), 67.2 (CH_2 Fmoc), 64.1 (C-5), 57.8 (CH Thr), 55.3, 55.2 (CH_3 PMB), 47.2 (CH Fmoc), 26.8 (CH_3 *t*Bu), 19.2 (Cq *t*Bu), 16.9 (CH_3 Thr). **HRMS:** $[C_{56}H_{61}NO_{11}Si + Na]^+$ found: 974.3910, calculated: 974.3906.

1-*O*-(2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl cysteine (30)

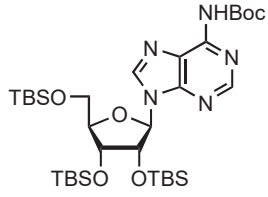
Compound **27** (1.08 g, 1.09 mmol, 1.0 eq.) was dissolved in DCM (11 mL, 0.1 M). DMBA (775 mg, 4.95 mmol, 4.5 eq.) and a catalytic amount of $Pd(PPh_3)_4$ were added and the reaction was stirred for 15 minutes before TLC showed full conversion of the starting material into a lower running product. The reaction was diluted with DCM, washed with sat. aq. $NaHCO_3$. The water layer was extracted six times with DCM. The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. After purification by flash column chromatography (0.5 -> 5% MeOH in DCM) all fractions containing the title compound were combined and concentrated *in vacuo*. The residue was taken up in DCM and washed with 10% aq. citric acid. The water layer was extracted twice with DCM and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo* yielding the title compound as an orange oil (943 mg, 0.988 mmol, 91%). **Rf:** 0.10 (25% EtOAc in pentane + 0.1% AcOH). **$[\alpha]_D^{20}$:** +62.8° (c = 1.0, $CHCl_3$). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.67 (d, *J* = 7.7 Hz, 2H, Fmoc arom.), 7.58 – 7.45 (m, 6H, Fmoc arom. + TBDPS arom.), 7.40 – 7.09 (m, 14H, Fmoc arom. + TBDPS arom. + PMB arom.), 6.81 (t, *J* = 8.3 Hz, 4H, PMB arom.), 6.59 (d, *J* = 8.3 Hz, 1H, NH), 5.40 (d, *J* = 5.4 Hz, 1H, H-1), 4.75 – 4.42 (m, 5H, CH Cys + 2x CH_2 PMB), 4.39 – 4.24 (m, 2H, H-4 + CH_{2a} Fmoc), 4.20 – 3.96 (m, 4H, CH_{2b} Fmoc + H-2 + H-3 + CH Fmoc), 3.81 – 3.67 (m, 7H, 2x CH_3 PMB + H-5_a), 3.60 (dd, *J* = 11.5, 2.9 Hz, 1H, H-5_b), 3.46 – 3.36 (m, 1H, CH_{2a} Cys), 2.96 (dd, *J* = 14.3, 4.0 Hz, 1H, CH_{2b} Cys), 0.93 (s, 9H, *t*Bu TBDPS).**^{13}C NMR:** (101 MHz, $CDCl_3$) δ 176.4 (C=O COOH), 159.4, 159.2 (Cq PMB), 156.5 (C=O Fmoc), 143.9, 143.8, 141.2 (Cq Fmoc), 135.6, 135.5 (CH arom. TBDPS), 133.1, 133.0 (Cq TBDPS), 130.1 (Cq PMB), 129.8, 129.8 (CH arom.), 129.6 (Cq PMB), 129.5, 127.7, 127.7, 127.6, 127.1, 127.1, 125.3, 125.2 (CH arom.), 119.9 (CH arom. Fmoc), 113.9, 113.8 (CH arom. PMB), 89.4 (C-1), 82.7 (C-4), 78.1 (C-2), 76.1 (C-3), 72.7, 72.2 (CH_2 PMB), 67.2 (CH_2 Fmoc), 63.4 (C-5), 55.3 (CH_3 PMB), 54.3 (CH Cys), 47.0 (CH Fmoc), 34.5 (CH_2 Cys), 26.8 (CH_3 *t*Bu), 19.2 (Cq *t*Bu). **HRMS:** $[C_{55}H_{59}NO_{10}SSi + Na]^+$ found: 976.3520, calculated: 976.3521.

2',3',5'-tri-*O*-*tert*-butyldimethylsilyl adenosine (31)



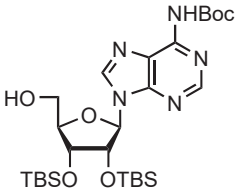
Adenosine (2.67 g, 10.0 mmol) was co-evaporated thrice with anhydrous pyridine, dissolved in anhydrous DMF (20 mL, 0.2 M) and heated to 50 °C after which imidazole (3.40 g, 50.0 mmol, 5.0 eq.) and *tert*-butyldimethylsilyl chloride (50 wt. % in toluene, 12.2 mL, 35.0 mmol, 3.5 eq.) were added. After stirring overnight, the reaction was quenched by the addition of H_2O (10 mL). The reaction was diluted with Et_2O and washed once with H_2O. The aqueous layer was extracted twice with Et_2O and the combined organic layers were washed with brine, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by column chromatography (20 -> 30% EtOAc in toluene) afforded the title compound as a crystalline white solid (6.11 g, 10.0 mmol, quant.). Spectral data was in accordance with literary precedence.[28] **Rf:** 0.19 (20% EtOAc in toluene). **^{1}H NMR** (400 MHz, $CDCl_3$): δ 8.35 (s, 1H, H-2), 8.18 (s, 1H, H-8), 6.30 (s, 2H, 6-NH_2), 6.05 (d, 1H, *J* = 5.2 Hz, H-1'), 4.70 (dd, 1H, *J* = 5.1, 4.3 Hz, H-2'), 4.33 (app. t, 1H, *J* = 3.9 Hz, H-3'), 4.14 (td, 1H, *J* = 3.9, 2.8 Hz, H-4'), 4.05 (dd, 1H, *J* = 11.3, 4.2 Hz, H-5'), 3.80 (dd, 1H, *J* = 11.3, 2.9 Hz, H-5'), 0.96 (s, 9H, CH_3 *t*Bu), 0.94 (s, 9H, CH_3 *t*Bu), 0.81 (s, 9H, CH_3 *t*Bu), 0.16 – 0.13 (m, 6H, CH_3 SiMe), 0.12 – 0.10 (m, 6H, CH_3 SiMe), -0.03 (s, 3H, CH_3 SiMe), -0.21 (s, 3H, CH_3 SiMe). **^{13}C NMR** (101 MHz, $CDCl_3$): δ 155.8 (C-6), 153.0 (C-2), 150.0 (C-4), 139.7 (C-8), 120.1 (C-5), 88.4 (C-1'), 85.5 (C-4'), 75.9 (C-2'), 72.1 (C-3'), 62.6 (C-5'), 26.2, 26.0, 25.8 (*t*Bu TBS), 18.6, 18.2, 18.0 (Cq *t*Bu), -4.3, -4.6, -4.6, -5.0, -5.3 (SiMe).

***N*^{6}-*tert*-butyloxycarbonyl-2',3',5'-tri-*O*-*tert*-butyldimethylsilyl adenosine (32)**



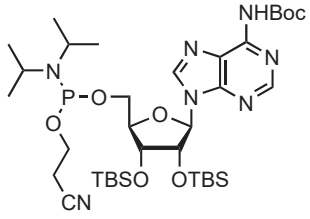
Compound **31** (5.40 g, 8.86 mmol) was co-evaporated thrice with anhydrous 1,4-dioxane, dissolved in anhydrous THF (89 mL, 0.1 M) and cooled to 0 °C, after which DMAP (217 mg, 1.77 mmol, 0.2 eq.) and di-*tert*-butyl dicarbonate (6.1 mL, 26.6 mmol, 3.0 eq.) were added. The mixture was heated to reflux and stirred for 2.5 hours. The reaction was allowed to warm to room temperature and concentrated *in vacuo*. The resulting residue was taken up in EtOAc, washed with brine, dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resulting red oil was dissolved in MeOH (89 mL, 0.1 M), cooled to 0 °C and methylamine (33 wt. % in EtOH, 3.3 mL, 35.4 mmol, 4.0 eq.) was added. After stirring overnight while gradually warming to room temperature, the resulting solution was concentrated *in vacuo*. Flash column chromatography (0 -> 30% EtOAc in pentane) furnished the title compound as a crystalline white solid (6.29 g, 8.66 mmol, quant.). Spectral data was in accordance with literary precedence.[28] **Rf:** 0.84 (20% EtOAc in toluene). **^{1}H NMR:** (400 MHz, $CDCl_3$): δ 8.74 (s, 1H, H-2), 8.33 (s, 1H, H-8), 8.22 (s, 1H, 6-NH), 6.09 (d, 1H, *J* = 5.3 Hz, H-1'), 4.65 (dd, 1H, *J* = 5.3, 4.3 Hz, H-2'), 4.29 (dd, 1H, *J* = 4.3, 3.4 Hz, H-3'), 4.14 (app. q, 1H, *J* = 3.5 Hz, H-4'), 4.02 (dd, 1H, *J* = 11.4, 3.9 Hz, H-5'), 3.80 (dd, 1H, *J* = 11.4, 2.7 Hz, H-5'), 1.55 (s, 9H, CH_3 *t*Bu Boc), 0.96 (s, 9H, CH_3 *t*Bu), 0.93 (s, 9H, CH_3 *t*Bu), 0.78 (s, 9H, CH_3 *t*Bu), 0.15 (s, 3H, CH_3 SiMe), 0.13 (s, 3H, CH_3 SiMe), 0.10 – 0.09 (m, 6H, CH_3 SiMe), -0.06 (s, 3H, CH_3 SiMe), -0.29 (s, 3H, CH_3 SiMe). **^{13}C NMR** (101 MHz, $CDCl_3$): δ 153.0 (C-2), 150.9 (C-6), 149.9 (C-4), 149.8 (C=O), 141.4 (C-8), 122.1 (C-5), 88.4 (C-2'), 85.8 (C-3'), 82.2 (Cq *t*Bu Boc), 76.2 (C-1'), 72.1 (C-4'), 62.7 (C-5'), 28.2 (*t*Bu Boc), 26.1, 25.9, 25.7 (*t*Bu TBS), 18.6, 18.1, 17.9 (Cq *t*Bu), -4.4, -4.6, -4.6, -5.1, -5.3, -5.3 (SiMe).

N^6-*tert*-butyloxycarbonyl-2',3'-di-*O*-*tert*-butyldimethylsilyl adenosine (33)



Compound **32** (6.29 g, 8.86 mmol) in THF (89 mL, 0.1 M) was cooled to 0 °C and a freshly prepared TFA:H_2O mixture (1:1; v/v, 65.8 mL, 50 eq.) was gradually added. After stirring for 4.5 hours at 0 °C, the solution was quenched by the careful addition of solid $NaHCO_3$ until pH ~ 7 was reached. The reaction was diluted with sat. aq. $NaHCO_3$ and extracted thrice with EtOAc. The resulting organic layers were combined, dried over $MgSO_4$ filtered and concentrated *in vacuo*. Flash column chromatography (30% -> 50% EtOAc in pentane) furnished the title compound as a white foam (4.98 g, 8.36 mmol, 94%). Spectral data was in accordance with literary precedence.[28] **Rf**: 0.79 (50% EtOAc in pentane). **^{1}H NMR:** (400 MHz, $CDCl_3$): δ 8.75 (s, 1H, H-2), 8.32 (s, 1H, 6-NH), 7.99 (s, 1H, H-8), 6.24 (br. s, 1H- 5'-OH), 5.83 (d, 1H, *J* = 7.8 Hz, H-1'), 5.03 (dd, 1H, *J* = 7.8, 4.6 Hz, H-2'), 4.34 (d, 1H, *J* = 4.5 Hz, H-3'), 4.18 (d, 1H, *J* = 1.7 Hz, H-4'), 3.96 (dd, 1H, *J* = 13.1, 1.8 Hz, H-5'), 3.73 (d, 1H, *J* = 13.1 Hz, H-5'), 1.57 (s, 9H, CH_3 *t*Bu Boc), 0.95 (s, 9H, CH_3 *t*Bu), 0.73 (s, 9H, CH_3 *t*Bu), 0.14 – 0.12 (m, 6H, CH_3 SiMe), -0.14 (s, 3H, CH_3 SiMe), -0.66 (s, 3H, CH_3 SiMe). **^{13}C NMR:** (101 MHz, $CDCl_3$): δ 152.5 (C-2), 150.8 (C-6), 149.7 (C=O), 149.5 (C-4), 142.9 (C-8), 123.3 (C-5), 91.2 (C-1'), 89.7 (C-4'), 82.6 (Cq *t*Bu Boc), 74.0 (C-2'), 74.0 (C-3'), 63.0 (C-5'), 28.2 (*t*Bu Boc), 25.9, 25.7 (*t*Bu TBS), 18.1, 17.8 (Cq *t*Bu), -4.5, -4.5, -4.6, -5.9 (SiMe).

5'-*O*-(N^6-*tert*-butyloxycarbonyl-2',3'-di-*O*-*tert*-butyldimethylsilyladenosine)-2-cyanoethyl-*N*-*N*-diisopropylphosphoramidite (10)



Compound **33** (1.49 g, 2.5 mmol) was co-evaporated thrice with toluene and dissolved in anhydrous DCM (25 mL, 0.1 M). DIPEA (1.34 mL, 7.5 mmol, 3.0 eq.) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (0.61 mL, 2.75 mmol, 1.1 eq.) were added to the reaction and after 3 hours of stirring, TLC indicated full conversion of starting material. The reaction was diluted with DCM and washed with brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (20 -> 30% EtOAc in pentane + 1% TEA) afforded the title compound as a white foam and a mixture of diastereomers (S_P/R_P). (1.77 g, 2.23 mmol, 89%). **Rf:** 0.76 (2:3 EtOAc in pentane). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 8.76 – 8.67 (m, 1H, H-2), 8.36 – 8.29 (m, 1H, H-8), 8.03 (s, 1H, 6-NH), 6.04 (dd, *J* = 15.4, 5.3 Hz, 1H, H-1'), 4.86 – 4.67 (m, 1H, H-2'), 4.36 – 4.18 (m, 2H, H-3' + H-4'), 4.12 – 3.75 (m, 4H, H-5' + OCH_2CH_2CN), 3.62 (m, 3H, H-5' + 2x RN$CH(CH_3)_2$), 2.68 (m, 2H, OCH_2CH_2CN), 1.55 (s, 9H), 1.20 (dq, *J* = 8.5, 3.6, 3.0 Hz, 12H, 2x RNCH(CH_3)$_2$), 0.92 (m, 9H, CH_3 *t*Bu), 0.79 – 0.74 (m, 9H, CH_3 *t*Bu), 0.14 – 0.03 (m, 6H, CH_3 SiMe), -0.02 – -0.14 (m, 3H, CH_3 SiMe), -0.18 – -0.39 (m, 3H, CH_3 SiMe). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 153.0, 152.9 (C-2), 151.0, 150.8 (C-6), 149.9, 149.9, 149.8 (C=O + C-4), 141.8 (C-8), 122.3, 122.2 (C-5), 117.5, 117.5 (Cq OCE), 89.1, 88.4 (C-1'), 85.2, 85.1, 84.4, 84.3 (C-4'), 82.3, 82.3 (Cq *t*Bu Boc), 75.6, 75.5 (C-2'), 73.0, 72.0 (C-3'), 63.0, 62.8, 62.2, 62.0 (C-5'), 58.8, 58.7, 58.6, 58.5 (OCH_2CH_2CN), 43.3, 43.2, 43.2, 43.1 (RNCH$(CH_3)_2$), 28.2 (*t*Bu Boc), 25.9, 25.9, 25.8, 25.8 (*t*Bu TBS), 24.9, 24.9, 24.8, 24.8, 24.7 (RNCH(CH_3)$_2$), 20.5, 20.5, 20.5 (OCH_2CH_2CN), 18.2, 18.1, 17.9 (Cq *t*Bu), -4.3, -4.4, -4.6, -4.6, -4.6, -4.9, -5.1 (SiMe). **^{31}P NMR:** (162 MHz, $CDCl_3$) δ 149.8, 149.7. **HRMS:** *mass was detected as its corresponding H-phosphonate* $[C_{30}H_{53}N_6O_8PSi_2 + H]^+$ found: 713.3272, calculated: 713.3274.

Solid phase synthesis

Peptide synthesis

The intermediate peptides were synthesized using standard, Fmoc-based solid phase peptide synthesis utilizing (pre-loaded) TentaGel® S AC purchased from Rapp Polymer Gmbh. Coupling cycles were as followed: Fmoc deprotection: 2x2 minutes, 1x5 minutes treatment with 20% piperidine in DMF.

Coupling: treatment of 6 eq. amino acid, 6 eq. HCTU (0.25 M in DMF) and 12 eq. DIPEA (1 M in DMF) for 30 minutes. Capping: 2x2 minutes treatment of the resin with a 10% Ac_2O solution in DMF and catalytic DIPEA. Washing between the steps was done with DMF. Ribosylated amino acids **28**, **29** and **30** were incorporated in the sequence by adding a solution of 2 eq. building block in a 0.25 M HCTU solution (2 eq.) in DMF and a 1 M DIPEA solution (4 eq.) in DMF to the resin in a fritted syringe. The resin was shaken overnight and thoroughly washed.

On-resin phosphorylation

The resin was treated with a sufficient amount of 1 M TBAF in THF (depending on the scale but enough so that the entire resin is submerged) for 30 minutes. The resin was thoroughly washed with DCM and DMF before the treatment was repeated once, furnishing the desilylated intermediate. The resin was then extensively washed with dry MeCN and flushed with nitrogen to remove traces of water before the resin was subjected to a solution of 5 eq. of $(FmO)_2PN(iPr)_2$ (as 0.25 M in MeCN) with 10 eq. ETT solution (as 0.25 M in MeCN). The resin was shaken for 30 minutes after which the resin was washed with MeCN. The resin was then treated with a sufficient amount of CSO solution (0.5 M in MeCN, depending on the scale but enough so that the entire resin is submerged) for 30 minutes. The resin was then treated with a 10% DBU solution in DMF (2x 15 minutes) to furnish the crude, immobilized and deprotected phosphoribosylated peptide.

Construction of the pyrophosphates

The resin was extensively washed with dry MeCN and flushed with nitrogen to remove traces of water. Phosphoramidite **10** (3 eq. as a 0.3 M solution in MeCN) and ETT (6 eq. as a 0.25 M in MeCN) were added to the resin and the reaction was shaken for 30 minutes. The resin was thoroughly washed with MeCN before a sufficient amount of CSO (0.5 M in MeCN, depending on the scale but enough so that the entire resin is submerged) was added to the resin and shaken for 30 minutes.

Final deprotection and cleavage

The resin was treated with a 10% DBU solution in DMF (2x 10 minutes) to remove the cyano ethyl protecting group. The resin was then treated with a 1 M TBAF solution in THF (2x 45 minutes) and washed with DMF followed by DCM. Final cleavage/deprotection occurred by treating the resin with a cleavage cocktail (2.5:10:87.5 TIS:TFA:DCM) for 4 hours. The crude products were collected by filtration and the resin was washed with a solution of 1:1:1 water:*t*BuOH:MeCN. The solvent was evaporated *in vacuo* and co-evaporated with a 1:1:1 water:*t*BuOH:MeCN solution.

Ac-Pro-Ala-Lys-Ser(5-O-adenosine diphosphate-α-D-ribosyl)-Ala-Pro-Ala-Pro-Lys-Lys-Gly-OH (39)

The general procedures described above were applied to 25 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Pro-OH, Fmoc-Ala-OH, Fmoc-Lys(Mtt)-OH and **28**. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (1.49 mg, 0.89 µmol, 3.6%). **^{1}H NMR:** (850 MHz, D_2O) δ 8.48 (s, 1H, H-2 adenine), 8.23 (s, 1H, H-8 adenine), 6.10 (d, J = 6.0 Hz, 1H, H-1' adenosine), 4.94 (d, J = 3.5 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -11.21, -11.31, -11.35, -11.46. **LC-MS:** (0 -> 50% B in A): Rt = 3.61. **HRMS:** $[C_{64}H_{105}N_{19}O_{27}P_2 + 2H]^{2+}$ found:817.8522, calculated: 817.8524.

Ac-Gly-Lys-Ser(5-*O*-adenosine-diphosphate-α-D-ribosyl)-Gly-Ala-Ala-Leu-Ser-Lys-Lys-Gly-OH (40)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Gly-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Ala-OH, Fmoc-Leu-OH, Fmoc-Ser(Trt)-OH and **28**. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (8.87 mg, 5.46 µmol, 11%). **^{1}H NMR:** (500 MHz, D_2O) δ 8.50 (s, 1H, H-2 adenine), 8.25 (s, 1H, H-8 adenine), 6.13 (d, *J* = 5.9 Hz, 1H, H-1' adenosine), 4.99 (d, *J* = 3.6 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -10.5, -10.6, -10.7, -10.8. **LC-MS:** (0 -> 20% B in A): Rt = 8.54. **HRMS:** $[C_{59}H_{101}N_{19}O_{28}P_2 + 2H]^{2+}$ found: 793.8334, calculated: 793.8342.

Ac-Gly-Lys-Ser-(5-*O*-adenosine-diphosphate-α-D-ribosyl)-Ser-Gly-Pro-Thr-Ser-Leu-Phe-Ala-Val-Thr-Val-Ala-Pro-Pro-Gly-Ala-Arg-Gly-OH (41)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Gly-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Ser(Trt)-OH, Fmoc-Pro-OH, Fmoc-Thr(Trt)-OH, Fmoc-Leu-OH, Fmoc-Phe-OH, Fmoc-Ala-OH, Fmoc-Val-OH, Fmoc-Arg(Alloc)$_2$-OH and **28**. The Alloc protecting group was removed by treating the resin with a freshly prepared solution of 10 mg $Pd(PPh_3)_4$ and 23 mg DMBA in 2.5 mL DCM. This procedure was then repeated twice prior to cleavage and final deprotection of remaining protecting groups. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (7.75 mg, 3.05 µmol, 6.1%). **^{1}H NMR:** (850 MHz, D_2O) δ 8.48 (s, 1H, H-2 adenine), 8.22 (s, 1H, H-8 adenine), 7.29 (t, J = 7.5 Hz, 2H, Phe arom.), 7.24 (t, J = 7.4 Hz, 1H, Phe arom.), 7.18 (d, J = 7.4 Hz, 2H, Phe arom.), 6.09 (d, J = 6.0 Hz, 1H, H-1' adenosine), 4.96 (d, J = 3.4 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -10.51, -10.62, -10.67, -10.77. **LC-MS:** (10 -> 90% B in A): Rt = 3.58. **HRMS:** $[C_{103}H_{164}N_{30}O_{41}P_2 + 2H]^{2+}$ found: 1270.5652, calculated: 1270.5646.

Ac-Gly-Lys-Ser(5-*O*-adenosine-diphosphate-α-D-ribosyl)-Gly-Ala-Ala-Leu-Ser-($^{13}C_6$-Lys)-Lys-Gly-OH (42)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Gly-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Ala-OH, Fmoc-Leu-OH, Fmoc-Ser(Trt)-OH, Fmoc-$^{13}C_6$-Lys(Boc)-OH and **28**. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (3.71 mg, 2.28 µmol, 4.6%). **HRMS:** $[C_{53}{}^{13}C_6H_{101}N_{19}O_{28}P_2 + 2H]^{2+}$ found: 796.8429, calculated: 796.8443.

Ac-Gly-Lys-Ser-(5-*O*-adenosine-diphosphate-α-D-ribosyl)-Ser-Gly-Pro-Thr-Ser-($^{13}C_6$-Leu)-Phe-Ala-Val-Thr-Val-Ala-Pro-Pro-Gly-Ala-Arg-Gly-OH (43)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Gly-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Ser(Trt)-OH, Fmoc-Pro-OH, Fmoc-Thr(Trt)-OH, Fmoc-Leu-OH, Fmoc-Phe-OH, Fmoc-Ala-OH, Fmoc-Val-OH, Fmoc-Arg(Alloc)$_2$-OH, Fmoc-$^{13}C_6$-Leu-OH and **28**. The Alloc protecting group was removed by treating the resin with a freshly prepared solution of 10 mg $Pd(PPh_3)_4$ and 23 mg DMBA in 2.5 mL DCM. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (2.02 mg, 0.78 µmol, 1.2%). **HRMS:** $[C_{97}{}^{13}C_6H_{164}N_{30}O_{41}P_2 + 2H]^{2+}$ found: 1273.5735, calculated: 1273.5746.

Ac-Gly-Lys-Ser-Ser-Gly-Pro-Thr(5-*O*-adensoine-diphosphate-α-D-ribosyl)-Ser-Leu-Phe-OH (44)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with phenylalanine. The amino acids used were: Fmoc-Gly-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Ser(Trt)-OH, Fmoc-Pro-OH, Fmoc-Thr(Trt)-OH, Fmoc-Leu-OH and **29**. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (7.57 mg, 4.74 µmol, 9.5%). **^{1}H NMR:** (850 MHz, D_2O) δ 8.50 (s, 1H, H-2 adenine), 8.21 (s, 1H, H-8 adenine), 7.28 (t, J = 7.5 Hz, 2H, Phe arom.), 7.22 (t, J = 7.3 Hz, 1H, Phe arom.), 7.16 (d, J = 7.0 Hz, 2H, Phe arom.), 6.10 (d, J = 5.6 Hz, 1H, H-1' adenosine), 4.96 (d, J = 4.5 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -11.15, -11.26, -11.33, -11.44. **LC-MS:** (10 -> 90% B in A): Rt = 3.33. **HRMS:** $[C_{60}H_{92}N_{16}O_{29}P_2 + 2H]^{2+}$ found: 782.2905, calculated: 782.2918.

Ac-Lys-Glu-Ser-Thr(5-*O*-adensoine-diphosphate-α-D-ribosyl)-Leu-His-Leu-Val-Leu-Arg-Leu-OH (45)
50 µmol TentaGel® S AC resin was loaded by treating the resin with 2.5 mL of a 0.2 M Fmoc-Leu-OH solution (10 eq.) and DIC (77 µL, 0.5 mmol, 10 eq.) in DMF together with a catalytic amount of DMAP for 2 hours. The resin was drained and washed with DMF. The general procedures described above were applied. The amino acids used were Fmoc-Lys(Mtt)-OH, Fmoc-Glu(O-2-Ph*i*Pr)-OH, Fmoc-Ser(Trt)-OH, Fmoc-Leu-OH, Fmoc-His(Trt)-OH, Fmoc-Val-OH, Fmoc-Arg(Alloc)$_2$-OH and **29**. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (0.91 mg. 0.47 µmol, 0.94%). ^{1}H NMR (850 MHz, D_2O) δ 8.47 (s, 1H, H-2 adenine), 8.21 (s, 1H, H-8 adenine), 6.09 (d, J = 6.0 Hz, 1H, H-1' adenosine), 4.95 (d, J = 4.0 Hz, 1H, H-1' ribosyl). **LC-MS:** (10 -> 90% B in A): Rt = 4.50. **HRMS:** $[C_{76}H_{128}N_{22}O_{30}P_2]^{2+}$ found: 946.4379, calculated: 946.4394.

Ac-Pro-Ala-Lys-Cys(5-O-adenosine diphosphate-α-D-ribosyl)-Ala-Pro-Ala-Pro-Lys-Lys-Gly-OH (46)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Pro-OH, Fmoc-Ala-OH, Fmoc-Lys(Mtt)-OH and **30.** The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (3.43 mg, 2.04 µmol, 4.1%). **^{1}H NMR:** (850 MHz, D_2O) δ 8.49 (s, 1H, H-2 adenine), 8.23 (s, 1H, H-8 adenine), 6.10 (d, J = 6.1 Hz, 1H, H-1' adenosine), 5.38 (d, J = 4.8 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -11.11, -11.21, -11.30, -11.40. **LC-MS:** (0 -> 50% B in A): Rt = 4.42. **HRMS:** $[C_{64}H_{105}N_{19}O_{26}P_2 + 2H]^{2+}$ found: 825.8408, calculated: 825.8410.

Biotin-Pro-Ala-Lys-Cys(5-O-adenosine diphosphate-α-D-ribosyl)-Ala-Pro-Ala-Pro-Lys-Lys-Gly-OH (47)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Pro-OH, Fmoc-Ala-OH, Fmoc-Lys(Mtt)-OH and **30.** Oxidation steps were carried out with a 0.5 M *t*BuOOH solution in MeCN. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (1.73 mg, 0.93 µmol, 1.9%). **^{1}H NMR:** (850 MHz, D_2O) δ 8.49 (s, 1H, H-2 adenine), 8.23 (s, 1H, H-8 adenine), 6.10 (d, J = 6.1 Hz, 1H, H-1' adenosine), 5.38 (d, J = 4.7 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -11.21, -11.31, -11.35, -11.46. **LC-MS:** (0 -> 50% B in A): Rt = 5.52. **HRMS:** $[C_{72}H_{117}N_{21}O_{27}P_2S_2 + 2H]^{2+}$ found: 917.8730, calculated: 917.8745.

Biochemical evaluation

Expression plasmids and protein purification

The construction of the expression plasmids and the purification procedures were described earlier.[17,34,41] Briefly, expression plasmids were transferred into Rossetta (DE3) cells and grown to an OD_{600} of 0.6 in LB medium supplemented with appropriate antibiotics. For metal-coordinating proteins the medium was further enriched either by addition of 2 mM $MgSO_4$ (ARHs) or 100 uM $ZnCl_2$ (*Spy*MacroD). Expression was induced with 0.4 mM isopropyl β-D-1-thioglactopyranoside (IPTG) and cultured were allowed to grow further overnight at 37 °C. Cultures were harvested by centrifugation, pellets resuspended in lysis buffer (50 mM TrisHCl [pH 8], 500 mM NaCl and 25 mM imidazole) and stored at -20 °C. Proteins were purified by Ni^{2+}-NTA chromatography (Jena Bioscience) according to the manufacturer's protocol using the following buffers: all buffers contained 50 mM TrisHCl (pH 8) and 500 mM NaCl; additionally, the lysis buffer contained 25 mM, the washing buffer 40 mM, and the elution buffer 500 mM imidazole. Proteins were dialyzed overnight against 50 mM TrisHCl (pH 8), 200 mM NaCl, 1 mM dithiothreitol and 5% (v/v) glycerol and stored at -80 °C. For the purification of ARH and ARH-like proteins all purification buffers were additionally supplemented with 10 mM $MgCl_2$.

(ADP-ribosyl)hydrolase activity assay

The peptide demodification assay was described earlier[14]. Briefly, peptide concentration for the assay were estimated using absorbance at l_{260nm} using the molar extinction coefficient of ADP-ribose (15,400 M^{-1} cm^{-1}). 20 µM indicated peptide were demodified by incubation with 1 µM hydrolase for 45 minutes at 30 °C in assay buffer (50 mM TrisHCl [pH 8], 200 mM NaCl, 10 mM MgCl2, 1 mM dithiothreitol and 0.2 µM human NUDT5[42]). Reactions were stopped and analyzed by performing the AMP-Glo™ assay (Promega) according to the manufacturer's protocol. Luminescence was recorded on a SpectraMax M5 plate reader (Molecular Devices) and data analyzed with GraphPad Prism 7. Control reactions were carried out in absence of peptide.

References

[1] C. G. Goff, *J. Biol. Chem.* **1974**, *249*, 6181–6190.
[2] J. Moss, M. Vaughan, *J. Biol. Chem.* **1977**, *252*, 2455–2457.
[3] S. K. Ueda, **1980**, *255*, 7616–7620.
[4] N. Ogata, K. Ueda, O. Hayaishi, **1980**, *255*, 7610–7615.
[5] Y. Zhang, J. Wang, M. Ding, Y. Yu, *Nat. Methods* **2013**, *10*, 981–984.
[6] R. Bredehorst, K. Wielckens, A. Gartemann, H. Lengyel, K. Klapproth, H. Hilz, *Eur. J. Biochem.* **1978**, *92*, 129–135.
[7] C. M. Daniels, S. E. Ong, A. K. L. Leung, *J. Proteome Res.* **2014**, *13*, 3510–3522.
[8] R. J. Collier, H. A. Cole, *Science* **1969**, *164*, 1179–1182.
[9] R. J. Collier, *Bacteriol. Rev.* **1975**, *39*, 54–85.
[10] R. E. West, J. Moss, M. Vaughan, T. Liu, T. Y. Liu, *J. Biol. Chem.* **1985**, *260*, 14428–14430.
[11] O. Leidecker, J. J. Bonfiglio, T. Colby, Q. Zhang, I. Atanassov, R. Zaja, L. Palazzo, A. Stockum, I. Ahel, I. Matic, *Nat. Chem. Biol.* **2016**, *12*, 998–1000.
[12] L. Palazzo, O. Leidecker, E. Prokhorova, H. Dauben, I. Matic, I. Ahel, *Elife* **2018**, *7*, 1–12.
[13] M. J. Suskiewicz, L. Palazzo, R. Hughes, I. Ahel, *FEBS J.* **2020**, 1–12.
[14] J. Voorneveld, J. G. M. Rack, I. Ahel, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2018**, *20*, 4140–4143.
[15] I. Gibbs-Seymour, P. Fontana, J. G. M. Rack, I. Ahel, *Mol. Cell* **2016**, *62*, 432–442.
[16] J. J. Bonfiglio, P. Fontana, Q. Zhang, T. Colby, I. Gibbs-Seymour, I. Atanassov, E. Bartlett, R. Zaja, I. Ahel, I. Matic, *Mol. Cell* **2017**, *0*, 1–9.
[17] P. Fontana, J. J. Bonfiglio, L. Palazzo, E. Bartlett, I. Matic, I. Ahel, *Elife* **2017**, *6*, 1–20.
[18] D. M. Leslie Pedrioli, M. Leutert, V. Bilan, K. Nowak, K. Gunasekera, E. Ferrari, R. Imhof, L. Malmström, M. O. Hottiger, *EMBO Rep.* **2018**, *19*, 1–11.
[19] E. Bartlett, J. J. Bonfiglio, E. Prokhorova, T. Colby, F. Zobel, I. Ahel, I. Matic, *Cell Rep.* **2018**, *24*, 3488–3502.
[20] L. J. McDonald, J. Moss, *Mol. Cell. Biochem.* **1994**, *138*, 221–226.
[21] F. Yan, C. Huang, F. Yan, C. Huang, X. Wang, J. Tan, S. Cheng, M. Wan, Z. Wang, *Mol. Cell* **2020**, 1–12.
[22] H. A. V Kistemaker, A. P. Nardozza, H. S. Overkleeft, G. A. van der Marel, A. G. Ladurner, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10634–10638.
[23] H. A. V Kistemaker, G. J. van der Heden van Noort, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2013**, *15*, 2306–2309.
[24] A. W. Chapman, *J. Chem. Soc. Perkin Trans. 1* **1925**, *127*, 1992–1998.
[25] K. Larsen, C. E. Olsen, M. S. Motawia, *Carbohydr. Res.* **2008**, *343*, 383–387.
[26] A. W. Chapman, *J. Chem. Soc. Perkin Trans. 1* **1925**, *127*, 1992–1998.
[27] H. Gold, P. van Delft, N. Meeuwenoord, J. D. C. Codée, D. V. Filippov, G. Eggink, H. S. Overkleeft, G. A. van der Marel, *J. Org. Chem.* **2008**, *73*, 9458–9460.
[28] C. Atdjian, L. Iannazzo, E. Braud, M. Ethève-Quelquejeu, *European J. Org. Chem.* **2018**, *2018*, 4411–4425.
[29] C. A. Vivelo, R. Wat, C. Agrawal, H. Y. Tee, A. K. L. Leung, *Nucleic Acids Res.* **2017**, *45*, D204–D209.
[30] V. Ayyappan, R. Wat, C. Barber, C. A. Vivelo, K. Gauch, P. Visanpattanasin, G. Cook, C. Sazeides, A. K. L. Leung, *Nucleic Acids Res.* **2020**, 1–5.
[31] F. Eisele, J. Kuhlmann, H. Waldmann, *Chem. - A Eur. J.* **2002**, *8*, 3362–3376.
[32] R. K. Garlick, R. W. Giese, *Biochem. J.* **1990**, *268*, 611–613.

[33] X. Zhou, J. Shearer, S. E. Rokita, *J. Am. Chem. Soc.* **2000**, *122*, 9046–9047.
[34] J. G. M. Rack, R. Morra, E. Barkauskaite, R. Kraehenbuehl, A. Ariza, Y. Qu, M. Ortmayer, O. Leidecker, D. R. Cameron, I. Matic, A. Y. Peleg, D. Leys, A. Traven, I. Ahel, *Mol. Cell* **2015**, *59*, 309–320.
[35] C. D. Appel, G. K. Feld, B. D. Wallace, R. S. Williams, *Protein Sci.* **2016**, *25*, 1682–1691.
[36] S. Buch-Larsen, I. Hendriks, J. Lodge, M. Rykær, B. Furtwängler, E. Shishkova, M. Westphall, J. Coon, M. Nielsen, *Cell Rep.* **2020**, *32*, 108176.
[37] L. Zhang, J. Cao, L. Dong, H. Lin, *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 13447–13456.
[38] M. E. Grunewald, M. G. Shaban, S. R. Mackin, A. R. Fehr, S. Perlman, *J. Virol.* **2020**, *94*, 1–17.
[39] A. Gomez, C. Bindesbøll, S. V. Satheesh, G. Grimaldi, D. Hutin, L. MacPherson, S. Ahmed, L. Tamblyn, T. Cho, H. I. Nebb, A. Moen, J. H. Anonsen, D. M. Grant, J. Matthews, *Biochem. J.* **2018**, *475*, 3827–3846.
[40] 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, *Nat. Commun.* **2021**, *12*, 1–18.
[41] D. Slade, M. S. Dunstan, E. Barkauskaite, R. Weston, P. Lafite, N. Dixon, M. Ahel, D. Leys, I. Ahel, *Nature* **2011**, *477*, 616–622.
[42] L. Palazzo, B. Thomas, A. S. Jemth, T. Colby, O. Leidecker, K. L. H. Feijs, R. Zaja, O. Loseva, J. C. Vert, I. Matic, T. Helleday, I. Ahel, *Biochem. J.* **2015**, *468*, 293–301.

Chapter 4

Solid-Phase Synthesis of Peptides with ADP-Ribosylated Tyrosine

Introduction

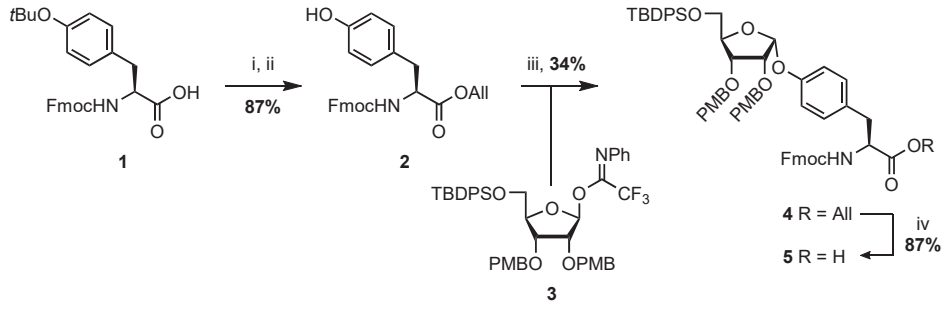
After the emergence of serine (Ser) as an ADPr acceptor,[1] it became apparent that ADPr as a post-translational modification extended to a wider variety of acceptor sources than previously thought. This prompted an unbiased re-evaluation of the LC-MS/MS datasets from past proteomics studies for the possible presence of the ADP-ribosylated amino acids that were not considered before as prevalent sites of modification, for example, Ser, tyrosine (Tyr) and threonine (Thr).[2] Besides redirection of the ADP-ribosylation site from lysine (Lys) to Ser in the apparent Lys-Ser (KS) recognition motif,[3] Tyr proved to be an acceptor of ADPr as well, albeit to a lesser extent than Ser.[4–7] It is noteworthy that this was observed as well in protein phosphorylation patterns, where Ser and Thr are modified more abundantly than Tyr which was estimated to take up less than 1% of the total phosphoproteome.[8–10] Furthermore, both Ser- and Tyr-ADP-ribosylation sites have been found to significantly overlap with phosphorylation sites,[4,11] indicating a possible crosstalk between ADP-ribosylation and phosphorylation of proteins. The importance of phosphoryl Tyr (P-Tyr) as a regulatory modification (excellently reviewed here[12]) is abundantly clear and it is suggested that the functional role of P-Tyr is quite distinct from that of P-Ser and P-Thr.[13] Given the possible interplay between Tyr-ADPr/Ser-ADPr and the phosphorylation of Ser one cannot exclude a regulatory role for Tyr-ADPr.

The identified Tyr-ADPr proteins appear to be mostly localized in the nucleus and are assembled by the PARP1:HPF1 complex.[4,14] ADP-ribosylation of Tyr by PARP1:HPF1 is also dependent on DNA damage thus bearing similarities with its Ser-ADPr counterpart. As described above, Tyr-ADPr occurs to a lesser extent, being responsible for approximately 3% of the ADPr proteome compared to 83% for Ser-ADPr.[5] Interestingly, HPF1 is ADP-ribosylated on Y238, a residue crucial for proper functioning of HPF1[15] which might indicate a regulatory role for the Tyr-ADPr modification on HPF1. Besides HPF1, PARP14 is modified on three Tyr residues.[5] PARP14 is a *mono*-ADP-ribosylating (MARylating) enzyme[16] and plays a regulatory role in RNA stability by modifying PARP13 and several other players in the RNA regulatory proteome.[17] The crosstalk between several members of the PARP family, HPF1 and Tyr-ADPr raises questions as to the exact physiological function of Tyr as ADPr acceptor. Synthetic, well-defined Tyr-ADP ribosylated peptides can greatly assist in elucidating the physiological role of Tyr-ADPr as is demonstrated by recent studies using synthetic peptides ADP-ribosylated at other amino acids.[18–20] This chapter describes a solid phase methodology developed for the synthesis of peptides MARylated at a Tyr residue and the biological evaluation thereof.

Results and Discussion

Building block synthesis

The synthetic strategy to obtain peptides modified with mono ADP-ribose at a predetermined Ser, Thr or cysteine (Cys) side chain as described in Chapter 3, was adopted to acquire peptides MARylated at a Tyr residue. Therefore, the efforts were firstly turned to synthesizing a ribosylated Tyr building block (**5**, Scheme 1) to be used in Fmoc-based solid phase peptide synthesis (SPPS). Commercially available Fmoc-Tyr(*t*Bu)-OH **1** was treated with allyl bromide under basic conditions to orthogonally protect the carboxylic acid as an allyl ester. The crude intermediate was then directly treated with TFA in DCM with TIS as a scavenger to cleave the *tert*-butyl aryl ether furnishing phenolic acceptor **2** in 87% yield over two steps.



Scheme 1. Synthesis of Tyr-ribosylated building block **5**, ready for SPPS. Reagents and conditions: i) All-Br, DIPEA, DMF. ii) 20% TFA in DCM, TIS. iii) TBSOTf, DCM, -50 °C. iv) $Pd(PPh_3)_4$, DMBA, DCM.

Next, the glycosylation reaction of this Tyr acceptor with known ribosyl donor **3** was investigated, starting with the same conditions as described in Chapter 2 for the Ser, Thr and Cys acceptors (Table 1, entry 1).[18] Under these conditions however, solubility issues arose as acceptor **2** did not dissolve properly in 0.1 M DCM at -50 °C. Therefore, a 1:1 mixture of DCM and 1,4-dioxane was used (entry 2) which proved to be effective in addressing similar issues in the past.[21] Unfortunately, the use of this solvent mixture did not improve the solubility of the Tyr acceptor. As the mixture of acceptor **2** and donor **3** is soluble in DCM up to 0.1 M concentration at room temperature, the following glycosylation procedure was attempted. At room temperature, a 0.1 M solution of compounds **2** and **3** in DCM was prepared that was subsequently slowly cooled to -20 °C during which a clear solution persisted. At this temperature, the glycosylation was performed (entry 3) as further lowering of the temperature resulted in the precipitation of acceptor **2**. However, addition of the activator resulted in a complex mixture of products which was inseparable by column chromatography. In line with the results of the similar glycosylation experiments

described in Chapter 3, the formation of multiple products can be explained by the loss of a-stereoselectivity and the unwanted acid catalyzed loss of the PMB protecting groups, at the temperature applied.[21]

Table 1. Optimization of the glycosylation conditions of acceptor **2** with donor **3**. All reactions were carried out at a 0.2 mmol scale. C (M) is the concentration of the donor in the solvent with 0.1 equivalents of activator relative to the donor. n.d. = not determined.

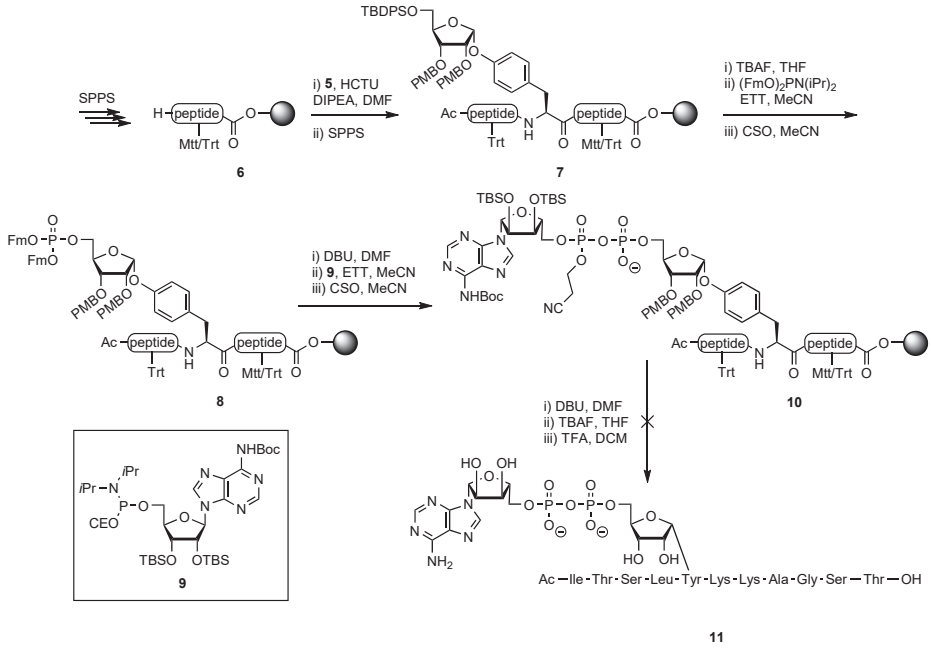
Entry	Solvent	C (M)	Activator	T (°C)	Reaction time	Yield
1	DCM	0.1	TMSOTf	-50	1 h	n.d.
2	DCM:dioxane	0.1	TMSOTf	-50	1 h	n.d.
3	DCM	0.1	TMSOTf	-20	1 h	n.d.
4	DCM	0.03	TMSOTf	-50	1 h	38%
5	DCM	0.03	TBSOTf	-50	2.5 h	47%

Decreasing the concentration of the reaction to 0.03 M allowed the temperature to be maintained at -50 °C without precipitation of acceptor **2** (entry 4). In this way, fully protected, ribosylated Tyr **4** could be obtained in a moderate 38% yield that was attributed to the slower reaction rate at this concentration. After increasing the reaction time to 2.5 hours (entry 5) and changing the activator to TBSOTf, the unwanted side reactions could be suppressed to give ribosylated Tyr **4** in 47% yield. However, scaling up the reaction to a 2.0 mmol scale proved difficult as the yield decreased to 34% (Scheme 1) but no attempts to further optimize this reaction were made as enough material was available to proceed. Pd-catalyzed cleavage of the allyl ester, using 1,3-dimethylbarbituric acid (DMBA) as scavenger, furnished orthogonally protected *O*-ribosylated Tyr **5**, suitable for the intended Fmoc-based SPPS strategy.

Solid phase synthesis of MARylated Tyr-peptide.

The SPPS strategy to acquire peptide **11**, MARylated at its Tyr residue, is depicted in Scheme 2. With the aid of standard Fmoc-chemistry, TentaGel® S AC resin pre-loaded with Thr, was elongated with commercially available Fmoc-amino acids to give immobilized oligopeptide **6.** Subsequent incorporation of Tyr-building block **5** and further elongation led to immobilized peptide **7**. At this stage, the phosphotriester at the 5-OH of the ribose moiety was introduced. Starting with TBAF mediated desilylation, phosphitylation of the released alcohol with Fm protected phosphoramidite and finally oxidation of the phosphite triester produced protected phosphotriester **8**. Next, DBU mediated removal of the Fmoc groups in **8** yielded the corresponding phosphate monoester which was followed by reaction with phosphoramidite **9**. Oxidation of the resulting P^{III} – P^{V} intermediate furnished immobilized, protected MARylated Tyr-peptide **10**. Similar to its Ser, Cys and Thr

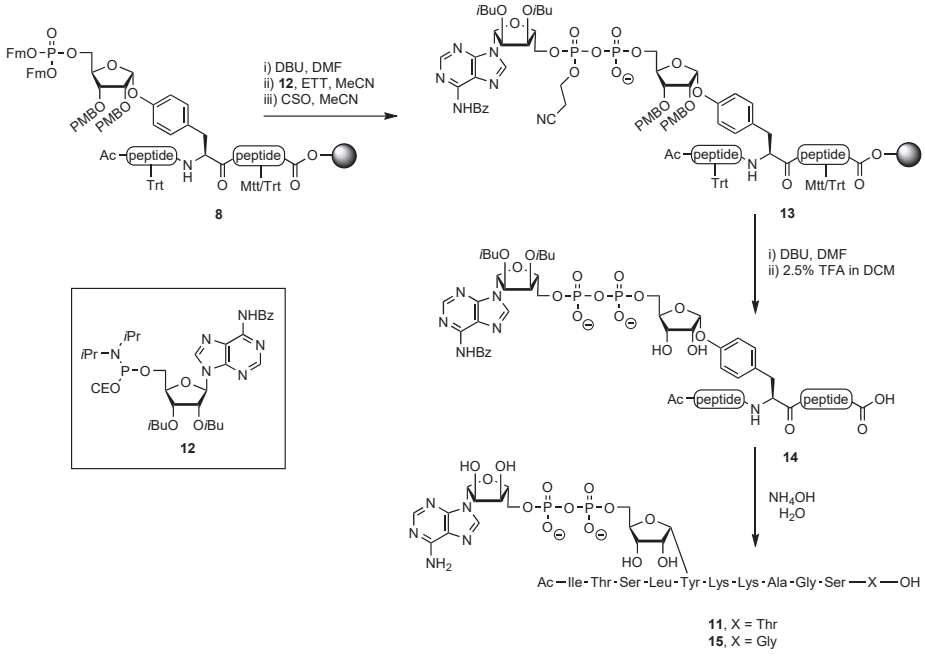
counterparts, the deprotection of **10** comprised of (i) removal of the cyanoethyl protecting group with DBU, (ii) desilylation of the 2'- and 3'-hydroxyls of the adenosine moiety and (iii) cleavage of the target peptide from the resin, together with the accompanying cleavage of the remaining protecting groups with 10% TFA in DCM. LC-MS analysis of the obtained crude mixture revealed a main product (~75% based on the UV-absorption measured in the LC-MS trace) that did not correspond to target **11** but with the free peptide in which the ADPr moiety was lacking. TFA-mediated acidolysis of the glycosidic bond of the tyrosyl ribofuranoside in **11** was suspected. Lowering the concentration of TFA in the cleavage cocktail to 5% did decrease the observed acidolysis to approximately 25% and further reducing the TFA concentration to 2.5% completely prevented this side reaction. Although the Trt, Mtt and PMB groups were completely removed in 2.5% TFA in DCM, the Boc protecting group on the exocyclic amine of the adenine was found to be too stable and required TFA levels of 5% or more for its removal.



Scheme 2. Attempted synthesis of Tyr-ADPr peptide **11** using the strategy developed in Chapter 3.

In order to prevent the unwanted acidolysis of the glycosidic linkage, an alternative for the Boc protecting group on the exocyclic amine of adenine was sought. As outlined in Chapter 2,[18,19] adenosine phosphoramidite **12** (Scheme 3), where the 2'- and 3' hydroxyl groups are protected with *iso*-butyryl and the exocyclic amine is masked with a benzoyl

group, was used for the synthesis of peptides MARylated at Ser. Application of **12** as a replacement for reagent **9** for the introduction of the pyrophosphate moiety using the same sequence of reactions, gave partially protected immobilized peptide **13** (Scheme 3). Removal of the protecting groups started with DBU-mediated cyanoethyl elimination, followed by treatment with 2.5% TFA in DCM, inducing both the removal of the Trt, Mtt, and PMB protecting groups and cleavage of the TentaGel® S AC linker, to give partially protected peptide **14**. Finally, removal of all base labile protecting groups on adenosine in **14** by treatment with a saturated, aqueous ammonium hydroxide solution resulted, after purification, in the isolation of Tyr-ADPr peptide **11** in 4.5% yield. Remarkably, LC-MS analysis showed that approximately 25% of the crude mixture to constitute a side product 18 Da lower in mass. The formation of this side product was attributed to treatment of the intermediate peptide **13** (and probably also **8**) with DBU, that allows for an E1cb reaction thereby eliminating the Trt group of the Thr residue. To prevent this side reaction, Thr was replaced with glycine (Gly) and the peptide synthesis was repeated. Gratifyingly, no dehydrated mass could be detected and *via* this method, peptide Ac-ITSLY^{ADPr}KKAGSG-OH (**15**) was obtained in 14% yield.



Scheme 3. Outline of the strategy in which adenosine amidite **9** is replaced by **12** and the accompanying deprotection sequence to provide **11** without acidolysis of the anomeric Tyr-ADPr bond.

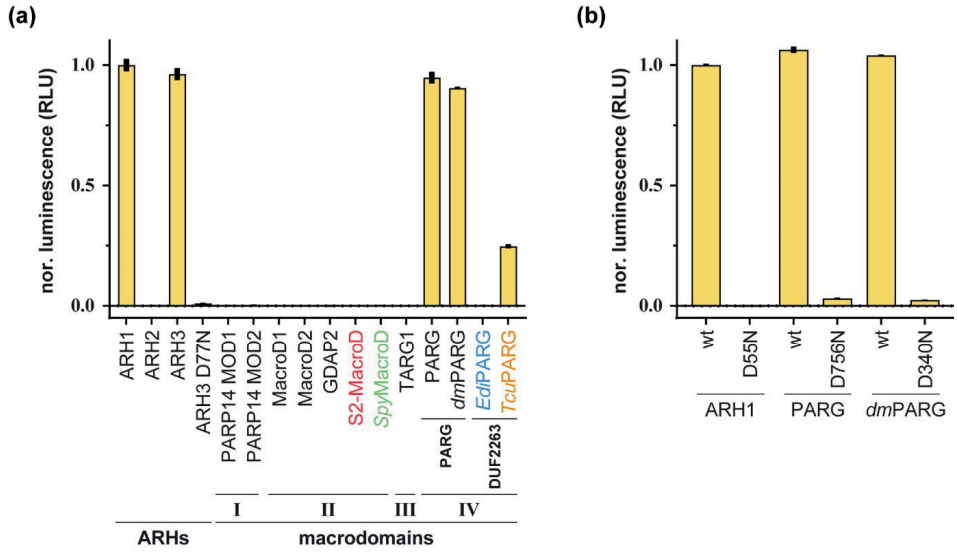


Figure 1. Hydrolysis of Tyr-MARylated peptide **11** by known (ADP-ribosyl)-hydrolases, performed by Johannes Gregor Matthias Rack affiliated with the Sir William Dunn School of Pathology, University of Oxford. Briefly, the ADP-ribose released in the hydrolysis reaction was converted by human NUDT5 to AMP, which in turn was detected via chemiluminescence using the AMP-Glo™ assay (Promega). Samples were background corrected, normalized to ARH1 wt and presented as mean values ± s.d. from triplicate measurements. (a) Hydrolysis of Tyr-MARylated peptide **11** by selected (ADP-ribosyl)-hydrolases of the ARH and macrodomain family. Protein from Animalia are indicated in black, Amoebozoa in blue, viral in red, bacterial in green and Archaea in orange. The macrodomains family is subdivided in several clades: (I) MacroH2A-like, (II) MacroD-type, (III) ALC1-like, and (IV) PARG-like. (b) Control experiment showing that Tyr-ADP-ribosylation reversal by ARH1 and PARG is enzyme-dependent.

Biological evaluation

As Tyr-ADPr is a newly found modification, the enzyme responsible for its reverse reaction is yet unknown. Therefore, a hydrolytic screening assay was performed with peptide **11**, testing a variety of ADP-ribosyl hydrolases in both the ARH- and macrodomain-family, the results of which are summarized in Figure 1a. Starting with the ARH family, ARH2 which is thought to be catalytically inactive,[22] shows no turnover of Tyr-ADPr. ARH3 and ARH1 are both efficient in hydrolyzing Tyr-ADPr which is interesting in the case of ARH1 as this hydrolase is generally thought to be selective towards *N*-linked Arg-ADPr.[22,23] This is the first report of ARH1 being able to hydrolyze an *O*-glycosidic bond despite the distinctly different modes of substrate recognition and ligand binding between ARH1 and ARH3.[24] Also, an unexpected turnover of Tyr-ADPr was encountered in the macrodomain family as PARG was able to hydrolyze peptide **11**. The surprising activity of ARH1 and PARG prompted

further experimentation with ARH1 and PARG mutants rendered catalytically inactive as control (Figure 1b). The catalytic mutants showed no enzymatic turnover of Tyr-ADPr, confirming that the displayed hydrolysis is indeed enzyme-dependent.

Conclusion

This chapter describes the synthesis of a ribosylated Fmoc-Tyr building block (**5**) and the accompanying optimization of the reaction conditions to couple ribosyl donor **3** to Tyr acceptor **2**. Building block **5** was applied in SPPS to produce two peptides MARylated at the Tyr residue. The observed acid lability of the glycosidic bond in Tyr-ADPr led to the development of a new strategy by a combination of the SPPS described in Chapters 2 and 3 by taking the temporary base-labile protecting groups on the adenosine moiety (Chapter 2) and the acid-labile protections on the ribosyl part and amino acid side chains (Chapter 3) to assemble Tyr-ADPr peptides **11** and **15**.

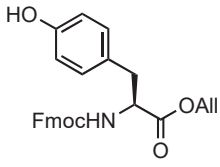
With peptide **11**, the hydrolytic properties of several known ADPr hydrolases towards the Tyr-ADPr modification was tested showing that ARH3 is able to hydrolyze Tyr-ADPr like Ser- and Thr-ADPr. However, an unexpected enzymatic activity of ARH1 and PARG towards **11** was encountered. Although ARH1 is now shown to posess a broader specificity and not limited to Arg-ADPr hydrolysis, the emergence of PARG as a Tyr-ADPr hydrolyzing enzyme is unexpected.

Experimental section

General synthetic procedures

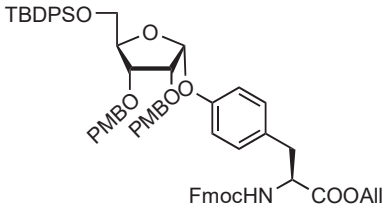
All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. Column chromatography was performed on silica gel 60 Å (40-63 µm, Macherey-Nagel). TLC analysis was performed on Macherey-Nagel aluminium sheets (silica gel 60 F_{254}). TLC was used to visualize compounds by UV at wavelength 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) or $KMnO_4$ spray (20 g/L $KMnO_4$ and 10 g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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. NMR spectra were recorded on a Bruker AV-400, AV-500 or AV-600 instrument. Chemical shifts (δ) are given in ppm relative to tetramethyl silane as internal standard. Coupling constants (*J*) are given in Hz. For compounds **11** and **15**, 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.

***N*-α-Fmoc-tyrosine allyl ester (2)**



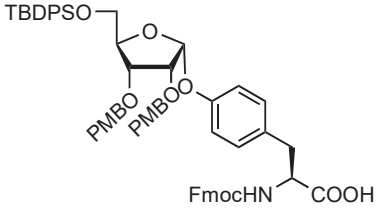
Commercially available Fmoc-Tyr(*t*Bu)-OH **1** (2.22 g, 5.00 mmol, 1 eq.) was co-evaporated thrice with 1,4-dioxane and dissolved in dry DMF (25 mL, 0.2 M). DIPEA (1.04 mL, 6 mmol, 1.2 eq.) was added followed by the addition of allyl-bromide (0.52 mL, 6 mmol. 1.2 eq.) and the reaction was stirred overnight. The reaction was quenched with H_2O (50 mL) and transferred into a separatory funnel. The reaction mixture was extracted thrice with Et_2O and the combined organic layers were washed three times with brine, dried over $MgSO_4$ and concentrated *in vacuo* furnishing the crude *tert*-butyl protected tyrosine allyl ester. The crude product was dissolved in a 4:1 v/v% DCM:TFA mixture (10 mL, 0.5 M), TIS (4.10 mL, 20 mmol, 4 eq.) was added and the reaction was stirred for 3 hours, after which it was diluted with toluene and concentrated *in vacuo*. The residue was co-evaporated thrice with toluene to remove the last traces of TFA. Flash column chromatography (10 -> 40% EtOAc in pentane) afforded the title compound as an off-white solid (1.93 g, 4.36 mmol, 87% over 2 steps) **Rf**: 0.70 in 40% EtOAc in pentane. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.76 (dd, *J* = 7.6, 1.1 Hz, 2H, Fmoc arom.), 7.65 – 7.55 (m, 2H, Fmoc arom.), 7.40 (td, *J* = 7.5, 1.1 Hz, 2H, Fmoc arom.), 7.31 (tt, *J* = 7.4, 1.3 Hz, 2H, Fmoc arom.), 5.91 (ddt, *J* = 16.4, 10.9, 5.7 Hz, 1H, OCH_2CHCH_2), 5.77 (d, *J* = 7.8 Hz, 1H, NH), 5.40 – 5.32 (m, 1H, OCH_2CHCH_{2a}), 5.26 (dd, *J* = 9.4, 1.2 Hz, 1H, OCH_2CHCH_{2b}), 4.69 (d, *J* = 5.7 Hz, 2H, OCH_2CHCH_2), 4.51 – 4.41 (m, 3H, CH_2 Fmoc + CH Tyr), 4.22 (t, *J* = 6.9 Hz, 1H, CH Fmoc), 3.98 (dd, *J* = 25.6, 11.0 Hz, 2H, CH_2 Tyr), 2.28 (s, 1H, OH). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 143.78, 141.47, 131.31 (OCH_2CHCH_2), 127.88, 127.24, 127.21, 125.22, 120.15, 119.18 (OCH_2CHCH_2), 67.34 (CH_2 Fmoc), 66.52 (OCH_2CHCH_2), 63.44 (CH_2 Tyr), 56.22 (CH Tyr), 47.25 (CH Fmoc).

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



Compound[21] **3** (1.60 g, 2.00 mmol, 1.1 eq.) and Fmoc-Tyr(OH)-OAll **2** (734 mg, 1.82 mmol, 1.0 eq. relative to the donor) were co-evaporated thrice with toluene and dissolved in DCM (70 mL, 0.03 M donor concentration). The reaction was cooled to -50 °C and TBSOTf (46 µL, 0.2 mmol, 0.1 eq. relative to the donor) was added. The reaction was stirred at -50 °C overnight after which TLC analysis showed near full conversion of the starting material. The reaction was quenched with TEA and concentrated *in vacuo*. Flash column chromatography (0.5 -> 3% acetone in DCM) yielded the title compound as a clear oil (651 mg, 0.617 mmol, 34%). **Rf:** 0.58 in 3% acetone in DCM. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.78 – 7.70 (m, 2H, Fmoc arom.), 7.67 – 7.51 (m, 6H, Fmoc arom. + TBDPS arom.), 7.48 – 7.24 (m, 14H, Fmoc arom. + TBDPS arom. + PMB arom.), 7.11 – 6.97 (m, 4H, Tyr arom.), 6.86 – 6.75 (m, 4H, PMB arom.), 5.88 (ddd, *J* = 16.5, 10.6, 5.2 Hz, 1H, CH_2CHCH_2), 5.51 (d, *J* = 4.3 Hz, 1H, H-1), 5.38 – 5.19 (m, 3H, NH + CH_2CHCH_2), 4.74 – 4.50 (m, 8H, CH Ser + CH_2CHCH_2), 4.44 (dd, *J* = 10.6, 7.1 Hz, 1H, CH_{2a} Fmoc), 4.40 – 4.29 (m, 1H, CH_{2b} Fmoc), 4.29 – 4.15 (m, 2H, H-4 + CH Fmoc), 4.09 (dd, *J* = 6.5, 2.8 Hz, 1H, H-3), 3.99 (dd, *J* = 6.4, 4.3 Hz, 1H, H-2), 3.78 (s, 3H, CH_3 PMB), 3.76 (s, 3H, CH_3 PMB), 3.58 (ddd, *J* = 53.4, 11.3, 3.1 Hz, 2H, H-5), 3.10 (d, *J* = 5.7 Hz, 2H, CH_2 Tyr), 0.95 (s, 9H, *t*Bu TBDPS).**^{13}C NMR:** (101 MHz, $CDCl_3$) δ 171.3 (C=O COOAll), 159.4, 159.3 (Cq PMB), 156.9 (C=O Fmoc), 155.7 (Cq Tyr), 143.9, 143.8, 141.4 (Cq Fmoc), 135.7, 135.6 (CH arom. TBDPS), 133.2, 133.1 (Cq TBDPS), 131.5 ($CHCHCH_2$), 130.4 (Cq PMB), 130.3 (CH arom. Tyr), 129.9, 129.8, 129.7, 129.7 (CH arom.), 128.9 (Cq PMB), 127.9, 127.9, 127.8, 127.8, 127.8, 127.2, 125.3, 125.2 (CH arom.), 120.1, 120.0 (CH arom. Fmoc), 119.3 (CH_2CHC, 117.5 (CH arom. Tyr), 113.9, 113.9, 113.8 (CH arom. PMB), 100.0 (C-1), 84.2 (C-4), 77.7 (C-2), 75.0 (C-3), 72.4, 72.1 (CH_2 PMB), 67.1 (CH_2 Fmoc), 66.2 (CH_2CHCH_2), 64.0 (C-5), 55.3, 55.3 (CH_3 PMB), 54.9 (CH Tyr), 47.2 (CH Fmoc), 37.5 (CH_2 Tyr), 26.9 (CH_3 *t*Bu), 19.3 (Cq *t*Bu). **HRMS:** $[C_{64}H_{67}NO_{11}Si + Na]^+$ found: 1076.4374, calculated: 1076.4376.

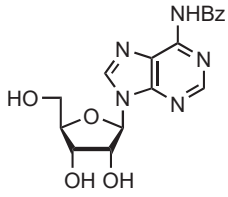
1-*O*-(2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl tyrosine (5)



Compound **4** (626 mg, 0.594 mmol, 1.0 eq.) was dissolved in DCM (6.0 mL, 0.1 M). DMBA (184 mg, 1.18 mmol, 2.0 eq.) and $Pd(PPh_3)_4$ (6.9 mg, 5.9 µmol, 0.01 eq.) were added and the reaction was stirred for 1 hour before TLC showed full conversion of the starting material into a lower running product. The reaction was diluted with DCM, washed with 1 M HCl and the organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography in (4 -> 6% MeOH in DCM) yielded the title compound as a white foam (518 mg, 0.511 mmol, 87%). **Rf:** 0.50 in 5% MeOH in DCM + 0.1% AcOH. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.72 (d, *J* = 7.5 Hz, 2H, Fmoc arom.), 7.65 – 7.49 (m, 6H, Fmoc arom. + TBDPS arom.), 7.47 – 7.20 (m, 14H, Fmoc arom. + TBDPS arom. + PMB arom.), 7.09 – 6.97 (m, 4H, Tyr arom.), 6.88 – 6.75 (m, 4H, PMB arom.), 5.48 (d, *J* = 4.3 Hz, 1H, H-1), 5.32 (d, *J* = 8.1 Hz, 1H, NH), 4.70 – 4.51 (m, 5H, 2x CH_2 PMB + CH Ser), 4.44 (dd, *J* = 10.5, 7.2 Hz, 1H, CH_{2a} Fmoc), 4.32 (dd, *J* = 10.8, 7.1 Hz, 1H, CH_{2b} Fmoc), 4.28 – 4.15 (m, 2H, H-4 + CH Fmoc), 4.08 (dd, *J* = 6.5, 2.6 Hz, 1H, H-3), 3.97 (dd, *J* = 6.5, 4.4 Hz, 1H, H-2), 3.76 (s, 3H, CH_3 PMB), 3.74 (s, 3H, CH_3 PMB), 3.63 (dd, *J* = 11.1, 3.3 Hz, 1H, H-5_a), 3.49 (dd, *J* = 11.2, 2.8 Hz, 1H, H-5_b), 3.18 – 3.01 (m, 2H, CH_2 Tyr), 0.94 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 174.8 (C=O COOH), 159.4, 159.3 (Cq PMB), 156.8 (C=O Fmoc), 155.9 (Cq Tyr), 143.9 9, 143.8, 141.4 (Cq Fmoc), 135.7, 135.6 (CH arom. TBDPS), 133.2,

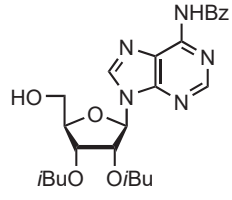
133.1 (Cq TBDPS), 130.4 (CH arom. Tyr), 130.2 (Cq PMB), 129.9, 129.9, 129.8, 129.8, 129.6 (CH arom.), 129.0 (Cq PMB), 127.9, 127.9, 127.8, 127.2, 125.3, 125.2, 120.0 (CH arom.), 117.5 (CH arom. Tyr), 114.0, 114.0, 113.8 (CH arom. PMB), 99.8 (C-1), 84.3 (C-4), 77.8 (C-2), 75.0 (C-3), 72.4, 72.1 (CH_2 PMB), 67.2 (CH_2 Fmoc), 64.0 (C-5), 55.3 (CH_3 PMB), 54.7 (CH Tyr), 47.2 (CH Fmoc), 37.0, (CH_2 Tyr), 26.9 (CH_3 *t*Bu), 19.3 (Cq *t*Bu). **HRMS:** $[C_{61}H_{63}NO_{11}Si + Na]^+$ found: 1036.4060, calculated: 1036.4063.

***N*⁶-benzoyl adenosine**



Commercially available adenosine (5.34 g, 20.0 mmol) was co-evaporated thrice with anhydrous pyridine (3 x 20 mL) and suspended in anhydrous pyridine (100 mL, 0.2 M) after which trimethylsilyl chloride (22.8 mL, 180 mmol, 9 eq.) was carefully added. After stirring for 30 minutes the solution had turned clear and benzoyl chloride (11.6 mL, 100 mmol, 5 eq.) was added dropwise over 5 minutes and stirring continued for another 3 hours. The solution was cooled in an ice bath and quenched by the addition of H_2O (20 mL). The mixture was allowed to warm to room temperature, followed by the addition of an aqueous ammonium hydroxide solution (30 wt. % NH_4OH in H_2O, 40 mL), resulting in the precipitation of the title compound as a white solid. The suspension was stirred for 1 hour, concentrated *in vacuo* and the white residue was partitioned between EtOAc (80 mL) and H_2O (280 mL). The aqueous layer was cooled to 0 °C to induce crystallization of the product which was collected by filtration, rinsed with ice cold Et_2O (30 mL) and dried under high vacuum at 60 °C to afford the title compound as a white crystalline solid (7.43 g, 20 mmol, quant.) Spectral data was in accordance with literary precedence. **Rf:** 0.25 (5% MeOH in DCM). **^{1}H NMR:** (400 MHz, DMSO-d_6): δ 10.50 (br. s, 1H, 6-NH), 8.76 (s, 1H, H-2), 8.73 (s, 1H, H-8), 8.09 – 8.02 (m, 2H, *o*-Bz), 7.70 – 7.61 (m, 1H, *p*-Bz), 7.60 – 7.51 (m, 2H, *m*-Bz), 6.04 (d, 1H, *J* = 5.8 Hz, H-1'), 5.58 (d, 1H, *J* = 6.1 Hz, 2'-OH), 5.27 (d, 1H, *J* = 4.9 Hz, 3'-OH), 5.15 (t, 1H, *J* = 5.6 Hz, 5'- OH), 4.66 (app. q, 1H, *J* = 5.6 Hz, H-2'), 4.19 (app. q, 1H, *J* = 4.9 Hz, H-3'), 3.99 (app. q, 1H, *J* = 3.9 Hz, H- 4'), 3.70 (app. dt, 1H, *J* = 11.9, 4.7 Hz, H-5'), 3.58 (ddd, 1H, *J* = 12.0, 6.1, 4.0 Hz, H-5'). **^{13}C NMR:** (101 MHz, DMSO-d_6): δ 166.2 (C=O), 152.7 (C-6), 152.1 (C-2), 150.9 (C-4), 143.7 (C-8), 133.8 (*ipso*-Bz), 133.0 (*p*-Bz), 129.0 (*o*-Bz, *m*-Bz), 126.3 (C-5), 88.1 (C-1'), 86.2 (C-4'), 74.2 (C-2'), 70.9 (C-3'), 61.8 (C-5').

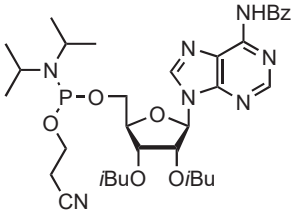
***N*⁶-benzoyl-2',3'-di-*O*-iso-butyryladenosine**



*N*⁶-benzoyl adenosine (1.86 g, 5.0 mmol) was co-evaporated thrice with anhydrous pyridine (3 x 10 mL) and dissolved in anhydrous pyridine (25 mL, 0.2 M) after which *tert*-butyldimethylsilyl chloride (50 wt. % in toluene, 1.91 mL, 5.5 mmol, 1.1 eq.) was added. After stirring overnight, isobutyric anhydride (3.0 mL, 18.1 mmol, 3.6 eq.) was added and stirring continued for a further 7 hours. The mixture was quenched by the addition of H_2O (0.36 mL, 20 mmol, 4 eq.), stirred for 30 minutes and concentrated *in vacuo*. The residue was taken up in EtOAc (100 mL) and washed successively with sat. aq. $NaHCO_3$ (50 mL), 1 M citric acid (50 mL) and H_2O (50 mL). The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by flash column chromatography (0 -> 50% EtOAc in pentane) furnished the intermediate product as a white foam (2.67 g, 4.27 mmol, 85%). The silylate compound was then dissolved in MeCN:H_2O (4:1; v/v, 43 mL, 0.1 M), after which *para*-toluene sulfonic acid monohydrate (1.14 g, 5.97 mmol, 1.4 eq.) was added. After stirring for 7 hours, the solution was neutralized by the addition of solid $NaHCO_3$ until pH ~ 7 and concentrated to dryness. The residue was partitioned between EtOAc (200 mL) and sat. aq. $NaHCO_3$ (50 mL) and the organic layer washed thrice with H_2O (3 x 50 mL). The combined water layers were extracted twice more with EtOAc (2 x 50 mL) and the resulting organic

layers combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by column chromatography (0 -> 1% MeOH in DCM) afforded the title compound as a white crystalline solid (2.22 g, 4.27 mmol, quant.). Spectral data was in accordance with literary precedence. **Rf:** 0.52 (3% MeOH in DCM). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 9.48 (br. s, 1H, 6-NH), 8.72 (s, 1H, H-2), 8.18 (s, 1H, H-8), 8.05 – 7.97 (m, 2H, *o*-Bz), 7.61 – 7.52 (m, 1H, *p*-Bz), 7.52 – 7.43 (m, 2H, m-Bz), 6.14 (d, 1H, *J* = 7.4 Hz, H-1'), 5.98 (dd, 1H, *J* = 7.5, 5.4 Hz, H-2'), 5.69 (dd, 1H, *J* = 5.4, 1.6 Hz, H-3'), 4.33 (app. q, 1H, *J* = 1.8 Hz, H-4'), 3.98 (dd, 1H, *J* = 12.9, 1.9 Hz, H-5'), 3.86 (dd, 1H, *J* = 12.8, 1.9 Hz, H-5'), 2.64 (hept, 1H, *J* = 7.0 Hz, $CH(Me)_2$*i*Bu), 2.49 (hept, 1H, J = 7.0 Hz, $CH(Me)_2$ *i*Bu), 1.24 – 1.20 (m, 6H, CH_3 *i*Bu), 1.09 (d, 3H, *J* = 7.0 Hz, CH_3 *i*Bu), 1.05 (d, 3H, *J* = 7.0 Hz, *i*Bu). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 175.8, 175.2 (C=O *i*Bu), 165.0 (C=O Bz), 152.3 (C-2), 151.0 (C-6), 150.3 (C-4), 142.4 (C-8), 133.4 (*ipso*-Bz), 132.9 (*p*-Bz), 128.8 (*m*-Bz), 128.0 (*o*-Bz), 124.3 (C-5), 88.3 (C-1'), 86.2 (C-4'), 73.0 (C-2'), 72.2 (C-3'), 62.4 (C-5'), 33.9, 33.6 (Ct *i*Bu), 19.0, 18.8, 18.8, 18.7 (Cp *i*Bu).

5'-*O*-(N^6-benzoyl-2',3'-di-*O*-iso-butyryladenosine)-2-cyanoethyl-*N*-*N*-diisopropylphosphoramidite (12)



N^6-benzoyl-2',3'-di-*O*-isobutyryladenosine (2.56 g, 5.0 mmol) was co-evaporated thrice in toluene and dissolved in dry DCM (20 mL, 0.3 M). DIPEA (2.28 mL, 12.5 mmol, 2.5 eq.) and 2-cyanoethyl-*N*,*N*-diis opropylchlorophosphoramidite (1.2 mL, 5.5 mmol, 1.1 eq.) were added and the reaction was stirred for 1 hour. The reaction was diluted with DCM washed with sat. aq. $NaHCO_3$ followed by brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (50% EtOAc in pentane + 1% TEA) yielded the title compound as a white foam (3.45 g, 4.84 mmol, 97%). Spectral data was in accordance with literary precedence[25]. **^{1}H NMR:** (500 MHz, $CDCl_3$) δ 9.26 (bs, 1H, NH), 8.84 – 8.74 (m, 1H, H-2), 8.57 – 8.50 (m, 1H, H-8), 8.07 – 7.97 (m, 2H, *o*-Bz), 7.60 (dt, 1H, *p*-Bz), 7.52 (dt, J = 9.0, 6.9, 1.7 Hz, 2H, *m*-Bz), 6.44 (dd, J = 22.9, 7.0 Hz, 1H, H-1'), 5.83 (ddd, J = 18.4, 7.0, 5.4 Hz, 1H, H-2'), 5.63 (ddd, J = 13.1, 5.4, 2.3 Hz, 1H, H-3'), 4.41 (p, J = 2.6 Hz, 1H, H-4'), 4.10 – 3.77 (m, 4H, H-5' + OCH_2CH_2CN), 3.74 – 3.55 (m, 2H, 2x CH N-*i*Pr), 2.76 – 2.63 (m, 3H, OCH_2CH_2CN + CH *i*Bu), 2.52 (pd, J = 7.0, 5.3 Hz, 1H, CH *i*Bu), 1.28 – 1.16 (m, 18H, 4x CH_3 N-*i*Pr + 2x CH_3 *i*Bu), 1.14 – 1.03 (m, 6H 2x CH_3 *i*Bu). **^{13}C NMR:** (126 MHz, $CDCl_3$) δ 176.0, 175.9, 175.5, 175.3 (C=O *i*Bu), 164.8 (C=O Bz), 153.0, 152.9 (C-6), 152.1, 152.0 (Cq Bz), 149.7, 141.2 (Cq-4), 141.3 (C-2, signal taken from HSQC) 133.8, 132.8 (CH *p*-Bz), 128.9, 128.9 (CH *m*-Bz), 127.9 (CH *o*-Bz), 123.1, 123.0 (C-5), 117.7, 117.6 (Cq OCE), 85.5, 85.1 (C-1'), 84.0, 83.9, 83.6, 83.5 (C-4'), 74.1, 74.0 (C-2'), 72.0, 71.7 (C-3'), 63.3, 63.1, 63.0, 62.9 (C-5'), 58.9, 58.7, 58.6 (OCH_2CH_2CN), 43.3, 43.2 (CH N-*i*Pr), 33.9, 33.9, 33.7, 33.6 (CH *i*Bu), 24.8, 24.8, 24.7, 24.7 (CH_3 N-*i*Pr), 20.5, 20.4, 20.4, 20.4 (OCH_2CH_2CN), 19.0, 19.0, 18.9, 18.8, 18.7 (CH_3 *i*Bu). **^{31}P NMR:** (202 MHz, $CDCl_3$): δ 149.6, 149.1.

General procedures for solid phase synthesis

Peptide synthesis

The intermediate peptides were synthesized using standard, Fmoc-based solid phase peptide synthesis utilizing (pre-loaded) TentaGel® S AC purchased from Rapp Polymer Gmbh. Coupling cycles were as followed: Fmoc deprotection: 2x2 minutes, 1x5 minutes treatment with 20% piperidine in DMF. Coupling: treatment of 6 eq. amino acid, 6 eq. HCTU (0.25 M in DMF) and 12 eq. DIPEA (1 M in DMF) for 30 minutes. Capping: 2x2 minutes treatment of the resin with a 10% Ac_2O solution in DMF and catalytic DIPEA. Washing between the steps was done with DMF. Ribosylated amino acid **5** was incorporated in the sequence by adding a solution of 2 eq. building block in a 0.25 M HCTU solution (2 eq.) in DMF and a 1 M DIPEA solution (4 eq.) in DMF to the resin in a fritted syringe. The resin was shaken overnight and thoroughly washed.

On-resin phosphorylation
The resin was treated with a sufficient amount of 1 M TBAF in THF (enough so that the entirety of the resin is submerged) for 30 minutes. The resin was thoroughly washed with DCM and DMF before the treatment was repeated once, furnishing the desilylated intermediate. The resin was then extensively washed with MeCN and flushed with nitrogen to remove traces of water before the resin was subjected to a solution of 5 eq. of $(FmO)_2PN(iPr)_2$ (0.25 M in MeCN) with 10 eq. ETT solution (0.25 M in MeCN). The resin was shaken for 30 minutes after which the resin was washed with MeCN. The resin was then treated with a sufficient amount of CSO solution (0.5 M in MeCN) for 30 minutes. The resin was then treated with a 10% DBU solution in DMF (2x 15 minutes) to furnish the crude, immobilized and deprotected phosphoribosylated peptide.

Construction of the pyrophosphate
The resin was extensively washed with MeCN and flushed with nitrogen to remove traces of water. The resin was then treated with a solution of compound **12** (3 eq., 0.3 M in MeCN) and ETT (6 eq., 0.25 M in MeCN) for 30 minutes. The resin was thoroughly washed with MeCN before a sufficient amount of CSO (0.5 M in MeCN) was added to the resin and shaken for 30 minutes.

Final deprotection and cleavage
The resin was then treated with a 10% DBU solution in DMF (2x 10 minutes) to remove the cyano ethyl protecting group. The resin was then treated with a 1 M TBAF solution in THF (2x 45 minutes) and washed with DMF followed by DCM. Final cleavage/deprotection occurred by treating the resin with a cleavage cocktail (2.5:2.5:95 TIS:TFA:DCM) for 1 hour. The crude products were collected by filtration and the resin was washed with a solution of 1:1:1 water:*t*BuOH:MeCN. The solvents were evaporated *in vacuo* and co-evaporated with a 1:1:1 water:*t*BuOH:MeCN solution.

Ac-Ile-Thr-Ser-Leu-Tyr(5-*O*-adenosine-diphosphate-α-D-ribosyl)-Lys-Lys-Ala-Gly-Ser-Thr-OH (11)
50 µmol TentaGel® S AC resin was loaded by treating the resin with 2.5 mL of a 0.2 M Fmoc-Thr(Trt)-OH solution (10 eq.) and DIC (77 µL, 0.5 mmol, 10 eq.) in DMF together with a catalytic amount of DMAP for 2 hours after which the general procedures described above were applied to the resin. The amino acids used were: Fmoc-Ile-OH, Fmoc-Thr(Trt)-OH), Fmoc-Ser(Trt)-OH, Fmoc-Leu-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Ala-OH, Fmoc-Gly-OH and **5**. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (3.98 mg, 2.28 µmol, 4.5%). **^{1}H NMR:** (500 MHz, D_2O) δ 8.35 (s, 1H, H-2 adenine), 8.10 (s, 1H, H-8 adenine), 6.88 (d, J = 8.5 Hz, 2H, Tyr arom.), 6.74 (d, J = 8.5 Hz, 2H, Tyr arom.), 5.97 (d, J = 5.5 Hz, 1H, H-1' adenosine), 5.38 (d, J = 4.5 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -11.1, -11.2, -11.3, -11.4. **LC-MS:** (0 -> 50% B in A) Rt = 5.87. **HRMS:** $[C_{69}H_{112}N_{18}O_{31}P_2 + 2H]^{2+}$ found: 876.3677, calculated: 876.3681.

Ac-Ile-Thr-Ser-Leu-Tyr(5-*O*-adenosine-diphosphate-α-D-ribosyl)-Lys-Lys-Ala-Gly-Ser-Gly-OH (15)
The general procedures described above were applied to 50 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were: Fmoc-Ile-OH, Fmoc-Thr(Trt)-OH), Fmoc-Ser(Trt)-OH, Fmoc-Leu-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Ala-OH, and **5** The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (12.4 mg, 7.13 µmol, 14%). **^{1}H NMR:** (500 MHz, D_2O) δ 8.43 (s, 1H, H-2 adenine), 8.18

(s, 1H, H-8 adenine), 6.96 (d, *J* = 8.6 Hz, 2H, H arom. Tyr), 6.83 (d, *J* = 8.7 Hz, 2H, H arom. Tyr.), 6.06 (d, *J* = 5.4 Hz, 1H, H-1' adenosine), 5.47 (d, *J* = 4.2 Hz, 1H, H-1' ribosyl). **^{31}P NMR:** (202 MHz, D_2O) δ -10.4, -10.5, -10.7, -10.8. **LC-MS:** (0 -> 50% B in A) Rt = 5.01. **HRMS:** $[C_{67}H_{108}N_{18}O_{30}P_2 + 2H]^{2+}$ found: 854.3545, calculated: 854.3550.

Biochemical evaluation

Expression plasmids and protein purification

The construction of the expression plasmids and the purification procedures were described earlier.[26–28] Briefly, expression plasmids were transferred into Rossetta (DE3) cells and grown to an OD_{600} of 0.6 in LB medium supplemented with appropriate antibiotics. For metal-coordinating proteins the medium was further enriched either by addition of 2 mM $MgSO_4$ (ARHs) or 100 µM $ZnCl_2$ (*Spy*MacroD). Expression was induced with 0.4 mM isopropyl β-D-1-thioglactopyranoside (IPTG) and cultured were allowed to grow further overnight at 17 °C. Cultures were harvested by centrifugation, pellets resuspended in lysis buffer (50 mM TrisHCl [pH 8], 500 mM NaCl and 25 mM imidazole) and stored at -20 °C. Proteins were purified by Ni^{2+}-NTA chromatography (Jena Bioscience) according to the manufacturer's protocol using the following buffers: all buffers contained 50 mM TrisHCl (pH 8) and 500 mM NaCl; additionally, the lysis buffer contained 25 mM, the washing buffer 40 mM, and the elution buffer 500 mM imidazole. Proteins were dialyzed overnight against 50 mM TrisHCl (pH 8), 200 mM NaCl, 1 mM dithiothreitol and 5% (v/v) glycerol and stored at -80 °C. For the purification of ARH and ARH-like proteins all purification buffers were additionally supplemented with 10 mM $MgCl_2$.

(ADP-ribosyl)-hydrolase activity assay

The peptide demodification assay was described earlier[18]. Briefly, peptide concentration for the assay was estimated using absorbance at I_{260nm} using the molar extinction coefficient of ADP-ribose (15,400 M^{-1} cm^{-1}). 20 µM indicated peptide were demodified by incubation with 1 µM hydrolase for 45 minutes at 30 °C in assay buffer (50 mM TrisHCl [pH 8], 200 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol and 0.2 µM human NUDT5[29]). Reactions were stopped and analyzed by performing the AMP-Glo™ assay (Promega) according to the manufacturer's protocol. Luminescence was recorded on a SpectraMax M5 plate reader (Molecular Devices) and data analyzed with GraphPad Prism 7. Control reactions were carried out in absence of peptide.

References

[1] O. Leidecker, J. J. Bonfiglio, T. Colby, Q. Zhang, I. Atanassov, R. Zaja, L. Palazzo, A. Stockum, I. Ahel, I. Matic, *Nat. Chem. Biol.* **2016**, *12*, 998–1000.

[2] R. Martello, M. Leutert, S. Jungmichel, V. Bilan, S. C. Larsen, C. Young, M. O. Hottiger, M. L. Nielsen, *Nat. Commun.* **2016**, *7*, 1–13.

[3] V. Bilan, M. Leutert, P. Nanni, C. Panse, M. O. Hottiger, *Anal. Chem.* **2017**, *89*, 1523–1530.

[4] D. M. Leslie Pedrioli, M. Leutert, V. Bilan, K. Nowak, K. Gunasekera, E. Ferrari, R. Imhof, L. Malmström, M. O. Hottiger, *EMBO Rep.* **2018**, *19*, 1–11.

[5] S. Buch-Larsen, I. Hendriks, J. Lodge, M. Rykær, B. Furtwängler, E. Shishkova, M. Westphall, J. Coon, M. Nielsen, *Cell Rep.* **2020**, *32*, 108176.

[6] J. J. Bonfiglio, P. Fontana, Q. Zhang, T. Colby, I. Gibbs-Seymour, I. Atanassov, E. Bartlett, R. Zaja, I. Ahel, I. Matic, *Mol. Cell* **2017**, *0*, 1–9.

[7] L. Palazzo, O. Leidecker, E. Prokhorova, H. Dauben, I. Matic, I. Ahel, *Elife* **2018**, *7*, 1–12.

[8] T. Hunter, B. M. Sefton, *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77*, 1311–1315.

[9] J. V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, M. Mann, *Cell* **2006**, *127*, 635–648.

[10] J. Villén, S. A. Beausoleil, S. A. Gerber, S. P. Gygi, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 1488–1493.

[11] S. C. Larsen, I. A. Hendriks, D. Lyon, L. J. Jensen, M. L. Nielsen, *Cell Rep.* **2018**, *24*, 2493-2505.e4.

[12] T. Hunter, *Curr. Opin. Cell Biol.* **2009**, *21*, 140–146.

[13] K. Sharma, R. C. J. D'Souza, S. Tyanova, C. Schaab, J. R. Wiśniewski, J. Cox, M. Mann, *Cell Rep.* **2014**, *8*, 1583–1594.

[14] E. Bartlett, J. J. Bonfiglio, E. Prokhorova, T. Colby, F. Zobel, I. Ahel, I. Matic, *Cell Rep.* **2018**, *24*, 3488–3502.

[15] I. Gibbs-Seymour, P. Fontana, J. G. M. Rack, I. Ahel, *Mol. Cell* **2016**, *62*, 432–442.

[16] S. Vyas, I. Matic, L. Uchima, J. Rood, R. Zaja, R. T. Hay, I. Ahel, P. Chang, *Nat. Commun.* **2014**, *5*, 4426.

[17] I. Carter-O'Connell, A. Vermehren-Schmaedick, H. Jin, R. K. Morgan, L. L. David, M. S. Cohen, *ACS Chem. Biol.* **2018**, *13*, 2841–2848.

[18] J. Voorneveld, J. G. M. Rack, I. Ahel, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2018**, *20*, 4140–4143.

[19] H. A. V Kistemaker, A. P. Nardozza, H. S. Overkleeft, G. A. van der Marel, A. G. Ladurner, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10634–10638.

[20] Y. Ando, E. Elkayam, R. L. McPherson, M. Dasovich, S.-J. Cheng, J. Voorneveld, D. V. Filippov, S.-E. Ong, L. Joshua-Tor, A. K. L. Leung, *Mol. Cell* **2019**, *73*, DOI 10.1016/j.molcel.2018.12.022.

[21] H. A. V Kistemaker, G. J. van der Heden van Noort, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2013**, *15*, 2306–2309.

[22] J. G. M. Rack, L. Palazzo, I. Ahel, *Genes Dev.* **2020**, *34*, 263–284.

[23] J. Moss, S. chen Tsai, R. Adamik, H. chia Chen, S. J. Stanley, *Biochemistry* **1988**, *27*, 5819–5823.

[24] J. G. M. Rack, A. Ariza, B. S. Drown, C. Henfrey, E. Bartlett, T. Shirai, P. J. Hergenrother, I. Ahel, *Cell Chem. Biol.* **2018**, *25*, 1533-1546.e12.

[25] G. J. van der Heden van Noort, M. G. van der Horst, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *J. Am. Chem. Soc.* **2010**, *132*, 5236–5240.

[26] P. Fontana, J. J. Bonfiglio, L. Palazzo, E. Bartlett, I. Matic, I. Ahel, *Elife* **2017**, *6*, 1–20.

[27] J. G. M. Rack, R. Morra, E. Barkauskaite, R. Kraehenbuehl, A. Ariza, Y. Qu, M. Ortmayer, O. Leidecker, D. R. Cameron, I. Matic, A. Y. Peleg, D. Leys, A. Traven, I. Ahel, *Mol. Cell* **2015**, *59*, 309–320.

[28] D. Slade, M. S. Dunstan, E. Barkauskaite, R. Weston, P. Lafite, N. Dixon, M. Ahel, D. Leys, I. Ahel, *Nature* **2011**, *477*, 616–622.

[29] L. Palazzo, B. Thomas, A. S. Jemth, T. Colby, O. Leidecker, K. L. H. Feijs, R. Zaja, O. Loseva, J. C. Vert, I. Matic, T. Helleday, I. Ahel, *Biochem. J.* **2015**, *468*, 293–301.

Chapter 5

A Methodology for the Preparation of Peptides ADP-Ribosylated on Arginine and Arg-ADPr Modified Human Ubiquitin

Introduction

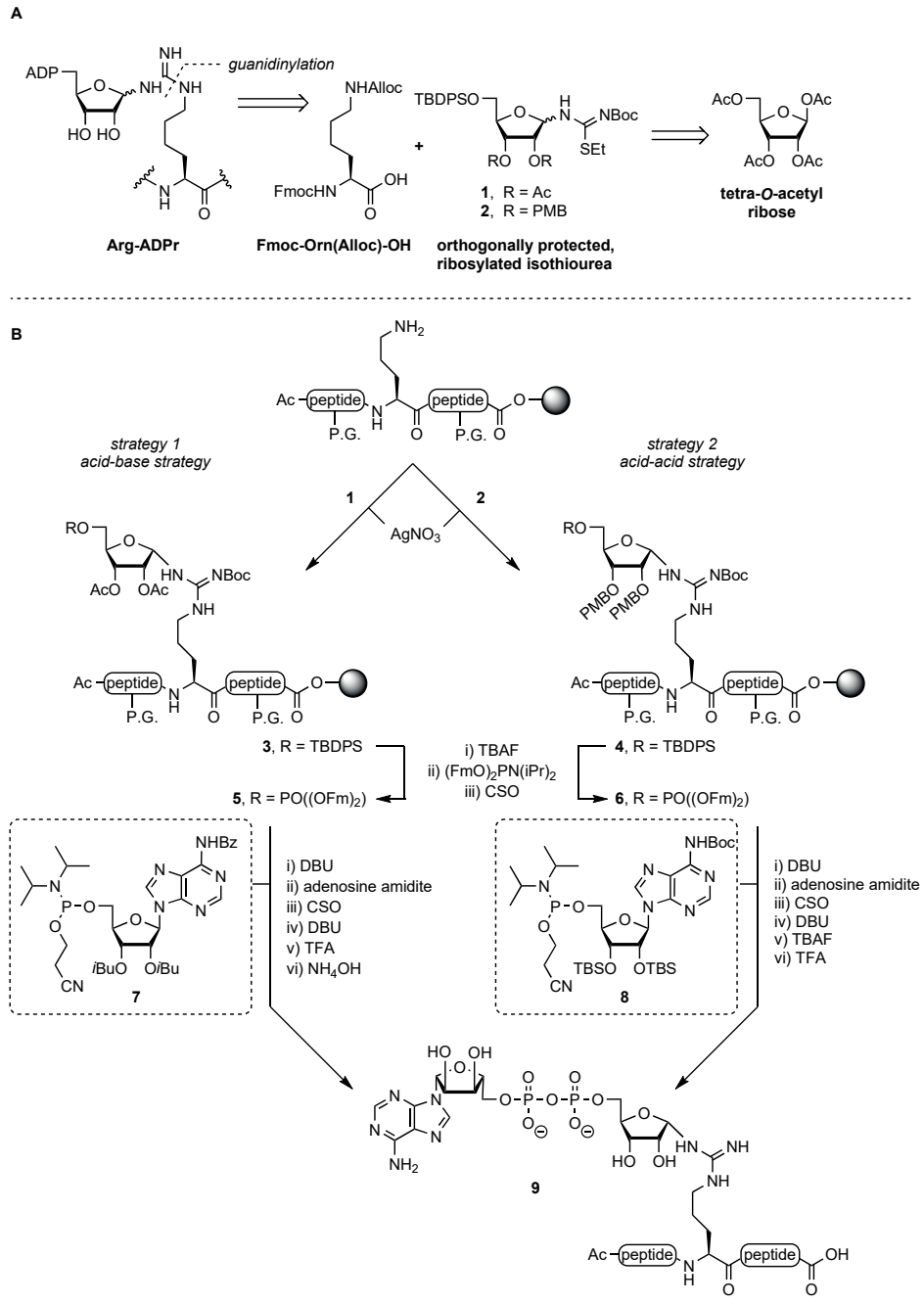
ADP-ribosyl-arginine, the first discovered ADP-ribosylated amino acid, was encountered in *E.Coli* after infection with T4 phages.[1] This finding encouraged a search for identifying more ADP-ribosylated arginine (Arg-ADPr) in proteins originating from both lower and higher organisms. For instance, site-directed mutagenesis of proteins led to the uncovering of several interesting targets to be specifically ADP-ribosylated on their arginine (Arg) residue.[2] However, the mutagenesis studies do not allow for a screening of a wide variety of biological systems for modification sites. Another approach for identifying Arg-ADP-ribosylation sites comprises the generation of antibodies specific for the Arg-ADPr modification.[3–6] Unfortunately, non-specific recognition of the generated antibodies may hinder the proper detection of Arg ADP-ribosylation.[7]

After the initial discovery of Arg-ADPr in 1974, many insights regarding its occurrence and physiological role in various organisms have been acquired.[7] The formation of Arg-ADPr in target proteins is mostly catalyzed by ADP-ribosyl transferases of the cholera toxin-like subfamily (ARTCs) which is characterized by the R-S-E catalytic triad.[8] Specificity of these ARTCs towards guanidinyl groups (Arg and agmatine) is found if the triad is extended with a glutamic acid, leading to an RSE-X-E recognition motif.[9–11] Although ARTCs are extracellular enzymes exclusively,[12] a significant amount of proteins ADP-ribosylated on Arg-residues, are found in the Golgi apparatus or the ER.[13,14] This finding is supported by the fact that the hydrolase ARH1, responsible for the reversal of Arg-ADPr,[15] is localized in the cytoplasm.[12] Despite the information acquired over the past decades, it is still unclear which enzyme is responsible for intracellular Arg ADP-ribosylation. One possibility is that an ARTC ADP-ribosylates Arg whilst being transported to the membrane.[16] Otherwise, an ARTD (ARTs of the subfamily diphtheria toxin-like) that permanently resides inside the cell is likely to be responsible for intracellular Arg ADP-ribosylation. For example ARTD10 which is able to catalyze the formation of Arg-ADPr, albeit that Arg does not appear to be its main target.[17] Up to now, no specific readers that recognize the Arg-ADPr modification are known. Furthermore, the chemical stability of the Arg-ADPr PTM is not properly studied. Knowledge about chemical stability is crucial as some proteomic conditions lead to biased read-outs, intentional[18] or not.[19] This chapter aims for the development of a strategy that allows the preparation of synthetic peptides, ADP-ribosylated on Arg residues as they can provide for useful molecular tools in the research of Arg ADP-ribosylation. To evaluate the ability of the obtained synthetic peptides to be recognized in a biological system a biochemical assay can be performed with DupA (an enzyme encountered in *Legionella pneumophila*) that specifically hydrolyzes the ADP-ribosylated site of Arg42 in human ubiquitin (hUb).

Synthetic strategy

All strategies for the synthesis of peptides, *mono*-ADP-ribosylated (MARylated) at a Ser- (Chapters 2 and 3), Thr- and Cys- (Chapter 3), or Tyr- (Chapter 4) residue relied on the stereoselective coupling of a ribosyl donor and an amino acid acceptor. However, introduction of a guanidinyl group at the anomeric centre of ribose, the linkage by which Arg is bound to ADPr, is usually not achieved via a direct glycosylation procedure. The ubiquity of the guanidine function in bio-molecules and the variety of natural human-made biopolymers has led to the availability of various guanidinylation reagents that also have been applied in the (glyco)peptide chemistry.[20–23] With respect to the synthesis of glycopeptides, two approaches can be discerned. Both strategies utilize the coupling of an *S*-alkyl-isothiourea, situated at the anomeric centre of the desired monosaccharide, with the side chain amine of ornithine (Orn) to synthesise the glycosyl-Arg linkage. In the first approach, the glycosyl-Arg building block is first obtained and then incorporated into a peptide sequence of choice via SPPS.[24,25] In the second approach, the SPPS of a specific peptide with an orthogonally protected Orn at a predetermined position is followed by the selective release of the masked amino group in its side chain. Subsequent on-resin coupling of the free amino group with the S-alkyl-isothiourea functionality yields the prospected Arg-modified glycopeptide.[26,27]

The approach designed to synthesize Arg-ADPr uses the second approach as starting point in which isothiourea function at the anomeric centre of ribose (**1** or **2**, Scheme 1A) is coupled on-resin to the primary amine of an ornithine residue incorporated in an immobilized peptide (Scheme 1A and B). A commercially available ornithine building block was chosen, the Alloc protecting group of which can be selectively removed on-resin relative to acid-labile side chain protecting groups and the TentaGel® S AC linker. Besides, two strategies differing in protecting group patterns on the adenosine and ribosyl moiety have been investigated (Scheme 1B). On the basis of the construction of Tyr-ADPr peptides (See Chapter 4), the first strategy (termed hybrid or more aptly 'acid-base' strategy) uses acetyl groups for the 2' and 3' hydroxyls in the ribosyl moiety and isobutyryl and benzoyl groups for adenosine (Scheme 1B). The side chains of the applied amino acid building blocks are protected with acid labile groups i.e., Trt for Ser- and Thr-residues and Mtt for lysine (Lys), as these are cleavable in low concentrations of TFA. As shown in Scheme 1B, the hybrid strategy entails the silver-promoted coupling of isothiourea **1** with the released amine of the ornithine in the immobilized peptide to give resin **3**. Next, replacement of the silyl-ether by an Fm protected phosphotriester, is followed by DBU mediated deprotection of the phosphate in **5**. With the thus obtained phosphate monoester, the pyrophosphate can be introduced via P^{III} - P^{V} coupling procedure with adenosine phosphoramidite **7** bearing isobutyryl and benzoyl as protecting groups. At the end of the synthesis the immobilized Arg-ADPr peptide is treated with TFA to remove the acid labile protecting groups with



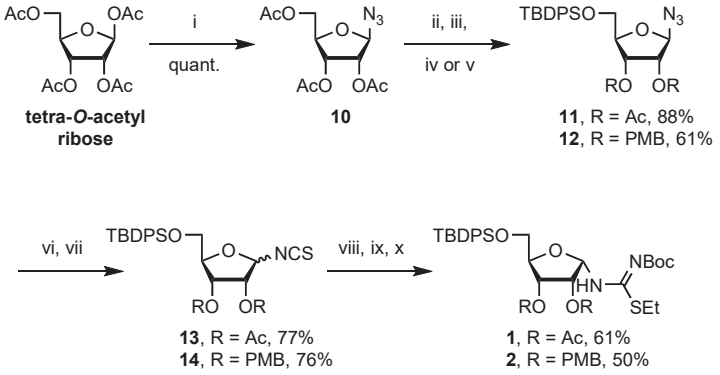
Scheme 1. A) Retrosynthetic analysis for the synthesis of Arg-ADPr via an on-resin guanidinylation approach. **B)** Proposed synthetic strategies to synthesize peptides, MARylated at their Arg-residue. P.G. = protecting group for amino acid side chains i.e., Trt for Ser/Thr/His and Mtt for Lys.

concomitant cleavage from the solid support. Finally, treatment with NH_4OH will result in deprotection of all ester functionalities to obtain MARylated Arg-ADPr peptide **9**. The second, termed here 'acid-acid' strategy, is described in Chapter 3 and circumvents the use of base labile protecting groups by the use of the PMB group for the 2' and 3' hydroxyls in the ribosyl isothiourea **2** (Scheme 1A). The on-resin guanidinylation of **2** with an ornithine residue will result in intermediate peptide **4**. The phosphate moiety is installed in a similar manner as for the hybrid strategy to furnish immobilized phosphotriester **6**. In the next step adenosine phosphoramidite **8**, provided with a Boc-group on the exocyclic amine and silyl ethers on the 2' and 3' hydroxyl, is used to for the introduction of the ADP moiety via the same P^{III} - P^{V} coupling procedure. After removal of the silyl groups, global deprotection and cleavage from the resin is accomplished by treatment with 10% TFA in DCM solution to furnish MARylated Arg-ADPr peptide **9**.

Results and discussion

Building block synthesis

Synthesis of the isothiocyanates **1** and **2** (Scheme 2) commenced with the treatment of the commercially available β-ribofuranose tetra-acetate with $SnCl_4$ as a Lewis acid[28] and TMS-N_3 as an azide donor to give azide **10** in quantitative yield. As described in Chapters 3 and 4, the primary acetate was replaced by the orthogonal TBDPS group by a three-step procedure, comprising deacetylation, silylation and acetylation to provide **11** in 88% overall yield.[29] By replacing the acetylation in this three-step procedure by methoxybenzylation, azide **10** was converted into **12** having the PMB group at the 2'- and 3'-positions with a yield of 61%. Next, the anomeric azides in **11** and **12** were reduced using Adam's catalyst and H_2. Attempts to work up the reaction proved cumbersome as the resulting amine is highly labile and concentration *in vacuo* led to full degradation of the product. Therefore, after filtration over a pad of Celite to rid the reaction mixture of the catalyst, the filtrate was directly treated with thiophosgene to furnish ribosyl isothiocyanates **13** and **14** as anomeric mixtures in 77% and 76% yield respectively. Although Arg-ADPr is known to undergo spontaneous anomerization to its β-anomer after initially being installed α-selectively,[7] the resulting anomers of the isothiocyanates **13** and **14** could easily be separated by column chromatography to obtain the individual anomers in a homogenous fashion. Next, the α-anomer of both **13** and **14** was subjected to ammonolysis by ammonia in THF, followed by protection of the amine in the resulting thiourea with the Boc group and finally treatment with ethyl iodide to furnish building blocks **1** (Ac protected) and **2** (PMB protected), ready for the on-resin, silver-promoted coupling to ornithine.

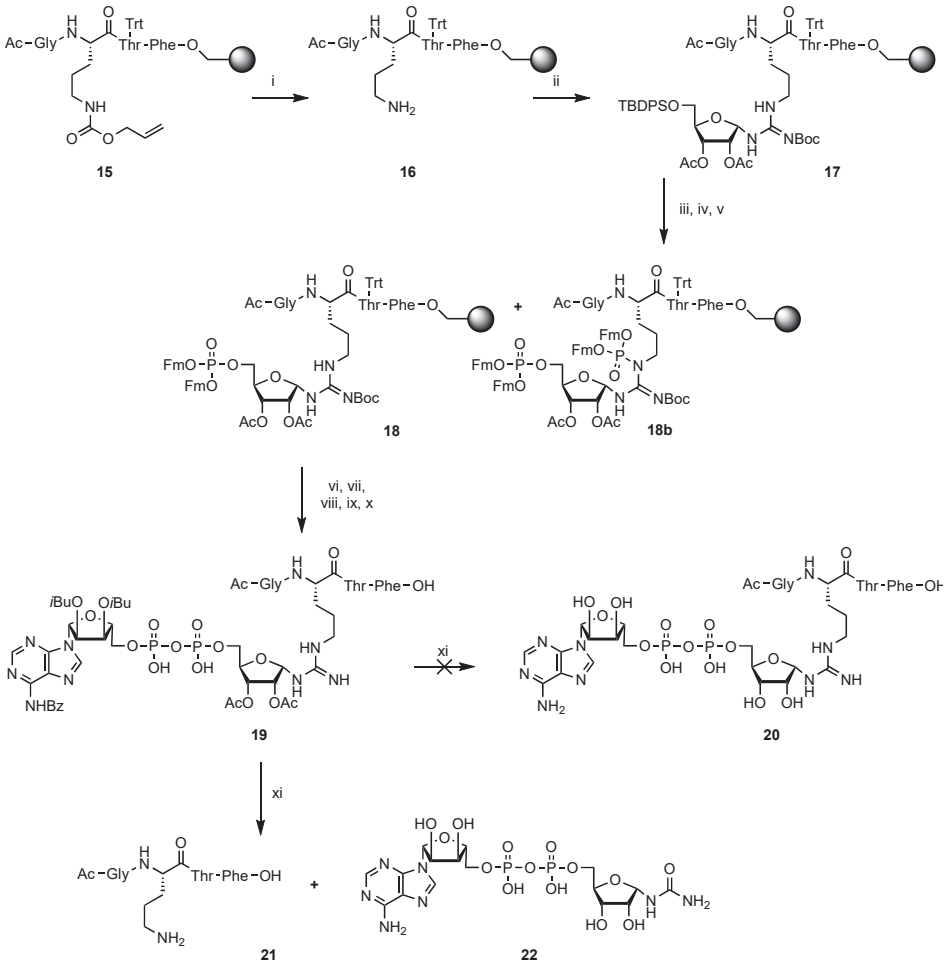


Scheme 2. Synthesis of building blocks **1** and **2**. Reagents and conditions: i) TMS-N_3, $SnCl_4$, DCM. ii) Na, MeOH, 0 °C. iii) TBDPS-Cl, pyr. iv) Ac_2O, DMAP, pyr. v) NaH, PMB-Cl, DMF, 0 °C. vi) PtO_2, H_2, EtOAc. vii) $SCCl_2$, K_2CO_3, water, EtOAc. viii) NH_3, THF. ix) Boc_2O, DMAP, DCM. x) EtI, K_2CO_3, MeCN.

Peptide synthesis

With the appropriate building blocks in hand, the on-resin synthesis of MARylated Arg-ADPr peptide was undertaken. Firstly, the hybrid strategy, using a combination of acid and base labile protecting groups for key building blocks **1** and **7** as outlined in Scheme 3 was investigated. Model tetrapeptide **15** was prepared by standard SPPS with incorporation of Fmoc-Orn(Alloc)-OH on the projected MARylation site. The selective, on-resin removal of the Alloc group with $Pd(PPh_3)_4$ and DMBA as a scavenger in DCM to furnish the primary amine. LC-MS monitoring of this reaction is imperative since, depending on the reagent quality, multiple treatments (up to 5 times) may be required to fully deprotect the ornithine residue. Degassing of the solvent and administering of the reaction solution to the resin in a nitrogen atmosphere did improve the efficiency of this deprotection. After LC-MS analysis showed full deprotection, immobilized peptide **16** was guanidinylated with isothiocyanate **1** using $AgNO_3$ as the Lewis acid. LC-MS analysis of the crude products after treatment with 10% TFA (Figure 1A) showed the lack of starting peptide derived from **16** and the formation of partially protected peptides (with and without the Boc group) derived from **17**. It is noteworthy that during cleavage and deprotection from the resin for a period of 2 hours spontaneous anomerization may occur as analyzed by LC-MS. Next, resin **17** was converted to immobilized phosphotriester **18** by desilylation, phosphitylation of the released alcohol with the bis-Fm-protected phosphoramidite reagent and oxidation of the intermediate phosphite triester. Treatment of resin **18** with 10% TFA and ensuing LC-MS analysis of the released crude product showed partially protected peptides (also with and without the Boc group) derived from **18** and major side products with masses that correspond with one additional Fm-protected phosphotriester (**18b**, Scheme 3, Figure 1B). The formation of **18b** can be explained by the residual nucleophilicity of the urethane protected guanidyl

function. Next, the ADP moiety is appended by deprotection of the phosphotriester in resin **18** (in the presence of **18b**) with a 10% DBU solution and subjecting the formed phosphate monoester to the P^{III} – P^{V} coupling procedure using *i*Bu/Bz protected adenosine amidite **7.** Oxidation was performed with CSO and deprotection of the cyanoethyl group was mediated by DBU. A sample of the resin was taken and treated with 10% TFA. LC-MS analysis of the crude products revealed a complex mixture (Figure 1C) in which peptides derived from **19** (with and without the Boc group) and peptides derived from **18** and **18b** (which was still partially protected with one Fm group) could be assigned. Also, some unreacted but fully deprotected phosphomonoester and the corresponding phosphoramidate were detected, co-eluting with a trityl cation explaining the intense UV-absorption in the LC-MS trace with a retention time of around 4 minutes. Quantification of the obtained products by UV-absorption proved to be cumbersome as products and side products are not base-separated. Therefore, the resin was subjected to a longer TFA treatment to ensure that all Boc protecting groups were removed and the crude product was treated overnight with 28 wt% NH_4OH in water and subsequently analyzed by LC-MS chromatography (Figure 1D). The mass of target Arg-ADPr peptide **20** was not observed, but a major product was detected with the same mass as tetrapeptide **21.** The possible hydrolysis of the guanidine function in **19** leading to **21** is supported by the finding of a similar hydrolysis of Arg-ADP-ribosylated HNP-1 when incubated at pH 9.[30] Moreover hydrolysis of guanidine function in **19** should result in the formation of ADP-ribosyl urea **22**, which indeed is highly abundant in the crude mixture. Due to the lability of the guanidinyl function towards the deprotection conditions of the ester protecting groups, the acid-base strategy was deemed not viable for the synthesis of MARylated Arg-ADPr peptides.



Scheme 3. Attempted synthesis of target tetrapeptide Ac-Gly-Arg^{ADPr}-Thr-Phe-OH **20** using the acid-base strategy. Reagents and conditions: i) $Pd(PPh_3)_4$, DMBA, DCM. ii) **1**, $AgNO_3$, TEA, DMF. iii) TBAF, THF. iv) $(FmO)_2PN(iPr)_2$, ETT, MeCN. v) CSO, MeCN. vi) DBU, DMF. vii) **7**, ETT MeCN. viii) CSO, MeCN. ix) DBU, DMF. x) TFA, DCM. xi) NH_4OH, H_2O.

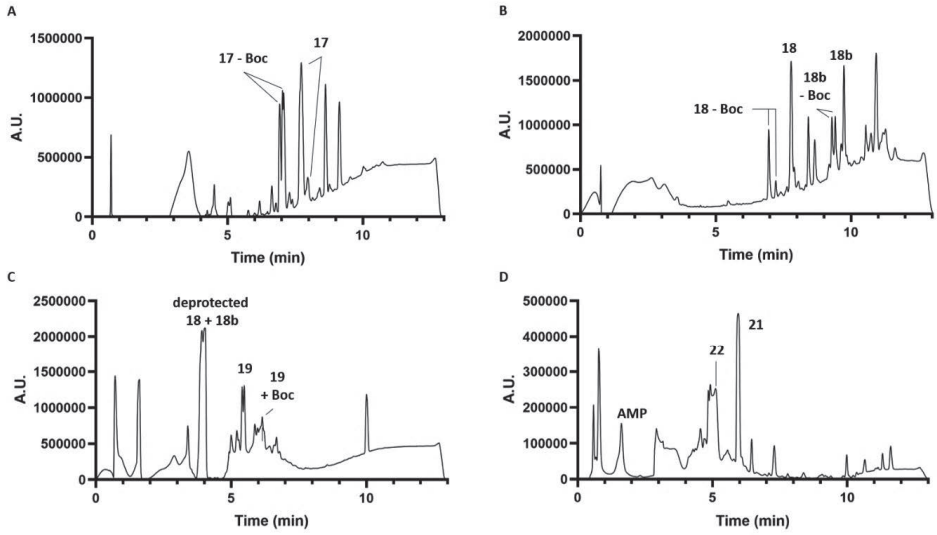
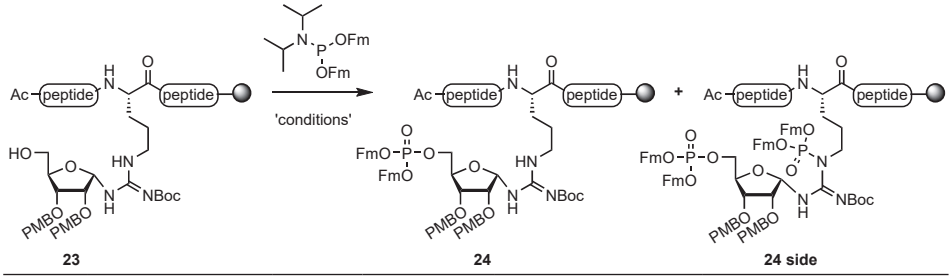


Figure 1. LC-MS traces of the crude intermediates *en route* to peptide **20** via the acid-base strategy as presented in Scheme 3. Resin samples were treated with 10% TFA for 2 hours prior to co-evaporation in a 1:1:1 *t*BuOH:H_2O:MeCN mixture. A) LC-MS analysis of intermediate peptide **17** after guanidinylation of **16** with **1**. B) LC-MS analysis after desilylation of **17** and subsequent phosphitylation with $(FmO)_2PN(iPr)_2$ and oxidation with CSO. C) LC-MS analysis after deprotection of **18** with DBU, P^{III} – P^{V} coupling with adenosine amidite **7** and subsequent treatment with DBU to remove the cyano-ethyl protecting group. D) LC-MS analysis after treatment of the crude product of the TFA cleavage of Figure 1C overnight with NH_4OH.

Therefore, attention was turned towards the acid-acid strategy, characterized by the presence of solely acid labile groups in the protected precursor of the target Arg-ADPr peptide at the final stage of the synthesis. As outlined in Scheme 4, the assembly of Arg-ADPr peptide Ac-Gly-Arg^{ADPr}-Thr-Phe-OH (**20**) started with the SPPS of immobilized peptide **16** as in the hybrid acid-base strategy. Guanidinylation of the amino group in the side chain of ornithine in **16** with isothiourea **2**, of which the cis-diol is protected with PMB ethers, was followed by removal of the primary TBDPS group yielding intermediate peptide **23**. The subsequent introduction of the phosphotriester was investigated more closely to suppress the phosphoramidate formation at the guanidyl function. A set of six reaction conditions was screened varying in activator (ETT, tetrazole and DCI) and equivalents of phosphorylating reagent (2.5 and 5.0 equivalents). The results of this screening are shown in Table 1. Firstly, the same reaction conditions as described above (Scheme 3) led mainly to the formation of phosphoramidate side product (**24 side**, 64%) compared to desired product **24** (36%, entry 1). Reducing the amount of phosphorylating reagent from 5.0 equivalents to 2.5 equivalents did suppress the side product formation in favor of formation of **24** (60% **24** and 40% side product, entry 2). Application of tetrazole

as activator was no improvement as 5.0 equivalents of phosphoramidite led to significant amounts of side product (35% **24**, 65% side product, entry 3) and by reducing the amount of reagent to 2.5 equivalent a ratio of 60% product vs. 40% side product was attained again (entry 4). When DCI was tried with 5.0 equivalents of the phosphitylating reagent, the ratio of product to side product did improve but still showed a slight preference for the side product which was formed in 56% (entry 5). Fortunately, when 2.5 equivalents of the phosphoramidite was used, the amount of product **24** being formed increased to 84% (entry 6).

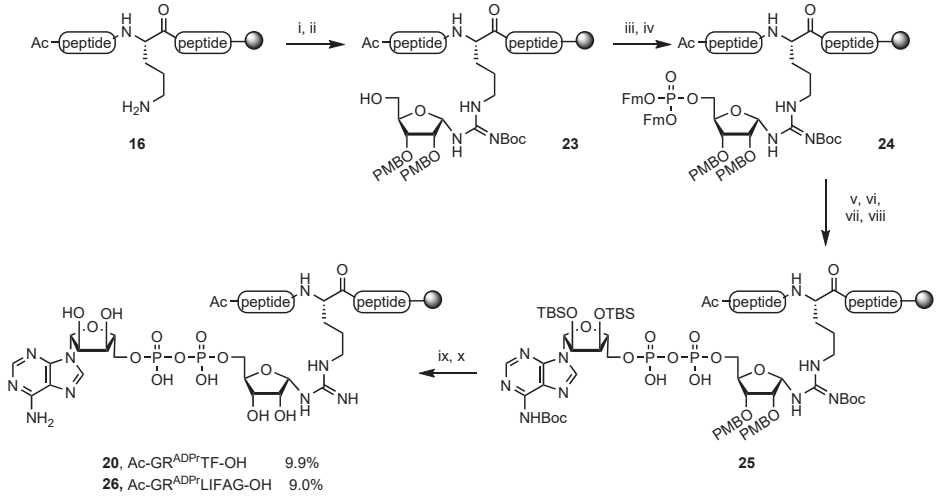
Table 1. Screening of the phosphitylation conditions for immobilized peptide **23** to give **24** and to suppress **24 side**. Product ratios were calculated by integration of peak area of UV absorptions in their LC-MS trace. Every reaction was carried out on a 10 µmol scale with resin **23**. The equivalents in the table are the equivalents of bis-Fm-protected amidite relative to **23**. The concentration is calculated relative to the phosphorylating reagent, 2.0 equivalents of activator was used relative to the phosphoramidite and all reactions were carried out for 30 minutes. After the phosphitylation, all reactions were treated with CSO for 30 minutes to oxidize the phosphite intermediates. The resin was then treated with 10% TFA and 2.5% TIS in DCM for 120 minutes to cleave the peptide from the resin. From the crude reaction mixture, a sample was prepared for LC-MS analysis.



Entry	Activator	Eq	C (mM)	Product 24 (%)	Side-product 24 (%)
1	ETT	5	88	36	64
2	ETT	2.5	44	60	40
3	Tetrazole	5	88	35	65
4	Tetrazole	2.5	44	60	40
5	DCI	5	88	44	56
6	DCI	2.5	44	84	16

Subsequently, the Fm protecting groups in **24**, obtained by applying the latter optimized phosphitylation conditions, were removed with DBU to give the phosphomonoester. The P^{III} - P^{V} coupling procedure to install the pyrophosphate started with the coupling of phosphomonoester was with TBS/Boc protected adenosine amidite **8** (Scheme 1B) using the same, optimized conditions (Table 1 entry 6) to avoid phosphoramidate formation.

Subsequent oxidation with CSO and removal of the cyanoethyl group in the pyrophosphate moiety led to protected and immobilized Arg-ADP-ribosylated target peptide **25**. The silyl ethers on the adenosine moiety were removed by treatment of the resin with a 1 M TBAF solution, as described in Chapter 3. Finally, the peptide was cleaved from the resin with concomitant loss of all remaining protecting groups (e.g., Boc, PMB and Trt) by administration of the resin to a 10% TFA in DCM solution. HPLC purification of the crude mixture led to the isolation of the first synthetic mono-Arg-ADPr linked peptide **20** in 9.9% overall yield. Encouraged by this result the same method was adopted for the synthesis of heptapeptide **26** (Scheme 4), sequence of which is derived from human hUb and known to be ADP-ribosylated on its Arg residue (Arg42 in human Ub). After prep-HPLC purification, peptide **26** was obtained in 9.0% yield.



Scheme 4. Synthesis of Arg-ADP-ribosylated peptides **20** and **26** using the acid-labile strategy. Reagents and conditions: i) **2**, $AgNO_3$, TEA, DMF. ii) TBAF, THF. iii) $(FmO)_2PN(iPr)_2$, DCI, MeCN. iv) CSO, MeCN. v) DBU, DMF. vi) **8**, DCI, MeCN. vii) CSO, MeCN. viii) DBU, DMF. ix) TBAF, THF. x) TFA, DCM.

A next challenge for the newly developed acid-acid strategy to obtain Arg-ADPr peptides comprised the synthesis of the full-length hUb with the ADPr modification on its Arg-42 residue. Therefore, a Ub mutant where Arg42 was replaced by an Alloc-protected ornithine residue was synthesized on a 6 µmol scale using a well-established SPPS approach.[31] Next, the installation of the Arg-ADPr moiety started with the removal of the orthogonal Alloc-protecting group on the resin-bound synthetic Ub. The efficiency of the subsequent coupling of the released amine in ornithine with ribosyl building block **2** was controlled by the following procedure. Cleavage of a test sample from the resin was performed and

LC-MS analysis of the crude product after coupling showed a clean LC-MS trace with the mass corresponding to the Ub-ribosylated intermediate. The protocol to synthesize Ub-ADPr was continued with slight changes as for the purpose of biochemical assays the synthetic Ub was equipped with a biotin tag on the *C*-terminus. Oxidation of the phosphite intermediate species was therefore performed using a 0.5 M *t*BuOOH solution in MeCN to avoid overoxidation of the biotin-tag as described in Chapter 3. Also, human Ub carries three additional Arg residues protected with the Pbf group in the SPPS protocol. These protecting groups require prolonged reaction times in high concentrations of TFA in order to be fully removed, typically 95% TFA for 2 hours. Therefore, a test cleavage, using 95% TFA was performed. Strikingly, the glycosidic guanidine moiety as well as the pyrophosphate bridge remained intact during a 1.5-hour treatment with 95% TFA which was sufficient to fully remove all protecting groups including the Pbf-groups. The remainder of the resin-bound Ub-ADPr was fully deprotected according to the renewed cleavage/deprotection protocol and HPLC purification yielded fully synthetic, human Ub, MARylated on its Arg42 residue in an overall yield of 0.5%.

Biochemical evaluation of synthetic Arg-ADPr peptides

Peptides **20** and **26** bearing the native ADPr-Arg linkage were assessed in their ability to be properly recognized in biologically relevant systems. Therefore, their ability to function as a substrate for DupA, an enzyme originating from *Legionella pneumophila*, was evaluated. In short, *Legionella p.* has multidomain containing SidE ligases which can ADP-ribosylate the Arg42 residue on human Ub with their MARylating domain. Subsequent translocation of the ADP-ribosylated Ub to their phosphodiesterase (PDE) domain allow the ligation of ubiquitin to a serine residue of host substrate proteins. The substrate protein is interconnected with ubiquitin via the phosphoribosyl moiety originating from ADPr and this process is an important factor in the bacterial infection, proliferation and host immune evasion.[32] Legionella carefully controls its host protein ubiquitination activity by the deployment of Dup enzymes, DupA and DupB, which release the ubiquitinated substrate.[33] DupA is also known to accept and hydrolyze the pyrophosphate bridge in Ub-ADPr with concomitant release of AMP, making DupA ideally suited for a hydrolytic assay with synthetic peptides **20** and **26**.

Tetramer **20** with a random sequence and heptamer **26**, a sequence representing the ADP-ribosylated Arg42 residue targeted by DupA, were compared in a DupA mediated hydrolysis assay (Figure 2) with enzymatically obtained Ub, ADP-ribosylated on Arg42[34] as a reference. The peptides were incubated in the presence or absence of DupA under buffered conditions and at indicated times analyzed using mass-spectrometry. Incubation of 5 µM 'nonsense' tetramer **20** and 3 µM DupA did not result in DupA mediated cleavage of the pyrophosphate bond. This indicates the ADP-ribosylated tetramer is not recognized

and processed by the enzyme. Incubation of 5 µM peptide **26** and 3 µM DupA resulted in 40% hydrolysis after 90 minutes. The observed hydrolysis of the pyrophosphate bond can be attributed to the catalytic activity of DupA. This confirms that, although slower than full-length Ub-ADPr, peptide **26** is recognized and processed by DupA. However, since **26** is a smaller fragment of Ub-ADPr as compared to native full-length Ub-ADPr (3 µM Ub-ADPr, 1 µM DupA, 100% hydrolyzed) it is likely that the full-length Ub-ADPr is recognized more efficiently by DupA.

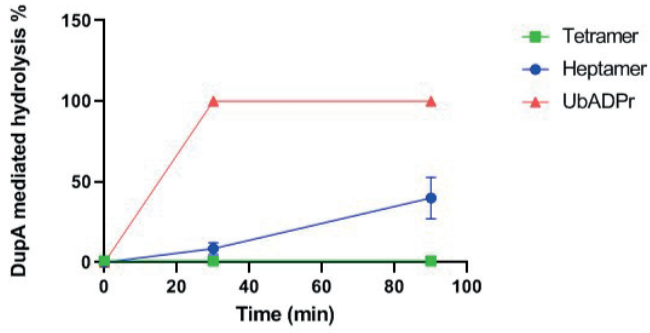


Figure 2. DupA mediated hydrolysis of tetramer **20**, heptamer **26** as compared to enzymatically obtained Ub-ADPr.[34] DupA mediated hydrolysis is measured over time as followed by mass-spectrometry.

Conclusion

This chapter describes the development of the first methodology for the synthesis of Arg-ADP-ribosylated peptides. Key event in the methodology is the coupling of an orthogonally protected, ribosylated isothiocyanate with the amine in the side chain of an ornithine residue furnishing the ribosylated guanidinyl functionality of Arg. The methodology was combined with the phosphoramidite method for the on-resin introduction of a phosphate monoester to give phosphoribosylated Arg. The phosphate monoester was subsequently employed in a P^{III} – P^{V} coupling procedure to install the pyrophosphate bridge to complete the ADPr moiety. With this new methodology, Arg ADP-ribosylated peptides **20** and **26** were obtained in 9.9 and 9.0% yield respectively. In an ultimate test for the applicability of this method for preparation of Arg-ADPr peptides, fully synthetic ubiquitin, ADP-ribosylated on Arg42 was prepared in 0.5% overall yield.

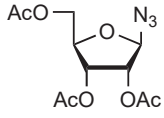
Besides NMR, LC-MS and HRMS analysis of peptides **20** and **26**, their ability to be recognized in biologically relevant settings was evaluated. Therefore, a hydrolytic assay was performed with DupA, an enzyme known to hydrolyze the pyrophosphate bridge in human ubiquitin ADP-ribosylated on Arg42. Tetramer **20** showed no hydrolytic turnover with DupA, most likely due to a mismatch of peptide sequence and recognition of the active site of DupA. Heptamer **26**, having the sequence flanking the Arg42 residue in hUb, in fact was recognized by DupA albeit less efficient than full-length Ub-ADPr. The fact that DupA can recognize heptamer **26**, together with analytical data (NMR, LC-MS, HRMS) concludes that this is the first report on the fully synthetic native Arg-ADP-ribosylated peptides.

Experimental section

General synthetic procedures

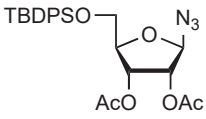
All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. Column chromatography was performed on silica gel 60 Å (40-63 µm, Macherey-Nagel). TLC analysis was performed on Macherey-Nagel aluminium sheets (silica gel 60 F_{254}). TLC was used to visualize compounds by UV at wavelength 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) or $KMnO_4$ spray (20 g/L $KMnO_4$ and 10 g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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. NMR spectra were recorded on a Bruker AV-400, AV-500 or AV-600 NMR. Chemical shifts (δ) are given in ppm relative to tetramethyl silane. Coupling constants (*J*) are given in Hz. All given ^{13}C-APT spectra are proton decoupled. In case of synthetic Ub-ADPr, HPLC purification was performed on a Shimadzu semi-preparative RP-HPLC system, equipped with a Waters C18-Xbridge 5 µm OBD (10 x 150 mm) column at a flowrate of 6.5 mL/min. using 2 mobile phases: A: MQ + 0.05% FA, B: MeCN + 0.05 % FA. Gradient: 0 -> 100% B. High resolution mass spectra were recorded on a Waters XEVO-G2 XS Q-TOF mass spectrometer equipped with an electrospray ion source in positive mode (source voltage 3.0 kV, desolvation gas flow 900 L/hr, temperature 250 °C) with resolution R = 22000 (mass range m/z = 50-2000) and 200 pg/uL Leu-Enk (m/z = 556.2771) as a "lock mass".

2,3,5-tri-*O*-acetyl-β-D-ribofuranosyl azide (10)



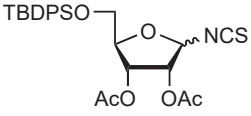
The title compound was prepared according to a modified, literature procedure.[28] Commercially available 1,2,3,5-tetra-*O*-acetyl-β-ribofuranose (7.96 g, 25 mmol) was dissolved in DCM (225 mL, 0.1 M). TMS-N_3 (3.65 mL, 27.5 mmol, 1.1 eq.) and a 1.0 M $SnCl_4$ solution in DCM (25 mL, 25 mmol, 1.0 eq.) were added to the reaction. The reaction was stirred for 2 hours after which TLC indicated full conversion. The reaction was carefully quenched with sat. aq. $NaHCO_3$ and transferred into a separatory funnel. The water layer was extracted with DCM and the combined organic layers were washed with brine, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (30% EtOAc in pentane) furnished the title compound as a colorless oil in quantitative yield. Spectral data was in accordance with literature.[28] **Rf**: 0.43 (30% EtOAc in pentane). **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 5.38 (d, *J* = 2.0 Hz, 1H, H-1), 5.34 (dd, *J* = 6.8, 4.8 Hz, 1H, H-3), 5.14 (dd, *J* = 4.8, 2.0 Hz, 1H, H-2), 4.42 (dd, *J* = 12.1, 3.2 Hz, 1H, H-5_a), 4.39 – 4.33 (m, 1H, H-4), 4.15 (dd, *J* = 12.1, 4.2 Hz, 1H, H-5_b), 2.13 (s, 6H, 2x Ac), 2.08 (s, 3H, Ac). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 170.5, 169.5, 169.4 (C=O Ac), 92.6 (C-1), 79.3 (C-4), 74.4 (C-2), 70.4 (C-3), 62.9 (C-5), 20.6, 20.5, 20.4 (CH_3 Ac).

2,3-di-*O*-acetyl-5-*tert*-butyldiphenylsilyl-β-D-ribofuranosyl azide (11)



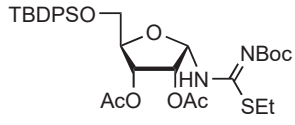
The title compound was prepared according to a modified, literature procedure.[29] Compound **10** (7.53 g, 25 mmol) was dissolved in MeOH (125 mL, 0.2 M) and the solution was cooled to 0 °C. A catalytic amount of Na(s) was added to the solution and the reaction was stirred for 2 hours after which an additional catalytic amount of Na(s) was added. After an additional 30 minutes of stirring, the reaction mixture was acidified with Amberlite H^+ resin till pH ≈ 6.0. The reaction was filtered and concentrated *in vacuo*. The crude triol (Rf = 0.36 in 10% MeOH in DCM) was co-evaporated with pyridine extensively, dissolved in pyridine (250 mL, 0.1 M) after which TBDPS-Cl (7.15 mL, 27.5 mmol, 1.1 eq.) was added and the reaction was stirred overnight. TLC indicated full conversion of the triol into a higher running product (Rf = 0.27 in 10% EtOAc in pentane). To the reaction mixture, DMAP (305 mg, 2.5 mmol, 0.1 eq.) and Ac_2O (11.8 mL, 125 mmol, 5.0 eq.) were added. After 1.5 hours of stirring, TLC indicated full conversion of the starting material. The reaction was quenched by the addition of MeOH and concentrated *in vacuo*. The crude residue was taken up in EtOAc and the resulting solution was washed with 1 M HCl, sat. aq. $NaHCO_3$ and brine successively. The organic layer was dried over $MgSO_4$, filtrated and concentrated *in vacuo*. Flash column chromatography (20 -> 30% Et_2O in pentane) furnished the title compound as a colorless oil (10.9 g, 21.9 mmol, 88% over three steps). **Rf:** 0.75 in 20% EtOAc in pentane. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.78 – 7.63 (m, 4H, TBDPS arom.), 7.48 – 7.35 (m, 6H, TBDPS arom.), 5.50 (t, *J* = 5.3 Hz, 1H, H-3), 5.42 (d, *J* = 3.0 Hz, 1H, H-1), 5.20 (dd, *J* = 4.9, 3.0 Hz, 1H, H-2), 4.21 (dt, *J* = 5.7, 3.6 Hz, 1H, H-4), 3.85 (dd, *J* = 11.4, 3.5 Hz, 1H, H-5_a), 3.72 (dd, *J* = 11.5, 3.6 Hz, 1H, H-5_b), 2.11 (s, 3H, Ac), 2.04 (s, 3H, Ac), 1.08 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 169.7, 169.6 (C=O Ac), 135.8, 135.8 (CH arom. TBDPS), 132.9 (Cq TBDPS), 130.0, 129.9, 127.9, 127.9 (CH arom. TBDPS), 92.9 (C-1), 82.6 (C-4), 74.9 (C-2), 71.0 (C-3), 63.4 (C-5), 26.8 (CH_3 *t*Bu), 20.7 (CH_3 Ac), 19.3 (Cq *t*Bu).

2,3-di-*O*-acetyl-5-*tert*-butyldiphenylsilyl-D-ribofuranosyl isothiocyanate (13)



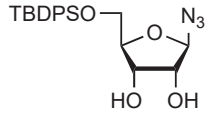
Compound **11** (10.0 g, 20.1 mmol) was dissolved in EtOAc (200 mL, 0.1 M) and the reaction was purged with nitrogen whilst sonicating for 10 minutes. To the stirred solution PtO_2 (913 mg, 4.02 mmol, 0.2 eq.) was added. The solution was purged with H_2 for 2 hours after which the reaction was filtered over a pad of Celite. To the filtrate water was added (200 mL) followed by K_2CO_3 (11.1 g, 80.5 mmol, 4.0 eq.) and thiophosgene (3.1 mL, 40.2 mmol, 2.0 eq.). The suspension was vigorously stirred overnight after which it was transferred into a separatory funnel and sat. aq. $NaHCO_3$ was added. The water layer was extracted thrice with EtOAc and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (20 -> 40% Et_2O in pentane) yielded the title compound as a colorless oil (α-anomer 3.26 g, 6.35 mmol. β-anomer 4.73 g, 9.21 mmol. Combined total yield of 77%). *α-anomer:* **Rf:** 0.71 in 20% Et_2O in pentane. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.66 (m, 4H, TBDPS arom.), 7.39 (m, 6H, TBDPS arom.), 5.66 (d, *J* = 5.2 Hz, 1H, H-1), 5.58 (dd, *J* = 6.0, 1.8 Hz, 1H, H-3), 5.34 (dd, *J* = 6.1, 5.2 Hz, 1H, H-2), 4.34 (q, *J* = 2.4 Hz, 1H, H-4), 3.80 (qd, *J* = 11.5, 2.7 Hz, 2H, H-5), 2.17 (s, 3H, Ac), 2.15 (s, 3H, Ac), 1.06 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 170.1, 169.4 (C=O Ac), 140.2 (NCS), 135.5, 135.5 (CH arom. TBDPS), 132.5, 132.4 (Cq TBDPS), 130.0, 129.9, 127.9, 127.8 (CH arom. TBDPS), 85.2 (C-1), 85.0 (C-4), 72.4 (C-2), 70.6 (C-3), 63.3 (C-5), 26.7 (CH_3 *t*Bu), 20.7, 20.1 (CH_3 Ac), 19.1 (Cq *t*Bu). *β-anomer:* **Rf:** 0.37 in 20% Et_2O in pentane. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.78 – 7.62 (m, 4H, TBDPS arom.), 7.49 – 7.35 (m, 6H, TBDPS arom.), 5.56 (d, *J* = 3.5 Hz, 1H, H-1), 5.49 (t, *J* = 5.0 Hz, 1H, H-3), 5.42 (dd, *J* = 4.9, 3.6 Hz, 1H, H-2), 4.19 (dt, *J* = 5.0, 3.5 Hz, 1H, H-4), 3.88 – 3.69 (m, 2H, H-5), 2.12 (s, 3H, Ac), 2.05 (s, 3H, Ac), 1.09 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 169.7, 169.4 (C=O Ac), 143.3 (NCS), 135.8, 135.8 (CH arom. TBDPS), 132.9, 132.7 (Cq TBDPS), 130.0, 129.9, 128.0 (CH arom. TBDPS), 88.0 (C-1), 83.1 (C-4), 75.6 (C-2), 71.3 (C-3), 63.5 (C-5), 26.8 (CH_3 *t*Bu), 20.7, 20.6 (CH_3 Ac), 19.3 (Cq *t*Bu). **HRMS:** $[C_{26}H_{31}NO_6SSi + Na]^+$ found: 536.1544, calculated: 536.1534.

1-(*tert*-butoxycarbonyl)-3-(2,3-di-*O*-acetyl-5-*tert*-butyldiphenylsilyl-α-D-ribofuranos-1-yl)-2-ethylisothiourea (1)



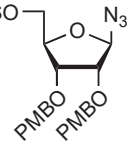
Compound **13** (3.26 g, 6.35 mmol) was dissolved in THF (32 mL, 0.2 M). The solution was saturated with ammonia by bubbling NH_3 (g) through the solution for 1 hour. The reaction was monitored by TLC and after complete conversion of starting material, the reaction was purged with N_2 for 1 minute. The crude thiourea was concentrated *in vacuo* until a white foam was obtained. The crude product was dissolved in DCM (64 mL, 0.1 M) and DMAP (78 mg, 0.64 mmol, 0.1 eq.), TEA (1.77 mL, 12.7 mmol, 2.0 eq.) and Boc_2O (1.60 mL, 6.99 mmol, 1.1 eq.) were added and the reaction was stirred overnight. The reaction was diluted with DCM and the solution was washed with water and brine consecutively. The water layers were extracted with DCM and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was dissolved in MeCN (64 mL, 0.1 M) and K_2CO_3 (8.78 g, 63.5 mmol, 10.0 eq.) and EtI (1.78 mL, 22.2 mmol, 3.5 eq.) were added to the solution. The suspension was stirred overnight and taken up in EtOAc. The organic layer was washed with brine and the water layer was extracted with EtOAc. The organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (10 -> 30% EtOAc in pentane) afforded the title compound as a white foam (2.54 g, 3.86 mmol, 61%). **Rf:** 0.26 in 10% EtOAc in pentane. **^{1}H NMR:** (500 MHz, $CDCl_3$) δ 10.66 (s, 1H, NH), 7.79 – 7.60 (m, 4H, TBDPS arom.), 7.47 – 7.31 (m, 6H, TBDPS arom.), 5.86 (dd, *J* = 8.5, 5.7 Hz, 1H, H-1), 5.62 (dd, *J* = 5.4, 1.8 Hz, 1H, H-3), 5.48 (t, *J* = 5.6 Hz, 1H, H-2), 4.22 (q, *J* = 2.6 Hz, 1H, H-4), 3.77 (d, *J* = 2.9 Hz, 2H, H-5), 3.13 (qd, *J* = 7.5, 1.1 Hz, 2H, CH_2 SEt), 2.22 (s, 3H, Ac), 2.10 (s, 3H, Ac), 1.50 (s, 9H, Boc), 1.31 (t, *J* = 7.5 Hz, 3H, CH_3 SEt), 1.09 (s, 9H, *t*Bu TBDPS).**^{13}C NMR:** (126 MHz, $CDCl_3$) δ 172.7 (RHN-*C*-(NBoc)(SEt)), 169.9, 169.1 (C=O Ac), 161.6 (C=O Boc), 135.6, 135.5 (CH arom. TBDPS), 132.6, 132.5 (Cq TBDPS), 129.8, 129.8, 127.8, 127.7 (CH arom. TBDPS), 82.5 (C-4), 81.7 (C-1), 79.2 (Cq *t*Bu Boc), 72.0 (C-3), 70.3 (C-2), 63.6 (C-5), 28.1 (CH_3 Boc), 26.7 (CH_3 *t*Bu TBDPS), 25.1 (CH_2 SEt), 20.4, 20.3 (CH_3 Ac), 19.1 (Cq *t*Bu TBDPS), 13.8 (CH_3 SEt). **HRMS:** $[C_{33}H_{46}N_2O_8SSi + Na]^+$ found: 659.2815, calculated: 659.2817.

5-*O*-*tert*-butyl-diphenylsilyl -β-D-ribofuranosyl azide



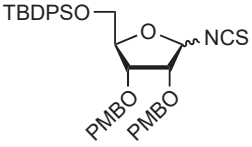
Compound **10** (2.99 g, 9.92 mmol was dissolved in MeOH (50 mL, 0.2 M). A 25 wt% solution of NaOMe in MeOH (0.23 mL, 0.99 mmol, 0.1 eq.) was added and the solution was stirred for 1 hour. The reaction was quenched with Amberlite H^+ resin, filtered and concentrated *in vacuo*. The resulting residue was co-evaporated extensively with pyridine and dissolved in pyridine (100 mL, 0.1 M). TBDPS-Cl (2.8 mL, 11 mmol, 1.1 eq.) was added and the reaction was stirred overnight. The reaction was quenched by the addition of MeOH and the reaction was concentrated *in vacuo*. The residue was taken up in EtOAc and the resulting solution was washed with 1 M HCl, sat. aq. $NaHCO_3$ and brine consecutively. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (20 -> 40% EtOAc in pentane) yielded the title compound as a colorless oil (3.65 g, 8.83 mmol, 89%). **Rf:** 0.54 in 40% EtOAc in pentane. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.77 – 7.62 (m, 4H, TBDPS arom.), 7.51 – 7.32 (m, 6H, TBDPS arom.), 5.31 (d, J = 1.9 Hz, 1H, H-1), 4.36 (t, J = 6.3, 4.8 Hz, 1H, H-3), 4.05 (dt, J = 6.3, 4.2 Hz, 1H, H-4), 3.96 (dd, J = 4.8, 2.0 Hz, 1H, H-2), 3.83 (qd, J = 11.2, 4.2 Hz, 2H, H-5), 3.06 (bs, 1H, OH), 2.60 (bs, 1H, OH), 1.08 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 135.7 (CH arom. TBDPS), 133.1, 133.0 (Cq TBDPS), 130.0, 130.0, 128.0, 127.9 (CH arom. TBDPS), 94.9 (C-1), 83.9 (C-4), 75.6 (C-2), 71.5 (C-3), 63.9 (C-5), 26.9 (CH_3 *t*Bu TBDPS), 19.3 (Cq *t*Bu TBDPS). **HRMS:** $[C_{21}H_{37}N_3O_4Si + Na]^+$ found: 436.1664, calculated: 436.1663.

5-*O*-((*tert*-butyl)-diphenylsilyl)-2,3-di-*O*-(4-methoxybenzyl)-β-D-ribofuranosyl azide (12)



5-*O*-((*tert*-butyl)-diphenylsilyl) -β-D-ribofuranosyl azide (3.65 g, 8.83 mmol) and TBABr (570 mg, 1.77 mmol, 0.2 eq.) were co-evaporated in toluene before they were dissolved in DMF (44 mL, 0.2 M). PMB-Cl (4.80 mL, 35.3 mmol, 4.0 eq.) was added and the solution was cooled to 0 °C. NaH (60 wt% dispersion in oil, 1.06 g, 26.5 mmol, 3.0 eq.) was added and the reaction was stirred for 15 minutes before the ice bath was removed. The reaction was stirred for an additional hour before TLC showed full conversion of the starting material. The reaction was cooled to 0 °C and carefully quenched by the addition of sat. aq. $NaHCO_3$. After bubbling seized, the emulsion was transferred into a separatory funnel and diluted with sat. aq. $NaHCO_3$. The water layer was extracted thrice with Et_2O and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (20 -> 30% Et_2O in pentane) furnished the title compound (3.91 g, 5.98 mmol, 68%) as a clear oil. **Rf:** 0.56 in 30% Et_2O in pentane. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.67 (ddt, J = 8.1, 6.5, 1.6 Hz, 4H, TBDPS arom.), 7.45 – 7.33 (m, 6H, TBDPS arom.), 7.32 – 7.25 (m, 2H, PMB arom.), 7.22 – 7.14 (m, 2H, PMB arom.), 6.92 – 6.78 (m, 4H, PMB arom.), 5.37 (d, J = 2.2 Hz, 1H, H-1), 4.55 (q, J = 11.8 Hz, 2H, CH_2 PMB), 4.49 – 4.34 (m, 2H, CH_2 PMB), 4.22 (dt, J = 6.6, 3.3 Hz, 1H, H-4), 4.16 (dd, J = 6.7, 4.5 Hz, 1H, H-3), 3.84 (dd, J = 11.5, 3.1 Hz, 1H, H-5_a), 3.78 (s, 3H, CH_3 PMB), 3.77 (s, 3H, CH_3 PMB), 3.73 (dd, J = 4.6, 2.2 Hz, 1H, H-2), 3.69 (dd, J = 11.5, 3.5 Hz, 1H, H-5_b), 1.04 (s, 9H, CH_3, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, CDCl3) δ 159.5, 159.4 (Cq PMB), 135.7, 135.7, 134.9 (CH arom. TBDPS), 133.2, 133.1 (Cq TBDPS), 129.8, 129.8, 129.8 (CH arom. TBDPS/PMB), 129.7 (Cq PMB), 129.7, 129.6 (CH arom. TBDPS/PMB), 129.5 (Cq PMB), 127.8, 127.8, 114.0, 113.9, 113.8 (CH arom. PMB), 93.2 (C-1), 83.0 (C-4), 79.7 (C-2), 76.0 (C-3), 72.2, 72.1 (CH_2 PMB), 63.1 (C-5), 55.3, 55.3 (CH_3 PMB), 26.8 (CH_3 TBDPS), 19.3 (Cq *t*Bu TBDPS). **HRMS:** $[C_{37}H_{43}N_3O_6Si + Na]^+$ found: 676.2812, calculated: 676.2813.

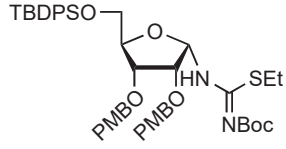
5-*O*-((*tert*-butyl)-diphenylsilyl)-2,3-di-*O*-(4-methoxybenzyl)-α,β-D-ribofuranosyl isothiocyanate (14)



Compound **12** (1.34 g, 2.04 mmol) was dissolved in EtOAc (20 mL. 0.1 M) and the reaction was purged with nitrogen whilst sonicating for 10 minutes. To the stirred solution PtO_2 (93 mg, 0.41 mmol, 0.2 eq.) was added. The solution was purged with H_2 for 2 hours after which the reaction was filtered over a pad of Celite. To the filtrate water was added (20 mL) followed by K_2CO_3 (11.3 g, 8.16 mmol, 4.0 eq.) and thiophosgene (313 µL, 4.08 mmol, 2.0 eq.). The suspension was vigorously stirred overnight after which it was transferred into a separatory funnel and sat. aq. $NaHCO_3$ was added. The water layer was extracted thrice with EtOAc and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (10 -> 15% Et_2O in pentane) yielded the title compound as a colorless oil (α-anomer 670 mg, 1.00 mmol. β-anomer 378 mg, 0.56 mmol. Combined total yield of 76%). *α-anomer:* **Rf:** 0.36 in 20% Et_2O in pentane. **^{1}H NMR:** (400 MHz, CDCl3) δ 7.57 (ddt, J = 8.3, 6.6, 1.5 Hz, 4H, TBDPS arom.), 7.47 – 7.33 (m, 6H, TBDPS arom.), 7.33 – 7.26 (m, 4H, PMB arom.), 6.90 – 6.81 (m, 4H, PMB arom.), 5.31 (d, J = 4.5 Hz, 1H, H-1), 4.74 (d, J = 11.7 Hz, 1H, CH_{2a} PMB), 4.61 (s, 2H, CH_2 PMB), 4.49 (d, J = 11.7 Hz, 1H, CH_{2b} PMB), 4.25 (q, J = 2.7 Hz, 1H, H-4), 4.09 – 4.00 (m, 2H, H-2 + H-3), 3.79 (s, 3H, CH_3 PMB), 3.78 (s, 3H, CH_3 PMB), 3.64 (dd, J = 11.5, 3.3 Hz, 1H, H-5_a), 3.55 (dd, J = 11.4, 2.8 Hz, 1H, H-5_b), 0.97 (s, 9H, *t*Bu TBPDS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 159.6, 159.3 (Cq PMB), 139.9 (NCS), 135.6, 135.6 (CH arom. TBDPS), 133.0, 132.8 (Cq TBDPS), 130.2 (Cq PMB), 130.0, 130.0, 129.7, 129.6 (CH arom TBDPS/PMB), 129.2 (Cq PMB), 127.9, 127.9, 114.1, 113.8 (CH arom. PMB), 86.1 (C-1), 85.7 (C-4), 79.6, 75.5 (C-2 + C-3), 72.8, 72.6 (CH_2 PMB), 63.7 (C-5), 55.3 (CH_3 PMB), 26.9 (CH_3 TBPDS), 19.3 (Cq *t*Bu TBDPS). *B-anomer:* **Rf:** 0.27 in 20% Et_2O in pentane. **^{1}H NMR:** (400 MHz, $CDCl_3$) δ 7.66 (ddt, J = 8.0, 6.3, 1.7 Hz, 4H, TBDPS arom.), 7.45 – 7.33 (m, 6H, TBDPS arom.), 7.30 – 7.16 (m, 4H, PMB arom.), 6.92 – 6.79 (m, 4H,

PMB arom.), 5.42 (d, J = 2.7 Hz, 1H, H-1), 4.55 (s, 2H, CH_2 PMB), 4.44 (q, J = 11.4 Hz, 2H, CH_2 PMB), 4.20 – 4.11 (m, 2H, H-3 + H-4), 3.96 (dd, J = 4.2, 2.8 Hz, 1H, H-2), 3.80 (dd, J = 11.5, 2.8 Hz, 1H, H-5_a), 3.77 (s, 3H, CH_3 PMB), 3.76 (s, 3H, CH_3 PMB), 3.68 (dd, J = 11.6, 3.2 Hz, 1H, H-5_b), 1.04 (s, 9H, *t*Bu TBDPS). **^{13}C NMR:** (101 MHz, $CDCl_3$) δ 159.6, 159.5 (Cq PMB), 140.7 (NCS), 135.7, 135.6 (CH arom. TBDPS), 133.2, 132.9 (Cq TBDPS), 129.8, 129.8, 129.6 (CH arom. TBDPS/PMB), 129.5, 129.1 (Cq PMB), 127.9, 127.8, 127.8 (CH arom. PMB), 114.1, 114.0, 113.9 (CH arom. PMB), 88.4 (C-1), 83.4 (C-4), 81.3 (C-2), 76.0 (C-3), 72.5, 72.3 (CH_2 PMB), 63.2 (C-5), 55.3, 55.3 (CH_3 PMB), 26.9 (CH_3 TBDPS), 19.3 (Cq *t*Bu TBDPS). **HRMS:** $[C_{38}H_{43}NO_6Si + Na]^+$ found: 692.2464, calculated: 692.2473.

1-(*tert*-butoxycarbonyl)-3-(5-*O*-((*tert*-butyl)-diphenylsilyl)-2,3-di-*O*-(4-methoxybenzyl)-α-D-ribofuranos-1-yl)-2-ethylisothiourea (2)



Compound **14** (670 mg, 1.05 mmol) was dissolved in THF (5.25 mL, 0.2 M). The solution was purged with NH_3 for 1 hour after which the reaction was purged with N_2 for 1 minute. The crude thiourea was concentrated *in vacuo* till a white foam. The crude product was dissolved in DCM (10.5 mL, 0.1 M). DMAP (13 mg, 0.11 mmol, 0.1 eq.) and Boc_2O (266 µL, 1.16 mmol, 1.1 eq.) were added and the reaction was stirred for 2 hours. The reaction was diluted with DCM and the resulting solution was washed with brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was dissolved in MeCN (10.5 mL, 0.1 M) and K_2CO_3 (1.45 g, 10.5 mmol, 10.0 eq.) and EtI (296 µL, 3.68 mmol, 3.5 eq.) were added to the solution. The suspension was stirred overnight and taken up in EtOAc. The organic layer was washed with brine, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (20 -> 40% Et_2O in pentane) furnished the title compound as a white foam (430 mg, 0.53 mmol, 50%). **Rf:** 0.42 in 30% Et_2O in pentane. **^{1}H NMR:** *measured at 59 °C* (500 MHz, $CDCl_3$) δ 7.67 – 7.57 (m, 4H, TBDPS arom.), 7.43 – 7.39 (m, 2H, TBDPS arom.), 7.39 – 7.33 (m, 4H, TBDPS arom.), 7.30 – 7.23 (m, 4H, PMB arom.), 6.86 – 6.79 (m, 4H, PMB arom.), 5.68 (bs, 1H, H-1), 4.65 (d, J = 11.3 Hz, 1H, CH_{2a} PMB), 4.56 – 4.45 (m, 3H, CH_2 PMB + CH_{2b} PMB), 4.18 (ddd, J = 4.4, 3.3, 2.4 Hz, 1H, H-4), 4.12 (dd, J = 5.0, 2.5 Hz, 1H, H-3), 4.09 (t, J = 5.1 Hz, 1H, H-2), 3.77 (s, 3H, CH_3 PMB), 3.76 (s, 3H, CH_3 PMB), 3.70 – 3.62 (m, 2H, H-5), 2.90 (bs, 2H, CH_2 Et), 1.49 (s, 9H, CH_3 Boc), 1.21 (t, J = 7.4 Hz, 3H, CH_3 Et), 1.03 (s, 9H, CH_3 TBDPS). **^{13}C NMR:** (126 MHz, $CDCl_3$) δ 161.1 (C=O Boc), 159.7, 159.5 (Cq PMB), 135.7, 135.6 (CH arom. TBDPS), 133.4, 133.3 (Cq TBDPS), 130.3 (Cq PMB), 129.9, 129.9 (CH arom. TBDPS/PMB), 129.8 (Cq PMB), 129.6 (CH arom. TBDPS/PMB), 127.8, 127.8, 114.1, 114.0 (CH arom. PMB), 83.0 (C-4), 82.1 (C-1), 79.1 (Cq *t*Bu Boc), 78.2 (C-3), 77.4 (C-2), 73.0, 72.6 (CH_2 PMB), 64.1 (C-5), 55.3 (CH_3 PMB), 28.3 (*t*Bu CH_3 Boc), 27.0 (*t*Bu CH_3 TBDPS), 25.2 (CH_2 Et), 19.3 (Cq *t*Bu TBDPS), 13.8 (CH_3 Et). **HRMS:** $[C_{45}H_{58}N_2O_8SSi + H]^+$ found: 815.3745, calculated: 815.3756.

Synthesis of peptides 20 and 26

The following procedures are applicable on 50 µmol resin and scalable for smaller scales.

Peptide synthesis

The intermediate peptides were synthesized using standard, Fmoc-based solid phase peptide synthesis utilizing (pre-loaded) TentaGel® S AC purchased from Rapp Polymer Gmbh. Coupling cycles were as followed: Fmoc deprotection: 2x2 minutes, 1x5 minutes treatment with 20% piperidine in DMF. Coupling: treatment of 6 eq. amino acid, 6 eq. HCTU (0.25 M in DMF) and 12 eq. DIPEA (1 M in DMF) for 30 minutes. Capping: 2x2 minutes treatment of the resin with a 10% Ac_2O solution in DMF and catalytic DIPEA. Washing between the steps was done with DMF. For the prospected Arg-ADPr site, commercially available Fmoc-Orn(Alloc)-OH was used in the coupling cycle.

Deprotection/building block coupling for Arg-ADPr peptides

The Alloc protecting group was removed by treating the resin with a freshly prepared solution of 10 mg $Pd(PPh_3)_4$ and 23 mg DMBA in 1 mL DCM (purged with nitrogen prior to use) for 15 minutes. This procedure was then repeated twice to ensure full deprotection. The resin was washed extensively with DCM and DMF. Coupling of the ribosyl building block was performed as followed: Ribosyl building block **2** (3 eq.) was dissolved in DMF (0.1 M) and added to the resin. TEA (30 eq.) followed by $AgNO_3$ (3 eq.) were added to the reaction and the syringe was wrapped in aluminum foil to protect it from light and shaken overnight. The resin was then extensively washed with DCM, DMF.

Deprotection/phosphorylation

The resin was washed with THF and treated with a 1 M TBAF in THF for 30 minutes. The resin was thoroughly washed with DCM and DMF before the treatment was repeated once, furnishing the desilylated intermediate. The resin was then extensively washed with MeCN and flushed with nitrogen to remove traces of water before the resin was subjected to a solution of $(FmO)_2PN(iPr)_2$ (2.5 eq. as 0.13 M in MeCN) and DCI (5.0 eq. as 0.25 M in MeCN) was added. The resin was shaken for 30 minutes after which the resin was washed with MeCN. The resin was then treated with a 0.5 M CSO solution in MeCN for 30 minutes. The resin was then treated with a 10% DBU solution in DMF (2x 15 minutes) to furnish the crude, immobilized and deprotected phosphoribosylated peptide.

Pyrophosphate synthesis

The resin was extensively washed with MeCN and flushed with nitrogen to remove traces of water. The resin was then treated with a solution of adenosine amidite **8** (3 eq. (0.13 M in MeCN)) and DCI (6 eq. (0.25 M in MeCN)) for 30 minutes. The resin was thoroughly washed with MeCN before a CSO solution (0.5 M in MeCN) was added to the resin and shaken for 30 minutes.

Final deprotection and cleavage

The resin was then treated with a 10% DBU solution in DMF (2x 10 minutes) to remove the cyano ethyl protecting group. The resin was then treated with a 1 M TBAF solution in THF (2x 45 minutes) and washed with DMF followed by DCM. Final cleavage/deprotection occurred by treating the resin with a cleavage cocktail (2.5:10:87.5 TIS:TFA:DCM) for 4 hours. The crude peptide was precipitated by flushing the cleavage cocktail in an ice-cold 1:1 mixture of Et_2O:pentane. The resin was washed twice with cleavage cocktail. The crude products were stored at -20 °C overnight to induce as much precipitation as possible before the products were centrifuged. The supernatant was removed, obtaining the solid crude peptide.

Ac-Gly-Arg(5-*O*-adenosine-diphosphate-α,β-D-ribosyl)-Thr-Phe-OH (20)

The general procedures described above were applied to 25 µmol TentaGel® S AC resin preloaded with Phe. The amino acids used were Fmoc-Gly-OH, Fmoc-Thr(Trt)-OH and Fmoc-Orn(Alloc)-OH. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (2.71 mg, 2.47 µmol, 9.9%). **^{1}H NMR:** (500 MHz, D_2O) δ 8.45 – 8.38 (m, 2H, H-2 adenine α/β), 8.20 – 8.10 (m, 2H, H-8 adenine α/β), 7.21 – 7.03 (m, 10H, Phe α/β), 6.02 (d, J = 5.8 Hz, 1H, H-1' adenosine α or β), 5.93 (d, J = 5.4 Hz, 1H, H-1' adenosine α or β), 5.25 (d, J = 4.3 Hz, 1H, H-1' ribosyl α or β), 5.05 (d, J = 5.8 Hz, 1H, H-1' adenosine α or β). **^{31}P NMR:** (202 MHz, D_2O) δ -10.2, -10.3, -10.3, -10.4, -10.7, -10.7, -10.8, -10.8. **LC-MS:** (0 -> 20% B in A): Rt = 7.15. **HRMS:** $[C_{38}H_{56}N_{12}O_{20}P_2 + H]^+$ found: 1063.3277, calculated: 1063.3282.

Ac-Gly-Arg(5-*O*-adenosine-diphosphate-α,β-D-ribosyl)-Leu-Ile-Phe-Ala-Gly-OH (26)

The general procedures described above were applied to 25 µmol TentaGel® S AC resin preloaded with Gly. The amino acids used were Fmoc-Gly-OH and Fmoc-Thr(Trt)-OH, Fmoc-Phe-OH, Fmoc-Ile-OH, Fmoc-Leu-OH and Fmoc-Orn(Alloc)-OH. The crude peptide was purified by RP-HPLC in NH_4OAc buffer. The pure fractions were concentrated, co-evaporated extensively with a 1:1 mixture of MeCN:Milli-Q water, redissolved in Milli-Q water and lyophilized to obtain the title compound as a white solid (6.09 mg, 4.51 µmol, 9.0%). **^{1}H NMR:** (850 MHz, D_2O) δ 8.54 – 8.43 (m, 2H, H-2 adenine α/β), 8.24 (s, 2H, H-8 adenine α/β), 7.28 – 7.25 (m, 4H, Phe arom. α/β), 7.25 – 7.20 (m, 2H, Phe arom. α/β), 7.20 – 7.15 (m, 4H, Phe arom. α/β), 6.09 (d, J = 5.8 Hz, 2H, H-1' adenosine), 5.31 (d, J = 4.3 Hz, 1H, H-1' ribosyl α or β), 5.11 (d, J = 5.8 Hz, 1H, H-1' ribosyl α or β). **^{31}P NMR:** (202 MHz, D_2O) δ -10.2, -10.3, -10.3, -10.4, -10.7, -10.7, -10.8, -10.8. **LC-MS:** (10 -> 90% B in A): Rt = 4.23. **HRMS:** $[C_{51}H_{79}N_{15}O_{22}P_2 + H]^+$ found: 1316.5073, calculated: 1316.5072.

Synthesis of full-length Ub-ADPr

Solid Phase Peptide Synthesis

SPPS was performed according to literature procedure[31] on a Syro II MultiSyntech Automated Peptide synthesizer using standard 9-fluorenylmethoxycarbonyl (Fmoc) based solid phase peptide chemistry at 20 µmol scale, using fourfold excess of amino acids relative to pre-loaded preloaded Fmoc amino acid chloro-trityl resin (0.2 mmol/g, Rapp Polymere GmbH). On position 42 in the peptide sequence arginine was replaced by Fmoc-Orn(Alloc)-OH. After SPPS, 6 µmol Ub1-76 (R42 → Alloc ornithine) on resin was treated with PyBOP (3.1 mg, 30 µmol, 5 eq) and Bt-PEG_2-COOH (16.1 mg, 30 µmol, 5 eq) in DMF (2 mL). After 5 minutes of shaking, DIPEA (16 µL, 90 µmol, 15 eq) was added. The reaction mixture was shaken overnight, after which a test cleavage confirmed full conversion of the conjugation.

Deprotection/building block coupling

The resin was then washed with DMF and DCM before resuspended in DCM. Alloc deprotection was performed by treating the resin with a solution of $(Pd(PPh_3)_4$ (1.4 mg, 1.2 µmol, 0.2 eq) and $PhSiH_3$ (15 µL, 120 µmol, 20 eq) in anhydrous DCM. This was done two times and a test cleavage confirmed complete deprotection. Coupling of the ribosyl building block was performed as follows: Ribosyl building block **2** (146 mg, 180 µmol, 30 eq.) was dissolved in DMF (1.8 mL, 0.1 M) and added to the resin. TEA (252 µL, 1800 µmol, 300 eq.) was added followed by the addition of $AgNO_3$ (31 mg, 180 µmol, 30 eq.). The syringe was wrapped in aluminum foil to protect it from light and shaken overnight. The resin was then extensively washed with DCM and DMF and transferred to a new syringe to rid the resin of residual silver salts.

Deprotection/phosphorylation

The resin was washed with THF and treated with a 1 M TBAF in THF for 30 minutes. The resin was thoroughly washed with DCM and DMF before the treatment was repeated once, furnishing the desilylated intermediate. The resin was then extensively washed with dry MeCN and flushed with nitrogen to remove traces of water. The resin was subjected to a solution of $(FmO)_2PN(iPr)_2$ (1.15 mL, 0.13 M in MeCN, 150 µmol, 25 eq.) and a DCI solution (1.20 mL, 0.25 M in MeCN, 50 eq.) was added. The resin was shaken for 30 minutes after which the resin was washed with MeCN. The resin was then treated with a *t*BuOOH solution (a 5.5 M solution in nonane was diluted ten times in MeCN to obtain the solution used for oxidation) for 30 minutes. The resin was thoroughly washed with DCM and DMF and subsequently treated with a 10% DBU solution in DMF (2x 15 minutes) to furnish the crude, immobilized and deprotected phosphoribosylated ubiquitin.

Pyrophosphate synthesis

The resin was extensively washed with MeCN and flushed with nitrogen to remove traces of water. The resin was then treated with a solution of adenosine amidite **8** (1.15 mL, 0.13 M in MeCN, 150 µmol, 25 eq.) and a DCI solution (1.20 mL, 0.25 M in MeCN, 50 eq.) for 30 minutes. The resin was thoroughly washed with MeCN before a *t*BuOOH solution (a 5.5 M solution in nonane was diluted ten times in MeCN to obtain the solution used for oxidation) was added to the resin and shaken for 30 minutes.

Final deprotection, cleavage and purification

The resin was thoroughly washed with DCM and DMF and subsequently treated with a 10% DBU solution in DMF (2x 15 minutes). The resin was washed with THF and treated with a 1 M TBAF in THF for 30 minutes. The resin was thoroughly washed with DCM and DMF before the treatment was repeated once. The resin was treated with TFA:TIS:H_2O:Phenol (90.5:2:5:2.5) for 1.5 hours before filtrated in an ice-cold solution of Et_2O:pentane (1:1). The precipitate formed was centrifuged (5 Minutes, 3500 rpm) and the supernatant decanted. The pellet was subsequently dried with N_2, taken up in warm DMSO and diluted in warm water before purified by RP-HPLC. Pure fractions were pooled and lyophilized affording UbR42-ADPr (301 µg, 0.031 µmol, 0.5%) as a white powder. **LC-MS**: Rt = 1.47 min. **HRMS:** $[C_{416}H_{688}N_{114}O_{139}P_2S + 5H]^{5+}$ found: 1922.0817, calculated: 1922.0005. $[C_{416}H_{688}N_{114}O_{139}P_2S + 6H]^{6+}$ found: 1601.7731, calculated: 1601.6680. $[C_{416}H_{688}N_{114}O_{139}P_2S + 7H]^{7+}$ found: 1373.0911, calculated: 1373.0022.

DupA mediated hydrolysis of peptides ADP-ribosylated at arginine

The peptides **20** and **26** (5 µM) in buffer (20 mM TRIS, 150 mM NaCl, pH 7.6) were incubated with DupA (3 µM) or without (background hydrolysis) at 37°C. At the indicated time points 15 µL sample was 4 times diluted before measuring HRMS. The ratio of product versus starting material was determined, corrected for t = 0 minutes and plotted as increase in pyrophosphate cleavage over time. The means of two individual measurements is depicted with standard deviation.

References

[1] C. G. Goff, *J. Biol. Chem.* **1974**, *249*, 6181–6190.

[2] A. K. Ganesan, D. W. Frank, R. P. Misra, G. Schmidt, J. T. Barbieri, *J. Biol. Chem.* **1998**, *273*, 7332–7337.

[3] T. Meyer, H. Hilz, *Eur. J. Biochem.* **1986**, *155*, 157–165.

[4] H. Osago, M. Terashima, N. Hara, K. Yamada, M. Tsuchiya, *J. Biochem. Biophys. Methods* **2008**, *70*, 1014–1019.

[5] K. D. Siegmund, F. Klink, *FEBS Lett.* **1992**, *312*, 139–142.

[6] B. Eide, P. Gierschik, A. Spiegel, *Biochemistry* **1986**, *25*, 6711–6715.

[7] S. Laing, M. Unger, F. Koch-Nolte, F. Haag, *Amino Acids* **2011**, *41*, 257–269.

[8] M. O. Hottiger, P. O. Hassa, B. Lüscher, H. Schüler, F. Koch-Nolte, *Trends Biochem. Sci.* **2010**, *35*, 208–219.

[9] N. Hara, M. Tsuchiya, M. Shimoyama, *J. Biol. Chem.* **1996**, *271*, 29552–29555.

[10] T. Maehama, S. I. Hoshino, T. Katada, *FEBS Lett.* **1996**, *388*, 189–191.

[11] M. Vogelsgesang, K. Aktories, *Biochemistry* **2006**, *45*, 1017–1025.

[12] B. Lüscher, M. Bütepage, L. Eckei, S. Krieg, P. Verheugd, B. H. Shilton, *Chem. Rev.* **2018**, *118*, 1092–1136.

[13] S. Buch-Larsen, I. Hendriks, J. Lodge, M. Rykær, B. Furtwängler, E. Shishkova, M. Westphall, J. Coon, M. Nielsen, *Cell Rep.* **2020**, *32*, 108176.

[14] D. M. Leslie Pedrioli, M. Leutert, V. Bilan, K. Nowak, K. Gunasekera, E. Ferrari, R. Imhof, L. Malmström, M. O. Hottiger, *EMBO Rep.* **2018**, *19*, 1–11.

[15] J. Moss, S. chen Tsai, R. Adamik, H. chia Chen, S. J. Stanley, *Biochemistry* **1988**, *27*, 5819–5823.

[16] R. J. Collier, *Toxicon* **2001**, *39*, 1793–1803.

[17] F. Rosenthal, P. Nanni, S. Barkow-Oesterreicher, M. O. Hottiger, *J. Proteome Res.* **2015**, *14*, 4072–4079.

[18] Y. Zhang, J. Wang, M. Ding, Y. Yu, *Nat. Methods* **2013**, *10*, 981–984.

[19] J. Voorneveld, J. G. M. Rack, I. Ahel, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2018**, *20*, 4140–4143.

[20] K. Feichtinger, H. L. Sings, T. J. Baker, K. Matthews, M. Goodman, *J. Org. Chem.* **1998**, *63*, 8432–8439.

[21] A. Hamze, J. Martinez, J.-F. Hernandez, *J. Org. Chem.* **2004**, *69*, 8394–8402.

[22] A. V. Rodríguez-mayor, G. J. Peralta-camacho, K. J. Cárdenas-martínez, J. E. García-castañeda, *Curr. Org. Chem.* **2020**, *24*, 2475–2497.

[23] D. E. DeMong, R. M. Williams, *J. Am. Chem. Soc.* **2003**, *125*, 8561–8565.

[24] S. Wang, L. Corcilius, P. P. Sharp, A. Rajkovic, M. Ibba, B. L. Parker, R. J. Payne, *Chem. Sci.* **2017**, *8*, 2296–2302.

[25] S. Wang, L. Corcilius, P. P. Sharp, R. J. Payne, *Bioorganic Med. Chem.* **2017**, *25*, 2895–2900.

[26] M. Pan, S. Li, X. Li, F. Shao, L. Liu, H. G. Hu, *Angew. Chemie - Int. Ed.* **2014**, *53*, 14517–14521.

[27] X. Li, R. Krafczyk, J. Macošek, Y. L. Li, Y. Zou, B. Simon, X. Pan, Q. Y. Wu, F. Yan, S. Li, J. Hennig, K. Jung, J. Lassak, H. G. Hu, *Chem. Sci.* **2016**, *7*, 6995–7001.

[28] A. Štimac, J. Kobe, *Carbohydr. Res.* **1992**, *232*, 359–365.

[29] G. J. van der Heden van Noort, M. G. van der Horst, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *J. Am. Chem. Soc.* **2010**, *132*, 5236–5240.

[30] L. A. Stevens, R. L. Levine, B. R. Gochuico, J. Moss, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 19796–19800.

[31] F. El Oualid, R. Merkx, R. Ekkebus, D. S. Hameed, J. J. Smit, A. De Jong, H. Hilkmann, T. K. Sixma, H. Ovaa, *Angew. Chemie - Int. Ed.* **2010**, *49*, 10149–10153.

[32] J. Qiu, M. J. Sheedlo, K. Yu, Y. Tan, E. S. Nakayasu, C. Das, X. Liu, Z. Luo, *Nature* **2016**, *533*, 120–124.

[33] D. Shin, R. Mukherjee, Y. Liu, A. Gonzalez, F. Bonn, Y. Liu, V. V. Rogov, M. Heinz, A. Stolz, G. Hummer, V. Dötsch, Z.-Q. Luo, S. Bhogaraju, I. Dikic, *Mol. Cell* **2020**, *77*, 164–179.

[34] R. Q. Kim, M. Misra, A. Gonzalez, I. Tomaskovic, D. Shin, H. Schindelin, D. V Filippov, H. Ovaa, I. Dikic, G. J. van der Heden van Noort, *Chem. - A Eur. J.* **2021**, *27*, 2506–2512.

Chapter 6

The Synthesis of ADP-Ribosylated Nucleic Acids

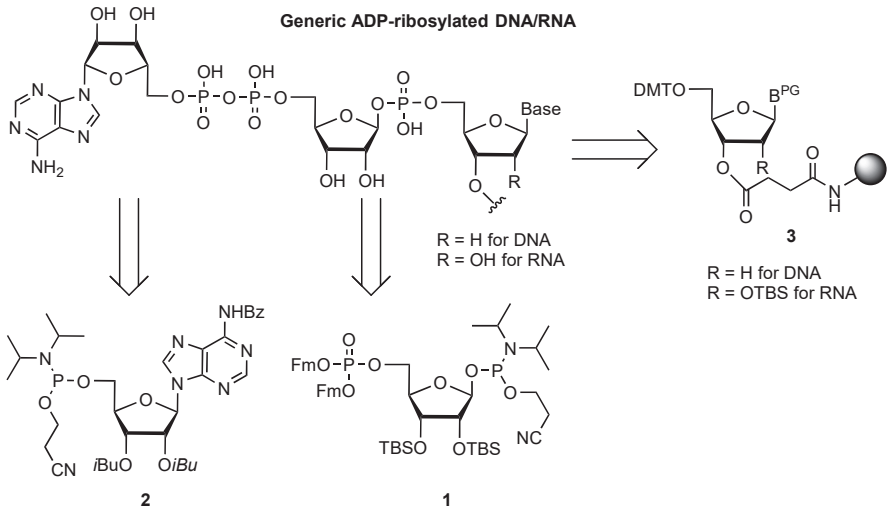
Introduction

Research in the field of ADP-ribosylation is mostly focused on the elucidation of the pathways and mechanisms that regulate ADP-ribosylation of specific amino acids in certain proteins. However, the ADP-ribosylation of biopolymers other than polypeptides, such as nucleic acids, starts to attract attention as well. An early example of ADP-ribosylation of DNA was discovered with proteins originating from the cabbage butterfly[1] and certain shellfish[2] that catalyze the formation of not only N^2-(a-ADP-ribos-1-yl)-2'-deoxyguanosine but also its b-anomer in duplex DNA. Interestingly these studies indicate an involvement of ADP-ribosylation in host-pathogen conflicts.[3] Another nucleobase modification is represented by the sequence specific modification of thymine in single stranded DNA by the bacterial toxin-antitoxin system DarT-DarG.[4] Moreover, recent *in vitro* studies showed that mammalian ARTs are able to ADP-ribosylate DNA[5–9] and RNA[10] oligomers, provided with a 5'- or 3'- terminal phosphate. In the resulting structures, the anomeric configurations of these terminal phosphates linked to the ADP-ribosyl moiety are not yet established. ADP-ribosylation of DNA can occur in both double and single stranded oligomers, seems to be dependent on the localization of the strand break and is probably independent of the nucleotide sequence. DNA oligomers are modified with ADPr by PARP1 and PARP2[6] but PARP3 appears to be most proficient and it is reported that nucleotide fragments, ADP-ribosylated by PARP3 can serve as primers for further modification to give PAR chains.[8] The reversal of the ADPr-modification can be accomplished by several hydrolases, such as MacroD1,[11] MacroD2, PARG, TARG and ARH3[6,10] that are able to remove the ADPr moiety from the terminal phosphate of DNA. The physiological significance of the *mono*-ADP-ribosylation of single-stranded breaks in DNA on the 5'-phosphorylated site is underscored by the fact that the modification is recognized by DNA ligases which ligate to the damaged site and allow DNA restoration by nucleophilic attack of the 3'-OH to the 5'-phosphorylated strand.

Although various *in vitro* studies have shown that terminal phosphorylated DNA/RNA can be ADP-ribosylated and several hydrolases can reverse this modification much is still unknown and cellular *in vitro* and subsequent *in vivo* studies are needed to establish the presence of ADP-ribosylated nucleic acids and to elucidate their functions.[5,9,12] This research will benefit from the presence of well-defined molecular tools, accessible by organic synthesis. For instance, structural elucidation of complexes of ADP-ribosylated DNA/RNA fragments and catalytic domains of both ARTDs and ARHs will provide valuable information as to their respective activity and specificity. This chapter describes the development of a synthetic procedure to ADP-ribosylated DNA and RNA.

Retrosynthetic analysis

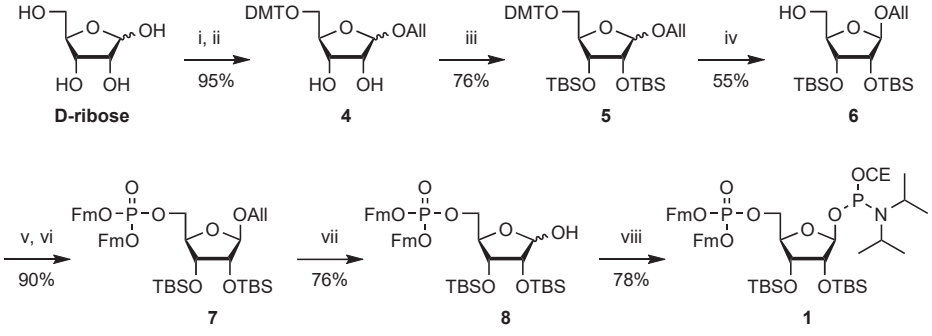
A generic structure of a DNA and RNA nucleotide with the ADP-ribosyl moiety at the 5'–phosphate in the b-configuration is depicted in Scheme 1. To make the building blocks compatible with standardized, automated oligonucleotide synthesis, phosphoramidite **2** was chosen for the adenosine moiety, a building block that was also applied in the automated synthesis of ADPr oligomers.[13] For the ribosyl moiety, key building block **1** was designed as (i) the TBS protecting group is often employed in RNA synthesis, (ii) the anomeric amidite enables coupling to the 5'-OH of the protected DNA/RNA oligomer and (iii) the Fm-protected phosphotriester on the 5-OH allows not only for a mild deprotection of the phosphate triester prior to coupling with amidite **2** but is also orthogonal to all protecting groups of the nucleobases. The standard solid support for oligonucleotide synthesis, controlled pore glass[14] (CPG) equipped with an amino alkyl linker was chosen as suitable. CPG can be functionalized via an ester linkage with the 3'(2') OH of any given (deoxy)ribonucleoside,[15] leading to **3** as solid support provided with the first (deoxy)nucleoside. In case of DNA building blocks, no additional protecting group is required and in case of RNA, the usual 2'-OTBS protected nucleotides can be used.



Scheme 1. Retrosynthetic analysis of generic, ADP-ribosylated nucleotides

Results and discussion

Functionalized CPG resin **3** is commercially available and building block **2** was synthesized by a known procedure[16] (see Chapter 4), leaving the synthesis of key ribosyl building block **1** (Scheme 2). The preparation commenced with Fischer glycosylation of D-ribose with allyl alcohol followed by the regioselective installation of the DMT protecting group on the 5'-OH, giving **4** in 95% yield. Protection of the 2'- and 3'-OH with silyl groups by treatment of **4** with TBS-Cl and imidazole provided **5**. Next, the DMT protecting group was removed by acidolysis with TFA which furnished **6** in 55% yield. It is noteworthy that only the β-anomer could be obtained after flash column chromatography, explaining the moderate yield. Next, a two-step phosphorylation reaction with known bis-Fm-protected phosphoramidite (see Chapter 3) followed by oxidation of the intermediate phosphite yielded triester **7** in a 90% yield. De-allylation of **7** was affected in a two-step procedure whereby the allyl group is isomerized by an iridium catalyst and the obtained enol ether is hydrolyzed by I_2/$NaHCO_3$/H_2O. Finally, **8** is treated with commercially available 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite and DIPEA to give crude **1**. Manual purification of **1** with standard silica gel chromatography, neutralized by adding TEA to the solvent to prevent hydrolysis of the anomeric amidite, proofed cumbersome as the addition of TEA triggered the cleavage of the Fm protecting groups. Therefore, purification of **1** was performed by automated column chromatography using a pH-neutral silica cartridge which furnished key building block **1** as the β-anomer in a total yield of 21% over 8 steps.

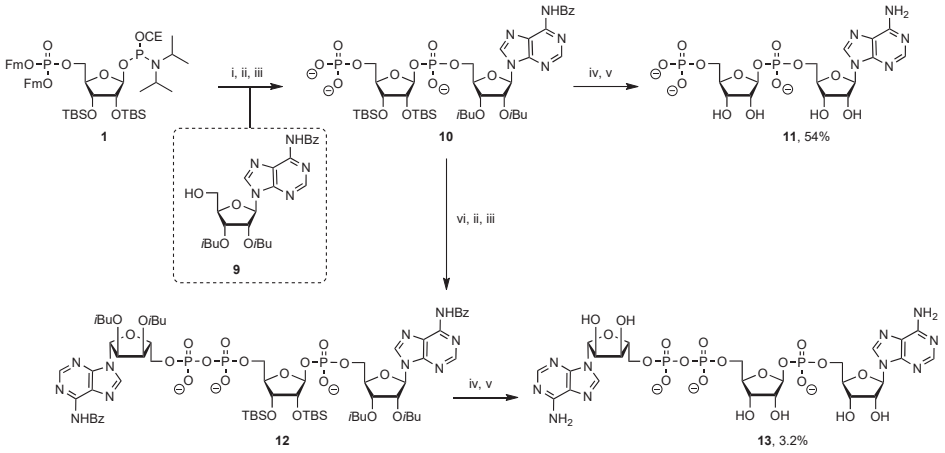


Scheme 2. Synthesis of key building block **1**, the ribosyl moiety suitable for on-resin nucleotide synthesis. Reagents and conditions: i) Ac-Cl, All-OH. ii) DMT-Cl, pyr. iii) TBS-Cl, imidazole, DMF. iv) TFA, TIS, DCM. v) $(FmO)_2PN(iPr)_2$, DCI, MeCN. vi) *t*BuOOH, nonane. vii) $Ir(COD)(Ph_2MeP)_2{\cdot}PF_6$, H_2, THF, then I_2, $NaHCO_3$, H_2O. viii) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM.

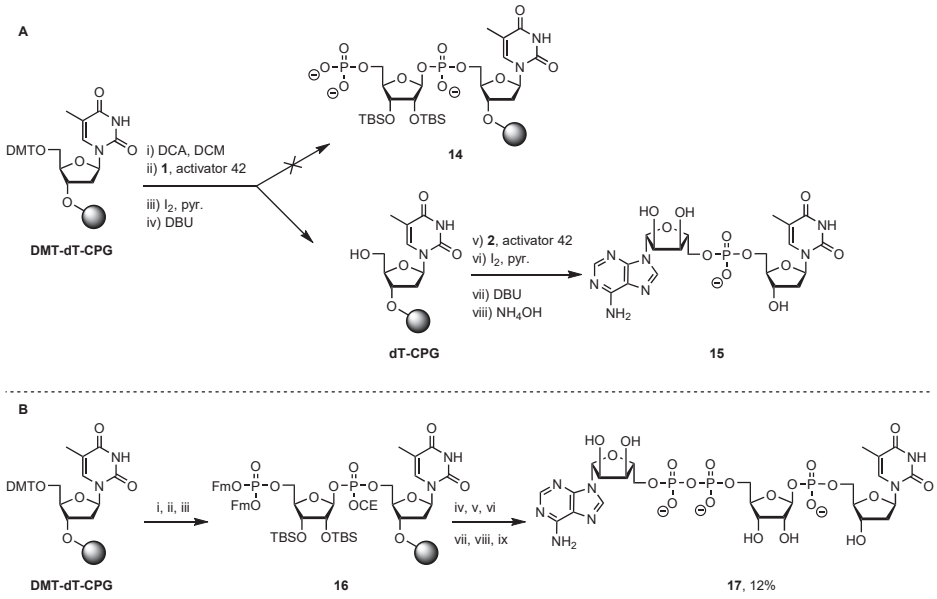
In order to establish whether building block **1** is indeed suitable for the synthesis of DNA/RNA oligomers ADP-ribosylated at their 5' end, a solution-phase exploratory study was performed (Scheme 3). Therefore, adenosine nucleoside **9** (an intermediate *en route* to amidite **2**) was chosen as a model since it mimics the protecting group pattern of the intermediates in the solid-phase synthesis. Amidite **1** was coupled to adenosine **9** using ETT as activator and the reaction progress was monitored by ^{31}P-NMR, showing quantitative conversion of **1** into the expected phosphite intermediate. Subsequent CSO-mediated oxidation led to the formation of a second phosphate triester and the obtained, fully protected crude product was not isolated but TEA was added to the reaction mixture. Monitoring of the reaction progress by LC-MS analysis showed that not only the cyanoethyl group but also one of the two Fm-groups were cleaved off almost instantly. On the other hand, the removal of the second Fm-group required overnight stirring to give trianion **10**. Next, the TBS groups were removed by concentration of the reaction mixture and treatment of the resulting amorphous oil with TEA·3HF in THF. Monitoring by LC-MS analysis showed that it took two days to completely remove the TBS protecting groups. Finally, ammonolysis of the remaining protecting groups (*i*Bu and Bz) with NH_4OH followed by HPLC purification furnished **11** in a 54% yield over 5 steps. Encouraged by this successful result, which indicated that the anomeric phosphodiester survived the applied reaction conditions, the pyrophosphate moiety was appended using the P^{III} – P^{V} procedure as described in earlier chapters. To achieve this, the phosphomonoester of **10** was coupled under influence of ETT with adenosine amidite **2**. Ensuing oxidation of the phosphite intermediate with CSO and the removal of the phosphate protecting groups with TEA furnished intermediate **12**. Analogously to the deprotection procedure for **11**, intermediate **12** was first treated with TEA·3HF and subsequently with NH_4OH to yield ADPr-P-RNA **13**. Although the LC-MS trace of the crude reaction mixture showed the predominant formation of product **13**, HPLC purification proved difficult and homogeneous target **13** was isolated in 3% yield only. Nevertheless, the exploratory study toward **11** and **13** shows that the chemistry for both the anomeric phosphodiester and the pyrophosphate function is viable, indicating that the solid phase synthesis to ribosylated nucleic acid fragments is feasible.

At this stage the solid phase synthesis of ribosylated DNA was investigated and it was decided to adopt the conditions of the standard automated DNA synthesis with regards to the type and concentration of the reagents and reaction times.[17] Commercially available DMT-dT-CPG (Scheme 4) was first treated with a DCA solution in DCM to cleave off the DMT protecting group. Next, anomeric phosphoramidite **1** was coupled to the 5'-OH of dT-CPG under the influence of 5-[3,5-Bis(trifluoromethyl)phenyl]-1*H*-tetrazole (Activator 42®). Subsequent oxidation with a I_2 solution in pyridine and water, standard in automated oligonucleotide synthesis,[17] and removal of the cyanoethyl and Fm phosphate protecting groups with DBU should result in immobilized **14**. Surprisingly, after performing a second

coupling cycle with adenosine amidite **2** and subsequent deprotection/cleavage, the LC-MS analysis of the crude product revealed **15** to be the major product, while the expected product was missing. In order to find a productive pathway to **14** several adjustments of the above procedure were made. However, increasing the reaction time of the coupling of building block **1**, repetition of the coupling cycle and use of the activator ETT did not lead to formation of the desired product. Fortunately, oxidation of the phosphite intermediate with CSO rather than I_2, resulted in the effective formation of immobilized **16**. Introduction of the pyrophosphate moiety by DBU mediated cleavage of the phosphate protecting groups, followed by coupling of adenosine amidite **2** with protected intermediate **16** and again, oxidation of the phosphate-phosphite intermediate with CSO proceeded without noteworthy side products as was observed by LC-MS analysis of the crude products. Deprotection and subsequent cleavage from the CPG support led to the silylated precursor of **17**. The remaining silyl ethers were removed by treatment of the crude product with a 4:3:2 v/v% solution of TEA·3HF:TEA:NMP solution. Direct injection of the reaction mixture on a HW40 size exclusion column (after quenching with an NH_4OAc buffer solution) resulted in homogenous **17** in 12% yield.



Scheme 3. Exploratory study for coupling of key building block **1** in solution. Reagents and conditions: i) **9**, ETT, MeCN. ii) CSO, MeCN. iii) TEA. iv) TEA·3HF, THF. v) NH_4OH, H_2O. vi) **2**, ETT, MeCN.



Scheme 4. A: Attempted coupling of building block **1** to DMT-dT-CPG using I_2 as oxidating agent and formation of side product **15**. **B:** Solid-phase synthesis of partially protected 5'-ADP-ribosylated thymidine **17**. Reagents and conditions: i) DCA, DCM. ii) **1**, activator 42®, MeCN. iii) CSO, MeCN. iv) DBU, MeCN. v) **2**, activator 42®, MeCN. vi) CSO, MeCN. vii) DBU, MeCN. viii) NH_4OH, H_2O. ix) TEA·3HF, TEA, NMP.

Conclusion

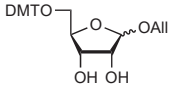
The important *in vitro* discoveries of various ADP-ribosylated nucleic acids raises questions how these DNA and RNA modifications occur *in vivo* and what biological role they play. The synthesis of well-defined fragments of ADP-ribosylated nucleic acid oligomers with and without labels will contribute to finding the answers. This chapter describes the first solution phase synthesis of ADP-ribosylated adenosine 5'-phosphate (**13**) and the first solid phase synthesis of ADP-ribosylated deoxythymidine 5'-phosphate (**17**). It turned out that oxidation of phosphite intermediates with CSO toward the target compounds is of prime importance for the successful synthesis of ADP-ribosylated nucleic acids. The developed procedure to ADP-ribosylated deoxythymidine 5'-phosphate lays down a basis for the future synthesis ADP-ribosylated nucleic acids.

Experimental section

General synthetic procedures

All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. 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 F_{254}). TLC was used to visualize compounds by UV at wavelength 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) or $KMnO_4$ spray (20 g/L $KMnO_4$ and 10 g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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. NMR spectra were recorded on a Bruker AV-400, AV-500 or AV-600 NMR. Chemical shifts (δ) are given in ppm relative to tetramethyl silane as internal standard. Coupling constants (*J*) are given in Hz. All given ^{13}C-APT spectra are proton decoupled.

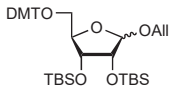
1-*O*-allyl-5-*O*-(4,4'-dimethoxytrityl)- α,β-D-ribofuranoside (4)



D-Ribose (7.51 g, 50 mmol) was suspended in allyl alcohol (125 mL, 0.4 M) and acetyl chloride (2.5 mL, 35 mmol, 0.7 eq.) was added. The reaction was stirred for 2 hours, then quenched with 6 mL pyridine and concentrated *in vacuo*. The residue was co-evaporated with pyridine, dissolved in pyridine (100 mL, 0.5 M) and DMT-Cl (17.8 g, 52.5 mmol, 1.05 eq.) was added. The reaction was stirred overnight after which the volume was reduced by evaporation *in vacuo* till about 20% of the original volume. The residue was taken up in EtOAc (1.000 mL) and carefully quenched with sat. aq. $NaHCO_3$. The suspension was transferred into a separatory funnel and the organic layer was washed with sat. aq. $NaHCO_3$, 1 M $CuSO_4$ and brine respectively. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (3 -> 10% acetone in DCM) yielded the title compound as an α:β mixture (23.48 g, 47.67 mmol, 95%). *α-anomer* **Rf:** 0.33 in 5% acetone in DCM. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.45 – 7.40 (m, 2H, DMT arom.), 7.33 – 7.25 (m, 5H, DMT arom.), 7.23 – 7.14 (m, 2H, DMT arom.), 6.86 – 6.78 (m, 4H, DMT arom.), 5.93 (dddd, *J* = 17.2, 10.3, 6.2, 5.2 Hz, 1H, OCH_2CHCH_2), 5.31 (dq, *J* = 17.2, 1.6 Hz, 1H, OCH_2CHCH_{2a}), 5.22 (dq, *J* = 10.4, 1.3 Hz, 1H, OCH_2CHCH_{2b}), 5.15 (d, *J* = 4.4 Hz, 1H, H-1), 4.33 (ddt, *J* = 12.9, 5.2, 1.5 Hz, 1H, $OCH_{2a}CHCH_2$), 4.29 – 4.22 (m, 1H, H-2), 4.20 – 4.07 (m, 2H, H-4 + $OCH_{2b}CHCH_2$), 4.02 – 3.96 (m, 1H, H-3), 3.77 (s, 6H, 2x CH_3 DMT), 3.31 (dd, *J* = 10.2, 3.8 Hz, 1H, H-5_a), 3.14 (dd, *J* = 10.2, 3.9 Hz, 1H, H-5_b), 3.02 (d, *J* = 9.9 Hz, 1H, 2' OH), 2.69 (d, *J* = 7.9 Hz, 1H, 3' OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 158.5, 144.8, 136.0, 135.9 (Cq DMT arom.), 133.9 (OCH_2CHCH_2), 130.1, 130.1, 129.2, 128.2, 127.9, 127.9, 126.9 (CH DMT arom.), 117.9 (OCH_2CHCH_2), 113.2 (CH DMT arom.), 101.2 (C-1), 86.2 (Cq DMT), 84.6 (C-4), 72.1 (C-2), 71.7 (C-3), 69.0 (OCH_2CHCH_2), 63.8 (C-5), 55.3, 55.3 (CH_3 DMT). *β-anomer* **Rf:** 0.13 in 5% acetone in DCM. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.50 – 7.44 (m, 2H, DMT arom.), 7.38 – 7.13 (m, 7H, DMT arom. + residual pyridine overlap), 6.84 – 6.78 (m, 4H, DMT arom.), 5.80 (dddd, *J* = 17.2, 10.3, 6.2, 5.2 Hz, 1H. OCH_2CHCH_2), 5.19 (dq, *J* = 17.2, 1.6 Hz, 1H, OCH_2CHCH_{2a}), 5.13 (dq, *J* = 10.4, 1.4 Hz, 1H, OCH_2CHCH_{2b}), 5.00 (d, *J* = 1.0 Hz, 1H, H-1), 4.28 (dd, *J* = 6.5, 4.8 Hz, 1H, H-3), 4.17 (ddt, *J* = 12.7, 5.1, 1.5 Hz, 1H, $OCH_{2a}CHCH_2$), 4.13 – 4.04 (m, 2H, H-3 + H-4), 3.94 (ddt, *J* = 12.7, 6.2, 1.4 Hz, 1H, $OCH_{2b}CHCH_2$), 3.77 (s, 6H, 2x CH_3 DMT), 3.32 – 3.20 (m, 2H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 158.5, 145.0,

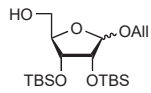
136.2 (Cq DMT arom.), 134.1 (OCH_2CHCH_2), 130.2, 130.2, 129.2, 128.3, 128.0, 126.9, 125.4 (CH arom. DMT), 117.5 (OCH_2CHCH_2), 113.3, 113.2 (CH arom. DMT), 106.4 (C-1), 86.2 (Cq DMT), 82.3 (C-4), 75.4 (C-2), 72.7 (C-3), 68.6 (OCH_2CHCH_2), 65.0 (C-5), 55.3 (CH_3 DMT). **HRMS:** $[C_{29}H_{32}O_7 + Na]^+$ found: 515.2042, calculated: 515.2040

1-*O*-allyl-5-*O*-(4,4'-dimethoxytrityl)-2,3-*O*-bis-*tert*-butyldimethylsilyl-α,β-D-ribofuranoside (5)



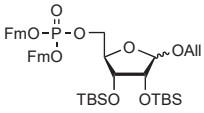
Compound **4** (23.48 g, 47.67 mmol) was co-evaporated with toluene and dissolved in DMF (240 mL, 0.2 M). Imidazole (9.74 g, 143 mmol, 3.0 eq.) and TBS-Cl (50 wt% in toluene, 50 mL, 143 mmol, 3.0 eq.) were added and the reaction was stirred overnight. The reaction was concentrated *in vacuo* and the residue was taken up in sat. aq. $NaHCO_3$ (1.000 mL). The water layer was extracted thrice with Et_2O and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (5 -> 10% Et_2O in pentane) yielded the title compound as an α:β mixture (26.05 g, 36.13 mmol, 76%). **Rf:** 0.63 (α-anomer) and 0.71 (β-anomer) in 10% Et_2O in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.57 – 7.51 (m, 2H, DMT arom. α + β), 7.48 – 7.32 (m, 5H DMT arom. α + β), 7.31 – 7.16 (m, 6H, DMT arom. α + β), 6.87 – 6.76 (m, 5H, DMT arom. α + β), 5.89 (dddd, *J* = 16.8, 10.3, 6.3, 5.1 Hz, 1H, OCH_2CHCH_2 α + β), 5.41 – 5.12 (m, 2H, OCH_2CHCH_2 α + β), 4.93 (s, 1H, H-1 β), 4.32 (ddt, *J* = 12.8, 5.2, 1.5 Hz, 1H, $OCH_{2a}CHCH_2$ α + β), 4.25 – 4.11 (m, 2H, H-3 + H-4 α + β), 4.11 – 4.03 (m, 1H, $OCH_{2b}CHCH_2$ α + β), 3.96 (dd, *J* = 3.6, 1.0 Hz, 1H, H-2 β), 3.83 – 3.73 (m, 7H, 2x CH_3 DMT α + β), 3.32 (dd, *J* = 10.2, 2.3 Hz, 1H, H-5 β), 3.02 (dt, *J* = 10.5, 5.3 Hz, 1H, H-5 β), 0.93 (s, 2H, *t*Bu TBS α), 0.91 (s, 9H, *t*Bu TBS β), 0.80 (s, 2H, *t*Bu TBS α), 0.73 (s, 9H, *t*Bu TBS β), 0.10 (dd, *J* = 8.3, 4.9 Hz, 7H 2x Me TBS α + β), -0.01 (s, 1H, Me TBS α), -0.05 (s, 3H, Me TBS β), -0.13 (s, 1H, Me TBS α), -0.20 (s, 3H, Me TBS β). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 158.5, 145.1, 136.5, 136.4 (Cq DMT arom.), 134.4 (OCH_2CHCH_2 β), 130.3, 130.3, 130.2, 128.5, 128.3, 127.9, 127.8, 126.8, 126.7 (CH DMT arom. α + β), 117.5 (OCH_2CHCH_2 β), 116.1 (OCH_2CHCH_2 α), 113.2, 113.1 (CH DMT arom. α + β), 106.4 (C-1 β), 102.0 (C-1 α), 85.8 (Cq DMT arom. β), 84.2 (C-4 α), 81.5 (C-4 β), 76.4 (C-2 β), 74.0 (C-2 α), 72.3 (C-3 β), 72.2 (C-3 α), 68.6 (OCH_2CHCH_2 β), 68.5 (OCH_2CHCH_2 α), 64.2 (C-5 β), 63.7 (C-5 α), 55.3 (CH_3 DMT α + β), 26.2, 25.9, 25.9 (CH_3 *t*Bu TBS α + β), 18.2, 18.1 (Cq *t*Bu TBS α + β), -4.0, -4.3, -4.5, -5.0 (Me TBS α + β). **HRMS:** $[C_{41}H_{60}O_7Si_2 + Na]^+$ found: 743.3767, calculated: 743.3770.

1-*O*-allyl-2,3-*O*-bis-*tert*-butyldimethylsilyl-β-D-ribofuranoside (6)



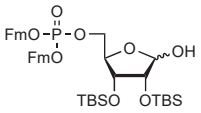
Compound **5** (26.05 g, 36.13 mmol) was dissolved in 5% TFA in DCM (180 mL total volume, 0.2 M). The reaction was stirred for 3 hours before the majority of the TFA was quenched with TEA. The reaction was diluted with DCM, washed with sat. aq. $NaHCO_3$ and brine consecutively. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (5 -> 10% EtOAc in pentane) yielded the title compound (only β-anomer, 8.49 g, 19.9 mmol, 55%). **Rf:** 0.37 in 10% Et_2O in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 6.02 – 5.80 (m, 1H, OCH_2CHCH_2), 5.38 – 5.16 (m, 2H, OCH_2CHCH_2), 4.81 (s, 1H, H-1), 4.28 (dd, *J* = 7.5, 4.2 Hz, 1H, H-3), 4.20 (ddt, *J* = 13.0, 5.3, 1.5 Hz, 1H, $OCH_{2a}CHCH_2$), 4.09 – 3.98 (m, 2H, H-4 + $OCH_{2b}CHCH_2$), 3.94 (d, *J* = 4.1 Hz, 1H, H-2), 3.83 (dd, *J* = 12.1, 2.5 Hz, 1H, H-5_a), 3.56 (dd, *J* = 12.1, 3.3 Hz, 1H, H-5_b), 0.96 – 0.86 (m, 18H, 2x *t*Bu TBS), 0.16 – 0.02 (m, 12H, 4x Me TBS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 134.1 (OCH_2CHCH_2), 117.8 (OCH_2CHCH_2), 106.7 (C-1), 82.8 (C-4), 76.9 (C-2), 70.9 (C-3), 69.1 (OCH_2CHCH_2), 61.6 (C-5), 26.0, 25.9 (CH_3 *t*Bu TBS), 18.2, 18.2 (Cq *t*Bu TBS), -4.1, -4.4, -4.5, -4.9 (Me TBS). **HRMS:** $[C_{20}H_{42}O_5Si_2 + Na]^+$ found: 441.2464, found: 441.2463.

1-*O*-allyl-2,3-*O*-bis-*tert*-butyldimethylsilyl-5-(*O*-difluorenylmethyl phosphate)-β-D-ribofuranoside (7)



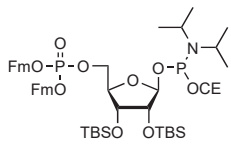
Compound **6** (570 mg, 1.36 mmol) was co-evaporated with toluene and dissolved in a 0.25 M solution of DCI in MeCN (10.9 mL, 2.73 mmol, 2.0 eq.). A 0.2 M stock solution of bis-(9H-fluoren-9-ylmethyl)-*N,N*-diisopropylamidophosphite in MeCN (8.15 mL, 1.63 mmol, 1.2 eq.) was added and the solution was stirred at room temperature. After 15 minutes of stirring, a 5.5 M solution of *t*BuOOH in nonane (2.5 mL, 13.6 mmol, 10 eq.) was added and the reaction was stirred for an additional hour. The reaction was quenched with sat. aq. $NaHCO_3$ and taken up in EtOAc. The resulting solution was washed with sat. aq. $NaHCO_3$ and brine consecutively, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (20% EtOAc in pentane) yielded the title compound as a white foam (1.04 g, 1.22 mmol, 90%). **Rf:** 0.44 in 20% EtOAc in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.71 (dddt, *J* = 9.7, 7.6, 2.0, 0.9 Hz, 4H, Fm arom.), 7.53 (ddq, *J* = 14.5, 7.5, 1.0 Hz, 4H, Fm arom.), 7.41 – 7.29 (m, 4H, Fm arom.), 7.29 – 7.19 (m, 6H, Fm arom.), 5.76 (dddd, *J* = 16.9, 10.3, 6.4, 5.0 Hz, 1H, OCH_2CHCH_2), 5.21 – 5.02 (m, 2H, OCH_2CHCH_2), 4.78 (s, 1H, H-1), 4.35 – 4.23 (m, 4H, 2x CH_2 Fm), 4.23 – 4.05 (m, 6H, H-3 + H-4 + H-5_a + 2x CH Fm + $OCH_{2a}CHCH_2$), 3.98 (dt, *J* = 10.6, 5.2 Hz, 1H, H-5_b), 3.91 (d, *J* = 3.8 Hz, 1H, H-2), 3.83 (ddt, *J* = 13.0, 6.4, 1.3 Hz, 1H, $OCH_{2b}CHCH_2$), 0.89 (s, 9H, *t*Bu TBS), 0.86 (s, 9H, *t*Bu TBS), 0.07 (s, 3H, Me TBS), 0.06 (s, 3H, Me TBS), 0.04 (s, 3H, Me TBS), 0.02 (s, 3H, Me TBS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.3, 143.2, 141.4, 141.4 (Cq Fm), 134.1 (OCH_2C, 127.9, 127.2, 125.4, 125.3, 125.3, 120.1, 120.1, 120.0 (CH arom. Fm), 117.5 (OCH_2CHCH_2), 106.0 (C-1), 80.4, 80.3 (C-4), 76.3 (C-2), 71.8 (C-3), 69.5, 69.5, 69.4, 69.4 (CH_2 Fm), 68.2 (OCH_2CHCH_2), 67.7, 67.7 (C-5), 48.1, 48.0 (CH Fm), 25.9, 25.9 (CH_3 *t*Bu TBS), 18.2, 18.1 (Cq *t*Bu TBS), -4.1, -4.4, -4.5, -4.9 (Me TBS). **^{31}P NMR** (162 MHz, $CDCl_3$) δ -0.8. **HRMS:** $[C_{48}H_{63}O_8Si_2 + Na]^+$ found: 877.3695, found: 877.3691.

2,3-*O*-bis-*tert*-butyldimethylsilyl-5-(*O*-difluorenylmethyl phosphate)-α,β-D-ribofuranoside (8)



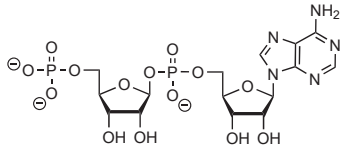
Compound **7** (1.04 g, 1.22 mmol) was dissolved in THF (6 mL, 0.2 M) and the reaction was purged with N_2. $Ir(COD)(Ph_2MeP)_2{\cdot}PF_6$ (15 mg, 0.018 mmol, 0.015 eq.) was added, followed by purging with H_2 for 15 seconds before the reaction was put under N_2 atmosphere again. The reaction was stirred for 1 hour before the reaction was diluted with 3.5 mL THF followed by the addition of 3.5 mL of sat. aq. $NaHCO_3$ and I_2 (464 mg, 1.83 mmol, 1.5 eq.). The reaction was stirred vigorously for an additional hour. The reaction was taken up in EtOAc and the organic layer was washed twice with sat. aq. $Na_2S_2O_3$ followed by brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (20 -> 40% EtOAc in pentane) yielded the title compound as an 2:1 α:β mixture (753 mg, 0.92 mmol,76%). **Rf:** 0.51 in 40% EtOAc in pentane. **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.75 – 7.66 (m, 8H, Fm arom. α + β), 7.59 – 7.43 (m, 8H, Fm arom. α + β), 7.42 – 7.18 (m, 16H, Fm arom. α + β), 5.08 (d, J = 1.8 Hz, 1H, H-1 β), 5.01 (dd, J = 11.2, 4.1 Hz, 1H, H-1 α), 4.41 (ddt, J = 32.1, 10.0, 6.3 Hz, 2H, CH_2 Fm β), 4.35 – 4.30 (m, 1H, H-3 β), 4.30 – 4.19 (m, 5H, H-5_a $β_a$ + 2x CH_2 Fm α), 4.19 – 4.05 (m, 7H, H-4 α + CH Fm α + β +CH_2 Fm β), 4.03 – 3.94 (m, 3H, H-3 α + H-4 β + H-5_b β), 3.93 – 3.87 (m, 2H, H-2 α + β), 3.84 (dd, J = 5.9, 4.4 Hz, 2H, H-5 α), 0.94 – 0.82 (m, 36H, *t*Bu TBS α + β), 0.13 – -0.01 (m, 24H, Me TBS α + β). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 143.3, 143.2, 143.1, 143.1, 143.0, 143.0, 141.5, 141.5, 141.4, 141.4, 141.4 (Cq Fm α + β), 128.0, 128.0, 127.9, 127.2, 127.2, 127.2, 125.3, 125.2, 125.1, 125.1, 120.1, 120.1, 120.1, 120.1, 120.0 (CH Fm arom. α + β), 102.7 (C-1 β), 97.7 (C-1 α), 82.7, 82.7 (C-4 α), 80.0, 80.0 (C-4 β), 77.2 (C-2β), 73.5 (C-3 α), 72.5 (C-2 α), 70.6 (C-3 β), 70.0, 70.0, 69.6, 69.5, 69.5, 69.5, 69.4, 69.4 (CH_2 Fm α + β), 67.4, 67.4 (C-5 β), 66.8, 66.7 (C-5 α), 48.1, 48.0, 48.0, 48.0, 47.9, 47.9, 47.9, 47.8 (CH Fm α + β), 25.9, 25.9, 25.8 (CH_3 *t*Bu TBS α +β), 18.3, 18.2, 18.1, 18.0 (Cq *t*Bu TBS α + β), -4.2, -4.4, -4.5, -4.6, -4.6, -4.7, -5.0, -5.0 (Me TBS α + β). **^{31}P NMR** (202 MHz, $CDCl_3$) δ -0.8 (β), -0.9 (α). **HRMS:** $[C_{45}H_{59}O_8Psi_2 + Na]^+$ found: 837.3382, calculated: 837.3378.

1'-*O*-(2',3'-*O*-bis-*tert*-butyldimethylsilyl-5'-(*O*-difluorenylmethyl phosphate)-β-D-ribofuranoside)-2-cyanoethyl-*N*-*N*-diisopropylphosphoramidite (1)



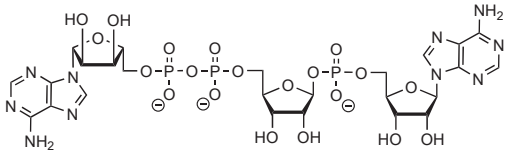
Compound **8** (2.45 g, 3.00 mmol, 1.0 eq.) was co-evaporated thrice with toluene and dissolved in anhydrous DCM (30 mL, 0.1 M). DIPEA (1.57 mL, 9.00 mmol, 3.0 eq.) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.00 mL, 4.50 mmol, 1.5 eq.) were added to the reaction and after 1.5 hours, TLC indicated full conversion of starting material. The reaction was diluted with DCM and washed with brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Automated column chromatography performed on a Biotage® Isolera™ Spektra Four system using neutral Biotage Zip (0 -> 20% Et_2O in pentane) afforded the title compound as a white foam and a mixture of diastereomers (S_P/R_P) (2.38 g, 2.34 mmol, 78%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.76 – 7.65 (m, 4H, Fm arom.), 7.57 – 7.45 (m, 4H, Fm arom.), 7.42 – 7.31 (m, 4H, Fm arom.), 7.30 – 7.20 (m, 4H, Fm arom.), 5.15 (d, J = 9.6 Hz, 1H, H-1), 4.39 – 4.08 (m, 10H, H-2/3 + H-4 + H-5 + CH_2 Fm + + CH_{2a} Fm + 2x CH Fm), 4.04 (tt, J = 10.8, 5.6 Hz, 1H, CH_{2b} Fm), 3.87 (dd, J = 31.7, 3.3 Hz, 1H, H-2/3), 3.79 – 3.61 (m, 2H, ROC$H_2$$CH_2$CN), 3.61 – 3.42 (m, 2H, 2x CH *i*Pr), 2.42 (dddd, J = 33.8, 15.5, 11.0, 6.0 Hz, 2H, ROCH$_2$CH_2CN), 1.18 – 1.05 (m, 12H, 4x CH_3 *i*Pr), 0.93 – 0.85 (m, 18H, *t*Bu TBS), 0.14 – 0.02 (m, 12H, Me TBS). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 143.3, 143.3, 143.2, 143.1, 141.5, 141.4, 141.4 (Cq Fm), 128.0, 127.9, 127.9, 127.2, 127.2, 127.2, 127.2, 125.4, 125.4, 125.3, 125.3, 125.3, 125.2 , 120.1, 120.1, 120.1, 120.0, 120.0, 120.0 (CH Fm arom.), 118.3, 117.5 (CN), 102.5, 102.4, 102.0, 101.8 (C-1), 80.3, 80.3, 79.9, 79.8 (C-4), 77.5, 77.3 (C-2/3), 71.8, 71.3 (C-2/3), 69.5, 69.4, 69.4, 69.3, 69.3 (C-5), 68.5, 68.5, 67.9, 67.8 (CH_2 Fm), 59.0, 58.9, 58.3, 58.1 (ROCH_2CH_2CN), 48.1, 48.1, 48.0, 48.0, 48.0 (CH Fm), 43.6, 43.6, 43.5, 43.5 (CH *i*Pr), 26.0, 25.9, 25.9 (CH_3 *t*Bu TBS), 24.7, 24.6, 24.6, 24.5, 24.4, 24.4 (CH_3 *i*Pr), 20.3, 20.3, 20.2 (ROCH$_2$CH_2CN), 18.1, 18.1, 18.1 (Cq *t*Bu TBS), -4.1, -4.3, -4.5, -4.5, -4.9, -4.9 (Me TBS). **^{31}P NMR** (202 MHz, $CDCl_3$) δ 149.1, 147.3, -0.9, -1.0.

5-*O*-phosphate-β-D-ribofuranoside-(5'-*O*-adenosine)-phosphate (11)



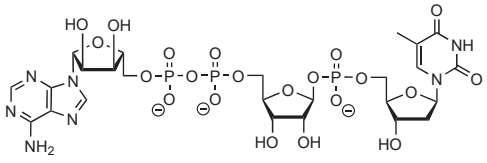
Compound **1** (51 mg, 50 µmol, 1.0 eq.) and adenosine **9** (31 mg, 60 µmol, 1.2 eq.) were co-evaporated thrice with toluene and a solution of ETT (0.25 M in MeCN, 400 µL, 100 µmol, 2.0 eq.) was added. The reaction was stirred for 15 minutes before a CSO solution (0.5 M in MeCN, 500 µL, 250 µmol, 5.0 eq.) was added. The solution was stirred for 30 minutes after which TEA (105 µL, 750 µmol, 15 eq.) was added and the solution was stirred an additional 16 hours. The reaction was concentrated *in vacuo* and co-evaporated with toluene. The resulting oil was dissolved in THF (0.5 mL, 0.1 M) and TEA·3HF (163 µL, 1.0 mmol, 20 eq.) was added. After 2 days, LC-MS analysis revealed full conversion of the starting material (Rt = 4.85 minutes in a 10 -> 90% MeCN gradient flow). To the solution, 3 mL of a 28% NH_4OH solution in water was added and the reaction was stirred overnight. The reaction was concentrated *in vacuo* and the resulting residue was taken up in Milli-Q water. The water layer was washed with EtOAc and the water layer was collected and concentrated *in vacuo*. Purification by prep-HPLC using a HILIC column and repeated lyophilization of the fractions containing product furnished the title compound as a white solid (16.4 mg, 27 µmol, 54%). **^{1}H NMR** (850 MHz, D_2O) δ 8.45 (s, 1H, adenine), 8.23 (s, 1H, adenine), 6.10 (d, J = 6.2 Hz, 1H, H-1'), 5.42 (d, J = 6.4 Hz, 1H, H-1), 4.74 (dd, J = 6.2, 5.1 Hz, 1H, H-2'), 4.50 (dd, J = 5.1, 3.1 Hz, 1H. H-3'), 4.36 (p, J = 2.8 Hz, 1H, H-4'), 4.25 (dd, J = 7.4, 4.5 Hz, 1H), 4.12 (ddd, J = 11.6, 4.7, 3.0 Hz, 1H, H-5'$_a$), 4.11 – 4.06 (m, 3H, H-2 + H-4 + H-5'$_b$), 4.03 (ddd, J = 11.3, 5.7, 3.4 Hz, 1H, H-5$_a$), 3.90 (dt, J = 11.3, 6.4 Hz, 1H. H-5$_b$). **^{13}C NMR** (214 MHz, D_2O) δ 156.0 (Cq adenine), 153.0 (CH adenine), 150.1 (Cq adenine), 140.9 (CH adenine), 119.6 (Cq adenine), 103.7 + 103.6 (C-1), 87.7 (C-1'), 85.3 + 85.2 (C-4'), 82.8 + 82.7 (C-4), 76.1 + 76.0 (C-2), 75.2 (C-2'), 71.5 (C-3'), 71.1 (C-3), 67.0 (C-5), 66.0 (C-5'). **^{31}P NMR** (202 MHz, D_2O) δ 1.1, -1.5.

(5-*O*-adenosine-diphosphate-β-D-ribosyl)-5'-*O*-adenosine phosphate (13)



Compound **1** (51 mg, 50 µmol, 1.0 eq.) and adenosine **9** (31 mg, 60 µmol, 1.2 eq.) were co-evaporated thrice with toluene and a solution of ETT (0.25 M in MeCN, 400 µL, 100 µmol, 2.0 eq.) was added. The reaction was stirred for 15 minutes before a CSO solution (0.5 M in MeCN, 500 µL, 250 µmol, 5.0 eq.) was added. The solution was stirred for 30 minutes after which TEA (105 µL, 750 µmol, 15 eq.) was added and the solution was stirred an additional 16 hours. The reaction was concentrated *in vacuo* and co-evaporated with toluene. The residue was co-evaporated thrice with toluene and compound **2** (71 mg, 100 µmol, 2.0 eq., separately co-evaporated with toluene) was added and a solution of ETT (0.25 M, 800 µL, 200 µmol, 4.0 eq.) was added. The reaction was stirred for 15 minutes before a CSO solution (0.5 M in MeCN, 800 µL, 400 µmol, 8.0 eq.) was added. The solution was stirred for 30 minutes after which DBU was added (112 µL, 750 µmol, 15 eq.) and the reaction was stirred 30 minutes. The reaction was concentrated *in vacuo* and co-evaporated in toluene. The resulting oil was dissolved in THF (0.5 mL, 0.1 M) and TEA·3HF (163 µL, 1.0 mmol, 20 eq.) was added. After 2 days, LC-MS analysis revealed full conversion of the starting material (Rt = 6.12 minutes in a 10 -> 90% MeCN gradient flow). To the solution, 3 mL of a 28% NH_4OH solution in water was added and the reaction was stirred overnight. The reaction was concentrated *in vacuo* and the resulting residue was taken up in Milli-Q water. The water layer was washed with EtOAc and the water layer was collected and concentrated *in vacuo*. Purification by prep-HPLC using a HILIC column and repeated lyophilization of the fractions containing product furnished the title compound as a white solid (1.44 mg, 1.62 µmol, 3.2%). **^{1}H NMR** (850 MHz, D_2O) δ 7.12 – 7.07 (m, 2H, adenine), 6.94 – 6.90 (m, 2H, adenine), 5.90 (d, J = 5.6 Hz, 1H, H-1 adenosine), 5.88 (d, J = 6.0 Hz, 1H, H-1 adenosine), 5.33 (d, J = 6.4 Hz, 1H, H-1 ribosyl), 4.60 – 4.55 (m, 2H, 2x H-2 adenosine), 4.40 – 4.35 (m, 2H, 2x H-3 adenosine), 4.24 – 4.19 (m, 2H, 2x H-4 adenosine), 4.11 – 4.05 (m, 3H, 2x CH_{2a} H-5 adenosine + H-3 ribosyl), 4.03 – 3.91 (m, 6H, 2x CH_{2b} adenosine + H-2 + H-4 + H-5 ribosyl). **^{13}C NMR** (peaks extracted from HSQC spectrum, 214 MHz, D_2O) δ 130.5, 130.5, 115.5, 115.5 (C-H adenine), 102.6 (C-1 ribosyl), 86.7, 86.3 (C-1 adenosine), 83.6 (C-4 adenosine), 81.5 (C-3 ribosyl), 75.0 (C-4 ribosyl), 74.4, 74.2 (C-2 adenosine), 70.1, 70.1, 70.1 (C-3 adenosine + C-2 ribosyl), 68.4, 66.8, 65.0, 65.0, 64.8 (C-5 adenosine + ribosyl). **^{31}P NMR** (202 MHz, D_2O) δ -1.4, -10.4. **HRMS:** $[C_{25}H_{35}N_{10}O_{20}P_3 + H]^+$ found: 889.1313, calculated: 889.1315.

(5-*O*-adenosine-diphosphate-β-D-ribosyl)-5'-*O*-thymidine phosphate (17)



In a fritted syringe, DMT-dT-CPG resin (300 mg, 10 µmol) was flushed with a 5 v/v% DCA solution in DCM until no yellow color appeared by addition of fresh DCA solution, indicating full removal of the DMT group. The resin was washed with anhydrous MeCN and flushed with N_2 to remove traces of water. Next, ribosyl building block **1** (0.1 M in MeCN, 400 µL, 40 µmol, 4.0 eq.) with a solution of activator 42® (0.25 M in MeCN, 600 µL, 150 µmol, 15 eq.). The resin was shaken for 10 minutes and the procedure was repeated to ensure maximum conversion. The phosphite intermediate was oxidized by treatment of the resin with a 0.3 M CSO solution (2x 10 minutes).The Fm protecting group was removed by treatment of the resin with a 10% DBU solution in MeCN (4x 5 minutes). The resin was washed with anhydrous MeCN and flushed with N_2 to remove traces of water. Next, adenosine amidite **2** (0.1 M in MeCN, 400 µL, 40 µmol, 4.0 eq.) was added to the resin, followed

by a solution of activator 42® (0.25 M in MeCN, 600 µL, 150 µmol, 15 eq.). The resin was shaken for 10 minutes and the procedure was repeated to ensure maximum conversion. The intermediate was oxidized by treatment of the resin with a 0.3 M CSO solution (2x 10 minutes). The cyano-ethyl protecting group was removed by treatment of the resin with a 10% DBU solution in MeCN (4x 5 minutes). Finally, the resin was shaken overnight in a 5 mL 28% NH_4OH solution in water. Afterwards, the resin was filtered and the filtrate was concentrated *in vacuo*. The remaining white solid was taken up in Mili-Q water (10 mL) and transferred in a tube followed by centrifugation. The supernatant was carefully collected and transferred into a round bottom flask and concentrated *in vacuo*. The residue was taken up in a 4:3:2 NMP:TEA·3HF:TEA solution (1.5 mL total volume) and the reaction mixture was transferred into a 15 mL tube. The reaction was shaken overnight and quenched by the addition of 1.5 mL of a 0.15 M NH_4OAc buffer. The resulting solution was directly injected on a HW40 column and eluted with NH_4OAc buffer. Fractions containing the product were concentrated and co-evaporated with 1:1 MeCN:Milli-Q water and lyophilized yielding the title compound as a white solid (1.07 mg, 1.24 µmol, 12%). **^{1}H NMR** (850 MHz, D_2O) δ 8.49 (s, 1H, adenine), 8.21 (s, 1H, adenine), 7.64 – 7.59 (m, 1H, thymine), 6.18 (dd, J = 8.3, 6.1 Hz, 1H, H-1' thymidine), 6.08 (d, J = 5.7 Hz, 1H, H-1' adenosine), 5.43 (d, J = 6.5 Hz, 1H, H-1' ribosyl), 4.70 (t, J = 5.4 Hz, 1H, H-2' adenosine), 4.57 (dt, J = 5.1, 2.3 Hz, 1H, H-3' thymidine), 4.49 (dd, J = 5.1, 3.7 Hz, 1H, H-3' adenosine), 4.35 (d, J = 3.4 Hz, 1H, H-4' adenosine), 4.30 (dd, J = 7.2, 4.4 Hz, 1H, H-3' ribosyl), 4.20 (s, 2H, H-5' adenosine), 4.18 – 4.08 (m, 4H, H-2' ribosyl + H-4' ribosyl + H-4' thymidine + CH_{2a} thymidine/ribosyl), 4.06 (dt, J = 11.4, 3.9 Hz, 1H, CH_{2a} thymidine/ribosyl), 4.04 – 3.98 (m, 2H, CH_{2b} thymidine + CH_{2b} ribosyl), 2.25 – 2.16 (m, 2H, H-2' thymidine), 1.83 (s, 3H, CH_3 thymine). **^{13}C NMR** (214 MHz, D_2O): δ 167.4 (C=O thymine), 155.6 (Cq adenine/thymine), 152.5 (CH adenine), 149.9 (Cq adenine/thymine), 141.1 (CH adenine, signal taken from HSQC) 138.1 (CH thymine), 119.5, 112.7, 112.5 (Cq adenine/thymine), 103.6 (C-1' ribosyl), 88 (C-1' adenosine).0, 86.9 (C-4' ribosyl), 86.0 (C-1' thymidine), 84.8 (C-4' adenosine), 82.5 (C-4' thymidine), 76.2 (C-2' ribosyl), 75.4 (C-2' adenosine), 72.5 (C-3' thymidine), 71.2, 71.2 (C-3' ribosyl + C-3' adenosine), 67.8, 66.4 (C-5' ribosyl + C5' thymidine), 66.1 (C-5' adenosine), 39.7 (C-2' thymidine), 12.6 (CH_3 thymine) . **^{31}P NMR** (202 MHz, D_2O, no EDTA was added): δ -1.7, -10.5. **HRMS:** $[C_{25}H_{36}N_7O_{21}P_3 + H]^+$ found: 864.1246, calculated: 864.1250.

References

[1] T. Takamura-Enya, M. Watanabe, Y. Totsuka, T. Kanazawa, Y. Matsushima-Hibiya, K. Koyama, T. Sugimura, K. Wakabayashi, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12414–12419.

[2] T. Nakano, Y. Matsushima-hibiya, M. Yamamoto, S. Enomoto, Y. Matsumoto, Y. Totsuka, M. Watanabe, T. Sugimura, K. Wakabayashi, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 13652–13657.

[3] L. Weixler, K. Schäringer, J. Momoh, B. Lüscher, K. L. H. Feijs, R. Zaja, *Nucleic Acids Res.* **2021**, *49*, 3634–3650.

[4] G. Jankevicius, A. Ariza, M. Ahel, I. Ahel, *Mol. Cell* **2016**, *64*, 1109–1116.

[5] E. Matta, A. Kiribayeva, B. Khassenov, B. T. Matkarimov, A. A. Ishchenko, *Sci. Rep.* **2020**, *10*, 1–11.

[6] D. Munnur, I. Ahel, *FEBS J.* **2017**, *284*, 4002–4016.

[7] I. Talhaoui, N. A. Lebedeva, G. Zarkovic, C. Saint-Pierre, M. M. Kutuzov, M. V. Sukhanova, B. T. Matkarimov, D. Gasparutto, M. K. Saparbaev, O. I. Lavrik, A. A. Ishchenko, *Nucleic Acids Res.* **2016**, *44*, 9279–9295.

[8] E. A. Belousova, A. A. Ishchenko, O. I. Lavrik, *Sci. Rep.* **2018**, *8*, 1–12.

[9] G. Zarkovic, E. A. Belousova, I. Talhaoui, C. Saint-Pierre, M. M. Kutuzov, B. T. Matkarimov, D. Biard, D. Gasparutto, O. I. Lavrik, A. A. Ishchenko, *Nucleic Acids Res.* **2018**, *46*, 2417–2431.

[10] D. Munnur, E. Bartlett, P. Mikolčević, I. T. Kirby, J. G. M. Rack, A. Mikoč, M. S. Cohen, I. Ahel, *Nucleic Acids Res.* **2019**, *47*, 5658–5669.

[11] T. Agnew, D. Munnur, K. Crawford, L. Palazzo, I. Ahel, *Front. Microbiol.* **2018**, *9*, 1–11.

[12] J. Groslambert, E. Prokorova, I. Ahel, *DNA Repair (Amst).* **2021**, *105*, 103144.

[13] H. A. V Kistemaker, L. N. Lameijer, N. J. Meeuwenoord, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2015**, *54*, 4915–4918.

[14] S. P. Adams, K. S. Kavka, E. J. Wykes, S. B. Holder, G. R. Galluppi, *J. Am. Chem. Soc.* **1983**, *105*, 661–663.

[15] A. P. Guzaev, R. T. Pon, *Curr. Protoc. Nucleic Acid Chem.* **2013**, *52*, 3.2.1-3.2.23.

[16] G. J. van der Heden van Noort, M. G. van der Horst, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *J. Am. Chem. Soc.* **2010**, *132*, 5236–5240.

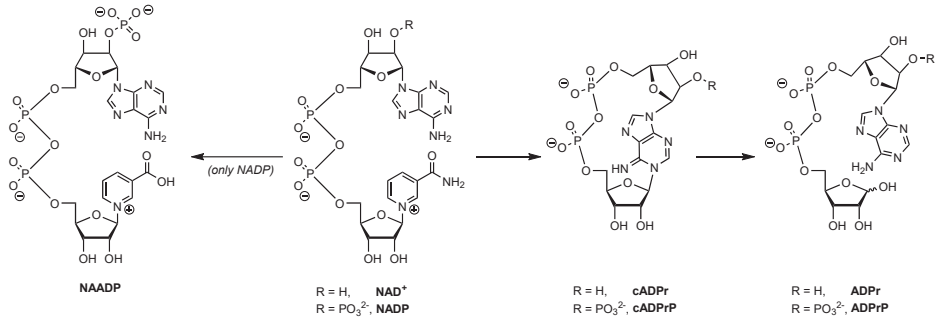
[17] M. H. Caruthers, *Science* **1981**, *230*, 281–285.

Chapter 7

Design and Synthesis of Potential Activity Based Probes for CD38 Based on Carba-Ribofuranose

Introduction

The cluster of differentiation 38 (CD38) proteins belong to the ADP-ribosyl cyclase family of which a type II, type III and a soluble type are known. While CD38 of type II is a glycoprotein with its catalytic site located outside the cell, type III is a non-glycosylated isoform with its catalytic site inside the cytosol. Although first discovered in thymocytes,[1] CD38 is ubiquitously expressed in most mammalian cells that have been examined.[2] CD38 can act as a membrane receptor[3,4] and binding with its ligand CD31 is, for example, involved in the cell adhesion of endothelial cells to lymphocytes and the modulation of cytoplasmic calcium fluxes.[5,6] CD38 is a multifunctional enzyme catalyzing not only the formation of cyclic ADPr (cADPr) out of nicotinamide adenine dinucleotide (NAD^+) but it can also cyclize NADP into cADPrP (Scheme 1). In addition, CD38 can hydrolyze cADPr and cADPrP into ADPr and ADPrP respectively.[7] Interestingly, it has also been established that CD38 can convert NAD phosphate (NADP) into nicotinic acid adenine dinucleotide (NAADP) in acidic environments.[8]



Scheme 1. Reactions catalyzed by CD38.

Both cADPr and NAADP are important signaling molecules that mobilize Ca^{2+},[9] a messenger that is involved in several pathways that regulate a variety of cellular processes.[10,11] In particular, Ca^{2+} homeostasis is of prime importance for immunity and the misregulation of Ca^{2+} signaling pathways in lymphocytes may lead to various autoimmune, inflammatory and immunodeficiency diseases. In this light, the importance of CD38 biology is underlined not only by its involvement in diseases such as diabetes,[12,13] AIDS[14,15] and chronic lymphocyte leukemia[16] but also in social behavior such as maternal care, which was severely reduced in CD38 knockout mice.[17,18] The importance of CD38 in health and disease and its role in multiple biological processes has resulted in a continued interest in elucidation of the corresponding biology that stands out for its complexity. In this framework, attention is directed

to the development of inhibitors of CD38. The catalysis by CD38 as an NAD^+ cyclase and hydrolase is thought to proceed *via* a non-covalent oxocarbenium intermediate. However, with substrates like 2-F-arabinose-nicotinamide mono nucleotide (2-ara-F-NMN), a covalent intermediate is formed whereby the glutamic acid (Glu) 226 residue in the CD38 active site performs a nucleophilic attack on 2-F-ara-NMN with concomitant expulsion of nicotinamide.[19–21] The latter proposed mechanism not only allows for the development of competitive[22,23] but also mechanism-based covalent inhibitors[24] modelled after its substrate NAD^+.[25] The availability of a covalent inhibitor allows for detection or even discovery of an active enzyme or enzyme family by activity-based protein profiling (ABPP) in the context of a biological system.[26–29] An activity-based probe (ABP) is a mechanism-based inhibitor provided with a tag to identify the corresponding enzyme. A fluorescent ABP has been reported by Hening Lin *et al.* that visualize CD38 on the cell membrane.[30] In addition, a cell-permeable ABP was developed that revealed the presence of CD38 in the cytoplasm albeit in significantly smaller quantities compared to its presence on the cell membrane.[31] These ABPs are based on the discovery of the group of Schramm that 2-F-ara-NMN can function as an electrophilic trap to form a covalent adduct with CD38. Likewise, the group of Withers has explored the use of 2'-deoxy-2'-fluoroglycosides as activity-based inhibitors and/or ABPs of glycosidases.[32–34] The group of Overkleeft discovered that differently configured cyclophellitol derivatives or similar configured cyclitols in which the epoxide is replaced by an aziridine moiety can function as efficient mechanism-based inhibitors and/or ABPs of several retaining glycosidases.[28,35–37]

Inspired by these achievements, candidate ABPs **1** and **2** derived from NAD^+ were designed (Figure 1A). The *carba*-ribosyl moiety in ABP **1** is functionalized with an aziridine electrophilic trap that is provided with an azide spacer as ligation handle. This enables the bio-orthogonal ligation with a reporter group such as a fluorophore or biotin tag via copper catalyzed, azide-alkyne cycloaddition (CuAAC). ABP **2** is functionalized with an epoxide as electrophilic trap and an alkyne spacer, positioned on the exocyclic amine of the adenine base, as ligation handle. The reporter groups ligated via CuAAC can provide a readout by in-gel fluorescence or mass spectrometry, applying either a pre-labeling protocol (Figure 1B) where the probe is already conjugated to the reporter group, or by a two-step labeling protocol in which the probe first reacts with the enzyme followed by an *in vitro* CuAAC with a suitable reporter tag.

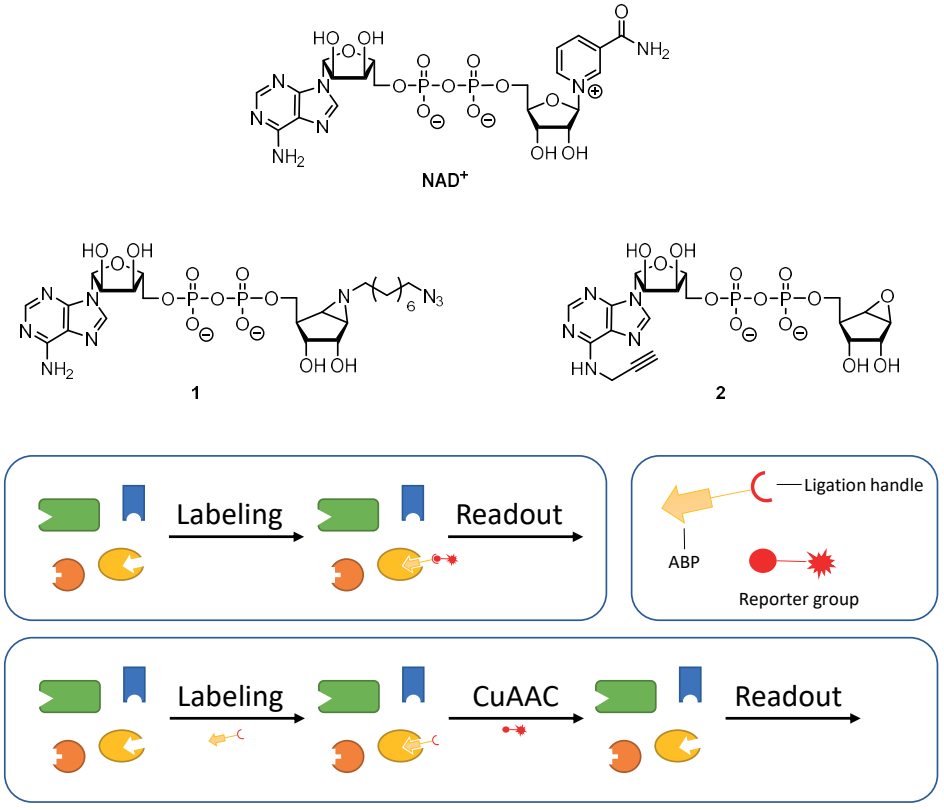
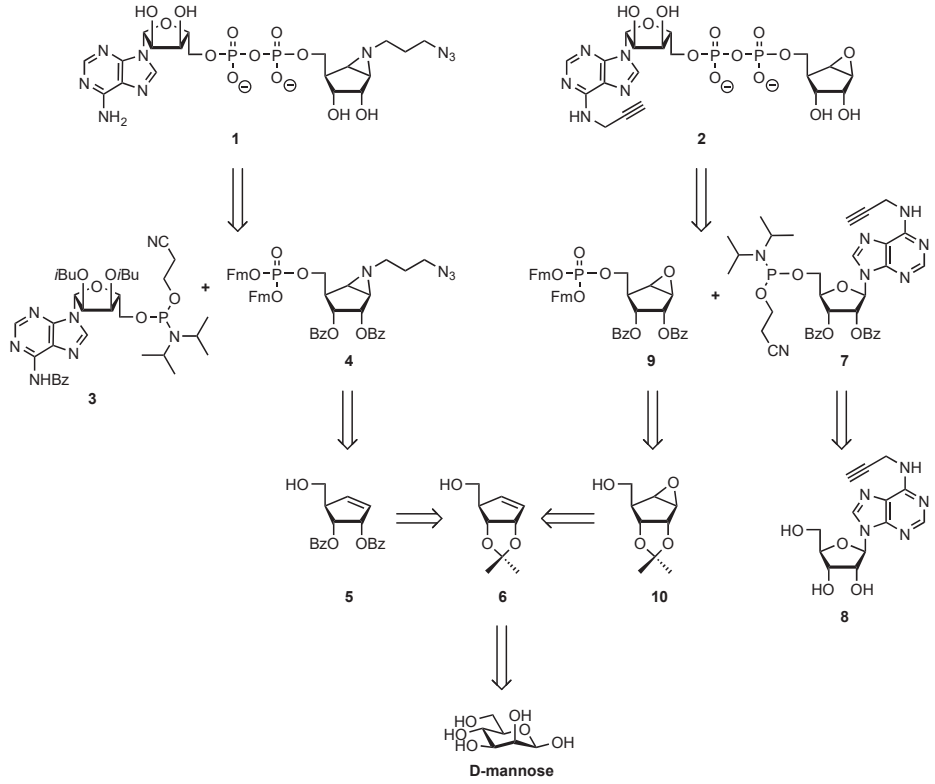


Figure 1. A) Design of ABPs **1** and **2** based on NAD^+. **B)** general workflow for one-step labeling (top) and two-step labeling (bottom). Readout is performed either by in-gel fluorescence or mass spectrometry analysis.

Results and discussion

The retrosynthetic analysis of ABPs **1** and **2** is depicted in Scheme 2. ABP **1** is provided with an azide as bio-orthogonal handle that is linked *via* a spacer at the aziridine moiety. ABP **2** bears a propargyl group placed at the exocyclic amine of adenosine. The pyrophosphate bridges in both **1** and **2** will be introduced using the established $P^{III} - P^{V}$ coupling method,[38] for which a suitably protected amidite and a phosphate monoester is needed. Thus, the four key building blocks required are: aziridine **4**, epoxide **9**, and adenosine phosphoramidites **3** and **7**. Known adenosine amidite **3**[39] is suitable to obtain aziridine-containing ABP **1**. Reagent **7**, decorated with the propargyl group, can be used to prepare ABP **2**. Phosphoramidite **7** is accessible *via* N^6-propargyladenosine **8**, which in turn can be obtained from adenosine by established procedures.[40,41] The phosphate monoesters needed for the pyrophosphate construction

are derived from warheads **4** and **9** after removal of the fluorenylmethyl (Fm) protecting groups. The phosphotriesters in **4** and **9** are accessible by phosphorylation with known bis-fluorenylmethyl phosphoramidite, followed by oxidation. Key to an efficient synthesis of aziridine **4** and epoxide **9** is the stereoselective introduction of the aziridine and the epoxide functional group. To achieve this, *carba*-ribose **6** is selected as suitable building block that can be efficiently obtained from D-mannose[42]. Compound **6** can be converted into epoxide **10** by $VO(acac)_2$ mediated epoxidation.[43,44] Upon standard protecting group manipulations on **6**, benzoylated cyclopentene **5** is conceivable. The homoallylic alcohol functionality in **5** can then be implemented in a four step reaction sequence, which has been successfully utilized before to synthesize several cyclophellitol-aziridines, [36,37,45] to give **4**.

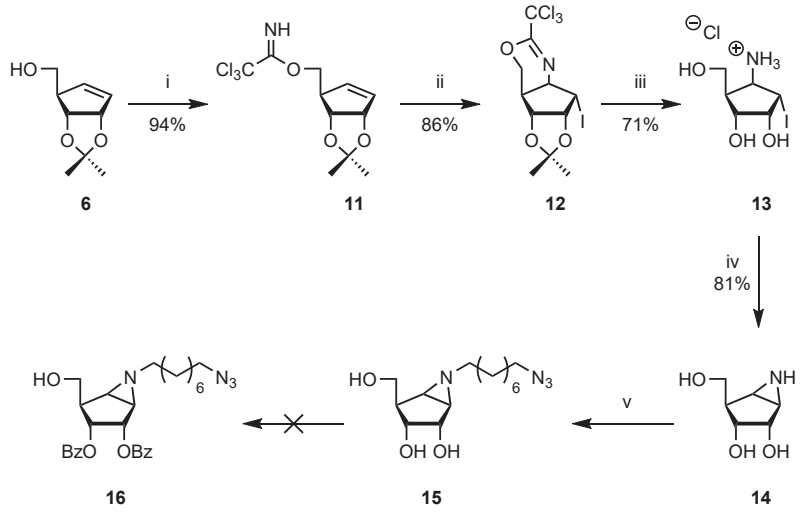


Scheme 2. Retrosynthetic analysis of ABPs **1** and **2** from D-mannose.

*Aziridine warhead (**4**)*

The first attempt to obtain *carba*-ribose **4**, containing an aziridine electrophilic trap is depicted in Scheme 3. The starting compound, isopropylidene protected **6**, was prepared from D-mannose according to a known 8-step procedure.[42] The installation of the aziridine

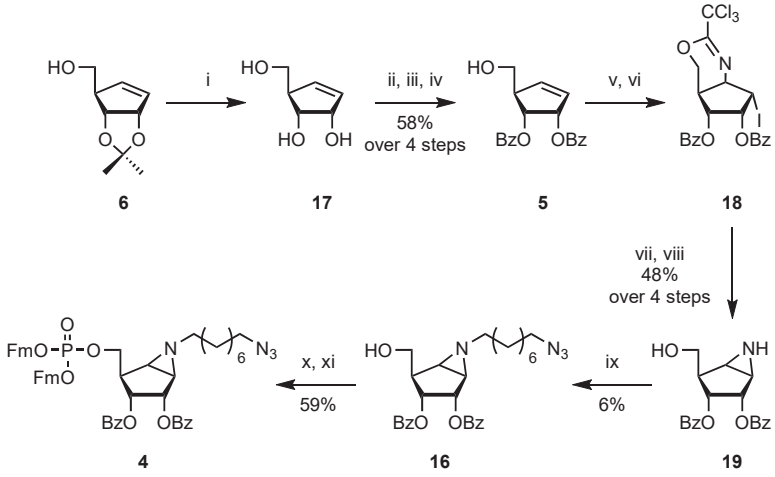
moiety was adopted from an established method to obtain aziridine cyclophellitols. [45] In the first step, the primary homo-allylic alcohol of **6** was transformed into trichloroacetimidate **11,** using DBU and trichloroacetonitrile. Subsequent iodocyclization by treatment of **11** with NIS gave imidate **12** in 86% yield. Simultaneous acid-mediated hydrolysis of both the imidate and the isopropylidene group resulted in HCl salt **13**. The crude product of this reaction was directly treated with basic Amberlite IRA67 resin to induce nucleophilic attack of the amine on the neighboring iodide, leading to the isolation of aziridine **14**. Subsequent chemoselective alkylation of the aziridine in **14** with 1-azido-8-iodooctane provided **15** with the prospected ligation handle. Unfortunately, all undertaken protecting group manipulations to make **15** suitable for the phosphorylation to give target **4** proofed to be cumbersome. Probably the aziridine group does not withstand the applied reaction conditions and thus it was decided to install the aziridine group later in the synthesis route.



Scheme 3. First attempt at the synthesis of the aziridine containing warhead. Reagents and conditions: i) Cl_3CCN, DBU, DCM, 0 °C. ii) NIS, MeCN, 0 °C. iii) HCl, MeOH, H_2O. iv) Amberlite IRA67 base resin, MeOH, 45 °C. v) 1-azido-8-iodooctane, K_2CO_3, DMF.

In order to facilitate the late-stage introduction of the aziridine moiety, the hydrolysis of isopropylidene group in **6** (Scheme 4) to afford triol **17** was followed by a three-step protecting group manipulation sequence, consisting of (i) regioselective introduction of a *tert*-butyl diphenyl silyl ether on the 5-OH, (ii) benzoylation of the 2- and 3-OH and (iii) F^- mediated removal of the silyl group to afford protected homoallylic alcohol **5**. Installation of the aziridine according to the procedure as described for the conversion of **11** into **14**,

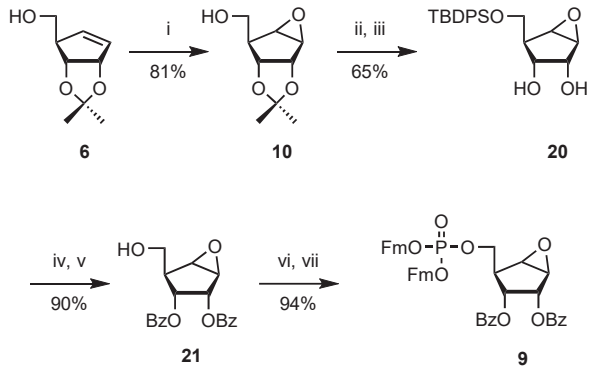
furnished the 2- and 3-*O*-Bz protected aziridine **19**. Subjection of **19** to the chemoselective alkylation as described in Scheme 3 provided **16** with the azide linker albeit in low yield. Finally, phosphitylation of the 5-OH with bis-fluorenylmethylphosphoramidite in the presence of activator ETT and oxidation of the phosphite intermediate with CSO yielded fully protected aziridine **4**, ready for pyrophosphate formation towards the target probe **1**.



Scheme 4. Synthesis of key aziridine containing warhead **4**. Reagents and conditions: i) AcOH, H_2O, 60 °C. ii) TBDPS-Cl, pyr. iii) BzCl, pyr. iv) TEA·3HF, THF. v) Cl_3CCN, DBU, DCM, 0 °C. vi) NIS, MeCN, 0 °C. vii) HCl, MeOH, H_2O. viii) TEA, MeOH, 45 °C. ix) 1-azido-8-iodooctane, K_2CO_3, DMF, 80 °C. x) $(FmO)_2PN(iPr)_2$, ETT, MeCN. xi) CSO, MeCN.

*Epoxide warhead (**9**)*

The route of synthesis towards epoxide building block **9** commenced with the $VO(acac)_2$ mediated, stereoselective epoxidation of homoallylic alcohol **6** to give **10** in 81% yield (Scheme 5). It is interesting to note that this epoxidation was unsuccessful with benzoylated alkene **5**. It can be postulated that the electron withdrawing character of the benzoyl protecting group inhibits the epoxidation. Acidolysis of the isopropylidene group in **10** after which the resulting triol was subjected to the same protecting group manipulations as described for the conversion of **17** into **5** (Scheme 4) furnished compound **21**. Phosphitylation and oxidation of **21** as described above, led to key epoxide **9** ready for use in the pyrophosphate formation towards target probe **2**.



Scheme 5. Synthesis of epoxide warhead **9**. Reagents and conditions: i) $VO(acac)_2$, *t*BuOOH, toluene, 60 °C. ii) AcOH, H_2O, 60 °C. iii) TBDPS-Cl, pyr, 65% over two steps. iv) BzCl, pyr. 90%. v) TEA·3HF, THF. vi) $FmO)_2PN(iPr)_2$, ETT, MeCN. vii) CSO, MeCN.

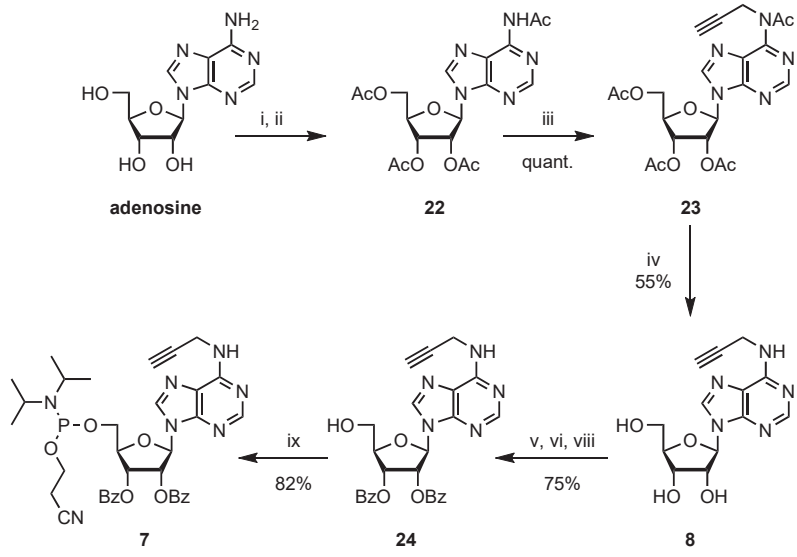
Adenosine building block

To execute the projected P^{III} – P^{V} coupling method for the pyrophosphate formation, the protected adenosine amidites **3** and **7** are needed for aziridine ABP **1** and epoxide APP **2**, respectively. Amidite **3** is a known compound[39] and the synthesis of amidite **7** is depicted in Scheme 6. The synthesis started with per-acetylation of commercially available adenosine with acetic anhydride in pyridine to give intermediate penta-acetate, bearing two acetyls on the exocyclic amine. One of the acetyl groups was removed by treatment with imidazole to give tetra-acetate **22**. After alkylation of **22** with propargyl bromide to afford **23** in quantitative yield, the remaining acetyl protecting groups were cleaved by sodium methoxide in methanol. The resulting triol **8** was subjected to the same three-step protecting group manipulation comprising of regioselective silylation of the 5-OH, benzoylation of the 2- and 3-OH and liberation of the 5-OH as described *en route* to aziridine **4** and epoxide **9** (Scheme 4 and 5 respectively), yielding primary alcohol **24** in 75% yield. Finally, reaction of the 5-OH in **24** with commercially available cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite in the presence of DIPEA furnished adenosine amidite **7**.

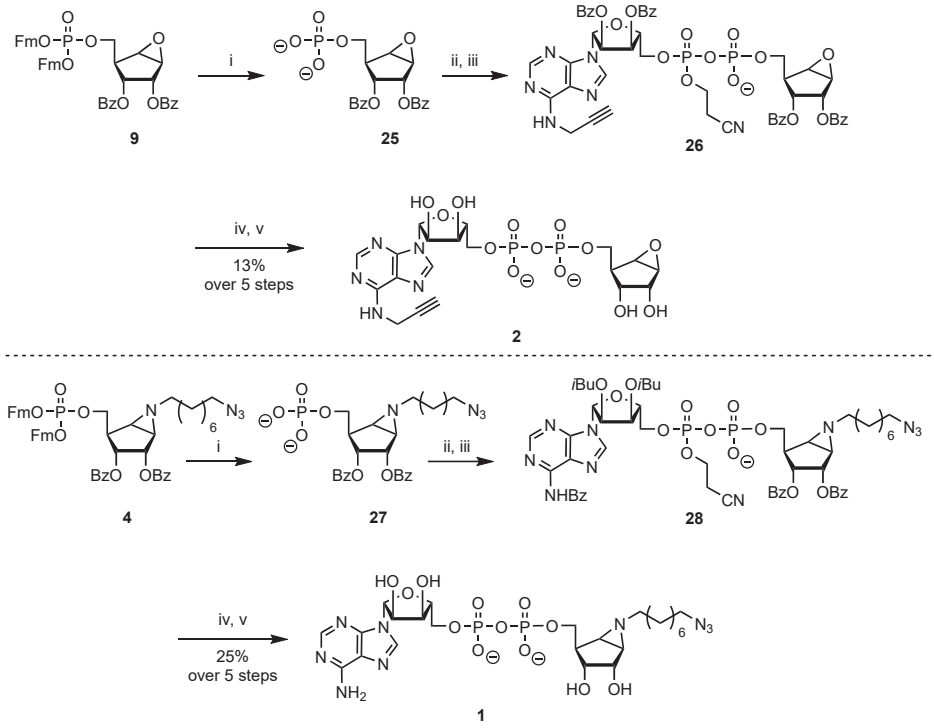
*Synthesis of ABPs **1** and **2***

With all building blocks in hand, the synthesis of ABPs **1** and **2** was undertaken by employing the P^{III} – P^{V} coupling strategy outlined in Scheme 7. Toward ABP **2**, the Fm protecting groups in phosphotriester **9** were removed by treatment with TEA to give the corresponding monoester **25**. The reaction was monitored by LC-MS analysis which indicated complete removal of the Fm groups after 24 hours. Co-evaporation of the reaction mixture with pyridine to convert the phosphomonoester into the more stable pyridinium salt. Next, **25** was treated with phosphoramidite **7** using ETT as an activator. The reaction was monitored

by ^{31}P-NMR, showing complete conversion into a phosphite-phosphate intermediate after 15 minutes. The intermediate was then oxidized by *t*BuOOH furnishing the partially protected pyrophosphate **26**. Finally, deprotection was performed by treating intermediate **26** with DBU to remove the cyanoethyl group and subsequently with ammonia to cleave all benzoyl groups. Purification with preparative HPLC led to the isolation of ABP **2** in 13% yield over 5 steps. For ABP **1** identical reaction conditions were applied with phosphotriester **4** and phosphoramidite **3**, furnishing ABP **1** in 25% yield after HPLC purification.



Scheme 6. Synthesis of adenosine amidite **7**. Reagents and conditions: i) Ac_2O, pyr. ii) Imidazole, MeOH. iii) propargyl bromide, DBU, MeCN. iv) NaOMe, MeOH. v) TBS-Cl, pyr. vi) BzCl, pyr. vii) PTSA, MeCN, H_2O. viii) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite, DIPEA, DCM.



Scheme 7. Synthesis of ABPs **1** and **2**. Reagents and conditions: i) TEA, MeCN. ii) **3** or **7**, ETT, MeCN. iii) *t*BuOOH. iv) DBU, DMF. v) NH_4OH, H_2O.

Conclusion

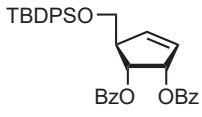
Two potential ABPs for the covalent inhibition of CD38 were designed and successfully synthesized. Both designs are based on NAD^+, the natural substrate of CD38. The probes bear electrophilic traps as a warhead to covalently bind the Glu residue in the active site of CD38. An aziridine was used as the warhead in case of ABP **1** and an epoxide-warhead was employed for ABP **2**. Both ABPs **1** and **2** are designed for one- or two-step labeling strategies (Figure 1B) and are equipped with bio-orthogonal ligation handles that would enable functionalization with a desired reporter group via CuAAC. Synthesis of both probes was executed *via* elaborate multistep procedures commencing with the deprotection of warheads **4** and **9**. Next, the 5-O-phosphates **25** and **27** were subject for the established P^{III} – P^{V} coupling procedure with appropriately protected adenosine phosphoramidites to furnish the prospected ABPs **1** and **2**. The activity of ABPs **1** and **2** towards CD38 will be evaluated *in vitro* in due course.

Experimental section

General synthetic procedures

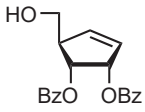
All reagents were of commercial grade and used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. 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 F_{254}). TLC was used to visualize compounds by UV at wavelength 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) or $KMnO_4$ spray (20 g/L $KMnO_4$ and 10 g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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. NMR spectra were recorded on a Bruker AV-400, AV-500 or AV-600 NMR. Chemical shifts (δ) are given in ppm relative to tetramethyl silane as internal standard. Coupling constants (*J*) are given in Hz. All given ^{13}C-APT spectra are proton decoupled. Proton and carbon numbering for NMR peak assignment, in case of the carbocyclic derivatives, was done similarly to their ribofuranosyl counterparts. Numbering starts at the 'anomeric' centre and progresses as for ribofuranose. 'H-6' or 'C-6' is used where the endocyclic oxygen is replaced for the substituted carbon.

(3*S*,4*R*,5*R*)-3,4,-*O*-dibenzoyl-5-(*O*-*tert*-butyldiphenylsilyl)-methylcyclopentene



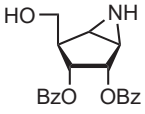
Compound **6**[42] (1.70 g, 10.0 mmol) was dissolved in a 80 v/v% solution of AcOH in H_2O (40 mL, 0.25 M). The reaction was stirred in an open flask at 60 °C for 2 hours. TLC indicated full conversion of the starting material to a lower running product (Rf = 0.3 in 10% EtOH in EtOAc). The reaction was concentrated *in vacuo* and thoroughly co-evaporated with a 1:1 mixture MeCN:H_2O followed by co-evaporation with 1,4-dioxane. The crude triol was than co-evaporated with pyridine thrice and dissolved in pyridine (100 mL, 0.1 M). TBDPS-Cl (2.86 mL, 11.0 mmol, 1.1 eq.) was added and the reaction was stirred overnight. Another portion of TBDPS-Cl (2.86 mL, 11.0 mmol, 1.1 eq.) was added and an additional 7 hours of stirring ensured full conversion of the triol into a higher running product. BzCl (4.64 mL, 40 mmol, 4.0 eq.) was added and the reaction was stirred overnight. The reaction was quenched with MeOH and concentrated *in vacuo*. The crude residue was taken up in EtOAc and the solution was washed with 1 M citric acid, sat. aq. $NaHCO_3$ and brine respectively. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (10 -> 15% Et_2O in pentane) yielded the title compound as a colorless oil (4.33 g, 7.51 mmol, 75%). **Rf:** 0.76 in 20% Et_2O in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.92 (ddd, J = 17.9, 8.1, 1.4 Hz, 4H, Bz arom.), 7.72 – 7.64 (m, 4H, TBDPS arom.), 7.54 – 7.44 (m, 2H, Bz arom.), 7.43 – 7.26 (m, 10H, Bz arom. + TBDPS arom.), 6.18 – 6.00 (m, 3H, H-1 + H-2 + H-6), 5.62 (dd, J = 6.6, 4.2 Hz, 1H, H-3), 3.87 (qd, J = 10.1, 5.0 Hz, 2H, H-5), 3.30 (q, J = 5.2 Hz, 1H, H-4), 1.05 (s, 9H, *t*Bu TBDPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.1, 166.0 (C=O Bz), 137.6 (C-6), 135.8, 135.7 (CH arom. TBDPS), 133.4, 133.4 (Cq arom. TBDPS), 133.0, 133.0 (CH arom. Bz), 130.2, 130.0 (Cq Bz), 129.9, 129.8, 129.7 (CH arom. Bz + CH arom. TBDPS), 129.4 (C-1), 128.3, 128.3, 127.9, 127.8 (CH arom. Bz + CH arom. TBDPS), 76.8 (C-2), 74.1 (C-3), 63.9 (C-5), 52.4 (C-4), 26.9 (CH_3 *t*Bu TBDPS), 19.4 (Cq *t*Bu TBDPS). **HRMS:** $[C_{36}H_{36}O_5Si + Na]^+$ found: 599.2221, calculated: 599.2224.

(3*S*,4*R*,5*R*)-3,4,-*O*-dibenzoyl-5-hydroxymethylcyclopentene (5)



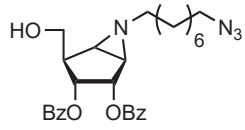
(3*S*,4*R*,5*R*)-3,4,-*O*-dibenzoyl-5-(*O*-*tert*-butyldiphenylsilyl)-methylcyclopentene (3.09 g, 5.36 mmol) was dissolved in THF (54 mL, 0.1 M) and the solution was cooled to 0 °C. TEA·3HF (17.5 mL, 107 mmol, 20.0 eq.) was added and the reaction was slowly allowed to warm to room temperature. After 4 days of stirring, the reaction was cooled to 0 °C and *carefully* quenched by the addition of sat. aq. $NaHCO_3$ until bubbling seized. The reaction was transferred into a separatory funnel and the water layer was extracted thrice with EtOAc. The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (30 -> 40% EtOAc in pentane) yielded the title compound as a colorless oil (1.40 g, 4.12 mmol, 77%). **Rf:** 0.41 in 30% EtOAc in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.99 – 7.93 (m, 2H, Bz arom.), 7.93 – 7.87 (m, 2H, Bz arom.), 7.53 – 7.43 (m, 2H, Bz arom.), 7.33 (t, *J* = 7.8 Hz, 2H, Bz arom.), 7.30 – 7.23 (m, 2H, Bz arom.), 6.13 (dd, *J* = 5.4, 2.0 Hz, 1H, H-6), 6.11 – 6.04 (m, 2H, H-1 + H-2), 5.63 – 5.48 (m, 1H, H-3), 3.91 – 3.74 (m, 2H, H-5), 3.34 – 3.22 (m, 1H. H-4), 3.04 (bs, 1H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.5, 166.0 (C=O Bz), 137.2 (C-6), 133.2, 133.0 (CH arom. Bz), 129.9 (Cq Bz), 129.7, 129.7 (CH arom. Bz), 129.5 (Cq Bz), 129.4 (C-1), 128.3, 128.3 (CH arom. Bz), 76.7 (C-2), 74.1 (C-3), 62.8 (C-5), 52.9 (C-4). **HRMS:** $[C_{20}H_{18}O_5 + Na]^+$ found: 361.1047, calculated: 361.1046.

(1*R*,2*S*,3*R*,4*R*)-1,2-aziridine-3,4-*O*-dibenzoyl-5-hydroxymethylcyclopentane (19)



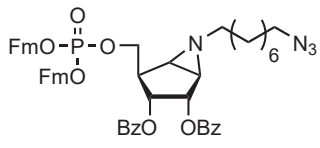
Compound **5** (1.25 g, 3.68 mmol) was dissolved in DCM (25 mL, 0.15 M) before DBU (0.22 mL, 1.47 mmol, 0.4 eq.) and trichloroacetonitrile (0.41 mL, 4.05 mmol, 1.1 eq.) were added. The reaction was stirred for 15 minutes before the reaction was diluted with DCM and the solution was washed with brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude imidate was dissolved in MeCN (12.5 mL, 0.3 M) and NIS (4.14 g, 18.4 mmol, 5.0 eq.) was added to the solution. The flask was wrapped in aluminum foil to protect it from light and stirred overnight. The reaction was diluted with EtOAc and the solution was washed with sat. aq. $Na_2S_2O_3$. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The resulting crude oxazoline was dissolved in 1,4-dioxane (37 mL, 0.1 M) and HCl (37% in H_2O, 11 mL) was added. The reaction was heated to 60 °C and stirred for 2 hours. The reaction was concentrated *in vacuo* and extensively co-evaporated with MeOH. The crude amine was dissolved in MeOH (37 mL, 0.1 M) and TEA (5.13 mL, 38.8 mmol, 10 eq.) was added. The reaction was stirred overnight at 55 °C. The reaction was concentrated *in vacuo* and re-dissolved in MeOH (37 mL, 0.1 M). Another portion of TEA (5.13 mL, 36.8 mmol, 10 eq.) was added and the reaction was refluxed for 14 hours to ensure full conversion of the amine to the aziridine. The reaction was concentrated *in vacuo* and flash column chromatography (5% MeOH in DCM) yielded the title compound as a 1:1.7 mixture product:TEA (calculated by ^{1}H-NMR) (938 mg, 1.78 mmol, 48%). **Rf**: 0.38 in 10% MeOH in DCM. **^{1}H NMR** (400 MHz, MeOD) δ 7.93 – 7.87 (m, 2H, Bz arom.), 7.87 – 7.79 (m, 2H, Bz arom.), 7.61 – 7.48 (m, 2H, Bz arom.), 7.40 (t, *J* = 7.8 Hz, 2H, Bz arom.), 7.31 (t, *J* = 7.8 Hz, 2H, Bz arom.), 5.68 (d, *J* = 5.1 Hz, 1H, H-2), 4.98 (dd, *J* = 8.6, 5.1 Hz, 1H, H-3), 3.92 – 3.78 (m, 2H, H-5), 3.07 (dd, *J* = 4.5, 2.6 Hz, 1H, H-6), 2.99 (d, *J* = 4.5 Hz, 1H, H-1), 2.76 (tdd, *J* = 8.3, 5.6, 2.6 Hz, 1H, H-4). **^{13}C NMR** (101 MHz, MeOD) δ 166.8, 166.7 (C=O Bz), 134.4, 134.3 (CH arom. Bz), 130.5 (Cq Bz), 130.3, 130.2, 129.5, 129.3 (CH arom. Bz), 74.8 (C-3), 73.1 (C-2), 62.0 (C-5), 46.4 (C-4), 36.3 (C-6), 35.7 (C-1).

(1*R*-2*S*-3*R*-4*R*)-1,2-(6-*N*-oct-(1-azide)-ane)aziridine-3,4-*O*-dibenzoyl-5-hydroxymethylcyclopentane (16)



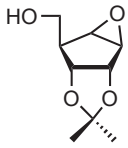
Compound **19** (930 mg, 1.78 mmol) was dissolved in DMF (6 mL, 0.3 M). 6-Iodo-1-azido-octane (1.00 g, 3.56 mmol, 2.0 eq.) and K_2CO_3 (296 mg, 2.14 mmol, 1.2 eq.) were added and the reaction was stirred overnight at 100 °C. The reaction was concentrated *in vacuo* and purification by flash column chromatography (3 -> 5% acetone in DCM) yielded the title compound (53 mg, 0.11 mmol, 6%). **Rf:** 0.48 in 5% acetone in DCM. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.02 – 7.93 (m, 2H, Bz arom.), 7.88 (dd, *J* = 8.3, 1.4 Hz, 2H, Bz arom.), 7.58 – 7.50 (m, 1H, Bz arom.), 7.50 – 7.44 (m, 1H, Bz arom.), 7.38 (t, *J* = 7.7 Hz, 2H, Bz arom.), 7.32 – 7.25 (m, 2H, Bz arom.), 5.69 (d, *J* = 5.2 Hz, 1H, H-2), 5.17 (dd, *J* = 8.3, 5.2 Hz, 1H, H-3), 4.03 (dd, *J* = 10.7, 4.5 Hz, 1H, H-5_a), 3.92 (dd, *J* = 10.7, 6.3 Hz, 1H, H-5_b), 3.27 (t, *J* = 6.9 Hz, 2H, RNCH$_2$CH$_2$(CH$_2$)$_4$CH$_2$C*H*$_2$N$_3$), 2.61 (dddd, *J* = 8.7, 6.7, 4.4, 2.6 Hz, 1H, H-4), 2.44 – 2.30 (m, 3H, H-1 + H-6 + RNC*H*$_{2a}$CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$), 2.21 – 2.02 (m, 1H, RNC*H*$_{2b}$CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$), 1.66 – 1.51 (m, 4H, RNC*H*$_2$CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$ + RNCH$_2$CH$_2$(CH$_2$)$_4$C*H*$_2$CH$_2$N$_3$), 1.38 – 1.29 (m, RNCH$_2$CH$_2$(C*H*$_2$)$_4$CH$_2$CH$_2$N$_3$ 8H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.0, 165.7 (C=O Bz), 133.2, 133.1 (CH arom. Bz), 129.9 (Cq Bz), 129.7, 129.7 (CH arom. Bz), 129.6 (Cq Bz), 128.4, 128.3 (CH arom. Bz), 74.1 (C-3), 72.3 (C-2), 62.2 (C-5), 58.4 (RNCH_2CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$), 51.5 (RNCH$_2$CH$_2$(CH$_2$)$_4$CH$_2$C*H*$_2$N$_3$), 45.5 (C-4), 44.1, 43.0 (C-1 + C-6), 29.7, 29.4, 29.1, 28.9 (RNCH$_2$CH$_2$(CH$_2$)$_4$C*H*$_2$CH$_2$N$_3$ + RNCH$_2$C*H*$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$ + RNCH$_2$CH$_2$(C*H*$_2$)$_4$CH$_2$CH$_2$N$_3$), 27.2, 26.7 (RNCH$_2$CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$). **HRMS:** $[C_{28}H_{34}N_4O_5 + H]^+$ found: 507.2607, calculated: 507.2602.

(1*R*-2*S*-3*R*-4*R*)-1,2- (6-*N*-oct-(1-azide)-ane)aziridine-3,4-*O*-dibenzoyl-5-(*O*-difluorenylmethyl phosphate)-methylcyclopentane (4)



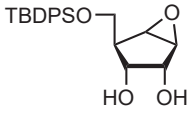
Compound **16** (53 mg, 0.11 mmol) was co-evaporated in toluene and dissolved in a solution of bis-(9H-fluoren-9-ylmethyl)-*N,N*-diisopropylamidophosphite (110 mg, 0.21 mmol, 2.0 eq.) and DCI (0.25 M, 0.42 mmol, 4.0 eq.) in MeCN (1.7 mL, 0.06 M). The reaction was stirred for 30 minutes. *t*BuOOH (5.5 M in nonane, 0.2 mL, 1.0 mmol, 10 eq.) was added and the reaction was stirred for an additional 45 minutes. The reaction was diluted with EtOAc and the solution was washed with sat. aq. $NaHCO_3$ and brine respectively. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (40% EtOAc in pentane) yielded the title compound (61 mg, 65 µmol, 59%). **Rf:** 0.64 in 40% EtOAc in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.98 – 7.90 (m, 2H, Bz arom.), 7.89 – 7.80 (m, 2H, Bz arom.), 7.74 – 7.63 (m, 4H, Fm arom.), 7.61 – 7.39 (m, 7H, Bz arom. + Fm arom.), 7.42 – 7.25 (m, 8H, Bz arom. + Fm arom.), 7.28 – 7.12 (m, 5H, Bz arom. + Fm arom.), 5.65 (d, *J* = 5.1 Hz, 1H, H-2), 4.90 (dd, *J* = 8.5, 5.1 Hz, 1H, H-3), 4.37 – 4.01 (m, 8H, H-5 + 2x CH Fm + 2x CH_2 Fm), 3.21 (t, *J* = 7.0 Hz, 2H, RNCH$_2$CH$_2$(CH$_2$)$_4$CH$_2$C*H*$_2$N$_3$), 2.69 (tdd, *J* = 8.7, 6.0, 2.6 Hz, 1H, H-4), 2.36 (d, *J* = 4.6 Hz, 1H, H-1), 2.31 – 2.15 (m, 2H, H-6 + RNC*H*$_{2a}$CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$), 2.09 – 1.93 (m, 1H, RNC*H*$_{2b}$CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$), 1.62 – 1.50 (m, 2H, RNCH$_2$CH$_2$(CH$_2$)$_4$C*H*$_2$CH$_2$N$_3$), 1.51 – 1.37 (m, 2H, RNCH$_2$C*H*$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$), 1.37 – 1.14 (m, 8H, RNCH$_2$CH$_2$(C*H*$_2$)$_4$CH$_2$CH$_2$N$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.6, 165.5 (C=O Bz), 143.1, 143.1, 143.0, 143.0, 141.4, 141.4 (Cq Fm), 133.2, 133.1 (CH arom. Bz), 129.8 (Cq Bz), 129.7, 129.7 (CH arom. Bz), 129.3 (Cq Bz), 128.4, 128.3, 127.9, 127.2, 127.1 (CH arom. Bz + CH arom. Fm), 125.1, 125.1, 125.1, 120.0, 120.0 (CH arom. Fm), 74.1 (C-3), 72.1 (C-2), 69.4, 69.3, 69.3, 69.3 (2x CH_2 Fm), 67.3, 67.2 (C-5), 58.2 (RNCH_2CH$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$), 51.5 (RNCH$_2$CH$_2$(CH$_2$)$_4$CH$_2$C*H*$_2$N$_3$), 48.0, 48.0, 47.9, 47.9 (2x CH Fm), 44.2, 44.2 (C-4), 43.6 (C-1), 43.0 (C-6), 29.6, 29.4, 29.1, 28.8 (RNCH$_2$CH$_2$(CH$_2$)$_4$C*H*$_2$CH$_2$N$_3$ + RNCH$_2$C*H*$_2$(CH$_2$)$_4$CH$_2$CH$_2$N$_3$ + RNCH$_2$CH$_2$(C*H*$_2$)$_4$CH$_2$CH$_2$N$_3$), 27.1, 26.7 (RNCH$_2$CH$_2$(C*H*$_2$)$_4$CH$_2$CH$_2$N$_3$). **^{31}P NMR** (162 MHz, $CDCl_3$) δ -1.1. **HRMS:** $[C_{56}H_{55}N_4O_8P + H]^+$ found: 943.3833, calculated: 948.3830.

(1*R*,2*R*,3*R*,4*R*,5*S*)-1,2-epoxy-3,4-*O*-isopropylidene-5-hydroxymethylcyclopentane (10)



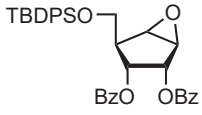
Compound **6**[42] (518 mg, 3.00 mmol) was co-evaporated with toluene and dissolved in toluene. To this solution, vanadyl acetylacetonate (80 mg, 0.3 mmol, 0.1 eq.) and *t*BuOOH (5.5 M in nonane, 2.2 mL, 12.0 mmol, 4.0 eq.) were added and the solution was stirred overnight at 60 °C. The reaction was concentrated *in vacuo*. Flash column chromatography (30 -> 50% EtOAc) furnished the title compound as a colorless oil (454 mg, 2.44 mmol, 81%). **Rf:** 0.21 in 30% EtOAc in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.65 (dd, *J* = 5.6, 0.8 Hz, 1H, H-2), 4.26 (dd, *J* = 5.7, 2.2 Hz, 1H, H-3), 3.91 – 3.76 (m, 2H, H-5), 3.70 (tt, *J* = 2.1, 1.0 Hz, 1H, H-6), 3.65 (d, *J* = 2.4 Hz, 1H, H-1), 2.41 (ddt, *J* = 7.3, 6.3, 2.3 Hz, 1H, H-4), 1.48 (s, 3H, CH_3 isopropylidene), 1.32 (s, 3H, CH_3 isopropylidene).**^{13}C NMR** (101 MHz, $CDCl_3$) δ 112.3 (Cq isopropylidene), 83.0 (C-3), 79.9 (C-2), 61.9 (C-5), 60.2 (C-6), 58.2 (C-1), 49.8 (C-4), 27.1, 24.7 (CH_3 isopropylidine).

(1*R*,2*R*,3*R*,4*R*,5*S*)-1,2-epoxy -5-(*O*-*tert*-butyldiphenylsilyl)-methylcyclopentane (20)



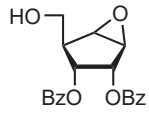
Compound **10** (393 mg, 2.11 mmol) was dissolved in an 80 v/v% solution of AcOH in H_2O (9 mL, 0.25 M). The reaction was stirred in an open flask at 60 °C for 2 hours. TLC indicated full conversion of the starting material to a lower running product (Rf = 0.3 in 10% EtOH in EtOAc). The reaction was concentrated *in vacuo* and thoroughly co-evaporated with a 1:1 mixture MeCN:H_2O followed by co-evaporation with 1,4-dioxane. The crude triol was than co-evaporated with pyridine thrice and dissolved in pyridine (21 mL, 0.1 M). TBDPS-Cl (1.08 mL, 4.22 mmol, 2.0 eq.) was added and the reaction was stirred overnight. The reaction was quenched by the addition of MeOH and the reaction was concentrated *in vacuo*. The resulting crude residue was taken up in EtOAc and the suspension was washed with 1 M HCl and brine respectively. The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The title compound was obtained after flash column chromatography (30 -> 40% EtOAc in pentane) as a colorless oil (529 mg, 1.37 mmol, 65% over two steps). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.73 – 7.62 (m, 4H, TBDPS arom.), 7.46 – 7.29 (m, 6H, TBDPS arom.), 4.11 (d, J = 5.2 Hz, 1H, H-2), 3.98 – 3.81 (m, 2H, H-5), 3.59 – 3.50 (m, 2H, H-3 + H-6), 3.48 (dd, J = 2.8, 0.8 Hz, 1H, H-1), 2.92 (bs, 1H, OH), 2.22 (tdd, J = 8.0, 6.6, 1.2 Hz, 1H, H-4), 1.07 (s, 9H, *t*Bu TBDPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 135.6 (CH arom. TBDPS), 133.3, 133.2 (Cq arom. TBDPS), 129.9, 127.9 (CH arom. TBDPS), 72.6 (C-3), 69.9 (C-2), 63.4 (C-5), 57.2 (C-6), 56.2 (C-1), 47.8 (C-4), 26.9 (CH_3 *t*Bu TBDPS), 19.2 (Cq *t*Bu TBDPS).

(1*R*,2*R*,3*R*,4*R*,5*S*)-1,2-epoxy-3,4-*O*-di-benzoyl-5-(*O*-*tert*-butyldiphenylsilyl)-methylcyclopentane



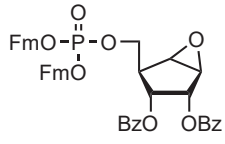
Compound **20** (442 mg, 1.15 mmol) was dissolved in pyridine (12 mL, 0.1 M). BzCl (0.53 mL, 4.6 mmol, 4.0 eq.) was added and the reaction was stirred 4 hours. The reaction was quenched by the addition of MeOH and concentrated *in vacuo*. The crude residue was taken up in EtOAc and the organic layer was washed with 1 M HCl and brine respectively, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (10 -> 20% Et_2O in pentane) yielded the title compound as a colorless oil (610 mg, 1.03 mmol, 90%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.95 – 7.89 (m, 2H, Bz arom.), 7.85 – 7.79 (m, 2H, Bz arom.), 7.72 – 7.62 (m, 4H, TBDPS arom.), 7.56 – 7.43 (m, 2H, Bz arom.), 7.43 – 7.31 (m, 8H, TBDPS arom. + Bz arom.), 7.31 – 7.24 (m, 2H, Bz arom.), 5.79 (d, J = 5.4, 0.8 Hz, 1H, H-2), 4.96 (dd, J = 8.2, 5.4 Hz, 1H, H-3), 4.04 – 3.93 (m, 2H, H-5), 3.87 (dd, J = 2.8, 1.2 Hz, 1H, H-6), 3.81 (dd, J = 2.7, 0.9 Hz, 1H, H-1), 2.84 – 2.73 (m, 1H, H-4), 1.06 (s, 9H, *t*Bu TBDPS). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.5, 165.5 (C=O Bz), 135.7, 135.6 (CH arom. TBDPS), 133.5 (Cq arom. TBDPS), 133.4 (CH arom. Bz), 133.3 (Cq arom. TBDPS), 133.2 (CH arom. Bz), 129.9, 129.8, 129.8, 129.7 (CH arom. Bz + CH arom. TBDPS), 129.5, 129.4 (Cq arom. Bz), 128.5, 128.3, 127.8, 127.8 (CH arom. Bz + CH arom. TBDPS), 72.6 (C-3), 70.9 (C-2), 62.4 (C-5), 57.0 (C-6), 54.8 (C-1), 46.2 (C-4), 26.9 (CH_3 *t*Bu TBDPS), 19.3 (Cq *t*Bu TBDPS). **HRMS:** $[C_{36}H_{36}O_6Si + H]^+$ found: 593.2350, calculated: 593.2354.

(1*R*,2*R*,3*R*,4*R*,5*S*)-1,2-epoxy-3,4-*O*-di-benzoyl-5-hydroxymethylcyclopentane (21)



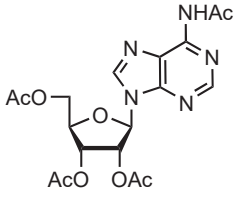
(3*S*,4*R*,5*R*)-1,2-epoxy-3,4-*O*-di-benzoyl-5-(*O*-tert-butyldiphenylsilyl)-methylcyclopentane (565 mg, 0.95 mmol) was dissolved in THF (9.5 mL, 0.1 M). TEA·3HF (1.56 mL, 9.53 mmol, 10 eq.) was carefully added and the reaction was stirred overnight. The reaction was cooled to 0 °C and carefully quenched with sat. aq. $NaHCO_3$. The reaction was transferred into a separatory funnel and the water layer was extracted thrice with EtOAc. The combined organic layers were dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (20 -> 40% EtOAc) yielded the title compound as a colorless oil in quantitative yield. **Rf:** 0.22 in 30% EtOAc in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.02 – 7.95 (m, 2H, Bz arom.), 7.89 – 7.82 (m, 2H, Bz arom.), 7.60 – 7.54 (m, 1H, Bz arom.), 7.49 (ddt, *J* = 8.8, 7.4, 1.3 Hz, 1H, Bz arom.), 7.44 – 7.37 (m, 2H, Bz arom.), 7.33 – 7.27 (m, 2H Bz arom.), 5.83 (d, *J* = 5.2 Hz, 1H, H-2), 5.13 (dd, *J* = 8.1, 5.4 Hz, 1H, H-3), 4.12 – 3.89 (m, 2H, H-5), 3.87 – 3.77 (m, 2H, H-1 + H-6), 2.79 – 2.68 (m, 1H, H-4), 2.13 (bs, 1H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.0, 165.5 (C=O Bz), 133.6, 133.4, 129.9, 129.8 (CH arom. Bz), 129.5, 129.3 (Cq Bz), 128.6, 128.4 (CH arom. Bz), 72.8 (C-3), 71.0 (C-2), 61.3 (C-5), 57.2, 54.4 (C1 + C-6), 46.2 (C-4). **HRMS:** $[C_{20}H_{18}O_6 + H]^+$ found: 355.1174, calculated: 355.1176.

(1*R*,2*R*,3*R*,4*R*,5*S*)-1,2-epoxy-3,4-*O*-di-benzoyl-5-(*O*-difluorenylmethyl phosphate)methylcyclopentane (9)



Compound **20** (431 mg, 1.22 mmol) was co-evaporated with toluene. A solution of bis-(9H-fluoren-9-ylmethyl)-*N,N*-diisopropyl amido phosphite (698 mg, 1.34 mmol, 1.1 eq.) in MeCN (12 mL, 0.1 M) was added. To this solution, a 0.25 M DCI solution in MeCN (9.76 mL, 2.44 mmol, 2.0 eq.) was added and the reaction was stirred for 15 minutes. A 5.5 M *t*BuOOH in nonane (2.2 mL, 12.2 mmol, 10 eq.) was added and the reaction was stirred for an additional 1.5 hours. The reaction was taken up in EtOAc and the organic layer was washed with sat. aq. $NaHCO_3$ and brine respectively. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (30 -> 50% EtOAc in pentane) yielded the title compound as a white foam (905 mg, 1.14 mmol, 94%). **Rf:** 0.4 in 40% EtOAc in pentane. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.96 – 7.88 (m, 2H, Bz arom.), 7.87 – 7.80 (m, 2H, Bz arom.), 7.72 – 7.63 (m, 4H, Fm arom.), 7.57 – 7.50 (m, 3H, Bz arom + Fm arom.), 7.50 – 7.43 (m, 4H, Bz arom + Fm arom.), 7.40 – 7.18 (m, 13H, Bz arom + Fm arom.), 5.79 (d, *J* = 5.4 Hz, 1H, H-2), 4.95 (dd, *J* = 8.1, 5.4 Hz, 1H, H-3), 4.41 – 4.18 (m, 4H, 2x CH_2 Fm), 4.18 – 3.97 (m, 4H, H-5 + 2x CH Fm), 3.77 (dd, *J* = 2.7, 0.9 Hz, 1H, H-1), 3.62 (dd, *J* = 2.8, 1.2 Hz, 1H, H-6), 2.87 – 2.71 (m, 1H, H-4). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.4, 165.4 (C=O Bz), 143.1, 143.0, 143.0, 143.0, 141.4, 141.4, 141.4 (Cq Fm), 133.5, 133.4, 129.8, 129.7 (CH arom. Bz), 129.3, 129.0 (Cq Bz), 128.5, 128.4 (CH arom. Bz), 128.0, 128.0, 128.0, 127.9, 127.3, 127.2, 127.2, 127.2, 125.2, 125.1, 125.1, 125.1, 120.2, 120.1, 120.1, 120.1, 120.0 (CH arom. Bz + CH arom. Fm), 72.2 (C-3), 70.5 (C-2), 69.4, 69.4, 69.3 (2x CH_2 Fm), 65.7, 65.6 (C-5), 56.1 (C-6), 54.6 (C-1), 48.0, 48.0, 47.9, 47.9 (2x CH Fm), 44.4, 44.3 (C-4). **^{31}P NMR** (162 MHz, $CDCl_3$) δ -1.29. **HRMS:** $[C_{48}H_{39}O_9P + H]^+$ found: 791.2397, calculated: 791.2405.

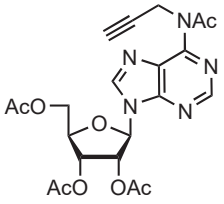
6-*N*-2',3',5'-tri-*O*-tetraacetyladenosine (22)



Adenosine (2.67 g, 10.0 mmol) was suspended in pyridine (50 mL, 0.2 M). Ac_2O (9.50 mL, 100 mmol. 10.0 eq.) was added and the reaction was stirred overnight. The now clear solution was quenched by the addition of EtOH and the reaction was concentrated *in vacuo* and co-evaporated with toluene. The resulting white solid was suspended in MeOH (25 mL, 0.4 M) and imidazole was added (408 mg, 6.0 mmol, 0.6 eq.) and the suspension was stirred overnight. The reaction was taken up in $CHCl_3$ and washed

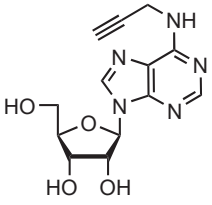
thrice with brine. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* yielding the title compound in quantitative yield as a white foam which was used without further purification. Spectral data was in accordance with literary precedence.[40]

6-*N*-acetyl-6-*N*-propargyl-2',3',5'-tri-*O*-acetyladenosine (23)



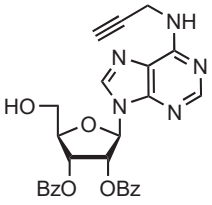
Compound **22** (1.31 g, 3.00 mmol) was dissolved in MeCN (10 mL, 0.3 M). DBU (0.95 mL, 6.34 mmol, 2.1 eq.) and propargyl bromide (0.49 mL, 4.52 mmol, 1.5 eq.) were added. The reaction was stirred for 1 hour after which the reaction was taken up in EtOAc. The organic layer was washed with 1 M HCl and brine respectively. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo* yielding the title compound as a white foam in quantitative yield. **Rf:** 0.64 in EtOAc. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.84 (s, 1H, H-2), 8.28 (s, 1H, H-8), 6.28 (d, *J* = 5.0 Hz, 1H, H-1'), 5.99 (t, *J* = 5.2 Hz, 1H, H-2'), 5.71 (t, *J* = 5.3 Hz, 1H, H-3'), 5.09 (d, *J* = 2.5 Hz, 2H, CH_2 propargyl), 4.56 – 4.36 (m, 3H, H-4'+ H-5'), 2.39 (s, 3H, NAc), 2.19 – 2.15 (m, 4H, Ac + CH propargyl), 2.14 (s, 3H, Ac), 2.12 (s, 3H, Ac). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.9, 170.3, 169.6, 169.4 (C=O Ac), 152.6, 152.5 (C-4 + C-6), 152.2 (C-2), 142.4 (C-8), 126.9 (C-5), 86.8 (C-1'), 80.3 (C-4'), 79.2 (Cq propargyl), 73.1 (C-2'), 71.7 (CH propargyl), 70.4 (C-3'), 63.0 (C-5'), 36.4 (CH_2 propargyl), 24.4, 20.8, 20.5, 20.4 (CH_3 Ac). **HRMS:** $[C_{21}H_{23}N_5O_8 + H]^+$ found: 474.1618, calculated: 474.1619.

6-*N*-propargyladenosine (8)



Compound **23** (2.57 g, 5.45 mmol) was dissolved in MeOH (55 mL, 0.1 M) and a catalytic amount of sodium was added. The reaction was stirred 4 hours and white precipitate was filtered yielding the title compound (908 mg, 2.97 mmol, 55%). Spectral data was in accordance with literary precedence.[41] **^{1}H NMR** (400 MHz, DMSO) δ 8.41 (s, 1H), 8.28 (s, 1H), 5.90 (d, J = 6.1 Hz, 1H), 5.36 (s, 2H), 4.60 (dd, J = 6.1, 4.9 Hz, 1H), 4.25 (s, 2H), 4.15 (dd, J = 5.0, 3.1 Hz, 1H), 3.96 (q, J = 3.5 Hz, 1H), 3.68 (dd, J = 12.1, 3.7 Hz, 1H), 3.55 (dd, J = 12.1, 3.7 Hz, 1H), 3.11 – 2.99 (m, 1H). **HRMS:** $[C_{13}H_{15}N_5O_4 + H]^+$ found: 306.1195, calculated: 306.1197.

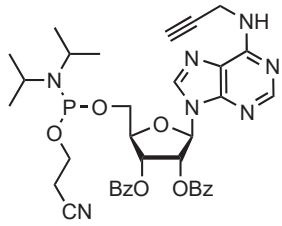
2',3'-*O*-dibenzoyl-6-*N*-propargyladenosine (24)



Compound **8** (445 mg, 1.46 mmol) was co-evaporated in pyridine and dissolved in pyridine (6 mL, 0.25 M). TBDMS-Cl (50 wt% in toluene, 0.76 mL, 2.19 mmol, 1.5 eq.) was added and the reaction was stirred for 2 hours after which TLC analysis showed full conversion of starting material into a higher running product (Rf = 0.35 in 5% MeOH in DCM). Bz-Cl (0.60 mL, 5.11 mmol, 3.5 eq.) was added to the reaction and after 1 hour of stirring TLC indicated full conversion of the intermediate. The reaction was taken up in EtOAc and the solution was washed with 1 M citric acid. The organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude residue was dissolved in a 1:4 H_2O:MeCN mixture (6 mL total volume, 0.25 M) and PTSA monohydrate (417 mg, 2.19 mmol, 1.5 eq.) was added. The reaction was stirred for 3 hours. The reaction was taken up in EtOAc and the organic layer was washed with sat. aq. $NaHCO_3$ and brine respectively. The water layers were extracted with EtOAc and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (40% EtOAc in pentane) yielded the title compound as a white foam (581 mg, 1.09 mmol, 75%). **Rf:** 0.64 (40% EtOAc in pentane) **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.46 (s, 1H, H-2), 8.11 – 8.01 (m, 2H, Bz arom.), 7.97 (s, 1H, H-8), 7.90 – 7.80 (m, 2H, Bz arom.), 7.59 (ddt, *J* = 8.7, 7.1, 1.3 Hz, 1H, Bz arom.), 7.54 – 7.39 (m, 3H, Bz arom.), 7.36 – 7.25 (m, 3H,

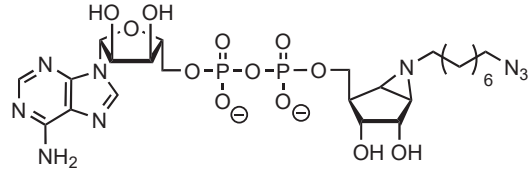
Bz arom.), 6.42 (dd, *J* = 7.5, 5.3 Hz, 1H, H-2'), 6.33 (d, *J* = 7.5 Hz, 1H, H-1'), 6.09 (dd, *J* = 5.3, 1.3 Hz, 1H, H-3'), 4.63 (q, *J* = 1.5 Hz, 1H, H-4'), 4.44 (s, 2H, CH_2 propargyl), 4.18 – 3.94 (m, 2H, H-5'), 2.25 (t, *J* = 2.5 Hz, 1H, CH propargyl). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.5, 164.8 (C=O Bz), 154.5 (C-4 + C-6), 152.9 (C-2), 139.9 (C-8), 133.7, 129.7 (CH arom. Bz), 129.2 (Cq Bz), 128.6, 128.5 (CH arom. Bz), 121.3 (C-5), 88.8 (C-1'), 86.5 (C-4'), 80.0 (Cq propargyl), 73.6 (C-3'), 73.5 (C-2'), 71.6 (CH propargyl), 62.7 (C-5'), 30.0 (CH_2 propargyl (signal taken from HSQC). **HRMS:** $[C_{27}H_{23}N_5O_6 + H]^+$ found: 514.1721, calculated: 514.1721.

5'-*O*-(*N*^{6}-propargyl-2',3'-di-*O*-benzoyladenosine)-2-cyanoethyl-*N*-*N*-diisopropylphosphoramidite (7)



Compound **24** (700 mg, 1.36 mmol) was co-evaporated with 1,4-dioxane and dissolved in DCM (4.5 mL, 0.3 M). DIPEA (0.59 mL, 3.40 mmol, 2.5 eq.) and 2-cyanoethyl-*N*,*N*-diisopropyl-chlorophosphoramidite (0.33 mL, 1.50 mmol, 1.1 eq.) were added and the reaction was stirred for 1 hour. The reaction was diluted with DCM and the solution was washed with sat. aq. $NaHCO_3$ and brine respectively. The water layers were extracted with DCM and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (40% EtOAc in pentane, neutralized with 1% TEA) yielded the title compound as a white foam and a mixture of diastereomers (S_P/R_P) (800 mg, 1.12 mmol, 82%). **Rf:** 0.42 in 40% EtOAc in pentane. **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.50 – 8.44 (m, 1H, CH adenine), 8.42 – 8.38 (m, 1H, CH adenine), 8.05 – 7.98 (m, 2H, Bz arom.), 7.95 – 7.86 (m, 2H, Bz arom.), 7.54 (dtd, J = 37.7, 7.4, 1.6 Hz, 2H, Bz arom.), 7.42 (td, J = 7.8, 2.4 Hz, 2H, Bz arom.), 7.32 (td, J = 7.9, 2.1 Hz, 2H, Bz arom.), 6.74 (bs, 1H, NH), 6.71 – 6.58 (m, 1H, H-1'), 6.25 – 6.14 (m, 1H, H-2'), 6.03 (ddd, J = 11.1, 5.5, 2.6 Hz, 1H, H-3'), 4.66 (q, J = 2.8 Hz, 1H, H-4'), 4.56 – 4.34 (m, 2H, CH_2 propargyl), 4.13 (ddq, J = 11.8, 6.9, 3.0 Hz, 1H, H-5'$_a$), 4.09 – 3.81 (m, 3H, H-5'$_b$ + O*C*H_2CH_2CN), 3.69 (dqq, J = 13.5, 6.8, 3.9, 2.7 Hz, 2H, CH N-*i*Pr), 2.73 (td, J = 6.4, 4.1 Hz, 2H, OCH_2CH_2CN), 2.28 (q, J = 2.6 Hz, 1H, CH propargyl), 1.29 – 1.17 (m, 12H, CH_3 N-*i*Pr). **^{13}C NMR** (126 MHz, $CDCl_3$): δ 165.5, 165.4, 165.0, 165.0 (C=O Bz), 154.1 (Cq adenine), 153.3, 153.3, (CH adenine), 138.5 (CH adenine, signal taken from HSQC) 133.7, 133.6, 129.8, 129.8, 129.8 (CH arom. Bz), 129.0, 128.9 (Cq Bz/adenine), 128.5 (CH arom. Bz), 128.5, 128.5 (Cq Bz/adenine), 128.4 (CH arom. Bz), 120.1, 120.0, 117.7, 117.6 (Cq Bz/adenine), 85.4, 85.0 (C-1'), 83.6, 83.5, 83.3, 83.2 (C-4'), 80.1 (Cq propargyl), 74.6, 74.5 (C-2'), 72.7, 72.6 (C-3'), 71.6 (CH propargyl), 63.3, 63.2, 63.0 (C-5'), 58.9, 58.7, 58.6 (OCH_2CH_2CN), 43.3, 43.3, 43.2, 43.2 (CH N-*i*Pr), 24.8, 24.7, 24.7, 24.7, 24.7 (CH_3 N-*i*Pr), 20.4, 20.4, 20.3 (OCH_2CH_2CN). **HRMS:** *mass was detected as its corresponding H-phosphonate* $[C_{30}H_{27}N_6O_8P + H]^+$ found: 631.1698, calculated: 631.1701.

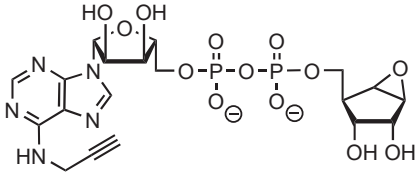
(1*R*,2*S*,3*R*,4*R*,5*R*)-(6-*N*-oct-(1-azide)-ane)azirdine-3,4-diol-5-*O*-(adenosine diphosphate) methylcyclopentane (1)



Aziridine **4** (61 mg, 65 µmol) was dissolved in MeCN (0.65 mL, 0.1 M) and TEA (127 µL, 0.91 mmol, 14 eq.) was added. The reaction was stirred overnight until LC-MS analysis showed full deprotection (gradient 10 -> 90% B in A, Rt = 6.09) from the Fm protecting groups. The reaction was co-evaporated with toluene, followed by co-evaporation with pyridine to convert the phosphate into the pyridinium salt followed by additional co-evaporation with toluene. Adenosine amidite **3** (56 mg, 78 µmol, 1.2 eq.) was co-evaporated separately with toluene and dissolved in 0.2 mL MeCN. This solution was added to the crude aziridine and the flask was rinsed with another 0.2 mL MeCN and added to the reaction. ETT (0.25 M in MeCN, 0.62 mL, 0.16 mmol, 2.4 eq.) was added

to the reaction (end volume 1.0 mL, 0.06 M). The reaction was stirred 15 minutes after which ^{31}P-NMR indicated full conversion of starting material (δ = 129.59, 127.36 P^{III} and -9.97 P^{V}). *t*BuOOH (5.5 M in nonane, 77 µL, 0.42 mmol, 6.5 eq.) was added and the reaction was stirred for 30 minutes after which ^{31}P-NMR (δ peaks shifted around -14) indicated full conversion of the phosphite species into the corresponding phosphate. DBU (0.5 M in DMF, 0.78 mL, 0.42 mmol, 6.0 eq.) was added and the reaction was stirred for 1 hour. Sat. aq. NH_4OH (1.6 mL) was added and the reaction was stirred overnight. The reaction was concentrated *in vacuo*. The residue was purified by gel filtration (HW40) followed by HPLC purification (buffered with NH_4OAc) where fractions containing product first were concentrated *in vacuo* and extensively co-evaporated with 1:1 MeCN:H_2O followed by lyophilization, yielding the title compound as a white solid (11.6 mg, 16.3 µmol, 25%). **^{1}H NMR** (850 MHz, D_2O) δ 8.50 (s, 1H, H-2), 8.20 (s, 1H, H-8), 6.08 (d, J = 5.6 Hz, 1H, H-1'), 4.68 (t, J = 5.3 Hz, 1H, H-2'), 4.47 (dd, J = 5.0, 3.8 Hz, 1H, H-3'), 4.35 (dt, J = 4.0, 2.4 Hz, 1H, H-4'), 4.29 – 4.23 (m, 2H, H-2 + H-5_a), 4.22 – 4.17 (m, 2H, H-5'), 4.15 (dt, J = 13.4, 5.0 Hz, 1H, H-5_b), 3.89 (s, 1H, H-3), 3.17 (t, J = 6.9 Hz, 2H, $RNCH_2CH_2(CH_2)_4CH_2CH_2N_3$), 2.95 (s, 1H, H-6), 2.59 (s, 1H, H-1), 2.46 (s, 1H, H-4), 1.54 – 1.44 (m, 2H, $RNCH_2CH_2(CH_2)_4CH_2CH_2N_3$), 1.39 (p, J = 7.0 Hz, 2H, $RNCH_2CH_2(CH_2)_4CH_2CH_2N_3$), 1.08 (p, J = 7.4 Hz, 2H, $RNCH_2CH_2(CH_2)_4CH_2CH_2N_3$), 1.03 – 0.88 (m, 8H, $RNCH_2CH_2(CH_2)_4CH_2CH_2N_3$). **^{13}C NMR** (214 MHz, D_2O) δ 155.5 (Cq adenine), 152.7 C-8 adenine), 148.9 (Cq adenine), 139.7 (C-2 adenine), 118.5 (Cq adenine), 87.0 (C-1'), 83.7, 83.7 (C-4'), 74.4 (C-2'), 70.2 (C-3'), 69.0 (C-2), 65.0, 65.0 (C-5'), 63.5 (C-5), 51.1 ($RNCH_2CH_2(CH_2)_4CH_2CH_2N_3$), 45.8, 44.4, 44.4 (C-1, C-4, C-6), 28.1, 28.0, 27.8, 26.7, 25.7, 25.7 ($RNCH_2CH_2(CH_2)_4CH_2CH_2N_3$). **^{31}P NMR** (202 MHz, D_2O) δ -10.2, -10.3, -10.6, -10.7. **LC-MS:** (0 -> 20% B in A) Rt = 8.49. **HRMS:** $[C_{24}H_{39}N_9O_{12}P_2 + H]^+$ found: 708.2263, calculated: 708.2266.

(1*R*,2*S*,3*R*,4*R*,5*R*)-1,2-epoxy-3,4-diol-5-*O*-((*N*^{6}-progargyladenosine)diphosphate)methylcyclopentane (2)



Epoxide **9** (158 mg, 0.2 mmol) was dissolved in MeCN (2.0 mL, 0.1 M) and TEA (0.4 mL, 2.8 mmol, 14 eq.) was added. The reaction was stirred overnight until LC-MS analysis showed full deprotection (gradient 10 -> 90% B in A, Rt = 5.13) from the Fm protecting groups. The reaction was co-evaporated with toluene, followed by co-evaporation with pyridine to convert the phosphate into the pyridinium salt followed by additional co-evaporation with toluene. Adenosine amidite **7** (171 mg, 0.24 mmol, 1.2 eq.) was co-evaporated separately with toluene and dissolved in 0.6 mL MeCN. This solution was added to the crude epoxide and the flask was rinsed with another 0.6 mL MeCN and added to the reaction. ETT (0.25 M in MeCN, 1.9 mL, 0.48 mmol, 2.4 eq.) was added to the reaction (end volume 3.1 mL, 0.06 M). The reaction was stirred 15 minutes after which ^{31}P-NMR indicated full conversion of starting material (δ = 127.49, 127.01 P^{III} and -11.30 P^{V}). *t*BuOOH (5.5 M in nonane, 0.24 mL, 1.3 mmol, 6.5 eq.) was added and the reaction was stirred for 30 minutes after which ^{31}P-NMR (δ = -12.95, -13.72, -13.86) and LC-MS analysis (gradient 10 -> 90% B in A, Rt = 6.50) indicated full conversion of the phosphite species into the corresponding phosphate. DBU (0.5 M in DMF, 2.4 mL, 1.2 mmol, 6.0 eq.) was added and the reaction was stirred 1 hour. Sat. aq. NH_4OH (5.0 mL) was added and the reaction was stirred overnight. The reaction was concentrated *in vacuo* and the residue was taken up in H_2O. The water layer was washed with EtOAc and the water layer was concentrated *in vacuo*. The residue was purified by gel filtration (HW40) followed by HPLC purification (buffered with NH_4OAc) where fractions containing product first were concentrated *in vacuo* and extensively co-evaporated with 1:1 MeCN:H_2O followed by lyophilization, yielding the title compound as a white solid (14.9 mg, 25.0 µmol, 13%). **^{1}H NMR** (850 MHz, D_2O) δ 8.37 (s, 1H, H-2 adenine), 8.18 (s, 1H, H-8 adenine), 6.00 (d, J = 5.7 Hz, 1H, H-1'), 4.64 (t, J = 5.4 Hz, 1H, H- 2'), 4.40 (dd, J = 5.1, 3.8 Hz, 1H, H-3'), 4.26 (dt, J = 3.8, 2.3 Hz, 1H, H-4'), 4.20 (d, J = 24.7 Hz, 2H, CH_2 propargyl),

4.13 – 4.06 (m, 2H, H-5'), 4.03 (dt, J = 10.7, 5.5 Hz, 1H, H-5_a), 3.96 (d, J = 5.2 Hz, 1H, H-2), 3.85 (ddd, J = 10.2, 8.4, 6.9 Hz, 1H, H-5_b), 3.53 (dd, J = 3.0, 1.2 Hz, 1H, H-6), 3.40 (dd, J = 2.9, 0.9 Hz, 1H, H-1), 3.30 (dd, J = 8.2, 5.2 Hz, 1H, H-3), 2.48 (t, J = 2.5 Hz, 1H, CH propargyl), 2.13 (tdd, J = 8.4, 5.3, 1.3 Hz, 1H, H-4). **^{13}C NMR** (214 MHz, D_2O) δ 153.7 (Cq adenine), 152.4 (C-8 adenine), 139.5 (C-2 adenine, signal taken from HSQC) 119.0 (Cq adenine), 86.8 (C-1'), 83.7, 83.6 (C-4'), 80.1 (Cq propargyl), 74.0 (C-2'), 71.6 (CH propargyl), 70.8 (C-3), 70.1, 70.1 (C-3'), 69.1 (C-2), 64.9, 64.9 (C-5'), 64.6, 64.6 (C-5), 57.4 (C-6), 55.9 (C-1), 44.8, 44.7 (C-4), 29.9 (CH_2 propargyl, signal taken from HSQC). **^{31}P NMR** (202 MHz, D_2O) δ -10.3, -10.4, -10.6, -10.7. **LC-MS:** (0 -> 20% B in A) Rt = 3.79. **HRMS:** $[C_{19}H_{25}N_5O_{13}P_2 + H]^+$ found: 594.0997, calculated: 594.0997.

References

[1] A. K. Bhan, E. L. Reinherz, S. Poppema, R. T. McCluskey, S. F. Schlossman, *J. Exp. Med.* **1980**, *152*, 771–782.

[2] F. Malavasi, S. Deaglio, A. Funaro, E. Ferrero, A. L. Horenstein, E. Ortolan, T. Vaisitti, S. Aydin, *Physiol. Rev.* **2008**, *88*, 841–886.

[3] L. Santos-Argumedo, C. Teixeira, G. Preece, P. A. Kirkham, R. M. E. Parkhouse, *J. Immunol.* **1993**, *151*, 3119–3130.

[4] A. D. A. Funaro, D. Delia, M. Zaccolo, F. Malavasi, C. Giulio, S. Roggero, F. Malavas, *J. Immunol.* **1990**, *145*, 2390–2396.

[5] U. Dianzani, A. Funaro, D. DiFranco, G. Garbarino, M. Bragardo, V. Redoglia, D. Buonfiglio, L. B. De Monte, A. Pileri, F. Malavasi, *J. Immunol.* **1994**, *153*, 952–959.

[6] S. Deaglio, M. Morra, R. Mallone, C. M. Ausiello, E. Prager, G. Garbarino, U. Dianzani, H. Stockinger, F. Malavasi, *J. Immunol.* **1998**, *160*, 395–402.

[7] M. Howard, J. C. Grimaldi, J. F. Bazan, F. E. Lund, L. Santos-Argumedo, R. M. E. Parkhouse, T. F. Walseth, H. C. Lee, *Science* **1993**, *262*, 1056–1059.

[8] R. Aarhus, R. M. Graeff, D. M. Dickey, T. F. Walseth, H. C. Lee, *J. Biol. Chem.* **1995**, *270*, 30327–30333.

[9] H. C. Lee, Y. J. Zhao, *J. Biol. Chem.* **2019**, *294*, 19831–19843.

[10] D. L. Clapper, T. F. Walseth, P. J. Dargie, Hon Cheung Lee, *J. Biol. Chem.* **1987**, *262*, 9561–9568.

[11] H. C. Lee, T. F. Walseth, G. T. Bratt, R. N. Hayes, D. L. Clapper, *J. Biol. Chem.* **1989**, *264*, 1608–1615.

[12] S. Takasawa, K. Nata, H. Yonekura, H. Okamoto, *Science* **1993**, *259*, 370–373.

[13] T. Matsuoka, Y. Kajimoto, H. Watada, Y. Umayahara, M. Kubota, R. Kawamori, Y. Yamasaki, T. Kamada, *Biochem Biophys Res Commun* **1995**, *214*, 239–246.

[14] H. N. Ho, L. E. Hultin, R. T. Mitsuyasu, J. L. Matud, M. A. Hausner, D. Bockstoce, C. C. Chou, S. O'Rourke, J. M. Taylor, J. V Giorgi, *J. Immunol.* **1993**, *150*, 3070–3079.

[15] L. Kestens, G. Vanham, P. Gigase, G. Young, I. Hannet, F. Vanlangendonck, F. Hulstaert, B. A. Bach, *Aids* **1992**, *6*, 793–797.

[16] R. N. Damle, T. Wasil, F. Fais, F. Ghiotto, A. Valetto, S. L. Allen, A. Buchbinder, D. Budman, K. Dittmar, J. Kolitz, S. M. Lichtman, P. Schulman, V. P. Vinciguerra, K. R. Rai, M. Ferrarini, N. Chiorazzi, *Blood* **1999**, *94*, 1840–1847.

[17] D. Jin, H. X. Liu, H. Hirai, T. Torashima, T. Nagai, O. Lopatina, N. A. Shnayder, K. Yamada, M. Noda, T. Seike, K. Fujita, S. Takasawa, S. Yokoyama, K. Koizumi, Y. Shiraishi, S. Tanaka, M. Hashii, T. Yoshihara, K. Higashida, M. S. Islam, N. Yamada, K. Hayashi, N. Noguchi, I. Kato, H. Okamoto, A. Matsushima, A. Salmina, T. Munesue, N. Shimizu, S. Mochida, M. Asano, H. Higashida, *Nature* **2007**, *446*, 41–45.

[18] A. B. Salmina, O. Lopatina, M. V. Ekimova, S. V. Mikhutkina, H. Higashida, *J. Neuroendocrinol.* **2010**, *22*, 380–392.

[19] A. A. Sauve, H. Deng, R. H. Angeletti, V. L. Schramm, *J. Am. Chem. Soc.* **2000**, *122*, 7855–7859.

[20] M. L. Love, D. M. E. Szebenyi, I. A. Kriksunov, D. J. Thiel, C. Munshi, R. Graeff, H. C. Lee, Q. Hao, *Structure* **2004**, *12*, 477–486.

[21] Q. Liu, I. A. Kriksunov, H. Jiang, R. Graeff, H. Lin, H. C. Lee, Q. Hao, *Chem. Biol.* **2008**, *15*, 1068–1078.

[22] Y. Zhou, K. Y. Ting, C. Mo, C. Lam, A. Ka, Y. Kwong, J. Xia, H. Jin, Z. Liu, L. Zhang, H. C. Lee, L. Zhang, *ChemMedChem* **2012**, *7*, 223–228.

[23] M. Dong, Y. Si, S. Sun, X. Pu, Z. Yang, L. Zhang, L. Zhang, P. Leung, M. Ching, K. Yee, J. Yue, Y. Zhou, I. A. Kriksunov, H. Cheung, *Org. Biomol. Chem.* **2011**, *9*, 3246–3257.

[24] A. a Sauve, V. L. Schramm, *Biochemistry* **2002**, *41*, 8455–63.

[25] A. Depaix, J. Kowalska, *Molecules* **2019**, *24*, 4187.
[26] A. C. M. van Esbroeck, A. P. A. Janssen, A. B. Cognetta, D. Ogasawara, G. Shpak, M. van der Kroeg, V. Kantae, M. P. Baggelaar, F. M. S. de Vrij, H. Deng, M. Allarà, F. Fezza, Z. Lin, T. van der Wel, M. Soethoudt, E. D. Mock, H. den Dulk, I. L. Baak, B. I. Florea, G. Hendriks, L. De Petrocellis, H. S. Overkleeft, T. Hankemeier, C. I. de Zeeuw, V. Di Marzo, M. Maccarrone, B. F. Cravatt, S. A. Kushner, M. van der Stelt, *Science* **2017**, *356*, 1084–1087.
[27] B. F. Cravatt, A. T. Wright, J. W. Kozarich, *Annu. Rev. Biochem.* **2008**, *77*, 383–414.
[28] N. Li, H. S. Overkleeft, B. I. Florea, *Curr. Opin. Chem. Biol.* **2012**, *16*, 227–233.
[29] J. Voorneveld, B. I. Florea, T. Bakkum, R. J. Mendowicz, M. S. van der Veer, B. Gagestein, S. I. van Kasteren, M. van der Stelt, H. S. Overkleeft, D. V. Filippov, *ChemBioChem* **2020**, DOI 10.1002/cbic.202000042.
[30] H. Jiang, J. Congleton, Q. Liu, P. Merchant, F. Malavasi, H. C. Lee, Q. Hao, A. Yen, H. Lin, *J. Am. Chem. Soc.* **2009**, *131*, 1658–1659.
[31] J. H. Shrimp, J. Hu, M. Dong, B. S. Wang, R. Macdonald, H. Jiang, Q. Hao, A. Yen, H. Lin, *J. Am. Chem. Soc.* **2014**, *136*, 5656–5663.
[32] S. J. Williams, S. G. Withers, *Carbohydr. Res.* **2000**, *327*, 27–46.
[33] J. D. Mccarter, M. J. Adamt, S. G. Withers, *Biochem. J.* **1992**, *286*, 721–727.
[34] I. P. Street, J. B. Kempton, S. G. Withers, *Biochemistry* **1992**, *31*, 9970–9978.
[35] M. D. Witte, M. T. C. Walvoort, K. Li, W. W. Kallemeijn, W. E. Donker-Koopman, R. G. Boot, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, *ChemBioChem* **2011**, *12*, 1263–1269.
[36] K. Y. Li, J. Jiang, M. D. Witte, W. W. Kallemeijn, H. van den Elst, C. S. Wong, S. D. Chander, S. Hoogendoorn, T. J. M. Beenakker, J. D. C. Codée, J. M. F. G. Aerts, G. A. van der Marel, H. S. Overkleeft, *European J. Org. Chem.* **2014**, *2014*, 6030–6043.
[37] W. W. Kallemeijn, K.-Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts, H. S. Overkleeft, *Angew. Chemie Int. Ed.* **2012**, *51*, 12529–12533.
[38] H. Gold, P. van Delft, N. Meeuwenoord, J. D. C. Codée, D. V. Filippov, G. Eggink, H. S. Overkleeft, G. A. van der Marel, *J. Org. Chem.* **2008**, *73*, 9458–9460.
[39] G. J. van der Heden van Noort, M. G. van der Horst, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *J. Am. Chem. Soc.* **2010**, *132*, 5236–5240.
[40] V. Tararov, S. Kolyachkina, C. Alexeev, S. Mikhailov, *Synthesis* **2011**, *15*, 2483–2489.
[41] I. Carter-O'Connell, H. Jin, R. K. Morgan, L. L. David, M. S. Cohen, *J. Am. Chem. Soc.* **2014**, *136*, 5201–5204.
[42] H. Ovaa, B. Lastdrager, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, J. H. van Boom, *J. Chem. Soc. Perkin Trans. 1* **2002**, 2370–2377.
[43] K. B. Sharpless, R. C. Michaelson, *J. Am. Chem. Soc.* **1973**, *95*, 6136–6137.
[44] T. Itoh, K. Jitsukawa, K. Kaneda, S. Teranishi, *J. Am. Chem. Soc.* **1979**, *101*, 159–169.
[45] J. Jiang, T. J. M. Beenakker, W. W. Kallemeijn, G. A. van der marel, H. Van Den Elst, J. D. C. Codée, J. M. F. G. Aerts, H. S. Overkleeft, *Chem. - A Eur. J.* **2015**, *21*, 10861–10869.

Chapter 8

Olaparib-Based Photoaffinity Probes for PARP-1 Detection in Living Cells

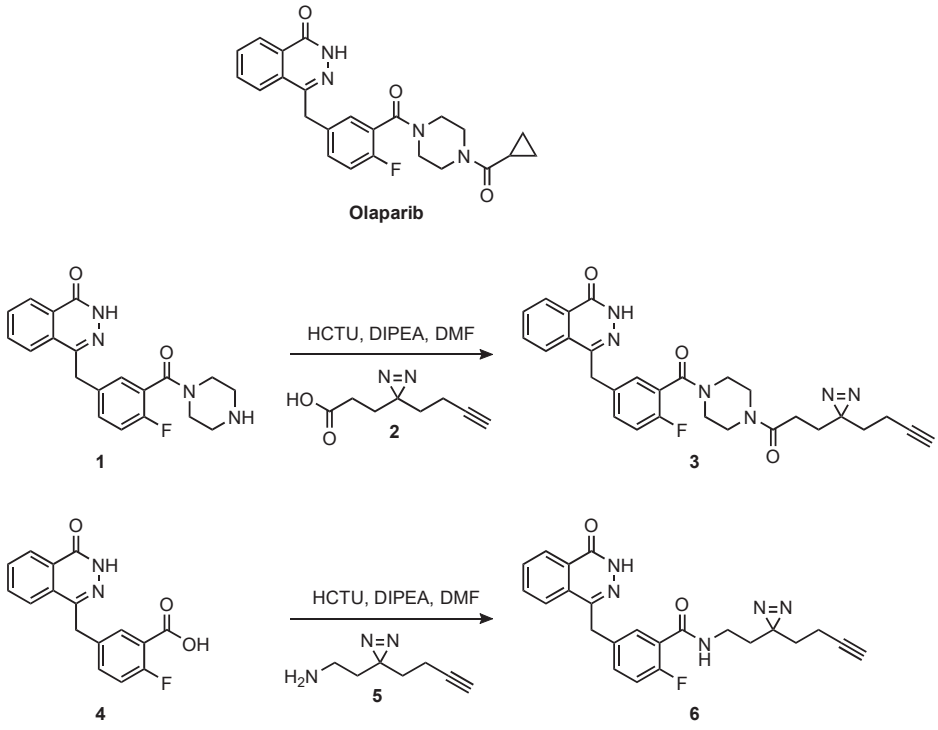
Introduction

One of the crucial steps in the repair of DNA double-strand breaks is ADP-ribosylation of chromatin,[1–6] a process driven by several members of the PARP family.[7–9] This family of transferases represents clinically relevant target proteins for drugs directed at tumors that are deficient in their DNA repair.[10,11] The ability to monitor PARPs *in vivo* is necessary in order to study the processes of the drug-target engagement and pharmacokinetics. One method of accomplishing this involves small molecule probes, which are capable to specifically and covalently bind to their target enzymes in living cells and in tissues. These probes have been developed for various classes of enzymes[12–16] and used in successful studies involving *in vivo* protein detection and quantification. One way to attain such a probe is to convert a known reversible enzyme inhibitor into a covalent probe suitable for affinity-based protein profiling (AfBPP).[16–19] AfBPP probes consist of a recognition element that reversibly interacts with the target of interest, to which a photoreactive group and bioorthogonal ligation handle are installed, without affecting the affinity of the parent structure. Irradiation by UV light activates the photoreactive group, creating a reactive intermediate which can form a covalent bond between the probe and the target protein in a process called photo-affinity labeling (PAL).[19–21] The bioorthogonal ligation handle is used to couple a reporter group, e.g. a fluorophore for in-gel analysis or a biotin for affinity enrichment with subsequent tandem mass spectrometry in chemical proteomics. This chapter presents such a covalent probe for the detection of PARP-1 in live cells using AfBPP (Figure 1). Prior work on the detection of PARP was focused solely on non-covalently binding molecules with the aim of either off-target finding[22,23] or *in vivo* imaging.[24,25] Therefore attention was directed to design, synthesize and biologically validate a PARP-1-selective photoaffinity probe for use in the standard proteomics protocols. Such well-established workflows make use of UV-induced PAL through a diazirine tag and copper-catalyzed alkyne-azide cycloaddition (CuAAC) that enable the pull-down of the protein for subsequent analysis by standard mass-spectrometry methods.

In this chapter, the design, synthesis and biological evaluation and validation of a PARP-1-selective photoaffinity probe for the use of chemical proteomics is presented. Such well-established workflows make use of UV-induced PAL through a diazirine tag and CuAAC that enable the pull-down of the protein for subsequent analysis by mass-spectrometry methods.

The clinical PARP inhibitor olaparib was chosen as the starting point for the design of the recognition element of photoaffinity probes **3** and **6** (Scheme 1), because it inhibits PARP-1 and PARP-2 with a nanomolar potency and shows exceptional selectivity over other cellular targets.[23,26,27] For PAL, a plethora of photoreactive groups have been

reported.[28] In order to limit the interference of the tag with the activity and selectivity of the probes, the compact diazirine-alkyne linkers **2** and **5** developed by Li *et al.*[29] were opted. Placement of the photoreactive group is of importance since it must be close to the peptide backbone of the protein of interest, yet not interfere with the affinity of the probe by steric hindrance. The nitrogen atom of the piperazine moiety distal from the phtalazinone substructure of olaparib is known to tolerate a wide range of modifications.[27,30] With this in mind, the photoaffinity tag can be attached to this nitrogen to minimize its influence on the binding efficiency to PARP-1, giving the design of probe **3**. A second, simplified probe **6**, in which the piperazine bridge was removed and the diazirine-alkyne tag was attached directly to benzylphtalazinone core of the olaparib structure was also synthesized.



Scheme 1. Structure of olaparib and the synthesis of the probes **3** and **6** derived from olaparib. Yields obtained are 33% for **3** and 75% for **6**.

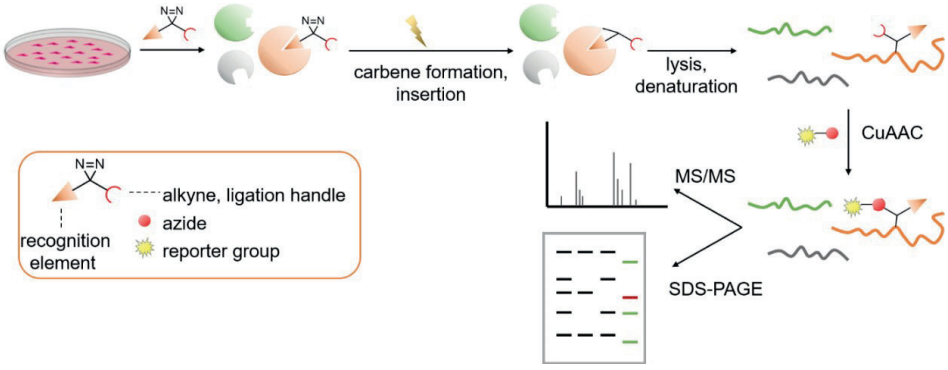


Figure 1. Schematic overview of AfBPP. After cells are treated with probe, irradiation at 350 nm leads to activation of the diazirine group, forming a highly reactive carbene that can rapidly insert into nearby N-H, O-H or C-H bonds. Protein extraction is done by cell lysis under denaturating conditions. Reporter or affinity handles are introduced by CuAAC click chemistry. Proteins are analyzed by SDS-PAGE and/or western blot analysis or by affinity purification, digested to peptides and LFQ proteomics analysis.

Results and discussion

Preparation of the probes was done via the HCTU-mediated condensation of either commercially available **1** or known compound **4**[24,27] with the linkers **2** and **5** respectively (Scheme 1). With the AfBPs **3** and **6** in hand, their ability to bind PARP covalently via PAL in live cells was tested on human B-cell-derived Raji cell line. Raji cells were chosen as the model because they have been previously used to study PARP[31] and have an advantageous nucleus to cytoplasm ratio that maximizes the PARP abundance. The cells were incubated with probes **3** or **6** (10 µM) and irradiated with UV light for 30 minutes to ensure complete activation of the diazirine and covalent labeling of PARP. During irradiation, the cells were cooled to 4 °C to counteract the heat generated from the irradiation. The cells were then harvested and lysed. Probe-bound proteins were ligated to azide-containing fluorophore AlexaFluor 647-N_3 under CuAAC conditions and further processed for in-gel fluorescence analysis. The gel revealed a fluorescent band at the expected height of PARP-1 around 116 kDa (Figure 2, lane 3), which was probe-dependent as cells treated with vehicle (lane 2) revealed no labeling. To further show probe-specific labeling, the UV irradiation step was omitted (lane 1) yielding no signal, showing that the observed labeling is UV-dependent rather than circumstantial activation.

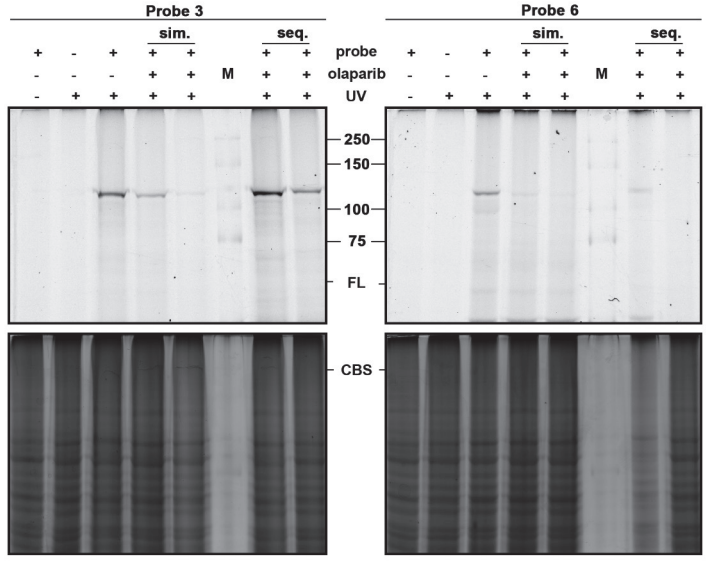


Figure 2. SDS-PAGE analysis of Raji cells exposed to probes **3** (left) and **6** (right) (10 μM) and no UV control. sim: simultaneous incubation with 1:1 (10 μM olaparib, 10 μM probe) or 5:1 (50 μM olaparib, 10 μM probe) ratio. seq: sequential incubation with olaparib, wash and probe at 1:1 ratio or 2:1 olaparib:probe ratio. FL: fluorescence channel. CBS: Coomassie blue staining for loading control of the gel in the upper panel. M: protein ladder

As an additional control, the gel used for in-gel fluorescence readout was analyzed by western blot with an anti-PARP-1 antibody confirming the fluorescent bands to be PARP-1 (Figure 3, bottom panel). To further verify the target, the experiments were repeated but the click reaction was performed with *N*-(3-azidopropyl)biotinamide, providing a handle to perform a pull-down with streptavidin. Probe-bound proteins from the lysates were enriched on streptavidin-coated beads and analyzed with SDS-PAGE followed by WB analysis with the anti-PARP-1 antibody to ascertain that the enriched protein is PARP-1 (data not shown). Having verified that probes **3** and **6** were able to label PARP in a UV-dependent, covalent manner, the binding site of the probe was investigated. If the probe retains the binding mode of the parent structure, labeling should be outcompeted by addition of olaparib. The probes and olaparib were incubated simultaneously in a 1:1 ratio after which a decrease of band intensity was observed (Figure 2, lane 4). After increasing the amount of olaparib to a 5:1 drug/probe ratio, labeling by the probe was abolished (lane 5). Interestingly, in a sequential competition experiment where the cells were first incubated with olaparib, washed, and then incubated with the probes in a 1:1 ratio, no apparent competition was observed (lane 7). Repeating the experiment with a 2:1 drug/probe ratio appeared to slightly increase competition of the probes (lane 8) but proved not as effective as simultaneous competition. This effect might be due to rapid dissociation of olaparib from the PARP-binding pocket during the washing step.

Next, attention was turned to demonstrate the ability of the probe to bind PARP in conditions frequently encountered in the field of ADP-ribosylation research. Treatment of cells with H_2O_2 generates DNA damage, which results in double strand breaks and subsequent activation of PARP. Therefore, Raji cells were treated with the probes as before but with addition of H_2O_2 during incubation. The addition of this oxidizing agent neither changed the ability of the probes to label PARP nor the efficiency of the labeling (Figure 3), suggesting that olaparib is able to bind PARP in the presence of extensive DNA-damage and appears consistent with previously reported findings.[22]

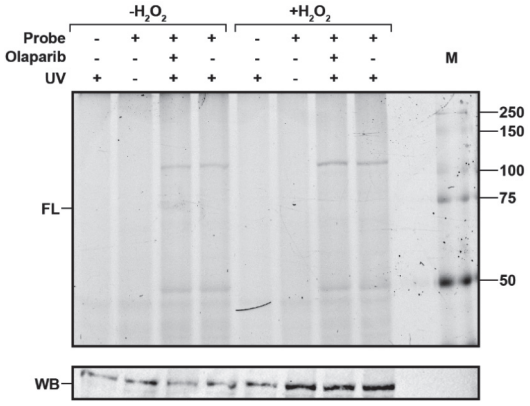


Figure 3. SDS-PAGE analysis with fluorescent imaging (FL) or western blot (WB) detection of Raji cell lysate either exposed or not exposed to H_2O_2 and probe 3, no-UV control, without or in sequential competition with olaparib (1:1 olaparib:probe ratio). WB: loading control and PARP1 detection of the gel in the upper panel by western blot analysis with a rabbit anti-PARP1 antibody. M: protein ladder

To identify and quantify the bands seen by SDS-PAGE and western blot analysis, an LC-MS/MS based label-free quantification (LFQ) method was used.[15] The results, depicted in Figure 4, show that PARP-1 is significantly UV-enriched by probe **3** indicating that the observed band at 116 kDa can be attributed to PARP-1 (Figures 4A and B). Competition analysis of significantly UV-enriched proteins shows PARP-1 is outcompeted by olaparib (Figure 4D) strongly indicating that the probe binds in the same protein pocket as olaparib. Several other proteins were selectively enriched by the probe upon UV irradiation having a nucleic acid binding motif as common feature, indicating that the probe targets these proteins possibly by their nucleic acid binding or perhaps by the presence of the diazirine group (Figure 4C).[32] PARP-1 was not significantly UV-enriched by probe **6** (data not shown) which agrees with the in-gel fluorescence data for this probe that showed fainter bands for PARP-1 compared with the background. Apparently, removal of the piperazine ring diminishes the binding affinity of probe **6** to PARP-1.

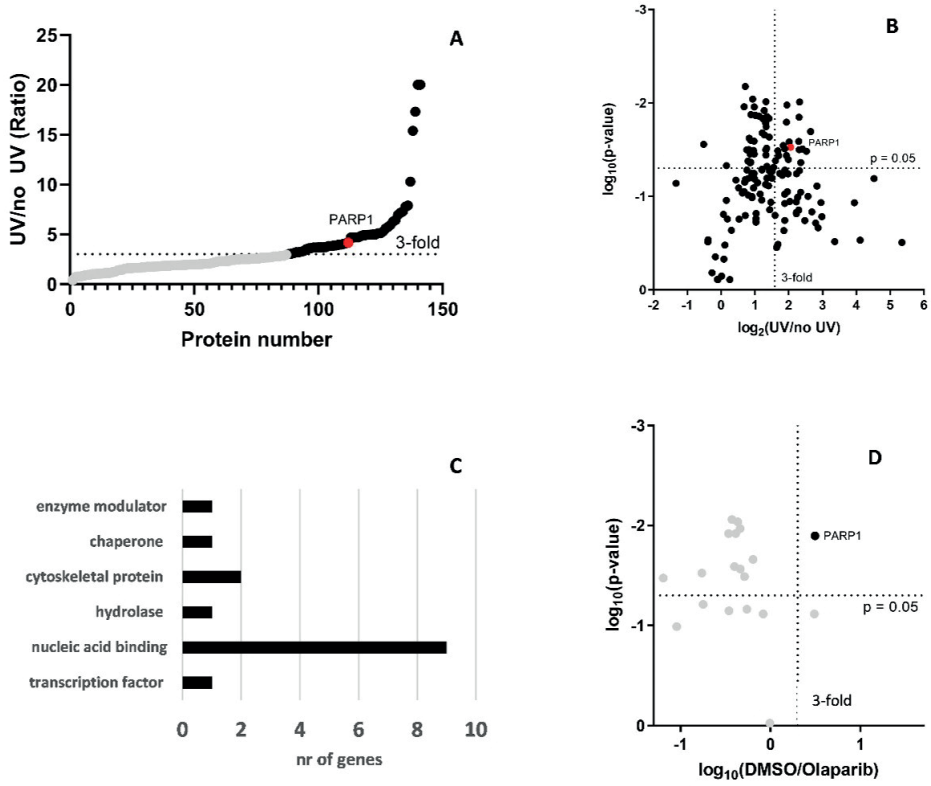


Figure 4. LFQ proteomics analysis of probe **3** in living Raji cells of proteins with at least 2 unique peptides. **A:** Waterfall plot of UV enrichment by probe **3** of proteins identified in the proteomics analysis, proteins showing >3-fold UV enrichment are in black. **B:** Volcano plot of UV enrichment of proteins by probe **3.** PARP-1 is marked in red. **C:** Significantly UV-enriched proteins sorted with the gene function analysis module of the PANTHER classification system (v.14.0) software.[33] **D:** Competition analysis of significantly UV-enriched probe targets, proteins outcompeted >3-fold by olaparib preaddition with a p-value < 0.05 using a Student's t-test are in black.

Conclusion

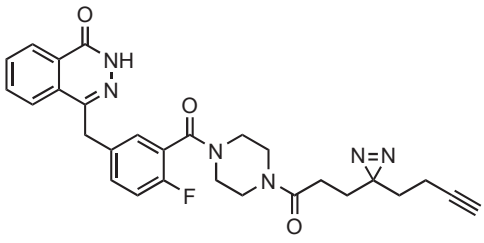
The design and synthesis of two synthetically accessible olaparib-based photoaffinity probes (**3** and **6**) for PARP-1 detection in living cells is presented in this chapter. Evidence for the efficacy of the probes is shown and a detailed protocol for application of these probes is provided therewith. Both **3** and **6** proved to be stable under oxidative conditions often employed for PARP activation while probe **3** was successfully used to label PARP-1 selectively in a live cell model and to detect its presence by in-gel fluorescence and LFQ LC-MS/MS methods.

Experimental section

General synthetic procedures

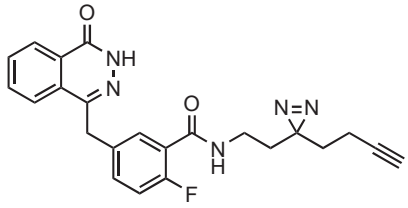
All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and ACN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. Column chromatography was performed on silica gel 60 Å (40-63 µm, Macherey-Nagel). TLC analysis was performed on Macherey-Nagel aluminium sheets (silica gel 60 F_{254}). TLC was used to visualize compounds by UV at wavelength 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) or $KMnO_4$ spray (20 g/L $KMnO_4$ and 10 g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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). NMR spectra were recorded on a Bruker AV-500 NMR. Chemical shifts (δ) are given in ppm relative to tetramethyl silane as internal standard. Coupling constants (*J*) are given in Hz. All given ^{13}C-APT spectra are proton decoupled.

4-(3-(4-(3-(3-(but-3-yn-1-yl)-3*H*-diazirin-3-yl)propanoyl)piperazine-1-carbonyl)-4-fluorobenzyl) phthalazin-1(2*H*)-one (3)



Compound **1** (38 mg, 0.10 mmol, 1.0 eq.), compound **2** (19 mg, 0.11 mmol, 1.1 eq.) and DIPEA (40 µL, 0.23 mmol, 2.2 eq.) were dissolved in DMF (1 mL). HCTU (47 mg, 0.11 mmol, 1.1 eq.) was added and the reaction was stirred at room temperature for 1.5 h. TLC showed full conversion, R_f = 0.73; DCM:MeOH=9:1 + three drops of TEA. The reaction was quenched by the addition of water and transferred into a separatory funnel. The water layer was extracted with EtOAc and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by silica gel column chromatography (1 -> 5% MeOH in DCM) to obtain the title compound **3** (17 mg, 0.033 mmol, 33%) as a white solid. **^{1}H-NMR** (500 MHz, $CDCl_3$) δ 12.12 – 11.89 (m, 1H, H-9), 8.48 (dd, J = 6.3, 3.2 Hz, 1H, H-6), 7.86 – 7.61 (m, 3H, H-1, H-2, H-3), 7.41 – 7.32 (m, 2H, H-14, H-18), 7.04 (td, J = 8.8, 2.4 Hz, 1H, H-15), 4.31 (s, 2H, H-12), 4.00 – 3.15 (m, 8H, H-22, H-23, H-25, H-26), 2.12 (t, J = 7.6 Hz, 1H, H-38), 2.07 – 1.92 (m, 4H, H-31, H-33), 1.88 (q, J = 7.4 Hz, 2H, H-36), 1.74 – 1.59 (m, 2H, H-29). **^{13}C-NMR** (126 MHz, $CDCl_3$) δ 169.8 + 169.7 (C-28, rotamers), 165.3 + 165.0 (C=O amide, rotamers), 161.1 (C-10), 157.9 + 156.0 (C-16, *J* = 239 Hz), 145.5 (C-7), 134.5 + 134.5 (C-13, *J* = 5.0 Hz), 133.6 (C-1/2/3), 131.8 + 131.8 (C-14, *J* = 7.5 Hz), 131.6 (C-1/2/3), 129.5 (C-4), 129.4 + 129.2 (C-18, rotamers), 128.2 (C-5), 127.1 (C-6), 125.0 (C-1/2/3), 123.6 + 123 (C-17, *J* = 17.5 Hz), 116.3 + 116.1 and 116.2 + 116.0 (C-15, *J* = 22.5 Hz) , 82.7 (C-37), 69.3 (C-38), 47.0 + 46.7 + 45.5 + 45.0 + 42.1 + 41.9 (C-22/23/25/26), 37.7 + 37.6 (C-12), 32.5 (C-29), 27.9 (C-31/C-33) 27.8 (C-31/C-33), 13.3 (C-36). **LC-MS:** 10 -> 90% B in A, Rt = 6.24. **HRMS:** $[C_{28}H_{27}FN_6O_3 + H]^+$: 515.2194 found, 515.2201 calculated

N-(2-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)ethyl)-2-fluoro-5-((4-oxo-3,4-dihydrophthalazin-1-yl)methyl)benzamide (6)



Compound **4** (60 mg, 0.20 mmol, 1.1 eq.), compound **5** (25 mg, 0.18 mmol, 1.0 eq.) and DIPEA (70 µL, 0.40 mmol, 2.2 eq.) were dissolved in DMF (1.5 mL). HCTU (83 mg, 0.20 mmol, 1.1 eq.) was added and the reaction was stirred for 16 h at room temperature. The reaction was quenched by the addition of water and transferred into a separatory funnel. The water layer was extracted with EtOAc and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by silica gel column chromatography (2->20% MeOH in DCM) to obtain the title compound **6** (57 mg, 0.14 mmol, 75%) as a white solid. **^{1}H NMR** (500 MHz, $CDCl_3$) δ 11.27 (s, 1H, H-9), 8.55 – 8.39 (m, 1H, H-6), 8.09 (dd, *J* = 7.4, 2.5 Hz, 1H, H-18), 7.81 – 7.70 (m, 3H, H-1, H-2, H-3), 7.37 (ddd, *J* = 8.3, 4.8, 2.5 Hz, 1H, H-14), 7.06 (dd, *J* = 11.6, 8.5 Hz, 1H, H-15), 6.87 (dt, *J* = 12.1, 5.7 Hz, 1H, H-21), 4.33 (s, 2H, H-12), 3.34 (q, *J* = 5.5 Hz, 2H, H-23), 2.03 (td, *J* = 7.3, 2.6 Hz, 2H, H-27), 1.99 (t, *J* = 2.6 Hz, 1H, H-29), 1.82 (t, *J* = 6.8 Hz, 2H, H-24), 1.69 (t, *J* = 7.3 Hz, 2H, H-26). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 163.3 + 163.3 (C-20), 160.9 (C-10), 160.7 + 158.7 (C-16, *J* = 246 Hz), 145.8 (C-7), 134.4 + 134.4 (C-13, *J* = 2.5 Hz), 133.8 (C1/2/3), 133.5 + 133.5 (C-14, *J* = 8.8 Hz), 132.0 + 132.0 (C-18, *J* = 2.5 Hz), 131.6 (C1/2/3), 129.6 (C-4), 128.4 (C-5), 127.2 (C-6), 125.2 (C1/2/3), 121.0 + 120.9 (C-17, *J* = 11 Hz), 116.8 + 116.6 (C-15, *J* = 25 Hz), 82.7 (C-28), 69.5 (C-29), 37.9 (C-12), 35.0 (C-23), 32.6 (C-24), 32.3 (C-26), 13.3 (C-27). **LC-MS:** 10 -> 90% B in A, Rt = 6.37. **HRMS:** $[C_{23}H_{20}FN_5O_2 + H]^+$: 418.1668 found, 418.1674 calculated.

Cell culture

Raji cells were cultured in phenol-red free RPMI-1640 medium buffered with sodium bicarbonate (Sigma-Aldrich) and freshly supplemented with new born calf serum (NBS, Hyclone), Glutamax, penicillin and streptomycin (200 µg/mL each, Duchefa) at 37°C in a humidified incubator with 7% CO_2. Phenol red was excluded from the medium to prevent UV light quenching by the pH indicator. Cells were growing happily and were passaged twice every week allowing for growth from 0.5 to 2-3 x 10^6 per mL in T175 flasks supplied with 100 mL fresh medium. For each experiment, 150 x 10^6 of logarithmically growing cells in 75 mL conditioned cell culture medium were seeded in T175 flasks and treated with 75 µL H_2O (control) or 75 µL 500 mM H_2O_2 (500 µM final) to induce double strand breaks for 30 minutes in the incubator. Cells were pelleted by centrifugation (5 minutes, 200 *g*), the supernatant was removed, the cell pellet was washed by trituration in 20 mL colourless RPMI-1640 medium supplemented with Glutamax (RPMI-g), spun down (5 minutes, 200 *g*) after which the cells were suspended in 3 mL RPMI-g and divided in 3 equal parts in 2 mL Eppendorf tubes.

Olaparib treatment

The cells were treated with 10 µM olaparib (2 mM in DMSO, Cayman Chemicals) or vehicle (5 µl of DMSO) and were incubated for 1 h at 37°C with shaking at 600 rpm. In case of sequential olaparib competition, the cells were pelleted, suspended in 2 mL RPMI-g, pelleted again, suspended in 1 mL RPMI-g containing 10 µM probe **3** or **6** or vehicle (5 µL DMSO) and were incubated for 2 h at 37 °C with shaking at 600 rpm. In case of simultaneous olaparib competition, µL 10 µM probe **3** or **6** or vehicle (5 µL DMSO) was added and cells were incubated for 2 h at 37°C with shaking at 600 rpm. The maximum DMSO concentration was 1%, probes were at 10 µM and olaparib was at 10 and 50 µM (1:1 or 5:1 of olaparib:probe ratios) end concentrations. After the incubation, cells were pelleted by centrifugation (5 minutes, 200 *g*), washed with 2 mL RPMI-g by pelleting (5 minutes, 200 *g*) and trituration in fresh RPMI-g, washed with 2 mL phosphate buffered saline (PBS) at 37 °C, washed with 2 mL ice cold PBS and suspended in 1 mL ice cold PBS.

UV-irradiation

Cells were plated on a 35 mm Petri dish and irradiated with 350 nm UV light for 30 minutes on a metal platform cooled at 2 °C (Caprotec box). The Caprobox contains 3 x 8.6 Watt UV lamps (Philips PL-S 9W BLB 2-Pin) that cover a surface of 0.01 m^2 and delivers 15.5 kJ in 30 minutes. A systematic variation to determine the optimal amount of UV light exposure was not performed, however pilot experiments with 10 and 20 minutes exposure yielded less pronounced PARP bands (data not shown). For no-UV controls, cells were plated on Petri dishes and incubated on ice for 30 minutes.

Cell lysis

Cells were collected, pelleted by centrifugation (4 °C, 5 minutes, 200 *g*), and the supernatant was removed. To prevent unwanted post-lysis processing of PARP and loss in reporter tag installation efficiency, the cell lysis and the CuAAC click chemistry were performed under denaturing conditions. Pilot experiments with lysis and CuAAC click in native protein environments (50 mM HEPES pH 7.4, 150 mM NaCl, protease inhibitors) showed less pronounced PARP bands after SDS-PAGE analysis, while native lysis followed by denaturation with 8 M urea prior to the click chemistry showed stronger bands (data not shown). The cell pellet was lysed under denaturing conditions by adding 300 µL lysis buffer (8 M urea, 10 mM HEPES, pH 7.4) and trituration until a highly viscous solution was reached due to the released DNA. The lysates were left on ice for 15 minutes, followed by DNA shearing with 5 pulses of 5 seconds by tip sonication on ice. Cell debris (small translucent pellet) was removed by centrifugation (4 °C, 15 minutes, 20.000 *g*), the supernatant transferred to a new tube and the protein concentration ranged from 8 — 15 µg/µL as determined with the Bradford colorimetric assay using a calibration curve of bovine serum albumin (BSA).

Copper catalysed click reaction

75 µg of protein (5 µL lysate) was mixed with 5 µL lysis buffer and 10 µL click cocktail and allowed to react at room temperature (RT) for 1 h in the dark. The click cocktail was made by first mixing 8 µL aq. $CuSO_4$ (100 mM) and 8 µl aq. sodium ascorbate (1 M) that resulted in a colour change from light blue to orange/brown due to reduction of Cu^{2+} to Cu^{+}. Addition of 8 µL THPTA (100 mM in DMSO) ligand that coordinates and stabilises the Cu^{+} species made the solution turn colourless. To prevent side reaction with amines of arginine residues, 8 µL aq. aminoguanidine (1 M) was added followed by 233 µL HEPES (100 mM, pH 7.4) and 0.8 µL AlexaFluor 647 azide (2 mM, Thermo Fisher Scientific) for fluorescent detection or diazo-biotin azide (2 mM, Click Chemistry Tools) for western blot analysis with Streptavidin-AlexaFluor 647 (Thermo Fisher Scientific) or affinity purification with MyOne C1 beads (Thermo Fisher Scientific). 7 µL 4X concentrated Laemmli sample buffer (SB) was added followed by heating at 95 °C for 5 minutes.

SDS PAGE analysis

Proteins were separated on a 1.5 mm, 10% SDS-PAGE gel at 100 V for 10 min and 130 V for another 3 hours, after which fluorescent signal was recorded on a Chemidoc MP system Bio-Rad). For reproducible PARP1 results, a minimum of 50 µg total protein loading per lane is necessary which justifies the use of the 1.5 mm thickness gels. In some cases, we have experienced unreproducible background staining possibly caused by the excess click cocktail reagents (data not shown). To eliminate background staining and clean the sample, a chloroform/methanol protein precipitation step between the click and SB steps was performed. After the click reaction, 80 µl lysis buffer was added followed by 400 µl methanol and vigorous mixing, then 100 µl chloroform, vigorous mixing and 300 µl water and vigorous mixing. Water addition leads to phase separation and with proteins precipitating out of solution as

white particles. Centrifugation at 10.000 g for 5 minutes affords a 3-layer system with water on top, organic layer at the bottom and a white protein interface as a thin film in the middle. The water layer is removed and 300 µl methanol is added followed by gentle mixing. Centrifuge at 10.000 g for 5 minutes to get the protein film at the bottom of the tube, remove the organic solvent completely and dry the tube for exactly 5 minutes at air in the fume hood. The protein film was solubilised in 30 µl 1x SB and heating at 95 °C for 5. This cleaning step is down scalable to protein amounts of minimum of 40 µg, less protein makes visibility of the protein film quite difficult and does not cause significant loss of PARP protein molecules. Coomassie served as loading control.

Western blot analysis

For PARP1 WB analysis, gels were first imaged by fluorescence followed by protein transfer to nitrocellulose membranes (Bio-Rad) using the high MW protocol (10 minutes at 2.5A, 25V) on a Trans-blot turbo transfer system (Bio-Rad) The precision plus dual color protein standard (Bio-Rad) was diluted 10x with 1xSB and 2 µL were loaded to assess the running of the gel and the electro transfer of the proteins to the nitrocellulose membranes. The membranes were washed with water and TBST (Tris buffered saline, 0.1% Tween 20) and subsequently blocked using 5% non-fat milk in TBST (30 minutes). To this, 4 µL of a rabbit-anti-PARP1 (9542, Cell Signalling Technology) antibody (1:2500 dilution) was added and incubated overnight at 4 °C. The blot was washed with 20 mL TBST (4X, 5 minutes), blocked again for 30 minutes in 10 mL 5% milk, and rolled with 5 µL (1:2000) secondary anti-rabbit IgG, HRP-linked antibody (7074, Cell Signalling Technology) for 1 h at rt. The membrane was washed with 20 mL TBST(4X, 5 minutes), TBS, water and subsequently imaged with Clarity max western ECL substrate (Bio-Rad) on the Chemidoc MP system. For biotin visualisation, the western blot procedure for PARP1 was used with several modifications. The blocking solution was 3% BSA in TBST, hybridisation with Streptavidin-AlexaFluor647 (1:2000) was used to tag biotinylated proteins for 1 h at room temperature or overnight at 4 °C.

Label free quantification by LC-MS/MS

For label-free quantification by LC-MS/MS-based proteomics (LFQ), biotin-N_3 was installed by CuAAC in 1 mg of protein in 100 µL lysis buffer reacted with 50 µL click cocktail for 1 h at RT in the dark followed by Bligh&Dyer precipitation to remove excess reagents. The protein film was taken in 100 µL 1% SDS/100 mM HEPES pH 7.4, allowed 30 minutes to dissolve with shaking at 600 rpm at RT, centrifuged at 20.000 g for 5 minutes at RT to remove particles and the solution transferred to a new tube. Disulfide bridges were reduced by addition of 5 µL 100 mM DTT (5 mM end concentration) for 30 minutes at 37°C and alkylated with 5 µL 400 mM iodoacetamide (20 mM end concentration) for 30 minutes at rt in the dark. Excess reagents were removed by Bligh&Dyer precipitation and the protein film was dispersed in 50 µL 2 % SDS/100 mM NH_4HCO_3 by vigorous vortexing. Stepwise dilution of 3 x 50 µl, 2 x 100 µl, 500 and 1000 µL pull down (PD) buffer (50 mM Tris pH7.4, 150 mM NaCl) followed by vigorous vortexing afforded a clear solution containing 0.05% SDS end concentration that is compatible with the streptavidin beads. After centrifugation at 20.000 g for 10 minutes to remove particles, the supernatant was transferred to a new tube. The pulldown was initiated by addition of 50 µL MyOne C1 streptavidin Dynabeads (Thermo Fisher Scientific) equilibrated in PD buffer and incubated for 1 h at RT or overnight at 4°C. After the incubation, the beads were collected for 2 minutes on the DynaMag (Thermo Fisher Scientific) and the supernatant was transferred to a new tube. The beads were washed with 2 × 500 µL of wash buffer I (4 M urea/50 mM NH_4HCO_3, 150 mM NaCl), 2 × 500 µL of wash buffer II (50 mM Tris pH = 7.4, 10 mM NaCl) and taken up in 200 µL on-bead digestion buffer (100 mM Tris pH 7.4, 10 mM NaCl, 1 mM $CaCl_2$, 2% MeCN in MilliQ water) followed by addition of 250 ng trypsin

(Promega). The beads were incubated at 37 °C overnight with shaking, collected on the DynaMag for 2 minutes and the supernatants containing the peptides were transferred to fresh tubes to which 5 µL of formic acid was added. Peptides were desalted using StageTips, concentrated to dryness, dissolved in 30 µL 96.9 % H_2O, 3% MeCN, 0.1% formic acid containing 10 fmol/µL enolase digest (Thermo Fisher Scientific) and 5 µL were analysed by nanoAcquity hyphenated to a Synapt G2Si mass spectrometer as previously described.[15] Peptide identification and LFQ was done by processing the raw data files with the Progenesis QIP software, proteins with at least 2 unique peptides were selected, normalised to enolase and the significance was determined with a Student's t-test and plotted versus the fold change for UV/no UV and with or without olaparib competition.

References

[1] S. Messner, M. O. Hottiger, *Trends Cell Biol.* **2011**, *21*, 534–542.

[2] K. Ueda, A. Omachi, M. Kawaichi, O. Hayaishi, *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 205–209.

[3] T. Boulikas, *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 3499–3503.

[4] A. Huletsky, C. Niedergang, A. Fréchette, R. Aubin, A. Gaudreau, G. G. Poirier, *Eur. J. Biochem.* **1985**, *146*, 277–285.

[5] N. A. Berger, G. W. Sikorski, *Biochemistry* **1981**, *20*, 3610–3614.

[6] E. G. Miller, *BBA Sect. Nucleic Acids Protein Synth.* **1975**, *395*, 191–200.

[7] S. Messner, M. Altmeyer, H. Zhao, A. Pozivil, B. Roschitzki, P. Gehrig, D. Rutishauser, D. Huang, A. Caflisch, M. O. Hottiger, *Nucleic Acids Res.* **2010**, *38*, 6350–6362.

[8] S. L. Rulten, A. E. O. Fisher, I. Robert, M. C. Zuma, M. Rouleau, L. Ju, G. Poirier, B. Reina-San-Martin, K. W. Caldecott, *Mol. Cell* **2011**, *41*, 33–45.

[9] B. Lüscher, M. Bütepage, L. Eckei, S. Krieg, P. Verheugd, B. H. Shilton, *Chem. Rev.* **2018**, *118*, 1092–1136.

[10] C. J. Lord, A. Ashworth, *Science* **2017**, *355*, 1152–1158.

[11] N. Curtin, *Biochem. Soc. Trans.* **2014**, *42*, 82–88.

[12] N. Liu, S. Hoogendoorn, B. van de Kar, A. Kaptein, T. Barf, C. Driessen, D. V. Filippov, G. A. van der Marel, M. van der Stelt, H. S. Overkleeft, *Org. Biomol. Chem.* **2015**, *13*, 5147–5157.

[13] M. P. Baggelaar, H. den Dulk, B. I. Florea, D. Fazio, N. Bernabò, M. Raspa, A. P. A. Janssen, F. Scavizzi, B. Barboni, H. S. Overkleeft, M. Maccarrone, M. van der Stelt, *ACS Chem. Biol.* **2019**, *14*, 2295–2304.

[14] A. C. M. van Esbroeck, A. P. A. Janssen, A. B. Cognetta, D. Ogasawara, G. Shpak, M. van der Kroeg, V. Kantae, M. P. Baggelaar, F. M. S. de Vrij, H. Deng, M. Allarà, F. Fezza, Z. Lin, T. van der Wel, M. Soethoudt, E. D. Mock, H. den Dulk, I. L. Baak, B. I. Florea, G. Hendriks, L. De Petrocellis, H. S. Overkleeft, T. Hankemeier, C. I. de Zeeuw, V. Di Marzo, M. Maccarrone, B. F. Cravatt, S. A. Kushner, M. van der Stelt, *Science* **2017**, *356*, 1084–1087.

[15] E. J. van Rooden, B. I. Florea, H. Deng, M. P. Baggelaar, A. C. M. van Esbroeck, J. Zhou, H. S. Overkleeft, M. van der Stelt, *Nat. Protoc.* **2018**, *13*, 752–767.

[16] S. Pan, H. Zhang, C. Wang, S. C. L. Yao, S. Q. Yao, *Nat. Prod. Rep.* **2016**, *33*, 612–620.

[17] M. P. Baggelaar, P. J. P. Chameau, V. Kantae, J. Hummel, K. L. Hsu, F. Janssen, T. van der Wel, M. Soethoudt, H. Deng, H. den Dulk, M. Allarà, B. I. Florea, V. Di Marzo, W. J. Wadman, C. G. Kruse, H. S. Overkleeft, T. Hankemeier, T. R. Werkman, B. F. Cravatt, M. van der Stelt, *J. Am. Chem. Soc.* **2015**, *137*, 8851–8857.

[18] X. Yang, T. J. M. Michiels, C. de Jong, M. Soethoudt, N. Dekker, E. Gordon, M. van der Stelt, L. H. Heitman, D. van der Es, A. P. Ijzerman, *J. Med. Chem.* **2018**, *61*, 7892–7901.

[19] M. Soethoudt, S. C. Stolze, M. V. Westphal, L. van Stralen, A. Martella, E. J. van Rooden, W. Guba, Z. V. Varga, H. Deng, S. I. Van Kasteren, U. Grether, A. P. Ijzerman, P. Pacher, E. M. Carreira, H. S. Overkleeft, A. Ioan-Facsinay, L. H. Heitman, M. van der Stelt, *J. Am. Chem. Soc.* **2018**, *140*, 6067–6075.

[20] Z. Wang, Z. Guo, T. Song, X. Zhang, N. He, P. Liu, P. Wang, Z. Zhang, *ChemBioChem* **2018**, *19*, 2312–2320.

[21] M. Jouanneau, B. McClary, J. C. P. Reyes, R. Chen, Y. Chen, W. Plunkett, X. Cheng, A. Z. Milinichik, E. F. Albone, J. O. Liu, D. Romo, *Bioorganic Med. Chem. Lett.* **2016**, *26*, 2092–2097.

[22] K. S. Yang, G. Budin, C. Tassa, O. Kister, R. Weissleder, *Angew. Chemie - Int. Ed.* **2013**, *52*, 10593–10597.

[23] C. E. Knezevic, G. Wright, L. L. Remsing Rix, W. Kim, B. M. Kuenzi, Y. Luo, J. M. Watters, J. M. Koomen, E. B. Haura, A. N. Monteiro, C. Radu, H. R. Lawrence, U. Rix, *Cell Chem. Biol.* **2016**, *23*, 1490–1503.

[24] F. Zmuda, G. Malviya, A. Blair, M. Boyd, A. J. Chalmers, A. Sutherland, S. L. Pimlott, *J. Med. Chem.* **2015**, *58*, 8683–8693.

[25] T. Reiner, S. Earley, A. Turetsky, R. Weissleder, *ChemBioChem* **2010**, *11*, 2374–2377.
[26] E. D. Deeks, *Drugs* **2015**, *75*, 231–240.
[27] K. A. Menear, C. Adcock, R. Boulter, X. L. Cockcroft, L. Copsey, A. Cranston, K. J. Dillon, J. Drzewiecki, S. Garman, S. Gomez, H. Javaid, F. Kerrigan, C. Knights, A. Lau, V. M. Loh, I. T. W. Matthews, S. Moore, M. J. O'Connor, G. C. M. Smith, N. M. B. Martin, *J. Med. Chem.* **2008**, *51*, 6581–6591.
[28] P. Geurink, L. Prely, G. A. van der Marel, R. Bischoff, H. S. Overkleeft, *Photoaffinity Labeling in Activity-Based Protein Profiling*, **2011**.
[29] Z. Li, P. Hao, L. Li, C. Y. J. Tan, X. Cheng, G. Y. J. Chen, S. K. Sze, H. M. Shen, S. Q. Yao, *Angew. Chemie - Int. Ed.* **2013**, *52*, 8551–8556.
[30] E. Wahlberg, T. Karlberg, E. Kouznetsova, N. Markova, A. Macchiarulo, A. G. Thorsell, E. Pol, Å. Frostell, T. Ekblad, D. Öncü, B. Kull, G. M. Robertson, R. Pellicciari, H. Schüler, J. Weigelt, *Nat. Biotechnol.* **2012**, *30*, 283–288.
[31] W. Ma, C. J. Halweg, D. Menendez, M. A. Resnick, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6590–6595.
[32] P. Kleiner, W. Heydenreuter, M. Stahl, V. S. Korotkov, S. A. Sieber, *Angew. Chemie - Int. Ed.* **2017**, *56*, 1396–1401.
[33] H. Mi, A. Muruganujan, X. Huang, D. Ebert, C. Mills, X. Guo, P. D. Thomas, *Nat. Protoc.* **2019**, *14*, 703–721.

Chapter 9

Summary and Future Prospects

Summary and Future Prospects

Although tremendous efforts have led to a significant improvement of synthetic methodologies to obtain ADP-ribosylated peptide fragments, work prior to that presented in this Thesis has solely led to ADPr-peptides where the natural aspartic acid (Asp) or glutamic acid (Glu) ADPr residues were substituted with asparagine (Asn) and glutamine (Gln) ADPr.[1,2] In the case of arginine (Arg), Arg- to citrulline-ADPr[2] or triazoles served as isosteric replacements for the native ADPr linkage.[3] The synthetic efforts mentioned above are discussed in more dept in Chapter 1. The work presented in this Thesis is mostly focused on synthesizing native ADP-ribosylated compounds. This final chapter will summarize the results and outline interesting and promising lines of research for future explorations.

Chapter 2

Chapter 2 is focused on synthesizing and investigating the native chemical structure of serine-(Ser) ADPr.[4] This modification was recently discovered and it turned out that serine is the main acceptor of ADP-ribosylation in the DNA damage response pathway.[5,6] Up to now it is generally believed that ARTs transfer NAD^+ in an α-fashion. However, for the formation of Ser-ADPr, a previously unencountered complex of PARP1 with HPF1 is formed,[7,8] giving rise to the question if the α-selectivity, which is observed for other acceptors, persists for Ser-ADPr. Therefore, two phosphoribosylated Ser building blocks were prepared with either the α (**1**)- or β (**2**)-configuration (Figure 1). These were then used to synthesize *mono*-ADP-ribosylated (MARylated) peptides containing an α- or β-glycosidic linkage between the ribosyl moiety of ADPr and the Ser side chain using a modified SPPS methodology.[2,4] It was hypothesized that substrate recognition by the human ADPr hydrolase 3 (hARH3), which is specific for Ser-ADPr,[9] would be selective for the native epimer. Treatment of α- and β-ADP-ribosylated peptides **3** and **4** with ARH3 indeed revealed α-Ser-ADPr to be the native epimer as α-linked Ser-ADPr was efficiently hydrolyzed whereas the β-isomer stayed intact.

After confirming that the Ser side chain is ADP-ribosylated in an α-selective fashion by the PARP1-HPF1 complex, the chemical stability of this modification is of particular interest. Ser-ADP-ribosylated peptide **3** was treated with reagents commonly encountered in the field of biochemistry and proteomics of ADP-ribosylation[10,11] i.e. aqueous acid (0.1 M TFA), alkali (0.1 M NaOH) and NH_2OH (0.5 M). It turned out that Ser-ADPr undergoes β-elimination when exposed to NaOH with over 50% breakdown into dehydroalanine and free ADPr after 5 hours.

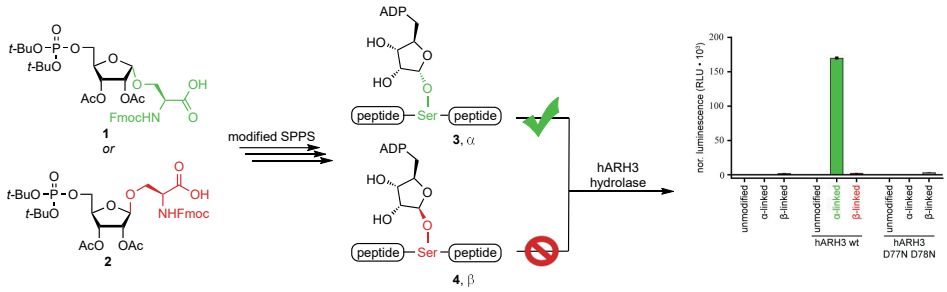
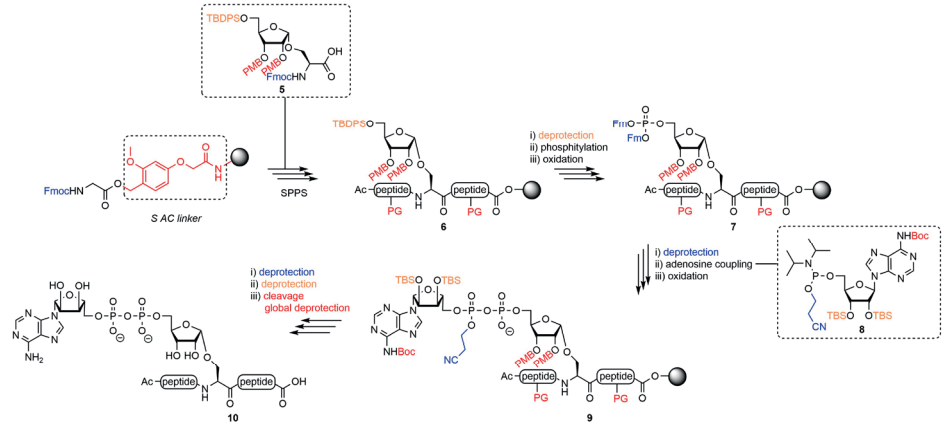


Figure 1. Schematic overview of Chapter 2.

Chapter 3

The development of a new methodology to synthesize ADP-ribosylated peptides is described in Chapter 3 and outlined in Scheme 1. The improvement this methodology brings about is two-fold: i) it eliminates the need for extensive protecting group manipulations during the solution phase synthesis of orthogonally protected, ribosylated amino acid building blocks like Ser **5** and ii) it harmonizes the protecting group patterns of the ADPr-moiety with that of the amino acids commonly applied in SPPS. For instance, Fmoc-Ser(Trt)-OH and Fmoc-Thr(Trt)-OH) are not compatible with the solid-phase strategy described in Chapter 2. Pursuing sequences with a Ser or Thr flanking the modification site, the trityl (Trt) protecting group of the hydroxyls in these amino acids must be replaced by an acetyl to allow further processing of the incorporated *t*Bu-protected phosphotriester. With the new methodology, acid sensitive protecting groups on amino acid side chains, can be kept throughout the entire synthetic route.



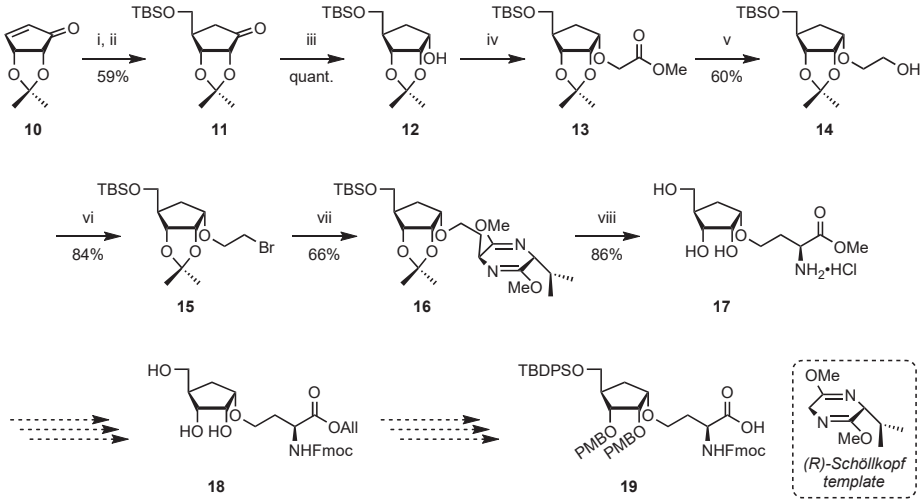
Scheme 1. Outline of the synthetic strategy developed in Chapter 3 to synthesize MARylated peptides.

With this newly developed strategy, a range of Ser-, Thr- and Cys-ADPr peptides have been synthesized and with these peptides a previously unknown hydrolase selective for *S*-ADPr (Cys) was discovered.

After demonstrating the effectiveness of the newly devised synthetic methodology, its scope can further be broadened by the synthesis of stabilized ADP-ribosylated analogues which can be used for antibody generation. The employment of ADPr specific antibodies has been explored in the past but was unproductive due to the inherent lability of the pyrophosphate moiety *in vivo* upon immunization of rabbits.[12–15] Furthermore, any crystal structures obtained from ARH3 with synthetic MARylated Ser-ADPr peptides are inherently derived from a catalytically inactive, mutated ARH3. The isosteres of ADP-ribosylated peptides which contain chemically stabilized mimics of the glycosidic linkage and pyrophosphate moieties can provide for a useful tool for *in vivo* studies and structural biology.

Stabilized mimic of the glycosidic ribosyl-Ser linkage in Ser-ADPr

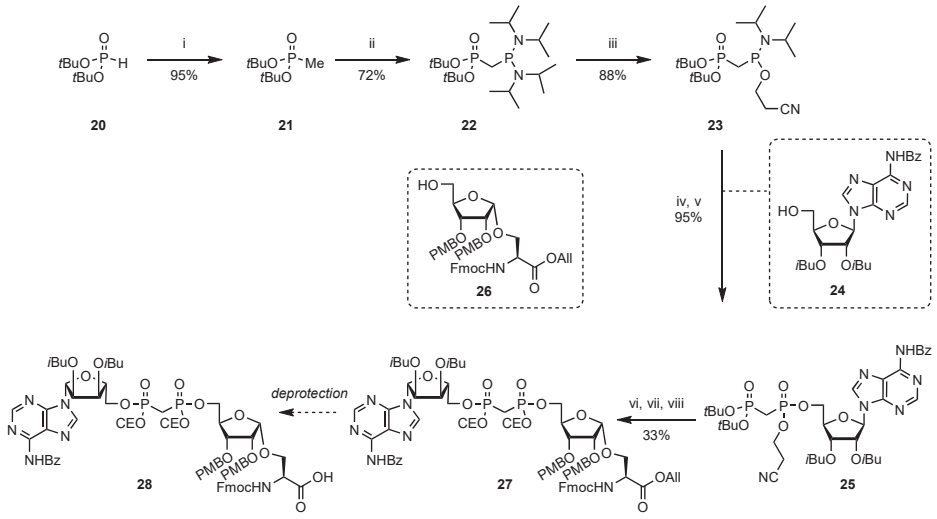
Stabilization of the glycosidic linkage by replacing the endocyclic oxygen with a methylene group to furnish *carba*-ribosyl Ser-ADPr will likely inhibit the hydrolytic activity of ARH3 towards this substrate. Synthetic routes to furnish *carba*-ribosyl precursors such as **10** (Scheme 2) are available in literature[16,17] and provide a starting point for the synthesis of *carba*-ribosyl homoserine **17**. This building block is stabilized in two ways: i) the Ser side chain is elongated with one additional methylene group, minimizing the possibility of elimination reactions described in Chapter 2[4] and ii) the endocyclic oxygen is replaced with a methylene group, rendering it unreactive towards ARH3 hydrolysis. The proposed synthetic route commences with known *carba*-ribose precursor **10** which is subjected to a 1,4-photoaddition with methanol, followed by silylation of the formed primary alcohol to yield **11** in 59% yield over two steps. Stereoselective reduction of ketone **11** forms α-*carba*-ribose **12**. Alkylation of the alcohol with methyl bromoacetate followed by complete reduction of the ester functionality furnished **14**. To allow decoration of the *carba*-ribose with Schöllkopf's template, alkyl bromide **15** was prepared. The efficiency of introduction of the template was enhanced by the addition of Ag_2CO_3 and hexamethyl phosphoramide (HMPA) to furnish **16** in 66% yield. Subsequent acidolysis of the bislactim with HCl and concomitant cleavage of the silyl ether and isopropylidene protecting groups, resulted in the isolation of homoserine derivative **17**. The proposed ensuing protecting group manipulation entails saponification of the methyl ester, protection of the amine with Fmoc and introduction of the allyl ester, leading to the conversion of **17** into **18**. Finally, introduction of earlier applied protecting groups would provide building block **19**, suitable for the same Fmoc based SPPS methodology to synthesize MARylated peptides as described in Chapter 3. With the aid of **19**, peptides analogous to Ser-ADPr but with a stabilized glycosidic linkage and not susceptible to β-elimination are within reach.



Scheme 2. Synthesis of *carba*-ribosylated homoserine building block compatible with Fmoc-based SPPS. Reagents and conditions: i) MeOH, Ph_2CO, *hv*, 60 °C. ii) TBS-Cl, imidazole, DCM. iii) $NaBH_4$, MeOH, 0 °C. iv) methyl bromoacetate, NaH, 15-crown-5, TBAI, THF. v) $LiAlH_4$, Et_2O, 0 °C. vi) Br_2, imidazole, PPh_3, 0 °C. vii) (*R*) — Schöllkopf template, *n*-BuLi, Ag_2CO_3, HMPA, THF, -78 °C. viii) aq. HCl, THF.

Stabilized isostere of the pyrophosphate moiety in Ser-ADPr

To obviate pyrophosphate cleavage in biological systems, methylene bisphosphonates analogues have been employed for pyrophosphate containing biomolecules.[18–21] In the field of ADPr research, such ADP-ribose analogues has been used to generate antibodies against ADPr but no specificity could be obtained thus far.[12] To overcome the notoriously difficult synthesis of non-symmetric methylene bisphosphonates, Engelsma *et al.*[22] developed phosphanylmethylphosphonate reagents that can be utilized in P^{III} – P^{V} chemistry and applied to the preparation of adenosine bisphosphonate ribose (ABPr). By employing the key reagent **23** using this methodology, Ser-conjugated ABPr **28** (Scheme 3) can be synthesized. The first step in this route comprises DCI mediated coupling of protected adenosine derivative **24** with **23** and subsequent oxidation with *t*BuOOH to give **25** in 96% yield. To introduce the ribosylated Ser moiety, both *t*Bu groups were removed with an HCl/HFIP cocktail.[23] Next, a double, PyNTP (3-nitro-1,2,4-triazol-1-yl-tris(pyrrolidin-1-yl)phosphonium hexafluorophosphate) mediated condensation step starting with a stoichiometric amount of 2-cyanoethanol followed by ribosyl building block **26** (easily prepared via synthons described in Chapter 3) furnished **27**. Cleavage of the allyl protecting groups will allow the isolation of ADP-ribosylated Ser isostere **28**, ready for Fmoc-based SPPS. Synthetic peptides, modified with the bisphosphonate stabilized bridge, are resistant to enzymatic turnover or spontaneous hydrolysis during immunization of animals and thus, possibly, Ser-ADPr specific antibodies can be generated.

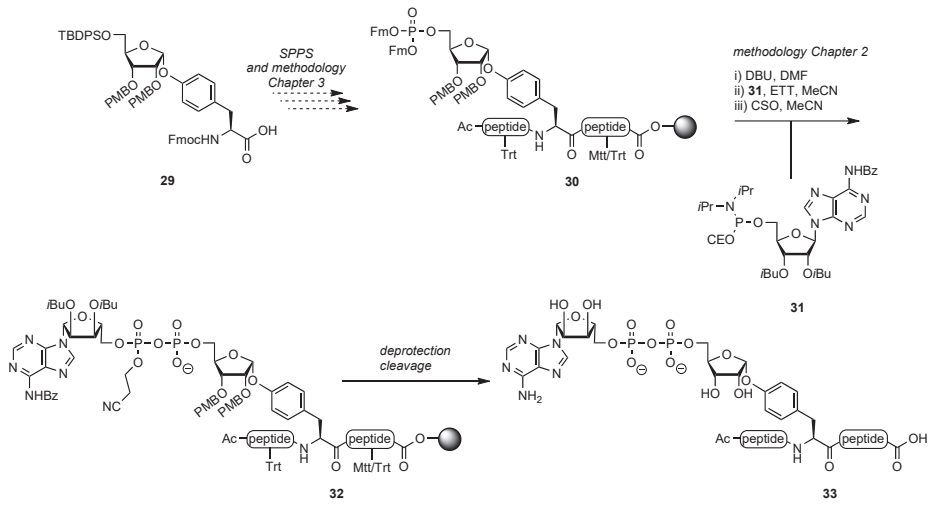


Scheme 3. Synthesis of Ser-ADPr building block **27**, stabilized with a methylene bisphosphonate bridge. Reagents and conditions: i) MeI, NaH, MeCN. 0 — 50 °C. ii) LDA, THF, -78 °C, then $((i\text{Pr})_2\text{N})_2\text{PCl}$. iii) 2-cyanoethanol, DCI, DCM. iv) **24**, DCI, MeCN, v) *t*BuOOH, nonane. vi) HCl, HFIP. vii) 2-cyanoethanol, PyNTP, MeCN. viii) **26**, PyNTP, MeCN.

Chapter 4

In Chapter 4 the strategy developed for the synthesis of Ser-, Thr- and Cys-ADPr peptides, was explored for Tyr-ADPr peptides. However, it turned out that the phenolic glycosidic linkage was cleaved when exposed to 10% TFA in DCM, which is required to remove the Boc protecting group of the adenine moiety at the end of the synthesis. Therefore, a strategy was invented that combined the approach presented in Chapter 3 with the one from Chapter 2. This alteration led to a methodology that allowed the efficient synthesis of Tyr-ADPr peptides. Such a 'hybrid' approach for the synthesis of Tyr-ADPr peptide making use of acid sensitive protections for the peptide and base sensitive ones for adenosine is described in detail in Chapter 4 and summarized here in Scheme 4.

First, building block **29** was prepared and used in SPPS as described in Chapter 3 to furnish intermediate immobilized peptide **30**. Next, the pyrophosphate moiety was installed using reagent **31**, described in Chapter 2, to furnish peptide **32**. Finally, a sequence of deprotection reactions which does include a treatment with TFA but in significantly lower concentrations (2.5% in DCM as opposed to 10%). Said conditions proved to be compatible with the phenolic glycosidic linkage yielding Tyr-ADPr peptides **33**. Biological studies with these peptides further demonstrated the permissiveness of ARH3 with respect to its substrates and revealed the unsuspected hydrolytic activity of PARG and ARH1 towards Tyr-ADPr.

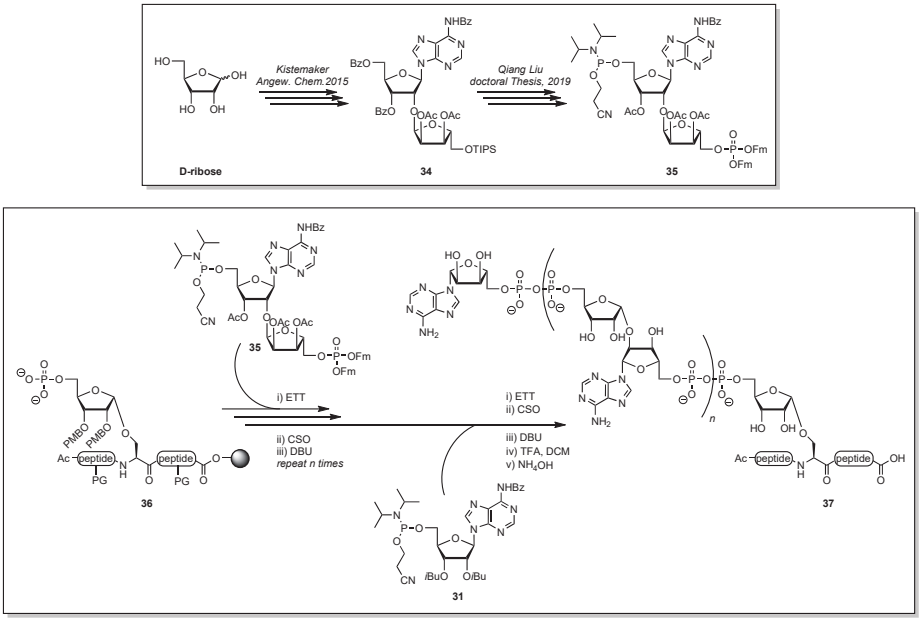


Scheme 4. Outline of the 'hybrid' strategy for Tyr-ADP-ribosylated peptides as described in Chapter 4.

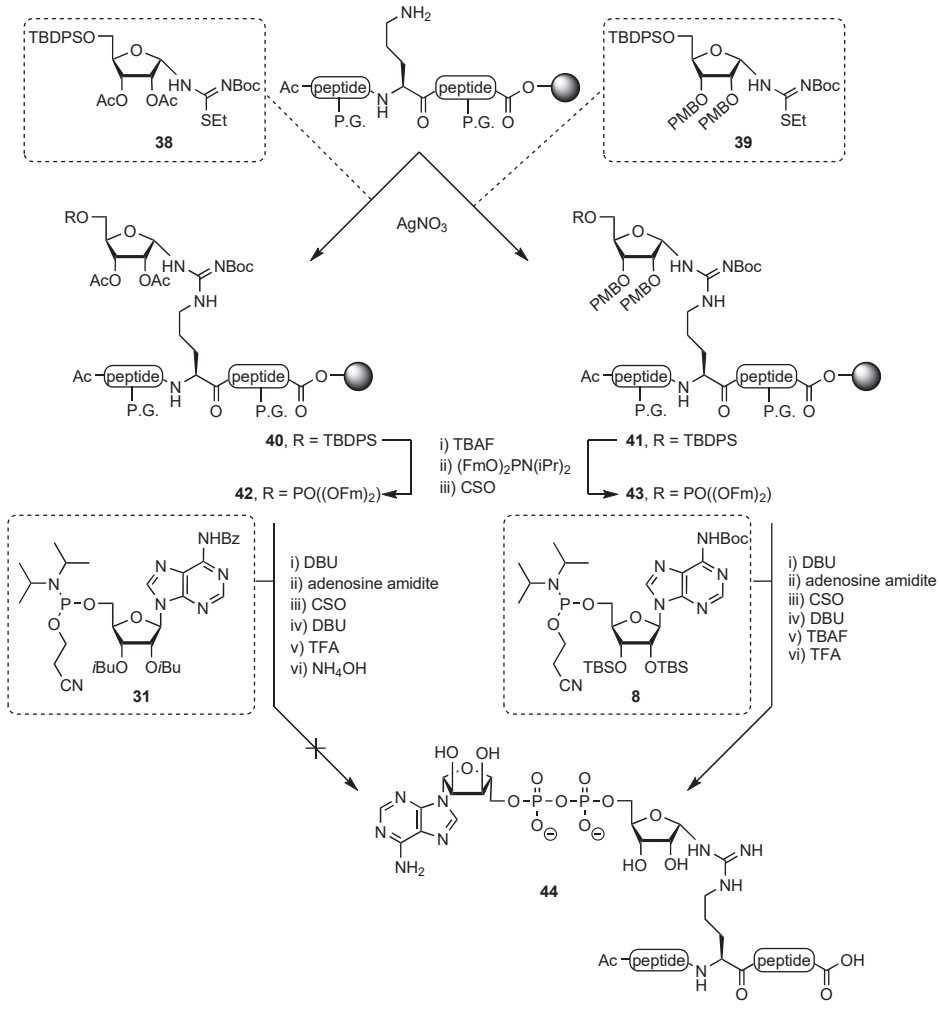
Beyond the thus assembled *O*-MARylated peptides, the next challenge can be found in PARylated peptides with a well-defined length of the PAR-chain. The synthesis of linear ADPr oligomers has been developed independently from the synthesis of MARylated peptides. To synthesize PAR chains on-resin, Kistemaker *et al.* designed a bifunctional disaccharide building block[24] containing a phosphoramidite and a phosphotriester. This methodology was refined by Liu *et al.*[25] by the design and application of disaccharide **35** (Scheme 5). It is proposed here to bring the strategies for the synthesis of ADP-ribosylated peptides and PAR-chains together. When bifunctional building block **35** is implemented in the synthetic strategy of MARylated peptides, described in Chapter 3, the coupling cycle can potentially be repeated *n* times starting from immobilized phosphomonoester **36**. When the desired chain length has been reached, adenosine phosphoramidite **31** can be used to terminate the PAR chain. Cleavage of the partially protected product from the resin followed by saponification of all remaining protecting groups will result in ADP-ribosylated peptides with a well-defined oligo-ADPr-chain, structures that remain inaccessible thus far.

Chapter 5

Having synthesized *O*- and *S*-ADPr peptides, one major native ADP-ribosylation site still remained synthetically inaccessible, namely Arg-ADPr. Chapter 5 focuses on developing a strategy to obtain the first fully synthetic peptides with this modification. Although an *N*-glycosidic linkage of the distal ribose to the Asn, Gln and Cit side chain could be built via glycosylation as reported by Kistemaker *et al.*[2] and as described in Chapters 2, 3 and 4 for the *O*- and S-glycosidic linkages, a different approach was opted for ADPr-Arg. Rather than glycosylation, a guanidyl riboside was appended at the side chain of ornithine using building blocks **38** or **39** (Scheme 6). As was demonstrated for Ser-, Thr-, Cys-, and Tyr-ADPr peptides, their glycosidic linkages show various reactivity towards basic, acidic, oxidative or nucleophilic conditions. Therefore, both the strategy described in Chapter 3 and the 'hybrid' strategy from Chapter 4 were investigated for their potential to synthesize Arg-ADPr peptides (Scheme 6). Indeed, the Arg-ADPr linkage proved labile towards nucleophilic bases such as aqueous ammonia which rendered the 'hybrid' strategy unsuitable for generating Arg-ADPr peptides. However, the ADP-ribosylated guanidyl group proved acid stable which not only allowed for the synthesis of Arg-ADPr peptides but also ubiquitin, MARylated on Arg42. This is the first report of a fully synthetic protein bearing ADPr on Arg.



Scheme 5. Synthesis of ADP-ribosylated peptides using a combination of the strategy for MARylated peptides presented in Chapter 3 and the work of Kistemaker[24] and Liu *et al.*[25]

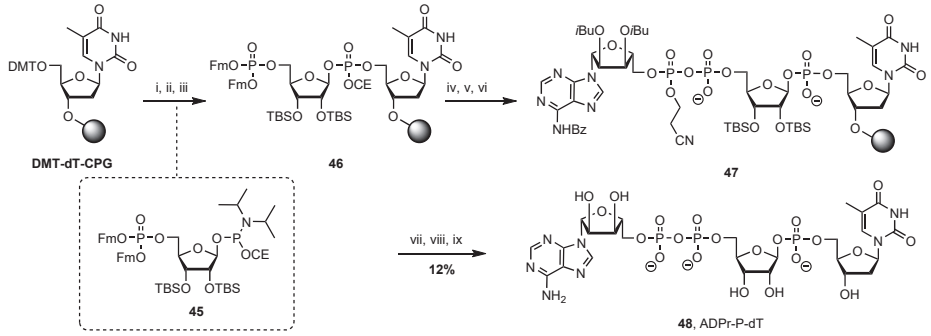


Scheme 6. Synthetic strategies described in Chapter 5 to synthesize Arg-ADP-ribosylated peptides.

Chapter 6

Apart from the ADP-ribosylation of specific amino acids in certain proteins, nucleic acids bearing a terminal phosphomonoester can undergo ADP-ribosylation.[26,27] Chapter 6 is dedicated to developing a methodology for the synthesis of ADP-ribosylated nucleotides, potentially suitable for an automated solid-phase oligonucleotide synthesis. Chapter 6 describes a novel, bifunctional ribosyl building block **45** (Scheme 7) for the synthesis of nucleotides, the 5'-phosphate of which is linked to ADP-ribose. The strategy was first explored in solution resulting in ADPr with a glycosidic linkage to AMP. After the successful solution-phase synthesis, the compatibility of ribosyl building block **45** with conventional

solid phase nucleotide synthesis was tested. After optimization of the chemistry, ADPr-P-dT **48** was manually synthesized on resin. This unprecedented synthesis of ADP-ribosylated nucleotide paves the way for automated ADPr oligonucleotide synthesis.

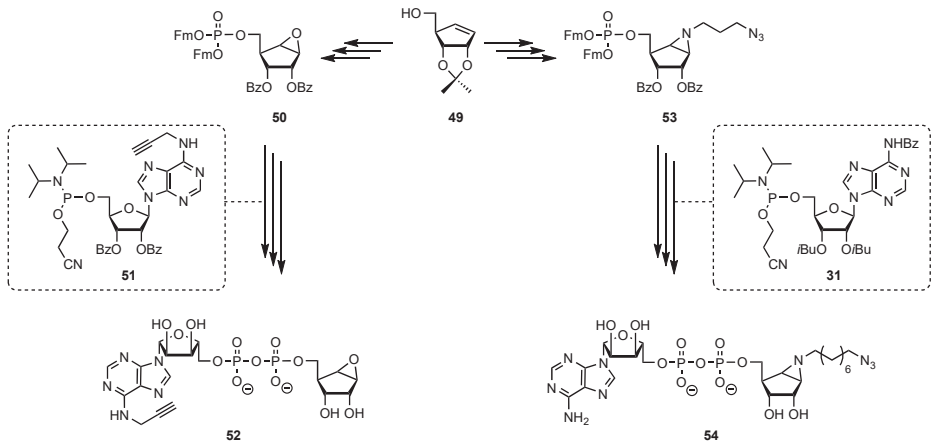


Scheme 7. Solid phase synthesis of ADPr 5'-linked to a nucleotide. Reagents and conditions: i) DCA, DCM. ii) **45**, activator 42®, MeCN. iii) CSO, MeCN. iv) DBU, MeCN. v) **31**, activator 42®, MeCN. vi) CSO, MeCN. vii) DBU, MeCN. viii) NH_4OH, H_2O. ix) TEA·3HF, TEA, NMP.

9

Chapter 7

Cluster of differentiation 38 (CD38) is a transmembrane protein responsible for the conversion of NAD^+ to free ADPr and cyclic ADPr (cADPr). Chapter 7 describes the synthesis of two potential activity-based probes (ABPs) for CD38 that utilize an electrophilic trap in the form of either an epoxide or aziridine to covalently bind with the catalytic Glu-residue in the active site of CD38. Similar mechanism-based probes employing epoxides and/or aziridines as electrophilic trap have been developed successfully for a variety of glycosidases.[28–31] The starting point of the synthetic route towards ABPs **52** and **54** is cyclopentene **49** (Scheme 8). In case of ABP **52**, precursor epoxide **50** bears an orthogonally protected 5'-phosphotriester. After its conversion into a phosphomonoester, the ribosyl moiety is suitable for a P^{III} – P^{V} coupling step with appropriately protected adenosine phosphoramidite **51**. The adenosine in ABP **52** is functionalized with a propargyl group on the exocyclic amine, enabling the installation of a reporter group via CuAAC for in-gel fluorescence or MS/MS proteomics. A similar route was pursued to obtain ABP **54**, the aziridine moiety of which is decorated with an azide containing linker for CuAAC ligation and thus previously described adenosine phosphoramidite **31** was ideally suited to synthesize ABP **54**. After the P^{III} – P^{V} coupling step, final removal of all protecting groups (i.e., *i*Bu and Bz) and HPLC purification, both ABP probes **52** and **54** were isolated in sufficient amounts to evaluate their binding affinity with CD38.



Scheme 8. Synthesis of potential ABPs **52** and **54** for CD38 as described in Chapter 7.

Chapter 8

Chapter 8 deals with the development of two affinity-based probes for PARP1 (ARTD1) that are based on the clinical and FDA approved PARP1 inhibitor olaparib. The photo-affinity based probes were designed by functionalization of olaparib derivatives with diazirine-alkyne linkers and are displayed in Figure 2, **55** and **56**. The probes were easily synthesized by HCTU mediated coupling of readily available olaparib derivatives with diazirine-alkyne containing linkers. The probes were incubated with living Raji cells enabling the non-covalent binding of the recognition element (olaparib 'part') of the probes with PARP1. Subjection of the cells to UV-irradiation activates the diazirine allowing carbene formation which then inserts itself in the peptide backbone, covalently labeling the inhibited protein. After cell lysis and denaturation, the lysate was labelled using CuAAC to introduce the reporter group. An azide containing fluorophore was used to allow in-gel fluorescence analysis whereas ligation with biotin allowed streptavidin affinity enrichment and pull-down of the labelled proteins. Analysis of the isolated proteins using an MS/MS based proteomics procedure showed that especially **56** was able to effectively target PARP1 in living cells. The ability of **56** to covalently bind PARP1 was also corroborated by in-gel fluorescence analysis. In addition, target validation was performed by Western-Blot analysis confirming the observed fluorescent band to be PARP1.

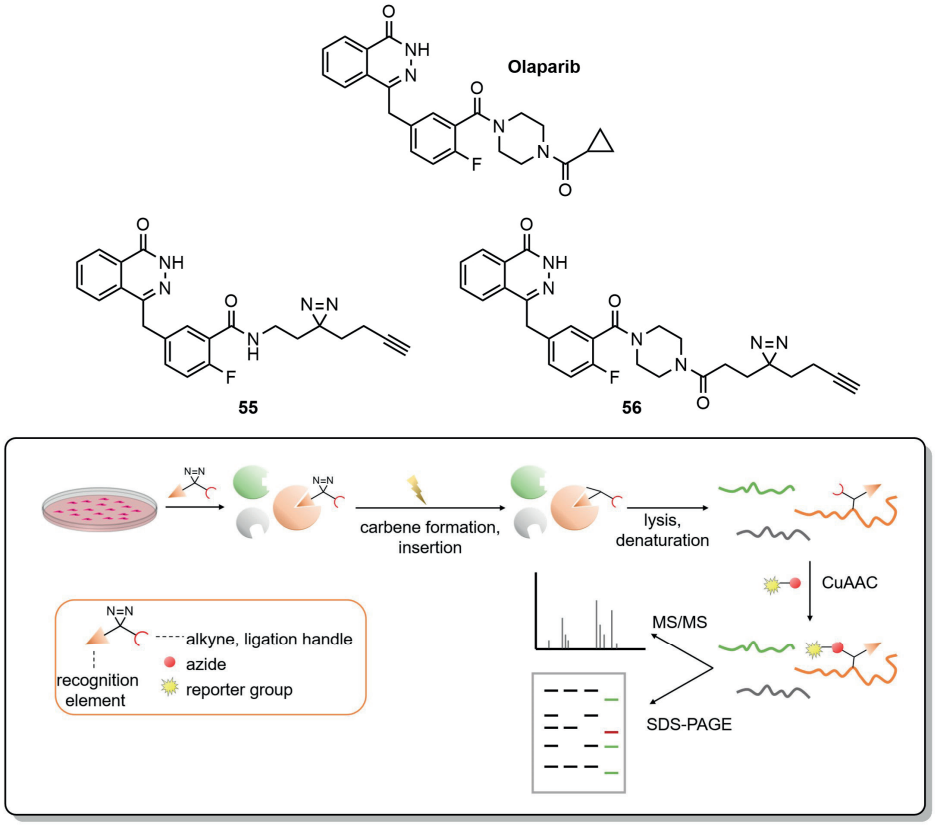


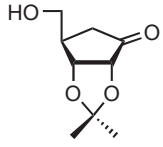
Figure 2. AfBPs **55** and **56** and schematic workflow of the proteomics performed with these probes.

Experimental section

General synthetic procedures

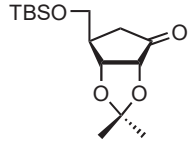
All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. Column chromatography was performed on silica gel 60 Å (40-63 µm, Macherey-Nagel). TLC analysis was performed on Macherey-Nagel aluminum sheets (silica gel 60 F_{254}). TLC was used to visualize compounds by UV at wavelength 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) or $KMnO_4$ spray (20 g/L $KMnO_4$ and 10 g/L K_2CO_3 in water) followed by charring at c.a. 250 °C. LC-MS 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 coupled to a Thermo LCQ Fleet Ion mass spectrometer with ESI. The method used was 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. NMR spectra were recorded on a Bruker AV-400, AV-500 or AV-600 NMR. Chemical shifts (δ) are given in ppm relative to tetramethyl silane as internal standard. Coupling constants (*J*) are given in Hz. All given ^{13}C-APT spectra are proton decoupled.

(2*R*,3*R*,4*R*)-2,3-O-isopropylidene-4-hydroxymethyl-cyclopentanone



A chromatography column was charged with enone **10** (1.23 g, 8.0 mmol) and benzophenone (262 mg, 1.44 mmol, 0.18 eq.) in anhydrous methanol (508 mL, 0.016 M) and purged with nitrogen gas for 1 hour, followed by additional purging with argon for 1 hour. The colorless solution was irradiated at 365 and 366 nm utilizing a Spectroline ENF-260C/FE (6 watts) and CAMAG 4 (14 watts) lamp. After irradiating overnight, the solution became yellow and the methanol was evaporated under reduced pressure to provide a crude amber residue. Purification by silica gel column chromatography (20 -> 60% EtOAc in pentane) furnished the title compound as a colorless oil (900 mg, 4.83 mmol, 60%). **Rf:** 0.22 (40% EtOAc in pentane). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 4.72 (d, 1H, *J* = 5.4 Hz, H-2), 4.29 (dt, 1H, *J* = 5.5, 1.4 Hz, H-3), 3.83 (dd, 1H, *J* = 10.3, 3.0 Hz, H-5), 3.69 (dd, 1H, *J* = 10.3, 3.6 Hz, H-5'), 2.82 (br. s, 1H, 5-OH), 2.74 (ddd, 1H, *J* = 18.4, 9.1, 1.0 Hz, H-6), 2.57 – 2.51 (m, 1H, H-4), 2.17 (dq, 1H, *J* = 18.5, 1.5 Hz, H-6'). **^{13}C NMR** (126 MHz, $CDCl_3$): δ 214.7 (C=O), 111.3 (Cq), 81.4 (C-2), 78.9 (C-3), 64.0 (C-5), 38.9 (C-4), 37.2 (C-6), 26.7, 24.7 (Me).

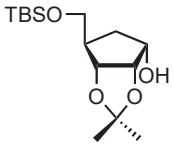
(2*R*,3*R*,4*R*)-2,3-O-isopropylidene-4-[((tert-butyldimethylsilyl)oxy)methyl]cyclopentanone (11)



Enone **10** (325 mg, 1.75 mmol) and imidazole (178 mg, 2.62 mmol, 1.5 eq.) were co-evaporated with anhydrous 1,4-dioxane (5 mL) and dissolved in anhydrous DCM (13.6 mL, 0.13 M) after which tert-butyldimethylsilyl chloride (50 wt. % in toluene, 0.73 mL, 2.1 mmol, 1.2 eq.) was added. The suspension was stirred for 4 hours, quenched by the addition of methanol (1 mL) and diluted with H_2O. The organic layer was separated and the aqueous layer extracted twice with DCM. The resulting organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by flash column chromatography (0 -> 20% EtOAc in pentane) afforded the title compound as a colorless oil (516 mg, 1.72 mmol, 98%). **Rf:** 0.86 (10% EtOAc in pentane). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 4.66 (d, 1H, *J* = 5.4 Hz, H-2), 4.24 (dd, 1H, *J* = 5.4, 1.4 Hz, H-3), 3.84 (dd, 1H, *J* = 9.8, 2.3 Hz, H-5), 3.64 (dd, 1H, *J* = 9.8, 2.8 Hz, H-5'), 2.74 (dd, 1H, *J* = 18.1, 9.1 Hz, H-6), 2.55 – 2.50 (m, 1H, H-4), 2.09 (dq, 1H, *J* = 18.1, 1.4 Hz, H-6'), 1.44 (s,

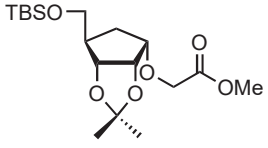
3H, CH_3 Me), 1.36 (s, 3H, CH_3 Me), 0.86 (s, 9H, CH_3 *t*Bu), 0.05 (s, 3H, CH_3 SiMe), 0.03 (s, 3H, CH_3 SiMe). **^{13}C NMR** (126 MHz, $CDCl_3$): δ 212.9 (C=O), 111.1 (Cq), 82.0 (C-2), 79.2 (C-3), 65.4 (C-5), 39.2 (C-4), 37.3 (C-6), 26.9 (Me), 25.9 (*t*Bu), 24.7 (Me), 18.30 (Cq tBu), -5.6, -5.7 (SiMe).

(1*S*,2*S*,3*R*,4*R*)-2,3-*O*-isopropylidene-4-[((tert-butyldimethylsilyl)oxy)methyl]-cyclopentanol (12)



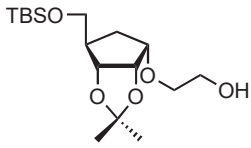
Compound **11** (360 mg, 1.2 mmol) was dissolved in methanol (36 mL, 0.03 M) and cooled to 0 °C after which sodium borohydride (55 mg, 1.45 mmol, 1.2 eq.) was added. After stirring for 15 minutes, the cooling bath was removed and the mixture stirred for another 15 minutes. The reaction was quenched by the addition of acetone (5 mL) and concentrated to dryness. The resulting orange residue was taken up in EtOAc and washed with sat. aq. NaCl. The aqueous layer was extracted twice with EtOAc and the organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by flash column chromatography (10 -> 20% EtOAc in pentane) afforded the title compound as a viscous colorless oil (362 mg, 1.2 mmol, quant.). **Rf:** 0.62 (10% EtOAc in pentane. **^{1}H NMR** (500 MHz, $CDCl_3$): δ 4.51 (dd, 1H, *J* = 5.9, 1.3 Hz, H-3), 4.45 (t, 1H, *J* = 5.7 Hz, H-2), 4.22 (td, 1H, *J* = 7.6, 5.6 Hz, H-1), 3.62 (dd, 1H, *J* = 10.1, 4.5 Hz, H-5), 3.50 (dd, 1H, *J* = 10.1, 4.6 Hz, H-5'), 2.47 (br. s, 1H, 1-OH), 2.23 – 2.17 (m, 1H, H-4), 1.90 – 1.84 (m, 2H, H-6), 1.51 (s, 3H, CH_3 Me), 1.37 (s, 3H, CH_3 Me), 0.90 (s, 9H, CH_3 tBu), 0.05 (s, 6H, CH_3 SiMe). **^{13}C NMR** (101 MHz, $CDCl_3$): δ 111.0 (Cq), 83.1 (C-3), 79.8 (C-2), 72.0 (C-1), 64.7 (C-5), 44.0 (C-4), 35.7 (C-6), 26.3 (Me), 26.0 (*t*Bu), 24.3 (Me), 18.3 (Cq *t*Bu), -5.5 (SiMe).

(1*S*,2*S*,3*R*,4*R*)-1-*O*-(2-methoxy-2-oxoethoxy)-2,3-*O*-isopropylidene-4- [((*tert*butyldimethylsilyl)oxy) methyl]-cyclopentane (13)



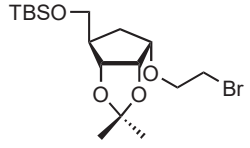
A 50 mL round bottom flask was flame-dried under vacuum, placed under an argon atmosphere and charged with cyclopentanol **12** (1.51 g, 5.0 mmol), which was subsequently co-evaporated thrice with anhydrous 1,4-dioxane. The residue was dissolved in anhydrous THF (25 mL, 0.2 M) after which sodium hydride (60 wt. % in mineral oil, 600 mg, 15 mmol, 3 eq.) and 15-crown-5 (1.2 mL, 6.0 mmol, 1.2 eq.) were added. The solution was stirred for 30 minutes until bubbling seized followed by the addition of TBABr (359 mg, 1.0 mmol, 0.2 eq.) and methyl bromoacetate (1.4 mL, 15 mmol, 3 eq.). The reaction was stirred for 9 days after which it was quenched by the careful addition of sat. aq. NH_4Cl (5 mL), diluted with H_2O and the aqueous layer was extracted thrice with Et_2O. The organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by silica gel column chromatography (5 -> 20% EtOAc in pentane) afforded the title compound as a colorless oil and an inseparable mixture with impurities and starting material (1.8 g, 4.81 mmol, 96%). **R_f:** 0.67 (20% EtOAc in pentane). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 4.56 (t, 1H, *J* = 5.2 Hz, H-2), 4.41 (dd, 1H, *J* = 5.5, 1.2 Hz, H-3), 4.22 - 415 (m, 2H, *J* = 16.6 Hz, CH_2COOMe), 4.07 (ddd, 1H, *J* = 9.8, 6.6, 4.9 Hz, H-1), 3.72 (s, 3H, CH_3 COOMe), 3.57 (dd, 1H, *J* = 10.0, 4.4 Hz, H-5), 3.47 (dd, *J* = 10.0, 4.0 Hz, H-5'), 2.15 – 2.07 (m, 2H, H-4, H-6), 1.83 (dd, 1H, *J* = 10.9, 6.5 Hz, H-6'), 1.48 (s, 3H, CH_3 Me), 1.31 (s, 3H, CH_3 Me), 0.86 (s, 9H, CH_3 *t*Bu), 0.01 (s, 6H, CH_3 SiMe). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.0 (C=O COOMe), 110.9 (Cq), 82.7 (C-3), 80.4 (C-1), 79.1 (C-2), 67.3 (C_s CH_2COOMe), 65.1 (C-5), 51.8 (COOMe), 44.1 (C-4), 31.6 (C-6), 26.5 (Me), 25.9 (*t*Bu), 24.5 (Me), 18.2 (Cq *t*Bu), -5.6 (SiMe). **LC-MS** (10 -> 90% B in A): Rt = 8.71.

(1*S*,2*S*,3*R*,4*R*)-1-*O*-hydroxyethoxy-2,3-*O*-isopropylidene-4-[((*tert*-butyldimethyl silyl)oxy)methyl]-cyclopentane (14)



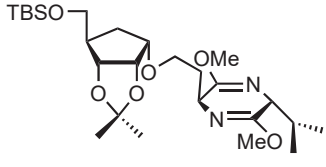
To an ice-cooled solution of crude methyl ester **13 (**1.8 mg,4.81 mmol) in anhydrous Et_2O (48.1 mL, 0.1 M) was added $LiAlH_4$ (456 mg, 12.0 mmol, 2.5 eq.). The suspension was stirred for 2 hours, quenched by the dropwise addition of sat. aq. NH_4Cl (5 mL) and stirred while gradually warming to room temperature. The mixture was diluted with 1 M aq. HCl, extracted six times with Et_2O and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The residue was purified by column chromatography (20 -> 40% EtOAc in pentane) to afford the title compound as a colorless oil (1.03 g, 2.98 mmol, 60% over 2 steps) and recovered starting material **2** (359 mg, 1.19 mmol, 24%). **R_f:** 0.17 (20% EtOAc in pentane). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 4.55 (t, 1H, *J* = 5.3 Hz, H-2), 4.47 (d, 1H, *J* = 5.6 Hz, H-3), 4.01 (ddd, 1H, *J* = 10.4, 6.7, 5.0 Hz, H-1), 3.79 – 3.66 (m, 3H, H-7, H-8 CH_2CH_2OH), 3.63 – 3.57 (m, 2H, H-5, H-7'), 3.49 (dd, 1H, *J* = 10.0, 4.6 Hz, H-5'), 3.01 (br. s, 1H, 8-OH), 2.15 (dt, 1H, *J* = 9.0, 4.7 Hz, H-4), 2.06 (ddd, 1H, *J* = 12.5, 10.4, 8.4 Hz, H-6), 1.83 (dd, 1H, *J* = 12.4, 6.7 Hz, H-6'), 1.50 (s, 3H, CH_3 Me), 1.33 (s, 3H, CH_3 Me), 0.89 (s, 9H, CH_3 *t*Bu), 0.04 (s, 6H, CH_3 SiMe). **^{13}C NMR** (126 MHz, $CDCl_3$): δ 110.9 (Cq), 83.2 (C-3), 80.4 (C-1), 78.8 (C-2), 71.8 (C-7), 65.1 (C-5), 62.0 (C-8), 43.7 (C-4), 32.3 (C-6), 26.4 (Me), 26.0 (*t*Bu), 24.6 (Me), 18.3 (Cq *t*Bu), -5.5 (SiMe).

(1*S*,2*S*,3*R*,4*R*)-1-*O*-bromoethoxy-2,3-*O*-isopropylidene-4-[((*tert*-butyldimethylsilyl)oxy)methyl]-cyclopentane (15)



Ethoxy alcohol **14** (520 mg, 1.5 mmol) was co-evaporated thrice with anhydrous 1,4-dioxane, dissolved in anhydrous DCM (15 mL, 0.1 M) and cooled to 0 °C. To this cooled solution imidazole (204 mg, 3.0 mmol, 2 eq.), PPh_3 (789 mg, 3.0 mmol, 2 eq.) and Br_2 (154 µL, 3.0 mmol 2.0 eq.) were added consecutively. After stirring for 1 hour, TLC indicated complete conversion and the solution was concentrated to dryness. The residue was taken up in Et_2O, washed with sat. aq. $NaHCO_3$ and sat. aq. $Na_2S_2O_3$ and the organic layer was dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by flash column chromatography (0 -> 10% EtOAc in pentane) afforded the title compound as a colorless oil (513 mg, 1.25 mmol, 84%). **R_f:** 0.81 (5% EtOAc in pentane). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 4.55 (t, 1H, *J* = 5.2 Hz, H-2), 4.45 (d, 1H, *J* = 5.5 Hz, H-3), 4.04 (ddd, 1H, *J* = 10.8, 6.6, 4.8 Hz, H-1), 3.92 (dt, 1H, *J* = 10.7, 6.7 Hz, H-7 OCH_2CH_2), 3.79 (dt, 1H, *J* = 10.7, 6.9 Hz, H-7' OCH_2CH_2), 3.61 (dd, 1H, *J* = 10.0, 4.6 Hz, H-5), 3.54 – 3.47 (m, 3H, H-5', H-8 CH_2CH_2Br), 2.20 – 2.06 (m, 2H, H-4, H-6), 1.83 (dd, 1H, *J* = 12.2, 6.5 Hz, H-6), 1.52 (s, 3H, CH_3 Me), 1.35 (s, 3H, CH_3 Me), 0.92 (s, 9H, CH_3 *t*Bu), 0.07 (s, 6H, CH_3 SiMe). **^{13}C NMR** (126 MHz, $CDCl_3$): δ 111.0 (Cq), 82.8 (C-3), 80.6 (C-1), 79.4 (C-2), 70.2 (C-7), 65.2 (C-5), 44.2 (C-4), 31.9 (C-6), 30.3 (C-8), 26.6 (Me), 26.0 (*t*Bu), 24.7 (Me), 18.3 (Cq *t*Bu), -5.4 (SiMe).

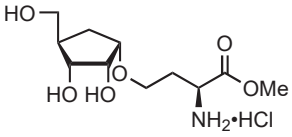
(1*S*,2*S*,3*R*,4*R*)-1-*O*-ethyl-((2*S*,5*R*)-5-isopropyl-3,6-dimethoxy-2,5- dihydro pyrazine)-2,3-*O*-isopropylidene-4-[((*tert*-butyldimethylsilyl)oxy) methyl] cyclopentane (16)



A 5 mL pear shaped flask was flame-dried and subsequently charged with (*R*)-Schöllkopf reagent (112.6 µL, 0.629 mmol, 1.05 eq.) under inert atmosphere. THF (0.5 mL) was added and the solution was cooled to -78 °C after which *n*-BuLi (2.5 M in hexanes, 251.3 µL, 1.05 eq.) was added dropwise over the duration of 3 minutes and the mixture stirred for an additional 30 minutes. A solution of co-evaporated ethoxybromide **15** (245 mg, 0.598 mmol, 1.0 eq.) in THF/HMPA (v/v; 5/2, 0.7 mL, 1.45 mL total, 0.43 M) was added to the yellow aza-enolate solution. Next, Ag_2CO_3 (165 mg, 0.598

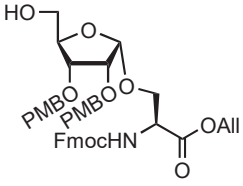
mmol, 1.0 eq.) was added and the mixture was stirred overnight while allowed to gradually warm to room temperature. The reaction was quenched by the addition of sat. aq. NH_4Cl (5 mL) and filtered over a pad of Celite. The filtrate was diluted with sat. aq. NH_4Cl, extracted thrice with Et_2O and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by silica gel column chromatography (2.5 -> 20% EtOAc in pentane) furnished an inseparable diastereomeric mixture of title dihydropyrazine **6** as a yellow oil (201 mg, 0.392 mmol, 66%, dr = 1/2.9). **R_f:** 0.32 (10% EtOAc in pentane). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 4.59 – 4.48 (m, 1H, H-2), 4.48 – 4.37 (m, 1H, 3-H), 4.08 (d, *J* = 7.6 Hz, 1H), 4.07 (dt, 1H, *J* = 7.6, 4.2 Hz, H-1), 3.95 – 3.87 (m, 3H, C^1-C*H*HCH$_2$, 2 x C*H* bislactim), 3.86 (s, 3H, OMe), 3.66 (s, 3H, OMe), 3.66 – 3.64 (m, 1H, C^1-C*H*HCH$_2$), 3.62 – 3.54 (m, 2H, H-5, C^1-CH$_2$C*H*H), 3.51 – 3.46 (m, 2H, H-5, C^1-CH$_2$C*H*H), 2.27 – 2.17 (m, 1H, C*H*(CH$_3$)$_2$ isopropyl), 2.13 – 1.98 (m, 1H, H-4), 1.96 – 1.83 (m, 1H, H-6), 1.78 – 1.73 (m, 1H, H-6), 1.49 (s, 3H, CH$_3$ Me), 1.31 (s, 3H, CH$_3$ Me), 1.04 (d, 1H, *J* = 6.9 Hz, CH$_3$ isopropyl), 0.89 (s, 9H, CH$_3$ *t*Bu), 0.70 (d, 1H, *J* = 6.9 Hz, CH$_3$ isopropyl), 0.05 (s, 6H, CH$_3$ SiMe). **^{13}C NMR** (126 MHz, $CDCl_3$): δ 163.94, 163.85, 163.6, 163.1, 111.1, 110.9, 110.8, 67.4, 66.8, 65.3, 65.23, 65.21, 64.8, 60.9, 36.02, 35.8, 34.4, 31.6, 31.5, 18.3, -5.40. **LC-MS** (10 -> 90% B in A): Rt = 9.44 and 9.83.

(1*S*,2*S*,3*R*,4*R*)-1-*O*-(*R*)-isothreonine-2,3-*O*-isopropylidene-4-[((*tert*-butyldimethyl silyl)oxy)methyl]-cyclopentane (17)



Compound **16** (0.300 mmol, 154 mg, dr = 1/2.9) was dissolved in THF (3 mL) after which a freshly prepared solution of 0.25 M aq. HCl (2.4 mL, 0.6 mmol, 2.0 eq., 5.4 mL total, 0.056 M) was added. The mixture was stirred at room temperature for 3 days, whereafter LC-MS indicated complete hydrolysis and the mixture was subsequently washed with Et_2O. The resulting aqueous layer was concentrated to dryness and co-evaporated thrice with anhydrous toluene (3 x 5 mL) to furnish the title compound as a 1:1 mixture with D-valine methyl ester, an amorphous colorless solid, as the corresponding hydrochloride salts (204 mg, 0.259 mmol, 86%). **^{1}H NMR** (500 MHz, D_2O): δ 4.28 (dt, *J* = 6.9, 3.7 Hz, 1H), 4.06 (td, *J* = 7.0, 3.8 Hz, 1H), 4.03 – 3.96 (m, 2H), 3.84 (dq, *J* = 12.7, 5.7, 5.0 Hz, 2H), 3.80 – 3.73 (m, 9H), 3.73 – 3.65 (m, 2H), 3.65 – 3.53 (m, 3H), 3.52 – 3.45 (m, 2H), 2.33 – 2.12 (m, 6H), 1.94 – 1.77 (m, 2H), 1.67 – 1.50 (m, 2H), 0.99 – 0.96 (m, 6H). **^{13}C NMR** (126 MHz, D_2O): δ 172.1, 171.9, 80.9, 80.8, 80.8, 75.7, 74.8, 74.7, 74.6, 74.2, 74.1, 74.1, 72.5, 67.3, 67.1, 67.0, 65.1, 65.0, 65.0, 60.0, 55.3, 55.3, 55.1, 53.9, 53.5, 46.1, 46.1, 46.0, 33.4, 31.0, 31.0, 30.9, 30.8, 30.6, 19.0, 18.7. **LC-MS** (10 -> 90% B in A): Rt = 0.765.

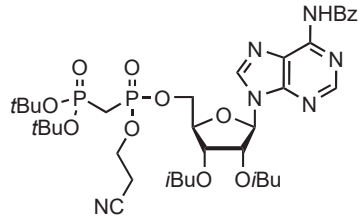
O^{β}-(2,3-di-*O*-*para*-methoxybenzyl-β-D-ribosyl)-N^{α}- fluorenylmethoxycarbonyl serine allyl ester (26)



1-*O*-(2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α-D-ribosyl)-*N*-fluorenylmethoxycarbonyl serine allyl ester (see compound **25** in Chapter 3, 250 mg, 0.255 mmol) was dissolved in anhydrous pyridine (2.56 mL), after which HF:pyr.(70 wt.%, 0.256 mL) was added. After stirring for 5 hours, TLC indicated full consumption of the starting material and the mixture was quenched by the addition of sat. aq. $NaHCO_3$ and the aqueous phase was extracted thrice with DCM. The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. Purification by column chromatography (4 -> 10% acetone in DCM) afforded the title compound as an off-white wax (180 mg, 0.243 mmol, 95%). **R_f:** 0.26 (5% acetone in DCM). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 7.73 (dd, 2H, *J* = 7.5, 2.0 Hz, Fmoc arom.), 7.57 (dd, 2H, *J* = 7.6, 4.1 Hz, Fmoc arom.), 7.36 (td, 2H, *J* = 7.5, 3.5 Hz, Fmoc arom.), 7.30 – 7.21 (m, 6H, Fmoc arom. + PMB arom.), 6.89 – 6.78 (m, 4H, PMB arom), 6.45 (d, 1H, *J* = 9.2 Hz, NH), 5.87 (ddq, 1H, *J* = 17.2, 12.0, 6.0 Hz, C*H*CH$_2$ allyl), 5.30 (dd, 1H, *J* = 17.2, 1.5 Hz, CHC*H*$_{2a}$ allyl), 5.18 (dd, 1H, *J* = 10.5, 1.3 Hz, CHC*H*$_{2b}$ allyl), 4.93 (d, 1H, *J* = 4.0 Hz, H-1), 4.69

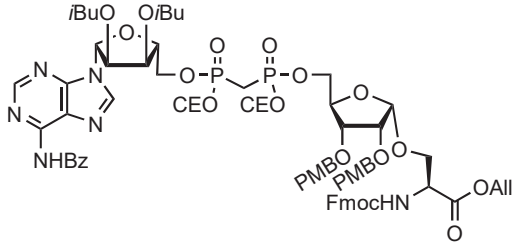
– 4.37 (m, 8H, CH_2CH allylic +2x CH_2 PMB + CH Ser + CH_{2a} Fmoc), 4.26 (dd, 1H, *J* = 10.5, 7.6 Hz, CH_{2b} Fmoc), 4.19 – 4.12 (m, 3H, H-4, CH, Fmoc), 3.92 (dd, 1H, *J* = 10.5, 3.2 Hz, CH_2), 3.87 (dd, 1H, *J* = 6.4, 3.5 Hz, H-3), 3.78 (s, 3H, OCH_3 PMB), 3.77 – 3.74 (m, 2H, H-2, 5-OH),3.73 (s, 3H, OCH_3 PMB), 3.66 (d, 1H, *J* = 11.6 Hz, H-5), 3.46 (d, 1H, *J* = 12.3 Hz, H-5'). **^{13}C NMR** (126 MHz, $CDCl_3$): δ 170.3 (C=O allyl ester), 159.5, 159.4 (Cq PMB), 156.5 (C=O carbamate), 144.2, 143.9, 141.4, 141.3 (Fmoc), 131.8 (*C*HCH_2 allyl), 130.3, 129.9, 129.8, 129.6 (PMB), 127.8, 127.7, 127.2, 127.1, 125.5, 125.3, 120.02, 120.00 (Fmoc), 118.6 (CH*C*H_2 allyl), 113.9, 113.8 (PMB), 101.6 (C-1), 83.6 (C-4), 78.4 (C-2), 75.0 (C-3), 72.5, 72.3 (*C*H_2 PMB), 68.1 (*C*H_2), 67.1 (*C*H_2 Fmoc), 66.2 (*C*H_2CH allylic), 62.8 (C-5), 55.4, 55.3 (OMe PMB), 54.6 (CH Ser), 47.2 (*C*H Fmoc).

N^6-benzoyl-2',3'-di-*O*-isobutyryl-(*R*$_P$/*S*$_P$)-5'-*O*-di-*tert*-butyl(((2-cyanoethoxy)phosphoryl) methyl) phosphonate adenosine (25)



A 25 mL round bottom flask was flame-dried under vacuum, placed under an argon atmosphere, charged with adenosine **24** (1.74 g, 3.41 mmol, 1.1 eq.) and phosphanyl methylphosphonate **23**[22] (1.27 g, 3.10 mmol) co-evaporated with anhydrous acetonitrile and the residue was dissolved in anhydrous acetonitrile (15.5 mL, 0.2 M). DCI (549 mg, 4.65 mmol, 1.5 eq.) was added and stirring commenced until ^{31}P NMR indicated complete phosphonitylation (δ 172.92) after 30 minutes. *t*BuOOH (5.5 M in nonane, 1.13 mL, 6.19 mmol, 2.0 eq.) was added and stirring commenced for an additional 15 minutes, whereafter ^{31}P NMR showed complete conversion to the phosphonate (δ 20.88). The mixture was diluted with EtOAc and the organic layer was washed with H_2O, sat. aq. $NaHCO_3$ and sat. aq. NaCl respectively. The organic layer was dried over Na_2SO_4, filtered and concentrated *in vacuo*. Purification by silica gel column chromatography (20 -> 40% acetone in DCM) afforded the title compound as a mixture of R_P/S_P diastereomers, as a crystalline white foam (2.45 g, 2.93 mmol, 95%). **R$_f$:** 0.25 (30% acetone in DCM). **^{1}H NMR** (500 MHz, $CDCl_3$): δ 9.24 (s, 2H, 6-NH), 8.81 (s, 1H, H-2), 8.80 (s, 1H, H-2), 8.50 (s, 1H, H-8), 8.41 (s, 1H, H-8), 8.06 – 7.99 (m, 4H, *o*-Bz), 7.93, 7.66 – 7.59 (m, 2H, *p*-Bz), 7.56 – 7.50 (m, 4H, *m*-Bz), 6.33 (d, 1H, *J* = 6.0 Hz, H-1'), 6.29 (d, 1H, *J* = 5.7 Hz, H-1'), 5.95 (t, 1H, *J* = 5.7 Hz, H-2'), 5.88 (t, 1H, *J* = 5.9 Hz, H-2'), 5.73 (dd, 1H, *J* = 5.7, 3.9 Hz, H-3'), 5.69 (dd, 1H, *J* = 5.7, 3.9 Hz, H-3'), 4.56 – 4.43 (m, 6H, H-4', H-5'), 4.40 – 4.26 (m, 4H, OCH_2CH_2CN), 2.85 – 2.69 (m, 4H, OCH_2CH_2CN), 2.69 – 2.61 (m, 2H, CH(Me)$_2$ *i*Bu), 2.59 – 2.44 (m, 6H, P-CH*H*-P, CH(Me)$_2$ *i*Bu), 1.55 – 1.49 (m, 36H, CH_3 ^{t}Bu), 1.25 – 1.21 (m, 12H, CH_3), 1.17 – 1.07 (m, 12H. **^{13}C-APT NMR** (126 MHz, $CDCl_3$): δ 175.9, 175.8, 175.52, 175.51 (C=O *i*Bu), 165.04, 165.03 (C=O Bz), 151.92 (C-2), 151.86 (C-6), 149.72 (C-4), 149.70 (C-8), 133.42 (*ipso*-Bz), 133.09 (*p*-Bz), 129.02 (*m*-Bz), 127.99 (*p*-Bz), 123.47, 123.36 (C-5), 116.92, 116.82 (Cq CH_2CH_2*C*N), 86.4, 86.0 (C-1), 84.61, 84.58, 84.54, 84.51, 84.47, 84.42 (Cq *t*Bu), 81.93, 81.88, 81.7, 81.6 (C-4'), 73.2, 73.1 (C-2'), 70.42, 70.35 (C-3'), 65.50, 65.45, 65.18, 65.13 (C-5'), 61.34, 61.30 (*C*H_2CH$_2$CN), 33.86, 33.84, 33.70, 33.68 (CH *i*Bu), 30.71, 30.41 (P-*C*-P), 30.39, 30.36, 30.33 (*t*Bu), 29.59, 29.42, 28.51, 28.30 (P-*C*-P), 19.96, 19.95, 19.91, 19.89 (CH$_2$*C*H$_2$CN), 18.97, 18.87, 18.84, 18.72 (CH_3 *i*Bu). **^{31}P NMR** (202 MHz, $CDCl_3$): δ 23.32 (d, *J* = 8.1 Hz, CH_2P(O)OCNE), 22.98 (d, *J* = 8.1 Hz, CH_2P(O)OCNE), 9.16 (d, *J* = 8.2 Hz, (*t*BuO)$_2$P(O)CH$_2$), 9.06 (d, *J* = 7.9 Hz, (*t*BuO)$_2$P(O)CH$_2$). **LC-MS** (30 -> 70% B in A): Rt = 7.25.

(R_p/S_p)-Ser-ABP ribosylated derivative (27)



To a solution of (di-*tert*-butyl)methylene bisphosphonate adenosine **25** (500 mg, 0.600 mmol) in hexafluoro isopropanol (1.43 mL, 0.42 M) was added a 6 M HCl 1,4-dioxane:H_2O (1:1 ; v/v, 116.4 µL, 0.7 mmol, 1.2 eq.). Reaction progress was gauged with ^{31}P NMR, which indicated complete acidolysis of the di-*tert*-butyl phosphonate moiety into the corresponding deprotected phosphonate (δ 16.37) after 4 hours. The mixture was subsequently diluted with 1,4-dioxane:toluene (1:1; v/v, 20 mL) and concentrated *in vacuo*. The co-evaporation process was repeated thrice to remove all traces of hexafluoro isopropanol and furnished deprotected methylene bisphosphonate adenosine as a crystalline white solid (433 mg, 0.599 mmol, quant.), which was used without further purification.

The resulting bisphosphonate adenosine was aliquoted and 73 mg (0.1 mmol) was co-evaporated thrice with anhydrous acetonitrile and dissolved in anhydrous acetonitrile (1 mL, 0.1 M). 2-Cyanoethanol (7.2 µL, 0.105 mmol, 1.05 eq.), anhydrous TEA (27.9 µL,0.2 mmol, 2.0 eq.) and anhydrous 2,6-lutidine (116 µL, 1.0 mmol, 10 eq.) and PyNTP (150 mg, 0.3 mmol, 3.0 eq.) were added successively and the resulting solution was stirred for 5 hours. Subsequently, ribosylated serine **26** (78 mg, 0.105 mmol, 1.05 eq.), which was co-evaporated thrice in anhydrous acetonitrile prior to use, as a solution in anhydrous acetonitrile (1 mL) and PyNTP (150 mg, 0.3 mmol, 3.0 eq.) were added. After stirring overnight, the mixture was quenched by the addition of 1 M NaOAc in MeOH (1 mL), stirred for 30 minutes and subsequently diluted with DCM (20 mL). The organic layer was washed with 1 M aq. HCl, dried over Na_2SO_4, filtered and concentrated *in vacuo*. Purification by column chromatography (0 -> 5% acetone in DCM) furnished crude title ADPr derivative **27** as a colorless residue. A second purification by column chromatography (0 -> 5% acetone in DCM) to remove residual traces of hexafluorophosphate afforded the title compound as a colorless wax (49 mg, 0.033 mmol, 33%). **^{31}P NMR** (202 MHz, $CDCl_3$): δ 26.97, 25.51, 25.15, 25.00. **LC-MS** (10 -> 90% B in A): Rt = 9.22 and 9.38.

References

[1] G. J. van der Heden van Noort, M. G. van der Horst, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *J. Am. Chem. Soc.* **2010**, *132*, 5236–5240.

[2] H. A. V Kistemaker, A. P. Nardozza, H. S. Overkleeft, G. A. van der Marel, A. G. Ladurner, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2016**, *55*, 10634–10638.

[3] Q. Liu, H. A. V. Kistemaker, S. Bhogaraju, I. Dikic, H. S. Overkleeft, G. A. van der Marel, H. Ovaa, G. J. van der Heden van Noort, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2018**, *57*, 1659–1662.

[4] J. Voorneveld, J. G. M. Rack, I. Ahel, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Org. Lett.* **2018**, *20*, 4140–4143.

[5] O. Leidecker, J. J. Bonfiglio, T. Colby, Q. Zhang, I. Atanassov, R. Zaja, L. Palazzo, A. Stockum, I. Ahel, I. Matic, *Nat. Chem. Biol.* **2016**, *12*, 998–1000.

[6] L. Palazzo, O. Leidecker, E. Prokhorova, H. Dauben, I. Matic, I. Ahel, *Elife* **2018**, *7*, 1–12.

[7] I. Gibbs-Seymour, P. Fontana, J. G. M. Rack, I. Ahel, *Mol. Cell* **2016**, *62*, 432–442.

[8] J. J. Bonfiglio, P. Fontana, Q. Zhang, T. Colby, I. Gibbs-Seymour, I. Atanassov, E. Bartlett, R. Zaja, I. Ahel, I. Matic, *Mol. Cell* **2017**, *0*, 1–9.

[9] P. Fontana, J. J. Bonfiglio, L. Palazzo, E. Bartlett, I. Matic, I. Ahel, *Elife* **2017**, *6*, 1–20.

[10] J. Moss, N. J. Oppenheimer, R. E. West, S. J. Stanley, *Biochemistry* **1986**, *25*, 5408–5414.

[11] Y. Zhang, J. Wang, M. Ding, Y. Yu, *Nat. Methods* **2013**, *10*, 981–984.

[12] T. Meyer, H. Hilz, *Eur. J. Biochem.* **1986**, *155*, 157–165.

[13] H. Osago, M. Terashima, N. Hara, K. Yamada, M. Tsuchiya, *J. Biochem. Biophys. Methods* **2008**, *70*, 1014–1019.

[14] K. D. Siegmund, F. Klink, *FEBS Lett.* **1992**, *312*, 139–142.

[15] B. Eide, P. Gierschik, A. Spiegel, *Biochemistry* **1986**, *25*, 6711–6715.

[16] S. Mashhood Ali, K. Ramesh, R. T. Borchardt, *Tetrahedron Lett.* **1990**, *31*, 1509–1512.

[17] H. Ovaa, B. Lastdrager, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, J. H. van Boom, *J. Chem. Soc. Perkin Trans. 1* **2002**, 2370–2377.

[18] T. C. Myers, K. Nakamura, J. W. Flesher, *J. Am. Chem. Soc.* **1963**, *85*, 3292–3295.

[19] L. M. Yates, D. Fiedler, *ACS Chem. Biol.* **2016**, *11*, 1066–1073.

[20] V. S. Borodkin, M. A. J. Ferguson, A. V. Nikolaev, *Tetrahedron Lett.* **2004**, *45*, 857–862.

[21] V. E. Marquez, C. K. H. Tseng, G. Gebeyehu, D. A. Cooney, G. S. Ahluwalia, J. A. Kelley, M. Dalai, R. W. Fuller, Y. A. Wilson, D. G. Johns, *J. Med. Chem.* **1986**, *29*, 1726–1731.

[22] S. B. Engelsma, N. J. Meeuwenoord, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2017**, *56*, 2955–2959.

[23] A. G. Volbeda, H. A. V. Kistemaker, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, J. D. C. Codée, *J. Org. Chem.* **2015**, *80*, 8796–8806.

[24] H. A. V Kistemaker, L. N. Lameijer, N. J. Meeuwenoord, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, *Angew. Chemie - Int. Ed.* **2015**, *54*, 4915–4918.

[25] Q. Liu, Ridderprint B.V., Leiden, **2019**, pp. 62–92.

[26] E. Matta, A. Kiribayeva, B. Khassenov, B. T. Matkarimov, A. A. Ishchenko, *Sci. Rep.* **2020**, *10*, 1–11.

[27] D. Munnur, I. Ahel, *FEBS J.* **2017**, *284*, 4002–4016.

[28] N. Li, H. S. Overkleeft, B. I. Florea, *Curr. Opin. Chem. Biol.* **2012**, *16*, 227–233.

[29] M. D. Witte, M. T. C. Walvoort, K. Li, W. W. Kallemeijn, W. E. Donker-Koopman, R. G. Boot, J. M. F. G. Aerts, J. D. C. Codée, G. A. van der Marel, H. S. Overkleeft, *ChemBioChem* **2011**, *12*, 1263–1269.

[30] K. Y. Li, J. Jiang, M. D. Witte, W. W. Kallemeijn, H. van den Elst, C. S. Wong, S. D. Chander, S. Hoogendoorn, T. J. M. Beenakker, J. D. C. Codée, J. M. F. G. Aerts, G. A. van der Marel, H. S. Overkleeft, *European J. Org. Chem.* **2014**, *2014*, 6030–6043.

[31] W. W. Kallemeijn, K.-Y. Li, M. D. Witte, A. R. A. Marques, J. Aten, S. Scheij, J. Jiang, L. I. Willems, T. M. Voorn-Brouwer, C. P. A. A. van Roomen, R. Ottenhoff, R. G. Boot, H. van den Elst, M. T. C. Walvoort, B. I. Florea, J. D. C. Codée, G. A. van der Marel, J. M. F. G. Aerts, H. S. Overkleeft, *Angew. Chemie Int. Ed.* **2012**, *51*, 12529–12533.

Appendix

Samenvatting

Samenvatting

ADP-ribosylering is een post-translationele modificatie van voornamelijk eiwitten waarbij NAD^+ wordt verbruikt om ADPr over te dragen op de zijketen van een specifiek aminozuur in het betreffende eiwit. Deze modificatie, die in de jaren 60 is ontdekt, wordt gekatalyseerd door een familie enzymen genaamd ARTs. De eerste chemische karakterisering van het gevormde product dateert uit 1974 waarbij arginine werd beschreven als substraat van ADP-ribosylering. Sinds die tijd is er veel onderzoek verricht naar de functie van deze modificatie, hetgeen resulteerde in belangrijke inzichten over de rol die ADP-ribosylering speelt met betrekking tot het behoud van de integriteit van het genoom, de regulering van de transcriptie, het verouderingsproces en cellulaire apoptose. In 2016 is ontdekt dat ook serine fungeert als acceptor van ADPr en sindsdien is duidelijk geworden dat, na het optreden van DNA-schade, de zijketen van serine het meest voorkomende en wellicht ook meest belangrijke doel van ADP-ribosylering is. Dit proefschrift is gericht op de ontwikkeling van een methodologie voor de organische synthese van ADP-geribosyleerde moleculen. Dergelijke moleculen vormen een zeer nuttig onderdeel in het arsenaal van technieken om ADP-ribosylering te onderzoeken, temeer daar biologisch gewonnen materiaal niet homogeen kan worden verkregen.

In hoofdstuk 2 wordt een methode, eerder gebruikt voor de vaste drager synthese van asparagine-, glutamine- en citrulline-ADP-geribosyleerde peptiden, toegepast om serine-ADP-geribosyleerde peptiden te synthetiseren. Kenmerkend voor deze methode is het gebruik van een beschermd aminozuur dat voorzien is van een gefosforyleerd ribose welke in oplossing wordt gesynthetiseerd. De pyrofosfaat functie wordt tijdens de vaste-drager synthese geïntroduceerd met behulp van een adenosine fosforamidiet. De gesynthetiseerde serine-ADP-geribosyleerde peptiden hebben zowel een α- als een β-anomere configuratie. Daardoor kon met een serine-ADPr hydrolase assay worden aangetoond dat ARTD1 (de meest onderzochte ART en beter bekend als PARP1) serine op een α-selectieve wijze ADP-ribosyleert. Dit komt overeen met het gegeven dat ARTD1 andere substraten zoals arginine, asparaginezuur en glutaminezuur, ook op een α-selectieve manier modificeert. Daarnaast is de chemische stabiliteit van de serine-ADPr verbinding onderzocht door het gebruik van condities die veelal worden gebruikt in biologische experimenten. Hieruit bleek dat behandeling van het peptide met een base zoals NaOH, het ingebouwde serine-ADPr β-eliminatie ondergaat waarbij dehydroalanine overblijft op de plek van modificatie.

Hoofdstuk 3 is gewijd aan het ontwikkelen van een nieuwe synthetische methode voor ADP-geribosyleerde peptiden. Deze methode verschilt met die van hoofdstuk 2 door het gebruik van geribosyleerde aminozuur bouwstenen, waardoor zowel de fosforylering van het ribose en de introductie van het pyrofosfaat aan de vaste drager plaats dient te vinden. Zo wordt de synthese in oplossing van deze bouwstenen vergemakkelijkt door vermindering van

het aantal benodigde en arbeidsintensieve reacties. Ook werd het beschermende groep patroon zoveel mogelijk aangepast aan de gangbare vaste-drager synthese van peptiden waardoor het gebruik van zuur-labiele beschermgroepen zoals trityl en 4-methyltrityl mogelijk werd. Met deze nieuwe methode zijn peptiden ADP-geribosyleerd op serine, threonine en cysteine gesynthetiseerd en vervolgens gebruikt voor het verkrijgen van kristalstructuren van ARH3. Ook is met dit materiaal de eerste cysteïne-ADPr hydrolase geïdentificeerd welke is gevonden in *Streptococcus pyogenes*.

De methode die in hoofdstuk 3 wordt behandeld is ook toegepast voor de synthese van peptiden ADP-geribosyleerd op tyrosine. In hoofdstuk 4 wordt besproken dat de glycoside binding tussen tyrosine en ribose onvoldoende stabiel bleek te zijn tijdens zure omstandigheden die bij deze methode gebruikt worden tijdens de laatste verwijdering van de beschermende groepen. Door de zuurgevoelige groepen in het adenosine fosforamidite te vervangen door basegevoelige groepen bleek het mogelijk om de tyrosine-ADPr modificatie in peptiden te introduceren. De gesynthetiseerde peptiden, ADP-geribosyleerd op tyrosine, zijn getest als substraat voor verschillende hydrolasen waaruit bleek dat tyrosine-ADPr gehydrolyseerd wordt door PARG. Dit enzym stond bekend als een hydrolase dat ADPr polymeren afbreekt tot de overeenkomstige mono-ADPr eiwitten. Deze nieuwgevonden activiteit van PARG ten opzichte van tyrosine-ADPr kan verstrekkende gevolgen hebben voor de juiste interpretatie van veel proteomics experimenten waar een hydrolyse stap van de lysaten met PARG, een onderdeel is van het protocol.

In hoofdstuk 5 wordt de synthese van peptiden ADP-geribosyleerd op arginine behandeld. De ontwikkelde methode onderscheidt zich van die beschreven in de hoofdstukken 2, 3 en 4, door het ontbreken van een geglycosyleerde arginine bouwsteen in de vaste drager synthese. In plaats daarvan werd orthogonaal beschermd ornithine ingebouwd, waarna aan de vaste drager de reactie van het amine van ornithine met een ribosyl isocyanaat bouwsteen een peptide met een geribosyleerd arginine residu opleverde. Echter het beschermende groep patroon in de bouwstenen en bijbehorende condities bleek van grote invloed op het succes van deze methode. Omdat het duidelijk is geworden dat elk aminozuur een eigen reactiviteit en stabiliteit heeft, zijn de methodologieën van zowel hoofdstuk 3 alsmede hoofdstuk 4 geprobeerd. Hieruit bleek dat een behandeling met een nucleofiele base, nodig voor de hybride methodologie van hoofdstuk 4, niet verenigbaar is met de aanwezigheid van ribosylguanidine functie. De methode ontwikkeld in hoofdstuk 3 bleek wel toepasbaar te zijn om peptiden te synthetiseren met de arginine-ADPr modificatie. Uiteindelijk werden verschillende peptiden met een arginine-ADPr modificatie als ook ubiquitine met een ADPr-modificatie op Arg42 gesynthetiseerd. Daarmee zijn niet alleen voor de eerste keer peptiden ADP-geribosyleerd op arginine toegankelijk gemaakt maar ook het eerste, volledig synthetische eiwit gemodificeerd met ADPr op arginine.

De hierboven genoemde hoofdstukken zijn allemaal gecentreerd om aminozuren als substraat voor ADPr. Daar nucleïnezuren ook ADP-geribosyleerd kunnen worden is hoofdstuk 6 gericht op de ontwikkeling van een vaste-drager methode om ADP-geribosyleerd DNA en RNA te synthetiseren. Hiervoor is een nieuwe, bifunctionele ribosyl bouwsteen ontworpen en het is vervolgens in oplossing getest of deze bouwsteen geschikt is voor het synthetiseren van RNA-ADPr. Deze ribosyl bouwsteen is daarna toegepast voor een synthese op de vaste drager en na optimalisatie is het eerste ADP-geribosyleerde nucleotide gesynthetiseerd.

Hoofdstukken 7 en 8 richten zich op het ontwikkelen van moleculen of probes, die het mogelijk maken om de activiteit van enzymen, betrokken bij ADP-ribosylering in kaart te brengen. Hoofdstuk 7 beschrijft het ontwerp en de synthese van twee zogenoemde “activity based probes” (ABPs) gericht op het enzym CD38. Deze ABPs zijn gebaseerd op NAD^+ en zodanig ontworpen om na interactie een covalente binding met een functionele groep in de actieve site van CD38 te vormen. Na vorming van deze binding zijn er verschillende mogelijkheden om de activiteit van het enzym te bepalen waaronder gelelektroforese en massa-spectrometrische analyse.

In hoofdstuk 8 wordt een foto-affiniteit techniek toegepast om de activiteit van ARTD1 (PARP1) in levende cellen te onderzoeken. Hiervoor werden twee foto-affiniteit ABPs ontworpen op basis van het medicijn olaparib, dat een specifieke ‘small molecule inhibitor’ voor ARTD1 is. De structuur van olaparib werd op twee verschillende wijzen voorzien van een zijketen met een diazirine als foto-reactieve functionaliteit en een alkyne als een handvat voor labeling. Na een voorspoedige synthese van deze ABPs werd hun effectiviteit getest door incubatie van levende cellen, behandeling met UV-licht, lysis van de cellen en analyse van het lysaat met SDS-PAGE analyse gekoppeld aan western blots alsmede tandem massa spectrometrie. Beide analyse-methoden wezen uit dat de probes in staat waren om ARTD1 op een significante en selectieve wijze zichtbaar te maken en hebben de potentie te worden gebruikt om de activiteit van ARTD1 in cellen te onderzoeken.

Appendix

List of publications

List of publications

Arginine ADP-ribosylation: Chemical Synthesis of Post-Translationally Modified Ubiquitin Proteins
J. Voorneveld, M.S. Kloet, R.Q. Kim, A. Moutsiopoulou, M. Misra, I. Dikic, D.V. Filippov, G.J. van der Heden van Noort
Manuscript submitted

Molecular Tools for the Study of ADP-ribosylation: A Unified and Versatile Method to Synthesise Native Mono-ADP-ribosylated Peptides
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
Chem. Eur. J. **2021**, *27*, 10621-10627

Olaparib-Based Photoaffinity Probes for PARP-1 Detection in Living Cells
J. Voorneveld, B.I. Florea, T. Bakkum, R.J. Mendowicz, M.S. van der Veer, B. Gagestein, S.I. van Kasteren, M. van der Stelt, H.S. Overkleeft, D.V. Filippov
ChemBioChem. **2020**. *21*, 2431-2434

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

Systematic analysis of ADP-ribose detection reagents and optimisation of sample preparation to detect ADP-ribosylation in vitro and in cells
L. Weixler, J. Voorneveld, G. Aydin, T.M.H.R. Bolte, J. Momoh, M. Bütepage, A. Golzmann, B. Lüscher, D.V. Filippov, R. Žaja, K.L.H. Feijs
Manuscript submitted

Mechanistic insights into the three steps of poly (ADP-ribosylation) reversal
J.G.M. Rack, Q. Liu, V. Zorzini, J. Voorneveld, A. Ariza, K.H. Ebrahimi, J.M. Reber, S.C. Krassnig, D. Ahel, G.A. van der Marel, A. Mangerich, J.S.O. McCullagh, D.V. Filippov, I. Ahel
Nat. Commun. **2021**. *12*, 1-14

Chemical synthesis of linear ADP-ribose oligomers up to pentamer and their binding to the oncogenic helicase ALC1
Q. Liu, G. Knobloch, J. Voorneveld, N.J. Meeuwenoord, H.S. Overkleeft, G.A. van der Marel, A.G. Ladurner, D.V. Filippov
Chem. Sci. **2021**. *12*, 12468-12475.

Synthesis of orthogonally protected and functionalized bacillosamines
J. van Mechelen, J. Voorneveld, H.S. Overkleeft, D.V. Filippov, G.A. van der Marel, J.D.C. Codée
Org. Biomol. Chem. **2020**. *18*, 2834-2837

ELTA: enzymatic labeling of terminal ADP-ribose
Y. Ando, E. Elkayam, R. L. McPherson, M. Dasovich, S. Cheng, J. Voorneveld, D. V. Filippov, S. Ong, L. Joshua-Tor, A.K.L. Leung
Mol.Cell. **2019**, *73*, 845-856

ADPr-peptide synthesis
H.A.V. Kistemaker, J. Voorneveld, D.V. Filippov
ADP-ribosylation and NAD+ Utilizing Enzymes, *Humana Press, New York NY*, **2018**. 345-369.

Appendix

Curriculum Vitae

Curriculum Vitae

Jim Voorneveld was born on December 3th, 1991 in Haarlem. From 2004 until 2009 he attended secondary school at Coornhert Lyceum, located in Haarlem. He received his HAVO diploma in 2009 after which he started at the University of Applied Sciences in Leiden where he obtained is B.A.Sc. degree, majoring in organic chemistry. During his final year, he conducted a research internship in the bio-organic synthesis group of prof.dr. H.S. Overkleeft under the supervision of prof.dr. G.A. van der Marel and dr. D.V. Filippov where he worked on the development of an activity based probe for CD38. After he received his bachelor's degree, he continued his studies at Leiden University where he majored in 'Research in Chemistry', specifically in the field of chemical biology. During his studies, a research internship in the bio-organic synthesis group of prof.dr. H.S. Overkleeft was performed under the supervision of prof.dr. J.D.C. Codée and prof.dr. G.A. van der Marel. His research focused on the synthesis of rare bacterial carbohydrates which are found on pathogenic bacterial strains such as *Neisseria gonorrhoeae* and *Neisseria meningitidis* with the goal to elucidate the mechanism of immune evasion in which these sugars play a role. His efforts resulted in a master thesis entitled 'The synthesis of orthogonally protected bacillosamines' after which he obtained his M.Sc. in September 2016.

In November 2016, he started his PhD studies in the bio-organic synthesis group under supervision of prof.dr. G.A. van der Marel and dr. D.V. Filippov. The results of his studies were orally presented at national and international conferences (Cold Spring Harbor, NY, USA 2018, CHAINS 2018) and in form of poster presentation (Cold Spring Harbor, NY, USA 2018, Eurocarb 2019 and Reedijk Symposium (2019). From May 2021 until January 2022 he worked as Formulation Scientist at SeraNovo B.V. in Leiden and is currently employed as Pharmaceutical Specialist at MSD in Haarlem.

Appendix

Dankwoord

Dankwoord

Hora est! De woorden waar elke promovendus zo hard naartoe streeft en ik ben hier geen uitzondering op geweest. Het was een intensieve periode en zonder twijfel de grootste uitdaging die ik ooit ben aangegaan. Dergelijke beproevingen kunnen niet gepaard gaan zonder de ondersteuning van veel mensen die ik hier ook graag wil benoemen.

Dima, ik wil je hartelijk danken voor het vertrouwen dat je mij vanaf de tijd dat ik student was hebt gegeven. Je hebt in mij geloofd en altijd achter mij gestaan. Dat is wat mij in staat heeft gesteld om dit proefschrift te kunnen schrijven en werk af te leveren waar ik met trots op kan terugkijken. Je deur stond altijd open voor hulp of een goed gesprek waarna je me met goede inzichten weer terug hebt laten keren naar het lab. Ook Gijs heeft hierbij een significante bijdrage kunnen leveren. Ik keek altijd uit naar een werkbespreking met zijn drieën en tijdens onze gesprekken hebben jullie mij altijd weten te motiveren en enthousiasmeren om er weer met volle moed tegenaan te gaan. Jullie waren nooit te druk voor hulp en kwamen zelfs op het lab buurten om te kijken hoe het ging, ook al werd de situatie die jullie daar aan troffen vaker wel dan niet bestempeld als 'the end of science'... Dima en Gijs, jullie zijn een bijzondere combinatie om te hebben als promotoren en ik zou het absoluut niet anders hebben gewild!

Bobby! Jij gekke en waanzinnige vent. Je hebt mijn tijd in het biolab een stuk leuker gemaakt dan ik mij als organisch chemicus ooit heb kunnen voorstellen. Zonder jouw hulp en doortastendheid zou ons werk dan ook nooit tot een mooie publicatie zijn gekomen en jouw bijdrage daarbij is dan ook van onmetelijk belang geweest.

Ook Hermen wil ik graag bedanken voor de kans die ik heb gekregen om in de Biosyn groep mijn promotie te doen. De maandagochtend meetings waren daarbij altijd een nuttige plek om inzichten te verwerven die ik anders had moeten missen en daar wil ik je voor bedanken. Ook binnen de Biosyn groep zijn er veel mensen die ik zou willen bedanken. Te veel zelfs om hier op te noemen en dat zal ik ook niet trachten te doen. In het bijzonder wil ik wel de mensen van DM.3.09 en, vooruit, ook de mensen van DM.3.17 bedanken voor de leuke en (soms misschien iets te) informele werksfeer. Ik heb een geweldige tijd met jullie gehad waar ik met veel plezier op terug kijk.

Tijdens mijn onderzoek heb ik ook de hulp gehad van een aantal studenten die ik hier apart wil benoemen. Rafal, you've done a tremendous job during your internship in Leiden. It was demanding but you stepped up and I'm proud that our names are side by side on our publication. Miriam, ook jij hebt hier natuurlijk hard aan gewerkt. Het was niet makkelijk maar je hebt laten zien wat je waard bent! Alex, je hebt een mooie hoeveelheid aan

materiaal weten te vervaardigen tijdens jouw stage. Hier is dankbaar gebruik van gemaakt om een kristalstructuur te verkrijgen van ARH3 en nu kun je jouw praktische kennis over nucleotide synthese mooi voortzetten bij ProQr. Het ga je goed! Marnix, jouw werk aan de altijd lastige functionele groep genaamd aziridines is van grote waarde geweest. Ik ben er trots op jou te hebben kunnen begeleiden in je bijzondere loopbaan en wil je bedanken voor alles dat je mij hebt geleerd, onder andere op het gebied van onze gezamelijke interesse in Warhammer! Luke, ook jouw naam staat verdiend te pronken op een publicatie. Je hebt de verwachtingen van een bachelor student ruim weten te overstijgen waarbij je hebt laten zien zelfstandig te kunnen werken en met een helder verhaal je resultaten in het lab aan mij uiteen kon zetten om zo tot de volgende stap te komen. Pascal, jouw werk aan de gestabiliseerde methyleenbisphosphonaten staat bij de future prospects en dat is niet zonder reden. Je hebt hiermee namelijk prachtig werk geleverd en het is niet meer dan terecht dat je hier dan ook zelf voor jouw eigen promotie een mooi vervolg aan mag geven!

Aan mijn familie, schoonfamilie en vrienden. Ook jullie weten dat het niet altijd even makkelijk is geweest en voor de morele steun die ik van jullie heb ontvangen kan ik jullie ook nooit genoeg bedanken. Aan mijn ouders die mij hebben aangemoedigd en gesteund om na mijn bachelor door te studeren. Hierom ben ik ook trots dat mijn vader als paranimf achter me staat wat voor mij symbolisch is voor het feit dat jullie me feilloos hebben gesteund in deze reis van zelfontwikkeling. Dat brengt mij tot de eervolle plek om als laatst te bedanken, mijn vrouw, Rosalyn, die tijdens deze reis ook veel te verduren heeft gekregen. Ik kan je nooit genoeg bedanken. De tweede positie als paranimf staat ook hier symbool voor de steun die jij me hebt gegeven in deze periode, zonder jou had ik het niet gekund.