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

Voorneveld, J.

Chapter 3

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

Introduction

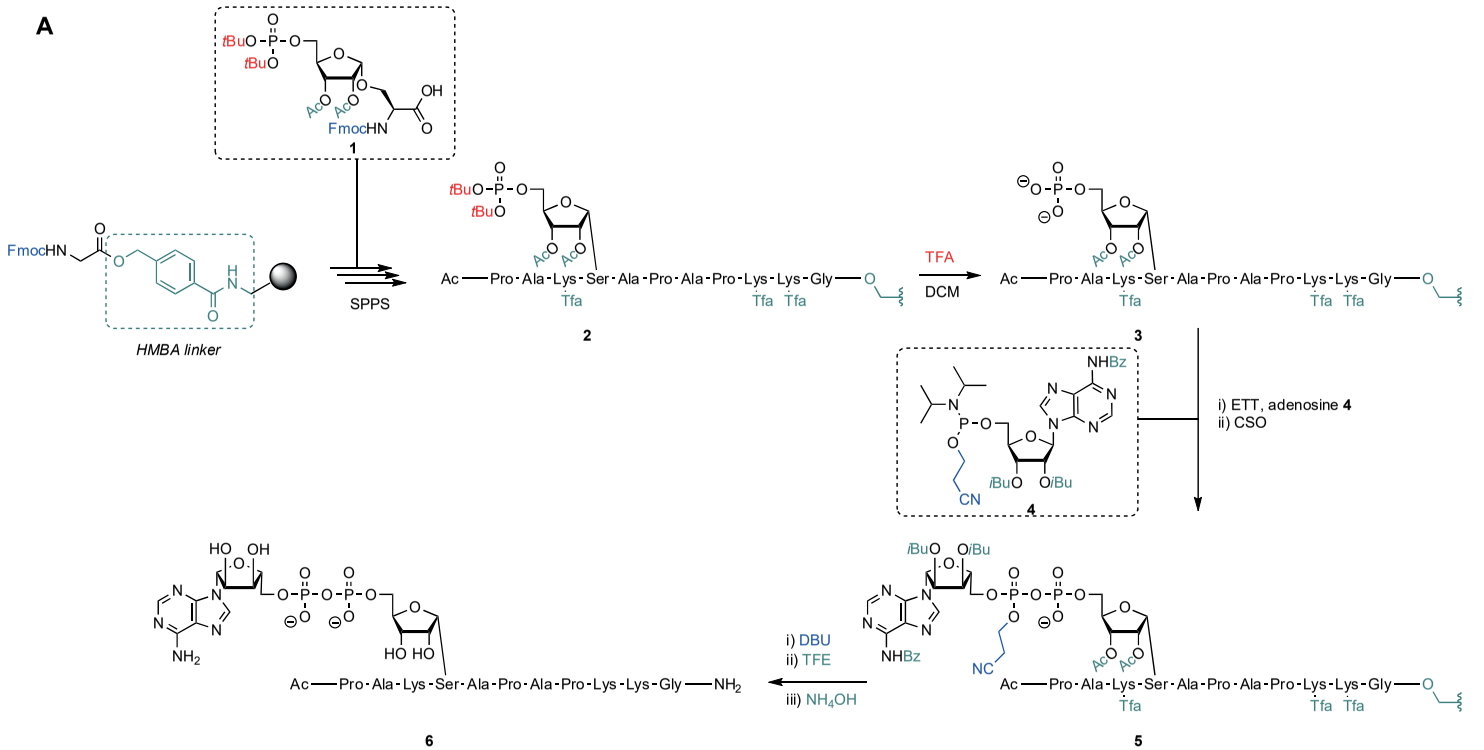
In the research on ADP-ribosylation, multiple types of amino acid side chains have been discovered to function as acceptors in this post-translational modification (PTM), that is introduced onto target proteins by (ADP-ribosyl)transferases (ARTs) from the PARP-family. Historically, arginine (Arg) was 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

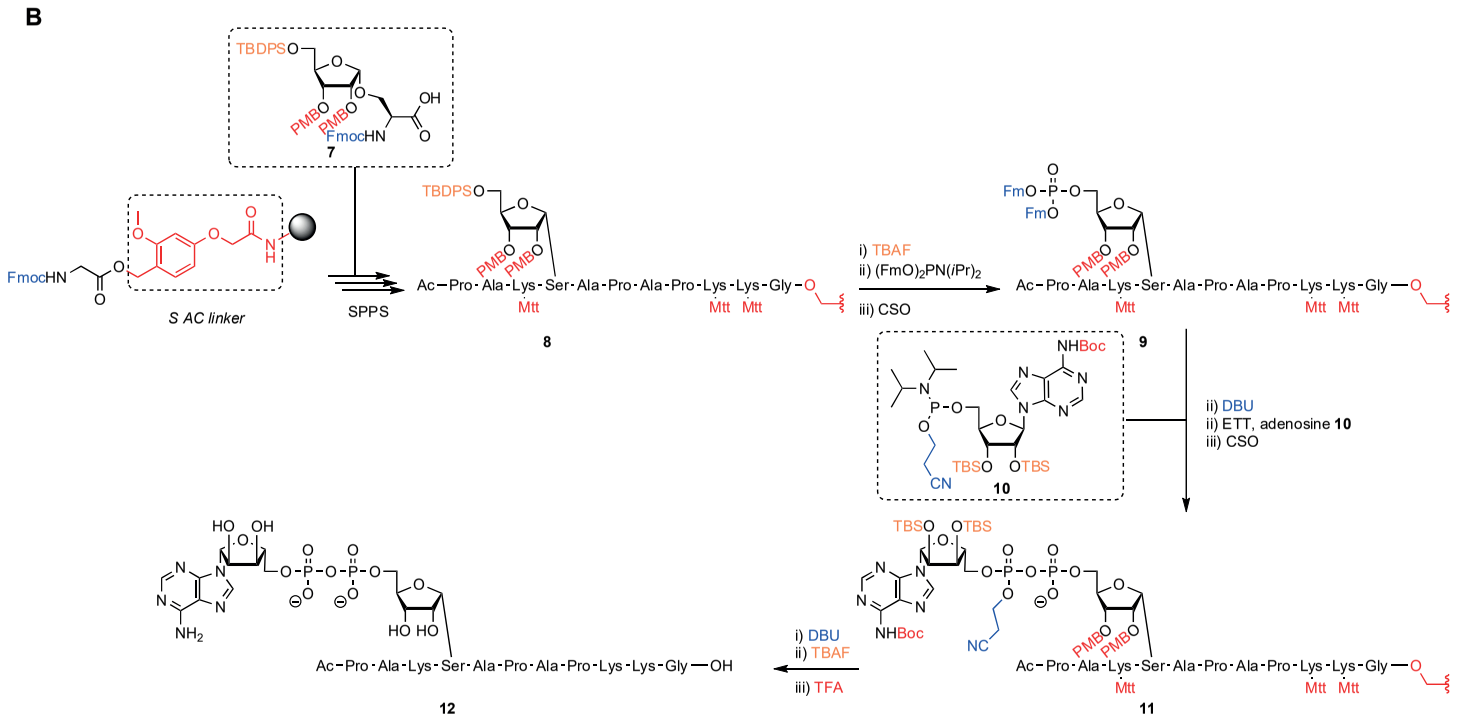
using a *t*Bu-protected phosphotriester that can be deprotected with TFA to give **3**. In the next step the pyrophosphate is constructed by coupling protected adenosine amidite **4** to phosphomonoester **3**, followed by oxidation of the resulting P^{III} – P^{V} intermediate to give protected ADPr-peptide 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

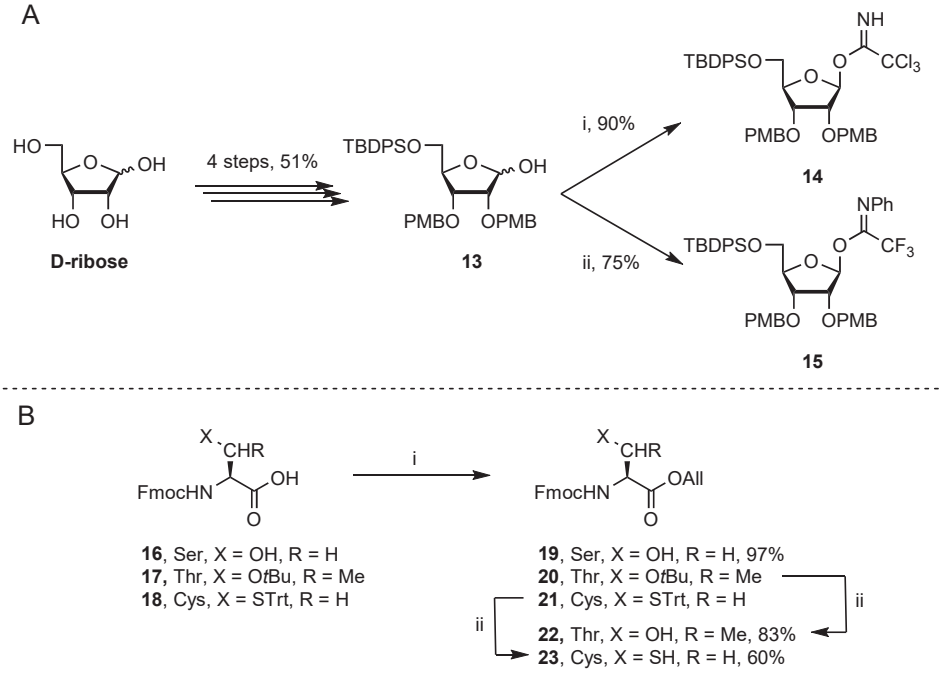
Ribofuranosylated Fmoc amino acids

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

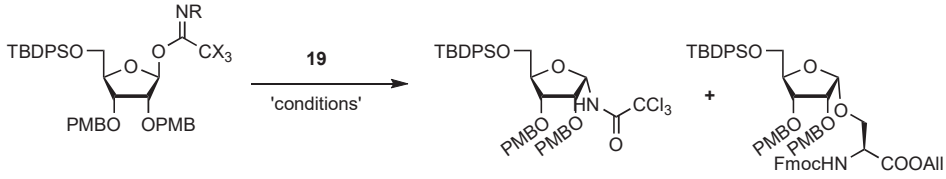
was reasoned that the softer character of the TBS cation would decrease the acidity of the glycosylation conditions and lead to less cleavage of the PMB groups. Although the use of TBSOTf significantly improved the yield of product **25**, side product **24** was still formed in 12%. A two-fold increase 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

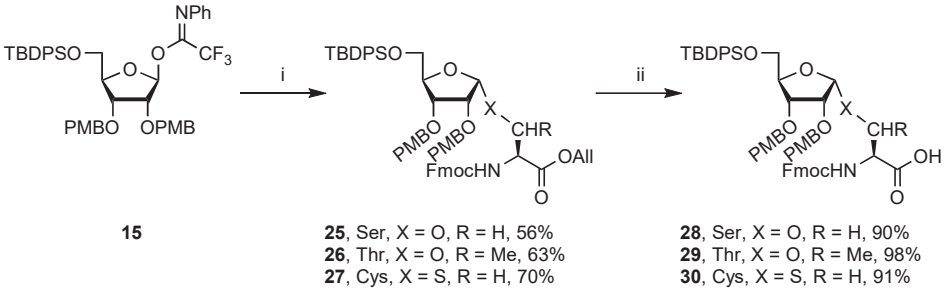
25

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

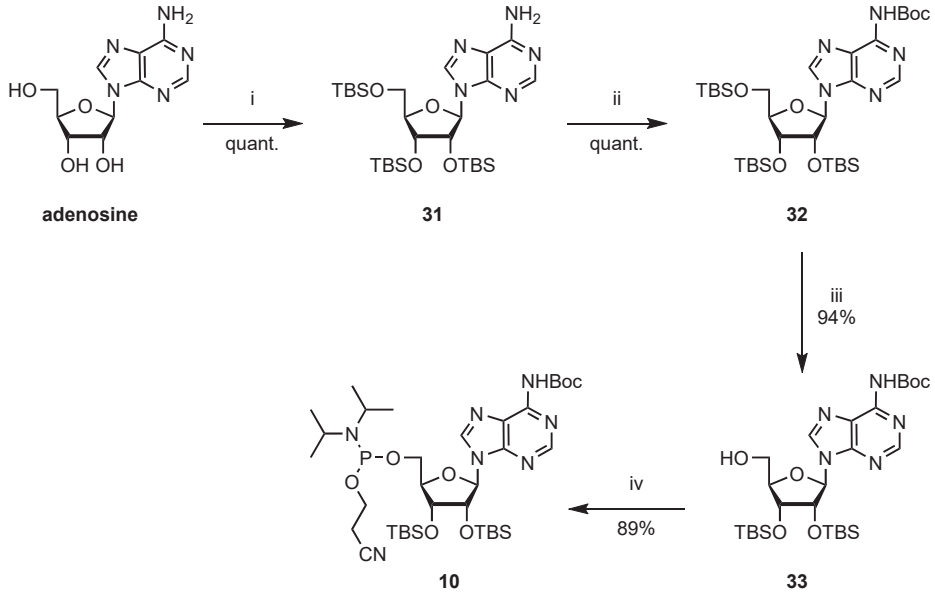
After the optimization of glycosylation conditions, the synthesis of ribosylated Fmoc amino acids **28**, **29** and **30** (Scheme 3) that are suitable building blocks for the solid phase synthesis of ADPr-peptides was undertaken. To this end, donor **15** was coupled to the appropriately protected 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.



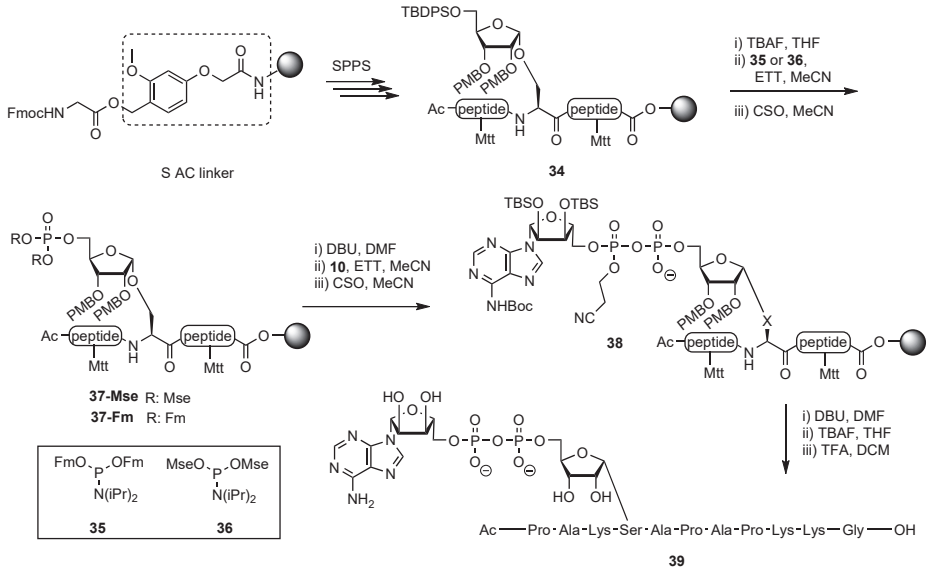
Scheme 4. Synthesis of adenosine amidite **10**. Reagents and conditions: i) TBS-Cl, imidazole, DMF, 50 °C. ii) Boc_2O, DMAP, THF, reflux. iii) TFA, H_2O, THF, 0 °C. iv) 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, DIPEA, DCM.

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.

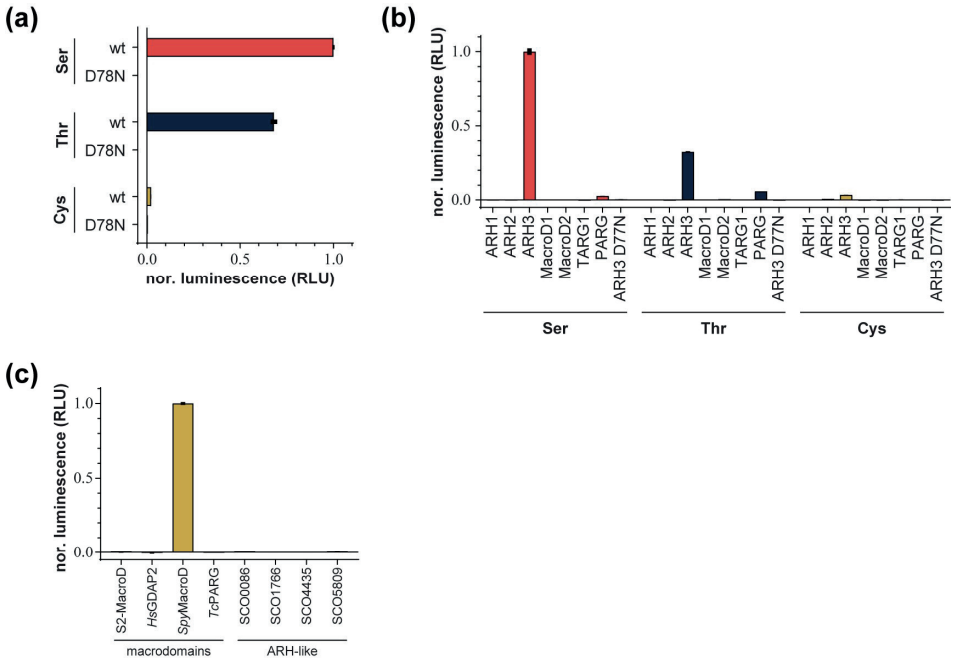


Figure 1. Enzymatic hydrolysis of glycosidic linkages in ADP-ribosylated Ser, Thr and Cys containing peptides **39**, **44** and **46**, respectively. (a-c) Measurements of hydrolase activity against the various ADPr-peptide linkages was assessed by converting released ADPr into AMP via NUDT5 and subsequently 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.

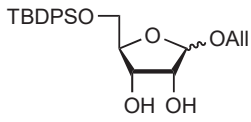
The availability of the Ser-ADPr, Thr-ADPr and Cys-ADPr peptides allowed the direct assessment and comparison of the sensitivity of these linkages towards enzymatic hydrolysis. It was found that the additional methyl group in Thr, as compared to Ser, leads to a pronounced reduction in turn-over by ARH3. 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

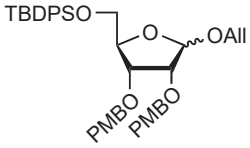
All reagents were used as received unless stated otherwise. Solvents used in synthesis were dried and stored over 4Å molecular sieves, except for MeOH and MeCN which were stored over 3Å molecular sieves. Triethylamine (TEA) and diisopropylethylamine (DIPEA) were stored over KOH pellets. Column chromatography 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



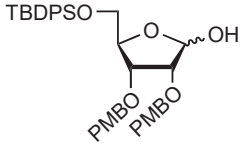
D-Ribose (11.3 g, 75 mmol) was suspended in allyl alcohol (190 mL, 0.4 M). Acetyl chloride (3.7 mL, 53 mmol, 0.7 eq.) was added and the suspension was stirred for 3 hours after which all starting material was dissolved. The reaction was quenched with pyridine and concentrated *in vacuo*. The residue 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).

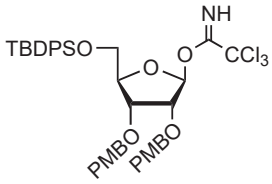
2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside (13)



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

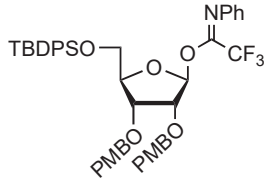
arom. α+β), 7.49 – 7.33 (m, 12H, TBDPS arom. α+β), 7.33 – 7.16 (m, 12H, PMB arom. α+β), 6.90 – 6.80 (m, 8H, PMB arom. α+β), 5.31 – 5.24 (m, 2H, H-1 α+β), 4.68 – 4.43 (m, 7H, 2x Ch_2 PMB α+β), 4.32 – 4.24 (m, 3H, H-3β + H-4α + CH_2 PMB), 4.21 (dt, *J* = 6.8, 2.6 Hz, 1H, H-4β), 4.05 (dd, *J* = 4.9, 1.8 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).

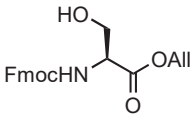
1-*O*-((*N*-phenyl)-2,2,2-trifluoroacetimido) 2,3-bis-*O*-(4-methoxybenzyl)-5-*O*-((*tert*-butyl)-diphenylsilyl)-α,β-D-ribofuranoside (15)



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),

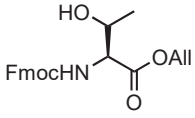
144.3, 143.9 (Cq NPh), 135.7, 135.6, 135.6, 135.5(CH arom. TBDPS/NPh), 133.3, 133.1, 133.0, 132.7, 130.4 (Cq TBDPS/PMB), 129.9, 129.9, 129.7, 129.7 (CH arom.), 129.6 (Cq), 129.5 (CH arom.), 129.4 (Cq), 129.4, 129.3, 129.2, 128.7, 128.7, 127.8, 127.8, 127.7 (CH arom.), 124.2, 120.6, 119.6, 114.3, 113.9, 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)**



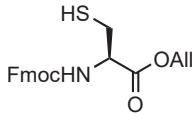
Fmoc-Ser-OH (3.45 g, 10 mmol, 1 eq.) was co-evaporated thrice with 1,4-dioxane and dissolved DMF (50 mL, 0.2 M). DIPEA (2.10 mL, 12 mmol, 1.2 eq.) was added followed by the addition of allyl-bromide (1.04 mL, 12 mmol, 1.2 eq.) and stirred overnight. The reaction was quenched with 100 mL water and transferred 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)**



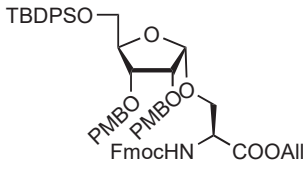
Fmoc-Cys(Trt)-OH (1.98 g, 5 mmol) was co-evaporated thrice with 1,4-dioxane and dissolved in DMF (25 mL, 0.2 M). DIPEA (1.04 mL, 6.0 mmol, 1.2 eq.) was added followed by allyl bromide (0.52 mL, 6.0 mmol, 1.2 eq.). The resulting solution was stirred overnight. The reaction was carefully quenched by the 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)



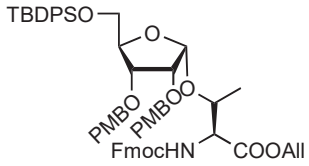
Fmoc-Cys(Trt)-OH (2.23 g, 3.81 mmol) was co-evaporated thrice with toluene and dissolved in DMF (19 mL, 0.2 M). DIPEA (0.80 mL, 4.6 mmol, 1.2 eq.) was added followed by allyl bromide (0.40 mL, 4.6 mmol, 1.2 eq.). The resulting solution was stirred overnight. The reaction was carefully quenched by the 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)



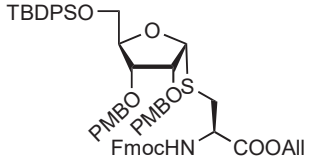
Compound **14** (1.60 g, 2.00 mmol, 1.1 eq.) and Fmoc-Ser-OAll **19** (669 mg, 1.82 mmol, 1.0 eq. relative to the donor) were co-evaporated thrice with toluene and dissolved in DCM (20 mL, 0.1 M relative to the donor). The reaction was cooled to -50 °C and TBSOTf (46 µL, 0.2 mmol, 0.1 eq. relative to 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)



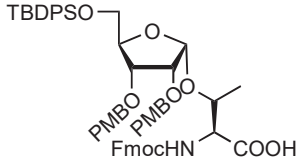
Compound **14** (1.60 g, 2.00 mmol, 1.3 eq.) and Fmoc-Cys-OAll **23** (595 mg, 1.55 mmol, 1.0 eq. relative to the donor) were co-evaporated thrice with toluene and dissolved in DCM (20 mL, 0.1 M relative to the donor). The reaction was cooled to -50 °C and TBSOTf (46 µL, 0.2 mmol, 0.1 eq. relative to 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)



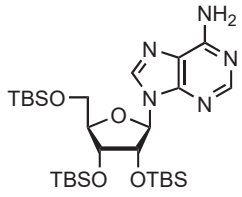
Compound **26** (525 mg, 0.529 mmol, 1.0 eq.) was dissolved in DCM (5.3 mL, 0.1 M). DMBA (372 mg, 2.38 mmol, 4.5 eq.) and a catalytic amount of $Pd(PPh_3)_4$ were added and the reaction was stirred for 15 minutes before TLC showed full conversion of the starting material into a lower running product. 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)

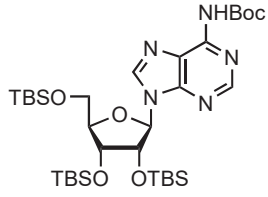
Compound **27** (1.08 g, 1.09 mmol, 1.0 eq.) was dissolved in DCM (11 mL, 0.1 M). DMBA (775 mg, 4.95 mmol, 4.5 eq.) and a catalytic amount of $Pd(PPh_3)_4$ were added and the reaction was stirred for 15 minutes before TLC showed full conversion of the starting material into a lower running product. The 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)



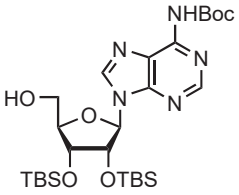
Adenosine (2.67 g, 10.0 mmol) was co-evaporated thrice with anhydrous pyridine, dissolved in anhydrous DMF (20 mL, 0.2 M) and heated to 50 °C after which imidazole (3.40 g, 50.0 mmol, 5.0 eq.) and *tert*-butyldimethylsilyl chloride (50 wt. % in toluene, 12.2 mL, 35.0 mmol, 3.5 eq.) were added. After 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)**



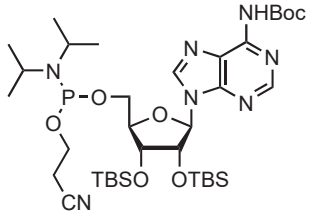
Compound **31** (5.40 g, 8.86 mmol) was co-evaporated thrice with anhydrous 1,4-dioxane, dissolved in anhydrous THF (89 mL, 0.1 M) and cooled to 0 °C, after which DMAP (217 mg, 1.77 mmol, 0.2 eq.) and di-*tert*-butyl dicarbonate (6.1 mL, 26.6 mmol, 3.0 eq.) were added. The mixture was heated to reflux 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.