



Universiteit
Leiden
The Netherlands

Synthetic methodology for the preparation of nucleic acid containing peptides

Heden-van Noort, G.J. van der

Citation

Heden-van Noort, G. J. van der. (2012, March 29). *Synthetic methodology for the preparation of nucleic acid containing peptides*. Retrieved from <https://hdl.handle.net/1887/18639>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/18639>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/18639> holds various files of this Leiden University dissertation.

Author: Heden-van Noort, Gerbrand Jan van der

Title: Synthetic methodology for the preparation of nucleic acid containing peptides

Date: 2012-03-29

Synthetic Methodology for the Preparation of Nucleic Acid Containing Peptides

Synthetic Methodology for the Preparation of Nucleic Acid Containing Peptides

Gerbrand J. van der Heden - van Noort

Gerbrand J. van der Heden - van Noort



Universiteit Leiden



Synthetic Methodology for the Preparation of Nucleic Acid Containing Peptides

PROEFSCHRIFT

Ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus prof. mr. P. F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 29 maart 2012
klokke 15.00 uur

door

Gerbrand Jan van der Heden – van Noort

Geboren te Alkemade in 1984

Promotiecommissie

Promotor: Prof. dr. G. A. van der Marel

Co-promotor: Dr. D. V. Filippov

Overige leden: Prof. dr. H. S. Overkleeft
Prof. dr. J. Brouwer
Prof. dr. C. A. A van Boeckel
Prof. dr. J. Lugtenburg
Dr. D. A. Stetsenko

Cover-design by Horizon Creative – Rotterdam (NL).

Colored artwork: Weinsberger Kreuz Schnittstelle für Gestaltung – Weinsberg (D) as commissioned by MARA-artist & events – Wuestenrot (D) for Himmelfahrt Festival.

Background artwork: Horizon Creative as commissioned by Make Up Your Mind, www.muym.nl.

The printing of this thesis was financially supported by the J. E. Jurriaanse Stichting.

Printed by Wöhrmann Printing Services.

‘Wij zijn uiteindelijk allemaal amateurs en dat is ook helemaal niet erg.’

Job Cohen – *De Volkskrant* – 2 oktober 2009

Table of Contents

List of Abbreviations	7
Chapter 1	9
General Introduction	
Chapter 2	35
A Versatile One-Pot Procedure to Phosphate Monoesters and Pyrophosphates	
Chapter 3	51
Synthesis of Mono-ADP-Ribosylated Oligopeptides Using Ribosylated Amino Acid Building Blocks	
Chapter 4	73
Ribosylation of Adenosine; An Orthogonally Protected Building Block for the Synthesis of ADP-Ribosyl Oligomers	

Chapter 5	89
A General Synthetic Method Towards Uridylylated Picornavirus VPg Proteins	
Chapter 6	103
Synthesis of Nucleotidylated Polio Virus VPg Proteins	
Chapter 7	121
Sequential Approach Towards RNA-Nucleopeptides	
Chapter 8	133
Summary and Future Prospects	
Nederlandse Samenvatting	142
List of Publications	144
CurriculumVitae (Nederlands)	146
CurriculumVitae (English)	147
Dankwoord	148

List of Abbreviations

A	adenine	dd	double doublet
Ac	acetyl	ddd	double double doublet
AcOH	acetic acid	DiPEA	<i>N,N</i> -diisopropyl- <i>N</i> -ethylamine
ADP-ribosyl	adenosine diphosphate ribosyl	DMAP	4-(<i>N,N</i> -dimethylamino)pyridine
Ala	L-alanine	DMF	<i>N,N</i> -dimethylformamide
All	allyl	Dmt	4,4'-dimethoxytrityl
Alloc	allyloxycarbonyl	DNA	deoxyribonucleic acid
ap.	apparent	DTT	dithiothreitol
aq.	aqueous	EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide
Arg	L-arginine	Et	ethyl
Arom.	aromatic	EtOAc	ethyl acetate
Asn	L-asparagine	EtOH	ethanol
Asp	L-aspartic acid	Fm	(9H-fluoren-9-yl)methyl
Bn	benzyl	Fmoc	(9H-fluoren-9-yl)methoxycarbonyl
Boc	<i>tert</i> -butoxycarbonyl	G	guanine
BOP	benzotriazol-1-yl-oxy-tris-(dimethylamino)phosphonium hexafluorophosphate	Gln	L-glutamine
Bz	benzoyl	Glu	L-glutamic acid
Bzl	benzyl	h	hour
C	cytosine	HATU	O-(7-azabenzotriazol-1-yl)- <i>N,N,N'</i> , <i>N'</i> -tetramethyluronium hexafluorophosphate
Cbz	benzyloxycarbonyl	HCP	highly crosslinked polystyrene
CNE	cyanoethyl	hept	heptet
Cq	quaternary carbon atom	HMBA	4-hydroxymethylbenzoic acid
CRE	cis-acting replication element	HMPB	4-(4-hydroxymethyl-3-methoxy-phenoxy)-butyric acid
Cys	L-cysteine	HOBt	<i>N</i> -hydroxybenzotriazole
d	doublet	HRMS	high resolution mass spectrometry
DAST	<i>N,N</i> -diethylaminosulfur trifluoride	Hs	L-homoserine
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene		
DCE	dichloroethane		
DCI	4,5-dicyanoimidazole		
DCM	dichloromethane		

Hz	herz	PyBOP	benzotriazol-1-yl-oxy-tris-
<i>i</i> Bu	<i>isobutyryl</i>		pyrrolidinophosphonium
<i>i</i> Pr	<i>isopropyl</i>		hexfluorophosphate
IR	infrared	RNA	ribonucleic acid
kD	kilo Dalton	RP-HPLC	reversed phase high
LC-MS	liquid chromatography mass spectrometry		performance liquid chromathography
Lev	levulinoyl	rt	room temperature
m	multiplet	s	singlet
M	molarity	sat.	saturated
MART	<i>mono</i> -ADP-ribosyl transferase	Ser	serine
MBHA	4-methylbenzhydrylamine	SPPS	solid phase peptide synthesis
mCPBA	<i>meta</i> -chloroperoxybenzoic acid	T	thymine
Me	methyl	t	triplet
MeCN	acetonitrile	TBAF	tetra- <i>n</i> -butylammonium fluoride
MeOH	methanol	TBDMS	<i>tert</i> -butyldimethylsilyl
min	minute	TBDPS	<i>tert</i> -butyldiphenylsilyl
NAD ⁺	nicotinamide adenine dinucleotide	<i>t</i> Bu	<i>tert</i> -butyl
NMP	<i>N</i> -methyl-2-pyrrolidone	TEA	triethylamine
NMR	nuclear magnetic resonance	TFA	trifluoroacetic acid
<i>n</i> PrOH	1-propanol	TFAc	trifluoroacetyl
NPS	2-nitrophenylsulfenyl	THF	tetrahydrofuran
Nuc	nucleophile	THP	tetrahydropyranyl
<i>p</i>	<i>para</i>	Thr	L-threonine
Pac	phenoxyacetyl	TIPDS	1,1,3,3-tetra- <i>isopropyl</i> disiloxane
PAR	<i>poly</i> adenosine diphosphate ribose	TIS	triisopropylsilane
PARP	PAR polymerase	TLC	thin layer chromatography
Pbf	2,2,4,6,7- pentamethyldihydro- benzofuran-5-sulfonyl	TMS	trimethylsilyl
PE	petroleum ether	Trt	trityl
Phac	phenacyl	Tyr	L-tyrosine
Phe	L-phenylalanine	U	uracil
PMB	<i>para</i> methoxybenzyl	VPg	viral protein genome-linked
POM	pivaloyloxymethyl		
ppm	parts per million		
Pro	L-proline		

Chapter 1

General Introduction

The complexity of life can be appreciated when looking at the structural diversity of biopolymers involved in biological processes. A wide variety of carbohydrates, peptides and nucleic acids plays crucial roles in an even wider range of important biochemical events. To gain insight in these biological processes the need for synthetically obtained fragments of biopolymers and derivatives thereof has risen in the last half of the twentieth century. Synthetic protocols towards oligonucleic acids and peptides have been developed allowing access to well-defined material for research in medicinal chemistry, chemical biology and structural biology. An enormous leap forward in gaining accessibility to such materials was the introduction of insoluble polymeric supports for the synthesis of peptides and oligonucleic acids.¹⁻³ The limits of solution phase assembly, such as insolubility and difficulties in purification after each sequential step, are greatly reduced by using solid support chemistry.² Automated synthesis based on this approach has proven to be a useful tool to obtain protein fragments and oligonucleotides. Hybrid structures between different classes of biopolymers, such as glycoproteins, glycolipids and nucleoproteins^{4,5} however are not always easily accessible using the currently used solid phase based approaches. Conditions that are used in the synthesis of one biopolymer fragment might not be compatible with the stability of the second biopolymer fragment.

The aim of the work described in this thesis is to develop synthetic methods that give access to hybrid molecules containing both a peptide part and a nucleic acid part. These methods should be applicable on solid support and ultimately open the way to the synthesis of more complex fragments derived from biomolecules.

An overview of naturally occurring hybrid molecules, consisting of a peptide and a nucleic acid and synthetic efforts to obtain fragments and derivatives of these interesting molecules is presented in this chapter.

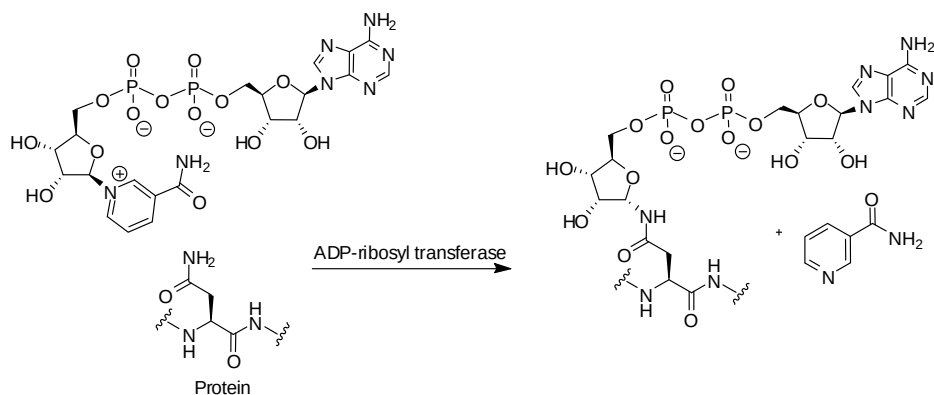
1.0 Post-translational modifications

Investigation of post-translational modifications of proteins is a major research topic since these modifications are key players in regulation of eukaryotic cellular processes.⁶ Such modifications range from the introduction of a simple acetyl⁷ or phosphate⁸ group to attachment of small proteins such as ubiquitin.⁹ Some post-translational modifications even cause alteration of the amino acid sequence of the target protein such as for instance deamidation of asparagine to aspartic acid residues.¹⁰ Introduction of nucleic acid monomers are also found amongst the diverse post-translational modifications, in processes such as ADP-ribosylation and adenylation.

1.1 ADP-ribosylation

Adenosine diphosphate ribosylation (ADP-ribosylation) is effected by two classes of enzymes called *poly*-ADP-ribosyl polymerases (PARPs) or *mono*-ADP-ribosyl transferases (MARTs).¹¹ These enzymes transfer ADP-ribose from β -NAD⁺ to nucleophilic functional groups in the side chain of amino acid residues of the target protein, a process that is accompanied by the release of nicotinamide (see **Figure 1**).¹²

Figure 1



ADP-ribosylation has an influence on a variety of processes ranging from the triggering of DNA repair mechanisms to regulation of immune responses.^{13,14}

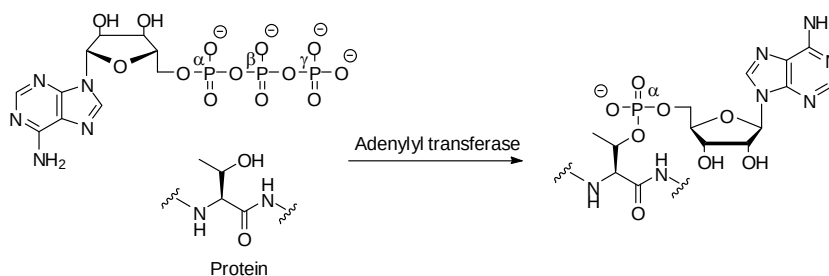
Schramm and co-workers have done extensive research on identifying the transition state of the ADP-ribosylation process effected by the bacterial pertussis toxin.¹⁵⁻¹⁸ For this work they treat a synthetic peptide fragment with pertussis toxin and NAD⁺ analogues to obtain ADP-ribosyl peptides. Muir and co-workers have reported the synthesis of aminooxy or *N*-methyl aminooxy

modified peptides that can be conjugated to ADP-ribose.¹⁹ These prepared conjugates carrying a non-natural linkage to ADP-ribose are recognized by *poly*-ADP-ribosylating enzymes in enzymatic essays. Although several other reports on modified NAD⁺ analogues have been published²⁰⁻²⁴, no fully chemical synthesis of ADP-ribosyl containing peptides is known to date.

1.2 Adenylylation

Adenylylation is another post translational modification process in which adenosine monophosphate (AMP) is transferred from adenosine triphosphate (ATP) to a nucleophilic residue in the protein backbone under the influence of adenylyl transferring enzymes (see **Figure 2**).²⁵ If this nucleophilic residue is a hydroxyl (in the case of serine, threonine or tyrosine) a stable phosphodiester bond is formed between adenosine and the protein. Adenylylation reactions involving the introduction of such a stable adenylyl adduct usually plays a role in the regulation of protein activity.²⁶ When carboxylate or amine side chains are adenylylated unstable mixed anhydrides or phosphoramidates are formed. The function of adenylylation in this type of reactions is the introduction of a good leaving group. In this transient process the energy stored in the α,β -phosphoric anhydride bond in ATP is released. This overall process facilitates thermodynamically unfavorable reactions to take place. Examples of adenylyl transferases are bacterial enzymes that adenylylate GTPases.²⁷⁻²⁹

Figure 2



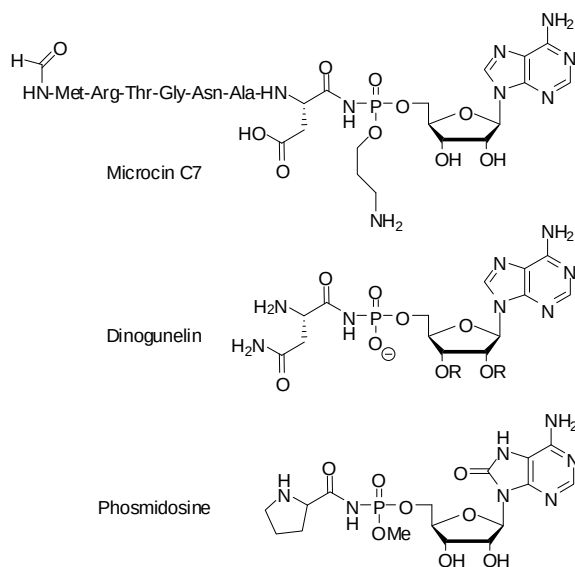
A synthetic route towards adenylylated peptides was published using a H-phosphonate approach.³⁰ Unprotected serine or threonine containing resin-bound peptides were transformed to their corresponding H-phosphonates and subsequently coupled to 2',3'-isopropyl adenosine and oxidized. Cleavage from the solid support and protective group removal using 95% trifluoroacetic acid yielded the target peptides. The yields, however, were very low, which might be due to depurination of the nucleotidyl peptides in acidic environment. Although small quantities of adenylyl serine or threonine peptides were obtained, this method might not be widely applicable since the adenylyl peptides seem to be unstable under the conditions used in this approach. Furthermore, a similar H-phosphonate approach was reported to be inefficient for the synthesis of tyrosine containing nucleotidyl peptides, thereby limiting the scope of such a H-phosphonate method to obtain adenylyl peptides.³¹ Another approach towards adenylylated peptides is based on the assembly of pre-adenylylated Fmoc protected amino acids.^{32,33} In this approach 2', 3'-O-isopropylidene protection is installed on the ribosyl part of adenosine and the exocyclic amine of adenine is protected with two Boc groups. The 5'-O-phosphoramidite of this

protected nucleoside is coupled to suitably protected serine, threonine and tyrosine. After oxidation and protective group manipulations these adenylyl amino acids are incorporated in peptides using an on-line solid phase peptide synthesis method. Acidic removal of the protective groups and simultaneous cleavage from the resin gives access to the adenylylated peptides accompanied by trace amounts of depurinated analogues.

2.0 Antibiotics

Some bacterial species are found to produce antibiotics, containing a peptide and nucleotide part, to defend themselves. *Escherichia coli* for instance produces the antibiotic Microcin C7 (see **Figure 3**).³⁴ The peptidyl part of this phosphoramidate mediates the transport of the antibiotic into the target bacterial cell, where processing leads to cleavage of the peptide bond between alanine and aspartic acid releasing the aspartic acid adenylyl phosphoramidate which acts as an inhibitor of protein synthesis in the targeted bacterium.^{35,36} A similar defense mechanism is employed by the eggs of the Japanese fish *Stichaeus grigorjewi*, that produce the Dinogunelin toxin (see **Figure 3**, R is either a proton or a unsaturated lipid residue).³⁷ Another example is Phosmidosine (see **Figure 3**), where 8-oxo-adenosine is attached to proline via a similar phosphoramidate linkage.³⁸ This antifungal antibiotic originating from *Streptomyces durhmeusie* is also found to possess anti-cancer activity. Phosmidosine and some analogues thereof have been prepared by coupling the methoxy-phosphoramidite of 8-oxo-adenosine to a protected proline amide.³⁹

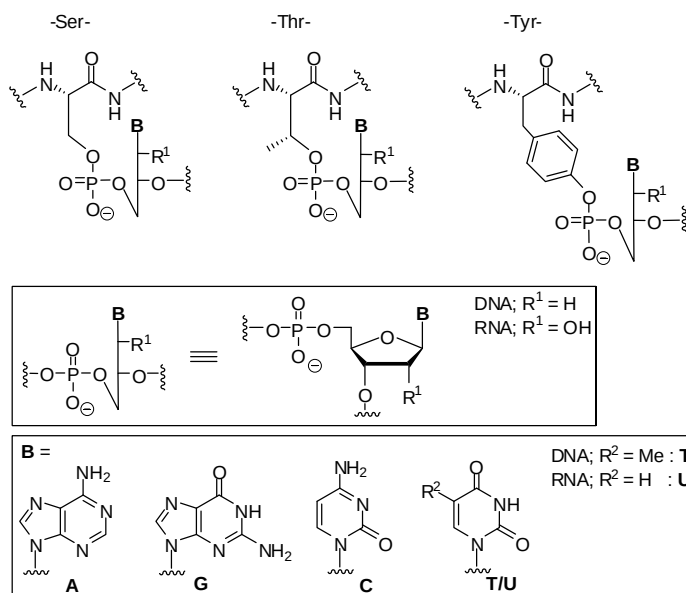
Figure 3



3.0 Nucleoproteins

Nucleoproteins are a class of hybrid biomolecules in which a protein is covalently linked to a nucleic acid chain via a phosphodiester bond between the 3'- or 5'-end of the DNA or RNA and the hydroxyl function of a serine, threonine or tyrosine residue in the protein backbone (See **Figure 4**).

Figure 4



Such nucleoproteins play important roles in the replication of RNA-viruses, DNA-viruses and bacteriophages:

RNA-viruses. In the case of poliovirus, a well investigated example of a RNA-virus, the small protein VPg (Viral Protein genome-linked), consisting of only 22 amino acids is attached via the tyrosine hydroxyl moiety to the 5'-end uridylic acid residue of the viral genome.⁴⁰ This protein acts as a cap-structure at the 5'-end of the viral RNA. Besides protecting the viral genome from exonuclease degradation VPg is believed to be a primer in progeny strand synthesis either as the free peptide or as its uridylylated form VPgpU or VPgpUpU.⁴¹

DNA-viruses. In adenovirus, an example of a DNA-virus, both 5'-ends of the double stranded viral genome are linked to a 55kD protein called TP (Terminal Protein) via a phosphodiester bond between serine residue 580 in the protein backbone and the 5'-deoxycytidylic acid of the genome. This nucleoprotein is believed to act as primer in DNA-replication.⁴²⁻⁴⁴

Bacteriophages. In a similar fashion bacteriophage Φ 29 DNA is covalently attached via the 5'-end deoxyadenosine to a serine residue in the 27kD protein called P3.^{45,46} This nucleoprotein is believed to act as primer in DNA-replication. Another example is the bacteriophage Φ X174

gene A protein, which is believed to cleave the double stranded DNA of the phage by the formation of a phosphodiester bond between a tyrosine residue and the 5'-end adenylic acid residue at the cleavage site, thereby creating a starting point for replication.⁴⁷

Topoisomerases. Nucleoproteins also occur as intermediates in unwinding and rewinding of supercoiled DNA in higher organisms. Topoisomerase type I and type II enzymes cleave the DNA backbone, reduce tension on the intertwined DNA to facilitate access for replication enzymes and then reconnect the cleaved DNA.⁴⁸ Topoisomerase type I enzymes induce single strand breaks via the formation of an intermediate in which a tyrosine residue in the protein's backbone is connected covalently to the 3'-phosphate of the broken DNA. Topoisomerase type II enzymes induce double strand breaks forming phosphodiester bond between a tyrosine or serine and the 5'-ends of both broken DNA strands.⁴⁹ Several drugs inhibiting topoisomerases are used in antibacterial and anticancer therapies. These drugs prevent the topoisomerase to reconnect the broken DNA, resulting in a permanent protein-bound DNA lesion, which eventually leads to apoptosis.⁵⁰

P53-protein. Another protein that is found to be covalently linked to RNA is tumor suppressor protein P53. This protein is involved in DNA repair and cell cycle regulation and the inability to express this protein leads to tumor formation.⁵¹ It is proposed that P53 is linked to the 5'-end of the RNA via Ser-389.⁵²

The synthesis of nucleoproteins or fragments thereof named nucleopeptides, is far more demanding than the synthesis of the individual nucleic acid and peptide fragments. When synthesizing such hybrid biopolymers, the inherent acid lability of the oligonucleotides prohibits the use of strong acid, which is commonly used in deprotection protocols in peptide chemistry. Likewise the strong basic conditions required for the removal of acyl protective groups at the exocyclic amines of nucleobases may be accompanied by unwanted β -elimination of phosphorylated serine and threonine residues.⁵³ In this process the proton on the α -position of serine or threonine is abstracted resulting in formation of dehydroalanine and the oligonucleotide 5'-phosphate. On top of this a complex protective group strategy is needed, requiring orthogonality of the permanent and temporary protective groups in the peptide and nucleic acid part of the projected nucleopeptide. One has to take all these properties of nucleopeptides and the desired protection scheme into account when devising a suitable synthetic route for fragments of nucleoproteins.

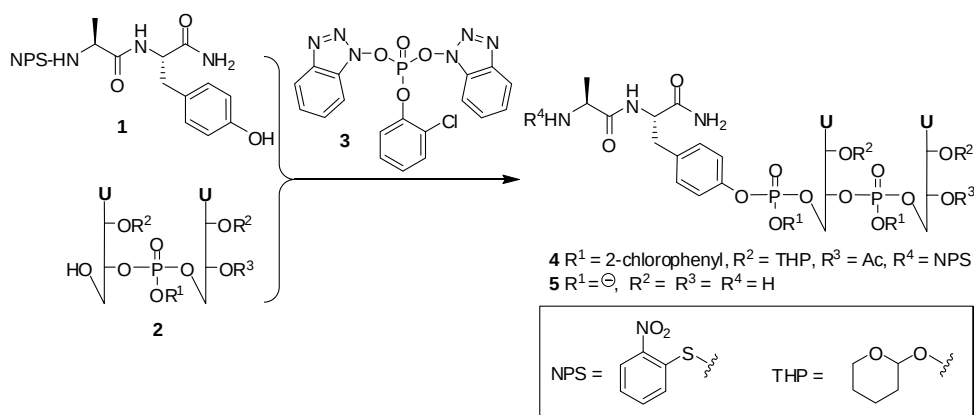
3.1 Synthesis of nucleopeptides in solution

Initial efforts to synthesize nucleopeptides take advantage of a balanced combination of solution approaches known from nucleic acid and peptide chemistry. A summary of different methods to install the crucial phosphodiester linkage between the peptide part and the nucleic acid part of nucleopeptides is described below, focusing on the phosphorus chemistry involved and the protective groups used in these strategies.

3.1.1 Phosphotriester approach

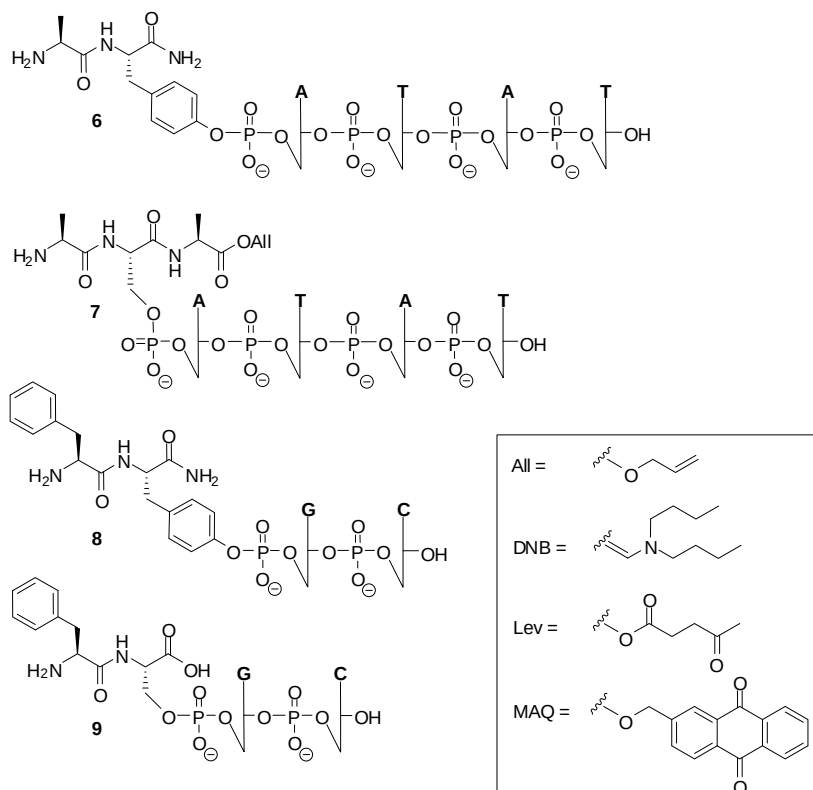
Van Boom and coworkers reported the first synthetically made ribonucleopeptide **5** (see **Figure 5**), which is prepared using a modified phosphotriester approach.⁵⁴ In this example bi-functional phosphorylating agent **3** equipped with two leaving groups (1-benzotriazolyl) and one permanent protective group (2-chlorophenyl) is first coupled to the hydroxyl functionality of dipeptide **1**, followed by coupling to dinucleotide **2** using *N*-methyl imidazole as activator. The protective groups used in this sequence of reactions are 2-nitrophenylsulfenyl (NPS) on the peptide N-terminus, 2-chlorophenyl on the phosphodiester, tetrahydropyranyl (THP) on the 2'-hydroxyl functions of the uridylic acid residues, acetyl on the 3'-hydroxyl of the terminal uridylic acid residue, while the C-terminal carboxylate of the peptide is capped as carboxamide. After construction of the protected phosphodiester linkage between the peptide part and the nucleotide part, removal of the protective groups was accomplished in a three step procedure. First oximate-reagent⁵⁵ was used to cleave the 2-chlorophenyl. Secondly a triethylamine-methanol-water treatment cleaved the acetyl group and a final acidic hydrolysis of the NPS and THP by treating with methanolic HCl (pH 2.5) for 3 hours followed by aqueous HCl (pH 2) for 14 hours. The order of removing protective groups in this sequence of reactions is carefully chosen to assure that migration or strand break in the RNA part is minimized.⁵⁶

Figure 5



The same phosphotriester methodology was applied in the preparation of deoxynucleopeptides **6** and **7**, having a phosphodiester between tyrosine or threonine residues and tetra-deoxynucleotide [5'-pATAT-3'], respectively (see **Figure 6**).⁵⁷ The exocyclic amine of adenine was protected using the NPS group, whereas the terminal thymidylic acid 3'-OH function was protected with the levulinoyl (Lev) group. Both protective groups could be removed without problems using tri-*n*-butylphosphine and hydrazine respectively, avoiding strong acidic conditions. In the case of **7** two issues deserve attention. First, the C-terminal allyl ester introduced to give access to the free acid, could not be removed using Pd(0) based procedures, rendering a partially protected nucleopeptide fragment. Second, the choice of protective groups in the preparation of **7** avoids the use of strong base as it was reported in a preliminary study that harsh alkaline conditions led to β -elimination at the amino acid side chain in syntheses of H-Ala-Ser[5'-pTT-3']-NH₂ and H-Ala-Thr[5'-pTT-3']-NH₂.⁵³ The phosphotriester approach was also employed to give nucleopeptides containing the dideoxynucleotide [5'-pGC-3'] (see **Figure 6**, compound **8** and **9**).⁵⁸ The protecting group used on the guanidine base was di-*n*-butylformamidinium (DNB), which is removable using hydrazinolysis, and NPS on cytidine. The 2-chlorophenyl groups, used as protection on the phosphodiester bonds, were removed using TBAF. The peptide C-terminal acid of **9** was protected as 2-(hydroxymethyl-9,10-anthraquinone (MAQ) ester, which could be removed using reductive cleavage.

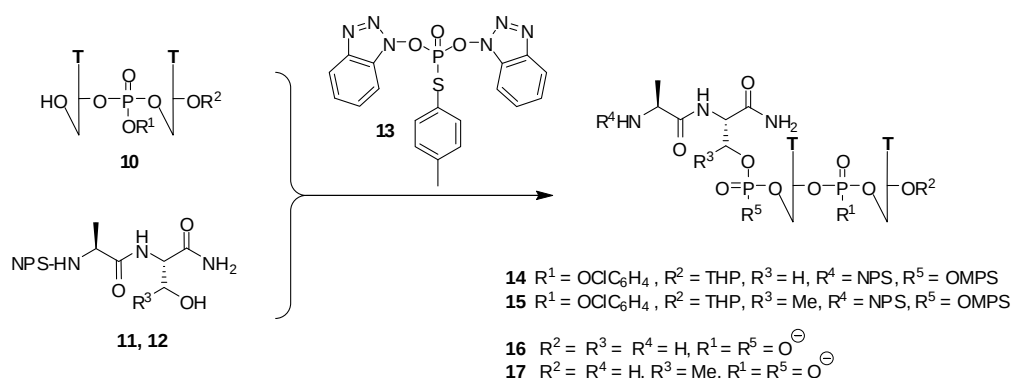
Figure 6



3.1.2 Phosphorothioate triesters

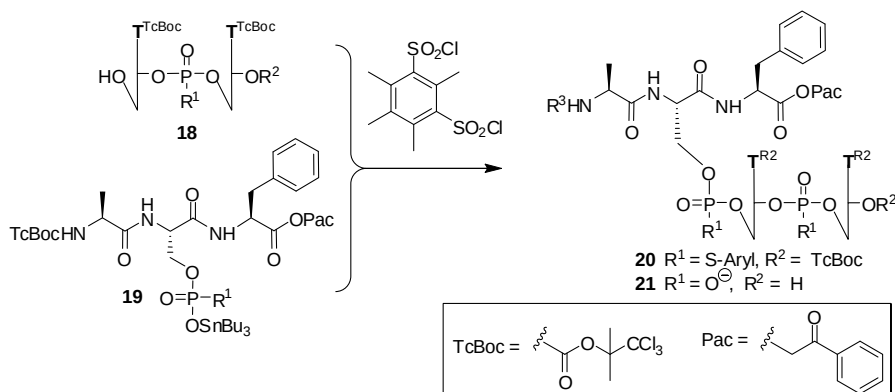
An alternative for the 2-chlorophenyl as phosphate protecting groups is a (substituted) phenylsulfenyl group.^{53, 59-62} Advantage of sulfenyl groups are the relatively mild deprotection procedures such as silver acetate, iodine in pyridine followed by water or (*n*-Bu₃Sn)₂O followed by TMSCl. These procedures avoid strong base, which is deleterious for serine and threonine containing nucleopeptides. Fully protected nucleopeptides **14** and **15** were prepared by reaction of phosphorylating agent **13** with dinucleotide **10** and subsequent condensation of the mono-hydroxybenzotriazole intermediate with dipeptides **11** or **12** (see **Figure 7**).^{53,59} The sulphenyl group in nucleopeptides **14** and **15** could be easily removed using silver acetate in pyridine-water, after which further deprotection gave access to **16** and **17**.

Figure 7



Another phosphorothioate approach makes use of phosphorothioate peptides, assembled from phosphorothioate amino acids⁶¹, that are activated and subsequently coupled to diribonucleotides to give diribonucleotide dipeptide Ala-Tyr[5'-pUU-3'] and dideoxyribonucleotide tripeptide Ala-Ser[5'-pTT-3']-Phe **21**, of which the synthesis is depicted in **Figure 8**.^{60,62}

Figure 8

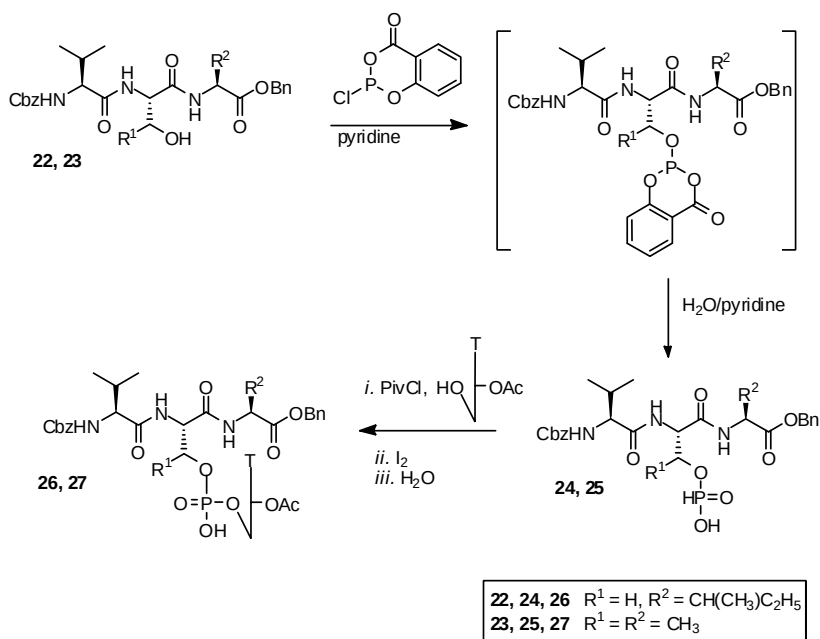


Deprotection of the phosphorothioates triesters using $(n\text{-Bu}_3\text{Sn})_2\text{O}$ resulted in formation of the unprotected phosphodiester bonds. The (2,2,2-trichloro-*tert*-butoxycarbonyl) (TcBoc) and phenacyl (Pac) ester groups could be removed using Zn/acetylacetone in pyridine under mild conditions.

3.1.3 H-phosphonate approach

The observation that oximate based procedures to remove 2-chlorophenyl protection on the phosphotriester moiety led to β -elimination in serine nucleopeptides, encouraged the investigation of the so-called H-phosphonate approach in nucleopeptide synthesis.³¹ In this approach the H-phosphonates of protected threonine and serine tripeptides (**22** and **23**, see **Figure 9**) were formed by reacting the peptide with salicylchlorophosphine to give an intermediate phosphite triester that was hydrolyzed to give H-phosphonates **24** and **25**, respectively. Subsequent reaction of H-phosphonate **24** or **25** with the free 5'-hydroxyl of 3'-O-acetyl thymidine under the agency of pivaloylchloride, followed by oxidation with iodine and immediate hydrolysis using water gave protected nucleopeptides **26** and **27**. Although the formation of the phosphodiester bond between thymidylic acid and peptides containing threonine or serine proved to be fruitful a similar coupling with tyrosine was not efficient, probably due to instability of the phosphite triester intermediate.

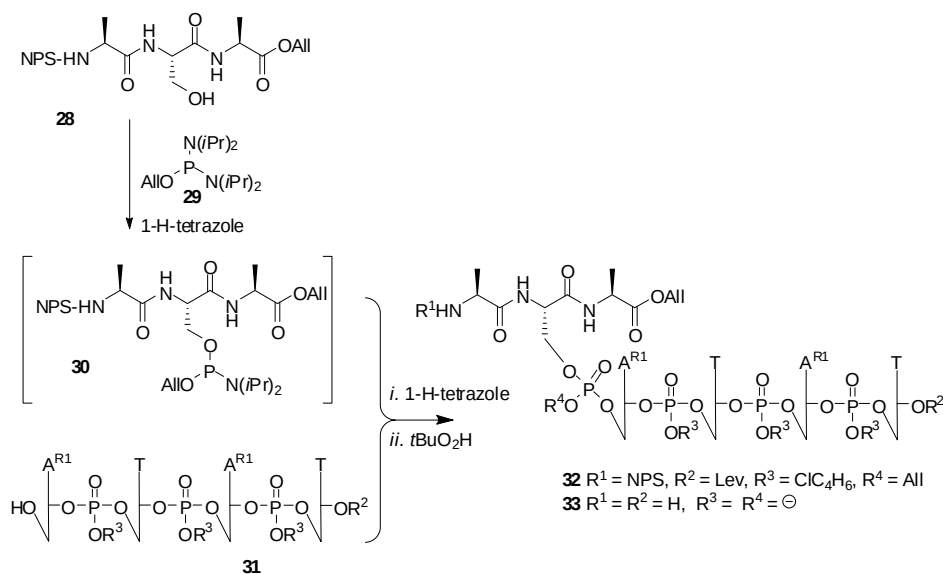
Figure 9



3.1.4 Phosphoramidite approach

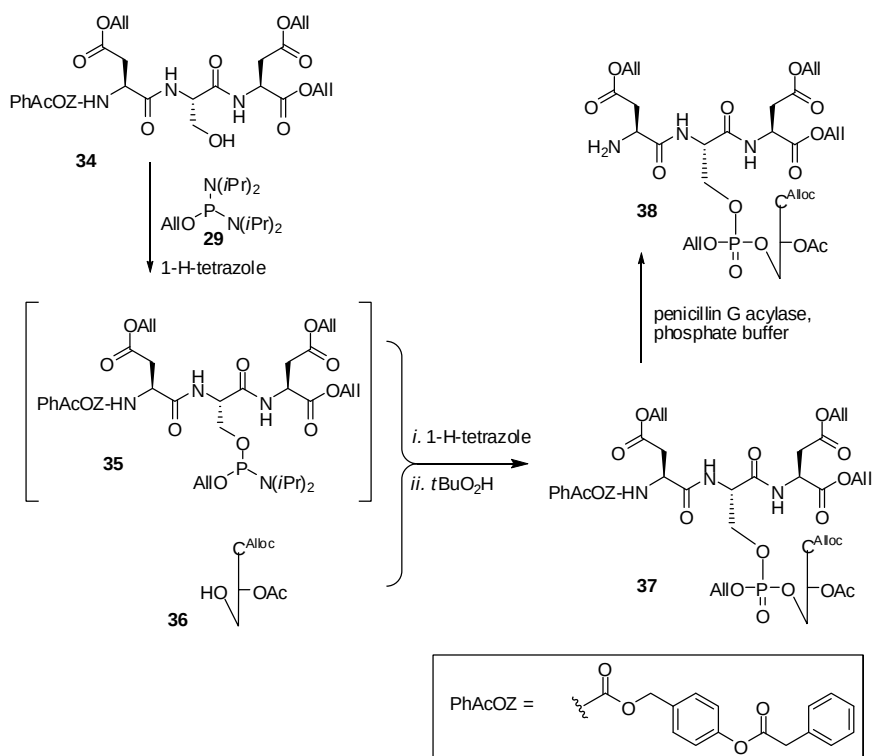
The phosphoramidite approach reported by Beaucage and Caruthers⁶³, based on the phosphite triester method first reported by Letsinger and Ogilvie⁶⁴, became the most used method in solid support based oligonucleotide synthesis. Therefore this method was also applied in nucleopeptide synthesis. Phosphitylation of peptide **28** using phosphordiamidite **29** gave rise to phosphoramidite **30** (see **Figure 10**).⁵⁷ The crude amidite was condensed under influence of tetrazole with separately prepared tetradeoxynucleotide **31**, followed by oxidation using peroxide to give protected nucleopeptide **32**. Using the phosphoramidite approach a higher yield (90%) is reported for the synthesis of **32** when compared to the phosphotriester method (64%).⁵⁷ Fully protected **32** was treated with TBAF in pyridine/water to cleave the phosphate protection 2-chlorophenyl and allyl groups, tri-*n*-butylphosphine in dioxane/water to cleave the NPS groups and hydrazine hydrate in pyridine/acetic acid to remove the Lev group, resulting in nucleopeptide **33**. No conditions however were found to cleave the C-terminal allyl ester.

Figure 10



Waldmann and co-workers have used a similar phosphoramidite approach to obtain nucleopeptides in which deoxycytidylic acid **36** is linked to a serine containing tripeptidyl phosphoramidite **35** (see **Figure 11**) to give fully protected nucleopeptide **37**.⁶⁵ The goal of this work was to extend the protective group repertoire used in nucleopeptide synthesis by the introduction of the phenylacetoxycarbonyl (PhAcOZ) group. This N-terminal PhAcOZ protective group in **37** can be removed by the enzyme penicillin G acylase under mild conditions (pH 6.8) to give **38**. No attempts were made however to remove all other protective groups on **38**.

Figure 11



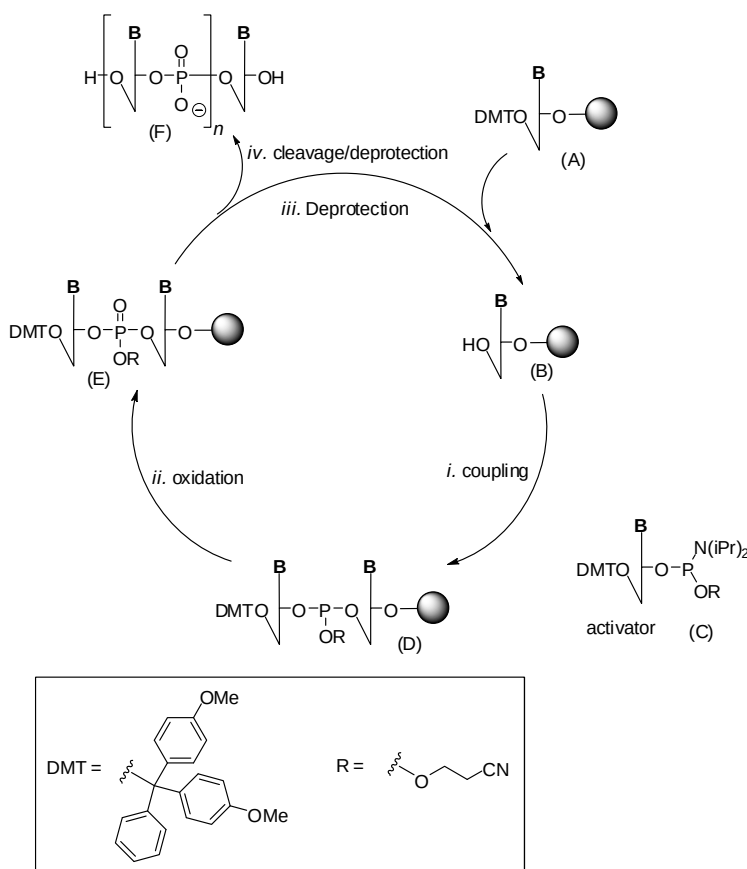
3.2 Syntheses of nucleopeptides on solid support

The (automated) solid phase synthesis of oligonucleotides and peptides eventually proved to be the best approach to obtain these biomolecules. The methodologies used in oligonucleotide and peptide chemistry were therefore adapted for the synthesis of nucleopeptides.

3.2.1 Synthesis of nucleopeptides starting from resin-bound oligonucleotides

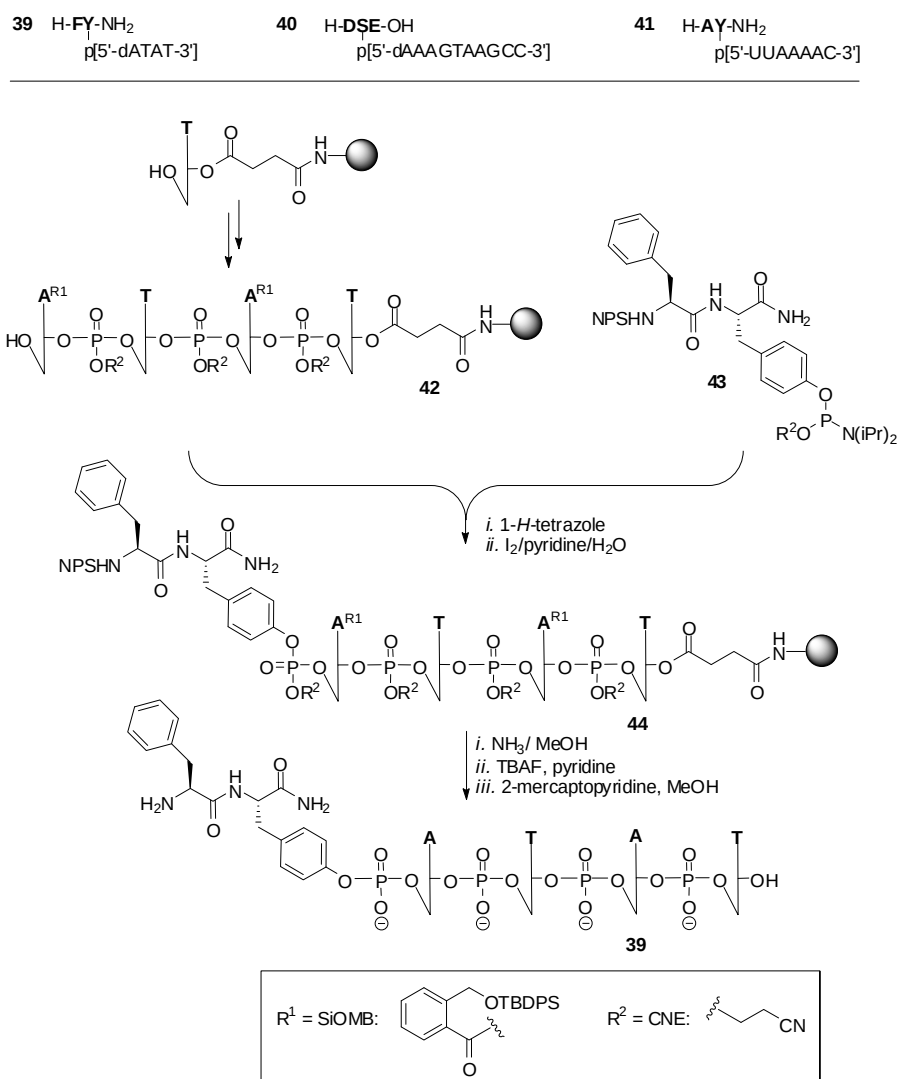
Applying the phosphoramidite methodology on a solid support was found to be most efficient in the preparation of oligonucleotides. In this method the first dimethoxytrityl (DMT) protected nucleotide is attached to the solid support (A) (see **Figure 12**) via a linker that is stable en route to the immobilized oligonucleotide (E), but can be cleaved in the final stage of the synthesis. The repetitive cycle in oligonucleotide synthesis starts with liberation of the primary hydroxyl function to give alcohol (B). Coupling with a 3'-phosphoramidite reagent (C) under agency of an activator gives phosphite triester intermediate (D), which is subsequently oxidized to give phosphotriester (E). Removal of the 5'- protecting group on the just attached nucleotide elongates the oligonucleotide with one nucleic acid residue. Repeating this cycle will elongate the resin-bound oligonucleotide chain from the 3'-end in the direction of the 5'-end, until final deprotection and release from the resin gives free oligonucleotide (F).

Figure 12



Nucleopeptides **39** - **41** were prepared via this method (exemplified for **39**, see **Figure 13**).⁶⁶⁻⁶⁸ First oligonucleotide **42** was prepared on solid support. Next, peptidylphosphoramidite **43** was coupled and subsequently oxidized to give fully protected **44**. Deprotection and release from the solid support gave access to nucleopeptide fragment **39**. The main feature in the protective group strategy used in the construction of **39** - **41** is the use of fluoride labile 2-(*tert*-butyldiphenylsilyloxymethyl)-benzoyl (SiOMB)⁶⁹ groups on the exocyclic amines of the nucleobases. Cyanoethyl protection, and the oxalylester anchor were cleaved simultaneously with the SiOMB groups using fluoride ions in water-pyridine, proving this method to be compatible with the integrity of the nucleopeptide fragments. The 3'-hydroxyl functions of ribonucleopeptide **41** were protected with TBDMS groups, which were also cleaved by the fluoride treatment. The N-termini of **39** and **41** were masked with the NPS group, which was cleaved using 2-mercaptopyridine.

Figure 13



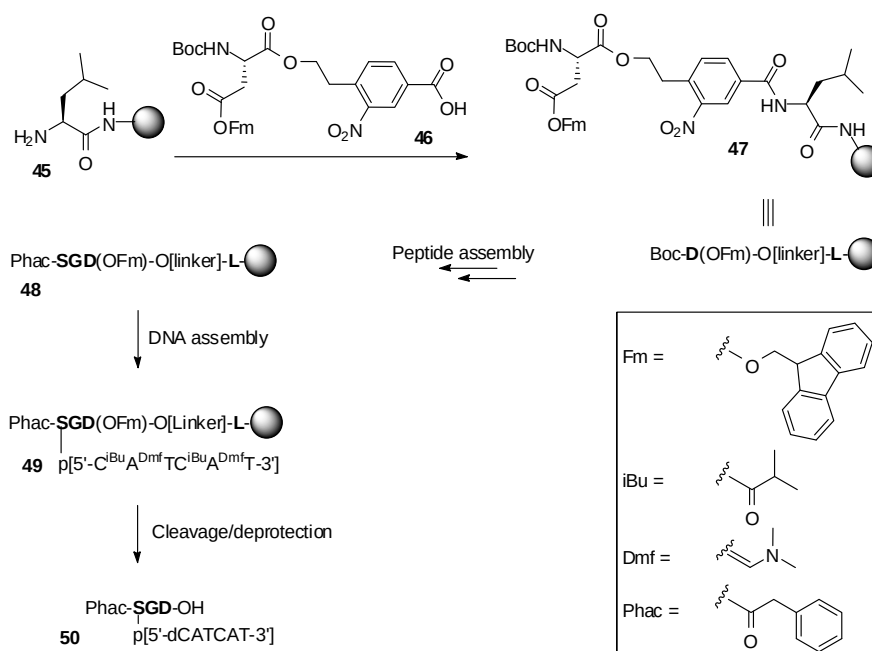
In the case of nucleopeptide **40** the peptide N-terminal amine as well as the acid functions were masked as nitro-benzyl carbamate/ester, which were removed under reductive conditions using sodium dithionite and sodium bicarbonate. Although the mild deblocking conditions needed for nitro-benzyl group removal prevented β -elimination, the deprotection was incomplete, complicating purification.

Peptidyl phosphoramidites allow access to larger nucleopeptide fragments than the solution based approaches, but there are also limitations to this technique. Low solubility of fully protected peptide fragments and the ensuing problems in preparation of peptidyl phosphoramidites⁷⁰, as well as decreased reactivity of such amidites are main disadvantages of this method.

3.2.2 Synthesis of nucleopeptides starting from resin-bound peptides

Reversing the order of nucleopeptide assembly by first preparing a resin bound peptide followed by repetitive coupling cycles of the appropriate nucleotide phosphoramidite to a hydroxyl function in the peptide backbone, might avoid the solubility and reactivity issues mentioned above. Extensive work by Grandas and co-workers has resulted in a wide repertoire of deoxynucleopeptides.⁷⁰⁻⁷⁸ An example of this approach is represented by the synthesis of **50** (see **Figure 14**).⁷¹ Leucyl amino methyl polystyrene resin **45** was coupled to α -(4-nitrophenyl)ethyl ester of Boc protected C-terminal aspartic acid β -fluorenylmethyl ester **46**.

Figure 14



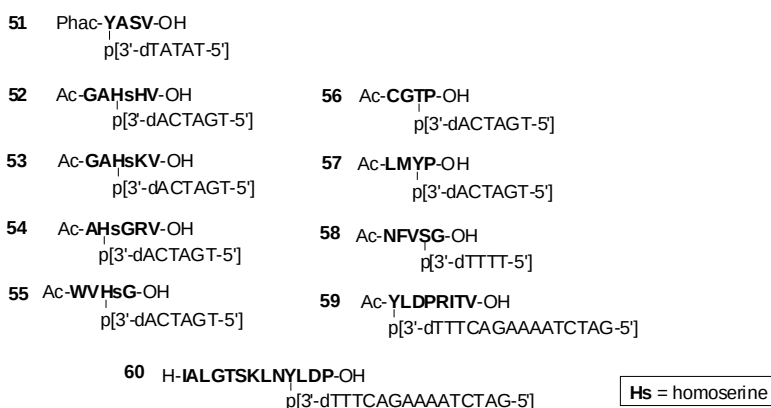
Subsequent sequential elongation of the peptide using Boc-based chemistry and final capping resulted in peptide **48**. DNA assembly on the free hydroxyl of serine was effected by repetitive coupling cycles with the appropriate 5'-phosphoramidite, capping, oxidation and liberation of the 3'-hydroxyl. In standard DNA synthesis 3'-phosphoramidites are used and the growing DNA chain is elongated from the 3'-end towards the 5'-end. In this particular case, 5'-phosphoramidites were prepared and used to elongate the growing oligonucleotide from the 5'-end towards the 3'-end. This non-standard approach proceeded without problems and final cleavage from the support by TBAF and removal of all nucleobase protective groups using ammonia in dioxane resulted in formation of the target nucleopeptide **50**.

Most of the results of the group of Grandas show examples of peptides linked to the 3'-end of DNA (in contrast to the example given in **Figure 14**, which is linked to the 5'-end). Such 3'-linked deoxyribonucleopeptides are for instance formed by topoisomerase I enzymes. Standard commercially available phosphoramidites can be used in the synthesis of such 3'-linked

nucleopeptides. Side chain protection in all naturally occurring amino acids used in this strategy are generally base-labile protective groups, of which an overview is described below.

The hydroxyl containing amino acid residues in nucleopeptide sequences, beside the linkage site for the oligonucleotide (see for instance **51**, **Figure 15**), were found to be efficiently protected as acetyl esters.⁷⁴

Figure 15



Incorporation of a histidine residue (see for instance **52**, **Figure 15**) was reported to proceed best by using the tosyl-protection group in peptides synthesis and removing this protection with HOBt prior to assembly of the nucleic acid part.⁷² It is not an option to use unprotected histidine during peptide synthesis since this causes epimerization at the α -carbon.

Lysine containing nucleopeptides (see for instance **53**, **Figure 15**) were reported to be best protected using either the fluorenylmethyloxycarbonyl (Fmoc) or trifluoroacetyl (TFAc) group on the ϵ -amine function.⁷⁶ However, the alkaline conditions needed to cleave the TFAc group limit its use.

Arginine containing nucleopeptides (see for instance **54**, **Figure 15**) were obtained by protecting the arginine side chain using Fmoc groups.⁷⁵ Due to the difficulties in the preparation of Fmoc-Arg(Fmoc)₂-OH it was decided in the synthesis of **59**, (see **Figure 15**) not to use protection on arginine. Ornithine side products were found indicating cleavage of the guanidine moiety in the arginine side chain. Other side reactions found at the guanidine group, when using unprotected arginine are acylation by amino acids in subsequent peptide coupling steps or nucleotide incorporation.⁷⁸

Tryptophane containing nucleopeptides (see for instance **55**, **Figure 15**) were obtained by protecting the indole ring with the formyl group, which could be efficiently removed by an ammonia treatment.⁷⁶

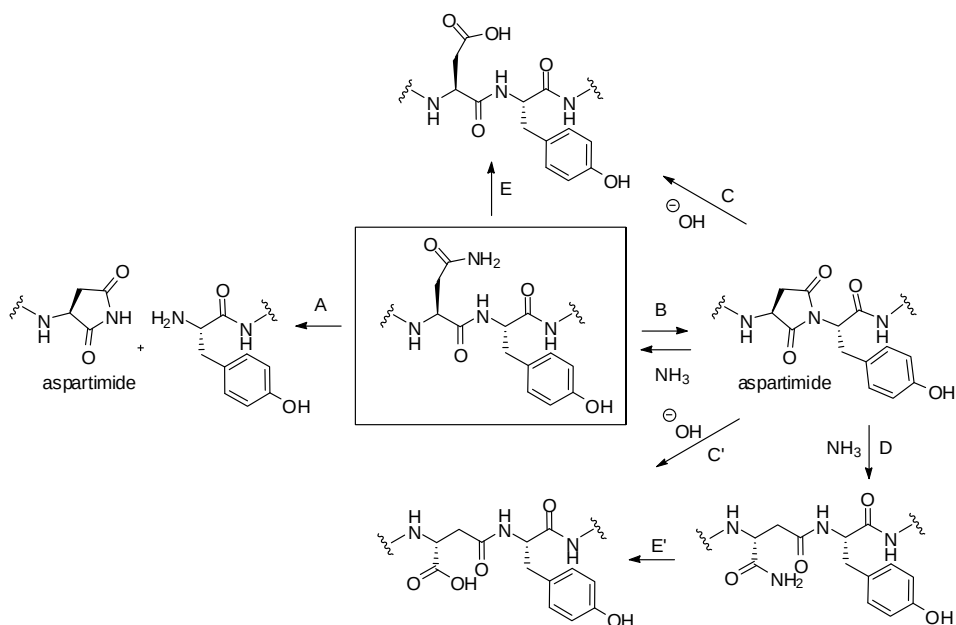
For cysteine containing peptides the standard protection of cysteine residues using the *tert*-butyl sulfonyl group is reported to give β -elimination in the final aqueous ammonia treatment.⁷⁹ The resulting dehydroalanine can be converted to serine or 2,3-diaminopropionic acid residues via conjugate addition. Prolonged heating of **56** (see **Figure 15**) in NH₃/dioxane indeed showed

complete conversion of cysteine to serine.⁷⁸ After cleavage of the S-*t*Bu group the free thiol is prone to disulfide formation or to alkylation by the acrylonitrile generated by cleavage of the cyanoethyl phosphate protective group. To avoid above mentioned problems the best method found so far is cleaving the S-*t*Bu using NH_3 / 1M DTT in dioxane for 8 hours at room temperature.

The synthesis of methionine containing nucleopeptides (see 57, **Figure 15**) can be accompanied by alkylation as a side reaction. Furthermore, aqueous iodine conditions used to oxidize phosphites to phosphotriesters might cause iodination and subsequent oxidation of the methionine residues. In peptide chemistry a well known method is to use methionine sulfoxide analogues, that are reduced after chain assembly, with for instance *N*-methylmercaptoacetamide. This strategy proved efficient for the synthesis of nucleopeptide 57.⁸⁰

Asparagine and glutamine containing nucleopeptides were synthesized using unprotected Asn or Gln building blocks.^{76,78} Although primary carboxamides are reported to react with phosphoramidites⁸¹ no significant amount of branched products were detected in the synthesis of an asparagine containing nucleopeptide (see 58, **Figure 15**). In the synthesis of 60 the aqueous ammonia treatment to remove acyl protection groups resulted in the formation of several side products.⁷⁸ Attack of the side chain carboxamide on the α -carbon of asparagine can result in the cleavage of the peptide backbone (pathway A, **Figure 16**), which indeed was observed.

Figure 16



On the other hand, attack of the NH in the amide bond between tyrosine and asparagine can result in the formation of an aspartimide (pathway B, **Figure 16**). Hydrolytic opening of this aspartimide can give rise to the aspartic acid containing peptide (pathway C) and to the β -peptide (pathway C'). Deamidation of the nucleopeptide can give rise to the aspartic acid containing peptide as well (pathway E). Opening of the aspartimide ring by ammonia can give rise to the β -peptide isomer (pathway D). Extensive characterization of the formed products however is difficult due to their size and complexity. The only performed analysis on the formed products is MALDI-TOF analysis which cannot cast a light on the identity of the isolated product and the side-reaction(s) that have occurred in the synthesis of **60**.

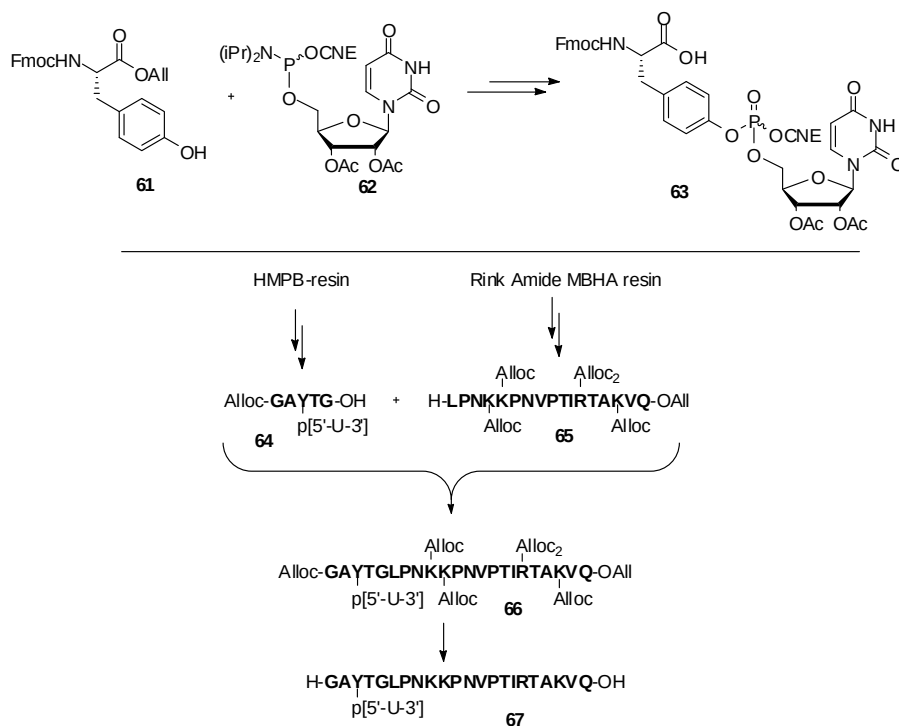
In studies on the incorporation of histidine, lysine, arginine and tryptophan in the nucleopeptides described above, it was opted to use an artificial linkage site between the peptide and deoxyoligonucleotide in order to circumvent problems such as β -elimination, associated with the labile nature of serine and threonine.⁷² The use of homoserine (Hs) instead of serine not only improved the stability under basic conditions⁷⁴ but also improves the stability of nucleopeptides against 3'-exonuclease activity in biological systems.⁷⁵

Although the above described solid phase based methodology is a powerful tool to create nucleopeptides, still the synthesis of each specific nucleopeptide is a challenge. Potential side reactions in both the peptide and nucleic acid parts and base lability are recurring problems, the severity of which varies from case to case.

3.2.3 Use of pre-uridylylated building blocks

Uridylylated VPg's are proposed to act as primers in viral replication and for this reason are interesting targets for nucleopeptide synthesis. The *mono*-uridylylated VPg (VPgpU) of polio virus type 1 (see **Figure 17**) was synthesized using a block coupling approach.⁸²

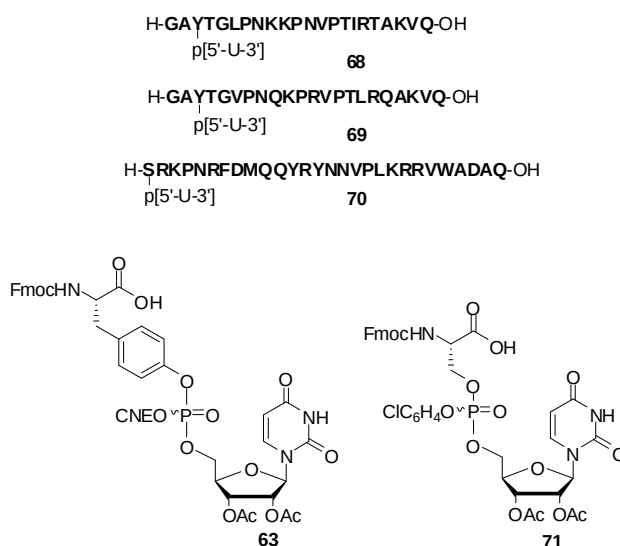
Figure 17



In this strategy pre-uridylylated tyrosine building **63** is of crucial importance. Coupling of protected tyrosine **61** and phosphoramidite **62**, followed by oxidation and protective group manipulations renders **63**. Pentameric peptide **64** was assembled on very acid-labile HMPB resin using pre-uridylylated tyrosine **65** via a Fmoc-based peptide synthesis protocol (in contrast to the Boc-based synthesis of nucleopeptides described in section 1.2.2). After release from the solid support and removal of all protective groups except the N-terminal Alloc, nucleopeptide **64** was coupled to partially protected heptadecamer **65**. This protected peptide was prepared separately on Rink amide MBHA resin. Condensation of **64** and **65** using the coupling mixture BOP/HOBt resulted in protected uridylylated VPgpU **66**, which was deprotected using Pd(0)-catalyzed hydrostannolysis of the Alloc groups and allyl ester to give the unprotected nucleopeptide **67**. The use of pre-uridylylated building blocks was also examined in the fully sequential solid support based synthesis of the VPgpU's of polio virus (**68**), coxsackie virus (**69**) and cowpea mosaic virus (**70**, see **Figure 18**).⁸³ Starting from Rink Amide MBHA resin the VPgpU's were constructed in a standard peptide solid phase approach using uridylylated

tyrosine **63** or uridylylated serine **71** (see **Figure 18**). Protective groups on the amino acid side chain functionalities used in this protocol were all acid labile. Acidic cleavage of protective groups on the peptide side chains and liberation of the peptides from the solid support followed by ammonia treatment to remove the acyl protection on the nucleoside as well as the N-terminal Fmoc gave tyrosine nucleopeptides **68** and **69** without problems. In the synthesis of serine nucleopeptide **70**, the phosphodiester moiety was protected using the 2-chlorophenyl protective group instead of the cyanoethyl, since the 2-chlorophenyl can be removed using fluoride ions in aqueous solvent (water/pyridine) to avoid strong base.

Figure 18

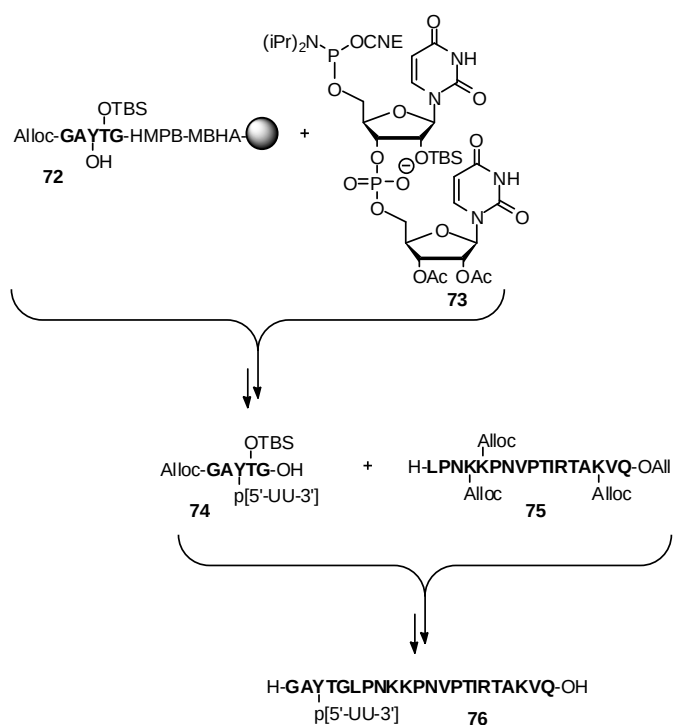


The advantage of this stepwise approach versus the block coupling method described earlier to obtain *mono*-uridylylated nucleoproteins is the ease of sequential synthesis and the use of protective groups that are cleaved using a simpler less complicated deprotection protocol. However this method does not allow the introduction of oligonucleic acid fragments since strong acid (95% trifluoroacetic acid) is used during the deprotection.

3.2.4 Solid phase convergent synthesis

Using a similar block coupling approach the double uridylylated nucleoproteins of coxsackie virus and polio virus type 1 (VPgpUpU's) were synthesized (as exemplified for polio virus VPgpUpU **76**, see **Figure 19**).⁸⁴ Pentapeptide **74** was prepared by reacting phosphoramidite **73** with resin bound **72**, followed by oxidation and cleavage of acyl protective groups and concomitant release from the solid support. Coupling of double uridylylated pentapeptide **74** to the appropriate, protected C-terminal heptadecamers of coxsackie virus or polio virus respectively followed by deprotection gives access to the respective VPgpUpU's. In a previously reported block coupling approach towards VPgpU nucleoproteins, the arginine residues were protected with Alloc groups.⁸²

Figure 19



During solid phase peptide synthesis however, the δ -Alloc group on arginine was found to be partially cleaved causing side reactions to occur and thereby complicating the purification process.⁸⁴ For this reason the arginine residue in the heptadecamer fragment en route to **75** was protected with the acid labile Pbf protective group which was removed prior to the block coupling step.

Outline of this thesis

Chapter 2 of this thesis describes the properties of a phosphoramidite equipped with two *p*-methoxybenzyl groups. Arbuzov like reactions allow the preparation of phosphates, phosphoramidates, phosphorothioates, pyrophosphates and phosphofluoridates in a one-pot procedure.

Chapter 3 describes the synthesis of pre-ribosylated glutamine and asparagine building blocks and their incorporation in peptides, that are natural sites of ADP-ribosylation. The methodology described in **Chapter 2** to create phosphates and pyrophosphates is applied in the synthesis of *mono*-ADP-ribosylated peptides both in solution and solid support based approaches.

Chapter 4 deals with the ribosylation of adenosine at the 2'-OH in an α -selective fashion. This glycosylated nucleoside is the repetitive unit of *poly*-ADP-ribose. The introduction of an orthogonal protective group pattern is described that would allow the use of this building block in stepwise *poly*-ADP-ribosylated peptide synthesis.

In **Chapter 5** a general method to obtain five picornavirus VPgpU's is reported using a pre-uridylylated tyrosine building block.

In **Chapter 6** the method from **Chapter 5** is extended to generate the other three canonical pre-nucleotidylated tyrosine building blocks. These are then used to generate the corresponding nucleotidylated polio virus VPg proteins.

Chapter 7 describes a linear solid support based synthesis of the nucleopeptide derived from the N-terminal region of picornaviral VPg. This approach demonstrates RNA synthesis using 5'-phosphoramidites on solid phase bound peptides.

References and Notes

- ¹ Merrifield, R. B.; *J. Am. Chem. Soc.*, **1963**, *85*, 2149 – 2154.
- ² Merrifield, B.; *Science*, **1986**, *232*, 341 – 347.
- ³ Caruthers, M. H.; *Science*, **1985**, *230*, 281 – 285.
- ⁴ Beaucage, S. L.; Iyer, R. P.; *Tetrahedron*, **1993**, *49*, 10441 – 10488.
- ⁵ Zubin, E. M.; Romanova, E. A.; Oretskaya, T. S.; *Russian Chem. Rev.*, **2002**, *71*, 239 – 264. 731.
- ⁶ Tate, E. W.; *J. Chem. Biol.*, **2008**, *1*, 17 – 26.
- ⁷ Collins, M. O.; Yu, L.; Choudhary, J. S.; *Proteomics*, **2007**, *7*, 2751 – 2768.
- ⁸ Mischerikow, N.; Heck, A. J. R.; *Proteomics*, **2011**, *11*, 571 – 589.
- ⁹ Welchman, R. L.; Gordon, C.; Mayer, R. J.; *Nat. Rev. Mol. Cell Biol.*, **2005**, *6*, 599 – 609.
- ¹⁰ Paleari, R.; Paglietti, E.; Mosca, A.; Mortarino, M.; Maccioni, L.; Satta, S.; Cao, A.; Galanello, *Clin. Chem.*, **1999**, *45*, 21 – 28.
- ¹¹ Corda, D.; Di Girolamo, M.; *EMBO J.*, **2003**, *22*, 1953 – 1958.
- ¹² Hassa, P. O.; Haenni, S. S.; Elser, M.; Hottiger, M. O.; *Microbiol. Mol. Biol. Rev.*, **2006**, *70*, 789 – 829.
- ¹³ Okana, S.; Lan, L.; Caldecott, K. W.; Mori, T.; Yasui, A.; *Mol. Cell. Biol.*, **2003**, 3974 – 3981.
- ¹⁴ Oliver, J.; Menissier-de Murcia, J.; Nacci, C.; Decker, P.; Andreantsitohaina, R.; Muller, S.; de la Rubia, G.; Stoclet, J. C.; de Murcia, G.; *EMBO J.*, **1999**, *18*, 4446 – 4454.
- ¹⁵ Scheuring, J.; Schramm, V. L.; *J. Am. Chem. Soc.*, **1995**, *117*, 12653 – 12654.
- ¹⁶ Scheuring, J.; Schramm, V. L.; *Biochemistry*, **1997**, *36*, 8215 – 8223.
- ¹⁷ Scheuring, J.; Berti, P. J.; Schramm, V. L.; *Biochemistry*, **1998**, *37*, 2748 – 2758.
- ¹⁸ Parikh, S. L.; Schramm, V. L.; *Biochemistry*, **2004**, *43*, 1204 – 1212.
- ¹⁹ Moyle, P. M.; Muir, T. W.; *J. Am. Chem. Soc.*, **2010**, *132*, 15878 – 15880.
- ²⁰ Comstock, L. R.; Denu, J. M.; *Org. Biomol. Chem.*, **2007**, *5*, 3087 – 3091.
- ²¹ Szczepankiewicz, B. G.; Koppetsch, K. J.; Perni, R. B.; *J. Org. Chem.*, **2011**, *76*, 6465 – 6474.
- ²² Zhou, G.; Parikh, S. L.; Tyler, P. C.; Evans, G. B.; Furneaux, R. H.; Zubkova, O. V.; Benjes, P. A.; Schramm, V. L.; *J. Am. Chem. Soc.*, **2004**, *126*, 5690 – 5698.
- ²³ Zhang, J.; Snyder, S. H.; *Biochemistry*, **1993**, *32*, 2228 – 2233.
- ²⁴ Jiang, H.; Kim, J. H.; Frizzell, K. M.; Krause, W. L.; Lin, H.; *J. Am. Chem. Soc.*, **2010**, *132*, 9363 – 9372.
- ²⁵ Yarbrough, M. L.; Orth, K.; *Nat. Chem. Biol.*, **2009**, *5*, 378 – 379.
- ²⁶ Itzen, A.; Blankenfeldt, W.; Goody, R. S.; *Trends in Biochemical Sciences*, **2011**, *36*, 221 – 228.

- ²⁷ Yarbrough, M. L.; Li, Y.; Kinch, L. N.; Grishin, N. V.; Ball, H. L.; Orth, K.; *Science*, **2009**, 323, 269 – 272.
- ²⁸ Worby, C. A.; Mattoo, S.; Kruger, R. P.; Corbeil, L. B.; Koller, A.; Mendez, J. C.; Zekarias, B.; Lazar, D.; Dixon, J. E.; *Mol. Cell*, **2009**, 34, 93 – 103.
- ²⁹ Muller, M. P.; Peters, H.; Blumer, J.; Blankenfeldt, W.; Goody, R. S.; Itzen, A.; *Science*, **2010**, 329, 946 – 949.
- ³⁰ Al-Eryani, R. A.; Li, Y.; Ball, H. L.; *Tetrahedron Lett.*, **2010**, 51, 1730 – 1731.
- ³¹ Kuyl-Yeheskiely, E.; Tromp, C. M.; Schaeffer, A. H.; van der Marel, G. A.; van Boom, J. H.; *Nucleic Acids Res.*, **1987**, 15, 1807 – 1818.
- ³² Albers, M. F.; van Vliet, B.; Hedberg, C.; *Org. Lett.*, **2011**, 13, 6014 – 6017.
- ³³ Smit, C.; Blumer, J.; Eerland, M. F.; Albers, M. F.; Muller, M. P.; Goody, R. S.; Itzen, A.; Hedberg, C.; *Angew. Chem. Int. Ed.*, **2011**, 50, 9200 – 9204.
- ³⁴ Guijarro, J. I.; Gonzales-Pastor, J. E.; Baleux, F.; San Millan, J. L.; Castilla, M.; Rico, M.; Moreno, F.; Delepiepierre, M.; *J. Biol. Chem.*, **1995**, 270, 23520 – 23532.
- ³⁵ Agarwal, V.; Metlitskaya, A.; Severinov, K.; Nair, S. K.; *J. Biol. Chem.*, **2011**, 286, 21295 – 21303.
- ³⁶ Roush, R. F.; Nolan, E. M.; Lohr, F.; Walsh, C. T.; *J. Am. Chem. Soc.*, **1998**, 120, 3603 – 3609.
- ³⁷ Matsunaga, S.; Takahashi, N.; Fusetani, N.; *Pure Appl. Chem.*, **2009**, 81, 1001 – 1008.
- ³⁸ Phillips, D. R.; Uramoto, M.; Isono, K.; Mc. Closkey, J. A.; *J. Org. Chem.*, **1993**, 58, 854 – 859.
- ³⁹ Moriguchi, T.; Asai, N.; Okada, K.; Seio, K.; Sasaki, T.; Sekine, M.; *J. Org. Chem.*, **2002**, 67, 3290 – 3300.
- ⁴⁰ Schein, C. H.; Volk, D. E.; Oezguen, N.; Paul, A.; *Proteins*, **2006**, 63, 719 – 726.
- ⁴¹ Steil, B. P.; Barton, D. J.; *Virus Res.*, **2009**, 139, 240 – 252.
- ⁴² Desiderio, S. V.; Kelly Jr., T. J.; *J. Mol. Biol.*, **1981**, 145, 319 – 337.
- ⁴³ Pronk, R.; Stuiver, M. H.; van der Vliet, P. C.; *Chromosoma*, **1992**, 102, S39 – S45.
- ⁴⁴ Smart, J. E.; Stillman, J. W.; *J. Mol. Biol.*, **1982**, 257, 13499 – 13506.
- ⁴⁵ Hermoso, J. M.; Salas, M.; *Proc. Natl. Acad. Sci. USA*, **1980**, 77, 6425 – 6428.
- ⁴⁶ Salas, M.; Mellado, R. P.; Vinuela, E.; *J. Mol. Biol.*, **1978**, 119, 269 – 291 .
- ⁴⁷ van Mansfeld, A. D. M.; van Teeffelen, H. A. A. M.; Baas, P. D.; Veeneman, G. H.; van Boom, J. H.; Jansz, H. S.; *FEBS Lett.*, **1984**, 173, 351 – 356.
- ⁴⁸ Wang, J. C.; *Annu. Rev. Biochem.*, **1996**, 65, 635 – 692.
- ⁴⁹ Connelly, J. C.; Leach, D. R. F.; *Mol. Cell*, **2004**, 13, 307 – 316.
- ⁵⁰ Chen, A. Y.; Liu, L. F.; *Annu. Rev. Pharmacol. Toxicol.*, **1994**, 34, 191 – 218.

- ⁵¹ Fontouro, B. M.; Atienza, C. A.; Sorokina, E. A.; Morimoto, T.; Carroll, R. B.; *Mol. Cel. Biol.*, **1997**, *17*, 3146 – 3154.
- ⁵² Samad, A.; Carroll, R. B.; *Mol. Cel. Biol.*, **1991**, *11*, 1598 – 1606.
- ⁵³ Kuyl-Yeheskiely, E.; van der Klein, P. A. M.; Visser, G. M.; van der Marel, G. A.; van Boom, J. H.; *Recl. Trav. Chim. Pays-Bas.*, **1986**, *105*, 67 – 70.
- ⁵⁴ Schattenkerk, C.; Wreesman, C. T. J.; de Graaf, M. J.; van der Marel, G. A.; van Boom, J. H.; *Tetrahedron Lett.*, **1984**, *25*, 5197 – 5200.
- ⁵⁵ Reese, C. B.; Zard, L.; *Nucleic Acids Res.*, **1981**, *9*, 4611 – 4626.
- ⁵⁶ Beaucage, S. L.; Iyear, R. P.; *Tetrahedron*, **1992**, *48*, 2223 – 2311.
- ⁵⁷ Kuyl-Yeheskiely, E.; Tromp, C. M.; Lefeber, A. W. M.; van der Marel, G. A.; van Boom, J. H.; *Tetrahedron*, **1988**, *20*, 6515 – 6523.
- ⁵⁸ Kuyl-Yeheskiely, E.; Dreef-Tromp, C. M.; Geluk, A.; van der Marel, G. A.; van Boom, J. H.; *Nucleic Acids Res.*, **1989**, *17*, 2897 – 2905.
- ⁵⁹ Kuyl-Yeheskiely, E.; Tromp, C. M.; Geluk, A.; van der Marel, G. A.; van Boom, J. H.; *Recl. Trav. Chim. Pays-Bas.*, **1988**, *107*, 567 – 569.
- ⁶⁰ Hotoda, H.; Ueno, Y.; Sekine, M.; Hata, T.; *Tetrahedron Lett.*, **1989**, *30*, 2117 – 2120.
- ⁶¹ Ueno, Y.; Saito, R.; Hata, T.; *Tetrahedron Lett.*, **1991**, *32*, 1347 – 1350.
- ⁶² Ueno, Y.; Saito, R.; Hata, T.; *Nucleic Acids Res.*, **1993**, *21*, 4451 – 4457.
- ⁶³ Letsinger, R. L.; Ogilvie, K. K.; *J. Am. Chem. Soc.*, **1967**, *89*, 4801 – 4803.
- ⁶⁴ Beaucage, S. L.; Caruthers, M. H.; *Tetrahedron Lett.*, **1981**, *22*, 1859 – 1862.
- ⁶⁵ Jeyaraj, D. A.; Prinz, H.; Waldmann, H.; *Chem. Eur. J.*, **2002**, *8*, 1879 – 1887.
- ⁶⁶ Dreef-Tromp, C. M.; van Dam, E. M. A.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H.; *Nucleic Acids Res.*, **1990**, *18*, 6491 – 6495.
- ⁶⁷ Dreef-Tromp, C. M.; van den Elst, H.; van den Boogaart, J. E.; van der Marel, G. A.; van Boom, J. H.; *Nucleic Acids Res.*, **1992**, *20*, 2435 – 2439.
- ⁶⁸ Dreef-Tromp, C. M.; van der Maarel, J. E.; van den Elst, H.; van den Boogaart, J. E.; van der Marel, G. A.; van Boom, J. H.; *Nucleic Acids Res.*, **1992**, *20*, 4015 – 4020.
- ⁶⁹ Dreef-Tromp, C. M.; Hoogerhout, P.; van der Marel, G. A.; van Boom, J. H.; *Tetrahedron Lett.*, **1990**, *31*, 427 – 430.
- ⁷⁰ Robles, J.; Pedroso, E.; Grandas, A.; *Nucleic Acids Res.*, **1995**, *23*, 4151 – 4161.
- ⁷¹ Robles, J.; Pedroso, E.; Grandas, A.; *Tetrahedron Lett.*, **1994**, *35*, 4449 – 4452.
- ⁷² Beltran, M.; Pedroso, E.; Grandas, A.; *Tetrahedron Lett.*, **1998**, *39*, 4115 – 4118.
- ⁷³ Robles, J.; Pedroso, E.; Grandas, A.; *J. Org. Chem.*, **1994**, *59*, 2482 – 2486.

- ⁷⁴ Beltran, M.; Perez, Y.; Robles, J.; Pedroso, E.; Grandas, A.; *Nucl. Nucl.*, **1997**, *16*, 1487 – 1488.
- ⁷⁵ Robles, J.; Maseda, M.; Beltran, M.; Concernau, M.; Pedroso, E.; Grandas, A.; *Bioconjugate Chem.*, **1997**, *8*, 785 – 788.
- ⁷⁶ Marchan, V.; Debethune, L.; Beltran, M.; Robles, J.; Travesset, I.; Fabregas, G.; Pedroso, E.; Grandas, A.; *Nucl. Nucl.*, **1999**, *18*, 1493 – 1494.
- ⁷⁷ Robles, J.; Beltran, M.; Marchan, V.; Perez, Y.; Travesset, I.; Pedroso, E.; Grandas, A.; *Tetrahedron*, **1999**, *55*, 13251 – 13264.
- ⁷⁸ Debethune, L.; Marchan, V.; Beltran, M.; Fabregas, G.; Pedroso, E.; Grandas, A.; *Tetrahedron*, **2002**, *58*, 6965 – 6978.
- ⁷⁹ Schwope, I.; Bledzinski, C. F.; Richert, C. J.; *J. Org. Chem.*, **1999**, *64*, 4749 – 4761.
- ⁸⁰ Marchan, V.; Rodriguez-Tanty, G.; Estrada, M.; Pedroso, E.; Grandas, E.; *Eur. J. Org. Chem.*, **2000**, 2495 – 2500.
- ⁸¹ Robles, J.; Pedroso, E.; Grandas, A.; *J. Org. Chem.*, **1995**, *60*, 4856 – 4861.
- ⁸² Filippov, D. V.; Kuyl-Yeheskiely, E.; van der Marel, G. A.; Tesser, G. I.; van Boom, J. H.; *Tetrahedron Lett.*, **1998**, *39*, 3597 – 3600.
- ⁸³ Kriek, N. M. A. J.; Filippov, D. V.; van den Elst, H.; Meeuwenoord, N. J.; Tesser, G. I.; van Boom, J. H.; van der Marel, G. A.; *Tetrahedron*, **2003**, *59*, 1589 – 1597.
- ⁸⁴ Kriek, N. M. A. J.; Meeuwenoord, N. J.; van den Elst, H.; Heus, H. A.; van der Marel, G. A.; Filippov, D. V.; *Org. Biomol. Chem.*, **2006**, *59*, 3576 – 3586.

Chapter 2

A Versatile One-Pot Procedure to Phosphate Monoesters and Pyrophosphates¹

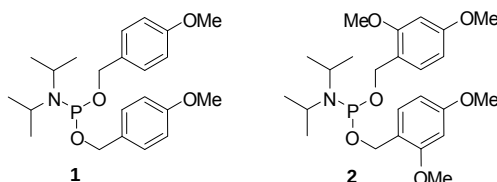
Introduction

Phosphorus containing functional groups are not only essential elements of nucleic acids but are also present in other naturally occurring compounds such as proteins², carbohydrates³ and lipids.⁴ Among these biomolecules phosphomono- and diesters^{2,3}, pyro- and triphosphates^{5,6}, phosphoramidates⁷ and C-phosphonates⁸ can be discerned. The pivotal role that these phosphoric acid derivatives play in a broad palette of biological processes has stimulated research into the synthesis of these classes of compounds.^{9,10} In addition, the pharmacological potential of phosphate ester analogues¹¹⁻¹³ led to the development of artificial phosphate derivatives such as phosphorothioates¹⁴ and phosphofluoridates.¹⁵ Initially developed to meet the needs of automated nucleic acid synthesis, the phosphoramidite approach is one of the most applied synthetic methodologies towards phosphate ester containing molecules.¹⁶⁻¹⁸ Introduction of phosphate esters is generally attained with the aid of amidites provided with protecting groups, which are often base-labile.^{19,20} Upon activation with a weakly acidic azole the amidite reagent reacts with an alcohol to give a phosphite triester. Subsequent oxidation to the phosphotriester and deprotection using alkaline conditions furnishes the target phosphodi- or monoester.²¹

As part of ongoing efforts²²⁻²⁴ aiming at the synthesis of naturally occurring phosphates and derivatives thereof via the phosphoramidite approach the phosphitylation properties of methoxybenzyl-*N,N*-diisopropylphosphoramidites **1** and **2** (**Figure 1**) were explored. One of the interesting envisioned properties of these substituted benzyl groups is their stability towards mild alkaline conditions and their lability towards mild acidic conditions. These reagents provide attractive entries to phosphomonoesters, since phosphitylation of the alcohol of choice with di-(*p*-methoxybenzyl)-*N,N*-diisopropyl phosphoramidite (**1**) followed by oxidation and

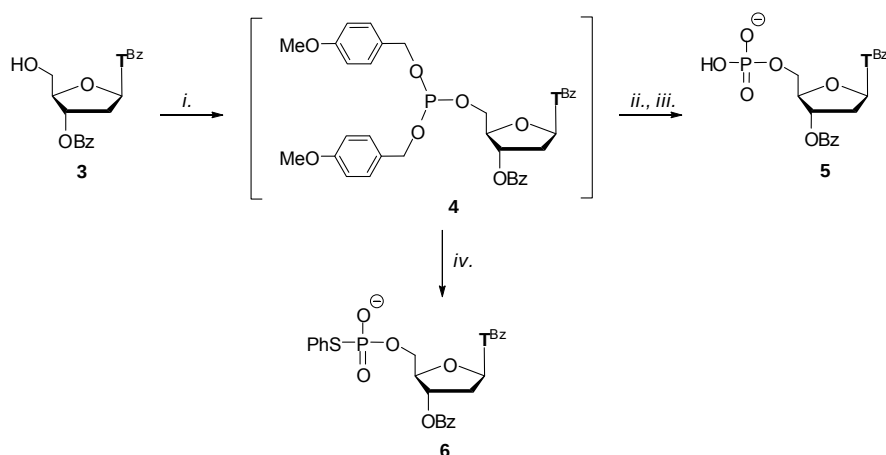
mild acidic removal of both *p*-methoxybenzyl groups would readily provide the corresponding phosphomonoester. Reagent **2** (**Figure 1**) equipped with *o,p*-dimethoxybenzyl groups would conceivably lead to a phosphotriester moiety convertible to a phosphomonoester under even milder conditions. An additional useful feature of reagents **1** and **2** is the potential of methoxybenzyl substituents to engage in Arbuzov-type reactions.²⁵ This would open the way to the synthesis of modified phosphates.

Figure 1

Methoxybenzyl phosphoramidites **1** and **2**

Results and Discussion

Phosphoramidites **1** and **2** were synthesized in good yields via a procedure reported for a related compound²⁶, comprising treatment of *p*-methoxybenzylalcohol and 2,4-dimethoxybenzylalcohol, respectively, with *N,N*-diisopropylaminodichlorophosphine in the presence of triethylamine. Reaction of reagent **1** (1.2 eq.) with thymidine **3** (1.0 eq.) using dicyanoimidazole (DCI) (1.2 eq.) as activator (**Scheme 1**) gave, as monitored by ³¹P-NMR spectroscopy (δ : 141.8 ppm), formation of phosphite triester **4** as the major product together with some minor phosphonates (δ : 14.1, 9.1 ppm), caused by the hydrolysis of excess reagent **1**. In the subsequent steps **4** was used without isolation, since silica gel column chromatography yielded only 29% of **4**, presumably due to its instability. Addition of anhydrous *t*BuO₂H to oxidize phosphite **4**, followed by *p*-methoxybenzyl deprotection using 3% TFA in DCM and chromatographic purification yielded phosphate **5** (³¹P-NMR, δ : 1.6 ppm) in 78% yield. In contrast, phosphitylation of **3** using agent **2** under identical conditions failed to yield the intended phosphite triester but instead produced unidentified C- and H-phosphonates in a 1:1 ratio (³¹P-NMR, δ : 31.9, 14.5 ppm).

Scheme 1. Synthesis of phosphate **5** and phosphorothioate **6**

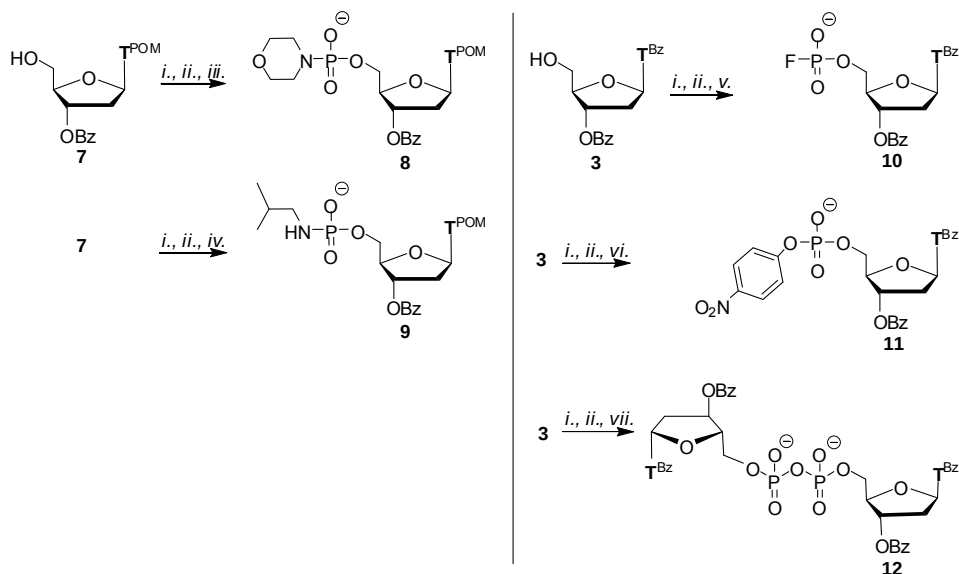
Reagents and conditions; *i.* **1**, DCI, MeCN, 0 °C, 20 min, *ii.* 5.5 M *t*BuO₂H, 2 h, r.t., *iii.* 3% TFA in DCM, 5 min, r.t., 78% (over two steps), *iv.* (PhS)₂, MeCN, 16 h, r.t., 79%.

Having demonstrated the effectiveness of phosphitylating agent **1** to introduce phosphate monoesters under mild acidic (deprotection) conditions the susceptibility of di-(*p*-methoxybenzyl) protected phosphite triester **4** to undergo an Arbuzov-type reaction was explored. Thymidine **3** was converted into phosphite triester **4** as described above, but now the reaction mixture was treated with diphenyldisulfide at the time the transformation of **3** to **4** was complete (as revealed by NMR) leading to phosphorothioate diester **6** in 79% yield (**Scheme 1**). Surprisingly complete in situ cleavage of both *p*-methoxybenzyl protecting groups was observed when the reaction was monitored by ³¹P-NMR spectroscopy, while it was expected to observe the corresponding triester still carrying one *p*-methoxybenzyl group. This one-pot procedure thus allows an immediate access to S-phenyl phosphorothioates, known precursors²⁷ in the synthesis of pyrophosphates. The question arose whether these results could be extended to other Arbuzov-type reactions using this one-pot procedure. The use of iodine as the oxidation reagent for **4** would allow introduction of a range of nucleophilic atoms at the phosphorus center via the putative phosphoryliodide.²⁸

An interesting class of compounds are the phosphoromorpholidates, the most well-known intermediates en route to pyrophosphates.²⁹ The solution phase synthesis of thymidine phosphoromorpholidate derivative **8** (**Scheme 2**) was undertaken by adaptation of an iodine mediated procedure to create phosphoromorpholidates on a solid support.³⁰ Thymidine **7** was phosphitylated with **1** in presence of DCI under the same conditions as described above and the obtained di-(*p*-methoxybenzyl) phosphite triester was oxidized with iodine in pyridine and subsequently treated with morpholine for 1 hour. After silica gel column chromatography morpholidate **8** was isolated in 86% yield.³¹ Following an analogous procedure using *iso*-butylamine instead of morpholine as nucleophile led to the clean formation of phosphoramidate **9** (79%).³¹

To broaden the scope of the one-pot procedure the use of non-nitrogen nucleophiles was investigated. Addition of DAST as fluoride source to the reaction mixture after iodine-mediated oxidation resulted in the formation of phosphofluoridate **10** (80%). Similarly, addition of the sodium salt of *p*-nitrophenol to the oxidized reaction mixture resulted in formation of the corresponding phosphodiester **11** (44%). Surprisingly, addition of water did not result in the formation of the expected phosphate monoester but furnished symmetric pyrophosphate **12** (78%).

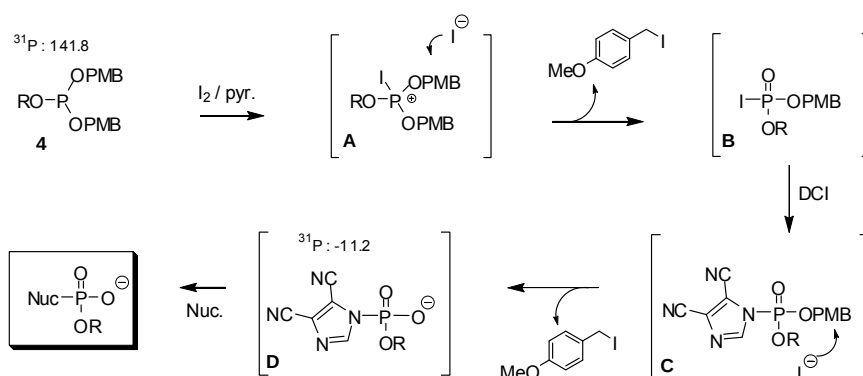
Scheme 2. Synthesis of phosphate monoesters **8** – **11** and symmetric pyrophosphate **12**



Reagents and conditions: *i.* **1**, DCI, MeCN, 0 °C, 20 min, *ii.* I₂, pyridine, 30 min, r.t., *iii.* morpholine, 1 h, r.t. (86% over three steps), *iv.* *iso*-butylamine, 1 h, r.t. (79% over three steps), *v.* DAST, 16 h, r.t. (80% over three steps), *vi.* *p*-nitrophenolate sodium salt, 16 h, r.t. (44% over three steps), *vii.* H₂O, 16 h, r.t. (78% over three steps).

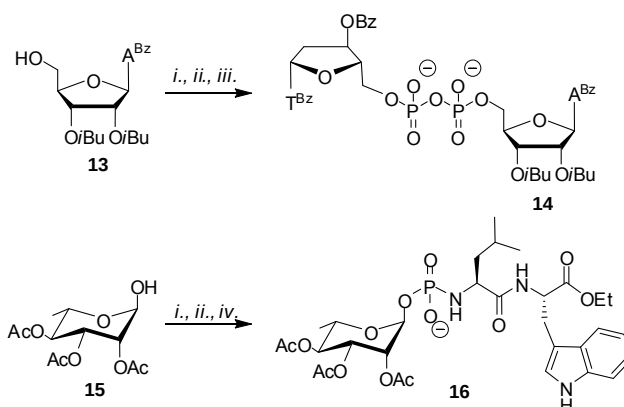
This unexpected formation of the pyrophosphate **12** upon addition of water as well as the fact that no special treatment is needed to remove the second *p*-methoxybenzyl group in the remaining reactions (**Scheme 2**) put forward the question on the identity of the reactive intermediates involved in the reaction sequence of this one-pot procedure. On the basis of the data that resulted from the analysis of selected reactions by ³¹P NMR spectroscopy the following mechanism is proposed (**Scheme 3**).

Scheme 3. Proposed reaction mechanism



Activation of amidite **1** (^{31}P -NMR, δ : 147.5) with DCI and subsequent coupling with alcohol **3** gave the phosphite triester **4** (^{31}P -NMR, δ : 141.8, see **Scheme 1**). Treatment of the phosphite triester **4** with iodine in pyridine (**Scheme 3**) would lead to phosphonium species **A** and the subsequent phosphoryliodide **B**, having a resonance around δ : -44, as described.²⁸ However, this resonance could not be detected, but instead a signal at δ : -11.2 was observed. This indicated that the putative phosphoryliodide **B** was transformed immediately into another intermediate with concomitant cleavage of the second *p*-methoxybenzyl group (PMB). This loss of the protective group can be explained by nucleophilic attack of iodide generated in step **C** on the benzylic position of the PMB group. On the basis of literature data that comparable phosphorazolidines have ^{31}P -NMR resonances at δ : -12.2 and -11.24 the resonance at δ : -11.2 was tentatively assigned to the dicyanoimidazolidine species **D**.^{29,32} Subsequent addition of a nucleophile would lead to displacement of dicyanoimidazole to give the corresponding phosphate derivative. The proposed dicyanoimidazolidine **D** as the reactive intermediate would explain not only the removal of both *p*-methoxybenzyl groups but also the formation of pyrophosphate **12**. Addition of water will partly and slowly hydrolyze dicyanoimidazolidine **D** to give the corresponding monophosphate that, in turn, reacts faster than water with **D** to give pyrophosphate (for example **12**, **Scheme 2**).

To further establish the applicability of this one pot procedure in the preparation of various phosphite derivatives, the syntheses of asymmetric pyrophosphate³³ **14** and phosphoramidon precursor³⁴ **16** (**Scheme 4**) were undertaken. Crude pyrophosphate **14** was obtained by phosphitylation of protected adenosine **13** with **1** under the influence of DCI, subsequent iodine/pyridine oxidation and, finally, addition of thymidine phosphate **5**. RP-HPLC and gel filtration yielded homogeneous pyrophosphate **14** in moderate yield (27%). Similarly phosphoramidon precursor **16** was synthesized in a moderate yield (19%) by phosphitylating β -rhamnose-triacetate **15** with **1**, iodine oxidation and addition of leucyltryptophan ethyl ester. Such modest yields, compared to synthesis of compounds **8** - **12** are probably caused by the losses in the purification process, as no major side reactions have been observed.

Scheme 4. Synthesis of asymmetric pyrophosphate **14** and phosphoramidon precursor **16**

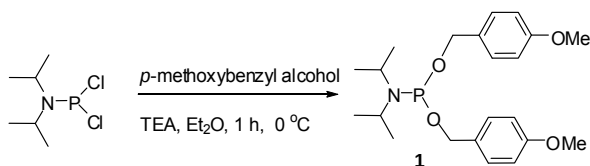
Reagents and conditions: *i.* **1**, DCI, MeCN, 0 °C, 20 min, *ii.* I₂, pyridine, 1 h, r.t., *iii.* **5** (Scheme 1), 16 h, r.t., 27% (over three steps), *iv.* H-Leu-Trp-OEt, 16 h, r.t. 19% (over three steps).

In conclusion, an efficient one-pot procedure to phosphoramidates, phosphorothioates, pyrophosphates and phosphofluoridates has been presented that comprises a sequence of reactions starting with phosphitylation of an alcohol with *di*-(*p*-methoxybenzyl)-*N,N*-diisopropylphosphoramidite **1**, followed by oxidation with iodine and treatment with a nucleophile of choice. The favourable properties of the *p*-methoxybenzyl groups allowed the use of reagent **1** to introduce phosphate monoesters under relatively mild acidic conditions. This feature as well as the potential to form asymmetric pyrophosphates will be further explored in the research as described in **Chapter 3**.

Experimental

General: All chemicals were used as received. Petroleum Ether with a boiling range of 40 - 60°C was used. THF and Et₂O were distilled over LiAlH₄ prior to use. DCM was distilled over CaH₂ prior to use. All other solvents used under anhydrous conditions were stored over molecular sieves (4Å) except for methanol which was stored over 3Å molecular sieves. Solvents used for workup and column chromatography were of technical grade and distilled before use. Unless stated otherwise, solvents were removed by rotary evaporation under reduced pressure at 40°C. Reactions were monitored by TLC-analysis using Merck 25 DC plastikfolien 60 F254 with detection by spraying with 20% H₂SO₄ in EtOH, (NH₄)₆Mo₇O₂₄·4H₂O (25 g/L) and (NH₄)₄Ce(SO₄)₄·2H₂O (10 g/L) in 10% sulfuric acid or by spraying with a solution of ninhydrin (3 g/L) in EtOH / AcOH (20/1 v/v), followed by charring at approx. 150°C. Column chromatography was performed using silica gel (0.04 – 0.063 mm, 60 Å, Screening Devices). For LC-MS analysis a JASCO HPLC-system (detection simultaneously at 214 and 254 nm) equipped with an analytical C18 column (4.6 mmD × 50 mmL, 3μ particle size) in combination with eluents A: H₂O, B: MeCN and C: 0.5% aq. TFA was used and coupled to a PE/SCIEX API 165 single quadrupole mass spectrometer (Perkin-Elmer). Alternatively a Thermo Finnigan LCQ Advantage MAX ion-trap mass spectrometer with an electrospray ion source coupled to Surveyor HPLC system (Thermo Finnigan) was used with the same analytical column. For reversed-phase HPLC purification of final compounds, an automated HPLC system supplied with a C18 column (10.0 mmD × 250 mmL, 5μ particle size) was used. High resolution mass spectra were recorded by direct injection (2 μL of a 2 μM solution in water/acetonitrile; 50/50; v/v and 0.1% formic acid) on a mass spectrometer (Thermo Finnigan LTQ Orbitrap) equipped with an electrospray ion source in positive mode (source voltage 3.5 kV, sheath gas flow 10, capillary temperature 250 °C) with resolution R = 60000 at m/z 400 (mass range m/z = 150-2000) and dioctylphthalate (m/z = 391.2842) as a “lock mass”. The high resolution mass spectrometer was calibrated prior to measurements with a calibration mixture (Thermo Finnigan). ¹H- and ¹³C-NMR spectra were measured on a Brüker AV-400 (400 MHz), a Brüker AV-500 (500 MHz) or a Brüker DMX-600 (600 MHz). Chemical shifts are given in ppm (δ) relative to TMS (0 ppm) and coupling constants are given in Hz. Phosphitylation reactions were monitored using ³¹P-NMR, by placing 400 μL of the crude reaction mixture in a NMR-tube. An acetone-cappillary filled with acetone-*d*₆ was used as external lock. Chemical shifts are given in ppm (δ) relative to 85% aq. H₃PO₄ (0 ppm). IR spectra of the neat compounds were recorded on a Shimadzu FTIR-8300 and are reported in cm⁻¹. Optical rotations are measured in CHCl₃ at a concentration of 10 mg/mL at 25 °C on a Propol automatic polarimeter (Sodium D-line, λ = 589 nm).

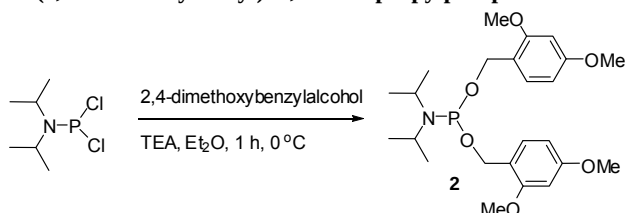
Di-(*p*-methoxybenzyl)-*N*, *N*-diisopropylphosphoramidite (**1**)



To a cooled (0 °C) solution of *N*, *N*-diisopropylphosphorodichloridite (2.4 g, 12.0 mmol) in Et₂O (50 mL) was added dropwise a solution of *p*-methoxybenzyl alcohol (3.8 mL, 30 mmol) and TEA (5.6 mL, 40 mmol) in Et₂O (50 mL). After 1 hour, the suspension was filtered and washed with sat. aq. NaHCO₃. The filtrate was dried using MgSO₄, concentrated and purified using silica gel column chromatography PE:TEA (99:1 → 23:2, v/v) as eluent to afford **1** as a colourless oil (3.98 g, 82%). ¹H-NMR (400 MHz, CDCl₃), δ: 7.26 (d, *J* = 8.6 Hz, 4H, Arom. PMB), 6.85 (d, *J* = 8.7 Hz, 4H, Arom. PMB), 4.71 - 4.59 (m, 4H, CH₂ PMB), 3.78 (s, 6H, CH₃ PMB), 3.72 - 3.63 (m, 2H, CH *i*Pr), 1.20, 1.19 (2x s, 12H, CH₃ *i*Pr). ¹³C-NMR (100 MHz, CDCl₃), δ: 158.8, 131.7 (Cq Arom.), 128.5 (Arom.), 113.7 (Arom.), 65.1, 65.0 (CH₂

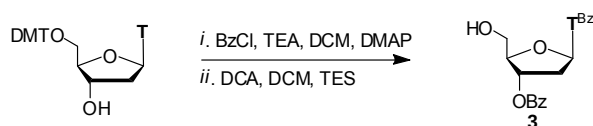
PMB), 55.2 (CH₃ PMB), 43.0, 42.9 (CH *i*Pr), 24.6, 24.5 (CH₃ *i*Pr). ³¹P-NMR (162 MHz, CDCl₃), δ: 147.52. IR: 2965, 1612, 1512, 1244, 1172, 1027, 967, 818, 729, 514. HRMS [C₂₂H₃₂NO₄P + H]⁺: calc. 406.2141, find. 406.2099.

Di-(2,4-dimethoxybenzyl)-*N*, *N*-diisopropylphosphoramidite (**2**)



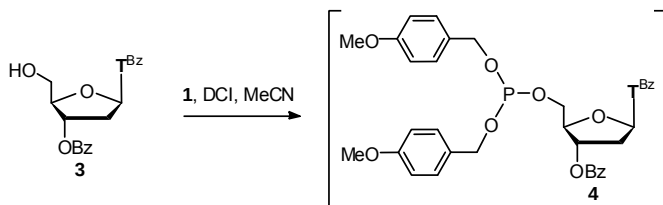
To a cooled (0 °C) solution of *N*, *N*-diisopropylphosphorodichloridite (0.5 g, 2.5 mmol) in Et₂O (10 mL) was added dropwise a solution of 2,4-dimethoxybenzylalcohol (0.9 g, 5.2 mmol) and TEA (1.2 mL, 8.3 mmol) in Et₂O (10 mL). The reaction was slowly warmed to room temperature and after 1 hour, the suspension was diluted in 40 mL pentane and filtered. The filtrate was concentrated and purified using silica gel column chromatography using Et₂O: PE:TEA as eluent (0:95:5 → 10:85:5, v/v) to afford **2** as a colourless oil (0.43 g, 0.9 mmol, 36%). ¹H-NMR (400 MHz, MeCN-*d*₃), δ: 7.25 (d, *J* = 8.3 Hz, 2H, Arom. DMB), 6.51 (d, *J* = 10.5, 2H, Arom. DMB), 6.48 (s, 2H, Arom. DMB), 4.73 - 4.54 (m, 4H, CH₂ DMB), 4.04 - 3.54 (2x s, 12H CH₃ DMB), 3.74 - 3.56 (m, 2H, CH *i*Pr), 1.20, 1.19 (2x s, 12H, CH₃ *i*Pr). ¹³C-NMR (100 MHz, MeCN-*d*₃), δ: 161.6, 159.1 (Cq Arom.), 130.6 (Arom.), 121.0 (Cq Arom.), 105.2, 99.1 (Arom.), 61.2, 61.0 (CH₂ DMB), 56.1, 56.0 (CH₃ DMB), 43.9, 43.8 (CH *i*Pr), 25.1, 25.0 (CH₃ *i*Pr). ³¹P-NMR (162 MHz, MeCN-*d*₃), δ: 148.59. HRMS [C₂₄H₃₆NO₆P + H]⁺: calc. 466.2353, find. 466.2342.

*N*³-Benzoyl-3'-*O*-benzoylthymidine (**3**)

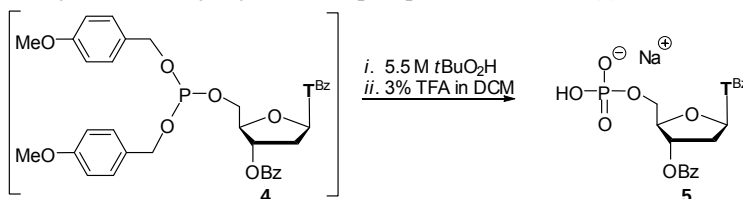


To a solution of 5'-(4',4-dimethoxytrityl)thymidine (5.45 g, 10 mmol) in DCM (50 mL) was added TEA (4.2 mL, 30 mmol), DMAP (0.24 g, 2 mmol) and benzoylchloride (3.0 mL, 26 mmol) after which the reaction mixture was stirred for 16 hours. The reaction mixture was concentrated and taken up in EtOAc followed by washing with sat. aq. NaHCO₃. The organic layer was dried (MgSO₄) and concentrated. The residue was dissolved in DCM (60 mL) and dichloroacetic acid (2.5 mL, 30 mmol) was added. After 15 min triethylsilane (16.1 mL, 100 mmol) was added and the reaction mixture was stirred at r.t. for 30 min. After washing with sat. aq. NaHCO₃, drying using MgSO₄ and concentration the residue was purified using silica gel column chromatography (0:100 → 10:90 MeOH/DCM v/v) to yield the title compound as a white foam (4.38 g, 9.7 mmol, 97%). ¹H-NMR (400 MHz, CDCl₃), δ: 8.03 (d, *J* = 1.1 Hz, 2H, Arom. Bz), 7.94 (d, *J* = 1.1 Hz, 2H, Arom. Bz), 7.81 (d, *J* = 1.1 Hz, 1H, H6), 7.74 - 7.54 (2x m, 2H, Arom. Bz), 7.46 (2x t, *J* = 10.8, 4.9 Hz, 4H, Arom. Bz), 6.40 (dd, *J* = 8.3, 6.0 Hz, 1H, H1'), 5.58 (dt, *J* = 5.9, 2.0 Hz, 1H, H3'), 4.24 (q, *J* = 2.3 Hz, 1H, H4'), 3.99 - 3.93 (m, 2H, H5'), 2.84 (bs, 1H, OH), 2.63 - 2.44 (m, 2H, H2'), 1.96 (d, *J* = 1.0 Hz, 3H, CH₃ T). ¹³C-NMR (100 MHz, CDCl₃), δ: 168.9 (CO Bz), 166.1, 162.8 (C4, C2), 149.4 (CO Bz), 136.1 (C6), 135.1, 133.6 (Arom. Bz), 131.4 (Cq Bz), 130.4, 129.6, 129.1, 128.5 (Arom. Bz), 111.3 (C5), 85.7 (C1'), 85.3 (C4'), 75.3 (C3'), 65.2 (C5'), 37.6 (C2'), 12.6 (CH₃T). IR: 3417, 1756, 1696, 1661, 1652 1644, 1447, 1262, 1229, 1096. HRMS [C₂₈H₃₀N₂O₇ + H]⁺: calc. 451.1499, find. 451.1499.

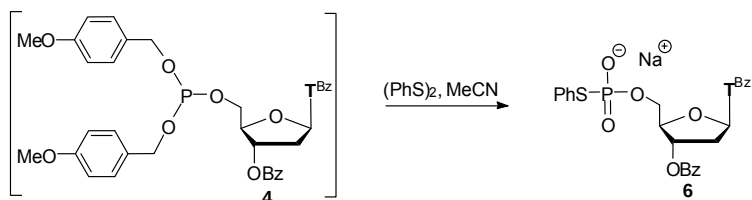
General procedure for synthesis of phosphite triester (4)



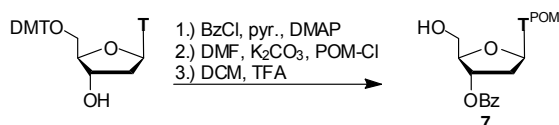
To a cooled (0 °C) solution of thymidine **3** (200 mg, 0.45 mmol) and **1** (0.20 mL, 0.50 mmol) in MeCN (3 mL) was slowly added a solution of DCI (60 mg, 0.5 mmol) in MeCN (1 mL). After 20 min. ^{31}P -NMR showed complete disappearance of **1** (148.5) and formation of triester **4** (^{31}P -NMR; 141.8). The solution thus obtained was used directly in the subsequent transformations. An analytical sample of **4** was concentrated and purified by column chromatography using a gradient of EA:PE:TEA 20:80:0.1 (v/v/v) to yield the title compound as a colourless oil (95 mg, 0.13 mmol, 29%). ^1H -NMR (400 MHz, CD_3CN), δ : 8.14 - 8.03 (m, 2H, Arom.), 8.03 - 7.94 (m, 2H, Arom.), 7.75 (s, 1H, H6), 7.67 - 7.64 (m, 1H, Arom.), 7.55 (2x t, J = 7.8 Hz, 4H, Arom.), 7.37 - 7.20 (m, 4H, Arom.), 6.91 (dd, J = 8.5, 1.2 Hz, 4H, Arom.), 6.36 (dd, J = 8.2, 6.0 Hz, 1H, H1'), 5.49 - 5.48 (m, 1H, H3'), 4.88 (d, J = 8.2 Hz, 4H, CH_2 PMB), 4.37 (q, J = 3.0 Hz, 1H, H4'), 4.23 - 4.03 (m, 2H, H5'), 3.86 - 3.71 (m, 6H, OMe PMB), 2.69 - 2.34 (m, 2H, H2'), 1.86 (s, 3H, CH_3 T). ^{13}C -NMR (100 MHz, $\text{MeCN}-d_3$), δ : 170.7 (CO Bz), 166.7, 163.8 (C4, C2), 150.5 (CO Bz), 137.1 (C6), 136.3, 134.4 (Arom. Bz), 132.6, 131.2 (Cq Bz), 130.8 (Cq PMB), 130.5, 130.40, 130.37, 129.6 (Arom. Bz), 114.8 (Arom. PMB), 111.8 (C5), 86.2 (C1'), 84.8, 84.7 (C4'), 76.2 (C3'), 65.6, 65.4, 65.1, 65.0 (CH_2 PMB), 63.3, 63.2 (C5'), 55.9 (CH_3 OMe), 38.2 (C2'), 12.7 (CH_3 T). ^{31}P -NMR (162 MHz, $\text{MeCN}-d_3$), δ : 141.87. IR: 2933, 1748, 170, 1659, 1612, 1515, 1449, 1246, 1175, 1096, 1026, 975, 713. HRMS [$\text{C}_{40}\text{H}_{39}\text{N}_2\text{O}_{11}\text{P} + \text{H}$] $^+$: calc. 755.2364, find. 755.2361.

 N^3 -Benzoyl-3'-O-benzoylthymidine-5'-phosphate sodium salt (5)

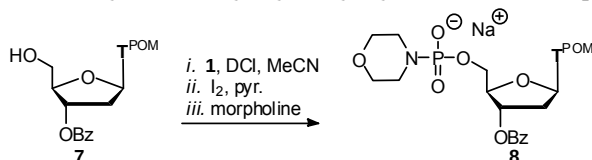
To a cooled (0 °C) solution of phosphite triester **4** (1.8 mmol) in MeCN (5 mL) was added $t\text{BuO}_2\text{H}$ (1.0 mL, 5.5 M in nonane) and the reaction mixture was stirred for 2 hours. The reaction mixture was concentrated and coevaporated three times with toluene. The residue was taken up in anhydrous DCM (20 mL) and TFA (0.6 mL) was added. After 5 min, toluene (5 mL) was added and the mixture was concentrated. The crude product was purified by silica gel column chromatography using DCM:MeOH as eluent (1:0 $\rightarrow$ 17:3 v/v) and treated with Amberlite Na^+ -resin to yield **5** as a sodium salt (0.91 g, 78 %). ^1H -NMR (400 MHz, $\text{MeOD}-d_4$), δ : 8.04 (s, 1H, H6), 8.03 (d, J = 7.5 Hz, 2H, Arom. Bz), 7.96 (d, J = 7.4 Hz, 2H, Arom. Bz), 7.81 - 7.37 (m, 6H, Arom. Bz), 6.56 - 6.31 (m, 1H, H1'), 5.68 (d, J = 5.2 Hz, 1H, H3'), 4.40 (s, 1H, H4'), 4.31 - 4.12 (m, 2H, H5'), 2.76 - 2.41 (m, 2H, H2'), 2.03 (s, 3H, CH_3 T). ^{13}C -NMR (100 MHz, $\text{MeOD}-d_4$), δ : 170.3 (CO Bz), 167.3, 164.9 (C2, C4), 151.1 (CO Bz), 138.4 (C6), 136.3, 134.6 (Arom. Bz), 133.0 (Cq, Bz), 131.5 (Arom. Bz), 130.6 (Cq Bz), 130.4, 129.7, 129.1 (Arom. Bz), 112.5 (C-5), 86.6 (C1'), 85.9 (C4'), 78.1 (C3'), 66.3 (C5'), 38.5 (C2'), 12.6 (CH_3 T). ^{31}P -NMR (162 MHz, $\text{CDCl}_3/\text{MeOD}-d_4$), δ : 1.60. IR: 3390, 1747, 1699, 1661, 1635, 1600, 1450, 1318, 1270, 1178, 1099, 1061, 916, 712. HRMS [$\text{C}_{24}\text{H}_{23}\text{N}_2\text{O}_{10}\text{P} + \text{H}$] $^+$: calc. 531.1163, find. 531.1160.

***N*³-Benzoyl-3'-*O*-benzoylthymidine-5'-(*S*-phenyl)-phosphorothioate sodium salt (**6**)**

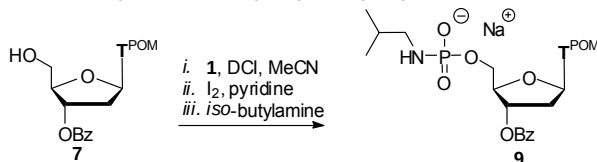
To a cooled (0 °C) solution of phosphite triester **4** (0.42 mmol) in MeCN (4 mL) was added diphenyldisulfide (0.22 g, 1.0 mmol) and the reaction mixture was stirred for an additional 16 hours. The reaction mixture was concentrated and the crude product was purified by silica gel column chromatography using DCM:MeOH as eluent (1:0 $\rightarrow$ 23:2 v/v) and treated with Amberlite Na⁺-resin to yield **6** (0.24 g, 79 %). ¹H-NMR (400 MHz, MeOD-*d*₄), δ : 8.12 (s, 1H, H₆), 8.04 (d, *J* = 7.8 Hz, 2H, Arom. Bz), 7.96 (d, *J* = 7.9 Hz, 2H, Arom. Bz), 7.79 - 7.24 (m, 6H, Arom. Bz), 7.20 - 7.12 (m, 5H, Arom. Ph), 6.44 (t, *J* = 7.6 Hz, 1H, H1'), 5.62 (d, *J* = 5.4 Hz, 1H, H3'), 4.30 (s, 1H, H4'), 4.03 - 3.88 (m, 2H, H5'), 2.73 - 2.44 (m, 2H, H2'), 2.00 (s, 3H, CH₃ T). ¹³C-NMR (100 MHz, MeOD-*d*₄), δ : 170.6 (CO Bz), 167.4, 166.2 (C₄, C₂), 151.0 (CO Bz), 138.3 (C₆), 136.3, 134.6 (CH Bz), 133.0 (C_q Bz), 131.5, 130.7, 130.4, 129.7 (Arom. Bz), 111.9 (C₅), 87.1 (C1'), 86.8 (C4'), 77.0 (C3'), 63.0 (C5'), 38.7 (C2'), 12.5 (CH₃, T). ³¹P-NMR (81 MHz, CDCl₃/MeOD-*d*₄), δ : 12.1. IR: 3406, 2925, 1760, 1700, 1661, 1644, 1447, 1267, 1096, 708. HRMS [C₃₀H₂₇N₂O₉PS + H]⁺: calc. 623.1247, fnd. 623.1263.

3'-*O*-Benzoyl-*N*³-Pivaloyloxymethylthymidine (7**)**

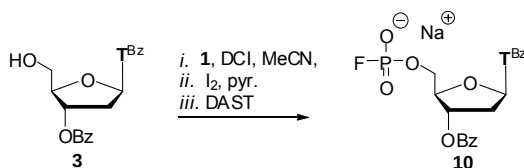
To a solution of 5'-(4',4'-dimethoxytrityl)thymidine (8.17 g, 15 mmol) in pyridine (75 mL) was added a catalytic amount of DMAP and benzoylchloride (2.3 mL, 20 mmol) followed by stirring for 2 hours at room temperature. After concentration the residue was taken up in EtOAc and washed with an aqueous CuSO₄ solution and the organic phase was dried (MgSO₄) and concentrated. The crude residue was dissolved in dry DMF (50 mL) and K₂CO₃ (10.3 g, 75 mmol) and pivaloyloxymethylchloride (5.5 mL, 37.5 mmol) were added and the mixture was stirred for 16 hours at room temperature. After concentration the residue was dissolved in DCM (50 mL) and TFA (1.5 mL) was added. The reaction mixture was quenched by the addition of MeOH (5 mL) and neutralized using TEA (2.8 mL, 20 mmol) after 1 hour. The reaction mixture was washed with water and sat. aq. NaCl, dried and concentrated. The crude residue was purified using silica gel column chromatography using a gradient of PE:EtOAc (80:20 $\rightarrow$ 40:60 v/v) to yield the title compound as a white foam (5.11 g, 11.1 mmol, 86%). ¹H-NMR (400 MHz, CDCl₃), δ : 7.95 (d, *J* = 8.0 Hz, 2H, Arom. Bz), 7.80 (s, 1H, H₆), 7.60 (t, *J* = 7.4 Hz, 1H, Arom. Bz), 7.46 (t, *J* = 7.7 Hz, 2H, Arom. Bz), 6.47 (dd, *J* = 8.4, 5.9 Hz, 1H, H1'), 6.08 - 5.84 (m, 2H, CH₂ POM), 5.69 - 5.57 (m, 1H, H3'), 4.30 (d, *J* = 2.0 Hz, 1H, H4'), 4.19 - 3.95 (m, 2H, H5'), 3.31 (bs, 1H, OH), 2.98 - 2.33 (m, 2H, H2'), 1.97 (s, 3H CH₃ T), 1.19 (s, 9H, *t*Bu). ¹³C-NMR (100 MHz, CDCl₃), δ : 177.6 (CO POM), 166.0, 162.5 (C₄, C₂), 150.3 (CO Bz), 135.2 (C₆), 133.7, 133.4 (Arom. Bz), 129.5 (C_q Bz), 129.8, 127.9 (Arom. Bz), 110.3 (C₅), 86.0 (C1'), 85.1 (C4'), 75.4 (C3'), 64.9 (CH₂ POM), 62.4 (C5'), 38.7 (C_q *t*Bu), 37.5 (C2'), 26.8 (*t*Bu), 13.0 (CH₃ T). IR: 3489, 2976, 1714, 1652, 1465, 1267, 1096, 711. HRMS [C₂₃H₂₈N₂O₈ + H]⁺: calc. 461.1918, fnd. 461.1916.

3'-O-Benzoyl-N³-Pivaloyloxymethylthymidine-5'-O-monophosphomorpholidate sodium salt (8)

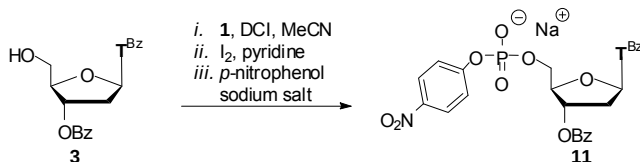
To a cooled (0 °C) solution of **7** (0.46 g, 1.0 mmol) and **1** (0.63 g, 1.3 mmol) in MeCN (15 mL) was slowly added a solution of DCI (0.15 g, 1.25 mmol) in MeCN (5 mL). After 20 min, a solution of iodine in pyridine (2 mL, 0.5 M) was added to the reaction mixture until the solution kept its brown colour. After 30 min of additional stirring, morpholine (0.26 mL, 3.0 mmol) was added and the mixture was stirred for an hour and then concentrated. The crude product was pre-purified by silica gel column chromatography using DCM:MeOH:TEA as eluent (98:0:2 → 8:2:0, v/v/v) to yield the slightly impure title compound (0.52 g, 86%). A part of this crude product was further purified using HPLC and treated with Amberlite Na⁺-resin to afford the sodium salt. ¹H-NMR (500 MHz, MeOD-*d*₄), δ : 8.07 (d, *J* = 7.4 Hz, 2H, Arom. Bz), 7.98 (s, 1H, H6), 7.64 - 7.44 (m, 3H, Arom. Bz), 6.47 (dd, *J* = 8.6, 5.9 Hz, 1H, H1'), 5.94 - 5.92 (m, 2H, CH₂ POM), 5.65 (d, *J* = 5.4 Hz, 1H, H3'), 4.38 (s, 1H, H4'), 4.22 - 3.99 (m, 2H, H5'), 3.71 - 3.58 (m, 4H, CH₂ Morph.), 3.13 (dd, *J* = 9.0, 4.6 Hz, 4H, CH₂ Morph.), 2.70 - 2.47 (m, 2H, H2'), 2.01 (s, 3H, CH₃ T), 1.17 (s, 9H, *t*Bu POM). ¹³C-NMR (125 MHz, MeOD-*d*₄), δ: 179.0 (CO POM), 167.3, 164.5 (C4, C2), 152.0 (CO Bz), 137.3 (C6), 134.6 (Arom. Bz), 131.0 (C_q Bz), 130.6, 129.6 (Arom. Bz), 111.4 (C5), 87.1 (C1'), 85.7 (C4'), 77.8 (C3'), 68.6 (CH₂N morpholine), 66.3 (CH₂ POM), 65.8 (C5), 46.7 (CH₂O, morpholine), 39.8 (C_q POM), 38.4 (C2'), 27.4 (CH₃ POM), 13.2 (CH₃ T). ³¹P-NMR (162 MHz, MeOD-*d*₄), δ: 6.42. IR: 1717, 1663, 1452, 1269, 1258, 1099, 712. HRMS [C₂₇H₃₆N₃O₁₁P + H]⁺: calc. 610.2160, fnd. 610.2157.

3'-O-Benzoyl-3-N-Pivaloyloxymethylthymidine-5'-O-monophospho-isobutylamidate sodium salt (9)

To a cooled (0 °C) solution of **7** (0.46 g, 1.0 mmol) and **1** (0.63 g, 1.3 mmol) in MeCN (15 mL) was slowly added a solution of DCI (0.15 g, 1.25 mmol) in MeCN (5 mL). After 20 min, a solution of iodine in pyridine (2.0 mL, 0.5 M) was added to the reaction mixture until the solution kept its brown colour. After 30 min of additional stirring, isobutylamine (0.3 mL, 3.0 mmol) was added and stirred for an hour. The reaction mixture was concentrated and the crude product was pre-purified by silica gel column chromatography using DCM:MeOH:TEA as eluent (98:0:2 → 8:3:0, v/v/v) to yield slightly impure **9** (0.24 g, 79 %). A part of this crude product was further purified using HPLC and treated with Amberlite Na⁺-resin to afford the sodium salt. ¹³C-NMR (125 MHz, MeOD-*d*₄), δ: 179.12 (CO POM), 167.4, 164.5 (C2, C4), 152.10 (CO Bz), 137.51 (C6), 134.58 (Arom. Bz), 130.93 (C_q, Bz), 130.64, 129.66 (Arom. Bz), 111.54 (C5), 87.08 (C1'), 85.73, 85.66 (C4', *J*_{CP} = 9.2 Hz), 77.74 (C3'), 66.25 (C5'), 65.50 (CH₂ POM), 49.52 (CH₂ isobutyl), 38.31 (C2'), 31.34 (CH isobutyl), 27.34 (CH₃ POM), 20.58, 19.37 (CH₃ isobutyl), 13.28 (CH₃ T). ³¹P-NMR (80.7 MHz, MeOD-*d*₄), δ: 7.42. MS: 596.2 [M + H]⁺. IR: 1716, 1666, 1469, 1446, 1311, 1269, 1257, 1111, 659, 644. HRMS [C₂₇H₃₈N₃O₁₀P + H]⁺: calc. 596.2367, fnd. 596.2366.

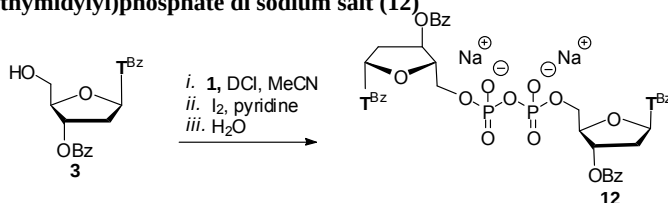
N³-Benzoyl-3'-O-benzoylthymidine-5'-phosphorofluoridate sodium salt (10)

To a cooled (0 °C) solution of **3** (0.45 mmol) and **1** (0.59 mmol) in MeCN (4 mL) was slowly added a solution of DCl (0.59 mmol) in MeCN (1 mL). After 20 min, a solution of iodine in pyridine was added to the reaction mixture until the solution kept its brown colour. After 30 min of additional stirring, DAST (0.07 mL, 1.17 mmol) was added followed by 16 hours of stirring. The reaction mixture was concentrated and the crude product was purified by silica gel column chromatography using DCM:MeOH as eluent (1:0 → 98:2 v/v) and treatment with Amberlite Na⁺-resin to yield **10** (0.19 g, 80 %). ¹H-NMR (400 MHz, MeOD-*d*₄), δ: 8.08 - 7.87 (m, 5H, Arom. Bz/H6), 7.70 (t, *J* = 7.5 Hz, 1H, Arom. Bz), 7.56 - 7.44 (m, 3H, Arom. Bz), 7.25 - 6.99 (m, 2H, Arom. Bz), 6.43 (t, *J* = 7.3 Hz, 1H, H1'), 5.62 (s, 1H, H3'), 4.39 (s, 2H, H4'), 4.37 - 4.17 (m, 2H, H5'), 2.57 (dd, *J* = 6.9, 3.4 Hz, 2H, H2'), 2.00 (s, 3H, CH₃ T). ¹³C-NMR (100 MHz, MeOD-*d*₄), δ: 170.3 (CO Bz), 167.3, 164.8 (C2, C4), 151.0 (CO Bz), 137.8 (C6), 136.3, 134.6 (Arom. Bz), 132.9 (Cq Bz), 131.5 (Arom. Bz), 130.8 (Cq Bz), 130.4, 129.9, 129.7, 129.2 (CH, Bz), 112.4 (C5), 86.6 (C1'), 85.4 (C4'), 77.4 (C3'), 67.6 (C5'), 38.4 (C2'), 12.5 (CH₃ T). ³¹P-NMR (162, MeOD-*d*₄), δ: -2.27, -8.01 (d, *J*_{PF} = 930 Hz). IR: 3391, 1747, 1698, 1652, 1656, 1599, 1449, 1250, 1100, 1061, 823, 714, 489. HRMS [C₂₄H₂₂FN₂O₉P + H]⁺: calc. 533.1119, fnd. 533.1118.

N³-Benzoyl-3'-O-benzoylthymidine-5'-(*p*-nitrophenyl)phosphate sodium salt (11)

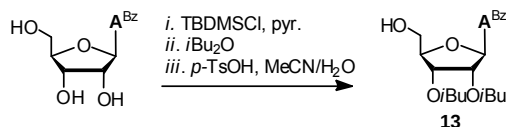
To a cooled (0 °C) solution of **3** (0.25 mmol) and **1** (0.33 mmol) in MeCN (2.5 mL) was slowly added a solution of DCl (0.33 mmol) in MeCN (1 mL). After 20 min, a solution of iodine in pyridine was added to the reaction mixture until the solution kept its brown colour. After 30 min of additional stirring, *p*-nitrophenol sodium salt (0.05 g, 0.33 mmol) was added followed by 16 hours of stirring. The reaction mixture was concentrated and the crude product was purified by silica gel column chromatography using DCM:MeOH as eluent (1:0 → 97:3 v/v) and treated with Amberlite Na⁺-resin to yield **11** (73 mg, 44%). ¹H-NMR (400 MHz, MeCN-*d*₃), δ: 8.14 - 7.31 (m, 16H, Arom. NO₂Ph/Arom. Bz/H6), 6.42 (bs, 1H, H1'), 5.54 (bs, 1H, H3'), 4.38 (bs, 1H, H4'), 4.27 (bs, 2H, H5'), 2.46 (bs, 2H, H2'), 1.89 (bs, 3H CH₃ T). ¹³C-NMR (100 MHz MeCN-*d*₃), δ: 170.7 (CO Bz), 166.7, 163.8 (C4, C2), 158.0 (Cq NO₂Ph), 150.4 (CO Bz), 143.8 (Cq NO₂Ph), 137.3 (C6), 136.3, 134.4 (Arom. Bz), 132.0 (Cq Bz), 131.2, 130.5, 130.4, 129.5 (Arom. Bz), 126.2, 121.4 (Arom. NO₂Ph), 111.9 (C5), 85.9 (C1'), 84.4 (C4'), 76.6 (C3'), 65.0 (C5'), 37.9 (C2'), 12.9 (CH₃ T). ³¹P-NMR (160 MHz, MeCN-*d*₃), δ: -6.32. IR: 1748, 1700, 1658, 1652, 1344, 1248, 1098, 897. HRMS [C₃₀H₂₆N₃O₁₂P + H]⁺: calc. 652.1326, fnd. 652.1330.

5'-(*N*³-Benzoyl-3'-*O*-benzoyl-thymidylyl-phosphoryl)-5'-(*N*³-Benzoyl-3'-*O*-benzoyl-thymidylyl)phosphate di sodium salt (12**)**



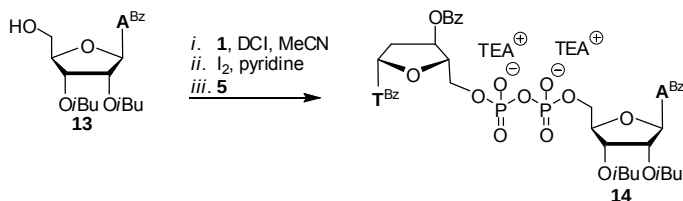
To a cooled (0 °C) solution of **3** (0.55 mmol) and **1** (0.71 mmol) in MeCN (5 mL) was slowly added a solution of DCl (0.71 mmol) in MeCN (1 mL). After 20 min, a solution of iodine in pyridine was added to the reaction mixture until the solution kept its brown colour. After 30 min of additional stirring, H₂O (70 µL, 3.5 mmol) was added followed by 16 hours of stirring. The reaction mixture was concentrated and the crude product was purified by silica gel column chromatography using DCM:MeOH:H₂O as eluent (1:0:0 → 6:3:1 v/v/v) followed by Amberlite Na⁺-resin treatment to yield **12** (0.45 g, 78 %). ¹H-NMR (400 MHz, MeOD-*d*₄), δ: 8.10 - 7.86 (m, 10H, Arom. Bz), 7.77 - 7.32 (m, 12H, Arom. Bz/H6), 6.54 - 6.37 (m, 2H, H1'), 5.74 (d, *J* = 4.4 Hz, 2H, H3'), 4.49 - 4.38 (m, 6H, H4', H5'), 2.86 - 2.46 (m, 4H, H2'), 2.05 (s, 6H, CH₃ T). ¹³C-NMR (100 MHz, MeOD-*d*₄), δ: 169.9 (CO Bz), 167.1, 164.2 (C2, C4), 150.5 (CO Bz), 137.5 (C6), 136.0, 134.4 (Arom. Bz), 132.2 (C_q Bz), 130.3, 129.9, 129.3, 129.2 (Arom. Bz), 112.2 (C5), 86.1 (C1'), 84.9 (C4'), 76.9 (C3'), 66.5 (C5'), 37.7 (C2'), 12.6 (CH₃ T). ³¹P-NMR (162 MHz, CDCl₃/MeOD-*d*₄), δ: -8.44. IR: 3384, 1748, 1699, 1661, 1645, 1635, 1448, 1249, 1108, 1066, 931, 712. HRMS [C₄₈H₄₄N₄O₁₉P₂ + H]⁺: calc. 1043.2147, find. 1043.2157.

***N*⁶-Benzoyl-2', 3'-di-*O*-iso-butyryladenosine (**13**)**



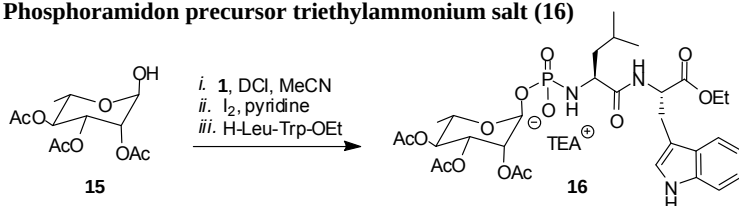
*N*⁶-Benzoyladenosine³⁵ (1.8 g, 4.8 mmol) was dissolved in pyridine (25 mL) after coevaporation with pyridine. TBDMS-Cl (10 mmol, 1.5 g) was added and the reaction mixture was stirred for 1 hour at room temperature. After this *iso*-butyric anhydride (20 mmol, 3.7 mL) was added and the reaction mixture was stirred for 48 hours. The reaction mixture was concentrated and taken up in DCM. After washing with 5% citric acid and H₂O the organic phase was dried (MgSO₄) and concentrated. The residue was taken up in MeCN/ H₂O (50 mL, 4:1, v/v) and *p*-TsOH (5 mmol, 0.8 g) was added after which the reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was concentrated and purified using silica gel column chromatography EA/PE (0:100 → 45:55 v/v) to afford the title compound as a white foam (1.9 g, 3.9 mmol, 78%). ¹H-NMR (400 MHz, CDCl₃); δ 9.26 (s, 1H, NH), 8.76 (s, 1H, H2), 8.14 (s, 1H, H8), 8.07 - 7.96 (m, 2H, Arom. Bz), 7.71 - 7.56 (m, 1H, Arom. Bz), 7.51 (dd, *J* = 10.4, 4.7 Hz, 2H, Arom. Bz), 6.13 (d, *J* = 7.5 Hz, 1H, H1'), 6.01 (dd, *J* = 7.5, 5.4 Hz, 1H, H2'), 5.88 (bs, 1H, OH), 5.71 (dd, *J* = 5.4, 1.3 Hz, 1H, H3'), 4.35 (d, *J* = 1.4 Hz, 1H, H4'), 4.10 (dd, *J* = 12.8, 1.6 Hz, 1H, H5_A'), 3.88 (d, *J* = 12.8 Hz, 1H, H5_B'), 2.65 (hept., *J* = 7.0 Hz, 1H, CH *i*Bu), 2.50 (hept., *J* = 7.0 Hz, 1H, CH *i*Bu), 1.24 (dd, *J* = 7.0, 2.2 Hz, 6H, CH₃ *i*Bu), 1.09 (dd, *J* = 14.0, 7.0 Hz, 6H, CH₃ *i*Bu). ¹³C-NMR (400 MHz, CDCl₃); δ 175.7, 175.1 (CO *i*Bu), 164.6 (CO Bz), 152.3 (C2), 150.9 (C6), 150.2 (C4), 142.2 (C8), 133.3 (C_q Bz), 132.8, 128.8, 127.9 (Bz), 124.4 (C5), 88.4 (C1'), 86.3 (C4'), 72.8 (C2'), 72.2 (C3'), 62.4 (C5'), 33.8, 33.5 (CH *i*Bu), 18.9, 18.8, 18.7, 18.6 (CH₃ *i*Bu). HRMS [C₂₅H₂₉N₅O₇ + H]⁺: calc. 512.2135, find. 512.2139.

5'-(*N*³-Benzoyl-3'-*O*-benzoyl-thymidylyl-phosphoryl)-5'-(*N*⁶-Benzoyl, 2',3'-*O*-di-isobutyryladenilyl)-phosphate di triethylammonium salt (14**)**



To a cooled solution of adenosine **13** (0.18 mmol, 0.92 mg) in MeCN (5 mL) and **1** (0.25 mmol, 97 μ L) was added DCl (0.22 mmol, 26 mg). After 20 min a solution of iodine in pyridine (0.5 M) was added to the reaction mixture and stirred for 1 hour. Then, phosphate **5** (0.25 mmol) in 2.5 mL MeCN was added and stirred at room temperature for 16 hours. RP-HPLC purification followed by gel filtration using a HW40 column (50 mM triethylammonium acetate in H₂O as eluent) afforded the title compound as a white solid (62 mg, 27 %). ¹H-NMR (400 MHz, MeOD-*d*₄), δ : 8.70 (s, 1H, H8-A), 8.50 (s, 1H, H2-A), 7.90 (d, *J* = 7.6 Hz, 2H, Arom. Bz), 7.76 (m, 3H, Arom. Bz/ H6-T), 7.68 (d, *J* = 7.2 Hz, 3H, Arom. Bz), 7.47 (dd, *J* = 13.4, 6.4 Hz, 4H, Arom. Bz), 7.38 (t, *J* = 7.5 Hz, 2H, Arom. Bz), 7.27 (d, *J* = 6.9 Hz, 2H, Arom. Bz), 6.29 (d, *J* = 5.8 Hz, 1H, H1'-A), 5.99 (d, *J* = 7.2 Hz, 1H, H1'-T), 5.73 (d, *J* = 5.5 Hz, 1H, H2'-A), 5.63 (s, 1H, H3'-A), 5.40 (s, 1H, H3'-T), 4.49 (s, 1H, H4'-A), 4.26 (s, 1H, H4'-T), 4.19 (s, 4H, H5'-A, H5'-T), 2.72 - 2.56 (m, 1H, CH *i*Bu), 2.50 - 2.40 (m, 1H, CH *i*Bu), 2.36 (s, 2H, H2'-T), 1.87 (s, 3H, CH₃-T), 1.11 (dd, *J* = 9.2, 7.3 Hz, 6H, CH₃ *i*Bu), 0.91 (dd, *J* = 16.6, 6.9 Hz, 6H, CH₃ *i*Bu). ³¹P-NMR (162 MHz, MeOD-*d*₄), δ : -10.69, -10.80, -10.96, -11.07 (dd, *J* = 18.3, 42.7 Hz). IR: 3358, 2927, 1746, 1699, 1661, 1652, 1455, 1247, 1099. HRMS [C₄₉H₅₁N₇O₁₉P₂ + H]⁺: calc. 1104.2787, fnd. 1104.2796.

Phosphoramidon precursor triethylammonium salt (16**)**



To a cooled solution of 2,3,4-tri-*O*-acetyl- β -L-rhamnopyranose³⁴ (0.5 mmol, 0.145 mg) in MeCN (4.5 mL) and **1** (0.6 mmol, 235 μ L) was added DCl (0.55 mmol, 65 mg). After 20 min a solution of iodine in pyridine (0.5 M) was added and the reaction mixture was stirred for 20 min. Then, H-Leu-Trp-OEt³⁴ (0.50 mmol) in 1.0 mL MeCN was added and stirred at room temperature for 48 hours. RP-HPLC purification followed by gel filtration using a HW40 column (50 mM triethylammonium acetate in H₂O as eluent) afforded the title compound as a white solid (74 mg, 19 %). ¹H-NMR (400 MHz, MeOD-*d*₄), δ : 9.07 (bs, 1H, Trp-indole NH), 7.52 (d, *J* = 7.8 Hz, 1H, Trp-NH), 7.45 (d, *J* = 7.6 Hz, 1H, Trp ArH), 7.34 (d, *J* = 8.1 Hz, 1H, Trp ArH), 7.23 (s, 1H, Trp CH indole), 7.13 (t, *J* = 7.4 Hz, 1H, Trp ArH), 7.05 (t, *J* = 7.4 Hz, 1H, Trp ArH), 5.44 (d, *J* = 7.5 Hz, 1H, Rha-H1'), 5.37 - 5.25 (m, 2H, Rha-H2', Rha-H3'), 5.06 (t, *J* = 10.0 Hz, 1H, Rha-H4'), 4.79 (dd, *J* = 13.2, 5.7 Hz, 1H, Trp-CH), 4.25 - 3.99 (m, 3H, Rha-H5', Trp-OCH₂CH), 3.78 - 3.69 (m, 1H, Leu-CH), 3.31 (dd, *J* = 5.2, 2.6 Hz, 2H, Trp-CH₂), 2.96 - 2.93 (m, 1H, Leu-NH), 2.12 (s, 3H, OAc), 2.02 (s, 3H, OAc), 1.97 (s, 3H, OAc), 1.83 - 1.68 (m, 2H, Leu-CH(CH₃)₂, Leu-CH₂), 1.63 (s, 1H, Leu-CH₂), 1.25 - 1.12 (m, 6H, Trp-OCH₂CH₃, Rha-CH₃), 0.88 (dd, *J* = 6.4, 2.5 Hz, 6H, Leu-CH(CH₃)₂). ³¹P-NMR (162 MHz, MeOD-*d*₄), δ : 3.16. HRMS [C₃₁H₄₄N₃O₁₃P + H]⁺: calc. 698.2684, fnd. 698.2685.

References and Notes

- ¹ Van der Heden van Noort, G. J.; Verhagen, C. P.; Van der Horst, M. G.; Overkleeft, H. S.; Van der Marel, G. A.; Filippov, D. V.; *Org. Lett.*, **2008**, *10*, 4461 - 4464.
- ² McMurray, J. S.; Coleman, D. R.; Wang, W.; Campbell, M. L.; *Biopolymers*, **2001**, *60*, 3 - 31.
- ³ Nikolaev, A. V.; Botvinko, I. V.; Ross, A. J.; *Carbohydr. Res.*, **2007**, *342*, 297 - 344.
- ⁴ Beaucage, S. L.; Iyer, R. P.; *Tetrahedron*, **1993**, *49*, 10441 - 10488.
- ⁵ Burgess, K.; Cook, D.; *Chem. Rev.*, **2000**, *100*, 2047 - 2059.
- ⁶ Hindsgaul, O. F., M.; *Molecular Glycobiology*; Oxford University Press: New York, **1994**.
- ⁷ Umezawa, S.; Tatsuta, K.; Izawa, O.; Tsuchiya, T.; *Tetrahedron Lett.*, **1972**, *13*, 97 - 100.
- ⁸ Fields, S. C.; *Tetrahedron*, **1999**, *55*, 12237 - 12272.
- ⁹ Kraszewski, A.; Stawinski, J.; *Pure Appl. Chem.*, **2007**, *79*, 2217 - 2227.
- ¹⁰ Michalski, J.; Dabkowski, W.; *Top. Curr. Chem.*, **2004**, *232*, 93 - 144.
- ¹¹ Moriguchi, T.; Asai, N.; Okada, K.; Seio, K.; Sasaki, T.; Sekine, M.; *J. Org. Chem.*, **2002**, *67*, 3290 - 3300.
- ¹² Burkart, M. D.; Vincent, S. P.; Duffels, A.; Murray, B. W.; Ley, S. V.; Wong, C. H.; *Bioorg. Med. Chem.*, **2000**, *8*, 1937 - 1946.
- ¹³ Klinman, D. M.; Yi, A. K.; Beaucage, S. L.; Conover, J.; Krieg, A. M.; *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2879-2883.
- ¹⁴ Stein, C. A.; Subasinghe, C.; Shinozuka, K.; Cohen, J. S.; *Nucleic Acids Res.*, **1988**, *16*, 3209 -3221.
- ¹⁵ Bollmark, M.; Stawinski, J.; *Nucleosides Nucleotides Nucleic Acids*, **1998**, *17*, 663 - 680.
- ¹⁶ Nifantiev, E. E.; Grachev, M. K.; Burmistrov, S. Y.; *Chem. Rev.*, **2000**, *100*, 3755 - 3799.
- ¹⁷ Beaucage, S. L.; Iyer, R. P.; *Tetrahedron*, **1993**, *49*, 6123 - 6194.
- ¹⁸ Beaucage, S. L.; Iyer, R. P.; *Tetrahedron*, **1992**, *48*, 2223 - 2311.
- ¹⁹ Ausin, C.; Grajkowski, A.; Cieslak, J.; Beaucage, S. L.; *Org. Lett.*, **2005**, *7*, 4201 - 4204.
- ²⁰ Lartia, R.; Asseline, U.; *Tetrahedron. Lett.*, **2004**, *45*, 5949 - 5952.
- ²¹ Nurminen, E.; Lonnberg, H.; *J. Phys. Org. Chem.*, **2004**, *17*, 1 - 17.
- ²² Ghisaidoobe, A. B. T.; de Koning, M. C.; Duynstee, H. I.; Ten Kortenaar, P. B. W.; Overkleeft, H. S.; Filippov, D. V.; van der Marel, G. A.; *Tetrahedron Lett.*, **2008**, *49*, 3129 - 3132.
- ²³ Kriek, N.; Meeuwenoord, N. J.; van den Elst, H.; Heus, H. A.; van der Marel, G. A.; Filippov, D. V.; *Org. Biomol. Chem.*, **2006**, *4*, 3576 - 3586.

- ²⁴ Filippov, D.; Timmers, C. M.; Roerdink, A. R.; van der Marel, G. A.; van Boom, J. H.; *Tetrahedron Lett.*, **1998**, 39, 4891 - 4894.
- ²⁵ Meyer, A.; Morvan, F.; Vasseur, J. J.; *Tetrahedron Lett.*, **2004**, 45, 3745 - 3748.
- ²⁶ Dreef, C. E.; Tuinman, R. J.; Lefeber, A. W. M.; Elle, C. J. J.; van der Marel, G. A.; van Boom, J. H.; *Tetrahedron*, **1991**, 47, 4709 - 4722.
- ²⁷ Shuto, S.; Shirato, M.; Sumita, Y.; Ueno, Y.; Matsuda, A.; *Tetrahedron Lett.*, **1998**, 39, 7341 - 7344.
- ²⁸ Nielsen, J.; Caruthers, M. H.; *J. Am. Chem. Soc.*, **1988**, 110, 6275 - 6276.
- ²⁹ Wittmann, V.; Wong, C. H.; *J. Org. Chem.*, **1997**, 62, 2144 - 2147.
- ³⁰ Cooke, L. A.; Frauendorf, C.; Gilea, M. A.; Holmes, S. C.; Vyle, J. S.; *Tetrahedron Lett.*, **2006**, 47, 719 - 722.
- ³¹ It became apparent that under the influence of morpholine or *iso*-butylamine the benzoyl protecting group on the nucleobase was removed. This problem was non-existent when the more stable POM group was used for protection of the thymine base.
- ³² Chandrasegaran, S.; Murakami, A.; Kan, L. S.; *J. Org. Chem.*, **1984**, 49, 4951 - 4957.
- ³³ Ohkubo, A.; Aoki, K.; Seio, K.; Sekine, M.; *Tetrahedron Lett.*, **2004**, 45, 979 - 982.
- ³⁴ Donahue, M. G.; Johnston, J. N.; *Bioorg. Med. Chem. Lett.*, **2006**, 16, 5602 - 5604.
- ³⁵ Ti, G. S.; Gaffney, B. L.; Jones, R. A.; *J. Am. Chem. Soc.*, **1982**, 104, 1316 - 1319.

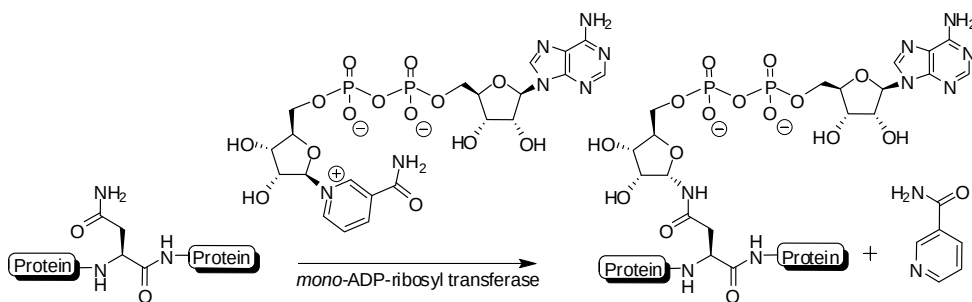
Chapter 3

Synthesis of *Mono*-ADP-Ribosylated Oligopeptides Using Ribosylated Amino Acid Building Blocks¹

Introduction

Adenosine diphosphate ribosylation (ADP-ribosylation) is a posttranslational modification of proteins that is effected by *poly*-ADP-ribosyl polymerases (PARPs) or *mono*-ADP-ribosyl transferases (MARTs).^{2,3} These enzymes transfer ADP-ribose from β -NAD⁺ to nucleophilic functional groups in the side chain of amino acid residues of the target protein, such as Asn, Glu, Asp, Arg or Cys. This process is accompanied by the release of nicotinamide (see **Figure 1**).²

Figure 1



Example of *mono*-ADP-ribosylation of a protein at an asparagine residue

PARP's interact with cellular proteins that play a role in DNA-repair mechanisms, transcription regulation, immune response and cell metabolism. For instance, *poly*-ADP-ribose synthesis mediated by PARP-1/2 on the N- or C-terminal part of histone H1- and H2B-proteins after single-strand DNA breaks, facilitates access to the DNA lesion by relaxation of the chromatin and induces recruitment of repair enzymes.⁴ In the absence of DNA breaks, however, *poly*-ADP-ribose polymer (PAR) mediated chromatin loosening is also observed, allowing transcriptional activation.⁵ PARP-1 is also reported to interact with and to stimulate the production of transcriptional factor NF- κ B, that regulates immune and inflammatory responses; PARP-1 deficient mice or mice treated with PARP inhibitors showed increased resistance against inflammatory stimuli.^{6,7} PAR and PARP's are important in the regulation of chromosomal segregation and possibly play a role in the regulation of cell division.⁸ Interestingly it was recently proposed that inhibiting ADP-ribosylation might be a novel approach to cancer therapy.⁹ Tumor cells of tumor-suppressor protein BRCA1 or BRCA2 mutation carriers have a specific DNA-repair defect and can be selectively killed in the absence of an exogenous DNA-damaging agent using PARP inhibitor olaparib. This treatment has little adverse effects when compared to conventional radio- or chemotherapy.

In contrast to *poly*-ADP-ribosylation, which is mostly effected by mammalian enzymes, the best studied *mono*-ADP-ribosylation events are of bacterial origin. The diphtheria toxin from *Corynebacterium diphtheria*, for instance, inhibits translation of eukaryotic elongation factor 2 (eEF2) blocking its functioning on the ribosome and thus intervening with protein synthesis, ultimately killing the cell.¹⁰ Toxins originating from *Bacillus cereus*, *Staphylococcus aureus*, and *Clostridium botulinum* also possess ADP-ribosyl transferase activity and target a part of the invaded hosts Rho-enzyme family, thereby inducing breakdown of the cytoskeleton.¹¹

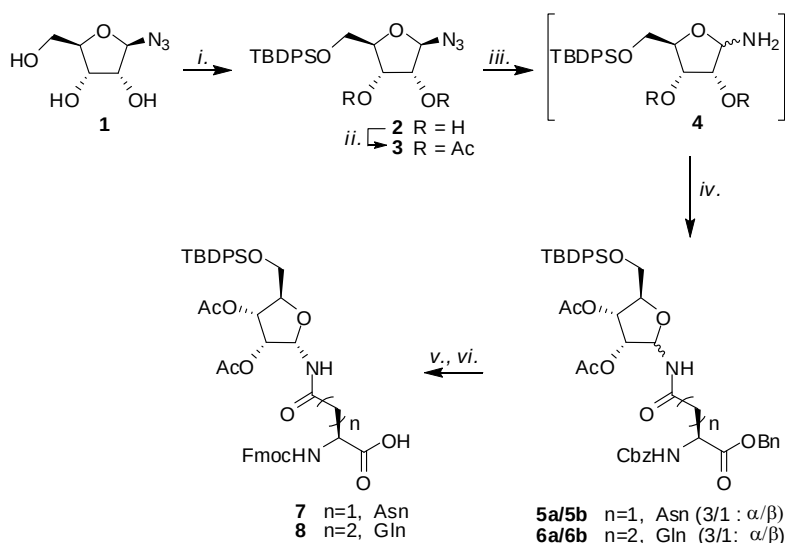
Although ADP-ribosylating enzymes have been identified in many prokaryotic and eukaryotic species as well as in viruses, most biological processes in which they partake and their exact mode of action in these processes have not been clarified completely.³ The elucidation of the role and mechanisms of ADP-ribosylation events would benefit greatly from the availability of well-defined ADP-ribosylated peptides and analogues thereof. The main issues addressed in this chapter towards the construction of *mono*-ADP-ribosylated peptides entail the synthesis of protected ribosylated amino acids that are compatible with peptide synthesis, development of a protective group strategy for the solid phase peptide synthesis that is compatible with the integrity of the ribosyl adenosine diphosphate moiety and an appropriate procedure for pyrophosphate formation.

Results and Discussion

Known ADP-ribosylation sites in naturally occurring proteins include Asn-41 of the mammalian Rho protein, the target of *mono*-ADP-ribosylation catalyzed by a toxin from *Clostridium botulinum* and Glu-3 of the human H2B histone, which is a substrate for PARP enzymes.¹²⁻¹⁴ In the later case, the ADP-ribose moiety is linked to the glutamic acid residue via an ester bond, which is known to be unstable under mild basic conditions.^{2,13} Therefore an amide bond, linking glutamine to the ADP-ribose moiety, comparable to the naturally occurring amide bond linking the asparagine residue in Rho proteins to the ADP-ribose moiety, was envisioned to be a better choice as synthetic target for histone H2B derived peptides. This amide linked glutamine ADP-ribose is isosteric to the natural glutamic acid ADP-ribosyl ester.

The key *N*-ribosylated Asn and Gln building blocks **7** and **8** (**Scheme 1**) were provided with the mutually orthogonal TBDPS and Fmoc protecting groups, allowing the introduction of a pyrophosphate moiety and elongation of the peptide chain in a sequential manner. The route of synthesis started with the silylation of the primary alcohol of known β -D-ribofuranosyl azide **1**¹⁵ and subsequent acylation of the secondary 2' and 3' hydroxyls in **2** to give fully protected azide **3** in a near quantitative yield. Reduction of β -azide **3** at 10°C using PtO_2/H_2 ¹⁶ resulted in the formation of hemi-aminal ether **4**, as a mixture of epimers at C-1. Others have prepared a similar ribosylamine in poor yield (6%), due to its instability.¹⁷ For this reason the hemi-aminal ether mixture was used immediately as a crude and EDC mediated coupling with either Cbz-Glu-OBn or Cbz-Asp-OBn gave the ribosylated amino acids **5** and **6** respectively, as anomeric mixtures. In addition a minor side reaction was detected in both cases, originating from migration of the acetyl group from the 2'O to the 1'NH in **4** followed by coupling of the amino acid to the 2'O position to form an ester linkage.

Scheme 1. Ribosylated Asn and Gln Fmoc amino acids



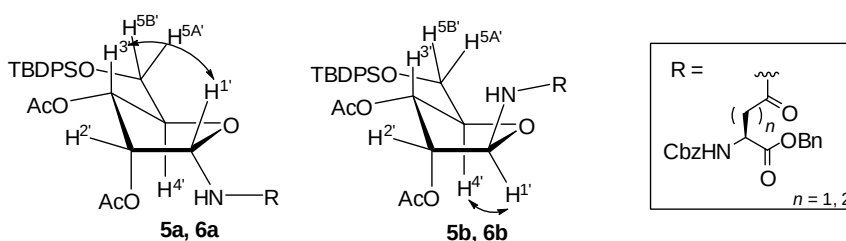
Reagents and conditions; *i.* TBDPSCl, pyridine, r.t., 16 h, quantitative, *ii.* Ac_2O , pyridine, r.t., 16 h, 98%, *iii.* PtO_2 , H_2 , EtOAc, 10 °C, 2 h, *iv.* Z-Asp-OBn or Z-Glu-OBn, EDC.HCl, DCM, 16 h, r.t., 47% or 61%, respectively (over two steps), *v.* 10% Pd/C, H_2 , MeOH, r.t., *vi.* FmocOSu, NaHCO_3 , water/acetone, 2 h, r.t., 84% or 97%, respectively (over two steps).

The anomeric mixture could be separated using silica gel column chromatography, but unlike in pyranoses, the coupling constants obtained from ^1H -NMR spectroscopy of the separated isomers could not be used to determine the anomeric configuration of **5a/5b** and **6a/6b** respectively, due to their small difference.¹⁸ In the case of ribosyl-asparagine mixture **5a/5b**, the $J_{1,2'}$ values found for the higher running product (TLC; $R_f = 0.5$) were 5.8 Hz in $\text{MeOD-}d_4$ and 5.2 Hz in CDCl_3 and the $J_{1,2'}$ values found for the lower running product (TLC; $R_f = 0.4$) were 4.9 Hz in $\text{MeOD-}d_4$ and 5.6 Hz in CDCl_3 . This small difference in the $J_{1,2'}$ values of ca 1 Hz indeed did not cast light on the identity of the anomers.

The H-1' was empirically found to resonate at lower field for 1,2-*cis* derivatives as compared to 1,2-*trans* compounds.¹⁸ In the case of **5a/5b**, the H-1' of the higher running product (TLC; R_f = 0.5) resonated at 5.61 ppm in MeOD-*d*₄ and 5.76 ppm in CDCl₃. The H-1' of the lower running product (TLC; R_f = 0.4) resonated at 4.87 ppm in MeOD-*d*₄ and 5.56 ppm in CDCl₃. This indicated that the higher running spot would be the β-anomer and the lower running spot the α-anomer.

These identities could be further confirmed by means of NOE NMR spectroscopy experiments. Presence of a cross-peak between H1' and H3' of ribose indicated the *cis* configuration **5a**, whereas for the corresponding *trans* epimer **5b** a cross-peak between H1' and H4' was observed in the NOESY spectra (see **Figure 2**). Similar results were obtained for ribosyl-glutamine **6**. Hydrogenation of the α-anomers **5a** and **6a** to cleave the benzyl ester and Cbz protective groups, and subsequent introduction of the Fmoc group under Schotten-Baumann conditions gave **7** and **8**.

Figure 2

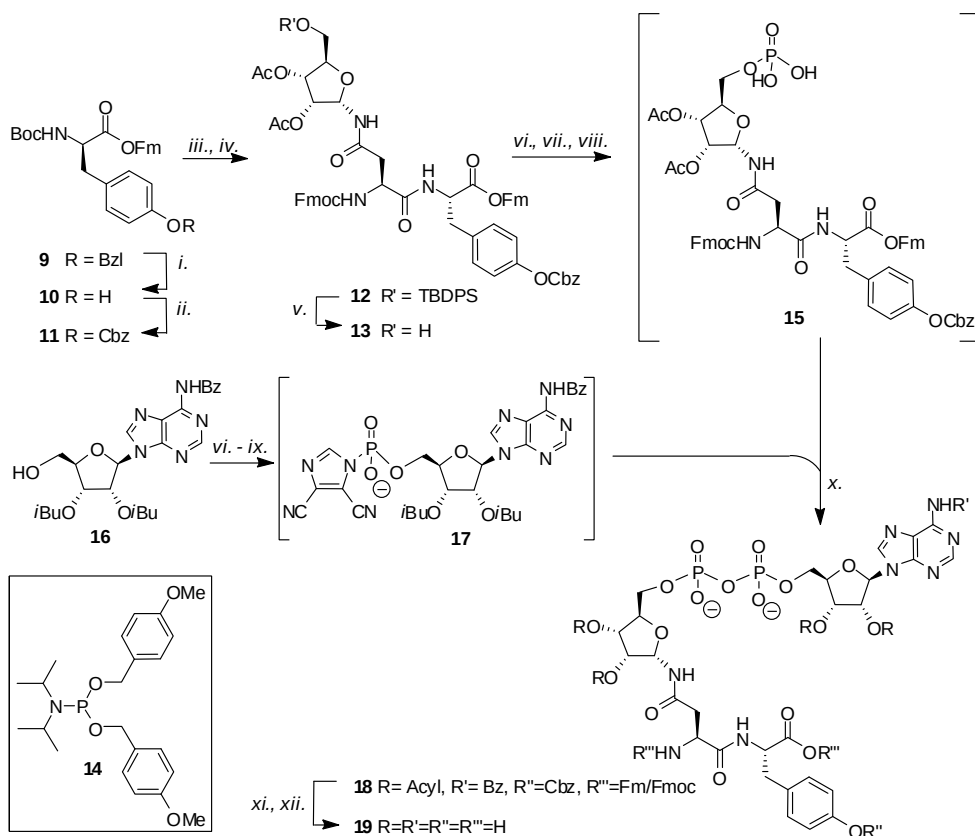


NOE observations.

With suitably protected ribosylated Gln and Asn building blocks in hand the synthesis of *mono*-ADP-ribosylated dipeptide **19** in solution was undertaken (see **Scheme 2**). Peptide fragment **19** is derived from the bovine Rho protein, that is ADP-ribosylated by the *Clostridium botulinum* C3 toxin at the Asn-41 residue.¹² The partially protected precursor **18** of target dipeptide **19** was provided with base labile protecting groups (Fm, Fmoc, Ac and Cbz carbonate) to permit a straightforward basic deprotection procedure in the final stage of the synthesis, since it was envisaged that both the glycosidic linkage and the pyrophosphate moiety in the target molecule **19** would be stable under such conditions. The route of synthesis started with the preparation of suitably protected tyrosine derivative **11** (**Scheme 2**) by conversion of commercially available Boc-Tyr(OBn)-OH into Fm ester **9**, followed by hydrogenolysis of the benzyl ether and subsequent installation of the Cbz group at the released phenolic hydroxyl function. Acidic removal of the Boc group in **11** and EDC assisted coupling to ribosylated Asn building block **7** resulted in the formation of dipeptide **12**. After unmasking the 5'-OH with HF/pyridine, the phosphate and the pyrophosphate functionality were introduced sequentially by adaptation of the procedure described in **Chapter 2**.¹⁹ Phosphate monoester **15** was obtained by phosphitylation of ribosylated dipeptide **13** with di(*p*-methoxybenzyl)-*N,N*-diisopropylphosphoramidite **14** and activator dicyanoimidazole (DCI), followed by oxidation of the intermediate phosphite triester with anhydrous *t*-BuO₂H and finally *p*-methoxybenzyl removal using 3% TFA. Similar phosphitylation of adenosine **16** followed by oxidation with

iodine afforded the putative intermediate **17**, an activated phosphorimidazolidate suitable for pyrophosphate formation. Addition of crude **15** to *in situ* formed **17** resulted in the partially protected *m*-ADP-ribosylated dipeptide **18**, which was purified using RP-HPLC. Removal of Fm and Fmoc groups using DBU, followed by cleavage of the remaining acyl functions using aqueous ammonia and gel filtration resulted in 4.1 mg (4.9 μ mol) of the homogeneous target compound **19**.

Scheme 2 Synthesis of *mono*-ADP-ribosylated dipeptide **19**



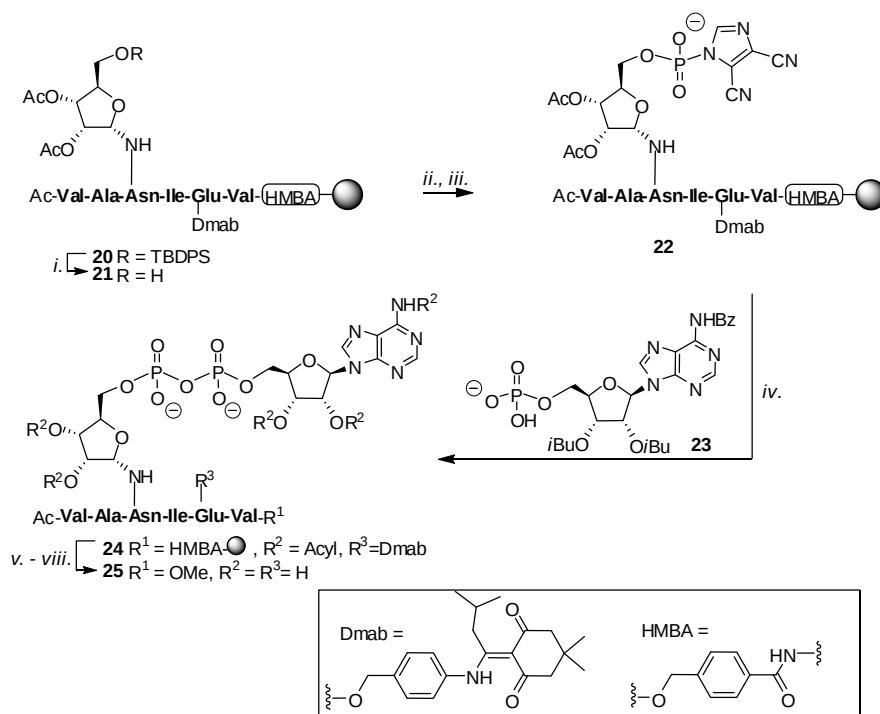
Reagents and conditions: *i.* Pd/C, H₂, *n*-PrOH/EtOAc, r.t., 1 h, 73%, *ii.* CbzCl, DiPEA, DCM, r.t., 2 h, 84%, *iii.* 3M HCl EtOAc, 30 min, r.t., *iv.* **7**, EDC.HCl, DiPEA, DCM, 1.5 h, r.t., (84% over three steps), *v.* HF/pyridine, r.t., 1 h, 85%, *vi.* **14**, DCl, MeCN/DCM, 0 °C, 1 h, *vii.* 5.5 M *t*BuO₂H, 2 h, r.t., *viii.* 3% TFA, DCM, 5 min, r.t., *ix.* I₂, pyr., 0 °C, 30 min, *x.* r.t., 16 h, (15%), *xi.* DBU, 30 min., r.t., *xii.* NH₃ aq., dioxane, 16 h, r.t., (58%).

The successful synthesis of the *m*-ADP-ribosylated dipeptide **19** prompted to investigate an approach towards elongated *m*-ADP-ribosyl peptide fragments. Having established an efficient solution phase method for the incorporation of ribosylated amino acid **7** as well as application of the method to prepare phosphate and pyrophosphate moieties with the aid of phosphitylating agent **14** described in **Chapter 2**, the question arose whether these methods would also be applicable to solid support based synthesis. Two complementary procedures were conceivable

to introduce the pyrophosphate function by reacting a phosphate monoester with an activated phosphorimidazolidate. The phosphorimidazolidate of an immobilized ribosylated peptide can be reacted with the phosphate monoester of a nucleotide or conversely a peptide-resin equipped with a phosphate monoester can be reacted with the nucleotide phosphorimidazolidate. *m*-ADP-ribosylated peptides, **25** and **32** (see **Scheme 3** and **Scheme 4**) were constructed each using one of these procedures.

The first procedure to introduce the pyrophosphate on resin involves introduction of the reactive phosphorimidazolidate at the 5' hydroxyl of immobilized ribosylated peptide and addition of the phosphate monoester. Hydroxymethylbenzoic acid (HMBA) was selected as a suitable linker for the on-resin synthesis of hexapeptide **25** (see **Scheme 3**) containing an ADP-ribosylated Asn residue. The HMBA-linker would allow the removal of the base labile protective groups on the peptide as well as the release of the ADP-ribosylated peptide from the resin in a single step. Furthermore HMBA is reported to be a good choice to afford thioesters, which will be convenient in the future for the synthesis of larger *m*-ADP-ribosylated peptides or proteins using native chemical ligation.

Hexapeptide **20** was assembled on commercially available Tentagel resin equipped with the HMBA linker using a BOP/HOBt Fmoc-based SPPS protocol.²⁰ En route to immobilized **20** the incorporation of ribosylated Asn building block **7** and ensuing repetition of the amino acid coupling cycle proceeded uneventfully as was determined by LC-MS analysis of the crude ribosyl peptide, obtained after cleavage of an analytical sample from the solid support and removal of the protecting groups in **20**. Unmasking of the 5'-hydroxyl on ribose **20** with TEA/HF in pyridine preceded the introduction of the ADP moiety. The activated phosphorimidazolidate was installed at the 5' hydroxyl of the immobilized peptide **21** using DCI assisted phosphitylation with reagent **14** and subsequent oxidation using iodine in pyridine. Addition of known *N*⁶-benzoyl-2',3'-di-*O*-iso-butyryladenyl-5'-monophosphate¹⁹ **23** led to formation of the immobilized and protected ADP-ribosylated peptide **24**. The Dmab group on the glutamic acid residue side chain was removed by a two step procedure²¹ and subsequent removal of the remaining protecting groups and concomitant cleavage from the resin was effected by treatment with methanolic ammonia to give C-terminal methyl ester **25**. LC-MS analysis of the crude reaction mixture showed besides the main product **25** the presence of a few side products, of which the C-terminal carboxamide, the terminal phosphate and the corresponding H-phosphonate could be identified.²² The formation of H-phosphonate was explained by invoking a partial degradation of the phosphite triester on the resin prior to iodine oxidation, the presence of the phosphate monoester was attributed to hydrolysis of the phosphorimidazolidate due to incomplete coupling with the incoming phosphate monoester. Nevertheless, 0.8 mg (0.6 μ mol) of the pure *m*-ADP-ribosylated hexapeptide **25** could be obtained after two RP-HPLC cycles and the identity of the product was ascertained by ³¹P-NMR spectroscopy and high-resolution mass-spectrometry.

Scheme 3. mono-ADP-ribosylated peptides **25**


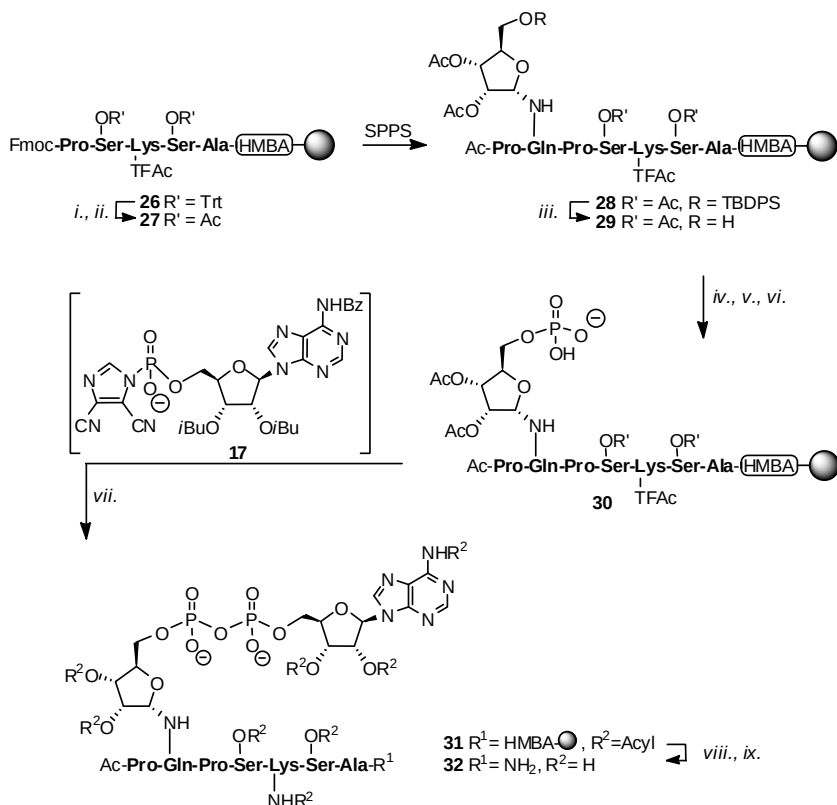
Reagents and conditions: i. HF/pyridine, r.t., 75 min, ii. **14** (Scheme 2), DCI, MeCN/DCM, 0 °C, 1 h, iii. I₂, pyr., 0 °C, 30 min, iv. **23**, v. 2% hydrazine, DMF, 3x 5 min, r.t., vi. 5 mM NaOH, MeOH, 3 h, r.t., vii. 7N NH₃/MeOH, 16 h, r.t., viii. RP-HPLC purification,

The second procedure to form the pyrophosphate linkage involved the introduction of the phosphate monoester at the 5' hydroxyl residue of immobilized ribosylated peptide **29** to give **30** while phosphorimidazolite **17** prepared in solution, was added to this solid support to obtain **31** (see Scheme 4). Such a procedure has the potential to be more effective as unwanted hydrolysis of the intermediate phosphorimidazolite can be counterbalanced by the addition of an excess of **17**. The target ADP-ribosylated heptapeptide **32** is an isosteric fragment of the naturally occurring N-terminus of human histone H2B ADP-ribosylated at E-3 (Scheme 4).^{13,14}

Precursor heptapeptide **26** was prepared using Tentagel resin equipped with the HMBA linker and a BOP/HOBt Fmoc-based SPPS protocol.²⁰ Removal of trityl protective groups on the serine residues followed by acetylation of the thus freed hydroxyls allows final basic deprotection conditions to remove all protective groups in a single step. Next ribosylated Gln building block **8** and Fmoc-Pro-OH were incorporated and LC-MS analysis of the crude ribosylated peptide **28**, obtained after cleavage of an analytical sample from the solid support and removal of the protecting groups, revealed that all previous steps had proceeded efficiently. Ensuing fluoride treatment of **28** released the 5' hydroxyl of ribose to give suitably protected ribosyl peptide **29**. On resin phosphate monoester formation was attained by phosphitylation with **14**, oxidation of the intermediate phosphite triester with *t*BuO₂H and *p*-methoxybenzyl removal using 3% TFA to obtain intermediate **30**. This phosphate equipped resin was treated

with an excess of phosphorimidazolide **17** to give the immobilized protected *m*-ADP-ribosylated **31**. Complete deprotection and cleavage from the solid support was effected by treatment with methanolic ammonia. Purification of the mixture by RP-HPLC followed by gel filtration afforded 1.3 mg (1.0 μ mol) homogeneous *m*-ADP-ribosylated heptapeptide **32**.

Scheme 4. *mono*-ADP-ribosylated peptides **32**



Reagents and conditions: i. 5% TFA, DCM, 3x 5 min, r.t., ii. Ac₂O, DMAP, NMP, 25 min, r.t., iii. HF/pyridine, r.t., 75 min, iv. **14** (Scheme 2), DCI, MeCN, 0 °C, 1 h, v. 5.5M *t*BuO₂H, 1 h, r.t., vi. 3% TFA, DCM, 3x 5 min, r.t., vii. **17** (Scheme 2), MeCN, r.t., 16 h, viii. 7N NH₃/MeOH, ix. RP-HPLC purification.

Analysis of the crude mixture by LC-MS showed besides the main product **32**, again the terminal phosphate monoester and the corresponding H-phosphonate as side products. The formation of the peptide H-phosphonate in course of the phosphorylation reaction proved to be a significant hindrance en route to the target ADP-ribosylated peptides and, in order to suppress this unwanted transformation, we reinvestigated the phosphorylation reaction of **14** on the immobilized ribosylated peptide. Unfortunately it proved to be impossible to suppress the formation of the H-phosphonate by variation of the excess of **14**, the temperature or the reaction time. Attempts to use an alternative phosphorylation method involving the treatment of the

immobilized peptide with POCl_3 in presence of pyridine did not result in an improved formation of the targeted phosphate monoester on resin.²³

A number of applications can be envisaged for the peptides accessible by the methodology described here. One line of research would be the attachment of ADP-ribosylated peptides, such as the ones prepared in this study, to a carrier protein as well-defined haptens to generate antibodies against the ADP-ribose modification. Such tools might prove invaluable for the identification of *mono*-ADP-ribosylated cellular proteins, an intriguing but little understood post-translational modification.²⁴ In this fashion peptide **32** may find use in the discovery of the hitherto unknown ribosylation on the Gln side chain. In an alternative line of research, ADP-ribosylated peptides **19**, **25**, and **32** or closely related analogues represent well-defined substrates for elongation by enzymes of the PARP family. As certain small peptides are reported to be accepted as substrates for PARP-1 (nonapeptide) and ADP-ribose protein lyase (pentapeptide, PE(ADP-ribosyl)PAK), a better insight into the mechanism of the action of PARP enzymes can be attained by using synthetic *mono*-ADP-ribosylated peptides as substrates in elongation assays.²⁵ When combined with established chemical ligation tactics to extend the polypeptide chain²⁶, *mono*-ADP-ribosylated peptides containing ADP-ribosylated glutamine (prepared using Fmoc amino acid **8**) allow preparation of *poly*-ADP-ribosylated proteins containing a stabilized ribose-protein linkage, if they would indeed be accepted by PARPs as substrates. Such constructs are unattainable by current biochemical techniques.

Conclusion

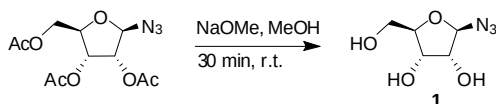
In conclusion a convenient method for the synthesis of protected ribosylated asparagine (**7**) and glutamine (**8**) building blocks is presented, based on the reduction of a protected furanose bearing an anomeric azide and the in situ coupling of the formed hemi-aminal ether to either aspartic or glutamic acid derivatives. After protective group manipulations these ribosylated amino acid building blocks are used in the synthesis of milligram amounts of well-defined ADP-ribosylated peptides both in solution and on solid support.

Experimental

General: All chemicals were used as received. EtOAc, pyridine, *n*-PrOH and MeCN were stored over activated 4Å molecular sieves and MeOH was stored over 3Å molecular sieves for at least 24 hours prior to use. DCM was distilled over CaH₂ and dioxane over LiAlH₄ prior to use. Solvents were removed by rotary evaporation under reduced pressure at a temperature not exceeding 40°C. Optical rotations are measured in CHCl₃ or MeOH at a concentration of 10 mg/mL at 25 °C. TLC analysis, silica gel column chromatography, NMR, LCMS, HPLC and HRMS techniques used are described in **Chapter 2**.

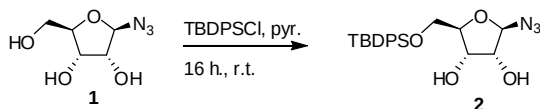
Procedure for solid phase peptide synthesis – Peptides **19** and **26** were synthesized on Tentagel S HMBA resin from Rapp Biopolymere, Weinstadt, Germany. The first amino acid was coupled using 5 eq. DIC, 5 eq. Fmoc-Ala-OH or Fmoc-Val-OH and 5% DMAP in DCM (30 mL/mmol) for 4 hours. Subsequent couplings were performed using 0.5M BOP, 0.5M HOBt, 1.25 M DiPEA and 5 eq. amino acid 0.5M in NMP for 75 minutes. Fmoc deprotection was accomplished using 2% DBU in NMP (3x 3 min.). Capping was done using 0.5M Ac₂O, 0.125M DiPEA and 0.015M HOBt in NMP (1x 1 min.). Fmoc amino acids used were: Fmoc-Pro-OH, Fmoc-Ser(OTrt)-OH, Fmoc-Lys(TFAc)-OH, Fmoc-Val-OH, Fmoc-Ala-OH, Fmoc-Ile-OH, Fmoc-Glu(ODmab)-OH.

β-D-Ribofuranosyl azide (**1**)¹⁵



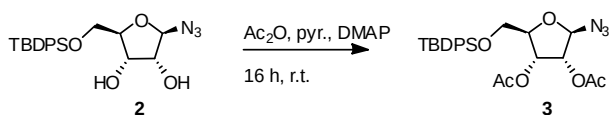
2,3,5-tri-*O*-acetyl- β- D- ribofuranosyl azide (2.60 g, 8.63 mmol), was dissolved in MeOH (40 mL) and NaOMe (1.54 g, 28.5 mmol) was added. After 30 min. TLC analysis (2 % MeOH/ DCM) indicated complete conversion. Amberlite IR120 H⁺ was added until pH = 6. Filtration and concentration gave the title compound as a colourless oil (1.27 g, 7.26 mmol, 85%). ¹H-NMR (400 MHz, MeOD-*d*₄) δ: 5.21 (s, 1H, H1'), 4.89 (s, 3H, OH's), 4.06 (dd, *J* = 6.7, 4.8 Hz, 1H, H3'), 4.03 - 3.93 (m, 1H, H4'), 3.83 (dd, *J* = 4.5, 1.3 Hz, 1H, H2'), 3.77 (dd, *J* = 12.0, 3.2 Hz, 1H, H5_A'), 3.59 (dd, *J* = 12.0, 6.0 Hz, 1H, H5_B'). ¹³C-NMR (100 MHz, MeOD-*d*₄) δ: 96.2 (C1'), 85.2 (C4'), 76.2 (C2'), 71.8 (C3'), 63.8 (C5'). [α_D MeOH] = -195.8°. IR: 3330.3, 2113.8, 1706.1, 1234.7, 1019.0, 925.9.

5-*tert*-Butyldiphenylsilyl-β-D-ribofuranosyl azide (**2**)



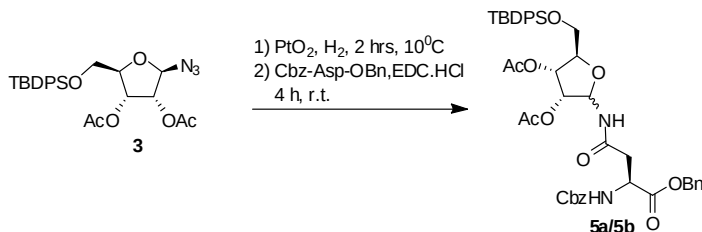
β-D-ribofuranosyl azide **1** (0.55 g, 3.15 mmol) was used without further purification and after coevaporation with pyridine dissolved in pyridine (40 mL). TBDPSCl (0.9 mL, 3.46 mmol) was added and the reaction mixture was stirred for 16 hours at room temperature. The residue was taken up in DCM, washed with 5 % citric acid, dried with MgSO₄ and concentrated. Silica gel column chromatography yielded the title compound (1.30 g, 3.15 mmol, quantitative) as a colourless oil. ¹H-NMR (400 MHz, CDCl₃) δ: 7.89 - 7.56 (m, 4H, Arom. Ph), 7.55 - 7.33 (m, 6H, Arom. Ph), 5.31 (d, *J* = 1.7 Hz, 1H, H1'), 4.37 (dd, *J* = 6.2, 5.0 Hz, 1H, H3'), 4.05 (dt, *J* = 6.3, 4.4 Hz, 1H, H4'), 3.98 (dd, *J* = 4.8, 1.7 Hz, 1H, H2'), 3.92 - 3.76 (m, 2H, H5'), 1.08 (s, 9H, *t*Bu). ¹³C-NMR (100 MHz, CDCl₃) δ: 135.6 (Arom.), 132.8 (Cq Arom.), 129.91, 129.87, 127.84, 127.80 (Arom.), 94.7 (C1'), 83.5 (C2'), 75.5 (C4'), 71.6 (C3'), 63.8 (C5'), 26.8 (CH₃ *t*Bu), 19.2 (Cq *t*Bu). [α_D CHCl₃] = -61.4°. IR: 3314.1, 2931.6, 2858.7, 2110.3, 1427.8, 1234.2, 1111.6, 988.8, 699.1, 501.6. HRMS [C₂₁H₂₇N₃O₄Si + Na]⁺: 436.1661 found, 436.1663 calc.

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



Ribofuranosyl azide **2** (4.13 g, 10 mmol) was dissolved in pyridine (20 mL) and a catalytic amount of DMAP and acetic anhydride (50 mmol, 4.69 mL) were added. The reaction mixture was stirred for 16 hours at room temperature. After concentration the residue was taken up in DCM and washed with 5 % citric acid, dried with MgSO_4 and purified using silica gel column chromatography (PE: EtOAc, 100/0 – 50/50, v/v). The title compound (4.88 g, 9.8 mmol, 98%) was obtained as a colourless oil. $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ : 7.70 (dd, J = 12.5, 4.6 Hz, 4H, Arom. Ph), 7.39 (dd, J = 13.9, 6.6 Hz, 6H, Arom. Ph), 5.51 (t, J = 5.3 Hz, 1H, H3'), 5.42 (d, J = 2.8 Hz, 1H, H1'), 5.21 (dd, J = 4.7, 3.0 Hz, 1H, H2'), 4.22 (dt, J = 5.7, 3.4 Hz, 1H, H4'), 3.86 (dd, J = 11.6, 3.6 Hz, 1H, H5A'), 3.73 (dd, J = 11.2, 3.2 Hz, 1H, H5B'), 2.09 (s, 3H, OAc), 2.03 (s, 3H, CH₃ OAc), 1.09 (s, 9H, *t*Bu). $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) δ : 169.4, 169.3 (CO OAc), 135.54, 134.50 (Arom.), 132.7 (Cq Arom.), 129.74, 129.67, 127.68, 127.66 (Arom.), 92.6 (C1'), 82.4 (C4'), 74.6 (C2'), 70.7 (C3'), 63.2 (C5'), 26.6 (CH₃ *t*Bu), 20.39, 20.37 (CH₃ OAc), 19.0 (Cq *t*Bu). $[\alpha]_{\text{D}}^{25}$ (CHCl_3) = -53.8°. IR: 2933.1, 2858.7, 2112.2, 1751.8, 1214.1, 1230.2, 1109.6, 701.6, 502.0. HRMS $[\text{C}_{25}\text{H}_{31}\text{N}_3\text{O}_6\text{Si} + \text{Na}]^+$: 520.1866 found, 520.1874 calc.

N^{β} -(5-*tert*-Butyldiphenylsilyl-2,3-di-*O*-acetyl- α/β -D-ribofuranosyl)- N^{α} -benzyloxycarbonyl asparagine benzyl ester (**5a/5b**)



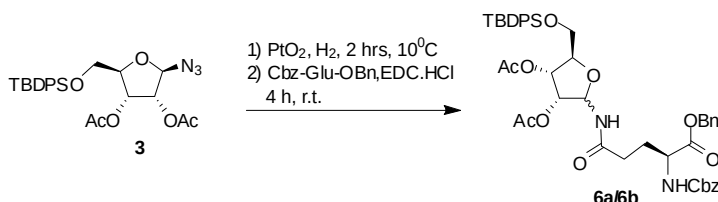
Platinum(IV)oxide (110 mg, 0.48 mmol) was added to a solution of azide **3** (1.24 g, 2.5 mmol) in EtOAc (20 mL) and purged with argon before hydrogen gas was bubbled through the reaction mixture for 2 hours at 10 °C, upon which TLC analysis (3:2, v/v, EtOAc/PE) showed complete disappearance of the starting material. The suspension was filtered through a pad of celite and concentrated. The residue was dissolved in DCM (30 mL) and EDC.HCl (0.53 g, 2.8 mmol) and Z-Asp-OBn (0.89 g, 2.5 mmol) were added under an argon atmosphere. After stirring for 4 hours the reaction mixture was concentrated. The residue was taken up in DCM and washed with sat. aq. NaHCO_3 and sat. aq. NaCl. The organic layer was dried with MgSO_4 , concentrated and the stereo isomers were separated using silica gel column chromatography (PE: EtOAc, 80/20 – 50/50). Compound **5a** was obtained as a white foam (0.65 g, 0.80 mmol, 32 %) along with the **5b** anomer (0.31 g, 0.38 mmol, 15%).

α -anomer **5a**: $^1\text{H-NMR}$ (400 MHz, $\text{MeOD-}d_4$) δ : 7.87 (s, 1H, NH), 7.68 (td, J = 7.4, 1.5 Hz, 4H, Arom.), 7.47 - 7.04 (m, 16H, Arom.), 6.04 (d, J = 4.9 Hz, 1H, H1'), 5.54 (t, J = 5.1 Hz, 1H, H3'), 5.44 (t, J = 5.0 Hz, 1H, H2'), 5.08, 5.16 (2x s, 4H, CH₂ Bn, CH₂ Cbz), 4.65 (t, J = 6.0 Hz, 1H, H4'), 4.24 - 3.96 (m, 1H, CH α -Asn), 3.77 (dd, J = 11.2, 2.8 Hz, 1H, H5A'), 3.72 (dd, J = 11.2, 2.8 Hz, 1H, H5B'), 2.96 - 2.70 (m, 2H, CH₂ β -Asn), 2.09, 2.06 (2x s, 6H, CH₃ OAc), 1.07 (s, 9H, *t*Bu). $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) δ : 170.6, 169.5 (CO α -Asn, CO β -Asn), 169.2, 169.0 (CO OAc), 158.2 (CO Cbz), 135.8 (Cq Arom.), 135.4 (Cq Arom.), 135.5 (Arom.), 132.6, 132.3 (Cq Arom.), 129.7, 128.4 - 127.7 (Arom.), 82.0 (C4'), 78.5 (C1'), 72.3 (C3'), 70.0 (C2'), 67.42, 66.9 (CH₂ Bn, CH₂ Cbz), 63.5 (C5'), 50.7 (CH α -Asn), 38.0 (CH₂ β -

Asn), 26.6 (CH₃ *t*Bu), 20.6, 20.2 (CH₃, OAc), 19.0 (Cq *t*Bu). α_D (CHCl₃) = + 35.8°. IR: 503.0, 698.4, 1112.8, 1211.9, 1505.7, 1733.7, 1747.7. LC-MS (50 - 90 % B in 15 min.) Rt = 9.70. HRMS [C₄₄H₅₀N₂O₁₁Si + H]⁺: 811.3263 found, 811.3211 calc.

β -anomer **5b**: ¹H-NMR (400 MHz, MeOD-*d*₄) δ : 7.88(s, 1H, NH), 7.76 - 7.55 (m, 4H, Arom.), 7.49 - 7.04 (m, 16H, Arom.), 5.60 (d, *J* = 5.8 Hz, 1H, H1'), 5.52 (t, *J* = 4.9 Hz, 1H, H3'), 5.16 (t, *J* = 5.6 Hz, 1H, H2'), 5.14, 5.07 (2x s, 4H, CH₂ Bn, CH₂ Cbz), 4.61 (t, *J* = 5.8 Hz, 1H, H4'), 4.05 (q, *J* = 4.0 Hz, 1H, CH α -Asn), 3.90 - 3.66 (m, 2H, H5'), 2.91 - 2.61 (m, 2H, CH₂ β -Asn), 2.07, 2.03 (2x s, 6H, CH₃ OAc), 1.04 (s, 9H, *t*Bu). ¹³C-NMR (100 MHz, CDCl₃) δ : 170.6, 170.1 (CO α -Asn, CO β -Asn), 169.8, 169.7 (CO OAc), 156.0 (CO Cbz), 135.4 (Cq Arom.), 135.3 (Arom.), 133.1, 132.0 (Cq Arom.), 130.0, 128.4 - 127.7 (Arom.), 82.0, 81.9 (C4', C1'), 74.1 (C2'), 71.2 (C3'), 67.2, 66.8 (CH₂ Bn, CH₂ Cbz), 63.3 (C5'), 50.2 (CH α -Asn), 37.1 (CH₂ β -Asn), 26.8 (CH₃ *t*Bu), 20.5, 20.4 (CH₃ OAc), 19.3 (Cq *t*Bu). α_D (CHCl₃) = +12.4°. IR: 503.3, 698.0, 1112.3, 1214.5, 1498.7, 1747.2. LC-MS (50 - 90 % B in 15 min.) Rt = 9.63. HRMS [C₄₄H₅₀N₂O₁₁Si + H]⁺: 811.3262 found, 811.3211 calc.

***N'*-(5-*tert*-Butyldiphenylsilyl-2,3-di-*O*-acetyl- α / β -D-ribose)-*N''*-benzyloxycarbonyl glutamine benzyl ester (6a/6b)**



Platinum(IV)oxide (440 mg, 1.9 mmol) was added to a solution of azide **3** (5 g, 9.6 mmol) in EtOAc (100 mL) and purged with argon before hydrogen gas was bubbled through the reaction mixture for 2 hours at 10 °C, upon which TLC analysis (3:2, v/v, EtOAc/PE) showed complete disappearance of the starting material. The suspension was filtered through a pad of celite and concentrated. The residue was dissolved in DCM (100 mL) and EDC.HCl (1.8 g, 9.4 mmol) and Z-Glu-OBn (3.0 g, 8.1 mmol) were added under an argon atmosphere. After stirring for 24 hours the reaction mixture was concentrated. The residue was taken up in DCM and washed with 1N HCl and sat. aq. NaCl. The organic layer was dried (MgSO₄), concentrated and the stereo isomers were separated using silica gel column chromatography (PE: EtOAc, 100/0 - 65/35). Compound **6a** was obtained as a white foam (3.0 g, 3.7 mmol, 46 %) along with **6b** (1.0 g, 1.2 mmol, 15%) and a side product due to migration of the acetyl to 1' NH and subsequent coupling of Z-Glu-OBn on the 2' OH.

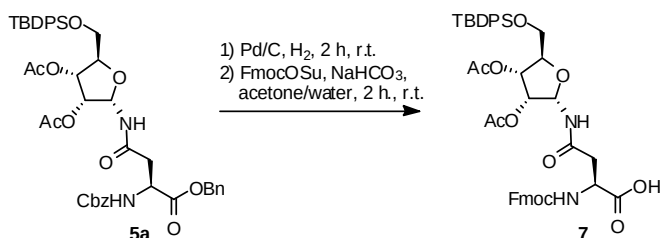
α -anomer **6a**: ¹H-NMR (CDCl₃) δ : 7.69 - 7.66, 7.42 - 7.37, 7.37 - 7.27 (3 x m, 20H, Arom.), 6.72 (d, *J* = 9.2 Hz, 1H, NH), 6.10 (dd, *J* = 5.2, 9.2 Hz, 1H, H1'), 5.82 (d, *J* = 7.6 Hz, 1H, NH), 5.58 (t, *J* = 3.2 Hz, 1H, H3'), 5.49 (t, *J* = 5.6 Hz, 1H, H2'), 5.14, 5.07 (2x s, 4H, CH₂ Bn, CH₂ Cbz), 4.42 (m, 1H, CH α -Gln), 4.10 (m, 1H, H4'), 3.71 (m, 2H, H5'), 2.33 (m, 2H, CH₂ β -Gln), 2.21 (m, 2H, CH₂ γ -Gln), 2.09, 2.07 (2x s, 6H, OAc), 1.07 (s, 9H, *t*Bu). ¹³C-NMR (CDCl₃) δ : 171.8, 171.5 (CO α -Gln, CO γ -Gln), 169.3, 169.0 (CO Ac), 156.0 (CO Cbz), 135.9, 135.4 (Cq Arom.), 135.5 (Arom.), 132.7, 132.4 (Cq Arom.), 129.6, 128.4 - 127.6 (Arom.), 81.6 (C4'), 78.6 (C1'), 72.2 (C3'), 70.1 (C2'), 67.1, 66.8 (CH₂ Bn, CH₂ Cbz), 63.5 (C5'), 53.4 (CH α -Gln), 32.4 (CH₂ β -Gln), 28.1 (CH₂ γ -Gln), 26.6 (*t*Bu), 20.5, 20.3 (CH₃ OAc), 18.9 (Cq *t*Bu). TLC R_f = 0.4 EtOAc/PE, 4:6, v/v). ESI-MS: *m/z* 825.20 [M+H]⁺. LC-MS (10 - 90 % B in 15 min), Rt = 9.94. IR: 3328, 2933, 2362, 1745, 1520, 1235, 1212, 1110, 1049, 741, 698, 604, 502. [α_D CHCl₃] = + 32.6°. HRMS [C₄₅H₅₂N₂O₁₁Si + H]⁺: 825.3419 found, 825.3413 calc.

β -anomer **6b**: ¹H-NMR (500 MHz, CDCl₃) δ : 7.68 - 7.65, 7.40 - 7.25 (2 x m, 20H, Arom.), 6.47, 6.46 (d, 1H, NH γ), 5.79 (dd, *J* = 5.7 Hz, 1H, H1'), 5.64 (s, 1H, NH α), 5.56 (dd, *J* = 5.3 Hz, 1H, H3'), 5.31 (t, *J* = 5.3 Hz, 1H, H2'), 5.12, 5.04 (2x s, 4H, CH₂ Bn), 4.31 - 4.27 (m, 1H, CH α), 4.12 - 4.10 (m, 1H, H4'),

3.87 (dd, $J = 11.2, 2.8$ Hz, 1H, H_{5A}'), 3.81 (dd, $J = 11.6, 2.0$ Hz, 1H, H_{5B}'), 2.08, 2.04 (2x s, 6H, OAc), 1.94 – 1.91 (m, 2H, $CH_2 \beta$), 2.13 – 2.10, 1.85 – 1.80 (2x m, 2H, $CH_2 \gamma$), 1.10 (s, 9H, *t*Bu). ^{13}C -NMR (100 MHz, CDCl_3) δ : 172.0, 171.6 (CO Ac), 169.8, 169.7 (CO α -Gln, CO γ -Gln), 156.1 (CO Cbz), 136.0 (Cq Arom.), 135.5, 135.4 (Arom.), 135.1, 133.1, 132.2 (Cq Arom.), 130.0, 129.93, 128.5, 128.3, 128.2, 128.0, 127.9 (Arom.), 81.9 ($C1'$, $C4'$), 74.0 ($C2'$), 71.4 ($C3'$), 67.2, 66.9 (CH_2 Bn, CH_2 Cbz), 63.5 ($C5'$), 53.3 (CH α -Gln), 31.8 ($CH_2 \beta$ -Gln), 27.7 ($CH_2 \gamma$ -Gln), 26.8 (*t*Bu), 20.6, 20.4 (OAc), 19.3 (Cq *t*Bu). TLC $R_f = 0.5$ EtOAc/PE, 4:6, v/v). ESI-MS: m/z 825.20 $[M+H]^+$. LC-MS (10 - 90 % B in 15 min), $R_t = 9.88$. IR: 3335.3, 2934.0, 1745.2, 1519.9, 1240.8, 1214.4, 1110.5, 1064.7, 738.9, 699.0, 502.2. $[\alpha_D^{25} \text{CHCl}_3] = -0.8^\circ$. HRMS $[C_{45}H_{52}N_2O_{11}Si + H]^+$: 825.3422 found, 825.3413 calc.

Migration side-product: ^1H -NMR (400 MHz, CDCl_3) δ : 7.61 (dd, 4H), 7.51 – 7.18 (m, 6H), 6.51 (d, $J = 9.0$ Hz, 1H, NH), 5.62 (d, $J = 14.1$ Hz, 1H, $H3'$), 5.34 (t, $J = 9.3$ Hz, 1H, $H1'$), 5.08 (2x s, 4H, CH_2 Bn/Cbz), 4.60 (d, $J = 9.5$ Hz, 1H, $H2'$), 4.33 (t, $J = 7.3$ Hz, 1H, CH α), 3.96 – 3.77 (m, 1H, $H4'$), 3.72 (t, $J = 10.8$ Hz, 1H, $H5A'$), 3.36 (dd, $J = 10.9, 5.1$ Hz, 1H, $H5B'$), 2.21 (s, 3H, NAc), 2.16 (s, 3H, $CH_2 \gamma$, $CH_2 \beta$), 1.90 (s, 3H, OAc), 1.85 (d, $J = 6.6$ Hz, 1H, $CH_2 \gamma$), 1.01 (s, 9H, *t*Bu). ^{13}C -NMR (100 MHz, CDCl_3) δ : 172.5, 171.7 (CO Ac), 170.7, 170.2 (CO α -Gln, CO γ -Gln), 156.4 (CO Cbz), 136.1, 135.8, 135.6, 135.1, 133.2, 132.5, 130.2, 128.72, 128.65, 128.59, 128.5, 128.3, 128.2, 128.0, 127.9 (Arom.), 75.3 ($C1'$), 70.6 ($C3'$), 68.8 ($C2'$), 67.5, 67.23 (CH_2 Bn), 67.15 ($C4'$), 65.6 ($C5'$), 53.3 (C α), 32.3 (C β), 28.5 (C γ), 26.8 (CH $_3$ *t*Bu), 21.0, 20.7 (CH $_3$ Ac), 19.1 (Cq *t*Bu). TLC $R_f = 0.6$ EtOAc/PE, 4:6, v/v). ESI-MS: m/z 825.33 $[M+H]^+$. LC-MS (10 - 90 % B in 15 min), $R_t = 12.09$. HRMS $[C_{45}H_{52}N_2O_{11}Si + H]^+$: 825.3420 found, 825.3413 calc.

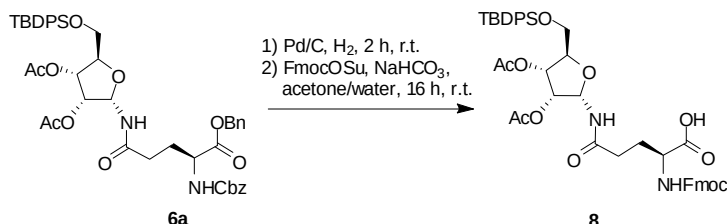
N^B -(5-*tert*-Butyldiphenylsilyl-2,3-di-*O*-acetyl- α -D-ribose)- N^α -fluorenylmethoxycarbonyl asparagine (7)



Amino acid **5a** (1.25 g, 1.54 mmol) was dissolved in MeOH (30 mL), Pd/C (10%, 190 mg) and a drop of conc. HCl were added after purging with argon. A flow of hydrogen gas was passed through the reaction mixture for 2 hours. LC-MS analysis (CN-Column, 10 - 90 % B in 26 min.) showed complete consumption of the starting material and formation of the unprotected amino acid. Next, NaHCO_3 (190 mg, 2.3 mmol) was added and the reaction mixture was filtered over a pad of celite. The solvent was removed and the residue was dissolved in H_2O /acetone (40 mL, 1:1, v/v). Fmoc-OSu (0.51 g, 1.5 mmol) and NaHCO_3 (0.13 g, 1.5 mmol) were added after which the reaction mixture was stirred at room temperature for 2 hours. The mixture was acidified to pH = 2 and acetone was removed under reduced pressure. The resulting solution was extracted with DCM, washed with 5% citric acid, H_2O , dried with MgSO_4 and concentrated. Purification using column chromatography (99.9 : 0.1 DCM/AcOH - 89.9/10.1 DCM/MeOH/AcOH) afforded the title compound as a white foam (1.1 g, 1.30 mmol, 84%). ^1H -NMR (400 MHz, $\text{MeOD}-d_4$) δ : 7.85 (s, 1H, NH), 7.77 – 6.91 (m, 18H, Arom.), 6.05 (d, $J = 4.9$ Hz, 1H, $H1'$), 5.51 (t, $J = 5.0$ Hz, 1H, $H3'$), 5.43 (t, $J = 5.0$ Hz, 1H, $H2'$), 4.41 (d, $J = 6.0$ Hz, 1H, CH α -Asn), 4.36 – 4.23 (m, 2H, CH_2 Fmoc), 4.14 (t, $J = 6.7$ Hz, 1H, CH Fmoc), 4.07 (s, 1H, $H4'$), 3.65 – 3.63 (m, 2H, $H5'$), 2.97 – 2.61 (m, 2H, $CH_2 \beta$ -Asn), 2.06, 2.05 (2x s, 6H, CH_3 OAc), 1.01 (s, 9H, *t*Bu). ^{13}C -NMR (100 MHz, $\text{MeOD}-d_4$) δ : 173.2, 171.0 (CO β -Asn, CO α -Asn), 170.2, 170.1 (CO OAc), 156.9 (CO Fmoc), 143.8, 143.8 (Cq Fmoc), 137.5 (Fmoc), 135.4, (Cq TBDPS), 129.7 – 119.6 (Arom. TBDPS, Fmoc), 81.2 ($C4'$), 79.2 ($C1'$), 71.5 ($C3'$), 70.7 ($C2'$), 66.8 ($C5'$), 63.5 (CH_2 Fmoc), 50.8 (CH α -Asn), 46.9 (CH Fmoc), 37.5 ($CH_2 \beta$ -Asn),

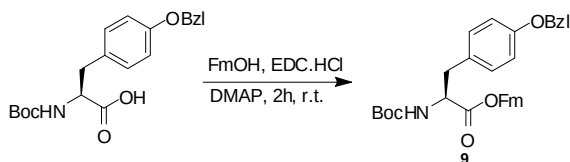
26.1 (CH₃ *t*Bu), 19.2, 19.1 (CH₃ OAc), 18.7 (Cq *t*Bu). α_D (CDCl₃) = + 56.2°. LC-MS (10 – 90 % B in 15 min.), Rt = 11.34 min. IR: 503.5, 701.4, 1112.5, 1212.2, 1505.7, 1747.5. HRMS [C₄₄H₄₈N₂O₁₁Si + H]⁺: 809.3106 found, 809.3055 calc.

***N*^ε-(5-*tert*-Butyldiphenylsilyl-2,3-di-*O*-acetyl- α -D-ribose)-*N*^α-fluorenylmethoxycarbonyl glutamine (8)**



Amino acid **6a** (0.2 g, 0.24 mmol) was dissolved in MeOH (5 mL), Pd-C (10%) (30 mg) and a drop of conc. HCl were added after which hydrogen gas was passed through the reaction mixture for 5 hours. LC-MS analysis (CN-Column, 10-90 % B in 26 min.) showed complete consumption of the starting material and formation of the unprotected amino acid. After neutralisation using NaHCO₃ and filtration over a pad of celite the solvent was concentrated and the residue was dissolved in H₂O/ acetone/ dioxane (3 mL, 1:1:2, v/v/v). Fmoc-OSu (80 mg, 0.24 mmol) and NaHCO₃ (20 mg, 0.24 mmol) were added after which the reaction mixture was stirred for 16 hours. The mixture was acidified to pH = 2 and acetone was removed under reduced pressure. The resulting solution was extracted with DCM, washed with 5% citric acid, H₂O, dried with MgSO₄ and concentrated. Purification using column chromatography (99.9 : 0.1 DCM/AcOH – 99.4 : 0.1 : 0.5 DCM/ AcOH/ MeOH) afforded the title compound as a white foam (0.19 g, 0.23 mmol, 97%). ¹H-NMR (400 MHz, MeOD-*d*₄) δ : 7.90 (s, 1H, NH), 7.81 - 7.14 (m, 18H, Arom.), 6.09 (d, *J* = 5.0 Hz, 1H, H1'), 5.57 (t, *J* = 5.0 Hz, 1H, H3'), 5.48 (t, *J* = 5.1 Hz, 1H, H3'), 4.35 - 4.22 (m, 5H, CH Fmoc, H5', H4', CH α), 3.79 - 3.71 (m, 2H, CH₂ Fmoc), 3.35 - 3.33 (m, 2H, CH₂ β), 2.24 - 2.22 (m, 1H, CH₂ γ), 2.09, 2.07 (2x s, 6H, CH₃ OAc), 2.08 - 1.89 (m, 1H, CH₂ γ), 1.06 (s, 9H, *t*Bu). ¹³C-NMR (100 MHz, MeOD-*d*₄) δ : 173.9, 173.7 (CO γ -Gln, CO α -Gln), 170.2, 170.0 (CO OAc), 157.2 (CO Fmoc), 143.9, 143.8, 141.2 (Cq Arom.), 135.4, 135.3 (Arom.), 132.7, 132.6 (Cq Arom.), 129.7 - 119.6 (Arom. Fmoc/TBDPS), 81.1 (C4'), 79.2 (C1'), 71.5 (C3'), 70.7 (C2'), 66.7 (C5'), 63.4 (CH₂ Fmoc), 53.4 (CH α -Gln), 47.0 (CH Fmoc), 32.0 (CH₂ β -Gln), 27.3 (CH₂ γ -Gln), 26.0 (CH₃ *t*Bu), 19.3, 19.1 (CH₃ OAc), 18.6 (Cq *t*Bu). ESI-MS: *m/z* 823.27 [M+H]⁺. LC-MS (50 - 90 % B in 15 min), Rt = 8.92. IR: 2933.1, 1742.5, 1519.6, 1212.6, 1108.1, 740.0, 702.8, 502.3. α_D (CHCl₃) = + 33.4°. HRMS [C₄₅H₅₀N₂O₁₁Si + H]⁺: 823.3263 found, 823.3256 calc.

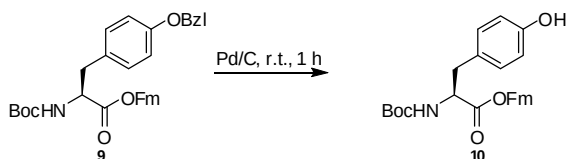
***N*^ε-*tert*-Butoxycarbonyl-4-*O*-benzyl-L-tyrosine fluorenylmethyl ester (9)**



Boc-Tyr(Bzl)-OH (0.74 g, 2.0 mmol) was coevaporated with MeCN and dissolved in DCM (40 mL) under an argon atmosphere. To this solution were added 9-fluorenylmethanol (0.43 g, 2.2 mmol), EDC.HCl (0.42 g, 2.2 mmol) and DMAP (32 mg, 0.2 mmol). The solution was stirred at room temperature for 2 hours, washed with 5% citric acid and H₂O. The combined organic layers were dried using MgSO₄ and concentrated. The crude compound was purified by silica gel column chromatography (PE/EtOAc, 95/5)

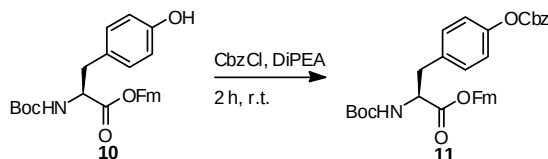
to give compound **9** as a white powder (0.84 g, 76%). ¹H-NMR (400 MHz, CDCl₃) δ: 7.80 - 7.69 (m, 2H, Arom. Bn/Fm), 7.50 (t, *J* = 8.1 Hz, 2H, Arom. Bn/Fm), 7.42 - 7.24 (m, 9H, Arom. Bn/Fm), 6.97 (d, *J* = 8.5 Hz, 2H, Arom. Tyr.), 6.83 (d, *J* = 8.5 Hz, 2H, Arom. Tyr.), 5.06 (s, 1H, NH α-Tyr), 4.91 (s, 2H, CH₂ Bn), 4.61 (dd, *J* = 13.9, 6.0 Hz, 1H, CH α-Tyr), 4.48 - 4.31 (m, 2H, CH₂ Fm), 4.13 (t, *J* = 6.1 Hz, 1H, CH Fm), 3.00 - 2.89 (m, 2H, CH₂ β-Tyr), 1.43 (s, 9H, *t*Bu). ¹³C-NMR (100 MHz, CDCl₃) δ: 171.9 (CO α-Tyr), 157.8 (Cq Tyr.), 155.1 (CO Boc), 143.5, 143.4, 141.2 (Cq Fm), 136.9 (Cq Bn), 130.24 (Arom. Tyr.), 128.5 (Arom. Bn/Fm), 127.9 (Cq Tyr.), 127.3, 127.1 (Arom. Bn/Fm), 125.0, 120.0 (Arom. Fm), 114.9 (Arom. Tyr.), 79.8 (Cq Boc), 69.8 (CH₂ Bn), 67.0 (CH₂ Fm), 54.6 (CH α Tyr.), 46.6 (CH Fm), 37.4 (CH₂ β Tyr.), 28.3 (CH₃ Boc). IR (cm⁻¹): 3348, 2931, 1733, 1686, 1527, 1512, 1242, 1163, 1048, 736, 549. α_D (CHCl₃) = 0.2°. LC-MS (50 - 90 % B in 15 min), Rt = 9.62. HRMS [C₃₅H₃₅NO₅ + H]⁺: 550.2585 found, 550.2588 calc.

N^α-*tert*-Butoxycarbonyl-L-tyrosine fluorenylmethyl ester (**10**)



Compound **9** (0.58 g, 1.1 mmol) was coevaporated with MeCN and dissolved in *n*-propanol/EtOAc (30 mL, 9:1 v/v). The solution was purged with argon prior to addition of Pd-C (10%, 100 mg) and conc. HCl (89 μL, 1.17 mmol). Hydrogen gas was bubbled through the solution for 1 hour after which the reaction mixture was filtered over celite, concentrated and purified by silica gel column chromatography (PE/EtOAc, 90/10 → 80/20) to give compound **10** as a white powder (0.36 g, 73%). ¹H-NMR (400 MHz, CDCl₃) δ: 7.70 (d, *J* = 6.7 Hz, 2H, Arom. Fm), 7.48 (t, *J* = 8.0 Hz, 2H, Arom. Fm), 7.36 - 7.32 (m, 2H, Arom. Fm), 7.28 - 7.23 (m, 2H, Arom. Fm), 6.92 - 6.86 (m, 2H, Arom. Fm), 6.74 - 6.69 (m, 2H, Arom. Tyr), 5.11 (d, *J* = 8.4 Hz, 1H, NH α-Tyr), 4.59 (dd, *J* = 13.8, 6.2 Hz, 1H, CH α-Tyr), 4.46 - 4.24 (m, 2H, CH₂ Fm), 4.11 (t, *J* = 6.7 Hz, 1H, CH Fm), 2.97 - 2.87 (m, 2H, CH₂ β-Tyr), 1.41 (s, 9H, *t*Bu). ¹³C-NMR (100 MHz, CDCl₃) δ: 172.2 (CO α-Tyr), 155.4, (CO Boc), 143.5, 143.3, 141.3, 141.3 (Cq Fm), 130.4 (Arom. Tyr), 127.9, 127.9 (Arom. Fm), 127.4 (Cq Tyr), 127.2, 125.1 (Arom. Fm), 120.1, 115.6 (Arom. Tyr), 80.3 (Cq Boc), 67.2 (CH₂ Fm), 54.7 (CH α Tyr), 46.7 (CH Fm), 37.5 (CH₂ β Tyr), 28.3 (CH₃ Boc). IR: 3367, 2979, 1686, 1516, 1244, 1163, 530. α_D (CHCl₃) = -1.0°. LC-MS (50 - 90 % B in 15 min), Rt = 5.62. HRMS [C₂₈H₂₉NO₅ + H]⁺: 460.2116 found, 460.2118 calc.

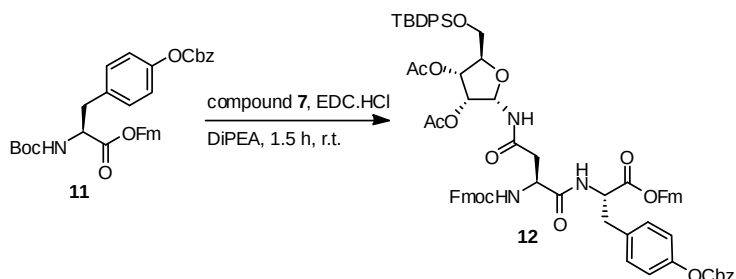
N^α-*tert*-Butoxycarbonyl-4-*O*-benzyloxycarbonyl-L-tyrosine fluorenylmethyl ester (**11**)



Compound **10** (0.31 g, 0.68 mmol) was coevaporated with MeCN and dissolved in DCM (3.5 mL) under an argon atmosphere. To this solution were added benzylchloroformate (0.1 mL, 0.75 mmol) and DiPEA (0.12 mL, 0.75 mmol). The solution was stirred at room temperature for 2 hours, washed with 5% citric acid and H₂O. The combined organic layers were dried and concentrated. The crude compound was purified by silica gel column chromatography (PE/EtOAc, 90/10 → 80/20) to give compound **11** as a white powder (0.34 g, 84%). ¹H-NMR (400 MHz, CDCl₃) δ: 7.73 (dd, *J* = 7.5, 2.0 Hz, 2H, Arom. Fm), 7.52 (t, *J* = 7.6 Hz, 2H, Arom. Fm), 7.43 - 7.32 (m, 7H, Arom. Fm/Cbz), 7.28 (t, *J* = 7.5 Hz, 2H, Arom. Fm/Cbz), 7.11 - 7.01 (m, 4H, Arom. Tyr), 5.22 (s, 2H, CH₂ Cbz), 5.00 (d, *J* = 8.0 Hz, 1H, NH α-Tyr), 4.61

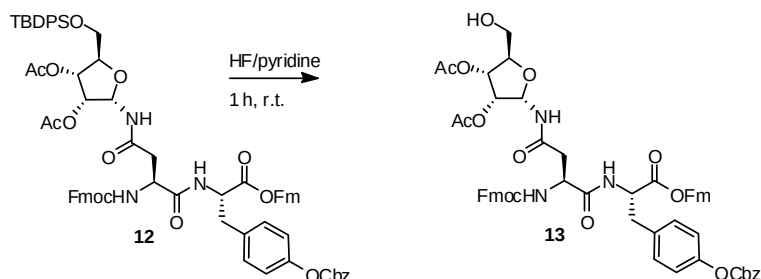
(dd, $J = 13.2, 6.1$ Hz, 1H, CH α -Tyr), 4.50 - 4.27 (m, 2H, CH₂ Fm), 4.14 (t, $J = 6.6$ Hz, 1H, CH Fm), 3.02 (dd, $J = 14.0, 6.0$ Hz, 1H, CH₂ β -Tyr), 2.89 (dd, $J = 14.0, 6.0$ Hz, 1H, CH₂ β -Tyr), 1.41 (s, 9H, tBu). ¹³C-NMR (100 MHz, CDCl₃): δ : 171.4 (CO α -Tyr), 154.9, 153.3 (CO Cbz, CO Boc), 149.9 (Cq Tyr) 143.2, 143.1, 141.0, 141.0 (Cq Fm), 134.5 (Cq Cbz.), 133.7 (Cq Tyr), 130.2 (Arom. Tyr), 128.5, 128.4, 127.6, 126.9 (Arom. Fm/Cbz), 124.7, 120.8 (Arom. Fm), 119.8 (Arom. Tyr), 79.7 (Cq Boc), 70.0 (CH₂ Cbz), 66.8 (CH₂ Fm), 54.1 (CH α Tyr), 46.4 (CH Fm), 37.1 (CH₂ β Tyr), 28.0 (CH₃ Boc). α_D (CHCl₃) = + 0.4°. IR: 3351, 1748, 1683, 523, 1267, 1161, 738, 545. LC-MS (50 - 90 % B in 15 min), Rt = 9.34. HRMS [C₃₆H₃₅NO₇ + H]⁺: 616.2301 found, 616.2305 calc.

***N*^β-(5-*tert*-Butyldiphenylsilyl-2,3-di-*O*-acetyl- α -D-ribose)-*N* ^{α} -fluorenylmethoxycarbonyl asparaginyl-(4-*O*-benzyloxycarbonyl)-L-tyrosine fluorenylmethyl ester (**12**)**



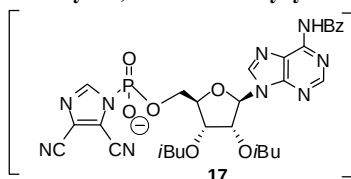
Compound **11** (0.33 g, 0.55 mmol) was dissolved in 3 M HCl/EtOAc and stirred at room temperature for 30 min. The reaction mixture was concentrated and coevaporated with toluene. The residue was dissolved in DCM (10 mL) and a solution of **7** (0.5 mmol) in DCM (2.5 mL) was added, followed by EDC·HCl (0.12 g, 0.6 mmol) and DiPEA (0.6 mmol, 0.1 mL). After stirring at room temperature for 1.5 hours the reaction mixture was washed with 5% citric acid and water. The organic layer was dried with MgSO₄ and concentrated. The crude compound was purified by silica gel column chromatography (PE/EtOAc, 100/0 → 35/65) to give compound **12** as a white powder (0.59 g, 0.46 mmol, 84%). ¹H-NMR (400 MHz, CDCl₃): δ 7.76 - 7.71 (m, 4H), 7.67 - 7.63 (m, 4H), 7.58 - 7.45 (m, 5H), 7.44 - 7.24 (m, 18H), 7.15 - 6.91 (m, 4H, Arom. Tyr), 6.69 (d, $J = 8.7$ Hz, 1H, NH β -Tyr), 6.23 (d, $J = 5.9$ Hz, 1H, NH α -Asn), 6.05 (dd, $J = 9.3, 5.6$ Hz, 1H, H1'), 5.61 - 5.52 (m, 1H, H3'), 5.47 (t, $J = 5.4$ Hz, 1H, H2'), 5.16 (s, 2H, CH₂ Cbz), 4.76 (dd, $J = 13.0, 6.5$ Hz, 1H, CH α -Asn), 4.53 - 4.51 (m, 1H, CH α -Tyr), 4.45 - 4.47 (m, 2H, CH₂ Fmoc), 4.34 - 4.31 (m, 2H, CH₂ Fm), 4.19 - 4.15 (m, 2H, CH Fmoc, CH Fm), 4.12 - 4.11 (m, 1H, H4'), 3.72 - 3.68 (m, 2H, H5'), 3.15 - 2.55 (m, 4H, CH₂ β -Asn, β -Tyr), 2.09, 2.07 (2x s, 6H, CH₃ OAc), 1.06 (s, 9H, tBu). ¹³C-NMR (100 MHz, CDCl₃): δ 171.3 (CO α -Tyr), 170.9, 170.6 (CO OAc), 169.6, 169.2 (CO α -Asn, CO β -Asn), 156.2 (CO Fmoc), 153.5 (CO Cbz), 150.3 (Cq Tyr), 143.8, 143.7, 143.5, 143.4, 141.4, 141.3 (Cq Fm, Cq Fmoc), 135.7 (Arom. TBDPS), 134.8 (Cq Cbz), 133.6 (Cq Tyr), 132.9, 132.7 (Cq TBDPS), 130.4, 129.9, 128.8, 128.7, 128.5, 128.0, 127.9, 127.8, 127.3, 127.2, 125.2, 125.0, 121.2, 120.0 (Arom. Fmoc, Arom. Fm, Arom. TBDPS, Arom. Cbz), 82.2 (C4'), 78.7 (C1'), 72.4 (C3'), 70.3 (CH₂ Cbz), 70.2 (C2') 67.5, 67.2 (CH₂ Fmoc, CH₂ Fm), 63.7 (C5'), 53.7 (C α Tyr), 51.3 (C α Asn), 47.1, 46.7 (CH Fmoc, CH Fm), 38.0 (CH₂ β -Asn), 37.0 (CH₂ β -Tyr), 26.8 (CH₃ TBDPS), 20.8, 20.6 (CH₃ OAc), 19.2 (Cq tBu). α_D (CHCl₃) = +29.6°. IR : 3278, 2933, 2360, 1747, 1696, 1648, 1534, 1211, 1104, 737, 699, 503. HRMS [C₇₅H₇₃N₃O₁₅Si + H]⁺: 1284.4903 found, 1284.4883 calc.

***N*^β-(2,3-di-*O*-acetyl- α -D-ribsyl)-benzyloxycarbonyl-L-tyrosine fluorenylmethyl ester (**12**)** ***N*^α-fluorenylmethoxycarbonyl asparaginy-(4-*O*-benzyloxycarbonyl)-L-tyrosine fluorenylmethyl ester (**13**)**



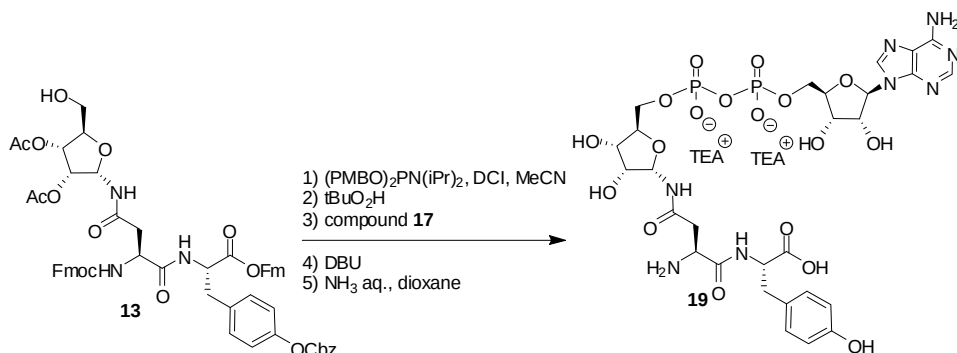
Fully protected **12** (0.61 g, 0.47 mmol) was dissolved in pyridine (10 mL) and treated with a HF/pyridine solution (1.0 mL) for 1 hour at room temperature. The reaction was quenched with sat. aq. NaHCO₃ and extracted with DCM. After washing with water the organic layer was dried with MgSO₄ and concentrated. The crude compound was purified by silica gel column chromatography (DCM/MeOH, 99/1) to give compound **13** as a white powder (0.42 g, 0.4 mmol, 85%). ¹H-NMR (400 MHz, CDCl₃): δ 7.77 - 7.72, 7.58 - 7.51, 7.42 - 7.29 (m, 21H, Arom. Fmoc, Arom. Fm, Arom. Cbz), 7.08 - 7.04 (m, 4H, Arom. Tyr.), 6.69 (d, *J* = 8.8 Hz, 1H, NH), 6.23 (d, *J* = 7.2 Hz, 1H, NH), 5.95 (m, 1H, H1'), 5.40 - 5.38 (m, 1H, H2'), 5.33 - 5.29 (m, 1H, H3'), 5.25 (s, 2H, CH₂ Cbz), 4.79 - 4.77 (m, 1H, H α -Asn), 4.56 - 4.53 (m, 1H, H α -Tyr), 4.51 - 4.49 (m, 2H, CH₂ Fmoc), 4.38 - 4.36 (m, 2H, CH₂ Fm), 4.24 - 4.20 (m, 2H, CH Fmoc, CH Fm), 4.12 (m, 1H, H4'), 3.87 - 3.64 (m, 2H, H5'), 3.03 - 2.56 (m, 4H, CH₂ β Asn, CH₂ β Tyr), 2.14, 2.12 (2x s, 6H, CH₃ OAc). ¹³C-NMR (100 MHz, DMSO-*d*₆): δ 172.0, 171.6, 170.0, 169.9, 169.8 (CO α -Tyr, CO OAc, CO α -Asn, CO γ -Asn), 156.2 (CO Fmoc), 153.4 (CO Cbz), 149.9 (Cq Tyr.), 144.3, 144.1, 143.9, 141.3, 141.2, 141.1 (Cq Fm, Cq Fmoc), 135.48 (Cq Cbz), 135.45 (Cq Tyr.), 130.7, 129.0, 128.8, 128.2, 128.1, 127.8, 127.7, 127.6, 125.8, 125.6, 121.4, 120.6 (Arom. Fmoc, Arom. Fm, Arom. Cbz), 80.5 (C4'), 79.1 (C1'), 71.5 (C3'), 70.7 (C2'), 70.2 (CH₂ Cbz), 66.7, 66.3 (CH₂ Fmoc, CH₂ Fm), 61.4 (C5'), 54.2 (C α Tyr), 51.6 (C α Asn), 47.0, 46.7 (CH Fmoc, CH Fm), 38.0 (CH₂ β -Asn), 36.0 (CH₂ β -Tyr), 20.9 (CH₃ OAc). α_D (CHCl₃) = + 26.8°. IR: 3288, 2924, 2360, 1746, 1652, 1538, 1269, 1212, 1046, 736. LC-MS (10 - 90 % B in 15 min), Rt = 11.30. HRMS [C₅₉H₅₅N₃O₁₅ + H]⁺: 1046.318 found, 1046.3705 calc.

***N*⁶-Benzoyl-2',3'-di-*O*-isobutyryladenyl-5-phosphorimidazolidate **17**¹⁹**

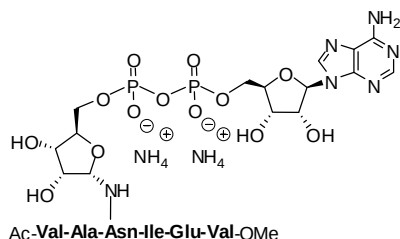


To a cooled (0 °C) solution of **16** (0.46 g, 1.0 mmol)¹⁹ and **14** (0.63 g, 1.3 mmol) in MeCN (15 mL) was slowly added a solution of DCI (0.15 g, 1.25 mmol) in MeCN (5 mL). After 20 min complete disappearance of **14** (³¹P-NMR; 148.5) and formation of the corresponding triester (³¹P-NMR; 141.8) was observed. Next, a solution of iodine in pyridine (2.0 mL, 0.5 M) was added to the reaction mixture until the solution kept its brown color. After 30 min of additional stirring the reaction mixture containing imidazolidate **17** was used without further purification.

Compound 19: H-Asn(^N₆ [5'-adenosine diphosphate- α -D-ribose])₂-Tyr-OH bis triethylammonium salt.



Ribosylated dipeptide **13** (100 mg, 0.1 mmol) was dissolved in dioxane/ MeCN (4 mL, 1:1 v/v) and cooled to 0 °C. Di(*p*-methoxybenzyl)-*N,N*-diisopropylphosphoramidite (43 μ L, 0.11 mmol) and a solution of DCI (13 mg, 0.11 mmol) in 200 μ L MeCN were added after which the reaction mixture was kept at 0 °C for 1 hour. 100 μ L of *t*BuO₂H (5.5 M solution in nonane) was added and stirred for an additional hour at 0 °C. After trituration with hexane, the residue was treated with 3% TFA in DCM to yield intermediate **15**. *N*⁶-Benzoyl-2', 3'-di-*O*-isobutyryladenyl-5-phosphorimidazolidate **17** (0.2 mmol) was added to the reaction mixture and stirred at room temperature for 16 hours. A silica column with isocratic elution (30% MeOH/DCM) and RP-HPLC purification 80 – 90 % MeOH in 20 mM NH₄OAc over 5 CV yielded **18** (26.2 mg, 15.4 μ mol, 15.4% based on **13**). To a solution of **18** (14.5 mg, 8.5 μ mol) in dioxane/MeCN (1:1, 2 ml) 5 eq. DBU (43 μ mol, 6.5 μ l) was added and reacted for 30 min at room temperature. A solution of aqueous ammonia /dioxane (2:1, 6 ml) was added and reacted for 16 hours at room temperature. After concentration the title compound was obtained by means of gel filtration using a HW40-column and aqueous triethyl ammoniumacetate as eluent, followed by lyophilisation of the appropriate fractions (4.1 mg, 4.9 μ mol, 58% based on **18**). ³¹P-NMR (161 MHz, D₂O): δ -10.45, -10.58, -10.75, -10.88. LC-MS (0 - 50 % MeCN in 10 mM NH₄OAc over 15 min), Rt = 4.0. HRMS [C₂₈H₃₈N₈O₁₈P₂ + H]⁺: 837.1861 found, 837.1852 calc.

Compound 25; Ac-Val-Ala-Asn(*N*^r [5'-adenosine diphosphate- α -D-ribosyl])-Ile-Glu-Val-OMe bis ammonium salt.**25**

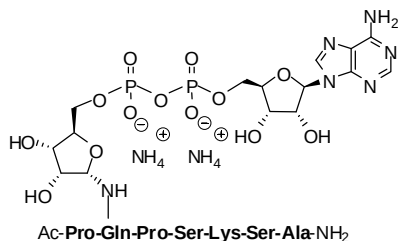
Intermediate **20** was prepared starting from Tentagel S HMBA resin (100 mg, 23 μ mol) via standard Fmoc based solid phase peptide synthesis, using BOP/HOBt as coupling reagent and a five fold excess of the appropriate amino acid, using Fmoc-Val-OH, Fmoc-Ala-OH, compound **7**, Fmoc-Ile-OH and Fmoc-Glu(Dmab)-OH. An analytical sample was deprotected and cleaved off the resin using the method described below, to determine quality of **20**. The resin was treated with 0.36 mL HF/pyridine diluted with pyridine (3 mL) (0 $^{\circ}$ C $\rightarrow$ r.t., 75 min). The resin was filtered and washed with pyridine and MeCN to give intermediate **21**.

The resin (80 mg, 18 μ mol) was coevaporated with MeCN and subsequently MeCN (3 mL), di(*p*-methoxybenzyl)-*N,N*-diisopropylphosphoramidite (0.23 mmol, 90 μ L) and a solution of DCI (0.23 mmol, 27 mg) in 1 mL MeCN were added after which the reaction mixture was kept at 0 $^{\circ}$ C for 1 hour. A 0.5 M solution of I_2 in pyridine was added until the reaction mixture kept the dark brown colour and stirred for an additional hour at 0 $^{\circ}$ C.

*N*⁶-benzoyl-2', 3'-di-*O*-isobutyryladenyl-5'-phosphate monoester **23** (0.23 mmol, 0.14 g) was added to the reaction mixture and stirred at room temperature for 24 hours.

The resin was filtered and washed with MeCN and treated with 2% hydrazine hydrate in DMF (3 times 5 min.), followed by 5 mM NaOH in MeOH/H₂O for 3 hrs. The resin was washed with MeCN and treated with 7N NH₃/MeOH (10 mL) for 16 hours. Subsequent filtration into cold Et₂O and centrifuging (6.5 min at 4000 RPM) followed by decantation of the supernatant resulted in the crude title compound. RP-HPLC using a gradient of 0 - 60% MeCN in 1% TFA in H₂O followed by a gradient RP-HPLC using a gradient of 0 - 60% MeCN of 10 mM NH₄OAc, followed by lyophilisation of the appropriate fractions yielded the title compound **25** (0.78 mg, 0.6 μ mol, 3.4% based on initial loading of the resin). ³¹P-NMR (161 MHz, D₂O): δ -11.04, -11.17, -11.31, -11.44. (EDTA was added to sharpen the signals). LC-MS (0 - 50 % MeCN [1% TFA] over 15 min), Rt = 8.07. HRMS [C₄₆H₇₄N₁₂O₂₄P₂ + H]⁺: 1241.4505 found, 1241.4486 calc.

Compound 32: Ac-Pro-Gln-(*N*^δ [5'-adenosine diphosphate- α -D-ribose])l)-Pro-Ser-Lys-Ser-Ala-NH₂ bis ammonium salt.



32

Starting from Tentagel S HMBA resin (150 mg, 65 μ mol) using standard Fmoc based solid phase peptide synthesis intermediate **26** was prepared using Fmoc-Pro-OH, Fmoc-Ser(Trt)-OH, Fmoc-Lys(TFAc)-OH and Fmoc-Ala-OH. The resin was repeatedly treated with 5% TFA in DCM until no yellow colour appeared upon addition of the fresh TFA solution. After washing with DCM, acetic anhydride (0.5 M, 200 μ L) and DMAP (0.05 M, 16 mg) in 2 mL NMP were added and the resin was shaken for 25 min. The resin was washed with NMP and DCM prior to coupling of the last two amino acids; compound **8**, and Fmoc-Pro-OH using the standard cycle to obtain **28**. An analytical sample was deprotected and cleaved off the resin using the method described below, to determine the quality.

The resin was coevaporated with MeCN and treated with a HF/pyridine (0.9 mL) diluted with pyridine (7.5 mL) (0 $^{\circ}$ C $\rightarrow$ r.t., 75 min). The resin was filtered and washed with pyridine and MeCN to obtain intermediate **29**.

The resin was coevaporated with MeCN and subsequently MeCN (8 mL), di(*p*-methoxybenzyl)-*N,N*-diisopropylphosphoramidite (10 eq., 250 μ L), dry 3 Å molsieves and a solution of DCI (10 eq., 75 mg) in 1 mL MeCN were added after which the reaction mixture was kept at 0 $^{\circ}$ C for 1 hour. Next the phosphite was oxidized using 5.5 M *t*BuO₂H (1.1 mL) in nonane for 1 hour at room temperature followed by addition of 3% TFA for 10 min. to deprotect the phosphotriester.

*N*⁶-Benzoyl-2', 3'-di-*O*-isobutyryladenyl-5-phosphorimidazolide **17** (0.4 mmol, 6 eq.) in 1 mL MeCN was added to the reaction mixture and stirred at room temperature for 16 hours.

The resin was filtered, washed with MeCN and treated with 7N NH₃/MeOH (10 mL) for 16 hours. Subsequent filtration into cold Et₂O and centrifuging (6.5 min at 4000 RPM) followed by decantation of the supernatant resulted in the crude title compound. Preparative RP-HPLC using a gradient of 2.5 - 40% MeCN in 10 mM aq. NH₄OAc, followed by lyophilisation of the appropriate fractions yielded the title compound (1.3 mg, 1.0 μ mol, 1.5% based on initial loading of the resin). ³¹P-NMR (161 MHz, D₂O): δ - 10.46, -10.59, -10.74, -10.87. LC-MS (0 - 50 % MeCN [1% TFA] over 15 min), Rt = 4.49. HRMS [C₄₇H₇₅N₁₅O₂₄P₂ + H]⁺: 1296.4675 found, 1296.4657 calc.

References and Notes

- ¹ Van der Heden van Noort, G. J.; Van der Horst, M. G.; Overkleeft, H. S.; Van der Marel, G. A.; Filippov, D. V.; *J. Am. Chem. Soc.*, **2010**, *132*, 5236 – 5240.
- ² Hassa, P. O.; Haenni, S. S.; Elser, M.; Hottiger, M. O.; *Microbiol. Mol. Biol. Rev.*, **2006**, *70*, 789 – 829.
- ³ Corda, D.; Di Girolamo, M.; *EMBO J.*, **2003**, *22*, 1953 – 1958.
- ⁴ Okana, S.; Lan, L.; Caldecott, K. W.; Mori, T.; Yasui, A.; *Mol. Cell. Biol.*, **2003**, 3974 – 3981.
- ⁵ Kim, M. Y.; Mauro, S.; Gevry, N.; Lis, J. T.; Kraus, W. L.; *Cell*, **2004**, 803 – 814.
- ⁶ Hassa, P. O.; Buerki, C.; Lombardi, C.; Imhof, R.; Hottiger, M. O.; *J. Biol. Chem.*, **2003**, *278*, 45145 – 45153.
- ⁷ Oliver, J.; Menissier-de Murcia, J.; Nacci, C.; Decker, P.; Andreantsitohaina, R.; Muller, S.; de la Rubia, G.; Stoclet, J. C.; de Murcia, G.; *EMBO J.*, **1999**, *18*, 4446 – 4454.
- ⁸ Augustin, A.; Spenlehauer, C.; Dumond, H.; Menissier-de Murcia, J.; Piel, M.; Schmit, A.; Apiou, F.; Vonesch, J.; Kock, M.; Bornens, M.; de Murcia, G.; *J. Cell Sci.*, **2003**, *116*, 1551 – 1562.
- ⁹ Fong, P. C.; Boss, D. S.; Yap, T. A.; Tutt, A.; Wu, P. J.; Mergui-Roelvink, M.; Mortimer, P.; Swaisland, H.; Lau, A.; O'Connor, M. J.; Ashworth, A.; Carmichael, J.; Kaye, S. B.; Schellens, J. H. M.; de Bono, J. S.; *N. Engl. J. Med.* **2009**, *361*, 123 – 134.
- ¹⁰ Yates, S. P.; Jorgensen, R.; Andersen, G. R.; Merrill, A. R.; *Trends Biochem. Sci.*, **2006**, *31*, 123 – 133.
- ¹¹ Wilde, C.; Chhatwal, G. S.; Schmalzing, G.; Aktories, K.; Just, I.; *J. Biol. Chem.*, **2001**, *276*, 9537 – 9542.
- ¹² Sekine, A.; Fujiwara, M.; Narumiya, S.; *J. Biol. Chem.*, **1989**, *264*, 8602 – 8605
- ¹³ Ogata, N.; Ueda, K.; Hayaishi, O.; *J. Biol. Chem.*, **1980**, *256*, 7610 – 7615.
- ¹⁴ Rouleau, M.; Aubin, R. A.; Poirier, G. G.; *J. Cell Sci.*, **2004**, *117*, 815 – 825.
- ¹⁵ 2,3,5-tri-O-Ac- β -D-ribofuranosyl azide was prepared according to Stimac, A.; Kobe, J.; *Carbohydr. Res.* **1992**, *232*, 359 – 365 and subsequently deacetylated using Zemplén conditions.
- ¹⁶ Boschelli, D. H.; Powell, D.; Sharky, V.; Semmelhack, M. F.; *Tetrahedron Lett.*, **1989**, *30*, 1599 – 1600.
- ¹⁷ Bonache, M. A.; Nuti, F.; Le Chevalier Isaad, A.; Real-Fernandez, F.; Chelli, M.; Rovero, P.; Papini, A. M.; *Tet. Lett.*, **2009**, *50*, 4151 – 4153.
- ¹⁸ Montgomery, J. A.; *Carbohydr. Res.*, **1974**, *33*, 184 – 187.
- ¹⁹ van der Heden van Noort, G. J.; Verhagen, C. P.; van der Horst, M. G.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *Org. Lett.*, **2008**, *10*, 4461 – 4464.
- ²⁰ Chan, W. C.; White, P. D., *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*. Oxford University Press, New York: 2000.

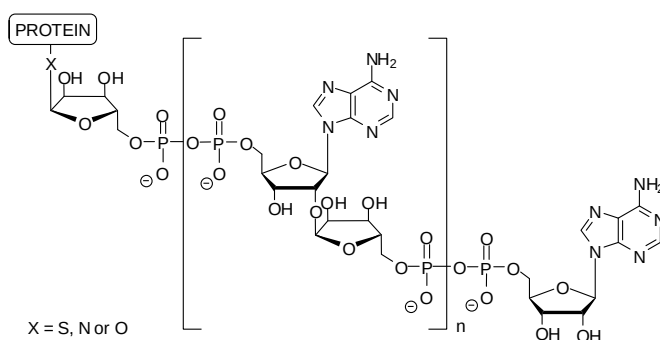
- ²¹ Conroy, T.; Jolliffe, K. A.; Payne, R. J.; *Org. Biomol. Chem.*, **2009**, 7, 2255 – 2258.
- ²² Story, S. C.; Aldrich, J. V.; *Int. J. Pept. Protein Res.*, **1992**, 39, 87 – 92. Methylester formation is known to occur when using MeOH/NH₃ to cleave HMBA.
- ²³ Zhang, B.; Bailey, V. C.; Potter, B. V. L.; *J. Org. Chem.*, **2008**, 73, 1693 – 1703.
- ²⁴ Di Girolamo, M.; Dani, N.; Stilla, A.; Corda, D.; *FEBS J.*, **2005**, 272, 4565 – 4575.
- ²⁵ Tao, Z. H.; Gao, P.; Liu, H. W.; *J. Am. Chem. Soc.*, **2009**, 131, 14258 – 14260.
- ²⁶ Hackenberger, C. P. R.; Schwarzer, D.; *Angew. Chem. Int. Ed.*, **2008**, 47, 10030 – 10074.

Ribosylation of Adenosine: An Orthogonally Protected Building Block for the Synthesis of ADP-Ribosyl Oligomers¹

Introduction

In **Chapter 3** the synthesis of *mono*-ADP-ribosylated peptides is described, in which ADP-ribose is linked to a peptide via an asparagine or glutamine residue.² In the literature some reports have appeared that describe the preparation of oxime linked analogues of *mono*-ADP-ribosylated peptides³ or the use of enzymatical techniques to access ADP-ribosyl peptides.⁴ The development of methods to prepare well defined ADP-ribosylated peptides and analogues is important to deepen the insight in the functioning of ADP-ribosylated proteins and associated enzymes. To enable the future synthesis of more advanced ADP-ribosyl derivatives containing more than one adenosine diphosphate ribose residue (see **Figure 1**, $n > 0$), the availability of a suitably protected ribosylated adenosine building block is crucial. Synthetic challenges in the construction of such a building block are the formation of the α -glycosidic bond between the ribofuranose and the 2'-hydroxyl of adenosine and the development of a suitable protective group strategy. *O*-Glycosylated nucleosides in general are difficult to synthesize^{5,6}, as in glycosylation reactions the nucleobase is often more nucleophilic than the target hydroxyl. As a result side reactions at the nucleobase may occur that can lead to depurination of the nucleoside in the case of adenosine and guanosine. Consequently, yields are often low for such glycosylation reactions.⁵ Shimofuridin,⁷⁻⁹ bearing a 2'-*O*- β -fucopyranosyl moiety, and adenophostin,¹⁰⁻¹⁶ bearing a 3'-*O*- α -glucopyranosyl moiety, are examples of glycosylated nucleosides that have been synthesized. Although several nucleosides decorated with *O*- β -linked furanosyl groups have been synthesized,^{6,17-19} only one synthesis of 2'-*O*- α -ribosyl adenosine is reported.²⁰

Figure 1



Poly-ADP-ribose polymer attached to target protein.

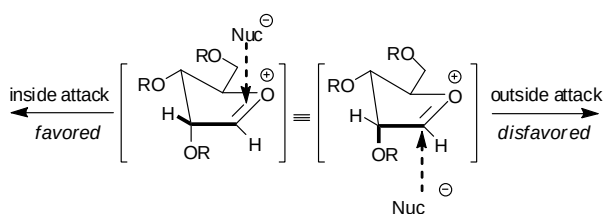
This procedure involves a 1,2-*trans* selective coupling of an 2'-*O*-acyl protected arabinofuranose building block to an adenosine derivative, followed by protective group manipulation and inversion of the stereochemistry at the 2'-position of the arabinofuranose moiety via an oxidation/reduction protocol. Although 2'-*O*- α -ribsyl adenosine was obtained for the first time, no strategy for the orthogonal protection of both primary hydroxyl groups was proposed, limiting the usefulness of this procedure in the synthesis of *poly*-ADP-ribosylated peptides.

Results and discussion

In order to synthesize 2'-*O*- α -ribsyl adenosine bearing orthogonal protective groups on the primary hydroxyl functions, an approach involving α -glycosylation of an adenosine derivative with a suitable ribofuranosyl donor is considered to be the most convenient. The use of 2',3'-isopropylidene²⁵ or 2',3'-benzylidene²⁶ protection groups on the ribofuranosyl donor, which are convenient protective groups due to the ease of introduction and cleavage, leads to anomeric mixtures. Application of acyl based protective groups would result in exclusive formation of the unwanted β -riboside due to neighboring group participation. Therefore, 2',3',5'-*tri*-*O*-benzyl-D-ribofuranose was chosen as glycosyl donor, since it is known that non-participating benzyl protection on the ribofuranosyl donor in glycosylation reactions are a prerequisite for the formation of mainly or exclusively the α -riboside.²¹⁻²⁴ Several explanations for the predisposition of ribofuranosyl donors, equipped with nonparticipating protective groups, to couple with high α -selectivity have been reported.^{22,23,24,27} The group of Woerpel studied the factors controlling the α -selective C-glycosylation of benzylated furanose donors and presented a model for preferential nucleophilic attack on furanosyl oxocarbenium ions, that occur as two inter-converting envelop conformers.^{23,24} Oxocarbenium ions are widely accepted as possible intermediates in Lewis acid promoted nucleophilic substitution on furanosyl donors.^{27,28} The stereo-chemical outcome of the glycosylation reaction is dictated by at least two factors; the face of attack of the incoming nucleophile and the lowest energy conformer of the two possible envelop forms of the oxocarbenium ion intermediate.

An incoming nucleophile attacks preferentially from inside the oxocarbenium ion envelope due to favorable staggered interactions between the substituents at C2' and the incoming nucleophile at C1', in contrast to the disfavored eclipsed interactions that accompany an attack from the outside-face of the envelope (see **Figure 2**).²⁴ In addition rehybridization of C1' from trigonal planar in the oxocarbenium ion to tetrahedral in the product proceeds via a envelope intermediate in the case of inside attack versus a less energetically favorable twist intermediate for outside attack. Taking both factors into account, outside attack is less favored than inside attack.

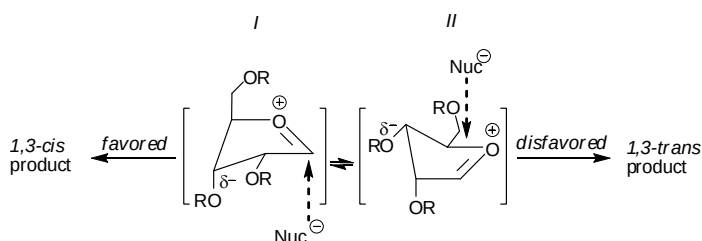
Figure 2



Nucleophilic attack on furanosyl oxocarbenium ions

Which of two accessible oxocarbenium ions is the lowest energy conformer is largely governed by the substitution pattern on the furanose scaffold. Alkoxy substituents on the C3'-position, for instance, prefer a pseudo-axial orientation thereby placing the oxygen of the substituent in close proximity to the cationic carbon of the oxocarbenium ion intermediate, stabilizing the ion (see **Figure 3**).²³ Conformer *I* will thus be lower in energy than conformer *II* and therefore be subjected to nucleophilic attack giving rise to mostly the 1,3-cis product. In the case of ribofuranose oxocarbenium ions the high 1,3-cis selectivity is mainly attributed to the pseudo-axial placement of 3'-alkoxy substituent and C2'-alkoxy substituents are believed to play a minor role.

Figure 3

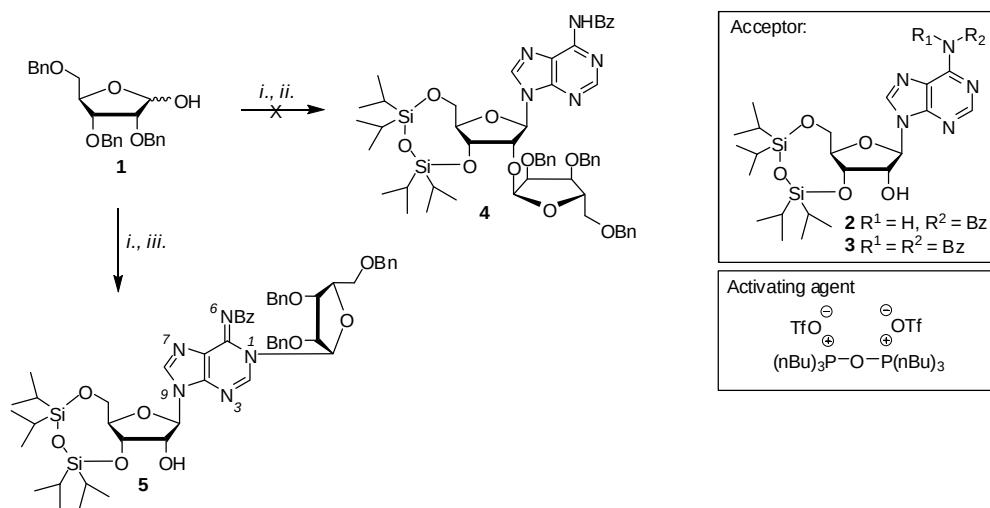


Ribofuranosyl oxocarbenium ion conformers

Having selected the donor, 3',5-O-TIPDS protected *N*⁶-benzoyl adenosine **2** seemed a suitable acceptor in the glycosylation reaction (see **Scheme 1**).³⁰ Since adenosine may depurinate under strongly acidic conditions³¹, a mild glycosylation protocol was sought. On this

basis attention was directed on the procedure of Mukaiyama²², who revealed that activation of **1** (see **Scheme 1**) with diphosphonium salts and ensuing base-assisted glycosylation resulted in a high yielding and α -selective ribofuranosylation of unhindered alcohols.³² However, the use of sterically demanding compound **2** as the acceptor under the reported conditions was unproductive.³³ When performing the glycosylation with a larger excess of base a product different from the expected compound **4** was formed. It appeared that *N*-glycosylation of the nucleobase had occurred since the ribose C1' was found to resonate at 86 ppm in NMR-analysis, which is typical for hemi-aminal ethers rather than around 100 ppm which is characteristic for a hemi-acetal such as in **4**. Vorbrüggen has postulated the involvement of N^1 , N^3 and N^7 glycosylated adenosine intermediates en route to the thermodynamic N^9 product in acid catalyzed adenosine synthesis.³⁴ It is reported that in ribosylation of the nucleobase adenine, N^1 glycosylation precedes isomerisation towards the N^9 -linked adenosine.³⁵ An intermediate in this process is the N^1, N^9 -bis-ribosyladenosine, which could be isolated.³⁶ Therefore the product isolated after coupling **1** with **2** under Mukaiyama conditions, is postulated to be **5**. Formation of this product can be explained by the increase in nucleophilicity of the N^1 position of the adenine moiety caused by deprotonation of the exocyclic amide by the excess of base, followed by nucleophilic attack on the oxocarbenium ion of 2',3',5'-tri-O-Bn-ribofuranose. Consistent with the structure of **5** are the NOESY NMR data supporting the presence of an α -glycosidic linkage and HMBC NMR data indicating the presence of the ribosyl moiety on N^1 of adenosine.³⁷ This interesting side reaction resulting in compound **5** may be valuable for the construction of cADPR analogues.³⁸ In an attempt to avoid *N*-glycosylation N^6, N^6 -dibenzoylated adenosine (**3**)³⁹ was used as acceptor (see **Scheme 1**). Acceptor **3** appeared to be insufficiently reactive under these conditions and the glycosylation proved to be unproductive.

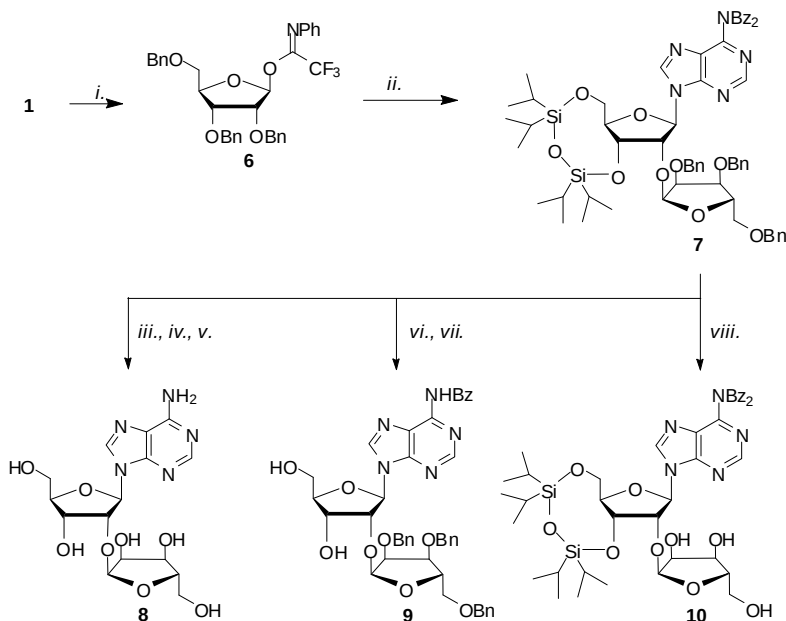
Scheme 1. Phosphonium salt mediated glycosylations



Dehydrative coupling conditions employing activation of the donor with Ph_2SO and Tf_2O gave intramolecular Friedel-Crafts alkylation with the 2'-OBn of the donor.⁴⁰ When using a carbonylimidazole donor⁴¹ or a di-butylphosphate donor⁴² only trace amounts of the product were obtained. After screening these methods, imidate donor **6** in combination with TMSOTf as promoter was found to be the most efficient glycosylating reagent (see **Scheme 2**).⁴³ Imidate **6** was prepared in 92% yield by reacting 2',3',5'-tri-*O*-Bn-ribofuranose with Cs_2CO_3 and $\text{Cl}(\text{C}=\text{NPh})\text{CF}_3$. Activation of imidate **6** with TMSOTf in the presence of acceptor **3** resulted in formation of **7**, containing the desired α -ribofuranosyl linkage in good yield.⁴⁴ Complete deprotection of **7** using TBAF and NH_3/MeOH and finally palladium catalyzed hydrogenation gave access to **8**, the spectroscopic and physical properties of which were in complete accordance with the published data.²⁰

The inherent orthogonality between silyl and benzyl ethers allows selective manipulation of the primary OH-groups in **7**. To exemplify this orthogonality the adenosine hydroxyls were deprotected by cleaving the silyl protective group using TBAF. Due to the alkaline nature of TBAF one of the two benzoyl groups on the exocyclic amine of adenosine was partially cleaved to give a monobenzoylated side product. In order to avoid this partial deprotection and the need for an extra purification step, compound **7** was first treated with MeNH_2 to completely remove one benzoyl while leaving the other benzoyl in place and then perform the TBAF treatment to give **9** (**Scheme 2**).⁴⁵

Scheme 2. Key glycosylation using imidate **6**

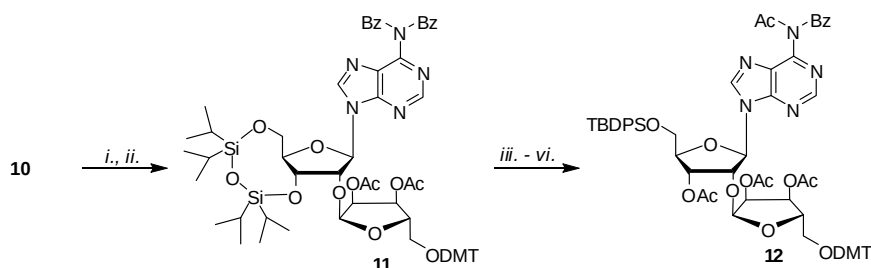


Reagents and conditions; *i*. Cs_2CO_3 , $\text{Cl}(\text{C}=\text{NPh})\text{CF}_3$, acetone/water, 3 h, r.t., (92%), *ii*. TMSOTf, acceptor **3** (**Scheme 1**), 0 °C, 1.5 h., (60%), *iii*. TBAF, *iv*. NH_3 , MeOH, *v*. Pd/C, H_2 , MeOH, 48 h, r.t., (16% over three steps), *vi*. MeNH_2 , 0 °C, 2 h, *vii*. TBAF, 1.5 h, r.t., (69% over two steps), *viii*. BCl_3 , -78 °C, 4.5 h, (45 %).

Similarly the hydroxyl groups on the ribose moiety were deprotected while leaving all the protective groups on adenosine in place. The obvious choice to remove benzyl ethers involves palladium catalyzed hydrogenation with Pd/C in an alcoholic solvent, which however, proceeded sluggishly. Therefore compound **7** was treated with BCl_3 at $-78\text{ }^\circ\text{C}$ to rapidly afford compound **10**.

To generate building block **12** equipped with orthogonal protective groups on the primary hydroxyls, compound **10** was chosen as starting material (see **Scheme 3**). Acyl protective groups that are orthogonal to the two primary hydroxyl protective groups were envisioned as protective groups for all other functionalities. Straightforward conversion of **10** by protecting the primary hydroxyl of ribose with DMT and subsequent acetylation of the remaining secondary hydroxyls (see **Scheme 3**) gave **11**. Cleavage of one benzoyl group in **11**, removal of the TIPDS protection, introduction of a TBDPS on the primary hydroxyl of adenosine, and subsequent acetylation of the secondary hydroxyls and the exocyclic nitrogen of adenosine gave fully protected compound **12**. The thus formed compound **12** bears orthogonal protective groups on both primary hydroxyl functions, which allows the future synthesis of poly-ADP-ribosylated peptides using the methodology as reported in **Chapter 3**.^{2,46}

Scheme 3. Introduction of orthogonal protective groups



Reagents and conditions; *i.* DMTCl, DMAP, pyr., 4 h, $40\text{ }^\circ\text{C}$, *ii.* Ac_2O , pyr., 16 h, $40\text{ }^\circ\text{C}$ (55% over two steps), *iii.* MeNH_2 , DCM/MeOH, 30 min, $0\text{ }^\circ\text{C}$, *iv.* TBAF, THF, 2 h, r.t., *v.* TBDPSCl, DMAP, pyr. 48 h, $50\text{ }^\circ\text{C}$, *vi.* Ac_2O , pyr. 16 h, r.t. (49% over four steps)

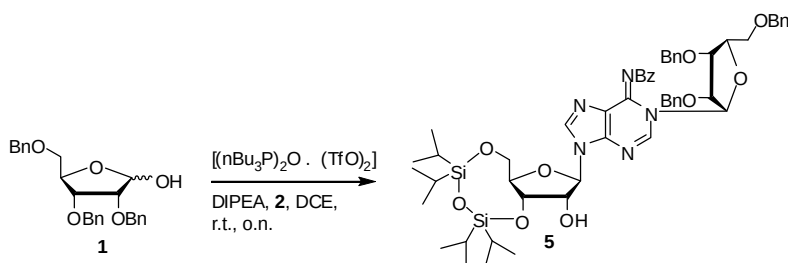
Conclusion

A method for α -selective ribosylation of adenosine at the 2'-OH function is presented using the anomeric trifluoroimidate of 2',3',5'-tri-*O*-benzyl ribose and trimethylsilyl triflate as activating agent. Protective groups on the ribose donor and adenosine acceptor not only dictated α -selectivity but also allowed selective deprotection of either the adenosine hydroxyls or the ribose hydroxyls. These properties were further exploited in the synthesis of compound **12**, which is equipped with a TBDPS and DMT on the two primary hydroxyls, respectively. This protective group pattern may allow a stepwise synthesis of poly-ADP-ribosyl modified peptides using the methodology described in **Chapter 3**.

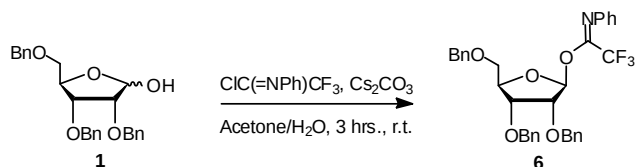
Experimental

General: All chemicals were used as received. EtOAc, pyridine, *n*-PrOH and MeCN were stored over activated 4 Å molecular sieves and MeOH was stored over activated 3 Å molecular sieves for at least 24 hours prior to use. DCM was distilled over CaH₂ and dioxane over LiAlH₄ prior to use. Solvents were removed by rotary evaporation under reduced pressure at a temperature not exceeding 40°C. Optical rotations are measured in CHCl₃ or MeOH at a concentration of 10 mg/mL at 25 °C. TLC analysis, silica gel column chromatography, NMR, LCMS, HPLC and HRMS techniques were used as described in **Chapter 2**. In the notation of NMR spectra ribose-proton and carbon atoms are reported as H' and C', whereas protons and carbons belonging to the adenosine sugar are reported as H'' and C'' and atoms belonging to the nucleobase as H and C.

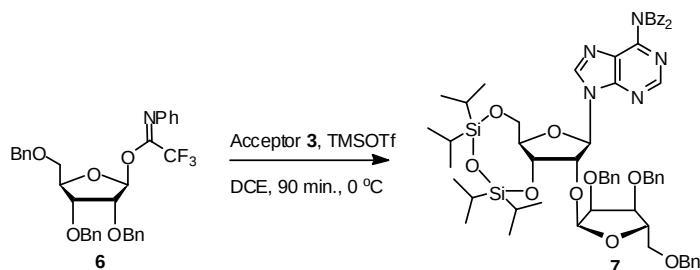
*N*⁶-Benzoyl-9-[3, 5-(1, 1, 3, 3-tetra-isopropylidisiloxane-1, 3-diyl)-*N*¹-(2, 3, 5-tri-*O*-benzyl- α -D-ribofuranosyl)- β -D-ribofuranosyl]-adenine (5)



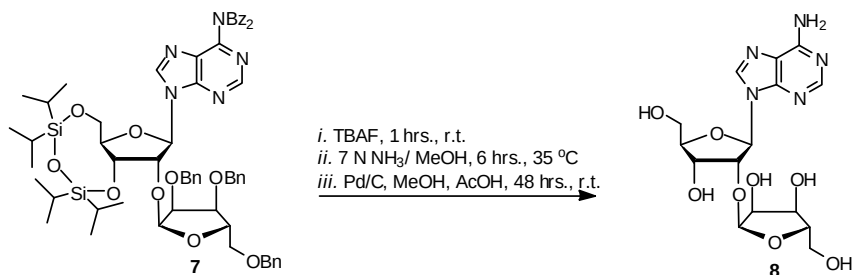
n-Bu₃PO (1.96 g, 9.0 mmol) was coevaporated with DCE three times prior to dissolving in freshly distilled DCE (40 mL) at 0 °C. Triflic anhydride (700 μ L, 4.2 mmol) was slowly added to the reaction mixture over a period of 10 min. After 30 min the reaction mixture was checked with ³¹P-NMR (161 MHz, 119.5 ppm, acetone-*d*₆ capillary was used as the external lock). Activated 4 Å molecular sieves were added prior to addition of 2,3,5-tri-*O*-Bn-ribose (0.77g, 1.83 mmol) and DIPEA (0.8 mL, 4.9 mmol) in DCE (40 mL), using a canule. After 2 hours ³¹P-NMR shows formation of the putative activated species (161 MHz, 98 ppm and 94 ppm, acetone-*d*₆ capillary was used as the external lock). Acceptor *N*⁶-Benzoyl-9-[3, 5-(1, 1, 3, 3-tetra-isopropylidisiloxane-1, 3-diyl)- β -D-ribofuranosyl]-adenine **2** (0.71 g, 1.15 mmol) and DIPEA (0.8 mL, 4.9 mmol) were dissolved in 40 mL DCE and added to the reaction mixture. The reaction mixture was stirred at room temperature over night and subsequently washed with 5% citric acid, dried with MgSO₄ and concentrated. The crude material was purified with silica gel column chromatography using a gradient of PE in EtOAc (100 : 0– 50/50, v/v) to afford the title compound as a white foam (865 mg, 0.85 mmol, 74% based on the donor). ¹H-NMR (400 MHz, CDCl₃) δ : 8.22 (s, 1H, H2), 8.16 – 8.03 (m, 2H, Arom. Bz), 7.75 (s, 1H, H8), 7.48 (t, *J* = 7.4 Hz, 1H, Arom. Bz), 7.39 (t, *J* = 7.5 Hz, 3H, Arom. Bz), 7.35 - 7.25 (m, 10H, Arom. Bn), 7.18 - 7.01 (m, 5H, Arom. Bn), 6.75 (d, *J* = 3.7 Hz, 1H, H1'), 5.84 (d, *J* = 1.2 Hz, 1H, H1''), 4.90 (t, *J* = 6.4 Hz, 1H, H3''), 4.68 - 4.33 (m, 9H, CH₂ Bn's, H2', H4', H2''), 4.21 (dd, *J* = 7.9, 4.3 Hz, 1H, H3'), 4.07 - 3.95 (m, 3H, H4'', H5''), 3.76 (dd, *J* = 12.0, 2.0 Hz, 1H, H5_A'), 3.59 (dd, *J* = 10.8, 3.6 Hz, 1H, H5_B'), 3.17 (s, 1H, OH), 1.13 - 0.95 (m, 28H, *i*Pr). ¹³C-NMR (101 MHz, CDCl₃) δ : 176.5 (CO Bz), 147.0 (C6), 145.4 (C2), 144.7 (C4), 138.6 (C8), 137.8, 137.2, 137.1 (Cq Bn), 135.8 (Cq Bz), 131.9 (Arom. Bz), 129.8, 128.6, 128.5, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7 (Arom. Bz/Bn), 122.0 (C5), 89.6 (C1''), 87.1 (C1'), 82.3 (C4''), 81.1 (C4'), 78.0 (C3'), 75.6 (C2'), 75.3 (C2''), 74.4, 73.6, 73.0 (CH₂ Bn), 71.2 (C3''), 69.0 (C5'), 62.1 (C5''), 17.5, 17.40, 17.35, 17.2, 17.1, 17.0 (CH₃ *i*Pr), 13.3, 13.1, 12.9, 12.7 (CH *i*Pr). α_D (CDCl₃) = -59 °. IR: 2866, 2924, 1636, 1454, 1246, 1124, 1035, 885, 695. HRMS [C₅₅H₆₉N₅O₁₀Si₂ + H]⁺: found 1016.4642, calc. 1016.4656.

(*N*-Phenyl)-2,2,2-trifluoroacetimido-2, 3, 5-tri-*O*-benzyl- β -D-ribofuranose (6)

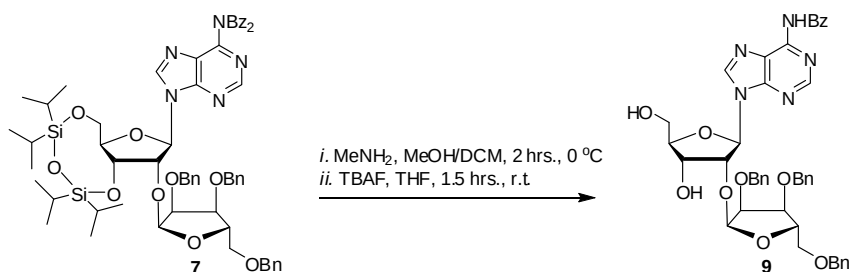
2,3,5-*O*-Benzyl-D-ribose (0.84 g, 2.0 mmol) was dissolved in acetone (8 mL) and water (100 μL). Cs_2CO_3 (0.72 g, 2.2 mmol) and ClC(=NPh)CF_3 (0.6 mL, 1.2 mmol) were added and the reaction mixture was stirred at room temperature for 3 hours. After filtration over celite, the solvents were removed and the residue was purified with silica gel column chromatography using a gradient of PE in EtOAc (100 : 0 – 90/10, v/v) to afford the title compound as a colourless oil (1.06 g, 1.80 mmol, 90%). $^1\text{H-NMR}$ (400 MHz, CDCl_3 , $T=303\text{K}$) δ : 7.60 - 7.16 (m, 17H, Arom. Bn/Ph), 7.10 (t, $J = 7.4$ Hz, 1H, Arom. Bn/Ph), 6.80 (d, $J = 7.7$ Hz, 2H, Arom. Bn/Ph), 6.29 (s, 1H, $\text{H1}'$), 4.79 - 4.37 (m, 7H, CH_2 Bn's, $\text{H4}'$), 4.15 (dd, $J = 7.3$, 4.7 Hz, 1H, $\text{H3}'$), 4.07 (d, $J = 4.2$ Hz, 1H, $\text{H2}'$), 3.70 (dd, $J = 11.2$, 3.2 Hz, 1H, $\text{H5}_\text{A}'$), 3.63 (dd, $J = 15.6$, 4.8 Hz, 1H, $\text{H5}_\text{B}'$). $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ : 143.8 (C=N), 138.1, 137.5, 137.3 (Cq Arom.), 128.7, 128.4, 128.4, 128.3, 128.1, 128.0, 127.9, 127.8, 127.5, 124.2, 119.6 (Arom.), 102.4 ($\text{C1}'$), 82.2 ($\text{C4}'$), 78.6 ($\text{C2}'$), 77.3 ($\text{C3}'$), 73.3, 72.7, 72.3 9 (CH_2 Bn), 70.0 ($\text{C5}'$). α_D (CDCl_3) = + 47.2°. IR: 3033, 2866, 1710, 1599, 1454, 1205, 1149, 1088, 906, 751, 694. HRMS [$\text{C}_{34}\text{H}_{32}\text{NO}_5 + \text{H}$] $^+$: found 592.2298, calc. 592.2305

***N*⁶,*N*⁶-di-Benzoyl-9-[3, 5-(1, 1, 3, 3-tetra-isopropylidisiloxane-1, 3-diyl)-2-*O*-(2, 3, 5-tri-*O*-benzyl- α -D-ribofuranosyl)]- β -D-ribofuranosyl]-adenine (7)**

Acceptor **3** (250 mg, 0.35 mmol) and donor **6** (300 mg, 0.5 mmol) were coevaporated with toluene three times prior to dissolving in DCE (10 mL). Freshly activated 4 Å molecular sieves were added and the solution was stirred under an argon atmosphere for 20 min. The reaction mixture was cooled to 0 °C and TMSOTf (65 μL , 0.35 mmol) was slowly added over a period of 15 min. After 90 min. the reaction mixture was allowed to warm to room temperature and quenched using 100 μL TEA. After evaporation of the solvent the crude material was purified with silica gel column chromatography using a gradient of DCM/PE/Acetone (80 : 20: 0 – 78/20/2, v/v/v) to afford the title compound as a white foam (230 mg, 0.21 mmol, 60% based on acceptor). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ : 8.61 (s, 1H, H2), 8.25 (s, 1H, H8), 7.85 (d, 4H, Arom. Bz), 7.45 (t, $J = 7.5$ Hz, 2H, Arom. Bz), 7.37 - 7.15 (m, 17H, Arom. Bn), 6.10 (s, 1H, $\text{H1}''$), 5.41 (d, $J = 3.9$ Hz, 1H, $\text{H1}'$), 4.71 - 4.54 (m, 7H, CH_2 Bn's, $\text{H3}''$, $\text{H2}''$), 4.51 - 4.36 (m, 4H, CH_2 Bn's, $\text{H4}''$), 4.28 (d, $J = 9.2$ Hz, 1H, $\text{H4}'$), 4.20 (d, $J = 14.4$ Hz, 1H, $\text{H5}_\text{A}''$), 3.99 (dd, $J = 13.2$, 2.0 Hz, 1H, $\text{H5}_\text{B}''$), 3.93 - 3.82 (m, 2H, $\text{H2}'$, $\text{H3}'$), 3.42 - 3.37 (m, 2H, $\text{H5}'$), 1.41 - 0.66 (m, 28H, *i*Pr). $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ : 172.3 (CO Bz), 152.17 (C2), 152.16 (C6), 151.7 (C4), 143.4 (C8), 138.5, 138.0, 137.8 (Cq Bn), 134.2 (Cq Bz), 133.0, 129.5, 128.8, 128.6, 128.4, 128.3, 128.06, 128.04, 128.01, 127.7, 127.65, 127.61 (CH Arom.), 101.2 ($\text{C1}'$), 89.7 ($\text{C1}''$), 82.2 ($\text{C4}'$), 81.6 ($\text{C4}''$), 79.8 ($\text{C2}''$), 78.1 ($\text{C2}'$), 74.8 ($\text{C3}'$), 73.5, 72.7, 72.1 (CH_2 Bn's), 69.8 ($\text{C5}'$), 68.0 ($\text{C3}''$), 59.6 ($\text{C5}''$), 17.6, 17.5, 17.4, 17.24, 17.18, 17.15, 17.07 (CH_3 *i*Pr), 13.5, 13.0, 12.9, 12.8 (CH *i*Pr). α_D (CDCl_3) = + 9.2°. IR: 2944, 2867, 1702, 1590, 1576, 1450, 1237, 1122, 1034. HRMS [$\text{C}_{62}\text{H}_{73}\text{N}_5\text{O}_{11}\text{Si}_2 + \text{H}$] $^+$: found 1120.4926, calc. 1120.4918

[2-O-(α -D-Ribofuranosyl)- β -D-ribofuranosyl]-adenine (8)

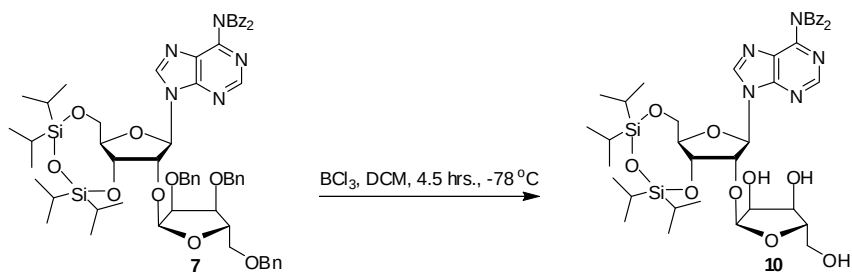
Compound **7** (0.23 g, 0.2 mmol) was dissolved in THF (10 mL) and treated with TBAF (1 M in THF, 0.5 mmol) at room temperature for 1 hour. The reaction mixture was concentrated, redissolved in DCM washed with brine, 1N HCl, dried with MgSO₄ and concentrated. The crude de-silylated intermediate was dissolved in 7N MeOH/NH₃ (20 mL) and heated for 6 hours at 35 °C. Concentration and coevaporation with toluene afforded the free amine. After dissolving in 15 mL MeOH and 0.5 mL AcOH, the reaction mixture was purged with argon, Pd/C (10% on activated charcoal) was added and the reaction mixture was brought under a H₂ atmosphere, where it was left at room temperature for 48 hours. Filtration over celite and RP-HPLC using a gradient of 0 – 40% MeCN in 0.1% aq. TFA, followed by lyophilisation of the appropriate fractions yielded the title compound in an overall yield of 16% (12.7 mg, 31.8 μ mol). ¹H-NMR (400 MHz, D₂O) δ : 8.25 (s, 1H, H8), 8.05 (s, 1H, H2), 6.11 (d, J = 6.2 Hz, 1H, H1''), 5.03 (d, J = 4.0 Hz, 1H, H1'), 4.81 (1H, H2''), 4.48 (dd, J = 5.0, 3.1 Hz, 1H, H3''), 4.26 (q, J = 2.9 Hz, 1H, H4''), 4.20 - 4.13 (m, 1H, H4'), 4.06 - 3.95 (m, 2H, H2', H3'), 3.87 (dd, J = 12.9, 2.6 Hz, 1H, H5_A''), 3.79 (dd, J = 12.9, 3.4 Hz, 1H, H5_B''), 3.63 (dd, J = 12.4, 3.5 Hz, 1H, H5_A'), 3.56 (dd, J = 12.4, 4.7 Hz, 1H, H5_B'). ¹³C-NMR (101 MHz, D₂O) δ 158.0 (C6), 155.0 (C2), 150.8 (C4), 143.1 (C8), 121.5 (C5), 104.4 (C1'), 89.4 (C1''), 88.7 (C4''), 87.9 (C4'), 81.6 (C2''), 74.0 (C2'), 73.4 (C3''), 72.2 (C3'), 64.04, 64.01 (C5', C5''). IR: 3318, 2932, 1652, 1480, 1418, 1334, 1090, 1032. HRMS [C₁₅H₂₁N₅O₈ + H]⁺: found 400.1462, calc. 400.1463.

N⁶,N⁶-di-Benzoyl-9-[2-O-(2, 3, 5-tri-O-benzyl- α -D-ribofuranosyl)- β -D-ribofuranosyl]-adenine (9)

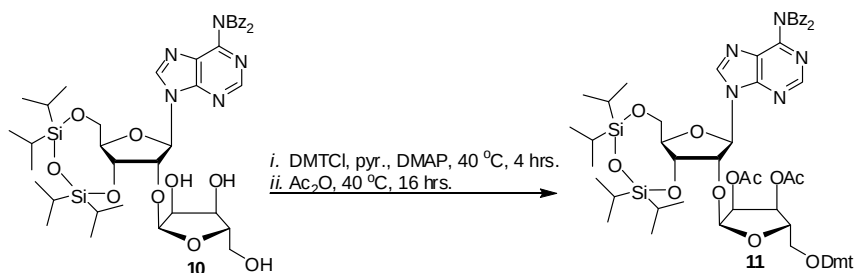
Compound **7** (0.13 g, 112 μ mol) was dissolved in DCM/MeOH (10 mL, 4:1 v/v) and treated with methylamine (8M solution in EtOH, 40 μ L) at 0 °C. After 2 hours the reaction mixture was concentrated and redissolved in THF (10 mL). Treatment with TBAF (1M sol. in THF, 300 μ L) at room temperature for 1.5 hours, concentration and purification using silica gel column chromatography with a gradient of PE/EtOAc/MeOH (100/ 0/ 0 – 0/95/5, v/v/v) afforded the title compound as a colourless syrup (60 mg, 78 μ mol, 69%). ¹H-NMR (400 MHz, CDCl₃) δ 9.16 (s, 1H, NH), 8.77 (s, 1H, H2), 8.03 (m, 2H, H8, Arom. Bz), 7.61 (t, J = 7.4 Hz, 1H, Arom. Bz), 7.52 (t, J = 7.6 Hz, 2H, Arom. Bz), 7.36 - 7.21 (m, 14H, Arom. Bn), 7.17 - 7.10 (m, 2H, Arom. Bn), 6.06 (d, J = 7.5 Hz, 1H, H1''), 6.01 (d, J = 10.5 Hz, 1H, 5'' OH), 4.98 (dd, J = 7.4, 4.7 Hz, 1H, H2''), 4.92 (d, J = 3.1 Hz, 1H, H1'), 4.75 (d, J = 12.0 Hz, 1H, CH₂ Bn), 4.61 (d, J = 12.0 Hz, 1H, CH₂ Bn), 4.57 - 4.54 (m, 1H, H3''), 4.52 (d, J = 12.4 Hz, 1H, CH₂ Bn), 4.49 (d, J = 12.0 Hz, 1H, CH₂ Bn), 4.41 - 4.33 (m, 2H, CH₂ Bn, H4''), 4.25 (t, J = 3.6 Hz, 1H, H4'), 4.08 (s, 1H, 3'' OH),

3.98 (d, $J = 12.9$ Hz, 1H, H5_B''), 3.93 - 3.84 (m, 2H, H2', H3'), 3.77 (t, $J = 11.8$ Hz, 1H, H5_A''), 3.37 - 3.20 (m, 2H, H5'). ¹³C-NMR (101 MHz, CDCl₃) δ : 164.6 (CO Bz), 152.3 (C2), 150.8 (C6), 150.4 (C4), 143.6 (C8), 137.7, 137.5, 137.2 (Cq Bn), 133.5 (Cq Bz), 133.1, 129.1, 128.6, 128.6, 128.58, 128.55, 128.24, 128.21, 128.0, 127.9, 127.6, 124.8 (CH Arom.), 100.7 (C1'), 89.7 (C1''), 88.3 (C4''), 83.5 (C4'), 79.1, 78.7, 76.8 (C2', C3', C2''), 73.6, 73.3, 73.0 (CH₂ Bn), 72.9 (C3''), 70.0 (C5'), 63.5 (C5''). IR: 3304, 2930, 1700, 1609, 1582, 1455, 1247, 1218, 1072, 1026, 744, 696. α_D (CDCl₃) = + 9.6°. HRMS [C₄₃H₄₃N₅O₉ + H]⁺: found 774.3140, calc. 774.3134.

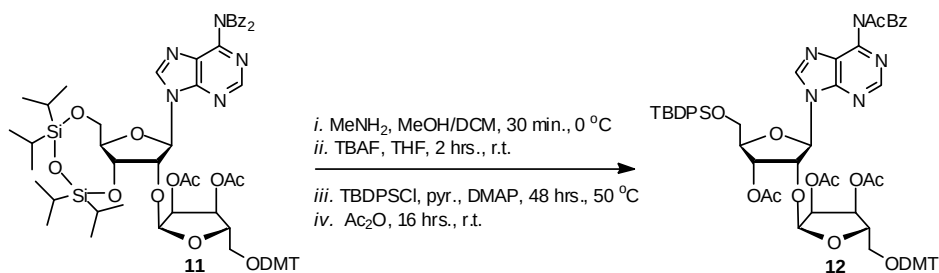
N⁶,N⁶-di-Benzoyl-9-[3, 5-(1, 1, 3, 3-tetra-isopropylidisiloxane-1, 3-diyl)-2-O-(α -D-ribofuranosyl)- β -D-ribofuranosyl]-adenine (10)



Ribosyl adenosine **7** (240 mg, 215 μ mol) was dissolved in dry DCM (15 mL) and cooled to -78 °C. Borontrichloride (1M solution in DCM, 1 mL) was slowly added and the reaction mixture was left to stir for 4.5 hours. The reaction was quenched with MeOH/TEA (400 μ L, 1:1 v/v) at -78 °C. Concentration of the crude product followed by coevaporation with toluene (3x) and purification using silica gel column chromatography with a gradient of PE/DCM/MeOH (20 : 80 : 0 – 20 : 79 : 1, v/v/v) afforded the title compound as a colourless syrup (82 mg, 96 μ mol, 45%). ¹H-NMR (400 MHz, CDCl₃) δ : 8.65 (s, 1H, H2), 8.31 (s, 1H, H8), 7.94 - 7.81 (m, 4H, Arom. Bz), 7.48 (t, $J = 7.5$ Hz, 2H, Arom. Bz), 7.36 (t, $J = 7.7$ Hz, 4H, Arom. Bz), 6.14 (s, 1H, H1''), 5.51 (d, $J = 3.6$ Hz, 1H, H1'), 4.79 (t, $J = 5.2$ Hz, 1H, H2''), 4.75 (t, $J = 5.2$ Hz, 1H, H3''), 4.35 - 4.08 (m, 4H, H5''^A, H4', H4'', H2'), 4.05 - 3.94 (m, 2H, H5''^B, H3'), 3.71 (dd, $J = 12.0, 3.2$ Hz, 1H, H5_A'), 3.66 (dd, $J = 12.0, 4.0$ Hz, 1H, H5_B'), 3.36 (d, $J = 10.5$ Hz, 1H, H5'-OH), 3.08 (s, 1H, OH), 2.02 (s, 1H, OH), 1.10 - 1.00 (m, 28H, *i*Pr). ¹³C-NMR (101 MHz, CDCl₃) δ : 172.5 (CO Bz), 152.5 (C2), 152.1, 152.0 (C6, C4), 142.9 (C8), 134.2 (Cq Bz), 133.2, 129.6, 128.9 (Arom. Bz), 128.3 (C5), 100.1 (C1'), 88.2 (C1''), 86.9 (C4'), 82.1 (C4''), 76.5 (C2''), 72.9 (C2'), 72.0 (C3'), 67.9 (C3''), 62.9 (C5'), 59.4 (C5''), 17.6, 17.5, 17.4, 17.2, 17.1, 16.9 (CH₃ *i*Pr), 13.7, 13.1, 13.0, 12.8 (CH *i*Pr). α_D (CDCl₃) = + 2.2°. IR: 3466, 2923, 2867, 1700, 1600, 1576, 1450, 1237, 1119, 1036, 885, 693. HRMS [C₄₁H₅₅N₅O₁₁Si₂ + H]⁺: found 850.3520, calc. 850.3509

***N*⁶,*N*⁶-di-Benzoyl-9-[3, 5-(1, 1, 3, 3-tetra-isopropylidisiloxane-1, 3-diyl)-2-*O*-(5-*O*-dimethoxytrityl)- 2, 3-di-*O*-acetyl- α -D-ribofuranosyl]- β -D-ribofuranosyl]-adenine (11)**

Nucleoside **10** (96 μ mol) was dissolved in pyridine (5 mL) and reacted with DMTCl (5 eq., 150 mg) and a catalytic amount of DMAP at 40 °C for 4 hours. Subsequently Ac₂O (500 μ L) was added and the reaction mixture was left at 40 °C for 16 hours. Concentration of the crude followed by purification using a silica gel column pre-neutralized using TEA and a gradient of Hexane/ EtOAc (100 : 0 – 60 : 40, v/v) afforded the title compound as a colourless syrup (64 mg, 52 μ mol, 55%). ¹H-NMR (400 MHz, CDCl₃) δ : 8.67 (s, 1H, H2), 8.33 (s, 1H, H8), 7.85 (d, *J* = 7.7 Hz, 4H, Arom.), 7.59 – 7.01 (m, 15H, Arom.), 6.81 (d, *J* = 8.8 Hz, 4H, Arom.), 6.01 (s, 1H, H1''), 5.78 (d, *J* = 4.3 Hz, 1H, H1'), 5.36 (dd, *J* = 7.2, 3.4 Hz, 1H, H3'), 5.13 (dd, *J* = 7.1, 4.3 Hz, 1H, H2'), 4.65 (dd, *J* = 9.2, 4.9 Hz, 1H, H3''), 4.60 (d, *J* = 4.9 Hz, 1H, H2''), 4.39 (dd, *J* = 6.9, 3.4 Hz, 1H, H4'), 4.23 (d, *J* = 13.3 Hz, 1H, H5_A''), 4.14 (d, *J* = 9.2 Hz, 1H, H4''), 4.02 (dd, *J* = 13.3, 2.2 Hz, 1H, H5_B''), 3.78 (s, 6H, OMe's DMT), 3.30 (dd, *J* = 10.0, 2.0 Hz, 1H, H5_A'), 3.17 (dd, *J* = 10.0, 3.6 Hz, 1H, H5_B'), 2.10, 2.09 (2x s, 6H, CH₃ OAc's), 1.28 – 0.84 (m, 28H, iPr). ¹³C-NMR (101 MHz, CDCl₃) δ : 172.4 (CO Bz), 170.4, 170.1 (CO Ac), 158.6 (Cq DMT), 152.4 (C4), 152.2 (C2), 151.9 (C6), 144.7 (Cq DMT), 143.1 (C8), 135.9, 135.8, 134.2, 133.0, 130.2, 129.6, 128.8, 128.4, 128.3, 128.0, 126.9 (Arom. Bz, DMT), 113.3 (Arom. DMT), 100.0 (C1'), 89.1 (C1''), 86.3 (Cq DMT), 81.9 (C4''), 81.0 (C4'), 79.4 (C2''), 71.7 (C2'), 70.5 (C3'), 68.1 (C3''), 63.0 (C5'), 59.8 (C5''), 55.4 (OMe DMT), 20.8, 20.7 (CH₃ OAc), 17.6, 17.4, 17.2, 17.1 (CH₃ iPr.), 13.6, 13.1, 12.9, 12.7 (CH iPr). IR: 2946, 2867, 1744, 1700, 1599, 1576, 1508, 1448, 1237, 1034, 726, 694. HRMS [C₆₆H₇₇N₅O₁₅Si₂ + H]⁺: found 1236.5048, calc. 1236.5028

***N*⁶-Benzoyl, *N*⁶-acetyl-9-[3-*O*-acetyl, 5-*O*-*tert*-butyldiphenylsilyl)-2-*O*-(5-*O*-dimethoxytrityl)- 2, 3-di-*O*-acetyl- α -D-ribofuranosyl]- β -D-ribofuranosyl]-adenine (12)**

Nucleoside **11** (52 μ mol) was dissolved in 4:1 DCM/MeOH (5 mL) at 0 °C and treated with 8M MeNH₂ in EtOH (30 μ L) for 30 min. After TLC analysis showed the complete conversion of the starting material into a lower running spot the reaction mixture was concentrated, redissolved in THF (5 mL) and treated with 1M TBAF in THF (300 μ L) for 2 hours. The reaction mixture was concentrated, coevaporated with pyridine and redissolved in pyridine (5 mL). TBDPSCl (10 eq., 140 μ L) and catalytic DMAP were added and the reaction mixture was stirred at 50 °C for 48 hours. Subsequently Ac₂O (200 μ L) was added and the reaction mixture was left at room temperature for 16 hours. Concentration of the crude was followed by

purification using a silica gel column pre-neutralized using TEA and a gradient of Hexane/ EtOAc (100 : 0 – 20 : 80) afforded a colourless syrup (42 mg) containing the title compound **12** (85% as determined by LCMS, 30 μ mol) and the product lacking the N-acetyl group (15% as determined by LCMS, 6 μ mol). This mixture was purified to allow complete analysis of the target compound **12** using RP-HPLC to afford 13 mg (11 μ mol) of pure **12** (32%). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ : 8.62 (s, 1H, H2), 8.33 (s, 1H, H8), 7.71 (d, J = 7.2 Hz, 2H, Arom.), 7.68 - 7.59 (m, 4H, Arom.), 7.52 - 7.12 (m, 18H, Arom.), 6.81 (d, J = 8.6 Hz, 4H, Arom.), 6.21 (d, J = 6.0 Hz, 1H, H1''), 5.63 - 5.53 (m, 1H, H3''), 5.35 (d, J = 4.6 Hz, 1H, H1'), 5.31 (dd, J = 6.9, 3.0 Hz, 1H, H3'), 5.06 (t, J = 5.6 Hz, 1H, H2''), 4.99 (dd, J = 6.9, 4.6 Hz, 1H, H2'), 4.25 - 4.22 (m, 1H, H4''), 4.18 - 4.17 (m, 1H, H4'), 4.00 (dd, J = 11.6, 2.8 Hz, 1H, H5_A''), 3.86 (dd, J = 11.6, 2.8 Hz, 1H, H5_B''), 3.77 (s, 6H, OMe DMT), 3.34 (dd, J = 10.4, 3.2 Hz, 1H, H5_A'), 3.13 (dd, J = 10.4, 3.2 Hz, 1H, H5_B'), 2.60 (s, 3H, NAc), 2.10 (s, 3H, OAc), 2.05 (s, 3H, OAc), 1.74 (s, 3H OAc), 1.09 (s, 9H, *t*Bu TBDPS). $^{13}\text{C-NMR}$ (101 MHz, CDCl_3) δ : 172.7 (CO Bz), 172.0 (CO N-Ac), 170.2, 169.8, 169.6 (CO O-Ac), 158.6 (Cq DMT), 153.3 (C4), 152.4 (C2), 151.1 (C6), 149.9, 144.6 (Cq Arom.), 143.6 (C8), 135.9, 135.7, 135.6, 135.6, 134.3, 132.9, 132.7, 132.4, 130.2, 130.2, 130.1, 129.4, 129.2, 128.7, 128.2, 128.09, 128.08, 128.0, 127.0 (Arom. Bz, DMT, TBDPS), 113.3 (Arom. DMT), 101.6 (C1'), 86.4 (Cq DMT, C1''), 83.8 (C4''), 81.8 (C4'), 78.7 (C2''), 71.7 (C3''), 71.50 (C2'), 70.48 (C3'), 63.3, 63.2 (C5', C5''), 55.3 (OMe DMT), 27.1 (CH_3 *t*Bu TBDPS), 25.6 (CH_3 NAc), 21.0, 20.9, 20.3 (CH_3 OAc), 19.4 (Cq *t*Bu). α_{D} (CDCl_3) = + 11.2°. IR: 2928, 2361, 2342, 1744, 1601, 1508, 1222, 1114, 1037, 910. HRMS [$\text{C}_{67}\text{H}_{69}\text{N}_5\text{O}_{15}\text{Si} + \text{H}$] $^+$: found 1212.4648, calc. 1212.4632.

Notes and references

- ¹ van der Heden van Noort, G. J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *Org. Lett.*, **2011**, *13*, 2920 – 2923.
- ² van der Heden van Noort, G. J.; van der Horst, M. G.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *J. Am. Chem. Soc.*, **2010**, *132*, 5236 – 5240.
- ³ Moyle, P. M.; Muir, T. W.; *J. Am. Chem. Soc.*, **2010**, *117*, 12653 – 12654.
- ⁴ Scheuring, J.; Schramm, V. L.; *J. Am. Chem. Soc.*, **1995**, *117*, 15878 – 15880.
- ⁵ Knapp, S.; Gore, V. K.; *J. Org. Chem.*, **1996**, *61*, 6744 – 6747.
- ⁶ Efimtseva, E. V.; Kulikova, I. V.; Mikhailov, S. N.; *Curr. Org. Chem.*, **2007**, *11*, 337 – 354.
- ⁷ Duynstee, H. I.; Wijsman, E. R.; van der Marel, G. A.; van Boom, J. H.; *Synlett*, **1996**, 313.
- ⁸ Kobayashi, J.; Doi, Y.; Ishibashi, M.; *J. Org. Chem.*, **1994**, *59*, 255 – 257.
- ⁹ Ning, J.; Xing, Y.; Kong, F.; *Carbohydr. Res.*, **2003**, *338*, 55 – 60.
- ¹⁰ de Kort, M.; Regenbogen, A. D.; Overkleeft, H. S.; Chaliss, R. A. J.; Iwata, Y.; Miyamoto, S.; van der Marel, G. A.; van Boom, J. H.; *Tetrahedron*, **2000**, *56*, 5915 – 5928.
- ¹¹ van Straten, N. C. R.; van der Marel, G. A.; van Boom, J. H.; *Tetrahedron*, **1997**, *53*, 6509 – 6522.
- ¹² Borissow, C. N.; Black, S. J.; Paul, M.; Tovey, S. C.; Dedos, S. G.; Taylor, C. W.; Potter, B. V. L.; *Org. Biomol. Chem.*, **2005**, *3*, 245 – 252.
- ¹³ Sureshan, K. M.; Trusselle, M.; Tovey, S. C.; Taylor, C. W.; Potter, B. V. L.; *J. Org. Chem.*, **2008**, *73*, 1682 – 1692.
- ¹⁴ Beecroft, M. D.; Marchant, J. S.; Riley, A. M.; van Straten, N. C. R.; van der Marel, G. A.; van Boom, J. H.; Potter, B. V. L.; Taylor, C. W.; *Mol. Pharmacol.*, **1999**, *55*, 109 – 117.
- ¹⁵ de Kort, M.; Correa, V.; Valentijn, A. R. P. M.; van der Marel, G. A.; Potter, B. V. L.; Taylor, C. W.; van Boom, J. H.; *J. Med. Chem.*, **2000**, *43*, 3295 – 3303.
- ¹⁶ de Kort, M.; Luijendijk, J.; van der Marel, G. A.; van Boom, J. H.; *Eur. J. Org. Chem.*, **2000**, *17*, 3085 – 3092.
- ¹⁷ Mikhailov, S. N.; Efimtseva, E. V.; Gurskaya, G. V.; Fomitcheva, M. V.; Meshkov, S. V.; *J. Carbohydr. Chem.*, **1997**, *16*, 75 – 92.
- ¹⁸ Efimtseva, E. V.; Bobkov, G. V.; Mikhailov, S. N.; Van Aerschot, A.; Schepers, G.; Busson, R.; Rozenski, J.; Herdewijn, P.; *Helv. Chim. Acta*, **2001**, *84*, 2387 – 2397.
- ¹⁹ Andreeva, O. I.; Golubeva, A. S.; Kochetkov, S. N.; Van Aerschot, A.; Herdewijn, P.; Efimtseva, E. V.; Ermolinsky, B. S.; Mikhailov, S. N.; *Bioorg. Med. Chem. Lett.*, **2002**, *12*, 681 – 684.
- ²⁰ Mikhailov, S. N.; Kulikova, I. V.; Nauwelaerts, K.; Herdewijn, P.; *Tetrahedron*, **2008**, *64*, 2871 – 2876.
- ²¹ Uchiro, H.; Mukaiyama, T.; *Chem. Lett.*, **1996**, 271 – 272.

- ²² Mukaiyama, T.; Suda, S.; *Chem. Lett.*, **1990**, 1143 – 1146.
- ²³ Larsen, C. H.; Ridgway, B. H.; Shaw, J. T.; Smith, D. M.; Woerpel, K. A.; *J. Am. Chem. Soc.*, **2005**, *127*, 10879 – 10884.
- ²⁴ Smith, D. M.; Tran, M. B.; Woerpel, K. A.; *J. Am. Chem. Soc.*, **2003**, *125*, 14149 – 14152.
- ²⁵ Hirano, S.; Ichikawa, S.; Matsuda, A.; *J. Org. Chem.*, **2007**, *72*, 9936 – 9946.
- ²⁶ Takahashi, H.; Isobe, M.; Goto, T.; *Tetrahedron*, **1991**, *47*, 6215 – 6222.
- ²⁷ Prevost, M.; St-Jean, O.; Guindon, Y.; *J. Am. Chem. Soc.*, **2010**, *132*, 12433 – 12439.
- ²⁸ Rhoad, J. S.; Cagg, B. A.; Carver, P. W.; *J. Phys. Chem. A*, **2010**, *114*, 5180 – 5186.
- ²⁹ Stalford, S. A.; Kilner, C. A.; Leach, A. G.; Turnbull, W. B. *Org. Biomol. Chem.* **2009**, *7*, 4842 – 4852.
- ³⁰ Markiewicz, W. T.; *J. Chem. Res-S.*, **1979**, 24 – 25.
- ³¹ van der Heden van Noort, G. J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *J. Org. Chem.*, **2010**, *75*, 5733 – 5736.
- ³² The α -selectivity in this reaction is proposed to originate from the equilibrium of the formed α - and β -phosphonium intermediates, of which the β -form would be more reactive towards S_N2 -substitution and thus govern 1', 2' - cis ribofuranoside formation. Although indeed the α - and β -phosphonium intermediates are detectable using ^{31}P -NMR (δ : 98.0, 94.5) this explanation does not exclude subsequent formation of the oxocarbenium ion and the related selectivity as explained by the Woerpel model.
- ³³ When this procedure was adopted for the glycosylation of 3,5-OTIPDS protected uridine, formation of 2'-O-(2',3',5'-tri- O-Bn- α -ribosyl)uridine was observed, however in low yield (15%).
- ³⁴ Vorbrüggen, H.; Höfle, G.; *Chem. Ber.*, **1981**, *114*, 1256 – 1268.
- ³⁵ Framski, G.; Gdaniec, Z.; Gdaniec, M.; Boryski, J.; *Tetrahedron*, **2006**, *62*, 10123 – 10129.
- ³⁶ Boryski, J.; *Curr. Org. Chem.*, **2008**, *12*, 309 – 325.
- ³⁷ The NOESY experiment shows crosspeaks between the ribose H1' and H3' and no crosspeak between the ribose H1' and H4', indicating an α -glycosidic linkage. The HMBC experiment shows crosspeaks between the ribose H1' and adenosine C6 and C2, indicating that the ribose moiety is attached to N¹.
- ³⁸ Moreau, C.; Ashamu, G. A.; Bailey, V. C.; Galione, A.; Guse, A. H.; Potter, B. V. L.; *Org. Biomol. Chem.*, **2011**, *9*, 278 – 290.
- ³⁹ Busca, P.; Etheve-Quelquejeu, M.; Valery, J.; *Tetrahedron Lett.*, **2003**, *44*, 9131 – 9134.
- ⁴⁰ Martin, O.R.; *Carbohydr. Res.*, **1987**, *171*, 211 – 222.
- ⁴¹ McCormick, J.; Li, Y.; McCormick, K.; Duynstee, H. I.; van Engen, A. K.; van der Marel, G. A.; Ganem, B.; van Boom, J. H.; Meinwald, J.; *J. Am. Chem. Soc.*, **1999**, *121*, 5661 – 5665.
- ⁴² Plante, O. J.; Palmacci, E. R.; Seeberger, P. H.; *Org. Lett.*, **2000**, *2*, 3841 – 3843.

- ⁴³ Yu, B.; Tao, H.; *Tetrahedron Lett.*, **2001**, 42, 2405 – 2407.
- ⁴⁴ The exocyclic amine of adenosine was dibenzoylated to avoid the side reaction observed when using the Mukaiyama protocol.
- ⁴⁵ Serebryany, V.; Beigelman, L.; *Tetrahedron Lett.*, **2002**, 43, 1983 – 1985.
- ⁴⁶ van der Heden van Noort, G. J.; Verhagen, C. P.; van der Horst, M. G.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *Org. Lett.*, **2008**, 10, 4461 – 4464.

Chapter 5

A General Synthetic Method Towards Uridylylated Picornavirus VPg Proteins

Introduction

Picornaviridae are single stranded, positive sense RNA viruses sub-classified in the genera: *Enterovirus*, *Rhinovirus*, *Hepatovirus*, *Parechovirus*, *Kobuvirus*, *Erbovirus*, *Cardiovirus*, *Aphthovirus* and *Teschovirus*.¹ Well-known examples of notorious human pathogens from this family of viruses are Rhinovirus, Hepatitis A virus and Poliovirus, which cause the common cold², acute liver infection³ and inflammation of the central neural system⁴, respectively. Other members of this family of viruses are known agricultural pathogens, such as Foot-and-mouth-disease virus which causes high fever and blisters in the mouth and on the feet of cloven-hoofed animals.⁵ Encephalomyocarditis virus⁶ and Theiler's virus⁷, which cause inflammation of the heart muscle and brain in rodents, are other examples.

All these viruses share the same genomic structure; a 5' non-translated region (NTR) with an internal ribosome entry site and specific cloverleaf structure, an open reading frame coding for all viral capsid and replication proteins, a 3' NTR and a 3' poly(A)-tail.¹ In addition, the most intriguing feature is that picornaviruses carry a small VPg-protein (Viral Protein genome-linked) at the 5'-end of their genome whereas in other species the 5'-end of the RNA is commonly protected with a cap structure such as the modified nucleotide 7-methyl guanylate, which is attached via an 5' - 5' triphosphate linkage.^{8,9} The VPg protein is a 22-24 amino acid sequence that is highly conserved in all picornaviruses, and plays a crucial role in viral replication.⁴

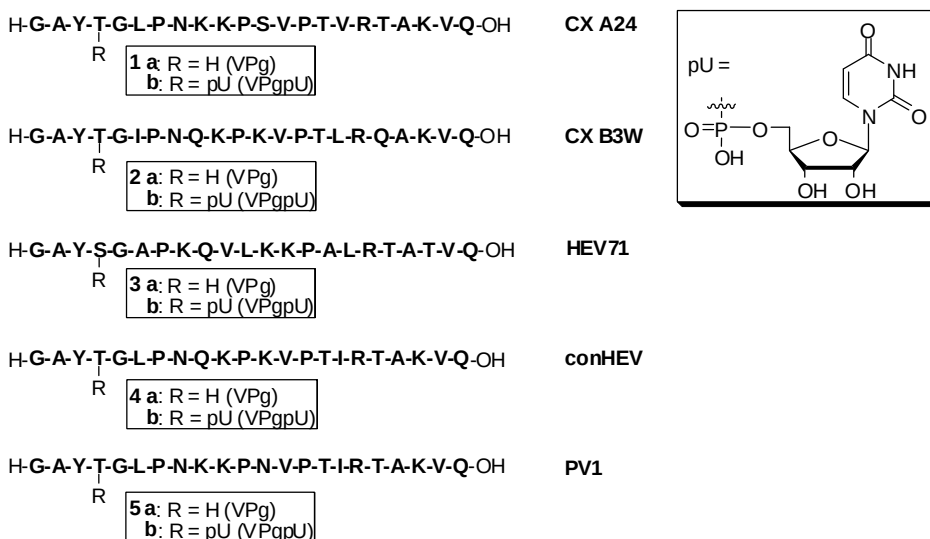
Upon infection by picornaviruses the viral genome is transcribed by the host ribosome generating viral replication proteins, amongst which the protein VPg. The positive sense genome is copied using its own replication proteins into a negative sense RNA strand, which acts as template for the formation of 40 to 70 copies of progeny positive sense RNA.¹¹ Since

more (+) strand RNA is formed than (-) strand RNA the replication is called asymmetric, indicating multiple replication mechanisms must be involved.

Since both strands are 5' capped with the VPg protein, it was long believed that uridylylated VPg was involved as primer in both (+) and (-) strand synthesis.¹² However, it is now proposed that this uridylylated VPg only acts as an essential primer in the synthesis of progeny (+) strand RNA, using the 3'- poly(A)-tail of the genome as template. The viral 3D polymerase is responsible for uridylylating VPg using a functional domain in the viral genome called the cis-replication element (CRE) as template.^{13,14} In contrast, (-) strand RNA synthesis is now believed to be primed by the non-uridylylated VPg-protein.^{15,16} These findings, suggesting a different primer for (-) and (+) strand formation, seem more in line with the asymmetric replication found for picornaviruses. The exact mode of action in this replication process, however, remains unclear.

Outbreaks of picornavirus related diseases are rapid due to their high virulence and ease of communicability. The consequences can be severe, as demonstrated by the outbreak of Foot-and-mouth-disease virus (FMDV) among United Kingdom farm animals in 2001.¹⁷ Understanding the role of the RNA linked VPg protein in viral replication in more detail might give insight in the development of new vaccines and drugs against diseases caused by these viruses. Synthetically made VPgpU and VPgpUpU proved to be useful tools in the elucidation of Poliovirus replication.^{18,19}

Figure 1. Coxsackie A24, Coxsackie B3W, Human Enterovirus 71, consensus Human Enterovirus and Poliovirus 1 VPg(pU)'s

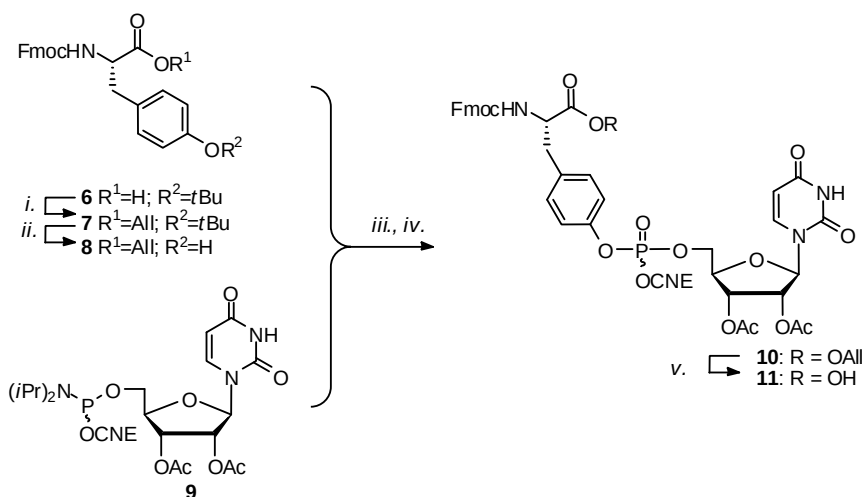


Previous reports describe the synthesis of Poliovirus VPgpU using a block condensation strategy²⁰ and the online SPPS synthesis of Poliovirus, Coxsackie B3 and Cowpea mosaic virus VPgpU's.²¹ As an extension of the latter procedure this chapter describes the solid phase synthesis of the VPg proteins as well as the VPgpU nucleoproteins from two additional coxsackie strands (CX A24 and CX B3W), human enterovirus 71 (HEV71), Poliovirus (PV1) and a consensus sequence for human enterovirus (conHEV)²² (see **Figure 1**). It was envisaged that the application of a pre-uridylylated tyrosine building block would result in a universal solid phase synthesis strategy for all Tyr-O⁴ linked VPgpU's.

Results and Discussion

Guided by an earlier synthesis, uridylylated tyrosine building block **11** was selected for the solid phase synthesis of the target VPgpU's (see **Scheme 1**).²⁰ A modified route of synthesis to **11** was undertaken that started with the protection of commercially available Fmoc-Tyr(*t*Bu)-OH **6** as its allyl ester. When cesium carbonate and allyl bromide were used to introduce the allyl ester, the Fmoc group in **6** was found to be cleaved. To avoid this unwanted side-reaction silver carbonate was used as a base and **7** was obtained in quantitative yield. Subsequent removal of the *t*-butyl ether using TFA gave phenol **8**. Phosphitylation of **8** with uridine-5'-phosphoramidite **9** under the agency of 4,5-dicyanoimidazole and subsequent oxidation of the intermediate phosphite using *t*-BuOOH yielded fully protected tyrosine **10**. Liberation of the acid functionality by palladium catalyzed removal of the allyl protecting group gave target tyrosine derivative **11**. In this reaction tributyltin hydride acts as a nucleophile, a process reported to proceed very efficient in allyl and Alloc cleavage in amino acid and peptide derivatives.²³

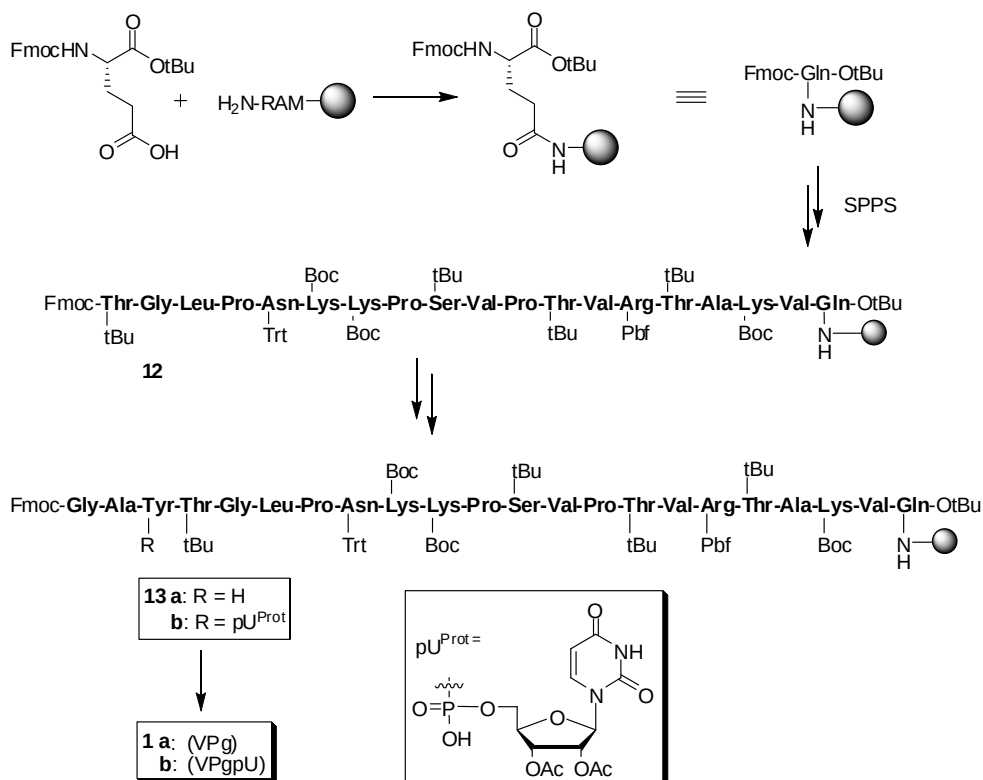
Scheme 1



Reagents and conditions: i. Ag_2CO_3 , DMF, r.t., 15 min; then All-Br, 2.5 h, r.t., (quantitative), ii. 70% TFA in DCM, 30 min, r.t., (quantitative), iii. 4,5-dicyanoimidazole, MeCN, 30 min, r.t., iv. *t*-BuOOH, 45 min, rt, (90% based on **8**), v. $Pd(PPh_3)_4$, Bu_3SnH , AcOH, THF/DCM, 1h, r.t., (62%).

The solid phase synthesis of VPgpU's **1b** – **5b** started with immobilization of the first glutamic acid residue via its side chain acid moiety to Rink amide resin (Tentagel RAM).²¹ This was achieved by HATU mediated coupling of Fmoc-Glu-OtBu to the resin. During the final acidic cleavage of the VPgpU from the resin a carboxamide was generated, converting the initially attached glutamic acid into the desired C-terminal glutamine residue (see **Scheme 2**).

Scheme 2. SPPS for Coxsackie A24 VPg **1a** and VPgpU **1b**



Preparation of Coxsackie A24 VPgpU **1b** is elaborated by removal of the Fmoc protection and repetitive elongation of the resin bound glutamine using standard automated Fmoc solid phase peptide synthesis and HATU as coupling reagent, to give the immobilized 19-mer peptide **12**. The immobilized peptide **12** was elongated performing a manual double coupling protocol using Fmoc-Tyr(pU)-OH **11** and BOP/HOBt as activating agents. Subsequent piperidine treatment liberated the terminal amine, and concomitantly cleaved the cyanoethyl group resulting in the formation of a phosphodiester. The presence of the phosphodiester gave no problems in the subsequent attachment of the last two N-terminal amino acids, as no side reactions were detected. Final removal of the Fmoc group followed by treatment with a mixture of 95% TFA, 2.5 % TIS and 2.5% H₂O, to effectuate removal of acid labile protective groups and release from the solid support yielded crude partially protected VPgpU. An alkaline treatment with NH₄OH in dioxane to remove 2'-O and 3'-O acetyl groups from the ribose

residue yielded the completely deprotected crude nucleopeptide, which was purified using RP-HPLC to afford homogeneous Cocksackie A24 VPgpU nucleoprotein **1b**.

Elongation of resin bound peptide **12** with Fmoc-Tyr(tBu)-OH, Fmoc-Ala-OH and Fmoc-Gly-OH gave protected VPg peptide **13a**. A final Fmoc deprotection step and acid treatment followed by RP-HPLC purification afforded Cocksackie A24 VPg protein **1a**.

In a similar fashion, using uridilylated tyrosine **11** or commercially available Fmoc-Tyr(tBu)-OH the preparation of nucleoproteins **2b** – **5b** and proteins **2a** – **5a**, respectively, proceeded uneventfully in a yield between 3 and 35%.²⁴

Conclusion

In this chapter suitably protected pre-uridilylated Fmoc tyrosine building block **11** is synthesized. The general applicability of a sequential solid phase based strategy is demonstrated by the successful synthesis of five picornavirus VPg proteins **1a** – **5a** and five uridilylated VPg proteins **1b** – **5b**.

Experimental

General: All chemicals were used as received. MeCN and THF were stored over activated 4Å molecular sieves at least 24 hours prior to use. DCM was distilled over CaH₂ prior to use. Solvents were removed by rotary evaporation under reduced pressure at a temperature not exceeding 40°C. TLC analysis, silica gel column chromatography, NMR, LCMS, HPLC and HRMS techniques used are described in **Chapter 2**.

Procedure for solid phase peptide synthesis: The peptides were prepared on an ABI-433A (Applied Biosystems, division of Perkin-Elmer) automated peptide synthesizer using the Fmoc based peptide synthesis protocol. The peptides were prepared on 50 µmol scale starting from Tentagel S RAM resin (with the exception of Coxsackie A24 VPg and VPgpU; 100 µmol). The first amino acid (Fmoc-Glu-OtBu) was coupled to the resin via the side-chain and converted in a Gln residue in course of the final cleavage off the resin. The consecutive steps performed in each cycle were:

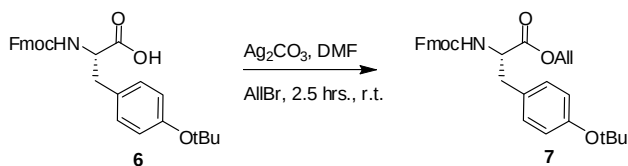
- Deprotection of the Fmoc-group with 20 % piperidine in NMP for 5 x 1 min.
- Coupling of the appropriate amino acid applying a five-fold excess. Generally the amino acid (0.25 mmol) was dissolved in NMP (0.5 mL) and subsequently 0.25 mmol of HATU (0.25 M in DMF/NMP, 1/1, v/v) and 0.32 mmol of DiPEA (1.0 M in NMP) were added. The resulting solution was transferred to the reaction vessel, which was shaken for 1 hour.
- The unreacted amine functions were capped using 0.5 M acetic anhydride, 0.125 M DiPEA and 0.015 M HOBt in NMP. The resulting suspension was shaken for 1 min.

Double couplings were conducted on β-branched amino acids and on Fmoc-Asn(Trt)-OH and Fmoc-Arg(Pbf)-OH.

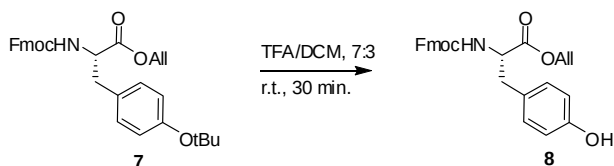
The 19-mer peptide containing resin was split in two equal portions of which one portion was elongated on the automated peptide synthesizer using a double coupling protocol for the last three AA's to obtain VPg peptides **1a** – **5a**. In the synthesis of the VPgpU's **1b** – **5b** the last three amino acids, Fmoc-Tyr(pU)-OH, Fmoc-Ala-OH and Fmoc-Gly-OH, were attached manually using a double coupling protocol. The consecutive steps performed in each cycle were:

- Deprotection of the Fmoc-group with 2 % DBU in DMF for 5 x 1 min.
- Coupling of the appropriate amino acid applying a five-fold excess. Generally the amino acid (0.25 mmol) was dissolved in NMP (0.5 mL) and subsequently 0.25 mmol of BOP/HOBt (0.5 M BOP/ 0.5 M HOBt in DMF/NMP, 1/1, v/v) and 0.32 mmol of DiPEA (1.25 M in NMP) were added. The resulting solution was transferred to the reaction vessel, which was shaken for 1 hour.
- The unreacted amine functions were capped using 0.5 M acetic anhydride, 0.125 M DiPEA and 0.015 M HOBt in NMP. The resulting suspension was shaken for 1 min.

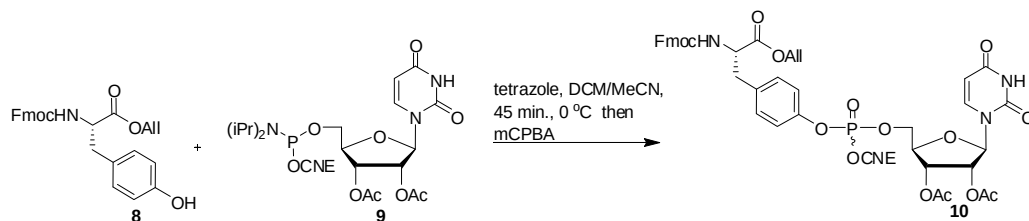
After final Fmoc deprotection, the peptide was cleaved off the resin using TFA/TIS/H₂O (95/5/5, v/v/v, 10 mL) by shaking for 4 hours. After filtration into cold Et₂O (35 mL) the precipitate was collected and coevaporated with toluene (2 x 5 mL) and then treated with 25% aq. ammonia (2.5 mL) in 1,4-dioxane (2.5 mL). After stirring for 4 hours concentration of the reaction mixture afforded the crude title compound. RP-HPLC purification was conducted on a Gilson automated HPLC supplied with a semi-preparative Phenomenex Gemini C₁₈ column (250 x 10.00 mm, 5 µ, flow: 4 mL/min). The applied eluent system was A: 0.5 % TFA in H₂O, B: MeCN and detection at 250 nm.

***N*^α- Fmoc-(*O*-*tert*-butyl)-tyrosine allyl ester (7)**

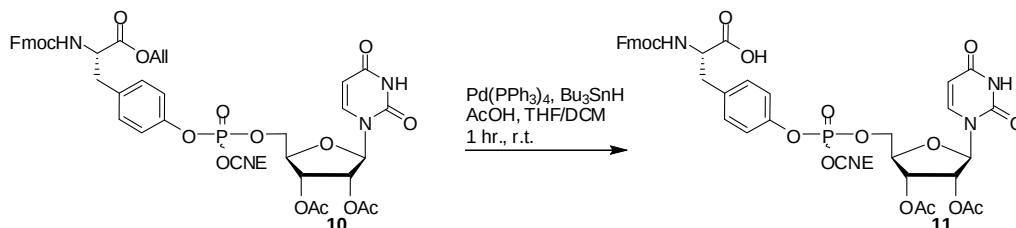
To a solution of *N*^α- Fmoc-(*O*-*tert*-butyl)-tyrosine **6** (2.3 g, 5.0 mmol) in DMF (25 mL) at 0 °C was added Ag₂CO₃ (2.7 g, 9.8 mmol) under exclusion of light. The reaction mixture was warmed to room temperature and stirred for 15 min. prior to addition of allylbromide (2.0 mL, 23 mmol). After 2.5 hours TLC analysis (70/29/1 EtOAc: PE: AcOH) indicated complete consumption of the starting material. The reaction mixture was filtered over celite and the filtrate was diluted in 100 mL EtOAc. The organic phase was subsequently washed with 10% KHSO₄ (2 x 25 mL) and water (2 x 30 mL), dried (MgSO₄) and concentrated. The residue was applied to a silica gel column and eluted with a gradient of EtOAc in PE (0/100 – 40/60, v/v) to yield the title compound quantitatively. ¹H-NMR (400 MHz, CDCl₃) δ 7.75 (d, *J* = 7.5 Hz, 2H, Arom. Fmoc), 7.61 - 7.52 (m, 2H, Arom. Fmoc), 7.38 (t, *J* = 7.4 Hz, 2H, Arom. Fmoc), 7.30 (t, *J* = 7.3 Hz, 2H, Arom. Fmoc), 7.11 - 6.95 (m, 2H, Arom. Tyr.), 6.89 (d, *J* = 7.7 Hz, 2H, Arom. Tyr.), 5.84 (ddd, *J* = 21.2, 10.2, 5.0 Hz, 1H, CH Allyl), 5.36 (d, *J* = 8.2 Hz, 1H, NH), 5.33 - 5.15 (m, 3H, CH₂ Allyl), 4.74 - 4.49 (m, 3H, CH α, CH₂ Allyl), 4.47 - 4.29 (m, 2H, CH₂ Fmoc), 4.19 (t, *J* = 6.8 Hz, 1H, CH Fmoc), 3.19 - 2.96 (m, 2H, CH₂ β), 1.31 (s, 9H, CH₃ *t*Bu). ¹³C-NMR (101 MHz, CDCl₃) δ 171.3 (CO α), 155.6, 154.5 (CO Fmoc, Cq Tyr.), 143.9, 143.8, 141.3 (Cq Arom. Fmoc), 131.5 (CH Allyl), 130.5 (Cq Tyr.), 129.87 (Arom. Tyr.), 129.82, 127.76, 127.1, 125.2, 125.1, 124.2, (Arom. Fmoc), 120.0 (CH₂ Allyl), 119.1 (Arom. Tyr.) 78.4 (Cq *t*Bu), 67.0 (CH₂ Fmoc), 66.1 (CH₂ Allyl), 54.9 (CH α, 47.2 (CH Fmoc), 37.7 (CH₂ β), 28.9 (CH₃ *t*Bu). IR: 2980, 1718, 1506, 1236, 1161, 757, 740. α_D (CHCl₃) = + 14.6°. HRMS: [C₃₁H₃₃NO₅ + H]⁺; calcd. 500.2432, found 500.2431.

***N*^α- Fmoc-tyrosine-allyl ester (8)**

Allyl ester **7** was dissolved in 30 mL TFA/DCM (7:3, v/v) and stirred for 30 min. at room temperature. Toluene (50 mL) was added and the solvents were removed. The residue was applied to a silica gel column and eluted with a gradient of EtOAc in toluene (0/100 – 10/90, v/v) to yield the title compound quantitatively. ¹H-NMR (400 MHz, CDCl₃) δ 7.73 (d, *J* = 7.5 Hz, 2H, Arom. Fmoc), 7.58 – 7.47 (m, 2H, Arom. Fmoc), 7.37 (t, *J* = 7.4 Hz, 2H, Arom. Fmoc), 7.28 (t, *J* = 7.3 Hz, 2H, Arom. Fmoc), 6.92 (d, *J* = 8.0 Hz, 2H, Arom. Tyr.), 6.70 (d, *J* = 8.0 Hz, 2H, Arom. Tyr.), 6.36 (s, 1H, OH), 5.90 - 5.81 (m, 1H, CH Allyl), 5.40 (d, *J* = 8.3 Hz, 1H, NH), 5.36 - 5.19 (m, 2H, CH₂ Allyl), 4.77 - 4.49 (m, 3H, CH₂ Allyl, CH α), 4.42 - 4.32 (m, 2H, CH₂ Fmoc), 4.17 (t, *J* = 6.8 Hz, 1H, CH Fmoc), 3.09 - 2.97 (m, 1H, CH₂ β). ¹³C-NMR (101 MHz, CDCl₃) δ 171.7 (CO α), 156.0, 155.3 (CO Fmoc, Cq Tyr.), 143.8, 143.7, 141.4 (Cq Arom. Fmoc), 130.5 (CH Allyl), 127.8 (Arom. Tyr.), 127.2, 125.2, 125.15, 120.1 (Arom. Fmoc), 119.4 (CH₂ Allyl), 115.6 (Arom. Tyr.), 67.2 (CH₂ Fmoc), 66.3 (CH₂ Allyl), 55.1 (CH α), 47.2 (CH Fmoc), 37.5 (CH₂ β). IR: 3317, 1689, 1516, 1450, 1214, 758, 739. α_D (CHCl₃) = + 16.8°. HRMS: [C₂₇H₂₅NO₅ + H]⁺; calcd. 444.1806, found 444.1804.

2-cyanoethyl-(*N*^α- Fmoc-tyrosin-4-yl-allyl ester)-(2', 3'-di-*O*-acetyluridyl-5'-yl) phosphate (10)


Phosphoramidite **9** was prepared by reacting 2', 3'-di-*O*-acetyluridine (0.98 g, 3 mmol), coevaporated with MeCN (2 x 7.5 mL) in DCM (15 mL), containing TEA (1.4 mL, 10 mmol) with 2-cyanoethoxy-*N,N*-diisopropylaminochlorophosphine (0.75 mL, 3.75 mmol) under an argon atmosphere. The reaction mixture was stirred for 45 min at room temperature. After ³¹P NMR showed complete consumption of the aminochlorophosphine and formation of phosphoramidite **9** (161 MHz, CDCl₃; δ 150.49, 148.68), 30 mL DCM was added and the reaction mixture was washed with water (20 mL) and 5% NaHCO₃ (10 mL). The organic phase was dried (MgSO₄) and concentrated. Approximately 0.6 mmol **9** was added to a solution of *N*^α- Fmoc-tyrosine-allyl ester (**8**) (0.21 g, 0.47 mmol) in 1 MeCN (10 mL). 4,5-Dicyanoimidazole (0.15 g, 0.6 mmol) was added and the reaction mixture was stirred for 30 min. until ³¹P NMR showed complete consumption of the phosphoramidite and formation of the phosphite (161 MHz, CDCl₃; δ 135.12, 134.97). A solution of *t*BuOOH (5.5 M in nonane) was added to the reaction mixture until after 45 minutes ³¹P NMR showed complete consumption of the phosphite and formation of the phospho triester (161 MHz, CDCl₃; δ -6.48, -6.83). The reaction mixture was diluted in 20 mL DCM and washed with 5% Na₂S₂O₃ and water. The organic phase was dried with MgSO₄ and concentrated. The residue was applied to a silica gel column and eluted with a gradient of MeOH in DCM (0/100 – 2/98, v/v), to yield the title compound as a mixture of epimers at P as a white foam in 90 % (0.37 g, 0.42 mmol) based on **8**. ¹H-NMR (400 MHz, CDCl₃); δ 9.78 (d, *J* = 10.0 Hz, 1H, NH U), 7.76 (ap. d, *J* = 7.5 Hz, 2H, Arom. Fmoc), 7.62 - 7.53 (m, 2H, Arom. Fmoc), 7.45 - 7.35 (m, 3H, Arom. Fmoc, H6), 7.34 - 7.26 (m, 2H, Arom. Fmoc), 7.20 - 7.04 (m, 4H, Arom. Tyr), 6.13 - 6.00 (m, 1H, H1'), 5.96 - 5.82 (m, 1H, CH Allyl), 5.79 - 5.58 (m, 2H, H5, NH α), 5.45 - 5.37 (m, 1H, H3'), 5.37 - 5.12 (m, 3H, H2', CH₂ Allyl), 4.71- 4.25 (m, 10H, 4', 5', CH₂ All, CH₂ CNEO, CH₂ Fmoc, CH α), 4.20 (t, *J* = 7.0 Hz, 1H, CH Fmoc), 3.21 - 3.01 (m, 2H, CH₂ β), 2.86 - 2.61 (m, 2H, CH₂ CNEO), 2.18 - 2.02 (m, 6H, OAc's). ¹³C-NMR (101 MHz, CDCl₃); δ 171.1, 171.0 (CO α), 169.83, 169.80, 169.73, 169.71 (CO Ac), 163.2, 163.1 (C4), 155.7 (CO Fmoc), 150.5 (C2), 149.03, 148.96 (Cq Tyr.), 143.84, 143.75, 141.29 (Cq Fmoc), 139.9, 139.7 (C6), 133.8 (Cq Tyr.), 131.4, 131.12 (Arom. Tyr.), 127.76, 127.09, 125.16, 125.11 (Aro. Fmoc), 120.0 (Arom. Tyr.), 119.3 (Arom. Fmoc), 117.9 (CH₂ Allyl), 117.87, 116.44 (CN), 103.5 (C5), 87.3, 87.0 (C1'), 80.59, 80.54, 80.52, 80.47 (C4'), 72.5 (C3'), 69.83, 69.82 (C2'), 67.36, 67.30, 67.24, 66.92, 66.88, 66.21, 66.17 (CH₂ Allyl, CH₂ CNEO, CH₂ Fmoc), 63.18, 63.08 (C5'), 54.8 (CH α), 47.1 (CH Fmoc), 37.3 (CH₂ β), 20.56, 20.54, 20.4 (CH₃ Ac), 19.7, 19.6 (CH₂ CNEO). ³¹P-NMR (81 MHz, CDCl₃); δ -6.48, -6.83. IR: 3067, 1694, 1506, 1452, 1378, 1211, 1037, 961. LCMS (10 - 90 % B in 15 min), Rt = 8.91. ESI-MS; *m/z* 887.40 [M + H]⁺. HRMS: [C₄₃H₄₃N₄O₁₅P + H]⁺; calcd. 887.2535, found 887.2542.

2-cyanoethyl-(*N*^ε- Fmoc-tyrosin-4-yl)-(2', 3'-di-*O*-acetyluridyl-5'-yl) phosphate (11)

To a stirred solution of **10** (2.15 g, 2.4 mmol) coevaporated with MeCN (7.5 mL x 2) in 1:1 THF/DCM (10 mL, v/v), containing AcOH (0.6 mL, 11 mmol), was added Bu₃SnH (1.3 mL, 4.9 mmol) and Pd(PPh₃)₄ (115 mg, 0.1 mmol) under an argon atmosphere. After one hour TLC analysis (10% MeOH/DCM) revealed complete consumption of the starting material. The reaction mixture was concentrated and purified using a silica gel column with a gradient of MeOH in DCM (0/100 – 4/96), to yield the title compound as a mixture of epimers as white foam in 62 % (1.26 g, 1.49 mmol). ¹H-NMR (400 MHz, CDCl₃) δ 9.99 (s, 1H), 7.73 – 7.13 (m, 13H), 6.00 (dd, *J* = 15.1, 5.3 Hz, 1H), 5.94 – 5.80 (m, 1H), 5.61 (dd, *J* = 56.2, 8.0 Hz, 1H), 5.49 – 5.36 (m, 1H), 5.35 – 5.27 (m, 1H), 4.68 – 4.62 (m, 1H), 4.51 – 4.37 (m, 3H), 4.37 – 4.23 (m, 4H), 4.21 – 4.08 (m, 1H), 3.26 – 3.00 (m, 2H), 2.82 – 2.55 (m, 2H), 2.19 – 1.95 (m, 6H). ¹³C-NMR (100 MHz, CDCl₃) δ 173.4 (CO α), 169.77, 169.75, 169.7, 169.6 (CO Acetyl), 164.2, 164.0 (Ur. 4), 155.7 (CO Fmoc), 150.24, 150.17 (Ur. 6), 148.83, 148.77 (Cq Tyr.), 143.8, 143.7, 141.2 (Cq Fmoc), 134.0 (Cq Tyr.), 131.2, 131.1 (Arom. Tyr.), 127.6, 127.0, 125.1 (Arom. Fmoc), 119.9, 119.7 (Arom. Fmoc, Arom. Tyr.), 116.5 (CN), 103.0, 102.9 (Ur. 5), 87.7, 87.3 (C1'), 80.5, 80.4 (C4'), 72.6, 72.5 (C2'), 69.7, 69.6 (C3'), 67.3, 67.1 (CH₂ Fmoc), 66.8 (C5'), 63.2, 63.1 (CH₂ CNEO), 54.6 (Cα), 47.1 (CH Fmoc), 37.0 (Cβ), 20.40, 20.37, 20.3 (CH₃ OAc), 19.53, 19.45 (CH₂ CNEO). ³¹P-NMR (162 MHz, CDCl₃) δ -6.59, -6.95. ESI-MS: *m/z* 847.40 [M+H]⁺. LCMS (10 – 90 B in 15 min.), Rt = 7.75. HRMS: [C₄₀H₃₉N₄O₁₅P + H]⁺; calcd. 847.2222, found 847.2229.

Coxsackie A24 Virus VPg, (H-Gly-Ala-Tyr-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Ser-Val-Pro-Thr-Val-Arg-Thr-Ala-Lys-Val-Gln-OH), (1a)

HPLC gradient: 10 - 60 % B in 3 CV. The appropriate fractions were lyophilized, yielding 22.47 mg, 7.79 μmol (16 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 – 50 % B in 15 min.), Rt = 6.54 min. ESI-MS: *m/z* 771.47 [M + 2H]²⁺, 1157.13 [M + 3H]³⁺. HRMS [C₁₀₃H₁₇₄N₃₀O₃₀ + 2H]²⁺; calcd. 1156.6578, found 1156.6577, [C₁₀₃H₁₇₄N₃₀O₃₀ + 3H]³⁺; calcd. 771.4410, found 771.4407, [C₁₀₃H₁₇₄N₃₀O₃₀ + 4H]⁴⁺; calcd. 578.8325, found 578.8323, [C₁₀₃H₁₇₄N₃₀O₃₀ + 5H]⁵⁺; calcd. 463.2675, found 463.2674. ¹H NMR (600 MHz, D₂O) δ 7.17 (d, *J* = 8.5 Hz, 2H), 6.83 (d, *J* = 8.6 Hz, 2H), 4.70 - 4.61 (m, 2H), 4.58 (ap. dd, *J* = 8.7, 5.5 Hz, 1H), 4.54 - 4.46 (m, 4H), 4.45 - 4.40 (m, 2H), 4.39 - 4.28 (m, 7H), 4.28 - 4.19 (m, 3H), 4.19 - 4.10 (m, 3H), 3.92 - 3.80 (m, 8H), 3.77 - 3.70 (m, 1H), 3.70 - 3.61 (m, 2H), 3.22 (ap. t, *J* = 7.0 Hz, 2H), 3.14 - 2.92 (m, 8H), 2.89 - 2.68 (m, 2H), 2.44 - 2.25 (m, 5H), 2.20 - 1.42 (m, 41H), 1.40 (ap. d, *J* = 7.2 Hz, 5H), 1.33 (ap. d, *J* = 7.2 Hz, 3H), 1.23 (ap. dd, *J* = 8.3, 6.5 Hz, 7H), 1.17 (ap. d, *J* = 6.4 Hz, 3H), 1.00 (ap. d, *J* = 6.7 Hz, 4H), 0.97 - 0.89 (m, 19H).

Coxsackie A24 Virus VPgU, (H-Gly-Ala-Tyr(pU)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Ser-Val-Pro-Thr-Val-Arg-Thr-Ala-Lys-Val-Gln-OH), (1b)

HPLC gradient: 10 - 60 % B in 3 CV. The appropriate fractions were lyophilized, yielding 54.72 mg, 17.19 μmol (34 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 – 50 % B in 15 min.), Rt = 6.32 min. ESI-MS: *m/z* 1310.13 [M+2H]²⁺, 873.73 [M+3H]³⁺. HRMS [C₁₁₂H₁₈₅N₃₂O₃₈P + 2H]²⁺; calcd. 1309.6705, found 1309.6705, [C₁₁₂H₁₈₅N₃₂O₃₈P + 3H]³⁺; calcd. 873.4494, found 873.4499, [C₁₁₂H₁₈₅N₃₂O₃₈P + 4H]⁴⁺; calcd. 655.3389, found 655.3401, [C₁₁₂H₁₈₅N₃₂O₃₈P + 5H]⁵⁺; calcd. 524.4725, found 524.4721. ³¹P-NMR (162 MHz, D₂O); δ -4.28. ¹H NMR (600 MHz, D₂O) δ 7.69 (d, *J* = 8.1 Hz, 1H), 7.19 (d, *J* = 8.5 Hz, 2H), 7.12 (d, *J* = 8.3 Hz, 2H), 5.94 (ap. d, *J* = 4.6 Hz, 1H), 5.68 (ap. d, *J* = 8.1

Hz, 1H), 4.73 - 4.67 (m, 1H), 4.67 - 4.62 (m, 1H), 4.62 - 4.56 (m, 1H), 4.54 - 4.38 (m, 6H), 4.38 - 4.25 (m, 12H), 4.25 - 4.11 (m, 6H), 3.97 - 3.79 (m, 9H), 3.76 - 3.69 (m, 1H), 3.69 - 3.62 (m, 2H), 3.21 (ap. t, J = 7.0 Hz, 2H), 3.12 (ap. dd, J = 13.9, 6.9 Hz, 1H), 3.01 (ap. dd, J = 17.6, 7.9 Hz, 7H), 2.89 - 2.72 (m, 2H), 2.37 (ap. t, J = 7.7 Hz, 4H), 2.35 - 1.42 (m, 42H), 1.40 (ap. d, J = 7.2 Hz, 7H), 1.31 (ap. d, J = 7.2 Hz, 3H), 1.22 (ap. dd, J = 8.1, 6.6 Hz, 6H), 1.19 (ap. d, J = 6.4 Hz, 3H), 0.99 (ap. d, J = 6.7 Hz, 4H), 1.04 - 0.79 (m, 24H).

Coxsackie B3 Virus Woodruff strain VPg, (H-Gly-Ala-Tyr-Thr-Gly-Ile-Pro-Asn-Gln-Lys-Pro-Lys-Val-Pro-Thr-Leu-Arg-Gln-Ala-Lys-Val-Gln-OH), (2a)

HPLC gradient: 10 - 30 % B in 3 CV. The appropriate fractions were lyophilized, yielding 11.47 mg, 3.87 μ mol (16 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), Rt = 6.54 min. ESI-MS: m/z 1198.20 $[M+2H]^{2+}$, 799.13 $[M+3H]^{3+}$. HRMS $[C_{107}H_{180}N_{32}O_{30} + 2H]^{2+}$; calcd. 1197.6844, found 1197.6846, $[C_{107}H_{180}N_{32}O_{30} + 3H]^{3+}$; calcd. 798.7920, found 798.7920, $[C_{107}H_{180}N_{32}O_{30} + 4H]^{4+}$; calcd. 599.3458, found 599.3457, $[C_{107}H_{180}N_{32}O_{30} + 5H]^{5+}$; calcd. 479.6781, found 479.6780. ^1H NMR (600 MHz, D_2O) δ 7.16 (d, J = 8.5 Hz, 2H), 6.83 (d, J = 8.5 Hz, 2H), 4.67 (ap. t, J = 8.0 Hz, 1H), 4.65 - 4.60 (m, 1H), 4.57 (ap. dd, J = 9.1, 5.1 Hz, 1H), 4.47 (ap. dd, J = 16.1, 7.8 Hz, 3H), 4.42 (ap. dd, J = 7.9, 6.6 Hz, 2H), 4.40 - 4.24 (m, 11H), 4.24 - 4.19 (m, 1H), 4.19 - 4.14 (m, 1H), 4.13 (ap. d, J = 7.9 Hz, 1H), 3.99 - 3.81 (m, 7H), 3.77 - 3.68 (m, 2H), 3.68 - 3.59 (m, 1H), 3.21 (ap. t, J = 7.0 Hz, 2H), 3.10 - 2.97 (m, 8H), 2.85 - 2.77 (m, 2H), 2.43 - 2.23 (m, 9H), 2.20 - 1.41 (m, 46H), 1.39 (ap. d, J = 7.2 Hz, 5H), 1.33 (ap. d, J = 7.2 Hz, 3H), 1.24 (ap. d, J = 6.4 Hz, 3H), 1.17 (ap. d, J = 6.4 Hz, 4H), 1.00 (ap. d, J = 6.8 Hz, 3H), 1.01 - 0.92 (m, 18H), 0.88 (ap. t, J = 6.4 Hz, 6H).

Coxsackie B3 Virus Woodruff strain VPgU, (H-Gly-Ala-Tyr(pU)-Thr-Gly-Ile-Pro-Asn-Gln-Lys-Pro-Lys-Val-Pro-Thr-Leu-Arg-Gln-Ala-Lys-Val-Gln-OH), (2b)

HPLC gradient: 15 - 25 % B in 3 CV. The appropriate fractions were lyophilized, yielding 2.76 mg, 0.84 μ mol (3 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), Rt = 6.39 min. ESI-MS: m/z 1351.13 $[M+2H]^{2+}$, 900.93 $[M+3H]^{3+}$. HRMS $[C_{116}H_{191}N_{34}O_{38}P + 2H]^{2+}$; calcd. 1350.6970, found 1350.6971, $[C_{116}H_{191}N_{34}O_{38}P + 3H]^{3+}$; calcd. 900.8004, found 900.8004. ^{31}P -NMR (162 MHz, D_2O); δ - 4.27. ^1H NMR (600 MHz, D_2O) δ 7.69 (d, J = 8.1 Hz, 1H), 7.19 (d, J = 8.5 Hz, 2H), 7.11 (d, J = 8.2 Hz, 2H), 5.94 (d, J = 4.6 Hz, 1H), 5.68 (d, J = 8.1 Hz, 1H), 4.70 (ap. t, J = 7.6 Hz, 1H), 4.64 - 4.55 (m, 2H), 4.47 (ap. dd, J = 14.8, 7.8 Hz, 3H), 4.44 - 4.40 (m, 2H), 4.40 - 4.26 (m, 13H), 4.23 - 4.14 (m, 4H), 4.13 (ap. d, J = 7.9 Hz, 1H), 3.99 - 3.78 (m, 7H), 3.78 - 3.67 (m, 2H), 3.67 - 3.60 (m, 1H), 3.21 (ap. t, J = 7.0 Hz, 2H), 3.12 (ap. dd, J = 13.9, 6.9 Hz, 1H), 3.01 (ap. dd, J = 16.6, 7.6 Hz, 7H), 2.89 - 2.73 (m, 2H), 2.47 - 2.23 (m, 10H), 2.21 - 1.41 (m, 47H), 1.39 (ap. d, J = 7.2 Hz, 5H), 1.31 (ap. d, J = 7.2 Hz, 3H), 1.24 (ap. d, J = 6.4 Hz, 3H), 1.19 (ap. d, J = 6.5 Hz, 4H), 1.00 (ap. d, J = 6.7 Hz, 3H), 0.98 - 0.92 (m, 14H), 0.88 (ap. dd, J = 13.2, 6.1 Hz, 7H).

Human Enterovirus 71 VPg, (H-Gly-Ala-Tyr-Ser-Gly-Ala-Pro-Lys-Gln-Val-Leu-Lys-Lys-Pro-Ala-Leu-Arg-Thr-Ala-Thr-Val-Gln-OH), (3a)

HPLC gradient: 10 - 30 % B in 3 CV. The appropriate fractions were lyophilized, yielding 10.89 mg, 3.98 μ mol (16 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), Rt = 6.36 min. ESI-MS: m/z 1143.13 $[M+2H]^{2+}$, 762.40 $[M+3H]^{3+}$. HRMS $[C_{102}H_{174}N_{30}O_{29} + 2H]^{2+}$; calcd. 1142.6604, found 1142.6608, $[C_{102}H_{174}N_{30}O_{29} + 3H]^{3+}$; calcd. 762.1093, found 762.1093. ^1H NMR (600 MHz, D_2O) δ 7.16 (d, J = 8.5 Hz, 2H), 6.84 (d, J = 8.5 Hz, 2H), 4.67 - 4.56 (m, 3H), 4.49 - 4.29 (m, 13H), 4.27 - 4.23 (m, 3H), 4.21 - 4.13 (m, 2H), 4.09 (ap. d, J = 8.3 Hz, 1H), 3.90 (ap. d, J = 5.0 Hz, 2H), 3.87 - 3.81 (m, 5H), 3.78 (ap. dd, J = 11.6, 5.2 Hz, 1H), 3.65 (ap. dd, J = 17.0, 7.0 Hz, 2H), 3.23 (ap. t, J = 6.9 Hz, 2H), 3.13 - 2.91 (m, 8H), 2.42 - 2.26 (m, 6H), 2.23 - 1.45 (m, 40H), 1.43 (ap. d, J = 7.2

Hz, 5H), 1.37 (ap. dd, $J = 13.1, 7.1$ Hz, 7H), 1.33 (ap. d, $J = 7.2$ Hz, 3H), 1.22 (ap. dd, $J = 6.3, 4.8$ Hz, 6H), 1.05 - 0.91 (m, 19H), 0.89 (ap. dd, $J = 11.0, 6.1$ Hz, 7H).

Human Enterovirus 71 VPgpU, (H-Gly-Ala-Tyr(pU)-Ser-Gly-Ala-Pro-Lys-Gln-Val-Leu-Lys-Lys-Pro-Ala-Leu-Arg-Thr-Ala-Thr-Val-Gln-OH), (3b)

HPLC gradient: 10 - 30 % B in 3 CV. The appropriate fractions were lyophilized, yielding 2.35 mg, 0.67 μmol (3 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), $R_t = 6.24$ min. ESI-MS: m/z 1296.07 $[\text{M}+2\text{H}]^{2+}$, 864.47 $[\text{M}+3\text{H}]^{3+}$. HRMS $[\text{C}_{111}\text{H}_{185}\text{N}_{32}\text{O}_{37}\text{P} + 2\text{H}]^{2+}$; calcd. 1295.6730, found 1295.6745, $[\text{C}_{111}\text{H}_{185}\text{N}_{32}\text{O}_{37}\text{P} + 3\text{H}]^{3+}$; calcd. 864.1178, found 864.1173. ^{31}P -NMR (162 MHz, D_2O); $\delta - 4.29$. ^1H NMR (600 MHz, D_2O) δ 7.69 (d, $J = 8.1$ Hz, 1H), 7.19 (d, $J = 8.6$ Hz, 2H), 7.12 (d, $J = 8.3$ Hz, 2H), 5.94 (d, $J = 4.8$ Hz, 1H), 5.67 (d, $J = 8.1$ Hz, 1H), 4.65 (ap. t, $J = 7.6$ Hz, 1H), 4.60 (ap. dt, $J = 9.6, 6.1$ Hz, 2H), 4.45 - 4.22 (m, 18H), 4.22 - 4.12 (m, 4H), 4.09 (ap. d, $J = 8.2$ Hz, 1H), 3.99 - 3.88 (m, 2H), 3.88 - 3.76 (m, 5H), 3.68 - 3.62 (m, 2H), 3.23 (ap. t, $J = 6.9$ Hz, 2H), 3.10 (ap. dd, $J = 13.9, 7.1$ Hz, 1H), 3.07 - 2.92 (m, 7H), 2.41 - 2.21 (m, 7H), 2.19 - 1.46 (m, 41H), 1.43 (ap. d, $J = 7.2$ Hz, 6H), 1.37 (ap. dd, $J = 13.7, 7.1$ Hz, 6H), 1.31 (ap., $J = 7.2$ Hz, 3H), 1.22 (ap. dd, $J = 6.3, 4.3$ Hz, 7H), 1.01 - 0.90 (m, 19H), 0.88 (ap. dd, $J = 11.1, 6.1$ Hz, 6H).

Consensus Human Enterovirus VPg, (H-Gly-Ala-Tyr-Thr-Gly-Leu-Pro-Asn-Gln-Lys-Pro-Lys-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Gln-OH), (4a)

HPLC gradient: 10 - 35 % B in 3 CV. The appropriate fractions were lyophilized, yielding 11.25 mg, 3.83 μmol (15 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), $R_t = 6.56$ min. ESI-MS: m/z 1184.07 $[\text{M}+2\text{H}]^{2+}$, 790.33 $[\text{M}+3\text{H}]^{3+}$. HRMS $[\text{C}_{106}\text{H}_{179}\text{N}_{31}\text{O}_{30} + 2\text{H}]^{2+}$; calcd. 1184.1789, found 1184.1789, $[\text{C}_{106}\text{H}_{179}\text{N}_{31}\text{O}_{30} + 3\text{H}]^{3+}$; calcd. 789.7884, found 789.7887, $[\text{C}_{106}\text{H}_{179}\text{N}_{31}\text{O}_{30} + 4\text{H}]^{4+}$; calcd. 592.5931, found 592.5937, $[\text{C}_{106}\text{H}_{179}\text{N}_{31}\text{O}_{30} + 5\text{H}]^{5+}$; calcd. 474.2759, found 474.2766. ^1H NMR (600 MHz, D_2O) δ 7.16 (d, $J = 8.6$ Hz, 1H), 6.83 (d, $J = 8.6$ Hz, 1H), 4.67 (t, $J = 8.0$ Hz, 1H), 4.65 - 4.59 (m, 1H), 4.57 (dd, $J = 9.1, 5.2$ Hz, 1H), 4.53 - 4.39 (m, 3H), 4.38 - 4.27 (m, 5H), 4.25 - 4.18 (m, 2H), 4.18 - 4.11 (m, 1H), 3.95 - 3.80 (m, 3H), 3.80 - 3.58 (m, 2H), 3.21 (ap. t, $J = 7.0$ Hz, 1H), 3.09 - 2.96 (m, 4H), 2.89 - 2.75 (m, 1H), 2.35 (ap. t, $J = 7.8$ Hz, 3H), 2.33 - 2.24 (m, 1H), 2.20 - 1.95 (m, 7H), 1.40 (ap. d, $J = 7.2$ Hz, 2H), 1.33 (ap. d, $J = 7.2$ Hz, 2H), 1.22 (ap. t, $J = 5.9$ Hz, 3H), 1.17 (ap. d, $J = 6.4$ Hz, 2H), 1.00 (ap. d, $J = 6.8$ Hz, 2H), 0.98 - 0.93 (m, 8H), 0.91 (ap. d, $J = 6.8$ Hz, 2H), 0.87 (ap. t, $J = 7.4$ Hz, 2H).

Consensus Human Enterovirus VPgpU, (H-Gly-Ala-Tyr(pU)-Thr-Gly-Leu-Pro-Asn-Gln-Lys-Pro-Lys-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Gln-OH), (4b)

HPLC gradient: 10 - 30 % B in 3 CV. The appropriate fractions were lyophilized, yielding 8.2 mg, 2.43 μmol (10 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), $R_t = 6.44$ min. ESI-MS: m/z 1338.13 $[\text{M}+2\text{H}]^{2+}$, 892.13 $[\text{M}+3\text{H}]^{3+}$. HRMS $[\text{C}_{115}\text{H}_{190}\text{N}_{33}\text{O}_{38}\text{P} + 2\text{H}]^{2+}$; calcd. 1337.1916, found 1337.1926, $[\text{C}_{115}\text{H}_{190}\text{N}_{33}\text{O}_{38}\text{P} + 3\text{H}]^{3+}$; calcd. 891.7968, found 891.7974, $[\text{C}_{115}\text{H}_{190}\text{N}_{33}\text{O}_{38}\text{P} + 4\text{H}]^{4+}$; calcd. 669.0994, found 669.0998. ^{31}P -NMR (162 MHz, D_2O); $\delta - 4.27$. ^1H NMR (600 MHz, D_2O) δ 7.69 (d, $J = 8.1$ Hz, 1H), 7.19 (d, $J = 8.5$ Hz, 2H), 7.12 (d, $J = 8.3$ Hz, 2H), 5.94 (d, $J = 4.6$ Hz, 1H), 5.68 (d, $J = 8.1$ Hz, 1H), 4.74 - 4.68 (m, 1H), 4.67 - 4.56 (m, 2H), 4.51 - 4.39 (m, 5H), 4.39 - 4.26 (m, 12H), 4.26 - 4.16 (m, 5H), 4.16 - 4.10 (m, 2H), 3.93 (s, 2H), 3.92 - 3.77 (m, 5H), 3.76 - 3.70 (m, 1H), 3.69 - 3.59 (m, 2H), 3.21 (ap. t, $J = 7.0$ Hz, 2H), 3.12 (ap. dd, $J = 14.0, 6.8$ Hz, 1H), 3.01 (ap. dd, $J = 15.8, 7.9$ Hz, 7H), 2.89 - 2.75 (m, 2H), 2.44 - 2.24 (m, 7H), 2.19 - 1.41 (m, 44H), 1.40 (ap. d, $J = 7.2$ Hz, 4H), 1.31 (ap. d, $J = 7.2$ Hz, 3H), 1.22 (ap. t, $J = 6.0$ Hz, 6H), 1.19 (ap. d, $J = 6.4$ Hz, 4H), 1.00 (ap. d, $J = 6.8$ Hz, 3H), 0.98 - 0.89 (m, 18H), 0.87 (ap. t, $J = 7.4$ Hz, 4H).

Poliovirus VPg, (H-Gly-Ala-Tyr-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Gln-OH), (5a)

HPLC gradient: 10 - 30 % B in 3 CV. The appropriate fractions were lyophilized, yielding 48.57 mg, 17.29 μmol (35 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), R_t = 6.48 min. ESI-MS: m/z 1178.13 $[\text{M}+2\text{H}]^{2+}$, 785.47 $[\text{M}+3\text{H}]^{3+}$. HRMS $[\text{C}_{105}\text{H}_{177}\text{N}_{31}\text{O}_{30} + 2\text{H}]^{2+}$; calcd. 1177.1711, found 1177.1690, $[\text{C}_{105}\text{H}_{177}\text{N}_{31}\text{O}_{30} + 3\text{H}]^{3+}$; calcd. 785.1165, found 785.1156, $[\text{C}_{105}\text{H}_{177}\text{N}_{31}\text{O}_{30} + 4\text{H}]^{4+}$; calcd. 589.0892, found 589.0900, $[\text{C}_{105}\text{H}_{177}\text{N}_{31}\text{O}_{30} + 5\text{H}]^{5+}$; calcd. 471.4728, found 471.4736. ^1H NMR (600 MHz, D_2O) δ 7.16 (d, J = 8.5 Hz, 2H), 6.83 (d, J = 8.5 Hz, 2H), 4.74 - 4.60 (m, 3H), 4.58 (ap. dd, J = 8.7, 5.4 Hz, 1H), 4.52 - 4.45 (m, 2H), 4.45 - 4.40 (m, 3H), 4.40 - 4.27 (m, 8H), 4.26 - 4.19 (m, 3H), 4.19 - 4.10 (m, 2H), 3.95 - 3.78 (m, 7H), 3.78 - 3.70 (m, 1H), 3.70 - 3.61 (m, 2H), 3.21 (ap. t, J = 7.0 Hz, 2H), 3.14 - 2.94 (m, 8H), 2.92 - 2.68 (m, 4H), 2.40 - 2.23 (m, 5H), 2.20 - 1.43 (m, 40H), 1.40 (ap. d, J = 7.2 Hz, 7H), 1.33 (ap. d, J = 7.2 Hz, 3H), 1.27 - 1.19 (m, 7H), 1.17 (ap. d, J = 6.4 Hz, 3H), 1.00 - 0.89 (m, 21H), 0.87 (ap. t, J = 7.4 Hz, 3H).

Poliovirus VPgpU, (H-Gly-Ala-Tyr(pU)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Gln-OH), (5b)

HPLC gradient: 10 - 50 % B in 3 CV. The appropriate fractions were lyophilized, yielding 13.29 mg, 4.26 μmol (17 % based on original loading of resin) of the nucleopeptide as TFA salt. LCMS (0 - 50 % B in 15 min.), R_t = 6.38 min. ESI-MS: m/z 1330.07 $[\text{M}+2\text{H}]^{2+}$, 887.40 $[\text{M}+3\text{H}]^{3+}$. HRMS $[\text{C}_{114}\text{H}_{189}\text{N}_{33}\text{O}_{38}\text{P} + 2\text{H}]^{2+}$; calcd. 1330.1838, found 1330.1826, $[\text{C}_{114}\text{H}_{189}\text{N}_{33}\text{O}_{38}\text{P} + 3\text{H}]^{3+}$; calcd. 887.1255, found 887.1248, $[\text{C}_{114}\text{H}_{189}\text{N}_{33}\text{O}_{38}\text{P} + 4\text{H}]^{4+}$; calcd. 665.5955, found 665.5960, $[\text{C}_{114}\text{H}_{189}\text{N}_{33}\text{O}_{38}\text{P} + 5\text{H}]^{5+}$; calcd. 532.6778, found 532.6768. ^{31}P -NMR (162 MHz, D_2O); δ - 4.28. ^1H NMR (600 MHz, D_2O) δ 7.69 (d, J = 8.1 Hz, 1H), 7.19 (d, J = 8.5 Hz, 2H), 7.12 (d, J = 8.3 Hz, 2H), 5.94 (d, J = 4.6 Hz, 1H), 5.68 (d, J = 8.1 Hz, 1H), 4.73 - 4.66 (m, 1H), 4.66 - 4.61 (m, 2H), 4.60 (ap. dd, J = 8.8, 5.3 Hz, 1H), 4.52 - 4.39 (m, 5H), 4.39 - 4.26 (m, 12H), 4.25 - 4.09 (m, 6H), 3.98 - 3.80 (m, 7H), 3.77 - 3.69 (m, 1H), 3.69 - 3.63 (m, 2H), 3.21 (ap. t, J = 7.1 Hz, 2H), 3.12 (ap. dd, J = 13.9, 6.9 Hz, 1H), 3.05 - 2.93 (m, 7H), 2.89 - 2.69 (m, 4H), 2.37 (ap. t, J = 7.6 Hz, 4H), 2.36 - 2.24 (m, 3H), 2.24 - 2.16 (m, 1H), 2.14 - 1.97 (m, 9H), 1.97 - 1.43 (m, 29H), 1.40 (ap. d, J = 7.2 Hz, 4H), 1.31 (ap. d, J = 7.2 Hz, 3H), 1.27 - 1.20 (m, 7H), 1.19 (ap. d, J = 6.4 Hz, 3H), 1.01 - 0.95 (m, 12H), 0.92 (ap. dd, J = 12.2, 6.6 Hz, 9H), 0.87 (ap. t, J = 7.4 Hz, 4H).

Notes and References

- ¹ Steil, B. P.; Barton, D. J.; *Virus Res.*, **2009**, 139, 240 - 252.
- ² Arden, K. E.; Mackay, I. M.; *Rev. Med. Virol.*, **2010**, 20, 156 - 176.
- ³ Ajmera, V.; Xia, G.; Vaughan, G.; Forbi, J. C.; Ganova-Raeva, L. M.; Khudyakov, Y.; Opio, C. K.; Taylor, R.; Restrepo, R.; Munoz, S.; Fontana, R. J.; Lee, W. M.; *J. Viral Hepatitis*, **2011**, 18, e167 - e174.
- ⁴ Nathanson, N.; *Advances in Virus Research*, **2008**, 71, 1 - 50.
- ⁵ Longjam, N.; Deb, R.; Sarmah, A. K.; Tayo, T.; Awachat, V. B.; Saxena, V. K.; *Veterinary Medicine International*, **2011**, 905768.
- ⁶ Doi, K.; *J. Comp. Path.*, **2011**, 144, 25 - 40.
- ⁷ Buckwalter, M. R.; Nga, P. T.; Gouilh, M. A.; Fiette, L.; Bureau, J.; Laird, M. E.; Buchrieser, J.; Ozden, S.; Cheval, J.; Eloit, M.; Manuguerra, J.; Gessain, A.; Brey, P. T.; Fontanet, A.; Albert, M. L.; *J. Virol.*, **2011**, 85, 6893 - 6905.
- ⁸ Rothberg, P. G.; Harris, T. J. R.; Nomoto, A.; Wimmer, E.; *Proc. Natl. Acad. Sci. USA*, **1978**, 75, 4868 - 2872.
- ⁹ Ambros, V.; Balimore, D.; *J. Biol. Chem.*, **1978**, 253, 5263 - 5266.
- ¹⁰ Schein, C. H.; Volk, D. E.; Oezguen, N.; Paul, A.; *Proteins*, **2006**, 63, 719 - 726.
- ¹¹ Novak, J. E.; Kirkegaard, K.; *J. Virol.*, **1991**, 65, 3384 - 3387.
- ¹² Steil, B. P.; Barton, D. J.; *J. Virol.*, **2009**, 83, 12660 - 12670.
- ¹³ Paul, A. V.; Rieder, E.; Kim, D. W.; van Boom, J. H.; Wimmer, E.; *J. Virol.*, **2000**, 74, 10359 - 10370.
- ¹⁴ Rieder, E.; Paul, A. V.; Kim, D. W.; van Boom, J. H.; Wimmer, E.; *J. Virol.*, **2000**, 74, 10371 - 10380.
- ¹⁵ Murray, K. E.; Barton, D. J.; *J. Virol.*, **2003**, 77, 4739 - 4750.
- ¹⁶ Morasco, B. J.; Nidhi, S.; Parilla, J.; Flanagan, J. B.; *J. Virol.*, **2003**, 77, 5136 - 5144.
- ¹⁷ Cottam, E. M.; Haydon, D. T.; Paton, D. J.; Gloster, J.; Wilesmith, J. W.; Ferris, N. P.; Hutchings, G. H.; King, D. P.; *J. Virol.*, **2006**, 80, 11274 - 11282.
- ¹⁸ Schein, C. H.; Oezguen, N.; van der Heden van Noort, G. J.; Filippov, D. V.; Paul, A.; Kumar, E.; Braun, W.; *Peptides*, **2010**, 31, 1441 - 1448.
- ¹⁹ Paul, A. V.; van Boom, J. H.; Filippov, D. V.; Wimmer, E.; *Nature*, **1998**, 393, 280 - 284.
- ²⁰ Filippov, D. V.; Kuyl-Yeheskiely, E.; van der Marel, G. A.; Tesser, G. I.; van Boom, J. H.; *Tetrahedron Lett.*, **1998**, 39, 3597 - 3600.
- ²¹ Kriek, N. M. A. J.; Filippov, D. V.; van den Elst, H.; Meeuwenoord, N. J.; Tesser, G. I.; van Boom, J. H.; van der Marel, G. A.; *Tetrahedron*, **2003**, 59, 1589 - 1597.
- ²² The consensus HEV VPg was designed using software that used 33 aligned HEV-VPg's to assign the conserved physicochemical properties and on the basis of these findings model a consensus sequence.

Schein, C. H.; *Proceedings of the Biotechnology and Bio-informatics Symposium (BIOT-2011)*, October 2011, Rice University, Houston (Tx), United States of America.

²³ Dangles, O.; Guibe, F.; Balavoine, G.; Lavielle, S.; Marquet, A.; *J. Org. Chem.*, **1987**, 52, 4984 - 4993.

²⁴ Preliminary results indicate that all synthetically prepared VPg's, including the consensus HEV protein, were recognized as substrate by viral 3D polymerases based on their ability to be uridylylated. Poliovirus and coxsackie A virus polymerases showed a preference for their own VPg with respect to other viral strand VPg's, but the consensus VPg was found to be uridylylated as well or even better as the wild type VPg.

Chapter 6

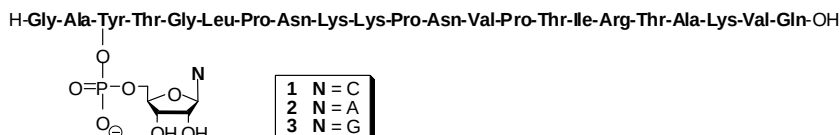
Synthesis of Nucleotidylated Poliovirus VPg Proteins¹

Introduction

Nucleotidylation of proteins is a widespread post-translational modification.²⁻⁵ For example, adenylation of tyrosine has been recognized as an important regulatory event, both in bacteria and in higher organisms.^{6,7} A number of RNA viruses use the uridylylated or guanylylated form of VPg (Viral Protein genome-linked) proteins to initiate the replication of the genome.⁸ Perhaps the most investigated protein nucleotidylation process takes place in *Picornaviridae* as described in **Chapter 5**.⁹ These viruses share the presence of a small VPg, that is covalently linked to the 5'-end of the viral genome.¹⁰ Although the mechanism of replication of picorna viruses has not been clarified completely, extensive research has shown that VPg and/or its uridylylated forms VPgpU or VPgpUpU can act as primers in both negative and positive strand RNA synthesis.¹¹⁻¹⁵ VPgpU and VPgpUpU, in turn, are presumably formed by uridylylation at the O-4 of tyrosine-3 of the 22 amino acid long VPg protein of poliovirus under influence of a cis-acting replication element (CRE) within the viral RNA.¹⁶⁻²⁰ It has been shown that the introduction of point mutations in this CRE motif led to 3D polymerase mediated reaction of CTP, ATP or GTP with O-4 tyrosine of VPg to produce VPgpC, VPgpA and VPgpG respectively.¹⁶ Although all four nucleosides can be incorporated, suggesting that the 3D polymerase is not specific for UTP as incoming nucleotide donor, VPgpG and VpgpA are not elongated towards the dinucleotide and VPgpC only in moderate amount when compared to VPgpU, resulting in a lower virulence.

As part of a program to develop synthetic routes towards nucleotidylated proteins and derivatives thereof a solid phase synthesis to obtain nucleotidylated VPg proteins, containing all four naturally occurring ribonucleotides was undertaken. The approach presented in **Chapter 5** to synthesize a pre-uridylylated tyrosine building block is adapted here towards the preparation of adenylylated, guanylylated and cytidylylated Fmoc-tyrosine building blocks that were subsequently used in the synthesis of Poliovirus VPgpA, VPgpC and VPgpG (see **Figure 1**).

Figure 1 Poliovirus VPgpC, VPgpA and VPgpG



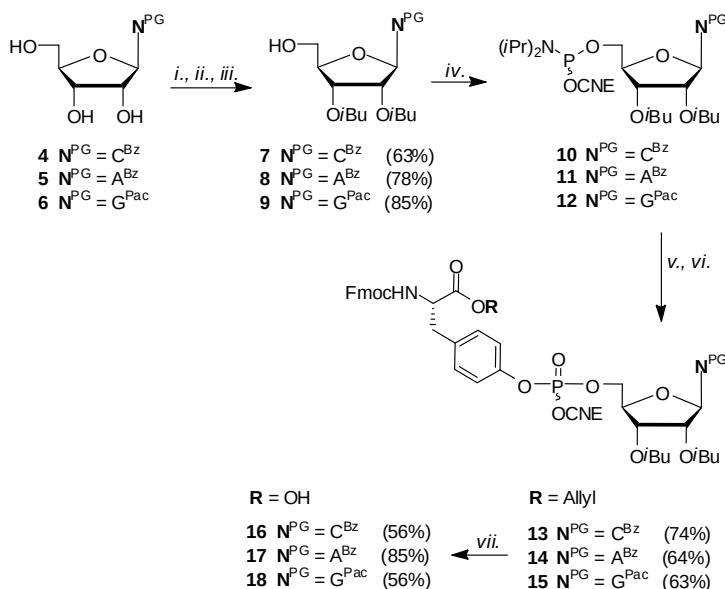
Results and Discussion

A solid phase synthesis using pre-nucleotidylated Fmoc-tyrosine building blocks, as was described for VPgpU in **Chapter 5**, is considered to be most convenient for the synthesis of poliovirus derived VPgpC, VPgpA and VPgpG.²¹ Key to the success of such a strategy is the choice of the protecting groups for the exocyclic amino functions of the nucleobases. Such protecting groups should be stable during the solid phase peptide synthesis and leave the peptide intact during deprotection at the end of the synthesis. Acyl protective groups are an obvious choice as these are well established in the field of nucleopeptide synthesis.^{22,23} Because the stability of acyl protecting groups depends on the nature of the nucleobase, the benzoyl group was selected to protect the exocyclic amino functions of both cytidine and adenosine while guanosine was equipped with the more labile phenoxyacetyl group.²⁴

The syntheses of protected nucleotidylated tyrosine building blocks **16-18** are shown in **Scheme 1**. Known *N*⁴-benzoylcytidine **4**²⁵, *N*⁶-benzoyladenine **5**²⁶ and *N*²-phenoxyacetylguanosine **6**²⁷ were selectively silylated with TBDMSCl in pyridine, treated with isobutyric anhydride and subsequently desilylated with acid to give partially protected nucleosides **7**, **8** and **9**, respectively.

Treatment of **7**, **8** and **9** with 2-cyanoethoxy-*N,N*-diisopropylaminochlorophosphine in the presence of triethylamine as base led to 5'-phosphoramidites **10**, **11** and **12**. Tetrazole mediated phosphitylation of the free hydroxyl function on the side chain of Fmoc-Tyr-OAll with amidites **10**, **11** and **12** and oxidation of the intermediate phosphite triesters with *m*-CPBA was followed by liberation of the acid by palladium catalyzed removal of the allyl protective group to give the respective nucleotidylated amino acid building blocks **16**, **17** and **18**.

Scheme 1. Synthesis of Fmoc-tyrosine derivatives

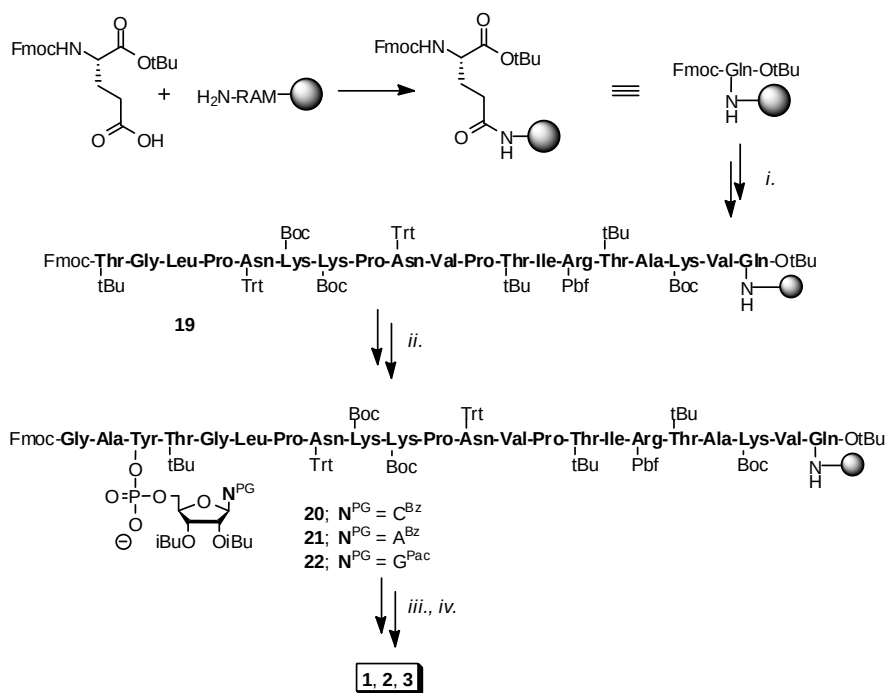


Reagents and conditions: *i.* TBDMSCl, pyr., *ii.* iBu_2O , pyr., *iii.* pTsOH, *iv.* 2-cyanoethoxy- N,N -diisopropylaminochlorophosphine, TEA, DCM, 45 min, *v.* Fmoc-Tyr-OAll, tetrazole, MeCN/DCM, 45 min, r.t., *vi.* mCPBA, 0 °C, *vii.* $Pd(PPh_3)_4$, $SnBu_3H$, THF/DCM, AcOH, r.t., 2h.

The solid phase synthesis of target compounds **1** - **3** started with the synthesis of the protected and immobilized 19-mer peptide **19**, a common intermediate en route to VPgpC, VPgpA and VPgpG (**Scheme 2**). Tentagel S RAM resin was functionalized with the first amino acid by attachment of Fmoc-Glu-OtBu via the side chain acid moiety.²² Automated solid phase peptide synthesis based on Fmoc chemistry, using commercially available amino acid building blocks and HCTU as condensing agent proceeded uneventfully.

Immobilized 19-mer **19** was manually elongated by the sequential incorporation of the nucleotidylated tyrosine **16**, Fmoc protected Ala and Gly using BOP/HOBt to give the fully protected VPgpC (**20**, **Scheme 2**). TFA mediated cleavage of the nucleopeptide from the resin and concomitant unmasking of the amino acid side chains, deprotection of exocyclic amine of cytosine with 25% aq. ammonia and final purification with RP-HPLC afforded homogeneous VPgpC **1** in 45% yield based on initial loading of the resin.

Scheme 2. Solid Phase Synthesis of VPgpC, VPgpA and VPgpG



Reagents and conditions: *i.* SPPS using a repetitive cycle of a) 20% piperidine/DMF, b) Fmoc-AA-OH, HCTU, DiPEA, NMP, c) Ac_2O , DiPEA, HOBt, NMP, *ii.* SPPS using a repetitive cycle of a) 2% DBU/DMF, b) Fmoc-AA-OH, BOP/HOBt or HATU, DiPEA, NMP, c) Ac_2O , DiPEA, HOBt, NMP, *iii.* TFA/TIS/ H_2O , 2.5 h, r.t., *iv.* 25% aq. ammonia/dioxane, 4 h, r.t.

A similar sequence of reactions as described for the formation of **20**, using in this case adenylylated tyrosine **17** gave immobilized partially protected VPgpA **21**. In the first instance, cleavage from the solid support, removal of all protecting groups and purification did not result in homogeneous VPgpA. It turned out that despite the relatively high acid stability of the glycosidic linkages of ribonucleosides, TFA treatment of fully protected VPgpA **21** was accompanied by depurination of the adenosine moiety.²⁹ This was unexpected because the 5'-phosphorylated adenosine residue was reported to be stable to TFA treatment³⁰ and 5'-adenylyl O-tyrosine, specifically, is known to withstand a treatment with strong acid (0.3 N HCl, 37 °C, 3h).³¹ Homogeneous VPgpA **2** could be obtained in 10% yield based on initial loading of the resin by acidolysis, RP-HPLC purification of the intermediate nucleopeptide **21**, ammonia treatment of the partially protected nucleopeptide and finally a second RP-HPLC purification.

Adoption of the procedure of VPgpC for the synthesis of VPgpG using guanylated tyrosine **18** and LCMS analysis of the crude product showed the presence of truncated sequences while depurination products were lacking. Guided by the work of Bae and Lakshman^{32,33}, who showed that BOP and PyBOP mediated condensations may lead to side reactions at the C-6 of the guanine base, the last three amino acids in the synthesis of VPgpG were incorporated using HATU as coupling reagent.³⁴ Subsequently, acid mediated cleavage from the solid support and

deprotection using TFA followed by a 25% aq. ammonia treatment and RP-HPLC purification, afforded VPgpG in 29% based on initial loading of the resin.

Conclusion

In conclusion, three unprecedented nucleotidylated amino acid building blocks were prepared and applied in the solid phase synthesis of poliovirus derived proteins VPgpC, VPgpA and VPgpG. This methodology opens the way to the synthesis of more elaborated nucleotidylated proteins such as VPgpG of Norovirus.³⁵ The synthetic, well defined constructs of this type may prove to be invaluable in the structural and biological studies of viral replication.

Experimental

General: All chemicals were used as received. MeCN and THF were stored over activated 4Å molecular sieves for at least 24 hours prior to use. DCM was distilled over CaH₂ prior to use. Solvents were removed by rotary evaporation under reduced pressure at a temperature not exceeding 40°C. TLC analysis, silica gel column chromatography, NMR, LCMS, HPLC and HRMS techniques were used as described in Chapter 2.

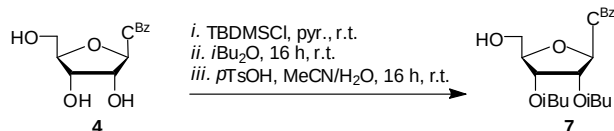
Procedure for solid phase peptide synthesis: The protected peptide **19** was prepared via solid phase peptide synthesis on an ABI-433A (Applied Biosystems, division of Perkin-Elmer) automated peptide synthesizer using the Fmoc based peptide synthesis protocol starting from Tentagel S RAM resin (200 mg, 50 µmol) to which Fmoc-Glu-OtBu was attached via the side-chain carboxylic acid functionality. As repetitive cycle were used:

- Fmoc cleavage using 20% piperidine in NMP.
- Coupling of the appropriate amino acid applying a five-fold excess, activation by 5 eq. HCTU in NMP (0.25M) and 12.5 eq. DiPEA in NMP (1.25 M) for 1 hour.
- Capping using 0.5 M acetic anhydride, 0.125 M DiPEA and 0.015 M HOBt in NMP for 1 min.

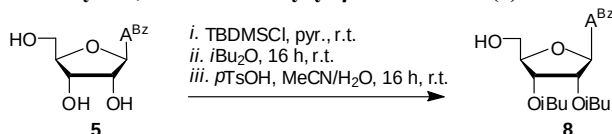
The last three amino acids were introduced manually on 25 µmol **19** by a double coupling protocol:

- Fmoc cleavage using 2% DBU in DMF.
- Coupling of the appropriate amino acid applying a five-fold excess, activation by 5 eq. BOP/HOBt (in the preparation of **1** and **2**) or HATU (in the preparation of **3**) in NMP and 12.5 eq. DiPEA in NMP (1.25 M) for 1 hour.
- Capping using 0.5 M acetic anhydride, 0.125 M DiPEA and 0.015 M HOBt in NMP for 1 min.

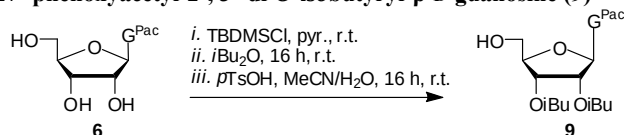
N⁴-Benzoyl-2', 3'-di-*O*-isobutyryl-β-D-cytidine (**7**)



Cytidine **4** (1.0 g, 2.9 mmol) was dissolved in pyridine (5 mL) after coevaporation with pyridine. TBDMSCl (0.87 g, 5.8 mmol) was added and the reaction mixture was stirred until it became a clear solution. After this, isobutyric anhydride (11.6 mmol, 1.93 mL) was added and the reaction mixture was stirred at room temperature for 16 hours, upon which TLC analysis (10% MeOH/DCM) indicated complete conversion of the starting material. The reaction mixture was concentrated and taken up in EtOAc. After washing with sat. aq. NaHCO₃, 5% citric acid and H₂O the organic phase was dried (MgSO₄) and concentrated. The residue was taken up in 4:1 v/v MeCN/ H₂O (30 mL), *p*-TsOH (0.55 g, 2.9 mmol) was added and the reaction mixture was stirred overnight at room temperature. The reaction mixture was concentrated and purified using silica gel column chromatography with a gradient of DCM/MeOH (100/0 – 98/2) to afford the title compound as a white solid (0.88 g, 1.8 mmol, 63 %). ¹H-NMR (400 MHz, CDCl₃) δ 8.36 (d, *J* = 7.5 Hz, 1H, H6), 7.88 (d, *J* = 7.5 Hz, 2H, Arom. Bz), 7.59 - 7.56 (m, 2H, H5, Arom Bz), 7.47 (t, *J* = 7.7 Hz, 2H, Arom. Bz), 6.17 (d, *J* = 5.1 Hz, 1H, H1'), 5.62 (t, *J* = 5.3 Hz, 1H, H2'), 5.58 - 5.51 (m, 1H, H3'), 4.38 - 4.17 (m, 1H, H4'), 4.01 (dd, *J* = 12.4, 1.9 Hz, 1H, H5_A'), 3.87 (dd, *J* = 12.4, 1.9 Hz, 1H, H5_B'), 2.74 - 2.42 (m, 2H, CH *i*Bu), 1.21 - 1.08 (m, 12H, CH₃ *i*Bu). ¹³C-NMR (100 MHz, CDCl₃); δ 175.6, 175.3 (CO *i*Bu), 167.0 (C2), 162.94 (CO Bz), 155.8 (C6), 145.0 (C4), 132.7 (Cq Bz), 132.6, 128.3, 127.6 (Bz), 97.6 (C5), 88.0 (C1'), 83.6 (C4'), 74.0 (C3'), 70.0 (C2'), 60.6 (C5'), 33.4, 33.3 (CH *i*Bu), 18.5, 18.3, 18.28, 18.2 (CH₃ *i*Bu). IR: 1743, 1653, 1624, 1558, 1481, 1311, 1244, 1099. LCMS (10-90% B in 15 min), Rt = 6.90. ESI-MS: *m/z* 487.9 [M]⁺. HRMS [C₂₄H₂₉N₃O₈ + H]⁺; calcd. 488.2027, found 488.2026.

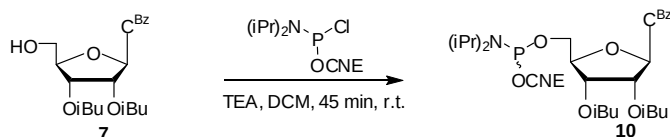
N⁶-Benzoyl-2', 3'-di-O-isobutyryl-β-D-adenosine (8)

Adenosine **5** (1.8 g, 4.8 mmol) was dissolved in pyridine (25 mL) after coevaporation with pyridine. TBDMSCl (1.5 g, 10 mmol) was added and the reaction mixture was stirred for 1 hour at room temperature. After this, isobutyric anhydride (20 mmol, 3.7 mL) was added and the reaction mixture was stirred for 48 hours. The reaction mixture was concentrated and taken up in DCM. After washing with 5% citric acid and H₂O the organic phase was dried (MgSO₄) and concentrated. The residue was taken up in 4:1 v/v MeCN/ H₂O (50 mL) and *p*-TsOH (0.8 g, 5 mmol) was added after which the reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was concentrated and purified with silica gel column chromatography using a gradient of EtOAc/PE (0/100 – 45/55) to afford the title compound as a white foam (1.9 g, 3.9 mmol, 78%). ¹H-NMR (400 MHz, CDCl₃) δ 9.82 (s, 1H, NH), 8.66 (s, 1H, H8), 8.31 (s, 1H, H2), 8.01 (d, *J* = 7.6 Hz, 2H, Arom. Bz), 7.53 (t, *J* = 7.3 Hz, 1H, Arom. Bz), 7.43 (t, *J* = 7.5 Hz, 2H, Arom. Bz), 6.21 (d, *J* = 7.1 Hz, 1H, H1'), 6.04 - 5.92 (m, 1H, H2'), 5.91 (s, 1H, 5'-OH), 5.69 (d, *J* = 5.2 Hz, 1H, H3'), 4.31 (s, 1H, H4'), 4.09 - 3.73 (m, 2H, H5'), 2.75 - 2.56 (m, 1H, CH *i*Bu), 2.56 - 2.38 (m, 1H, CH *i*Bu), 1.21 - 1.07 (m, 12H, CH₃ *i*Bu). ¹³C-NMR (100 MHz, CDCl₃) δ 175.6, 175.1 (CO *i*Bu), 164.6 (CO Bz), 152.3 (C2), 150.9, 150.2 (C4, C6), 142.2 (C8), 133.3 (Cq Bz), 132.8 (Arom. Bz), 128.8 (Arom. Bz), 127.9 (Arom. Bz), 124.4 (C5), 88.4 (C1'), 86.3 (C4'), 72.8 (C2'), 72.2 (C3'), 62.4 (C5'), 33.8, 33.5 (CH *i*Bu), 18.89, 18.75, 18.67, 18.59 (CH₃ *i*Bu). IR: 3274, 2977, 1743, 1701, 1609, 1582, 1454, 1246, 1153, 1095. LCMS (10-90% B in 15 min), Rt = 6.57. ESI-MS: *m/z* 512.2 [M + H]⁺. HRMS [C₂₅H₂₉N₅O₇ + H]⁺; calcd. 512.2135, found 512.2139.

N²-phenoxyacetyl-2', 3'-di-O-isobutyryl-β-D-guanosine (9)

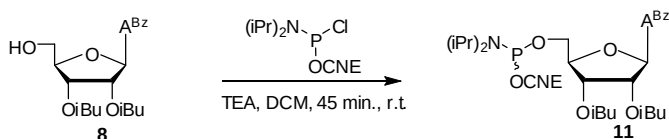
Guanosine **6** (4.17 g, 10 mmol) was dissolved in pyridine (50 mL) after coevaporation with pyridine. TBDMSCl (3.0 g, 20.0 mmol) was added and the reaction mixture was stirred at room temperature until TLC analysis indicated complete consumption of the starting material. Isobutyric anhydride (6.6 mL, 40.0 mmol) was added and the reaction mixture was stirred overnight. After concentration the residue was redissolved in EtOAc (100 mL) and washed with sat. aq. NaHCO₃, 5% citric acid and water. The organic layer was dried (MgSO₄) and directly desilylated without further purification. Towards this end the residue was dissolved in MeCN/water (100 mL, 4:1, v/v) and excess *p*-TsOH (2.4 g, 14.8 mmol) was added. After concentration the residue was taken up in EtOAc and washed with sat. aq. NaHCO₃ and water. Silica gel column chromatography purification using a gradient of EtOAc in toluene (0/100 – 70/30) yielded the title compound as an off-white powder (4.5 g, 8.5 mmol, 85 %). ¹H-NMR (400 MHz, CDCl₃) δ 7.78 (s, 1H, H8), 7.37 (dd, *J* = 8.5, 7.6 Hz, 2H, Arom. Pac), 7.09 (t, *J* = 7.4 Hz, 1H, Arom. Pac), 7.05 (d, *J* = 7.9 Hz, 2H, Arom. Pac), 5.95 (d, *J* = 7.2 Hz, 1H, H1'), 5.89 (dd, *J* = 7.1, 5.5 Hz, 1H, H2'), 5.71 (dd, *J* = 5.3, 2.2 Hz, 1H, H3'), 4.92 (s, 1H, NH), 4.72 (s, 2H, CH₂ Pac), 4.31 (d, *J* = 1.9 Hz, 1H, H4'), 4.12 - 3.79 (m, 2H, H5'), 2.76 - 2.58 (m, 1H, CH *i*Bu), 2.58 - 2.45 (m, 1H, CH *i*Bu), 1.23 - 1.11 (m, 12H, CH₃ *i*Bu). ¹³C-NMR (100 MHz, CDCl₃) δ 175.7, 175.2 (CO *i*Bu), 169.9 (CO Pac), 156.4 (Cq Pac), 155.0 (C6), 147.0, 146.5 (C2, C4), 138.9 (C8), 129.7, 122.6, 114.7 (Arom. Pac), 87.5 (C1'), 85.0 (C4'), 72.6 (C2'), 71.4 (C3'), 66.5 (CH₂ Pac), 61.9 (C5'), 33.7, 33.5 (CH *i*Bu), 18.8, 18.7, 18.6, 18.5 (CH₃ *i*Bu). IR: 2974, 1684, 1608, 1558, 1471, 1407, 1188, 1150, 1096. LCMS (10-90% B in 15 min), Rt = 7.94. ESI-MS: *m/z* 558.0 [M+H]⁺. HRMS: [C₂₆H₃₁N₅O₉ + H]⁺ calcd. 558.2194, found 558.2194.

2-cyanoethoxy-*N,N'*-di-isopropylamino-(*N*⁴-benzoyl-2', 3'-di-*O*- isobutyryl-β-D-cytidyl-5'-yl) phosphine (10)



To a stirred solution of **7** (0.89 g, 1.8 mmol), coevaporated with MeCN, in DCM (5 mL), containing distilled TEA (0.84 mL, 5.9 mmol), 2-cyanoethoxy-*N,N'*-di-isopropylaminochlorophosphine (0.50 mL, 2.25 mmol) was added under an argon atmosphere. The reaction mixture was stirred for 45 min at room temperature. After ³¹P-NMR showed complete consumption of the 2-cyanoethoxy-*N,N'*-di-isopropylaminochlorophosphine (80 MHz, CDCl₃; δ 180.4) and formation of the phosphoramidite (80 MHz, CDCl₃; δ 150.3, 149.6), 30 mL DCM was added and the reaction mixture was washed with water and 5% NaHCO₃. The organic phase was dried (MgSO₄) and concentrated. The residue was applied on a pre-neutralized (1%TEA in PE) silica gel column and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as an off-white foam. ¹H-NMR (400 MHz, CDCl₃); δ 8.18 (t, *J* = 7.8 Hz 1H, H6), 7.94 – 7.76 (m, 2H, Bz), 7.44 – 7.32 (m, 3H, H5, Bz), 6.24 (d, *J* = 5.7 Hz, 1H, H1'), 6.17 (d, *J* = 4.5 Hz, 1H, H1'), 5.38 (dd, *J* = 5.1, 3.7 Hz, 1H, H2'), 5.33 – 5.28 (m, 1H, H3'), 4.08 – 3.80 (m, 4H, H5', CH₂ CNEO), 3.69 – 3.62 (m, 2H, CH *iPr*), 2.76 – 2.69 (m, 2H, CH₂ CNEO), 2.63 – 2.55 (m, 2H, CH *iBu*), 1.27 – 1.14 (m, 24H, CH₃ *iBu*, CH₃ *iPr*). ¹³C-NMR (100 MHz, CDCl₃); δ 175.6, 175.5, 175.3, 175.2 (CO *iBu*), 167.3 (C2), 162.7 (CO Bz), 154.6, 144.4, 144.3 (C6, C4), 133.6 (Cq Bz), 133.1, 128.7, 127.8 (Bz), 117.7 (CN), 97.4 (C5), 87.6, 87.1 (C1'), 82.8, 82.7, 82.4, 82.3 (C4'), 74.1, 74.0 (C2'), 70.8, 70.5 (C3'), 62.9, 62.7, 62.6, 62.4 (C5'), 58.8, 58.6, 58.4 (CH₂ CNEO), 43.2, 43.2, 43.1, 43.0 (CH *iPr*), 33.69, 33.66, 33.55 (CH *iBu*), 24.7, 24.6, 24.5, 24.5 (CH₃ *iPr*), 20.4, 20.3, 20.2 (CH₂ CNEO), 19.0 (CH₃ *iBu*). ³¹P-NMR (161 MHz, CDCl₃); δ 150.3, 149.6. IR: 2980, 1746, 1666, 1624, 1556, 1481, 1310, 1245, 1153. LCMS (10-90% B in 15 min), Rt = 6.80, ESI-MS: *m/z* 605.1 [M + H]⁺ of the corresponding H-Phosphonate. TLCMS: 688.5 [M + H]⁺ HRMS of the corresponding H-Phosphonate: [C₂₇H₃₃N₄O₁₀ + H]⁺ calcd. 605.2007, found 605.2005.

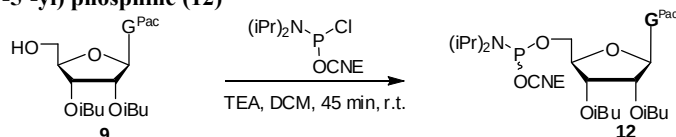
2-cyanoethoxy-*N,N'*-di-isopropylamino-(*N*⁶-benzoyl-2',3'-di-*O*-isobutyryl-β-D-adenyl-5'-yl) phosphine (11)



To a stirred solution of **8** (1.3 g, 2.6 mmol), coevaporated with MeCN, in DCM (15 mL), containing distilled TEA (1.2 mL, 8.5 mmol), 2-cyanoethoxy-*N,N'*-di-isopropylaminochlorophosphine (0.7 mL, 3.2 mmol) was added under an argon atmosphere. The reaction mixture was stirred for 45 min at room temperature. After ³¹P-NMR showed complete consumption of the 2-cyanoethoxy-*N,N'*-di-isopropylaminochlorophosphine (161 MHz, CDCl₃; δ 180.4) and formation of the product (161 MHz, CDCl₃; δ 150.0, 149.5), 30 mL DCM was added and the reaction mixture was washed with water and 5% NaHCO₃. The organic phase was dried (MgSO₄) and concentrated. The residue was applied on a pre-neutralized (1%TEA in PE) silica gel column and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as an off-white foam (1.1 g, 1.6 mmol, 60%). ¹H NMR (400 MHz, CDCl₃) δ 9.64 (s, 1H, NH), 8.75 (d, *J* = 4.1 Hz, 1H, H2), 8.53 (d, *J* = 8.1 Hz, 1H, H8), 8.02 (d, *J* = 7.3 Hz, 2H, Arom. Bz), 7.56 (t, *J* = 7.2 Hz, 1H, Arom. Bz), 7.47 (t, *J* = 7.3 Hz, 2H, Arom. Bz), 6.45, 6.40 (2x d, *J* = 6.6 Hz, 1H, H1'), 6.00 – 5.79 (m, 1H, H2'), 5.76 – 5.50 (m, 1H, H3'), 4.42 (m, 1H, H4'), 4.09 – 3.77 (m, 4H, H5', CH₂ CNEO), 3.74 – 3.56 (m, 2H, CH *iPr*), 2.79 – 2.43 (m, 4H, CH₂ CNEO, CH *iBu*), 1.31 – 1.00 (m, 24H, CH₃ *iBu*, CH₃ *iPr*). ¹³C-NMR (100 MHz, CDCl₃); δ 175.91, 175.86, 175.4, 175.3 (CO *iBu*), 164.6 (CO Bz), 154.9 (C2), 152.0, 151.9 (C6), 149.6 (C4), 141.4, 141.3 (C8), 133.7 (Cq Bz), 132.8, 128.9,

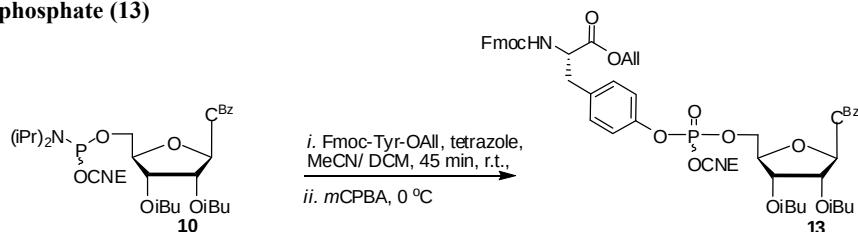
128.8, 127.8 (Arom. Bz), 122.9 (C5), 117.7, 117.6 (CN), 85.4, 85.0 (C1'), 83.9, 83.8, 83.6, 83.5 (C4'), 74.0, 73.9 (C2'), 72.0, 71.7 (C3'), 63.2, 63.0, 63.0, 62.8 (C5'), 58.8, 58.6, 58.4 (CH₂ CNEO), 43.2, 43.1 (CH *i*Pr), 33.9, 33.8, 33.59, 33.57 (CH *i*Bu), 24.7, 24.6 (CH₃ *i*Pr), 20.40, 20.36, 20.32, 20.29 (CH₂ CNEO), 18.9, 18.8, 18.7, 18.6 (CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ 150.0, 149.5. IR: 2968, 1746, 1702, 1610, 1582, 1508, 1458, 1245, 1184, 1153. LCMS (10-90% B in 15 min), Rt = 7.37. ESI-MS: *m/z* 629.1 [M + H]⁺ of the corresponding H-Phosphonate. TLCMS: 712.6 [M + H]⁺. HRMS: [C₃₅H₄₈N₇O₁₀P + H]⁺; 712.3218 calcd., 712.3229 found.

2-cyanoethoxy-*N,N'*-di-isopropylamino-(*N*²-phenoxyacetyl-2',3'-di-*O*-isobutyryl-β-D-guanyl-5'-yl) phosphine (12)



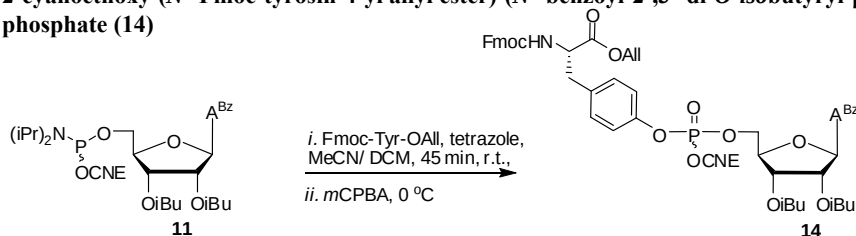
To a stirred solution of **9** (0.94 g, 1.77 mmol), coevaporated with MeCN in DCM (10 mL), containing freshly distilled TEA (0.8 mL, 5.8 mmol), 2-cyanoethoxy-*N,N*-di-isopropylaminochlorophosphine (0.5 mL, 2.2 mmol) was added under an argon atmosphere. The reaction mixture was subsequently stirred for 45 min at room temperature. After ³¹P-NMR showed complete consumption of the 2-cyanoethoxy-*N,N*-diisopropylaminochlorophosphine (80 MHz, CDCl₃; δ 180.4) and formation of the phosphoramidite (³¹P-NMR, 80 MHz: δ 150.0, 149.1), 20 mL DCM was added and the reaction mixture was washed with water and 5% NaHCO₃. The organic phase was dried (MgSO₄) and concentrated. The residue was applied on a pre-neutralized (1% TEA in PE) silica gel column and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as an off-white foam (1.1 g, 1.4 mmol, 78%). ¹H-NMR (400 MHz, CDCl₃); δ 8.14 (s, 0.5 H, C2), 8.13 (s, 0.5 H, C2), 7.49 – 7.31 (m, 2H, Arom. Pac), 7.13 (t, *J* = 7.4 Hz, 1H, Arom. Pac), 7.08 – 6.94 (m, 2H, Arom. Pac), 6.11 (dd, *J* = 9.5, 6.1 Hz, 1H, H1'), 5.88 (dd, *J* = 10.5, 4.9 Hz, 1H, H2'), 5.72 (td, *J* = 5.6, 3.4 Hz, 1H, H3'), 4.73 (s, 2H, CH₂ Pac), 4.39 – 3.38 (m, 3H, H4'), 3.99 – 3.80 (m, 4H, H5', CH₂ CNEO), 3.67 – 2.60 (m, 2H, CH *i*Pr), 2.75 – 2.54 (m, 3H, CH *i*Bu, CH₂ CNEO), 1.24 – 1.05 (m, 24H, CH₃ *i*Bu, CH₃ *i*Pr). ¹³C-NMR (100 MHz, CDCl₃); δ 175.82, 175.76, 175.2, 175.1 (CO *i*Bu), 169.6, 169.45 (CO Pac), 156.5, 156.4 (C6), 155.4 (Cq Pac), 148.0, 147.9 (C4), 146.4, 138.1, 137.7, 130.02 (Arom. Pac), 129.97 (Arom. Pac), 123.05 (Arom. Pac), 122.98 (Arom. Pac), 117.7, 114.9 (Arom. Pac), 85.7, 85.3 (C1'), 83.3, 83.2, 82.9, 82.8 (C4'), 73.44, 73.37 (C2'), 71.4, 71.3 (C3'), 67.0 (CH₂ Pac), 62.92, 62.85, 62.75, 62.67 (C5'), 58.7, 58.5, 58.3 (CH₂ CNEO), 43.26, 43.21, 43.13, 43.08 (CH *i*Pr), 33.8, 33.6, (CH *i*Bu), 24.71, 24.67, 24.64, 24.60 (CH₃ *i*Pr), 20.39, 20.35, 20.31, 20.28 (CH₂ CNEO), 18.91, 18.87, 18.8, 18.6 (CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ 150.1, 149.7. IR: 2972, 1745, 1685, 1599, 1560, 1497, 1388, 1364, 1183, 1154. LCMS (10-90% B in 15 min), Rt = 7.93. ESI-MS: *m/z* 674.9 [M+H]⁺ of the corresponding H-Phosphonate. TLCMS: 758.5 [M+H]⁺. HRMS: [C₃₅H₄₈N₇O₁₀P + H]⁺; 758.3273 calcd., 758.3280 found.

2-cyanoethoxy-(*N*^α-Fmoc-tyrosin-4-yl allyl ester)(*N*⁴-benzoyl-2', 3'-di-*O*-isobutryrlyl-β-D-cytidyl- 5'-yl) phosphate (13)



N^α- Fmoc-tyrosine-allyl ester (0.33 g, 0.75 mmol) was added to compound **10** (0.50 g, 0.72 mmol) and the two compounds were coevaporated with MeCN and DCE and then dissolved in 1:1 v/v MeCN/DCM (5 mL). Tetrazole (200 mg, 2.88 mmol) was added and the reaction mixture was stirred for 45 min until ³¹P-NMR (161 MHz: δ 135.36, 134.91) showed complete consumption of the phosphoramidite. The reaction mixture was cooled to 0 °C and *m*-CPBA was added in portions until ³¹P-NMR showed complete consumption of the phosphite and formation of the phosphotriester (161 MHz: δ -6.17, -6.31). The reaction mixture was diluted in 20 mL DCM and washed with 5% Na₂S₂O₃ and water. The organic phase was dried (MgSO₄) and concentrated. The residue was applied to a silica gel column and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as a colourless oil in 74 % (0.55 g, 0.53 mmol). ¹H-NMR (400 MHz, CDCl₃); δ 8.18 - 7.06 (m, 18H, H₆, Arom. Fmoc, Arom. Bz, Arom. Tyr.), 6.39 (d, *J* = 8.2 Hz, 0.5H, H1'), 6.30 (d, *J* = 8.2 Hz, 0.5H, H1'), 6.19 (d, *J* = 6.5 Hz, 1H, H5), 6.10 - 5.79 (m, 1H, CH Allyl), 5.73 (d, *J* = 8.2 Hz, 0.5H, NH), 5.57 (d, *J* = 8.2 Hz, 0.5H, NH), 5.41 - 5.18 (m, 4H, H3', H2', CH₂ Allyl), 4.74 (m, 1H, CH α), 4.40 - 4.26 (m, 10H, H4', H5', CH₂ Fmoc, CH₂ Allyl, CH₂ CNEO, CH Fmoc), 3.16 - 3.06 (m, 2H, CH₂ β), 2.59 - 2.55 (m, 2H, CH₂ CNEO), 2.61 - 2.53 (m, 2H, CH *i*Bu), 1.21 - 1.14 (m, 12H, CH₃ *i*Bu). ¹³C-NMR (100 MHz, CDCl₃); δ 175.62, 175.55, 175.4 (CO *i*Bu), 171.2 (CO α), 167.2 (C2), 163.1, 163.0 (CO Bz), 155.9, 155.8 (CO Fmoc), 154.8 (C6), 149.1, 149.0 (Cq Arom. Tyr.), 144.6, 144.3 (C4), 143.8, 143.7, 141.2 (Cq Fmoc), 134.3, 134.1 (Cq Arom. Tyr.), 133.03, 132.99 (Bz), 132.7 (Cq Bz), 131.5 (CH Allyl), 131.0 (Arom. Tyr.), 128.7 (Bz), 128.0 (Arom. Fmoc), 127.7 (Bz), 127.0, 125.1 (Arom. Fmoc), 119.9 (Arom. Fmoc, Arom. Tyr.), 118.9 (CH₂ Allyl), 116.7 (CN), 97.7 (C5), 89.5, 88.9 (C1'), 80.8, 80.7 (C4'), 73.5 (C2'), 69.63, 69.55 (C3'), 67.1, 66.9, 66.0 (CH₂ Fmoc, CH₂ Allyl, C5'), 63.3, 63.2 (CH₂ CNEO), 55.0 (CH α), 47.0, 46.4 (CH Fmoc), 37.1 (CH₂ β), 33.7 (CH *i*Bu), 19.62, 19.55 (CH₂ CNEO), 18.80, 18.78, 18.75, 18.68 (CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ -5.86, -5.89. IR: 2980, 1740, 1700, 1668, 1626, 1556, 1506, 1481, 1250, 1188, 1036. LCMS (50 – 90 % B in 15 min), Rt = 6.26. ESI-MS: *m/z* 1046.5 [M+H]⁺. HRMS: [C₅₄H₅₆N₅O₁₅P + H]⁺; calcd. 1046.3583, found 1046.3593.

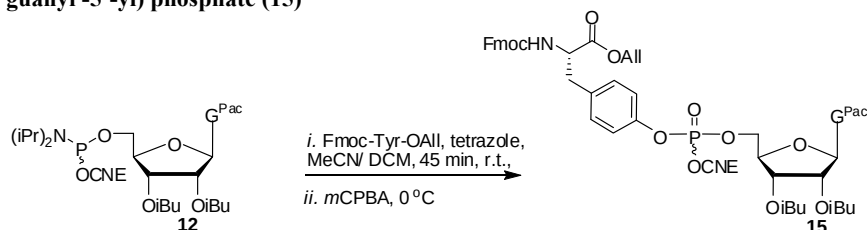
2-cyanoethoxy-(*N*^α-Fmoc-tyrosin-4-yl allyl ester)-(*N*⁶-benzoyl-2',3'-di-*O*-isobutryrlyl-β-D-adenyl- 5'-yl) phosphate (14)



N^α-Fmoc-tyrosine-allyl ester (0.26 g, 0.58 mmol) was added to amidite **11** (0.40 g, 0.58 mmol) and the two compounds were coevaporated with MeCN and DCE and then dissolved in 1:1 v/v MeCN/DCM (10 mL). Tetrazole (0.12 g, 1.74 mmol) was added and the reaction mixture was stirred for 45 minutes until ³¹P-NMR (161 MHz: δ 135.36, 134.91) showed complete consumption of the phosphoramidite (161 MHz, CDCl₃; δ 150.31, 148.83). The reaction mixture was cooled to 0 °C and *m*-CPBA was added in portions until ³¹P-NMR (161 MHz: δ -6.17, -6.31) showed complete consumption of the phosphite. The reaction

mixture was diluted in DCM and washed with 5% Na₂S₂O₃ and water. The organic phase was dried (MgSO₄) and concentrated. The residue was applied to a silica gel column and eluted with a gradient of EtOAc in PE (50/50 – 100/0) to yield the title compound as a colourless oil (0.39 g, 0.37 mmol, 64%) ¹H-NMR (400 MHz, CDCl₃); δ 9.30, 9.28 (m, 1H, H₂), 8.78 (2x s, 1H, H₈), 7.95 - 6.89 (m, 17H), 6.53 (d, *J* = 8.0 Hz, 1H, NH), 6.37 - 6.33 (m, 1H, H1'), 5.92 - 5.68 (m, 1H, CH Allyl, H3'), 5.54 - 5.51 (m, 1H, H2'), 5.28 - 5.15 (m, 2H, CH₂ Allyl), 4.65 - 4.60 (m, 1H, CH α), 4.56 - 4.32 (m, 9H, C4', C5', CH₂ Fmoc, CH₂ Allyl, CH₂ CNEO), 4.27 - 3.98 (m, 1H, CH Fmoc), 3.12 - 3.03 (m, 2H, CH₂ β), 2.92 - 2.88 (m, 2H, CH₂ CNEO), 2.82 - 2.47 (m, 2H, CH iBu), 1.26 - 1.04 (m, 12H, CH₃ iBu). ¹³C-NMR (100 MHz, CDCl₃); δ 175.7, 175.5, 175.2, 175.0 (CO iBu), 171.0, 170.8 (CO α), 165.0, 164.8 (CO Bz), 155.6 (CO Fmoc), 152.7 (C2), 151.3, 151.1 (C4, C6), 149.5, 148.8 (Cq Tyr.), 143.6, 143.2, 141.22, 141.15, 140.9, 140.5 (Cq Fmoc), 139.6 (C8), 134.0, 133.6, 133.4, 133.2 (Cq Arom. Tyr.), 132.3 – 123.9 (Arom.), 122.7, 122.0 (C5), 119.9 (Arom.) 119.1 (CH₂ Allyl), 116.1 (CN), 85.9, 85.5 (C1'), 81.7, 81.6, 81.44, 81.36 (C4'), 73.5, 73.1 (C2'), 70.3, 70.2 (C3'), 67.1, 66.0, 65.9, 65.3 (CH₂ Fmoc, CH₂ Allyl, CH₂ CNEO), 62.9 (C5'), 54.4, 54.0 (CH α), 47.0, 46.9 (CH Fmoc), 36.6, 36.2 (CH₂ β), 33.6, 33.4 (CH iBu), 19.5, 19.4 (CH₂ CNEO), 18.8, 18.7, 18.5 (CH₃ iBu). ³¹P-NMR (161 MHz, CDCl₃); δ -6.97, 6.98. IR: 2976, 1738, 1699, 1608, 1583, 1505, 1456, 1247, 1189, 1155, 1029, 961. LCMS (50 – 90 % B in 15 min), Rt = 10.25. ESI-MS: *m/z* 1070.2 [M+H]⁺. HRMS: [C₅₅H₅₆N₇O₁₄P + H]⁺; 1070.3695 calcd., 1070.3706 found.

2-cyanoethoxy-(*N*^a-Fmoc-tyrosin-4-yl-allyl guanyl -5'-yl) phosphate (15)



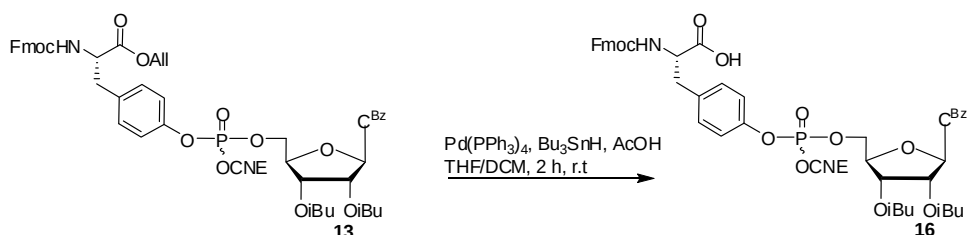
N^a-Fmoc-tyrosine-allyl ester (0.78 g, 1.77 mmol) was added to amidite **12** (1.0 eq., 1.77 mmol) and the two compounds were coevaporated with MeCN and DCE and then dissolved in 1:1 MeCN/DCM (10 mL, v/v). Tetrazole (295 mg, 4.2 mmol) was added and the reaction mixture was stirred for 45 minutes until ³¹P NMR showed complete consumption of the phosphoramidite (³¹P-NMR, 161 MHz: δ 134.18, 134.06). The reaction mixture was cooled to 0 °C and *m*-CPBA was added in portions until ³¹P NMR showed complete oxidation of the phosphite (³¹P-NMR, 161 MHz: δ -6.56, -6.98). The reaction mixture was diluted in 20 mL DCM and washed with 5% Na₂S₂O₃ and water. The organic phase was dried (MgSO₄) and concentrated. The residue was applied to a silica gel using a gradient of EtOAc in PE (50/50 – 70/30) to yield the separate epimers at phosphorus as white foams, 233 mg (0.21 mmol, 12 %) (Rt = 0.5, EtOAc) and 476 mg (0.47 mmol, 27 %) (Rt = 0.6, EtOAc) of which analytical samples were repurified using silica gel column chromatography using a gradient of 2 % MeOH in DCM, and a mix of both epimers (0.43 g, 0.38 mmol, 24 %) respectively. A total of 1.14 g, 1.0 mmol, 56 %.

epimer A (Rt = 0.6), (93:7 epimeric ration). ¹H-NMR (400 MHz, CDCl₃); δ 11.74, 10.33 (2 x s, 1H, NH H1, NH H2), 7.77 (s, 1H, H₈), 7.76 – 7.71 (m, 2H, Arom.), 7.56 – 7.53 (m, 2H, Arom.), 7.37 – 7.33 (m, 2H, Arom.), 7.31 – 7.24 (m, 4H, Arom.), 7.01 – 6.95 (m, 6H, Arom.), 6.12 – 6.09 (m, 1H, H2'), 5.97 – 5.95 (m, 2H, H1'), 5.89 – 5.78 (m, 2H, CH Allyl, H3'), 5.26 (m, 2H, CH₂ Allyl), 4.60 – 4.52 (m, 7H, CH₂ Allyl, CH₂ Fmoc, H5', Ha), 4.43 (s, 1 H, H4'), 4.36 – 4.23 (m, 2H, CH₂ CNEO), 4.19 – 4.15 (m, 1H, CH Fmoc), 3.12 – 2.97 (m, 2H, CH₂ β), 2.73 – 2.60 (m, 2H, CH₂ CNEO), 2.56 – 2.49 (m, 2H, CH iBu), 1.22 – 1.10 (m, 12H, CH₃ iBu). ¹³C-NMR (100 MHz, CDCl₃); δ 175.6, 175.1 (CO iBu), 170.8, 170.1 (CO Pac, CO α), 156.9, 155.5 (Cq Pac, C6), 155.0 (CO Fmoc), 148.4, 148.3 (Cq Tyr.), 147.4, 146.5 (C4, C2), 143.6, 140.9 (Cq Fmoc), 134.0 (Cq Tyr), 131.2 (CH Allyl), 130.3 (Arom. Tyr.), 129.4 (Arom. Pac), 127.4, 126.8, 124.9 (Arom. Fmoc), 122.6, 121.9 (C5), 119.8, 119.7, 119.6, (Arom. Fmoc, Arom. Tyr.), 118.8 (CH₂ Allyl),

116.1 (CN), 114.6 (Arom. Pac), 87.0 (C1'), 81.0, 80.9 (C4'), 71.2 (C2'), 70.0 (C3'), 67.0, 66.63, 66.60 (CH₂ Fmoc, CH₂ Pac, C5'), 65.8 (CH₂ Allyl), 62.8, 62.7 (CH₂ CNEO), 54.6 (CH α), 46.8 (CH Fmoc), 36.8 (CH₂ β), 33.5, 33.3 (CH *i*Bu), 19.3, 19.2 (CH₂ CNEO), 18.9, 18.6, 18.5, 18.4, 18.3 (CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ - 6.94. IR: 2970, 1695, 1610, 1558, 1539, 1497, 1472, 1448, 1188, 1153, 1041. LCMS (50 - 90% B in 15 min), Rt = 7.56 ESI-MS: *m/z* 1116.6 [M+H]⁺. α_D (CHCl₃): -17.4. HRMS: C₅₆H₅₈N₇O₁₆P + H⁺ calcd. 1116.3750, found 1116.3765.

epimer B (Rt = 0.5), (96:4 *epimer* ratio). ¹H-NMR (400 MHz, CDCl₃); δ 11.84, 10.42 (2 x s, 1H, NH α , NH H2), 7.75 - 7.73 (m, 3H, Arom.), 7.57 (m, 2H, Arom.), 7.39 - 7.35 (m, 2H, Arom.), 7.29 - 7.25 (m, 4H, Arom.), 7.09 - 6.97 (m, 6H, Arom.), 6.01 - 5.95 (m, 2H, H1', H2'), 5.93 - 5.82 (m, 1H, CH Allyl), 5.79 - 5.76 (m, 1H, H3'), 5.33 - 5.23 (m, 2H, CH₂ Allyl), 4.69 - 4.68 (m, 2H, CH₂ Pac), 4.64 - 4.60 (m, 3H, CH α , CH₂ Allyl), 4.49 - 4.47 (m, 2H, CH₂ Fmoc), 4.38 - 4.23 (m, 2H, H4', H5'), 4.21 - 4.15 (m, 3H, CH Fmoc, CH₂ CNEO), 3.16 - 3.06 (m, 2H, CH₂ β), 2.65 - 2.49 (m, 4H, CH *i*Bu, CH₂ CNEO), 1.20 - 1.09 (m, 12H, CH₃ *i*Bu). ¹³C-NMR (100 MHz, CDCl₃); δ 175.6, 175.1 (CO *i*Bu), 170.7, 170.1 (CO Pac, CO α), 156.9, 155.5 (Cq Pac, C6), 155.0 (CO Fmoc), 148.4, 148.3 (Cq Tyr.), 147.4, 146.5 (C4, C2), 143.6, 140.9 (Cq Fmoc), 134.0 (Cq Tyr.), 131.2, 130.3 (CH Allyl), 129.4, 127.4, 126.8, 124.9 (Arom.), 122.6 (C5), 122.0, 119.7, 119.6 (Arom.), 118.8 (CH₂ Allyl), 116.1, 114.6 (CN), 87.0 (C1'), 81.0, 80.9 (C4'), 71.2 (C2'), 69.9 (C3'), 67.0, 66.6, 66.4, 65.8, 65.87 (CH₂ Allyl, C5', CH₂ Fmoc, CH₂ Pac), 62.8, 62.7 (CH₂ CNEO), 54.6 (CH α), 46.8 (CH Fmoc), 36.8 (CH₂ β), 33.5, 33.3 (CH *i*Bu), 19.3, 19.2 (CH₂ CNEO), 18.9, 18.6, 18.5, 18.4 (CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ -7.3. IR: 2970, 1695, 1610, 1558, 1539, 1497, 1471, 1450, 1387, 1190, 1153, 1033. LCMS (50 -90% B in 15 min), Rt = 7.61 ESI-MS: *m/z* 1116.5 [M+H]⁺. α_D (CHCl₃): -24.8. HRMS: [C₅₆H₅₈N₇O₁₆P + H]⁺ calcd. 1116.3750, found 1116.3764.

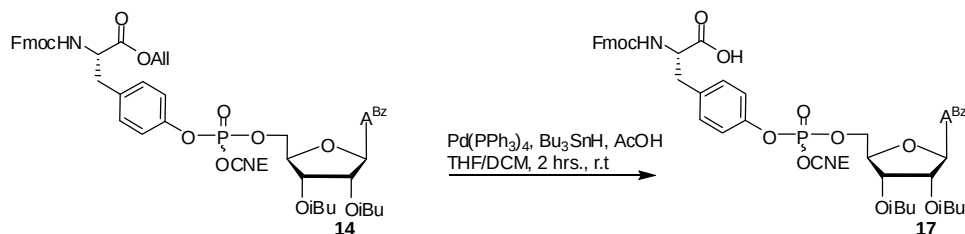
2-cyanoethoxy-(*N*⁴-Fmoc-tyrosin-4-yl)(*N*⁴-benzoyl-2',3'-di-*O*-isobutryl- β -D-cytidin-5'-yl) phosphate (16)



Compound **13** (0.55 g, 0.53 mmol) was coevaporated with MeCN and dissolved in 1:1 THF/DCM (10 mL, v/v). After addition of AcOH (2.6 mmol, 0.15 mL), Bu₃SnH (1.05 mmol, 0.32 mL) and Pd(PPh₃)₄ (0.02 mmol, 20 mg) the reaction was stirred for 2 hours. The reaction mixture was concentrated and the residue was applied to a silica gel column and eluted with a gradient of MeOH in DCM (0/100 - 5/95) to yield the title compound as a white foam (0.30 g, 0.29 mmol, 56 %). ¹H-NMR (400 MHz, CDCl₃); δ 7.92 - 7.87, 7.75 - 7.73, 7.54 - 7.47, 7.39 - 7.35, 7.29 - 7.25, 7.21 - 7.14 (m, 17H, H6, H5, Arom.), 6.07 (m, 1H, C1'), 5.72 (m, 1H, NH), 5.46 - 5.42 (m, 2H, C3', C2'), 4.70 (m, 1H, CH α), 4.48 - 4.43 (m, 2H, CH₂ Fmoc), 4.33 - 4.31 (m, 3H, C4', C5'), 4.30 - 4.23 (m, 2H, CH₂ CNEO), 4.13 (m, 1H, CH Fmoc), 3.18 - 3.11 (m, 2H, CH₂ β), 2.75 - 2.63 (m, 2H, CH₂ CNEO), 2.62 - 2.58 (m, 2H, CH *i*Bu), 1.17 - 1.14 (m, 12H, CH₃ *i*Bu). ¹³C-NMR (100 MHz, CDCl₃); δ 175.50, 175.46, 175.4 (CO *i*Bu), 174.3 (CO α), 167.2 (C2), 163.2, 163.1 (CO Bz), 155.6 (CO Fmoc), 154.3 (C4), 148.8, 148.7 (Cq Arom. Tyr.), 145.0 (C6), 143.7, 143.6, 141.0 (Cq Arom. Fmoc), 134.0 (Cq Tyr.), 133.0 (Bz), 132.48 (Cq Bz), 131.0 (Arom. Tyr.), 128.50 (Bz), 128.0 (Arom. Fmoc), 127.5 (Bz), 126.9, 125.0, 124.9 (Arom. Fmoc), 119.8, 119.6 (Arom. Fmoc, Arom. Tyr.), 116.3 (CN), 97.6 (C5), 89.8, 89.29 (C1'), 80.7, 80.6 (C4'), 73.4 (C2'), 69.4, 69.3 (C3'), 66.95, 66.86, 66.68 (CH₂ Fmoc, C5'), 63.02, 62.97 (CH₂ CNEO), 54.6 (C α), 46.9 (CH Fmoc), 37.0 (C β), 33.5 (CH *i*Bu), 19.43, 19.35 (CH₂ CNEO), 18.64, 18.62 (CH₃ *i*Bu). ³¹P-NMR (161 MHz, CDCl₃); δ -6.37, -6.77. IR: 1699,

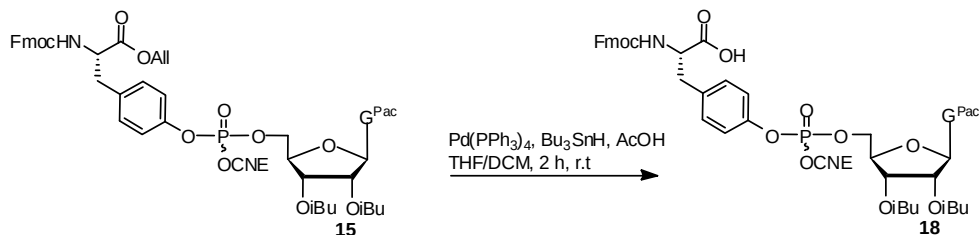
1668, 1624, 1558, 1481, 1450, 1249. LCMS (10 – 90 % B in 15 min) Rt = 9.06. ESI-MS: m/z 1006.7 $[M+H]^+$. HRMS: $[C_{51}H_{52}N_5O_{15}P + H]^+$; calcd. 1006.3270, found 1006.3281.

2-cyanoethoxy-(*N*^ε-Fmoc-tyrosin-4-yl)-(N⁶-benzoyl-2',3'-di-*O*-isobutyryl-β-D-adenyl-5'-yl) phosphate (17)



Compound **14** (0.39 g, 0.37 mmol) was coevaporated with MeCN and dissolved in 1:1 THF/DCM (10 mL, v/v). After addition of AcOH (1.7 mmol, 0.10 mL), Bu₃SnH (0.74 mmol, 0.21 mL) and Pd(PPh₃)₄ (0.015 mmol, 18 mg) the reaction was stirred for 2 hours. The reaction mixture was concentrated and the residue was applied to a silica gel column and eluted with a gradient of MeOH in DCM (0/100 – 5/95) to yield the title compound as a white foam (0.32 g, 0.31 mmol, 85 %). ¹H-NMR (400 MHz, CDCl₃); δ 8.78 (2x s, 1H, H8), 8.12, 7.86 (2x s, 1H, H2), 7.98 – 6.89 (m, 17H), 6.43, 6.38 (2x d, *J* = 7.2 Hz, 1H, H1'), 5.97, 5.95 (d, *J* = 7.6 Hz, 1H, NH), 5.75 – 5.59 (m, 2H, C3', C2'), 4.52 – 4.36 (m, 7H, CH₂ Fmoc, CH α, C4', CH₂ CNEO, CH Fmoc), 4.17 – 4.15 (m, 2H, H5'), 3.27 – 3.21 (m, 2H, CH₂ β), 2.98 – 2.92 (m, 2H, CH₂ CNEO), 2.72 – 2.48 (2x m, 2H, CH iBu), 1.29 – 1.06 (m, 12H, CH₃ iBu). ¹³C-NMR (100 MHz, CDCl₃); δ 175.5, 175.3 (CO iBu), 174.4, 174.0 (CO α), 165.7 (CO Bz), 155.4 (CO Fmoc), 153.1 (C2), 151.6 (Cq Arom. Tyr.), 149.8 (C6), 149.0 (C4), 143.75, 143.68, 143.61, 141.2 (Cq Arom. Fmoc), 140.4 (C8), 134.0 (Cq Tyr.), 133.5 (Cq Bz), 132.7 (Arom. Bz), 131.2, 131.0 (Arom. Tyr.), 128.5, 128.5 (Arom. Bz), 128.2, 128.2 (Arom. Fmoc), 127.7, 127.6 (Arom. Bz), 127.0, 125.0, 125.0, 124.9, 124.7 (Arom. Fmoc), 122.0, 121.8 (C5), 119.9, 119.9, 119.7 (Arom. Fmoc, Arom. Tyr.), 116.2, 116.1 (CN), 85.2, 84.8 (C1'), 82.0, 81.9 (C4'), 73.6, 73.3 (C2'), 70.4, 70.0 (C3'), 67.6, 67.1, 66.6, 66.4 (CH₂ Fmoc, CH₂ CNEO), 63.0, 63.0, 62.8, 62.8 (C5'), 55.0, 54.8 (CH α), 47.2 (CH Fmoc), 38.3, 37.9 (CH₂ β), 33.79, 33.77, 33.56, 33.53 (CH iBu), 19.7, 19.6 (CH₂ CNEO), 18.9, 18.8, 18.7, 18.6, 18.60, 18.57 (CH₃ iBu). ³¹P-NMR (161 MHz, CDCl₃); δ -6.8, -7.1. IR: 2976, 2360, 1738, 1714, 1700, 1584, 1505, 1455, 1247, 1155, 1032. LCMS (10 – 90 % B in 15 min) Rt = 9.39. ESI-MS: m/z 1030.1 $[M+H]^+$. HRMS: $[C_{52}H_{52}N_7O_{14}P + H]^+$; 1030.3383 calcd., 1030.3394 found.

2-cyanoethoxy-(*N*^ε-Fmoc-tyrosin-4-yl)-(N²-phenoxyacetyl-2',3'-di-*O*-isobutyryl-β-D-guanyl-5'-yl) phosphate (18)



A mixture epimers of **15** (0.43 g, 0.38 mmol) was coevaporated with MeCN and dissolved in 1:1 v/v THF/DCM (10 mL). After addition of AcOH (2.0 mmol, 0.11 mL), Bu₃SnH (0.36 mL, 0.76 mmol) and Pd(PPh₃)₄ (0.02 mmol, 20 mg) the reaction was stirred for 2 hours. The reaction mixture was concentrated and the residue was applied to a silica gel column and eluted with a gradient of MeOH in DCM (0/100 – 4/96) containing 1% of AcOH to yield the title compound as a white foam (0.23 g, 0.21 mmol, 56 %). ¹H-

NMR (400 MHz, CDCl_3); δ 7.88 (s, 1H, H8), 7.75 - 6.99 (m, 2H, Arom.), 7.59 - 7.57 (m, 2H, Arom), 7.39 - 7.35 (m, 2H, Arom), 7.31 - 7.27 (m, 4H, Arom), 7.13 (m, 2H, Arom), 7.04 - 6.99 (m, 5H, Arom), 6.00 - 5.92 (m, 2H, H1', H3'), 5.88 - 5.85 (m, 1H, H2'), 4.68 (m, 2H, CH_2 Pac), 4.48 - 4.17 (m, 9H, CH α , CH_2 Fmoc, H4', H5', CH Fmoc, CH_2 CNEO), 3.15 (m, 1H, CH_2 β), 2.67 - 2.55 (m, 4H, CH *i*Bu, CH_2 CNEO), 1.27 - 1.10 (m, 12H, CH_3 *i*Bu). ^{13}C -NMR (100 MHz, CDCl_3); δ 176.2, 175.9, 175.5 (CO *i*Bu), 174.0 (CO Pac), 170.3 (CO α), 156.8 (Cq Pac), 155.37, 155.36 (C6, CO Fmoc), 148.9, 148.8, 148.5 (Cq Tyr.), 147.8 (C4), 146.9 (C8), 143.9, 143.8, 141.3 (Cq Fmoc), 134.2 (Cq Tyr.), 131.2 (Arom. Tyr.), 129.8 (Arom. Pac), 127.7, 127.1, 125.2 (Arom. Fmoc), 122.6 (C5), 121.6 (Arom. Pac), 119.9, 119.7 (Arom. Fmoc, Arom. Tyr.), 116.6 (CN), 114.9 (Arom. Pac), 87.0 (C1'), 80.9, 80.8 (C4'), 72.5 (C2'), 69.8 (C3'), 67.1, 66.9, 66.8 (C5', CH_2 PAC, CH_2 Fmoc), 63.1, 63.03 (CH_2 CNEO), 54.9 (CH α), 47.2 (CH Fmoc), 37.1 (CH_2 β), 33.8, 33.7 (CH *i*Bu), 19.5, 19.4 (CH_2 CNEO), 18.9, 18.84, 18.81, 18.7 (CH_3 *i*Bu). ^{31}P -NMR (161 MHz, CDCl_3); δ -7.0, -7.1. IR: 2951, 1686, 1574, 1560, 1412, 1248, 1209, 1153, 1070. LCMS (10-90% B in 15 min), Rt = 8.96, ESI-MS: m/z 1076.6 $[\text{M}+\text{H}]^+$. HRMS: $\text{C}_{53}\text{H}_{54}\text{N}_7\text{O}_{16}\text{P} + \text{H}]^+$ calcd. 1076.3437, found 1076.3449.

VPgpC, (H-Gly-Ala-Tyr(pC)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile- Arg-Thr-Ala-Lys-Val-Glu-OH) (1)

The crude protected nucleopeptide was cleaved off the resin using TFA/TIS/ H_2O (95/5/5, v/v/v, 5 mL) by shaking for 2.5 hours. After filtration into cold Et_2O the precipitate was coevaporated with toluene (2 x 5 mL) and then treated with 25% aq. ammonia (2.5 mL) in 1,4-dioxane (2.5 mL). After stirring for 4 hours, concentration of the reaction mixture afforded the crude title compound. RP-HPLC purification was conducted on an automated HPLC system supplied with a semi-preparative C₁₈ column (250 x 10.00 mm, 5 μM , flow: 4 mL/min). The applied eluent system was; A: H_2O , B: MeCN, C: 0.5 % TFA and detection at 280 nm. Gradient: 10 - 25 % B in 3 CV. The collected fractions were lyophilized, yielding 14.89 mg, 5.6 μmol (45 % based on original loading of resin) of the nucleopeptide. ^{31}P -NMR (161 MHz, D_2O); δ -4.24. LCMS (00 - 90 % B in 15 min.), Rt = 4.18 min. ESI-MS: m/z 2659.2 $[\text{M}]^+$, 1330.4 $[\text{M}]^{2+}$, 887.6 $[\text{M}]^{3+}$. HRMS: $[\text{C}_{114}\text{H}_{189}\text{N}_{34}\text{O}_{37}\text{P} + \text{H}]^{2+}$ calcd. 1329.6918, found 1329.6937, $[\text{M} + \text{H}]^{3+}$ calcd. 886.7970, found 886.7981, $[\text{M} + \text{H}]^{4+}$ calcd. 665.3495, found 665.3497, $[\text{M} + \text{H}]^{5+}$ calcd. 532.4811, found 532.4807. ^1H -NMR (600 MHz, D_2O) δ 7.97 (d, J = 8.0 Hz, 1H), 7.20 (d, J = 8.4 Hz, 2H), 7.10 (d, J = 8.1 Hz, 2H), 6.02 (d, J = 7.9 Hz, 1H), 5.90 (d, J = 3.9 Hz, 1H), 4.67 (ap. dd, J = 15.2, 7.9 Hz, 1H), 4.61 (ap. dd, J = 12.8, 5.9 Hz, 1H), 4.57 (ap. dd, J = 8.6, 5.4 Hz, 1H), 4.51 - 4.37 (m, 5H), 4.37 - 4.20 (m, 14H), 4.21 - 4.05 (m, 7H), 3.97 - 3.77 (m, 8H), 3.70 (ap. dd, J = 16.8, 7.0 Hz, 1H), 3.64 (ap. dd, J = 16.3, 6.8 Hz, 2H), 2.35 (ap. t, J = 7.6 Hz, 3H), 2.30 - 2.25 (m, 4H), 2.17 - 1.93 (m, 14H), 1.93 - 1.40 (m, 43H), 1.37 (ap. d, J = 7.2 Hz, 7H), 1.28 (ap. d, J = 7.2 Hz, 5H), 1.22 - 1.18 (m, 9H), 1.16 (ap. d, J = 6.4 Hz, 5H), 0.98 - 0.84 (m, 35H).

VPgpA, (H-Gly-Ala-Tyr(pA)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Glu-OH) (2)

The crude partially protected nucleopeptide was cleaved of the resin using TFA/TIS/ H_2O (95/5/5, v/v/v, 5 mL) by shaking for 4 hours. After filtration into cold Et_2O the precipitate was concentrated and RP-HPLC purification was conducted on an automated HPLC system supplied with a semi-preparative C₁₈ column (250 x 10.00 mm, 5 μM , flow: 4 mL/min). The applied eluent system was; A: 0.5 % TFA in H_2O , B: MeCN, Gradient: 20- 50 % B in 3 CV and detection at 280 nm. to give partially protected **21**. This nucleopeptide was then treated with 25% aq. ammonia (1.25 mL) in 1,4-dioxane (1.25 mL). After stirring for 5 hours careful concentration of the reaction mixture afforded the crude title compound. RP-HPLC purification was conducted on an automated HPLC system supplied with a semi-preparative C₁₈ column (250 x 10.00 mm, 5 μM , flow: 4 mL/min). The applied eluents system was; A: 0.5 % TFA in H_2O , B: MeCN and detection at 280 nm. Gradient: 10- 25 % B in 3 CV. The collected fractions were lyophilized, yielding the target nucleopeptide **21**; 6.1 mg, 2.3 μmol (9% based on initial loading of the resin). ^{31}P -NMR (161 MHz, D_2O); δ -4.21. LCMS (00 - 90 % B in 15 min.), Rt = 4.18 min. ESI-MS: m/z 1345.0 $[\text{M}]^{2+}$, 896.6 $[\text{M}]^{3+}$, 672.7 $[\text{M}]^{4+}$. HRMS: $[\text{C}_{115}\text{H}_{189}\text{N}_{36}\text{O}_{36}\text{P} + \text{H}]^{2+}$ calcd. 1341.6974, found 1341.6949, $[\text{M} + \text{H}]^{3+}$ calcd. 894.8007, found 894.8008, $[\text{M} + \text{H}]^{4+}$ calcd. 671.3524, found 671.3532, $[\text{M} + \text{H}]^{5+}$ calcd. 537.2833, found 537.2839. ^1H NMR (400 MHz, D_2O) δ 8.36 (s, 1H), 8.33 (s, 1H), 6.96 (q, J = 8.8 Hz, 4H), 6.13 (d, J = 5.6 Hz, 1H), 4.70 - 4.56 (m, 1H), 4.56 - 4.39 (m, 4H), 4.39 - 4.27 (m, 6H), 4.27 - 4.08 (m, 5H), 4.02 - 3.78 (m,

5H), 3.78 - 3.71 (m, 1H), 3.67 (ap. d, $J = 7.6$ Hz, 2H), 3.22 (ap. t, $J = 6.9$ Hz, 2H), 3.13 - 2.95 (m, 6H), 2.95 - 2.67 (m, 4H), 2.41 - 2.22 (m, 4H), 2.09 - 2.07 (m, 8H), 1.97 - 1.45 (m, 22H), 1.41 (ap. d, $J = 7.1$ Hz, 4H), 1.29 (ap. d, $J = 7.2$ Hz, 3H), 1.23 (ap. dd, $J = 6.3, 3.2$ Hz, 5H), 1.19 (ap. d, $J = 6.4$ Hz, 3H), 1.04 - 0.81 (m, 19H).

VPgG, (H-Gly-Ala-Tyr(pG)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Glu-OH) (3)

The crude partially protected nucleopeptide was cleaved of the resin using TFA/TIS/H₂O (95/5/5, v/v/v, 5 mL) by shaking for 2.5 hours. After filtration into cold Et₂O the precipitate was coevaporated with toluene (2 x 5 mL) and then treated with 25% aq. ammonia (1.25 mL) in 1,4-dioxane (1.25 mL). After stirring for 5 hours careful concentration of the reaction mixture afforded the crude title compound. RP-HPLC purification was conducted on an automated HPLC system supplied with a semi-preparative C18 column (250 x 10.00 mm, 5 μ M, flow: 4 mL/min). The applied eluent system was; A: H₂O, B: MeCN, C: 0.5 % TFA in H₂O, in a gradient of 10- 30 % B (10% C as stationary phase) in 3 CV and detection at 280 nm. The collected fractions were lyophilized, yielding the target nucleopeptide 22; 9.8 mg, 3.6 μ mol (29% based on initial loading of the resin). ³¹P-NMR (161 MHz, D₂O); δ - 4.21. LCMS (00 – 90% B in 5 min.), Rt = 4.21 min., ESI-MS: m/z 2700.2 [M + H]⁺, 1350.2[M+H]²⁺, 900.6 [M+H]³⁺. HRMS: [C₁₁₅H₁₈₉N₃₆O₃₇P + H]²⁺ calcd. 1349.6948, found 1349.6952, [M + H]³⁺ calcd. 900.1323, found 900.1327, [M + H]⁴⁺ calcd. 675.3511, found 675.3508, [M + H]⁵⁺ calcd. 540.4823, found 540.4816. ¹H-NMR (600 MHz, D₂O) δ 7.85 (s, 1H), 7.08 - 6.84 (m, 4H), 5.86 (d, $J = 5.9$ Hz, 1H), 4.90 (ap. t, $J = 5.6$ Hz, 1H), 4.66 (ap. t, $J = 7.1$ Hz, 1H), 4.63 - 4.59 (m, 2H), 4.57 (ap. d, $J = 11.2$ Hz, 1H), 4.50 - 4.37 (m, 9H), 4.37 - 4.26 (m, 13H), 4.26 - 4.08 (m, 15H), 3.91 (ap. d, $J = 13.7$ Hz, 3H), 3.89 - 3.79 (m, 5H), 3.77 (ap. d, $J = 3.6$ Hz, 3H), 3.70 (ap. d, $J = 7.7$ Hz, 2H), 3.64 (s, 4H), 3.19 (ap. t, $J = 6.0$ Hz, 5H), 3.09 - 2.92 (m, 15H), 2.89 (ap. dd, $J = 13.8, 8.9$ Hz, 2H), 2.83 - 2.62 (m, 8H), 2.38 - 2.20 (m, 11H), 2.13 - 1.39 (m, 92H), 1.37 (ap. d, $J = 7.2$ Hz, 10H), 1.27 (ap. d, $J = 7.2$ Hz, 7H), 1.21 - 1.11 (m, 23H), 0.95 (ap. dd, $J = 13.3, 6.8$ Hz, 23H), 0.90 (ap. dd, $J = 13.5, 5.4$ Hz, 25H), 0.85 (ap. t, $J = 7.3$ Hz, 10H).

Notes and References

- ¹ Van der Heden van Noort, G. J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *J. Org. Chem.*, **2010**, 75, 5733 – 5736.
- ² Puustinen, P., Makinen, K., *J. Biol. Chem.*, **2004**, 279, 38103 – 38110.
- ³ Pan, J., Lin, L., Tao, Y. J., *Virology*, **2009**, 395, 87 – 96.
- ⁴ Shapiro, B. M., Stadtman, E. R., *J. Biol. Chem.*, **1968**, 243, 3769 – 3771.
- ⁵ Belliot, G., Sosnovtsev, S. V., Chang, K., McPhie, P., Green, K. Y., *Virology*, **2008**, 374, 33 – 49.
- ⁶ Broberg, C. A., Orth, K., *Curr. Opin. Microbiol.*, **2010**, 13, 34 – 40.
- ⁷ Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., Zekarias, B., Lazar, C., Dixon, J. E., *Mol. Cell*, **2009**, 34, 93 – 103.
- ⁸ Rothberg, P. G., Harris, T. J., Nomoto, A., Wimmer, E., *Proc. Natl. Acad. Sci. U.S.A.*, **1978**, 75, 4868 – 4872.
- ⁹ Steil, B. P., Barton, D. J., *Virus Res.*, **2009**, 139, 240 – 252.
- ¹⁰ Ambros, V., Baltimore, D., *J. Biol. Chem.*, **1978**, 352, 5263 – 5266.
- ¹¹ Steil, B. P., Barton, D. J., *J. Virol.*, **2009**, 83, 12660 – 12670.
- ¹² Van Ooij, M. J. M., Vogt, D. A., Paul, A., Castro C., Kuijpers, J., van Kuppeveld, F., J. M., Cameron, C. E., Wimmer, E., Andino, R., Melcher, W. J. G., *J. Gen. Virol.*, **2006**, 87, 103 – 113.
- ¹³ Murray, K., Barton, D., *J. Virol.*, **2003**, 77, 4739 – 4750.
- ¹⁴ Pathak, H. B., Arnold, J. J., Wiegand, P. N., Hargittai, M. R. S., Cameron, C. E., *J. Biol. Chem.*, **2007**, 282, 16202 – 16213.
- ¹⁵ Goodfellow, I. G., Kerrigan, D., Evans, D. J., *RNA*, **2003**, 9, 124 – 137.
- ¹⁶ Paul, A. V., Yin, J., Mugavero, J., Rieder, E., Liu, Y., Wimmer, E., *J. Biol. Chem.*, **2003**, 278, 43951 – 43960.
- ¹⁷ Gerber, K., Wimmer, E., Paul, A. V., *J. Virol.*, **2001**, 75, 10979 – 10990.
- ¹⁸ Paul, A. V., Rieder, E., Kim, D. W., van Boom, J. H., Wimmer, E., *J. Virol.*, **2000**, 74, 10359 – 10370.
- ¹⁹ Schein, C. H., Volk, D. E., Oezguen, N., Paul, A., *Proteins: Struct.,Funct., Bioinf.*, **2006**, 63, 719 – 726.
- ²⁰ Rieder, E., Paul, A. V., Kim, D. W., van Boom, J. H., Wimmer, E., *J. Virol.*, **2000**, 74, 10371 – 10380.
- ²¹ Van der Heden van Noort, G. J., van der Horst, M. G., Overkleeft, H. S., van der Marel, G. A., Filippov, D. V., *J. Am. Chem. Soc.*, **2010**, 132, 5236 – 5240.
- ²² Kriek, N. M. A., Fillipov, D. V., van den Elst, H., Meeuwenoord, N. J., Tesser, G. I., van Boom, J. H., van der Marel, G. A., *Tetrahedron*, **2003**, 59, 1589 – 1597.

- ²³ Robles, J., Beltran, M., Marchan, V., Perez, Y., Travasset, I., Pedroso, E., Grandas, A., *Tetrahedron*, **1999**, 55, 13251 – 13264.
- ²⁴ Dreef-Tromp, C. M., van Dam, E. M. A., van den Elst, H., van der Marel, G. A., van Boom, J. H., *Nucleic Acids Res.*, **1990**, 18, 6491 – 6495.
- ²⁵ Sung, W. L.; Narang, S. A.; *Can. J. Chem.*, **1982**, 60, 111 – 120.
- ²⁶ van der Heden van Noort, G. J.; Verhagen, C. P.; van der Horst, M. G.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *Org. Lett.*, **2008**, 20, 4461 – 4464.
- ²⁷ Fan, Y.; Gaffney, B. L.; Jones, R. A.; *Org. Lett.*, **2004**, 6, 2555 – 2557.
- ²⁸ Schulhof, J. C., Molko, D., Teoule, R., *Nucleic Acids Res.*, **1987**, 15, 397 – 416.
- ²⁹ See supporting information of Van der Heden van Noort, G. J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; *J. Org. Chem.*, **2010**, 75, 5733 – 5736 for a crude LCMS trace of **18**.
- ³⁰ Al-Eryani, R. A., Ball, H. L., *Tetrahedron Lett.*, **2010**, 51, 1730 – 1731.
- ³¹ Medzihradsky, D., Chen, S. L., Kenyon, G. L., Gibson, B. W., *J. Am. Chem. Soc.*, **1994**, 116, 9413 – 9419.
- ³² Bae, S., Lakshman, M. K., *J. Am. Chem. Soc.*, **2007**, 129, 782 – 789.
- ³³ Bae, S., Lakshman, M. K., *Org. Lett.*, **2008**, 10, 2203 – 2206.
- ³⁴ Pritz, S., Wolf, Y., Klemm, C., Bienert, M., *Tetrahedron Lett.*, **2006**, 47, 5893 – 5896.
- ³⁵ Belliot, G.; Sosnovtsev, S. V.; Chang, K. O.; McPhie, P.; Green, K. Y.; *Virology*, **2008**, 374, 33 – 49.

Solid Phase Approaches Towards RNA-Nucleopeptides

Introduction

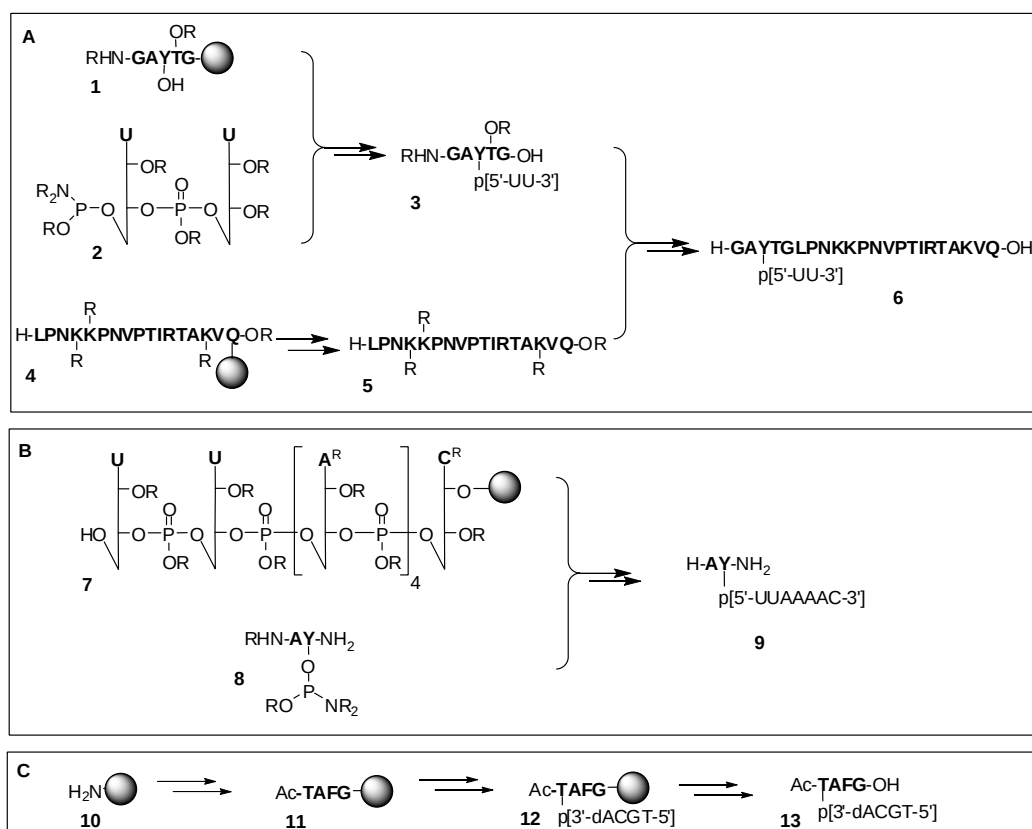
The synthesis of fragments of nucleoproteins is far more challenging than synthesis of the individual nucleic acid or peptide components. Although excellent protocols for (automated) synthesis of peptides and nucleic acids are available, the synthesis of nucleopeptides asks for the development of a new approach, in which the integrity of the peptide, the nucleic acid and the connecting phosphodiester linkage is kept intact. Not all protective groups used in nucleic acid and peptide chemistry can be used for nucleopeptide synthesis. Basic conditions commonly used in nucleic acid synthesis can cause β -elimination in serine or threonine containing nucleopeptides and cleave the crucial phosphodiester linkage between the nucleic acid and peptide part.¹ Strongly acidic conditions commonly used in peptide chemistry on the other hand can give rise to side reactions in the nucleic acid portion of nucleopeptides, such as depurination in the case of DNA² and phosphate migration in RNA.³ Synthetic routes towards DNA-nucleopeptides have been extensively explored⁴, while the more challenging RNA-nucleopeptides have received less attention. Nucleopeptides, containing a short RNA fragment, have been synthetically obtained and have proven to be useful tools in the study, of for instance, viral replication pathways.^{5,6} To gain further insight in the reaction mechanism and structure of the proteins involved in these pathways, availability of nucleopeptides carrying longer RNA fragments can be very helpful.

Block-coupling strategies in which the final coupling is executed in solution (see **Figure 1A**) have been reported to give access to RNA-nucleopeptides carrying one or two uridylyl groups.^{7,8} In such an approach the phosphoramidite of dinucleotide **2** is coupled to a short resin bound peptide **1** and subsequently oxidized to give the desired phosphotriester bond. After deprotection and release from the support the nucleopeptide fragment **3** can be condensed with another separately prepared peptide fragment **5** to give the desired full length VPgpUpU nucleoprotein **6**. No successful block-coupling synthesis of nucleopeptides containing oligoribonucleotides has been reported.

Another approach involves the synthesis of an oligonucleotide on solid support and the subsequent coupling of a separately prepared peptide fragment (see **Figure 1B**). It has been established that coupling of peptidyl phosphoramidite **8** to a RNA fragment anchored on a solid support (**7**) is a viable way to obtain nucleopeptide **9**. To date several nucleopeptides have been prepared using this methodology.⁹⁻¹² Peptide sequences having more than 2 amino acids attached to a 7-mer RNA fragment have not been reported, due to the limited reactivity of peptidyl phosphoramidites.

An approach in which both peptide and nucleic acid are stepwise elongated on the same solid support would be ideal. An example of such an approach starts with the construction of a peptide on solid support (**11**) followed by the construction of the oligonucleotide onto this immobilized peptide (**12**) (see **Figure 1C**). Final cleavage of the protective groups and release from the resin results in the nucleopeptide **13**. This approach has been reported to be successful for the synthesis of 3'-end linked DNA-nucleopeptide fragments.⁴ Examples of 5'-end linked DNA-nucleopeptides however are scarce¹³⁻¹⁵ and both 3'-end linked and 5'-end linked RNA-nucleopeptides have never been prepared using such a methodology.

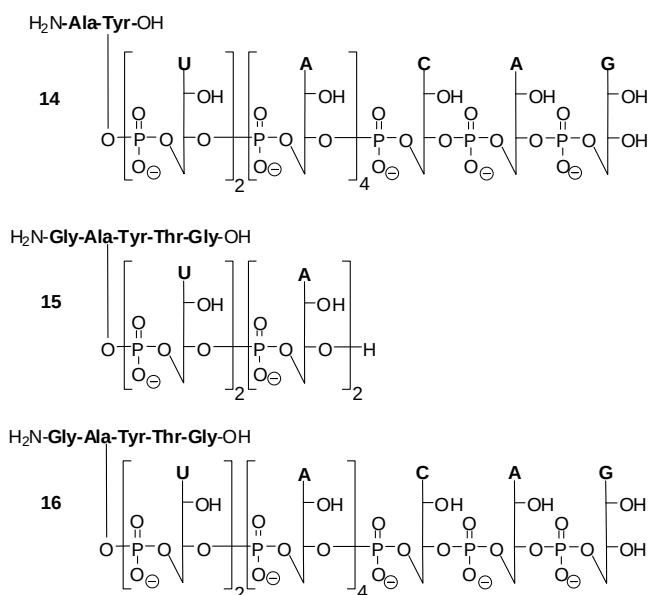
Figure 1



Results and Discussion

The poliovirus VPg protein is a 22 amino acid protein in which tyrosine-3 is linked to the 5'-end of the viral (+)-strand RNA genome. It has been shown that VPg nucleopeptides containing one or two uridylyl moieties act as primers in viral replication pathways. Research on these replication processes benefits from well-defined VPg nucleopeptide fragments. Such fragments cannot be isolated from natural sources and must be prepared by synthetic procedures. In **Chapter 5** and **Chapter 6** the synthesis of *mono*-nucleotidylated VPg proteins is described and the synthesis of VPg proteins carrying two uridylyl moieties has been reported earlier.⁸ However, the assembly of VPg derived peptides carrying longer RNA fragments is not yet published. This chapter describes a study toward the synthesis of a dipeptidyl nonaribonucleotide **14**, a pentapeptidyl tetraribonucleotide **15** and a pentapeptidyl nonaribonucleotide **16** derived from the VPg nucleoprotein (see **Figure 2**).

Figure 2

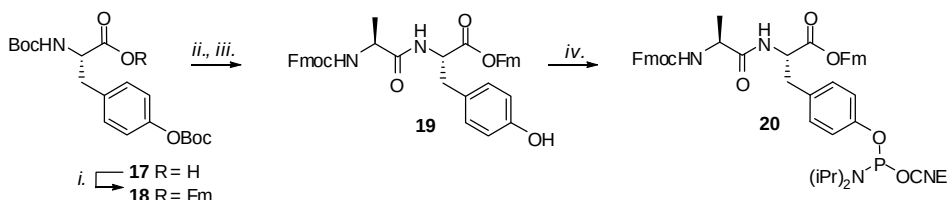


In an earlier report the poliovirus derived RNA-nucleopeptide H-Ala-Tyr[5'-pUUAAAAC-3']-NH₂ has been prepared using the methodology depicted in **Figure 1B**.⁹ A similar approach was selected to gain access to the VPg nucleoprotein derived H-Ala-Tyr[5'-pUUAAAACAG-3']-OH fragment **14**. In the reported example the RNA fragment was equipped with silyl protective groups on the exocyclic amines of the nucleosides, the dipeptide bore a nitrophenylsulfonyl (NPS) protective group on the N-terminus, whereas the C-terminus was permanently blocked as a carboxamide.¹⁰ Altering the protective group strategy on the peptide part of this dipeptidylribonucleotide fragment to Fmoc and Fm (see compound **20**, **Scheme 1**) would reduce the number of deprotection steps, as Fmoc and Fm are readily cleaved under the conditions that are used to eliminate cyanoethyl protection on the phosphodiester moieties. In addition this protective group strategy would give access to the C-terminal acid instead of the carboxamide.

Commercially available nucleotide phosphoramidites for RNA synthesis, equipped with acyl protective groups on the exocyclic amines were chosen, to avoid the preparation of the previously used silyl protected phosphoramidites.

First peptidylphosphoramidite **20** was prepared in solution (see **Scheme 1**). Boc protected tyrosine **17** was esterified with 9-fluorenylmethanol (FmOH) using EDC·HCl and DMAP to give Fm ester **18**. Acidolysis of both Boc groups followed by coupling to Fmoc-Ala-OH resulted in dipeptide **19**, which was phosphitylated to yield the crucial phosphoramidite **20**.

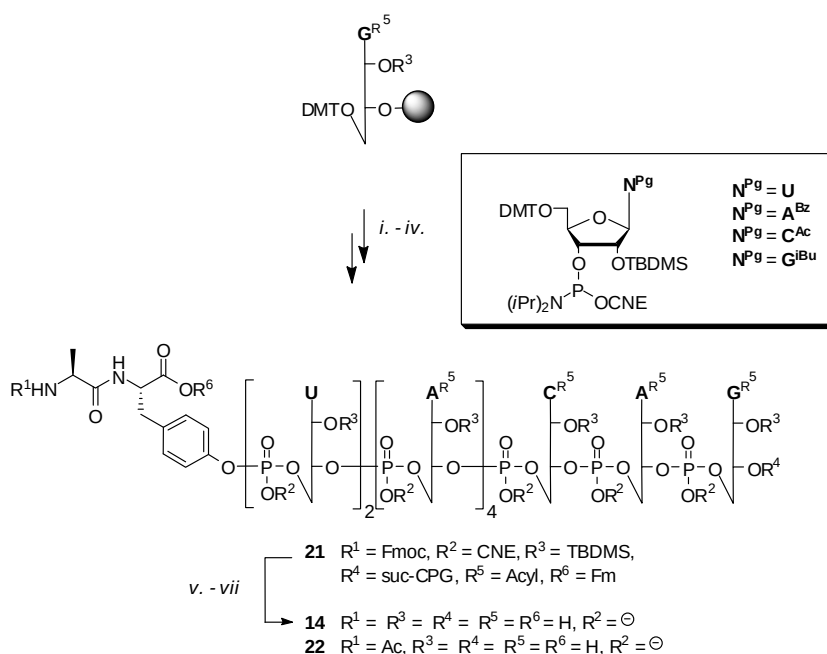
Scheme 1



Conditions and reagents: *i.* FmOH, EDC·HCl, DMAP, DCM, 16 h, r.t. (62%), *ii.* DCM, TFA, TIS, 90 min, r.t., *iii.* Fmoc-Ala-OH, DiPEA, EDC·HCl, DCM, 3 h, r.t. (44% over two steps), *iv.* 2-cyanoethoxy-*N,N*-diisopropylaminochlorophosphine, TEA, MeCN/Dioxane, 45 min, 0°C (62%).

The protected RNA fragment 5'-pUUAAAACAG-3' was prepared using commercially available 5'-*O*-DMT RNA-phosphoramidites and a standard solid phase synthesis protocol in an automated nucleic acid synthesizer (see **Scheme 2**). In the final coupling cycle the immobilized nonribonucleotide was elongated with amidite **20** to give fully protected resin bound nucleopeptide **21**. Deprotection was accomplished by removing the Fmoc/Fm and CNE groups using DBU, cleavage of the acyl functionalities and release from the solid support using NH₃ in MeOH and final TEA·HF treatment to remove silyl protective groups.

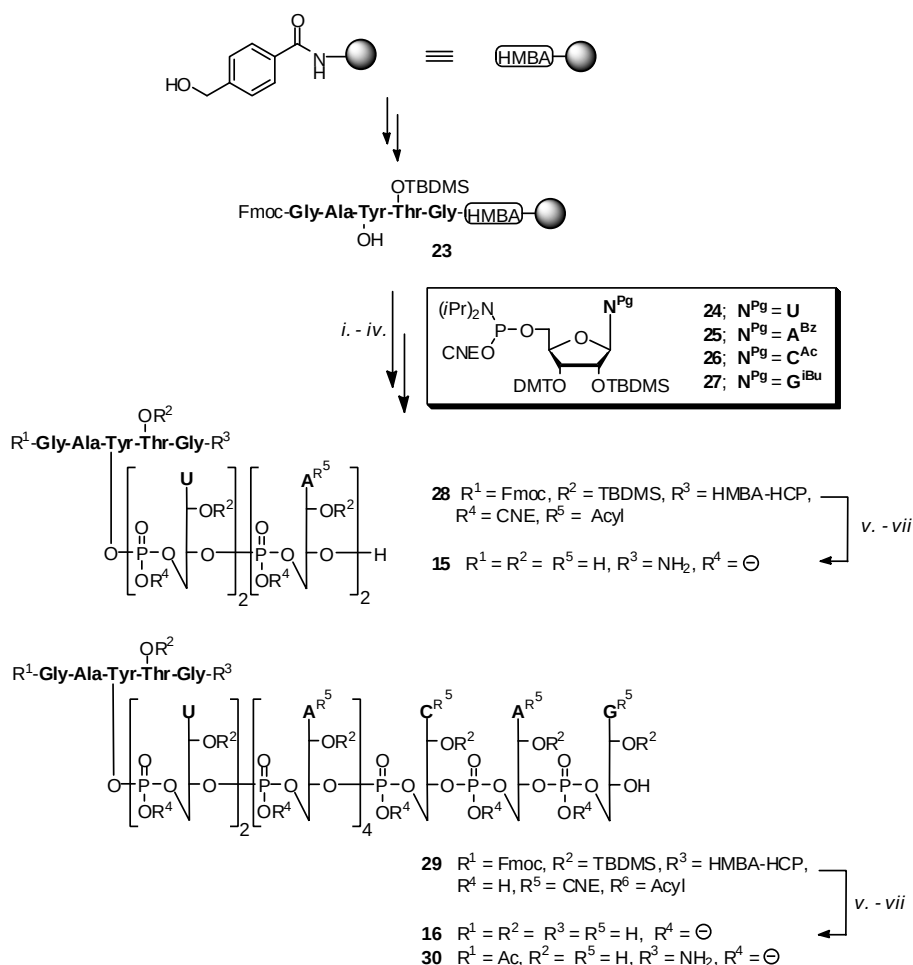
Analysis of the crude product by LC-MS showed the presence of an additional acetyl group that remained on the molecule even after treatment with NH₃. The capping step after phosphoramidite couplings presumably resulted in acetylation of the mono-acyl protected exocyclic amines in the oligonucleotide. One of those acetyl groups might possibly migrate to the N-terminus after cleavage of the Fmoc group giving rise to **22** instead of **14**.



Conditions and reagents: for each nucleotide the following cycle was used *i.* 3% DCA in toluene, *ii.* 5eq. appropriate phosphoramidite, 5-benzylmercaptotetrazole, MeCN, *iii.* I₂, pyridine/H₂O, *iv.* *N*-methylimidazole, Ac₂O, lutidine, MeCN, *v.* 2% DBU, DMF, 3x 5 min. *vi.* NH₃/MeOH, 16 h, 50 °C. *vii.* TEA·HF, DMSO, 16 h, r. t.

The synthesis of nucleopeptides containing longer peptide and RNA fragments (such as **15** and **16** in **Figure 2**) however would not be feasible using the above described approach. On the basis of a reported, fully sequential solid phase synthesis of partially protected DNA nucleopeptide fragments^{14,15} it seemed appropriate to start with construction of a peptide on solid support, followed by RNA synthesis on this immobilized peptide (**figure 1C**). In this approach the oligonucleotide is elongated from the 5'-end towards the 3'-end, in contrast to reverse elongation that is usually applied in solid phase nucleic acid synthesis. The use of 5'-phosphoramidites bearing acyl protective groups on the exocyclic amines and a 3'-O-DMT as temporary protective group seemed appropriate for such a synthesis. To assess the viability of this approach, the synthesis of nucleopeptides **15** and **16** was undertaken (see **Scheme 3**). Pentapeptidyl tetra ribonucleotide **15** and pentapeptidyl nona ribonucleotide **16** are derived from the genome bound VPg protein found in picornaviruses and contains the five N-terminal amino acids of the native VPg.

Scheme 3



Conditions and reagents: *i.* 5eq. phosphoramidite **24** – **27**, *ii.* $\text{I}_2/\text{pyridine}/\text{H}_2\text{O}$, *iii.* Ac_2O , *N*-methylimidazole, lutidine, MeCN, *iv.* 3% DCA, toluene, *v.* 2% DBU, DMF, 3x 5 min, *vi.* NH_3/TFE , 16 h, 50 °C, *vii.* TEA.HF, DMSO, 16 h, r. t.

In the first stage immobilized pentapeptide **23** was prepared. The side chain protective groups of the amino acid building blocks were chosen to be orthogonal to the DMT group used on the phosphoramidites. For this reason TBDMS and Fmoc were selected to protect the threonine hydroxyl and N-terminus, respectively. These protective groups are simultaneously cleaved with the TBDMS and nucleobase protective groups on the RNA fragment, in the final stage towards the target nucleopeptide. Highly cross-linked polystyrene (HCP) was chosen as the solid support, since oligonucleotide synthesis is known to be optimal using this type of resin. The hydroxymethylbenzoic acid (HMBA) linker was applied, allowing simultaneous cleavage with the acyl protective groups on the oligonucleotide by NH_3 treatment. The first amino acid was immobilized by esterification of Fmoc-Gly-OH with the HMBA-HCP resin. Repetitive cycles of Fmoc-deprotection and coupling of the appropriate amino acid resulted in formation

of pentapeptide **23**. Resin **23** was then transferred to an automated oligonucleotide synthesizer. The tyrosine hydroxyl in **23** was phosphitylated using the 5'-phosphoramidite of 3'-O-DMT protected uridine (**24**) followed by oxidation, capping and liberation of the 3'-hydroxyl. Repetition of this cycle using the appropriate phosphoramidites **24** - **27** resulted in the formation of fully protected, resin bound ribonucleopeptides **28** and **29**. Resin bound **28** was treated with DBU to cleave the Fmoc and cyanoethyl groups and subsequently treated with NH_3 in trifluoroethanol.¹⁶ Final treatment with triethylamine hydrofluoride followed by RP-HPLC purification and gel filtration gave pentapeptidyl tetraribonucleotide **15**. Application of the same deprotection protocol to resin bound peptide **29** afforded target pentapeptidyl nonaribonucleotide **16** which was found to be accompanied by its N-acetylated form **30**.

Conclusion

In this chapter the synthesis of three RNA-nucleopeptide fragments derived from polio virus VPg is demonstrated. In the first approach a dipeptidyl nonaribonucleotide is assembled using a methodology where the oligonucleotide is assembled on solid support prior to coupling of a dipeptide phosphoramidite, prepared separately in solution. A pentapeptidyl tetraribonucleotide and a pentapeptidyl nonaribonucleotide are prepared by an approach in which a solid phase peptide synthesis is combined with a solid phase oligonucleotide synthesis in an automated nucleic acid synthesizer. First the peptide fragment is assembled on solid support followed by stepwise elongation of the RNA using 5'-phosphoramidites starting from the 5'-end towards the 3'-end. This is the first time that RNA nucleopeptides of this complexity are assembled in a fully stepwise manner.

Experimental

General: 3'-O-DMT-2'-O-TBDMS-*N*⁹-benzoyl-adenosine-5'-(*N,N*-di-isopropylamino-2-cyanoethoxy) phosphoramidite, 3'-O-DMT-2'-O-TBDMS-uridine-5'-(*N,N*-di-isopropylamino-2-cyanoethoxy) phosphoramidite, -O-DMT-2'-O-TBDMS – *N*⁴-acetyl – cytidine-5'-(*N,N*-di-isopropylamino-2-cyanoethoxy) phosphoramidite and 3'-O-DMT-2'-O-TBDMS – *N*²-isobutyryl- guanosine -5'-(*N,N*-di-isopropylamino-2-cyanoethoxy) phosphoramidite were purchased from ChemGenes Corporation, Wilmington, USA and used as received. Solvents and reagents for peptide and oligonucleotide synthesis were used as received.

Procedure for solid phase oligonucleotide synthesis:

The oligo-ribonucleotides were prepared on a 10 μmol scale using an Acta Oligo Pilot (GE) automated nucleic acid synthesizer starting from peptidyl-HMBA-HCP resin **8** or commercially available amino-HCP resin. Couplings were performed using 5 eq. of commercially available amidites at a 0.1 M concentration. 5-benzylmercaptotetrazole (BMT, 0.3 M) was used as the activating agent in 15 minute coupling cycles. Oxidation was performed by means of standard procedure using I₂ (50 mM) in pyridine/H₂O (90:10, v/v). Capping was performed using a standard protocol employing cap A (40% *N*-methylimidazole in MeCN) and cap B (40% Ac₂O, 40% lutidine in MeCN) 1:1 (v/v). Cleavage of the DMT group was accomplished using 3% DCA in toluene.

HCP-HMBA resin

Amino HCP resin (1 g, 200 μmol) was treated with 4-hydroxymethylbenzoic acid (HMBA) (92 mg, 600 μmol), BOP (266 mg, 600 μmol), HOBT (82 mg, 600 μmol) and DiPEA (200 μl, 1.2 mmol) in 10 ml DMF for 5 hours at room temperature. The resin was then washed three times with DMF, DCM and air dried.

Fmoc-Gly-Ala-Tyr-Thr-Gly-HMBA-HCP resin

The resin bound pentapeptide **8** was prepared using a manual coupling protocol starting from HCP-HMBA resin (500 mg, 100 μmol). Fmoc-Gly-OH (150 mg, 500 μmol) was coupled using DIC (101 μL, 500 μmol) and catalytic DMAP in DCM (5 mL) for 4 hours at room temperature. After washing with DCM the unreacted resin was capped using the procedure used for oligonucleotide capping as described above. For subsequent peptide couplings of Fmoc-Thr(OTBDMS)-OH, Fmoc-Ala-OH, Fmoc-Tyr-OH and Fmoc-Gly-OH the following cycle was repeated:

- Fmoc cleavage using 2% DBU in DMF.
- Coupling of the appropriate amino acid applying a five-fold excess, activation by 5 eq. HATU in NMP (0.22M) and 12.5 eq. DiPEA in NMP (1.25 M) for 1 hour.

Boc-Tyr(OBoc)-OFm (18)

Boc-Tyr(OBoc)-OH (1.5 mmol, 0.63 g) was dissolved in DCM (20 mL). FmOH (1.47 mmol, 287 mg), EDC.HCl (2.25 mmol) and DMAP (0.15 mmol, 24 mg) were added and the reaction mixture was stirred for 16 hours at room temperature. After washing with 5% citric acid the organic layer was dried using MgSO₄ and concentrated. The crude product was purified using a silica column chromatography using a 5% EA in PE. The title compound was obtained as a colourless syrup (520 mg, 0.93 mmol, 62 %). ¹H-NMR (400 MHz, CDCl₃) δ 7.79 - 7.68 (m, 2H, Arom. Fm), 7.54 (t, *J* = 7.4 Hz, 2H, Arom. Fm), 7.42 - 7.36 (m, 2H, Arom. Fm), 7.30 (t, *J* = 7.5 Hz, 2H, Arom. Fm), 7.12 - 6.94 (m, 4H, Arom. Tyr.), 4.98 (d, *J* = 8.2 Hz, 1H, NH), 4.61 (dd, *J* = 13.4, 6.1 Hz, 1H, CH α), 4.50 - 4.36 (m, 2H, CH₂ Fm), 4.16 (t, *J* = 6.7 Hz, 1H, CH Fm), 2.99 - 2.95 (m, 2H, CH₂ β), 1.53 (s, 9H, *t*Bu), 1.42 (s, 9H, *t*Bu). ¹³C-NMR (101 MHz, CDCl₃) δ 171.8 (CO α), 155.2 (CO Boc *N*^α), 151.8 (CO Boc *O*), 150.2 (Cq Tyr *O*), 143.54, 143.45, 141.39, 141.35

(Cq Fm), 133.6 (Cq Tyr), 130.3, 127.96, 127.93, 127.26, 127.24, 125.04, 125.02, 121.4, 120.1 (Arom. Fm/Tyr.), 83.5, 80.1 (Cq tBu), 67.1 (CH₂ Fm), 54.4 (α), 46.8 (CH Fm), 37.5 (β), 28.4, 27.7 (tBu).

Fmoc-Ala-Tyr(OH)-OFm (19)

Boc-Tyr(OBoc)-OFm (0.9 mmol) was dissolved in DCM (40 mL) and treated with TFA (6 mL) and TIS (0.5 mL) for 90 minutes at room temperature. Toluene was added (20 mL) and the reaction mixture was concentrated. The crude product was dissolved in DCM (20 mL) and Fmoc-Ala-OH (0.91 mmol, 283 mg), EDC.HCl (0.95 mmol, 120 mg) and DiPEA (1.9 mmol, 315 μ L) were added. After stirring for 3 hours at room temperature the reaction mixture was extracted using sat. aq. NaH₂CO₃. The organic layer was dried using MgSO₄ and concentrated. The crude product was dissolved in a small amount of DCM and crystallised from hexane to yield the title compound as a colourless gel (260 mg, 0.4 mmol, 44%). ¹H-NMR (400 MHz, DMSO) δ 9.19 (s, 1H, OH), 8.31 (s, 1H, NH), 8.23 (d, J = 7.4 Hz, 1H, NH), 7.96 - 7.83 (m, 4H, Arom. Fmoc/Fm), 7.76 - 7.55 (m, 4H, Arom. Fmoc/Fm), 7.49 - 7.22 (m, 8H, Arom. Fmoc/Fm), 6.78 (dd, J = 8.2 Hz, 4H, Arom. Tyr.), 4.50 - 4.29 (m, 4H, CH₂ Fmoc, CH₂ Fm), 4.16 (m, 6H, CH α Tyr, CH α Ala, CH Fmoc, CH Fm), 2.73 (d, J = 7.0 Hz, 2H, CH₂ β Tyr), 1.18 (d, J = 7.1 Hz, 3H, CH₂ β Ala). ¹³C-NMR (101 MHz, CDCl₃/MeOD-*d*₄) δ 172.8, 171.3 (CO α Ala, CO α Tyr), 155.6 (CO Fmoc), 143.1, 143.0, 140.9, 129.9, 127.5, 127.3, 126.8, 126.7, 124.7, 124.6, 119.6, 115.0 (Arom.), 66.8, 66.7 (CH₂ Fm, CH₂ Fmoc), 53.6 (α Tyr), 50.1 (α Ala), 46.7, 46.3 (CH Fm, CH Fmoc), 36.3 (β Tyr), 17.7 (β Ala).

Fmoc-Ala-Tyr(O-[*N,N*-di-isopropyl-amino-2-cyanoethoxy-phosphine])-OFm (20)

Fmoc-Ala-Tyr(OH)-OFm (240 mg, 0.37 mmol) was coevaporated with MeCN prior to being dissolved in MeCN/dioxane (10 mL, 1:1, v/v) at 0 °C. 2-cyanoethoxy-*N,N*-diisopropylaminochlorophosphine (90 μ L, 0.4 mmol) and TEA (154 μ L, 1.2 mmol) were added and the reaction mixture was stirred for 45 min. The reaction mixture was extracted using DCM and sat. aq. NaH₂CO₃. The organic layer was dried using MgSO₄ and purified using a pre-neutralized (1 % TEA) silica column using a gradient of EA: PE, 0 $\rightarrow$ 50. The title compound was obtained as a colourless syrup (213 mg, 0.23 mmol, 62%). ¹H-NMR (400 MHz, CDCl₃) δ 7.75 (dd, J = 7.3, 3.3 Hz, 4H, Arom. Fmoc/Fm), 7.62 - 7.44 (m, 4H, Arom. Fmoc/Fm), 7.42 - 7.22 (m, 8H, Arom. Fmoc/Fm), 7.03 - 6.69 (m, 4H, Arom. Tyr.), 6.54 (d, J = 6.0 Hz, 1H, NH), 5.50 (s, 1H, NH), 4.85 (dd, J = 13.0, 6.1 Hz, 1H, CH α Ala), 4.63 - 4.32 (m, 4H, CH₂ Fmoc, CH₂ Fm), 4.26 - 4.04 (m, 3H, CH α Tyr, CH Fmoc, CH Fm), 3.96 - 3.72 (m, 2H, CH₂ CNEO), 3.70 - 3.52 (m, 2H, CH₂ β Tyr), 2.99 (m, 2H, CH *i*Pr), 2.68 - 2.29 (m, 2H, CH₂ CNEO), 1.33 (d, J = 6.5 Hz, 3H, CH₂ β Ala), 1.15 (dd, J = 23.3, 6.8 Hz, 12H, CH₃ *i*Pr). ¹³C-NMR (101 MHz, CDCl₃) δ 172.0, 171.2 (CO α Ala, CO α Tyr), 156.0 (CO Fmoc), 153.5, 143.9, 143.5, 141.39, 141.35, 130.4, 129.7, 129.6, 128.0, 127.8, 127.3, 127.2, 125.2, 125.0, 120.3, 120.3, 120.2, 120.1, 117.6 (Arom.), 67.3 (CH₂ Fm, CH₂ Fmoc), 58.8, 58.6 (CH₂ CNEO), 53.5 (CH α Tyr), 50.5 (CH α Ala), 47.2, 46.7 (CH Fm, CH Fmoc), 43.8, 43.7 (CH *i*Pr), 37.1 (CH₂ β Tyr), 24.7, 24.6, 24.5, 24.4 (CH₃ *i*Pr), 20.4, 20.3 (CH₂ CNEO), 18.7 (CH₂ β Ala). ³¹P-NMR (161 MHz, CDCl₃); δ 146.6, 146.4

Ac-Ala-Tyr(pUpUpApApApApCpApG)-OH (22)

The oligonucleotide was prepared on 10 μ mol scale as described above starting from amino-HCP resin. The last coupling was performed using amidite **20**. The fully protected product was treated with 2% DBU in dry MeCN to cleave Fmoc/Fm and cyanoethoxy protective groups (3x 5 min). Subsequent treatment with an aqueous solution of concentrated NH₄OH/MeNH₂ (1 : 1 v/v) overnight at room temperature for 16 hours cleaved the peptidylated oligo nucleotide of the resin as well as removed all the acyl protective groups on the nucleosides. The crude peptidylated oligo nucleotide was concentrated and redissolved in DMSO (150 μ L) and treated with TEA·HF (850 μ L) for 16 hours at room temperature. Precipitation from cold *n*BuOH and HPLC purification resulted in the title compound **22**. Yield: 2.0 mg, 0.6 μ mol, 6%. LC-

MS (0 - 25 % MeCN in 10 mM NH₄OAc in 25 min), Rt = 7.84. ESI-MS m/z : 1602.6 [M + H]²⁺, 1068.6 [M + H]³⁺. ³¹P-NMR (161 MHz, CDCl₃); δ -0.47, -4.16. HRMS [C₁₀₁H₁₂₄N₃₉O₆₅P₉ + H]⁺: found 3202.5111, calc. 3202.5307, [C₁₀₁H₁₂₄N₃₉O₆₅P₉ + 2H]²⁺: found 1601.7715, calc. 1601.7690, [C₁₀₁H₁₂₄N₃₉O₆₅P₉ + 3H]³⁺: found 1068.1828, calc. 1068.1817.

H-Gly-Ala-Tyr(pUpUpApA)-Thr-Gly-NH₂ (15)

The RNA-fragment was assembled on pentapeptide resin **23** using the automated oligonucleotide protocol as described above on 10 μ mol scale. The fully protected product was treated with 2% DBU in dry MeCN to cleave the cyanoethoxy protective groups on the phosphates (3x 5 min.). Subsequent treatment with sat. NH₃ in trifluoroethanol (750 μ L) at room temperature for 16 hours cleaved the peptidylated oligo nucleotide of the resin as well as removed all the acyl protective groups on the nucleosides. The resin was washed with H₂O/EtOH, 1:1, v/v and the filtrate was concentrated. The crude peptidylated oligo nucleotide was dissolved in DMSO (150 μ L) and treated with TEA·HF (850 μ L) for 16 hours at room temperature. Precipitation from cold *n*-BuOH and HPLC purification resulted in the title compound **15** (1.65 mg, 0.95 μ mol, 9.5 %). LC-MS (0 - 20 % MeCN in 10 mM NH₄OAc in 15 min), Rt = 6.20. ESI-MS m/z : 1737.8 [M + H]⁺, 869.6 [M + H]²⁺. ³¹P-NMR (161 MHz, CDCl₃); δ -0.32, -0.39, -0.42, -4.40. HRMS [C₅₈H₇₆N₂₀O₃₅P₄ + H]⁺: found 1737.3841, calc. 1737.3850, [C₅₈H₇₆N₂₀O₃₅P₄ + Na]⁺: found 1759.3611, calc. 1759.3624

H-Gly-Ala-Tyr(pUpUpApApApApCpApG)-Thr-Gly-NH₂ (16) and Ac-Gly-Ala-Tyr(pUpUpApApApApCpApG)-Thr-Gly-NH₂ (30)

The RNA-fragment was assembled on pentapeptide **23** using the automated oligonucleotide protocol as described above on 10 μ mol scale. The fully protected product was treated with 2% DBU in dry MeCN to cleave the cyanoethoxy protective groups on the phosphates (3x 5 min.). Subsequent treatment with sat. NH₃ in trifluoroethanol (750 μ L) at 50 °C for 16 hours cleaved the peptidylated oligo nucleotide of the resin as well as removed all the acyl protective groups on the nucleosides. The resin was washed with H₂O/EtOH, 1:1, v/v and the filtrate was concentrated. The crude peptidylated oligo nucleotide was dissolved in DMSO (150 μ L) and treated with TEA·HF (850 μ L) for 16 hours at 50 °C. Precipitation from cold *n*-BuOH and HPLC purification resulted in the title compound **16** (0.85 mg, 0.25 μ mol, 2.5 %). LC-MS (0 - 25 % MeCN in 10 mM NH₄OAc in 15 min), Rt = 5.70. ESI-MS m/z : 1126.1 [M + H]⁺. ³¹P-NMR (161 MHz, CDCl₃); δ -0.46, -0.51, -0.65, -0.69, -0.85, -4.42. HRMS [C₁₀₇H₁₃₆N₄₃O₆₇P₉ + H]⁺: found 3374.6152, calc. 3374.6267, [C₁₀₇H₁₃₆N₄₃O₆₇P₉ + 2H]²⁺: found 1687.8194, calc. 1687.8174, [C₁₀₇H₁₃₆N₄₃O₆₇P₉ + 3H]³⁺: found 1125.5483, calc. 1125.5471.

and side-product (30)

(1.1 mg, 0.32 μ mol, 3.2 %). LC-MS (0 - 25 % MeCN in 10 mM NH₄OAc in 15 min), Rt = 5.53. ESI-MS m/z : 1140.4 [M + H]⁺. ³¹P-NMR (161 MHz, CDCl₃); δ -0.33, -0.46, -0.54, -0.64, -0.81, -4.42. HRMS [C₁₀₉H₁₃₈N₄₃O₆₈P₉ + H]⁺: found 3416.6242, calc. 3416.6374, [C₁₀₉H₁₃₈N₄₃O₆₈P₉ + 2H]²⁺: found 1708.8249, calc. 1708.8223, [C₁₀₉H₁₃₈N₄₃O₆₈P₉ + 3H]³⁺: found 1139.5520, calc. 1139.5506.

References and Notes

- ¹ Kuyl-Yeheskiely, E.; van der Klein, P. A. M.; Visser, G. M.; van der Marel, G. A.; van Boom, J. H.; *Recl. Trav. Chim. Pays-Bas.*, **1986**, *105*, 67 – 70.
- ² Suzuki, T.; Ohsumi, S.; Makino, K.; *Nucleic Acids Res.*, **1996**, *22*, 4997 – 5003.
- ³ Reese, C. B.; *Tetrahedron*, **1978**, *34*, 3143 – 3179.
- ⁴ Debethune, L.; Marchan, V.; Fabregas, G.; Pedroso, E.; Grandas, A.; *Tetrahedron*, **2002**, *58*, 6965 – 6978.
- ⁵ Paul, A. V.; van Boom, J. H.; Filippov, D. V.; Wimmer, E.; *Nature*, **1998**, *393*, 280 – 284.
- ⁶ Schein, C. H.; Oezguen, N.; van der Heden van Noort, G. J.; Filippov, D. V.; Paul, A.; Kumar, E.; Braun, W.; *Peptides*, **2010**, *31*, 1441– 1448.
- ⁷ Filippov, D. V.; Kuyl-Yeheskiely, E.; van der Marel, G. A.; Tesser, G. I.; van Boom, J. H.; *Tetrahedron Lett.*, **1998**, *39*, 3597 – 3600.
- ⁸ Kriek, N. M. A. J.; Meeuwenoord, N. J.; van den Elst, H.; Heus, H. A.; van der Marel, G. A.; Filippov, D. V.; *Org. Biomol. Chem.*, **2006**, *59*, 3576 – 3586.
- ⁹ Dreef-Tromp, C. M.; van den Elst, H.; van den Boogaart, J. E.; van der Marel, G. A.; van Boom, J. H.; *Nucl. Ac. Res.*, **1992**, *20*, 2435 – 2439.
- ¹⁰ Dreef-Tromp, C. M.; van der Maarel, J. E.; van den Elst, H.; van den Boogaart, J. E.; van der Marel, G. A.; van Boom, J. H.; *Nucl. Ac. Res.*, **1992**, *20*, 4015 – 4020.
- ¹¹ Dreef-Tromp, C. M.; van Dam, E. M. A.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H.; *Nucl. Ac. Res.*, **1990**, *18*, 6491 – 6495.
- ¹² Robles, J.; Pedroso, E.; Grandas, A.; *Tetrahedron Lett.*, **1991**, *32*, 4389 – 4392.
- ¹³ Robles, J.; Pedroso, E.; Grandas, A.; *Tetrahedron Lett.*, **1994**, *35*, 4449 – 4452.
- ¹⁴ Robles, J.; Pedroso, E.; Grandas, A.; *Nucleic Acids Res.*, **1995**, *23*, 4151 – 4161.
- ¹⁵ Debethune, L.; Marchan, V.; Beltran, M.; Fabregas, G.; Pedroso, E.; Grandas, A.; *Tetrahedron*, **2002**, *58*, 6965 – 6978.
- ¹⁶ The in **Chapter 3** used method to cleave the ester linkage between peptides and the HMBA linker using NH₃ in MeOH, resulted in formation of a mixture of the C-terminal carboxamide and C-terminal methyl ester. Adapting this method, and using NH₃ in trifluoroethanol only gives rise to formation of the C-terminal carboxamide.

Chapter 8

Summary and Future Prospects

This thesis describes the development of new methodologies directed to the syntheses of fragments of naturally occurring hybrid biomolecules composed of a peptide and nucleic acid. The two classes of peptide nucleic acid hybrids that are pursued are the class of nucleoproteins, which is found to play crucial roles in viral replication pathways and DNA replicating machinery in higher organisms, and the class of ADP-ribosylated proteins, which is a result of the complicated and non-understood post-translational modification.

In **Chapter 1** an overview is given of known methods to prepare naturally occurring hybrid biomolecules, consisting of a peptide fragment and a nucleic acid fragment. Emphasis is placed on the way the crucial linkage between the peptide and nucleic acid fragment is constructed.

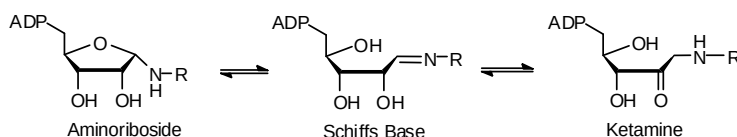
In **Chapter 2** the scope of a one-pot phosphorylation procedure using a phosphoramidite equipped with two acid labile *p*-methoxybenzyl groups is described. Using this reagent a wide variety of phosphoric acid analogues can be synthesized, such as phosphomonoesters, phosphorothioates, phosphoramidates, phosphofluoridates, phosphodiester and pyrophosphates. An Arbuzov-like reaction mechanism is proposed to explain the formation of a reactive intermediate that allows introduction of several nucleophiles.

In **Chapter 3** a method to introduce a ribosyl moiety on the side chain of glutamine and asparagine is reported. Reduction of a ribofuranose carrying an anomeric azide followed by a peptide coupling of the formed ribosylamine to the side chain carboxylic acid of the appropriate amino acid resulted in the target modified amino acids. Subsequently, these ribosylated amino acids were incorporated in small, biological relevant peptides. The methodology described in **Chapter 2** was used to obtain the required pyrophosphate linkage to adenosine. These first examples of synthetically made ADP-ribosylated peptides open the way to more complex structures.

Other sites of ADP-ribosylation have been identified, where the nucleophilic side chains of lysine, arginine or cysteine residues form the linkage with the anomeric centre of ADP-ribose. The development of synthetic routes to the ribosylated analogues of these amino acids and their

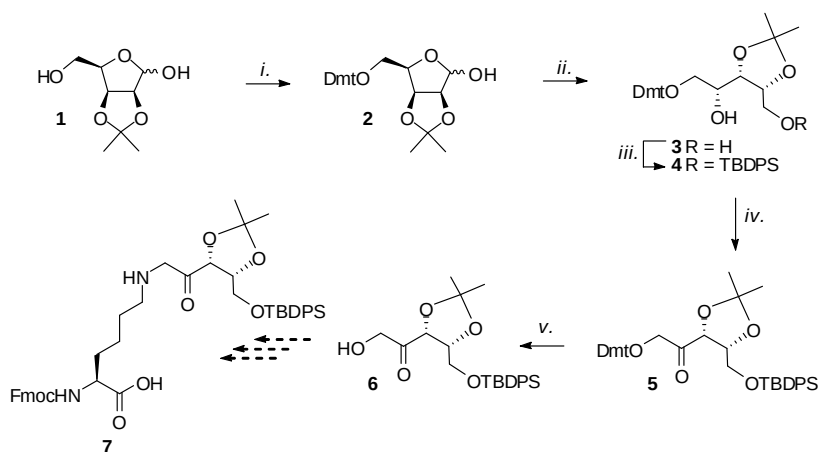
subsequent incorporation in ADP-ribosyl peptides can be attained by adaption of the methodology from **Chapter 3**. Hottiger and coworkers have identified lysine residues to be the site of ADP-ribosylation in the auto-ADP-ribosylation of the PARP-1 enzyme.¹ They have proposed a reaction mechanism in which the initial aminoriboside, formed by enzyme catalyzed condensation of the lysine ϵ -amine with NAD⁺, forms a Schiff's base, which then undergoes an Amadori rearrangement to give the proposed ketamine linkage between lysine and ribose (see **Figure 1**).

Figure 1



The synthesis of ADP-ribosylated peptides containing such a linkage raises the need for a different strategy towards a suitably protected lysine-ribosyl ketamine (such as **7**, see **Scheme 1**). In order to secure the correct stereochemistry in **7**, D-lyxose derivative **1** was used to synthesize β -hydroxyketone **6**. In this approach orthogonal protection of the lyxose 1' and 5' hydroxyl functions, followed by opening of the furanose ring and oxidation of the liberated hydroxyl eventually can give rise to **6**.

Scheme 1

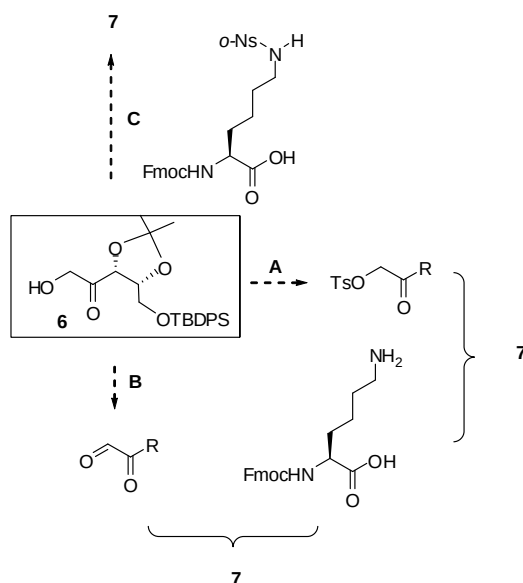


Reagents and conditions: *i.* DMTCl, pyridine, 72 h, r.t., (80%), *ii.* NaBH₄, EtOH, 16 h, r.t., (68%), *iii.* TBDPSCl, pyr., 16 h, r.t., (99 %), *iv.* DCC, TFA, DCM, 16 h, r.t., (85 %), *v.* TFA, pyrrole, 15 min, r.t., (68 %).

Several methods to link a lysine derivative to β -hydroxyketone **6** to obtain **7** can be envisaged (see **Figure 2**). Converting the hydroxyl function in **6** into a good leaving group such as a tosylate followed by nucleophilic substitution of the ϵ -amine of lysine may render **7** (Pathway A). Partial oxidation of the hydroxyl function in **6** to a glyoxal moiety followed by reductive

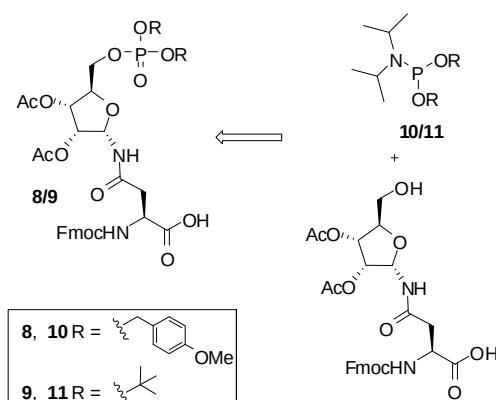
amination to the ϵ -amine of lysine is also an option (Pathway B). Alternatively a Mitsunobu reaction using ϵ -nosyl protected lysine and alcohol **6** can be explored (Pathway C).

Figure 2



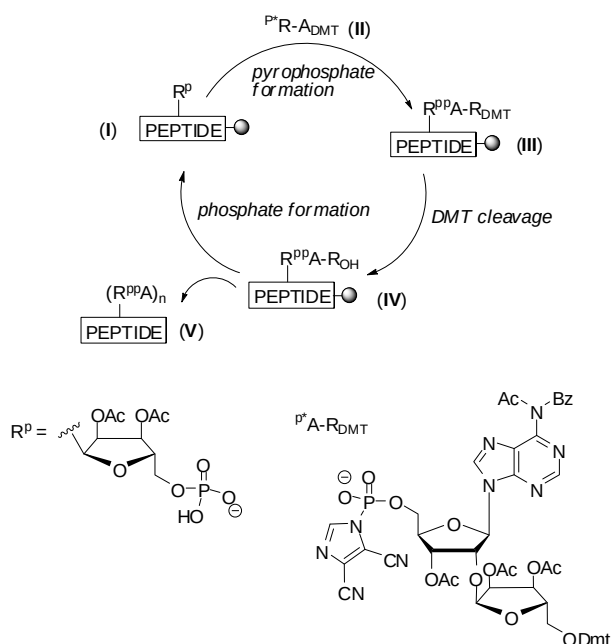
Another point of attention is the moderate yield of the introduction of the phosphate monoester on resin bound ribosylated peptides as shown in **Chapter 3**. Approaches using harsh conditions to introduce the phosphate monoester were unsuccessful, and a new approach has to be developed. On resin phosphorylation can be circumvented by the design and synthesis of an amino acid equipped with a phosphotriester (such as **8** or **9**, see **Figure 3**). Subsequent application of phosphorylated amino acid building blocks in a solid phase synthesis followed by mild removal of the protective groups is proposed to allow an efficient incorporation of the phosphate monoester. Since solid phase peptide synthesis, using the Fmoc-group is projected, acid-labile protection for the phosphate group is required. Compound **8** can be prepared using phosphoramidite **10**, described in **Chapter 2**. However, the *p*-methoxybenzyl groups are prone to undergo nucleophilic displacement and may therefore be too labile to function as semi-permanent protection. An alternative can be the use of *tert*-butyl protection as indicated for **9** and **11**.²

Figure 3



Chapter 4 deals with the formation of an α -glycosidic linkage between ribofuranose and adenosine. The choice of benzyl ethers as protective groups on the ribofuranose donor not only steers the α -selectivity in the glycosylation reaction, but also allows the subsequent introduction of orthogonal protective groups on both primary hydroxyls functions in the glycosylation product. Such a protective group pattern is a prerequisite for sequential elongation of a *poly*-ADP-ribosyl chain using the methods described in **Chapter 2** and **Chapter 3** to give access to *poly*-ADP-ribosyl peptides (see **Figure 4**). Reaction of an immobilized peptide anchored to a solid support and equipped with 5'-phosphorylribose (I), with the reactive imidazolide species of suitably protected ribosyl-adenosine (II) will give a pyrophosphate linkage (III). After removing the DMT protection on the newly introduced ribose group (IV) and phosphorylation of the liberated alcohol a new cycle can start. After a number of repetitions of the coupling cycle the elongation will be ended (step IV) by using the imidazolide of adenosine instead of ribosyl-adenosine. Finally, removal of the protecting groups and cleavage from the solid support will result in the *poly*-ADP-ribosyl peptide (V).

Figure 4



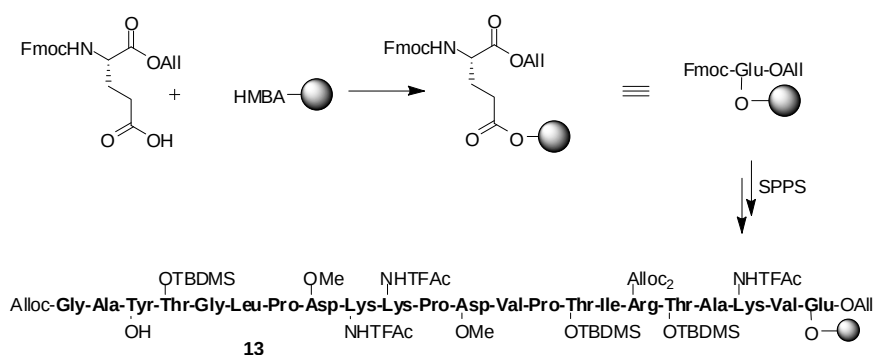
In **Chapter 5** the synthesis of a protected pre-uridylylated tyrosine building block is described. This building block is used in a solid phase peptide synthesis protocol to give several naturally occurring nucleoproteins, originating from mammalian pathogens. These compounds are presently investigated to gain insight in the replication mechanisms of picornaviruses.

Chapter 6 reports the synthesis of three protected ribonucleotidylated tyrosine building blocks. These adenylylated, guanylylated and cytidylylated tyrosine derivatives are incorporated in the VPg protein of polio virus using the solid phase synthesis protocol as described in **Chapter 5**.

Commonly, oligonucleotides and nucleopeptide fragments are prepared by the use of 3'-phosphoramidite reagents. **Chapter 7** reports an approach to create nucleopeptide fragments by elongation of the (oligo)nucleotide from the 5'-end towards the 3'-end using 5'-phosphoramidite RNA building blocks. Immobilized oligopeptides, prepared by solid phase peptide synthesis, were extended to give nucleopeptides of unprecedented length with the aid of these 5'-phosphoramidites.

Having established that phosphitylation using 5'-phosphoramidites is a viable entry to sequentially synthesize oligo-ribonucleotides on a resin bound peptide, the application of this approach to construct an oligoribonucleotide on the full length VPg protein is the next challenge. Since standard protective groups on amino acid residues require harsh acidic deprotection conditions which are incompatible with the stability of the RNA part of the nucleopeptide, the protective group pattern of the amino acid building blocks have to be

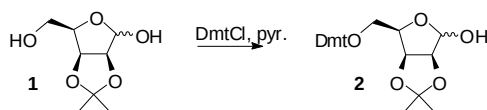
adapted. For this reason, the lysine residues in the VPg proteins are envisaged to be protected with the trifluoroacetyl protective group (see **Figure 5**), which will be simultaneously cleaved with the acyl protective groups on the nucleobases using NH_3 treatment. Threonine hydroxyl functions are envisaged to be protected using the TBDMS group, which will be cleaved in the final $\text{TEA} \cdot \text{HF}$ treatment of the nucleoprotein. Introduction of aspartic acid β -methylesters may lead to formation of the desired asparagine residues in the VPg protein during the NH_3 treatment. Although aspartic esters are notorious for aspartimide formation, this side reaction is highly sequence dependent. It has been shown that the -Asp(OX)-Val- and -Asp(OX)-Lys- sequences, which are occurring in the polio virus VPg proteins, are not highly susceptible to aspartimide formation in contrast to for instance -Asp(OX)-Gly- or -Asp(OX)-Ala-.^{3,4}

Figure 5

As the results on nucleopeptide synthesis from **Chapter 7** indicate that capping during the RNA synthesis causes cleavage of Fmoc and subsequent acetylation of the free amine, an alternative protective group for the N-terminus is desired. Taking into account that in the polio virus VPg an arginine residue is present, and non-acid labile protective groups for this residue are limited, a preliminary study was executed using the Alloc protective group both on the N-terminus and arginine side chain and the allyl ester as C-terminal protective group. Previous results indicate that removal of Alloc/allyl protection is compatible with the integrity of the VPgUpU nucleoprotein.⁵ In this study, the not routinely used HMBA-HCP resin was selected for the construction of fully protected VPg protein **13** (see **Figure 5**). Deprotection of an aliquot of **13** to determine the quality of the synthesis was achieved by Pd catalyzed removal of Alloc and allyl, subsequent NH_3 treatment to remove acyl protection and cleave the peptide from the resin and finally fluoride treatment to remove silyl ethers. LCMS analysis revealed that the molecular mass of the major product formed corresponds to the fully deprotected polio virus VPg protein. Encouraged by these results future RNA oligo-nucleotide synthesis using the 5'-phosphoramidite approach on protected resin bound VPg protein **13** is a potential approach to oligonucleotidylated VPg nucleoproteins.

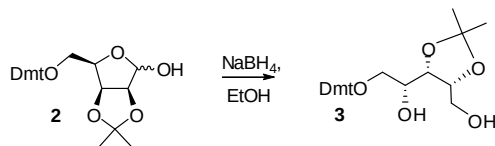
Experimental

5'-O-DMT-2,3-O-isopropylidene- α/β -D-lyxofuranose (2)



2,3-O-isopropyl- α/β -D-lyxofuranose (7.4 g, 39 mmol) was dissolved in pyridine (80 mL) and DMTCl (15.6 g, 46 mmol) was added. The reaction mixture was stirred for 72 hours at room temperature. Concentration of the crude and purification using silica gel column chromatography (EtOAc/PE; 0:100 – 30:70) afforded the title compound (15.3 g, 31 mmol, 80%). $^1\text{H-NMR}$ (400 MHz, CDCl_3); δ 7.49 (d, J = 7.5 Hz, 2H), 7.37 (d, J = 8.7 Hz, 4H), 7.26 (t, J = 7.6 Hz, 3H), 7.17 (dd, J = 8.4, 6.1 Hz, 2H), 6.81 (d, J = 8.1 Hz, 5H), 5.37 (s, 1H, H1'), 4.95 (d, J = 7.6 Hz, 0.1H, H1'), 4.71 (dd, J = 5.8, 3.8 Hz, 1H, H3'), 4.67 (dd, J = 6.0, 3.3 Hz, 0.1H, H3'), 4.60 (d, J = 5.9 Hz, 0.1H, H2'), 4.56 (d, J = 5.9 Hz, 1H, H2'), 4.45 (dd, J = 6.0, 3.6 Hz, 0.1H, H4'), 4.34 (dd, J = 9.5, 5.7 Hz, 1H, H4'), 3.75 (s, 6H, OMe), 3.40 (d, J = 5.9 Hz, 2H, H5'), 3.24 (s, 1H, OH), 1.40 (s, 0.3, *i*Pr), 1.33 (s, 0.3H, *i*Pr), 1.30 (s, 3H, *i*Pr), 1.25 (s, 3H, *i*Pr). $^{13}\text{C-NMR}$ (101 MHz, CDCl_3); δ 158.4 (Cq DMT), 145.0 (Cq DMT), 136.3 (Cq DMT), 136.2, 130.28, 130.26, 128.4, 127.8, 126.73, 113.1, 112.5 (Cq *i*Pr), 101.2 (C1'), 96.8 (C1'A), 86.4 (Cq DMT), 85.5 (C2'), 80.2 (C3'), 79.80, 79.7 (C4'), 78.7, 75.3, 61.8 (C5'), 61.1 (C5'A), 55.3 (OMe), 26.1 (*i*Pr), 25.9 (*i*Pr), 25.3 (*i*Pr), 25.1 (*i*Pr).

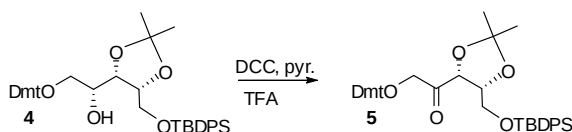
5'-O-DMT-2,3-O-isopropylidene-D-lyxitol (3)



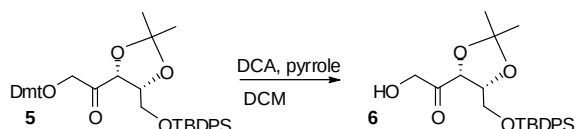
Compound **2** (6.15 g, 12.5 mmol) was dissolved in EtOH (100 mL) at 0 °C. NaBH_4 (1.2 g, 32 mmol) was slowly added. The reaction mixture was stirred for 16 hours at room temperature, upon which TLC analysis (EtOAc/PE, 70:30) revealed complete consumption of the starting material. The reaction mixture was quenched using AcOH until pH = 7.5. The reaction mixture was concentrated and the crude redissolved in DCM and washed with aq. NaHCO_3 . The organic layer was dried (MgSO_4) and concentrated. Silica gel column chromatography (EtOAc/PE; 0:100 – 50:50) using a pre-neutralized column (1% TEA in PE) afforded the title compound (4.2 g, 8.5 mmol, 68%). $^1\text{H-NMR}$ (400 MHz, CDCl_3); δ 7.44 (d, J = 7.9 Hz, 2H), 7.33 (d, J = 8.8 Hz, 4H), 7.26 (t, J = 7.6 Hz, 2H), 7.18 (t, J = 7.2 Hz, 1H), 6.81 (d, J = 8.9 Hz, 4H), 4.30 (dd, J = 6.9, 2.0 Hz, 1H, H3'), 4.19 (dd, J = 11.6, 4.9 Hz, 1H, H2'), 3.91 - 3.77 (m, 1H, H4'), 3.80 - 3.65 (m, 8H, H1', OMe), 3.27 - 3.23 (m, 2H, H5'), 3.17 (s, 2H, OH), 1.46 (s, 3H, *i*Pr), 1.36 (s, 3H, *i*Pr). $^{13}\text{C-NMR}$ (101 MHz, CDCl_3); δ 158.4 (Cq DMT), 144.7 (Cq DMT), 135.9 (Cq DMT), 135.9, 130.02, 130.00, 128.1, 127.8, 126.8, 113.1, 108.2 (Cq *i*Pr), 86.3 (Cq DMT), 77.3 (C2'), 76.2 (C3'), 68.2 (C4'), 65.0 (C5'), 61.2 (C1'), 55.1 (OMe DMT), 27.0, 25.0 (*i*Pr). IR: 3420, 2934, 1608, 1508, 1247, 1174, 1031, 827. HRMS [$\text{C}_{29}\text{H}_{34}\text{O}_7 + \text{Na}$] $^+$: found 517.2194, calc. 517.2196.

5'-O-DMT-1'-O-*tert*-butyl diphenyl silyl-2,3-O-isopropylidene- D-lyxitol (4)

Compound 3 (8.5 mmol) was coevaporated with pyridine and subsequently dissolved in pyridine (50 mL). After addition of TBDPSCl (9.3 mmol, 2.4 mL) the reaction mixture was stirred at room temperature for 16 hours. The reaction mixture was concentrated and purification using silica gel column chromatography (EtOAc/PE; 0:100 – 5:95) afforded the title compound (6.2 g, 8.4 mmol, 99%). ¹H-NMR (400 MHz, CDCl₃); δ 7.67 (dd, *J* = 10.7, 4.1 Hz, 4H), 7.45 (d, *J* = 7.4 Hz, 2H), 7.41 - 7.33 (m, 9H), 7.26 (t, *J* = 7.5 Hz, 2H), 7.22 - 7.13 (m, 2H), 6.80 (d, *J* = 8.8 Hz, 4H), 4.44 (dd, *J* = 6.6, 2.4 Hz, 1H, H_{3'}), 4.25 (dd, *J* = 11.6, 6.7 Hz, 1H, H_{2'}), 4.10 (dd, *J* = 6.2, 4.4 Hz, 1H, H_{4'}), 4.00 (dd, *J* = 10.7, 7.1 Hz, 1H, H_{1'}), 3.82 - 3.68 (m, 7H, H_{1'}, OMe), 3.40 - 3.17 (m, 2H, H_{5'}), 2.75 (d, *J* = 6.8 Hz, 1H, OH), 1.42 (s, 3H, *i*Pr), 1.36 (s, 3H, *i*Pr), 1.05 (s, 9H, *t*Bu). ¹³C-NMR (101 MHz, CDCl₃); δ 158.5 (Cq DMT), 145.0 (Cq DMT), 136.23, 136.19 (Cq DMT), 135.7 (Arom. TBDPS), 133.1, 133.0 (Cq DMT), 130.19, 130.17, 129.91, 129.89, 128.3 (Arom. DMT), 127.9 (Arom. TBDPS), 126.8 (Arom. DMT), 113.2 (Arom. DMT), 108.3 (Cq DMT), 86.2 (Cq DMT), 77.2 (C_{2'}), 76.6 (C_{3'}), 68.2 (C_{4'}), 65.0 (C_{5'}), 63.0 (C_{1'}), 55.3 (OMe), 27.2 (*t*Bu), 27.0, 25.1 (*i*Pr), 19.3 (Cq *t*Bu). IR: 2933, 1608, 1508, 1248, 1112, 1034, 825. HRMS [C₄₅H₅₂O₇Si + Na]⁺: found 755.3379, calc. 755.3374.

1'-O-DMT-5'-O-*tert*-butyl diphenyl silyl-3,4-O-isopropylidene-D-ribulose (5)

Compound 4 (1.1 g, 1.5 mmol) was dissolved in DMSO (7 mL) and dicyclohexylcarbodiimide (4.2 mmol, 870 mg), pyridine (1.4 mmol, 115 μL) and trifluoroacetic acid (0.7 mmol, 52 μL) were added. After stirring at room temperature for 16 hours the reaction mixture was diluted in EtOAc (50 mL) and water (10 mL). After stirring for 10 min. the organic layer was washed with water five times, dried (MgSO₄) and concentrated. NMR-spectroscopy revealed circa 50% conversion, so the procedure was repeated once more. Final purification using silica gel column chromatography (EtOAc/PE; 0:100 – 5:95) afforded the title compound (0.9 g, 1.3 mmol, 85%). ¹H-NMR (400 MHz, CDCl₃); δ 7.73 - 7.55 (m, 4H), 7.46 - 7.39 (m, 4H), 7.36 (dd, *J* = 11.5, 5.0 Hz, 4H), 7.33 - 7.27 (m, 4H), 7.27 - 7.14 (m, 3H), 6.83 - 6.60 (m, 4H), 4.59 (d, *J* = 7.6 Hz, 1H, H_{3'}), 4.47 (dt, *J* = 7.6, 4.7 Hz, 1H, H_{2'}), 4.07 (m, 2H, H_{5'}), 3.85 - 3.69 (m, 7H, 1H', OMe), 3.61 (dd, *J* = 11.2, 4.8 Hz, 1H, H_{1'}), 1.38 (s, 3H, *i*Pr), 1.27 (s, 3H, *i*Pr), 0.97 (s, 9H, *t*Bu). ¹³C-NMR (101 MHz, CDCl₃); δ 204.6 (CO), 158.7 (Cq DMT), 144.4 (Cq DMT), 135.77, 135.76 (Arom. TBDPS), 135.61, 135.59 (Cq DMT), 133.2 (Cq), 130.13, 130.11, 129.79, 129.75, 128.2, 128.1, 127.79, 127.76, 127.1 (Arom. DMT/TBDPS), 113.4 (Arom. DMT), 109.7 (Cq *i*Pr), 87.2 (Cq DMT), 79.8 (C_{3'}), 78.5 (C_{2'}), 69.3 (C_{5'}), 62.3 (C_{1'}), 55.3 (OMe), 26.9 (*t*Bu), 26.8, 24.9 (*i*Pr), 19.3 (Cq *t*Bu). IR: 2932, 2857, 2117, 1733, 1608, 1508, 1249, 1175, 1112, 1033, 825. HRMS [C₄₅H₅₀O₇Si + Na]⁺: found 753.3218, calc. 753.3223.

5'-O-tert-butyl diphenyl silyl-3,4-O-isopropylidene-D-ribose (6)

Compound **5** (1.1 mmol) was dissolved in DCM (15 mL) at 0 °C and treated with pyrrole (700 µL) and dichloroacetic acid (300 µL). After 15 minutes the reaction mixture was washed with sat. aq. NaHCO₃, dried (MgSO₄) and concentrated. Purification using silica gel column chromatography (EtOAc/PE; 0:100 – 10:90) afforded the title compound (0.3 g, 0.8 mmol, 68%). ¹H-NMR (400 MHz, CDCl₃); δ 7.85 - 7.68 (m, 2H), 7.64 (dd, *J* = 7.6, 1.6 Hz, 2H), 7.50 - 7.33 (m, 6H), 4.67 (d, *J* = 8.3 Hz, 1H, H3'), 4.56 (m, 2H, H5'), 4.44 (dd, *J* = 8.2, 3.5 Hz, 1H, H2'), 3.76 (m, 2H, H1'), 2.98 (s, 1H, OH), 1.59 (s, 3H, *i*Pr), 1.36 (s, 3H, *i*Pr), 1.02 (s, 9H, *t*Bu). ¹³C-NMR (101 MHz, CDCl₃); δ 209.7 (CO), 135.84, 135.75, 135.65 (Arom. TBDPS), 133.0, 132.8 (Cq TBDPS), 129.92, 129.89, 127.9, 127.8 (Arom. TBDPS), 110.1 (Cq *i*Pr), 79.5 (C3'), 78.8 (C2'), 67.5 (C5'), 61.4 (C1'), 26.8 (*t*Bu), 26.5, 24.4 (*i*Pr), 19.2 (Cq *t*Bu).

References

- ¹ Altmeyer, M.; Messner, S.; Hasso, P. O.; Fey, M.; Hottiger, M. O.; *Nucleic Acids Res.*, **2009**, *37*, 3723 – 3738.
- ² Gomez, C.; Chen, J.; Wang, S.; *Tetrahedron Lett.*, **2009**, *50*, 6691 – 6692.
- ³ Stathopoulos, P.; Papas, S.; Kostidis, S.; Tsikaris, V.; *J. Pept. Sci.*, **2005**, *11*, 658 – 664
- ⁴ Lapeyre, M.; Leprince, J.; Massonneau, M.; Oulyadi, H.; Renard, P.; Romieu, A.; Turcatti, G.; Vaudry, H.; *Chem. Eur. J.*, **2006**, *12*, 3655 – 3671.
- ⁵ Kriek, N. M. A. J.; Meeuwenoord, N. J.; van den Elst, H.; Heus, H. A.; van der Marel, G. A.; Filippov, D. V.; *Org. Biomol. Chem.*, **2006**, *59*, 3576 – 3586.

Nederlandse samenvatting

Dit proefschrift beschrijft de ontwikkeling van nieuwe methoden voor de synthese van hybride biomoleculen die samengesteld zijn uit een peptide- en een nucleïnezuurfragment. Zulke hybride moleculen komen in de natuur voor en hebben belangrijke functies. In dit proefschrift wordt aandacht besteed aan twee klassen van hybride biopolymeren; de klasse van nucleoproteïnen, die een belangrijke rol speelt in virale replicatie processen en de klasse van ADP-geribosyleerde eiwitten, die het resultaat zijn van een post-translationeel modificatie proces.

In **Hoofdstuk 1** wordt een overzicht gegeven van bestaande synthetische methoden, die toegang geven tot natuurlijk voorkomende nucleïnezuur bevattende peptiden. Speciale aandacht wordt besteed aan de methoden waarmee de cruciale en labiele binding tussen het nucleïnezuur-fragment en het peptide kan worden geïntroduceerd.

In **Hoofdstuk 2** worden de mogelijkheden beschreven van een eenpots fosforyleringsprocedure, die gebaseerd is op het gebruik van een fosforamidiet voorzien van twee zuur-labele *p*-methoxybenzyl groepen. Met behulp van dit fosforamidiet is het mogelijk gebleken een aantal fosforzuur analoga te maken, zoals fosfaten, fosforothioaten, fosforamidaten, fosforofluoridaten, fosfodiësters en pyrofosfaten. Er wordt een mechanisme voor de vorming van een reactief intermediair gepostuleerd dat de introductie van verschillende nucleofielen mogelijk maakt en de vorming van fosforzuur analoga verklaard.

In **Hoofdstuk 3** wordt een methode beschreven om een ribosyl groep aan te brengen op de zijketen van de aminozuren glutamine en asparagine. Een azide, aangebracht op de anomere positie van de ribofuranose, werd gereduceerd en vervolgens werd een peptide koppeling uitgevoerd tussen het gevormde amine en de zuurgroep in de zijketen van glutaminezuur of asparaginezuur. De verkregen gemodificeerde aminozuren werden ingebouwd in verschillende biologisch relevante peptiden met behulp van vaste drager chemie. De geribosyleerde peptiden konden worden verlengd met een pyrofosfaat binding naar adenosine door toepassing van de fosforyleringsmethode uit **Hoofdstuk 2**. De aldus ontwikkelde methodologie naar ge-ADP-ribosyleerde peptiden opent de weg naar meer complexe structuren.

Hoofdstuk 4 omvat een studie naar het introduceren van een α -glycosidische binding tussen het anomere centrum van ribofuranose en de 2'-hydroxyl van adenosine. Geribosyleerd adenosine is de repeterende eenheid die wordt gevonden in ge-*poly*-ADP-ribosyleerde eiwitten. De keuze van benzylethers als beschermgroep op de ribofuranose donor bepaalde niet alleen de α -selectiviteit van de glycosyleringsreactie maar maakte ook de introductie van orthogonale beschermgroepen op de beide primaire hydroxylgroepen in het reactieproduct mogelijk. De aldus vervaardigde bouwsteen kan in de toekomst gebruik gaan worden voor de synthese van ge-*poly*-ADP-ribosyleerde peptiden.

Hoofdstuk 5 behandelt de synthese van een geuridylyleerde tyrosine bouwsteen. Deze bouwsteen werd vervolgens gebruikt in de vaste drager synthese van natuurlijk voorkomende

genoomeiwitten van de zoogdier-pathogenen; poliovirus type 1, coxsackievirus type A, coxsackievirus type B en humaan enterovirus type 71. Deze gemodificeerde eiwitten zijn mogelijk belangrijk gereedschap voor de biologie om inzicht te verkrijgen in de replicatiemechanismen van picornavirussen.

In **Hoofdstuk 6** wordt de synthese beschreven van de overige drie ribonucleotide derivaten van tyrosine. Deze geadenylyleerde, geguanylyleerde en gecytidylyleerde tyrosine bouwstenen werden vervolgens via een vaste drager synthese ingebouwd in het genoomeiwit VPg van polio virus door gebruik te maken van de methodes die in **Hoofdstuk 5** wordt beschreven. De zure condities nodig om de beschermgroepen te verwijderen en het eiwit van de vaste drager af te splitsen bleken in het geval van adenosine ook depurineren te veroorzaken en daarmee voor een lagere opbrengst te zorgen.

Hoofdstuk 7 behandelt de synthese van nucleïnezuur bevattende peptide fragmenten op de vaste drager met behulp van 5'-fosforamidiëten. De RNA fragmenten die met deze amiditen kunnen worden bereid, worden in de ongebruikelijke 5' naar 3'-richting verlengd. Geïmmobiliseerde peptiden, gesynthetiseerd door middel van een vaste drager synthese, werden met behulp van de 5'-fosforamidiëten voorzien van een RNA-fragment. Op deze manier werden een pentapeptidyl tetranucleotide en een pentapeptidyl nonanucleotide vervaardigd. Deze synthetische RNA-peptide hybriden hebben een tot nu toe ongeëvenaarde lengte.

List of Publications

Ribosylation of Adenosine: An Orthogonally Protected Building Block for the Synthesis of ADP-Ribosyl Oligomers

van der Heden van Noort, G.J.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D.V.; *Organic Letters*, **2011**, 13, 2920 - 2923.

Synthesis of Nucleotidylated Polio Virus VPg Proteins

Van der Heden van Noort, G.J.; Overkleeft, H.S.; van der Marel, G.A.; Filippov, D.V.; *Journal of Organic Chemistry*, **2010**, 75, 5733 - 5736.

NMR solution structure of poliovirus uridylylated peptide linked to the genome (VPgpU)

Schein, C. H.; Oezguen, N.; Van der Heden van Noort, G. J.; Filippov, D. V.; Paul, A.; Kumar, E.; Braun, W.; *Peptides*, **2010**, 31, 1441 - 1448.

Synthesis of mono-ADP-ribosylated oligopeptides using ribosylated amino acid building blocks

Van der Heden van Noort, G.J.; van der Horst, M.G.; Overkleeft, H.S.; van der Marel, G.A.; Filippov, D.V.; *Journal of the American Chemical Society*, **2010**, 132, 5236 - 5240.

2-Azidoalkoxy-7-hydro-8-oxoadenine derivatives as TLR7 agonists inducing dendritic cell maturation

Weterings, J. J.; Khan, S.; van der Heden van Noort, G.J.; Melief, C.J.M.; Overkleeft, H.S.; van der Burg, S.H.; Ossendorp, F.; van der Marel, G.A.; Filippov, D.V.; *Bioorganic & Medicinal Chemistry Letters*, **2009**, 19, 2249 - 2251.

A Versatile One-Pot Procedure to Phosphate Monoesters and Pyrophosphates Using Di(p-methoxybenzyl)-N,N-diisopropylphosphoramidite

Van der Heden van Noort, G.J.; Verhagen, C.P.; van der Horst, M.G.; Overkleeft, H.S.; van der Marel, G.A.; Filippov, D.V.; *Organic Letters*, **2008**, *10*, 4461 - 4464.

Synthesis of 2-alkoxy-8-hydroxyadenylpeptides: Towards synthetic epitope-based vaccines

Weterings, J. J.; Khan, S.; van der Heden, G.J.; Drijfhout, J.W.; Melief, C.J.M.; Overkleeft, H.S.; van der Burg, O.H.; Ossendorp, F.; van der Marel, G.A.; Filippov, D.V.; *Bioorganic & Medicinal Chemistry Letters*, **2006**, *16*, 3258 - 3261.

Curriculum vitae in het Nederlands

Gerbrand Jan van der Heden werd op 21 januari 1984 geboren te Alkemade. Nadat in 2002 het VWO diploma aan het Herbert Vissers College te Nieuw Vennep met het profiel 'Natuur en Gezondheid' werd behaald, werd in hetzelfde jaar begonnen met de bachelor opleiding Scheikunde aan de Universiteit Leiden. Het propedeutisch examen werd in augustus 2003 afgelegd en de bachelor opleiding werd in oktober 2005 afgerond met een onderzoeksstage met als onderwerp 'Sugar amino acids' in de vakgroep Bio Organische Synthese onder leiding van Prof. H. S. Overkleeft en Prof. Dr. G. A. van der Marel. Aansluitend werd begonnen met de master opleiding Chemistry met als specialisatie 'Design & Synthesis' aan de Universiteit Leiden. De hoofdvakstage met als onderwerp 'Toll Like Receptor Ligands' werd uitgevoerd in de vakgroep Bio Organische Synthese onder leiding van Prof. H. S. Overkleeft en Prof. Dr. G. A. van der Marel. De bijvakstage werd verricht aan het Leids Universitair Medisch Centrum in de vakgroep Immunohematologie en bloedtransfusie onder begeleiding van Prof. Dr. F. Ossendorp en Dr. S. Khan. Het master diploma werd behaald in augustus 2007.

Van juli 2007 tot september 2011 werd als assistent in opleiding het in dit proefschrift beschreven onderzoek verricht in de vakgroep Bio Organische Synthese aan de Universiteit Leiden onder de dagelijkse supervisie van Dr. D. V. Filippov. Gedeelten van dit onderzoek werden mondeling gepresenteerd op het 'Dutch Peptide Symposium' (de Bilt, Nederland, maart 2009) en in de vorm van poster presentaties op de jaarlijkse NWO-CW 'Design and Synthesis meetings' te Lunteren. Tevens werd in juli 2009 deelgenomen aan de HRMSC summer school 'New Horizons in Synthetic Methodolgy' te Maastricht. In december 2010 werd deelgenomen aan het '5th International Peptide Symposium' in Kyoto, Japan, waar de presentatie van de auteur werd beloond met de 'Good Stone Award' voor sprekers die een opvallende presentatie verzorgden op het 'Young Investigators Symposium'.

Vanaf oktober 2011 is de auteur van dit proefschrift werkzaam als post-doctoraal onderzoeker voor Dynomics BV, Rotterdam.

Curriculum vitae in English

Gerbrand Jan van der Heden was born on the 21st of January 1984 in Alkemade. After completing his secondary education at the Herbert Vissers College, Nieuw Vennep in 2002, he started his academic studies in Chemistry at Leiden University. The propedeutic phase was completed in August 2003 and between August 2003 and October 2005 he performed his bachelor education. The bachelor research project was carried out in the Bio Organic Synthesis group of Prof. Dr. H. S. Overkleeft and Prof. Dr. G. A. van der Marel on the topic of “Sugar amino acids”. From October 2005 until August 2007 he performed his master education with the specialization ‘Design & Synthesis’. The major research internship with the topic “Toll Like Receptor Ligands” took place in the Bio Organic Synthesis group of Prof. Dr. H. S. Overkleeft and Prof. Dr. G. A. van der Marel. His minor research project was conducted at the University of Leiden Medical Centre (LUMC) in the department of Immunohematology and Blood Transfusion under the guidance of Prof. Dr. F. Ossendorp and Dr. S. Khan.

In the period from July 2007 until September 2011 he conducted the PhD research described in this thesis in the Bio Organic Synthesis group at Leiden University under the daily guidance of Dr. D. V. Filippov. Parts of the work described herein were presented at the Dutch Peptide Symposium (De Bilt, The Netherlands, March 2009, oral presentation) and by means of poster presentations at the annual NWO-CW ‘Design and Synthesis meetings’ (Lunteren, The Netherlands). Furthermore he participated in the HRMSC summer school ‘New Horizons in Synthetic Methodology’ (Maastricht, The Netherlands, July 2009). He attended the ‘5th International Peptide Symposium’ (Kyoto, Japan, December 2010) where he was awarded the ‘Good Stone Award’ for those speakers who have made outstanding presentations at the Young Investigators' Symposium.

Since October 2011 the auteur of this dissertation is working as post-doctoral researcher for Dynomics BV, Rotterdam.

Dankwoord

Bedankt!