

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**Synthesis of hypermodified
nucleosides and examination of
nucleic acid templated stereoselective
peptide formation in an RNA-peptide
world**

Ewa Maria Węgrzyn

aus

Poznań, Polen

2024

Erklärung

Diese Dissertation wurde in Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Thomas Carell betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfsmittel erarbeitet.

München, den 02.11.2024

..... Ewa Maria Węgrzyn.....

Dissertation eingereicht am 12.09.2024

1. Gutachter Prof. Dr. Thomas Carell

2. Gutachter Dr. Pavel Kielkowski

Mündliche Prüfung am 18.10.2024

Dedicated to Babu

“It is the time you have wasted for your rose that makes your rose so important.”

- *Antoine de Saint-Exupéry, “The Little Prince”*

Acknowledgements

My PhD has been an amazing journey, that would not have been the same without the people that surrounded me during that time.

In the first place, I would like to thank my Doktorvater Prof. Dr. Thomas Carell, who has been an inspiring and supportive supervisor during the entirety of my PhD. I am highly grateful for being part of his fascinating research group and for everything that I learned from him.

I thank Dr. Pavel Kielkowski for the second evaluation of my thesis.

I thank Dr. Markus Müller for interesting discussions and helpful research inputs as well as for proofreading of my thesis. Thanks for launching the AK cinema, that brought forward many fun evenings exploring the world of cinema.

Thanks to Dr. Nädä Raddaoui for taking care of paperwork and for organising events.

I thank Ms. Slava Gärtner for being always a help in bureaucratic tasks and for letting me rent her accommodation at the start of my PhD.

I thank Kerstin, for her friendship and support, for always having an open ear for private and for work-related topics. Thank you for being our loyal lunch-break partner.

I thank the whole protoribosome subgroup, both the old cast (Dr. Luis Escobar, Dr. Felix Müller, Chun-Yin (Jamie) Chan, Alexander Pichler, Johannes Singer) and the new (Dr. Ivana Mejdrová, Kathrin Halter, Tobias Kernmayr, Dr. Giacomo Ganazzoli) for the years spent doing amazing research together and supporting each other.

I thank our Breakfast Club: Ivana Mejdrová, Elsa Peev, Nikolai Diukarev, A Hyeon Lee, Hanife Sahin and Szymon Drewniak for a good start into each working day and for making me laugh, with your great sense of humour. I am grateful for your friendship.

I thank the whole Carell group, especially my lab mates from F4.001b and from L03.012.

I thank my parents, my grandparents and my siblings, who have accompanied me throughout this journey, supported me and believed in me all my life long. Without your support, I would most likely not be where I am now. Particularly my grandma, who was waiting impatiently for this moment, and whom I would like to dedicate this work.

Last, but by no means least, I thank my wife, for being the best lab and life partner, and our son, according to whom my work is located on Neptune and I where I spend the days playing with Playmobil fire trucks. You are both always there with and for me, you are my sunshines, who always give me a sense of hope and happiness even on the difficult days. I am looking forward to the next chapter in our life and to everything what the future has planned for us.

Publications within PhD

E. Węgrzyn*; I. Mejdrová*; T. Carell, Gradual evolution of a homo-L-peptide world on homo-D-configured RNA and DNA. *Chem. Sci.* **2024**, 15, 14171-14176. DOI: <https://doi.org/10.1039/D4SC03384A>

E. Węgrzyn*; I. Mejdrová*; F. Müller; M. Nainytė; L. Escobar; T. Carell, RNA-Templated Peptide Bond Formation Promotes L-Homochirality. *Angew. Chem. Int. Ed.* **2024**, e202319235. DOI: <https://doi.org/10.1002/anie.202319235> or <https://doi.org/10.1002/ange.202319235> (german version)

J.N. Singer*; F.M. Müller*; E. Węgrzyn; C. Hölzl; H. Hurmiz; C. Liu; L. Escobar; T. Carell, Loading of Amino Acids onto RNA in a Putative RNA-Peptide World. *Angew. Chem. Int. Ed.* **2023**, 62, e202302360. DOI: <https://doi.org/10.1002/anie.202302360> or <https://doi.org/10.1002/ange.202302360> (german version)

F. Müller*; L. Escobar*; F. Xu; E. Węgrzyn; M. Nainytė; T. Amatov; C.Y. Chan; A. Pichler; T. Carell, A prebiotically plausible scenario of an RNA-peptide world. *Nature* **2022**, 605, 279-284. DOI: <https://doi.org/10.1038/s41586-022-04676-3>

* These authors contributed equally.

Other publications

S. Epple; A. Modi; Y.R. Baker; E. Węgrzyn; D. Traoré; P. Wanat; A. E. S. Tyburn; A. Shivalingam; L. Taemaitree; A. H. El-Sagheer; T. Brown, A New 1,5-Disubstituted Triazole DNA Backbone Mimic with Enhanced Polymerase Compatibility. *J. Am. Chem. Soc.* **2021**, 143 (39), 16293-16301. DOI: <https://doi.org/10.1021/jacs.1c08057>

Conferences

JCF Frühjahrssymposium 2021 (GDCh, online), 29.03 – 01.04.2021

Nucleic Acid Forum (RSC, online), 09.07.2021

XVIIIth Symposium on Chemistry of Nucleic Acid Components SCNAC (Český Krumlov, Czech Republic), 05. – 10.06.2022; Poster presentation

Molecular Origins of Life (München), 16. – 17.06.2022; Poster presentation

New RNA Therapeutics, (München), 24.06.2022

Bioinspired Complex Systems from Basic Science to Practical Applications, (Neve Ilan, Israel), 6. – 11.11.2022; Poster presentation

Oligo conference 2023, (Oxford, UK), 27. – 28.03.2023

Oligo conference 2024, (Oxford, UK), 20. – 21.03.2024

Molecular Origins of Life (München), 19. – 20.07.2024

Table of Contents

Acknowledgements.....	ii
Publications	iii
Conferences	iii
Abstract	vii
1. Introduction	1
1.1 Nucleic acids, proteins and the central dogma of molecular biology.....	1
1.2 Ribosomal translation.....	3
1.3 Nucleoside modifications of the tRNA and their properties.....	4
1.4 Complexity of life and the last universal common ancestor	5
1.5 Prebiotic chemistry and origin of life research	7
1.6 “RNA world” hypothesis	9
1.7 Chirality.....	11
1.8 Theories on the origin of biological homochirality	12
1.9 RNA-mediated peptide synthesis and chirality in the context of translation	14
2. Aim of the project	20
3. Published results.....	21
3.1 A prebiotically plausible scenario of an RNA-peptide world.....	21
3.2 Loading of Amino Acids onto RNA in a Putative RNA-Peptide World.....	33
3.3 RNA-Templated Peptide Bond Formation Promotes L-Homochirality.....	43
3.4 Gradual evolution of a homo-L-peptide world on homo-D-configured RNA and DNA ..	53
4. Unpublished results.....	61
4.1 Rate constants of cleavage kinetics	61
5. Experimental part	64
6. Abbreviations	72
7. References.....	74
Appendix I.....	79
Appendix II.....	161

Appendix III.....	201
Appendix IV	249

Abstract

Life on Earth in its full complexity could not exist without the ability to replicate and maintain its essential genetic and metabolic functions. The question of “how” and “when” life emerged, is one of the biggest unsolved mysteries of humankind. We can follow the evolution of living species by analysing the fossils excavated by scientist around the globe, but tracing back the beginnings of life to understand the transition from abiotic Earth to living matter is a much more difficult riddle to address. Modern life is based on the interactions between the nucleic acids (DNA and RNA) and the proteins, also known as the “central dogma of molecular biology”. One of the most important processes of contemporary biology is the translation, which can be seen as a link between the genotype and phenotype. In this process, the genetic information is decoded and translated into a chain of amino acids, which, upon folding, forms catalytically active proteins. The translation itself poses a chicken-and-egg conundrum, since it involves both nucleic acids and proteins in form of tRNA and the ribosome (a nucleic acid-protein hybrid). The RNA-world hypothesis suggests that initially the RNA molecule carried the dual function of being the keeper of genetic information as well as a catalyst. It is suspected that this versatility was supported by the presence of an extended genetic alphabet of modified nucleosides. Interestingly, a collection of modified nucleosides such as t^6A , m^6t^6A and $(m)nm^5U$, which may be considered to be “molecular fossils”, can be found in close proximity of the anticodon loop of modern tRNA.

Another burning question in the prebiotic science field is the emergence and maintenance of biological homochirality. Modern life, with the complex protein folding and specific interactions, would not be able to maintain its functions if the building blocks were composed of mixed diastereomers. Life requires the molecules of life to be homochiral, with RNA built exclusively out of D-ribose and the proteins out of L-amino acids.

In this work, I address the chicken-and-egg problem, postulating an RNA-peptide world, in which RNA can self-decorate with peptides. I show that the nucleoside-amino acid hybrid structure $(m^6)aa^6A$, (analogous to $(m^6)t^6A$), can be formed by loading an amino acid onto N^6 -methylurea adenosine under prebiotically plausible conditions. The modified nucleosides $(m^6)aa^6A$ and $(m)nm^5U$ can be incorporated into RNA, forming strands that can act as a donor and an acceptor, respectively. The complementary oligonucleotides hybridise through hydrogen bonds, which brings the modified nucleosides into close proximity and facilitates peptide bond formation after prior activation of the amino acid (**Figure 1**). Subsequently, the formed RNA-peptide hairpin strand can be thermally cleaved at the urea moiety, releasing the amino acid-loaded acceptor strand. Upon encounter with another amino acid-carrying donor strand, the cycle can be repeated and the peptide will grow longer. This iterative cycle can be

conducted under one-pot conditions, imitating a prebiotic world where the separation of the intermediates was not required. I present results with a range of different amino acid and various activation methods, leading to the formation of up to decapeptides on RNA.

I demonstrate in my thesis that the peptide coupling reaction in RNA, but also in DNA, exhibits high stereoselectivity towards the naturally occurring L-amino acids. I report that the close proximity of the amino acid to RNA has strong influence on the stereoselectivity, which indicates that the D-ribose sugar present in RNA induces the L-homochirality of peptides. This stereochemical preference is based on the kinetic rather than thermodynamic aspects of the reactions, since the rate constants are the highest for the L-L amino acid coupling. The transfer of di- and tri- peptides is also possible with clear preference for homochirality. I also show a temperature-driven one-pot peptide synthesis with clear selectivity to form the homo-L-peptides.

This thesis provides data that suggests a plausible alternative to the RNA world, namely the RNA-peptide world, in which RNA can self-decorate with peptides and these hybrids can then perform a peptide synthesis cycle, selectively transferring L-amino acids. The repetition of this cycle leads to the growth of longer peptides and can be considered a prototype of a prebiotic, primitive, stereoselective translation mechanism, leading to homochiral peptides.

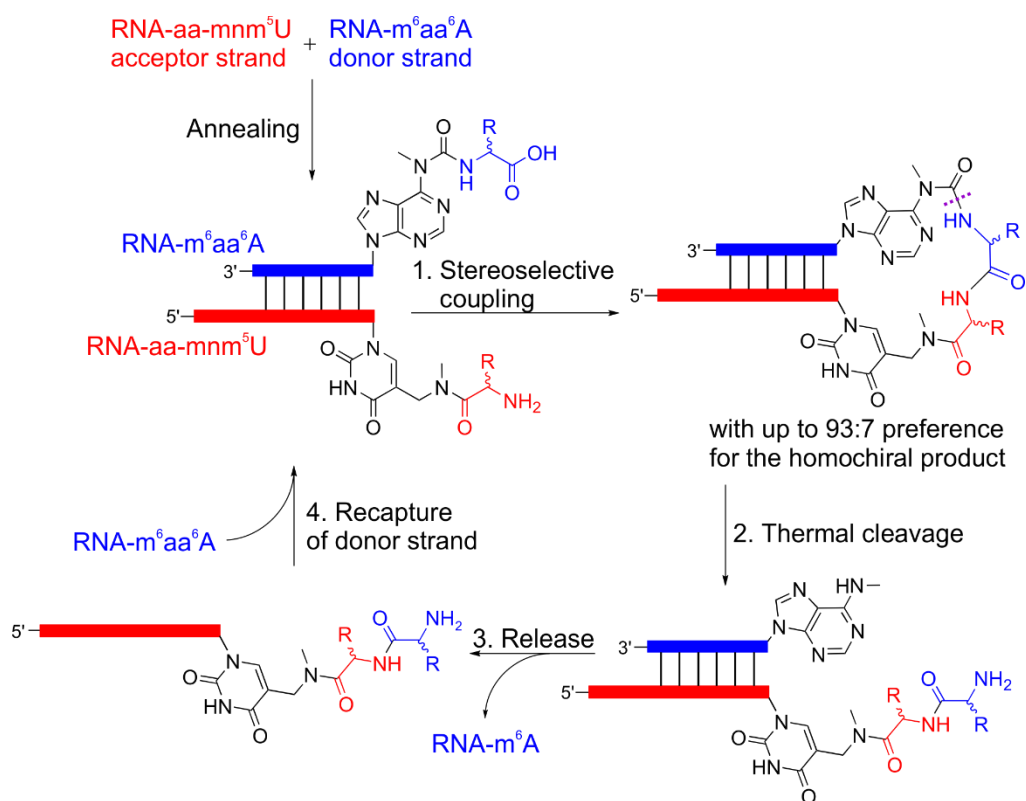


Figure 1. The RNA-peptide synthesis cycle.

1. Introduction

Modern life is based on processes sustained by the crucial biopolymers: ribonucleic acid (RNA), deoxyribonucleic acid (DNA) and proteins. Nucleic acids and amino acids are the basic building blocks of life without which no living organisms could exist in the form as we know them today.

1.1 Nucleic acids, proteins and the central dogma of molecular biology

The structure of DNA was first described by James D. Watson, Francis H. C. Crick^[1] and Rosalind Franklin^[2] in 1953. This was one of the first steps to the basic understanding of the structure and function of our genetic material. The nucleic acids, DNA and RNA, are responsible for storage and transfer of genetic information. They are comprised of nucleotides, connected via phosphodiester bonds at the 3'- and 5'- hydroxy groups, with each nucleotide being built out of three units: a (deoxy)ribose sugar, a phosphate group and a nucleobase. The nucleobases are divided into purine (adenine, guanine) and pyrimidine bases (cytosine and thymine (in DNA) or uracil (in RNA)) (**Figure 2a**). Due to the ability to form hydrogen bonds, the nucleotides create Watson-Crick base pairs (A-T/U and C-G) between two complementary strands of DNA/RNA to form an antiparallel nucleic acid duplex (**Figure 2b**).

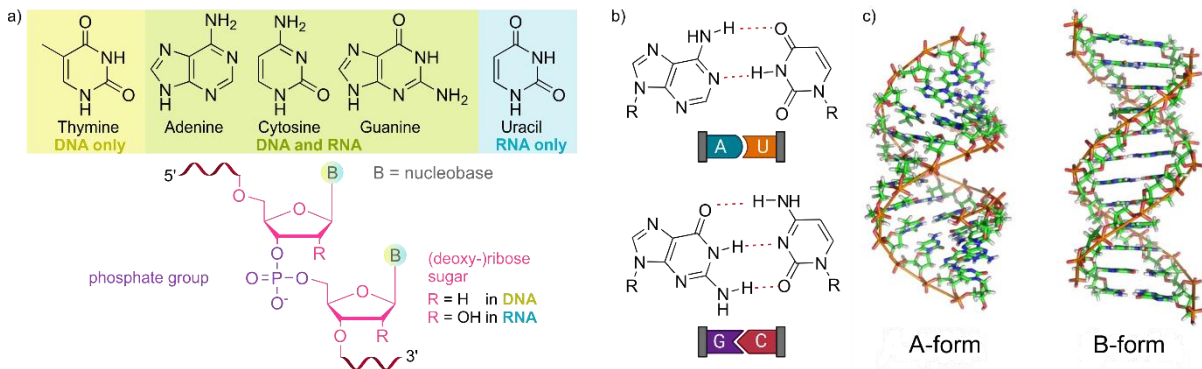


Figure 2: a) Structure of the nucleobases and general structure of a nucleoside. b) Base pairing between A-U and G-C bases. c) Secondary structure of nucleic acids: A- and B-form (adapted from Cerreta 2013)^[3]

The pentose ring, present in RNA (as ribose) and DNA (as deoxyribose) is twisted out-of-plane (sugar pucker), rather than being a planar structure. This changes the relative orientation and thus reduces the steric interactions between the sugar substituents.^[4] The DNA adopts the C2'-endo sugar pucker, while the RNA prefers the C3'-endo pucker – more favourable due to the gauche effect caused by the presence of the 2'OH group. The sugar puckering has impact on the secondary structure of the double-helix formed by these nucleic acids, leading to a B-form DNA and A-form RNA structure (**Figure 2c**).^[5] Due to the chiral nature and optical activity

of nucleic acids, their secondary structure and spatial conformation can be examined by circular dichroism (CD) spectroscopy, involving circularly polarized light (CPL). In a typical CD spectrum of the A-form, a strong negative band appears around 210 nm and a strong positive band at around 260 nm. The B-form causes a characteristic negative signal at ca. 245 nm and a positive one at 280 nm.^[6]

Besides the nucleic acids, the other key players of contemporary biology are proteins, which exhibit a very wide range of structural features and thereof resulting functions. Proteins are built of long polymeric chains of α -amino acids, occurring naturally in the L- configuration. The 20 types of proteinogenic α -amino acids differ in the structure of their side chains (R), which have impact on their functional properties (**Figure 3a**). Currently there is a consensus of 11 amino acids being considered “prebiotically plausible”, such as glycine (Gly), alanine (Ala), valine (Val), leucine (Leu), isoleucine (Ile), serine (Ser), cysteine (Cys), threonine (Thr), proline (Pro), aspartic acid (Asp) and glutamic acid (Glu) (**Figure 3b**).^{[7],[8],[9]} The structure of proteins can be divided into primary (sequence), secondary (folding of the peptide chain), tertiary (three-dimensional) and quaternary (assembly of subunits, as in **Figure 3c**) structure.

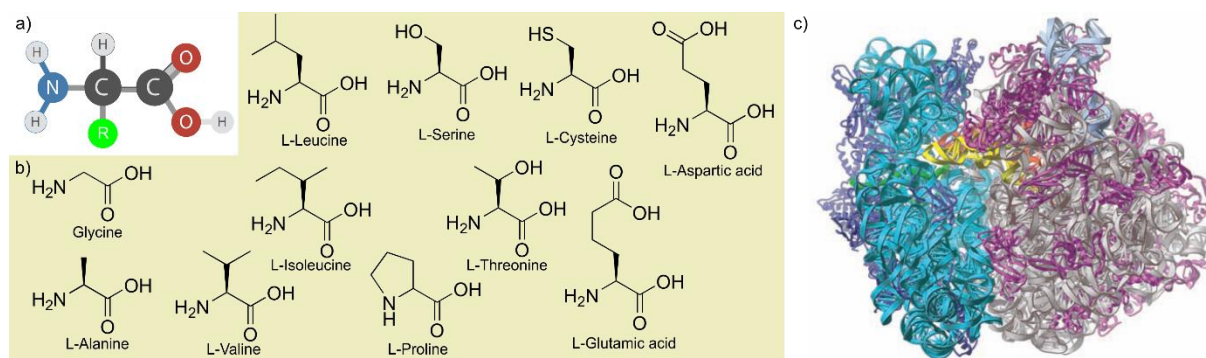


Figure 3: a) General structure of an amino acid, R = side chain; b) amino acids considered to be prebiotically plausible;^{[7],[8],[9]} and c) the quaternary crystal structure of a ribosome (adapted from Noller 2005).^[10]

The modern biology requires a very complex machinery, reliant on the multifaceted processes of replication, transcription and translation – involving the macromolecules DNA, RNA and proteins. The functions and interdependences between these molecules of life were summarized by F. H. C. Crick in 1958^[11] and are nowadays known as the “central dogma of molecular biology”.^[12] It describes the information flow from the genomic DNA, through RNA, into proteins (**Figure 4**) – in other words the transition from genotype (keeper of genetic information) into phenotype (the expression of this gene and thereof resulting structural and catalytic functions). The DNA, which mainly functions as a storage of genetic information and can be replicated with the help of DNA polymerase, is then transcribed into messenger RNA (mRNA) by enzymes called RNA polymerases. The mRNA serves as a template for the protein synthesis and is translated into a chain of amino acids, that then folds into higher order

structures to yield proteins. Even though the process is not entirely unidirectional (RNA can be reverse-transcribed into DNA), an information transfer from proteins back into nucleic acids has not been found.^[11]

The modern translation is aided by the ribosome^[13] (consisting of ribosomal RNA (rRNA) and proteins) with the help of another type of RNA, the transfer RNA (tRNA). Ergo, the process of making proteins involves a protein-containing machinery (ribosome) itself as a catalyst for the peptide bond formation. This chicken-and-egg conundrum was noted by Dr. A.L Dounce in 1956^[14] and is the subject of contemporary research ongoing in the field of nucleic acids.

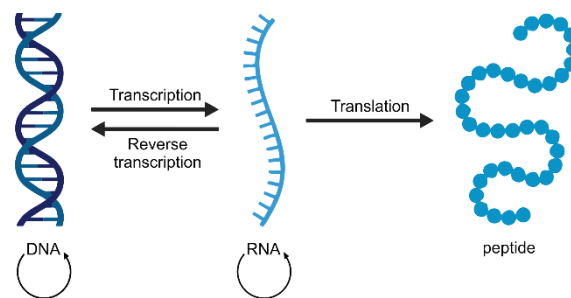


Figure 4: Central dogma of molecular biology describing the flow of information from DNA to RNA to peptides.

1.2 Ribosomal translation

For a long time, the understanding of the mechanism of ribosomal translation was limited. Therefore, the thorough examination of this process as well as its key players, the transfer RNA (tRNA) and the ribosome, has been of utmost importance in the past few decades.

The transfer RNA (tRNA) is a molecule of 76-90 nucleotides length.^[15] Its secondary structure, resembling a cloverleaf, can be divided into five distinct sub-structures: the D arm, the anticodon arm, the variable loop, the TΨC arm, and the acceptor stem.^[15] The nucleotide triplet (anticodon) in the anticodon loop of the tRNA dictates which amino acid will be attached to the tRNA. The tRNA folds into an L-shaped tertiary structure, that is crucial for proper fitting into the ribosomal binding site.^[16]

The first step of translation is the aminoacylation, during which the amino acid is attached via an ester-bond to the 3' end of the acceptor stem of the cognate tRNA by aminoacyl-tRNA synthetases (aaRSs).^[17] The activation of the amino acids occurs by the reaction with adenosine triphosphate (ATP).^[18] The resulting aminoacyl-tRNA acts as an adapter molecule, responsible for recognition and decoding the mRNA sequence by anticodon-codon interactions^[19] in the second step of the translation (peptidyl transfer), which takes place at the peptidyl transferase centre (PTC) of the ribosome.^[20]

The ribosome is a highly complex moiety, therefore it took many years to elucidate the X-ray crystal structure of the full ribosome macromolecule,^{[10],[21],[22]} which was necessary for better understanding of its function and the mechanism of translation. The ribosome is a ribonucleoprotein hybrid structure, built of ribosomal RNA (rRNA) and ribosomal proteins. It can be divided into two distinct subunits: the small and the large subunit, each of them containing three tRNA binding sites.^{[13],[20],[23]} The smaller subunit controls the fidelity of translation by facilitating the interactions between the codon of the mRNA and the anticodon of the aminoacyl-tRNA, while the large subunit, containing the PTC, is responsible for catalysis of the peptide bond formation on the acceptor arm of the tRNA.^[13] The translation proceeds on the mRNA in 5' → 3' direction, binding the complementary anticodon of the tRNA bearing the cognate amino acid.

It has been demonstrated that the PTC is an all-RNA site, so the ribosomal proteins are not indispensable for the actual peptide formation.^[24] These results suggest that the PTC can be seen as a ribozyme^[25] and a potential remnant of proto-ribosomal peptide synthesis.^[26] Next to catalysing peptide bond formation, the ribosome is also protecting the base-labile ester bond from hydrolysis and by this preventing deacylation,^[27] therefore alternative, more stable linkages in form of nucleoside modifications should be considered in the context of a prebiotic translation mechanism.

1.3 Nucleoside modifications of the tRNA and their properties

Nucleoside modifications are highly abundant in the tRNA, with the greatest variety found in the anticodon loop.^[15] They contribute to the maintenance of correct structural conformation and functionality,^[15] thermal stabilization of the tRNA^[28] as well as diversification of the genetic code and aiding the decoding process.^[29] Inside and in immediate proximity of the anticodon, a rich variety of hypermodified nucleosides like m³C, k²C, agm²C, i⁶A, m⁶A, t⁶A, m⁶t⁶A, nm⁵U, mnm⁵U, cmnm⁵U, 2'O-methylations (**Figure 5**) can be found.^{[15],[30],[31],[32],[33]}

Some of these modifications are of particular interest to us due to their functionalities, such as for example the prebiotically plausible modifications (m)nm⁵U₃₄ at the Wobble position 34 and (m⁶)t⁶A₃₇ at position 37 of the anticodon loop of the tRNA.

The modified uridines like nm⁵U, mnm⁵U or cmnm⁵U all contain a reactive amino group at the residue connected to the nucleobase. They play a role in decoding and contribute to binding specificity to the mRNA.^[33] Bacteria and archaea contain modified cysteine residues lysidine (k²C or L)^[34] and agmatidine (agm²C or C⁺)^[35] at position 34, respectively, which aid codon recognition and increase decoding efficiency in tRNA^{lle}.^{[36],[37]}

Interestingly, the t^6A or m^6t^6A modifications, essential for codon recognition,^[38] are nucleoside-amino acid hybrids, containing an amino acid connected to the nucleobase via a urea bridge. The urea linkage is characterised by high stability in aqueous solution both at basic and acidic pH^[39] and by thermal cleavability.^{[40],[41]}

One of the most common modifications, highly abundant in rRNA, are methylations. 2'OMe nucleotides increase the stability of the RNA^{[42],[43]} by reducing the phosphodiester backbone hydrolysis^[44] and preventing base excision by blocking the nucleophilic attack at C1'. The modification m^6A , adenosine methylated at the N⁶-position of adenine, is thought to be involved in mRNA processing and gene expression.^{[45],[46],[47]} Both methylations are thought to be prebiotically plausible.

The above mentioned non-canonical nucleotides are conserved in tRNAs and rRNAs across almost all kingdoms of life^[48] and are therefore considered 'living molecular fossils' that represent relics of an RNA world.^{[49],[50]}

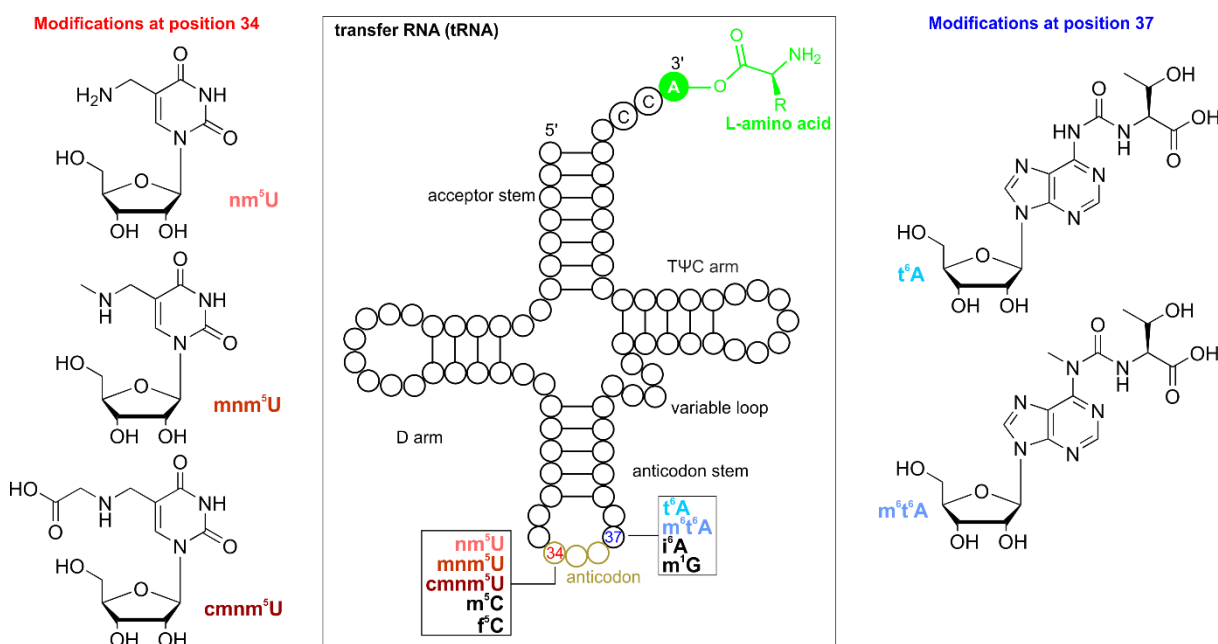


Figure 5: Schematic representation of a cloverleaf structure of tRNA and the location of selected non-canonical modification at the anticodon loop.

1.4 Complexity of life and the last universal common ancestor

How to define life in the first place? Especially for the first form of life, the line between being a bunch of atoms and coming alive is a particularly thin one. The National Aeronautics and Space Agency (NASA) adopted the definition of life originally suggested by Joyce in 1994: "Life is a self-sustained chemical system capable of undergoing Darwinian evolution".^[51] Even though this is now an accepted scientific consensus, the definition faces criticism.^[52] In

particular due to the inclusion of the term “Darwinian evolution”, which should be considered a consequence, rather than the requirement for the formation of minimal life.^[51] For the prebiotic research field, more important than the exact definition of life, is the understanding of how the transition from chemistry to biology happened.^[53] Life requires a general mechanism for self-replication, resistant to environmental changes and structural variations, with a net excess of creative over degradative processes.^[54]

The early, primitive life was naturally different from the highly complex life forms existing today. It is believed that the origin of all species can be traced back to the Last Universal Common Ancestor (LUCA), a theoretical model organism that existed ca. 3.5 – 3.8 billion years ago,^[55] which with time differentiated and evolved, ultimately leading to the broad variety of species that we know today (**Figure 6**). It is suggested that some features of contemporary biology must have developed at a very early stage in the history of life, due to the necessity of continuous evolution from one organism to the next one.^[56] LUCA is believed to have used DNA for genetic information storage, polymerases to copy DNA and ribosome to make proteins, all surrounded by a cell membrane.^[57] Phylogenetic studies endeavour to trace back the genetic and evolutionary connections of organisms.^[58] The classification of organisms into three “domains” of life (Archaea, Bacteria and Eukarya),^[59] is now being superseded by the theory that Eukaryotes in fact originated from Archaea and Bacteria, reducing the “tree of life” of LUCA to two branches^[57]. This archaeobacterial origin is hypothesized due to the analysis of the small subunit of ribosomal RNA, and suggests that the eukaryotic cell components evolved later from the more primitive, but already diversified prokaryotic cell.^[60] The biggest number of essential genes preserved in all domains of life are related to translation^[61] and they can be seen as a relic of LUCA.^{[62],[63],[64]} It has been reported that the minimal cell can function with less than 300 genes, and more than half of them are connected to translation.^[65]

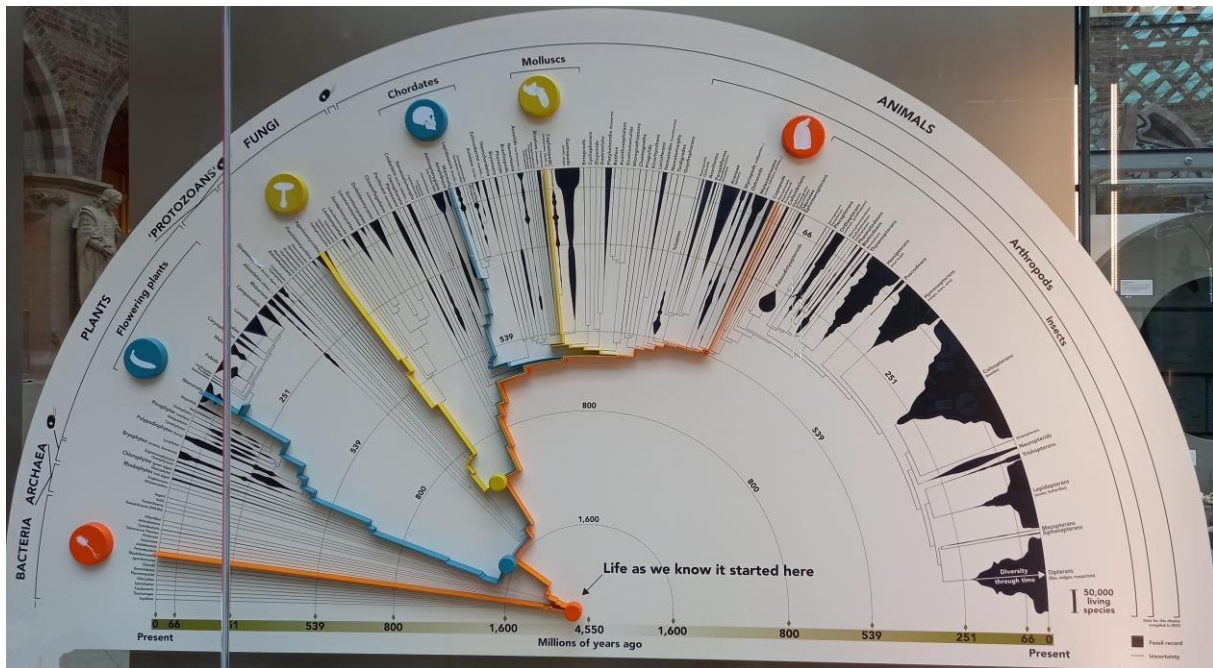


Figure 6: A representation of the phylogenetic tree of life, evolving from a common ancestor around four billion years ago. Picture taken by the author at the Natural History Museum in Oxford, UK.

Yet, the appearance of LUCA is not marking the emergence of life, but rather the development of an already complexified prokaryotic organism,^[66] with features, more resembling a primitive form of modern life, than what we could consider first life.^[67] LUCA, which is by definition too complex to emerge spontaneously, is believed to have been preceded by a simpler molecular self-replicator, like the hypothesised “initial Darwinian ancestor” (IDA).^[68] IDA could have been a nucleopeptide replicator consisting of amino acids and nucleosides – biological building blocks present on Early Earth.^[69] But what preceded IDA, at what point did the transition from non-living matter to life happen?

1.5 Prebiotic chemistry and origin of life research

The exact date of the formation of first primitive life is subject of a substantial scientific argument – some scientists claim that life could be around 3.5 billion years old^[70] or older^[71], while others criticize these findings due to strong metamorphism/recrystallization of the found carbonaceous inclusions during the geological periods following their formation.^[72] The “when” of life formation cannot be predicted with full certainty, neither does the “where”. The place must have been protected from the harsh, inhabitable conditions that were prevalent on the Earth at that time. Geothermal sites^[73] are potential places where first molecules of life could have assembled due to the chemically favourable conditions as well as the beneficial spatial separation of non-orthogonal reactions. The components could have formed in wet/dry cycles^{[74],[75]} in separated rivers, that eventually merged into each other and flowed into ponds.

On the way, they were exposed to sunlight irradiation, which caused evaporation of water leading to dry-down process. Various mineral surfaces may have aided the reactions.^[76] Hydrothermal vents could have been the source of elevated temperatures of up to ca. 350°C required for some reactions.^[77] Another option are freezing-thawing cycles,^[78] in which the reactants were exposed to concentration increase in the interstitial spaces of ice crystals under eutectic conditions, and then dilution through the thawing of the ice.^[79] Current research confirms the possible presence of frozen areas on early Earth.^[80] Due to reduced solubility of CO₂ in ice, the eutectic conditions could have also been beneficial for the regulation of the pH of the prebiotic reaction mixtures.^[81] Generally, variations in temperature could have been caused by the day-night cycle, volcanic eruptions or the impact of meteorites on the Early Earth.^{[82],[83]}

But before answering the big question of how the inhabitable place, that our planet was at its beginnings, brought forward first self-sustaining organisms, one has to examine the plausible formation pathways of the building blocks of life from the abiotic environment present on Early Earth.

Eschenmoser pointed out, that “the origin of life cannot be ‘discovered’, it has to be ‘re-invented’”.^[84] The questions of reinvention of the mechanism of biogenesis are being addressed by researchers in the field of prebiotic chemistry, pioneered by the contemporaries A. I. Oparin^[85], J. B. S. Haldane^[86] and S. L. Miller^[87] in the 20th century, examining possible pathways of how the life on Earth has started from the “primordial soup” of building blocks of life. The common approaches are either the geochemical “bottom up” or biological “top down” approach – either trying to decipher how the inorganic compounds turned into molecules of life (transition from inanimate to animate matter), or how to reduce biology to the most primitive version of itself, in order to build a hypothetical model of a most primitive living organism.^[88] Both approaches are difficult to address.

The concept of a “primordial soup”^[89] as the prebiotic source of the first molecules of life has been coined, assuming that the compartmentalization is a later aspect of biology. Early on, in mid-19th century, the well-known abiotic Strecker synthesis^{[90],[91]} of alanine has been demonstrated, followed by the formation of ribose from formaldehyde in the autocatalytic process known as the formose reaction.^{[92],[93]} The problem of low stability of ribose^[94] has been alleviated by stabilization with borate minerals.^[95] Later on, Stanley Miller demonstrated with the Miller-Urey experiment, that in an electrical discharge experiment under conditions considered prebiotically plausible at that time, various amino acids can be produced from CH₄, NH₃, H₂O and H₂.^{[87],[96]} Eventually, it was shown that also the nucleobases^{[97],[98],[99],[100]}, and finally nucleotides^{[74],[101],[102],[103],[104]} can be produced under primordial conditions, for example

through HCN chemistry.^{[105],[106]} The nucleotides can then assemble into oligonucleotide chains by thermal gradient^[107] or on glass surfaces.^[108] Further, the spontaneous formation of peptides from amino acids on mineral surfaces has been reported,^[109] leading to the abiotic synthesis of proteinoids (protein-like molecules) synthesised via thermal condensation reactions under the removal of water (evaporation).^{[110],[111]} Eventually, the enzyme-free replication of RNA on RNA template could be shown.^{[112],[113]}

In addition to the terrestrial synthesis pathways, the extra-terrestrial source of organic molecules^{[114],[115]} cannot be excluded. Meteorites like the Murchison meteorite^{[116],[117]} are sources of amino acids, providing a composition of amino acid similar to those produced in spark discharge experiments.^[118]

It has been demonstrated how amino acids, nucleobases and sugars can be formed from inorganic precursors^[119] and then assembled into peptides and oligonucleotides under prebiotically plausible conditions. The next step is to understand how these molecules could have gained self-replicative and catalytic abilities and how these processes were linked to an evolutionary circuit.

1.6 “RNA world” hypothesis

As pointed out before, the mechanism of modern translation represents a chicken-and-egg conundrum. Today, we see that RNA makes protein and proteins make RNA. To solve this chicken-and-egg puzzle it was hypothesized that the early predecessors may have consisted of one component only.^[120] Supported by the discovery of the catalytic properties of RNA by Cech and Altman,^{[121],[122]} as well as the fact that the PTC – which is the core of translation – contains exclusively RNA, the “RNA world” hypothesis has been proposed. It is suggesting that at the prime of life ca. 4.2 – 3.6 billion years ago,^[123] the abiotically synthesized RNA, which is even in modern biology a highly versatile molecule,^[124] could have acted both as storage of genetic information and as metabolic catalyst.^{[120],[56],[125],[126]} In that scenario, the predecessor of modern protein-based enzymes were ribozymes (RNA enzymes), RNA structures with catalytic properties.^{[125],[127]}

As suggested by Gilbert in 1986,^[120] the RNA building blocks would first assemble from a nucleotide soup into oligonucleotides, which would then develop into self-replicators by recombining and mutating its own structure through intron splicing (as the first stage of evolution). With the aid of simple cofactors, they would exhibit a range of primitive enzymatic functions and finally start to synthesize proteins.^[128] Eventually, the proteins would then evolve into better, more efficient catalyst than the RNA counterparts, even though the range of the available functions would be analogous to those of the ribozymes.^[120] The function of storage

of genetic information would be later transferred to DNA, likely due to the fact that DNA is generally more stable against hydrolytic cleavage^[129] and is less prone to secondary structure formation by self-folding, compared to RNA^[130] which simultaneously makes DNA a worse candidate for catalytic properties.

Ribozymes can therefore be seen as molecular fossils, a remainder of the autocatalytic function of RNA.^[131] They have been described to catalyse the formation of nucleotides from the prebiotically accessible ribose and nucleobases,^{[132],[133],[134]} the aminoacylation of the tRNA^{[26],[135],[136]} and to mediate the peptide bond formation.^{[137],[138],[139],[140]} This is important in the light of the origin of translation, as they could have served as a primitive version of the peptidyl transferase centre.^[141]

The RNA-catalysed replication of RNA is one essential requirement for the existence of the RNA world, but also a simple form of metabolism is thought to have accompanied this process, providing a source of the feedstock molecules.^[54] Equally important is the aspect of prebiotic compartmentalization and molecular organization. Some form of a primitive cell (protocell) might have existed, forming vesicles needed to provide packaging and therefore protection of the molecules of life as well as the increase of reagent concentration.^{[51],[131]} It could have been an hydrothermally formed abiotic, inorganic compartment^[142] or a form of acellular organization by the RNA, based on covalent or non-covalent interactions.^[143] A membrane formation from a self-complementary peptide has also been shown.^[144]

The “RNA world” is not an implausible theory,^[126] although it has its shortcomings related to the difficulty of the continuous evolution and the transfer of the catalytic abilities to proteins at a later stage of evolution. The catalytic abilities of RNA are limited due to low substrate affinity and rather non-specific interactions, so the feasibility of RNA forming a universal, substrate-independent replicase or polymerase, without presence of any protein is difficult to believe. Also, the “RNA world” theory is focusing on canonical nucleosides only, excluding the rich variety of the hypermodified nucleosides, which could have had significant influence on the functionality of the RNA. It is believed that the RNA could have co-existed with cofactors, which would assist the RNA-catalysed reactions, enhancing their functionality.^[145] These cofactors could have been either inorganic molecules like metal ions or Fe-S clusters, or organic cofactors, some of which could have been synthesised under prebiotically plausible conditions.^[146] Recent studies suggest the possibility of a “DNA-first” model,^[147] with DNA being the first carrier of genetic information or alternatively a coevolution of RNA with DNA into a “heterogenous nucleic acid genetic system”^[148] characterised by higher stability against hydrolysis.^[129]

The “RNA world” hypothesis is accompanied by the theory of having a hybrid “RNA-protein world”, where the RNA and proteins would share the functions of the catalysis and heredity. The RNA would be able to catalyse peptide bond formation,^[149] and then co-evolve with the peptides^[150] into an “RNA-peptide world” of improved functionality. This system could provide a continuous evolution without the need to transfer the catalytic functions from RNA to proteins on the way to the contemporary DNA-RNA-protein system.^[64]

Despite its structural complexity, RNA is assumed to have been able to arise prebiotically.^[151] However, the presence of chiral centres, which are necessary for the duplex formation, complicates the stereoselective prebiotic formation of RNA significantly.^[152] Going one step back, it is hypothesized that the RNA could have been preceded by more primitive, non-chiral predecessor.^[153] A potential candidate is the glycol nucleic acid (GNA),^[154] or peptide nucleic acid (PNA), consisting of N-(2-aminoethyl)glycine units.^[155] A more complex alternative is the threose nucleic acid (TNA) made of α -L-threofuranosyl building blocks.^[156] All of these polymeric structures have the advantage of a greater degree of simplicity compared to the ribofuranosyl analogue, while being able to maintain the correct base pairing, forming stable duplexes with itself and with RNA. They are possible candidates as templates for the assembly of ribonucleotides into polymeric RNA structures.

1.7 Chirality

Chirality is a property of molecules, already observed by Louis Pasteur in 1848 on sodium-ammonium salt of tartaric acid crystals,^{[157].^[158]} later, in 1894, described by Lord Kelvin^[159] and subsequently specified by Cahn, Ingold and Prelog.^[160] The requirement for chirality is the presence of a chiral centre, which creates two mirror images that are non-superimposable. Even though the physical and chemical properties of the enantiomers are identical,^[114] their way of interacting with other molecules might differ.^[161] All amino acids (except for the achiral glycine) contain an asymmetric carbon centre, so they exist in two different enantiomeric forms, named historically with the stereo descriptors L- (left-handed) and D- (right-handed) in the Fischer projection.^[162] Even though the CIP-nomenclature^[160] is more commonly used in stereochemistry nowadays, the L- and D- nomenclature remains still the more prevalent for amino acids.

The measure for the purity of a mixture of enantiomers is the enantiomeric excess (ee) (or diastereomeric excess (de) for a mixture for diastereomers). By definition, a pure enantiomer has the ee of 100%, while a racemic mixture has none (0%).

1.8 Theories on the origin of biological homochirality

Life, as we know it, is homochiral, with proteins compiled of L-amino acids and nucleic acids of D-(deoxy)ribose. This property influences the folding into structures of higher order and their functionality.^[163] Homochirality is particularly crucial for the proper folding of DNA and RNA, due to the requirement of complementarity.^[164] Traces of D-amino acids, which are thought to be molecular markers of the aging process, have been found in living organisms,^[165] although not in any of the functional molecules.

The study on how these molecules of life became homochiral (homochirogenesis)^[166] has been of increasing interest over the past decades. The fact that the homochirality is a property of all kingdoms of life, is a supporting argument for the existence of a common ancestor of all species. It is of great interest to know the mechanism through which the enantiomeric bias was created. The work of Breslow et al. showed that in the formose reaction L-amino acids are able to promote excess formation of D-glyceraldehyde,^[167] and Levine et al. report prebiotically plausible enantiomeric amplification of amino acids.^[168] Despite these promising results, we know, that prebiotic reactions do not display particularly good diastereoselectivity, usually producing a mixture of enantiomers. It is assumed that at the prebiotic stage, the primordial soup consisted of roughly equal amounts of L- and D-enantiomers, with only a minimal enantiomeric excess if at all,^[169] or that the mixture was entirely racemic. In that case, the enantiomeric bias, that eventually led to homochiral biomolecules, would have emerged at a later stage and would have been amplified over the course of evolution.^{[114],[170],[171]} Having enantiopure solution of monomers is particularly important when it comes to the process of assembling higher-order structures. Joyce reported that the presence of the opposite enantiomer obstructed the process of template-directed homochiral polymerization of nucleic acids,^[112] also referred to as “enantiomeric cross-inhibition”.^[172]

The path to biological homochirality starts with ways to produce enantiomeric enriched mixtures of biomolecules of the “correct” chirality. There is a range of possible chemical^[173] and physical^{[174],[175]} mechanistic models for chiral amplification and enantiomeric enrichment, both of terrestrial and extra-terrestrial origin (**Figure 7**).^[176]

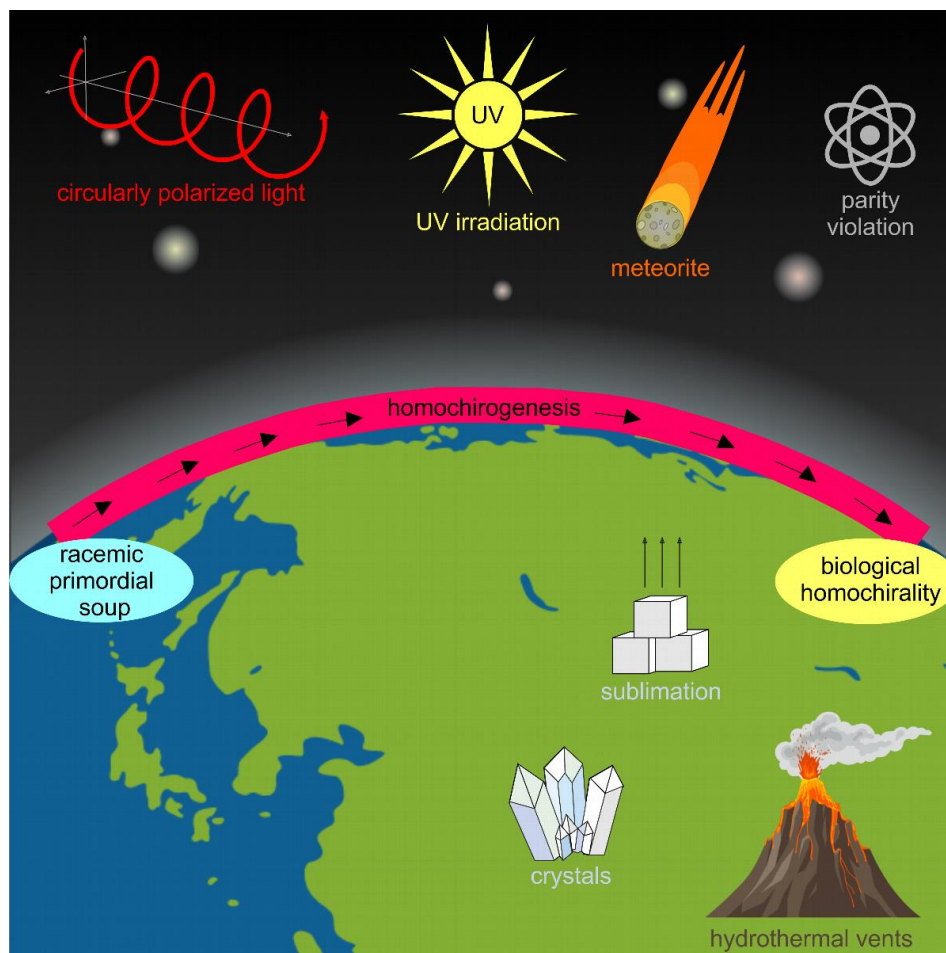


Figure 7: Possible impacts on the formation of biological homochirality.

On Earth, the autocatalytic amplification of one enantiomer could have taken place in hydrothermal vents.^[75] There are examples of spontaneous chiral amplification of serine by sublimation^[177] or enantioselective adsorption on calcite crystals^[178] or quartz.^[179]

There are theories that the initial enantiomeric imbalance came from outside our planet,^{[180],[181]} with models indicating possible induction of a slight enantiomeric excess of one enantiomer through circularly polarized light (CPL),^{[182],[183],[184]} or UV photodecomposition^[185] with subsequent amplification through autocatalytic processes. Also sublimation is a plausible process,^[186] even though this claim is criticized for being based on experiments carried out under out-of-equilibrium conditions,^[187] not really suitable as a prebiotically plausible explanation. Traces of enantiomeric excess of L-amino acids has been found on meteorites,^{[188],[189]} which could have been a basis for initial imbalance of enantiomers. Some mathematically-physical approaches link the emergence of homochirality to parity violation,^{[163],[190]} due to a very small energy difference between the two enantiomers.^[191] These theories discuss the possible source of the initial enantiomeric imbalance, but they do not

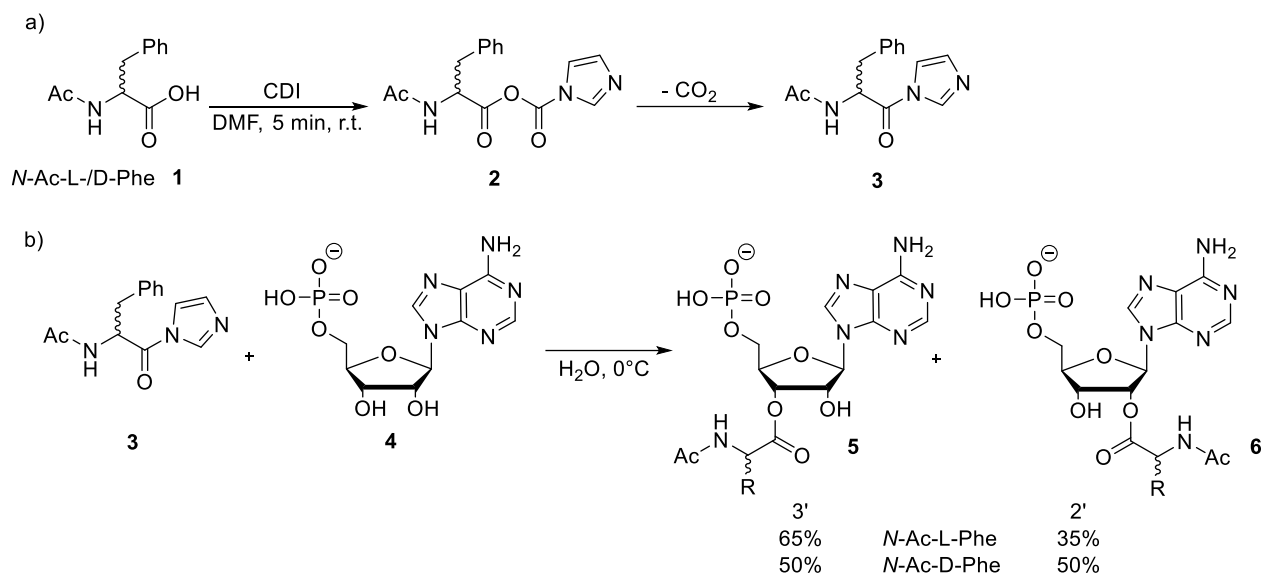
explain the emergence of a reliable amplification mechanism, which is a requirement for the maintenance of the obtained homochiral bias.^[192]

The alternative scenario would be, that the homochirality emerged and gained relevance at the level of polymers, where the structural uniformity would stipulate the functionality.^{[193],[161]} Not only the origin of homochiral properties, but even more the amplification and the maintenance thereof are crucial. Every enantiopure compound is with time subjected to racemization, so the question is, how the biomolecules circumvented this problem. Theories, like the Frank model,^[194] suggest that processes like autocatalysis and mutual antagonism could have started and then helped maintaining enantiomeric imbalance.^[161]

1.9 RNA-mediated peptide synthesis and chirality in the context of translation

The question of prebiotically plausible peptide formation and its stereoselectivity has been addressed by several research groups, focusing on different stages of translation.

The group of **Lacey** has focused on the first step of translation (aminoacylation) and examined the stereoselective formation of bis(α -aminoacyl)esters of 5'-AMP, as potential intermediates in the protein synthesis in the prebiotic context.^[195] *N*-Ac-L-Phe or *N*-Ac-D-Phe **1** was activated with carbonyldiimidazole CDI to form a mixed anhydride **2**, which then degraded spontaneously to give the imidazolide of *N*-Ac-Phe **3** upon CO₂ release (**Scheme 1a**). The imidazolide **3** was then reacted with the commercially available 5'-AMP **4**, to give a mixture of 3'- and 2'- monoesters of 5'-AMP **5** + **6** (**Scheme 1b**). The reaction was quenched with hydrochloric acid to decompose the unreacted excess imidazolide **3**. Lowered pH also prevented the degradation of the formed ester and its migration between the 2' and 3' positions.



Scheme 1: a) Synthesis of *N*-Ac-Phe imidazolide from *N*-Ac-Phe; b) Reaction of 5'-AMP with the *N*-Ac-Phe imidazolide to give a mixture of 2'- and 3'- monoesters of 5'-AMP.

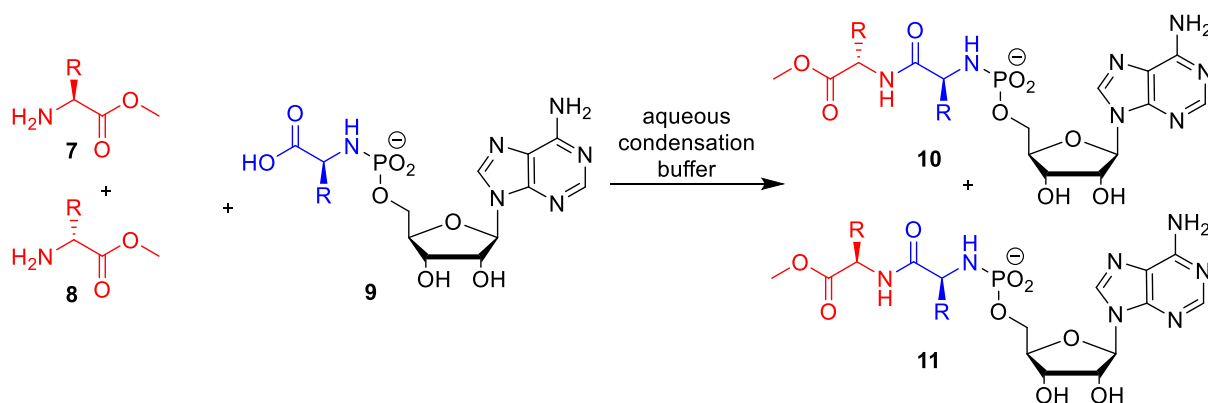
It has been observed that the amino acid was continuously migrating between positions 2' and 3' at neutral pH, with the L-amino acid favouring to 65% the 3' position (which is the relevant position in the process of translation), while for the D-amino acid did not show significant preference for either position.^[196] Since it has been reported, that the esterification preferentially takes place at the 2' position,^[197] one can suspect that the L-Phe monoester which has 65% of free 2'-OH groups, will react faster to form a bis(α -aminoacyl)ester than the D-counterpart. To examine this hypothesis, the mixture of *N*-Ac-L-Phe or *N*-Ac-D-Phe monoesters of 5'-AMP was reacted in a competition reaction with either L- or D-Phe imidazolide. The outcome was evaluated by monitoring the disappearance of the monoesters via integration of HPLC signals. Kinetic studies have shown, that the L-, L- bis(α -aminoacyl) ester was formed 2.7 times faster than the D-, D- bis(α -aminoacyl) ester.^[195]

Additionally, it was shown that at pH 5 the amino acid in an adenylate anhydride (5'-aminoacyl-AMP) migrated spontaneously from the 5'-phosphate group to 2' or 3'-OH group of the ribose. The L-amino acids were transferred nearly quantitatively, while the D-amino acids formed the ester in only 50% yield.^[198] The synthesis of 5'-aminoacyl phosphate mixed anhydrides is described to be accessible with prebiotic activation methods.^[199] Previously, it was demonstrated that amino acids could be transferred to the 2'/3'-OH of a ribonucleotide strand in an imidazole catalysed reaction.^[200] The question of peptide bond formation has not been addressed in this work.

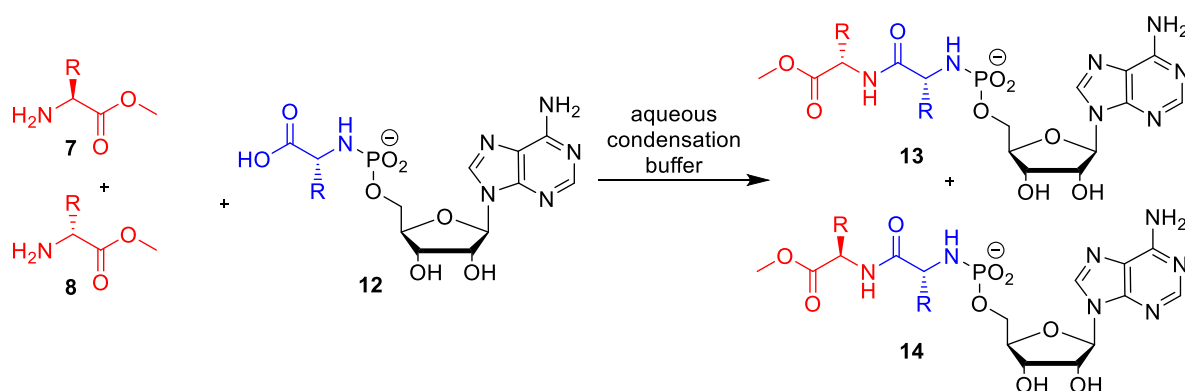
Researchers from the Richert group have reported the formation of phosphoramidates from ribonucleotides and amino acids in a condensation buffer.^[201] These so called "amino acyl nucleotides" are prepared by reaction of 2-methylimidazolide of AMP and free amino acid and

analysed by nuclear magnetic resonance (NMR) spectroscopy.^[202] A variety of L-amino acids, and also their non-naturally occurring D-counterparts was used to create a library of amino acidyl nucleotides. These amino acidyl nucleotides could be transformed into dipeptide compounds by coupling with amino acid methyl esters in an aqueous PIPES condensation buffer at pH 7.5, containing MgCl₂ and EDC as activator, at 0 °C.^[203]

With these compounds in hand, they questioned the diastereoselectivity of the dipeptide forming reactions. For this purpose, an equimolar mixture of the L- and D-amino acid methyl ester (**7** and **8**) was prepared and incubated with L- (**Scheme 2**) or D- (**Scheme 3**) amino acidyl nucleotide (**9** or **12**) under the described condensation conditions for 72 h to yield a mixture of diastereoisomers (**10+11** or **13+14**). The diastereomeric ratio could be determined from the integration of the ¹H, ³¹P or COSY NMR peaks in the spectrum of the crude reaction mixture. The limiting factor, impeding the quantification in the analysis, was the occasional lack of peak separation.^[203]



Scheme 2: Chirality competition reactions with L-amino acidyl nucleotide by Doppleb et al.^[203]



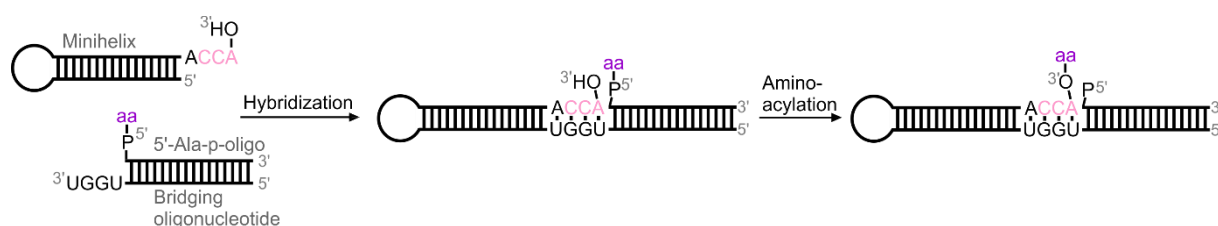
Scheme 3: Chirality competition reactions with D-amino acidyl nucleotide by Doppleb et al.^[203]

A consistent preference for the formation of the homochiral D/D or L/L dipeptide product was reported, among a range of amino acids (in particular for valine and tryptophan). Interestingly, the D/D homochiral product was in more cases formed with a higher diastereomeric excess

over the L/L product. Control coupling reactions with amino acid methyl esters and acetyl protected L-amino acids were performed. As expected, these coupling reactions exhibited no diastereoselectivity, indicating that the homochiral preference is prompted by the D-ribose of the nucleoside. These results suggest, that the nucleotide plays a crucial role in the induction of the stereoselectivity of the peptide bond formation.

Similar levels of diastereoselectivity were observed by the researchers from Pascal lab, working with the 5(4*H*)-oxazolones^[204] – prebiotically plausible analogues of activated α -amino acid – which are known for their fast epimerisation and therefore the loss of stereochemical information. They demonstrated a kinetic preference for adapting the homochiral oxazolone form through the C-terminal epimerization of a dipeptide in presence of EDC. The reaction of the oxazolone with a nucleophilic amino group of another amino acid yielded a preferential formation of the homochiral product. The diastereomeric ratios (determined by integration of HPLC peaks) of homo- vs. heterochiral products were growing with increasing length of the peptide chain leading to the enrichment of homochiral structures.^[205]

Tamura and Schimmel have worked on selective aminoacylation of an RNA minihelix with acyl phosphate mixed anhydrides.^{[206],[207],[208],[209],[210]} They designed an RNA minihelix (that can be considered a potential forerunner of the amino acid attachment site in tRNA^[211]) containing naturally occurring D-ribose moieties, with the ability to form a hairpin structure through self-complementarity. The minihelix contained a four-nucleotide single-stranded overhang at the 3' end, which could hybridize to a duplex of a bridging oligonucleotide and a 5' aminoacyl phosphate oligonucleotide through Watson-Crick base pairing (**Scheme 4**). The binding led to a spontaneous aminoacylation (energetically a downhill reaction^[209]) of the minihelix in 15% yield with a stereochemical preference for L-Ala over D-Ala of ca. 4:1 (determined by denatured polyacrylamide gel electrophoresis).^[206] The same level of preference could be shown for leucine and phenylalanine. Exchanging D-ribose by L-ribose building blocks in the RNA caused selectivity for D-amino acids with a reciprocal preference of 1:3.6.^{[206],[208]} It has been suggested that the homochirality of amino acids was established at the step of aminoacylation, dictated by the chirality to the RNA.^[207]

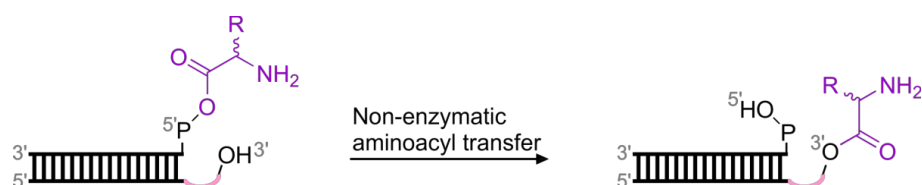


Scheme 4: Stereoselective aminoacylation of an RNA minihelix with an aminoacyl phosphate oligonucleotide. Scheme adapted from Tamura and Schimmel (2004).^[206]

John Sutherland's group follow-up research emphasized the shortcomings of this model from the evolutionary perspective, such as the missing rationale for the length and sequence of the chosen 3'-overhang, or the explanation of the chemoselectivity of the aminoacyl transfer.^[212]

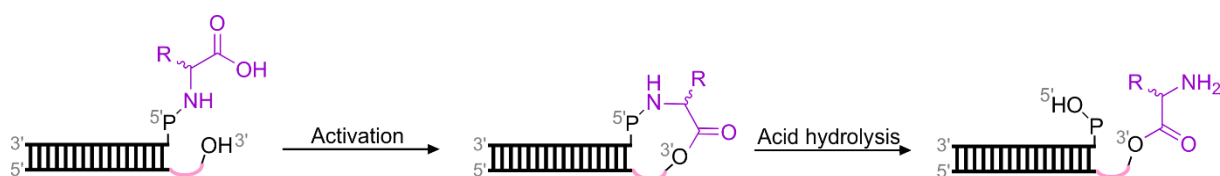
To address these issues, the researchers first envisioned a screening experiment through various lengths and sequences of the 3'-overhang. They changed the experimental setup by removing the "superfluous" bridging oligonucleotide and instead performing the aminoacyl transfer directly in between two complementary RNA strands forming a nicked loop (**Scheme 5**).^[213] Due to lower stability of the aminoacyl phosphate mixed anhydride over the aminoacyl ester, the reaction occurred spontaneously.

Their next step was the examination of stereoselectivity of the inter-strand aminoacyl transfer in RNA using L- and D-amino acids.^{[213],[81]} Depending on the sequence of the overhang, a stereoselectivity of up to 10:1 in favour of the transfer of the L-amino acid over the D-amino acid was obtained.



Scheme 5: Schematic representation of the nicked loop aminoacyl transfer.

The downside of this aminoacyl transfer strategy is that the aminoacyl-RNA mixed anhydride is unstable and susceptible to decomposition in presence of CO_2 .^[81] As an alternative, Sutherland's group employed the use of RNA-amino acid phosphoramidates, that could react with the 2'/3'-OH of canonical oligonucleotides. Upon activation of the carboxylic acid, the ester formation could take place and subsequently the phosphoramidate could be cleaved under mildly acidic conditions (**Scheme 6**). When L- and D-Ala phosphoramidates were used, a significant stereoselectivity of 9.5:1 in favour of the L-amino acid was noted. The reaction was repeated with a range of amino acid and the preference for the L-amino acid was consistent. Interestingly, yields and stereoselectivity were greater when the reaction was performed under eutectic conditions. Under these conditions, they could observe the stereoselectivity towards L-Val of 48:1 compared to D-Val was observed.^[81]



Scheme 6: Schematic representation of the nicked loop phosphoramidate transfer.

In summary, it has been shown that the RNA, or more specifically the D-ribose dictates the stereoselectivity of either the aminoacylation or the peptide coupling step. This preference for the L-amino acids over D-amino acids is a universal observation and indicates an interdependence between the stereochemistry of the nucleic acid sugar and the amino acids.

2. Aim of the project

Nucleoside modifications are highly abundant in parts of the tRNA and rRNA, playing an indispensable role in structural folding and the maintenance of functionality. Even though the exact moment of their appearance during evolution is unknown, many of them are considered to be “molecular fossils”, due to their suspected presence in the prebiotic world. Nucleosides bearing an amino acid bound to their nucleobase via a urea bridge (e.g. t⁶A), are found in close proximity of the anticodon loop of the tRNA. These nucleoside-amino acid hybrids, an intersection point between genotype and phenotype, are a basis for possible concepts of RNA-peptide co-evolution. Not only are these conjugates relevant in the aspect of the origin of life, but also in the context of the origin of biological homochirality, due to the fact that they consist exclusively of D-ribose and L-amino acids. The requirement of homochirality is predominant across all biology, in order to ensure proper base-pairing and helix formation in nucleic acids and proper folding of proteins into functional higher order structures.

The aim of the work was first to show how the amino acids could have been linked to the nucleobase or RNA strand under conditions, that are compatible with the environment prevalent on the primordial Earth. Once this modified nucleoside could be formed under prebiotically plausible conditions, we were planning to incorporate it into RNA oligonucleotides. For pragmatic reasons, phosphoramidite chemistry and solid-phase oligonucleotide synthesis should be employed to produce these RNA-amino acid hybrid strands in sufficient quantity. Our next goal was to establish a model of RNA-templated amino acid (or further even peptide) transfer, with the amino acid bearing strand acting as a “donor” strand, able to transfer the amino acid onto a complementary “acceptor” strand. In order to eventually release the amino acid from the donor strand and to ensure full amino acid transfer between strands, we were planning to take advantage of the thermolytic properties of the urea linkage.

Furthermore, we aimed to explore the questions of homochirality and its prebiotic origin, particularly in regard to the RNA-peptide synthesis. We were interested in how the amino acid/peptide transfer would be influenced by an exchange of the L-amino acid by the D-counterpart. The idea was to subject the L- vs. D-amino acid modified RNA to direct competition in peptide bond formation between donor and acceptor strands. The thermodynamic and kinetic aspects of this reaction should be examined and also the impact of increased length of the peptide. Finally, we were interested how the stereoselectivity would change if we reacted dipeptides of mixed stereoselectivity or if the RNA was exchanged for DNA. Eventually, the RNA-peptide coupling/cleavage chemistry could be combined with the stereochemical competition reactions into a prebiotically plausible, temperature-driven one-pot reaction cycle.

3. Published results

3.1 A prebiotically plausible scenario of an RNA-peptide world

Authors

Felix Müller*, Luis Escobar*, Felix Xu, Ewa Węgrzyn, Milda Nainytė, Tynchtyk Amatov, Chun-Yin Chan, Alexander Pichler, Thomas Carell[°]

Nature, **2022**, 605, 279-284

DOI: <https://doi.org/10.1038/s41586-022-04676-3>

Supplementary Information included in the Appendix I

* These authors contributed equally [°] corresponding author

Summary

The RNA-world theory is the central concept in the origin of life research, although it does not provide a rationale for the origin of the ribosome and the mechanism of evolution of the translational machinery. Considering the peptide-decorated nucleoside modifications found in the contemporary transfer RNAs (tRNAs), we developed a cycle consisting of a donor/acceptor RNA strand hybrid, able to transfer single or multiple amino acids between the oligonucleotide strands, by forming a hairpin-type intermediate and subsequent thermal urea bond cleavage. The transfer of longer peptides allowed formation of up to a decapeptide on an 11mer oligonucleotide. With this, we were able to perform consecutive one-pot cycles with up three peptide couplings as a model for a primitive peptide-forming machinery.

Personal contribution

Synthesis of modified phosphoramidites and RNA strands, performing the peptide coupling reactions and urea cleavage experiments, analysis of the results and writing of the manuscript.

A prebiotically plausible scenario of an RNA–peptide world

<https://doi.org/10.1038/s41586-022-04676-3>

Received: 13 July 2021

Accepted: 22 February 2022

Published online: 11 May 2022

Open access

 Check for updates

Felix Müller^{1,2}, Luis Escobar^{1,2}, Felix Xu¹, Ewa Węgrzyn¹, Milda Nainytė¹, Tynchtyk Amatov¹, Chun-Yin Chan¹, Alexander Pichler¹ & Thomas Carell¹✉

The RNA world concept¹ is one of the most fundamental pillars of the origin of life theory^{2–4}. It predicts that life evolved from increasingly complex self-replicating RNA molecules^{1,2,4}. The question of how this RNA world then advanced to the next stage, in which proteins became the catalysts of life and RNA reduced its function predominantly to information storage, is one of the most mysterious chicken-and-egg conundrums in evolution^{3–5}. Here we show that non-canonical RNA bases, which are found today in transfer and ribosomal RNAs^{6,7}, and which are considered to be relics of the RNA world^{8–12}, are able to establish peptide synthesis directly on RNA. The discovered chemistry creates complex peptide-decorated RNA chimeric molecules, which suggests the early existence of an RNA–peptide world¹³ from which ribosomal peptide synthesis¹⁴ may have emerged^{15,16}. The ability to grow peptides on RNA with the help of non-canonical vestige nucleosides offers the possibility of an early co-evolution of covalently connected RNAs and peptides^{13,17,18}, which then could have dissociated at a higher level of sophistication to create the dualistic nucleic acid–protein world that is the hallmark of all life on Earth.

A central commonality of all cellular life is the translational process, in which ribosomal RNA (rRNA) catalyses peptide formation with the help of transfer RNAs (tRNA), which function as amino acid carrying adapter molecules^{14,19,20}. Comparative genomics²¹ suggests that ribosomal translation is one of the oldest evolutionary processes^{15,16,22,23}, which dates back to the hypothetical RNA world^{1–4}. The questions of how and when RNA learned to instruct peptide synthesis is one of the grand unsolved challenges in prebiotic evolutionary research^{3–5}.

The immense complexity of ribosomal translation¹⁴ demands a step-wise evolutionary process¹¹. From the perspective of the RNA world, at some point RNA must have gained the ability to instruct and catalyse the synthesis of, initially, just small peptides. This initiated the transition from a pure RNA world¹ into an RNA–peptide world¹³. In this RNA–peptide world, both molecular species could have co-evolved to gain increasing ‘translation’ and ‘replication’ efficiency¹⁷.

To gain insight into the initial processes that may have enabled the emergence of an RNA–peptide world¹³, we analysed the chemical properties of non-canonical nucleosides^{6,7}, which can be traced back to the last universal common ancestor and, as such, are considered to be ‘living molecular fossils’ of an early RNA world^{8–12}.

This approach, which can be called ‘palaeochemistry’, enabled us to learn about the chemical possibilities that existed in the RNA world and, therefore, sets the chemical framework for the emergence of life. In contrast to earlier investigations of the origin of translation^{24–29}, we used naturally occurring non-canonical vestige nucleosides and conditions compatible with aqueous wet–dry cycles^{30,31}.

Peptide synthesis on RNA

In modern tRNAs (Fig. 1a), the amino acids that give peptides are linked to the CCA 3′ terminus via a labile ester group³². Some tRNAs, however,

contain additional amino acids in the form of amino acid-modified nucleosides, for example, g⁶A (ref. 33), t⁶A (ref. 34) and m⁶t⁶A (ref. 35), which are found directly next to the anticodon loop at position 37. Other non-canonical vestige nucleosides often present in the wobble position 34 are nm⁵U and mnm⁵U (refs. 36–38).

Close inspection of their chemical structures (Fig. 1b) suggests that if they are in close proximity (step 1), an RNA-based peptide synthesis may be able to start (step 2), which would create, via a hairpin-type intermediate, a peptide attached by a urea linkage to the nucleobase (m⁶)aa⁶A. Cleavage of the urea^{39,40} (step 3) would furnish RNA with a peptide connected to a (m)nm⁵U (step 4). Subsequently, strand displacement with a new (m⁶)aa⁶A strand may finally enable the next peptide elongation step.

To investigate the potential evolution of an RNA–peptide world, we synthesized two complementary sets of RNA strands, **1a–1j** and **2a–2c** (Fig. 2). The first set contained various m⁶aa⁶A nucleotides⁴¹ at the 5′ end (**1a–1j**) as RNA donor strands. The complementary RNA acceptor strands were prepared with an (m)nm⁵U nucleotide at the 3′ terminus (**2a–2c**). Figure 2a shows the reactions between **1a** and **2a**. The analytical data are presented in Fig. 2b. We hybridized **1a** with **2a** and activated the carboxylic acid of **1a** using reagents such as EDC⁴²/ Sulfo-NHS⁴³, DMTMM·Cl⁴³ or methyl isonitrile⁴⁴ (pH 6, 25 °C). In all cases we observed high yielding product formation (Fig. 2c).

A kinetic analysis shows that the nature of the amino acid affects the coupling rate (Fig. 2d). For example, G (in **1a**) couples to **2c** with an apparent rate constant (k_{app}) of 0.1 h^{−1}. For the amino acids L (in **1d**), I (in **1e**) and M (in **1h**) a fourfold higher rate constant (≈0.4 h^{−1}) was determined, and the highest rate was measured for F (in **1g**) with $k_{app} > 1$ h^{−1}. These differences establish a pronounced amino acid selectivity in the coupling reaction, probably as a result of distinct pre-organizations. We next reduced the length of the RNA donor strand to five, and finally to three, nucleotides (Supplementary Information). We detected coupling even with a trimer

¹Department of Chemistry, Ludwig-Maximilians-Universität (LMU) München, Munich, Germany. ²These authors contributed equally: Felix Müller, Luis Escobar. ✉e-mail: thomas.carell@lmu.de

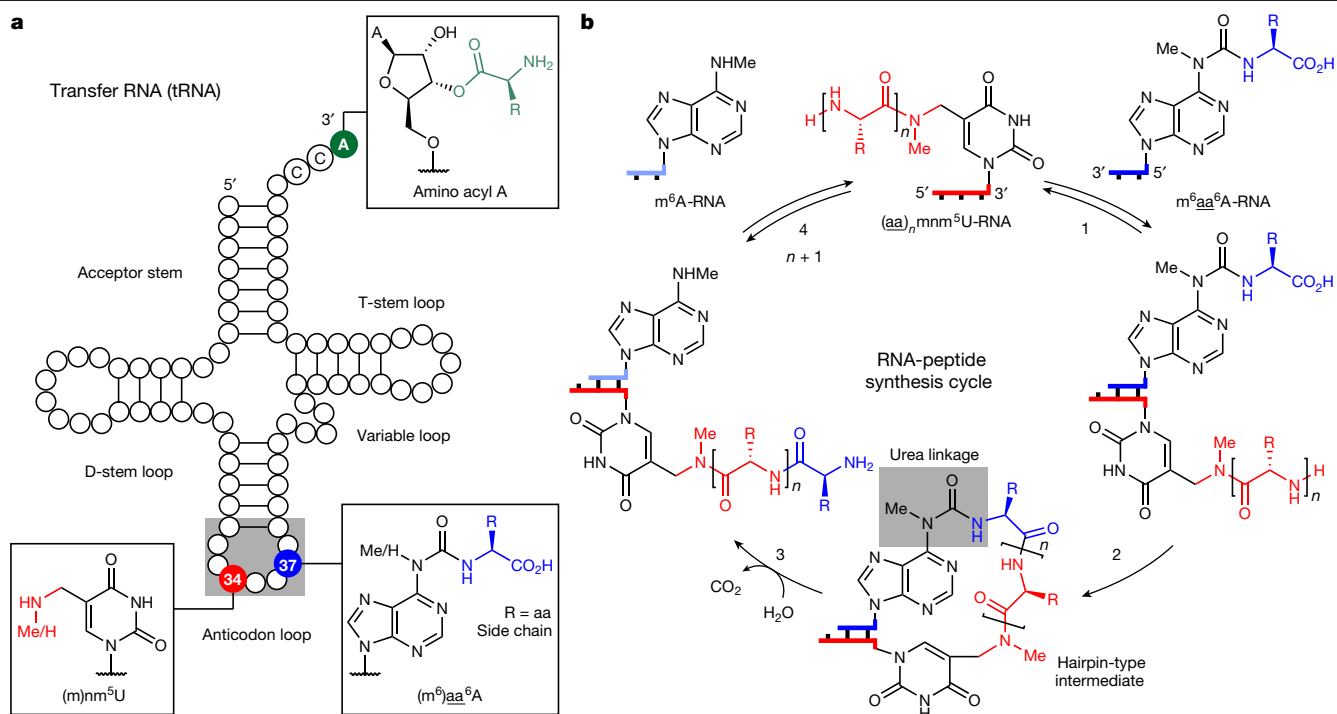


Fig. 1 | Concept of how nucleoside relics of the RNA world enable RNA-based peptide synthesis. a, tRNA structure showing selected ribose and nucleobase modifications. The 3'-amino acid-acylated adenosine is located at the CCA 3' end in contemporary tRNAs. 5-Methylaminomethyl uridine, mnm⁵U, is found in

the wobble position 34. The amino acid-modified carbamoyl adenosine, (m⁶)aa⁶A (aa, amino acid), is present at position 37 in certain tRNAs. **b**, General RNA-peptide synthesis cycle based on mnm⁵U and m⁶aa⁶A. The structures of oligonucleotides are simplified and only terminal nucleobases are drawn.

RNA donor strand, although it required duplex-enforcing high salt and low temperature conditions (1M NaCl and 0 °C). The interaction of three nucleotides on the donor with the corresponding triplet on the acceptor seems to be the lower limit for productive coupling. Interestingly, this is the size of the codon-anticodon interaction in contemporary translation^{11,18}.

We next investigated coupling of the nitrile derivative of **1a** (m⁶g_{CN}⁶A, **1j**) with the different acceptors **2a–2c** under the recently described prebiotically plausible thiol activation conditions⁴⁵ (DTT, pH 8, 25 °C). Here also, the coupling products were obtained within a few hours (Fig. 2c). For example, the combination of nm⁵U **2b** with **1a** gives coupling yields of 64% and 66% using EDC/Sulfo-NHS or DMTMM-Cl, respectively. Coupling of **1a** and **2a**, featuring a secondary amine, afforded **3a** in 16% and 33% yields. The nitrile of **1j** afforded yields of up to 65% after thiol activation coupling.

We next measured the stability of the hairpin-type intermediates. For the hairpin **3a** (Fig. 2a), a melting temperature (T_m) of approximately 87 °C was determined, which in comparison to the starting duplex (approximately 30 °C for **1a-2a**, see Supplementary Information), proves that the peptide formation reaction generated thermally more stable structures. This could have been an advantage during wet-dry cycling under early Earth conditions.

The discovered concept also enabled the synthesis of longer peptides. When we used 3'-ymnm⁵U-RNA-5' **2c** as the acceptor, we observed, on reaction with **1a–1j**, peptide bond formation with up to 77% yield (Fig. 2c, d and Fig. 3a).

We next studied the cleavage of the urea linkage and found that this reaction was possible at elevated temperatures (90 °C) in water at pH 6 (Fig. 2a, b). After 6 h, the products, m⁶A-containing RNA **4** and RNA **5a** were formed already with a yield of 15%.

Longer peptide structures on RNA

We next investigated how the length of the generated peptides influences the coupling reaction (Fig. 3 and Extended Data Fig. 1). For this

study we used synthetic 3'-peptide-mnm⁵U-RNA-5' acceptor strands as starting materials (Supplementary Information). The synthesized acceptor strands were hybridized to the donor strand **1a**. After carboxylic acid activation, rapid formation of elongated hairpin-type intermediates with yields between 40% and 60% was observed (Fig. 3b). We found that the coupling yields did not drop substantially with increasing peptide length, suggesting that other factors, such as the RNA hybridization kinetics, are rate limiting. In all cases, the subsequent urea cleavage (pH 4, 90 °C) affords dipeptide- to hexapeptide-decorated RNAs in 10–15% yield. These modest yields are the result of substantial RNA degradation, driven by the pH and temperature conditions that were used. The decomposition of RNA, however, can be overcome by using 2'-O-methyl nucleotides (see 'Stepwise growth of peptides on RNA'), which are also vestiges of the early RNA world⁴⁶.

During urea cleavage we detected competing formation of hydantoin side products⁴⁷, depending on the pH and temperature (Fig. 3a). Under mildly acidic conditions (pH 6, 90 °C), exclusive formation of the hydantoin product, cyclic-**5c**, was observed. Reducing the temperature and a shift to higher acidity (pH 4, 60 °C) led to the preferential formation of the peptide product, **5c** (approximately 7:1 **5c**:cyclic-**5c** ratio).

Fragment coupling on RNA

We investigated whether longer peptides can also be generated by fragment coupling chemistry with RNA donor strands containing an already longer peptide (m⁶peptide⁶A). This is essential because an RNA-peptide world, with initially low chemical efficiency, might have been limited to the synthesis of smaller peptides. We found that the required adenosine nucleosides, containing a whole peptide attached to the N⁶-position, are available if the peptides that are produced by RNA degradation of the RNA-peptide chimeras, for example, can react with nitrosated N⁶-methylurea adenosine (Fig. 4a). When we treated N⁶-methylurea adenosine with NaNO₂ (5% H₃PO₄) and added the solution to triglycine (pH 9.5), we obtained the peptide-coupled adenosine nucleoside ggg⁶A

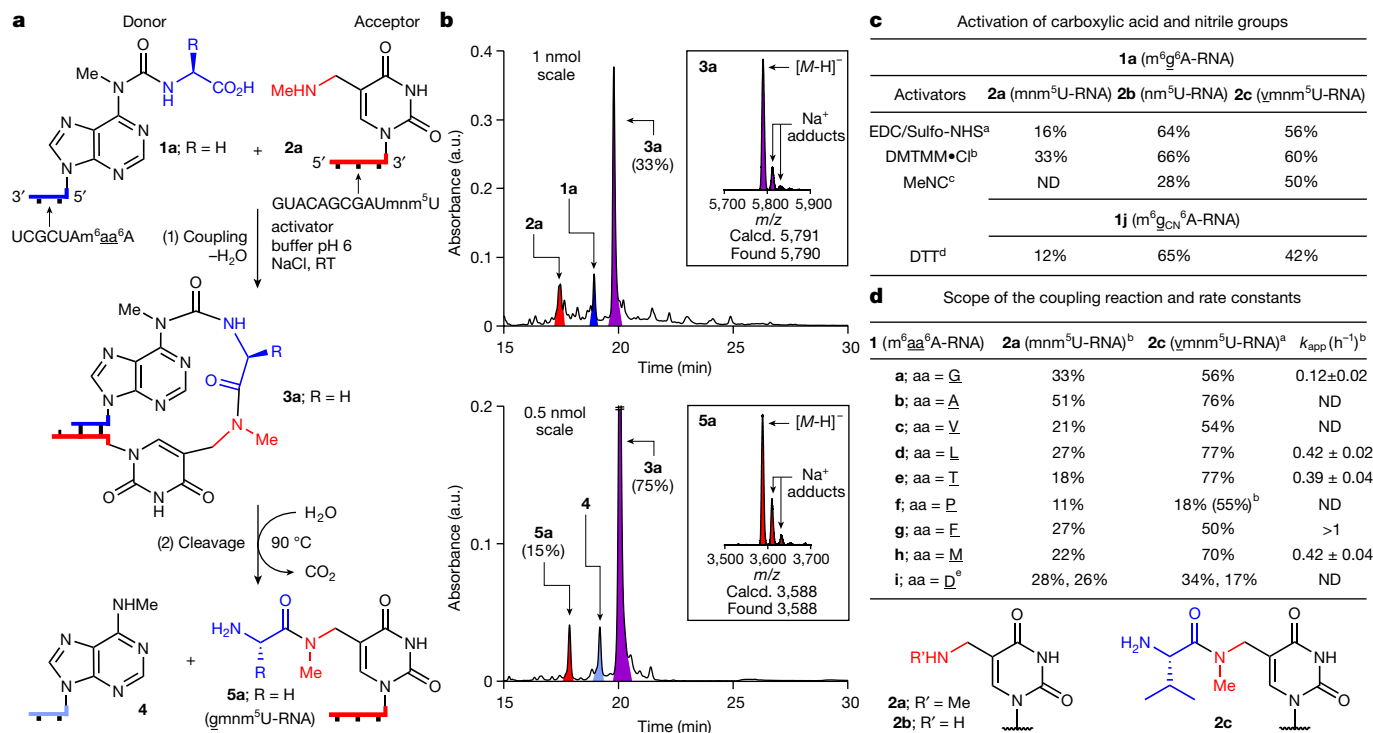


Fig. 2 | Peptide synthesis on RNA with terminal (m)nm⁵U and m⁶aa⁶A nucleotides.

a, Reaction scheme for **1a** (5'-m⁶g⁶A-RNA-3') and **2a** (3'-mnm⁵U-RNA-5') with coupling (1) and cleavage (2). **b**, HPLC chromatograms of the crude reaction mixtures, obtained after coupling of **1a** with **2a** using DMTMM•Cl (see reaction condition b) and cleavage of **3a** (100 mM MES buffer pH 6, 100 mM NaCl, 90 °C, 6 h). HPLC peaks of RNAs are coloured: donor in blue; acceptor in red; hairpin-type intermediate in purple; and cleaved donor strand in pale blue. The insets show MALDI-TOF data (negative mode) of the isolated products **3a** and **5a**. Calcd., calculated. **c**, Coupling results obtained with different activators for **1a** and **1j** with **2a–2c**. **d**, Coupling reactions with

different donors **1a–1i** and acceptors **2a, 2c**, and apparent rate constants (k_{app}) of selected coupling reactions with **2c**. All coupling reactions were carried out using a concentration of 50 μ M for **1a–1j** and 50 μ M for **2a–2c** (100 mM NaCl, 25 °C). ^a50 mM EDC/Sulfo-NHS (100 mM MES buffer pH 6, 24 h). ^b50 mM DMTMM•Cl (100 mM MES buffer pH 6, 24 h). ^c50 mM MeNC (50 mM DCI buffer pH 6, 5 days). ^d50 mM DTT (100 mM borate buffer pH 8, 24 h). ^eThe two yields with **1i** (aa, D) describe the reaction of the aspartic acid α -COOH and of the side chain COOH. An assignment was not performed. RT, room temperature; ND, not determined.

in approximately 65% yield. Incorporation of (m⁶ggg⁶A into RNA and hybridization of this donor strand with a 3'-ggymnm⁵U-RNA-5' acceptor strand furnished, after coupling and urea cleavage, the RNA-peptide chimera 3'-ggggymnm⁵U-RNA-5' (53% coupling, approximately 10% cleavage; Fig. 4b, left). We could also directly transfer longer peptides. When we hybridized the 5'-m⁶gagg⁶A-RNA-3' donor with the 3'-agggymnm⁵U-RNA-5' acceptor, 3'-gagggagggymnm⁵U-RNA-5' was obtained as the product (56% coupling, approximately 9% cleavage; Fig. 4b, right). These experiments suggest the possibility of generating highly complex RNA-peptide chimeras with just a small number of reaction steps⁴⁸.

Multiple peptide growth on RNA

We next investigated whether peptide growth is possible at different RNA positions simultaneously. To this end, we examined the simultaneous binding of different donor strands to one or two acceptor strands. We hybridized two donor strands (7-mer: 5'-m⁶g⁶A-RNA-3' and 10-mer: 5'-m⁶v⁶A-RNA-3') to a single RNA acceptor strand (21-mer) with a central gnm⁵U and a 3' terminal nm⁵U (Fig. 5a, left). On activation of the carboxylic acids, a GG-dipeptide was synthesized in the centre of the RNA, whereas a valine amino acid was attached to the 3' end of the acceptor strand. In a different experiment, we hybridized an RNA donor strand (22-mer), containing both a 3'-m⁶g⁶A and a 5'-m⁶v⁶A, to two different acceptor RNAs, containing a central ynm⁵U (21-mer) and a 3' terminal ynm⁵U (11-mer) (Fig. 5a, right). On activation, we observed formation of a central GV- and a terminal VV-dipeptide.

Effect of base pairing

To investigate the importance of sequence complementarity, we added two RNA donor strands of different lengths (7-mer: 5'-m⁶g⁶A-RNA-3' and 11-mer: 5'-m⁶v⁶A-RNA-3') to an acceptor strand with a ynm⁵U at the 3' end (11-mer: **2c**) (Fig. 5b, left). On the basis of the melting temperatures of the two possible duplexes (approximately 30 °C for the 7-mer-11-mer and 59 °C for the 11-mer-11-mer, see Supplementary Information), only formation of the VV-dipeptide RNA conjugate, derived from the thermodynamically more stable duplex, was observed. Finally, we mixed two RNA donor strands of identical length (7-mer). The first contained a 5'-m⁶l⁶A and the second a 5'-m⁶g⁶A, together with two mismatches. We added this mixture to an RNA acceptor strand (11-mer: **2c**) with a 3'-ymnm⁵U nucleotide (Fig. 5b, right). In this experiment, exclusive formation of the LV-dipeptide was found, generated from the fully complementary strands and thus the more stable duplex. Collectively, these results support that full complementarity is needed for efficient peptide synthesis.

Stepwise growth of peptides on RNA

We finally investigated whether one-pot stepwise growth of a peptide on RNA is possible (Fig. 5c). To increase the stability of the RNA towards phosphodiester hydrolysis, as needed for this experiment, we used the RNA acceptor strand **2g**, in which the contemporary canonical bases were replaced by the non-canonical 2'-OMe nucleotides: A_m, C_m, G_m and U_m. The strand **2g** was equipped with an additional 3'-mnm⁵U

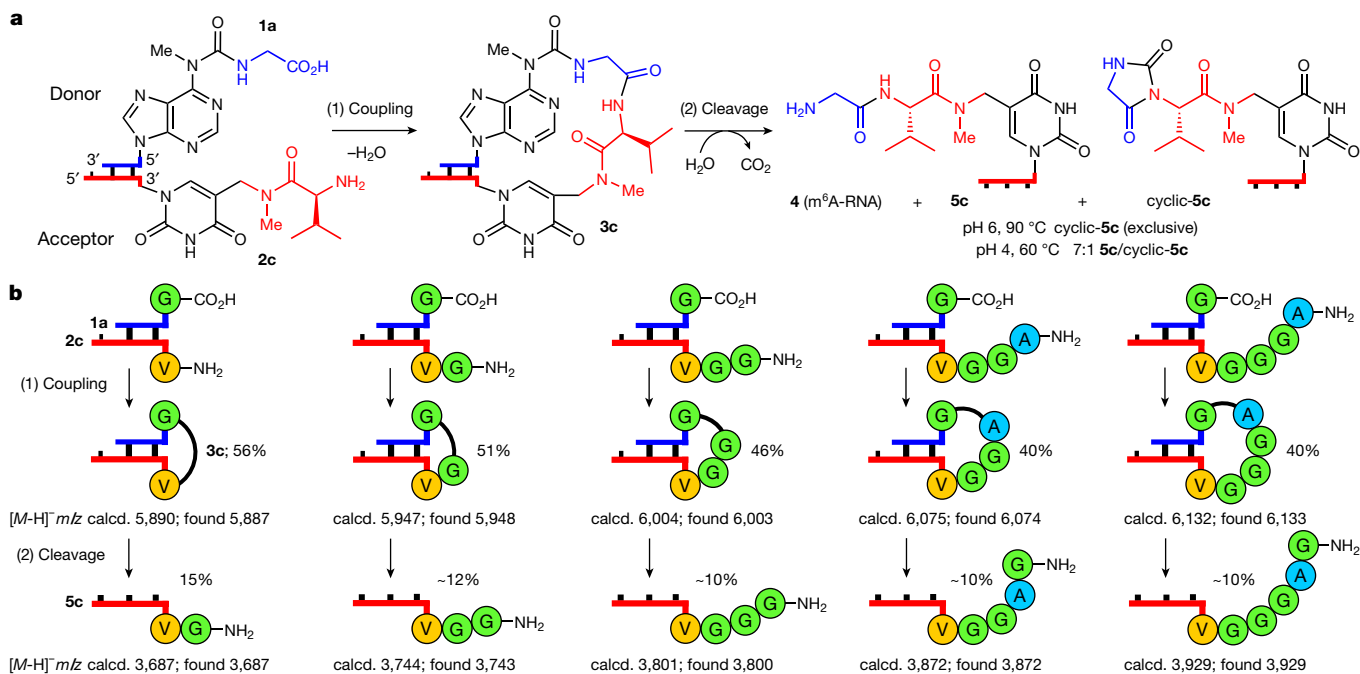


Fig. 3 | Growth of longer peptide structures on RNA. **a**, Scheme for the reaction of **1a** ($5'\text{-m}^6\text{g}^{\text{A}}\text{-RNA-3}'$) with **2c** ($3'\text{-ymmm}^3\text{U-RNA-5}'$) including coupling (1) and cleavage (2). **b**, Coupling reactions between **1a** and RNA-peptide acceptor strands using EDC/Sulfo-NHS (see reaction condition a in Fig. 2) and

cleavage reactions of the coupled compounds (100 mM acetate buffer pH 4, 100 mM NaCl, 90 °C, 6 h). MALDI-TOF data (negative mode) of the isolated products are given.

nucleotide. For the experiment we used the same amount of donor strand for all coupling steps and performed filtration steps to remove remaining activator. After two couplings, two urea cleavages and two

filtrations, we observed, by high-performance liquid chromatography (HPLC) analysis, the presence of the product $3'\text{-ggmm}^5\text{U-RNA-5}'$ **7g** (Fig. 5c, left). The circumvented material consuming isolation steps

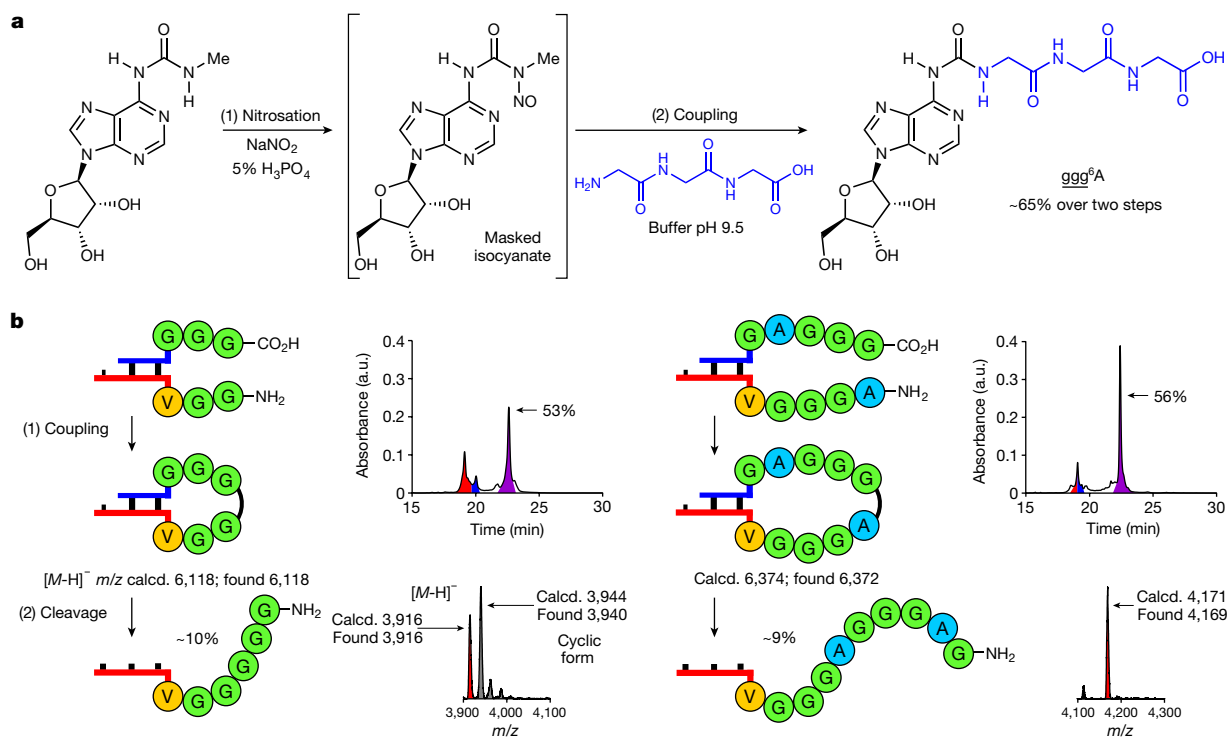


Fig. 4 | Capture of peptides by nitrosated N^6 -methylurea adenosine for fragment condensation. **a**, Prebiotically plausible formation of peptide^A structures, such as ggg^6A . **b**, Coupling reactions between RNA-peptide conjugates using EDC/Sulfo-NHS (see reaction condition a in Fig. 2) and cleavage reactions of the coupled compounds (see reaction conditions in

Fig. 3). HPLC chromatograms show the crude mixtures of the coupling reactions. The RNA signals are coloured: donor in blue; acceptor in red; and hairpin-type intermediate in purple. MALDI-TOF data (negative mode) are shown for the isolated products, together with the $5'\text{-m}^6\text{A-RNA-3}'$ strand **4** and the hydantoin side product (cyclic form) in the case indicated.

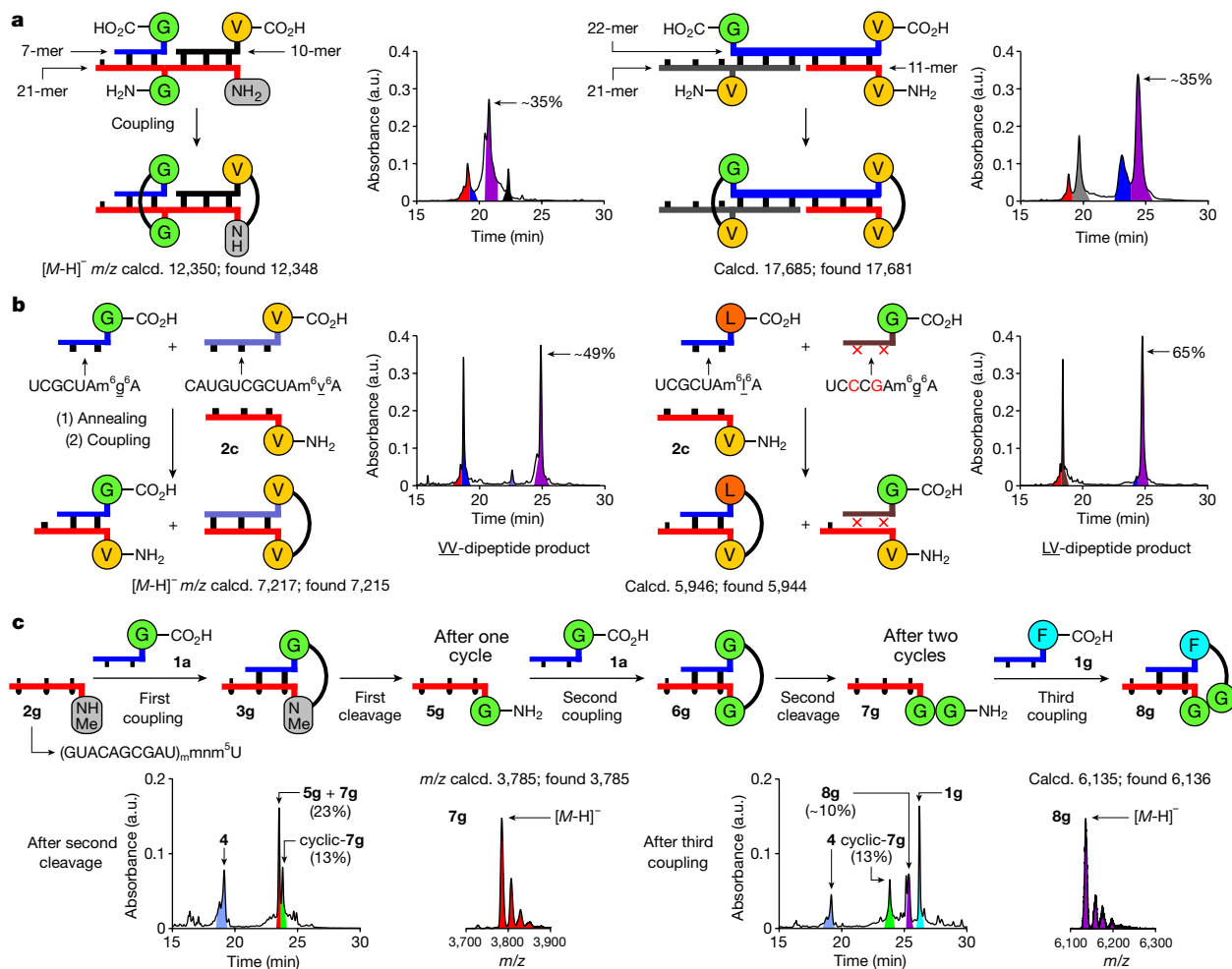


Fig. 5 | Parallel growth of peptides at various positions on RNA, effect of base pairing and RNA-peptide synthesis cycles. a, Coupling of oligonucleotides containing multiple donor or acceptor units (EDC/Sulfo-NHS, see reaction condition a in Fig. 2). **b**, Annealing followed by coupling (EDC/Sulfo-NHS, see reaction condition a in Fig. 2) of an acceptor strand with donor strands of different length (left) or base sequence (right). **c**, Two RNA-peptide synthesis cycles with a third coupling step using a 2'-OMe acceptor strand

and performed under one-pot conditions with intermediary filtration to remove the remaining activator (coupling: DMTMM·Cl, see reaction condition b in Fig. 2; cleavage: 100 mM acetate buffer pH 4, 100 mM NaCl, 90 °C, 24 h; MES buffer pH 6 was used in the first cleavage reaction). HPLC chromatograms show the crude mixtures of the coupling and cleavage reactions. Peaks of RNA strands are coloured as in the reaction scheme. MALDI-TOF data (negative mode) of the isolated products are given.

(Extended Data Fig. 2) enabled us to obtain the product in an overall yield of about 18%. A final, third coupling reaction with the 5'-m⁶f⁶A donor strand **1g** furnished the FGG-hairpin intermediate **8g** in approximately 10% overall yield (Fig. 5c, right).

We next studied fragment condensation with the 5'-m⁶ggg⁶A-RNA-3' donor strand and the complementary 3'-aggmm⁵U-RNA-5' acceptor strand, consisting only of 2'-OMe nucleotides. Here, coupling with approximately 50% and urea cleavage with approximately 85% generated the product 3'-gggaggmm⁵U-RNA-5', together with some of the hydantoin side product (Supplementary Information). Together these data show that, with the help of 2'-OMe nucleotides, peptides can grow on RNA in a stepwise fashion and via fragment condensation to generate higher complexity.

Discussion

The plausible formation of catalytically competent and self-replicating RNA structures without the aid of proteins is one of the major challenges for the model of the RNA world¹⁻⁴. It is difficult to imagine how an RNA world with complex RNA molecules could have emerged without the help of proteins and it is hard to envision how such an RNA world

transitions into the modern dualistic RNA and protein world, in which RNA predominantly encodes information whereas proteins are the key catalysts of life.

We found that non-canonical vestige nucleosides⁸⁻¹², which are key components of contemporary RNAs^{6,7}, are able to equip RNA with the ability to self-decorate with peptides. This creates chimeric structures, in which both chemical entities can co-evolve in a covalently connected form¹³, generating gradually more and more sophisticated and complex RNA-peptide structures. Although, in this study, we observe peptide coupling on RNA in good yields, the efficiency will certainly improve if we allow optimization of the structures and sequences of the RNA-peptides by chemical evolution. The simultaneous presence of the chemical functionalities of RNA and amino acids certainly increases the chance of generating catalytically competent structures. The stabilization of RNA by incorporation of 2'-OMe nucleotides significantly improved the urea cleavage yield.

Interestingly, in the coupling step we observed large differences in the rate constants, which suggests that our system has the potential to preferentially generate certain peptides. We also found that peptides can simultaneously grow at multiple sites on RNA on the basis of rules determined by sequence complementarity, which is the indispensable requirement for efficient peptide growth.

All these data together support the idea that non-canonical vestige nucleosides in RNA have the potential to create peptide self-decorating RNAs and hence an RNA–peptide world. The formed RNA–peptide chimeras are comparatively stable, and so it is conceivable that some of these structures learned, at some point, to activate amino acids by adenylation⁴⁹ and to transfer them onto the ribose OH groups⁵⁰ to capture the reactivity in structures that were large and hydrophobic enough to exclude water. This would then have been the transition from the non-canonical nucleoside-based RNA–peptide world to the ribosome-centred translational process that is a hallmark of all life on Earth today.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-04676-3>.

- Gilbert, W. Origin of life: the RNA world. *Nature* **319**, 618 (1986).
- Orgel, L. E. Evolution of the genetic apparatus. *J. Mol. Biol.* **38**, 381–393 (1968).
- Crick, F. H. C., Brenner, S., Klug, A. & Piecznik, G. A speculation on the origin of protein synthesis. *Orig. Life Evol. Biosph.* **7**, 389–397 (1976).
- Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **418**, 214–221 (2002).
- Bowman, J. C., Hud, N. V. & Williams, L. D. The ribosome challenge to the RNA world. *J. Mol. Evol.* **80**, 143–161 (2015).
- Decatur, W. A. & Fournier, M. J. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **27**, 344–351 (2002).
- Carell, T. et al. Structure and function of noncanonical nucleobases. *Angew. Chem. Int. Ed. Engl.* **51**, 7110–7131 (2012).
- Wong, J. T.-F. Origin of genetically encoded protein synthesis: a model based on selection for RNA peptidation. *Orig. Life Evol. Biosph.* **21**, 165–176 (1991).
- Di Giulio, M. Reflections on the origin of the genetic code: a hypothesis. *J. Theor. Biol.* **191**, 191–196 (1998).
- Rios, A. C. & Tor, Y. On the origin of the canonical nucleobases: an assessment of selection pressures across chemical and early biological evolution. *Isr. J. Chem.* **53**, 469–483 (2013).
- Grosjean, H. & Westhof, E. An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.* **44**, 8020–8040 (2016).
- Beenstock, J. & Sicheri, F. The structural and functional workings of KEOPS. *Nucleic Acids Res.* **49**, 10818–10834 (2021).
- Di Giulio, M. On the RNA world: evidence in favor of an early ribonucleopeptide world. *J. Mol. Evol.* **45**, 571–578 (1997).
- Ramakrishnan, V. Ribosome structure and the mechanism of translation. *Cell* **108**, 557–572 (2002).
- Fox, G. E. Origin and evolution of the ribosome. *Cold Spring Harb. Perspect. Biol.* **2**, a003483 (2010).
- Bowman, J. C., Petrov, A. S., Frenkel-Pinter, M., Penev, P. I. & Williams, L. D. Root of the tree: the significance, evolution, and origins of the ribosome. *Chem. Rev.* **120**, 4848–4878 (2020).
- Eigen, M. & Schuster, P. A principle of natural self-organization. *Naturwissenschaften* **64**, 541–565 (1977).
- Szathmáry, E. Coding coenzyme handles: a hypothesis for the origin of the genetic code. *Proc. Natl Acad. Sci. USA* **90**, 9916–9920 (1993).
- Noller, H. F. RNA structure: reading the ribosome. *Science* **309**, 1508–1514 (2005).
- Steitz, T. A. A structural understanding of the dynamic ribosome machine. *Nat. Rev. Mol. Cell Biol.* **9**, 242–253 (2008).
- Koonin, E. V. Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat. Rev. Microbiol.* **1**, 127–136 (2003).
- Woese, C. The universal ancestor. *Proc. Natl Acad. Sci. USA* **95**, 6854–6859 (1998).
- Becerra, A., Delaye, L., Islas, S. & Lazcano, A. The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains. *Annu. Rev. Ecol. Evol. Syst.* **38**, 361–379 (2007).
- Kuhn, H. Self-organization of molecular systems and evolution of the genetic apparatus. *Angew. Chem. Int. Ed. Engl.* **11**, 798–820 (1972).
- Kuhn, H. & Waser, J. Molecular self-organization and the origin of life. *Angew. Chem. Int. Ed. Engl.* **20**, 500–520 (1981).
- Tamura, K. & Schimmel, P. Oligonucleotide-directed peptide synthesis in a ribosome- and ribozyme-free system. *Proc. Natl Acad. Sci. USA* **98**, 1393–1397 (2001).
- Tamura, K. & Schimmel, P. Peptide synthesis with a template-like RNA guide and aminoacyl phosphate adaptors. *Proc. Natl Acad. Sci. USA* **100**, 8666–8669 (2003).
- Turk, R. M., Chumachenko, N. V. & Yarus, M. Multiple translational products from a five-nucleotide ribozyme. *Proc. Natl Acad. Sci. USA* **107**, 4585–4589 (2010).
- Jash, B., Tremmel, P., Jovanovic, D. & Richert, C. Single nucleotide translation without ribosomes. *Nat. Chem.* **13**, 751–757 (2021).
- Forsythe, J. G. et al. Ester-mediated amide bond formation driven by wet–dry cycles: a possible path to polypeptides on the prebiotic Earth. *Angew. Chem. Int. Ed. Engl.* **54**, 9871–9875 (2015).
- Becker, S. et al. Wet-dry cycles enable the parallel origin of canonical and non-canonical nucleosides by continuous synthesis. *Nat. Commun.* **9**, 163 (2018).
- Tetzlaff, C. N. & Richert, C. Synthesis and hydrolytic stability of 5'-aminoacylated oligouridylic acids. *Tetrahedron Lett.* **42**, 5681–5684 (2001).
- Schweizer, M. P., McGrath, K. & Baczyński, L. The isolation and characterization of N-[9-(β-D-ribofuranosyl)-purin-6-ylcarbamoyl]glycine from yeast transfer RNA. *Biochem. Biophys. Res. Commun.* **40**, 1046–1052 (1970).
- Perrochia, L. et al. *In vitro* biosynthesis of a universal t⁶A tRNA modification in Archaea and Eukarya. *Nucleic Acids Res.* **41**, 1953–1964 (2012).
- Kimura-Harada, F., Von Minden, D. L., McCloskey, J. A. & Nishimura, S. N-[(9-β-D-Ribofuranosyl)purin-6-yl]-N-methylcarbamoyl]threonine, a modified nucleoside isolated from *Escherichia coli* threonine transfer ribonucleic acid. *Biochemistry* **11**, 3910–3915 (1972).
- Robertson, M. & Miller, S. Prebiotic synthesis of 5-substituted uracils: a bridge between the RNA world and the DNA-protein world. *Science* **268**, 702–705 (1995).
- Murphy, F. V., Ramakrishnan, V., Malkiewicz, A. & Agris, P. F. The role of modifications in codon discrimination by tRNA^{Val}. *Nat. Struct. Mol. Biol.* **11**, 1186–1191 (2004).
- Kitamura, A. et al. Characterization and structure of the *Aquifex aeolicus* protein DUF752: a bacterial tRNA-methyltransferase (MnmC2) functioning without the usually fused oxidase domain (MnmC1). *J. Biol. Chem.* **287**, 43950–43960 (2012).
- Hutchby, M. et al. Hindered ureas as masked isocyanates: facile carbamoylation of nucleophiles under neutral conditions. *Angew. Chem. Int. Ed. Engl.* **48**, 8721–8724 (2009).
- Ohkubo, A. et al. New thermolytic carbamoyl groups for the protection of nucleobases. *Org. Biomol. Chem.* **7**, 687–694 (2009).
- Nainytė, M. et al. Amino acid modified RNA bases as building blocks of an early Earth RNA-peptide world. *Chem. Eur. J.* **26**, 14856–14860 (2020).
- Schimpl, A., Lemmon, R. M. & Calvin, M. Cyanamide formation under primitive Earth conditions. *Science* **147**, 149–150 (1965).
- Gartner, Z. J., Kanan, M. W. & Liu, D. R. Expanding the reaction scope of DNA-templated synthesis. *Angew. Chem. Int. Ed. Engl.* **41**, 1796–1800 (2002).
- Liu, Z. et al. Harnessing chemical energy for the activation and joining of prebiotic building blocks. *Nat. Chem.* **12**, 1023–1028 (2020).
- Foden, C. S. et al. Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water. *Science* **370**, 865–869 (2020).
- Schneider, C. et al. Noncanonical RNA nucleosides as molecular fossils of an early Earth—generation by prebiotic methylations and carbamoylations. *Angew. Chem. Int. Ed. Engl.* **57**, 5943–5946 (2018).
- Danger, G., Plasson, R. & Pascal, R. Pathways for the formation and evolution of peptides in prebiotic environments. *Chem. Soc. Rev.* **41**, 5416–5429 (2012).
- Bondalapati, S., Jbara, M. & Brik, A. Expanding the chemical toolbox for the synthesis of large and uniquely modified proteins. *Nat. Chem.* **8**, 407–418 (2016).
- Berg, P. The chemical synthesis of amino acyl adenylates. *J. Biol. Chem.* **233**, 608–611 (1958).
- Wu, L.-F., Su, M., Liu, Z., Bjork, S. J. & Sutherland, J. D. Interstrand aminoacyl transfer in a tRNA acceptor stem-overhang mimic. *J. Am. Chem. Soc.* **143**, 11836–11842 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

Methods

General method for the peptide coupling reactions

The RNA donor and acceptor strands (1:1 ratio, 5 nmol of each strand) were annealed with NaCl (5 μ l from a 1 M aqueous solution) by heating at 95 °C for 4 min, followed by cooling down slowly to room temperature. After that, MES buffer pH 6 (25 μ l from a 400 mM aqueous solution) and NaCl (5 μ l from a 1 M aqueous solution) were added to the oligonucleotide solution. Finally, carboxylic acid or nitrile activator/s (10 μ l of each component from a 500 mM aqueous solution) and water (100 μ l of total reaction volume) were added to the solution mixture. The peptide coupling reaction was incubated at 25 °C for 24 h. The crude reaction mixtures were analysed by HPLC and MALDI-TOF mass spectrometry.

General method for the urea cleavage reactions

The hairpin-type intermediate (0.5 nmol) was diluted with MES buffer pH 6 or acetate buffer pH 4 (12.5 μ l from a 400 mM aqueous solution), NaCl (5 μ l from a 1 M aqueous solution) and water (50 μ l of total reaction volume). The urea cleavage reaction was incubated at 60–90 °C at different time intervals. The crude reaction mixtures were analysed by HPLC and MALDI-TOF mass spectrometry.

Data availability

The data that support the findings of this study are available within the paper and its Supplementary Information.

Acknowledgements We thank the Deutsche Forschungsgemeinschaft for supporting this research through the DFG grants: CA275/11-3 (ID: 326039064), CRC1309 (ID: 325871075, A4), CRC1032 (ID: 201269156, A5) and CRC1361 (ID: 393547839, P2). We thank the Volkswagen Foundation for funding this research (grant EvoRib). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement no. 741912 (EPiR) and under the Marie Skłodowska-Curie grant agreement no. 861381 (Nature-ETN). L.E. thanks the Alexander von Humboldt Foundation for a postdoctoral fellowship (ESP 1214218 HFST-P).

Author contributions F.M., L.E., F.X. and E.W. synthesized the modified phosphoramidites and RNA strands and performed the peptide coupling and urea cleavage experiments. M.N. synthesized RNA donor strands and performed preliminary experiments. T.A. refined and developed mechanistic concepts and performed initial proof-of-principle studies. C.-Y.C. and A.P. synthesized modified phosphoramidites. T.C. conceived the project and directed the research. All authors contributed to the analysis of the results and writing of the manuscript.

Competing interests The authors declare no competing interests.

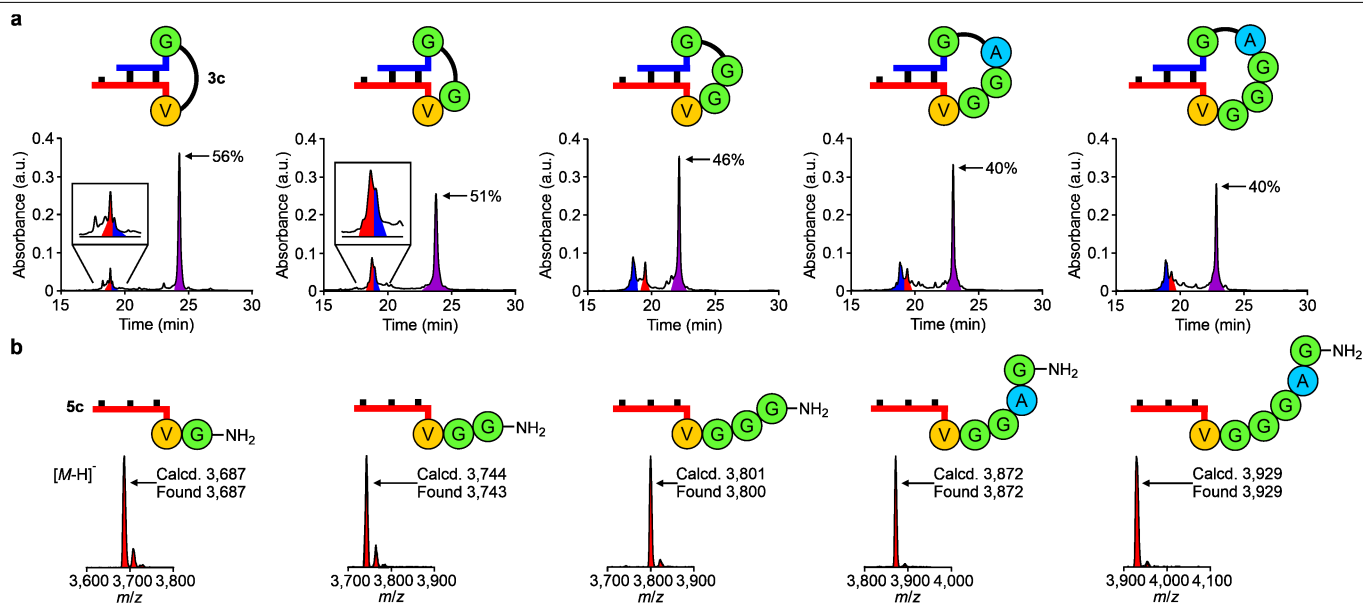
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-04676-3>.

Correspondence and requests for materials should be addressed to Thomas Carell.

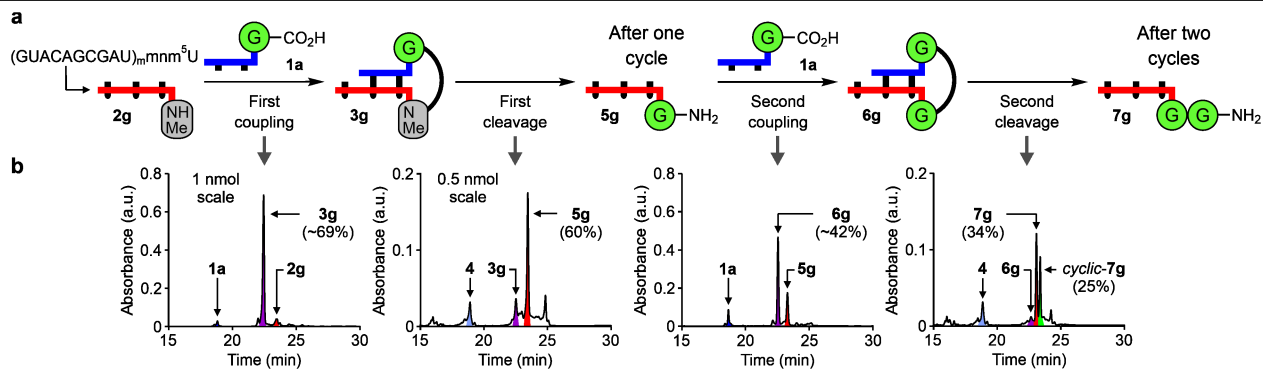
Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Analytical data of the growth of longer peptides on RNA. **a**, HPLC chromatograms show the crude mixtures of the coupling reactions (100 mM MES buffer pH 6, 100 mM NaCl, 50 mM EDC/Sulfo-NHS, 25 °C, 24 h) between 5'-m⁶g⁶A-RNA-3' **1a** and RNA-peptide acceptor strands. **b**, MALDI-TOF mass spectra (negative mode) are shown for the isolated

products obtained after the cleavage reactions (100 mM acetate buffer pH 4, 100 mM NaCl, 90 °C, 6 h) of the coupled compounds. In the HPLCs, the RNA strands are coloured: donor in blue; acceptor in red and hairpin-type intermediate in purple.



Extended Data Fig. 2 | RNA-peptide synthesis cycles using a 2'-OMe acceptor strand. **a**, Two RNA-peptide synthesis cycles in which the product of each step was separated and added into the next reaction (coupling conditions: 100 mM MES buffer pH 6, 100 mM NaCl, 50 mM DMTMM•Cl, 25 °C, 24 h;

cleavage conditions: 100 mM acetate buffer pH 4, 100 mM NaCl, 90 °C, 24 h). **b**, HPLC chromatograms show the crude mixtures of the coupling and cleavage reactions. In the HPLCs, peaks of RNA strands are coloured as in the reaction scheme. The product 3'-ggmnm⁵U-RNA-5' **7g** was obtained in ≈ 6% overall yield.

3.2 Loading of Amino Acids onto RNA in a Putative RNA-Peptide World

Authors

Johannes N. Singer*, Felix M. Müller*, Ewa Węgrzyn, Christina Hölzl, Hans Hurmiz, Chuyi Liu, Luis Escobar[°] und Thomas Carell[°]

Angew. Chem. Int. Ed., **2023**, 62, e202302360

DOI: <https://doi.org/10.1002/anie.202302360>

Supplementary Information included in the Appendix II

* These authors contributed equally [°] corresponding authors

Summary

We have previously shown the relevance of non-canonical nucleotides in the context of RNA-based peptide synthesis. The RNA can self-decorate with peptides through an RNA-templated transfer on amino acids, but in order to be able to perform this reactions, amino acid modified nucleosides have to be present. In this work, we show a prebiotically plausible pathway of attaching amino acids to N⁶-methylcarbamoyl adenosine, which would represent the donating part in the peptide synthesis cycle. We show that a variety on amino acids can be loaded onto a nucleoside or even onto RNA. We make use of fluctuating temperatures and pH values, providing an efficient way to create those amino acid-RNA hybrids.

Personal contribution

Synthesis of Lys compound and measurement of the calibration curve, participation in the writing of the supporting information and proofreading of the manuscript.



Loading of Amino Acids onto RNA in a Putative RNA-Peptide World

Johannes N. Singer[†], Felix M. Müller[†], Ewa Węgrzyn, Christina Hölzl, Hans Hurmiz, Chuyi Liu, Luis Escobar,* and Thomas Carell*

Abstract: RNA is a molecule that can both store genetic information and perform catalytic reactions. This observed dualism places RNA into the limelight of concepts about the origin of life. The RNA world concept argues that life started from self-replicating RNA molecules, which evolved toward increasingly complex structures. Recently, we demonstrated that RNA, with the help of conserved non-canonical nucleosides, which are also putative relics of an early RNA world, had the ability to grow peptides covalently connected to RNA nucleobases, creating RNA-peptide chimeras. It is conceivable that such molecules, which combined the information-coding properties of RNA with the catalytic potential of amino acid side chains, were once the structures from which life emerged. Herein, we report prebiotic chemistry that enabled the loading of both nucleosides and RNAs with amino acids as the first step toward RNA-based peptide synthesis in a putative RNA-peptide world.

acid to its cognate tRNA (aminoacylation) requires the activation of the amino acid as an adenylate^[3] and a dedicated aminoacyl tRNA synthetase,^[4–6] which transfers the amino acid to the 3'-CCA end of the tRNA by catalyzing the formation of an ester bond. Although tRNAs are considered to be highly conserved structures,^[7–10] the emergence of this complex process is still one of the main unsolved questions in origin of life research.

The main problem for chemical evolution is that complex chemical structures, such as mRNAs, ribosomes and amino acid loaded tRNAs, are required to generate a machinery that can efficiently form peptide bonds. But how could such complex systems have evolved in the absence of efficient peptide synthesis machines? To solve this chicken-and-egg conundrum, we recently postulated that RNA-amino acid conjugates could have been a starting point for evolving more complex systems.^[11] We learned that RNA was capable of self-decorating with peptides with the help of non-canonical nucleosides,^[12] that is, *N*⁶-methylated derivatives of glycine- and threonine-modified *N*⁶-carbamoyl adenosine, **1a** (*g*⁶A)^[13] and **1b** (*t*⁶A).^[14] In addition, **1a–b** are found in contemporary tRNAs in all three kingdoms of life^[15,16] as potential relics of an early RNA world.^[17,18] In aqueous solution, these compounds exhibited a good stability over a wide pH range at moderate temperatures due to the urea bond.^[19,20] In the presence of a second non-canonical nucleoside, namely, 5-methylaminomethyl uridine (*mm*⁵U),^[21,22] and when both nucleosides were placed in close proximity by RNA hybridization,^[23] we observed the formation of peptides. This RNA-templated reaction established a primitive peptide synthesis cycle that afforded RNAs decorated with peptides. Although our system is far away from a translational machinery, it showed that RNA-peptide chimeras, having, in principle, information-encoding properties and an expanded repertoire of functional groups, could have existed in a prebiotic world.

We would like to emphasize that the obtained RNA-peptide chimeras are not considered by us to be the direct precursors of the contemporary ribosome. In this regard, seminal works by other research groups showed that amino acids could be connected to 5'-phosphate groups as acyl phosphate mixed anhydrides^[24] or phosphoramidates,^[25–31] and transferred to the ribose 2'/3'-hydroxy group^[32] of terminal nucleotides in RNAs.^[33–40]

The first step of the described peptide growth on RNA, however, required the efficient loading of amino acids onto *N*⁶-carbamoyl adenosine nucleotides. Previously, we reported that amino acids, namely, Gly and Thr, could react

Introduction

The essential biological process of translation converts a genotype into a phenotype. It is based on information-encoding messenger RNA (mRNA), the ribosome as the catalytic entity, and transfer RNAs (tRNAs) that carry specific amino acids and bind to the information-encoding units on the mRNA.^[1,2] The attachment of an specific amino

[*] J. N. Singer,[†] F. M. Müller,[†] E. Węgrzyn, Dr. C. Hölzl, H. Hurmiz, C. Liu, Dr. L. Escobar, Prof. Dr. T. Carell
 Department of Chemistry, Ludwig-Maximilians-Universität (LMU) München
 Butenandtstrasse 5–13, 81377 Munich (Germany)
 E-mail: luisescobar1992@hotmail.es
 thomas.carell@lmu.de

Dr. C. Hölzl

Present address: AbbVie Deutschland GmbH & Co. KG
 Knollstrasse 50, 67061 Ludwigshafen am Rhein (Germany)

[†] These authors contributed equally to this work.

© 2023 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

with methylisocyanate (MIC) affording *N*-methylcarbamoyl amino acids in quantitative yields (Scheme 1a).^[41] After nitrosation of the *N*-methylcarbamoyl amino acids, the obtained *N*-nitroso derivatives decomposed at pH 8 and 70 °C to give *N*-isocyanates, which reacted with adenosine to provide the required amino acid modified *N*⁶-carbamoyl (**1a** and **1b**) and *O*-carbamoyl nucleosides. While this chemistry rested on prebiotically plausible starting materials, it had the drawback that the yields were sensitive to the exact reaction conditions and fairly low.

Herein, we show that a prebiotically more plausible loading of amino acids is possible with *N*⁶-methylcarbamoyl adenosine **2** (Scheme 1b). The loading reaction was so efficient that it was also possible when **2** was embedded in RNA strands. The reported chemistry could be expanded to *N*²-methylcarbamoyl guanosine **3** and *N*⁴-methylcarbamoyl cytidine **4**. The obtained results suggest that the loading of amino acids onto RNAs constitutes an additional property, along with information-encoding, which needs to be considered in the context of prebiotic chemistry.

Results and Discussion

Prebiotically Plausible Synthesis of *N*⁶-Methylcarbamoyl Adenosine

Initially, we investigated the formation of the urea-modified adenosine **2** from either 1,3-dimethylurea (DMU) or methylisocyanate (MIC) under prebiotically plausible reaction conditions (Scheme 1b). Both DMU and MIC are prebiotically plausible starting materials. MIC could be generated from CH₄ and HNCO by vacuum-UV light irradiation at 20 K.^[42] Moreover, MIC was detected in the comet 67P/Churyumov-Gerasimenko^[43] and in the protostar IRAS 16293-2422.^[42,44] In turn, DMU was formed by continuous flow plasma discharge experiments from gas mixtures (N₂/CO/CO₂/H₂).^[45] It is also possible that DMU was produced

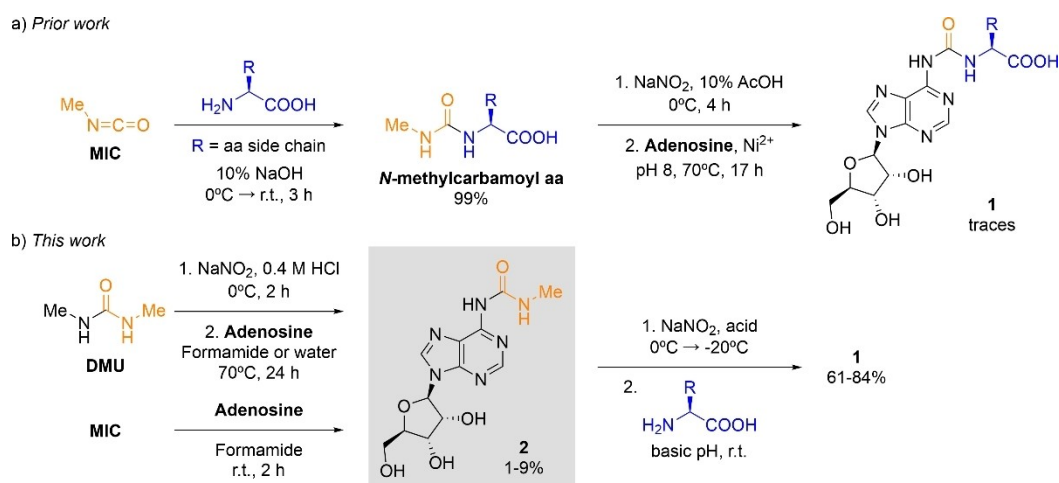
by the combination of MIC with methylamine, which was also present on the above-mentioned comet.^[43]

In a 0.4 M HCl aqueous solution at 0 °C, the reaction of DMU with 1 equiv of NaNO₂^[46–48] gave its *N*-nitroso derivative in 98 % yield.^[49] Subsequently, the addition of adenosine to a formamide^[43,50] or an aqueous solution, containing 50 equiv of the *N*-nitroso derivative of DMU, followed by heating at 70 °C afforded the *N*⁶-carbamoylated adenosine **2** in up to 2 % yield. The high-performance liquid chromatography–mass spectrometry (HPLC-MS) analyses of the crude reaction mixtures showed the formation of **2**, as well as other adenosine derivatives having either methyl or *N*-methylcarbamoyl substituents, or both (Figure S1a–b). Most likely, the isomeric products of **2** had the *N*-methylcarbamoyl substituent at one of the ribose hydroxy groups, that is, *O*-methylcarbamoyl adenosine derivatives. In formamide solution at room temperature (r.t.), the reaction of adenosine with 5.5 equiv of MIC yielded **2** in a maximum amount of 9 % (Figure S2). Taken together, these results showed that two prebiotically plausible synthetic pathways could provide the urea-modified adenosine **2**.

Next, we probed the stability of the *N*⁶-methylcarbamoyl adenosine **2** at 70 °C in aqueous solution at different pH values within the range 6–9.5. Under all these conditions, the HPLC-MS analyses, after 24 h, indicated that **2** was stable in aqueous solution. Thus, heating at 70 °C the above obtained crude reaction mixtures for 24 h in 50 mM borate buffer pH 9.5 led to the hydrolysis of some of the *O*-methylcarbamoyl adenosine derivatives, whereas **2** persisted in the aqueous solution (Figure S3).

Prebiotic Loading of Amino Acids onto Nucleosides

Based on our previous study on the prebiotic synthesis of amino acid modified *N*⁶-carbamoyl adenosine **1** from *N*-methylcarbamoyl amino acids (Scheme 1a),^[41] we envisaged that the analogous nitrosation of the urea-modified adenosine **2**, followed by its conversion into the corresponding *N*⁶-



Scheme 1. Prebiotic synthesis of amino acid modified *N*⁶-carbamoyl adenosine nucleosides (**1**) from: a) *N*-methylcarbamoyl amino acids and b) *N*⁶-methylcarbamoyl adenosine (**2**). aa = amino acid.

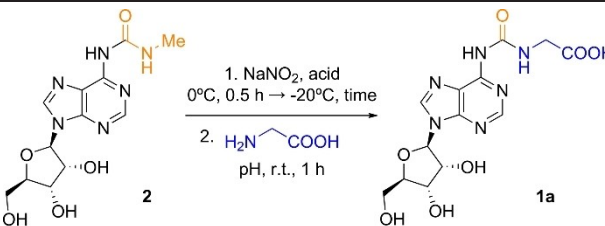
isocyanate could also lead to the formation of **1** (Scheme 1b). First, we optimized the conditions for the reaction of **1** with Gly using HPLC-MS analyses (Table 1 and Tables S1–S4). In the first reaction step, we performed the nitrosation of **2** with 12.5 equiv of NaNO₂ in an acidic aqueous solution. Among the acids tested, 5% H₃PO₄^[51] in water gave the best results, indicating that the formation of the *N*-nitroso derivative of **2** required rather harsh acidic conditions (pH < 1.5). Next, we investigated the reaction under freezing conditions, and we found that the yield strongly improved at –20 °C for 22 h. Most likely, under these conditions, the reactants were excluded from the growing ice crystals and concentrated in the interstitial (eutectic) phase.^[52–55] In the second reaction step, the thawed aqueous solution, containing initially **2** and NaNO₂, was adjusted to near neutral or basic pH. At r.t. and in the presence of 10 equiv of Gly, the loading of the amino acid onto **2** was very efficient at pH 9.5,^[56] affording **1a** (g⁶A) in a remarkable 84% yield. When the amino acid was added in the first step, the nitrosation reaction, the amino acid modified *N*⁶-carbamoyl adenosine was also obtained, yet in lower yield (Figure S6 and Table S5). Note that the prebiotic

formation of **1a** required dramatic changes in the pH of the aqueous solution. In this respect, it is worth mentioning that the second reaction step, coupling of the amino acid, also occurred under slightly acidic conditions in 5% yield (Table 1). We also attempted the loading of Gly using the *N*⁶-methylated version of **2**. However, under the optimized reaction conditions, the yield of the expected nucleoside product (m⁶g⁶A) was reduced to 8% (Figure S9).

To evaluate the scope of the prebiotic synthesis of amino acid modified *N*⁶-carbamoyl adenosines **1**, we carried out the reaction with a series of L-amino acids, having aliphatic (Ala, Pro), aromatic (Phe), neutral polar (Thr) and ionizable (Asp, Lys) side chains (Table 2).^[57] Under the reaction conditions optimized above for **1a** (g⁶A), we observed, in all cases, the formation of the nucleoside products **1b–g** in yields between 61–78% (Figures S4–S5). We also detected, in all cases, the presence of ca. 10% adenosine and the remaining starting material **2** (Figure 1a). Interestingly, the reaction of **2** with Lys, with its α - and ϵ -amino groups, provided a selectivity of 80:20 for the product α -**1f** (k⁶A) versus ϵ -**1f**.

Next, we studied the loading of an amino nitrile onto *N*⁶-methylcarbamoyl adenosine **2**. Amino nitriles were suggested as prebiotically plausible precursors for

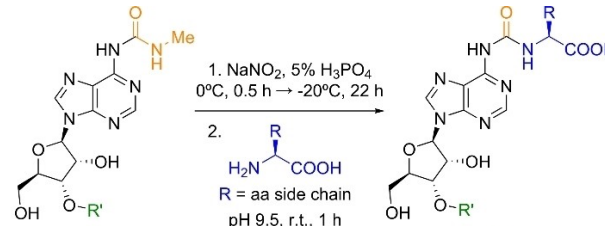
Table 1: Optimization of reaction conditions using *N*⁶-methylcarbamoyl adenosine (**2**) and Gly.



Reaction step 1		Reaction step 2	Yield [%] ^[a]
Acid	Freezing time [h]	pH	
neat acetic acid			n.d.
neat formic acid			18
1 M HCl	22	9.5	78
1 M H ₂ SO ₄			64
5% H ₃ PO ₄			84
	0		4
5% H ₃ PO ₄	3	9.5	42
	22		84
		6.2	5
		7.4	34
5% H ₃ PO ₄	22	8.6	60
		9.5	84

[a] Yields determined by HPLC analysis using the calibration curve of **1a** (Figure S10). n.d. = not detected.

Table 2: Amino acid scope using *N*⁶-methylcarbamoyl adenosine (**2**) and RNA oligonucleotide (**ON2**).



Amino acid	Nucleoside product	Yield [%] ^[a]	RNA product	Yield [%] ^[a,b]
Gly	1a (g ⁶ A)	84	ON1a	18
Thr	1b (t ⁶ A)	61	ON1b	18
Phe	1c (f ⁶ A)	78	ON1c	23
Pro	1d (p ⁶ A)	74	ON1d	15
Asp	1e (d ⁶ A)	64	ON1e	10
Lys	α - 1f (k ⁶ A)	70	α - ON1f	10
	ϵ - 1f	16	ϵ - ON1f	n.d.
Ala	1g (a ⁶ A)	75	ON1g	16
Gly _{cn} ^[c]	1h (g _{cn} ⁶ A)	26	ON1h	n.d.

[a] Yields determined by HPLC analysis using the calibration curves of reference compounds (Figures S10 and S19). [b] Reactions performed in the presence of GdmCl. [c] The structure of the Gly derivative contained a nitrile group instead of a carboxylic acid. A = adenosine. n.d. = not detected.

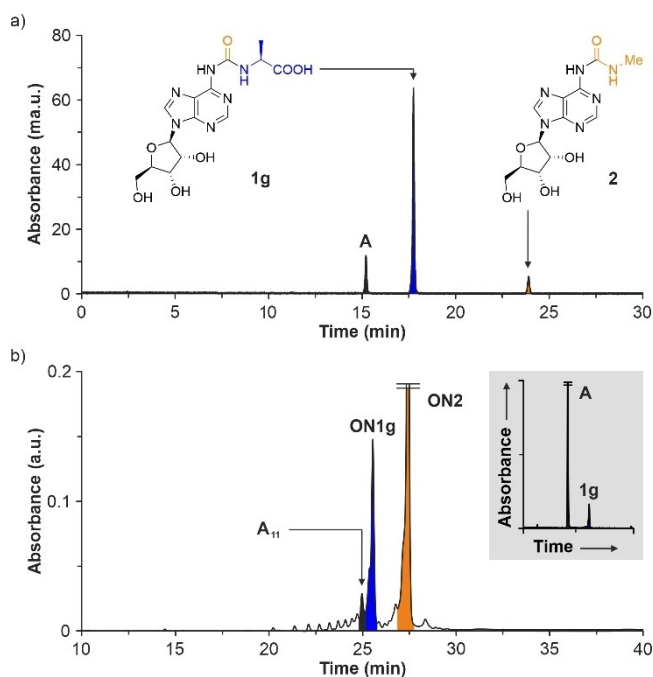


Figure 1. HPLC chromatograms of the crude reaction mixtures of Ala with: a) **2** and b) **ON2**. In b, inset shows enzymatic digestion of **ON1g**. A = adenosine and **A₁₁** = 5'-AAAAAAAAA-3'.

peptides.^[11,58–61] Indeed, the reaction of **2** with the nitrile derivative of Gly, that is, Gly_{CN}, generated the nucleoside product **1h** (g_{CN}⁶A) in 26 % yield (Table 2).

Encouraged by the reaction of **2** with amino acids, we explored the loading of Gly with other urea-modified nucleosides: *N*²-methylcarbamoyl guanosine **3** and *N*⁴-methylcarbamoyl cytidine **4** (Table 3). We found that the addition

Table 3: Base scope using different nucleosides and RNA oligonucleotides.

Base	Nucleoside	Product	Yield [%] ^[a]	RNA	Product	Yield [%] ^[a,b]
A	2	1a (g ⁶ A)	84	ON2'	ON1a'	18
G	3	5 (g ² G)	78	ON3	ON5	35
C	4	6 (g ⁴ C)	58	ON4	ON6	5

[a] Yields determined by HPLC analysis using the calibration curves of reference compounds (Figures S10, S11 and S19). [b] Reactions performed in the presence of GdmCl. A = adenosine; C = cytidine, G = guanosine, U = uridine and X = xanthosine.

of Gly to an aqueous solution of **3** at pH 9.5, after its nitrosation, afforded **5** (g²G) in 78 % yield (Figure 2a). Similarly, the combination of **4** with Gly gave **6** (g⁴C) in 58 % yield (Figure S8). These results suggested that the putative RNA-peptide synthesis might not be limited to RNAs containing the amino acid modified *N*⁶-carbamoyl adenosine **1**, although the non-canonical nucleosides **5** (g²G) and **6** (g⁴C) were so far not found in contemporary RNAs.^[62]

All together, these data showed that *N*-methylcarbamoyl nucleosides, **2–4**, could be efficiently loaded with amino acids and amino nitriles in aqueous solution under prebiotically plausible conditions to form amino acid modified *N*-carbamoyl nucleosides.

Prebiotic Loading of Amino Acids onto RNA

We investigated whether the loading of amino acids was possible when *N*⁶-methylcarbamoyl adenosine **2** was incorporated into RNA. For these experiments, and the subsequent RNA-based amino acid transfer, we synthesized the phosphoramidite derivative of **2** in four synthetic steps with a 35 % overall yield (Scheme S19). Then, we used solid-phase RNA synthesis to incorporate the urea-modified nucleoside at the 5'-end of the 11-mer homo-A RNA strand **ON2** (Table 2).

We started with the homo-A RNA strand **ON2** and analyzed the loading of a series of amino acids. For example, when we treated **ON2** with 500 equiv of NaNO₂ in 5 % H₃PO₄ at –20 °C and then, we added, at r.t., an excess of Ala, followed by basification to pH 9.5, the HPLC chroma-

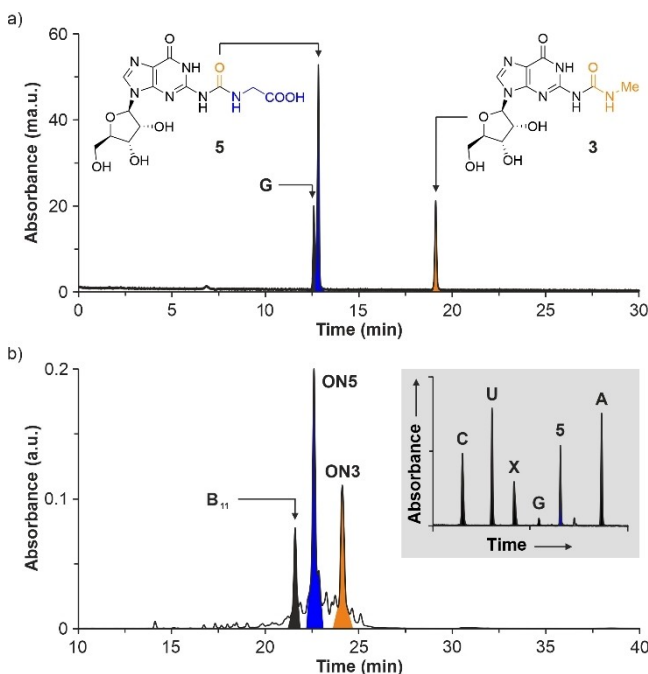


Figure 2. HPLC chromatograms of the crude reaction mixtures of Gly with: a) **3** and b) **ON3**. In b, inset shows enzymatic digestion of **ON5**. A = adenosine; C = cytidine; G = guanosine; U = uridine; X = xanthosine and **B₁₁** = 5'-GAUCXCUXUAC-3'.

togram of the crude reaction mixture, as well as the matrix-assisted laser desorption/ionization–time-of-flight (MALDI-TOF) mass spectrum of the purified compound revealed the formation of the RNA strand **ON1g** with a terminal ⁶A nucleotide (Figure S14). Next to the remaining starting material, we noted the formation of three major side products. Using MALDI-TOF MS, we assigned the degradation products to 10-mer and 11-mer RNA strands, **A₁₀** and **A₁₁**, that had lost the terminal urea-modified nucleoside and the *N*⁶-methylcarbamoyl substituent, respectively. We also detected a 10-mer RNA strand with a terminal ⁶A nucleotide, **A₉-a⁶A**. It seems that the nitrosation reaction provoked the partial degradation of the RNA strands. However, the addition of 100 mM of guanidinium chloride (GdmCl) salt reduced the side reactions dramatically (Figure 1b), potentially by formation of salt-bridges between the (nitrosated)^[63,64] guanidinium cation and the phosphodiester backbone.^[65] Other salts, such as NaCl and NaClO₄, had little effect on the RNA degradation (Figure S14).

In order to further prove the loading of Ala onto the RNA strand **ON2**, we performed an enzymatic digestion of the isolated RNA product **ON1g** into the nucleosides, and analyzed the digest with the help of the corresponding nucleoside reference compounds by HPLC-MS.^[66] We treated **ON1g** with a nucleoside digestion mixture at 37 °C for 2 h. Analysis of the obtained nucleoside mixture showed two peaks corresponding to **1g** (a⁶A) and adenosine in a ca. 1:10 ratio (Figure 1b), which was in agreement with the nucleotide composition of the digested RNA strand.

In the presence of GdmCl and for all the amino acids studied, we observed the formation of the RNA strands **ON1a–g** in 10–23 % yield (Table 2 and Figures S15–S16). It is worth mentioning that the reaction of Lys with the RNA strand **ON2** yielded predominantly the α -regioisomer, **α -ON1f**, of the k⁶A nucleotide. For the loading with the amino nitrile Gly_{CN}, we did not detect the expected RNA product **ON1h**.

In analogy to the loading of amino acids onto different nucleosides, we evaluated the loading of Gly with three 11-mer RNA strands, **ON2'–ON4** (Table 3), having all four canonical nucleotides, but bearing distinct *N*-methylcarbamoyl nucleotides at the 5'-end. The RNA strands were synthesized from the corresponding phosphoramidite derivatives of **2–4** (Schemes S19–S21). In all cases, the reactions of the three RNA strands with Gly, under the conditions described above with GdmCl, gave the respective products, **ON1a'**, **ON5–ON6**, in 5–35 % yield (Table 3, Figure 2b and Figure S12). MALDI-TOF MS analyses of the obtained RNA products showed the successful loading of Gly onto the parent RNA strands. In addition, we detected, after the enzymatic digestion of the RNA products, the partial conversion of guanosine to xanthosine by nitrosative deamination.^[67,68]

Moreover, we investigated the loading of Phe onto an RNA strand **ON2''** bearing two *N*⁶-methylcarbamoyl adenosine nucleotides, placed at the 5'-end and an internal position of the sequence (Scheme S8). Using similar reaction conditions to those described above, we observed the

formation of the double-loaded RNA strand **ON1c'** as the major species in the crude reaction mixture (Figure S13).

Collectively, these data showed that RNA strands containing *N*-methylcarbamoyl nucleotides at either terminal or internal positions, or both, could be loaded with amino acids in a sufficient extent to perform RNA-based amino acid transfer (see below).

RNA-Based Amino Acid Transfer Following the Loading Reaction

Finally, we wanted to demonstrate that the prebiotic loading of an amino acid onto an RNA strand, functionalized with a urea-modified adenosine nucleotide at the 5'-end, could undergo RNA-based amino acid transfer (Figure 3a). The loading reaction provides a terminal amino acid modified *N*⁶-carbamoyl adenosine (donor) as the basis for the synthesis reaction. For the amino acid transfer, we hybridize the obtained donor RNA strand with a complementary acceptor RNA strand, containing a 5-aminomethyl uridine at the 3'-end, forming a duplex. Activation of the carboxylic acid in the donor RNA strand leads to the formation of a hairpin intermediate, which is opened at elevated temperatures by urea cleavage, transferring the loaded amino acid to the acceptor RNA strand. The obtained acceptor RNA strand can be involved in subsequent reactions, enabling peptide growth on RNA.^[11] Since the nitrosation reaction converts guanosine to xanthosine, we employed a binary code of adenosine and uridine nucleotides in the initial RNA strand **ON7**. We also selected Phe for the experiment because of its high reaction efficiency (Table 2).^[11]

In this experiment, the individual reaction steps were monitored by HPLC (Figure 3b). In the first step, the prebiotic loading of Phe onto the RNA strand **ON7**, under the reaction conditions described above with GdmCl, gave the donor RNA strand **ON8** in ca. 38 % yield. In the second step, we combined the isolated donor RNA strand **ON8** with an equimolar amount of the acceptor RNA strand **ON9** in 100 mM 2-(*N*-morpholino)ethanesulfonate (MES) buffer pH 6 and 1 M NaCl. The RNA strand **ON9** contained 2'-methoxy nucleotides, which are found in contemporary ribosomal RNA (rRNA)^[69] and minimize degradation. The addition of 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM-Cl) to the solution mixture of **ON8** and **ON9** at 15 °C led to the formation of the hairpin intermediate **ON10** in ca. 40 % yield. Since the melting temperature of the reactive duplex **ON8-ON9** was ca. 35 °C in the buffered aqueous solution (Figure S20a–b), it was assembled almost quantitatively in the reaction mixture. As reported previously, the activation of the carboxylic acid could also be performed prebiotically with methylisonitrile (MeNC) and 4,5-dicyanoimidazole (DCI).^[11,70] In the third step, we warmed the aqueous solution of the hairpin intermediate **ON10** to 90 °C at pH 6 to cleave the urea bond,^[71,72] which furnished the RNA-peptide chimera **ON11** in ca. 48 % yield. In contrast to our previous experiments with donor RNA strands containing amino acid modified *N*⁶-methyl-*N*⁶-carbamoyl adenosine at

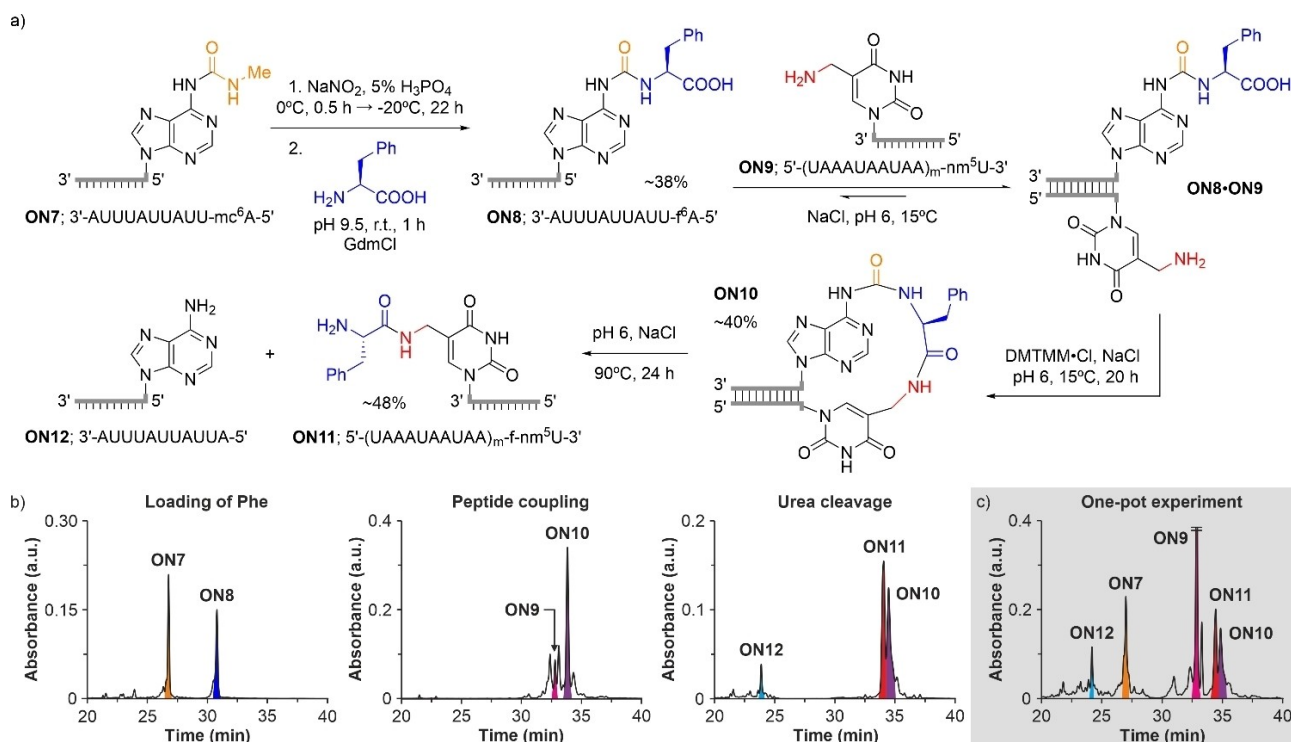


Figure 3. a) Reaction scheme for RNA-based amino acid transfer: loading of Phe, peptide coupling and urea cleavage. HPLC chromatograms of the crude reaction mixtures of: b) the individual steps and c) the one-pot experiment. mc⁶A = N⁶-methylcarbamoyl adenosine; nm⁵U = 5-aminomethyl uridine; A = adenosine and U = uridine. m (subscript) indicates that the RNA strands contained 2'-methoxy nucleotides.

the 5'-end,^[11] the absence of the N⁶-methyl substituent in **ON8** prevented the formation of a hydantoin side-product in the urea cleavage step.^[73]

Finally, we performed the loading of Phe and the RNA-based amino acid transfer in a one-pot experiment (Figure 3c). The crude reaction mixtures were filtered after each step to remove the excess of salts and low-molecular weight compounds but retaining the RNA strands. Indeed, analysis of the one-pot experiment by HPLC showed the formation of the RNA-peptide chimera **ON11** in ca. 11% overall yield (Figure 3c), which was in line with the result obtained for the three individual reaction steps (ca. 7%).

Conclusion

Recently, we reported that RNA with the help of two non-canonical nucleosides, that is, amino acid modified N⁶-carbamoyl adenosine and 5-methylaminomethyl uridine, gained the ability to carry amino acids at certain nucleobase positions. The subsequent transfer of amino acids from one nucleobase to another enabled the synthesis of small peptides on RNA, affording RNA-peptide chimeras. The ability of RNA to decorate itself with the help of non-canonical nucleosides constitutes an additional function, which places RNA in a prime position for enabling the chemical evolution of life.

In this work, we described a new method to load RNA strands with amino acids. While this loading reaction was so

far only possible with adenosine in low yields, we could now improve the loading efficiencies to yields up to 84%. Most importantly, the loading reaction was also possible in RNA strands with acceptable efficiencies between 10–23%. These results suggested that RNA-peptide chimeras are molecules that could have formed under early Earth conditions. In addition, we expanded the loading reaction to N-methylcarbamoyl guanosine and cytidine, and to their corresponding RNA strands. A caveat of the reported process is that we detected deamination of guanosine under the loading conditions. In contrast, the loading with Phe and the subsequent RNA-based amino acid transfer was found to be surprisingly efficient. This observation suggests that loading preferences, for yet unknown reasons, might exist. Thus, we believe that different amino acids might have distinct preferences to be loaded onto specific sequences. This is the prerequisite for RNA-encoded peptide synthesis. Our study also showed that the absence of a N⁶-methyl substituent in N⁶-methylcarbamoyl adenosine enhanced the loading efficiency and prevented the formation of hydantoin side-products.

The reported chemistry also required big changes in the pH of the aqueous solution. While the nitrosation of the N⁶-methylcarbamoyl nucleoside took place under acidic conditions (5% H₃PO₄), the reaction with the amino acid was performed at the optimal pH 9.5. Therefore, we need to assume that the first reaction step occurred in an acidic environment, such as an acidic pond, from which a mixture of the nitrosated nucleoside and the amino acids flow out,

e.g. along a carbonate containing river bed or into a carbonate dominated second basin. Repetitive loading and RNA-based peptide synthesis cycles would require continuous fluctuations of the pH between acidic and moderately basic in a dynamic process. However, it is worth mentioning that a low amino acid loading efficiency is not necessarily a disadvantage for an RNA-peptide world, since a high turnover of canonical nucleotides into *N*-carbamoyl derivatives would lead to the loss of the information-encoding property of RNA, hindering replication. We are just at the beginning of learning about the possibilities that are offered by RNA-peptide conjugates and how they could be integrated into coding schemes, e.g. via Hoogsteen-type base-pairing.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for supporting this research through the DFG grants: CA275/11-3 (ID: 326039064), CRC1309 (ID: 325871075, A4), CRC1032 (ID: 201269156, A5) and CRC1361 (ID: 393547839, P2). We thank the Volkswagen Foundation for funding this research (grant EvoRib). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No. 741912 (EPiR). L. E. thanks the Alexander von Humboldt Foundation for a postdoctoral fellowship (ESP 1214218 HFST-P). We also thank M.Sc. Johann de Graaff for helpful discussions. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the Supporting Information of this article.

Keywords: Amino Acids · Nucleosides · Origin of Life · Prebiotic Chemistry · RNA-Peptide World

- [1] V. Ramakrishnan, *Cell* **2002**, *108*, 557–572.
- [2] T. A. Steitz, *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 242–253.
- [3] M. B. Hoagland, E. B. Keller, P. C. Zamecnik, *J. Biol. Chem.* **1956**, *218*, 345–358.
- [4] M. Delarue, *Curr. Opin. Struct. Biol.* **1995**, *5*, 48–55.
- [5] M. Ibba, D. Söll, *Annu. Rev. Biochem.* **2000**, *69*, 617–650.
- [6] R. B. Ganesh, S. J. Maerkl, *Front. Bioeng. Biotechnol.* **2022**, *10*, 918659.
- [7] M. Di Giulio, *J. Mol. Evol.* **1997**, *45*, 571–578.
- [8] K. Tamura, *J. Biosci.* **2011**, *36*, 921–928.
- [9] A. Harish, G. Caetano-Anollés, *PLoS One* **2012**, *7*, e32776.
- [10] J. C. Bowman, A. S. Petrov, M. Frenkel-Pinter, P. I. Penev, L. D. Williams, *Chem. Rev.* **2020**, *120*, 4848–4878.

- [11] F. Müller, L. Escobar, F. Xu, E. Węgrzyn, M. Nainytė, T. Amatov, C. Y. Chan, A. Pichler, T. Carell, *Nature* **2022**, *605*, 279–284.
- [12] T. Carell, C. Brandmayr, A. Hienzsch, M. Muller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs, M. Wagner, *Angew. Chem. Int. Ed.* **2012**, *51*, 7110–7131.
- [13] M. P. Schweizer, K. McGrath, L. Baczynskij, *Biochem. Biophys. Res. Commun.* **1970**, *40*, 1046–1052.
- [14] L. Perrochia, E. Crozat, A. Hecker, W. Zhang, J. Bareille, B. Collinet, H. van Tilbeurgh, P. Forterre, T. Basta, *Nucleic Acids Res.* **2013**, *41*, 1953–1964.
- [15] H. Grosjean, E. Westhof, *Nucleic Acids Res.* **2016**, *44*, 8020–8040.
- [16] J. Beenstock, F. Sicheri, *Nucleic Acids Res.* **2021**, *49*, 10818–10834.
- [17] W. Gilbert, *Nature* **1986**, *319*, 618.
- [18] G. F. Joyce, *Nature* **2002**, *418*, 214–221.
- [19] K. R. Lynn, *J. Phys. Chem.* **1965**, *69*, 687–689.
- [20] J. Akester, J. Cui, G. Fraenkel, *J. Org. Chem.* **1997**, *62*, 431–434.
- [21] F. V. Murphy, V. Ramakrishnan, A. Malkiewicz, P. F. Agris, *Nat. Struct. Mol. Biol.* **2004**, *11*, 1186–1191.
- [22] A. Kitamura, M. Nishimoto, T. Sengoku, R. Shibata, G. Jäger, G. R. Björk, H. Grosjean, S. Yokoyama, Y. Bessho, *J. Biol. Chem.* **2012**, *287*, 43950–43960.
- [23] M. Nainytė, F. Müller, G. Ganazzoli, C. Y. Chan, A. Crisp, D. Globisch, T. Carell, *Chem. Eur. J.* **2020**, *26*, 14856–14860.
- [24] P. Berg, *J. Biol. Chem.* **1958**, *233*, 608–611.
- [25] R. Lohrmann, L. E. Orgel, *Nature* **1973**, *244*, 418–420.
- [26] J. L. Shim, R. Lohrmann, L. E. Orgel, *J. Am. Chem. Soc.* **1974**, *96*, 5283–5284.
- [27] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14564–14569.
- [28] H. Griesser, M. Bechthold, P. Tremmel, E. Kervio, C. Richert, *Angew. Chem. Int. Ed.* **2017**, *56*, 1224–1228.
- [29] H. Griesser, P. Tremmel, E. Kervio, C. Pfeffer, U. E. Steiner, C. Richert, *Angew. Chem. Int. Ed.* **2017**, *56*, 1219–1223.
- [30] M. Rächle, G. Leveau, C. Richert, *Eur. J. Org. Chem.* **2020**, 6966–6975.
- [31] B. Jash, P. Tremmel, D. Jovanovic, C. Richert, *Nat. Chem.* **2021**, *13*, 751–757.
- [32] N. S. M. D. Wickramasinghe, M. P. Staves, J. C. Lacey, Jr., *Biochemistry* **1991**, *30*, 2768–2772.
- [33] M. Illangasekare, G. Sanchez, T. Nickles, M. Yarus, *Science* **1995**, *267*, 643–647.
- [34] K. Tamura, P. Schimmel, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8666–8669.
- [35] K. Tamura, P. Schimmel, *Science* **2004**, *305*, 1253.
- [36] K. Tamura, P. R. Schimmel, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13750–13752.
- [37] N. V. Chumachenko, Y. Novikov, M. Yarus, *J. Am. Chem. Soc.* **2009**, *131*, 5257–5263.
- [38] R. M. Turk, N. V. Chumachenko, M. Yarus, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4585–4589.
- [39] R. M. Turk, M. Illangasekare, M. Yarus, *J. Am. Chem. Soc.* **2011**, *133*, 6044–6050.
- [40] L. F. Wu, M. Su, Z. W. Liu, S. J. Bjork, J. D. Sutherland, *J. Am. Chem. Soc.* **2021**, *143*, 11836–11842.
- [41] C. Schneider, S. Becker, H. Okamura, A. Crisp, T. Amatov, M. Stadlmeier, T. Carell, *Angew. Chem. Int. Ed.* **2018**, *57*, 5943–5946.
- [42] N. F. W. Ligterink, A. Coutens, V. Kofman, H. S. P. Müller, R. T. Garrod, H. Calcutt, S. F. Wampfler, J. K. Jørgensen, H. Linnartz, E. F. van Dishoeck, *Mon. Not. R. Astron. Soc.* **2017**, *469*, 2219–2229.
- [43] F. Goesmann, H. Rosenbauer, J. H. Bredehöft, M. Cabane, P. Ehrenfreund, T. Gautier, C. Giri, H. Krüger, L. Le Roy, A. J.

- MacDermott, S. McKenna-Lawlor, U. J. Meierhenrich, G. M. M. Caro, F. Raulin, R. Roll, A. Steele, H. Steininger, R. Sternberg, C. Szopa, W. Thiemann, S. Ulamec, *Science* **2015**, *349*, aab0689.
- [44] R. Martín-Doménech, V. M. Rivilla, I. Jiménez-Serra, D. Quénard, L. Testi, J. Martín-Pintado, *Mon. Not. R. Astron. Soc.* **2017**, *469*, 2230–2234.
- [45] M. N. Heinrich, W. R. Thompson, C. Sagan, *Bull. Am. Astron. Soc.* **1991**, *23*, 1211.
- [46] Y. L. Yung, M. B. McElroy, *Science* **1979**, *203*, 1002–1004.
- [47] R. L. Mancinelli, C. P. McKay, *Origins Life Evol. Biospheres* **1988**, *18*, 311–325.
- [48] S. Ranjan, Z. R. Todd, P. B. Rimmer, D. D. Sasselov, A. R. Babbín, *Geochim. Geophys. Geosyst.* **2019**, *20*, 2021–2039.
- [49] G. W. Breton, M. Turlington, *Tetrahedron Lett.* **2014**, *55*, 4661–4663.
- [50] R. Saladino, C. Crestini, S. Pino, G. Costanzo, E. Di Mauro, *Phys. Life Rev.* **2012**, *9*, 84–104.
- [51] D. J. Ritson, S. J. Mojzsis, J. D. Sutherland, *Nat. Geosci.* **2020**, *13*, 344–348.
- [52] T. Vajda, *Cell. Mol. Life Sci.* **1999**, *56*, 398–414.
- [53] C. Menor-Salván, M. R. Marín-Yaseli, *Chem. Soc. Rev.* **2012**, *41*, 5404–5415.
- [54] K. Kitada, Y. Suda, N. Takenaka, *J. Phys. Chem. A* **2017**, *121*, 5383–5388.
- [55] N. Kitadai, S. Maruyama, *Geosci. Front.* **2018**, *9*, 1117–1153.
- [56] B. T. Golding, C. Bleasdale, J. McGinnis, S. Müller, H. T. Rees, N. H. Rees, P. B. Farmer, W. P. Watson, *Tetrahedron* **1997**, *53*, 4063–4082.
- [57] M. Kimura, S. Akanuma, *J. Mol. Evol.* **2020**, *88*, 372–381.
- [58] S. Islam, M. W. Powner, *Chem* **2017**, *2*, 470–501.
- [59] P. Canavelli, S. Islam, M. W. Powner, *Nature* **2019**, *571*, 546–549.
- [60] C. S. Foden, S. Islam, C. Fernandez-Garcia, L. Maugeri, T. D. Sheppard, M. W. Powner, *Science* **2020**, *370*, 865–869.
- [61] J. Singh, D. Whitaker, B. Thoma, S. Islam, C. S. Foden, A. E. Aliev, T. D. Sheppard, M. W. Powner, *J. Am. Chem. Soc.* **2022**, *144*, 10151–10155.
- [62] S. P. Dutta, C. I. Hong, G. P. Murphy, A. Mittelman, G. B. Chheda, *Biochemistry* **1975**, *14*, 3144–3151.
- [63] A. F. McKay, *Chem. Rev.* **1952**, *51*, 301–346.
- [64] I. Fernández, P. Hervés, M. Parajó, *J. Phys. Org. Chem.* **2008**, *21*, 713–717.
- [65] J. Vušurović, E.-M. Schneeberger, K. Breuker, *ChemistryOpen* **2017**, *6*, 739–750.
- [66] F. Xu, A. Crisp, T. Schinkel, R. C. A. Dubini, S. Hübner, S. Becker, F. Schelter, P. Rovó, T. Carell, *Angew. Chem. Int. Ed.* **2022**, *61*, e202211945.
- [67] J. Xu, N. J. Green, D. A. Russell, Z. Liu, J. D. Sutherland, *J. Am. Chem. Soc.* **2021**, *143*, 14482–14486.
- [68] S. Mair, K. Erharter, E. Renard, K. Brillet, M. Brunner, A. Lusser, C. Kreutz, E. Ennifar, R. Micura, *Nucleic Acids Res.* **2022**, *50*, 6038–6051.
- [69] W. A. Decatur, M. J. Fournier, *Trends Biochem. Sci.* **2002**, *27*, 344–351.
- [70] Z. W. Liu, L. F. Wu, J. F. Xu, C. Bonfio, D. A. Russell, J. D. Sutherland, *Nat. Chem.* **2020**, *12*, 1023–1028.
- [71] M. Hutchby, C. E. Houlden, J. G. Ford, S. N. G. Tyler, M. R. Gagné, G. C. Lloyd-Jones, K. I. Booker-Milburn, *Angew. Chem. Int. Ed.* **2009**, *48*, 8721–8724.
- [72] A. Ohkubo, R. Kasuya, K. Miyata, H. Tsunoda, K. Seio, M. Sekine, *Org. Biomol. Chem.* **2009**, *7*, 687–694.
- [73] G. Danger, R. Plasson, R. Pascal, *Chem. Soc. Rev.* **2012**, *41*, 5416–5429.

Manuscript received: February 15, 2023

Accepted manuscript online: March 7, 2023

Version of record online: April 17, 2023

3.3 RNA-Templated Peptide Bond Formation Promotes L-Homochirality

Authors

Ewa Węgrzyn^{*}, Ivana Mejdrová^{*}, Felix M. Müller, Milda Nainytė, Luis Escobar[°] und Thomas Carell[°]

Angew. Chem. Int. Ed., **2024**, e202319235

DOI: <https://doi.org/10.1002/anie.202319235>

Supplementary Information included in the Appendix III

^{*} These authors contributed equally [°] corresponding authors

Summary

The origin of biological homochirality, relevant for proper function of nucleic acids and proteins, is one of the biggest unanswered questions of the prebiotic chemistry field. We observe the exclusive abundance of D-(deoxy)ribose and L-amino acid in all the functional biomolecules of living organisms. We recently reported an RNA-peptide synthesis cycle, which makes use of amino acid-decorated nucleotides incorporated into RNA and which, with help of RNA-templation, enables amino acid or peptide transfer between complementary RNA strands. In this work, we explore the effect of exchanging the L-amino acid for D-amino acid and its influence on the stereochemical preference for homo- vs. heterochiral peptide formation. We perform competitive coupling reactions, in which an equimolar mixture of L- and D-amino acid modified RNA competes for reaction with the amino acid attached to the counterpart strand. We examine the thermodynamics and kinetics of this reaction with a selected number of amino acids and different activation methods.

Personal contribution

Synthesis of modified phosphoramidites and oligonucleotides, characterization of coupling products and performance of competition reactions, measurement of the calibration curves for quantification, preparation of the manuscript and supporting information, planning of the project and development of concepts.

Homochirality

RNA-Templated Peptide Bond Formation Promotes L-Homochirality

Ewa Węgrzyn⁺, Ivana Mejdrová⁺, Felix M. Müller, Milda Nainytė, Luis Escobar,* and Thomas Carell*

Abstract: The world in which we live is homochiral. The ribose units that form the backbone of DNA and RNA are all D-configured and the encoded amino acids that comprise the proteins of all living species feature an all-L-configuration at the α -carbon atoms. The homochirality of α -amino acids is essential for folding of the peptides into well-defined and functional 3D structures and the homochirality of D-ribose is crucial for helix formation and base-pairing. The question of why nature uses only encoded L- α -amino acids is not understood. Herein, we show that an RNA-peptide world, in which peptides grow on RNAs constructed from D-ribose, leads to the self-selection of homo-L-peptides, which provides a possible explanation for the homo-D-ribose and homo-L-amino acid combination seen in nature.

Introduction

Folding of peptides into well-defined and functional protein structures,^[1,2,3] as well as the duplex formation of RNA and DNA,^[4,5,6] in which information can be stored due to Watson-Crick base-pairing,^[7] requires homochirality.^[8,9] The living nature that surrounds us evolved based on homo-D-configured ribose, which forms the backbone of RNA and DNA, and homo-L-configured α -amino acids, which are indispensable in the world of proteins. While homochirality is the prerequisite for folding of these biopolymers into 3D structures, the question of why D-sugars are combined with L-amino acids and not with the D-counterparts and hence the diastereoselectivity of nature is unknown.^[10,11,12]

The current idea of how life may have started is based on the RNA world hypothesis, in which RNAs self-replicated and folded into catalytically active structures.^[13,14,15] At some point in evolution, these RNA structures gained the potential to connect amino acids to form peptides and proteins with increasing catalytic capabilities.^[16] Even though there is more than one model to explain how RNA learned, at some point, to make peptides and proteins, the amino acids were initially likely connected to each other in close proximity to the RNA structures, with RNAs acting as peptide forming catalysts.

We recently reported the idea that, instead of a pure RNA world, chimeric RNA-amino acid/peptide structures may have created an RNA-peptide world, in which RNA and peptides were covalently connected until the structures got large enough to replace covalent bonding by non-covalent interactions.^[17,18,19] We observed that RNA strands equipped with ubiquitous non-canonical nucleosides, such as 5-methylaminomethyl uridine (mnm⁵U) and threonine-modified N⁶-carbamoyl adenosine (Thr⁶A), which are found in contemporary transfer RNAs (tRNAs)^[20,21,22] and which can be considered to be molecular fossils,^[23,24] allow RNA to self-decorate with peptides.^[18]

The synthesis cycle that allows peptides to form directly on RNA is depicted in Figure 1a, but it should be mentioned again that this is just one example of how peptides can be formed on and by RNA. The key features of our model involve: 1. the reaction of two amino acids, attached via the non-canonical nucleosides m⁶aa⁶A and mnm⁵U to complementary RNA strands, to give a stable hairpin structure. 2. The subsequent cleavage of the urea bond by simple heating, which breaks the hairpin. This is followed by a potential exchange (3. release/4. annealing) of the formed m⁶A-containing donor RNA strand for another amino acid-containing oligonucleotide. Repeating this cycle allows the growing of longer peptides on the acceptor RNA strand.

Although it is impossible to prove that these particular reactions were indeed involved in creating an RNA-peptide world on the early Earth, it is a prebiotically plausible concept of how RNA could have initially templated the synthesis of peptides. The models of how RNA encoded peptide synthesis have in common that the peptide forming reactions take place on RNA. Consequently, our non-canonical nucleoside-based model allows us to investigate the stereochemical preferences of RNA-templated peptide growth.

Previous studies by the research groups of Lacey,^[25,26] Tamura and Schimmel,^[27,28,29,30,31] Sutherland^[32,33,34] and

[*] E. Węgrzyn,⁺ Dr. I. Mejdrová,⁺ Dr. F. M. Müller, Dr. M. Nainytė, Dr. L. Escobar, Prof. Dr. T. Carell
 Department of Chemistry, Institute for Chemical Epigenetics (ICE-M)
 Ludwig-Maximilians-Universität (LMU) München
 Butenandtstrasse 5-13, 81377 Munich (Germany)
 E-mail: luisescobar1992@hotmail.es
 thomas.carell@lmu.de

[†] These authors contributed equally.

© 2024 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

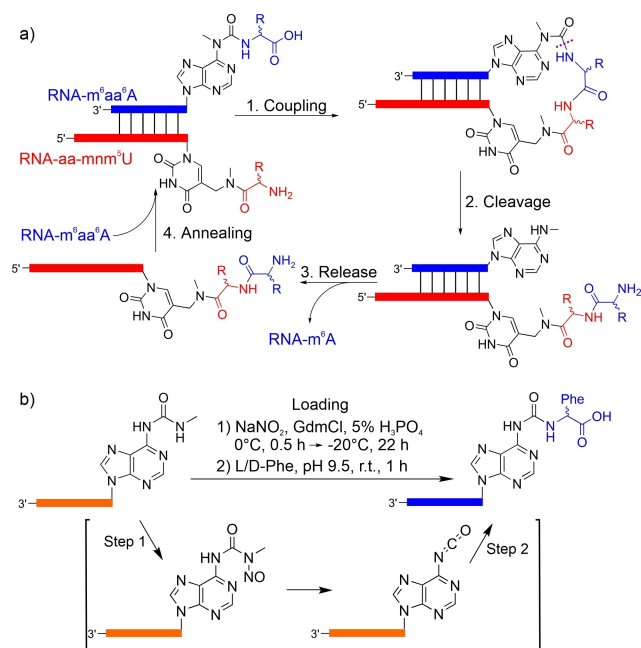


Figure 1. a) RNA-peptide synthesis cycle involving donor and acceptor RNA strands with m^6aa^6A and $aa-mmm^5U$, respectively, and b) prebiotic loading reaction of a racemic mixture of L/D-Phe onto an RNA strand containing a terminal N^6 -methylcarbamoyl adenosine nucleotide. GdmCl = guanidinium chloride.

Richert^[35] showed already that α -amino acids connected to 5'-phosphorylated nucleotides and RNAs, either via the carboxylic acid group as acyl phosphate mixed anhydrides,^[36] or with the α -amino group in the form of phosphoramidates,^[37] react preferentially when they are L-configured. Herein, we investigated how our RNA-peptide synthesis cycle (Figure 1a) is influenced by stereochemistry.

Results and Discussion

The first step of the RNA-templated peptide growth cycle involves the loading of an amino acid onto the donor RNA strand to give a non-canonical amino acid-modified adenosine nucleotide, (m^6) aa^6A . For example, Gly⁶A and (m^6)Thr⁶A are known to exist in the anticodon loop of contemporary tRNAs.^[38,39,40,41] Previously, we showed that this loading step could be achieved in a prebiotically plausible way by the reaction of an N^6 -methylurea adenosine nucleotide with NO^+ in the presence of amino acids or even peptides via an N^6 -isocyanate intermediate (Figure 1b).^[19,42] Therefore, we initially investigated whether the loading reaction displayed any stereochemical preference. We performed this reaction onto an RNA strand containing an N^6 -methylcarbamoyl adenosine nucleotide at the 5'-end with a racemic mixture of L/D-Phe. The aromatic amino acid was chosen because it facilitates the analysis of this crude reaction mixture by high-performance liquid chromatography (HPLC) using UV/Vis detection. In this loading reaction, we did not observe any stereochemical preference (Figure S12). The L- and D-Phe were loaded onto the RNA

strand with the same preference, which gives a 1:1 isomeric mixture of the terminal L- and D-Phe⁶A nucleotides in 35% overall yield.

We next investigated how the peptide growth reaction (step 1, Figure 1a) is influenced by either L- or D-amino acids. For these experiments, we synthesized a series of donor and acceptor RNA strands (**ON1** and **ON2**, respectively) using an automated solid-phase RNA synthesizer and 2'-OMe nucleosides (Figure 2 and SI). We used the 2'-OMe nucleosides, which are very prevalent in contemporary ribosomal RNA (rRNA), to increase the stability of the phosphodiester backbone toward hydrolysis and to achieve higher duplex melting temperatures with short RNA strands.^[43,44,45] In particular, we prepared the 7-mer donor RNA strands **ON1a-e** with either L- and D-amino acid-modified N^6 -methylcarbamoyl adenosine nucleotides ($m^6aa^6A_m$) at the 5'-end. We also synthesized the complementary 11-mer acceptor RNA strands **ON2a** with either L- and D-amino acid-modified 5-methylaminomethyl uridine nucleotides ($aa-mmm^5U_m$) at the 3'-end. We focused at Val, Ala and Thr, which are thought to have been early amino acids. They are found in meteorites and they are formed under prebiotically plausible conditions.^[46,47,48,49,50,51,52,53] All RNA strands were purified by HPLC and characterized by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) (Tables S1-S2).

Next, peptide formation reactions leading to the corresponding hairpin intermediates were performed (Figure 2). In order to enable accurate analysis of the stereochemical outcome of the coupling reactions by HPLC, we first synthesized independently the hairpins **ON3a-e** as standards from the complementary donor and acceptor RNA strands **ON1a-e** and **ON2a**. After annealing of **ON1a-e** with **ON2a**, we added in each case 50 mM of an activator,^[54] either 1-

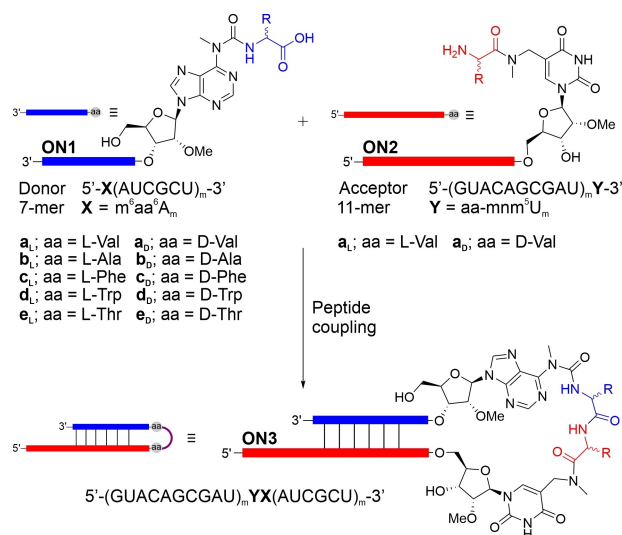


Figure 2. Peptide coupling reactions between donor and acceptor RNA strands, **ON1** and **ON2**, respectively, to yield hairpin products, **ON3**. The RNA strands used in this work contained L- and D-amino acids. aa = amino acid and R = aa side chain. m (subscript) indicates that the RNA strands were composed of 2'-OMe nucleotides.

ethyl-3-(3-dimethylaminopropyl) carbodiimide with *N*-hydroxysulfosuccinimide (EDC/Sulfo-NHS) or 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM•Cl), in 100 mM 2-(*N*-morpholino)ethanesulfonate (MES) buffer at pH 6 containing 100 mM of NaCl. After 2–6 h at r.t., the obtained hairpin reference compounds **ON3a-e** were isolated by HPLC and characterized by MALDI-TOF MS (Table S3).

With these standards in hand, we next investigated the diastereoselectivity of the RNA-based peptide synthesis (Figure 2). To this end, we studied the coupling of the donor RNA strands **ON1**, connected to either an L- (**ON1_L**) or a D-amino acid (**ON1_D**), with the complementary acceptor RNA strands **ON2**, containing either an L-Val (**ON2_{a_L}**) or a D-Val (**ON2_{a_D}**), in direct competition experiments (Figure 3a–c). First, we started with Val in the donor RNA strand **ON1a**. For the experiment, we prepared an equimolar solution of **ON1a_L** (L-Val) and **ON1a_D** (D-Val) in water. HPLC analysis of the mixture confirmed the 1:1 relationship (Figure 3a). Next, we performed the peptide coupling reaction after the addition of 1 equiv. of the acceptor RNA strand **ON2a_L** (L-Val) using the activator EDC/Sulfo-NHS

(MES buffer pH 6, NaCl, r.t., 2 h). HPLC analysis of the crude reaction mixture revealed the formation of the two possible hairpin products, **ON3a_{LL}** (L-Val,L-Val) and **ON3a_{DL}** (D-Val,L-Val), in 69% overall yield (Figure 3b). To our surprise, we immediately noted a significant selectivity for the L,L-product. **ON3a_{LL}** formed with a remarkable diastereoselectivity of 93:7 (L,L:D,L). For comparison, we carried out the same competition experiment (**ON1a_L** and **ON1a_D**) in the presence of the acceptor RNA strand **ON2a_D** (D-Val). In this case, the HPLC analysis showed a lower yield for **ON3a_{LD}** (L-Val,D-Val) plus **ON3a_{DD}** (D-Val,D-Val) of only 36% and a lower diastereoselectivity of only 66:34 (L,D:D,D) (Figure 3c). Similar results were obtained when we repeated the competition experiments with **ON1a** and **ON2a** at lower temperature (5 °C, Table S9). These data showed that the L,L-hairpin products are the preferred reaction outcomes. If we assume that life started with the formation of peptides close to RNA, we can surmise that the D-ribose in RNA promotes L-homochirality of the peptide formed in close proximity.

To further strengthen this argument, we repeated the experiment with other amino acids using the L/D-donor

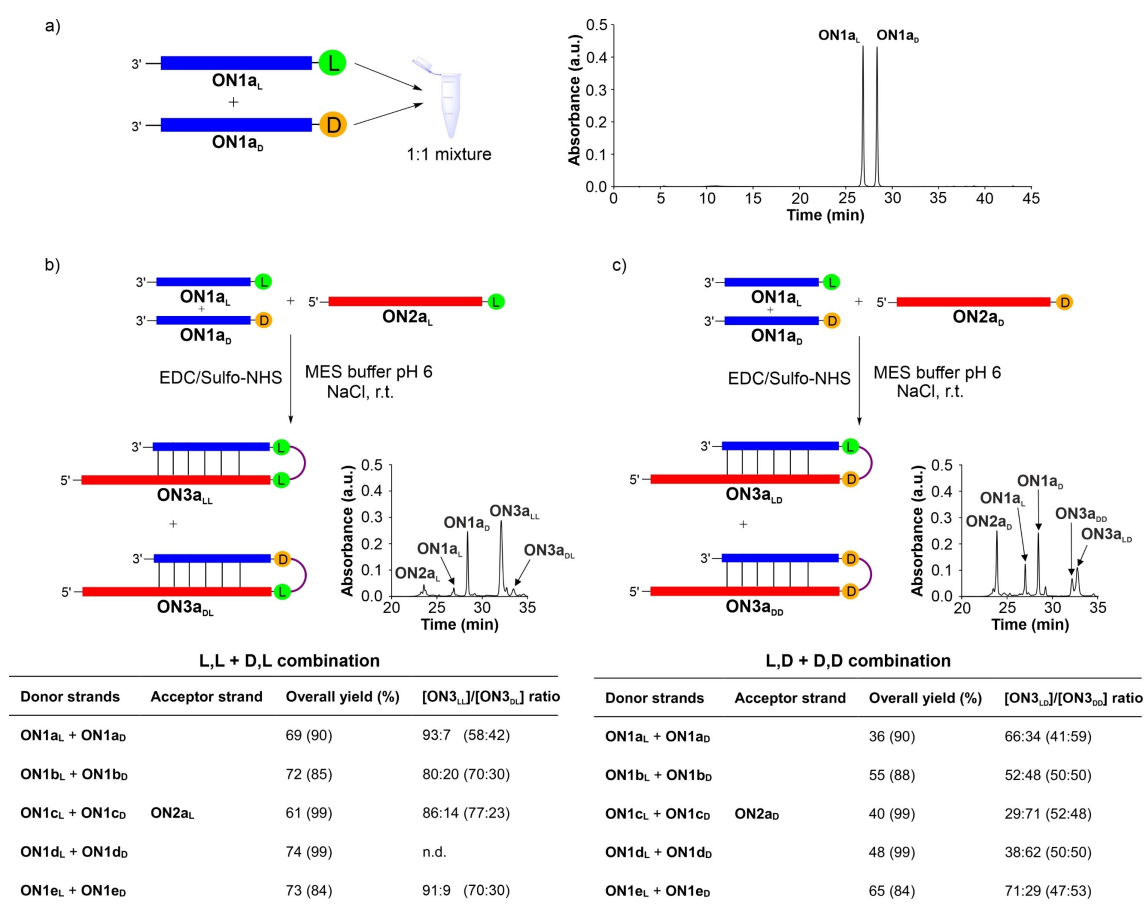


Figure 3. a) Equimolar mixture of **ON1a_L** and **ON1a_D**, and competitive peptide coupling reactions between **ON1a_L** and **ON1a_D** with: b) **ON2a_L** and c) **ON2a_D**. The HPLC chromatograms correspond to the analyzed crude reaction mixtures and the tables summarize the results obtained for the coupling reactions of **ON1a-e** and **ON2a** using EDC/Sulfo-NHS and DMTMM•Cl (in parenthesis) as activators. n.d. = not determined due to product overlap. Reaction conditions: [**ON1**] = 50 μM; [**ON2**] = 50 μM; [buffer] = 100 mM; [NaCl] = 100 mM and [activator] = 50 mM. Yields determined by HPLC analysis using the calibration curve of a reference compound (Figure S2).

RNA strands **ON1b-e** loaded with Ala, Phe, Trp and Thr (Figure 3 and Figures S4-S5). The competition experiments with the acceptor RNA strands, **ON2a_L** and **ON2a_D**, provided, in all cases, the expected hairpin products **ON3b-e**. Again, with EDC/Sulfo-NHS activator, we observed better yields with **ON2a_L** (61-74 %) compared to the reaction with **ON2a_D** (40-65 %). For the L,L:D,L-selectivity using **ON2a_L**, we measured diastereoselectivities from 80:20 (L,L:D,L, **ON1b**) to 91:9 (L,L:D,L, **ON1e**) (Figure 4a). In contrast, turning the selectivity experiment around (L,D:D,D-combinations) with the acceptor strand **ON2a_D**, bearing a D-amino acid, provided unpredictable and largely fluctuating selectivities ranging from 29:71 (L,D:D,D, **ON1c**) to 71:29 (L,D:D,D, **ON1e**) (Figure 4b).

When we changed the activation method to DMTMM•Cl, we observed diastereoselectivities from 58:42

(L,L:D,L, **ON1a**) to 77:23 (L,L:D,L, **ON1c**) for the preferred L:D,L-combinations with **ON2a_L** and only ca. 50:50 (L,D:D,D) mixtures for the L:D,D-combinations with **ON2a_D** (Figure 3b-c). In all cases, the L,L-combinations with **ON2a_L** gave high yields and better diastereoselectivities.

We next wanted to know whether these selectivities change when we use a prebiotically more suitable activation method and performed experiments with methyl isonitrile (MeNC) in 4,5-dicyanoimidazole (DCI) buffer at pH 6.^[55] We observed in the experiments with selected donor RNA strands, **ON1b** (Ala) and **ON1c** (Phe), the clean formation of the hairpin products **ON3b-c** in less than 16 % overall yield (Table S6). However, the diastereoselectivities were very high with >72 % in favor of the homochiral dipeptide products. Only for the competitive reactions with the acceptor RNA strand **ON2a_D**, we obtained different diastereoselectivity ratios for the same amino acid depending on the activator used (EDC/Sulfo-NHS, DMTMM•Cl or MeNC).

In order to study how the selectivities evolve when larger peptides grow on RNA, we next prepared donor RNA strands containing L-homochiral di- and tripeptides with Val, **ON1f_L** and **ON1g_L**, respectively (see SI). In the first set of competition experiments with the activator EDC/Sulfo-NHS, we used an equimolar mixture of **ON1a_L** with **ON1f_L** or **ON1g_L** (Figure 5a and Figure S7). The competitive coupling reactions with **ON2a_L** indicated that the mono-L-Val donor RNA strand **ON1a_L** reacted preferentially with the acceptor strand to give the L,L-dipeptide hairpin **ON3a_{LL}** with high selectivities of 73:27 and 89:11 over the larger tri- and tetrapeptide hairpin products, **ON3f_{LL}** and **ON3g_{LL}**. In contrast, the same competition reactions with the acceptor strand carrying a D-amino acid, **ON2a_D**, showed low selectivity (ca. 50:50).

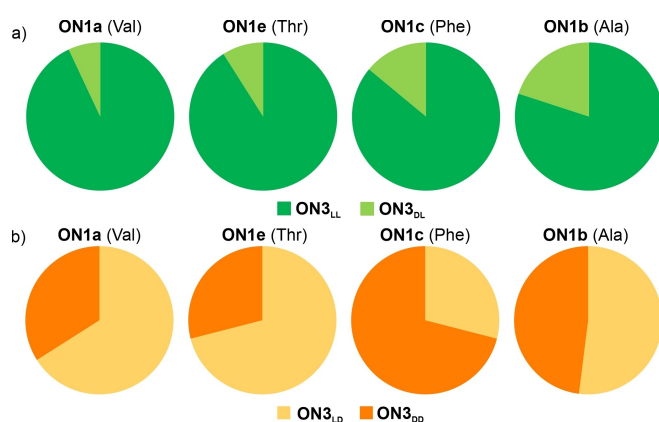


Figure 4. Relative isomeric composition of the hairpin products, **ON3**, obtained in the competitive peptide coupling reactions between **ON1** with: a) **ON2a_L** and b) **ON2a_D**.

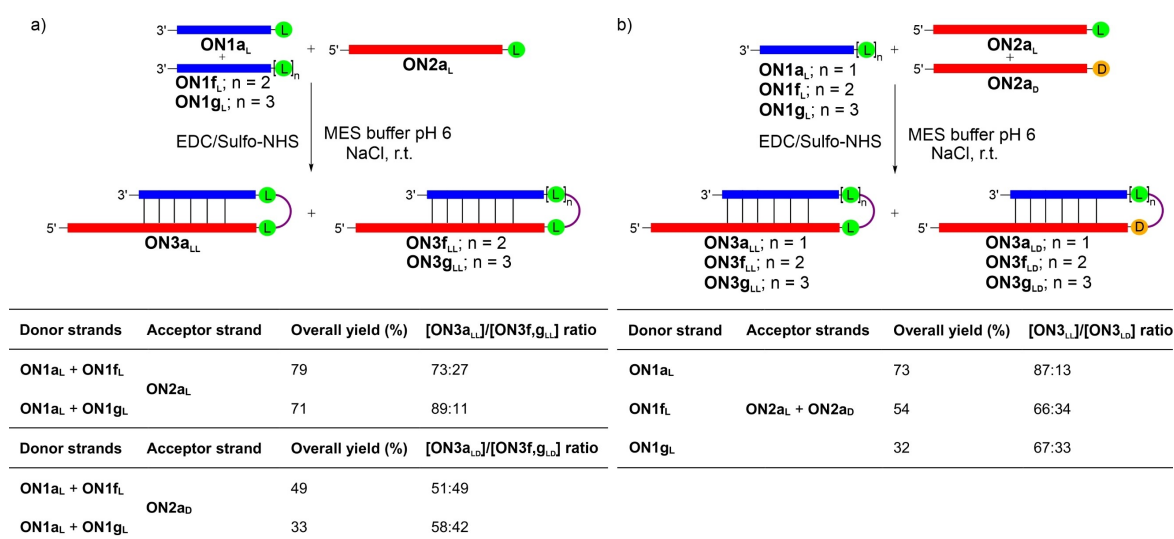


Figure 5. Competitive peptide coupling reactions a) between **ON2a_L** with **ON1a_L** and **ON1f_L**, **ON1g_L**, and b) between **ON2a_L** and **ON2a_D** with **ON1a_L**, **ON1f_L** or **ON1g_L**. The tables summarize the results obtained using EDC/Sulfo-NHS as activator. All amino acids and peptides contain Val. Reaction conditions: [ON1] = 50 μM; [ON2] = 50 μM; [buffer] = 100 mM; [NaCl] = 100 mM and [activator] = 50 mM. Yields determined by HPLC analysis using the calibration curve of a reference compound (Figure S2).

In a second set of experiments, we conducted reverse competition reactions with the activator EDC/Sulfo-NHS, in which an equimolar mixture of **ON2a_L** and **ON2a_D** was allowed to couple with either **ON1a_L**, **ON1f_L** or **ON1g_L** (Figure 5b and Figure S8, see SI for a competition experiment with **ON1a_D**). In all these cases, the results showed a clear preference for the formation of the L,L-homochiral products **ON3_{LL}**. The highest diastereoselectivity was obtained with the mono-L-Val donor RNA strand **ON1a_L** (87:13 L,L:L,D). The longer di- and tripeptides in **ON1f_L** and **ON1g_L** led to decreased selectivities of 66:34 and 67:33 (L,L:L,D), respectively. All together, these data suggest a strong tendency for the preferred formation of L-homochiral peptides, when the peptide synthesis occurs in close proximity to RNA.

In order to investigate whether the tendency to grow all-L-homochiral peptides is the result of different thermodynamic stabilities of the initially formed RNA duplexes, we measured melting temperatures (T_m) and we determined rate constants (k_{app}) for the peptide coupling reactions using selected donor and acceptor RNA strands. For reasons of prebiotic plausibility, we decided to perform these experiments with the amino acid Val in **ON1a** and **ON2a**, respectively.

First, we measured the melting temperatures of all possible duplex combinations (**ON1a_L•ON2a_L**, **ON1a_D•ON2a_L**, **ON1a_L•ON2a_D** and **ON1a_D•ON2a_D**) using temperature-dependent UV spectroscopy experiments at 260 nm (Figure 6a). All recorded melting curves displayed a mono-sigmoidal shape and were fit to a two-state model.^[56] The obtained data showed that the T_m values were

indistinguishable for the four donor•acceptor RNA duplex combinations with $T_m = 37.1 \pm 0.8^\circ\text{C}$. In addition, we measured the same T_m value for an RNA duplex composed of **ON1a_L** and **ON2b**, which lacked the amino acid at the 3'-terminal mm⁵U nucleotide (Figure 6a). These data showed that different duplex stabilities cannot explain the observed diastereoselectivities.

To determine the rate constants, we monitored separate peptide coupling reactions between the complementary donor and acceptor RNA strands at different time intervals using EDC/Sulfo-NHS (Figure 6b and Figure S10) and DMTMM•Cl (Figures S9-S10) as activators. The kinetic data for the formation of the hairpin products (**ON3a_{LL}**, **ON3a_{DL}**, **ON3a_{LD}** and **ON3a_{DD}**) were fit to a pseudo-first order equation^[57] (see SI). The calculated apparent rate constant values were, as expected, different and dependent on the donor RNA strand, **ON1a_L** and **ON1a_D**, as well as the activator used (Table 1). To our delight, we noted that the hairpin product ratios, calculated using exclusively the k_{app} values ($k_{app}(\text{ON1a}_L)/k_{app}(\text{ON1a}_D) = [\text{ON3a}_L]/[\text{ON3a}_D]$), are in good agreement with the values obtained in the competition experiments (Figure 3b-c). Taken together, these data support the idea that the L-amino acids are better aligned for reaction on RNA (constructed from D-ribose). The D-ribose creates a right-handed A-form RNA helix that appears to promote L-homochirality if the amino acids react directly on the RNA to form peptides.

Conclusions

Nature utilizes homo-D-ribose to build the backbone of DNA and RNA and homo-L-configured α -amino acids to create well-defined and functional protein structures. Herein, we investigated the question of how RNA can help to

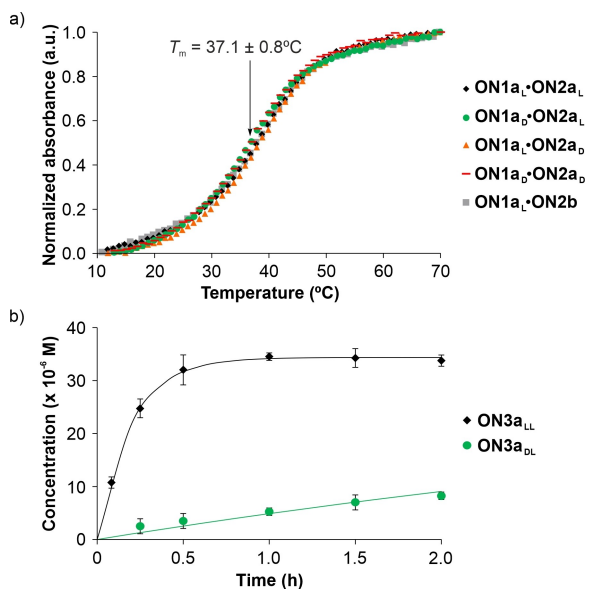


Figure 6. a) UV (260 nm) melting curves of the four possible duplexes formed between the donor RNA strands **ON1a_L** and **ON1a_D** with the acceptor RNA strands **ON2a_L** and **ON2a_D**. The melting curve of **ON1a_L** with **ON2b** (acceptor RNA strand without amino acid) is also shown. b) Kinetic plots for the separate peptide coupling reactions of **ON2a_L** with **ON1a_L** and **ON1a_D**, respectively, using EDC/Sulfo-NHS as activator. Error bars are shown as standard deviations.

Table 1: Apparent rate constant values (k_{app}) determined for the peptide coupling reactions of **ON1a** with **ON2a** using EDC/Sulfo-NHS and DMTMM•Cl as activators. Errors are indicated as standard deviations.

Activator	Donor strand	Acceptor strand	Hairpin strand	k_{app} (h ⁻¹)	Calcd. ratio ^a
EDC/ Sulfo-NHS	ON1a_L	ON2a_L	ON3a_{LL}	5.04 ± 0.12	87
			ON3a_{DL}	0.15 ± 0.01	13
	ON1a_D	ON2a_D	ON3a_{LD}	n.d.	n.d.
			ON3a_{DD}	0.32 ± 0.01	n.d.
DMTMM•Cl	ON1a_L	ON2a_L	ON3a_{LL}	0.91 ± 0.07	59
			ON3a_{DL}	0.43 ± 0.01	41
	ON1a_D	ON2a_D	ON3a_{LD}	0.54 ± 0.01	47
			ON3a_{DD}	0.68 ± 0.03	53

^a Product ratio calculated from the measured apparent rate constants. n.d. = not determined due to signal overlap with activated species.

establish an L-homochiral peptide world. The basic assumption was that initially peptides were formed close to RNA or directly connected to RNA in an RNA-peptide world. To simulate this process, we used our recently introduced concept of peptide growth directly on RNA as a model.^[18]

First, we studied the stereochemical preference of the loading reaction of an amino acid onto RNA. We observed that this reaction provided donor RNA strands with a 1:1 relationship of the L- and D-amino acids. Specifically, we studied how L/D-Phe reacts with an RNA strand, equipped with a terminal *N*⁶-methylcarbamoyl adenosine nucleotide, which creates in situ an *N*⁶-isocyanate. Both L- and D-amino acids reacted with the same preference and hence no diastereoselectivity was observed. Next, we investigated the coupling reaction with a set of selected amino acids (Val, Ala, Phe, Trp and Thr), and we observed a high preference for the coupling of L-configured amino acids connected to RNAs. In all cases, we found that the L,L-homochiral dipeptide hairpin products were formed in better yields and with higher diastereoselectivities. This preference was maintained even when larger tri- and tetrapeptide hairpins, for the growth of longer peptides, were formed. On the other hand, the coupling reactions with a D-amino acid on the acceptor RNA strand provided variable selectivities depending on the amino acid and the activator, potentially because of an unfavorable steric alignment of the amino acids for coupling.

These data allow us to formulate a scenario that yields an all-L-homochiral peptide world. If we assume that small RNAs (or nucleosides) captured amino acids and peptides out of the primordial soup to bring them in close proximity within an RNA duplex, the connection of the amino acids to each other is strongly favored when the amino acids are L-configured, which leads with time to an enrichment of the L,L-coupling products. Repetitive cycles of peptide growth and peptide release followed by recapturing of the peptides by RNA and connection of the recaptured peptides with either amino acids or peptides could then lead to a possible outgrowth process that provides a homo-L-configured peptide world.

What it is next needed is a scenario that leads to the preferred recapturing of peptides over single amino acids and fragment condensation reactions, which were already observed,^[18] that quickly grow longer peptide structures in just a few chemical steps.

The most important result of this study is that D-ribose-based nucleotides, which give rise to right-handed A-form RNA duplexes, preferentially connect L-amino acids to form homo-L-peptides. We do not address in this study the question of the origin of homo-D-RNA.

Supporting Information

Reaction procedures, characterization data, peptide coupling reactions, melting curves, kinetic experiments and additional references are included in the Supporting Information (Ref. [17,18,19,56,57,58,59,60,61,62,63]).

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for supporting this research through the DFG grants: CA275/11-3 (ID: 326039064), CRC1309 (ID: 325871075, A4), CRC1032 (ID: 201269156, A5) and CRC1361 (ID: 393547839, P2). We thank the Volkswagen Foundation for funding this research (grant EvoRib). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No. 741912 (EPiR). L. E. thanks the Alexander von Humboldt Foundation for a postdoctoral fellowship (ESP 1214218 HFST-P). Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Amino Acids · Diastereoselectivity · Homochirality · RNA-Peptide World · Origin of Life

- [1] C. M. Dobson, *Nature* **2003**, *426*, 884–890.
- [2] D. Balchin, M. Hayer-Hartl, F. U. Hartl, *Science* **2016**, *353*, aac4354.
- [3] B. Kuhlman, P. Bradley, *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 681–697.
- [4] R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. V. Fratini, M. L. Kopka, *Science* **1982**, *216*, 475–485.
- [5] P. Belmont, J.-F. Constant, M. Demeunynck, *Chem. Soc. Rev.* **2001**, *30*, 70–81.
- [6] A. Travers, G. Muskhelishvili, *FEBS J.* **2015**, *282*, 2279–2295.
- [7] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737–738.
- [8] W. A. Bonner, N. E. Blair, F. M. Dirbas, *Origins Life Evol. Biospheres* **1981**, *11*, 119–134.
- [9] K. Tamura, *J. Mol. Evol.* **2019**, *87*, 143–146.
- [10] S. Toxvaerd, *Int. J. Mol. Sci.* **2009**, *10*, 1290–1299.
- [11] S. F. Ozturk, D. D. Sasselov, *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2204765119.
- [12] S. F. Ozturk, Z. Liu, J. D. Sutherland, D. D. Sasselov, *Sci. Adv.* **2023**, *9*, eadg8274.
- [13] L. E. Orgel, *J. Mol. Biol.* **1968**, *38*, 381–393.
- [14] W. Gilbert, *Nature* **1986**, *319*, 618.
- [15] G. F. Joyce, *Nature* **2002**, *418*, 214–221.
- [16] J. C. Bowman, N. V. Hud, L. D. Williams, *J. Mol. Evol.* **2015**, *80*, 143–161.
- [17] M. Nainytė, F. Müller, G. Ganazzoli, C.-Y. Chan, A. Crisp, D. Globisch, T. Carell, *Chem. Eur. J.* **2020**, *26*, 14856–14860.
- [18] F. Müller, L. Escobar, F. Xu, E. Węgrzyn, M. Nainytė, T. Amatov, C. Y. Chan, A. Pichler, T. Carell, *Nature* **2022**, *605*, 279–284.
- [19] J. N. Singer, F. M. Müller, E. Węgrzyn, C. Hölzl, H. Hurmiz, C. Liu, L. Escobar, T. Carell, *Angew. Chem. Int. Ed.* **2023**, *62*, e202302360.
- [20] P. F. Agris, *Nucleic Acids Res.* **2004**, *32*, 223–238.

- [21] T. Carell, C. Brandmayr, A. Hienzsch, M. Muller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs, M. Wagner, *Angew. Chem. Int. Ed.* **2012**, *51*, 7110–7131.
- [22] H. Grosjean, E. Westhof, *Nucleic Acids Res.* **2016**, *44*, 8020–8040.
- [23] M. P. Robertson, S. L. Miller, *Science* **1995**, *268*, 702–705.
- [24] M. Di Giulio, *J. Theor. Biol.* **1998**, *191*, 191–196.
- [25] J. C. Lacey, R. D. Thomas, M. P. Staves, C. L. Watkins, *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* **1991**, *1076*, 395–400.
- [26] N. S. M. D. Wickramasinghe, M. P. Staves, J. C. Lacey Jr., *Biochemistry* **1991**, *30*, 2768–2772.
- [27] K. Tamura, P. Schimmel, *Science* **2004**, *305*, 1253.
- [28] K. Tamura, P. R. Schimmel, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13750–13752.
- [29] K. Tamura, *BioSystems* **2008**, *92*, 91–98.
- [30] K. Tamura, *Int. J. Mol. Sci.* **2011**, *12*, 4745–4757.
- [31] K. Tamura, *Life* **2015**, *5*, 1687–1699.
- [32] L.-F. Wu, M. Su, Z. Liu, S. J. Bjork, J. D. Sutherland, *J. Am. Chem. Soc.* **2021**, *143*, 11836–11842.
- [33] S. J. Roberts, Z. Liu, J. D. Sutherland, *J. Am. Chem. Soc.* **2022**, *144*, 4254–4259.
- [34] M. Su, C. Schmitt, Z. Liu, S. J. Roberts, K. C. Liu, K. Röder, A. Jäschke, D. J. Wales, J. D. Sutherland, *J. Am. Chem. Soc.* **2023**, *145*, 15971–15980.
- [35] O. Doppleb, J. Bremer, M. Bechthold, C. Sánchez Rico, D. Göhringer, H. Griesser, C. Richert, *Chem. Eur. J.* **2021**, *27*, 13544–13551.
- [36] S. A. Kauffman, N. Lehman, *Interface Focus* **2023**, *13*, 20230009.
- [37] J. J. Petkowski, W. Bains, S. Seager, *Molecules* **2019**, *24*, 866.
- [38] G. B. Chheda, R. H. Hall, D. I. Magrath, J. Mozejko, M. P. Schweizer, L. Stasiuk, P. R. Taylor, *Biochemistry* **1969**, *8*, 3278–3282.
- [39] M. P. Schweizer, K. McGrath, L. Baczynskyj, *Biochem. Biophys. Res. Commun.* **1970**, *40*, 1046–1052.
- [40] F. Kimura-Harada, D. L. Von Minden, J. A. McCloskey, S. Nishimura, *Biochemistry* **1972**, *11*, 3910–3915.
- [41] L. Perrochia, E. Crozat, A. Hecker, W. Zhang, J. Bareille, B. Collinet, H. van Tilbeurgh, P. Forterre, T. Basta, *Nucleic Acids Res.* **2013**, *41*, 1953–1964.
- [42] C. Schneider, S. Becker, H. Okamura, A. Crisp, T. Amatov, M. Stadlmeier, T. Carell, *Angew. Chem. Int. Ed.* **2018**, *57*, 5943–5946.
- [43] L. L. Cummins, S. R. Owens, L. M. Risen, E. A. Lesnik, S. M. Freier, D. McGee, C. J. Guinosso, P. D. Cook, *Nucleic Acids Res.* **1995**, *23*, 2019–2024.
- [44] S. M. Freier, K.-H. Altmann, *Nucleic Acids Res.* **1997**, *25*, 4429–4443.
- [45] W. A. Decatur, M. J. Fournier, *Trends Biochem. Sci.* **2002**, *27*, 344–351.
- [46] L. M. Longo, J. Lee, M. Blaber, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2135–2139.
- [47] S. L. Miller, H. C. Urey, *Science* **1959**, *130*, 245–251.
- [48] D. Ring, Y. Wolman, N. Friedmann, S. L. Miller, *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 765–768.
- [49] Y. Wolman, W. J. Haverland, S. L. Miller, *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 809–811.
- [50] C. Huber, G. Wächtershäuser, *Science* **2006**, *314*, 630–632.
- [51] L. M. Longo, M. Blaber, *Arch. Biochem. Biophys.* **2012**, *526*, 16–21.
- [52] M. Kimura, S. Akanuma, *J. Mol. Evol.* **2020**, *88*, 372–381.
- [53] R. Root-Bernstein, A. G. Baker, T. Rhinesmith, M. Turke, J. Huber, A. W. Brown, *Life* **2023**, *13*, 265.
- [54] Z. J. Gartner, M. W. Kanan, D. R. Liu, *Angew. Chem. Int. Ed.* **2002**, *41*, 1796–1800.
- [55] Z. Liu, L.-F. Wu, J. Xu, C. Bonfio, D. A. Russell, J. D. Sutherland, *Nat. Chem.* **2020**, *12*, 1023–1028.
- [56] S. G. J. Mochrie, *Am. J. Phys.* **2011**, *79*, 1121–1126.
- [57] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, U. Kummer, *Bioinformatics* **2006**, *22*, 3067–3074.
- [58] F. Himmelsbach, B. S. Schulz, T. Trichtinger, R. Charubala, W. Pfeleiderer, *Tetrahedron* **1984**, *40*, 59–72.
- [59] F. Ferreira, F. Morvan, *Nucleosides Nucleotides Nucleic Acids* **2005**, *24*, 1009–1013.
- [60] A. V. Tataurov, Y. You, R. Owczarzy, *Biophys. Chem.* **2008**, *133*, 66–70.
- [61] G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176–2179.
- [62] X. Li, J. Wu, X. Li, W. Mu, X. Liu, Y. Jin, W. Xu, Y. Zhang, *Bioorg. Med. Chem.* **2015**, *23*, 6258–6270.
- [63] D. L. Usanov, A. I. Chan, J. P. Maianti, D. R. Liu, *Nat. Chem.* **2018**, *10*, 704–714.

Manuscript received: December 13, 2023

Accepted manuscript online: February 26, 2024

Version of record online: April 5, 2024

3.4 Gradual evolution of a homo-L-peptide world on homo-D-configured RNA and DNA

Authors

Ewa Węgrzyn^{*}, Ivana Mejdrová^{*} und Thomas Carell[°]

Chem. Sci. **2024**, 15, 14171-14176

DOI: <https://doi.org/10.1039/D4SC03384A>

Supplementary Information included in the Appendix IV

^{*} These authors contributed equally [°] corresponding author

Summary

In our current work, we combine the stereoselective RNA-mediated peptide coupling with the prebiotically plausible cold/hot-directed synthesis cycles, resembling the temperature variations prevalent on early Earth. We show that the peptides can couple stereoselectively, followed by thermal cleavage of the hairpin intermediate and release of the formed peptide from the donor strand. These reactions can be then repeated under one-pot conditions and the iterative cycles can eventually lead to a gradual enrichment of the homochiral L-peptides. We also show, that even if dipeptides with mixed L-/D-stereochemistry are formed, they get outcompeted by the homochiral counterparts. Additionally, we demonstrate that the preference for homochiral peptide formation is also maintained in DNA.

Personal contribution

Synthesis of modified phosphoramidites and oligonucleotides, characterization of coupling products and performance of competition reactions, including one-pot reactions, measurement of the calibration curves for quantification, measurement of CD spectra, preparation of the manuscript and supporting information, planning of the project and development of concepts.

Cite this: *Chem. Sci.*, 2024, 15, 14171

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 23rd May 2024

Accepted 26th July 2024

DOI: 10.1039/d4sc03384a

rsc.li/chemical-science

Gradual evolution of a homo-L-peptide world on homo-D-configured RNA and DNA†

Ewa Węgrzyn,[‡] Ivana Mejdrová,[‡] and Thomas Carell^{‡*}

Modern life requires the translation of genetic information – encoded by nucleic acids – into proteins, which establishes the essential link between genotype and phenotype. During translation, exclusively L-amino acids are loaded onto transfer RNA molecules (tRNA), which are then connected at the ribosome to give homo-L-proteins. In contrast to the homo-L-configuration of amino acids and proteins, the oligonucleotides involved are all D-configured (deoxy)ribosides. Previously, others and us have shown that if peptide synthesis occurs at homo D-configured oligonucleotides, a pronounced L-amino acid selectivity is observed, which reflects the D-sugar/L-amino acid world that evolved in nature. Here we further explore this astonishing selectivity. We show a peptide-synthesis/recapture-cycle that can lead to a gradual enrichment and hence selection of a homo-L-peptide world. We show that even if peptides with a mixed L/D-stereochemistry are formed, they are not competitive against the homo-L-counterparts. We also demonstrate that this selectivity is not limited to RNA but that peptide synthesis on DNA features the same L-amino acid preference. In total, the data bring us a step closer to an understanding of how homochirality on Earth once evolved.

Introduction

Life, as we know it, requires nucleic acid biomolecules to encode genetic information and amino acid-based proteins to catalyse biochemical reactions, that are essential for the maintenance of life. While the genetic code contains the blueprint for the multitude of vital proteins, the proteins in turn are essential for metabolism and required for decoding the sequence information and its replication. This strong interdependence of the genotype established by nucleic acids and the phenotype established by proteins is a hallmark of all life on Earth.¹ The element where the genotype “meets” the phenotype is the ribosome.^{2–4} Consequently, the evolution of the ribosome is a chicken-and-egg conundrum and one of the most challenging mysteries of the origin of life.^{5,6} The translation machinery that is in place today shows an extremely high stereoselectivity towards L-amino acids. First, a specific transfer RNA (tRNA) consisting of D-ribose is selectively loaded with the corresponding L-amino acid.⁷ In the step of translation, these L-amino acids are connected to other L-amino acids in a messenger RNA (mRNA)-based templation process, catalysed by the RNA components (all L-configured) of the ribosome. The exact time-point when homochirality emerged is unknown. Mechanistically it was suggested that it was caused

by a small initial enantiomeric imbalance on the monomer level,⁸ followed by processes of chiral amplification,^{9,10} leading to an enantiomeric induction between homochiral nucleic acids and peptides.^{11,12}

Numerous ideas exist of how the decoding¹³ could have evolved,^{14–16} which are all connected to the RNA world concept.^{17–19} This model predicts that in prebiotic times it was only RNA which encoded information and which also catalysed the essential reactions such as loading itself with specific amino acids.^{20–22} In the RNA world, the process of connecting amino acids was catalysed by RNA.^{22,23} The fact that the peptidyl transferase center of the ribosome is composed of RNA,²⁴ provides strong support for this “RNA-only” idea.^{25,26}

Currently, the idea of an RNA-only world is questioned. Recent studies²⁷ suggest instead the coexistence/coevolution of RNA with other entities like peptides,²⁸ cofactors^{29,30} or DNA.^{31–33} Even a DNA-first theory was formulated.^{33,34} The fact that most known RNA catalysts have rather low turn-over numbers strongly suggests that amino acids, peptides and cofactors might have been early on involved.³⁵ Already in 1976, White pointed out that cofactors alone could be considered as remnants of early life ribozymes,^{29,30} which acted in concert with sulfur-containing modified nucleosides.³⁶ Recently the idea of an RNA-peptide world was brought forward,^{37,38} with the peptides adding stability and catalytic capabilities to the RNA.³⁹ Such an early partnership could have been the foundation for the development of early amino acid-tRNA synthetases.⁴⁰

We recently introduced the idea, that the high number of non-canonical nucleosides present in RNA, are living fossils^{41,42}

Department of Chemistry, Center for Nucleic Acids Therapies at the Institute for Chemical Epigenetics (ICE-M), Ludwig-Maximilians-Universität (LMU), München Butenandtstrasse 5-13, 81377 Munich, Germany. E-mail: thomas.carell@lmu.de

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4sc03384a>

‡ These authors contributed equally.



of an early RNA/DNA world.⁴³ We could show that these non-canonical nucleosides add functions to RNA that would strongly benefit the evolution of life.⁴⁴ For example, non-canonical nucleosides t⁶A, g⁶A, as well as (m)nm⁵U,^{45–47} enable RNA to grow peptides on itself in a process which can establish an RNA-peptide world (Fig. 1).⁴⁸ The loading with amino acids and peptides and their connection to more complex structures is in principle possible under prebiotically plausible freezing/thawing conditions.^{49,50} It was shown, that this loading reaction exhibits no stereochemical bias, meaning that L- and D-amino acids are loaded onto RNA with no preference.^{12,51} However, when the amino acids are allowed to react with each other to give peptides, a large kinetic advantage was observed for the L-configured amino acids.^{11,52–57}

Based on this observation, we can envision a primitive peptide synthesis cycle as depicted in Fig. 1, that is driven by cold/hot cycles (−20 °C to +90 °C). In this model, amino acids (aa) and small peptides react (*via* nitrosation of N⁶-urea-A) with RNA (step I and II) to give (m⁶)aa⁶A-containing RNA-donor strands ((m⁶)aa⁶A-RNA),⁵¹ which can hybridize to complementary RNA acceptor strands containing the non-canonical

nucleoside (m)nm⁵U (step 0) in the cold. Upon activation and reaction of the amino acids with the (m)nm⁵U nucleoside (step 1), a hairpin structure with a high thermal stability is formed.⁴⁸ During a phase of increased temperature (90 °C), cleavage of the thermally labile urea bond will occur (step 2) together with an immediate dissociation of the strands (step 3). Upon cooling of the solution, a new donor strand can bind (step 4) to continue the hot/cold-governed synthesis cycle.⁵⁸ In the prebiotic context, these temperature variations could have originated from volcanic eruptions, meteorite impact or even day/night-dependent presence of sunlight.^{59,60} Current research reports on one hand cold, even frozen environments on early Earth,^{61,62} and on the other hot temperatures provided by the hydrothermal vents.⁶³

We hypothesise that the initially formed peptides were short and released into the environment upon degradation of the RNA. And such they could have been taken up again by donor strands to continue the process at another RNA “host” strand (step II). Here we show that such a scenario is realistic and we show that homo-L peptides would feature a clear competitive advantage over peptide strands contaminated with D-amino acids.

Results and discussion

In order to establish a stereochemical self-selecting peptide synthesis cycle, we intended to investigate the one-pot growth of a peptide using a racemic mixture of L- and D-Val RNA donor strands (Fig. 2). We selected valine because it is one of the simplest chiral amino acids for which a prebiotic existence is plausible.^{64,65} Valine has also been observed in meteorites^{66,67} and its formation has been demonstrated in Urey–Miller-type spark discharge experiments.^{68–71}

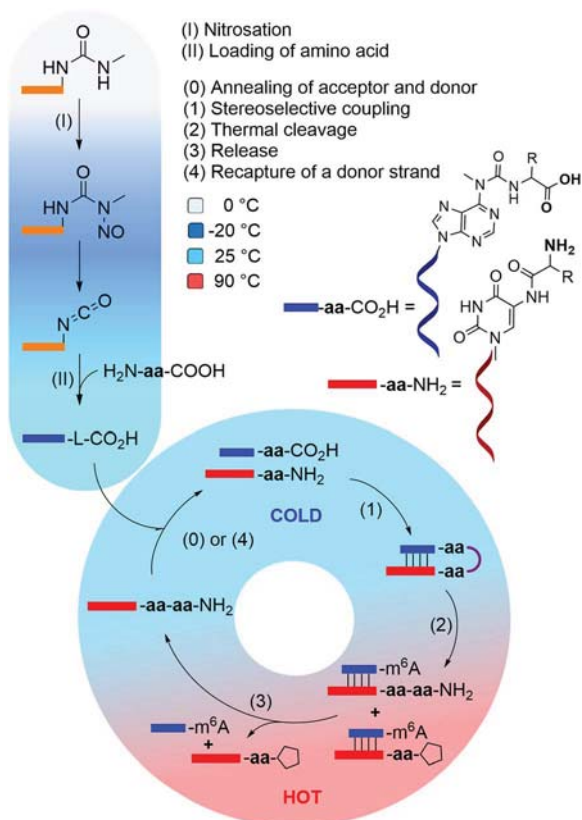


Fig. 1 Cold/hot cycle of amino acid loading onto RNA (steps I and II) and iterative peptide synthesis cycle, that could lead to homochiral peptide formation in an outgrowth process. The steps of the cycle include: annealing of the complementary acceptor and donor strands (step 0), stereoselective peptide coupling (step 1), thermal cleavage of the peptide from the donor strand (step 2), followed by a release of the peptide-RNA acceptor conjugate (step 3), that can then recapture another amino acid-loaded donor strand (step 4), and enter another cycle. R stands for L- or D-Val.

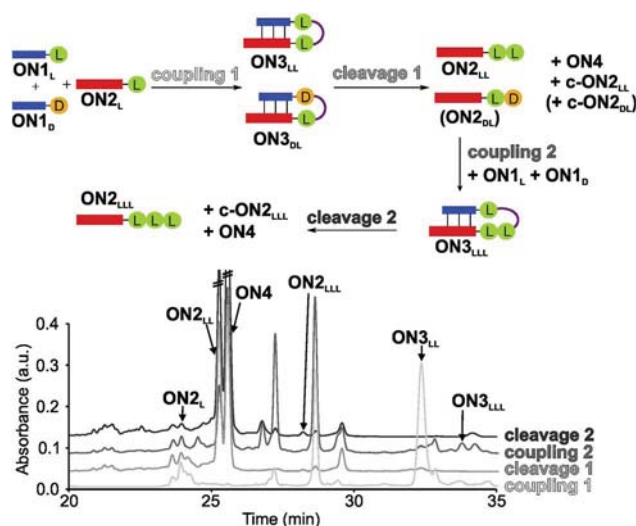


Fig. 2 One-pot synthesis cycle scheme and HPLC-chromatograms. For reasons of simplicity, only the products formed in detectable amounts and relevant for further reactions are indicated in the scheme. Yields were determined by HPLC using the calibration curves of reference compounds (Fig. S2 and S3†).



For the experiment, we first synthesized several 7-mer RNA donor strands with either L- or D-Val **ON1_L** (L = L-Val) or **ON1_D** (D = D-Val), or a stereochemically mixed Val-dipeptide, connected to the RNA *via* 5'-N⁶-methylcarbamoyl adenosine nucleosides (m⁶-L/D-Val⁶A_m). As the complementary RNA acceptor strand we prepared an 11-mer RNA strand with a 3'-L- or D-Val-modified 5-methylaminomethyl uridine (mm⁵U) nucleotide (**ON2_L** or **ON2_D**). 2'OMe nucleotides were used for the construction of the RNA to increase the thermal stability.^{72,73} We next combined all the donor **ON1_L** or **ON1_D** with all the acceptor **ON2_L** or **ON2_D** strands in the cold part of the cycle and performed the peptide coupling with EDC/Sulfo-NHS as activator (MES buffer pH 6, NaCl, r.t., 2 h), to obtain the different hairpin products **ON3_{LL}**, **ON3_{DL}** and **ON3_{LD}**, **ON3_{DD}** (not shown). In the hot part of the cycle, the hairpin products were cleaved (acetate buffer pH 4, NaCl, 90 °C, 48 h). As the result of the cycle we obtained in all cases the expected dipeptide products **ON2_{LL}**, **ON2_{DL}** and **ON2_{LD}**, **ON2_{DD}** (not shown), and the peptide-free m⁶A-RNA **ON4**, together with a small amount of the corresponding hydantoin by-product **c-ON2_{LL}**, **c-ON2_{LD}**, **c-ON2_{DL}** or **c-ON2_{DD}** (see Scheme S6 in ESI†).⁴⁸ All the oligonucleotides formed were purified by high-performance liquid chromatography (HPLC) and characterized using MALDI-ToF mass spectrometry. The experiment showed that the hot/cold cycle is in all cases effective and the obtained products were next used as standards for the analysis of potential diastereoselectivities. It is known from literature, that the measured diastereoselectivity of peptide formation in the absence of the sugar is very low, arguing that it is sugar and the helical chirality that induces the selectivity.⁵⁷

In order to analyse the diastereoselectivities, we next prepared an equimolar solution of 1 eq **ON1_L** and 1 eq **ON1_D** donor strands (always with Val) and added 1 eq of **ON2_L** (Fig. 2 and S4†). The coupling reaction was performed as described above. By using the generated oligonucleotide products from above as references, we determined by HPLC a diastereoselectivity of 94 : 6 in favour of the LL-homochiral product **ON3_{LL}** over the heterochiral DL-product **ON3_{DL}** with an overall yield of 58%. After simple filtration of the crude reaction mixture to remove unreacted activator and exchange of the buffer solution, the crude product was introduced into the hot-phase cleavage reaction (48 h, 90 °C). We indeed obtained the cleaved product **ON2_{LL}** with a diastereoselectivity of 98 : 2 for the LL-dipeptide over the second product **ON2_{DL}**. In addition, the hydantoin by-products **c-ON2_{LL}** and **c-ON2_{DL}** were formed. The overall yield was 42% over the two steps.

The reaction mixture was next lyophilized (in analogy to a dry-down step) and the buffer was exchanged (comparable to fresh solution floating in a tidal process). We then added another portion of the equimolar mixture of **ON1_L** and **ON1_D**. After activation, we detected the formation of the hairpin product **ON3_{LLL}** by HPLC and MALDI-ToF, albeit in a low yield of 3% (over three steps). The heterochiral product **ON3_{LLD}** was not detected. We then went through another hot-phase and performed the cleavage reaction, which provided in 1% yield (over 4 steps) the RNA strand **ON2_{LLL}** (Fig. 2 and S5; Table S8†). The main reason for the low yield is the large amount of

possible side reactions due to a high number of reactants present in the one-pot mixture.¹²

Although we detected a high selectivity for the formation of the LL- and LLL-products, we next wanted to learn how the presence of a D-contamination would influence the process. To study this, we synthesized the L,L-Val (**ON1_{LL}**), L,D- (**ON1_{LD}**), D,L- (**ON1_{DL}**) and D,D-Val (**ON1_{DD}**) RNA donor strands and performed the coupling reactions with either the **ON2_L** or the **ON2_D** acceptor strand. Again, in order to enable precise characterization, we first performed the coupling reactions with each donor-acceptor strand combination separately to obtain the tripeptide hairpin products, as reference compounds.

To study the diastereoselectivities (Fig. 3), we next prepared equimolar mixtures of the dipeptide RNA donor strands (referred to as **ON1_X** and **ON1_Y**, where X and Y corresponds to either LL, LD, DL or DD, while X ≠ Y).

We allowed the **ON1_X** and **ON1_Y** RNA oligonucleotides with the different dipeptides to compete for reactions with **ON2_L** (Fig. 3A and B) or **ON2_D** (Fig. 3A and C).

First, we allowed the equimolar mixtures of **ON1_X** and **ON1_Y** to react with **ON2_L**, using EDC/Sulfo-NHS as the activator (MES buffer pH 6, NaCl, r.t., 2 h). The reaction mixtures were then analysed by HPLC (Fig. 3 and Table S9†). We observed in the reactions yields between 25–56%. Next, we determined the ratio between **ON3_{XL}** and **ON3_{YL}**. Again, high diastereoselectivities in favor of products that were formed with an L-Val in close proximity of the RNA were observed. The selectivity for **ON3_{LLL}** over **ON3_{LDL}** was only 59 : 41 (Fig. 3Ba), but for **ON3_{LLL}** over **ON3_{DL}** we detected 81 : 19 (Fig. 3Bc), suggesting that the exchange of the L-amino acid directly at the nucleobase against the D-counterpart strongly decreases the diastereoselectivity. For the reaction of an equimolar mixture of **ON1_{LD}** and **ON1_{DD}** we observed a diastereoselectivity of 62 : 38 for the **ON3_{LDL}** product over **ON3_{DDL}** (Fig. 3Bg). Again, the more L-amino acids are present, the better is the peptide formation reaction. The ratios for the other combinations could not be calculated with certainty due to massive peak overlap.

In order to gain deeper insight how a D-amino acids positioned directly at the nucleobase influences the selectivities, we next repeated the competition reactions with **ON2_D** (Fig. 3C and Table S9†). Now, we detected highly fluctuating results, but again, even for **ON2_D** there is some selectivity to react with a donor strand containing a L-Val attached to the RNA-donor strand over a D-Val. We saw a rather limited influence of the second amino acid present in the donor strand. The highest stereoselectivity of 80 : 20 was observed for the products **ON3_{LLD}** vs. **ON3_{DL}** (Fig. 3Cd). The overall yields were ranging from 24 to 50%.

We noticed during the study that the second amino acid of *e.g.* **ON1_{LL}** undergoes a *ca.* 25% racemization when incubated under the described coupling conditions for 2 h. The mono-Val donors **ON1_L** and **ON1_D**, however, did not racemize when subjected to the same conditions. This phenomenon makes the analysis and interpretation of the described reactions difficult, but the general result is firmly established, that the homo-L situation is always winning.



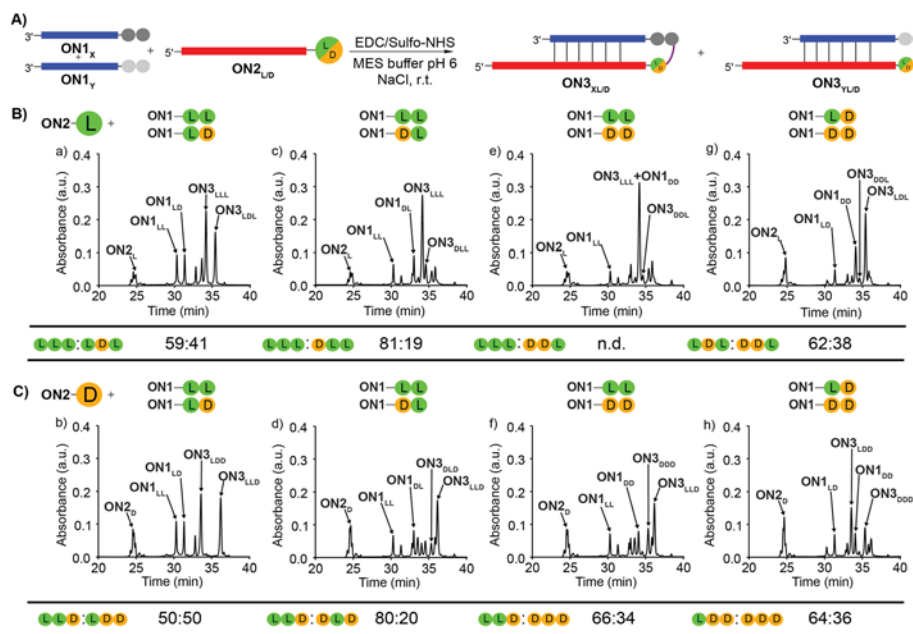


Fig. 3 (A) General scheme for the competitive peptide coupling reactions of equimolar mixture of ON1_x and ON1_y with (B) ON2_L or (C) ON2_D . X and Y can correspond to either LL, LD, DL or DD, while $X \neq Y$. The HPLC-chromatograms show the analysed crude reaction mixtures and the tables summarize the results obtained for the competitive coupling reactions, specifically product ratios of (B) $[\text{ON3}_{xLD}]/[\text{ON3}_{yLD}]$ and (C) $[\text{ON3}_{xD}]/[\text{ON3}_{yD}]$. Reaction conditions: $[\text{ON1}] = 50 \mu\text{M}$; $[\text{ON2}] = 50 \mu\text{M}$; [buffer] = 100 mM; [NaCl] = 100 mM and [activator] = 50 mM. Yields were determined by HPLC analysis using the calibration curve of a reference compound (Fig. S2†).

To complete the set of experiments with dipeptide donors, we also analysed the competition reactions of ON1_{LL} , ON1_{LD} , ON1_{DL} and ON1_{DD} with a 1 : 1 mixture of ON2_L and ON2_D (Fig. 4 and S7†). For this experiment we prepared an equimolar mixture of ON2_L and ON2_D and allowed them to react with the dipeptide RNA donor strand ON1_X under the described coupling conditions. The reaction with ON1_{LL} showed

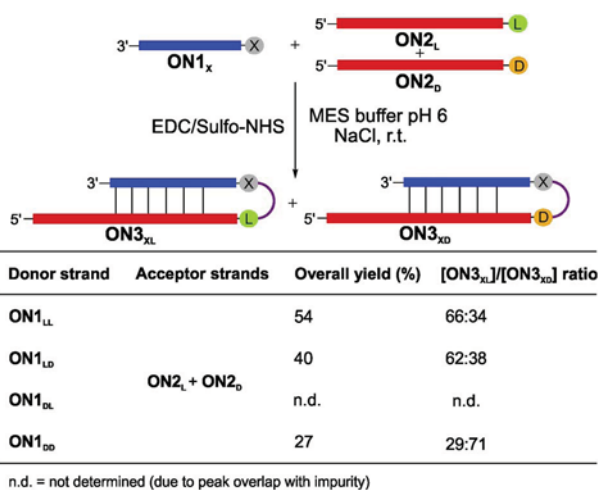


Fig. 4 Competitive peptide coupling reactions of equimolar mixture of ON2_L and ON2_D with ON1_X . X can correspond to either LL, LD, DL or DD. The table summarizes the results obtained for the competitive coupling reactions. Reaction conditions: $[\text{ON1}] = 50 \mu\text{M}$; $[\text{ON2}] = 50 \mu\text{M}$; [buffer] = 100 mM; [NaCl] = 100 mM and [activator] = 50 mM. Yields were determined by HPLC analysis using the calibration curve of a reference compound (Fig. S2†).

a selectivity of 66 : 34 in favour of the homochiral LLL -hairpin product ON3_{LLL} over the product ON3_{LLD} (Fig. 4). This diastereoselectivity decreased slightly when we used ON1_{LD} (62 : 38, $\text{ON3}_{LDL} : \text{ON3}_{LDD}$). The D,D -Val donor ON1_{DD} showed a reversed selectivity towards the homochiral all- D product ON3_{DDD} of 71 : 29 compared to ON3_{DDL} (Fig. S7†). The coupling reaction with ON1_{DL} could not be analysed due to a HPLC peak overlap of the product with an impurity.

Finally, we wanted to investigate if the observed tendency to form homo- L products prevails when one moves from RNA to DNA in light of the recent discussion that life may have started with a DNA-world³³ or a mixed RNA–DNA world.³¹ We therefore synthesized the DNA analogues of the donor strands dON1_L and dON1_D and the acceptor strands dON2_L and dON2_D by solid-phase oligonucleotide synthesis to form the hairpin products dON3 . We conducted the competitive coupling reactions and observed the formation of the dON3_{LL} with a high diastereoselectivity of 85 : 15 over the heterochiral product dON3_{DL} (Fig. 5a). The reverse competition reaction yielded a 31 : 69 stereoselectivity for dON3_{LD} over dON3_{DD} (Fig. 5b). These results suggest that DNA will induce a similar diastereoselectivity towards the homo- L products.

In order to gain more information about the conformation of the RNA duplexes, we finally measured the circular dichroism (CD) spectra of three different solutions: (1) annealed RNA duplex of ON1_L and ON2_L , (2) hairpin duplex ON3_{LL} and (3) the canonical analogue annealed RNA duplex of ON5 and ON6 (Fig. S9†). We observed CD-spectra typical for double-stranded A-form RNA with no significant changes introduced by the



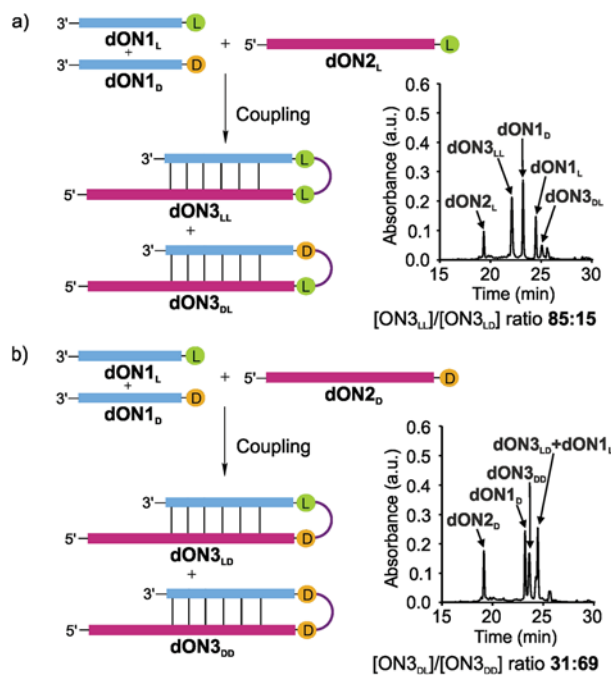


Fig. 5 DNA-competitive peptide coupling reactions between dON1_L and dON1_D with (a) dON2_L or (b) dON2_D using EDC/Sulfo-NHS as activator. Yields were determined by HPLC analysis using the calibration curve of a reference compound (Fig. S2†).

nucleobase modification or the hairpin formation between them.

Conclusions

In this work we explored the possibility of a primitive temperature-driven one-pot peptide synthesis cycle, in which stereoselective peptide coupling followed by thermal cleavage can lead to an enriched formation, albeit in low yields, of the homochiral all *L*-peptides.

We investigated how different stereochemical dipeptide combinations, namely the presence of *L*- and *D*-amino acids in the cycle, would influence the peptide-forming chemistry. We found that whatever combination we investigated, in almost all cases we observed faster reactions and higher yields with *L*-amino acids and homo-*L*-peptides. The *L*-situation outcompeted reactions with *D*-amino acids present on the donor or on the acceptor strand, especially when the (*L*-) amino acid was directly attached to the nucleobase and therefore in close contact with the nucleic acid strand. Important is the observation that this stereochemical “preferential handshake” (*D*-ribose and *L*-amino acid) is also true with *D*-deoxyribose, which forms DNA. It shows us that even in a putative prebiotic world in which DNA-only or a mixed DNA–RNA combination were the carrier of genetic information, the stereochemical preference for *L* amino acids prevails.

Data availability

The data supporting this article have been included as part of the ESI.†

Author contributions

E. W. and I. M. were responsible for data curation, formal analysis, investigation and visualisation. T. C. had the research idea and was in charge of funding acquisition, project administration and supervision. All authors were involved in the conceptualization and the writing of the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft for supporting this research through the DFG grants: CA275/11-3 (ID: 326039064), CRC1309 (ID: 325871075, A4), CRC1032 (ID: 201269156, A5) and CRC1361 (ID: 393547839, P2). We thank the Volkswagen Foundation for funding this research (grant EvoRib). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program under grant agreement No. 741912 (EPIR) and the Marie Skłodowska-Curie grant agreement No. 861381 (Nature-ETN). The research was supported by the BMBF in the framework of the Cluster4Future program (Cluster for Nucleic Acid Therapeutics Munich, CNATM) (Project ID: 03ZU1201AA).

References

- 1 F. Crick, *Nature*, 1970, **227**, 561–563.
- 2 V. Ramakrishnan, *Cell*, 2002, **108**, 557–572.
- 3 T. A. Steitz, *Nat. Rev. Mol. Cell Biol.*, 2008, **9**, 242–253.
- 4 A. Yonath, M. A. Saper, I. Makowski, J. Müssig, J. Piefke, H. D. Bartunik, K. S. Bartels and H. G. Wittmann, *J. Mol. Biol.*, 1986, **187**, 633–636.
- 5 F. H. C. Crick, *Symp. Soc. Exp. Biol.*, 1958, **12**, 138–163.
- 6 F. H. C. Crick, S. Brenner, A. Klug and G. Pieczenik, *Origins Life*, 1976, **7**, 389–397.
- 7 R. Giegé and G. Eriani, in *eLS*, John Wiley & Sons, Ltd, Chichester, 2014, pp. 1007–1030.
- 8 J. E. Hein and D. G. Blackmond, *Acc. Chem. Res.*, 2012, **45**, 2045–2054.
- 9 R. Breslow and M. S. Levine, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 12979–12980.
- 10 R. Breslow and Z.-L. Cheng, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 9144–9146.
- 11 K. Tamura and P. Schimmel, *Science*, 2004, **305**, 1253.
- 12 E. Węgrzyn, I. Mejdrová, F. Müller, M. Nainyè, L. Escobar and T. Carell, *Angew. Chem., Int. Ed.*, 2024, **63**, e202319235.
- 13 P. F. Agris, *Nucleic Acids Res.*, 2004, **32**, 223–238.
- 14 X. Y. Guo and M. Su, *Int. J. Mol. Sci.*, 2023, **24**, 197.
- 15 Y. I. Wolf and E. V. Koonin, *Biol. Direct*, 2007, **2**, 14.
- 16 M. Seki, *Genes Genet. Syst.*, 2023, **98**, 9–24.
- 17 W. Gilbert, *Nature*, 1986, **319**, 618.
- 18 G. F. Joyce, *Nature*, 2002, **418**, 214–221.



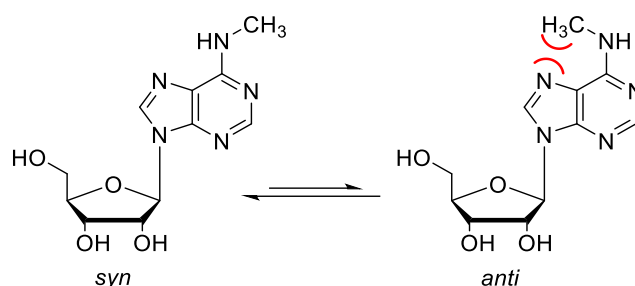
- 19 J. C. Bowman, N. V. Hud and L. D. Williams, *J. Mol. Evol.*, 2015, **80**, 143–161.
- 20 B. Jash and C. Richert, *Chem. Sci.*, 2020, **11**, 3487–3494.
- 21 B. Jash, P. Tremmel, D. Jovanovic and C. Richert, *Nat. Chem.*, 2021, **13**, 751–757.
- 22 N. Lee, Y. Bessho, K. Wei, J. W. Szostak and H. Suga, *Nat. Struct. Biol.*, 2000, **7**, 28–33.
- 23 I. Agmon, *Int. J. Mol. Sci.*, 2022, **23**, 15756.
- 24 P. Nissen, J. Hansen, N. Ban, P. B. Moore and T. A. Steitz, *Science*, 2000, **289**, 920–930.
- 25 I. Agmon, *Int. J. Mol. Sci.*, 2009, **10**, 2921–2934.
- 26 T. Bose, G. Fridkin, C. Davidovich, M. Krupkin, N. Dinger, A. H. Falkovich, Y. Peleg, I. Agmon, A. Bashan and A. Yonath, *Nucleic Acids Res.*, 2022, **50**, 1815–1828.
- 27 S. Tagami and P. Li, *Dev., Growth Differ.*, 2023, **65**, 167–174.
- 28 D. Kunnev, *Life*, 2020, **10**, 269.
- 29 H. B. White III, *J. Mol. Evol.*, 1976, **7**, 101–104.
- 30 A. D. Goldman and B. Kacar, *J. Mol. Evol.*, 2021, **89**, 127–133.
- 31 J. Xu, V. Chmela, N. J. Green, D. A. Russell, M. J. Janicki, R. W. Góra, R. Szabla, A. D. Bond and J. D. Sutherland, *Nature*, 2020, **582**, 60–66.
- 32 T. Lindahl, *Nature*, 1993, **362**, 709–715.
- 33 J. S. Teichert, F. M. Kruse and O. Trapp, *Angew. Chem., Int. Ed.*, 2019, **58**, 9944–9947.
- 34 J. Oró, S. L. Miller and A. Lazcano, *Annu. Rev. Earth Planet. Sci.*, 1990, **18**, 317–356.
- 35 H. S. Bernhardt, *Biol. Direct*, 2012, **7**, 23.
- 36 J. C. Cochrane and S. A. Strobel, *RNA*, 2008, **14**, 993–1002.
- 37 Y. Kim, K. Opron and Z. F. Burton, *Life*, 2019, **9**, 37.
- 38 A. S. Burton, J. C. Stern, J. E. Elsilá, D. P. Glavin and J. P. Dworkin, *Chem. Soc. Rev.*, 2012, **41**, 5459–5472.
- 39 P. T. S. v. d. Gulik, *Life*, 2015, **5**, 1629–1637.
- 40 C. W. Carter Jr, *Nat. Hist.*, 2016, **125**, 28–33.
- 41 M. P. Robertson and S. L. Miller, *Science*, 1995, **268**, 702–705.
- 42 C. Schneider, S. Becker, H. Okamura, A. Crisp, T. Amatov, M. Stadlmeier and T. Carell, *Angew. Chem., Int. Ed.*, 2018, **57**, 5943–5946.
- 43 S. Becker, C. Schneider, H. Okamura, A. Crisp, T. Amatov, M. Dejmek and T. Carell, *Nat. Commun.*, 2018, **9**, 163.
- 44 I. Agmon, *Life*, 2024, **14**, 277.
- 45 T. Carell, C. Brandmayr, A. Hienzsch, M. Müller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs and M. Wagner, *Angew. Chem., Int. Ed.*, 2012, **51**, 7110–7131.
- 46 J.-B. Zhou, E.-D. Wang and X.-L. Zhou, *Cell. Mol. Life Sci.*, 2021, **78**, 7087–7105.
- 47 T. Suzuki, Y. Yashiro, I. Kikuchi, Y. Ishigami, H. Saito, I. Matsuzawa, S. Okada, M. Mito, S. Iwasaki, D. Ma, X. Zhao, K. Asano, H. Lin, Y. Kirino, Y. Sakaguchi and T. Suzuki, *Nat. Commun.*, 2020, **11**, 4269.
- 48 F. Müller, L. Escobar, F. Xu, E. Węgrzyn, M. Nainytė, T. Amatov, C. Y. Chan, A. Pichler and T. Carell, *Nature*, 2022, **605**, 279–284.
- 49 K. Kitada, Y. Suda and N. Takenaka, *J. Phys. Chem. A*, 2017, **121**, 5383–5388.
- 50 C. Menor-Salván and M. R. Marín-Yaseli, *Chem. Soc. Rev.*, 2012, **41**, 5404–5415.
- 51 J. N. Singer, F. M. Müller, E. Węgrzyn, C. Hölzl, H. Hurmiz, C. Liu, L. Escobar and T. Carell, *Angew. Chem., Int. Ed.*, 2023, **62**, e202302360.
- 52 J. C. Lacey, R. D. Thomas, M. P. Staves and C. L. Watkins, *Biochim. Biophys. Acta*, 1991, **1076**, 395–400.
- 53 K. Tamura, *BioSystems*, 2008, **92**, 91–98.
- 54 L. F. Wu, M. Su, Z. Liu, S. J. Bjork and J. D. Sutherland, *J. Am. Chem. Soc.*, 2021, **143**, 11836–11842.
- 55 S. J. Roberts, Z. Liu and J. D. Sutherland, *J. Am. Chem. Soc.*, 2022, **144**, 4254–4259.
- 56 M. Su, C. Schmitt, Z. Liu, S. J. Roberts, K. C. Liu, K. Röder, A. Jäschke, D. J. Wales and J. D. Sutherland, *J. Am. Chem. Soc.*, 2023, **145**, 15971–15980.
- 57 O. Doppleb, J. Bremer, M. Bechthold, C. Sánchez Rico, D. Göhringer, H. Griesser and C. Richert, *Chem.–Eur. J.*, 2021, **27**, 13544–13551.
- 58 D. Ross and D. Deamer, *Life*, 2023, **13**, 1749.
- 59 S. Becker, C. Schneider, A. Crisp and T. Carell, *Nat. Commun.*, 2018, **9**, 5174.
- 60 B. Charnay, G. Le Hir, F. Fluteau, F. Forget and D. C. Catling, *Earth Planet. Sci. Lett.*, 2017, **474**, 97–109.
- 61 K. Le Vay, E. Salibi, E. Y. Song and H. Mutschler, *Chem.–Asian J.*, 2020, **15**, 214–230.
- 62 M. J. de Wit and H. Furnes, *Sci. Adv.*, 2016, **2**, e1500368.
- 63 W. Martin, J. Baross, D. Kelley and M. J. Russell, *Nat. Rev. Microbiol.*, 2008, **6**, 805–814.
- 64 L. M. Longo and M. Blaber, *Arch. Biochem. Biophys.*, 2012, **526**, 16–21.
- 65 L. M. Longo, J. Lee and M. Blaber, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 2135–2139.
- 66 J. R. Cronin and C. B. Moore, *Science*, 1971, **172**, 1327–1329.
- 67 A. Shimoyama, C. Ponnampereuma and K. Yanai, *Nature*, 1979, **282**, 394–396.
- 68 S. L. Miller, *Science*, 1953, **117**, 528–529.
- 69 Y. Wolman, W. J. Haverland and S. L. Miller, *Proc. Natl. Acad. Sci. U. S. A.*, 1972, **69**, 809–811.
- 70 A. P. Johnson, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, A. Lazcano and J. L. Bada, *Science*, 2008, **322**, 404.
- 71 E. T. Parker, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, M. Callahan, A. Aubrey, A. Lazcano and J. L. Bada, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 5526–5531.
- 72 L. L. Cummins, S. R. Owens, L. M. Risen, E. A. Lesnik, S. M. Freier, D. McGee, C. J. Guinosso and P. D. Cook, *Nucleic Acids Res.*, 1995, **23**, 2019–2024.
- 73 W. A. Decatur and M. J. Fournier, *Trends Biochem. Sci.*, 2002, **27**, 344–351.



4. Unpublished results

4.1 Rate constants of cleavage kinetics

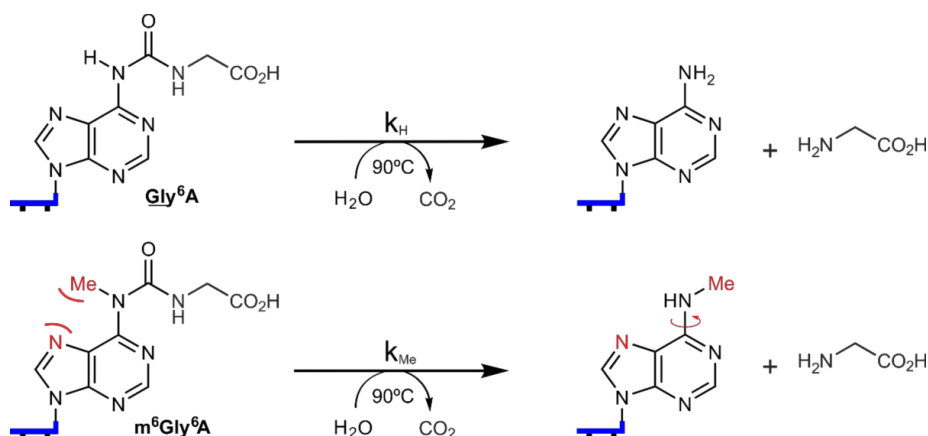
Previously, we reported that the urea cleavage step in the RNA-peptide synthesis cycle proceeds at a faster rate for donor strands containing m^6aa^6A , compared to aa^6A RNA strands.^[214] The rationale behind it, is that in the presence of the N^6 -methyl group on the adenosine, the *syn*-conformation is energetically favoured over the *anti*-conformation, due to the steric clash between the N^7 position and the methyl group (**Scheme 7**).^[47]



Scheme 7: The *syn*- and *anti*-conformation of m^6A nucleoside.

In the m^6aa^6A nucleoside, the methyl group would be forced into the *anti*-conformation, to minimize the steric interactions caused by the bulky amino acid substituent. This strain and the urge to twist back to *syn* is a driving force for the cleavage reaction and influences the reaction rate.^[47]

To quantify the difference in the rate constants of the cleavage step, kinetic experiments of this reaction with oligonucleotide strands containing g^6A (ON_H) and m^6g^6A (ON_{Me}) were conducted (**Scheme 8**). 2'OMe nucleosides were used to reduce the RNA decomposition caused by prolonged exposure to high temperature in the acidic milieu.



Scheme 8: The cleavage reaction of the RNA oligonucleotides ON_H and ON_{Me} at $90^\circ C$.

The oligonucleotide ON was mixed with MES buffer (pH 6) or acetate buffer (pH 4), NaCl and H₂O and the reaction mixture was heated to 90°C for different periods of time (0 – 30 h). The concentration of the components was: [ON] = 10 μM, [buffer] = 100 mM, [NaCl] = 100 mM. The progress of the reaction was measured by injecting the reaction mixture into the HPLC at different time points. The concentration of the ON could be determined by integration of the oligonucleotide peak.

The experimental cleavage data for concentration of ON (μM) vs. time (h) were plotted and the data fitted to the corresponding theoretical kinetic model using the Parameter Estimation Module of COPASI software Version 4.29 (**Figure 8**).^[215]

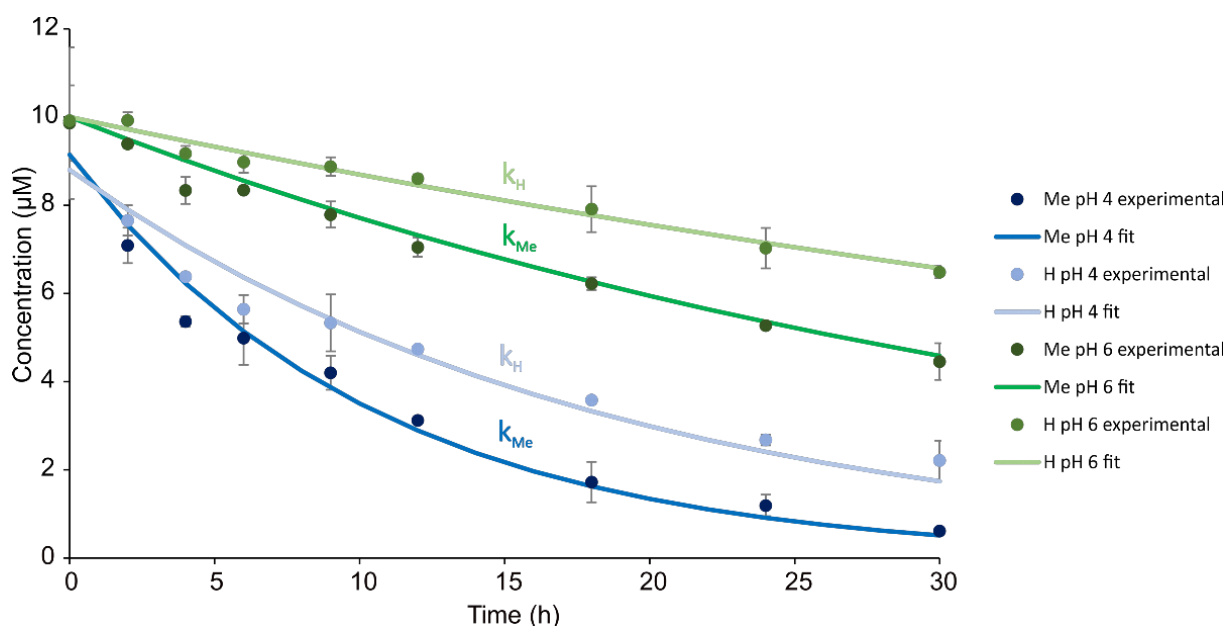


Figure 8: Plots of concentration of ON vs. time. Cleavage rate constant of ON_H at pH 6 (k_H) and pH 4 (k_H) and of ON_{Me} at pH 6 (k_H) and pH 4 (k_H).

The rate constants, which depend only on the change in concentration of the starting material ON, were calculated using the equation for a first-order reaction:

$$rate = - \frac{d [ON]}{dt} = k [ON]$$

The equation can be rearranged to:

$$\ln[ON] = -k t + \ln[ON]_0$$

So, if we plot the ln of the oligonucleotide concentration ([ON]) vs. time (h), we can calculate the gradient m of the linear function which will correspond to the rate constant of the reaction (**Table 1**).

$$y = m x + c$$

From the rate constant, we can then calculate the Gibbs' free energy of the reaction:

$$\Delta G^\circ = -RT \ln k$$

$$k = e^{-\frac{\Delta G^\circ}{RT}}$$

It is more accurate to use the Eyring equation, due to the temperature dependence.

$$k = \frac{k_B T}{h} e^{-\frac{\Delta G^\ddagger}{RT}}$$

We can calculate the difference in Gibbs' free energy ΔG for cleavage of ON_{Me} vs. ON_H (**Table 1**).

Table 1: The rate constants of ON_H or ON_{Me} at pH 4 or pH 6, and the difference in the Gibbs' free energy between these reactions.

	k_H (s ⁻¹)	k_{Me} (s ⁻¹)	$\Delta G_{Me} - \Delta G_H$ (kcal/mol)
pH 4	1.50×10^{-5}	2.67×10^{-5}	0.42
pH 6	3.88×10^{-6}	7.22×10^{-6}	0.45

We can conclude that the presence of the N⁶ methyl group and the reduction of the pH of the reaction mixture, significantly increase the rate constant of the cleavage reaction.

The cleavage reaction of a hairpin strand with more than one amino acid (dipeptide or longer), is accompanied by the formation of a cyclic hydantoin by-product.^[204] This side product was forming increasingly at higher pH and higher temperature.^[214] This means that the cleavage reaction at pH 4 is not only faster than at higher pH, but also reduces the extent of the side reaction. Due to the presence of 2'OMe modification, the RNA degradation is minimal even at low pH.

5. Experimental part

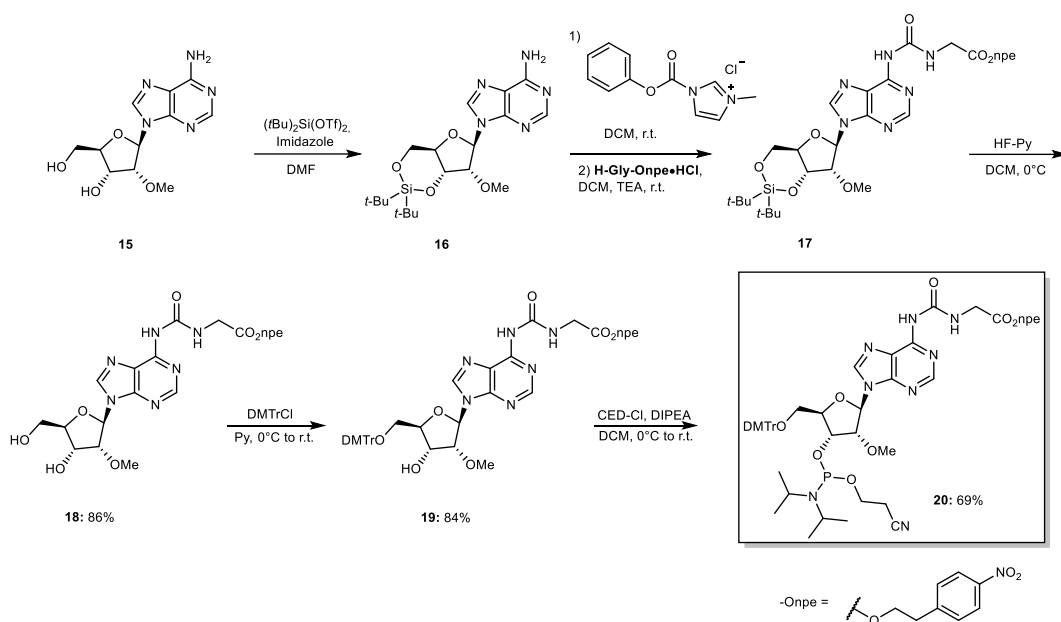
5.1 General information and instruments for nucleosides and phosphoramidites

Reagents were purchased from commercial suppliers and used without further purification. Anhydrous solvents, stored under inert atmosphere, were also purchased. All reactions involving air/moisture sensitive reagents/intermediates were performed under inert atmosphere using oven-dried glassware. Routine ^1H NMR, ^{13}C NMR and ^{31}P NMR spectra were recorded on a Bruker Ascend 400 spectrometer (400 MHz for ^1H NMR, 100 MHz for ^{13}C NMR and 162 MHz for ^{31}P NMR), Bruker Ascend 500 spectrometer (500 MHz for ^1H NMR, 125 MHz for ^{13}C NMR and 202 MHz for ^{31}P NMR) or Bruker ARX 600 spectrometer (600 MHz for ^1H NMR, 150 MHz for ^{13}C NMR and 243 MHz for ^{31}P NMR). Deuterated solvents used are indicated in the characterization and chemical shifts (δ) are reported in ppm. Residual solvent peaks were used as reference.^[216] All NMR J values are given in Hz. COSY, HMQC and HMBC NMR experiments were recorded to help with the assignment of ^1H and ^{13}C signals. NMR spectra were analyzed using MestReNova software version 10.0. High Resolution Mass Spectra (HRMS) were measured on a Thermo Finnigan LTQ-FT with ESI as ionization mode. IR spectra were recorded on a Perkin-Elmer Spectrum BX II FT-IR instrument or Shimadzu IRSpirit FT-IR instrument. Both equipped with an ATR accessory. Column chromatography was performed with technical grade silica gel, 40-63 μm particle size. Reaction progress was monitored by Thin Layer Chromatography (TLC) analysis on silica gel 60 F254 and stained with *para*-anisaldehyde, potassium permanganate or ninhydrin solution.

5.2 Synthesis of phosphoramidites

Most of the compounds synthesized during my PhD have been published.^{[214],[217],[218]} For unpublished synthesis, see below.

Nucleobase-modified 2'-methoxy *N*⁶-carbamoyl adenosine phosphoramidite



Scheme 9: Synthesis of nucleobase-modified 2'OMe *N*⁶-methylcarbamoyl adenosine phosphoramidite

Compounds **15**, **16**, **17** have been synthesized according to literature.^[214]

General procedure for the synthesis of 18: A solution of compound **17** (1.0 equiv.) in CH₂Cl₂/pyridine (9:1 v/v) inside a plastic reaction vessel was cooled to 0°C. Subsequently, a solution of 70% HF-pyridine (5.0 equiv.) was slowly added and the reaction mixture was stirred at 0°C for 2 h. The reaction mixture was diluted with aq. sat. NaHCO₃ solution and extracted with CH₂Cl₂. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated under reduced pressure. The crude product was purified by silica gel column chromatography to isolate the product **18** as a white foam.

18: Yield: 86%; *R*_f = 0.26 (95:5 CH₂Cl₂/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3227 (w); 2936 (w); 1744 (m); 1694 (m); 1612 (m); 1588 (m); 1514 (s); 1468 (m); 1343 (s); 1255 (m); 1186 (s); 1088 (m); 1054 (m); 856 (m). ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 9.85 (s, 1H); 8.75 (s, 1H); 8.53 (s, 1H); 8.28 (s, 1H); 8.16 – 8.06 (m, 2H); 7.43 – 7.35 (m, 2H); 6.00 (d, *J* = 6.9 Hz, 1H); 4.66 (dd, *J* = 6.9 Hz, *J* = 4.7 Hz, 1H); 4.61 (dd, *J* = 4.6 Hz, *J* = 1.3 Hz, 1H); 4.45 (t, *J* = 6.6 Hz, 2H); 4.37 (q, *J* = 1.6 Hz, 1H); 4.28 – 4.11 (m, 2H); 3.99 (dd, *J* = 13.0 Hz, *J* = 1.8 Hz, 1H); 3.80 (dd, *J* = 13.0 Hz, *J* = 1.7 Hz, 1H); 3.38 (s, 3H); 3.09 (t, *J* = 6.6 Hz, 2H). ¹³C NMR (101 MHz, CDCl₃, 298 K) δ (ppm): 169.7; 153.9; 149.3; 146.9; 145.3; 143.5; 129.8; 123.8; 89.4; 88.0; 82.9; 77.2; 70.3; 64.7; 63.1; 58.9; 42.1; 34.8. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₂H₂₆N₇O₉ 532.1786; Found 532.1786.

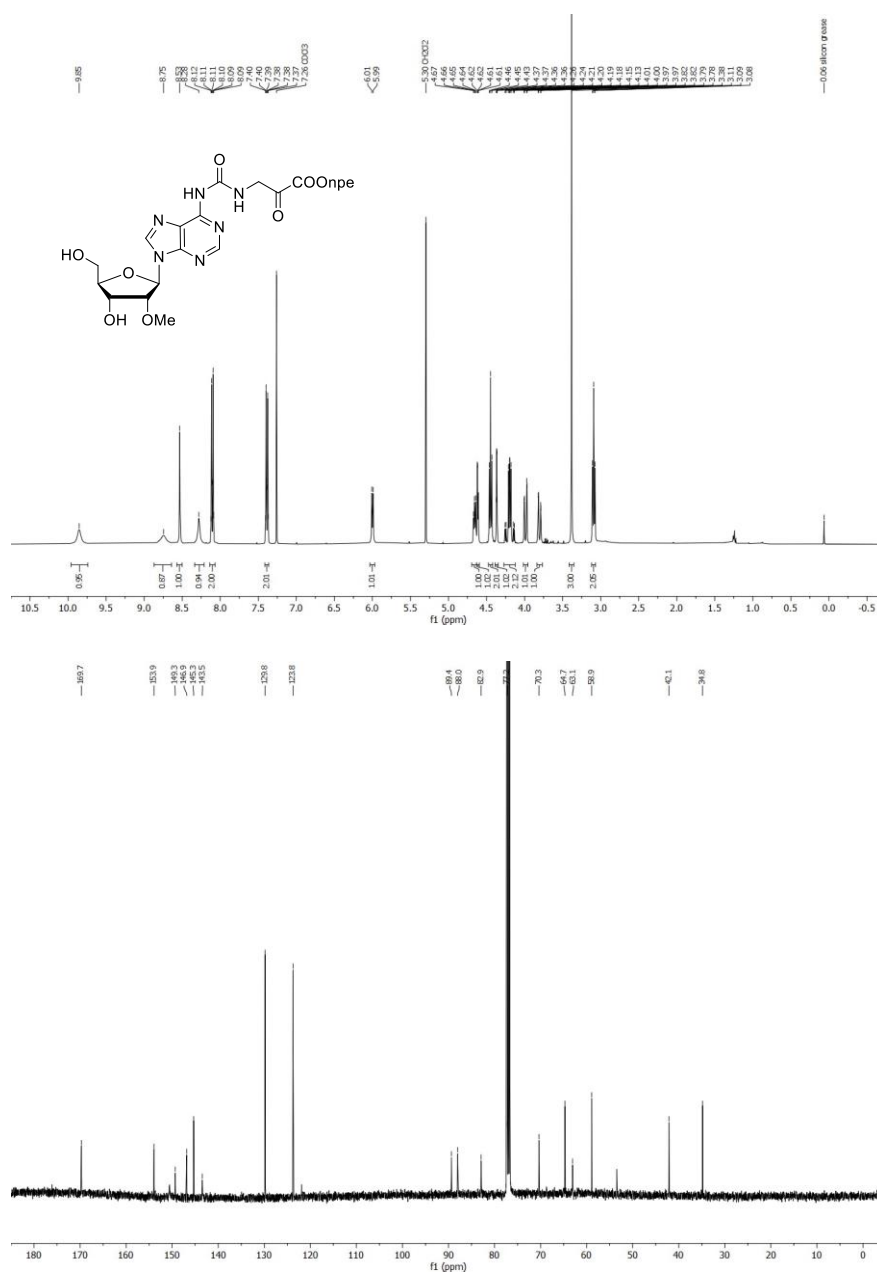
General procedure for the synthesis of 19: Compound **18** (1 equiv.) was dissolved in dry pyridine and stirred under nitrogen atmosphere at r.t. 4,4-Dimethoxytrityl chloride (1.5 equiv.) was added in two portions and the reaction was stirred at r.t. overnight. After that, the crude was concentrated and purified by silica gel column chromatography (eluent containing 0.1% pyridine) affording the product **19** as a white foam.

19: Yield: 84%; R_f = 0.40 (96:4 CH₂Cl₂/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3234 (w); 2934 (w); 1747 (s); 1698 (m); 1607 (m); 1586 (m); 1508 (s); 1466 (m); 1343 (s); 1247 (s); 1175 (s); 1030 (s); 826 (m); 701 (s). ¹H NMR (400 MHz, acetone-d₆, 298 K) δ (ppm): 9.89 (t, J = 5.8 Hz, 1H); 8.80 (s, 1H); 8.50 (s, 1H); 8.45 (s, 1H); 8.14 – 8.06 (m, 2H); 7.63 – 7.55 (m, 2H); 7.50 – 7.43 (m, 2H); 7.36 – 7.30 (m, 3H); 7.31 – 7.22 (m, 2H); 7.24 – 7.15 (m, 1H); 6.88 – 6.78 (m, 4H); 6.24 (d, J = 3.9 Hz, 1H); 4.70 (dt, J = 6.4, 5.2 Hz, 1H); 4.63 (dd, J = 5.0 Hz, J = 4.0 Hz, 1H); 4.45 (t, J = 6.4 Hz, 2H); 4.28 – 4.20 (m, 2H); 4.14 (d, J = 5.8 Hz, 2H); 3.76 (d, J = 2.2 Hz, 6H); 3.53 (s, 3H); 3.52 – 3.39 (m, 2H); 3.14 (t, J = 6.4 Hz, 2H). ¹³C NMR (101 MHz, acetone-d₆, 298 K) δ (ppm): 170.6; 159.6; 159.6; 154.5; 151.7; 151.2; 151.2; 150.7; 147.6; 147.4; 146.1; 143.3; 136.7; 136.7; 136.6; 131.1; 131.0; 130.9; 129.0; 128.6; 127.5; 124.6; 124.2; 121.7; 113.8; 87.8; 87.0; 85.0; 83.8; 70.7; 65.2; 64.4; 58.8; 55.5; 42.6; 35.3. HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₄₃H₄₄N₇O₁₁ 834.3093; Found 834.3079.

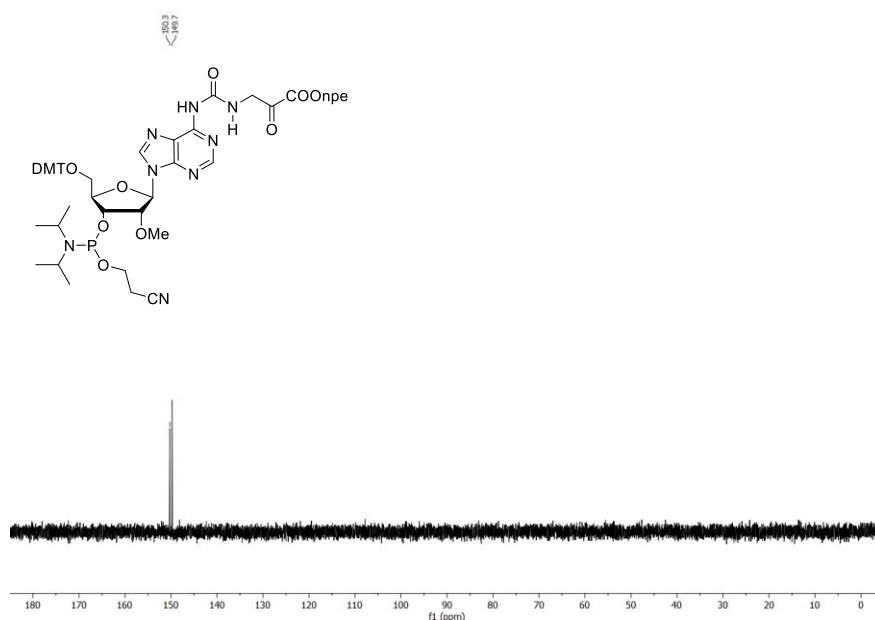
General procedure for the synthesis of compound 20: Compound **19** (1 equiv.) was added to a dry-oven round-bottom flask and dissolved in dry CH₂Cl₂. The solution was stirred under Argon atmosphere at 0°C. DIPEA (4 equiv.) was added dropwise. Finally, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (2.5 equiv.) was added dropwise. The reaction was stirred at r.t. for 3 h. After that, the reaction was stopped and diluted with CH₂Cl₂. The crude was washed with aqueous saturated NaHCO₃ and the organic layer was separated. The crude was further extracted with CH₂Cl₂. The combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography (eluent containing 0.1% pyridine). The products were isolated as a mixture of diastereoisomers as a white foam. Finally, the product **20** was lyophilized from benzene.

20: Yield = 69%. R_f = 0.63 (1:4 iHex/EtOAc). ³¹P NMR (202 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 150.3; 149.7. HRMS (ESI) m/z : [M+H]⁺ Calcd for C₅₂H₆₁O₁₂N₉P 1034.4171; Found 1034.4154.

NMR of compound 18



NMR of compound **20**



5.3 General information and instruments for oligonucleotides

Synthesis and purification of oligonucleotides

Phosphoramidites of 2'-O-Me ribonucleosides (2'-OMe-Bz-A-CE, 2'-OMe-Dmf-G-CE, 2'-OMe-Ac-C-CE and 2'-OMe-U-CE) were purchased from LinkTech and Sigma-Aldrich. Oligonucleotides (ONs) were synthesized on a 1 μ mol scale using RNA SynBase™ CPG 1000/110 and High Load Glen UnySupport™ as solid supports using an RNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with a standard phosphoramidite chemistry. ONs were synthesized in DMT-OFF mode using DCA as a deblocking agent in CH₂Cl₂, BTT or Activator 42® as activator in MeCN, Ac₂O as capping reagent in pyridine/THF and I₂ as oxidizer in pyridine/H₂O.

Deprotection of Onpe and teoc groups

For the deprotection of the *para*-nitrophenylethyl (Onpe) group in ONs containing amino acid-modified carbamoyl adenosine nucleosides, the solid support beads were suspended in a 9:1 THF/DBU solution mixture (1 mL) and incubated at r.t. for 2 h.^[219] After that, the supernatant was removed, and the beads were washed with THF (3x1 mL).

For the deprotection of the 2-(trimethylsilyl)ethoxycarbonyl (teoc) group in ONs containing 5-methylaminomethyl uridine nucleosides, the solid support beads were suspended in a saturated solution of ZnBr₂ in 1:1 MeNO₂/IPA (1 mL) and incubated at r.t. overnight.^[220] After that, the supernatant was removed, and the beads were washed with 0.1 M EDTA in water (1 mL) and water (1 mL).

Coupling of amino acids to ONs anchored to the solid support beads

The solid support beads (1 μmol) in an Eppendorf tube were washed with dry DMF (0.3 mL). In a separate Eppendorf tube, Boc-protected L- or D-Valine, DMTMM•BF₄ (100 μmol) as activator and dry DIPEA (200 μmol) were dissolved in dry DMF (0.6 mL). Subsequently, the amino acid solution was added to the solid support beads and the reaction was incubated in an orbital shaker at r.t. for 1 h. The suspension was centrifuged and the supernatant was removed. The solid support beads were washed with dry DMF (2×0.3 mL) and dry MeCN (2×0.3 mL). Finally, the beads were dried using a SpeedVac concentrator.

For the deprotection of the *tert*-butyloxycarbonyl (Boc) group in ONs after the coupling of a Boc-protected amino acid or peptide, the solid support beads were suspended in a 1:1 TFA/CH₂Cl₂ solution mixture (0.5 mL) and incubated for 5 min at r.t.^[221] After that, the supernatant was removed and the solid support beads were washed with CH₂Cl₂ (2×0.5 mL).

Cleavage from beads and precipitation of the synthesized ON

The solid support beads were suspended in a 1:1 aqueous solution mixture (0.6 mL) of 30% NH₄OH and 40% MeNH₂. The suspension was heated at 65°C (8 min for SynBase™ CPG 1000/110 and 60 min for High Load Glen UnySupport™). Subsequently, the supernatant was collected, and the beads were washed with water (2×0.3 mL). The combined aqueous solutions were concentrated under reduced pressure using a SpeedVac concentrator. After that, the crude was dissolved in DMSO (100 μL) and the ON was precipitated by adding 3 M NaOAc in water (25 μL) and *n*-butanol (1 mL). The mixture was kept at -80°C for 2 h and centrifuged at 4°C for 1 h. The supernatant was removed, and the white precipitate was lyophilized.

Purification of the synthesized ON by HPLC and desalting

The crude was purified by semi-preparative HPLC (1260 Infinity II Manual Preparative LC System from Agilent equipped with a G7114A detector) using a reverse-phase (RP) VP 250/10 Nucleodur 100-5 C18ec column from Macherey-Nagel (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O

and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-25% of B in 45 min; Flow rate = 5 mL·min⁻¹). The purified ON was analyzed by RP-HPLC (1260 Infinity II LC System from Agilent equipped with a G7165A detector) using an EC 250/4 Nucleodur 100-3 C18ec from Macherey-Nagel (Gradient: 0-30% of B in 45 min; Flow rate = 1 mL·min⁻¹). Finally, the purified ON was desalted using a C18 RP-cartridge from Waters.

Determination of the concentration and the mass of the synthesized ON

The absorbance of the synthesized ON in H₂O solution was measured using an IMPLEN NanoPhotometer® N60/N50 at 260 nm. The extinction coefficient of the ON was calculated using the OligoAnalyzer Version 3.0 from Integrated DNA Technologies. For ONs incorporating non-canonical bases, the extinction coefficients were assumed to be identical to those containing only canonical counterparts.

The synthesized ON (2-3 µL) was desalted on a 0.025 µm VSWP filter (Millipore), co-crystallized in a 3-hydroxypicolinic acid matrix (HPA, 1 µL) and analyzed by matrix-assisted laser desorption/ionization – time-of-flight (MALDI-TOF) mass spectrometry (negative mode).

5.4 Synthesis of oligonucleotides

The synthesis of most oligonucleotides from my PhD has been published.^{[214],[218]}

The oligonucleotides described in Section 4 are characterized below.

Table 2: HPLC retention times (0-40% of B in 45 min) and MALDI-ToF mass spectrometric analysis (negative mode) of oligonucleotide strands **ON_H** and **ON_{Me}** (described in Section 4) containing g⁶A or m⁶g⁶A at the 5'-end. The subscript _m stands for 2'OMe nucleotides.

Sequence	Donor strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
5'-X(AUCGCU) _m -3'	ON_H ; X = g ⁶ A _m	23.1	2361.4	2360.6
	ON_{Me} ; X = m ⁶ g ⁶ A _m	24.0	2375.4	2374.8

6. Abbreviations

A	adenine (nucleobase), adenosine (nucleoside)
aaRS	aminoacyl-tRNA synthetase
Ac	acetyl
AMP	adenosine 5'-monophosphate
ATP	adenosine 5'-triphosphate
Boc	<i>tert</i> -butyloxycarbonyl
C	cytosine (nucleobase), cytidine (nucleoside)
CDI	carbonyldiimidazole
COSY	correlation spectroscopy
D	stereodescriptor for the D-enantiomer (lat. dexter)
DBU	1,8-diazabicyclo[5.4.0]undec-7-en
DCM	dichloromethane
de	diastereomeric excess
DIPEA	diisopropylethylamine
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ee	enantiomeric excess
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	ethylenediaminetetraacetic acid
equiv.	equivalents
ESI	electron spray ionization
G	guanine (nucleobase), guanosine (nucleoside)
h	hour(s)
HPA	3-hydroxypicolinic acid
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
IDA	initial Darwinian ancestor
IR	infrared spectroscopy
L	stereodescriptor for the L-enantiomer (lat. laevus)
LUCA	last universal common ancestor
MALDI-TOF	matrix assisted laser desorption-ionisation – time-of-flight
MeCN	acetonitrile
MES	2-(<i>N</i> -morpholino)-ethane sulfonate

MIC	methyl isocyanate
mRNA	messenger ribonucleic acid
MS	mass spectrometry
NMR	nuclear magnetic resonance
npe	<i>p</i> -nitrophenyl ethyl
ON	oligonucleotide
Phe	phenylalanine
PIPES	1,4-piperazinediethanesulfonic acid
ppm	parts per million
PTC	peptidyl transferase centre
R _f	retardation factor
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
T	thymine (nucleobase), thymidine (nucleoside)
TBS	<i>tert</i> -butylmethylsilyl
TEA	triethylamine
teoc	(trimethylsilyl)ethoxycarbonyl
THF	tetrahydrofuran
TLC	thin-layer chromatography
tRNA	transfer ribonucleic acid
U	uracil (nucleobase), uridin (nucleoside)

7. References

- [1] J. D. Watson, F. H. Crick, *Nature* **1953**, *171*, 737-738.
- [2] R. E. Franklin, R. G. Gosling, *Nature* **1953**, *171*, 740-741.
- [3] A. Cerreta, *PhD thesis* **2013**.
- [4] M. Huang, T. J. Giese, T. S. Lee, D. M. York, *J. Chem. Theory Comput.* **2014**, *10*, 1538-1545.
- [5] R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. V. Fratini, M. L. Kopka, *Science* **1982**, *216*, 475-485.
- [6] A. M. Chauca-Diaz, Y. J. Choi, M. J. Resendiz, *Biopolymers* **2015**, *103*, 167-174.
- [7] L. M. Longo, J. Lee, M. Blaber, *Proc. Natl. Acad. Sci. U.S.A* **2013**, *110*, 2135-2139.
- [8] L. M. Longo, M. Blaber, *Arch. Biochem. Biophys.* **2012**, *526*, 16-21.
- [9] C. S. Foden, S. Islam, C. Fernández-García, L. Maugeri, T. D. Sheppard, M. W. Powner, *Science* **2020**, *370*, 865-869.
- [10] H. F. Noller, *Science* **2005**, *309*, 1508-1514.
- [11] F. H. C. Crick, *Symp. Soc. Exp. Biol.* **1958**, *12*, 138-163.
- [12] F. Crick, *Nature* **1970**, *227*, 561-563.
- [13] V. Ramakrishnan, *Cell* **2002**, *108*, 557-572.
- [14] A. L. Dounce, G. Gamow, S. Spiegelman, P. Newmark, D. Harker, M. Soodak, *Round-Table Discussion* **1956**, *47*, 103-112.
- [15] J.-B. Zhou, E.-D. Wang, X.-L. Zhou, *Cell. Mol. Life Sci.* **2021**, *78*, 7087-7105.
- [16] H. A. Nguyen, E. D. Hoffer, C. M. Dunham, *J. Biol. Chem.* **2019**, *294*, 5281-5291.
- [17] M. Ibbá, D. Soll, *Annu. Rev. Biochem.* **2000**, *69*, 617-650.
- [18] M. B. Hoagland, E. B. Keller, P. C. Zamecnik, *J. Biol. Chem.* **1956**, *218*, 345-358.
- [19] R. Giegé, G. Eriani, in *eLS. John Wiley & Sons, Ltd: Chichester*, **2014**, pp. 1007-1030.
- [20] T. A. Steitz, *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 242-253.
- [21] J. H. Cate, M. M. Yusupov, G. Z. Yusupova, T. N. Earnest, H. F. Noller, *Science* **1999**, *285*, 2095-2104.
- [22] M. M. Yusupov, G. Z. Yusupova, A. Baucom, K. Lieberman, T. N. Earnest, J. H. D. Cate, H. F. Noller, *Science* **2001**, *292*, 883-896.
- [23] A. Yonath, M. A. Saper, I. Makowski, J. Müssig, J. Piefke, H. D. Bartunik, K. S. Bartels, H. G. Wittmann, *J. Mol. Biol.* **1986**, *187*, 633-636.
- [24] P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 920-930.
- [25] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 905-920.
- [26] I. Agmon, *Int. J. Mol. Sci.* **2022**, *23*, 15756.
- [27] T. Martin Schmeing, K. S. Huang, S. A. Strobel, T. A. Steitz, *Nature* **2005**, *438*, 520-524.
- [28] J. A. Kowalak, J. J. Dalluge, J. A. McCloskey, K. O. Stetter, *Biochemistry* **1994**, *33*, 7869-7876.
- [29] P. F. Agris, *Nucleic Acids Res.* **2004**, *32*, 223-238.
- [30] T. Suzuki, Y. Yashiro, I. Kikuchi, Y. Ishigami, H. Saito, I. Matsuzawa, S. Okada, M. Mito, S. Iwasaki, D. Ma, X. Zhao, K. Asano, H. Lin, Y. Kirino, Y. Sakaguchi, T. Suzuki, *Nat. Commun.* **2020**, *11*, 4269.
- [31] M. P. Robertson, S. L. Miller, *Science* **1995**, *268*, 702-705.
- [32] F. V. Murphy, V. Ramakrishnan, A. Malkiewicz, P. F. Agris, *Nat. Struct. Mol. Biol.* **2004**, *11*, 1186-1191.
- [33] P. J. McCown, A. Ruszkowska, C. N. Kunkler, K. Breger, J. P. Hulewicz, M. C. Wang, N. A. Springer, J. A. Brown, *Wiley Interdiscip. Rev.* **2020**, *11*, e1595.
- [34] T. Muramatsu, S. Yokoyama, N. Horie, A. Matsuda, T. Ueda, Z. Yamaizumi, Y. Kuchino, S. Nishimura, T. Miyazawa, *J. Biol. Chem.* **1988**, *263*, 9261-9267.
- [35] Y. Ikeuchi, S. Kimura, T. Numata, D. Nakamura, T. Yokogawa, T. Ogata, T. Wada, T. Suzuki, T. Suzuki, *Nat. Chem. Biol.* **2010**, *6*, 277-282.
- [36] N. Akiyama, K. Ishiguro, T. Yokoyama, K. Miyauchi, A. Nagao, M. Shirouzu, T. Suzuki, *Nat. Struct. Mol. Biol.* **2024**, *31*, 817-825.
- [37] M. Nainyte, T. Carell, *Helv. Chim. Acta* **2020**, *103*, e2000016.
- [38] B.-i. Kang, K. Miyauchi, M. Matuszewski, G. S. D'Almeida, Mary Anne T. Rubio, J. D. Alfonso, K. Inoue, Y. Sakaguchi, T. Suzuki, E. Sochacka, T. Suzuki, *Nucleic Acids Res.* **2016**, *45*, 2124-2136.
- [39] S. P. Dutta, C. I. Hong, G. P. Murphy, A. Mittelman, G. B. Chheda, *Biochemistry* **1975**, *14*, 3144-3151.
- [40] A. Ohkubo, R. Kasuya, K. Miyata, H. Tsunoda, K. Seio, M. Sekine, *Org. Biomol. Chem.* **2009**, *7*, 687-694.

- [41] L. Escobar, *ChemSystemsChem* **2024**, *n/a*, e202400030.
- [42] G. Suresh, U. D. Priyakumar, *J. Mol. Graph.* **2015**, *61*, 150-159.
- [43] L. L. Cummins, S. R. Owens, L. M. Risen, E. A. Lesnik, S. M. Freier, D. McGee, C. J. Guinosso, P. D. Cook, *Nucleic Acids Res.* **1995**, *23*, 2019-2024.
- [44] W. A. Decatur, M. J. Fournier, *Trends Biochem. Sci.* **2002**, *27*, 344-351.
- [45] P. Boccaletto, M. A. Machnicka, E. Purta, P. Piatkowski, B. Baginski, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm, J. M. Bujnicki, *Nucleic Acids Res.* **2018**, *46*, D303-d307.
- [46] J. Y. Roignant, M. Soller, *Trends Genet.* **2017**, *33*, 380-390.
- [47] C. Roost, S. R. Lynch, P. J. Batista, K. Qu, H. Y. Chang, E. T. Kool, *J. Am. Chem. Soc.* **2015**, *137*, 2107-2115.
- [48] H. Grosjean, E. Westhof, *Nucleic Acids Res* **2016**, *44*, 8020-8040.
- [49] M. Di Giulio, *J. Theor. Biol.* **1998**, *191*, 191-196.
- [50] A. C. Rios, Y. Tor, *Isr. J. Chem.* **2013**, *53*, 469-483.
- [51] P. L. Luisi, *Origins Life Evol. Biospheres* **1998**, *28*, 613-622.
- [52] S. A. Benner, *Astrobiology* **2010**, *10*, 1021-1030.
- [53] J. W. Szostak, *J. Biomol. Struct. Dyn.* **2012**, *29*, 599-600.
- [54] G. F. Joyce, *Nature* **2002**, *418*, 214-221.
- [55] N. Arndt, E. Nisbet, *Annu. Rev. Earth Planet. Sci.* **2012**, *40*, 521-549.
- [56] L. E. Orgel, *J. Mol. Biol.* **1968**, *38*, 381-393.
- [57] M. C. Weiss, F. L. Sousa, N. Mrnjavac, S. Neukirchen, M. Roettger, S. Nelson-Sathi, W. F. Martin, *Nat. Microbiol.* **2016**, *1*, 16116.
- [58] W. F. Doolittle, *Science* **1999**, *284*, 2124-2128.
- [59] C. R. Woese, O. Kandler, M. L. Wheelis, *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 4576-4579.
- [60] C. J. Cox, P. G. Foster, R. P. Hirt, S. R. Harris, T. M. Embley, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20356-20361.
- [61] J. K. Harris, S. T. Kelley, G. B. Spiegelman, N. R. Pace, *Genome Res.* **2003**, *13*, 407-412.
- [62] E. V. Koonin, *Nat. Rev. Microbiol.* **2003**, *1*, 127-136.
- [63] Y. I. Wolf, E. V. Koonin, *Biol. Direct* **2007**, *2*, 14.
- [64] I. Agmon, *Life* **2024**, *14*, 277.
- [65] J. I. Glass, C. Merryman, K. S. Wise, C. A. Hutchison, 3rd, H. O. Smith, *Cold Spring Harbor Perspect. Biol.* **2017**, *9*.
- [66] A. Kirschning, *Chem. Eur. J.* **2022**, *28*, e202201419.
- [67] N. Glansdorff, Y. Xu, B. Labedan, *Biol. Direct* **2008**, *3*, 29.
- [68] M. Yarus, *Cold Spring Harbor Perspect. Biol.* **2011**, *3*.
- [69] B. M. A. G. Piette, J. G. Heddle, *Trends Ecol. Evol.* **2020**, *35*, 397-406.
- [70] J. W. Schopf, *Philos. Trans. R. Soc., B* **2006**, *361*, 869-885.
- [71] S. J. Mojzsis, G. Arrhenius, K. D. McKeegan, T. M. Harrison, A. P. Nutman, C. R. L. Friend, *Nature* **1996**, *384*, 55-59.
- [72] S. Moorbath, *Nature* **2005**, *434*, 155-155.
- [73] T. Djokic, M. J. Van Kranendonk, K. A. Campbell, M. R. Walter, C. R. Ward, *Nat. Commun.* **2017**, *8*, 15263.
- [74] S. Becker, C. Schneider, H. Okamura, A. Crisp, T. Amatov, M. Dejmek, T. Carell, *Nat. Commun.* **2018**, *9*, 163.
- [75] D. Ross, D. Deamer, *Life* **2023**, *13*, 1749.
- [76] B. H. Patel, C. Percivalle, D. J. Ritson, C. D. Duffy, J. D. Sutherland, *Nat. Chem.* **2015**, *7*, 301-307.
- [77] W. Martin, J. Baross, D. Kelley, M. J. Russell, *Nat. Rev. Microbiol.* **2008**, *6*, 805-814.
- [78] R. Sanchez, J. Ferris, L. E. Orgel, *Science* **1966**, *153*, 72-73.
- [79] C. Menor-Salván, M. R. Marín-Yaseli, *Chem. Soc. Rev.* **2012**, *41*, 5404-5415.
- [80] K. Le Vay, E. Salibi, E. Y. Song, H. Mutschler, *Chem. Asian J.* **2020**, *15*, 214-230.
- [81] S. J. Roberts, Z. Liu, J. D. Sutherland, *J. Am. Chem. Soc.* **2022**, *144*, 4254-4259.
- [82] B. Charnay, G. Le Hir, F. Fluteau, F. Forget, D. C. Catling, *Earth and Planetary Science Letters* **2017**, *474*, 97-109.
- [83] S. Becker, C. Schneider, A. Crisp, T. Carell, *Nat. Commun.* **2018**, *9*, 5174.
- [84] A. Eschenmoser, *Tetrahedron* **2007**, *63*, 12821-12844.
- [85] A. Lazcano, *J. Mol. Evol.* **2016**, *83*, 214-222.
- [86] S. Tirard, *J. Genet.* **2017**, *96*, 735-739.
- [87] S. L. Miller, *Science* **1953**, *117*, 528-529.
- [88] J. D. Sutherland, *Angew. Chem. Int. Ed.* **2016**, *55*, 104-121.

- [89] J. L. Bada, A. Lazcano, *Science* **2003**, *300*, 745-746.
- [90] A. Strecker, *Justus Liebigs Ann. Chem.* **1854**, *91*, 349-351.
- [91] L. Legnani, A. Darù, A. X. Jones, D. G. Blackmond, *J. Am. Chem. Soc.* **2021**, *143*, 7852-7858.
- [92] A. Butlerow, *Justus Liebigs Ann. Chem.* **1861**, *120*, 295-298.
- [93] R. Breslow, *Tetrahedron Lett.* **1959**, *1*, 22-26.
- [94] R. Larralde, M. P. Robertson, S. L. Miller, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8158-8160.
- [95] A. Ricardo, M. A. Carrigan, A. N. Olcott, S. A. Benner, *Science* **2004**, *303*, 196-196.
- [96] S. L. Miller, H. C. Urey, *Science* **1959**, *130*, 245-251.
- [97] J. Oró, *Nature* **1961**, *191*, 1193-1194.
- [98] M. Levy, S. L. Miller, J. Oró, *J. Mol. Evol.* **1999**, *49*, 165-168.
- [99] J. P. Ferris, P. C. Joshi, E. H. Edelson, J. G. Lawless, *J. Mol. Evol.* **1978**, *11*, 293-311.
- [100] M. P. Callahan, H. James Cleaves, II, *ACS Earth Space Chem.* **2023**, *7*, 2321-2326.
- [101] C. Schneider, S. Becker, H. Okamura, A. Crisp, T. Amatov, M. Stadlmeier, T. Carell, *Angew. Chem. Int. Ed.* **2018**, *57*, 5943-5946.
- [102] M. W. Powner, B. Gerland, J. D. Sutherland, *Nature* **2009**, *459*, 239-242.
- [103] M. W. Powner, J. D. Sutherland, J. W. Szostak, *J. Am. Chem. Soc.* **2010**, *132*, 16677-16688.
- [104] S. Becker, J. Feldmann, S. Wiedemann, H. Okamura, C. Schneider, K. Iwan, A. Crisp, M. Rossa, T. Amatov, T. Carell, *Science* **2019**, *366*, 76-82.
- [105] J. Oró, *Biochem. Biophys. Res. Commun.* **1960**, *2*, 407-412.
- [106] J. P. Ferris, L. E. Orgel, *J. Am. Chem. Soc.* **1966**, *88*, 1074-1074.
- [107] C. B. Mast, S. Schink, U. Gerland, D. Braun, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8030-8035.
- [108] C. A. Jerome, H.-J. Kim, S. J. Mojzsis, S. A. Benner, E. Biondi, *Astrobiology* **2022**, *22*, 629-636.
- [109] J. P. Ferris, A. R. Hill, R. Liu, L. E. Orgel, *Nature* **1996**, *381*, 59-61.
- [110] S. W. Fox, T. V. Waehneltd, *Biochim. Biophys. Acta, Protein Struct.* **1968**, *160*, 246-249.
- [111] K. Dose, *Origins Life* **1974**, *5*, 239-252.
- [112] G. F. Joyce, G. M. Visser, C. A. A. van Boeckel, J. H. van Boom, L. E. Orgel, J. van Westrenen, *Nature* **1984**, *310*, 602-604.
- [113] G. Leveau, D. Pfeffer, B. Altaner, E. Kervio, F. Welsch, U. Gerland, C. Richert, *Angew. Chem. Int. Ed.* **2022**, *61*, e202203067.
- [114] M. G. Weller, *Life* **2024**, *14*, 341.
- [115] B. N. Khare, C. Sagan, H. Ogino, B. Nagy, C. Er, K. H. Schram, E. T. Arakawa, *Icarus* **1986**, *68*, 176-184.
- [116] S. Pizzarello, J. R. Cronin, *Geochim. Cosmochim. Acta* **2000**, *64*, 329-338.
- [117] S. Pizzarello, E. Shock, *Cold Spring Harbor Perspect. Biol.* **2010**, *2*, a002105.
- [118] Y. Wolman, W. J. Haverland, S. L. Miller, *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 809-811.
- [119] C. Ponnampereuma, N. W. Gabel, *Space Life Sci.* **1968**, *1*, 64-96.
- [120] W. Gilbert, *Nature* **1986**, *319*, 618-618.
- [121] T. R. Cech, *Science* **1987**, *236*, 1532-1539.
- [122] S. Altman, *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 749-758.
- [123] G. F. Joyce, *New Biol.* **1991**, *3*, 399-407.
- [124] J. A. Doudna, T. R. Cech, *Nature* **2002**, *418*, 222-228.
- [125] J. C. Bowman, N. V. Hud, L. D. Williams, *J. Mol. Evol.* **2015**, *80*, 143-161.
- [126] H. S. Bernhardt, *Biol. Direct* **2012**, *7*, 23.
- [127] K. D. E. James, A.D., *Mol. Origins Life* **1998**, *Volume 28*, pp. 269-294.
- [128] A. Kirschning, *Angew. Chem. Int. Ed.* **2021**, *60*, 6242-6269.
- [129] T. Lindahl, *Nature* **1993**, *362*, 709-715.
- [130] J. K. Bashkin, *Curr. Biol.* **1997**, *7*, R286-R288.
- [131] J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **2001**, *409*, 387-390.
- [132] W. K. Johnston, P. J. Unrau, M. S. Lawrence, M. E. Glasner, D. P. Bartel, *Science* **2001**, *292*, 1319-1325.
- [133] L. E. Orgel, R. Lohrmann, *Acc. Chem. Res.* **1974**, *7*, 368-377.
- [134] J. R. Lorsch, J. W. Szostak, *Nature* **1994**, *371*, 31-36.
- [135] N. Lee, Y. Bessho, K. Wei, J. W. Szostak, H. Suga, *Nat. Struct. Biol.* **2000**, *7*, 28-33.
- [136] R. M. Turk, N. V. Chumachenko, M. Yarus, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 4585-4589.
- [137] B. Zhang, T. R. Cech, *Nature* **1997**, *390*, 96-100.
- [138] T. Bose, G. Fridkin, C. Davidovich, M. Krupkin, N. Dinger, Alla H. Falkovich, Y. Peleg, I. Agmon, A. Bashan, A. Yonath, *Nucleic Acids Res.* **2022**, *50*, 1815-1828.
- [139] D. Xu, Y. Wang, *Biochem. Biophys. Res. Commun.* **2021**, *544*, 81-85.

- [140] I. Agmon, *Int. J. Mol. Sci.* **2018**, *19*, 4021.
- [141] I. Agmon, *Int. J. Mol. Sci.* **2009**, *10*, 2921-2934.
- [142] E. V. Koonin, W. Martin, *Trends Genet.* **2005**, *21*, 647-654.
- [143] T. J. Gibson, A. I. Lamond, *J. Mol. Evol.* **1990**, *30*, 7-15.
- [144] S. Zhang, T. Holmes, C. Lockshin, A. Rich, *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 3334-3338.
- [145] H. B. White, 3rd, *J. Mol. Evol.* **1976**, *7*, 101-104.
- [146] A. D. Goldman, B. Kacar, *J. Mol. Evol.* **2021**, *89*, 127-133.
- [147] J. S. Teichert, F. M. Kruse, O. Trapp, *Angew. Chem. Int. Ed.* **2019**, *58*, 9944-9947.
- [148] J. Xu, V. Chmela, N. J. Green, D. A. Russell, M. J. Janicki, R. W. Góra, R. Szabla, A. D. Bond, J. D. Sutherland, *Nature* **2020**, *582*, 60-66.
- [149] L. E. Orgel, *J. Mol. Evol.* **1989**, *29*, 465-474.
- [150] S. Tagami, P. Li, *Dev., Growth Differ.* **2023**, *65*, 167-174.
- [151] S. A. Benner, H.-J. Kim, M. A. Carrigan, *Acc. Chem. Res.* **2012**, *45*, 2025-2034.
- [152] C. Anastasi, F. F. Buchet, M. A. Crowe, A. L. Parkes, M. W. Powner, J. M. Smith, J. D. Sutherland, *Chem. Biodiversity* **2007**, *4*, 721-739.
- [153] A. Eschenmoser, *Science* **1999**, *284*, 2118-2124.
- [154] L. Zhang, A. Peritz, E. Meggers, *J. Am. Chem. Soc.* **2005**, *127*, 4174-4175.
- [155] K. E. Nelson, M. Levy, S. L. Miller, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3868-3871.
- [156] K. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, A. Eschenmoser, *Science* **2000**, *290*, 1347-1351.
- [157] L. Pasteur, *Recherches sur les propriétés spécifiques des deux acides qui composent l'acide racémique*, Bachelier, **1850**.
- [158] G. Vantomme, J. Crassous, *Chirality* **2021**, *33*, 597-601.
- [159] S. W. T. L. Kelvin, Clarendon Press, **1894**.
- [160] R. S. Cahn, C. Ingold, V. Prelog, *Angew. Chem. Int. Ed.* **1966**, *5*, 385-415.
- [161] D. G. Blackmond, *Cold Spring Harbor Perspect. Biol.* **2010**, *2*, a002147.
- [162] S. R. Buxton, S. M. Roberts, in *Einführung in die Organische Stereochemie* (Eds.: S. R. Buxton, S. M. Roberts), Vieweg+Teubner Verlag, Wiesbaden, **1999**, pp. 37-48.
- [163] L. D. Barron, *Space Sci. Rev.* **2008**, *135*, 187-201.
- [164] W. A. Bonner, *Origins Life Evol. Biospheres* **1995**, *25*, 175-190.
- [165] N. Fujii, *Origins Life Evol. Biospheres* **2002**, *32*, 103-127.
- [166] P. Cintas, *Angew. Chem. Int. Ed.* **2008**, *47*, 2918-2920.
- [167] R. Breslow, Z.-L. Cheng, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5723-5725.
- [168] M. Levine, C. S. Kenesky, D. Mazori, R. Breslow, *Org. Lett.* **2008**, *10*, 2433-2436.
- [169] J. E. Hein, D. G. Blackmond, *Acc. Chem. Res.* **2012**, *45*, 2045-2054.
- [170] N. Fujii, T. Saito, *Chem. Rec.* **2004**, *4*, 267-278.
- [171] S. Toxvaerd, *Symmetry* **2023**, *15*, 155.
- [172] G. F. Joyce, A. W. Schwartz, S. L. Miller, L. E. Orgel, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 4398-4402.
- [173] K. Soai, T. Shibata, H. Morioka, K. Choji, *Nature* **1995**, *378*, 767-768.
- [174] M. Klusmann, H. Iwamura, S. P. Mathew, D. H. Wells, U. Pandya, A. Armstrong, D. G. Blackmond, *Nature* **2006**, *441*, 621-623.
- [175] C. Viedma, J. E. Ortiz, T. d. Torres, T. Izumi, D. G. Blackmond, *J. Am. Chem. Soc.* **2008**, *130*, 15274-15275.
- [176] V. Goldanskii, V. V. Kuzmin, *Nature* **1991**, *352*, 114-114.
- [177] R. H. Perry, C. Wu, M. Nefliu, R. G. Cooks, *Chem. Commun.* **2007**, 1071-1073.
- [178] R. M. Hazen, T. R. Filley, G. A. Goodfriend, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5487-5490.
- [179] W. A. Bonner, P. R. Kavasmaneck, F. S. Martin, J. J. Flores, *Science* **1974**, *186*, 143-144.
- [180] J. Oró, *Nature* **1961**, *190*, 389-390.
- [181] A. S. Sato, Mitsuo; Watanabe, Natsuki; Boero, Mauro; Shigeta, Yasuteru; Umemura, Masayuki, *Astrobiology* **2023**, *23*, 1019-1026.
- [182] C. M. P. Modica, P. de Marcellus, L. Nahon, U. J. Meierhenrich, L. Le Sergeant d'Hendecourt, *Astrophys. J.* **2014**, *788*, 79.
- [183] A. Brandenburg, in *Prebiotic Chemistry and the Origin of Life* (Eds.: A. Neubeck, S. McMahon), Springer International Publishing, Cham, **2021**, pp. 87-115.
- [184] J. Bailey, A. Chrysostomou, J. H. Hough, T. M. Gledhill, A. McCall, S. Clark, F. Menard, M. Tamura, *Science* **1998**, *281*, 672-674.
- [185] U. J. Meierhenrich, L. Nahon, C. Alcaraz, J. H. Bredehöft, S. V. Hoffmann, B. Barbier, A. Brack, *Angew. Chem. Int. Ed.* **2005**, *44*, 5630-5634.
- [186] S. P. Fletcher, R. B. C. Jagt, B. L. Feringa, *Chem. Commun.* **2007**, 2578-2580.

- [187] D. G. Blackmond, M. Klussmann, *Chem. Commun.* **2007**, 3990-3996.
- [188] J. R. Cronin, S. Pizzarello, *Science* **1997**, *275*, 951-955.
- [189] S. Pizzarello, *Acc. Chem. Res.* **2006**, *39*, 231-237.
- [190] F. Devinsky, *Symmetry* **2021**, *13*, 2277.
- [191] M. Quack, *Angew. Chem. Int. Ed.* **2002**, *41*, 4618-4630.
- [192] W. A. Bonner, *Origins Life Evol. Biospheres* **1999**, *29*, 615-624.
- [193] H. Kuhn, *Angew. Chem. Int. Ed.* **1972**, *11*, 798-820.
- [194] F. C. Frank, *Biochim. Biophys. Acta* **1953**, *11*, 459-463.
- [195] J. C. Lacey, R. D. Thomas, M. P. Staves, C. L. Watkins, *Biochim. Biophys. Acta* **1991**, *1076*, 395-400.
- [196] J. C. Lacey, A. F. Hawkins, R. D. Thomas, C. L. Watkins, *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4996-5000.
- [197] J. C. Lacey, Jr., R. D. Thomas, N. S. Wickramasinghe, C. L. Watkins, *J. Mol. Evol.* **1990**, *31*, 251-256.
- [198] N. S. M. D. Wickramasinghe, M. P. Staves, J. C. Lacey, Jr., *Biochemistry* **1991**, *30*, 2768-2772.
- [199] R. Lohrmann, L. E. Orgel, *Nature* **1973**, *244*, 418-420.
- [200] A. L. Weber, J. C. Lacey, *J. Mol. Evol.* **1975**, *6*, 309-320.
- [201] M. Jauker, H. Griesser, C. Richert, *Angew. Chem. Int. Ed.* **2015**, *54*, 14559-14563.
- [202] D. Jovanovic, P. Tremmel, P. S. Pallan, M. Egli, C. Richert, *Angew. Chem. Int. Ed.* **2020**, *59*, 20154-20160.
- [203] O. Doppleb, J. Bremer, M. Bechthold, C. Sánchez Rico, D. Göhringer, H. Griesser, C. Richert, *Chem. Eur. J.* **2021**, *27*, 13544-13551.
- [204] G. Danger, R. Plasson, R. Pascal, *Chem. Soc. Rev.* **2012**, *41*, 5416-5429.
- [205] D. Beaufils, G. Danger, L. Boiteau, J.-C. Rossi, R. Pascal, *Chem. Commun.* **2014**, *50*, 3100-3102.
- [206] K. Tamura, P. Schimmel, *Science* **2004**, *305*, 1253-1253.
- [207] K. Tamura, P. R. Schimmel, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13750-13752.
- [208] K. Tamura, *BioSystems* **2008**, *92*, 91-98.
- [209] K. Tamura, *Int. J. Mol. Sci.* **2011**, *12*, 4745-4757.
- [210] K. Tamura, *Life (Basel, Switz.)* **2015**, *5*, 1687-1699.
- [211] N. Maizels, A. M. Weiner, *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 6729-6734.
- [212] S. Altman, *Philos. Trans. R. Soc., B* **2011**, *366*, 2936-2941.
- [213] L. F. Wu, M. Su, Z. Liu, S. J. Bjork, J. D. Sutherland, *J. Am. Chem. Soc.* **2021**, *143*, 11836-11842.
- [214] F. Müller, L. Escobar, F. Xu, E. Węgrzyn, M. Nainytė, T. Amatov, C. Y. Chan, A. Pichler, T. Carell, *Nature* **2022**, *605*, 279-284.
- [215] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, U. Kummer, *Bioinformatics (Oxford, England)* **2006**, *22*, 3067-3074.
- [216] G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176-2179.
- [217] J. N. Singer, F. M. Müller, E. Węgrzyn, C. Hölzl, H. Hurmiz, C. Liu, L. Escobar, T. Carell, *Angew. Chem. Int. Ed.* **2023**, *62*, e202302360.
- [218] E. Węgrzyn, I. Mejdrová, F. Müller, M. Nainytė, L. Escobar, T. Carell, *Angew. Chem. Int. Ed.* **2024**, *63*, e202319235.
- [219] F. Himmelsbach, B. S. Schulz, T. Trichtinger, R. Charubala, W. Pfeleiderer, *Tetrahedron* **1984**, *40*, 59-72.
- [220] F. P. Ferreira, F. Morvan, *Nucleosides, Nucleotides Nucleic Acids* **2005**, *24*, 1009 - 1013.
- [221] D. L. Usanov, A. I. Chan, J. P. Maianti, D. R. Liu, *Nat. Chem.* **2018**, *10*, 704-714.

Appendix I

Supplementary information

A prebiotically plausible scenario of an RNA-peptide world

In the format provided by the authors and unedited

A prebiotically plausible scenario of an RNA-peptide world

Felix Müller^{1*}, Luis Escobar^{1*}, Felix Xu¹, Ewa Węgrzyn¹, Milda Nainytė¹, Tynchtyk Amatov¹, Chun-Yin Chan¹, Alexander Pichler¹ and Thomas Carell^{1#}

¹ Department of Chemistry, Ludwig-Maximilians-Universität (LMU) München, Butenandtstrasse 5-13, 81377 München, Germany.

* These authors contributed equally.

E-Mail: thomas.carell@lmu.de

Supplementary Information

Table of Contents

1.	General information and instruments for phosphoramidites, amino acids and peptides	S3
2.	Synthesis and characterization data	S3
2.1	Nucleobase-modified 5-methyluridine phosphoramidites	S3
2.2	Npe-protected amino acids and peptides	S6
2.3	Nucleobase-modified <i>N</i> ⁶ -carbamoyl adenosine phosphoramidites	S8
2.4	Nucleobase-modified <i>N</i> ⁶ -triglycylcarbamoyl adenosine nucleoside	S16
2.5	Nucleobase-modified <i>N</i> ⁶ -methylurea adenosine nucleoside	S17
2.6	Nucleobase-modified <i>N</i> ⁶ -triglycylcarbamoyl adenosine nucleoside under prebiotic conditions	S18
2.7	Nucleobase-modified 5-methyluridine 2'-methoxy phosphoramidite	S20
2.8	Nucleobase-modified 2'-methoxy <i>N</i> ⁶ -carbamoyl adenosine phosphoramidite	S22
3.	General information and instruments for oligonucleotides	S24
3.1	Synthesis and purification of oligonucleotides	S24
3.2	Analysis of coupling and cleavage reactions by HPLC and MALDI-TOF mass spectrometry	S25
3.3	Coupling of amino acids and peptides to ONs anchored to the solid support beads	S25
4.	Synthesized oligonucleotides using a DNA/RNA automated synthesizer	S26
4.1	Canonical oligonucleotides (CON)	S26
4.2	Donor oligonucleotides (ON1) with a complementary sequence	S26
4.3	Acceptor oligonucleotides (ON2) with a complementary sequence	S27
4.4	Donor oligonucleotides with non-complementary sequences	S28
5.	HPLC calibration curves using canonical oligonucleotides (CON1-6) and hairpin-type intermediate (ON3a)	S28
6.	Coupling reactions between donor and acceptor oligonucleotides, ON1 and ON2	S31
6.1	Control experiments	S31
6.2	Screening of activators using ON1a (m ⁶ g ⁶ A) and ON2a (mnm ⁵ U)	S31
6.3	Screening of activators using ON1a (m ⁶ g ⁶ A) and ON2b (nm ⁵ U)	S33
6.4	Screening of activators using ON1a (m ⁶ g ⁶ A) and ON2c (vmnm ⁵ U)	S34
6.5	Coupling reactions of ON1j (m ⁶ g ⁶ A, amino nitrile) with ON2a-c	S35
6.6	Coupling reactions of ON1b-i (m ⁶ aa ⁶ A) with ON2a	S37
6.7	Coupling reactions of ON1b-i (m ⁶ aa ⁶ A) with ON2c	S41
7.	Synthesized peptide-oligonucleotides using solid support beads	S45
7.1	Donor peptide-oligonucleotides with a complementary sequence	S45
7.2	Acceptor peptide-oligonucleotides with a complementary sequence	S46
8.	Coupling reactions between donor and acceptor peptide-oligonucleotides	S47
8.1	Coupling reactions of donor peptide-oligonucleotides with ON2c	S47

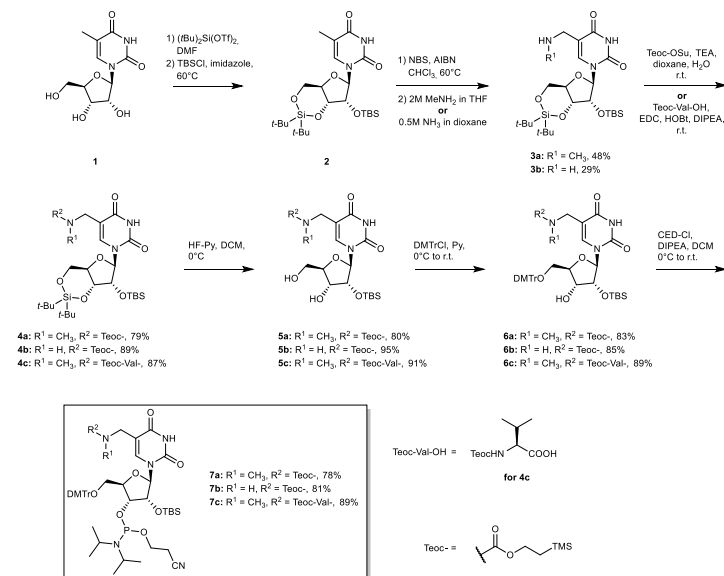
8.2	Coupling reactions of ON1a (m ⁶ g ⁶ A) with acceptor peptide-oligonucleotides	S49
8.3	Coupling reactions of donor and acceptor peptide-oligonucleotides	S51
9.	Concentration of the product versus time in selected coupling reactions	S52
10.	Coupling reactions between oligonucleotides containing multiple donor or acceptor units	S53
11.	Coupling reactions between ON2c and donor oligonucleotides with non-complementary sequences	S55
12.	Coupling reactions between ON2c and donor oligonucleotides with different lengths	S56
13.	Stability of selected acceptor oligonucleotides (ON2)	S58
14.	Cleavage of urea in selected oligonucleotides and cyclic peptide products	S59
14.1	Cleavage reactions of ON1c (m ⁶ v ⁶ A) and ON1k (v ⁶ A) at pH 5	S59
14.2	Cleavage reaction of ON3a (m ⁶ g ⁶ A coupled with mnm ⁵ U)	S60
14.3	Cleavage reactions of ON3c (m ⁶ g ⁶ A coupled with vmnm ⁵ U)	S62
14.4	Cleavage reactions of peptide-oligonucleotides at pH 4	S65
15.	Coupling and cleavage reactions between donor and acceptor oligonucleotides containing 2'-OMe nucleosides	S68
15.1	Coupling and cleavage reactions of ON1a (m ⁶ g ⁶ A) with ON2g	S68
15.2	Coupling and cleavage reactions of ON1o (m ⁶ g ⁶ Am) with ON2h	S71
15.3	Coupling and cleavage reactions of donor and acceptor-peptide oligonucleotides	S73
15.4	Coupling reactions between ON2g and donor oligonucleotides of different length	S74
16.	Determination of melting temperatures by UV spectroscopic experiments	S76
16.1	Melting temperature of a double strand from canonical oligonucleotides	S76
16.2	Melting temperatures of double strands from donor and acceptor oligonucleotides	S76
16.3	Melting temperatures of double strands from donor and acceptor peptide-oligonucleotides	S78
16.4	Melting temperatures of selected cyclic peptide products	S79
17.	NMR spectra of synthesized compounds	S81
18.	References	S156

1. General information and instruments for phosphoramidites, amino acids and peptides

Reagents were purchased from commercial suppliers and used without further purification unless otherwise stated. All anhydrous solvents stored under inert atmosphere were also purchased. All reactions involving air/moisture sensitive reagents/intermediates were performed under inert atmosphere using oven-dried glassware. Routine ¹H NMR, ¹³C{¹H} NMR and ³¹P{¹H} NMR were recorded on a Bruker Ascend 400 spectrometer (400 MHz for ¹H NMR, 100 MHz for ¹³C NMR and 162 MHz for ³¹P NMR) or a Bruker ARX 600 spectrometer (600 MHz for ¹H NMR, 150 MHz for ¹³C NMR and 243 MHz for ³¹P NMR). Deuterated solvents used are indicated in the characterization and chemical shifts (δ) are reported in ppm. Residual solvent peaks were used as reference. ¹All NMR J values are given in Hz. COSY, HMQC and HMBC experiments were recorded to help with the assignment of ¹H and ¹³C signals. NMR spectra were analyzed using MestReNova software version 10.0. High Resolution Mass Spectra (HRMS) were measured on a Thermo Finnigan LTQ-FT with ESI as ionization mode. IR spectra were recorded on a Perkin-Elmer Spectrum BX II FT-IR instrument equipped with an ATR accessory. Column chromatography was performed with silica gel technical grade (Macherey-Nagel), 40-63 μm particle size. Reaction progress was monitored by Thin Layer Chromatography (TLC) analysis on silica gel 60 F254 and stained with *para*-anisaldehyde, potassium permanganate or cerium ammonium molybdate solution.

2. Synthesis and characterization data

2.1 Nucleobase-modified 5-methyluridine phosphoramidites



Scheme S1. Synthesis of nucleobase-modified 5-methyluridine phosphoramidites.

General procedure for the synthesis of **3a,b**:

Silyl-protected 5-methyluridine **2** was synthesized starting from 5-methyluridine **1** following a procedure previously described in literature.² A solution of **2** (1.0 equiv.) in dry CHCl₃ was heated at 60°C. *N*-bromosuccinimide (NBS) (1.2 equiv., previously purified by recrystallization) and azobisisobutyronitrile (AIBN) (0.12 equiv.) were added and the reaction was stirred under reflux for 1.5 h. After that, the reaction mixture was cooled to r.t. and either MeNH₂ (2 M in THF, 5.0 equiv.) for **3a** or NH₃ (0.5 M in 1,4-dioxane, 5.0 equiv.) for **3b** were added. The resulting suspension was stirred for 2 h at r.t. and, subsequently, it was diluted with aq. sat. NaHCO₃ solution. The crude was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated. The crude was purified by silica gel column chromatography to furnish **3a,b** as a yellow foam.

3a: Yield: 48%; $R_f = 0.11$ (9:1 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2931 (w), 2858 (w), 2359 (w), 1682 (s), 1462 (m), 1386 (w), 1254 (m), 1202 (w), 1167 (w), 1115 (m), 1057 (s), 1000 (m), 938 (w), 882 (m), 827 (s), 778 (s), 754 (m), 685 (w); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 7.35 (s, 1H), 5.66 (s, 1H), 4.47 (dd, $J = 9.5, 4.7$ Hz, 1H), 4.28 (d, $J = 4.7$ Hz, 1H), 4.18-4.06 (m, 1H), 4.05-3.97 (m, 1H), 3.92 (dd, $J = 9.5, 4.7$ Hz, 1H), 3.58-3.47 (m, 2H), 2.41 (s, 3H), 1.03 (s, 9H), 1.01 (s, 9H), 0.91 (s, 9H), 0.15 (s, 3H), 0.12 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 164.2, 150.1, 138.1, 111.1, 94.1, 76.1, 75.3, 74.6, 67.7, 47.6, 35.0, 27.6, 27.1, 26.0, 22.9, 20.5, 18.4, -4.2, -4.9; HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₂₅H₄₈N₃O₆Si₂ 542.3076; Found 542.3076.

3b: Yield: 29%; $R_f = 0.25$ (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3052 (w), 2934 (w), 2858 (w), 2363 (w), 1687 (m), 1471 (w), 1422 (w), 1388 (w), 1264 (s), 1204 (w), 1168 (w), 1115 (m), 1059 (m), 999 (m), 938 (w), 896 (w), 882 (m), 828 (s), 780 (m), 731 (s), 702 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 7.29 (s, 1H), 5.69 (s, 1H), 4.49 (dd, $J = 9.1, 5.0$ Hz, 1H), 4.28 (d, $J = 4.8$ Hz, 1H), 4.18-4.08 (m, 1H), 4.01 (dd, $J = 10.6, 9.1$ Hz, 1H), 3.92 (dd, $J = 9.5, 4.8$ Hz, 1H), 3.60 (s, 2H), 1.05 (s, 9H), 1.02 (s, 9H), 0.92 (s, 9H), 0.16 (s, 3H), 0.13 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 163.2, 149.6, 136.3, 94.1, 76.2, 75.4, 74.6, 67.7, 39.2, 27.6, 27.1, 26.0, 22.9, 20.5, 18.4, -4.2, -4.9; HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₂₄H₄₆N₃O₆Si₂ 528.2920; Found 528.2921.

General procedures for the synthesis of 4a-c:

Procedure A (for compounds 4a,b): To a solution of **3a,b** (1.0 equiv.) in 1,4-dioxane and H₂O (1:1 v/v) were added teoc-OSu (1.1 equiv.) and triethylamine (TEA) (1.5 equiv.). The mixture was stirred at r.t. for 16 h. After that, the crude was diluted with water and extracted three times with Et₂O. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated. The obtained residue was purified by silica gel column chromatography to yield the teoc-protected compound **4a,b** as a white solid.

Procedure B (for compound 4c): Teoc-protected valine was synthesized following a previously reported procedure in literature.³ Teoc-Val-OH (1.2 equiv.) was dissolved in dry DCM and DMF (99:1 v/v). To the solution, 1-hydroxybenzotriazole hydrate (HOBt·H₂O) (1.2 equiv.), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl) (1.2 equiv.) and *N,N*-diisopropylethylamine (DIPEA) (1.2 equiv.) were added. After stirring at r.t. for 30 min, a solution of **3a** (1.0 equiv.) in DCM was added and the reaction was stirred for 24 h. The reaction mixture was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated. Purification by silica gel column chromatography furnished the amino acid conjugate **4c** as a white foam.

4a: Yield: 79%; $R_f = 0.34$ (4:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3054 (w), 2956 (w), 2359 (w), 1692 (m), 1463 (w), 1422 (w), 1264 (s), 1214 (w), 1167 (w), 1146 (w), 1059 (w), 1000 (w), 938 (w), 895 (m), 838 (m), 730 (s), 702 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 9.14 (s, 1H), 7.54 (s, 1H), 5.65 (s, 1H), 4.48 (dd, $J = 9.2, 4.2$ Hz, 1H), 4.28 (d, $J = 4.2$ Hz, 1H), 4.23-3.98 (m, 6H), 3.91 (dd, $J = 9.2, 4.2$ Hz, 1H), 2.96 (s, 3H), 1.05 (s, 9H), 1.03-0.96 (m, 11H), 0.93 (s, 9H), 0.18 (s, 3H), 0.13 (s, 3H), 0.04 (s, 9H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 163.6, 157.1, 149.7, 139.4, 110.8, 93.8, 76.0, 75.5, 74.9, 67.6, 63.8, 45.0, 35.6, 27.7, 27.1, 26.0, 22.8, 20.5, 18.4, 17.9, -1.3, -4.2, -4.9; HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₃₁H₆₀N₃O₆Si₃ 686.3683; Found 686.3683.

4b: Yield: 89%; $R_f = 0.23$ (4:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2937 (w), 2359 (w), 2167 (w), 1690 (m), 1470 (w), 1251 (m), 1213 (w), 1127 (w), 1061 (m), 999 (m), 831 (m), 779 (m), 730 (s); ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) 8.13 (s, 1H), 7.46 (s, 1H), 5.65 (s, 1H), 5.23 (t, $J = 5.9$ Hz, 1H), 4.50 (dd, $J = 9.0, 4.9$ Hz, 1H), 4.28 (d, $J = 4.6$ Hz, 1H), 4.20-4.05 (m, 4H), 3.98 (d, $J = 6.3$ Hz, 2H), 3.90 (dd, $J = 9.5, 4.6$ Hz, 1H), 1.06 (s, 9H), 1.02 (s, 9H), 0.99-0.88 (m, 11H), 0.18 (s, 3H), 0.14 (s, 3H), 0.03 (s, 9H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 162.9, 156.9, 149.4, 138.4, 111.4, 93.9, 76.0, 75.5, 74.9, 67.6, 63.4, 37.7, 27.7, 27.1, 26.0, 22.9, 20.5, 18.4, 17.8, -1.3, -4.1, -4.9; HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₃₀H₅₈N₃O₆Si₃ 672.3526; Found 672.3535.

4c: Yield: 87%; $R_f = 0.29$ (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3053 (w), 2956 (w), 2859 (w), 2359 (w), 1689 (m), 1648 (w), 1586 (w), 1536 (w), 1471 (m), 1382 (w), 1366 (m), 1311 (w), 1264 (s), 1168 (w), 1114 (m), 1059 (m), 1002 (w), 938 (w), 835 (m), 732 (s), 702 (s); For major rotamer: ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 9.22 (s, 1H), 7.66 (s, 1H), 5.69 (s, 1H), 5.40 (d, $J = 9.0$ Hz, 1H), 4.51-4.42 (m, 2H), 4.25 (d, $J = 14.3$ Hz, 1H), 4.21-4.01 (m, 6H), 3.97 (dd, $J = 9.0, 4.8$ Hz, 1H), 3.21 (s, 3H), 1.92-1.87 (m, 1H), 1.09 (s, 9H), 1.05-0.98 (m, 11H), 0.94-0.89 (m, 12H), 0.80 (d, $J = 6.7$ Hz, 3H), 0.14 (s, 3H), 0.11 (s, 3H), 0.02 (s, 9H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 172.7, 163.6, 157.0, 149.6, 141.3, 110.1, 93.9, 76.1, 75.6, 74.8, 67.6, 63.4, 55.4, 44.5, 37.3, 31.3, 27.7, 27.1, 26.0, 22.8, 20.5, 19.6, 18.4, 17.8, 17.1, -1.3, -4.2, -5.0; HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₃₆H₆₉N₄O₆Si₃ 785.4367; Found 785.4363.

General procedure for the synthesis of 5a-c:

The modified 5-methyluridine **4a-c** (1.0 equiv.) was dissolved in DCM/pyridine (9:1 v/v) and cooled to 0°C in a plastic reaction vessel. Subsequently, a solution of 70% HF-pyridine (5.0 equiv.) was slowly added, and the reaction mixture was stirred at 0°C for 2 h. The reaction was quenched by adding aq. sat. NaHCO₃ and the crude was extracted three times with DCM. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated. The crude product was purified by silica gel column chromatography to afford the diol compound **5a-c** as a white foam.

5a: Yield: 80%; $R_f = 0.42$ (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3417 (w), 3060 (w), 2949 (w), 2856 (w), 2359 (w), 1673 (s), 1462 (m), 1401 (w), 1362 (w), 1250 (m), 1214 (w), 1144 (m), 1088 (m), 1060 (m), 1005 (w), 938 (w), 833 (s), 777 (s), 693 (w); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 9.47 (s, 1H), 8.19 (s, 1H), 5.87 (d, $J = 5.2$ Hz, 1H), 4.48 (t, $J = 5.1$ Hz, 1H), 4.30-3.85 (m, 7H), 3.83-3.74 (m, 1H), 2.97 (s, 3H), 2.78 (br s, 1H), 1.03-0.92 (m, 2H), 0.88 (s, 9H), 0.06 (s, 6H), 0.02 (s, 9H) (some proton signals appeared too broad for an unequivocal assignment); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 163.8, 157.4, 150.5, 141.9, 111.4, 90.2, 85.8, 75.3, 71.3, 64.1, 62.2, 44.5, 35.6, 25.8, 18.1, -1.4, -4.7 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₂₃H₄₄N₃O₈Si₂ 546.2661; Found 546.2666.

5b: Yield: 95%; $R_f = 0.23$ (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3386 (w), 2950 (w), 2854 (w), 2362 (w), 1674 (s), 1524 (m), 1470 (m), 1390 (w), 1333 (w), 1248 (s), 1179 (w), 1115 (m), 1086 (w), 1060 (s), 1001 (w), 938 (w), 902 (w), 857 (m), 833 (s), 779 (s), 694 (w); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 8.51 (s, 1H), 8.09 (s, 1H), 5.80 (s, 1H), 5.35 (t, $J = 6.2$ Hz, 1H), 4.49 (t, $J = 4.8$ Hz, 1H), 4.33-4.22 (m, 1H), 4.17-4.05 (m, 4H), 4.02-3.92 (m, 3H), 3.81 (dd, $J = 12.0, 5.3$ Hz, 1H), 3.57 (t, $J = 5.3$ Hz, 1H), 2.70 (d, $J = 4.3$ Hz, 1H), 0.99-0.92 (m, 3H), 0.90 (s, 9H), 0.11-0.08 (m, 6H), 0.02 (s, 9H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 163.0, 157.2, 150.2, 141.0, 111.8, 90.9, 85.7, 75.1, 71.0, 63.6, 62.1, 37.2, 25.8, 18.1, 17.8, -1.3, -4.6, -5.0; HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₂₂H₄₂N₃O₈Si₂ 532.2505; Found 532.2509.

5c: Yield: 91%; $R_f = 0.18$ (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3440 (w), 3054 (w), 2953 (w), 2857 (w), 2359 (w), 1677 (s), 1463 (m), 1401 (w), 1362 (w), 1264 (s), 1250 (m), 1215 (w), 1137 (m), 1112 (w), 1089 (w), 1060 (w), 1005 (w), 937 (w), 836 (s), 779 (m), 733 (s), 701 (s); For major rotamer: ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 8.80 (s, 1H), 8.06 (s, 1H), 5.94 (d, $J = 5.0$ Hz, 1H), 5.45 (d, $J = 9.9$ Hz, 1H), 4.59 (d, $J = 15.0$ Hz, 1H), 4.49 (dd, $J = 9.9, 5.4$ Hz, 1H), 4.43-4.05 (m, 5H), 3.91 (d, $J = 15.0$ Hz, 2H), 3.80 (d, $J = 12.0$ Hz, 1H), 3.19 (s, 3H), 2.71 (d, $J = 3.1$ Hz, 1H), 2.01-1.94 (m, 1H), 1.01-0.93 (m, 5H), 0.93-0.85 (m, 13H), 0.07 (s, 3H), 0.06 (s, 3H), 0.03 (s, 9H) (some proton signals appeared too broad for an unequivocal assignment); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 172.9, 163.2, 157.5, 150.2, 139.8, 110.0, 89.5, 85.7, 75.9, 71.3, 63.9, 61.8, 55.8, 44.5, 36.7, 31.0, 25.8, 19.7, 18.1, 17.9, 17.1, -1.4, -4.7, -5.1 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) m/z : [M+H]⁺ Calcd. for C₂₈H₅₃N₄O₉Si₃ 645.3346; Found 645.3349.

General procedure for the synthesis of 6a-c:

To a solution of the 3',5'-deprotected 5-methyluridine derivative **5a-c** (1.0 equiv.) in pyridine was added 4,4'-dimethoxytrityl chloride (DMTr-Cl) (1.5 equiv.). After stirring at r.t. for 16 h, the reaction mixture was concentrated and purified by silica gel column chromatography with an addition of 0.1% of pyridine to the eluent to afford the DMTr-protected compound **6a-c** as a white foam.

6a: Yield: 83%; $R_f = 0.57$ (1:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3444 (w), 3055 (w), 2953 (w), 2857 (w), 2359 (w), 1678 (s), 1608 (w), 1583 (w), 1508 (m), 1463 (m), 1401 (w), 1342 (w), 1297 (w), 1264 (m), 1248 (s), 1175 (m), 1150 (m), 1113 (w), 1089 (w), 1034 (m), 1006 (w), 938 (w), 910 (w), 830 (s), 780 (m), 733 (s), 701 (s); For major rotamer: ¹H NMR (400 MHz, acetone-*d*₆, 298 K) δ (ppm): 10.21 (s, 1H), 7.76 (s, 1H), 7.57-7.46 (m, 2H), 7.45-7.37 (m, 4H), 7.37-7.29 (m, 2H), 7.28-7.19 (m, 1H), 6.90 (d, $J = 8.9$ Hz, 4H), 5.94 (s, 1H), 4.44 (br s, 1H), 4.22-4.03 (m, 3H), 3.84-3.71 (m, 8H), 3.44 (br s, 2H), 2.90 (br s, 3H), 1.08-0.81 (m, 11H), 0.15 (s, 6H), 0.03 (s, 9H) (some proton signals appeared too broad for an unequivocal assignment); ¹³C{¹H} NMR (100 MHz, acetone-*d*₆, 298 K) δ (ppm): 163.8, 159.6, 156.7, 151.2, 146.1, 136.7, 131.1, 131.0, 129.0, 128.7, 114.0, 114.0, 113.6, 111.4, 89.7, 87.3, 84.3, 76.5, 71.5, 64.6, 63.6, 55.5, 46.5, 35.7, 26.2, 18.7, 18.3, -1.4, -4.6, -4.6 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) m/z : [M-H]⁻ Calcd. for C₄₄H₆₀N₃O₁₀Si₂ 846.3823; Found 846.3825.

6b: Yield: 85%; $R_f = 0.33$ (2:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3342 (w), 2950 (w), 2855 (w), 2358 (w), 1708 (s), 1607 (w), 1582 (w), 1508 (m), 1462 (m), 1390 (w), 1248 (s), 1175 (m), 1116 (m), 1089 (w), 1063 (m), 1035 (m), 969 (w), 937 (w), 859 (m), 835 (s), 780 (s), 754 (m), 726 (w), 699 (m); ¹H NMR (400 MHz, acetone-*d*₆, 298 K) δ (ppm): 10.18 (s, 1H), 7.78 (s, 1H), 7.52 (d, $J = 7.6$ Hz, 2H), 7.41 (d, $J = 8.8$ Hz, 4H), 7.34 (t, $J = 7.6$ Hz, 2H), 7.24 (t, $J = 7.6$ Hz, 1H), 6.92 (d, $J = 8.8$ Hz, 4H), 6.05 (t, $J = 5.0$ Hz, 1H), 5.95 (d, $J = 4.6$ Hz, 1H), 4.46 (t, $J = 5.0$ Hz, 1H), 4.28-4.22 (m, 1H), 4.17-4.12 (m, 1H), 3.82 (d, $J = 5.8$ Hz, 1H), 3.79 (s, 3H), 3.64 (dd, $J = 14.5, 5.5$ Hz, 1H), 3.56 (dd, $J = 14.5, 5.9$ Hz, 1H), 3.45 (dd, $J = 10.8, 4.2$ Hz, 1H), 3.39 (dd, $J = 10.8, 2.5$ Hz, 1H),

0.97-0.83 (m, 1H), 0.15 (s, 3H), 0.14 (s, 3H), 0.02 (s, 9H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 163.5, 159.6, 157.0, 151.2, 146.0, 138.9, 136.7, 136.6, 131.0, 129.0, 128.8, 127.6, 114.0, 112.2, 89.6, 87.4, 84.3, 76.7, 71.6, 64.4, 62.8, 55.5, 38.5, 26.2, 18.7, 18.4, -1.4, -4.6, -4.7; HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{43}\text{H}_{60}\text{N}_3\text{O}_{10}\text{Si}_2$ 834.3812; Found 834.3801.

6c: Yield: 89%; R_f = 0.42 (1:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3054 (w), 2954 (w), 2930 (w), 2857 (w), 2359 (w), 1687 (s), 1644 (w), 1608 (w), 1508 (m), 1463 (m), 1389 (w), 1263 (m), 1249 (s), 1175 (m), 1115 (w), 1083 (w), 1061 (w), 1035 (m), 967 (w), 935 (w), 914 (w), 858 (m), 833 (s), 780 (w), 733 (s), 700 (s); For major rotamer: ^1H NMR (400 MHz, CDCl_3 , 298 K) δ (ppm): 9.10 (s, 1H), 7.75 (s, 1H), 7.45 (d, J = 7.3 Hz, 2H), 7.35 (d, J = 8.8 Hz, 4H), 7.31-7.24 (m, 2H), 7.20 (t, J = 7.3 Hz, 1H), 6.82 (d, J = 8.8 Hz, 4H), 5.88 (d, J = 3.8 Hz, 1H), 5.41 (d, J = 8.9 Hz, 1H), 4.43-4.37 (m, 1H), 4.34-4.29 (m, 1H), 4.18-4.07 (m, 4H), 4.00 (d, J = 14.3 Hz, 1H), 3.77 (s, 6H), 3.54-3.40 (m, 3H), 3.11 (s, 3H), 2.59 (d, J = 6.3 Hz, 1H), 1.92-1.78 (m, 1H), 1.04-0.96 (m, 2H), 0.94-0.86 (m, 12H), 0.79 (d, J = 6.7 Hz, 3H), 0.12 (s, 3H), 0.11 (s, 3H), 0.03 (s, 9H) (some proton signals appeared too broad for an unequivocal assignment); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3 , 298 K) δ (ppm): 172.0, 163.2, 158.7, 156.9, 150.0, 149.8, 144.8, 141.2, 136.2, 135.8, 130.4, 130.3, 128.3, 128.0, 127.0, 123.9, 113.3, 110.3, 89.9, 86.8, 83.7, 75.6, 70.8, 63.6, 63.3, 55.3, 45.5, 37.1, 31.4, 25.8, 19.6, 18.1, 17.9, 17.3, -1.4, -4.5, -5.1 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) m/z : $[\text{M}-\text{H}]^-$ Calcd. for $\text{C}_{48}\text{H}_{69}\text{N}_4\text{O}_{11}\text{Si}_2$ 945.4507; Found 945.4508.

General procedure for the synthesis of phosphoramidites **7a-c**:

A solution of 5'-DMTR-protected compound **6a-c** (1.0 equiv.) and DIPEA (4.0 equiv.) in dry DCM was cooled to 0°C. To this solution was slowly added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (CED-Cl) (2.5 equiv.) and the reaction mixture was stirred at r.t. for 5 h. The reaction was quenched by addition of aq. sat. NaHCO_3 and the crude was extracted three times with DCM. The combined organic layers were dried (MgSO_4), filtered and concentrated under reduced pressure. After purification by silica gel column chromatography with an addition of 0.1% pyridine and co-lyophilization from benzene, the desired phosphoramidite **7a-c** was obtained as a mixture of diastereoisomers and rotamers as a white foam.

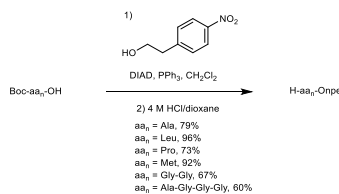
7a: Yield: 78%; R_f = 0.17 (2:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.3, 150.2, 148.8, 148.3; HRMS (ESI) m/z : $[\text{M}-\text{H}]^-$ Calcd. for $\text{C}_{53}\text{H}_{77}\text{N}_5\text{O}_{11}\text{PSi}_2$ 1046.4901, Found 1046.4896.

7b: Yield: 81%; R_f = 0.57 (1:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.2, 148.3; HRMS (ESI) m/z : $[\text{M}-\text{H}]^-$ Calcd. for $\text{C}_{52}\text{H}_{77}\text{N}_5\text{O}_{11}\text{PSi}_2$ 1032.4744; Found 1032.4745.

7c: Yield: 89%; R_f = 0.45 (1:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.0, 149.9, 149.4, 149.3; HRMS (ESI) m/z : $[\text{M}-\text{H}]^-$ Calcd. for $\text{C}_{58}\text{H}_{86}\text{N}_6\text{O}_{12}\text{PSi}_2$ 1145.5585; Found 1145.5595.

2.2 Npe-protected amino acids and peptides

The npe-protected amino acids Gly, Val, Thr, Phe and Asp were synthesized following previously reported procedures in the literature.⁴



Scheme S2. Synthesis of npe-protected amino acids and peptides.

Procedure A (for Ala, Leu, Pro and Met): Step 1. **Boc-aa-OH** (1.0 equiv.), 2-(4-nitrophenyl)ethanol (1.3 equiv.) and PPh₃ (1.3 equiv.) were dissolved in dry CH_2Cl_2 and stirred at 0°C under nitrogen atmosphere. Diisopropyl azodicarboxylate (DIAD) (1.3 equiv.) was added dropwise and the reaction was stirred at r.t. overnight. Afterwards, the reaction was stopped and the crude was washed two times with water. The organic layer was dried (Na_2SO_4), filtered and concentrated *in vacuo*. The crude was purified by silica gel column chromatography affording the **Boc-aa-Onpe** as a white solid. Step 2. **Boc-aa-Onpe** (1.0 equiv.) was dissolved in 4 M HCl in 1,4-dioxane at 0°C. After the reaction was stirred at r.t. for 1 h, the mixture was concentrated obtaining a white solid. The white solid was triturated with Et_2O , filtered and washed with additional Et_2O . The npe-protected amino acid **H-aa-Onpe** chloride salt was isolated as a white solid. For **Pro**, an oil was obtained which was washed with aq. sat. NaHCO_3 and the

crude was extracted with EtOAc. The organic layer was dried (Na_2SO_4), filtered and concentrated under reduced pressure. The crude was purified by silica gel column chromatography affording the **H-Pro-Onpe** as a pale-yellow oil.

H-Ala-Onpe-HCl: Yield: 79% over two steps; IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2843 (m), 1730 (s), 1598 (m), 1345 (s), 1269 (w), 1233 (s), 1195 (m), 1115 (m), 820 (m), 746 (m); ^1H NMR (400 MHz, DMSO- d_6 , 298 K) δ (ppm): 8.59 (br s, 3H), 8.19-8.17 (m, 2H), 7.61-7.59 (m, 2H), 4.50-4.36 (m, 2H), 4.00 (q, J = 7.2 Hz, 1H), 3.10 (t, J = 6.3 Hz, 2H), 1.33 (d, J = 7.2 Hz, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, DMSO- d_6 , 298 K) δ (ppm): 169.9, 146.4, 146.3, 130.4, 123.5, 65.3, 47.8, 33.9, 15.7; HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{11}\text{H}_{15}\text{N}_2\text{O}_4$: 239.1026; Found 239.1027.

H-Leu-Onpe-HCl: Yield: 96% over two steps; IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3663 (w), 2871 (m), 1737 (s), 1589 (m), 1516 (s), 1503 (s), 1380 (s), 1260 (w), 1207 (m), 1109 (w), 959 (w), 856 (m), 812 (m), 735 (s); ^1H NMR (400 MHz, DMSO- d_6 , 298 K) δ (ppm): 8.63 (s, 3H), 8.17 (d, J = 8.7 Hz, 2H), 7.60 (d, J = 8.7 Hz, 2H), 4.51-4.38 (m, 2H), 3.82 (t, J = 6.5 Hz, 1H), 3.11 (t, J = 6.5 Hz, 2H), 1.54-1.43 (m, 3H), 0.75 (t, J = 5.3 Hz, 6H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, DMSO- d_6 , 298 K) δ (ppm): 169.8, 146.4, 130.4, 123.4, 65.3, 50.4, 33.8, 23.7, 22.2, 21.8; HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{14}\text{H}_{21}\text{N}_2\text{O}_4$: 281.1496; Found 281.1495.

H-Pro-Onpe: Yield: 73% over two steps; R_f = 0.30 (9:1 $\text{CH}_2\text{Cl}_2/\text{IPA}$); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3400 (w), 2879 (w), 1649 (s), 1513 (s), 1432 (m), 1318 (s), 1159 (w), 1048 (m), 856 (m), 747 (m); ^1H NMR (400 MHz, CDCl_3 , 298 K) δ (ppm): 8.18-8.16 (m, 2H); 7.40-7.38 (m, 2H); 4.40-4.37 (m, 2H); 3.74-3.71 (m, 1H); 3.07 (t, J = 6.7 Hz, 2H); 3.05-2.87 (m, 2H); 2.29 (br s, 1H); 2.13-2.02 (m, 1H); 1.77-1.67 (m, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3 , 298 K) δ (ppm): 175.4, 147.0, 145.6, 129.9, 123.9, 64.3, 59.8, 47.1, 35.0, 30.4, 25.6; HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{13}\text{H}_{17}\text{N}_2\text{O}_4$: 265.1183; Found 265.1179.

H-Met-Onpe-HCl: Yield: 92% over two steps; IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2852 (w), 1756 (m), 1743 (m), 1598 (w), 1567 (w), 1509 (s), 1347 (s), 1279 (m), 1256 (w), 1230 (w), 1206 (m), 1194 (m), 1148 (w), 1109 (w), 1066 (m), 1000 (w), 856 (m), 827 (m), 793 (w), 769 (m), 744 (s), 694 (m); ^1H NMR (400 MHz, DMSO- d_6 , 298 K) δ (ppm): 8.77 (s, 3H), 8.18-8.15 (m, 2H), 7.62-7.59 (m, 2H), 4.50-4.41 (m, 2H), 4.01 (br s, 1H), 3.11 (t, J = 6.3 Hz, 2H), 2.55-2.43 (m, 1H), 2.38-2.27 (m, 1H), 2.03-1.86 (m, 5H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, DMSO- d_6 , 298 K) δ (ppm): 169.1, 146.3, 146.3, 130.4, 123.5, 65.4, 50.8, 33.8, 29.3, 28.2, 14.1; HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{13}\text{H}_{19}\text{N}_2\text{O}_4\text{S}$: 299.1060; Found 299.1058.

Procedure B (for Gly-Gly and Ala-Gly-Gly-Gly): Step 1. **Boc-aa_n-OH** (1.0 equiv.) and 2-(4-nitrophenyl)ethanol (1.3 equiv.) were suspended in dry ACN. Dry pyridine was added giving a solution. The solution was stirred at 0°C under nitrogen atmosphere. *N,N*-dicyclohexylcarbodiimide (DCC) (1.3 equiv.) and HOBt (1.3 equiv.) were added and the reaction was stirred at r.t. overnight. After that, the reaction was quenched with 1 M aq. citric acid solution at r.t. for 30 min. The crude was diluted with EtOAc and filtered. The precipitate was washed with EtOAc two times. The organic layer was washed with aq. sat. NaHCO_3 solution and water. The organic layer was dried (Na_2SO_4), filtered and concentrated. The crude was purified by silica gel column chromatography affording the boc- and npe-protected peptide. Step 2. **Boc-aa_n-Onpe** (1.0 equiv.) was dissolved in 4 M HCl in 1,4-dioxane at 0°C. After the reaction was stirred at r.t. for 2 h, the mixture was concentrated obtaining a white solid. The white solid was triturated with Et_2O , filtered and washed with additional Et_2O . The npe-protected peptide **H-aa_n-Onpe** chloride salt was isolated as a white solid.

H-Gly-Gly-Onpe-HCl: Yield: 67% over two steps; IR (ATR) $\tilde{\nu}$ (cm^{-1}): 1741 (s), 1677 (m), 1660 (s), 1570 (w), 1512 (s), 1479 (w), 1229 (w), 1207 (s), 1199 (s), 1109 (m); ^1H NMR (400 MHz, DMSO- d_6 , 298 K) δ (ppm): 8.91 (t, J = 5.7 Hz, 1H), 8.21 (br s, 3H), 8.19-8.17 (m, 2H), 7.58-7.56 (m, 2H), 4.34 (t, J = 6.4 Hz, 2H), 3.91 (d, J = 5.7 Hz, 2H), 3.59 (s, 2H), 3.06 (t, J = 6.4 Hz, 2H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, DMSO- d_6 , 298 K) δ (ppm): 169.4, 166.6, 146.5, 146.3, 130.3, 123.5, 64.3, 40.6, 39.7 (the signal overlaps with that of the solvent), 34.0; HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{12}\text{H}_{16}\text{N}_3\text{O}_5$: 282.1084; Found 282.1086.

H-Ala-Gly-Gly-Gly-Onpe-HCl: Yield: 60% over two steps; IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3222 (w), 2931 (w), 1743 (w), 1654 (s), 1514 (s), 1188 (s), 1117 (m), 871 (m), 856 (m), 697 (m); ^1H NMR (400 MHz, DMSO- d_6 , 298 K) δ (ppm): 8.80 (t, J = 5.5 Hz, 1H), 8.38-8.33 (m, 2H), 8.25 (br s, 3H), 8.18-8.16 (m, 2H), 7.58-7.55 (m, 2H), 4.31 (t, J = 6.4 Hz, 2H), 3.91-3.87 (m, 1H), 3.84-3.79 (m, 4H), 3.76-3.72 (m, 2H), 3.04 (t, J = 6.4 Hz, 2H), 1.36 (d, J = 6.9 Hz, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, DMSO- d_6 , 298 K) δ (ppm): 169.9, 169.7, 169.3, 168.6, 146.5, 146.3, 130.3, 123.5, 64.2, 48.2, 42.0, 41.7, 40.6, 34.0, 17.1; HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{17}\text{H}_{24}\text{N}_5\text{O}_7$: 410.1670; Found: 410.1671.

145.6, 139.3, 129.9, 123.8, 122.8, 92.5, 77.4, 76.0, 75.7, 74.8, 68.0, 64.6, 50.0, 35.0, 34.7, 27.6, 27.1, 26.0, 22.9, 20.5, 18.5, 18.4, -4.1, -4.9; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₇H₅₅N₇O₉Si₂ 800.3829; Found 800.3836.

11c: Yield: 90%; *R*_f = 0.34 (1:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3246 (m), 2977 (w), 1732 (m), 1686 (s), 1524 (s), 1469 (w), 1372 (m), 1254 (s), 1177 (m), 1147 (w), 1107 (s), 1050 (s), 1020 (m), 926 (m), 853 (w), 790 (m), 744 (w); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 11.07 (d, *J* = 7.6 Hz, 1H), 8.48 (s, 1H), 8.09 (d, *J* = 8.7 Hz, 2H), 7.97 (s, 1H), 7.38 (d, *J* = 8.7 Hz, 2H), 5.99 (s, 1H), 4.59 (d, *J* = 4.6 Hz, 1H), 4.54-4.47 (m, 2H), 4.46-4.39 (m, 3H), 4.27-4.22 (m, 1H), 4.03 (dd, *J* = 10.5, 9.2 Hz, 1H), 3.97 (s, 3H), 3.09 (t, *J* = 6.6 Hz, 2H), 2.25-2.18 (m, 1H), 1.07 (s, 9H), 1.04 (s, 9H), 0.99 (d, *J* = 6.8 Hz, 3H), 0.95-0.93 (m, 12H), 0.18 (s, 3H), 0.16 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 172.4, 156.0, 153.3, 151.6, 150.0, 146.9, 145.7, 139.3, 129.9, 123.8, 122.8, 92.5, 76.0, 75.6, 74.9, 68.0, 64.4, 59.8, 35.0, 34.7, 30.8, 27.6, 27.1, 26.0, 22.9, 20.5, 19.6, 18.5, 18.2, -4.1, -4.8; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₉H₆₂N₇O₉Si₂ 828.4142; Found 828.4143.

11d: Yield: 76%; *R*_f = 0.47 (5:3 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3230 (w), 2933 (w), 1740 (s), 1690 (s), 1580 (s), 1520 (s), 1469 (s), 1345 (s), 1259 (s), 1134 (s), 1057 (s), 1013 (s), 900 (w), 826 (s), 780 (s), 750 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 10.92 (d, *J* = 6.8 Hz, 1H), 8.47 (s, 1H), 8.09 (d, *J* = 8.7 Hz, 2H), 7.97 (s, 1H), 7.38 (d, *J* = 8.7 Hz, 2H), 6.00 (s, 1H), 4.59 (d, *J* = 4.6 Hz, 1H), 4.57-4.48 (m, 2H), 4.47-4.39 (m, 3H), 4.25 (td, *J* = 10.1, 5.1 Hz, 1H), 4.03 (dd, *J* = 10.5, 9.2 Hz, 1H), 3.96 (s, 3H), 3.08 (t, *J* = 6.6 Hz, 2H), 1.73-1.61 (m, 3H), 1.07 (s, 9H), 1.05 (s, 9H), 0.95-0.94 (m, 12H), 0.93 (s, 3H), 0.18 (s, 3H), 0.16 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 173.4, 155.7, 153.3, 151.6, 150.1, 146.9, 145.7, 139.3, 129.9, 123.8, 122.8, 92.5, 76.0, 75.6, 74.9, 68.0, 64.5, 53.0, 41.2, 35.0, 34.7, 27.6, 27.1, 26.0, 25.3, 23.0, 22.9, 22.1, 20.5, 18.5, -4.1, -4.8; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₄₀H₆₄N₇O₉Si₂ 842.4299; Found 842.4296.

11e: Yield: 73%; *R*_f = 0.34 (4:3 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3237 (w), 2931 (s), 2857 (s), 1737 (s), 1701 (s), 1610 (s), 1520 (s), 1465 (s), 1345 (s), 1250 (s), 1136 (w), 1057 (s), 998 (w), 894 (w), 840 (s), 777 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 11.00 (d, *J* = 8.7 Hz, 1H), 8.43 (s, 1H), 8.01-7.96 (m, 3H), 7.32 (d, *J* = 8.7 Hz, 2H), 6.02 (s, 1H), 4.60 (d, *J* = 4.6 Hz, 1H), 4.58 (dd, *J* = 8.7, 1.7 Hz, 1H), 4.53-4.46 (m, 3H), 4.43-4.30 (m, 2H), 4.29-4.22 (m, 1H), 4.04 (t, *J* = 9.5 Hz, 1H), 3.98 (s, 3H), 3.03 (t, *J* = 6.5 Hz, 2H), 1.23 (d, *J* = 6.2 Hz, 3H), 1.08 (s, 9H), 1.05 (s, 9H), 0.95 (s, 9H), 0.88 (s, 9H), 0.19 (s, 3H), 0.16 (s, 3H), 0.05 (s, 3H), -0.05 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 171.3, 156.4, 153.4, 151.6, 150.2, 146.8, 145.7, 139.4, 129.9, 123.7, 122.8, 92.5, 76.0, 75.7, 74.9, 68.9, 68.0, 64.6, 60.6, 34.9, 27.6, 27.2, 26.0, 25.7, 22.9, 21.3, 20.5, 18.5, 18.0, -4.1, -4.2, -4.9, -5.3; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₄₄H₇₄N₇O₁₀Si₃ 944.4799; Found: 944.4793.

11f: Yield: 73%; *R*_f = 0.30 (8:2 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2933 (w), 1744 (w), 1683 (m), 1583 (m), 1392 (m), 1345 (s), 1166 (m), 1056 (s), 1002 (m), 783 (m); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 8.46 (s, 1H), 8.17-8.15 (m, 2H), 7.84 (s, 1H), 7.39-7.38 (m, 2H), 5.93 (s, 1H), 4.58-4.57 (m, 1H), 4.54 (br s, 2H), 4.49 (dd, *J* = 9.8, 5.1 Hz, 2H), 4.39-4.35 (m, 1H), 4.22 (ddd, *J* = 9.8, 9.8, 5.1 Hz, 1H), 4.03 (dd, *J* = 9.8, 9.8 Hz, 1H), 3.55 (br s, 3H), 3.06 (br s, 2H), 2.18-2.16 (m, 1H), 1.88-1.86 (m, 3H), 1.08 (s, 9H), 0.92 (s, 9H), 0.15 (s, 3H), 0.14 (s, 3H) (some proton signals of proline appeared too broad for an unequivocal assignment); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 172.1, 156.9, 153.3, 152.5, 150.8, 147.0, 145.6, 139.4, 129.9, 123.9, 92.6, 76.0, 75.6, 74.8, 67.9, 64.6, 60.0, 48.0, 35.0, 34.9, 27.6, 27.1, 26.0, 24.2, 22.9, 20.5, 18.5, -4.2, -4.8 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₉H₆₀N₇O₉Si₂ 826.3985; Found 826.3991.

11g: Yield: 85%; *R*_f = 0.50 (2:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2931 (w), 2857 (w), 1738 (w), 1682 (s), 1568 (s), 1518 (s), 1469 (s), 1344 (s), 1261 (s), 1166 (s), 1134 (s), 1056 (s), 1011 (s), 895 (w), 826 (s), 778 (s); ¹H NMR (400 MHz, CD₂Cl₂, 298 K) δ (ppm): 10.87 (d, *J* = 6.9 Hz, 1H), 8.30 (s, 1H), 8.04 (d, *J* = 8.7 Hz, 2H), 7.99 (s, 1H), 7.34 (d, *J* = 8.7 Hz, 2H), 7.29-7.24 (m, 3H), 7.20-7.12 (m, 2H), 6.02 (s, 1H), 4.77 (q, *J* = 6.5 Hz, 1H), 4.59 (d, *J* = 4.6 Hz, 1H), 4.53-4.42 (m, 2H), 4.38 (td, *J* = 6.5, 3.9 Hz, 2H), 4.25 (td, *J* = 10.0, 5.1 Hz, 1H), 4.06 (dd, *J* = 10.5, 9.2 Hz, 1H), 3.88 (s, 3H), 3.13 (dd, *J* = 6.4, 2.1 Hz, 2H), 3.03 (t, *J* = 6.4 Hz, 2H), 1.09 (s, 9H), 1.06 (s, 9H), 0.96 (s, 9H), 0.19 (s, 3H), 0.17 (s, 3H); ¹³C{¹H} NMR (100 MHz, CD₂Cl₂, 298 K) δ (ppm): 172.4, 155.9, 153.5, 152.1, 150.3, 147.3, 146.5, 139.9, 137.2, 130.4, 129.9, 129.0, 127.6, 124.0, 123.1, 92.8, 76.5, 76.1, 75.3, 68.3, 65.0, 56.3, 38.3, 35.3, 34.9, 27.8, 27.4, 26.2, 23.1, 20.8, 18.8, -4.0, -4.7; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₄₃H₆₂N₇O₉Si₂ 876.4142; Found 876.4148.

11h: Yield: 78%; *R*_f = 0.30 (2.5:1 *i*-Hexane/EtOAc); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 11.08 (d, *J* = 7.2 Hz, 1H), 8.49 (s, 1H), 8.09 (d, *J* = 8.6 Hz, 2H), 7.97 (s, 1H), 7.38 (d, *J* = 8.6 Hz, 2H), 6.00 (s, 1H), 4.70 (td, *J* = 7.2, 5.1 Hz, 1H), 4.59 (d, *J* = 4.6 Hz, 1H), 4.52 (dd, *J* = 9.2, 5.1 Hz, 1H), 4.46-4.41 (m, 3H), 4.25 (td, *J* = 10.1, 5.1 Hz, 1H), 4.03 (dd, *J* = 10.5, 9.2 Hz, 1H), 3.96 (s, 3H), 3.09 (t, *J* = 6.5 Hz, 2H), 2.52 (td, *J* = 7.3, 1.9 Hz, 2H), 2.24-1.99 (m, 2H), 2.07 (s, 3H), 1.07 (s, 8H), 1.04 (s, 9H), 0.95 (s, 9H), 0.18 (s, 3H), 0.16 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 172.3, 155.7, 153.2, 151.7, 150.1, 147.0, 145.6, 139.4, 129.9, 123.8, 122.8, 92.5, 76.0, 75.7, 74.9, 68.0, 64.8, 53.5, 35.0, 34.8, 31.7, 30.3, 27.6, 27.1, 26.0, 22.9, 20.5, 18.5, 15.6, -4.1, -4.8; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₉H₆₂N₇O₉Si₂ 860.3863; Found 860.3858.

11i: Yield: 77%; *R*_f = 0.40 (8:2 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2933 (w), 1737 (m), 1683 (m), 1569 (m), 1518 (s), 1344 (s), 1166 (m), 1057 (m), 1000 (m), 780 (m); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 11.32 (d, *J* = 7.5 Hz, 1H), 8.40 (s, 1H), 8.09-8.06 (m, 2H), 8.02-7.99 (m, 2H), 7.98 (s, 1H), 7.34-7.32 (m, 4H), 6.01 (s, 1H), 4.87 (dt, *J* = 7.5, 4.5 Hz, 1H), 4.61 (d, *J* = 4.5 Hz, 1H), 4.52 (dd, *J* = 9.2, 4.5 Hz, 1H), 4.46-4.40 (m, 3H), 4.40-4.23 (m, 3H), 4.04 (dd, *J* = 9.2, 9.2 Hz, 1H), 3.94 (s, 3H), 3.06-2.99 (m, 4H), 2.97-2.95 (m, 2H), 1.07 (s, 9H), 1.05 (s, 9H), 0.95 (s, 9H), 0.19 (s, 3H), 0.16 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 171.0, 170.9, 155.6, 152.9, 151.7, 149.9, 146.9, 146.8, 145.6, 145.5, 139.5, 129.8 (x2), 123.8, 123.7, 122.6, 92.5, 76.0, 75.6, 74.9, 68.0, 65.0, 64.4, 50.5, 36.6, 34.9, 34.8, 34.7, 27.6, 27.1, 26.0, 22.9, 20.5, 18.5, -4.1, -4.9; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₄₆H₆₅O₁₃N₈Si₂ 993.4203; Found 993.4215.

11j: Yield: 74%; *R*_f = 0.34 (4:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3121 (w), 2933 (m), 2896 (w), 2857 (m), 2168 (w), 1692 (s), 1570 (s), 1525 (s), 1469 (s), 1422 (w), 1360 (m), 1328 (m), 1308 (w), 1299 (w), 1278 (m), 1249 (m), 1218 (w), 1198 (w), 1165 (s), 1141 (s), 1111 (m), 1062 (s), 1024 (s), 1001 (s), 968 (w), 889 (m), 825 (s), 784 (s); ¹H NMR (600 MHz, CDCl₃, 298 K) δ (ppm): 11.14 (t, *J* = 5.6 Hz, 1H), 8.53 (s, 1H), 7.99 (s, 1H), 6.01 (s, 1H), 4.56 (d, *J* = 4.6 Hz, 1H), 4.52 (dd, *J* = 9.4, 4.6 Hz, 1H), 4.39 (dd, *J* = 9.4, 4.6 Hz, 1H), 4.34 (dd, *J* = 5.2, 2.0 Hz, 2H), 4.26 (td, *J* = 9.4, 5.2 Hz, 1H), 4.04-4.02 (m, 1H), 4.02 (s, 3H), 1.07 (s, 9H), 1.05 (s, 9H), 0.95 (s, 9H), 0.18 (s, 3H), 0.16 (s, 3H); ¹³C{¹H} NMR (150 MHz, CDCl₃, 298 K) δ (ppm): 156.0, 152.9, 151.9, 150.1, 139.6, 122.8, 116.7, 92.5, 76.1, 75.7, 74.9, 67.9, 35.0, 29.2, 27.6, 27.2, 26.0, 22.9, 20.5, 18.5, -4.1, -4.8; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₈H₄₈N₇O₉Si₂ 618.3250; Found 618.3256.

General procedure for the synthesis of 12a-j:

A solution of the modified adenosine derivative **11a-j** (1.0 equiv.) in DCM/pyridine (9:1 v/v) inside a plastic reaction vessel was cooled to 0°C. Subsequently, a solution of 70% HF-pyridine (5.0 equiv.) was slowly added and the reaction mixture was stirred at 0°C for 2 h. The reaction mixture was diluted with aq. sat. NaHCO₃ solution and extracted three times with DCM. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated under reduced pressure. The crude product was purified by silica gel column chromatography to isolate the 3',5'-deprotected adenosine derivative **12a-j** as a white foam.

12a: Yield: 97%; *R*_f = 0.37 (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2932 (w), 2857 (w), 1738 (w), 1688 (m), 1606 (w), 1581 (m), 1571 (m), 1518 (s), 1471 (m), 1445 (w), 1345 (s), 1253 (m), 1219 (m), 1135 (m), 1083 (m), 1031 (m), 858 (w), 838 (s), 780 (s), 750 (m); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 10.89 (t, *J* = 5.3 Hz, 1H), 8.52 (s, 1H), 8.14 (d, *J* = 8.6 Hz, 2H), 7.96 (s, 1H), 7.39 (d, *J* = 8.6 Hz, 2H), 5.90 (d, *J* = 11.6 Hz, 1H), 5.81 (d, *J* = 7.3 Hz, 1H), 5.12 (dd, *J* = 7.3, 4.9 Hz, 1H), 4.43 (td, *J* = 6.7, 1.7 Hz, 2H), 4.39-4.33 (m, 2H), 4.24-4.10 (m, 2H), 4.03-3.91 (m, 4H), 3.82-3.70 (m, 1H), 3.09 (t, *J* = 6.7 Hz, 2H), 2.81 (s, 1H), 0.80 (s, 9H), -0.18 (s, 3H), -0.39 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 170.2, 155.9, 153.8, 151.2, 149.7, 147.0, 145.4, 141.6, 129.9, 123.9, 123.7, 91.4, 87.7, 74.2, 72.8, 64.8, 63.5, 43.0, 35.0, 25.6, 17.9, -5.2, -5.3; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₈H₄₀N₇O₉Si 646.2651; Found 646.2645.

12b: Yield: 95%; *R*_f = 0.40 (100:5 DCM/IPA); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3191 (w), 2927 (w), 2856 (w), 1739 (m), 1681 (s), 1610 (m), 1568 (s), 1519 (s), 1469 (m), 1344 (s), 1261 (m), 1211 (w), 1143 (w), 1018 (m), 998 (m), 836 (s), 779 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 10.73 (br s, 1H), 8.58 (s, 1H), 8.15 (d, *J* = 8.7 Hz, 2H), 8.11 (s, 1H), 7.41 (d, *J* = 8.7 Hz, 2H), 5.84 (d, *J* = 6.9 Hz, 1H), 5.13-5.06 (m, 1H), 4.66-4.53 (m, 1H), 4.49-4.40 (m, 2H), 4.39-4.35 (m, 2H), 4.00-3.93 (m, 4H), 3.78 (d, *J* = 12.9 Hz, 1H), 3.10 (t, *J* = 6.6 Hz, 2H), 1.47 (d, *J* = 7.2 Hz, 3H), 0.82 (s, 9H), -0.14 (s, 3H), -0.34 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 173.4, 155.1, 153.9, 151.2, 149.7, 147.1, 145.6, 141.5, 130.0, 123.9, 123.8, 91.5, 87.7, 74.2, 72.9, 64.8, 63.5, 50.1, 35.0, 34.9, 25.6, 18.4, 18.0, -5.2, -5.3; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₈H₄₂N₇O₉Si 660.2808; Found: 660.2807.

12c: Yield: 95%; *R*_f = 0.16 (100:3 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3244 (w), 2952 (w), 2929 (w), 2359 (w), 1736 (w), 1681 (m), 1571 (m), 1518 (s), 1469 (m), 1422 (w), 1345 (s), 1255 (m), 1187 (m), 1145 (m), 1089 (m), 1046 (w), 1016 (m), 907 (m), 857 (m), 837 (s), 780 (s), 746 (m); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 11.02 (d, *J* = 7.4 Hz, 1H), 8.54 (s, 1H), 8.15 (d, *J* = 8.6 Hz, 2H), 7.95 (s, 1H), 7.41 (d, *J* = 8.6 Hz, 2H), 5.93 (dd, *J* = 11.9, 1.5 Hz, 1H), 5.81 (d, *J* = 7.4 Hz, 1H), 5.15 (dd, *J* = 7.4, 4.8 Hz, 1H), 4.51-4.34 (m, 5H), 4.02-3.91 (m, 4H), 3.81-3.71 (m, 1H), 3.10 (t, *J* = 6.6 Hz, 2H), 2.81 (s, 1H), 2.29-2.15 (m, 1H), 0.98 (d, *J* = 6.8 Hz, 3H), 0.94 (d, *J* = 6.9 Hz, 3H), 0.81 (s, 9H), -0.16 (s, 3H), -0.37 (s, 3H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 172.4, 155.7, 154.0, 151.2, 149.6, 147.0, 145.6, 141.5, 129.9, 123.9, 123.7, 91.5, 87.7, 74.1, 72.8, 64.6, 63.5, 59.8, 35.0, 34.9, 30.8, 25.6, 19.6, 18.1, 17.9, -5.2, -5.3; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₁H₄₆N₇O₉Si 688.3121; Found 688.3120.

12d: Yield: 98%; *R*_f = 0.52 (9:1 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3244 (w), 2952 (w), 2929 (w), 2856 (w), 1736 (w), 1695 (s), 1610 (s), 1588 (s), 1520 (s), 1469 (s), 1345 (s), 1313 (w), 1250 (s), 1129 (w), 1093 (s), 835 (s), 760 (s); ¹H NMR (400 MHz, acetone-d₆, 298 K) δ (ppm): 10.86 (d, *J* = 7.1 Hz, 1H), 8.63 (s, 1H), 8.58 (s, 1H), 8.13 (d, *J* = 8.7 Hz, 2H), 7.59 (d, *J* = 8.7 Hz, 2H), 6.13 (d, *J* = 5.8 Hz, 1H), 5.04 (dd, *J* = 8.3, 3.7 Hz, 1H), 4.98 (t, *J* = 4.7 Hz, 1H), 4.49-4.41 (m, 3H), 4.40-4.37 (m, 1H), 4.21 (dd, *J* = 2.6 Hz, 1H), 3.98 (d, *J* = 4.0 Hz, 1H), 3.92 (s, 3H), 3.90-

3.87 (m, 1H), 3.80-3.75 (m, 1H), 3.15 (t, $J = 6.4$ Hz, 2H), 1.76-1.53 (m, 3H), 0.91 (dd, $J = 6.4, 3.3$ Hz, 6H), 0.81 (s, 9H), -0.05 (s, 3H), -0.18 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 173.4, 155.9, 154.0, 152.7, 150.5, 147.6, 147.5, 142.5, 131.1, 124.1, 123.6, 90.5, 87.5, 76.6, 72.4, 65.1, 62.8, 53.7, 41.7, 35.3, 34.8, 26.0, 25.8, 23.1, 22.1, 18.6, -4.9, -5.1; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{32}\text{H}_{48}\text{N}_7\text{O}_9\text{Si}$ 702.3277; Found: 702.3279.

12e: Yield: 94%; $R_f = 0.37$ (9:1 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3244 (w), 2952 (w), 2929 (w), 2856 (w), 1736 (w), 1695 (s), 1610 (s), 1588 (s), 1520 (s), 1469 (s), 1345 (s), 1313 (w), 1250 (s), 1129 (w), 1093 (s), 835 (s), 760 (s); ^1H NMR (400 MHz, CDCl_3 , 298 K) δ (ppm): 10.90 (d, $J = 8.6$ Hz, 1H), 8.48 (s, 1H), 8.10 (d, $J = 8.6$ Hz, 2H), 7.96 (s, 1H), 7.37 (d, $J = 8.6$ Hz, 2H), 5.82 (d, $J = 7.3$ Hz, 1H), 5.14 (dd, $J = 7.3, 4.8$ Hz, 1H), 4.56 (dd, $J = 8.6, 1.8$ Hz, 1H), 4.50-4.41 (m, 2H), 4.39-4.35 (m, 2H), 4.29-4.23 (m, 1H), 3.99 (s, 3H), 3.96 (dd, $J = 13.0, 1.8$ Hz, 1H), 3.76 (dd, $J = 13.0, 1.8$ Hz, 1H), 3.06 (t, $J = 6.7$ Hz, 2H), 1.22 (d, $J = 6.2$ Hz, 3H), 0.87 (s, 9H), 0.81 (s, 9H), 0.03 (s, 3H), -0.06 (s, 3H), -0.16 (s, 3H), -0.37 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3 , 298 K) δ (ppm): 171.2, 156.1, 154.1, 151.2, 149.8, 147.0, 145.6, 141.5, 130.0, 123.9, 91.5, 87.7, 74.2, 72.9, 68.8, 64.9, 63.5, 60.7, 35.1, 35.0, 25.7, 25.6, 21.3, 18.0, 17.9, -4.1, -5.2, -5.3, -5.3; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{36}\text{H}_{58}\text{N}_7\text{O}_{10}\text{Si}_2$ 804.3778; Found: 804.3768.

12f: Yield: 82%; $R_f = 0.15$ (98:2 DCM/IPA); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2929 (w), 1743 (w), 1679 (m), 1585 (s), 1519 (m), 1391 (m), 1344 (s), 1090 (m), 1046 (m), 780 (m); ^1H NMR (400 MHz, CDCl_3 , 298 K) δ (ppm): 8.47 (s, 1H), 8.17-8.14 (m, 2H), 7.81 (s, 1H), 7.39 (br s, 2H), 6.35 (d, $J = 12.1$ Hz, 1H), 5.76 (d, $J = 7.4$ Hz, 1H), 5.15 (dd, $J = 7.4, 4.8$ Hz, 1H), 4.61-4.46 (m, 2H), 4.38 (s, 1H), 4.35 (d, $J = 4.8$ Hz, 1H), 3.96 (d, $J = 12.1$ Hz, 1H), 3.75 (dd, $J = 12.1, 12.1$ Hz, 1H), 3.58 (br s, 3H), 3.06 (br s, 2H), 2.80 (s, 1H), 2.18-2.13 (m, 1H), 1.92-1.88 (m, 3H), 0.79 (s, 9H), -0.19 (s, 3H), -0.41 (s, 3H) (some proton signals of proline appeared too broad for an unequivocal assignment); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3 , 298 K) δ (ppm): 171.9, 156.3, 153.7, 152.0, 150.1, 147.0, 141.3, 129.9, 123.9, 91.3, 87.8, 74.1, 73.0, 64.7, 63.5, 59.9, 47.9, 35.0, 34.9, 25.6, 24.2, 17.9, -5.2, -5.4 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{31}\text{H}_{44}\text{N}_7\text{O}_9\text{Si}$ 686.2964; Found 686.2963.

12g: Yield: 90%; $R_f = 0.50$ (98:2 *n*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3391 (w), 3194 (w), 2951 (w), 2855 (w), 1738 (s), 1681 (s), 1568 (s), 1516 (s), 1469 (s), 1344 (s), 1261 (s), 1171 (s), 1128 (s), 1091 (s), 1016 (s), 836 (s), 779 (s); ^1H NMR (400 MHz, CD_2Cl_2 , 298 K) δ (ppm): 10.81 (d, $J = 6.7$ Hz, 1H), 8.31 (s, 1H), 8.11 (d, $J = 8.7$ Hz, 2H), 7.99 (s, 1H), 7.39 (d, $J = 8.7$ Hz, 2H), 7.33-7.20 (m, 3H), 7.17-7.13 (m, 2H), 5.83 (d, $J = 7.3$ Hz, 1H), 5.67 (dd, $J = 11.9, 2.0$ Hz, 1H), 5.10 (dd, $J = 7.3, 4.7$ Hz, 1H), 4.78 (td, $J = 6.8, 5.7$ Hz, 1H), 4.46-4.36 (m, 2H), 4.36-4.32 (m, 2H), 3.93-3.89 (m, 4H), 3.79-3.68 (m, 1H), 3.13 (dd, $J = 6.3, 3.6$ Hz, 2H), 3.09-2.99 (m, 2H), 2.81 (s, 1H), 0.80 (s, 9H), -0.18 (s, 3H), -0.38 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CD_2Cl_2 , 298 K) δ (ppm): 172.3, 155.6, 154.1, 151.7, 150.3, 149.9, 147.4, 146.4, 142.1, 137.1, 130.4, 129.9, 129.1, 127.7, 124.1, 91.7, 88.2, 74.7, 73.3, 65.2, 63.7, 56.3, 38.3, 35.3, 35.0, 25.8, 18.2, -5.1, -5.2; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{39}\text{H}_{46}\text{N}_7\text{O}_9\text{Si}$ 736.3121; Found 736.3118.

12h: Yield: 96%; $R_f = 0.40$ (100:5 DCM/MeOH); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 10.93 (d, $J = 7.1$ Hz, 1H), 8.63 (s, 1H), 8.58 (s, 1H), 8.12 (d, $J = 8.7$ Hz, 2H), 7.59 (d, $J = 8.7$ Hz, 2H), 6.12 (d, $J = 5.8$ Hz, 1H), 5.02 (dd, $J = 8.3, 3.7$ Hz, 1H), 4.97 (dd, $J = 5.9, 4.7$ Hz, 1H), 4.60 (td, $J = 7.5, 5.3$ Hz, 1H), 4.46 (td, $J = 6.5, 1.9$ Hz, 2H), 4.38 (td, $J = 4.4, 2.9$ Hz, 1H), 4.22-4.18 (m, 1H), 3.96 (d, $J = 4.0$ Hz, 1H), 3.92-3.87 (m, 4H), 3.82-3.74 (m, 1H), 3.16 (t, $J = 6.3$ Hz, 2H), 2.52 (t, $J = 7.9$ Hz, 2H), 2.15-1.97 (m, 5H), 0.80 (s, 9H), -0.05 (s, 3H), -0.18 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 172.6, 155.9, 154.0, 152.8, 150.5, 147.7, 147.4, 142.5, 131.1, 124.2, 123.6, 90.4, 87.5, 76.6, 72.4, 65.4, 62.8, 54.2, 35.3, 34.9, 32.2, 30.6, 26.0, 18.6, 15.1, -4.9, -5.1; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{31}\text{H}_{46}\text{N}_7\text{O}_9\text{Si}$ 720.2841; Found 720.2833.

12i: Yield: 89%; $R_f = 0.15$ (97:3 DCM/IPA); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2930 (w), 1735 (m), 1682 (m), 1570 (m), 1516 (s), 1468 (m), 1261 (m), 1018 (m), 837 (m), 781 (m); ^1H NMR (400 MHz, CDCl_3 , 298 K) δ (ppm): 11.31 (d, $J = 7.5$ Hz, 1H), 8.48 (s, 1H), 8.13-8.10 (m, 4H), 7.96 (s, 1H), 7.38-7.33 (m, 4H), 5.87 (d, $J = 12.5$ Hz, 1H), 5.81 (d, $J = 7.3$ Hz, 1H), 5.13 (dd, $J = 7.3, 4.8$ Hz, 1H), 4.87 (dt, $J = 7.3, 4.8$ Hz, 1H), 4.49-4.25 (m, 6H), 3.97 (s, 3H), 3.97-3.94 (m, 1H), 3.73 (dd, $J = 12.5, 12.5$ Hz, 1H), 3.06 (t, $J = 6.7$ Hz, 2H), 3.01 (t, $J = 6.7$ Hz, 2H), 2.99-2.94 (m, 2H), 2.80 (s, 1H), 0.81 (s, 9H), -0.16 (s, 3H), -0.38 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3 , 298 K) δ (ppm): 170.9, 170.8, 155.4, 153.7, 151.3, 150.0, 149.6, 147.0, 145.5, 145.3, 141.7, 129.9, 129.8, 123.9, 123.8, 123.7, 91.5, 87.7, 74.2, 72.8, 65.2, 64.5, 63.5, 50.6, 36.6, 34.9, 34.8 (x2), 25.6, 17.9, -5.2, -5.3; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{38}\text{H}_{49}\text{O}_{13}\text{N}_8\text{Si}$ 853.3182; Found 853.3187.

12j: Yield: 92%; $R_f = 0.44$ (10:1 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3347 (w), 2929 (m), 2857 (m), 1731 (w), 1681 (s), 1570 (s), 1515 (s), 1462 (s), 1422 (m), 1360 (w), 1329 (w), 1262 (s), 1217 (m), 1126 (s), 1035 (s), 994 (m), 901 (m), 866 (m), 835 (s), 779 (s); ^1H NMR (600 MHz, CDCl_3 , 298 K) δ (ppm): 11.02 (t, $J = 5.7$ Hz, 1H), 8.55 (s, 1H), 7.98 (s, 1H), 5.82 (d, $J = 7.4$ Hz, 1H), 5.77 (d, $J = 11.4$ Hz, 1H), 5.13 (dd, $J = 7.4, 4.8$ Hz, 1H), 4.42-4.29 (m, 4H), 4.04 (s, 3H), 4.00-3.92 (m, 1H), 3.82-3.73 (m, 1H), 2.79 (s, 1H), 0.81 (s, 9H), -0.16 (s, 3H), -0.38 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$

NMR (150 MHz, CDCl_3 , 298 K) δ (ppm): 155.7, 153.5, 151.5, 149.6, 141.9, 123.8, 116.6, 110.2, 91.5, 87.7, 74.2, 72.9, 63.5, 35.2, 29.2, 25.6, 18.0, -5.1, -5.3; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{20}\text{H}_{32}\text{N}_7\text{O}_5\text{Si}$ 478.2229; Found 478.2231.

General procedure for the synthesis of 13a-j:

The 3',5'-deprotected adenosine derivative **12a-j** (1.0 equiv.) was dissolved in pyridine and DMTrCl (1.5 equiv.) was added. The reaction mixture was stirred at r.t. for 16 h and afterwards the solvents were removed *in vacuo*. Purification by silica gel column chromatography with an addition of 0.1% pyridine afforded the DMTr-protected adenosine derivative **13a-j** as a white or pale-yellow foam.

13a: Yield: 86%; $R_f = 0.16$ (1:1 *n*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3320 (w), 2929 (w), 2853 (w), 1749 (w), 1681 (m), 1606 (w), 1568 (m), 1510 (s), 1466 (m), 1345 (s), 1300 (w), 1250 (s), 1213 (m), 1176 (s), 1066 (w), 1034 (s), 1005 (w), 916 (w), 856 (m), 834 (s), 782 (m), 699 (m); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 10.85 (t, $J = 5.6$ Hz, 1H), 8.49 (s, 1H), 8.46 (s, 1H), 8.12 (d, $J = 8.7$ Hz, 2H), 7.58 (d, $J = 8.7$ Hz, 2H), 7.53-7.47 (m, 2H), 7.37 (dd, $J = 9.0, 2.3$ Hz, 4H), 7.32-7.19 (m, 3H), 6.86 (dd, $J = 9.0, 2.3$ Hz, 4H), 6.18 (d, $J = 4.4$ Hz, 1H), 5.07 (t, $J = 4.4$ Hz, 1H), 4.54-4.50 (m, 1H), 4.43 (t, $J = 6.4$ Hz, 2H), 4.31-4.26 (m, 1H), 4.11 (d, $J = 5.8$ Hz, 2H), 3.98 (d, $J = 5.8$ Hz, 1H), 3.93 (s, 3H), 3.77 (s, 6H), 3.49-3.43 (m, 2H), 3.13 (t, $J = 6.4$ Hz, 2H), 0.86 (s, 9H), 0.07 (s, 3H), -0.03 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 170.8, 159.6, 156.5, 153.7, 153.2, 150.7, 147.6, 146.1, 141.8, 136.7, 131.1, 131.0, 129.0, 128.6, 127.6, 124.2, 123.2, 113.9, 89.9, 87.1, 84.7, 76.5, 71.9, 65.1, 64.3, 55.5, 43.4, 35.3, 34.8, 26.1, 18.7, -4.6, -4.8; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{49}\text{H}_{58}\text{N}_7\text{O}_{11}\text{Si}$ 948.3958; Found 948.3949.

13b: Yield: 85%; $R_f = 0.70$ (4:1 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2928 (w), 1741 (w), 1681 (w), 1610 (m), 1568 (m), 1508 (s), 1463 (m), 1344 (m), 1251 (s), 1174 (m), 1018 (m), 835 (s), 781 (m); ^1H NMR (400 MHz, CD_2Cl_2 , 298 K) δ (ppm): 10.91 (d, $J = 6.5$ Hz, 1H), 8.44 (s, 1H), 8.17 (s, 1H), 8.10 (d, $J = 8.7$ Hz, 2H), 7.49-7.44 (m, 2H), 7.42 (d, $J = 8.7$ Hz, 2H), 7.35 (d, $J = 8.9$ Hz, 4H), 7.32-7.20 (m, 3H), 6.82 (d, $J = 8.9$ Hz, 4H), 6.08 (d, $J = 4.9$ Hz, 1H), 4.97 (t, $J = 4.9$ Hz, 1H), 4.58-4.45 (m, 1H), 4.45-4.33 (m, 3H), 4.25-4.20 (m, 1H), 3.92 (s, 3H), 3.77 (s, 6H), 3.49 (dd, $J = 10.7, 3.1$ Hz, 1H), 3.39 (dd, $J = 10.7, 4.2$ Hz, 1H), 3.09 (t, $J = 6.5$ Hz, 2H), 2.64 (br s, 1H), 1.44 (d, $J = 7.2$ Hz, 3H), 0.86 (s, 9H), 0.02 (s, 3H), -0.09 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CD_2Cl_2 , 298 K) δ (ppm): 173.7, 155.7, 153.7, 152.8, 150.5, 147.4, 146.5, 145.4, 140.6, 136.2, 130.6, 130.4, 128.6, 128.4, 127.4, 124.0, 123.1, 113.7, 89.2, 87.1, 84.6, 76.1, 71.9, 65.0, 63.9, 55.8, 50.5, 35.4, 34.9, 25.9, 18.5, 18.4, -4.6, -4.9; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{50}\text{H}_{60}\text{N}_7\text{O}_{11}\text{Si}$ 962.4115; Found 962.4128.

13c: Yield: 75%; $R_f = 0.15$ (2:1 *n*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2950 (w), 2850 (w), 1730 (w), 1670 (w), 1607 (m), 1577 (s), 1508 (s), 1464 (w), 1347 (s), 1250 (s), 1177 (s), 1150 (w), 1090 (s), 1035 (m), 981 (w), 913 (s), 866 (s), 839 (s), 701 (s); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 11.03 (d, $J = 7.7$ Hz, 1H), 8.52 (s, 1H), 8.49 (s, 1H), 8.09 (d, $J = 8.7$ Hz, 2H), 7.58 (d, $J = 8.7$ Hz, 2H), 7.52-7.48 (m, 2H), 7.41-7.34 (m, 4H), 7.29-7.16 (m, 3H), 6.86 (dd, $J = 9.0, 2.7$ Hz, 4H), 6.19 (d, $J = 4.3$ Hz, 1H), 5.06 (t, $J = 4.3$ Hz, 1H), 4.54-4.37 (m, 4H), 4.32-4.28 (m, 1H), 3.97 (d, $J = 5.9$ Hz, 1H), 3.93 (s, 3H), 3.77 (s, 6H), 3.48-3.44 (m, 2H), 3.14 (t, $J = 6.3$ Hz, 2H), 2.81-2.80 (m, 2H), 0.98 (d, $J = 6.8$ Hz, 3H), 0.94 (d, $J = 6.8$ Hz, 3H), 0.87 (s, 9H), 0.08 (s, 3H), -0.01 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 172.4, 159.5, 156.2, 153.8, 153.1, 150.5, 147.5, 147.4, 146.1, 141.8, 136.7, 136.6, 131.0, 129.1, 128.9, 128.6, 127.5, 124.1, 123.2, 113.9, 90.0, 87.1, 84.6, 76.5, 71.8, 65.0, 64.3, 60.5, 55.5, 35.3, 34.8, 31.4, 26.1, 19.7, 18.7, 18.4, -4.6, -4.8; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{52}\text{H}_{64}\text{N}_7\text{O}_{11}\text{Si}$ 990.4428; Found 990.4430.

13d: Yield: 72%; $R_f = 0.20$ (2:1 *n*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2950 (w), 2852 (w), 1729 (w), 1670 (w), 1607 (s), 1577 (s), 1508 (s), 1464 (w), 1347 (s), 1250 (s), 1177 (s), 1152 (w), 1091 (s), 1035 (s), 981 (w), 913 (s), 866 (s), 839 (s), 699 (s); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 10.89 (d, $J = 7.1$ Hz, 1H), 8.50 (s, 1H), 8.49 (s, 1H), 8.10 (d, $J = 8.8$ Hz, 2H), 7.58 (d, $J = 8.8$ Hz, 2H), 7.50 (d, $J = 7.2$ Hz, 2H), 7.38 (dd, $J = 9.0, 2.5$ Hz, 4H), 7.31-7.25 (m, 2H), 7.25-7.19 (m, 1H), 6.86 (dd, $J = 9.0, 2.5$ Hz, 4H), 6.18 (d, $J = 4.3$ Hz, 1H), 5.05 (t, $J = 4.3$ Hz, 1H), 4.53-4.37 (m, 4H), 4.29 (dd, $J = 4.3, 4.3$ Hz, 1H), 3.97 (d, $J = 5.9$ Hz, 1H), 3.91 (s, 3H), 3.77 (s, 6H), 3.51-3.42 (m, 2H), 3.14 (t, $J = 6.3$ Hz, 2H), 1.74-1.56 (m, 3H), 0.92 (d, $J = 2.0$ Hz, 3H), 0.92 (d, $J = 2.0$ Hz, 3H), 0.86 (s, 9H), 0.07 (s, 3H), -0.02 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 173.4, 159.6, 156.0, 153.8, 153.1, 150.6, 147.6, 147.5, 146.1, 141.9, 136.7, 131.1, 131.0, 131.0, 129.1, 129.0, 128.6, 127.6, 124.1, 123.3, 113.9, 90.0, 87.1, 84.7, 76.5, 71.9, 65.1, 64.3, 55.5, 53.7, 41.7, 35.3, 34.8, 26.1, 25.8, 23.1, 22.2, 18.7, -4.6, -4.8; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{53}\text{H}_{66}\text{N}_7\text{O}_{11}\text{Si}$ 1004.4584; Found 1004.4579.

13e: Yield: 68%; $R_f = 0.22$ (2:1 *n*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2908 (w), 1757 (w), 1718 (w), 1670 (w), 1608 (w), 1507 (s), 1441 (w), 1294 (w), 1248 (s), 1177 (s), 1090 (s), 1034 (s), 975 (s), 913 (s), 869 (s), 776 (s), 703 (s); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 10.89 (d, $J = 8.7$ Hz, 1H), 8.50 (s, 1H), 8.42 (s, 1H), 8.03 (d, $J = 8.7$ Hz, 2H), 7.52 (d, $J = 8.8$ Hz, 4H), 7.41-7.32 (m, 4H), 7.28 (t, $J = 7.4$ Hz, 2H), 7.25-7.17 (m, 1H), 6.85 (dd, $J = 8.8, 1.8$ Hz, 4H), 6.20 (d, $J = 4.5$ Hz, 1H), 5.11 (t, $J = 4.5$ Hz, 1H), 4.54-4.48 (m, 3H), 4.45-4.31 (m, 2H), 4.31-4.26

(m, 1H), 3.98 (d, $J = 5.7$ Hz, 1H), 3.95 (s, 3H), 3.76 (s, 6H), 3.48 (qd, $J = 10.5$, 4.1 Hz, 2H), 3.12 (t, $J = 6.3$ Hz, 2H), 1.25 (d, $J = 6.3$ Hz, 3H), 0.88 (s, 9H), 0.86 (s, 9H), 0.07 (s, 6H), 0.03 (s, 6H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 171.7, 159.6, 153.1, 150.4, 142.0, 136.7, 131.1, 131.0, 129.0, 128.6, 127.6, 124.1, 123.4, 113.9, 90.0, 87.1, 84.8, 76.4, 71.9, 69.7, 65.5, 64.4, 61.2, 55.5, 35.3, 35.1, 26.1, 26.0, 21.6, 18.7, 18.4, -4.2, -4.6, -4.8, -5.2; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{57}\text{H}_{76}\text{N}_7\text{O}_{12}\text{Si}_2$ 1106.5085; Found: 1106.5103.

13f: Yield: 80%; $R_f = 0.30$ (6:4 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2930 (w), 1743 (w), 1680 (m), 1582 (s), 1509 (m), 1391 (m), 1345 (s), 1249 (s), 1174 (s), 782 (m); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 8.37 (s, 1H), 8.33 (s, 1H), 8.18-8.16 (m, 2H), 7.56 (br s, 2H), 7.49 (d, $J = 7.3$ Hz, 2H), 7.37-7.35 (m, 4H), 7.28 (dd, $J = 7.3$, 7.3 Hz, 2H), 7.21 (t, $J = 7.3$ Hz, 1H), 6.88-6.84 (m, 4H), 6.11 (d, $J = 4.8$ Hz, 1H), 5.10 (dd, $J = 4.8$, 4.8 Hz, 1H), 4.53-4.49 (m, 1H), 4.45-4.38 (m, 2H), 4.25 (dd, $J = 8.2$, 4.8 Hz, 1H), 3.94 (d, $J = 4.8$ Hz, 1H), 3.78 (s, 6H), 3.50-3.39 (m, 6H), 3.12-3.10 (m, 2H), 1.84-1.79 (m, 2H), 1.71 (br s, 1H), 0.83 (s, 9H), 0.04 (s, 3H), -0.08 (s, 3H) (some proton signals of proline appeared too broad for an unequivocal assignment); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 172.6, 159.6, 153.8, 152.9, 152.2, 147.7, 147.3, 146.0, 141.7, 136.7, 136.6, 131.0, 130.9, 129.1, 128.6, 127.6, 124.3, 113.9, 89.5, 87.1, 84.8, 84.7, 76.3, 72.0, 71.9, 65.2, 64.3, 60.7, 55.5, 48.4, 35.3, 34.6, 30.4, 26.1, 18.7, -4.7, -4.9; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{52}\text{H}_{62}\text{N}_7\text{O}_{11}\text{Si}$ 988.4270; Found 988.4280.

13g: Yield: 90%; $R_f = 0.50$ (5:1 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3538 (w), 2953 (w), 2855 (w), 1738 (w), 1681 (w), 1568 (s), 1508 (s), 1463 (s), 1344 (s), 1249 (s), 1174 (s), 1031 (s), 1016 (s), 834 (s), 781 (s); ^1H NMR (400 MHz, CD_2Cl_2 , 298 K) δ (ppm): 10.89 (d, $J = 6.8$ Hz, 1H), 8.22 (s, 1H), 8.15 (s, 1H), 8.07 (d, $J = 8.7$ Hz, 2H), 7.50-7.45 (m, 2H), 7.40-7.16 (m, 12H), 7.14 (dd, $J = 7.3$, 2.1 Hz, 2H), 6.82 (d, $J = 8.9$ Hz, 4H), 6.06 (d, $J = 5.0$ Hz, 1H), 4.97 (t, $J = 5.0$ Hz, 1H), 4.79-4.75 (m, 1H), 4.41-4.35 (m, 3H), 4.26-4.20 (m, 1H), 3.89 (s, 3H), 3.76 (s, 6H), 3.48 (dd, $J = 10.7$, 3.1 Hz, 1H), 3.38 (dd, $J = 10.7$, 4.2 Hz, 1H), 3.12 (d, $J = 6.3$ Hz, 2H), 3.03 (t, $J = 6.5$ Hz, 2H), 2.64 (d, $J = 4.8$ Hz, 1H), 0.86 (s, 9H), 0.02 (s, 3H), -0.10 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CD_2Cl_2 , 298 K) δ (ppm): 172.4, 159.2, 155.9, 153.5, 152.7, 150.3, 150.2, 147.3, 146.5, 145.4, 140.6, 140.1, 137.2, 136.2, 136.2, 130.6, 130.6, 130.4, 129.9, 129.6, 129.0, 128.6, 128.4, 128.3, 128.2, 127.6, 127.4, 124.0, 123.0, 113.7, 113.6, 89.1, 87.1, 84.6, 76.1, 71.9, 65.1, 64.0, 56.3, 55.7, 38.3, 35.3, 34.9, 25.9, 18.4, -4.6, -4.9; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{58}\text{H}_{64}\text{N}_7\text{O}_{11}\text{Si}$ 1038.4428; Found 1038.4447.

13h: Yield: 92%; $R_f = 0.25$ (100:5 DCM/EtOAc); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 10.98 (d, $J = 7.1$ Hz, 1H), 8.50 (s, 1H), 8.49 (s, 1H), 8.14-8.07 (d, $J = 8.9$ Hz, 2H), 7.59 (d, $J = 8.6$ Hz, 2H), 7.50 (d, $J = 7.2$ Hz, 2H), 7.37 (dd, $J = 9.0$, 2.7 Hz, 4H), 7.29 (t, $J = 7.4$ Hz, 2H), 7.22 (t, $J = 7.2$ Hz, 1H), 6.86 (dd, $J = 8.9$, 3.2 Hz, 4H), 6.18 (d, $J = 4.3$ Hz, 1H), 5.07 (t, $J = 4.6$ Hz, 1H), 4.61 (td, $J = 7.5$, 5.3 Hz, 1H), 4.52 (q, $J = 5.4$ Hz, 1H), 4.47 (td, $J = 6.1$, 4.3 Hz, 2H), 4.29 (q, $J = 4.4$ Hz, 1H), 3.98 (d, $J = 5.9$ Hz, 1H), 3.91 (s, 3H), 3.77 (s, 6H), 3.47 (dd, $J = 4.1$, 2.1 Hz, 2H), 3.16 (t, $J = 6.3$ Hz, 2H), 2.53 (td, $J = 7.2$, 1.5 Hz, 2H), 2.17-1.95 (m, 5H), 0.86 (s, 9H), 0.07 (s, 3H), -0.03 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 172.6, 159.6, 159.6, 156.0, 153.7, 153.1, 150.7, 150.6, 147.6, 147.5, 146.1, 141.9, 136.7, 136.7, 131.0, 131.0, 131.0, 129.0, 128.6, 127.6, 124.6, 124.1, 123.3, 113.9, 90.0, 87.1, 84.7, 76.4, 71.9, 65.3, 64.3, 55.5, 54.1, 35.3, 34.8, 32.2, 30.7, 26.1, 18.7, 15.2, -4.6, -4.8; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{52}\text{H}_{64}\text{N}_7\text{O}_{11}\text{Si}$ 1022.4148; Found 1022.4137.

13i: Yield: 80%; $R_f = 0.25$ (95:5 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 2932 (w), 1737 (m), 1682 (m), 1571 (m), 1518 (s), 1464 (s), 1251 (s), 1176 (s), 1018 (m), 836 (s); ^1H NMR (400 MHz, acetone- d_6 , 298 K) δ (ppm): 11.21 (d, $J = 7.5$ Hz, 1H), 8.49 (s, 1H), 8.39 (s, 1H), 8.09-8.04 (m, 4H), 7.55-7.49 (m, 6H), 7.39-7.35 (m, 4H), 7.29 (dd, $J = 7.5$, 7.5 Hz, 2H), 7.21 (t, $J = 7.5$ Hz, 1H), 6.88-6.83 (m, 4H), 6.18 (d, $J = 4.2$ Hz, 1H), 5.04 (dd, $J = 4.2$, 4.2 Hz, 1H), 4.82 (dt, $J = 7.5$, 5.3 Hz, 1H), 4.53 (dd, $J = 4.2$, 4.2 Hz, 1H), 4.44-4.27 (m, 5H), 3.96 (d, $J = 6.1$ Hz, 1H), 3.90 (s, 3H), 3.77 (s, 6H), 3.51-3.44 (m, 2H), 3.11 (t, $J = 6.2$ Hz, 2H), 3.06 (t, $J = 6.2$ Hz, 2H), 2.96-2.94 (m, 2H), 0.86 (s, 9H), 0.08 (s, 3H), -0.01 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K) δ (ppm): 171.5, 171.2, 159.6, 155.9, 153.5, 153.1, 150.4, 147.6, 147.5, 147.4, 147.2, 146.1, 141.9, 136.7, 131.0 (x2), 130.9, 129.0, 128.6, 127.6, 124.1, 124.0, 123.2, 113.9, 90.1, 87.1, 84.5, 76.5, 71.8, 65.6, 65.0, 64.2, 55.5, 51.4, 37.1, 35.3, 35.2, 34.8, 26.1, 18.7, -4.6, -4.8; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{59}\text{H}_{67}\text{O}_{15}\text{N}_6\text{Si}$ 1155.4489; Found 1155.4504.

13j: Yield: 86%; $R_f = 0.21$ (5:2 *n*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm^{-1}): 3397 (w), 2954 (m), 2926 (s), 2854 (m), 2168 (w), 1682 (m), 1607 (m), 1569 (s), 1508 (s), 1462 (s), 1445 (m), 1362 (w), 1297 (w), 1249 (s), 1174 (s), 1134 (m), 1032 (s), 994 (m), 904 (w), 833 (s), 781 (m), 700 (m); ^1H NMR (600 MHz, CDCl_3 , 298 K) δ (ppm): 11.02 (t, $J = 5.7$ Hz, 1H), 8.50 (s, 1H), 8.48 (s, 1H), 7.50 (d, $J = 7.2$ Hz, 2H), 7.38-7.36 (m, 4H), 7.29 (t, $J = 7.6$ Hz, 2H), 7.24-7.21 (m, 1H), 6.87-6.83 (m, 4H), 6.18 (d, $J = 4.6$ Hz, 1H), 5.08 (dd, $J = 4.6$, 4.6 Hz, 1H), 4.54-4.49 (m, 1H), 4.39 (d, $J = 5.7$ Hz, 2H), 4.28 (td, $J = 4.7$, 3.4 Hz, 1H), 3.98 (s, 3H), 3.96 (d, $J = 5.8$ Hz, 1H), 3.78 (s, 6H), 3.48 (dd, $J = 10.2$, 3.8 Hz, 1H), 3.45 (dd, $J = 10.2$, 4.7 Hz, 1H), 0.85 (s, 9H), 0.06 (s, 3H), -0.05 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (150 MHz, CDCl_3 , 298 K) δ (ppm): 159.6, 159.6, 156.5, 153.5, 153.4, 150.7, 146.1, 142.2, 142.2, 136.7, 131.0, 131.0, 129.0, 129.0, 128.6, 127.6, 123.3, 118.2, 113.9, 89.9, 84.8, 76.4, 71.9, 64.4, 55.5, 55.5, 34.9, 29.7, 26.1, 18.7, -4.6, -4.8; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{41}\text{H}_{50}\text{N}_7\text{O}_7\text{Si}$ 780.3535; Found 780.3538.

General procedure for the synthesis of 14a-j:

To a solution of 5'-DMTr-protected adenosine derivative **13a-j** (1.0 equiv.) in anhydrous DCM, *N,N*-diisopropylethylamine (DIPEA) (4.0 equiv.) was added. After cooling down to 0°C, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (CED-Cl) (2.5 equiv.) was added dropwise and the reaction mixture was stirred at r.t. for 5 h. Afterwards aq. sat. NaHCO_3 solution was added to the reaction mixture and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (MgSO_4), filtered and concentrated *in vacuo*. The crude product was purified by silica gel column chromatography with addition of 0.1% pyridine and copolymerized from benzene to afford the desired phosphoramidite **14a-j** as a mixture of diastereoisomers, as a white or pale-yellow foam.

14a: Yield: 85%; $R_f = 0.15$ (2:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.1, 148.7; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{58}\text{H}_{75}\text{N}_9\text{O}_{12}\text{PSi}$ 1148.5037; Found 1148.5052.

14b: Yield: 85%; $R_f = 0.50$ (1:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, CD_2Cl_2 , 298 K) δ (ppm): 150.6, 149.2; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{59}\text{H}_{77}\text{N}_9\text{O}_{12}\text{PSi}$ 1162.5193; Found 1162.5221.

14c: Yield: 77%; $R_f = 0.35$ (1:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.1, 148.7; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{61}\text{H}_{81}\text{N}_9\text{O}_{12}\text{PSi}$ 1190.5506; Found 1190.5492.

14d: Yield: 75%; $R_f = 0.38$ (1:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.1, 148.7; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{62}\text{H}_{83}\text{N}_9\text{O}_{12}\text{PSi}$ 1204.5663; Found 1204.5682.

14e: Yield: 62%; $R_f = 0.43$ (1:1 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.2, 148.5; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{66}\text{H}_{93}\text{N}_9\text{O}_{13}\text{PSi}_2$ 1306.6164; Found 1306.6189.

14f: Yield: 80%; $R_f = 0.30$ (6:4 DCM/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.2, 148.6; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{61}\text{H}_{79}\text{N}_9\text{O}_{12}\text{PSi}$ 1188.5350; Found 1188.5388.

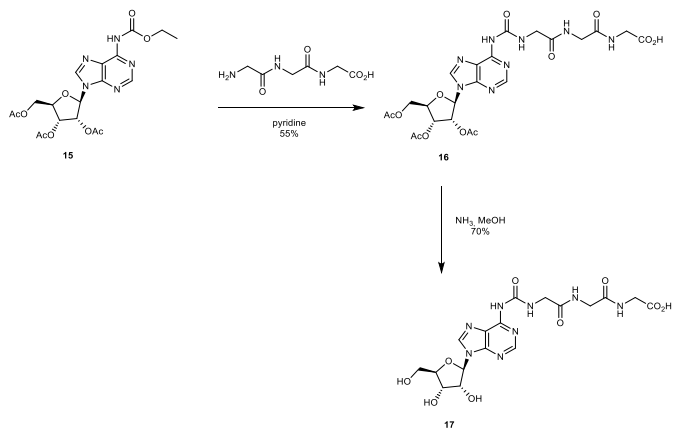
14g: Yield: 90%; $R_f = 0.30$ (5:3 *i*-Hexane/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.7, 149.1; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{65}\text{H}_{81}\text{N}_9\text{O}_{12}\text{PSi}$ 1238.5506; Found 1238.5530.

14h: Yield: 89%; $R_f = 0.30$ (6:4 DCM/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.1, 148.6; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{61}\text{H}_{81}\text{N}_9\text{O}_{12}\text{PSi}$ 1222.5227; Found 1222.5215.

14i: Yield: 65%; $R_f = 0.15$ (92:8 DCM/EtOAc); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.1, 148.7; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{68}\text{H}_{84}\text{O}_{16}\text{N}_{10}\text{PSi}$ 1355.5567; Found 1355.5590.

14j: Yield: 89%; $R_f = 0.39$ (2:1 EtOAc/*i*-Hexane); $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K) δ (ppm): 150.3, 148.6; HRMS (ESI) m/z : [M+H] $^+$ Calcd. for $\text{C}_{59}\text{H}_{67}\text{N}_9\text{O}_{13}\text{PSi}$ 980.4614; Found 980.4614.

2.4 Nucleobase-modified *N*⁶-triglycylcarbamoyl adenosine nucleoside



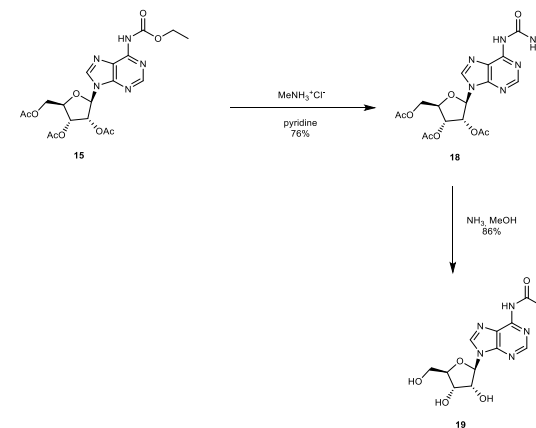
Scheme S4. Synthesis of *N*⁶-triglycylcarbamoyl adenosine **17**.

The compound **15** was synthesized according to a procedure previously described in the literature.^{6,7}

Acetyl protected *N*⁶-triglycylcarbamoyl adenosine **16:** Carbamate derivative **15** (0.25 g, 0.54 mmol, 1.0 equiv.) was dissolved in dry pyridine and **H-Gly-Gly-Gly-OH** (0.20 g, 1.1 mmol, 2.0 equiv.) was added. The mixture was stirred under reflux for 7 h and at r.t. overnight. After that, the crude was filtered and concentrated. The crude was resuspended in toluene and concentrated. Finally, the crude was crystallized from EtOH affording the product **16** as a white solid (0.18 g, 0.30 mmol, 55% yield). IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3353 (w), 1745 (m), 1732 (m), 1697 (m), 1608 (w), 1590 (w), 1515 (s), 1216 (s), 1038 (s), 902 (w); ¹H NMR (400 MHz, DMSO-*d*₆, 298 K) δ (ppm): 12.58 (br s, 1H), 9.94 (s, 1H), 9.66 (t, *J* = 5.3 Hz, 1H), 8.65 (s, 1H), 8.59 (s, 1H), 8.38 (t, *J* = 5.8 Hz, 1H), 8.21 (t, *J* = 5.8 Hz, 1H), 6.30 (d, *J* = 5.4 Hz, 1H), 6.03 (dd, *J* = 5.4, 5.4 Hz, 1H), 5.63 (dd, *J* = 5.4, 5.4 Hz, 1H), 4.44-4.38 (m, 2H), 4.29-4.24 (m, 1H), 3.98 (d, *J* = 5.3 Hz, 2H), 3.77-3.75 (m, 4H), 2.12 (s, 3H), 2.04 (s, 3H), 2.01 (s, 3H); ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K) δ (ppm): 171.2, 170.1, 169.5, 169.3, 169.2 (x2), 153.5, 151.1, 150.5, 150.1, 142.7, 120.5, 85.8, 79.6, 72.0, 70.0, 62.7, 43.1, 41.8, 40.6, 20.5, 20.4, 20.2; HRMS (ESI) *m/z*: [M+Na]⁺ Calcd. for C₂₃H₂₈O₁₂N₆Na 631.1718; Found 631.1721.

***N*⁶-triglycylcarbamoyl adenosine **17**:** Protected adenosine derivative **16** (0.12 g, 0.20 mmol, 1.0 equiv.) was dissolved in 7 N NH₃ in MeOH. The reaction was heated at 40°C for 1.5 h and at r.t. overnight. After that, the crude was concentrated. Finally, the crude product was recrystallized from EtOH (5 mL) affording the product as a white solid (67 mg, 0.14 mmol, 70% yield). IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3281 (m), 2936 (w), 1691 (s), 1658 (s), 1551 (s), 1470 (s), 1240 (s), 1058 (s), 794 (m), 690 (s); ¹H NMR (400 MHz, DMSO-*d*₆, 298 K) δ (ppm): 9.70 (t, *J* = 5.0 Hz, 1H), 8.68 (s, 1H), 8.56 (s, 1H), 8.41 (t, *J* = 5.7 Hz, 1H), 7.83 (t, *J* = 5.0 Hz, 1H), 5.98 (d, *J* = 5.6 Hz, 1H), 4.59 (dd, *J* = 5.6, 5.6 Hz, 1H), 4.18 (dd, *J* = 5.6, 5.6 Hz, 1H), 3.99-3.98 (m, 3H), 3.74 (d, *J* = 5.7 Hz, 2H), 3.71-3.56 (m, 4H); ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K) δ (ppm): 171.2, 169.2, 168.5, 153.6, 150.9, 150.4, 150.3, 142.2, 120.4, 87.7, 85.7, 73.8, 70.3, 61.3, 43.0, 42.1, 42.0; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₇H₂₃O₉N₆ 483.1582; Found 483.1583.

2.5 Nucleobase-modified *N*⁶-methylurea adenosine nucleoside

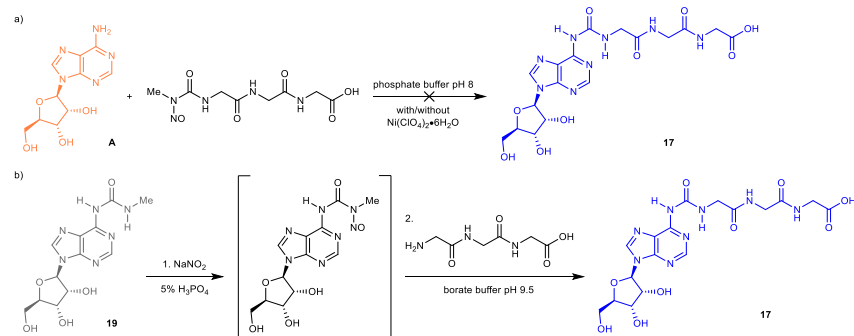


Scheme S5. Synthesis of *N*⁶-methylurea adenosine **19**.

Acetyl protected *N*⁶-methylurea adenosine **18:** Carbamate derivative **15** (0.40 g, 0.86 mmol, 1.0 equiv.) was dissolved in dry pyridine and methylammonium chloride (0.17 g, 2.6 mmol, 3.0 equiv.) was added. The mixture was stirred under reflux overnight. After that, the crude was filtered, washed with EtOAc and concentrated. The crude was suspended in toluene and concentrated. Finally, the crude was purified by silica gel column chromatography (20 g, 95:5 DCM/IPA) affording the product **18** as a white foam (0.30 g, 0.65 mmol, 76% yield). IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3246 (w), 1744 (m), 1699 (m), 1590 (m), 1544 (m), 1469 (w), 1365 (w), 1212 (s), 1046 (m), 797 (w); ¹H NMR (400 MHz, DMSO-*d*₆, 298 K) δ (ppm): 9.74 (s, 1H), 9.20 (c, *J* = 4.6 Hz, 1H), 8.63 (s, 1H), 8.57 (s, 1H), 6.29 (d, *J* = 5.3 Hz, 1H), 6.03 (dd, *J* = 5.3, 5.3 Hz, 1H), 5.64 (dd, *J* = 5.3, 5.3 Hz, 1H), 4.43-4.23 (m, 3H), 2.83 (d, *J* = 4.6 Hz, 3H), 2.12 (s, 3H), 2.04 (s, 3H), 2.01 (s, 3H); ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K) δ (ppm): 170.1, 169.5, 169.3, 153.9, 151.1, 150.5, 150.0, 142.6, 120.3, 85.8, 79.6, 72.0, 70.0, 62.7, 26.3, 20.5, 20.4, 20.2; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₈H₂₃O₉N₆ 451.1571; Found 451.1573.

***N*⁶-methylurea adenosine **19**:** Protected *N*⁶-methylurea adenosine **18** (0.26 g, 0.58 mmol, 1.0 equiv.) was dissolved in 7 N NH₃ in MeOH. The reaction was heated at 40°C for 1.5 h and at r.t. overnight. After that, the crude was concentrated. Finally, the crude was triturated in EtOH, filtered and washed with EtOH affording the product as a white solid (0.16 g, 0.50 mmol, 86% yield). IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3360 (w), 1703 (m), 1584 (m), 1537 (m), 1462 (m), 1297 (m), 1245 (s), 1103 (m), 1057 (m), 795 (m); ¹H NMR (400 MHz, DMSO-*d*₆, 298 K) δ (ppm): 9.57 (br s, 1H), 9.24 (c, *J* = 4.5 Hz, 1H), 8.65 (s, 1H), 8.54 (s, 1H), 5.97 (d, *J* = 5.7 Hz, 1H), 5.53 (d, *J* = 5.7 Hz, 1H), 5.24 (d, *J* = 4.8 Hz, 1H), 5.15 (dd, *J* = 5.7, 5.7 Hz, 1H); 4.62-4.58 (m, 1H), 4.19-4.15 (m, 1H), 3.98-3.95 (m, 1H), 3.71-3.54 (m, 2H), 2.83 (d, *J* = 4.5 Hz, 3H); ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K) δ (ppm): 154.0, 150.8, 150.3, 150.2, 142.2, 120.2, 87.7, 85.7, 73.8, 70.3, 61.3, 26.3; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₂H₁₇O₉N₆ 325.1254; Found 325.1257.

2.6 Nucleobase-modified *N*⁶-triglycylcarbamoyl adenosine nucleoside under prebiotic conditions



Scheme S6. Synthesis of *N*⁶-triglycylcarbamoyl adenosine **17** under prebiotic conditions using: a) nitroso derivative of the *N*-methylurea peptide and b) *N*⁶-methylurea adenosine **19**.

Method A:⁹ Adenosine **A** (2.67 mg, 10 μmol , 1.0 equiv.) was dissolved in 30 mM phosphate buffer pH 8 (370 μL). The nitroso derivative of the *N*-methylurea peptide (5.50 mg, 20 μmol , 2.0 equiv.) was dissolved in water (40 μL) and added to the adenosine's solution. Either water (40 μL) or $\text{Ni}(\text{ClO}_4)_2 \cdot 6\text{H}_2\text{O}$ (91.34 mg, 250 μmol , 25 equiv.) in water (40 μL) was added and the reaction was heated at 70°C for 24 h in a ThermoMixer. Finally, an aliquot (50 μL) of the reaction crude was diluted with water (up to 1 mL), filtered and analyzed by LC-MS (Buffer A: 2 mM HCOONH_4 pH 5.5 in H_2O and buffer B: 2 mM HCOONH_4 pH 5.5 in 20:80 $\text{H}_2\text{O}/\text{MeCN}$; Gradient: 0-20% of B in 30 min; Flow rate = 0.15 $\text{mL} \cdot \text{min}^{-1}$ and Injection: 5 μL).

This prebiotic synthetic method did not afford the *N*⁶-triglycylcarbamoyl adenosine **17**. We only detected the formation of traces of inosine when using the $\text{Ni}(\text{II})$ salt.

Method B: Step 1. *N*⁶-methylurea adenosine **19** (1 mg, 3.08 μmol , 1.0 equiv.) was dissolved in 5% H_3PO_4 in water (140 μL) and cooled to 0°C in an ice bath. NaNO_2 (2.66 mg, 38.54 μmol , 12.5 equiv.) was dissolved in water (10 μL) and added to the previous solution. The reaction was incubated at 0°C for 2 h and -20°C for 22 h. After that, the adenosine's solution was allowed to reach 0°C. Step 2. The peptide (5.83 mg, 30.84 μmol , 10 equiv.) was dissolved in 30 mM borate buffer pH 9.5 (3 mL) and cooled down to 0°C. The adenosine's solution was added to the peptide's solution and the pH was adjusted to 9.5 with 4 N NaOH (60 μL). The reaction was stirred at r.t. for 1 h. Finally, an aliquot (25 μL) of the reaction crude was diluted with water (up to 1 mL), filtered and analyzed by LC-MS (Buffer A: 2 mM HCOONH_4 pH 5.5 in H_2O and buffer B: 2 mM HCOONH_4 pH 5.5 in 20:80 $\text{H}_2\text{O}/\text{MeCN}$; Gradient: 0-20% of B in 30 min; Flow rate = 0.15 $\text{mL} \cdot \text{min}^{-1}$ and Injection: 5 μL).

This prebiotic synthetic method afforded the *N*⁶-triglycylcarbamoyl adenosine **17** in 65% yield. The assignment and amount of the compounds observed in the HPL-chromatogram (Figure S1) was performed by analyzing separate solutions of those synthesized using non-prebiotic methods. Mass spectrometry analyses confirmed the assignments.

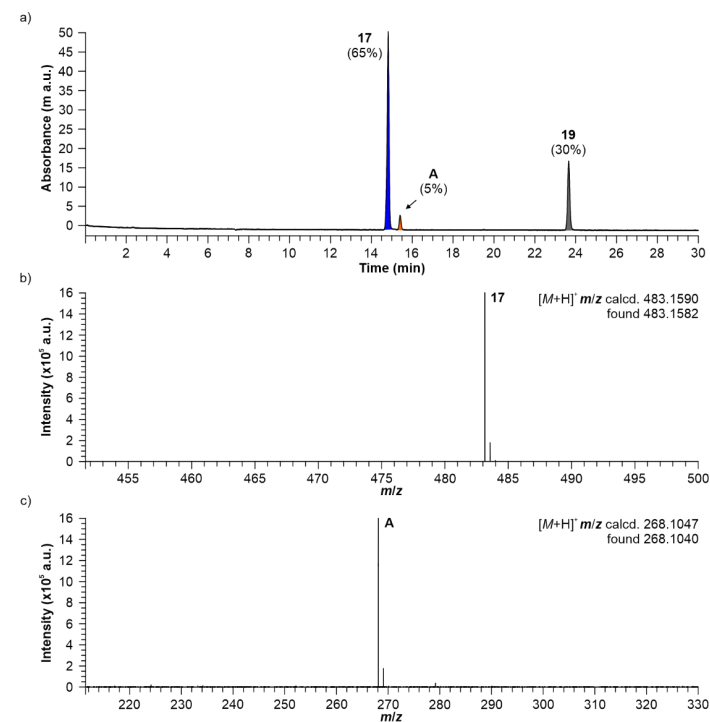
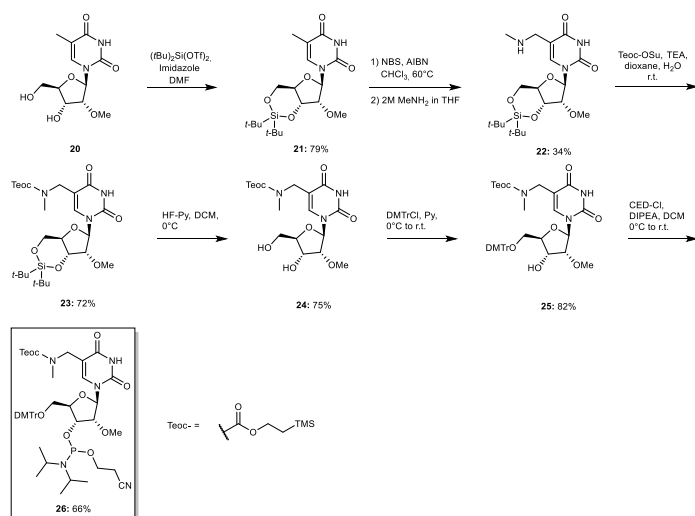


Figure S1. a) HPL-chromatogram of the reaction crude using the **Method B** shown in Scheme S6; mass spectra of the chromatographic peaks observed at: b) 14.8 and c) 15.4 min. The mass spectra confirmed the formation of the compounds **17** and **A**.

2.7 Nucleobase-modified 5-methyluridine 2'-methoxy phosphoramidite



Scheme S7. Synthesis of nucleobase-modified 5-methyluridine 2'-methoxy phosphoramidite.

General procedure for the synthesis of 21:

A suspension of 2'-OMe m⁵U **20** (1.0 equiv.) in DMF was cooled to 0°C. Di-*tert*-butylsilyl bis(trifluoromethanesulfonate) (1.1 equiv.) was added dropwise and the mixture was stirred at r.t. for 30 min. To the reaction was added imidazole (2.5 equiv.) and the resulting solution was stirred at r.t. for 16 h. The crude was concentrated under reduced pressure and the residue was redissolved in EtOAc and washed with water, aq. sat. NaHCO₃ solution and brine. The organic layer was dried (MgSO₄), filtered and concentrated. The crude was purified by silica gel column chromatography to yield **21** as a white foam.

21: Yield: 79%; *R*_f = 0.29 (5:1 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2933 (w), 2859 (w), 1681 (m), 1471 (m), 1365 (w), 1323 (w), 1266 (w), 1148 (m), 1131 (m), 1064 (s), 1038 (m), 1919 (w), 952 (w), 907 (s), 826 (s), 727 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 7.02 (s, 1H), 5.63 (d, *J* = 0.8 Hz, 1H), 4.48-4.45 (m, 1H), 4.07-3.95 (m, 3H), 3.94-3.90 (m, 1H), 3.61 (s, 3H), 1.93 (s, 3H), 1.07 (s, 9H), 1.03 (s, 9H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 163.6, 149.7, 136.1, 111.2, 91.9, 82.2, 77.4, 74.5, 67.4, 59.3, 27.5, 27.2, 22.9, 20.5, 12.8; HRMS (ESI) *m/z* [M+H]⁺ Calcd. for C₁₉H₃₃N₃O₆Si 413.2102; Found 413.2106.

General procedure for the synthesis of 22:

A solution of **21** (1.0 equiv.) in dry CHCl₃ was heated at 60°C. *N*-bromosuccinimide (NBS) (1.2 equiv., previously purified by recrystallization) and azobisisobutyronitrile (AIBN) (0.12 equiv.) were added and the reaction was stirred under reflux for 1.5 h. After that, the reaction mixture was cooled to r.t. and MeNH₂ (2 M in THF, 5.0 equiv.) was added. The resulting suspension was stirred for 2 h at r.t. and, subsequently, it was diluted with aq. sat. NaHCO₃ solution. The crude was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated. The crude was purified by silica gel column chromatography to furnish **22** as a yellow foam.

22: Yield: 34%; *R*_f = 0.30 (9:1 DCM/IPA); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2934 (w), 2859 (w), 1680 (s), 1468 (m), 1245 (s), 1201 (w), 1132 (m), 1064 (m), 1034 (m), 961 (w), 852 (w), 826 (s), 735 (w); ¹H NMR (400 MHz, acetone-*d*₆, 298 K) δ (ppm): 8.02 (s, 1H), 5.81 (s, 1H), 4.43-4.32 (m, 2H), 4.25-4.18 (m, 1H), 4.11 (d, *J* = 5.0 Hz, 1H), 4.08-3.98 (m, 2H), 3.84 (d, *J* = 7.0 Hz, 2H), 3.58 (s, 3H), 2.62 (s, 3H), 1.07 (s, 9H), 1.03 (s, 9H); ¹³C{¹H} NMR (100 MHz, acetone-*d*₆, 298 K) δ (ppm): 163.9, 150.5, 142.6, 107.5, 91.7, 82.9, 77.8, 75.4, 67.8, 59.2, 45.8, 33.3, 27.8, 27.5, 23.1, 20.9; HRMS (ESI) *m/z* [M+H]⁺ Calcd. for C₂₀H₃₆N₃O₆Si 442.2368; Found 442.2370.

General procedure for the synthesis of 23:

To a solution of **22** (1.0 equiv.) in 1,4-dioxane and H₂O (1:1 v/v) were added teoc-OSu (1.1 equiv.) and triethylamine (TEA) (1.5 equiv.). The mixture was stirred at r.t. for 16 h. After that, the crude was diluted with water and extracted three times with Et₂O. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated. The obtained residue was purified by silica gel column chromatography to yield the teoc-protected compound **23** as a white solid.

23: Yield: 72%; *R*_f = 0.53 (95:5 DCM/IPA); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2948 (w), 2894 (w), 2859 (w), 1725 (m), 1464 (w), 1384 (w), 1280 (w), 1245 (s), 1198 (m), 1139 (m), 1057 (m), 1029 (m), 955 (w), 920 (w), 826 (s), 744 (m), 691 (w); For major rotamer: ¹H NMR (400 MHz, acetone-*d*₆, 298 K) δ (ppm): 10.27 (br s, 1H), 7.54 (s, 1H), 5.76 (s, 1H), 4.47 (d, *J* = 4.1 Hz, 1H), 4.27-3.95 (m, 8H), 3.59 (s, 3H), 2.94 (s, 3H), 1.08 (s, 9H), 1.04-1.00 (m, 11H), 0.06 (s, 9H); ¹³C{¹H} NMR (100 MHz, acetone-*d*₆, 298 K) δ (ppm): 150.6, 139.7, 111.1, 91.4, 83.0, 77.7, 75.4, 68.1, 63.8, 59.2, 45.5, 35.3, 27.8, 27.5, 23.2, 20.9, 18.4, -1.3; HRMS (ESI) *m/z* [M+H]⁺ Calcd. for C₂₆H₄₈N₃O₆Si₂ 586.2975; Found 586.2981.

General procedure for the synthesis of 24:

The modified 2'-OMe 5-methyluridine **23** (1.0 equiv.) was dissolved in DCM/pyridine (9:1 v/v) and cooled to 0°C in a plastic reaction vessel. Subsequently, a solution of 70% HF-pyridine (5.0 equiv.) was slowly added, and the reaction mixture was stirred at 0°C for 2 h. The reaction was quenched by adding aq. sat. NaHCO₃ and the crude was extracted three times with DCM. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated. The crude product was purified by silica gel column chromatography to afford the diol compound **24** as a white foam.

24: Yield: 75%; *R*_f = 0.22 (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3060 (w), 2951 (w), 1710 (m), 1463 (m), 1401 (m), 1249 (s), 1214 (m), 1114 (m), 1086 (m), 1062 (m), 988 (w), 938 (w), 838 (s), 769 (m), 694 (w); For major rotamer: ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 10.17 (br s, 1H), 8.09 (s, 1H), 5.99 (d, *J* = 4.3 Hz, 1H), 4.34 (s, 1H), 4.28-4.12 (m, 3H), 4.11-3.92 (m, 5H), 3.92-3.74 (m, 2H), 3.47 (s, 3H), 2.95 (s, 3H), 1.02 (s, 2H), 0.04 (s, 9H) (some proton signals appeared too broad for an unequivocal assignment); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 163.9, 157.2, 151.2, 140.3, 138.6, 111.0, 87.9, 84.4, 69.9, 63.9, 62.1, 58.5, 45.8, 35.3, 18.3, -1.4 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) *m/z* [M+H]⁺ Calcd. for C₁₈H₃₂N₃O₆Si 446.1953; Found 446.1954.

General procedure for the synthesis of 25:

To a solution of the 2'-OMe 3',5'-deprotected 5-methyluridine derivative **24** (1.0 equiv.) in pyridine was added 4,4'-dimethoxytrityl chloride (DMTrCl) (1.5 equiv.). After stirring at r.t. for 16 h, the reaction mixture was concentrated and purified by silica gel column chromatography with an addition of 0.1% of pyridine to the eluent to afford the DMTr-protected compound **25** as a white foam.

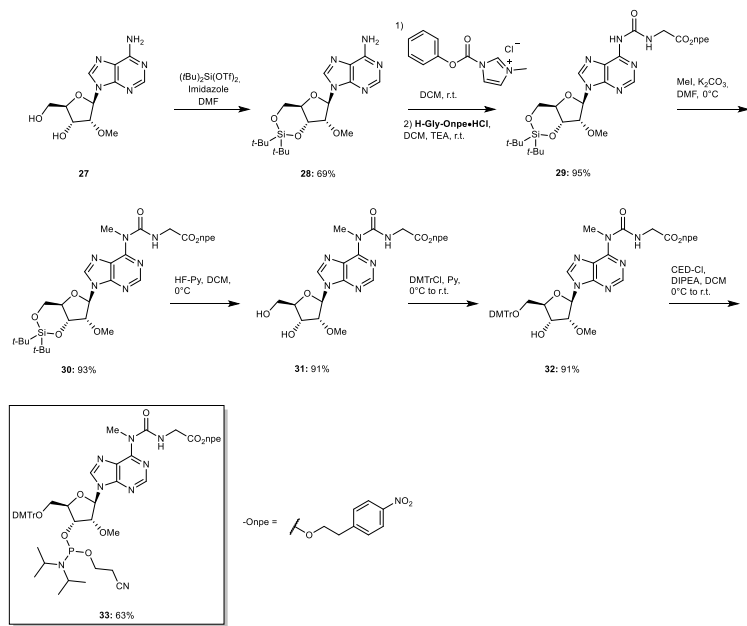
25: Yield: 82%; *R*_f = 0.34 (1:1 DCM/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2953 (w), 1694 (m), 1607 (w), 1508 (m), 1461 (m), 1397 (w), 1344 (w), 1298 (w), 1245 (s), 1175 (m), 1166 (m), 1063 (m), 1032 (s), 962 (w), 832 (s), 756 (w), 726 (w); For major rotamer: ¹H NMR (400 MHz, acetone-*d*₆, 298 K) δ (ppm): 10.20 (br s, 1H), 7.74 (s, 1H), 7.60-7.50 (m, 2H), 7.46-7.38 (m, 4H), 7.32 (t, *J* = 7.8 Hz, 2H), 7.25-7.20 (m, 1H), 6.89 (d, *J* = 8.9 Hz, 4H), 5.95 (s, 1H), 4.45-4.21 (m, 1H), 4.17-3.90 (m, 4H), 3.87-3.66 (m, 8H), 3.59-3.37 (m, 5H), 2.88 (s, 3H), 1.00-0.91 (m, 2H), 0.02 (s, 9H) (some proton signals appeared too broad for an unequivocal assignment); ¹³C{¹H} NMR (100 MHz, acetone-*d*₆, 298 K) δ (ppm): 163.9, 159.6, 156.8, 151.0, 146.1, 140.4, 136.9, 131.1, 129.1, 128.7, 127.5, 114.0, 88.5, 87.2, 84.1, 70.3, 64.5, 63.6, 58.7, 55.5, 46.3, 35.6, 18.4, -1.4 (some carbon signals appeared too broad for an unequivocal assignment); HRMS (ESI) *m/z* [M-H]⁻ Calcd. for C₃₆H₅₀N₃O₁₀Si 746.3114; Found 746.3113.

General procedure for the synthesis of phosphoramidite 26:

A solution of 5'-DMTr-protected compound **25** (1.0 equiv.) and DIPEA (4.0 equiv.) in dry DCM was cooled to 0°C. To this solution was slowly added 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (CED-Cl) (2.5 equiv.) and the reaction mixture was stirred at r.t. for 5 h. The reaction was quenched by addition of aq. sat. NaHCO₃ and the crude was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. After purification by silica gel column chromatography with an addition of 0.1% pyridine and co-lyophilization from benzene the desired phosphoramidite **26** was obtained as a mixture of diastereoisomers and rotamers as a white foam.

26: Yield: 66%; *R*_f = 0.19 (1:1 DCM/EtOAc); ³¹P{¹H} NMR (162 MHz, acetone-*d*₆, 298 K) δ (ppm): 150.0, 149.9, 149.8, 149.7; HRMS (ESI) *m/z* [M-H]⁻ Calcd. for C₄₈H₆₇N₅O₁₁PSi 948.4338; Found 948.4333.

2.8 Nucleobase-modified 2'-methoxy *N*⁶-carbamoyl adenosine phosphoramidite



Scheme S8. Synthesis of nucleobase-modified 2'-methoxy *N*⁶-carbamoyl adenosine phosphoramidite.

General procedure for the synthesis of 28:

A suspension of 2'-OMe adenosine **27** (1.0 equiv.) in DMF was cooled to 0°C. Di-*tert*-butylsilyl *bis*(trifluoromethanesulfonate) (1.1 equiv.) was added dropwise and the mixture was stirred at r.t. for 30 min. To the reaction was added imidazole (2.5 equiv.) and the resulting solution was stirred at r.t. for 16 h. The reaction was concentrated under reduced pressure and the residue was redissolved in EtOAc and washed with water, aq. sat. NaHCO₃ solution and brine. The organic layer was dried (MgSO₄), filtered and concentrated. The crude was purified by silica gel column chromatography to yield **28** as a white foam.

28: Yield: 69%; *R*_f = 0.21 (100:1 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3319 (w), 3161 (m), 2933 (w), 1669 (s), 1600 (s), 1472 (s), 1367 (m), 1328 (m), 1260 (m), 1207 (m), 1133 (s), 1069 (s), 1027 (s), 966 (s), 907 (w), 829 (s), 739 (s), 653 (s); ¹H NMR (400 MHz, DMSO-*d*₆, 298 K) δ (ppm): 8.32 (s, 1H), 8.13 (s, 1H), 7.36 (s, 2H), 6.01 (s, 1H), 4.89 (dd, *J* = 9.0, 4.8 Hz, 1H), 4.34 (d, *J* = 4.8 Hz, 1H), 4.31 (d, *J* = 4.8 Hz, 1H), 4.03-3.94 (m, 2H), 3.54 (s, 3H), 1.08 (s, 9H); ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K) δ (ppm): 156.6, 153.2, 149.2, 140.4, 119.6, 88.6, 82.1, 76.9, 74.5, 67.3, 58.8, 27.7, 27.4, 22.7, 20.4; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₉H₃₂N₅O₅Si 422.2218; Found 422.2220.

General procedure for the synthesis of 29:

To a solution of silyl-protected 2'-OMe adenosine **28** (1.0 equiv.) in DCM was added 1-*N*-methyl-3-phenoxy-carbonyl-imidazolium chloride (2.0 equiv.). The resulting suspension was stirred at r.t. for 16 h and then **H-aa-Omppe-HCl** (2.0 equiv.) together with NEt₃ (2.0 equiv.) was added. After stirring for 16 h, the reaction mixture was quenched by the addition of aq. sat. NaHCO₃ and the crude was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated *in vacuo*. Purification by silica gel column chromatography furnished the amino acid-modified adenosine derivative **29** as a white foam.

29: Yield: 95%; *R*_f = 0.23 (100:1 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3235 (w), 2934 (w), 2856 (m), 1747 (m), 1702 (s), 1587 (m), 1518 (s), 1467 (s), 1343 (s), 1257 (m), 1187 (s), 1138 (s), 1062 (s), 1014 (m), 825 (s), 736 (m), 651 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 9.99 (t, *J* = 5.6 Hz, 1H), 8.77 (s, 1H), 8.51 (s, 1H), 8.26 (s, 1H),

8.08 (d, *J* = 8.7 Hz, 2H), 7.38 (d, *J* = 8.7 Hz, 2H), 6.01 (s, 1H), 4.65 (dd, *J* = 9.6, 4.6 Hz, 1H), 4.50-4.38 (m, 3H), 4.27 (d, *J* = 4.6 Hz, 1H), 4.22-4.14 (m, 3H), 4.05 (dd, *J* = 9.6, 9.6 Hz, 1H), 3.69 (s, 3H), 3.09 (t, *J* = 6.6 Hz, 2H), 1.09 (s, 9H), 1.06 (s, 9H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 170.0, 154.3, 151.3, 150.3, 149.9, 146.9, 145.5, 142.1, 129.9, 123.8, 121.1, 89.7, 82.4, 77.3, 74.9, 67.6, 64.7, 59.5, 42.2, 35.0, 27.5, 27.2, 22.9, 20.5; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₀H₄₂N₇O₅Si 672.2808; Found 672.2808.

General procedure for the synthesis of 30:

The amino acid-modified 2'-OMe adenosine derivative **29** (1.0 equiv.) was dissolved in DMF and cooled to 0°C. To the solution were added K₂CO₃ (3.0 equiv.) together with MeI (2.0 equiv.) and the reaction was stirred at r.t. for 2 h. The reaction mixture was diluted with H₂O and extracted three times with EtOAc. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated. The obtained residue was purified by silica gel column chromatography to give **30** as a white foam.

30: Yield: 93%; *R*_f = 0.32 (1:1 *i*-Hexane/EtOAc); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2932 (w), 1857 (w), 1746 (m), 1682 (s), 1567 (s), 1517 (s), 1467 (s), 1343 (s), 1266 (m), 1192 (m), 1135 (s), 1062 (s), 1027 (s), 826 (s), 735 (m), 651 (s); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 10.95 (t, *J* = 5.4 Hz, 1H), 8.51 (s, 1H), 8.10 (d, *J* = 8.7 Hz, 2H), 7.98 (s, 1H), 7.37 (d, *J* = 8.7 Hz, 2H), 6.02 (s, 1H), 4.62-4.54 (m, 1H), 4.48 (dd, *J* = 9.2, 5.0 Hz, 1H), 4.43 (t, *J* = 6.6 Hz, 2H), 4.27-4.12 (m, 4H), 4.03 (d, *J* = 10.5 Hz, 1H), 3.98 (s, 3H), 3.69 (s, 3H), 3.08 (t, *J* = 6.6 Hz, 2H), 1.09 (s, 9H), 1.05 (s, 9H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 170.3, 156.2, 153.2, 151.7, 150.3, 147.0, 145.6, 139.6, 129.9, 123.8, 122.8, 89.7, 82.3, 77.3, 74.8, 67.6, 64.6, 59.5, 43.0, 35.0, 34.8, 27.5, 27.2, 22.9, 20.5; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₁H₄₄N₇O₅Si 686.2964; Found 686.2967.

General procedure for the synthesis of 31:

A solution of the modified 2'-OMe adenosine derivative **30** (1.0 equiv.) in DCM/pyridine (9:1 v/v) inside a plastic reaction vessel was cooled to 0°C. Subsequently, a solution of 70% HF-pyridine (5.0 equiv.) was slowly added and the reaction mixture was stirred at 0°C for 2 h. The reaction mixture was diluted with aq. sat. NaHCO₃ solution and extracted three times with DCM. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated under reduced pressure. The crude product was purified by silica gel column chromatography to isolate the 3',5'-deprotected adenosine derivative **31** as a white foam.

31: Yield: 91%; *R*_f = 0.25 (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 3201 (w), 2935 (w), 1743 (m), 1677 (m), 1568 (s), 1514 (s), 1464 (m), 1343 (s), 1268 (m), 1209 (m), 1110 (m), 1036 (m), 856 (m), 795 (s), 697 (m), 645 (m); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 10.85 (t, *J* = 5.4 Hz, 1H), 8.51 (s, 1H), 8.13 (d, *J* = 8.8 Hz, 2H), 8.01 (s, 1H), 7.39 (d, *J* = 8.8 Hz, 2H), 5.94-5.91 (m, 2H), 4.72 (dd, *J* = 7.4, 4.7 Hz, 1H), 4.60 (d, *J* = 4.7 Hz, 1H), 4.43 (t, *J* = 6.6 Hz, 2H), 4.37 (d, *J* = 1.0 Hz, 1H), 4.25-4.09 (m, 2H), 4.01 (s, 3H), 4.00-3.92 (m, 1H), 3.84-3.74 (m, 1H), 3.37 (s, 3H), 3.09 (t, *J* = 6.6 Hz, 2H), 2.69 (d, *J* = 1.7 Hz, 1H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 170.2, 156.0, 153.8, 151.2, 149.6, 147.0, 145.5, 141.6, 129.9, 123.9, 123.9, 89.7, 88.2, 82.3, 70.6, 64.7, 63.4, 59.0, 43.1, 35.0; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₃H₂₈N₇O₅ 546.1943; Found 546.1943.

General procedure for the synthesis of 32:

The 3',5'-deprotected 2'-OMe adenosine derivative **31** (1.0 equiv.) was dissolved in pyridine and DMTrCl (1.5 equiv.) was added. The reaction mixture was stirred at r.t. for 16 h and afterwards the solvents were removed *in vacuo*. Purification by silica gel column chromatography with an addition of 0.1% pyridine afforded the DMTr-protected adenosine derivative **32** as a pale-yellow foam.

32: Yield: 91%; *R*_f = 0.45 (100:5 DCM/MeOH); IR (ATR) $\tilde{\nu}$ (cm⁻¹): 2358 (w), 1682 (m), 1568 (m), 1509 (s), 1463 (m), 1344 (s), 1249 (m), 1174 (m), 1033 (s), 701 (w), 667 (w); ¹H NMR (400 MHz, CDCl₃, 298 K) δ (ppm): 10.84 (t, *J* = 5.4 Hz, 1H), 8.48-8.42 (m, 2H), 8.10 (d, *J* = 8.6 Hz, 2H), 7.57 (d, *J* = 8.6 Hz, 2H), 7.48 (d, *J* = 7.4 Hz, 2H), 7.39-7.32 (m, 4H), 7.28 (t, *J* = 7.4 Hz, 2H), 7.24-7.20 (m, 1H), 6.91-6.78 (m, 4H), 6.27 (d, *J* = 4.0 Hz, 1H), 4.74-4.64 (m, 1H), 4.59 (t, *J* = 4.5 Hz, 1H), 4.43 (t, *J* = 6.4 Hz, 2H), 4.28-4.21 (m, 2H), 4.11 (d, *J* = 5.6 Hz, 2H), 3.92 (s, 3H), 3.77 (s, 6H), 3.53 (s, 3H), 3.45 (d, *J* = 4.6 Hz, 2H), 3.13 (t, *J* = 6.4 Hz, 2H); ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K) δ (ppm): 170.7, 159.6, 156.5, 153.7, 153.0, 150.7, 147.6, 147.4, 146.0, 141.7, 136.7, 131.0, 130.9, 129.0, 128.6, 127.6, 124.1, 123.2, 113.8, 87.6, 87.1, 84.9, 83.8, 70.6, 65.0, 64.3, 58.8, 55.5, 43.4, 35.3, 34.8; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₄₄H₄₆N₇O₁₁ 848.3249; Found 848.3234.

General procedure for the synthesis of 33:

To a solution of 5'-DMTr-protected 2'-OMe adenosine derivative **32** (1.0 equiv.) in anhydrous DCM, *N,N*-diisopropylethylamine (DIPEA) (4.0 equiv.) was added. After cooling down to 0°C, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (CED-Cl) (2.5 equiv.) was added dropwise and the reaction mixture was stirred

at r.t. for 5 h. After that, aq. sat. NaHCO₃ solution was added to the reaction mixture and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated *in vacuo*. The crude product was purified by silica gel column chromatography with addition of 0.1% pyridine and co-lyophilized from benzene to afford the desired phosphoramidite **33** as a mixture of diastereoisomers and as a white foam.

33: Yield: 63%; *R*_f = 0.25 (1:1 *i*-Hexane/EtOAc); ³¹P{¹H} NMR (162 MHz, acetone-*d*₆, 298 K) δ (ppm): 150.2, 149.7; HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₃H₆₃N₉O₁₂P 1048.4328; Found 1048.4309.

3. General information and instruments for oligonucleotides

3.1 Synthesis and purification of oligonucleotides

Phosphoramidites of canonical ribonucleosides (Bz-A-CE, Dmf-G-CE, Ac-C-CE and U-CE) were purchased from LinkTech and Sigma-Aldrich. Oligonucleotides (ONs) were synthesized on a 1 μmol scale using RNA SynBase™ CPG 1000/110 and High Load Glen UnySupport™ as solid supports for strands containing amino acid-modified carbamoyl adenosine and 5-(methyl)aminomethyl uridine derivatives, respectively, using an RNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with a standard phosphoramidite chemistry. ONs were synthesized in DMT-OFF mode using DCA as a deblocking agent in CH₂Cl₂, BTT or Activator 42® as activator in MeCN, Ac₂O as capping reagent in pyridine/THF and I₂ as oxidizer in pyridine/H₂O.

Deprotection of npe and teoc groups

For the deprotection of the *para*-nitrophenylethyl (npe) group in ONs containing amino acid-modified carbamoyl adenosine derivatives, the solid support beads were suspended in a 9:1 THF/DBU solution mixture (1 mL) and incubated at r.t. for 2 h.⁹ After that, the supernatant was removed and the beads were washed with THF (3x1 mL).

For the deprotection of the 2-(trimethylsilyl)ethoxycarbonyl (teoc) group in ONs containing 5-(methyl)aminomethyl uridine derivatives, the solid support beads were suspended in a saturated solution of ZnBr₂ in 1:1 MeNO₂/IPA (1 mL) and incubated at r.t. overnight.¹⁰ After that, the supernatant was removed and the beads were washed with 0.1 M EDTA in water (1 mL) and water (1 mL).

Cleavage from beads, deprotection of TBS groups and precipitation of the synthesized ON

The solid support beads were suspended in a 1:1 aqueous solution mixture (0.6 mL) of 30% NH₄OH and 40% MeNH₂. The suspension was heated at 65°C (8 min for SynBase™ CPG 1000/110 and 60 min for High Load Glen UnySupport™). Subsequently, the supernatant was collected and the beads were washed with water (2x0.3 mL). The combined aqueous solutions were concentrated under reduced pressure using a SpeedVac concentrator. After that, the crude was dissolved in DMSO (100 μL) and triethylamine trihydrofluoride (125 μL) was added. The solution was heated at 65°C for 1.5 h. Finally, the ON was precipitated by adding 3 M NaOAc in water (25 μL) and *n*-butanol (1 mL). The mixture was kept at -80°C for 2 h and centrifuged at 4°C for 1 h. The supernatant was removed and the white precipitate was lyophilized.

Purification of the synthesized ON by HPLC and desalting

The crude was purified by semi-preparative HPLC (1260 Infinity II Manual Preparative LC System from Agilent equipped with a G7114A detector) using a reverse-phase (RP) VP 250/10 Nucleodur 100-5 C18ec column from Macherey-Nagel. Buffers: A) 0.1 M AcOH/Et₃N in H₂O at pH 7 and B) 0.1 M AcOH/Et₃N in 80% (v/v) MeCN in H₂O. Gradient: 0-25% of B in 45 min. Flow rate = 5 mL·min⁻¹. The purified ON was analyzed by RP-HPLC (1260 Infinity II LC System from Agilent equipped with a G7165A detector) using an EC 250/4 Nucleodur 100-3 C18ec from Macherey-Nagel. Gradient: 0-30% or 0-40% of B in 45 min. Flow rate = 1 mL·min⁻¹. Finally, the purified ON was desalted using a C18 RP-cartridge from Waters.

Determination of the concentration and the mass of the synthesized ON

The absorbance of the synthesized ON in H₂O solution was measured using an IMPLEN NanoPhotometer® N60/N50 at 260 nm. The extinction coefficient of the single stranded ONs was calculated using the OligoAnalyzer Version 3.0 from Integrated DNA Technologies. For ONs incorporating non-canonical bases, the extinction coefficients were assumed to be identical to those containing only canonical counterparts.

The synthesized ON (2-3 μL) was desalted on a 0.025 μm VSWP filter (Millipore), co-crystallized in a 3-hydroxyisobutyric acid matrix (HPA, 1 μL) and analyzed by MALDI-TOF mass spectrometry (negative mode).

3.2 Analysis of coupling and cleavage reactions by HPLC and MALDI-TOF mass spectrometry

The crudes of the coupling and cleavage reactions were analyzed by RP-HPLC using an EC 250/4 Nucleodur 100-3 C18ec column from Macherey-Nagel. Buffers: A) 0.1 M AcOH/Et₃N in H₂O at pH 7 and B) 0.1 M AcOH/Et₃N in 80% (v/v) MeCN in H₂O. Gradient: 0-40% of B in 45 min. Flow rate = 1 mL·min⁻¹. Injection: 20 μL (1 nmol). The same HPLC method was used for the purification of the products obtained in the coupling and cleavage reactions. The yields of the reactions were calculated by integration of the chromatographic peaks of the products and the use of the calibration curves of the corresponding canonical ONs (see Section 5). In order to simplify the calculations, we assumed that the formed products and the canonical oligonucleotides used for calibration featured identical extinction coefficients, which were calculated for single stranded RNAs. It is expected that double strands and/or secondary structures are disrupted under the HPLC conditions used.

The crudes of the reactions and the isolated products (2-3 μL) were desalted on a 0.025 μm VSWP filter (Millipore), co-crystallized in a 3-hydroxyisobutyric acid matrix (HPA, 1 μL) and analyzed by MALDI-TOF mass spectrometry (negative mode).

3.3 Coupling of amino acids and peptides to ONs anchored to the solid support beads

Oligonucleotides (ONs) were synthesized on a 4 μmol scale using the High Load Glen UnySupport™ for strands containing glycine-modified carbamoyl adenosine and 5-valine-methylaminomethyl uridine derivatives using an RNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with a standard phosphoramidite chemistry. The npe and teoc protecting groups were removed as described in Section 3.1 and the solid support beads were dried using a SpeedVac concentrator.

The solid support beads (1 μmol) in an Eppendorf tube were washed with dry DMF (0.3 mL). In a separate Eppendorf tube, Boc-protected amino acid (for altering of the mnm⁵U derivatives), npe-protected amino acid (for altering of the m⁶g⁶A derivatives) or protected peptide (100 μmol), DMTMM·BF₄ (100 μmol) as activator and dry DIPEA (200 μmol) were dissolved in dry DMF (0.6 mL). Subsequently, the amino acid or peptide solution was added to the solid support beads and the reaction was incubated in an orbital shaker at r.t. for 1 h. The suspension was centrifuged and the supernatant was removed. The solid support beads were washed with dry DMF (2x0.3 mL) and dry MeCN (2x0.3 mL). Finally, the beads were dried using a SpeedVac concentrator.

For the deprotection of the *tert*-butyloxycarbonyl (Boc) group in ONs after the coupling of a Boc-protected amino acid or peptide, the solid support beads were suspended in a 1:1 TFA/CH₂Cl₂ solution mixture (0.5 mL) and incubated for 5 min at r.t.¹¹ After that, the supernatant was removed and the solid support beads were washed with CH₂Cl₂ (2x0.5 mL). The deprotection of the npe-protected adenosine derivatives was performed as described in Section 3.1.

The ONs containing 5-peptide-methylaminomethyl uridine derivatives were cleaved from the solid support beads using a 1:1 aqueous solution mixture (0.6 mL) of 30% NH₄OH and 40% MeNH₂ at 65°C for 60 min. The ONs containing peptide-modified carbamoyl adenosine derivatives were cleaved from the solid support beads using a 30% NH₄OH aqueous solution (0.6 mL) at r.t. overnight. The following work-up and purification steps were identical to those described in Section 3.1. Based on HPLC analyses, we calculated that the coupling reaction using the solid support beads and DMTMM·BF₄ as activator proceeded in an extent larger than 70%.

4. Synthesized oligonucleotides using a DNA/RNA automated synthesizer

4.1 Canonical oligonucleotides (CON)

RNA sequences:

CON1: 5'-AAU CGC U-3'

CON2: 5'-GUA CAG CGA UU-3'

CON3: 5'-GUA CAG CGA UUA AUC GCU-3'

CON4: 5'-AmAmUm CmGmCm Um-3'

CON5: 5'-GmUmCm AmGmUm AmCmAm GmCmGm AmUmUm-3'

CON6: 5'-GmUmCm AmGmUm AmCmAm GmCmGm AmUmUm AmAmUm CmGmCm Um-3'

Table S1. HPLC retention times (0-30% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of canonical oligonucleotides.

Strand	t_R (min)	m/z calcd. for $[M-H]^-$	found
CON1	23.6	2162.3	2162.0
CON2	23.1	3487.5	3486.9
CON3	23.9	5712.8	5711.7

Table S2. HPLC retention times (0-40% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of canonical oligonucleotides.

Strand	t_R (min)	m/z calcd. for $[M-H]^-$	found
CON4	23.3	2261.6	2260.1
CON5	18.8	4772.7	4772.8
CON6	18.6	6998.0	6995.1

The sequences of **CON1-6** are similar to those of the modified ONs used in the coupling reactions. These canonical ONs were used for the development of HPLC calibration curves in Section 5.

4.2 Donor oligonucleotides (ON1) with a complementary sequence

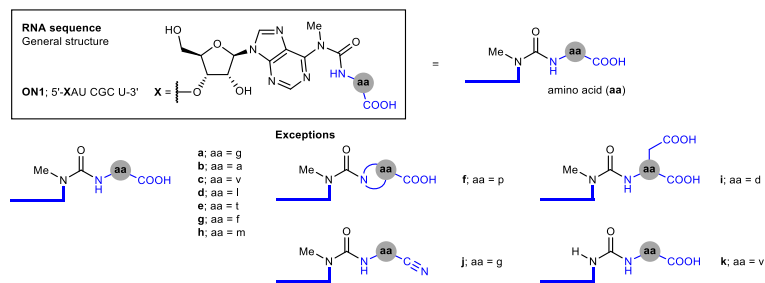


Figure S2. RNA sequence and general structure of amino acid-modified carbamoyl adenosine derivatives.

Other RNA donor strands with longer sequences:

ON1i: 5'-XCU AUU GAG U-3'; X = m⁶v⁶A

ON1m: 5'-X¹AU CGC UGU ACC CUA UUG AGU X²-3'; X¹ = m⁶v⁶A; X² = m⁶g⁶A

ON1n: 5'-XAU CGC UGU AC-3'; X = m⁶v⁶A

ON1o: 5'-XAmUm CmGmCm Um-3'; X = m⁶g⁶Am

ON1p: 5'-XAmUm CmGm-3'; X = m⁶g⁶Am

ON1q: 5'-XAmUm-3'; X = m⁶g⁶Am

Table S3. HPLC retention times (0-40% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of **ON1**.

Strand	t_R (min)	m/z calcd. for $[M-H]^-$	found
ON1a: X = m ⁶ g ⁶ A	18.8	2277.4	2278.4
ON1b: X = m ⁶ g ⁶ A	20.2	2291.4	2290.0
ON1c: X = m ⁶ v ⁶ A	22.2	2319.4	2317.8
ON1d: X = m ⁶ v ⁶ A	24.3	2333.4	2331.6
ON1e: X = m ⁶ v ⁶ A	18.9	2321.4	2320.0
ON1f: X = m ⁶ g ⁶ A	18.0	2317.4	2316.8
ON1g: X = m ⁶ v ⁶ A	24.5	2368.6	2365.4
ON1h: X = m ⁶ m ⁶ A	23.2	2351.4	2350.4
ON1i: X = m ⁶ g ⁶ A	17.2	2335.4	2334.3
ON1j: X = m ⁶ g ⁶ A (amino nitrile)	21.2	2258.4	2258.5
ON1k: X = v ⁶ A (non-methylated)	20.6	2305.4	2302.2
ON1l: X = m ⁶ v ⁶ A	22.3	3300.5	3301.1
ON1m: X ¹ = m ⁶ v ⁶ A and X ² = m ⁶ g ⁶ A	23.1	7231.0	7233.7
ON1n: X = m ⁶ v ⁶ A	23.2	3604.6	3603.4
ON1o: X = m ⁶ g ⁶ Am	23.8	2375.5	2374.4
ON1p: X = m ⁶ g ⁶ Am	23.6	1736.4	1735.1
ON1q: X = m ⁶ g ⁶ Am	23.1	1058.2	1058.2

4.3 Acceptor oligonucleotides (ON2) with a complementary sequence

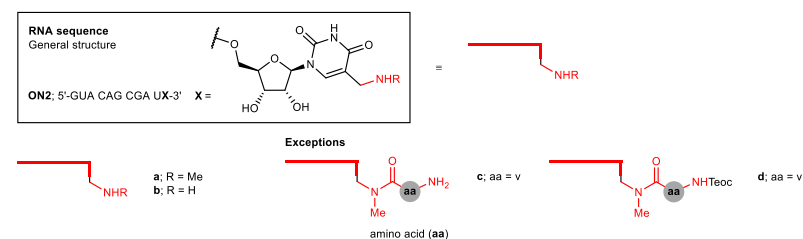


Figure S3. RNA sequence and general structure of (methyl)aminomethyl uridine derivatives.

Other RNA acceptor strands with longer sequences:

ON2e: 5'-GUA CAG CGA UX¹A CUC AAU AGX²-3'; X¹ = gmmn⁶U; X² = nm⁶U

ON2f: 5'-GUA CAG CGA UXA CUC AAU AGG-3'; X = ymmn⁶U

ON2g: 5'-GmUmAm CmAmGm CmGmAm UmX-3'; X = mnm⁶U

ON2h: 5'-GmUmCm AmGmUm AmCmAm GmCmGm AmUmX-3'; X = mnm⁶Um

Table S4. HPLC retention times (0-40% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of **ON2**.

Strand	t_R (min)	m/z calcd. for $[M-H]^-$	found
ON2a: X = mnm ⁶ U	17.4	3530.5	3529.7
ON2b: X = nm ⁶ U	17.8	3516.5	3515.9
ON2c: X = ymmn ⁶ U	18.6	3629.6	3627.2
ON2d: X = Teoc-ymnm ⁶ U	37.7	3773.7	3776.9
ON2e: X ¹ = gmmn ⁶ U and X ² = nm ⁶ U	18.7	6806.0	6806.4
ON2f: X = ymmn ⁶ U	19.9	6858.0	6857.7
ON2g: X = mnm ⁶ U	23.0	3670.5	3670.4
ON2h: X = mnm ⁶ Um	24.2	5025.9	5026.0

4.4 Donor oligonucleotides with non-complementary sequences

RNA sequences that are not fully complementary to the acceptor **ON2**:

ON1r: 5'-XAU **AGC** U-3'; X = m⁶g⁶A (one mismatch marked in red)

ON1s: 5'-XAG **CCC** U-3'; X = m⁶g⁶A (two mismatches marked in red)

Table S5. HPLC retention times (0-40% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of **ON1r** and **ON1s**.

Strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
ON1r : X = m ⁶ g ⁶ A	20.0	2301.4	2301.2
ON1s : X = m ⁶ g ⁶ A	19.5	2276.4	2275.7

5. HPLC calibration curves using canonical oligonucleotides (**CON1-6**) and hairpin-type intermediate (**ON3a**)

Canonical oligonucleotides, **CON1-6**, and hairpin-type intermediate, **ON3a**, were used for the development of HPLC calibration curves. Separate stock solutions of **CON1-6** and **ON3a** were prepared in water (100 μM). Separate standard solutions containing 1.2; 1.0; 0.8; 0.6; 0.4; 0.2 and 0.1 nmol of **CON1-6** and **ON3a** were prepared in a final volume of 20 μL. The standard solutions were injected in an analytical HPLC equipped with a C18 column and using buffers A and B (gradient: 0-30% or 0-40% of B in 45 min; flow rate = 1 mL·min⁻¹). The absorbance was monitored at 260 nm and the areas of the chromatographic peaks were determined by integration of the HPLC-chromatograms. The plot of the chromatographic area (a.u.) versus the amount (nmol) of each oligonucleotide followed a linear relationship.

Calibration curve of **CON1**

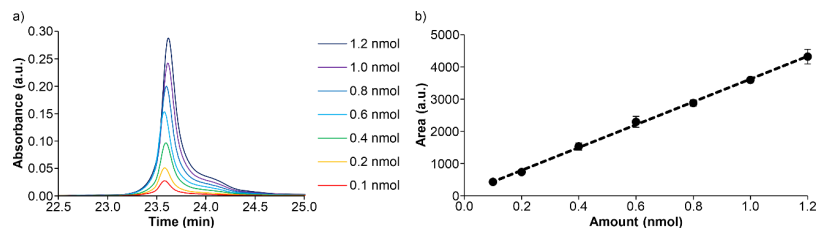


Figure S4. a) Selected region of the HPLC-chromatograms upon the injection of incremental amounts (nmol) and b) chromatographic area (a.u.) vs. amount (nmol) of **CON1**. In b) the line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

Calibration curve of **CON2**

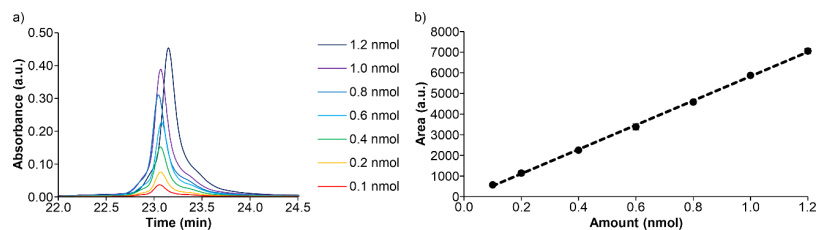


Figure S5. a) Selected region of the HPLC-chromatograms upon the injection of incremental amounts (nmol) and b) chromatographic area (a.u.) vs. amount (nmol) of **CON2**. In b) the line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

Calibration curve of **CON3**

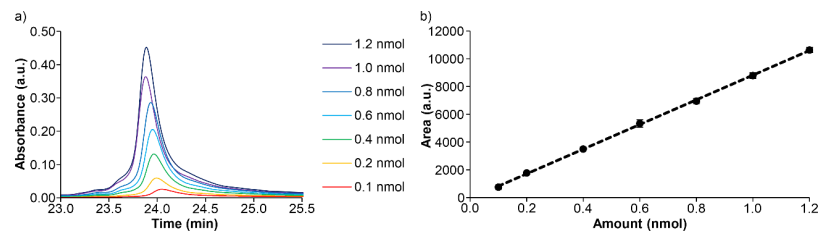


Figure S6. a) Selected region of the HPLC-chromatograms upon the injection of incremental amounts (nmol) and b) chromatographic area (a.u.) vs. amount (nmol) of **CON3**. In b) the line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

Calibration curve of **ON3a**

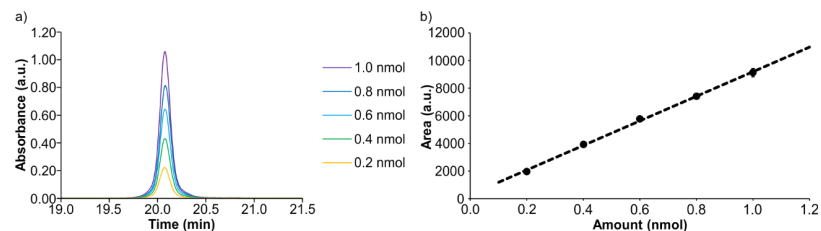


Figure S7. a) Selected region of the HPLC-chromatograms upon the injection of incremental amounts (nmol) and b) chromatographic area (a.u.) vs. amount (nmol) of **ON3a**. In b) the line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

The results of the calibration curves of **CON3** (canonical oligonucleotide) and **ON3a** (hairpin-type intermediate) were very similar (Table S6).

Calibration curve of **CON4**

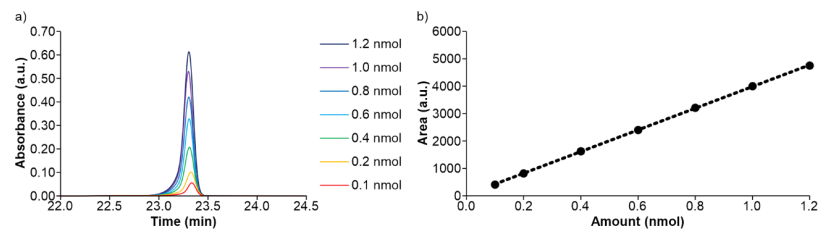


Figure S8. a) Selected region of the HPLC-chromatograms upon the injection of incremental amounts (nmol) and b) chromatographic area (a.u.) vs. amount (nmol) of **CON4**. In b) the line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

Calibration curve of CON5

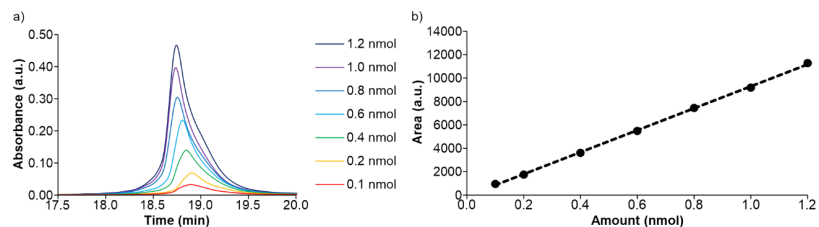


Figure S9. a) Selected region of the HPLC-chromatograms upon the injection of incremental amounts (nmol) and b) chromatographic area (a.u.) vs. amount (nmol) of **CON5**. In b) the line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

Calibration curve of CON6

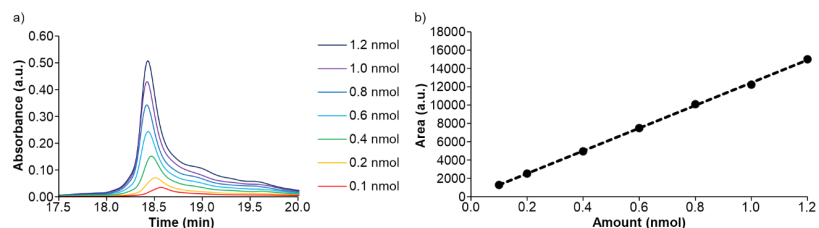


Figure S10. a) Selected region of the HPLC-chromatograms upon the injection of incremental amounts (nmol) and b) chromatographic area (a.u.) vs. amount (nmol) of **CON6**. In b) the line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

Table S6. Calibration curves ($y = mx + n$) obtained by HPLC analyses of **CON1-6** and **ON3a** and calculated extinction coefficients of **CON1-6** using the OligoAnalyzer Version 3.0 from Integrated DNA Technologies.

Strand	Slope, m (nmol ⁻¹)	Intercept, n	r^2	ϵ (M ⁻¹ ·cm ⁻¹)
CON1	3534.2	82.3	0.9989	65500
CON2	5903.7	-73.3	0.9994	107200
CON3	8885.4	-64.6	0.9997	170700
ON3a	8890.5	299.2	0.9986	170700 ^a
CON4	3952.4	32.36	0.9999	68800
CON5	9376.2	-83.62	0.9995	153800
CON6	12405.0	41.49	0.9996	221900

^a In order to simplify the calculations, the extinction coefficient of **ON3a** was assumed to be identical to that of **CON3**.

6. Coupling reactions between donor and acceptor oligonucleotides, ON1 and ON2

Stock solutions of pH buffer (400 mM), NaCl (1 M) and activator (500 mM, Figure S11) were prepared in water. Subsequently, equimolar amounts of **ON1** and **ON2** (3-5 nmol) were annealed at 95°C for 4 min in water containing NaCl (half of the volume required for the reaction). Finally, buffer, NaCl, activator solutions and water were added to the ONs' solution and the reaction was incubated in a ThermoMixer at 25°C for 24 h.

Concentration of the components in the reaction mixture: 50 μ M of **ON1**, 50 μ M of **ON2**, 100 mM of buffer, 100 mM of NaCl and 50 mM of activator (see figure footnotes for details).

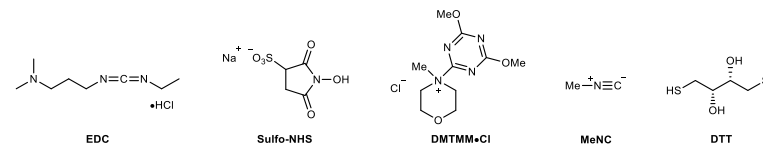


Figure S11. Activators of carboxylic acid and nitrile groups.

The crudes of the reactions (20 μ L, 1 nmol) were analyzed as indicated in Section 3.2.

6.1 Control experiments

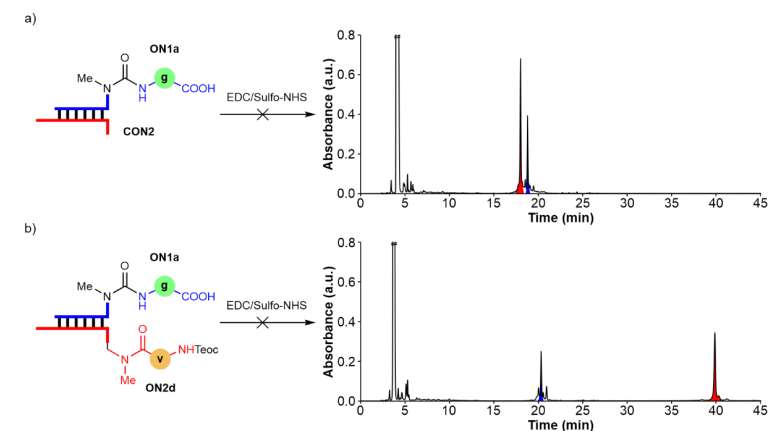
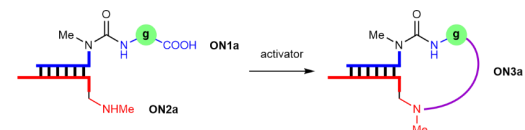


Figure S12. HPLC-chromatograms of the reactions of **ON1a**; $X = m^6g^6A$ with: a) **CON2** (complementary canonical ON) and b) **ON2d**; $X = \text{Teoc-}\gamma\text{mm}^5\text{U}$ in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Control reactions using the donor strand **1a** and the RNA strand lacking the mnm group on the 3'-terminal uridine base **CON2** or the protected 3'- $\gamma\text{mm}^5\text{U}$ -RNA-5' acceptor strand **ON2d** did not provide noticeable evidence for the formation of the corresponding hairpin-type intermediate products.

6.2 Screening of activators using ON1a (m^6g^6A) and ON2a (mnm^5U)



Scheme S9. Coupling of **ON1a**; $X = m^6g^6A$ with **ON2a**. The formed peptide bond is marked in purple.

MES buffer at pH 6 (adjusted with NaOH)

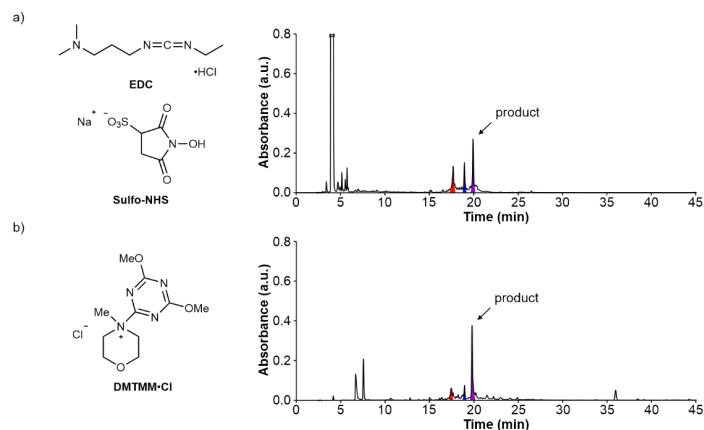


Figure S13. HPL-chromatograms of the reactions of **ON1a**; X = m⁶g⁶A with **ON2a** using: a) EDC/Sulfo-NHS and b) DMTMM-Cl as activators.

Table S7. Results obtained in the coupling reactions of **ON1a**; X = m⁶g⁶A with **ON2a** (average of, at least, two experiments).

Activators	pH	Time (h)	Average Yield ± Error (%) ^a
EDC/Sulfo-NHS	6	24	16±4
DMTMM-Cl	6	24	33±2

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

MOPS buffer at pH 7 (adjusted with NaOH)

Table S8. Results obtained in the coupling reactions of **ON1a**; X = m⁶g⁶A with **ON2a** (average of, at least, two experiments).

Activators	pH	Time (h)	Average Yield ± Error (%) ^a
EDC/Sulfo-NHS	7	24	20±2

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

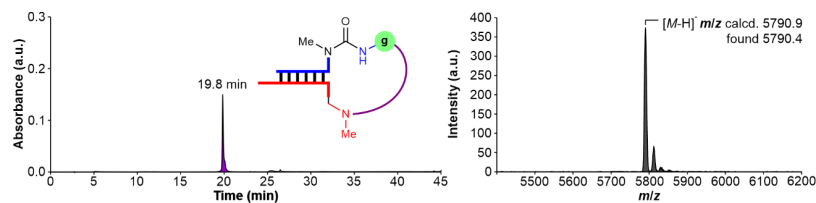


Figure S14. Left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the isolated product **ON3a**.

6.3 Screening of activators using ON1a (m⁶g⁶A) and ON2b (nm⁵U)



Scheme S10. Coupling of **ON1a**; X = m⁶g⁶A with **ON2b**. The formed peptide bond is marked in purple.

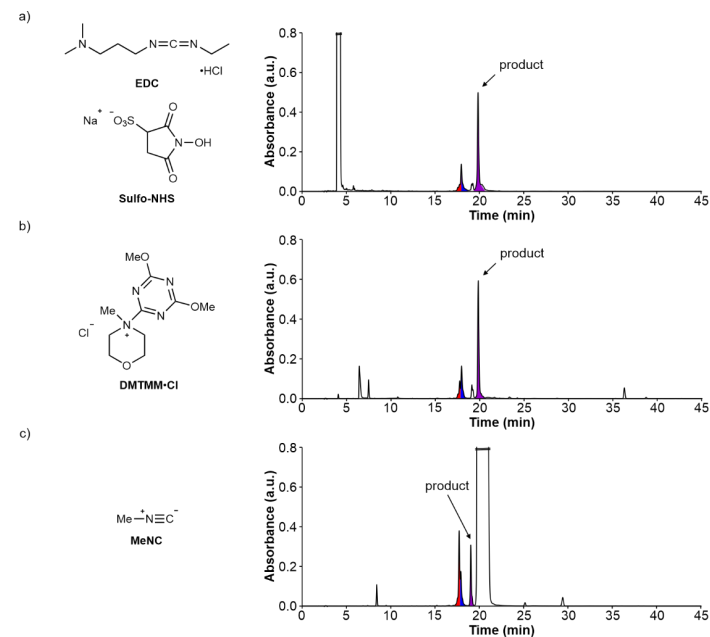


Figure S15. HPL-chromatograms of the reactions of **ON1a**; X = m⁶g⁶A with **ON2b** using: a) EDC/Sulfo-NHS; b) DMTMM-Cl and c) MeNC as activators. MES buffer (100 mM) at pH 6 in a) and b). DCl buffer (50 mM) at pH 6 in c).

Table S9. Results obtained in the coupling reactions of **ON1a**; X = m⁶g⁶A with **ON2b** (average of, at least, two experiments).

Activators	pH	Time (h)	Average Yield ± Error (%) ^a
EDC/Sulfo-NHS	6	24	64±2
DMTMM-Cl	6	24	66±2
MeNC	6	120	28±4

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

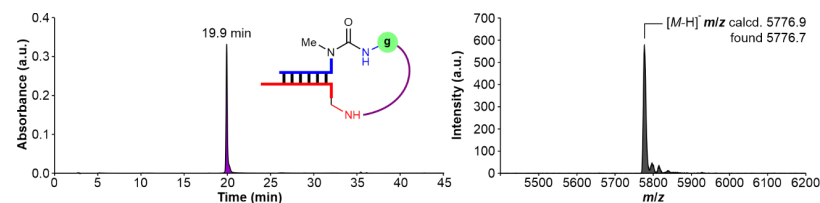
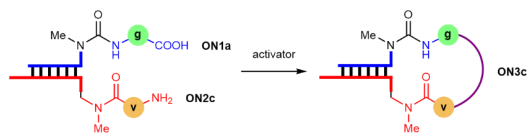


Figure S16. left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the isolated product **ON3b**.

6.4 Screening of activators using ON1a (m⁶g⁶A) and ON2c (vmm⁵U)



Scheme S11. Coupling of **ON1a**; X = m⁶g⁶A with **ON2c**. The formed peptide bond is marked in purple.

Buffer at pH 6

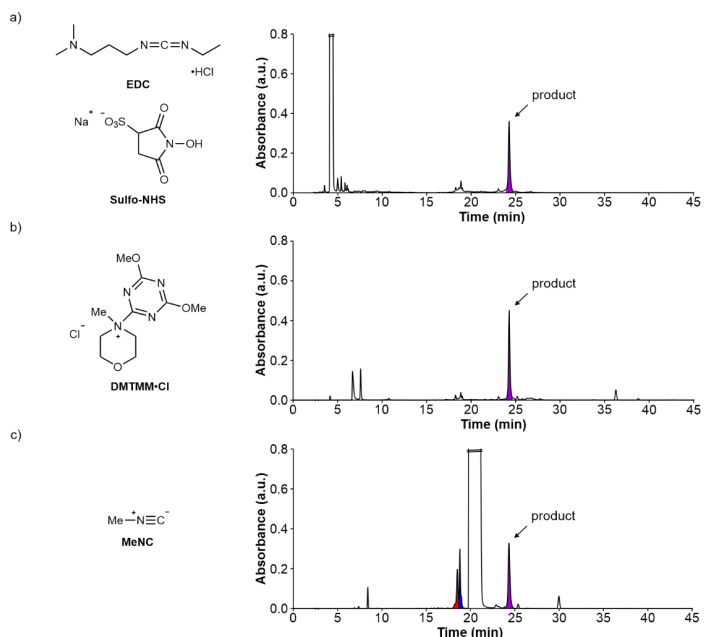


Figure S17. HPL-chromatograms of the reactions of **ON1a**; X = m⁶g⁶A with **ON2c** using: a) EDC/Sulfo-NHS; b) DMTMM-Cl and c) MeNC as activators. MES buffer (100 mM) at pH 6 in a) and b). DCl buffer (50 mM) at pH 6 in c).

Table S10. Results obtained in the coupling reactions of **ON1a**; X = m⁶g⁶A with **ON2c** (average of, at least, two experiments).

Activators	pH	Time (h)	Average Yield ± Error (%) ^a
EDC/Sulfo-NHS	6	24	56±1
DMTMM-Cl	6	24	60±2
MeNC	6	120	50±5

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

MOPS buffer at pH 7 (adjusted with NaOH)

Table S11. Results obtained in the coupling reactions of **ON1a**; X = m⁶g⁶A with **ON2c** (average of, at least, two experiments).

Activators	pH	Time (h)	Average Yield ± Error (%) ^a
EDC/Sulfo-NHS	7	24	50±5
DMTMM-Cl	7	24	23±1

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

MOPS buffer at pH 8 (adjusted with NaOH)

Table S12. Results obtained in the coupling reactions of **ON1a**; X = m⁶g⁶A with **ON2c** (average of, at least, two experiments).

Activators	pH	Time (h)	Average Yield ± Error (%) ^a
EDC/Sulfo-NHS	8	24	34±1
DMTMM-Cl	8	24	5±2

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

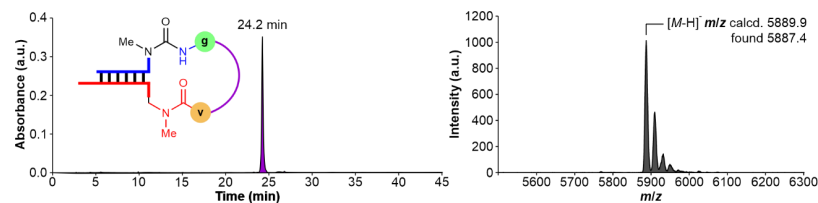


Figure S18. left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the isolated product **ON3c**.

6.5 Coupling reactions of ON1j (m⁶g⁶A, amino nitrile) with ON2a-c

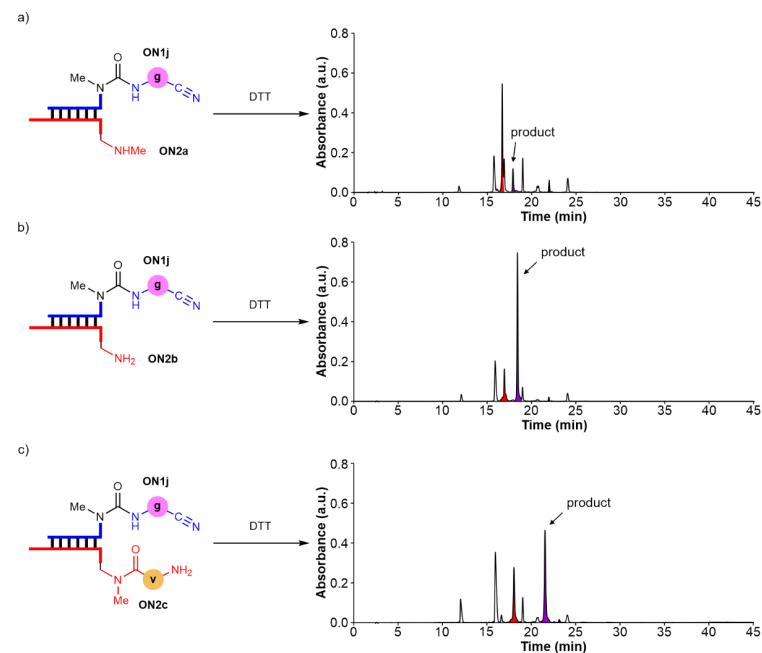


Figure S19. HPL-chromatograms of the reactions of **ON1j**; X = m⁶g⁶A (amino nitrile) with: a) **ON2a**; X = mnm⁵U; b) **ON2b**; X = nm⁵U and c) **ON2c**; X = vmm⁵U in boric acid buffer at pH 8 using DTT as activator.

Table S13. Results obtained in the coupling reactions of **ON1j**; **X** = m^6g^6A (amino nitrile) with **ON2a-c** using DTT as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error (%) ^a
	ON2a ; X = mnm^5U	12 \pm 1
ON1j ; X = m^6g^6A (amino nitrile)	ON2b ; X = nm^5U	65 \pm 2
	ON2c ; X = $ymnm^5U$	42 \pm 1

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

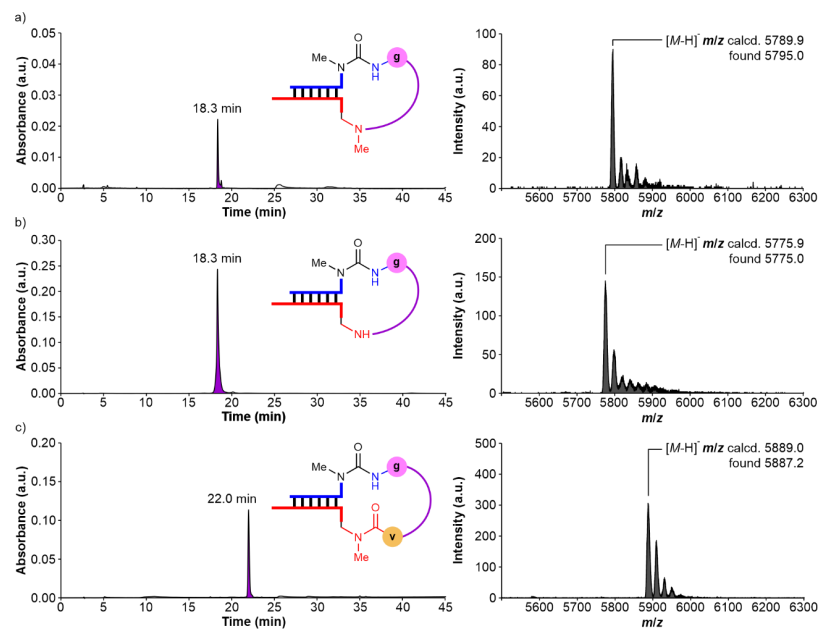


Figure S20. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of **ON1j**; **X** = m^6g^6A (amino nitrile) with: a) **ON2a**; **X** = mnm^5U ; b) **ON2b**; **X** = nm^5U and c) **ON2c**; **X** = $ymnm^5U$.

6.6 Coupling reactions of **ON1b-i** (m^6aa^6A) with **ON2a**

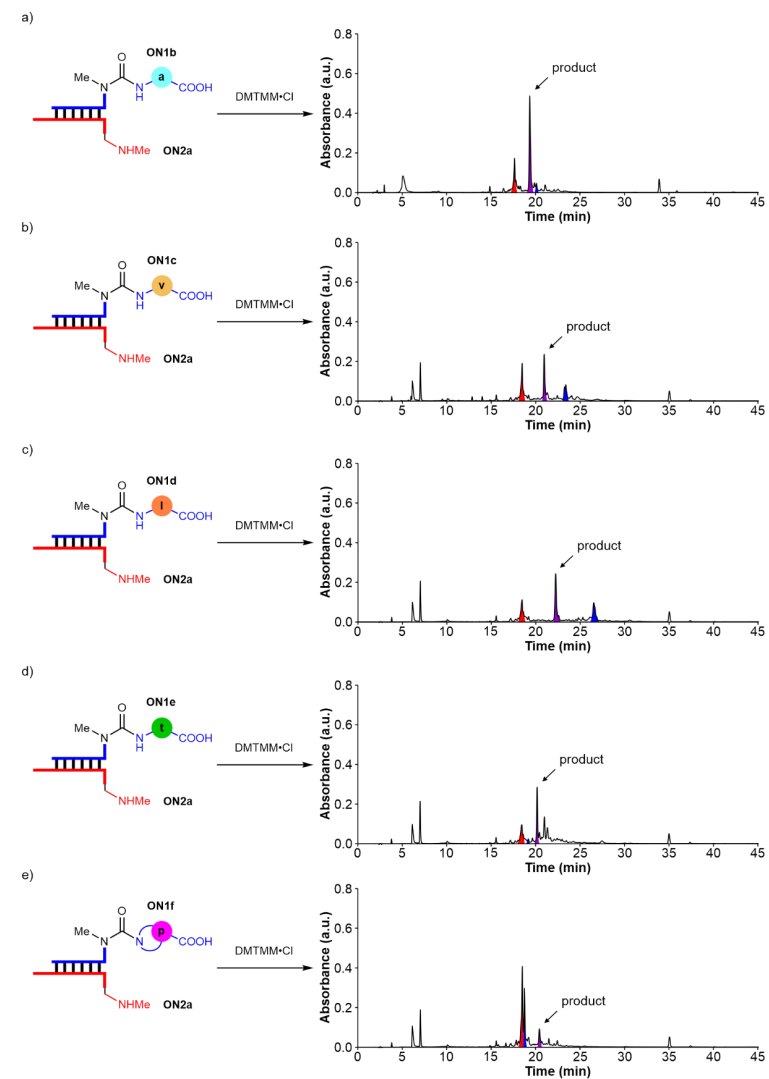


Figure S21. HPL-chromatograms of the reactions of **ON2a**; **X** = mnm^5U with: a) **ON1b**; **X** = m^6a^6A ; b) **ON1c**; **X** = m^6v^6A ; c) **ON1d**; **X** = m^6t^6A ; d) **ON1e**; **X** = m^6l^6A and e) **ON1f**; **X** = m^6p^6A in MES buffer at pH 6 using DMTMM-Cl as activator.

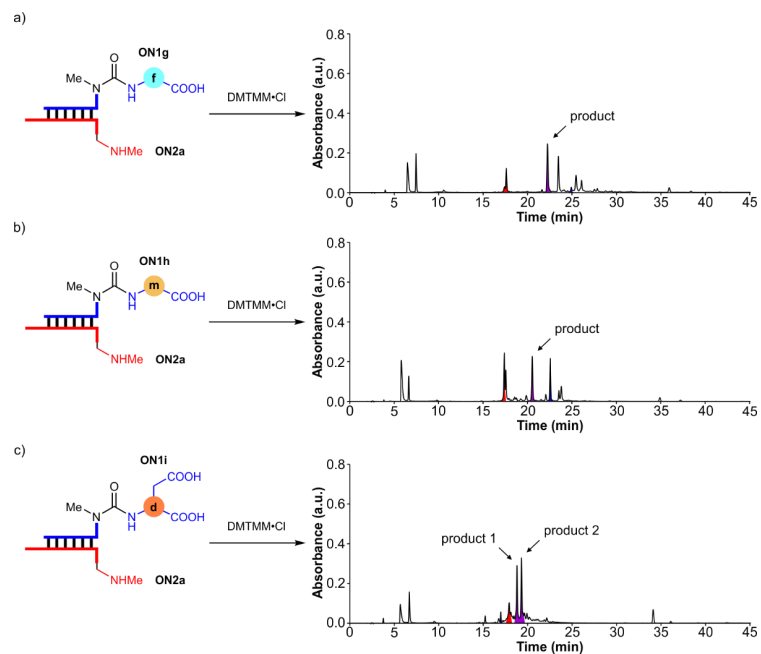


Figure S22. HPL-chromatograms of the reactions of **ON2a**; **X** = mnm^3U with: a) **ON1g**; **X** = $m^{6f}A$; b) **ON1h**; **X** = $m^{6m}A$ and c) **ON1i**; **X** = $m^{6d}A$ in MES buffer at pH 6 using DMTMM-Cl as activator. For **ON1i**, the two peaks corresponded to the products of the reaction of the Asp α -COOH and of the side chain COOH. An assignment was not performed.

Table S14. Results obtained in the coupling reactions of **ON1b-i**; **X** = $m^{6aa}A$ with **ON2a** using DMTMM-Cl as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error (%) ^a
ON1b ; X = $m^{6a}A$	ON2a ; X = mnm^3U	51 \pm 1
ON1c ; X = $m^{6v}A$		21 \pm 1
ON1d ; X = $m^{6f}A$		27 \pm 1
ON1e ; X = $m^{6e}A$		18 \pm 5
ON1f ; X = $m^{6p}A$		11 \pm 1
ON1g ; X = $m^{6g}A$		27 \pm 1
ON1h ; X = $m^{6m}A$		22 \pm 1
ON1i ; X = $m^{6d}A$		28 \pm 4; 26 \pm 3 ^b

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**. ^b For **ON1i**, the two yields describe the reaction of the Asp α -COOH and of the side chain COOH. An assignment was not performed.

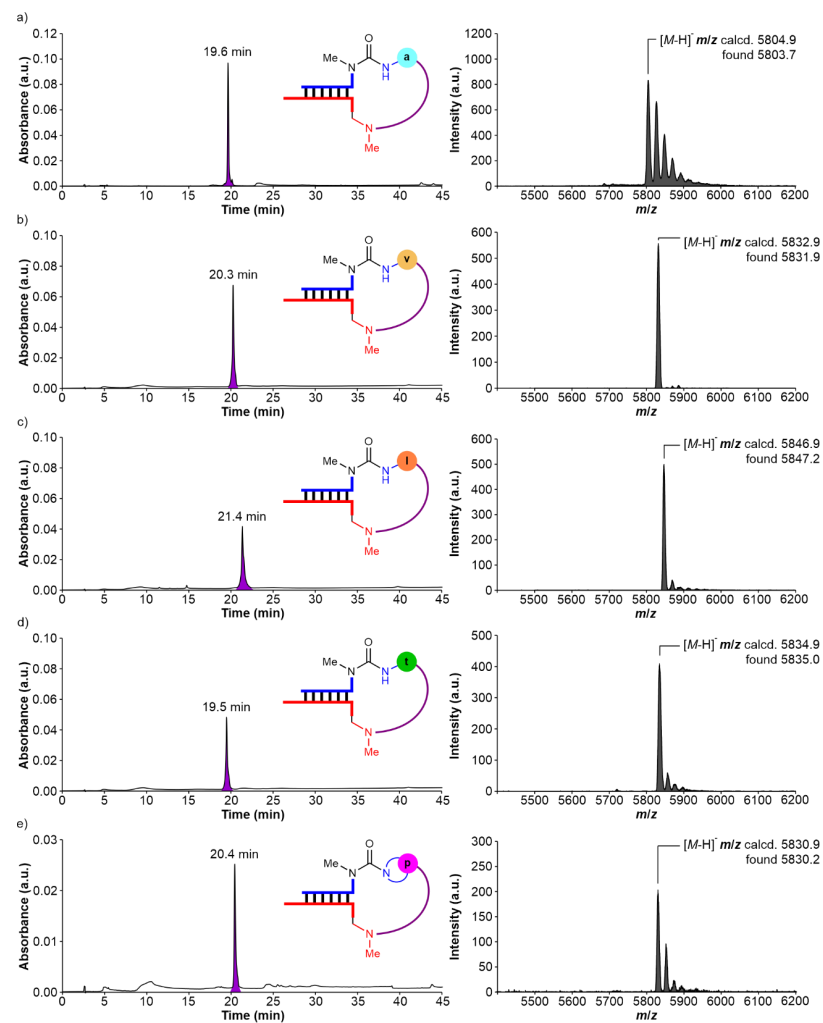


Figure S23. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of **ON2a**; **X** = mnm^3U with: a) **ON1b**; **X** = $m^{6a}A$; b) **ON1c**; **X** = $m^{6v}A$; c) **ON1d**; **X** = $m^{6f}A$; d) **ON1e**; **X** = $m^{6e}A$ and e) **ON1f**; **X** = $m^{6p}A$.

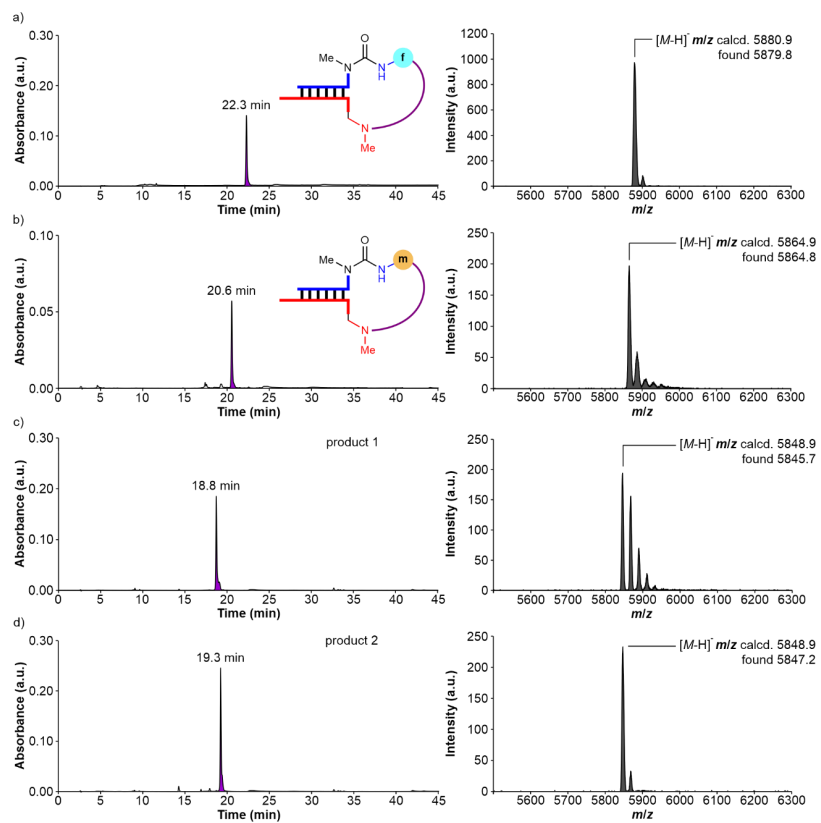


Figure S24. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of ON2a; X = mnm³U with: a) ON1g; X = m¹⁰A; b) ON1h; X = m¹⁰A; c) ON1i; X = m⁶A and d) ON1j; X = m⁶A.

6.7 Coupling reactions of ON1b-i (m⁶aa⁶A) with ON2c

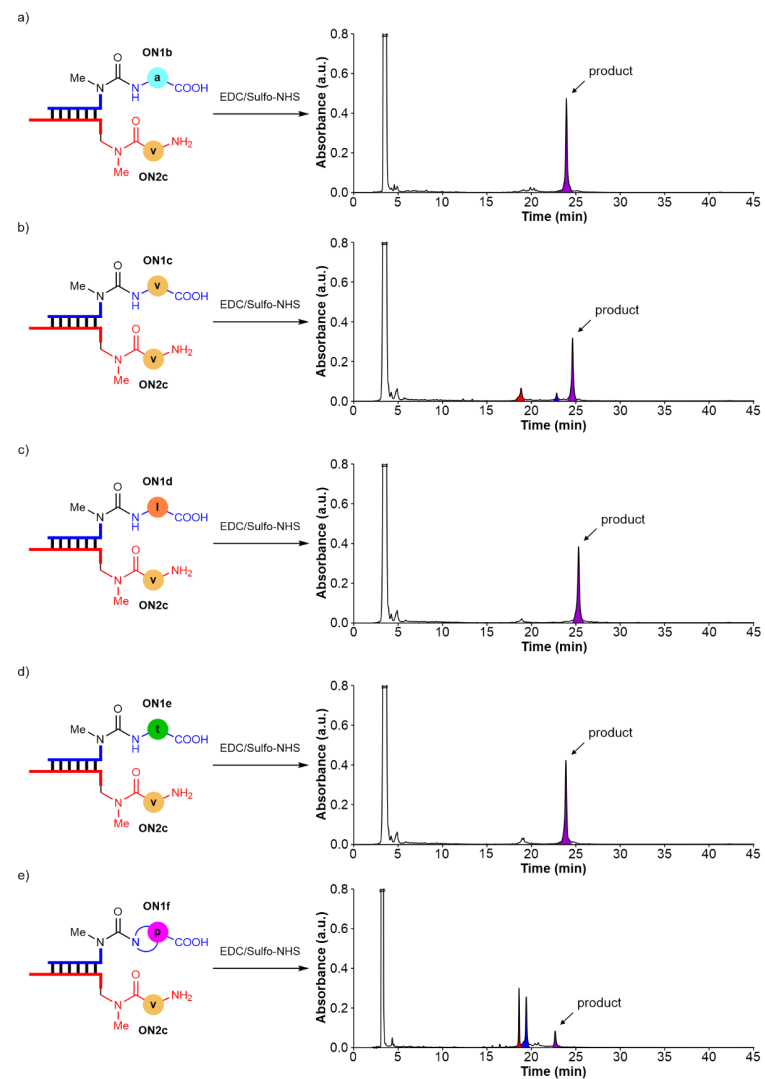


Figure S25. HPL-chromatograms of the reactions of ON2c; X = γ mm⁵U with: a) ON1b; X = m⁶A; b) ON1c; X = m⁶A; c) ON1d; X = m⁶A; d) ON1e; X = m⁶A and e) ON1f; X = m⁶A in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

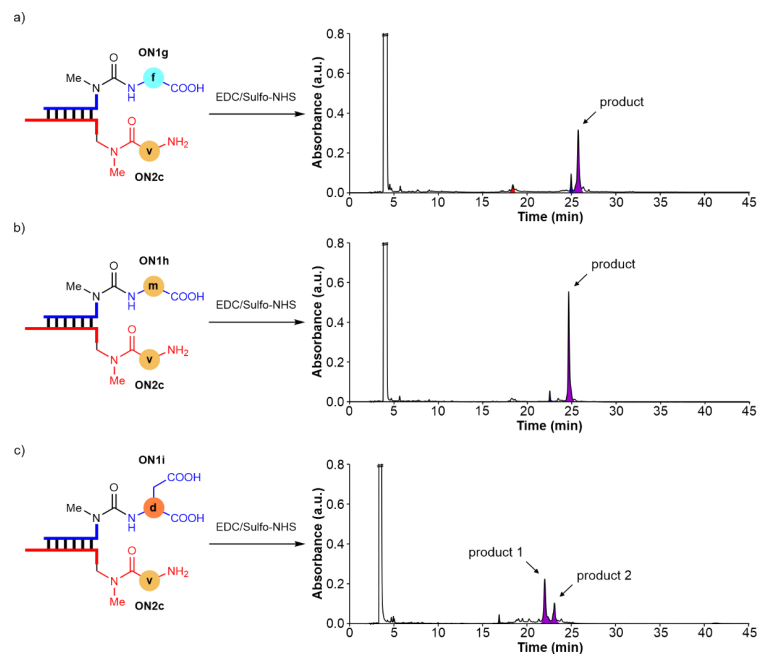


Figure S26. HPL-chromatograms of the reactions of **ON2c**; **X** = γ mm^gU with: a) **ON1g**; **X** = m^gf^gA; b) **ON1h**; **X** = m^gm^gA and c) **ON1i**; **X** = m^gg^gA in MES buffer at pH 6 using EDC/Sulfo-NHS as activator. For **ON1i**, the two peaks corresponded to the products of the reaction of the Asp α -COOH and of the side chain COOH. An assignment was not performed.

Table S15. Results obtained in the coupling reactions of **ON1b-i**; **X** = m^gaa^gA with **ON2c** using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error (%) ^a
ON1b ; X = m ^g g ^g A	ON2c ; X = γ mm ^g U	76 \pm 2
ON1c ; X = m ^g v ^g A		54 \pm 1
ON1d ; X = m ^g f ^g A		77 \pm 1
ON1e ; X = m ^g f ^g A		77 \pm 1
ON1f ; X = m ^g g ^g A		18 \pm 4 (55 \pm 5) ^b
ON1g ; X = m ^g f ^g A		50 \pm 1
ON1h ; X = m ^g m ^g A		70 \pm 2
ON1i ; X = m ^g g ^g A		34 \pm 1; 17 \pm 2 ^c

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**. ^b Using DMTMM-Cl as activator. ^c For **ON1i**, the two yields describe the reaction of the Asp α -COOH and of the side chain COOH. An assignment was not performed.

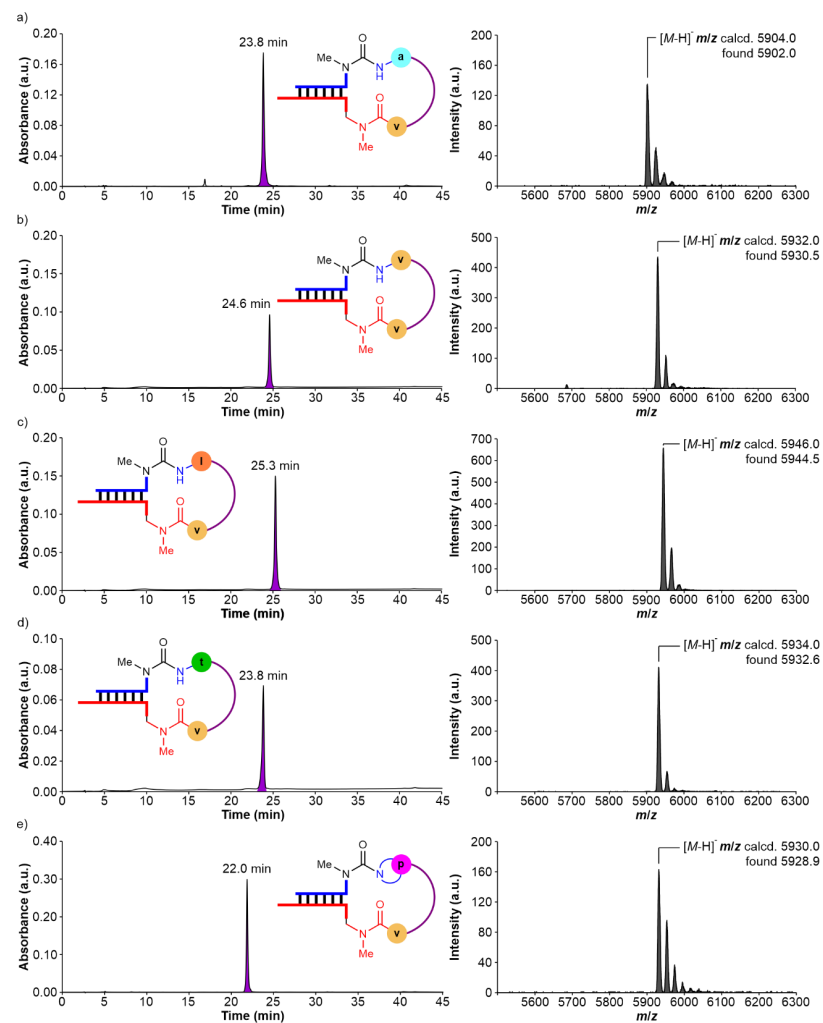


Figure S27. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of **ON2c**; **X** = γ mm^gU with: a) **ON1b**; **X** = m^gg^gA; b) **ON1c**; **X** = m^gv^gA; c) **ON1d**; **X** = m^gf^gA; d) **ON1e**; **X** = m^gf^gA and e) **ON1f**; **X** = m^gg^gA.

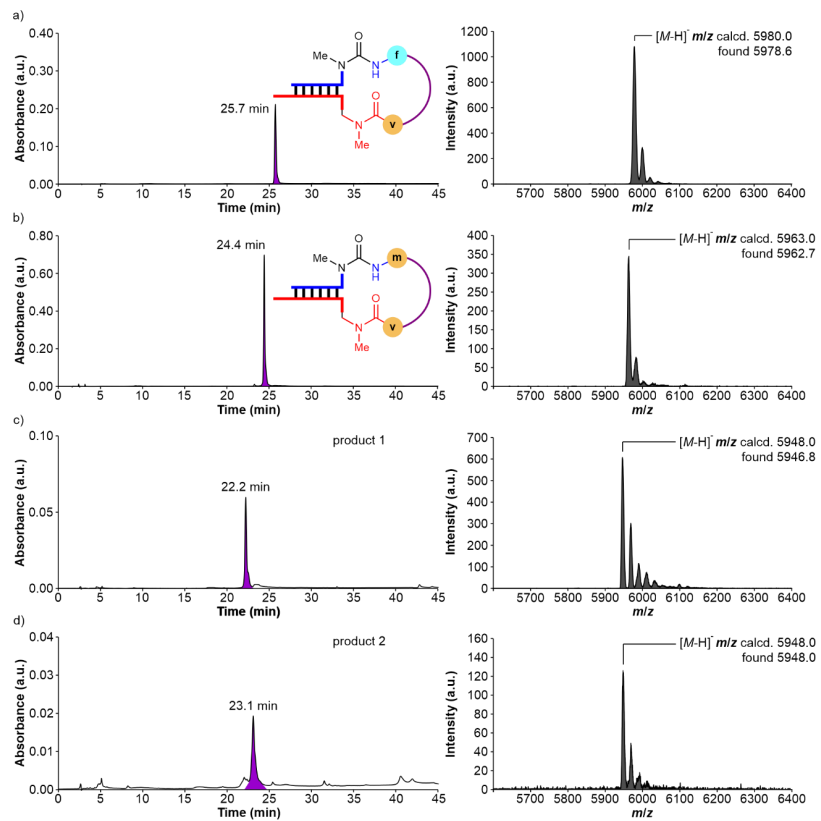


Figure S28. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of ON2c; X = γ mmⁿU with: a) ON1g; X = m²f⁶A; b) ON1h; X = m²m²A; c) ON1i; X = m²d²A and d) ON1j; X = m²d²A.

7. Synthesized peptide-oligonucleotides using solid support beads

7.1 Donor peptide-oligonucleotides with a complementary sequence

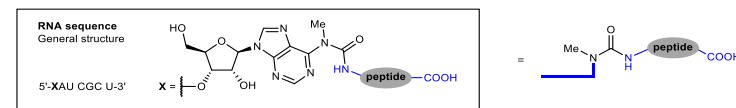


Figure S29. RNA sequence and general structure of peptide-modified carbamoyl adenosine derivatives.

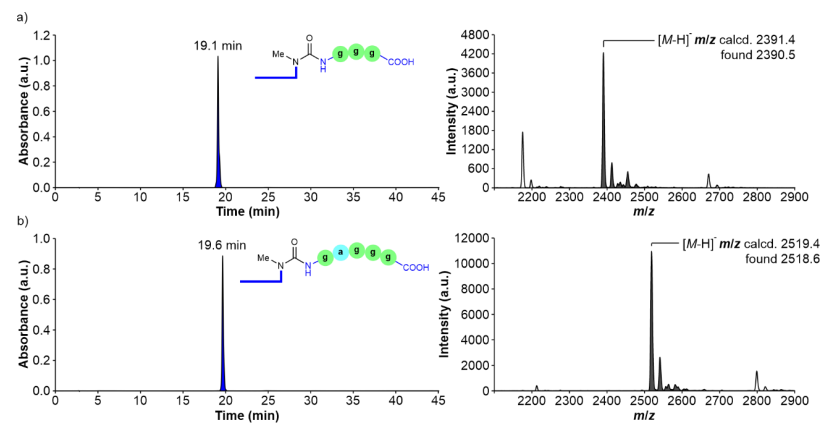


Figure S30. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the synthesized peptide-oligonucleotides: a) 5'-m⁶(ggg)⁶A-RNA-3' and b) 5'-m⁶(gagg)⁶A-RNA-3'.

7.2 Acceptor peptide-oligonucleotides with a complementary sequence

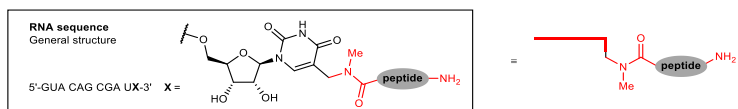


Figure S31. RNA sequence and general structure of peptide-modified methylaminomethyl uridine derivatives.

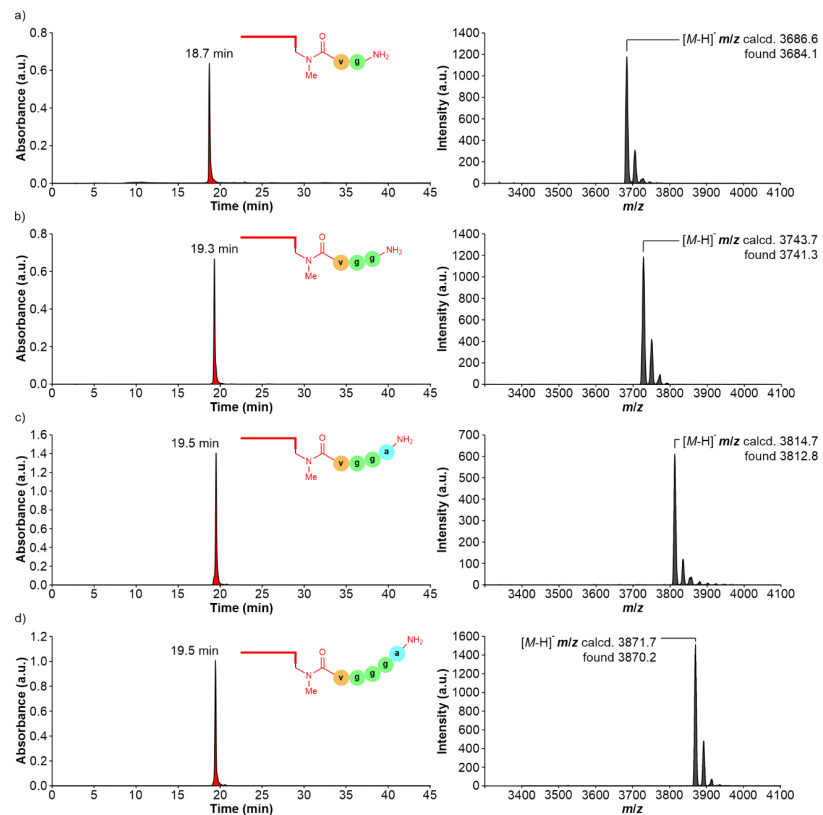


Figure S32. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the synthesized peptide-oligonucleotides: a) 3'-gvmm⁵U-RNA-5'; b) 3'-gvmm⁵U-RNA-5'; c) 3'-aggvmm⁵U-RNA-5' and d) 3'-aggvmm⁵U-RNA-5'.

5'-GmUmAm CmAmGm CmGmAm UmX-3'; X = aggmm⁵U

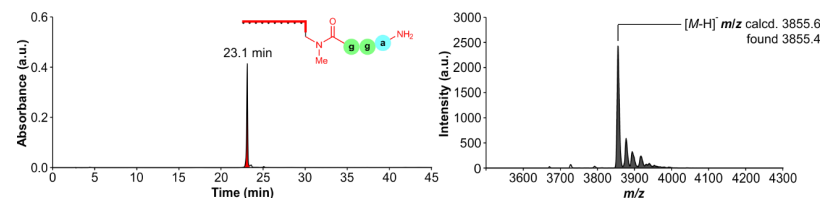


Figure S33. left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the synthesized peptide-oligonucleotide 3'-aggmm⁵U-RNA-5' containing 2'-OMe nucleosides.

8. Coupling reactions between donor and acceptor peptide-oligonucleotides

The peptide coupling reactions were carried out under identical conditions to those described in Section 0.

8.1 Coupling reactions of donor peptide-oligonucleotides with ON2c

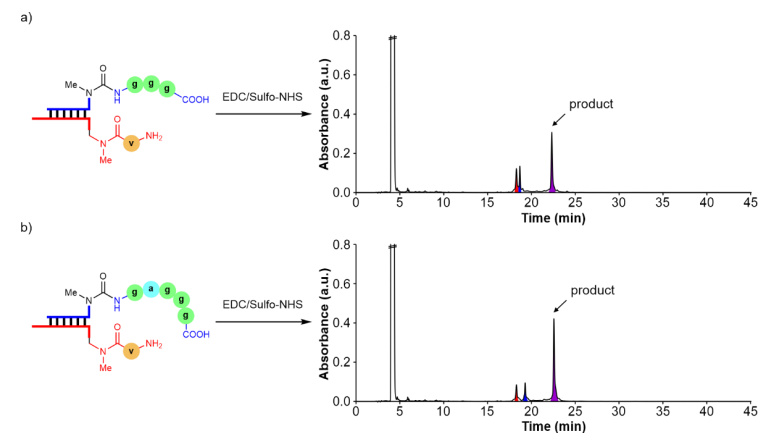


Figure S34. HPL-chromatograms of the reactions of ON2c; X = vmm⁵U with: a) 5'-m⁶(ggg)⁵A-RNA-3' and b) 5'-m⁶(gaggg)⁵A-RNA-3' in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Table S16. Results obtained in the coupling reactions of **ON2c**; $X = \text{ymnm}^5\text{U}$ with peptide-modified donor oligonucleotides using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error (%) ^a
5'-m ⁶ (ggg) ⁶ A-RNA-3'	ON2c ; $X = \text{ymnm}^5\text{U}$	35 \pm 1
5'-m ⁶ (gagg) ⁶ A-RNA-3'		43 \pm 1

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

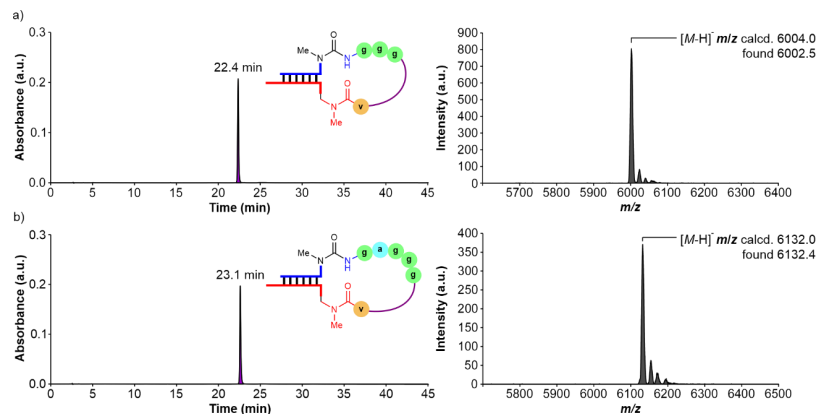


Figure S35. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of **ON2c**; $X = \text{ymnm}^5\text{U}$ with: a) 5'-m⁶(ggg)⁶A-RNA-3' and b) 5'-m⁶(gagg)⁶A-RNA-3'.

8.2 Coupling reactions of **ON1a** (m⁶g⁶A) with acceptor peptide-oligonucleotides

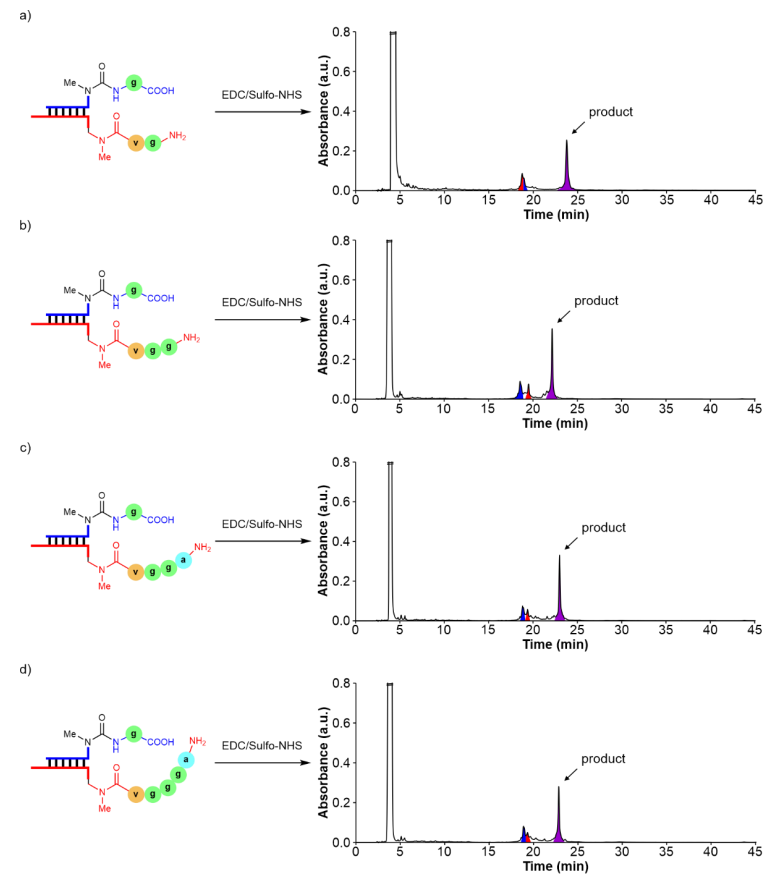


Figure S36. HPL-chromatograms of the reactions of **ON1a**; $X = \text{m}^6\text{g}^6\text{A}$ with: a) 3'-gvymnm⁵U-RNA-5'; b) 3'-ggvymnm⁵U-RNA-5'; c) 3'-aggvymnm⁵U-RNA-5' and d) 3'-aggvymnm⁵U-RNA-5' in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Table S17. Results obtained in the coupling reactions of **ON1a**; $X = \text{m}^6\text{g}^6\text{A}$ with peptide-modified acceptor oligonucleotides using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error (%) ^a
ON1a ; $X = \text{m}^6\text{g}^6\text{A}$	3'-gvymnm ⁵ U-RNA-5'	51 \pm 1
	3'-ggvymnm ⁵ U-RNA-5'	46 \pm 4
	3'-aggvymnm ⁵ U-RNA-5'	40 \pm 1
	3'-aggvymnm ⁵ U-RNA-5'	40 \pm 3 (57 \pm 2) ^b

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**. ^b Using DMTMM-Cl as activator.

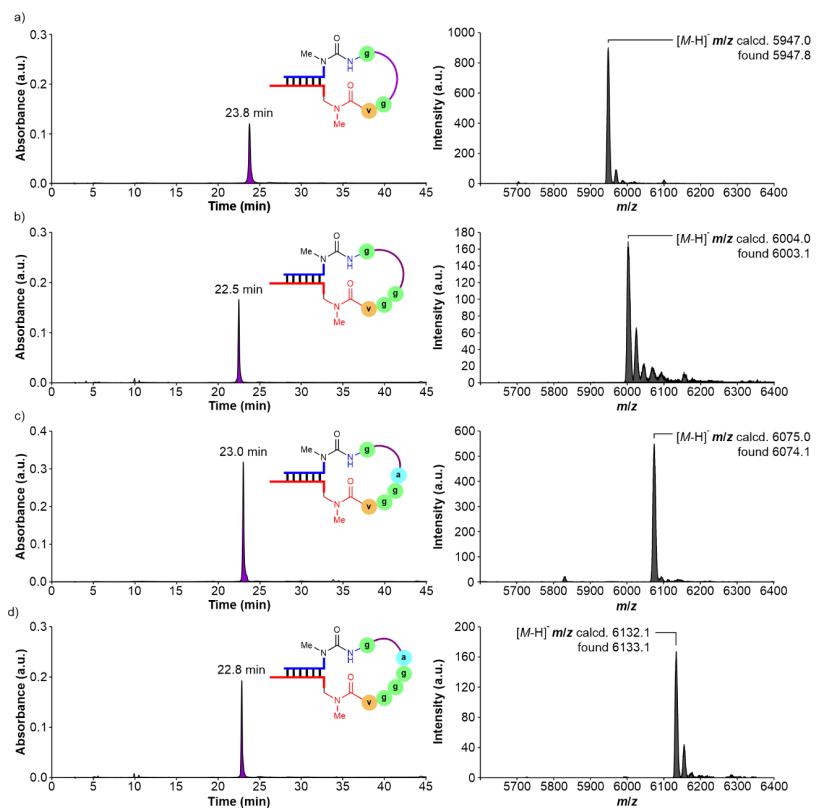


Figure S37. Left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of **ON1a**; $X = m^6g^3A$ with: a) 3'-*gvmmn*⁵U-RNA-5'; b) 3'-*gvmmn*⁵U-RNA-5'; c) 3'-*agggvmmn*⁵U-RNA-5' and d) 3'-*agggvmmn*⁵U-RNA-5'.

8.3 Coupling reactions of donor and acceptor peptide-oligonucleotides

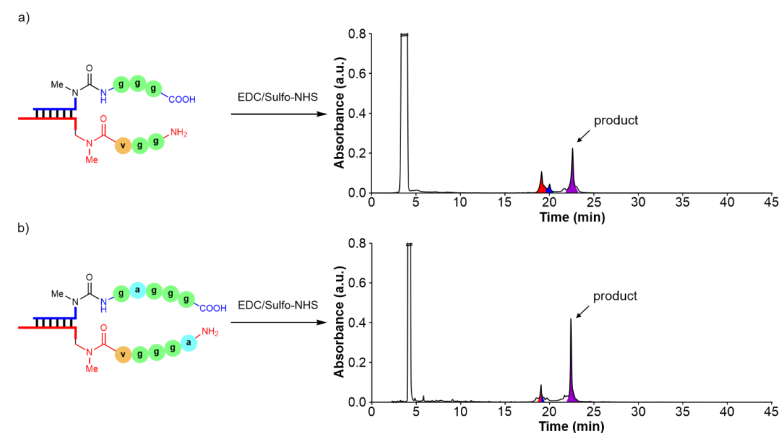


Figure S38. HPL-chromatograms of the reactions of: a) 5- $m^6(ggg)A$ -RNA-3' with 3'-*gvmmn*⁵U-RNA-5' and b) 5- $m^6(gagg)A$ -RNA-3' with 3'-*agggvmmn*⁵U-RNA-5' in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Table S18. Results obtained in the coupling reactions of peptide-modified donor and acceptor oligonucleotides using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error (%) ^a
5- $m^6(ggg)A$ -RNA-3'	3'- <i>gvmmn</i> ⁵ U-RNA-5'	53 \pm 1
5- $m^6(gagg)A$ -RNA-3'	3'- <i>agggvmmn</i> ⁵ U-RNA-5'	56 \pm 3

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

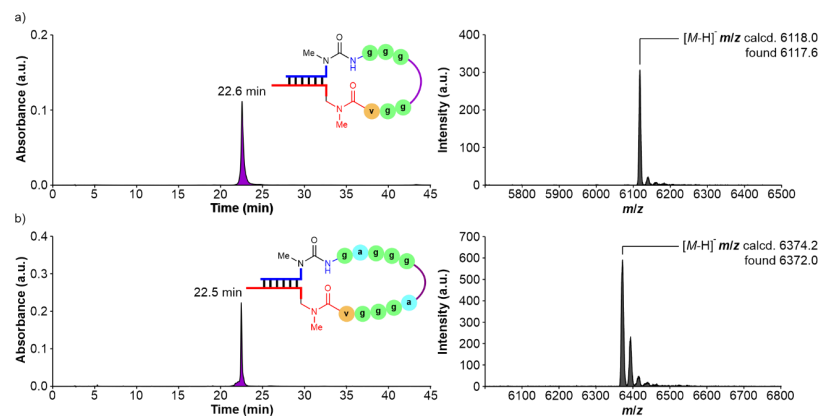


Figure S39. Left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of: a) 5- $m^6(ggg)A$ -RNA-3' with 3'-*gvmmn*⁵U-RNA-5' and b) 5- $m^6(gagg)A$ -RNA-3' with 3'-*agggvmmn*⁵U-RNA-5'.

9. Concentration of the product versus time in selected coupling reactions

The peptide coupling reactions were carried out under identical conditions to those described in Section 0 using DMTMM·Cl as activator.

The data (concentration of product vs. time) was fit to the corresponding theoretical kinetic model using the Parameter Estimation Module of COPASI software Version 4.29.¹² We introduced the theoretical kinetic model shown below:

Double strand → Hairpin-type Intermediate; k_{app}

The initial concentration of the double strand was refined as variable but constrained between 30 and 50×10^{-6} M. The fit of the data returned the rate constant value k_{app} . This fitting procedure is similar to that reported by others in the literature.¹³

In all cases, the fit of the experimental data was good based on the residual values, reported as sum of squared residuals (SSR), and the visual inspection of the curves.

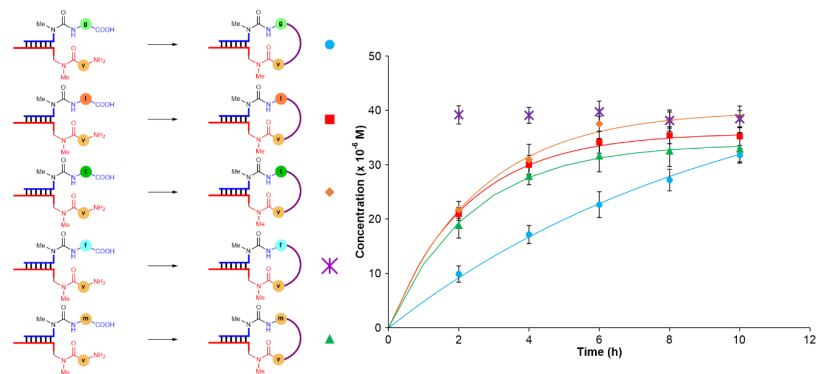


Figure S40. Concentration of the product (M) vs. time (h) in selected peptide coupling reactions using DMTMM-Cl as activator. Lines show fit of the data to the theoretical kinetic model. Error bars are the standard deviations.

Table S19. Calculated rate constant values for selected coupling reactions (average of, at least, two experiments).

Donor strand	Acceptor strand	k_{app} (h^{-1}) ^a	SSR ^b
ON1a; X = m ³ g ² A		0.12±0.02	2.00 × 10 ⁻¹²
ON1d; X = m ³ l ⁶ A		0.42±0.02	8.20 × 10 ⁻¹³
ON1e; X = m ² l ⁶ A	ON2c; X = <u>ymnm</u> ⁵ U	0.39±0.04	2.50 × 10 ⁻¹²
ON1g; X = m ² l ⁶ A		>1	n.d.
ON1h; X = m ² m ³ A		0.42±0.04	5.80 × 10 ⁻¹³

^a Errors are indicated as standard deviations. ^b SSR = Sum of squared residuals.

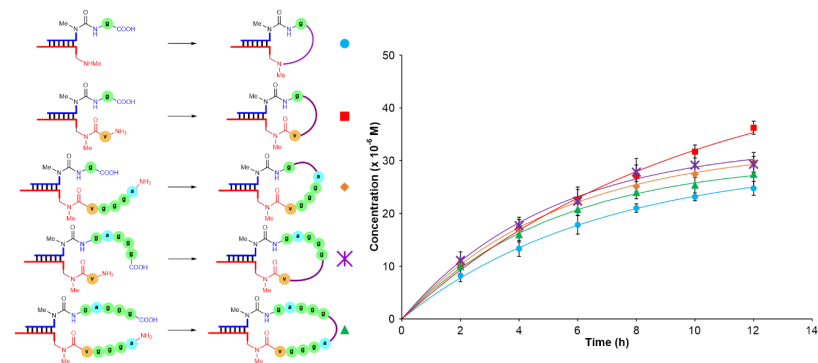


Figure S41. Concentration of the product (M) vs. time (h) in selected peptide coupling reactions using DMTMM-Cl as activator. Lines show fit of the data to the theoretical kinetic model. Error bars are the standard deviations.

Table S20. Calculated rate constant values for selected coupling reactions (average of, at least, two experiments).

Donor strand	Acceptor strand	k_{app} (h^{-1}) ^a	SSR ^b
ON1a; X = m ³ g ² A	ON2a; X = mnm ⁵ U	0.14±0.02	3.36 × 10 ⁻¹³
ON1a; X = m ³ g ² A	ON2c; X = <u>ymnm</u> ⁵ U	0.12±0.02	2.00 × 10 ⁻¹²
ON1a; X = m ³ g ² A	3'-aggymnm ⁵ U-RNA-5'	0.18±0.02	5.59 × 10 ⁻¹³
5'-m ³ (gaggg) ⁶ A-RNA-3'	ON2c; X = <u>ymnm</u> ⁵ U	0.19±0.02	3.82 × 10 ⁻¹²
5'-m ³ (gaggg) ⁶ A-RNA-3'	3'-aggymnm ⁵ U-RNA-5'	0.19±0.01	4.45 × 10 ⁻¹³

^a Errors are indicated as standard deviations. ^b SSR = Sum of squared residuals.

10. Coupling reactions between oligonucleotides containing multiple donor or acceptor units

The peptide coupling reactions were carried out under identical conditions to those described in Section 0.

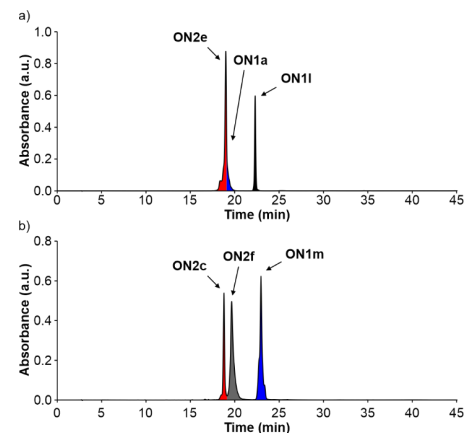


Figure S42. HPL-chromatograms of equimolar mixtures of: a) ON1a; X = m³g²A, ON1l; X = m³v⁴A and ON2e; X¹ = gnm⁵U and X² = nm⁵U, and b) ON1m; X¹ = m³v⁴A and X² = m³g²A, ON2c; X = ymnm⁵U and ON2f; X = ymnm⁵U.

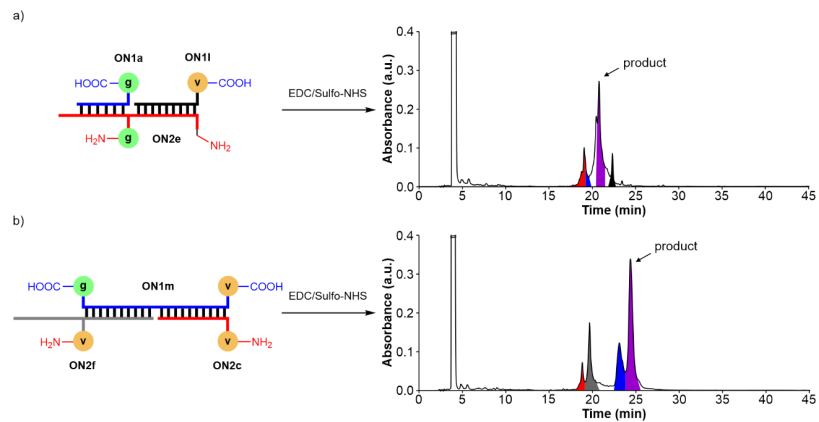


Figure S43. HPL-chromatograms of the reactions of: a) **ON1a**; X = m⁶g⁶A, **ON1l**; X = m⁶v⁶A and **ON2e**; X¹ = gmm⁵U and X² = nm⁵U and b) **ON1m**; X¹ = m⁶v⁶A and X² = m⁶g⁶A, **ON2c**; X = vmm⁵U and **ON2f**; X = vmm⁵U in MES buffer at pH 6 using EDC/Sulfo-NHS as activator. The terminal functional groups of the ONs (urea and amide) are omitted for clarity.

Table S21. Results obtained in the coupling reactions of oligonucleotides containing multiple donor or acceptor units using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield ± Error (%) ^a
ON1a ; X = m ⁶ g ⁶ A	ON2e ; X ¹ = gmm ⁵ U and X ² = nm ⁵ U	35±2 (29±1) ^b
ON1m ; X ¹ = m ⁶ v ⁶ A and X ² = m ⁶ g ⁶ A	ON2c ; X = vmm ⁵ U and ON2f ; X = vmm ⁵ U	35±3 (32±2) ^b

^a Calculated yield from the chromatographic peak of the product based on the total area of the initial components (Figure S42). ^b Using DMTMM-Cl as activator.

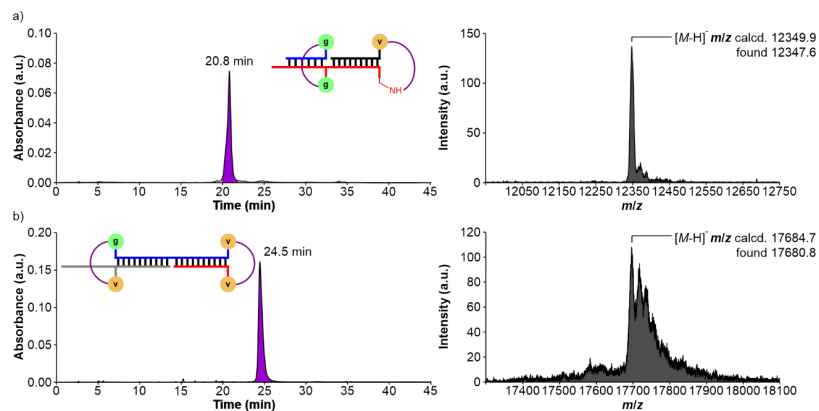


Figure S44. left) HPL-chromatograms and right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions of: a) **ON1a**; X = m⁶g⁶A, **ON1l**; X = m⁶v⁶A and **ON2e**; X¹ = gmm⁵U and X² = nm⁵U and b) **ON1m**; X¹ = m⁶v⁶A and X² = m⁶g⁶A, **ON2c**; X = vmm⁵U and **ON2f**; X = vmm⁵U.

11. Coupling reactions between ON2c and donor oligonucleotides with non-complementary sequences

The peptide coupling reactions were carried out under identical conditions to those described in Section 0.

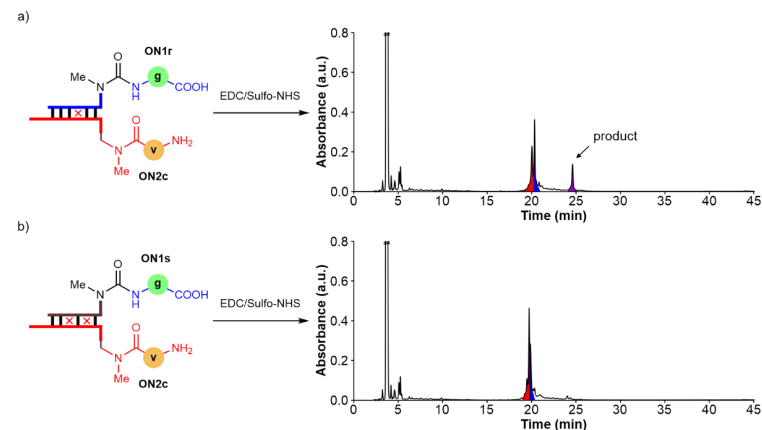


Figure S45. HPL-chromatograms of the reactions of **ON2c**; X = vmm⁵U with: a) **ON1r**; X = m⁶g⁶A and b) **ON1s**; X = m⁶g⁶A in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Table S22. Results obtained in the coupling reactions of **ON2c**; X = vmm⁵U with **ON1r**; X = m⁶g⁶A or **ON1s**; X = m⁶g⁶A using EDC/Sulfo-NHS as activator.

Donor strand	Acceptor strand	Yield (%) ^a
ON1r ; X = m ⁶ g ⁶ A	ON2c ; X = vmm ⁵ U	~14 (~35) ^b
ON1s ; X = m ⁶ g ⁶ A	ON2c ; X = vmm ⁵ U	< 3 (~12) ^b

^a Estimated yield from the chromatographic peak of the product using the calibration curve of **CON3**. Note that we assumed that the formed product features an extinction coefficient similar to that of **CON3**. ^b Using DMTMM-Cl as activator.

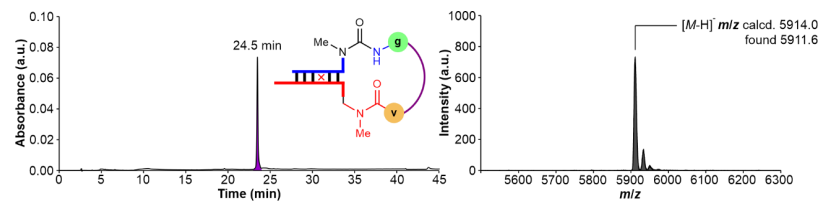
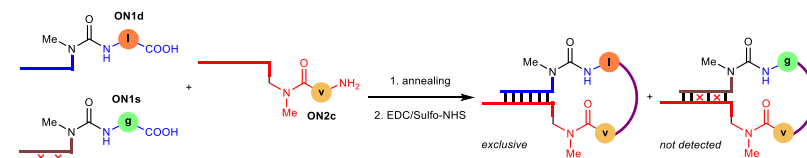


Figure S46. left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the isolated product from the reaction of **ON2c**; X = vmm⁵U with **ON1r**; X = m⁶g⁶A.



Scheme S12. Annealing and coupling reaction of **ON1d**; X = m⁶l⁶A, **ON1s**; X = m⁶g⁶A and **ON2c**; X = vmm⁵U. The formed peptide bonds are marked in purple.

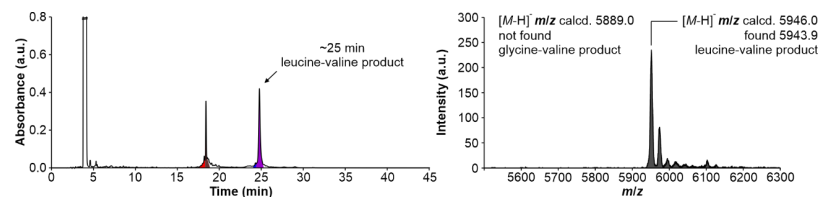


Figure S47. left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the reaction of **ON1d**; **X** = m^6g^6A , **ON1s**; **X** = m^5g^6A and **ON2c**; **X** = $\underline{v}mm^5U$ in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Table S23. Results obtained in the coupling reaction of **ON1d**; **X** = m^6g^6A , **ON1s**; **X** = m^5g^6A and **ON2c**; **X** = $\underline{v}mm^5U$ using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error of ly -peptide (%) ^a	Yield of gv -peptide (%)
ON1d ; X = m^6g^6A	ON2c ; X = $\underline{v}mm^5U$	65 \pm 2	not detected
ON1s ; X = m^5g^6A			

^a Calculated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

12. Coupling reactions between ON2c and donor oligonucleotides with different lengths

The peptide coupling reactions were carried out under identical conditions to those described in Section 0.

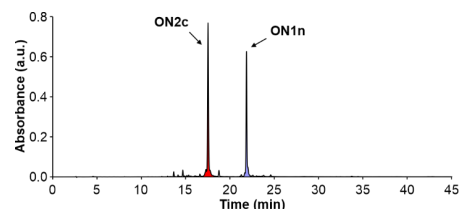


Figure S48. HPL-chromatogram of an equimolar mixture of **ON1n**; **X** = m^5v^6A and **ON2c**; **X** = $\underline{v}mm^5U$.

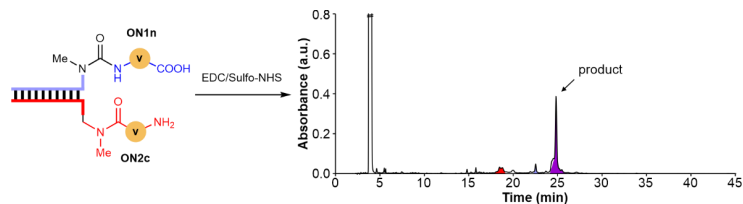


Figure S49. HPL-chromatogram of the reaction of **ON1n**; **X** = m^5v^6A with **ON2c**; **X** = $\underline{v}mm^5U$ in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Table S24. Result obtained in the coupling reaction of **ON1n**; **X** = m^5v^6A with **ON2c**; **X** = $\underline{v}mm^5U$ using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error (%) ^a
ON1n ; X = m^5v^6A	ON2c ; X = $\underline{v}mm^5U$	49 \pm 1

^a Calculated yield from the chromatographic peak of the product based on the total area of the initial components (Figure S48).

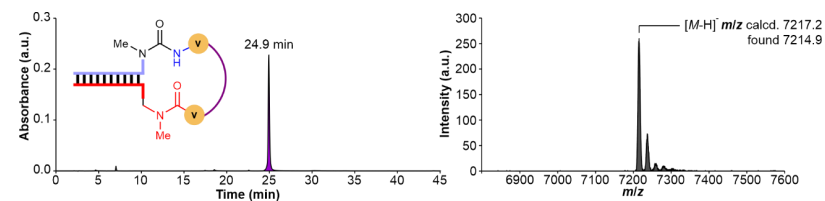
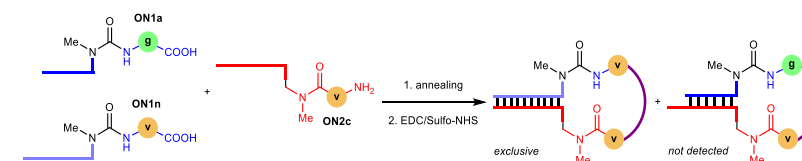


Figure S50. left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the isolated product from the reaction of **ON1n**; **X** = m^5v^6A with **ON2c**; **X** = $\underline{v}mm^5U$.



Scheme S13. Annealing and coupling reaction of **ON1a**; **X** = m^6g^6A , **ON1n**; **X** = m^5v^6A and **ON2c**; **X** = $\underline{v}mm^5U$. The formed peptide bonds are marked in purple.

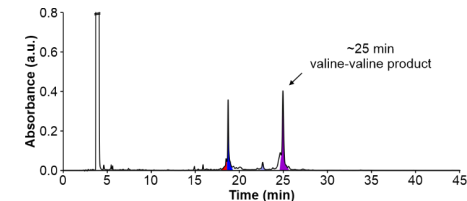


Figure S51. HPL-chromatogram of the reaction of **ON1a**; **X** = m^6g^6A , **ON1n**; **X** = m^5v^6A and **ON2c**; **X** = $\underline{v}mm^5U$ in MES buffer at pH 6 using EDC/Sulfo-NHS as activator.

Table S25. Results obtained in the coupling reaction of **ON1a**; **X** = m^6g^6A , **ON1n**; **X** = m^5v^6A and **ON2c**; **X** = $\underline{v}mm^5U$ using EDC/Sulfo-NHS as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield \pm Error of vy -peptide (%) ^a	Yield of gv -peptide (%) ^a
ON1a ; X = m^6g^6A	ON2c ; X = $\underline{v}mm^5U$	49 \pm 2	not detected
ON1n ; X = m^5v^6A			

^a Calculated yield from the chromatographic peak of the product based on the total area of the initial components (Figure S48).

13. Stability of selected acceptor oligonucleotides (ON2)

The oligonucleotide (0.5 nmol) was added to an Eppendorf tube. Buffer, NaCl and water were added to the ON's solution and the reaction was heated in a Thermocycler.

Concentration of the components in the reaction mixture: 10-50 μ M of oligonucleotide, 100 mM of buffer and 100 mM of NaCl (see figure footnotes for details).

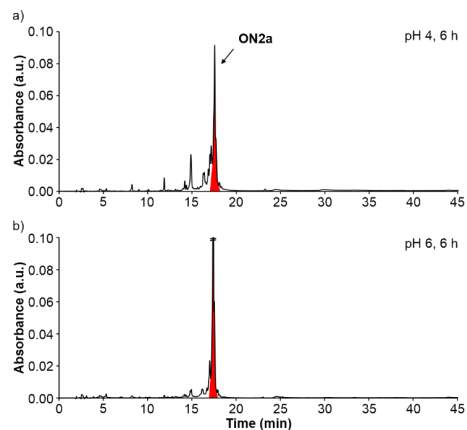


Figure S52. HPL-chromatograms of the stability of **ON2a**; X = mm⁵U in: a) acetate buffer at pH 4 and b) MES buffer at pH 6 after 6 h at 90°C.

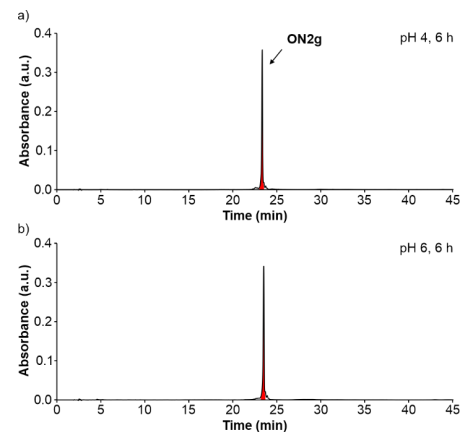


Figure S53. HPL-chromatograms of the stability of **ON2g**; X = mm⁵U in: a) acetate buffer at pH 4 and b) MES buffer at pH 6 after 6 h at 90°C.

Table S26. Results obtained in the stability of **ON2a** and **ON2g** (average of, at least, two experiments).^a

pH	Time (h)	Average Amount \pm Error (%)	
		ON2a	ON2g
4	6	40 \pm 3	>95
6	6	70 \pm 5	>95

^a Calculated amounts from the chromatographic peaks using the corresponding calibration curves.

14. Cleavage of urea in selected oligonucleotides and cyclic peptide products

The cleavage reactions were carried out under identical conditions to those described in Section 13.

14.1 Cleavage reactions of ON1c (m⁶v⁶A) and ON1k (v⁶A) at pH 5

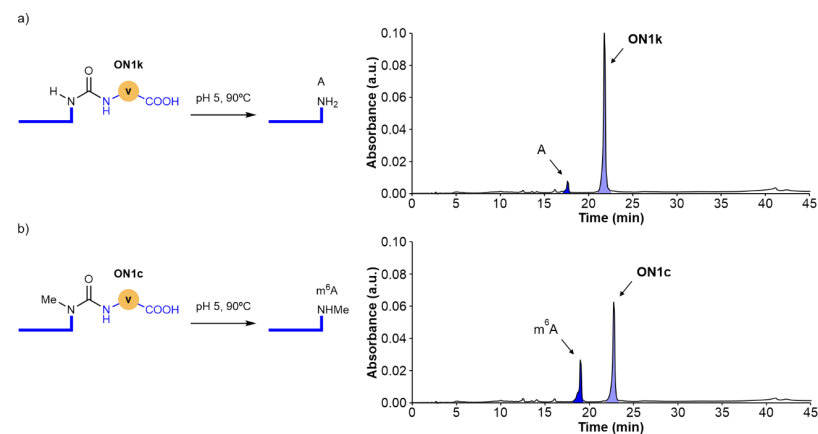


Figure S54. HPL-chromatograms of the cleavage reactions of: a) **ON1k**; X = v⁶A and b) **ON1c**; X = m⁶v⁶A in acetate buffer at pH 5 after 12 h at 90°C.

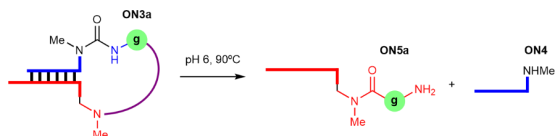
These experiments indicated that the urea cleavage reaction of the unmethylated aa⁶A-RNA donor strand **ON1k** was slower than that of the methylated version, m⁶aa⁶A-RNA **ON1c**.

Table S27. Results obtained in the cleavage reactions of **ON1c** and **ON1k** (average of, at least, two experiments).^a

pH	Time (h)	Average Amount \pm Error (%)	
		ON1k	A-strand
5	12	85 \pm 3	10 \pm 1
		ON1c	m ⁶ A-strand
		65 \pm 1	20 \pm 1

^a Calculated amounts from the chromatographic peaks using the corresponding calibration curves.

14.2 Cleavage reaction of ON3a (m⁶g⁶A coupled with mmm⁵U)



Scheme S14. Cleavage of urea in ON3a. The peptide bond is marked in purple.

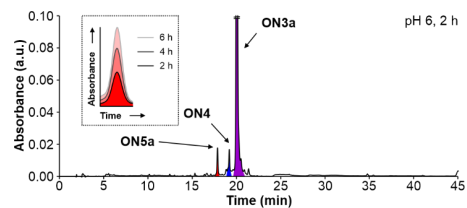


Figure S55. HPL-chromatogram of the cleavage reaction of ON3a in MES buffer at pH 6 after 2 h at 90°C. Inset shows the selected region of the HPL-chromatograms after 2, 4 and 6 h.

Table S28. Results obtained in the cleavage reaction of ON3a (average of, at least, two experiments).^a

pH	Time (h)	Average Amount ± Error (%)		
		ON3a	ON4 (m ⁶ A)	ON5a (gmm ⁵ U)
6	6	75±2	15±1	15±1 (t _R = 17.5 min)

^a Calculated amounts from the chromatographic peaks using the corresponding calibration curves.

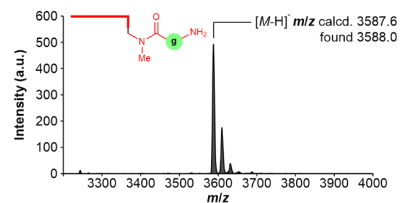


Figure S56. MALDI-TOF mass spectrum (negative mode) of the isolated ON5a (gmm⁵U).

Additional experiments at pH 4 and pH 6

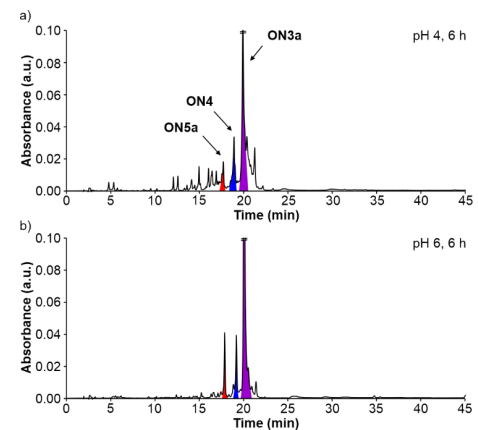


Figure S57. HPL-chromatograms of the cleavage reactions of ON3a in: a) acetate buffer at pH 4 and b) MES buffer at pH 6 after 6 h at 90°C.

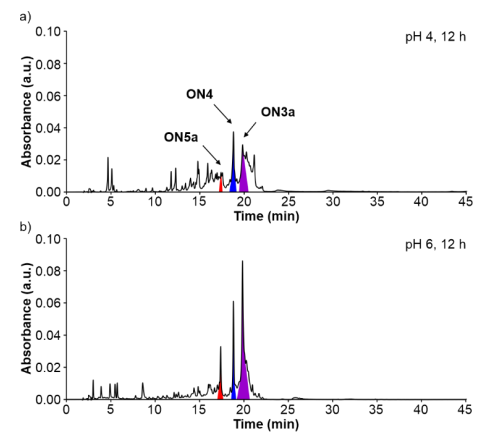


Figure S58. HPL-chromatograms of the cleavage reactions of ON3a in: a) acetate buffer at pH 4 and b) MES buffer at pH 6 after 12 h at 90°C.

Table S29. Results obtained in the cleavage reaction of ON3a (average of, at least, two experiments).^a

pH	Time (h)	Average Amount ± Error (%)
		ON5a (gmm ⁵ U)
4	6	10±2
	12	n.d.
6	6	15±1
	12	10±1

^a Calculated amounts from the chromatographic peak using the calibration curve of CON2. n.d. = not determined.

14.3 Cleavage reactions of ON3c (m⁶g⁶A coupled with vmm⁵U)

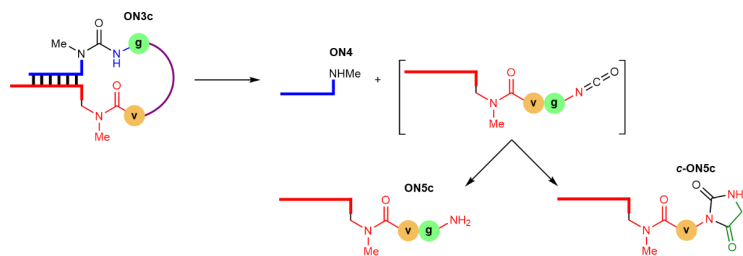


Figure S59. Cleavage of urea in ON3c. The peptide bond is marked in purple.

Cleavage reactions at 60°C

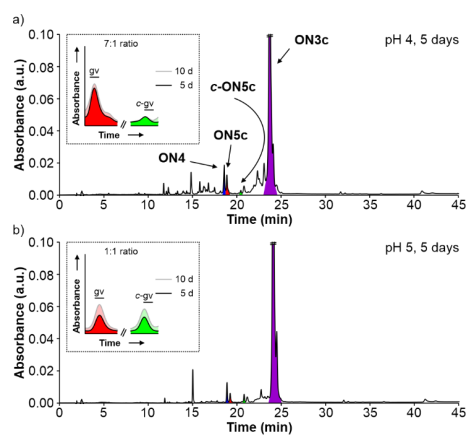


Figure S60. HPL-chromatograms of the cleavage reactions of ON3c in acetate buffer at: a) pH 4 and b) pH 5 after 5 days at 60°C. Inset shows the selected region of the HPL-chromatograms after 5 and 10 days.

Table S30. Results obtained in the cleavage reactions of ON3c at 60°C (average of, at least, two experiments).^a

pH	Time (days)	Average Amount ± Error (%)				Ratio (ON5c/c-ON5c)
		ON3c	ON4 (m ⁶ A)	ON5c (gvmm ⁵ U)	c-ON5c (c-gvmm ⁵ U)	
4	10	50±2	10.5±1	9±1 (t _r = 19.5 min)	1.5±0.5 (t _r = 21.0 min)	~7:1
5	10	80±3	6±1	3±1	3±1	~1:1

^a Calculated amounts from the chromatographic peaks using the corresponding calibration curves.

Cleavage reactions at 90°C

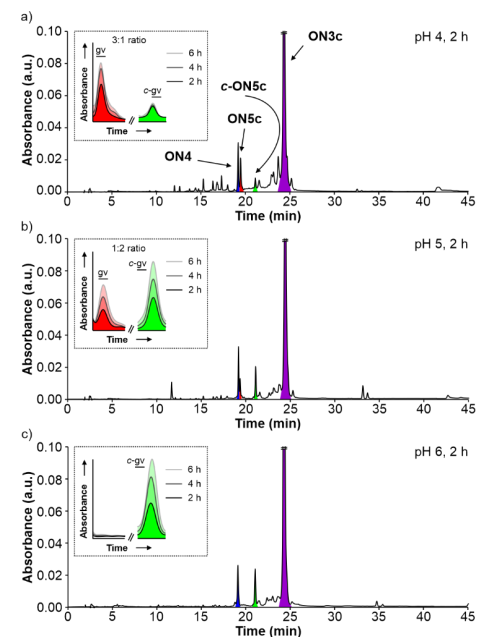


Figure S61. HPL-chromatograms of the cleavage reactions of ON3c in: a) acetate buffer at pH 4; b) acetate buffer at pH 5 and c) MES buffer at pH 6 after 2 h at 90°C. Inset shows the selected region of the HPL-chromatograms after 2, 4 and 6 h.

Table S31. Results obtained in the cleavage reactions of ON3c at 90°C (average of, at least, two experiments).^a

pH	Time (h)	Average Amount ± Error (%)				Ratio (ON5c/c-ON5c)
		ON3c	ON4 (m ⁶ A)	ON5c (gvmm ⁵ U)	c-ON5c (c-gvmm ⁵ U)	
4	6	30±3	20±2	15±2 (t _r = 19.5 min)	5±1 (t _r = 21.0 min)	~3:1
5	6	55±3	25±2	8±1	17±3	~1:2
6	6	60±2	25±1	-	25±1	-

^a Calculated amounts from the chromatographic peaks using the corresponding calibration curves.

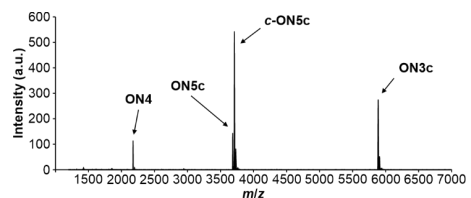


Figure S62. MALDI-TOF mass spectrum (negative mode) of the cleavage reaction of **ON3c** in acetate buffer at pH 5 after 2 h at 90°C. A similar MALDI-TOF mass spectrum was obtained at pH 4. The indicated peaks correspond to the $[M-H]^-$ ions.

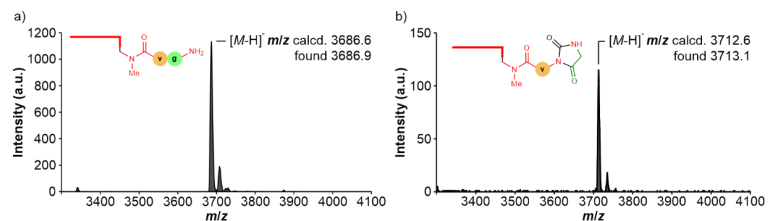
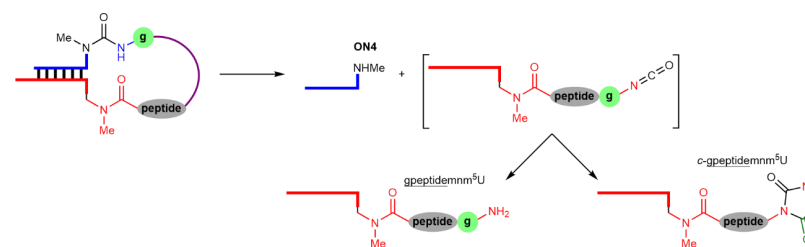


Figure S63. MALDI-TOF mass spectrum (negative mode) of the isolated: a) **ON5c** ($gymnm^5U$) and b) **c-ON5c** ($c-gymnm^5U$).

14.4 Cleavage reactions of peptide-oligonucleotides at pH 4



Scheme S15. Cleavage of urea in peptide-oligonucleotides. The peptide bond is marked in purple.

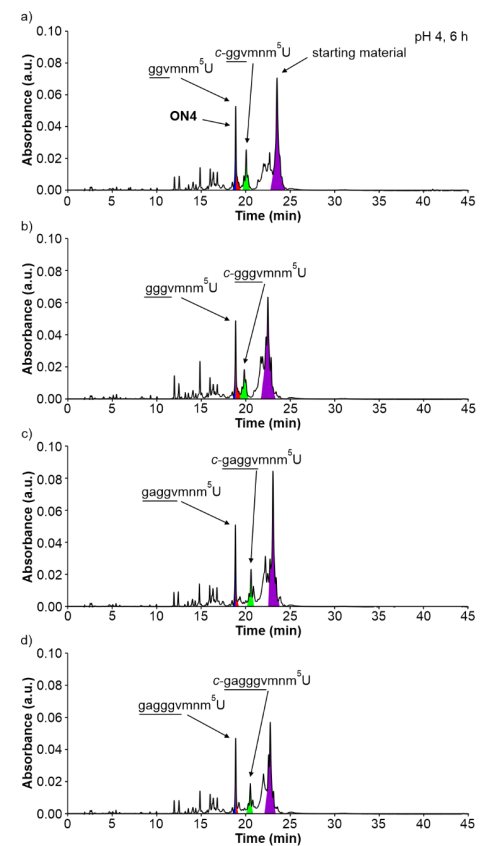


Figure S64. HPL-chromatograms of the cleavage reactions of peptide-oligonucleotides (Section 8.2) in acetate buffer at pH 4 to give: a) $gymnm^5U$; b) $ggymnm^5U$; c) $gagymnm^5U$ and d) $gaggymnm^5U$ oligonucleotides, together with hydantoin side products, after 6 h at 90°C.

The 3'-H₂N-peptidemnm⁵U-RNA-5' and m⁵A products overlap in the HPL-chromatograms. Therefore, they were isolated as a mixture in a single fraction.

Table S32. Results obtained in the cleavage reactions of peptide-oligonucleotides (Section 8.2).^a

3'-H ₂ N-peptidemnm ⁵ U-RNA-5'	Amount (%)
3'-ggvmm ⁵ U-RNA-5'	~12
3'-gggvmm ⁵ U-RNA-5'	~10
3'-gaggvmm ⁵ U-RNA-5'	~10
3'-gaggvmm ⁵ U-RNA-5'	~10

^a Estimated amounts assuming that the 3'-H₂N-peptidemnm⁵U-RNA-5' products and the hydantoin counterparts were formed in a similar extent.

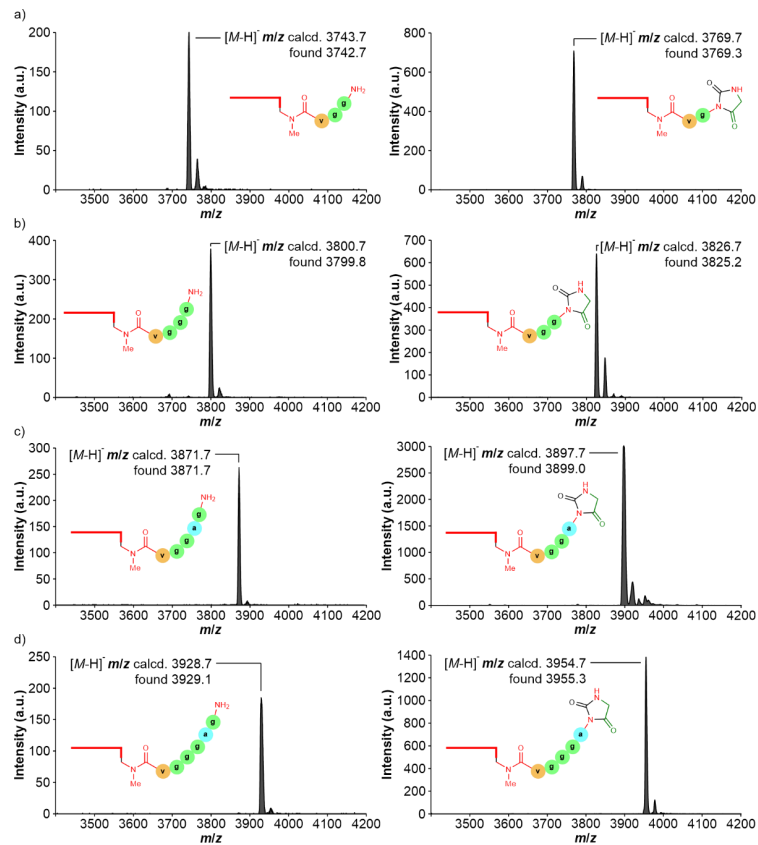


Figure S65. MALDI-TOF mass spectra (negative mode) of the isolated: a) ggvmm⁵U; b) gggvmm⁵U; c) gaggvmm⁵U and d) gaggvmm⁵U oligonucleotides (left) and hydantoin side products (right). Note that the analyzed 3'-H₂N-peptidemnm⁵U-RNA-5' samples (left) contained the m⁵A product (*m/z* region not shown).

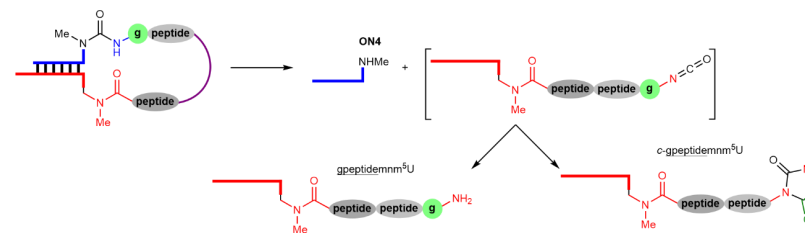


Figure S66. Cleavage of urea in gpeptide-peptidemnm⁵U-oligonucleotides. The peptide bond is marked in purple.

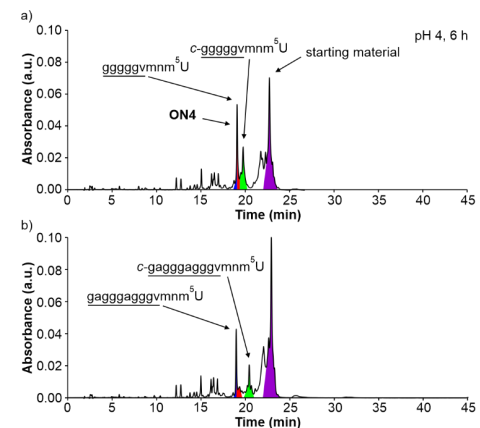


Figure S67. HPL-chromatograms of the cleavage reactions of peptide-oligonucleotides (Section 8.3) in acetate buffer at pH 4 to give: a) ggggvmm⁵U and b) gaggvmm⁵U oligonucleotides, together with hydantoin side products, after 6 h at 90°C.

The 3'-H₂N-peptidemnm⁵U-RNA-5' and m⁵A products overlap in the HPL-chromatograms. In addition, the 3'-H₂N-peptidemnm⁵U-RNA-5' in a) overlaps with the hydantoin side product.

Table S33. Results obtained in the cleavage reactions of peptide-oligonucleotides (Section 8.3).^a

3'-H ₂ N-peptidemnm ⁵ U-RNA-5'	Amount (%)
3'-ggggvmm ⁵ U-RNA-5'	~10
3'-gaggaggvmm ⁵ U-RNA-5'	~9

^a Estimated amounts assuming that the 3'-H₂N-peptidemnm⁵U-RNA-5' products and the hydantoin counterparts were formed in a similar extent.

One pot reaction

The one pot reaction was performed with 15 nmol of **ON2g** as starting acceptor strand. 15 nmol of donor strand **ON1a** or **ON1g** were added for each coupling reaction. After each coupling reaction and the second cleavage, the crude was filtered using an Amicon® ultra centrifugal filter (3 kDa Nominal Molecular Weight Cut-Off) to remove the remaining activator and exchange the buffer solution. The volume of the solution was maintained constant throughout the five reaction steps. 20 µL of the crude (1 nmol) were analyzed by HPLC after the second coupling, the second cleavage and the third coupling reactions.

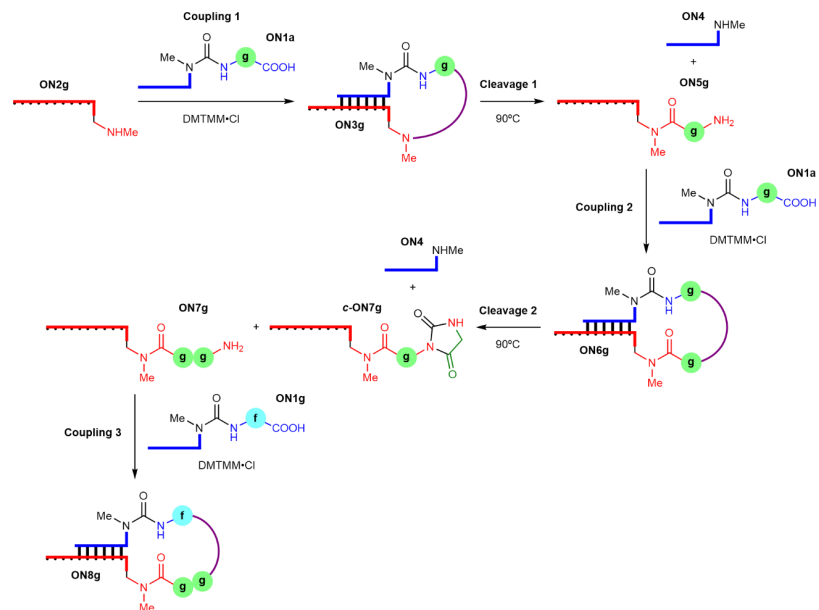


Figure S70. Coupling and cleavage of **ON1a**; X = m⁶g⁶A and **ON1g**; X = m⁶f⁶A with **ON2g**. The formed peptide bond is marked in purple.

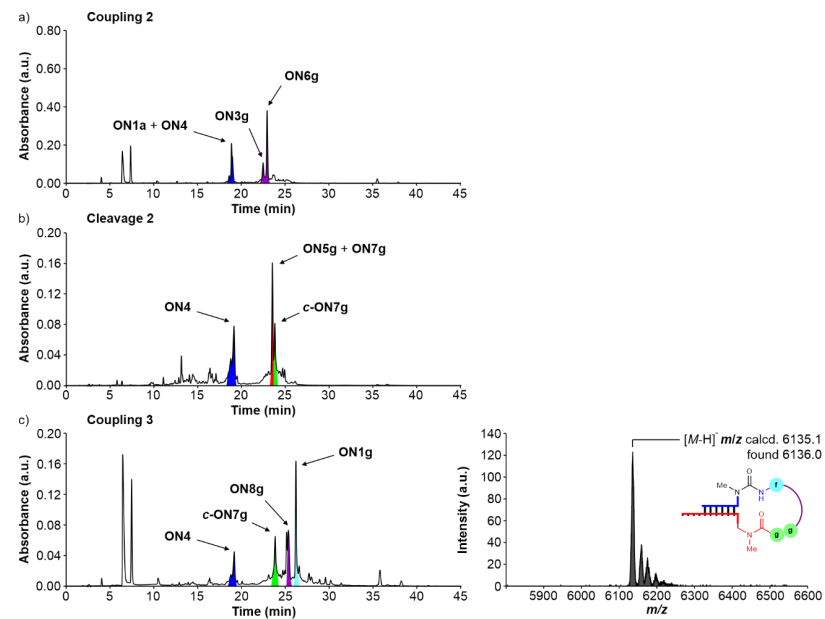


Figure S71. HPLC-chromatograms of the one pot reaction of **ON1a**; X = m⁶g⁶A and **ON1g**; X = m⁶f⁶A with **ON2g**: a) coupling 2; b) cleavage 2 and c) coupling 3.

Table S35. Results obtained in the one pot reaction of **ON1a**; X = m⁶g⁶A and **ON1g**; X = m⁶f⁶A with **ON2g**.

Steps	Activators	pH	T (°C)	Time (h)	Yield (%) ^a
Coupling 2 (ON6g)	DMTMM-Cl	6	25	24	~36 in three steps
Cleavage 2 (ON5g + ON7g)	-	4	90	24	23 in four steps
Coupling 3 (ON8g)	DMTMM-Cl	6	25	24	~10 in five steps

^a Calculated/estimated amounts from the chromatographic peaks using the corresponding calibration curves.

15.2 Coupling and cleavage reactions of **ON1o** (m⁶g⁶Am) with **ON2h**

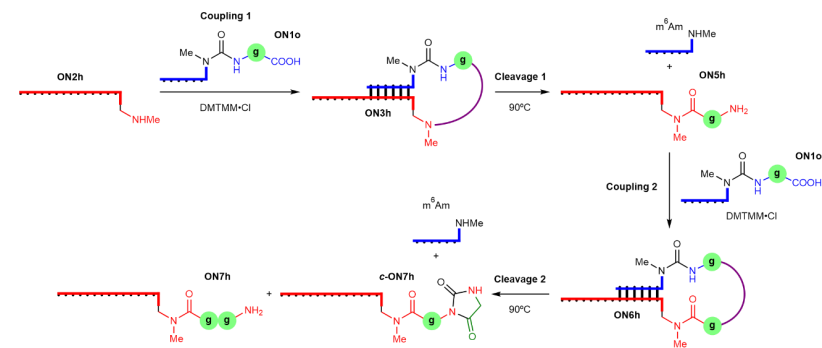


Figure S72. Coupling and cleavage of **ON1o**; X = m⁶g⁶Am with **ON2h**. The formed peptide bond is marked in purple.

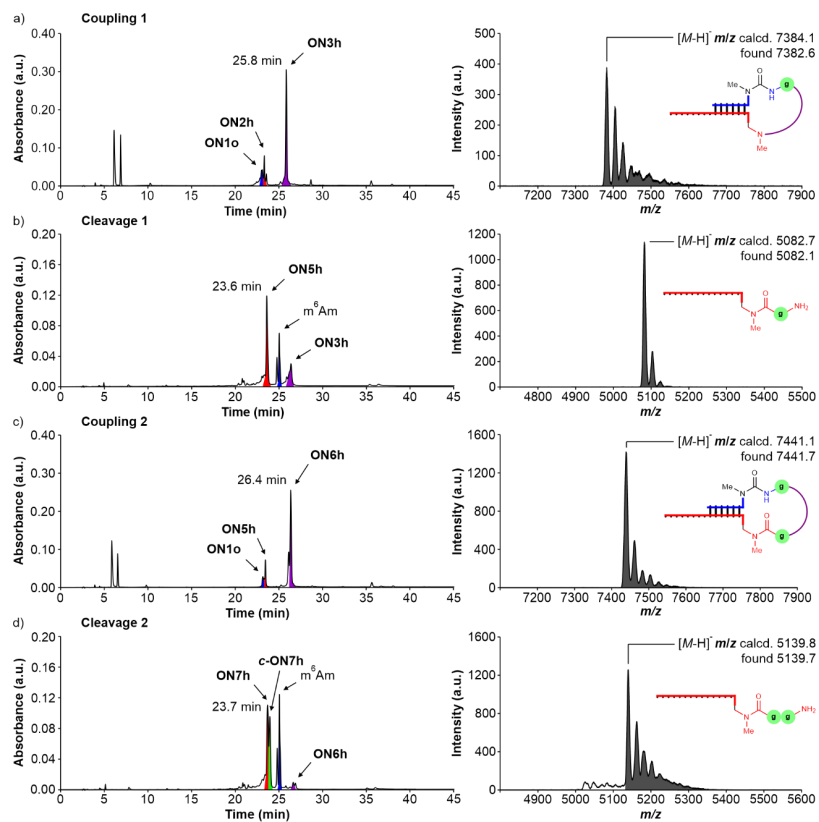


Figure S73. Left) HPL-chromatograms of the reactions of **ON1o**; X = m⁶g⁶Am with **ON2h**: a) coupling 1; b) cleavage 1; c) coupling 2 and d) cleavage 2. The product of each step was separated by HPLC and added into the next reaction. Right) MALDI-TOF mass spectra (negative mode) of the isolated products from the reactions a)-d).

Table S36. Results obtained in the coupling and cleavage reactions of **ON1o**; X = m⁶g⁶Am with **ON2h**.

Steps	Activators	pH	T (°C)	Time (h)	Yield (%) ^a
Coupling 1 (ON3h)	DMTMM-Cl	6	25	24	46
Cleavage 1 (ON5h)	-	4	90	48	30
Coupling 2 (ON6h)	DMTMM-Cl	6	25	24	41
Cleavage 2 (ON7h)	-	4	90	48	28

^a Calculated amounts from the chromatographic peaks using the corresponding calibration curves.

15.3 Coupling and cleavage reactions of donor and acceptor-peptide oligonucleotides

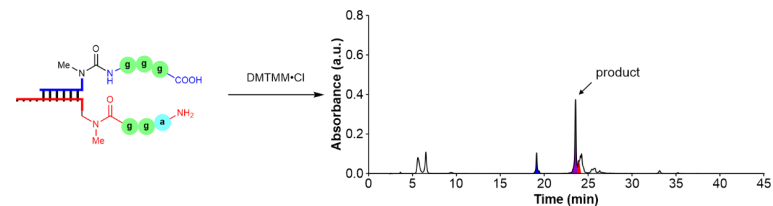


Figure S74. HPL-chromatogram of the reaction of 5'-m⁶(ggg)⁶A-RNA-3' with 3'-aggmm⁶U-RNA-5' containing 2'-OMe nucleosides in MES buffer at pH 6 using DMTMM-Cl as activator.

Table S37. Result obtained in the coupling reaction of peptide-modified donor and acceptor oligonucleotides using DMTMM-Cl as activator.

Donor strand	Acceptor strand	Yield (%) ^a
5'-m ⁶ (ggg) ⁶ A-RNA-3'	3'-aggmm ⁶ U-RNA-5'	~50

^a Estimated yield from the chromatographic peak of the product using the calibration curve of **CON3**.

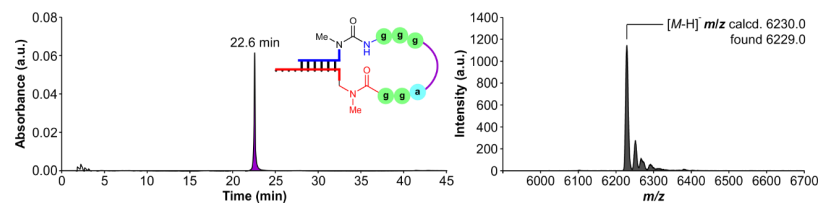
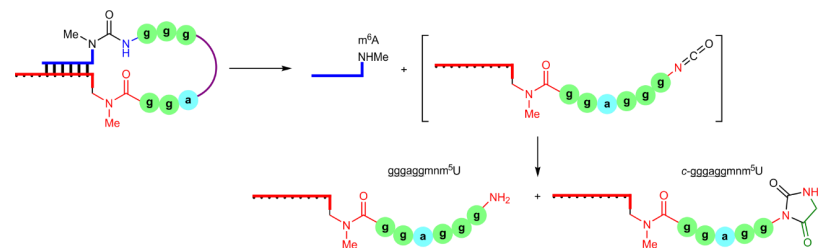


Figure S75. Left) HPL-chromatogram and right) MALDI-TOF mass spectrum (negative mode) of the isolated product from the reaction of 5'-m⁶(ggg)⁶A-RNA-3' with 3'-aggmm⁶U-RNA-5' containing 2'-OMe nucleosides.



Scheme S17. Cleavage of urea in peptide-peptidemm⁶U-oligonucleotide. The peptide bond is marked in purple.

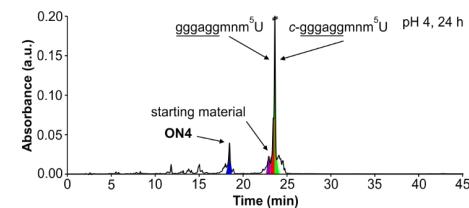


Figure S76. HPL-chromatogram of the cleavage reaction of peptide-oligonucleotide in acetate buffer at pH 4 to give gggaggmm⁶U and c-gggaggmm⁶U oligonucleotides after 24 h at 90°C.

The 3'-H₂N-peptidemm⁶U-RNA-5' and hydantoin side products overlap in the HPL-chromatogram.

Table S38. Result obtained in the cleavage reaction of peptide-oligonucleotide.^a

Product oligonucleotides containing 2'-OMe nucleosides	Amount (%)
3'-ggagagmm ⁵ U-RNA-5' and 3'-c-gggagagmm ⁵ U-RNA-5'	~85 (t _R = 23.6 min)

^a Estimated amount from the chromatographic peak using the calibration curve of **CON2**.

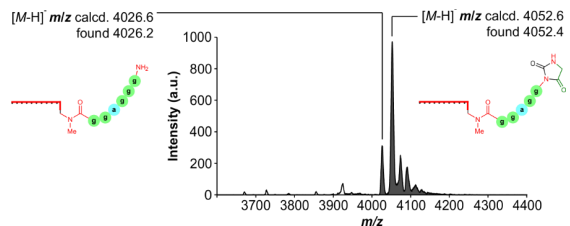


Figure S77. MALDI-TOF mass spectra (negative mode) of the isolated gggagagmm⁵U and hydantoin side product.

15.4 Coupling reactions between ON2g and donor oligonucleotides of different length

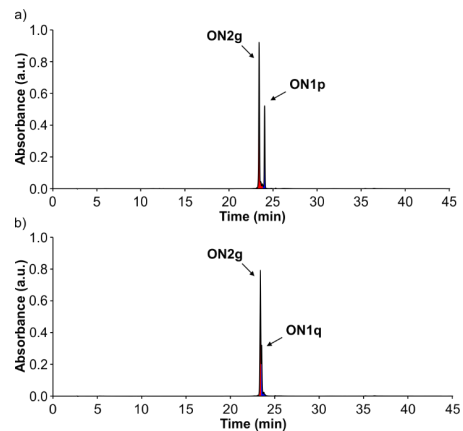
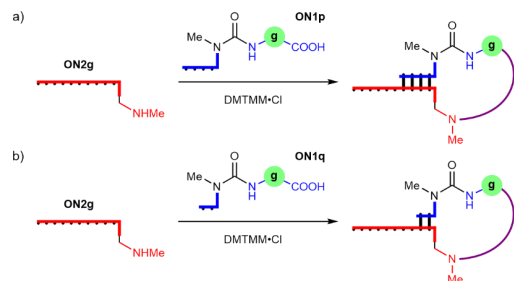


Figure S78. HPL-chromatograms of an equimolar mixture of **ON2g**; **X** = mnm⁵U with: a) **ON1p**; **X** = m⁶g⁶Am and b) **ON1q**; **X** = m⁶g⁶Am.



Scheme S18. Coupling of **ON2g**; **X** = mnm⁵U with: a) **ON1p**; **X** = m⁶g⁶Am and b) **ON1q**; **X** = m⁶g⁶Am. The formed peptide bond is marked in purple.

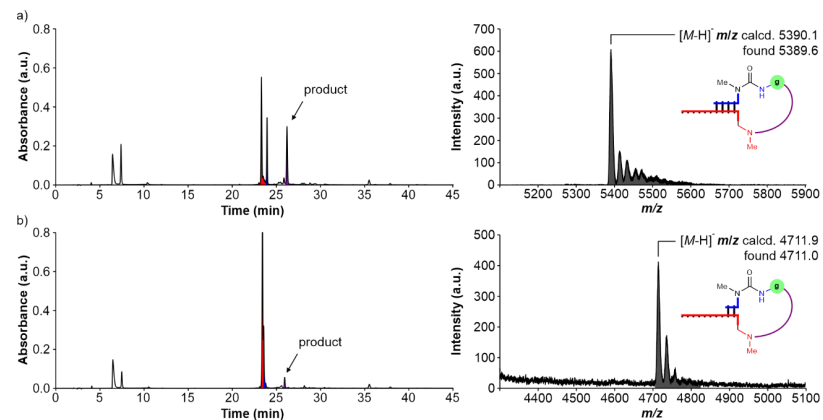


Figure S79. left) HPL-chromatograms of the reactions of: a) **ON1p**; **X** = m⁶g⁶Am and b) **ON1q**; **X** = m⁶g⁶Am with **ON2g**; **X** = mnm⁵U in MES buffer at pH 6 using DMTMM-Cl as activator. The reaction b) was carried out at 0°C using 1 M NaCl. right) MALDI-TOF mass spectra (negative mode) of the isolated products.

Table S39. Results obtained in the coupling reactions of **ON2g**; **X** = mnm⁵U with **ON1p**; **X** = m⁶g⁶Am or **ON1q**; **X** = m⁶g⁶Am using DMTMM-Cl as activator (average of, at least, two experiments).

Donor strand	Acceptor strand	Average Yield ± Error (%) ^a
ON1p ; X = m ⁶ g ⁶ Am	ON2g ; X = mnm ⁵ U	19±2 (t _R = 26.2 min)
ON1q ; X = m ⁶ g ⁶ Am	ON2g ; X = mnm ⁵ U	5±1 ^b (t _R = 26.0 min)

^a Calculated yield from the chromatographic peak of the product based on the total area of the initial components (Figure S78). ^b Using 1 M NaCl at 0°C.

16. Determination of melting temperatures by UV spectroscopic experiments

The UV melting curves were measured on a JASCO V-650 spectrometer at 260 nm using 10 mm QS cuvettes with a scanning rate of $1^{\circ}\text{C}\cdot\text{min}^{-1}$. The obtained UV spectroscopic data were fit to the corresponding function to determine the melting temperature/s.

For double strands of non-self-complementary oligonucleotides, the data were fit to a two-state melting model, *i.e.* double strand – random coil equilibrium, using a mono-sigmoidal Boltzmann function.¹⁴ On the contrary, the data were fit to a three-state melting model, *i.e.* double strand – hairpin – random coil equilibria, for single strands of self-complementary oligonucleotides using a double-sigmoidal Boltzmann function.^{15,16}

For the experiments, we prepared aqueous solutions containing equimolar amounts of the oligonucleotides (5 μM), 10 mM phosphate buffer at pH 7 and 150 mM NaCl. The oligonucleotides were annealed by heating to 95°C for 4 min and, subsequently, by cooling down slowly to 5°C before the variable-temperature UV spectroscopic experiment.

16.1 Melting temperature of a double strand from canonical oligonucleotides

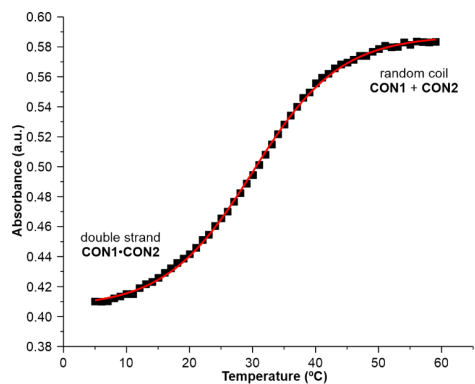


Figure S80. Melting curve of **CON1** and **CON2**. Line shows the fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function. $T_m = 30.1^{\circ}\text{C}$.

16.2 Melting temperatures of double strands from donor and acceptor oligonucleotides

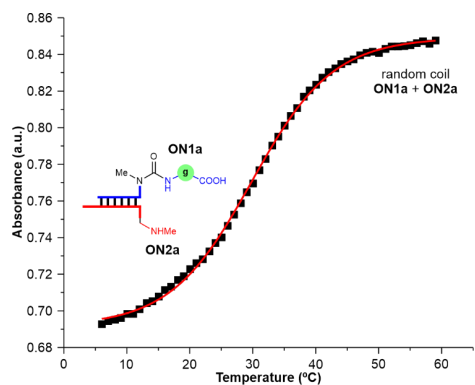


Figure S81. Melting curve of **ON1a**; $X = m^6g^6A$ and **ON2a**; $X = mnm^5U$. Line shows the fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function. $T_m = 30.4^{\circ}\text{C}$.

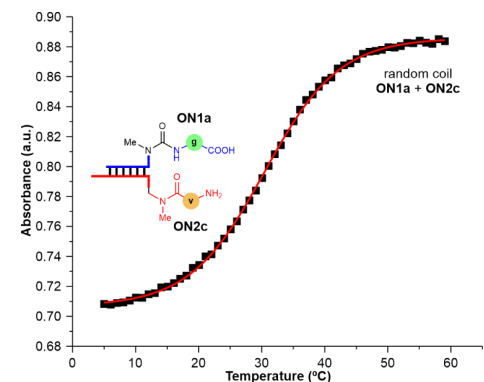


Figure S82. Melting curve of **ON1a**; $X = m^6g^6A$ and **ON2c**; $X = \underline{y}mm^5U$. Line shows the fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function. $T_m = 30.5^{\circ}\text{C}$.

The melting temperatures of the double strands containing modified A and U bases, **ON1a**; $X = m^6g^6A$, **ON2a**; $X = mnm^5U$ and **ON2c**; $X = \underline{y}mm^5U$, were very similar to those determined for canonical oligonucleotides, **CON1** and **CON2**.

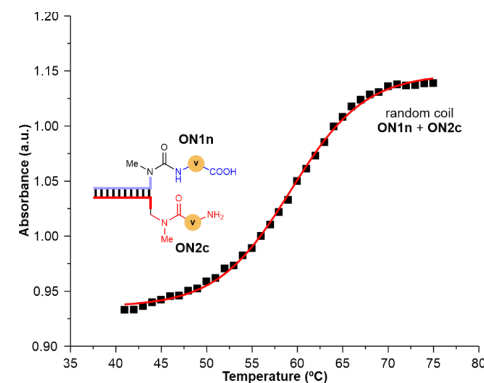


Figure S83. Melting curve of **ON1n**; $X = m^6v^6A$ and **ON2c**; $X = \underline{y}mm^5U$. Line shows the fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function. $T_m = 59.2^{\circ}\text{C}$.

16.3 Melting temperatures of double strands from donor and acceptor peptide-oligonucleotides

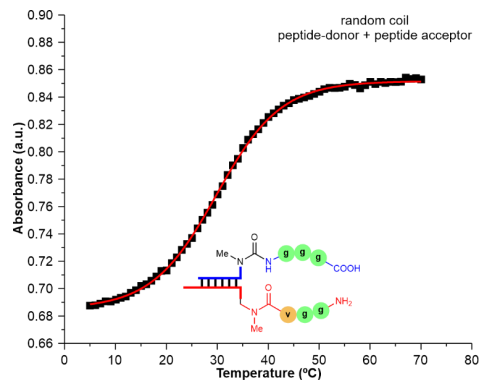


Figure S84. Melting curve of 5'-m⁶(ggg)⁶A-RNA-3' with 3'-ggvmm⁵U-RNA-5'. Line shows the fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function. $T_m = 30.0^\circ\text{C}$.

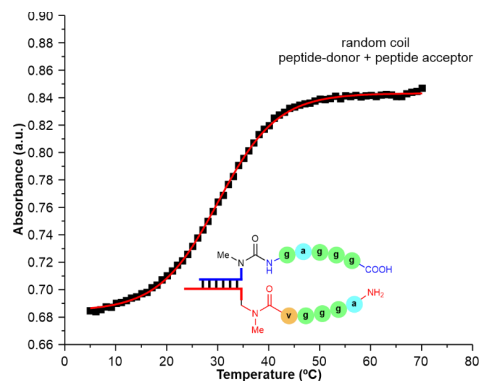


Figure S85. Melting curve of 5'-m⁶(gaggg)⁶A-RNA-3' with 3'-aggvmm⁵U-RNA-5'. Line shows the fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function. $T_m = 29.8^\circ\text{C}$.

16.4 Melting temperatures of selected cyclic peptide products

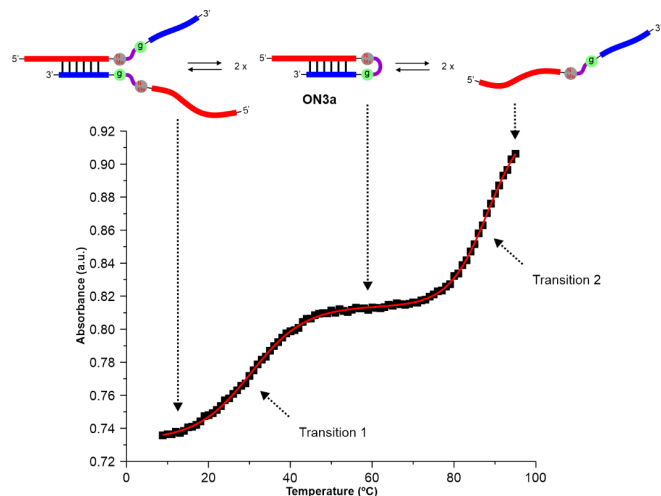


Figure S86. Melting curve of ON3a. Line shows the fit of the data to a three-state melting model using a double-sigmoidal Boltzmann function. $T_{m1} = 30.8^\circ\text{C}$ and $T_{m2} = 87.5^\circ\text{C}$. Top panel shows representation of the three states involved in the two transitions.

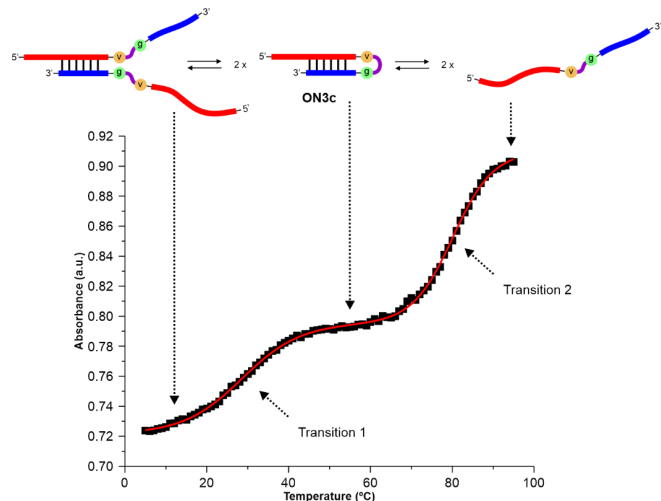


Figure S87. Melting curve of ON3c. Line shows the fit of the data to a three-state melting model using a double-sigmoidal Boltzmann function. $T_{m1} = 28.4^\circ\text{C}$ and $T_{m2} = 80.1^\circ\text{C}$. Top panel shows representation of the three states involved in the two transitions.

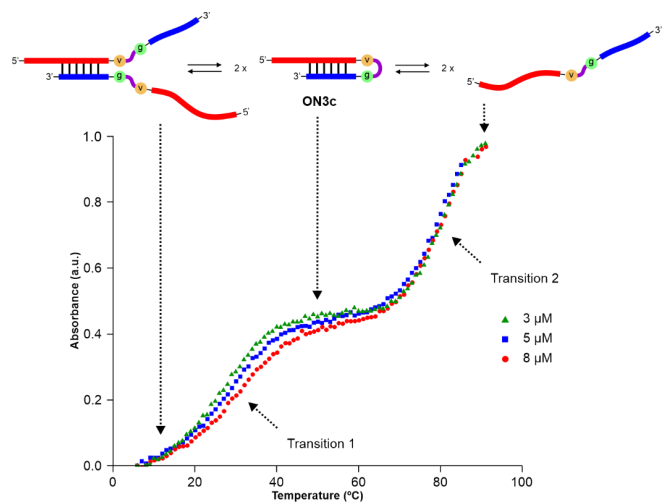
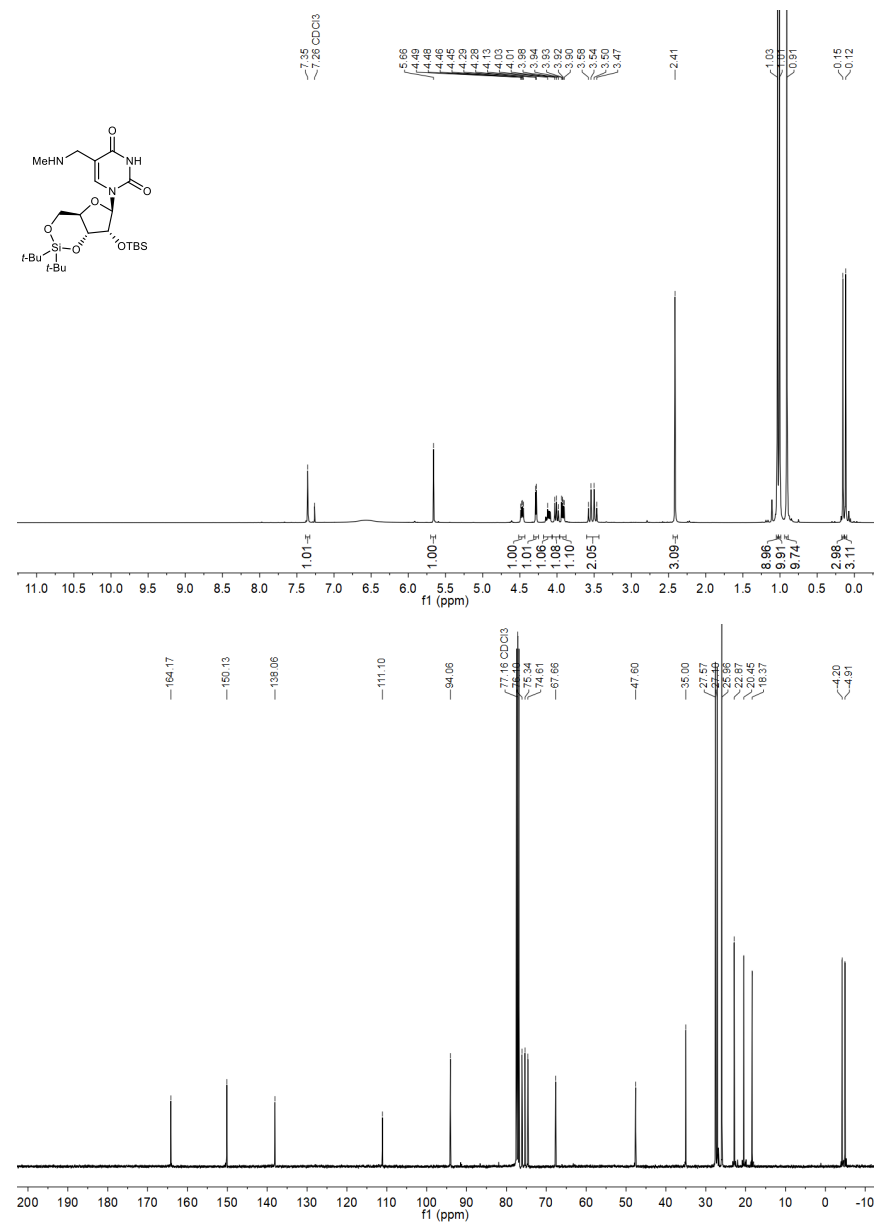


Figure S88. Normalized melting curves of **ON3c** at 3, 5 and 8 μM concentration. Top panel shows representation of the three states involved in the two transitions.

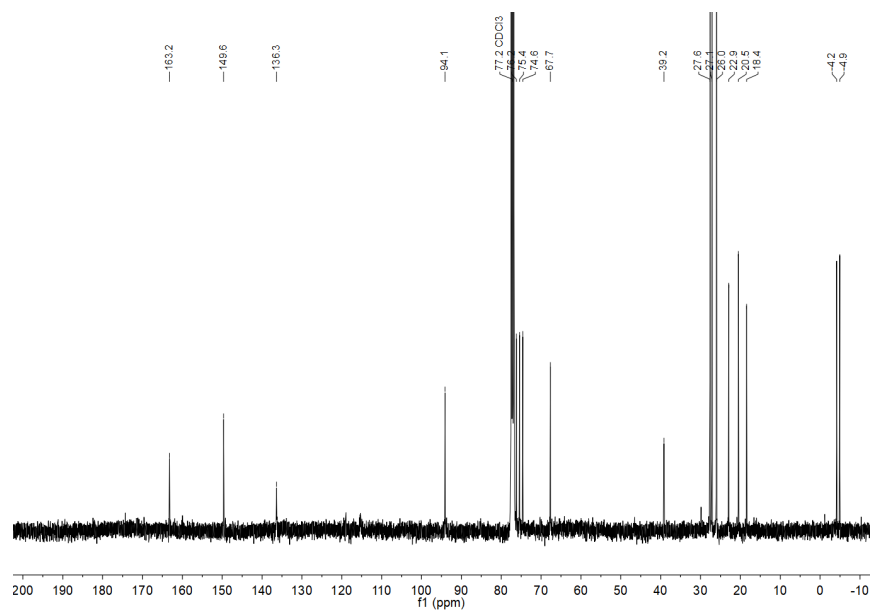
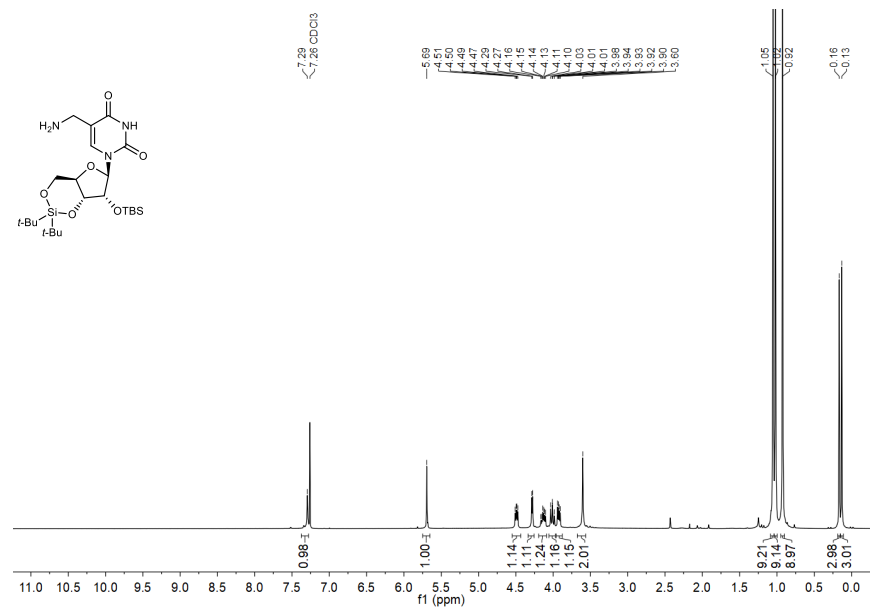
We observed a two-step melting profile in the experiments performed with the RNA oligonucleotides **ON3a** and **ON3c** (Figure S86, Figure S87 and Figure S88). At low temperature (transition 1), the double strand (duplex) is transformed into the hairpin. At high temperature (transition 2), the hairpin is converted into the random coil. The intermolecular and intramolecular dissociation of the base pairs, *i.e.* breaking of hydrogen-bonding and π -stacking interactions, is induced by the increase in temperature over the course of the experiments.^{15,16}

17. NMR spectra of synthesized compounds

^1H and $^{13}\text{C}\{^1\text{H}\}$ NMR spectra of compound 3a

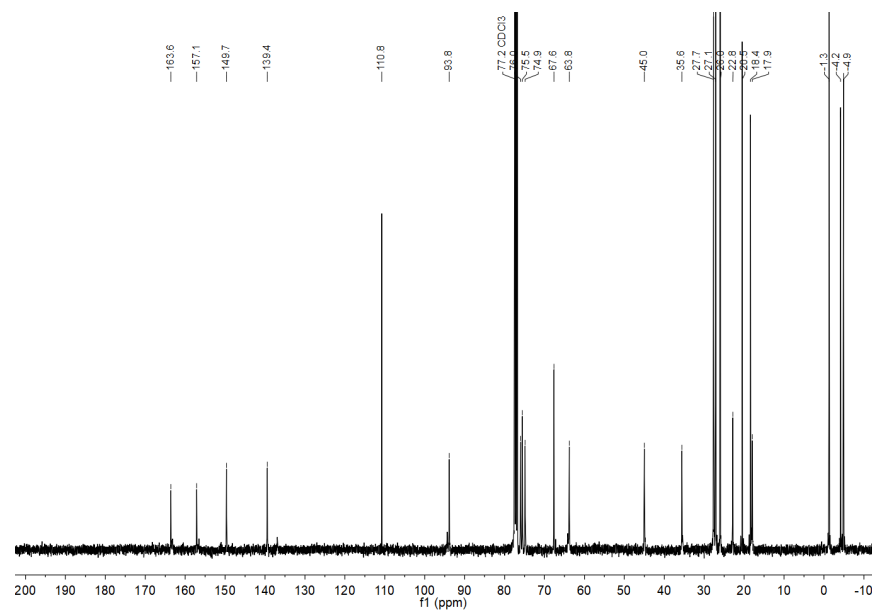
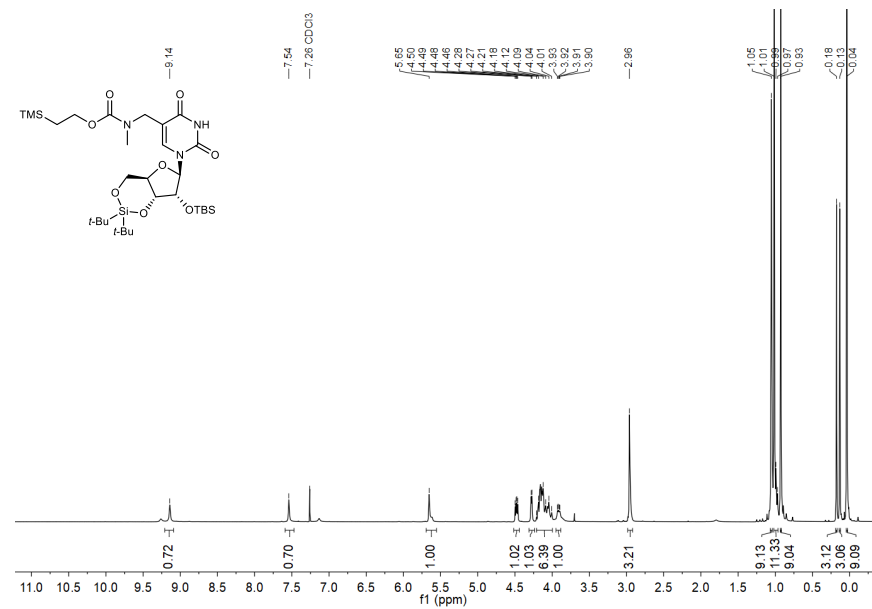


¹H and ¹³C{¹H} NMR spectra of compound 3b



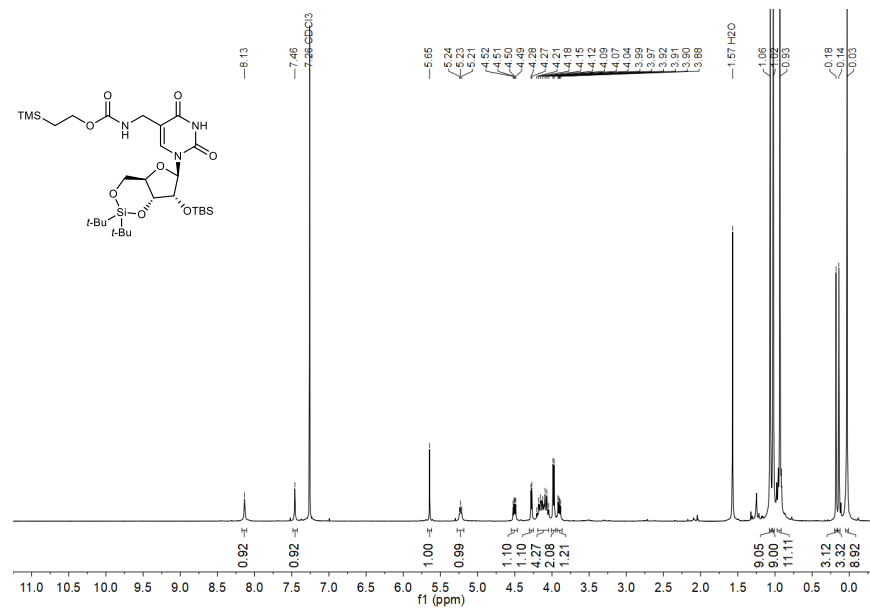
S82

¹H and ¹³C{¹H} NMR spectra of compound 4a



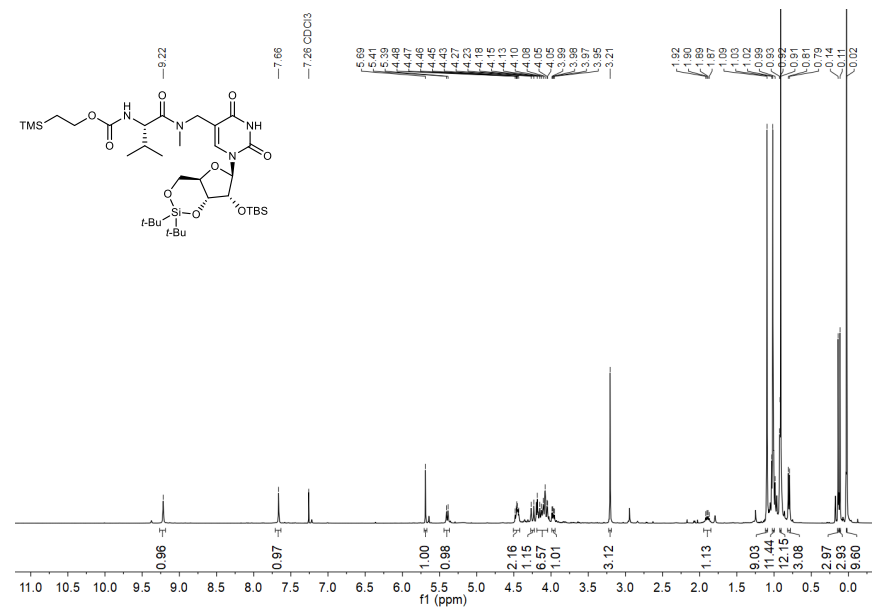
S83

¹H and ¹³C(¹H) NMR spectra of compound 4b



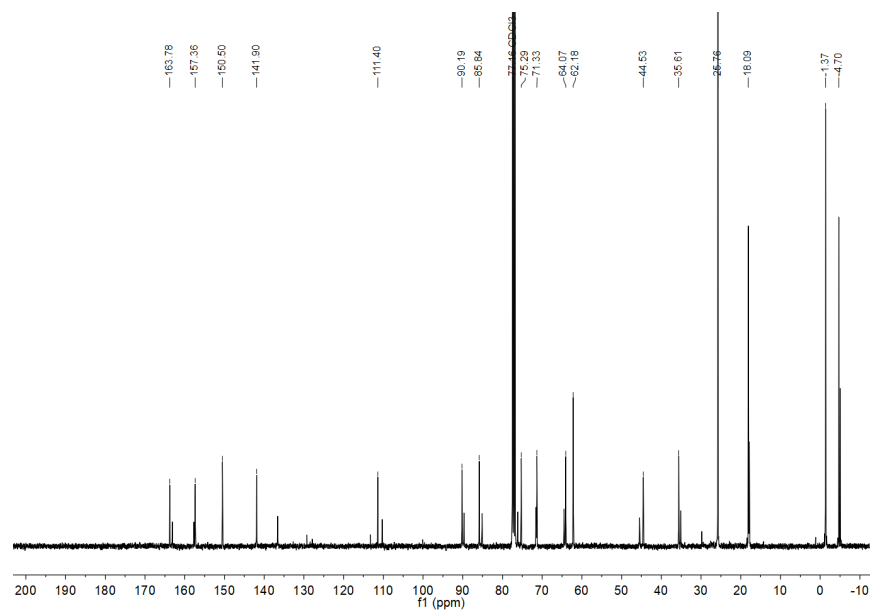
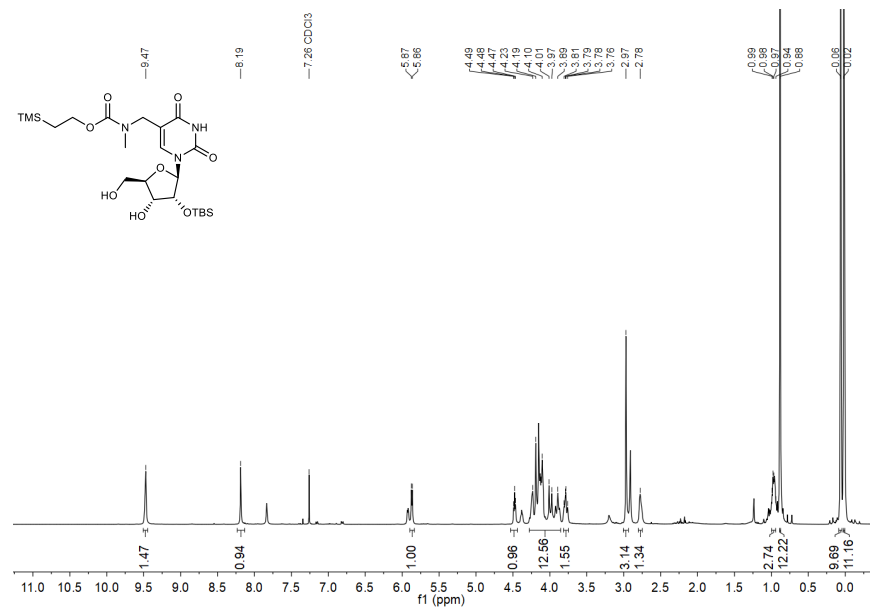
S84

¹H and ¹³C(¹H) NMR spectra of compound 4c

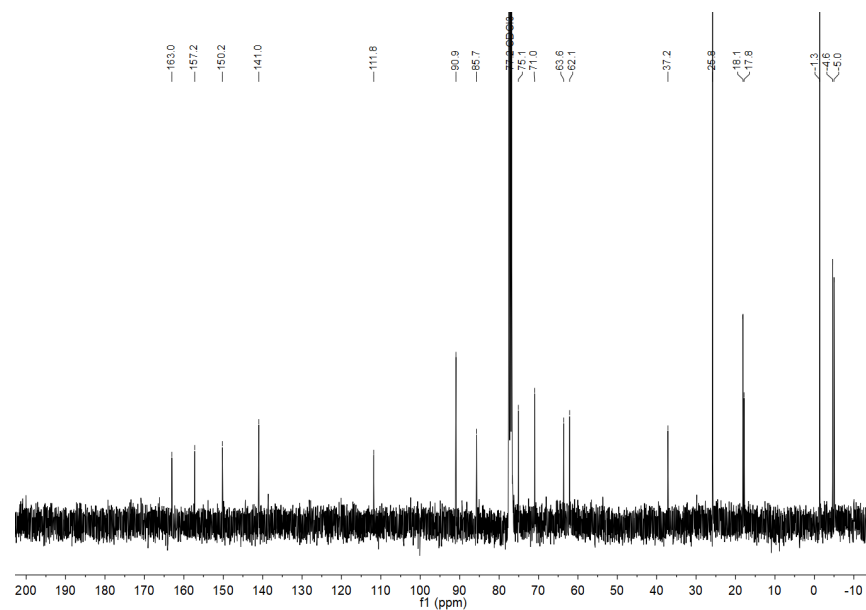
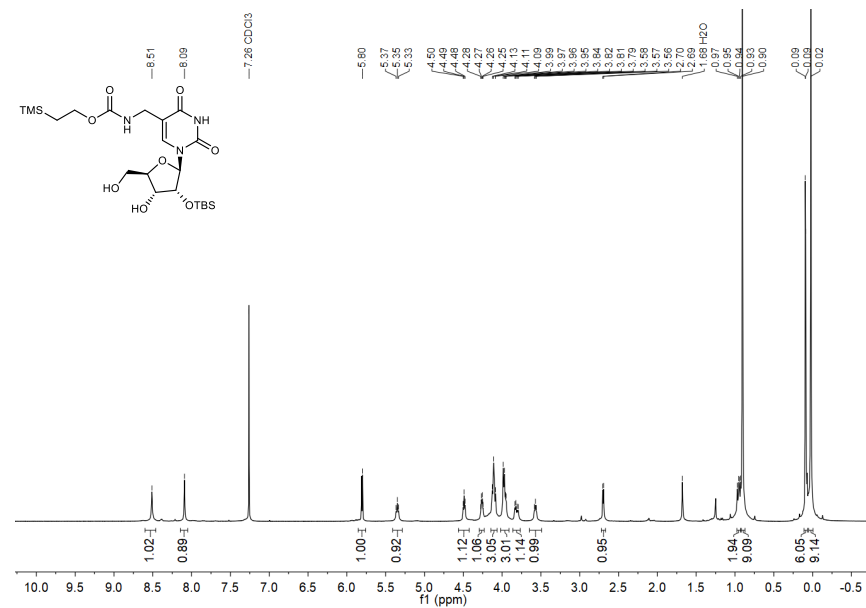


S85

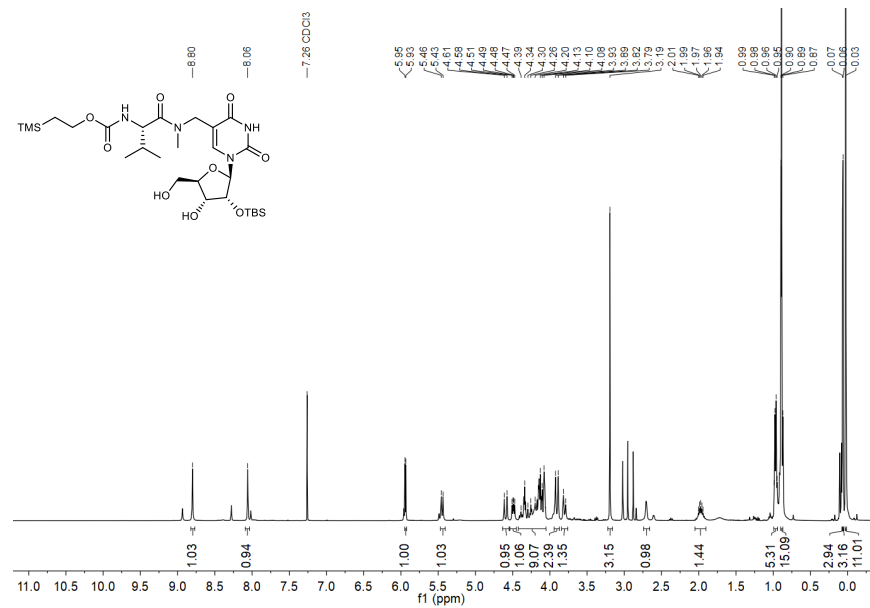
¹H and ¹³C{¹H} NMR spectra of compound 5a



¹H and ¹³C{¹H} NMR spectra of compound 5b

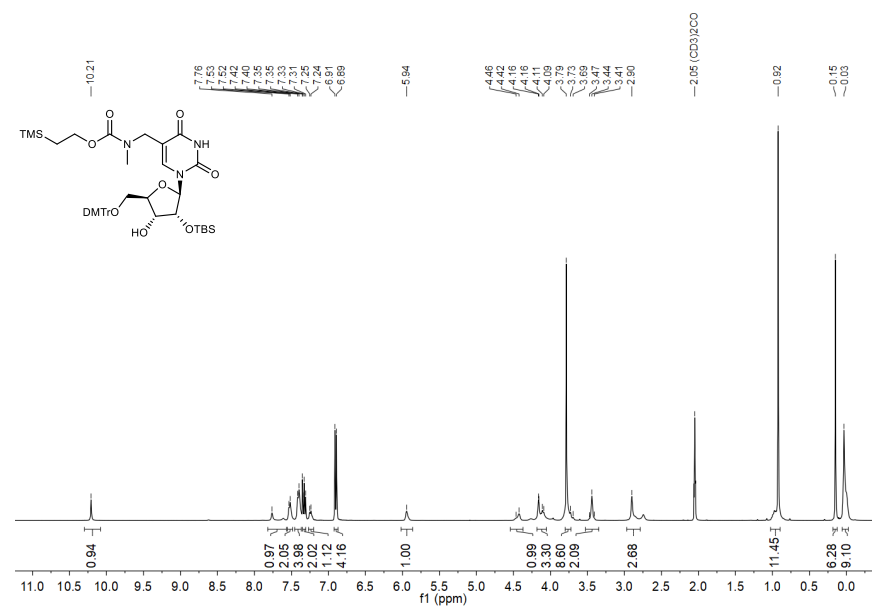


¹H and ¹³C{¹H} NMR spectra of compound 5c



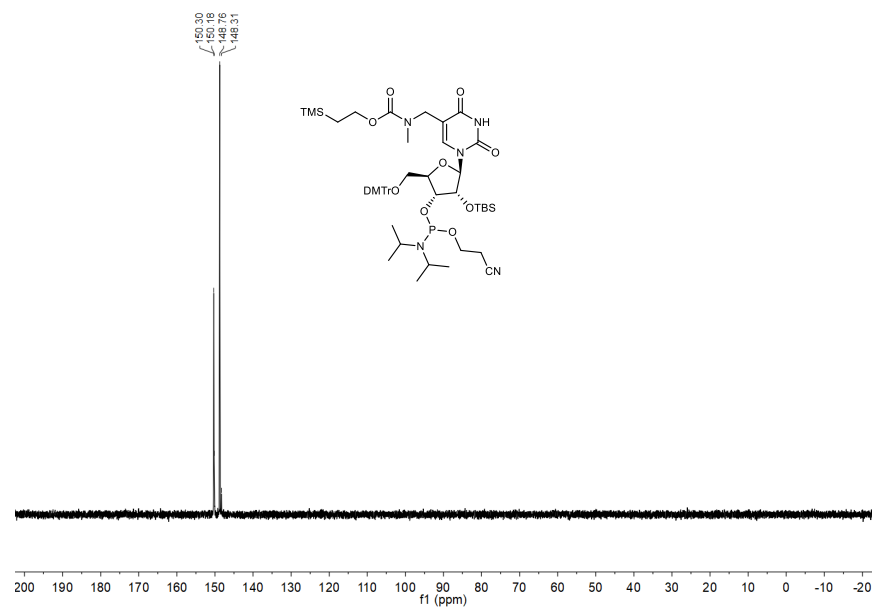
S88

¹H and ¹³C{¹H} NMR spectra of compound 6a

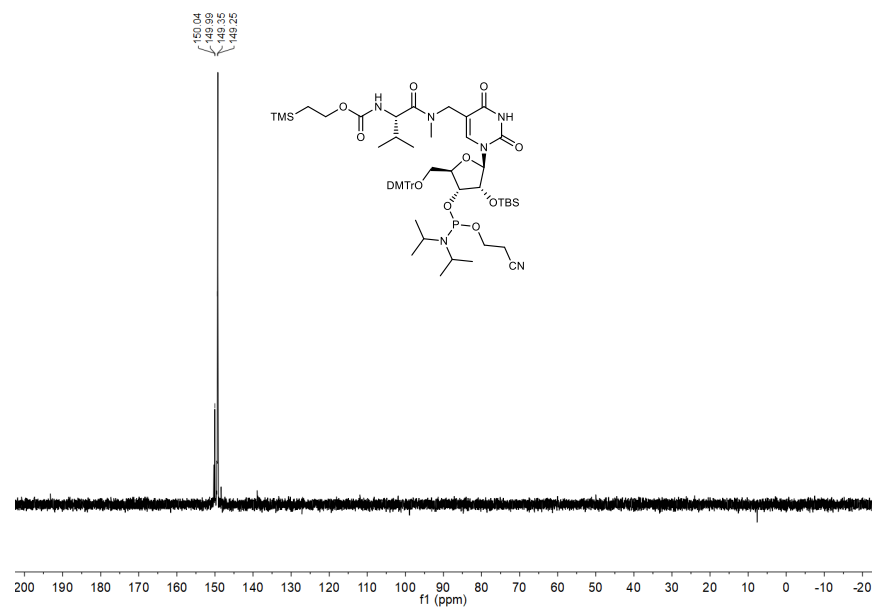


S89

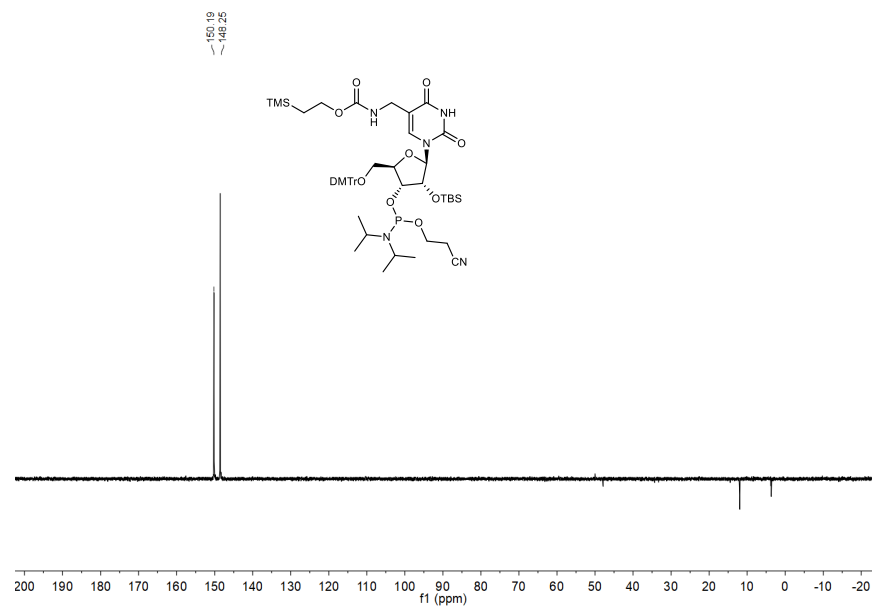
$^{31}\text{P}\{^1\text{H}\}$ NMR spectrum of compound 7a



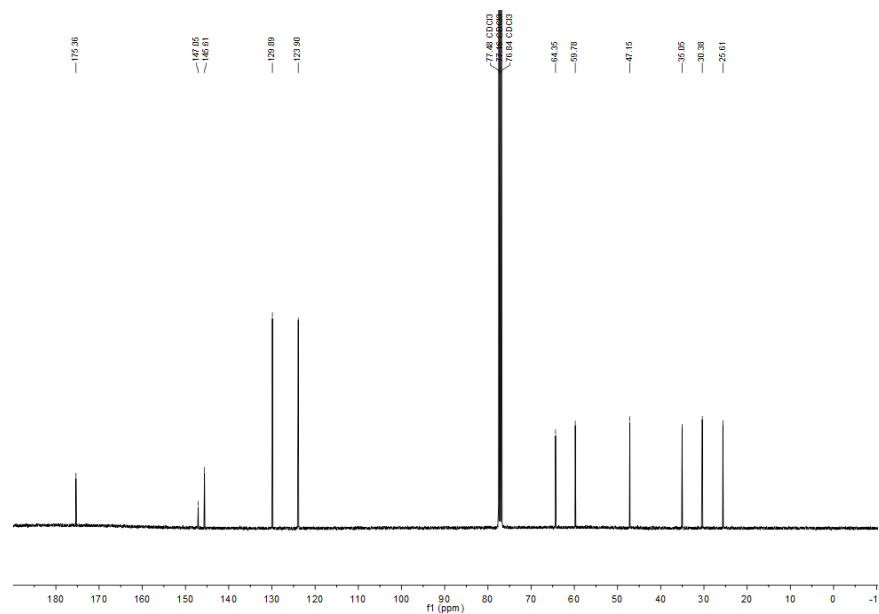
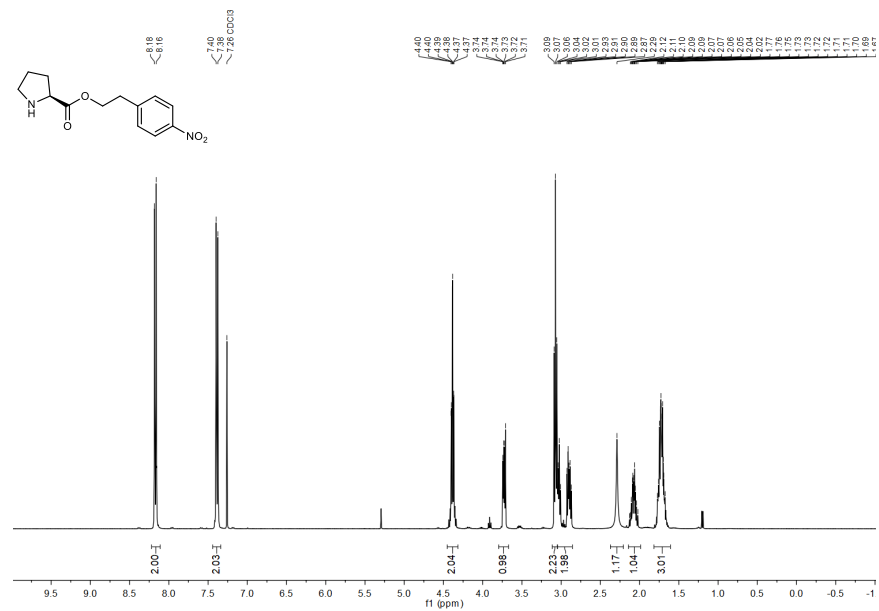
$^{31}\text{P}\{^1\text{H}\}$ NMR spectrum of compound 7c



$^{31}\text{P}\{^1\text{H}\}$ NMR spectrum of compound 7b

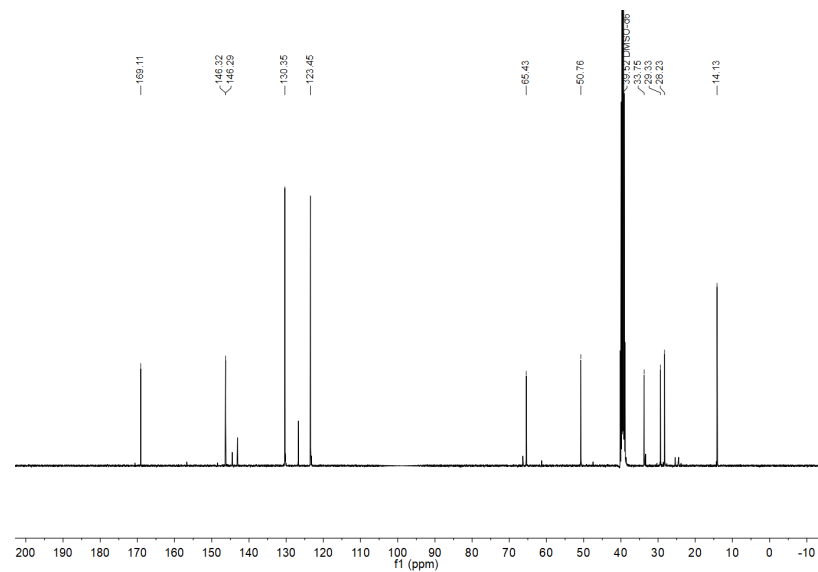
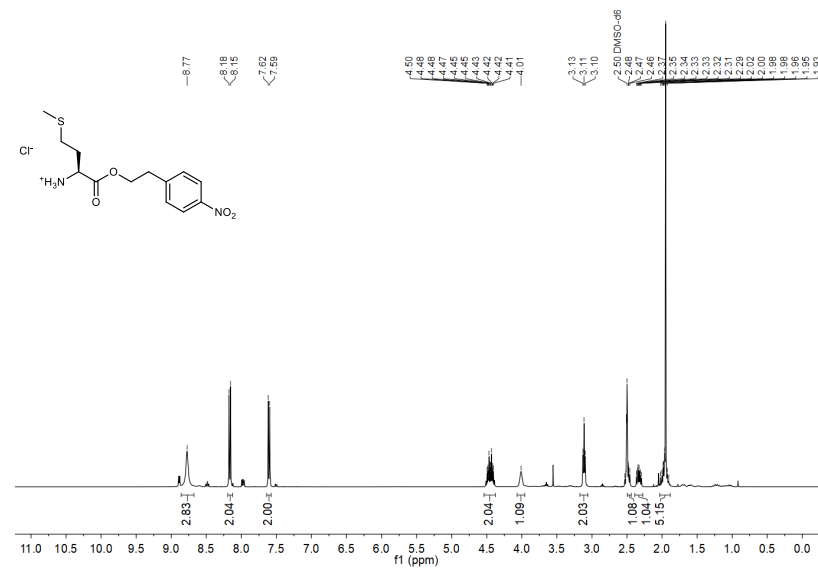


¹H and ¹³C{¹H} NMR spectra of compound H-Pro-Onpe



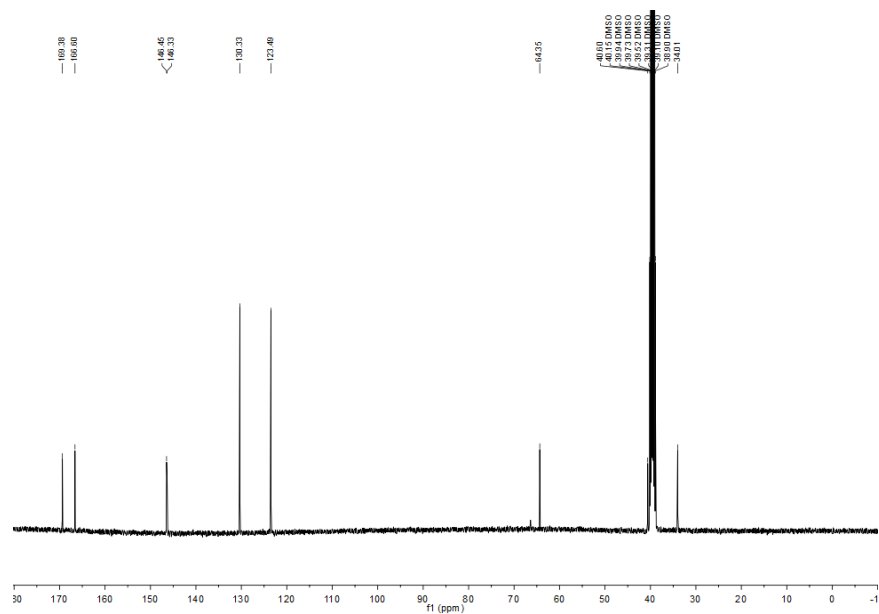
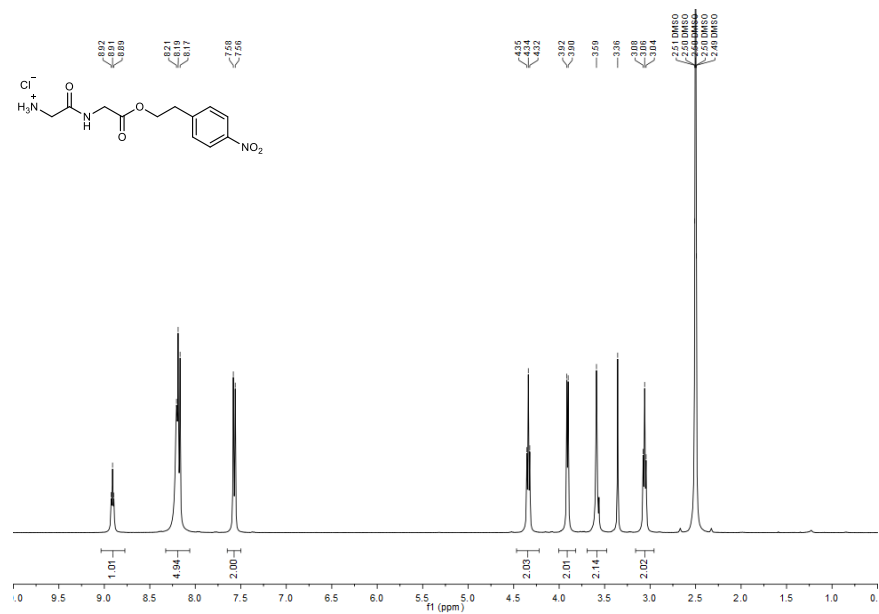
S96

¹H and ¹³C{¹H} NMR spectra of compound H-Met-Onpe·HCl



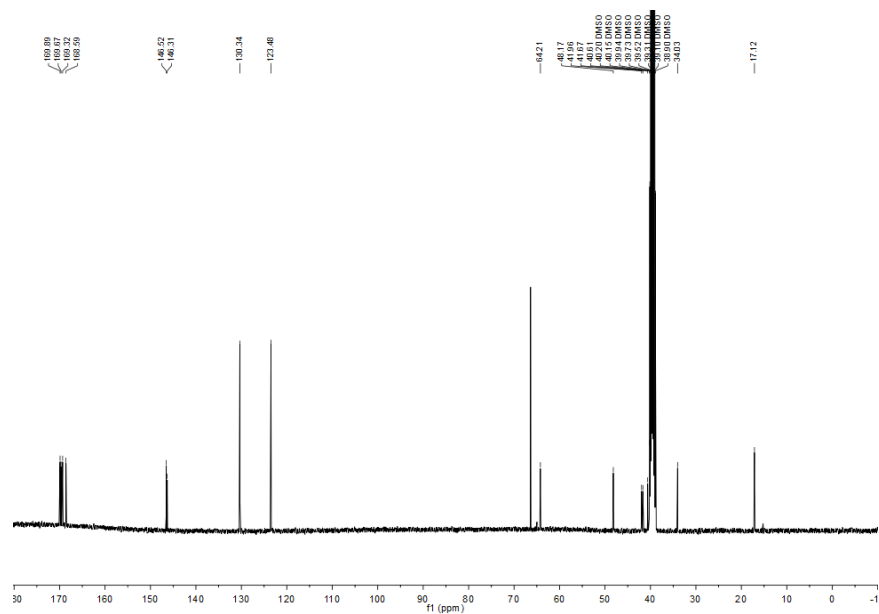
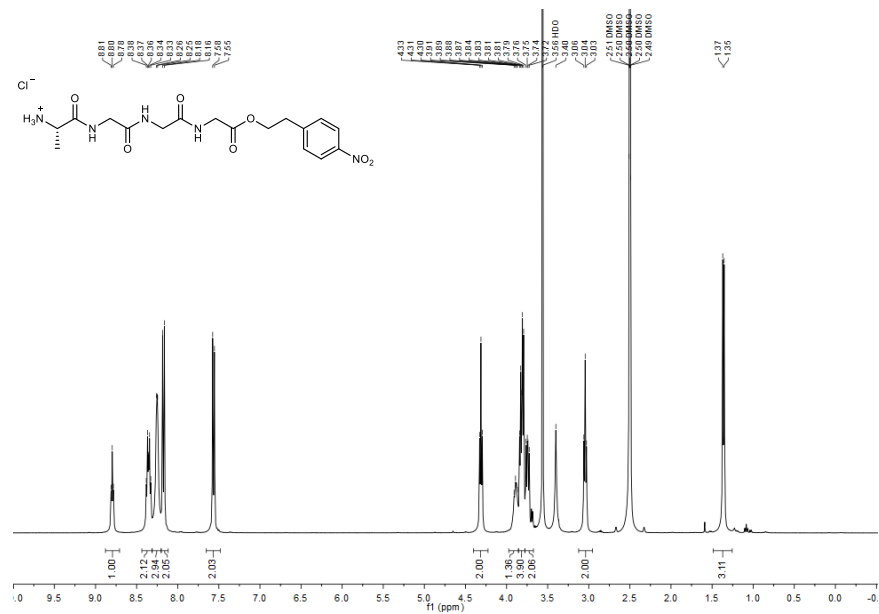
S97

¹H and ¹³C{¹H} NMR spectra of compound H-Gly-Gly-Onpe-HCl



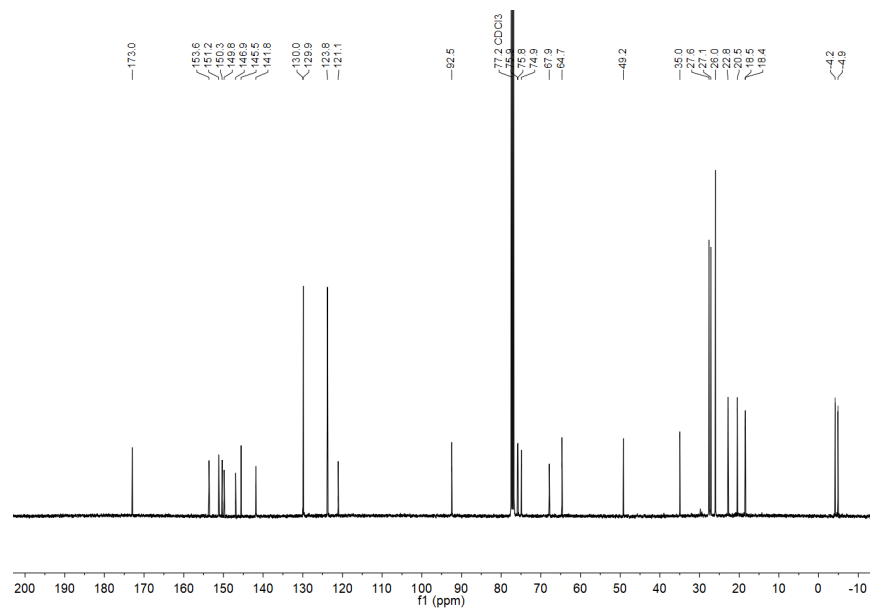
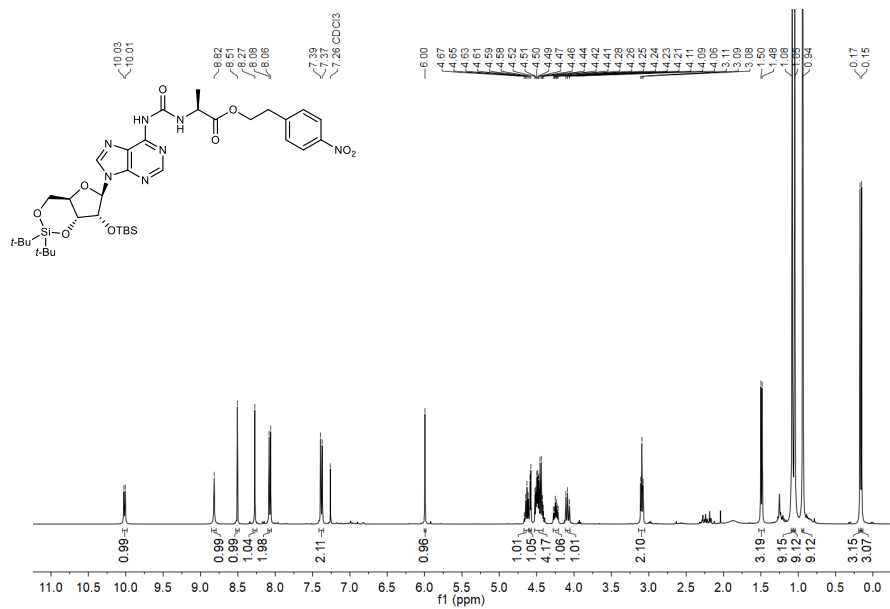
S98

¹H and ¹³C{¹H} NMR spectra of compound H-Ala-Gly-Gly-Gly-Onpe-HCl



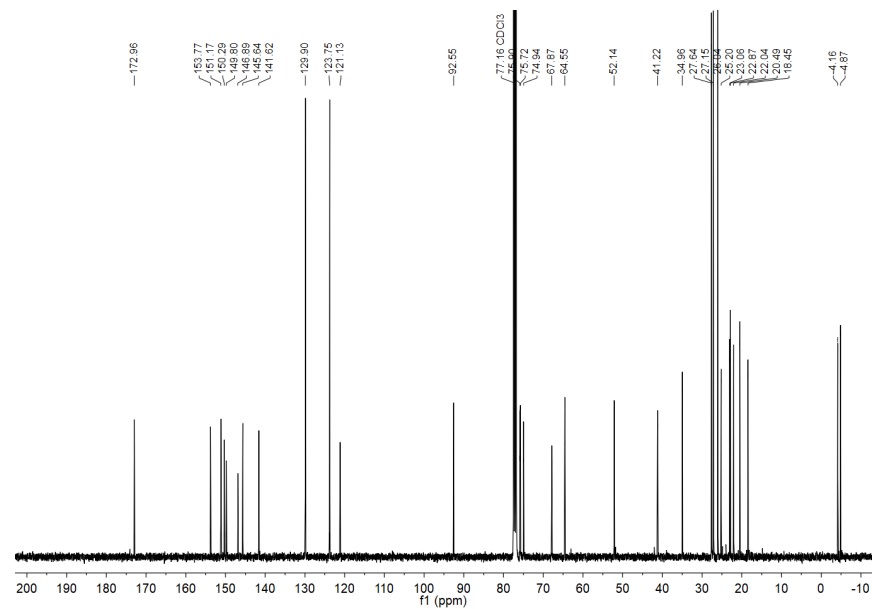
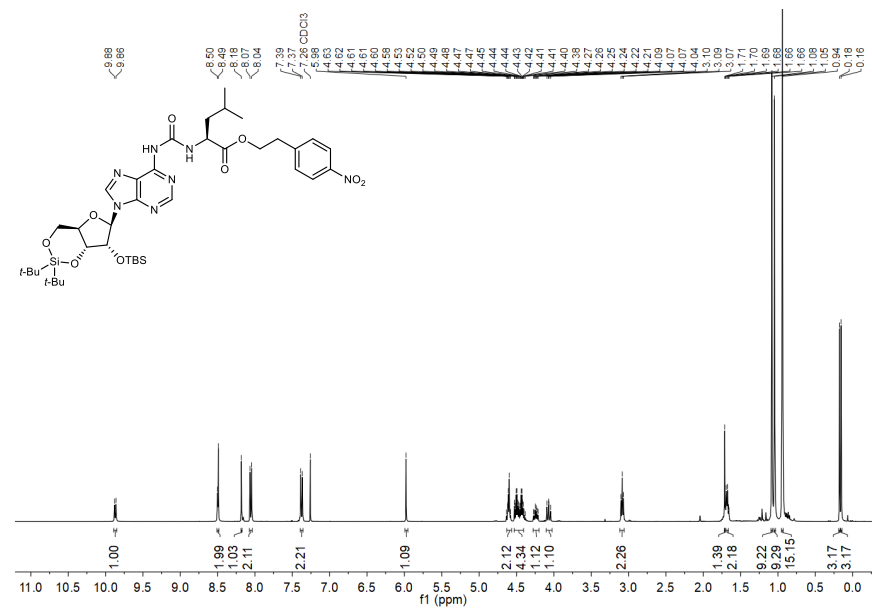
S99

¹H and ¹³C{¹H} NMR spectra of compound 10b



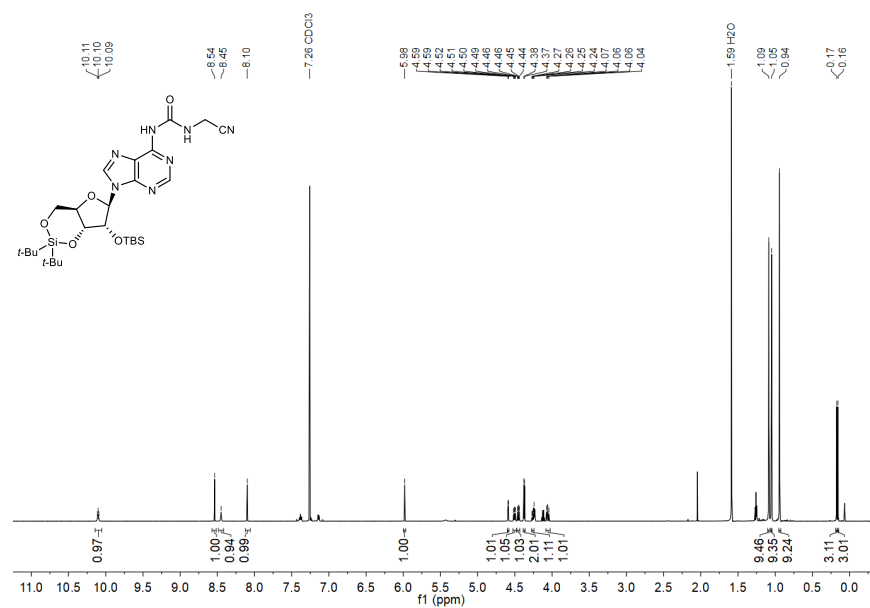
S100

¹H and ¹³C{¹H} NMR spectra of compound 10d



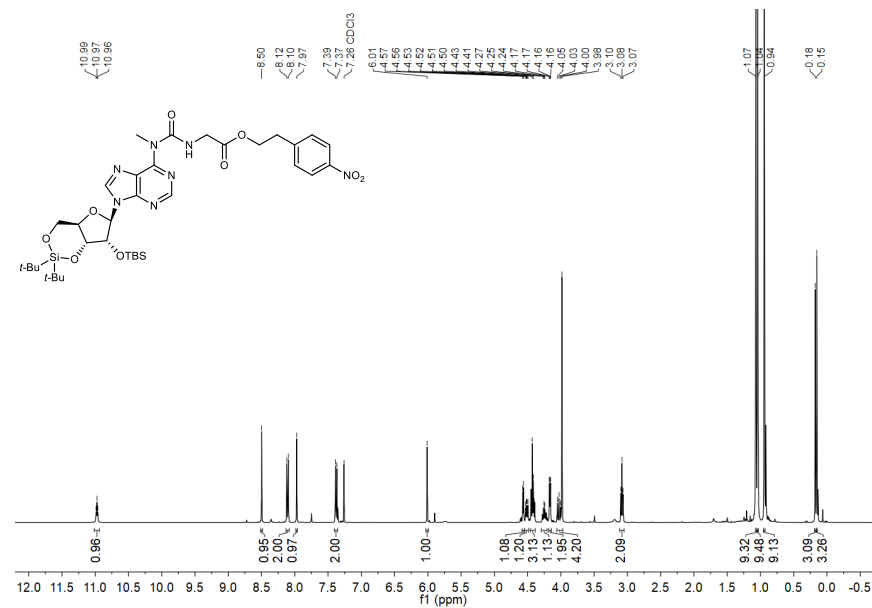
S101

¹H and ¹³C{¹H} NMR spectra of compound 10j



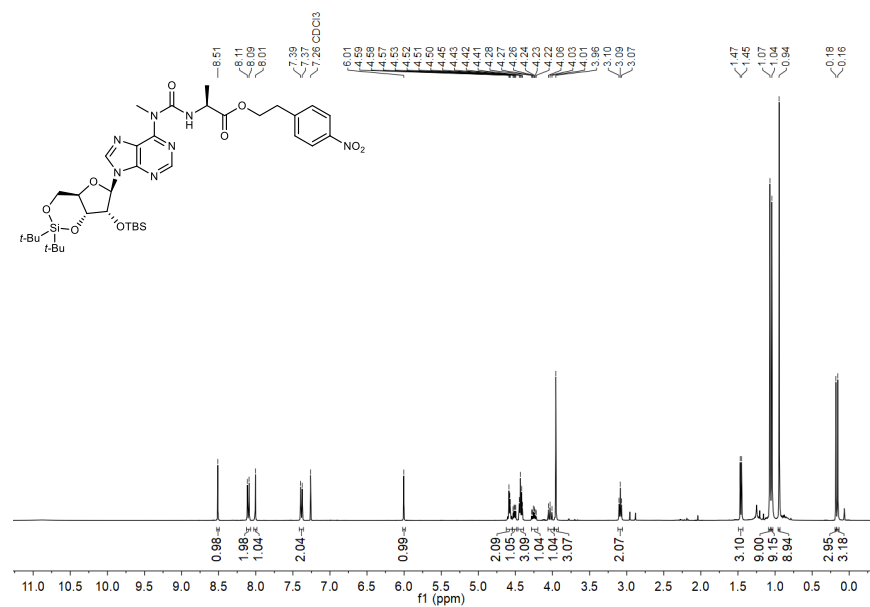
S104

¹H and ¹³C{¹H} NMR spectra of compound 11a



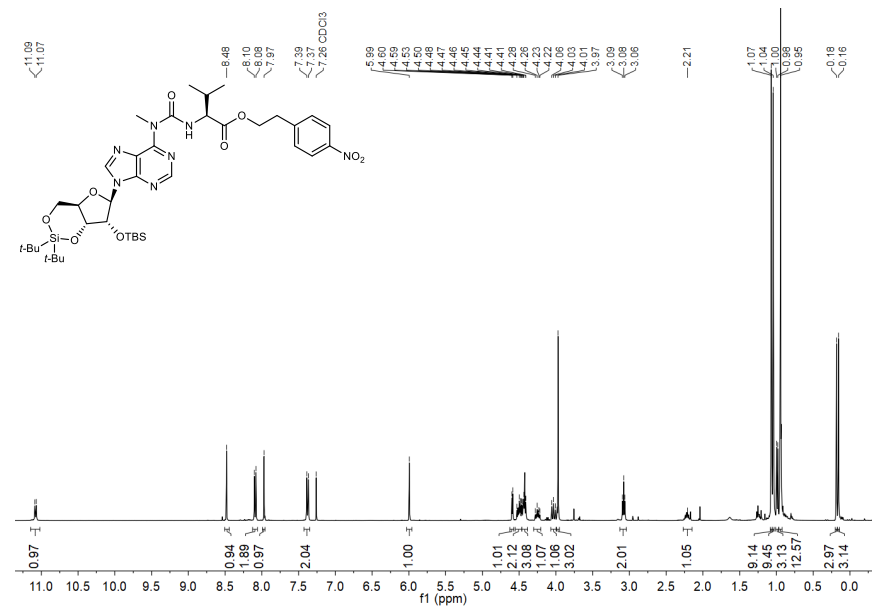
S105

¹H and ¹³C{¹H} NMR spectra of compound 11b



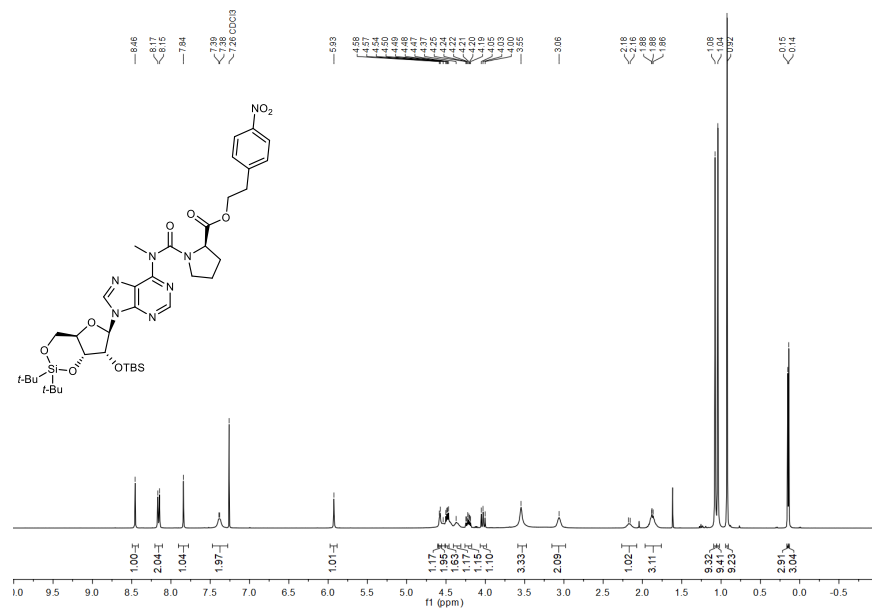
S106

¹H and ¹³C{¹H} NMR spectra of compound 11c



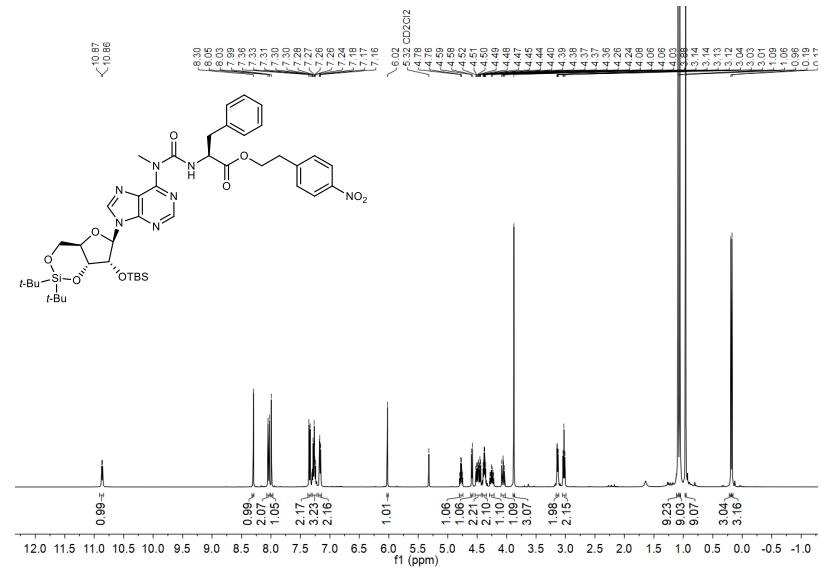
S107

¹H and ¹³C{¹H} NMR spectra of compound 11f



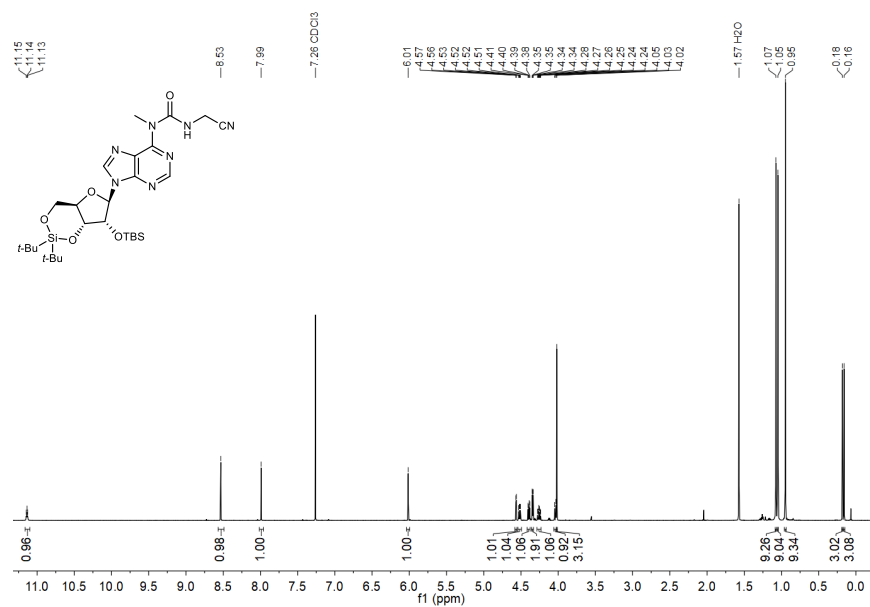
S110

¹H and ¹³C{¹H} NMR spectra of compound 11g



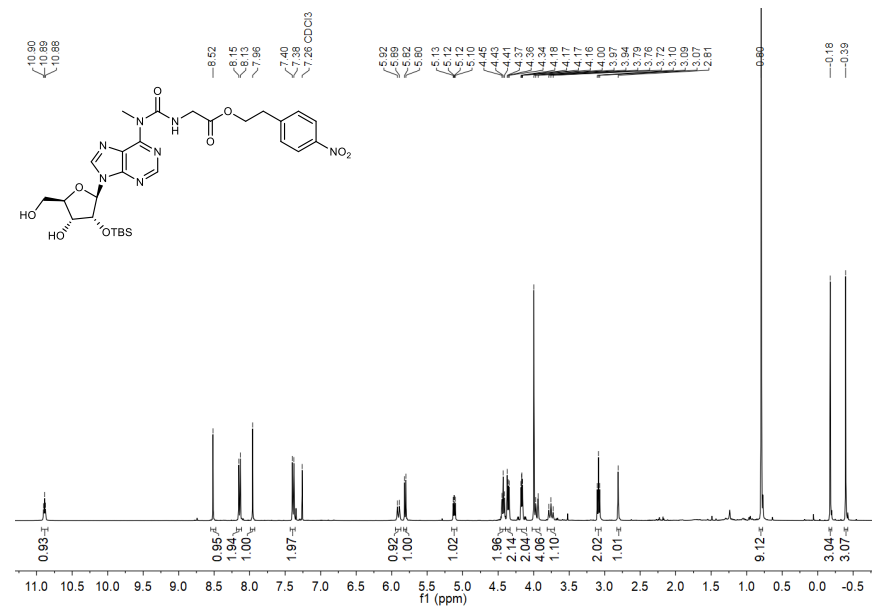
S111

¹H and ¹³C(¹H) NMR spectra of compound 11j



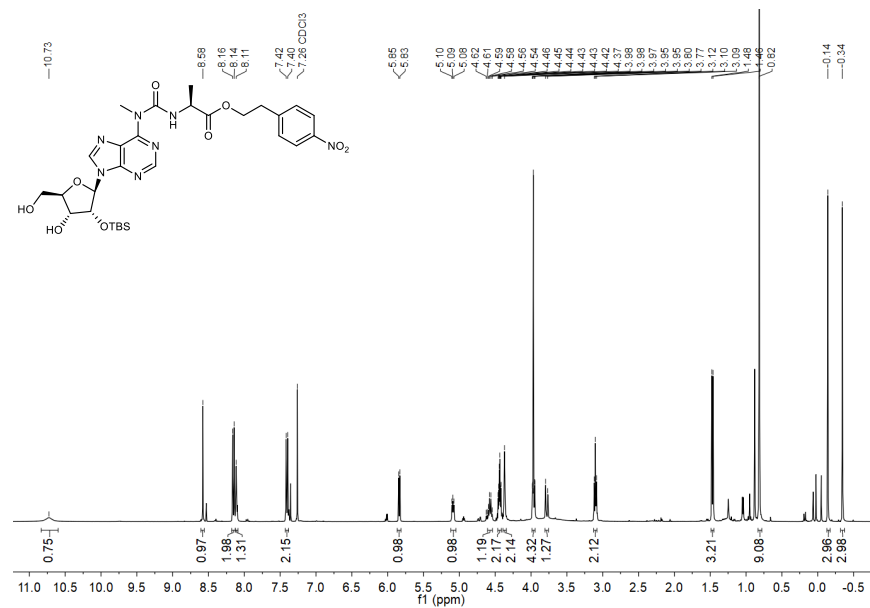
S114

¹H and ¹³C(¹H) NMR spectra of compound 12a



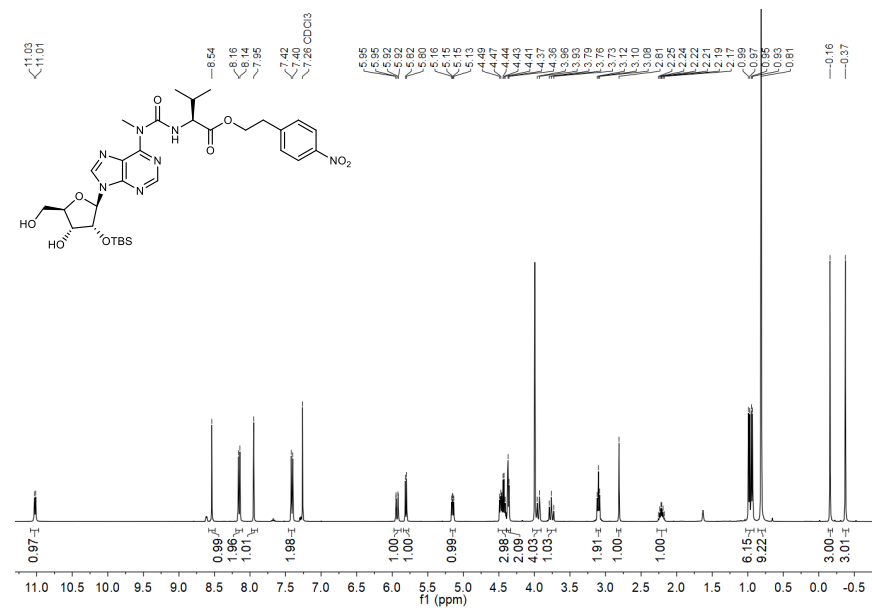
S115

¹H and ¹³C{¹H} NMR spectra of compound 12b



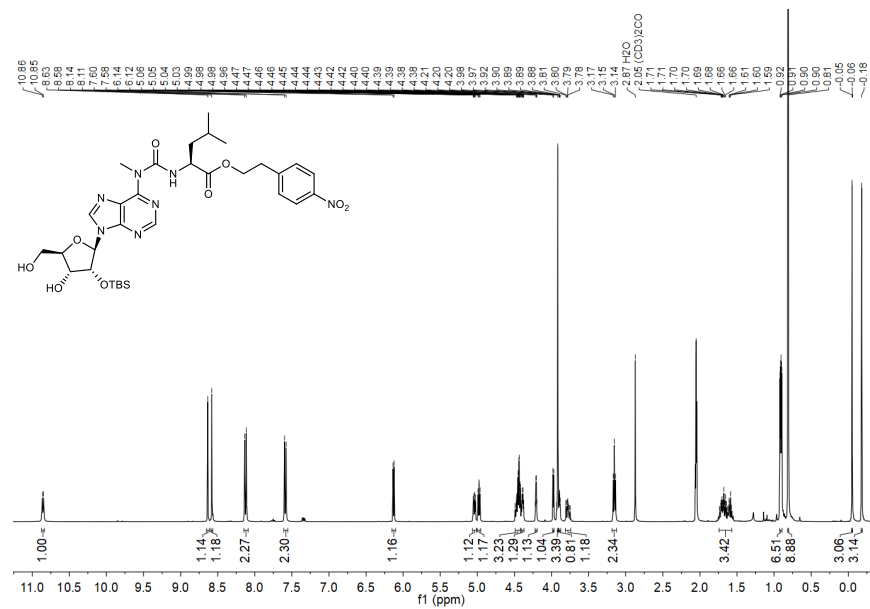
S116

¹H and ¹³C{¹H} NMR spectra of compound 12c



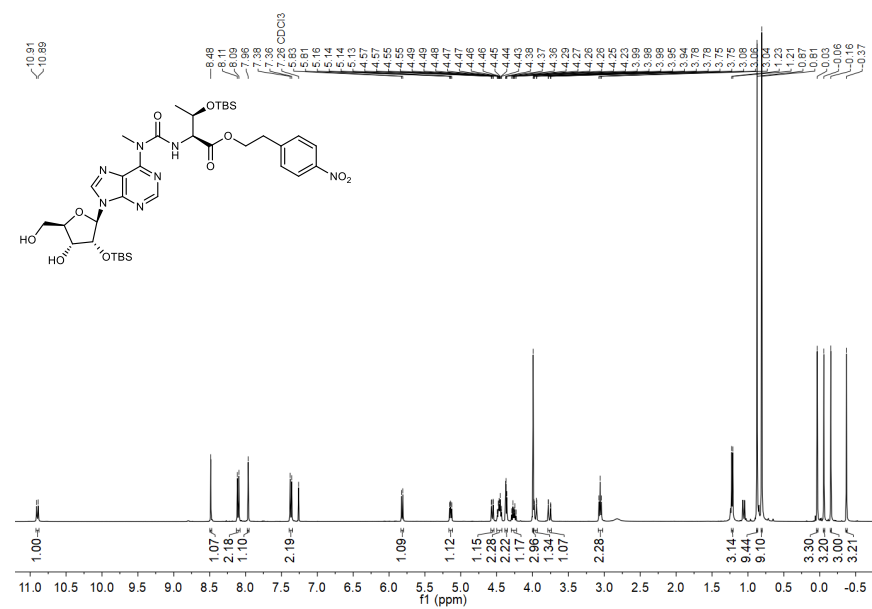
S117

¹H and ¹³C{¹H} NMR spectra of compound 12d



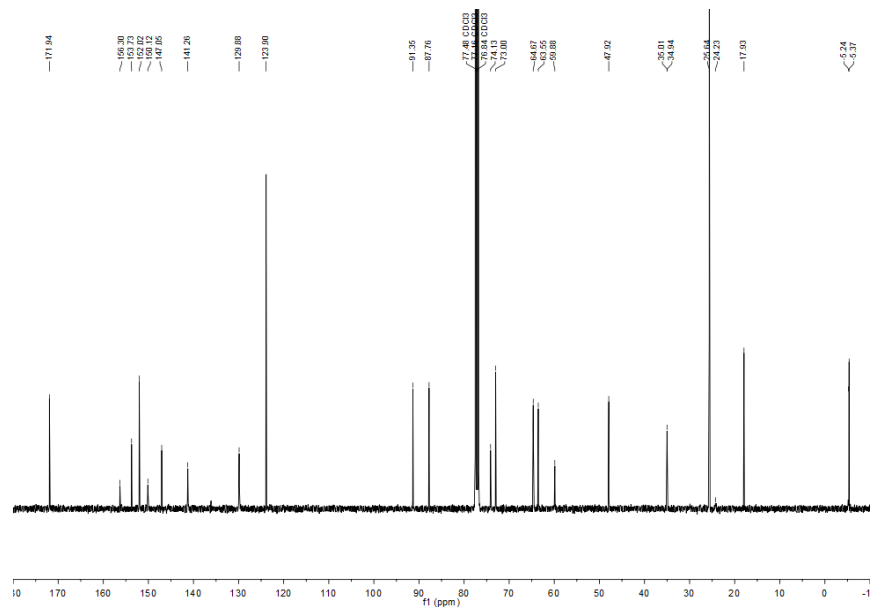
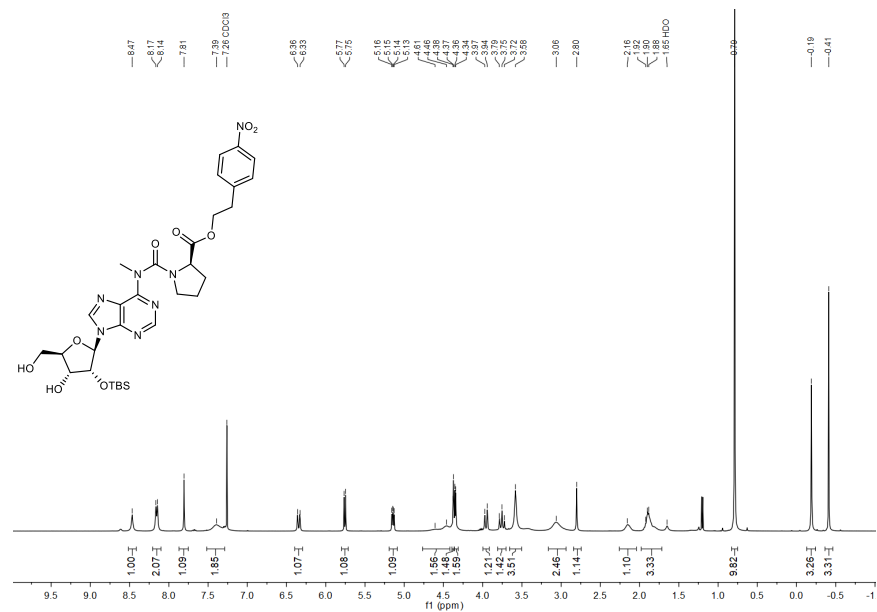
S118

¹H and ¹³C{¹H} NMR spectra of compound 12e



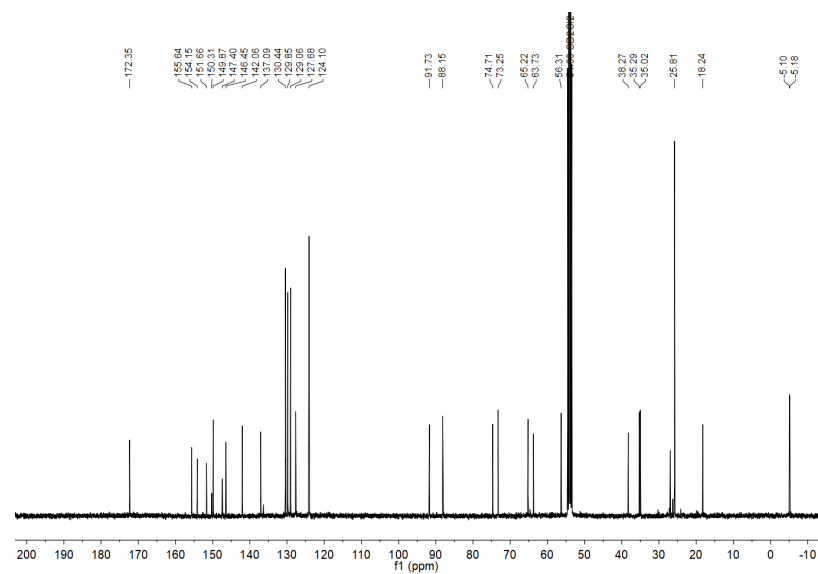
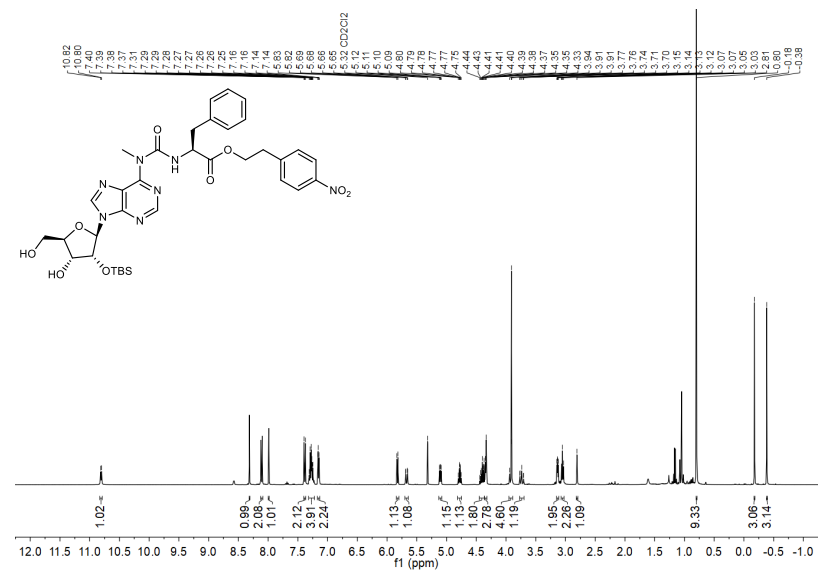
S119

¹H and ¹³C{¹H} NMR spectra of compound 12f



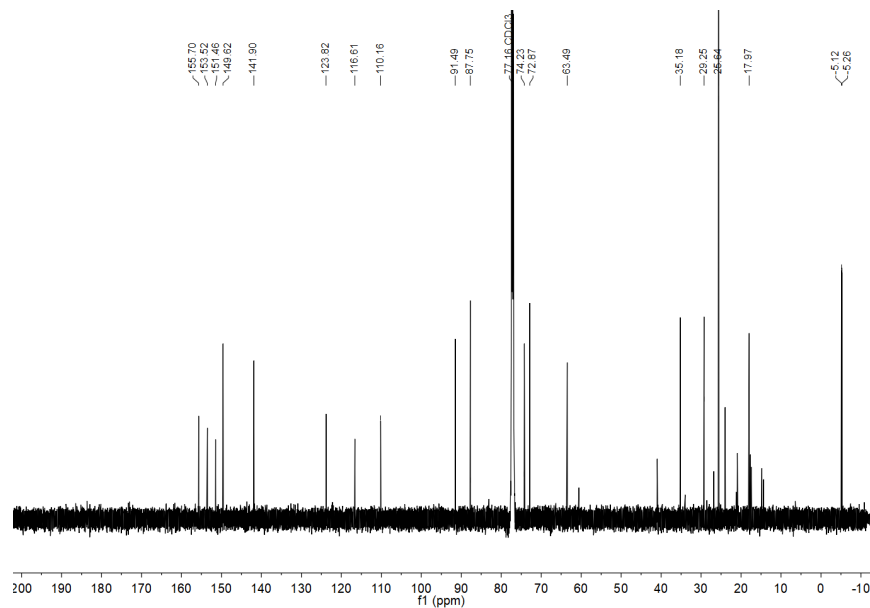
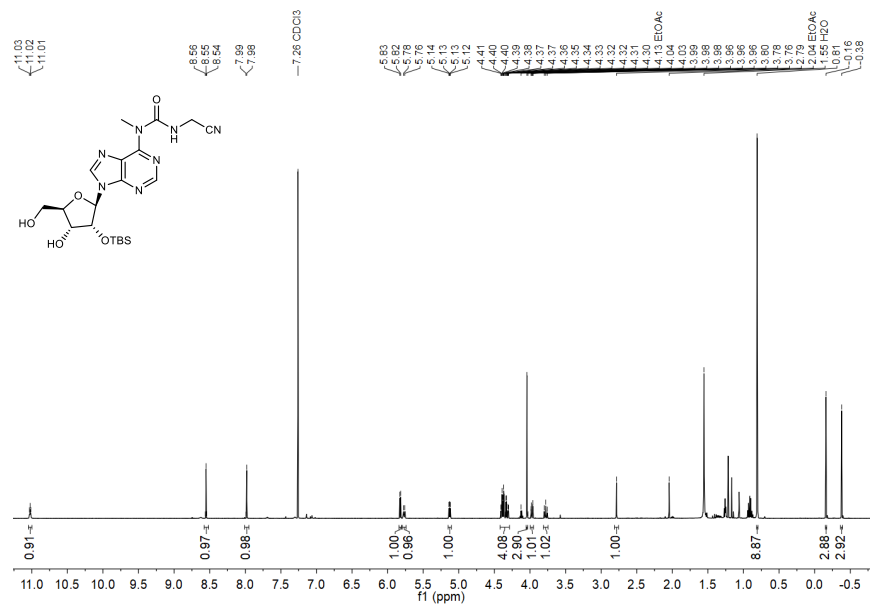
S120

¹H and ¹³C{¹H} NMR spectra of compound 12g

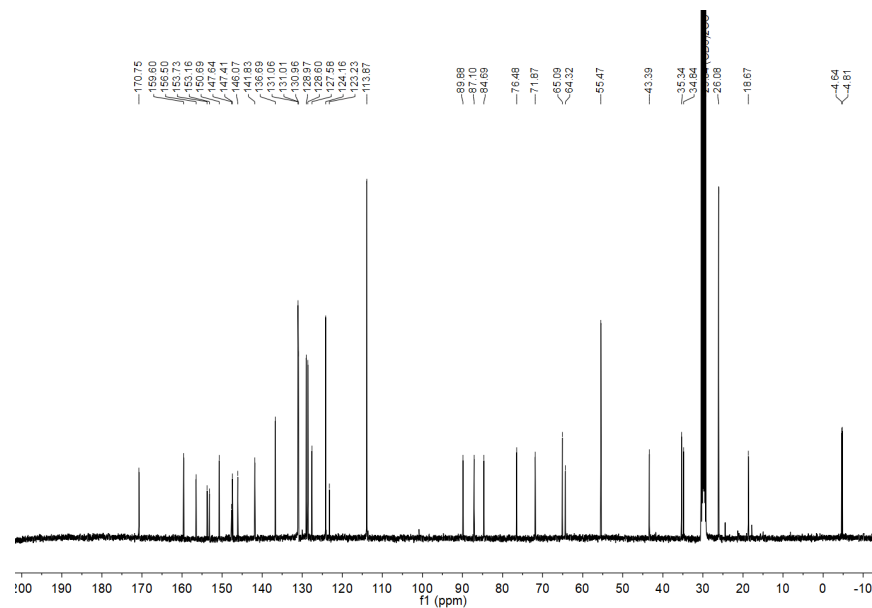
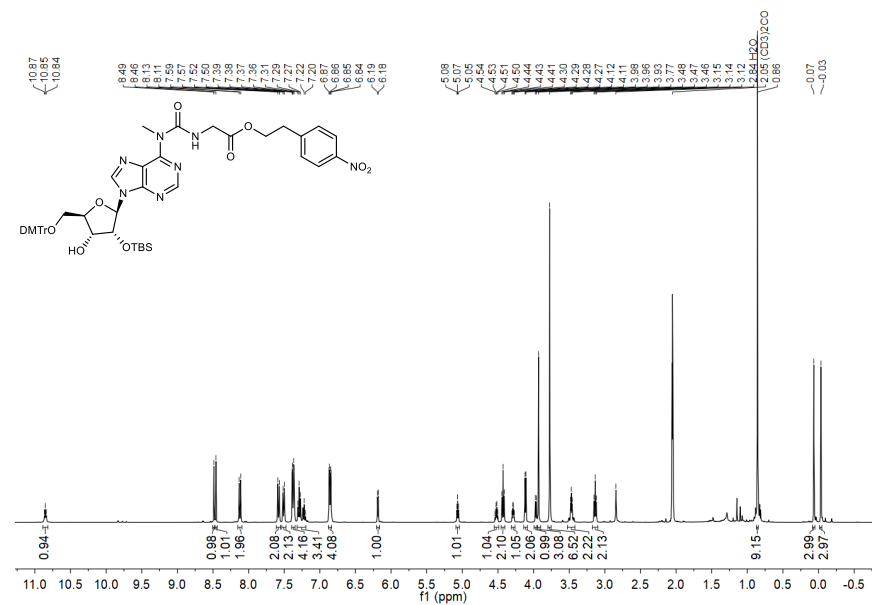


S121

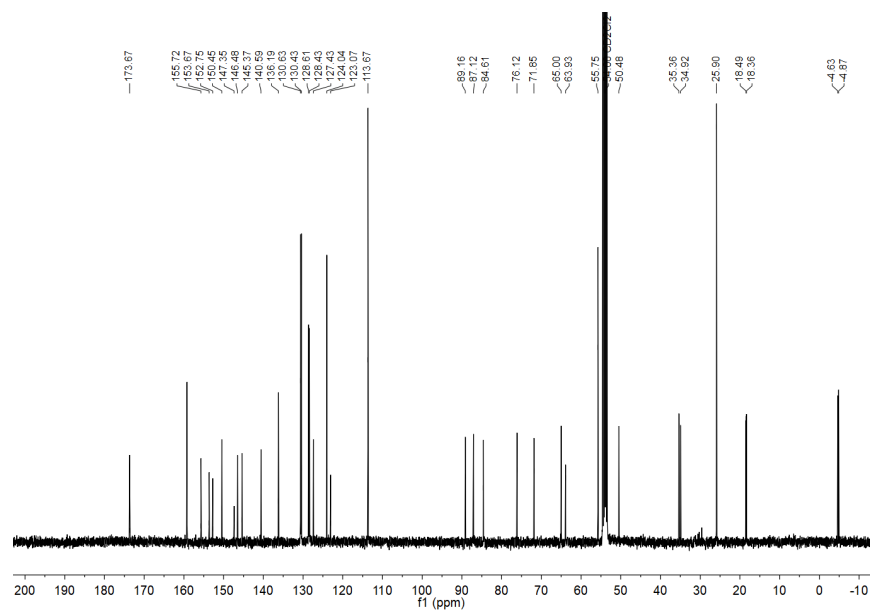
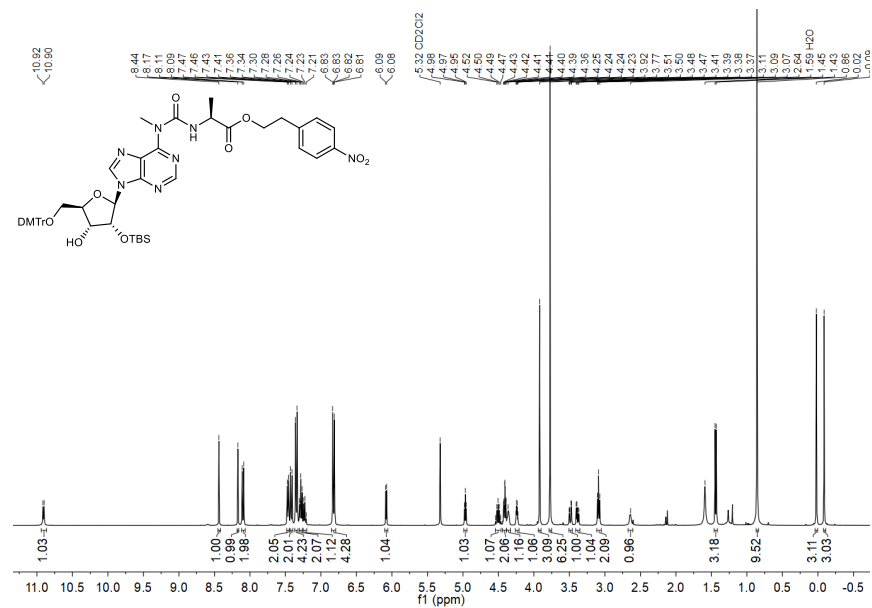
¹H and ¹³C{¹H} NMR spectra of compound 12j



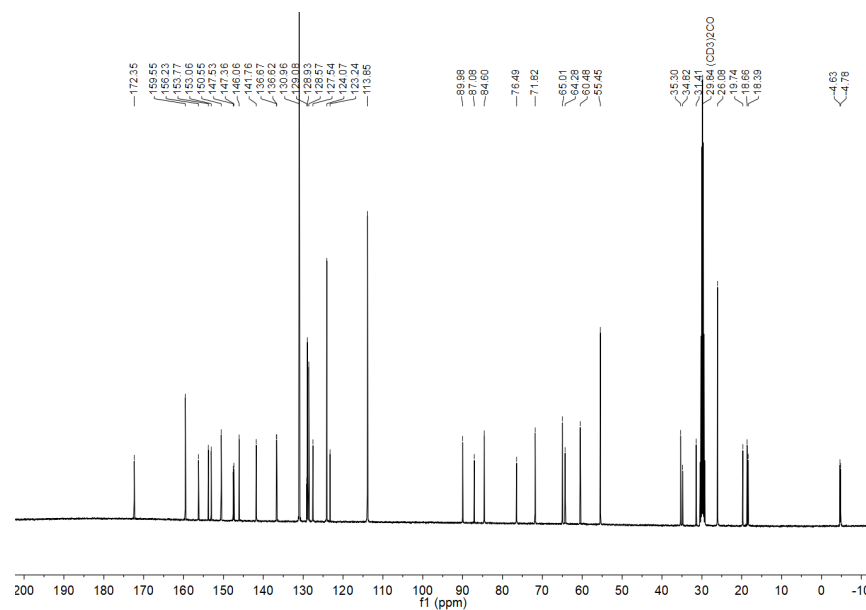
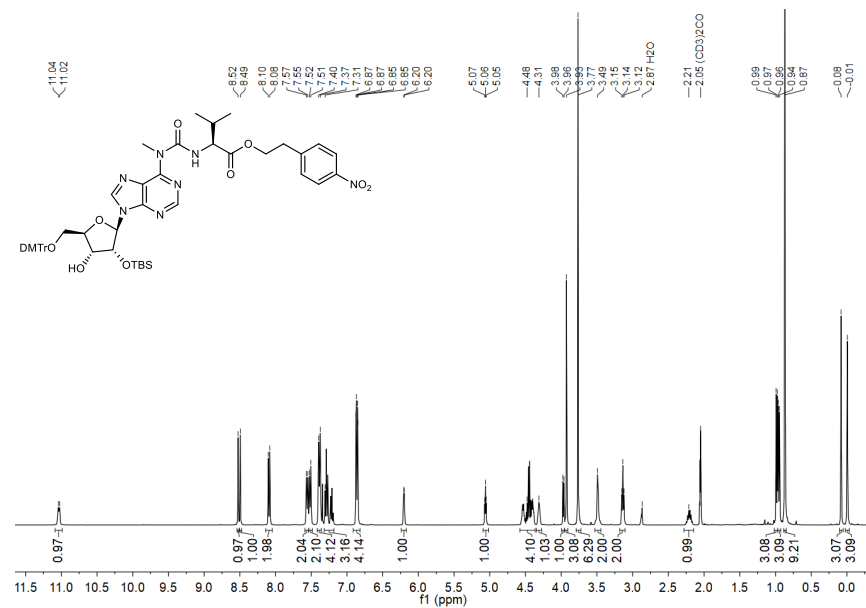
¹H and ¹³C{¹H} NMR spectra of compound 13a



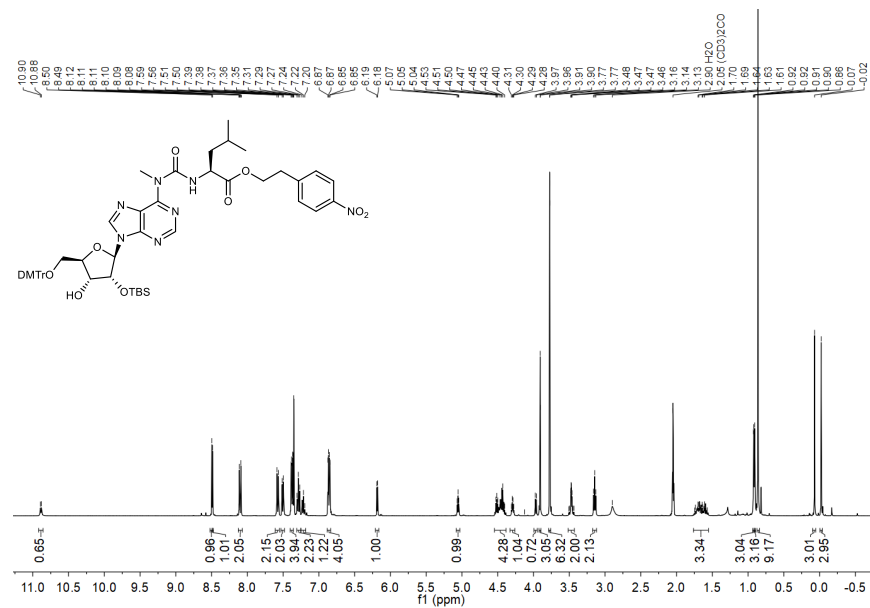
¹H and ¹³C{¹H} NMR spectra of compound 13b



¹H and ¹³C{¹H} NMR spectra of compound 13c

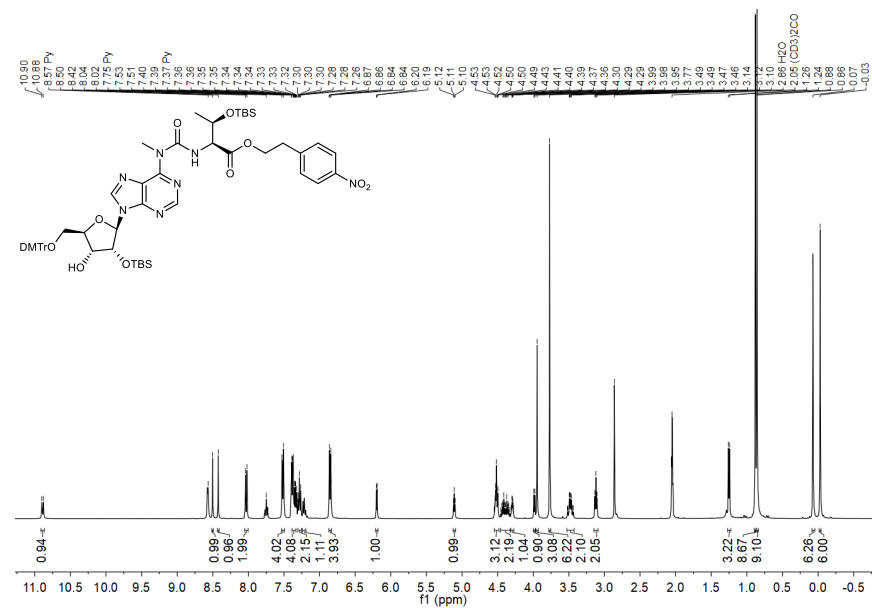


¹H and ¹³C{¹H} NMR spectra of compound 13d



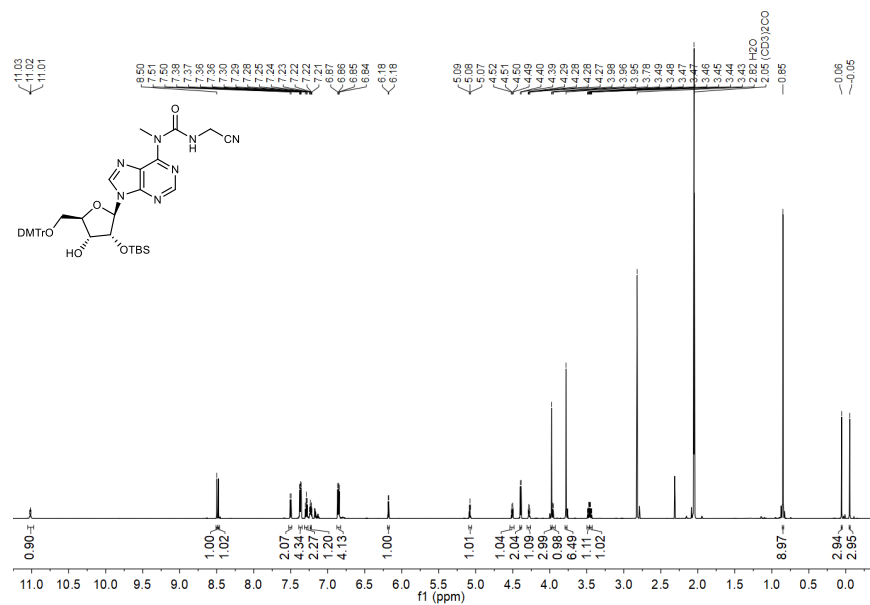
S128

¹H and ¹³C{¹H} NMR spectra of compound 13e



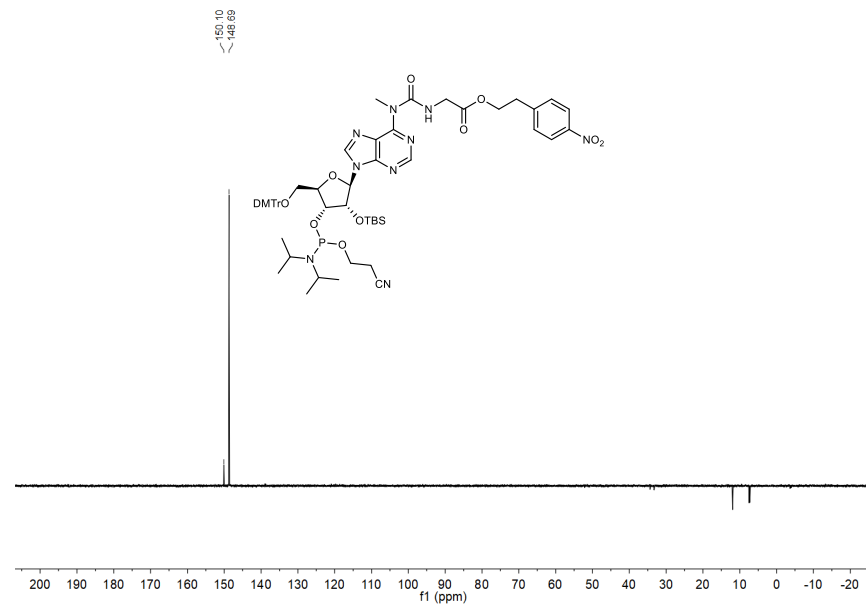
S129

¹H and ¹³C(¹H) NMR spectra of compound 13j

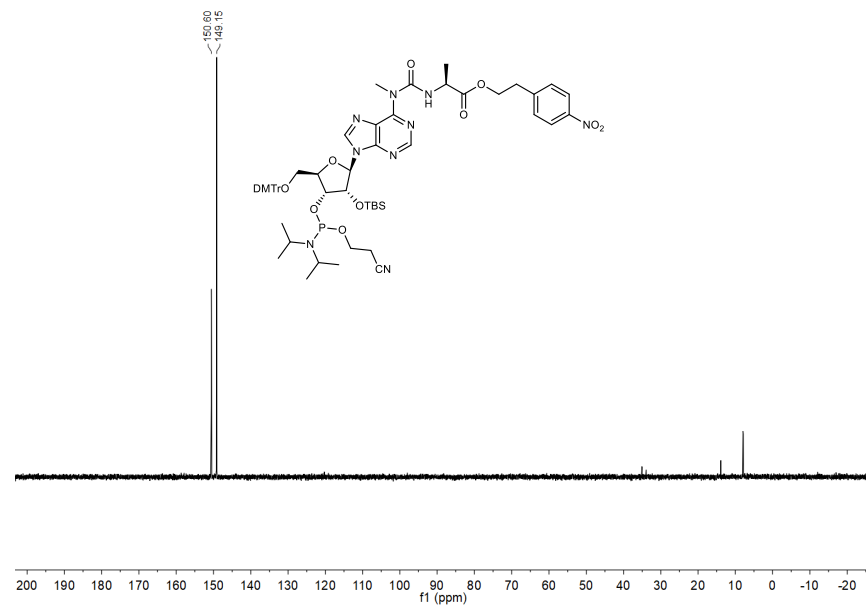


S134

³¹P{¹H} NMR spectrum of compound 14a

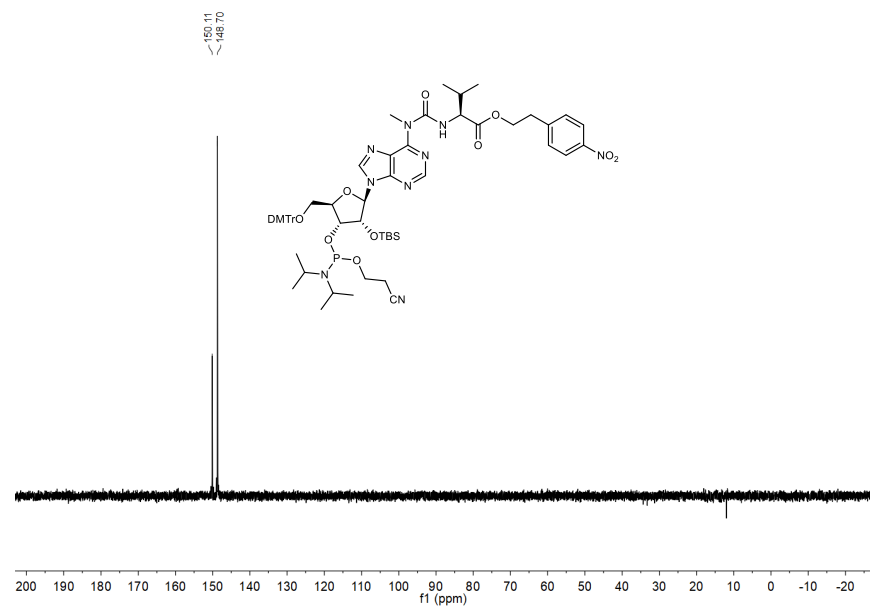


³¹P{¹H} NMR spectrum of compound 14b

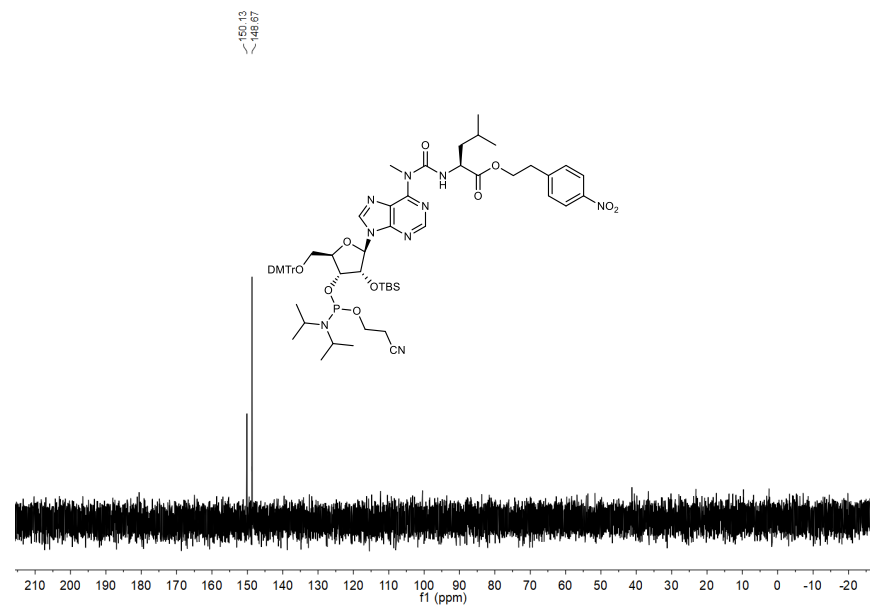


S135

$^{31}\text{P}\{^1\text{H}\}$ NMR spectrum of compound 14c

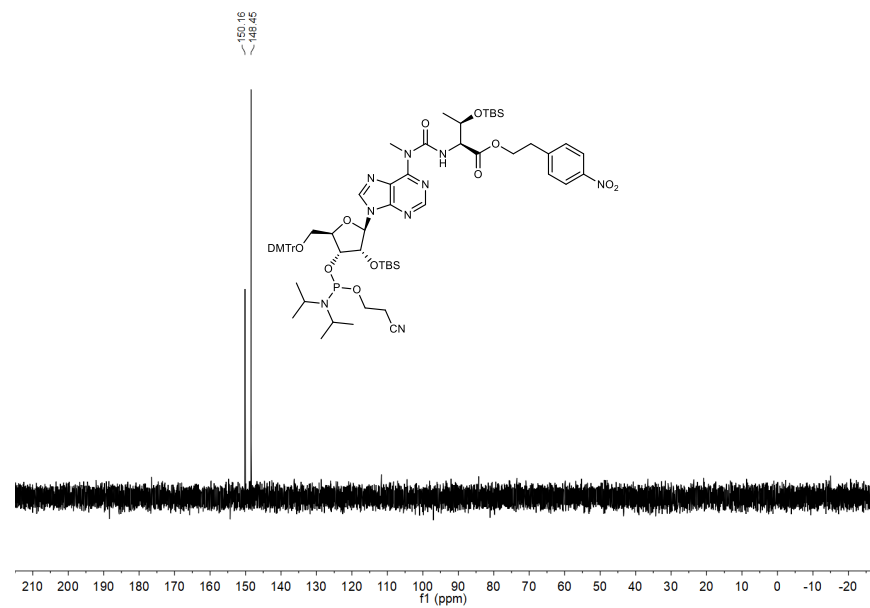


$^{31}\text{P}\{^1\text{H}\}$ NMR spectrum of compound 14d

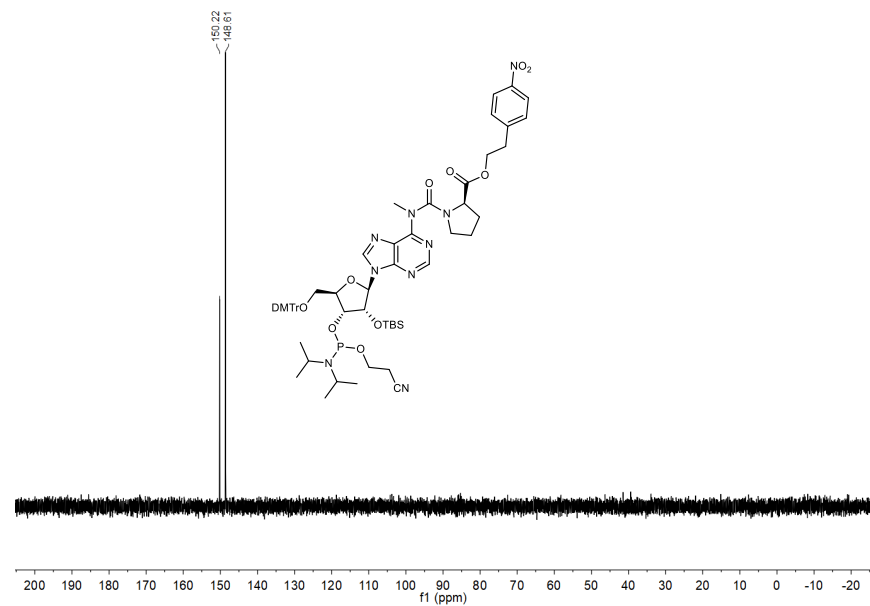


S136

$^{31}\text{P}\{^1\text{H}\}$ NMR spectrum of compound 14e

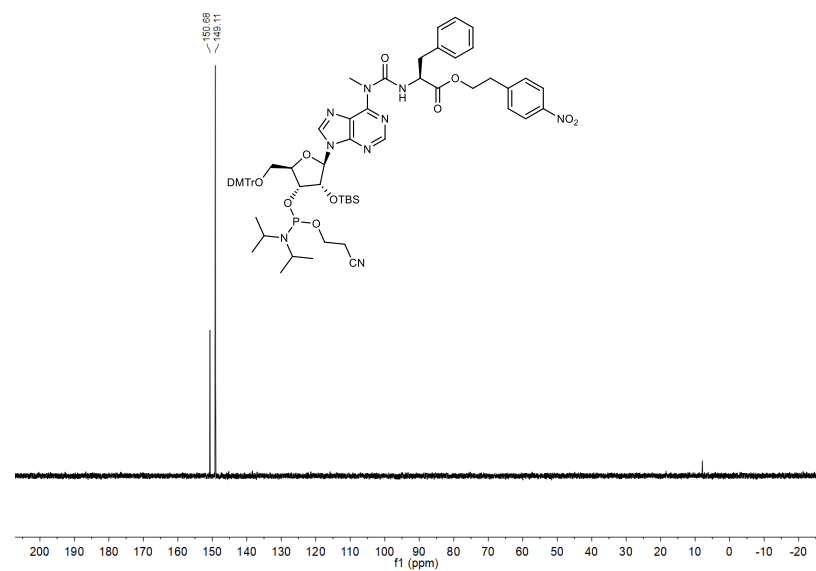


$^{31}\text{P}\{^1\text{H}\}$ NMR spectrum of compound 14f

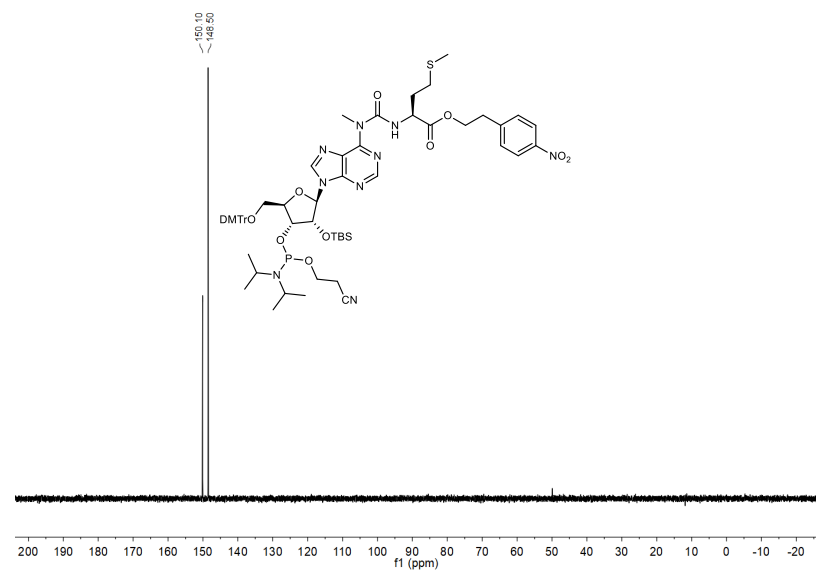


S137

³¹P{¹H} NMR spectrum of compound 14g

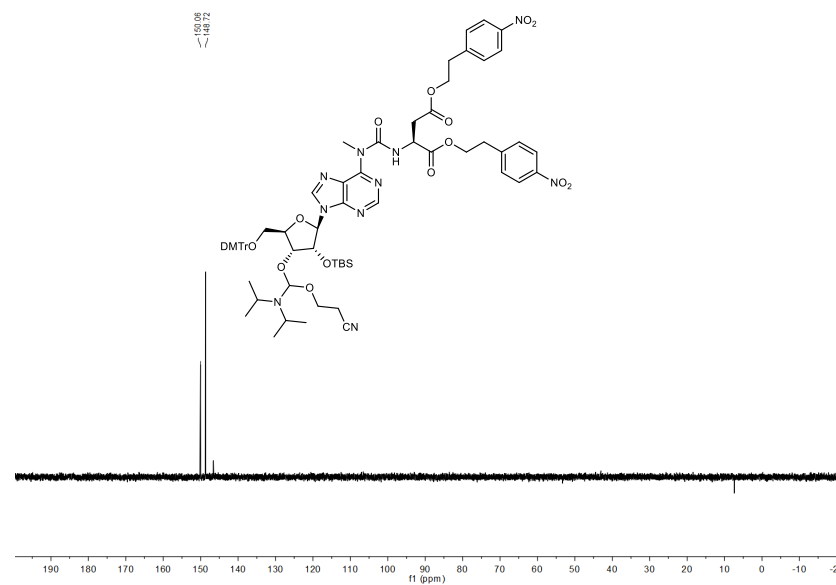


³¹P{¹H} NMR spectrum of compound 14h

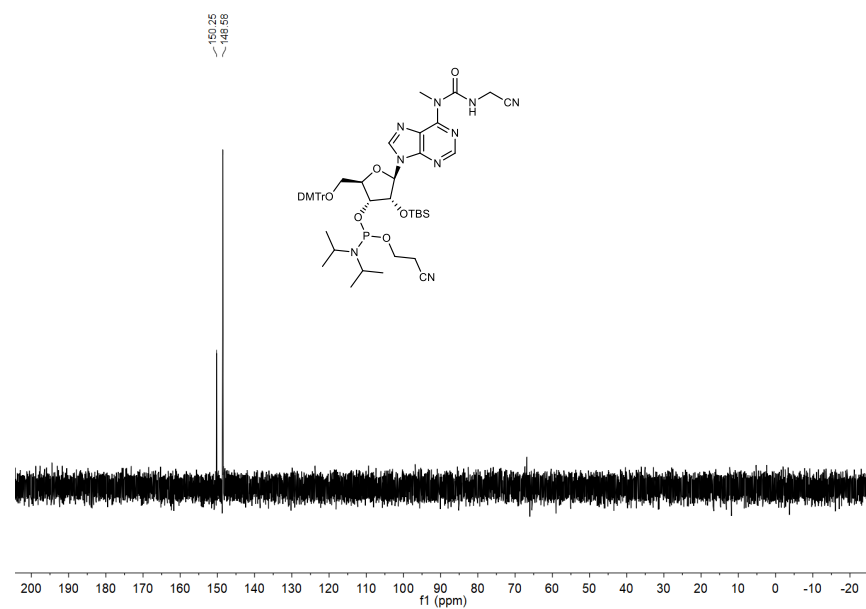


S138

³¹P{¹H} NMR spectrum of compound 14i

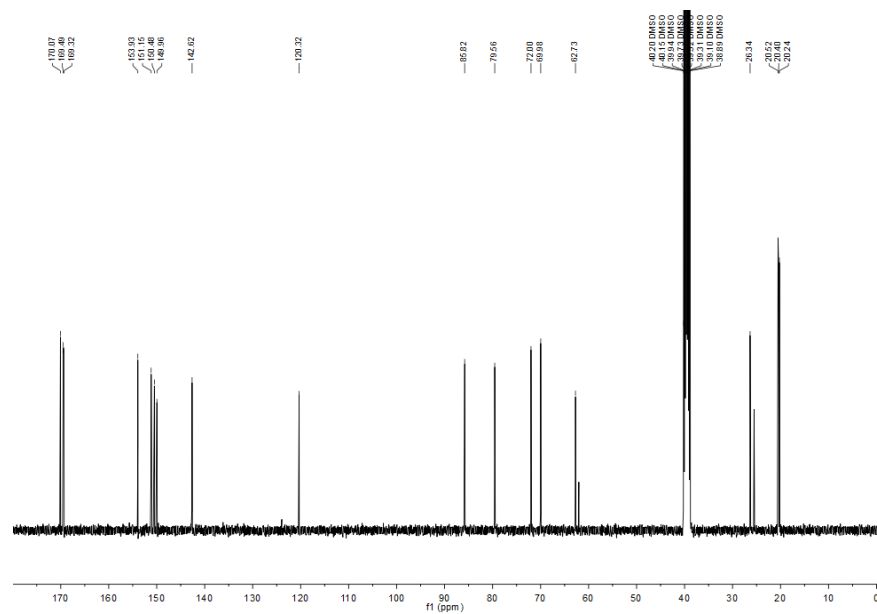
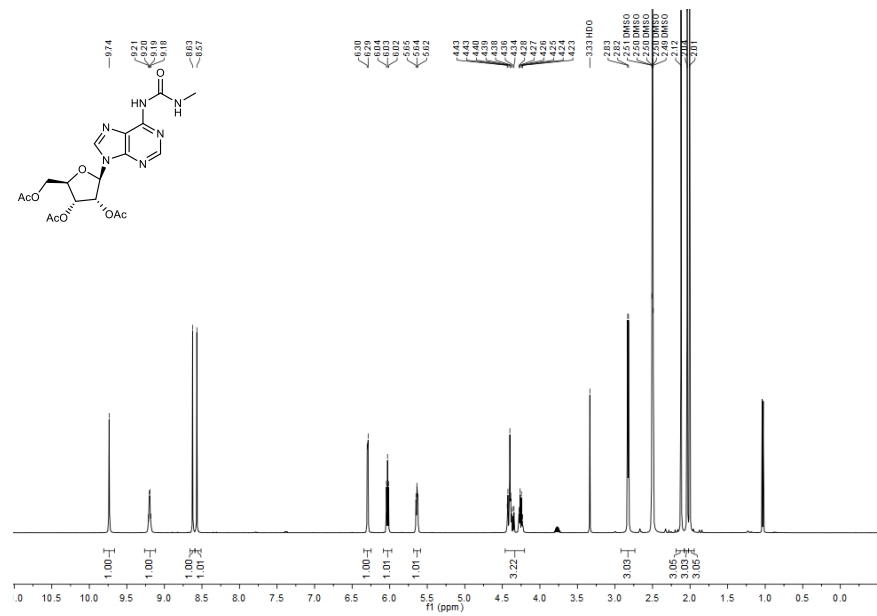


³¹P{¹H} NMR spectrum of compound 14j



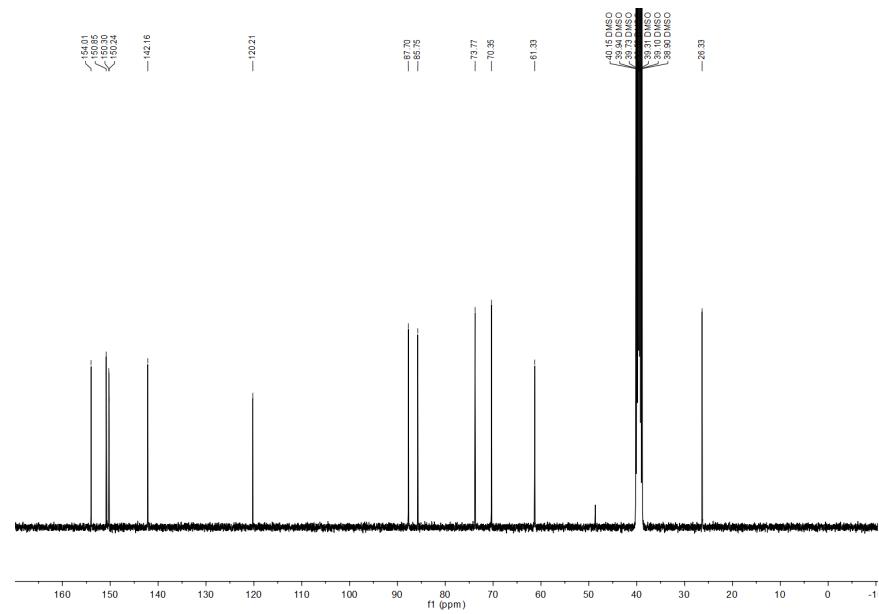
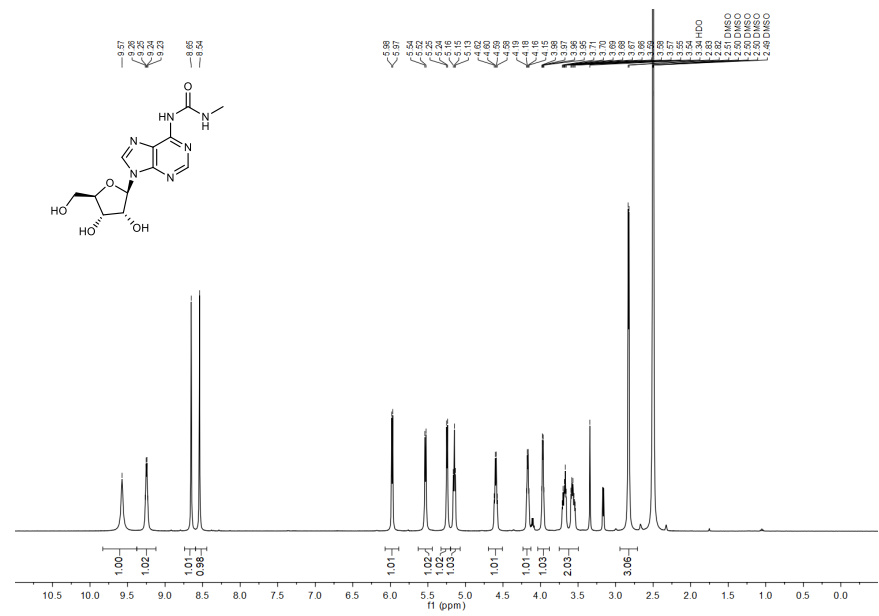
S139

¹H and ¹³C(¹H) NMR spectra of compound 18



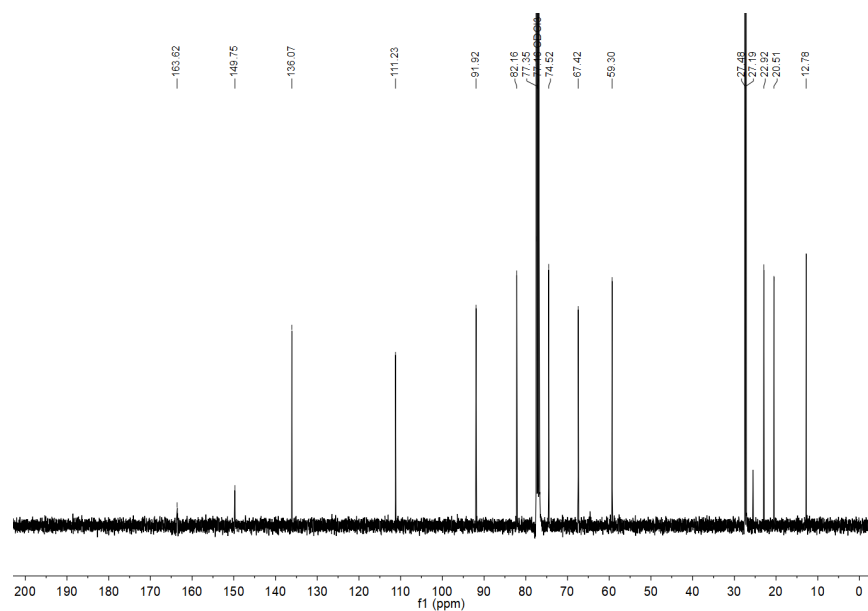
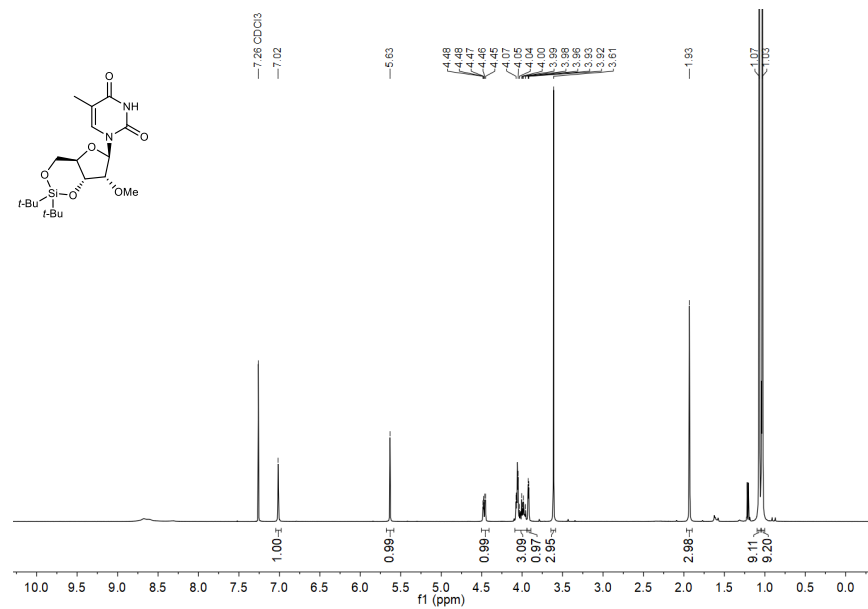
S142

¹H and ¹³C(¹H) NMR spectra of compound 19



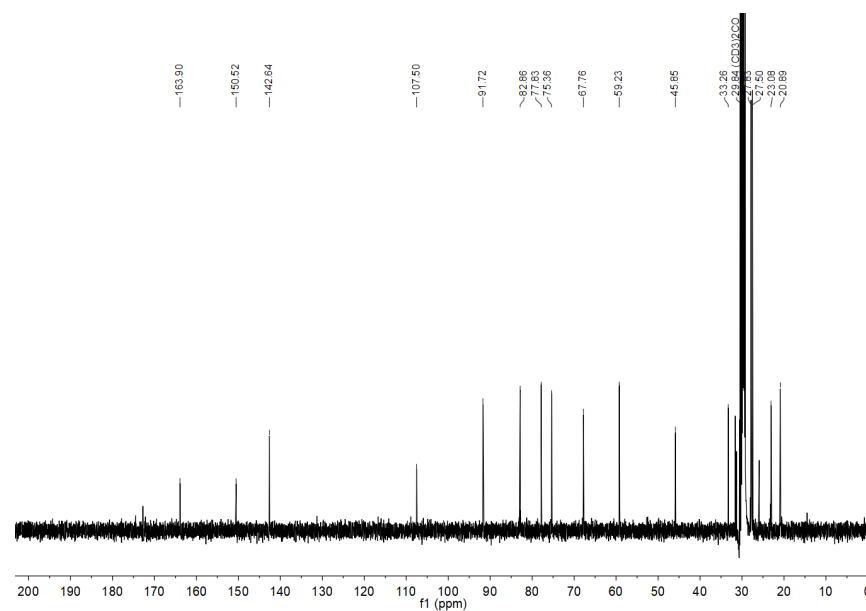
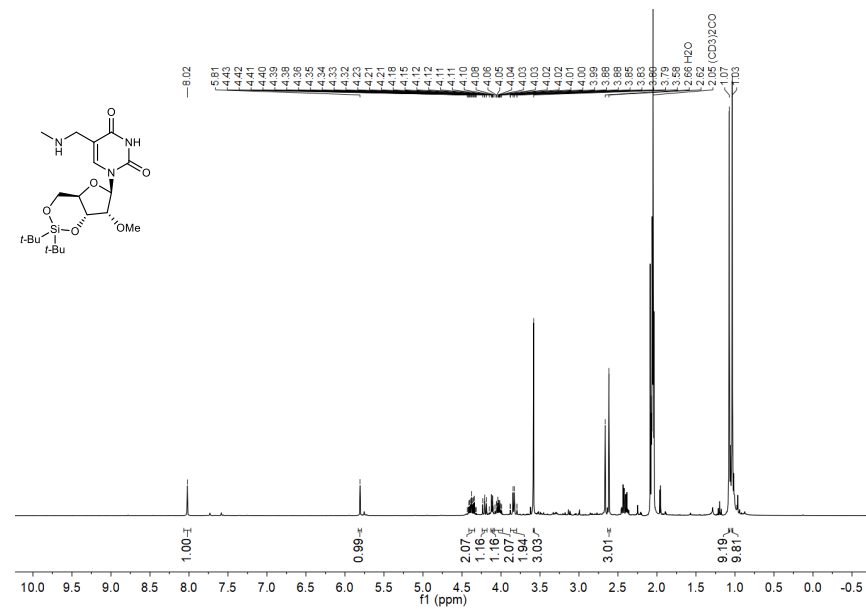
S143

¹H and ¹³C{¹H} NMR spectra of compound 21



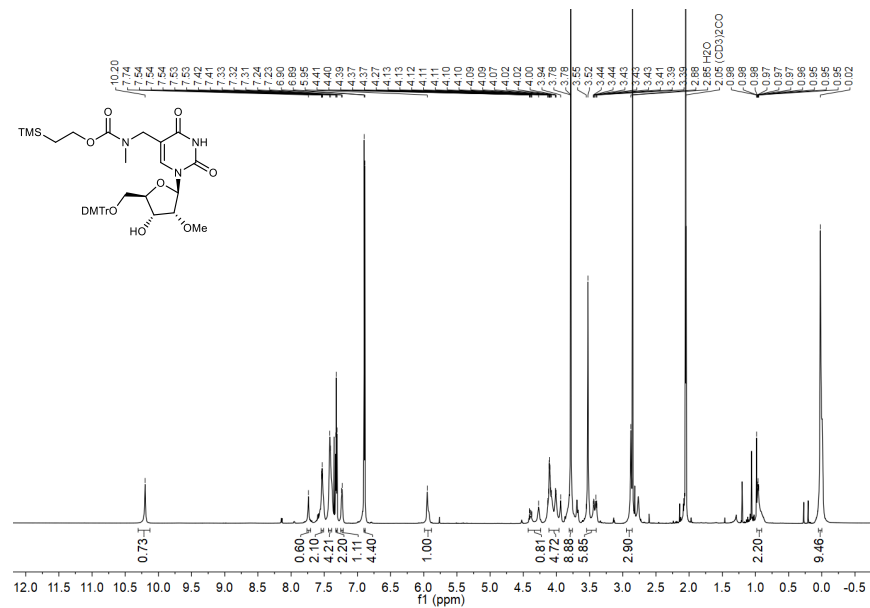
S144

¹H and ¹³C{¹H} NMR spectra of compound 22

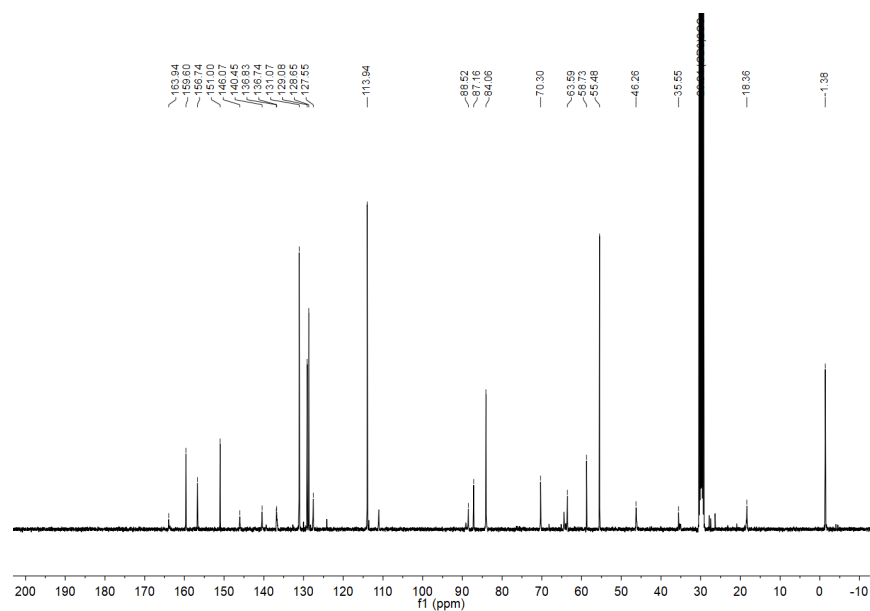
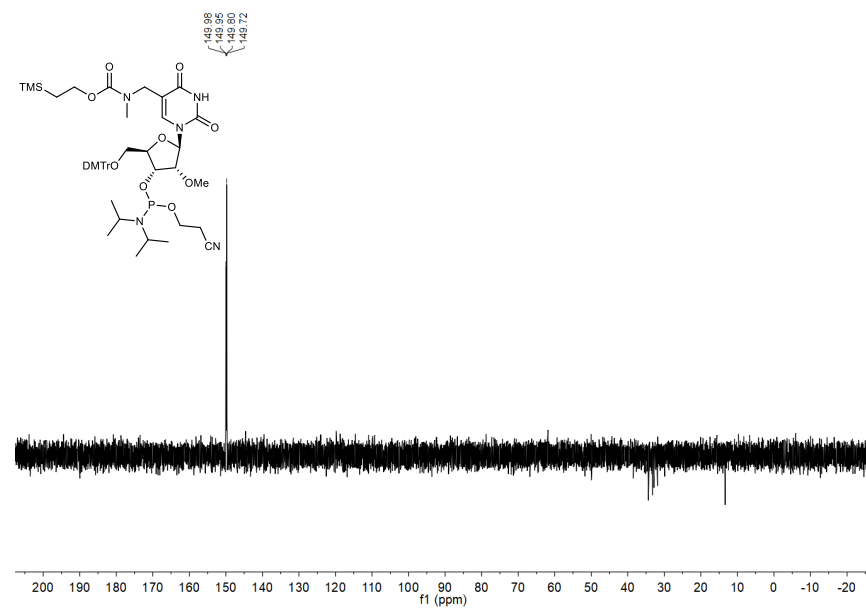


S145

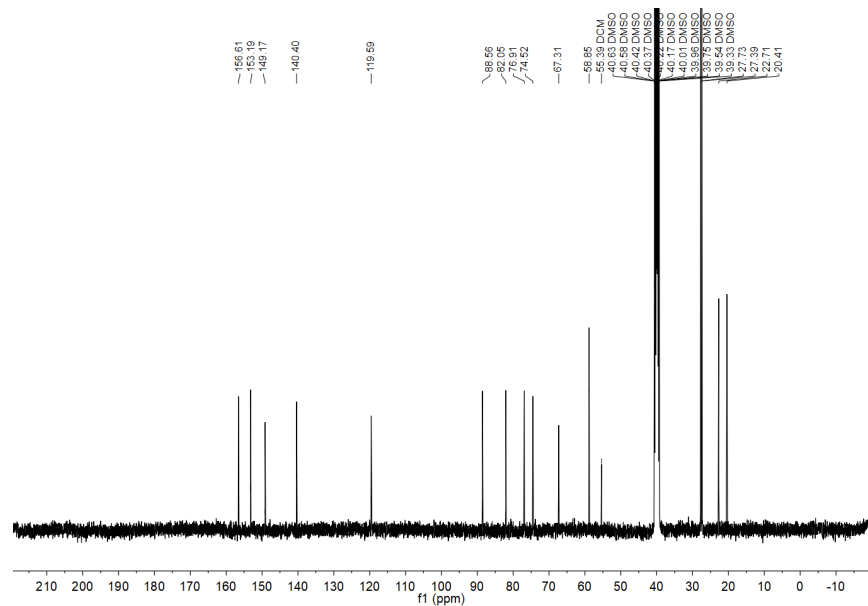
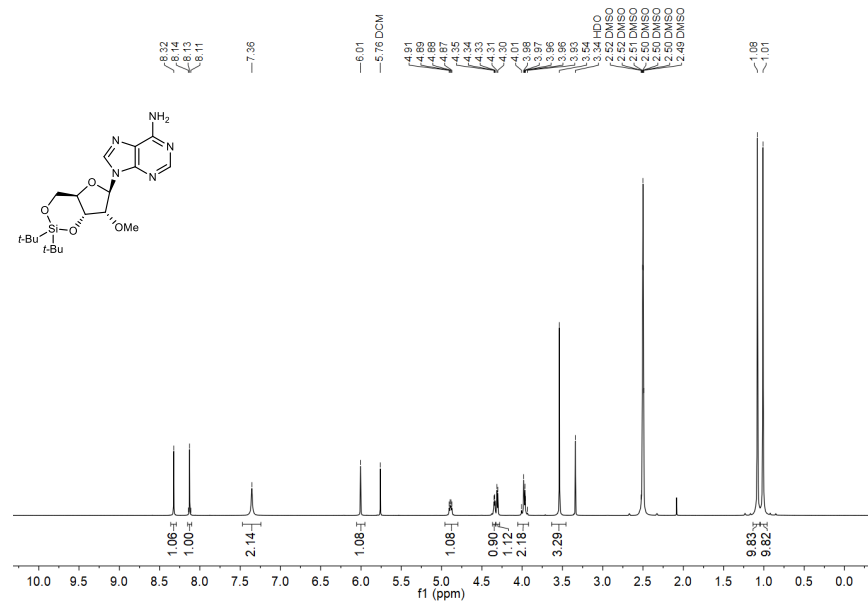
¹H and ¹³C{¹H} NMR spectra of compound 25



³¹P{¹H} NMR spectrum of compound 26

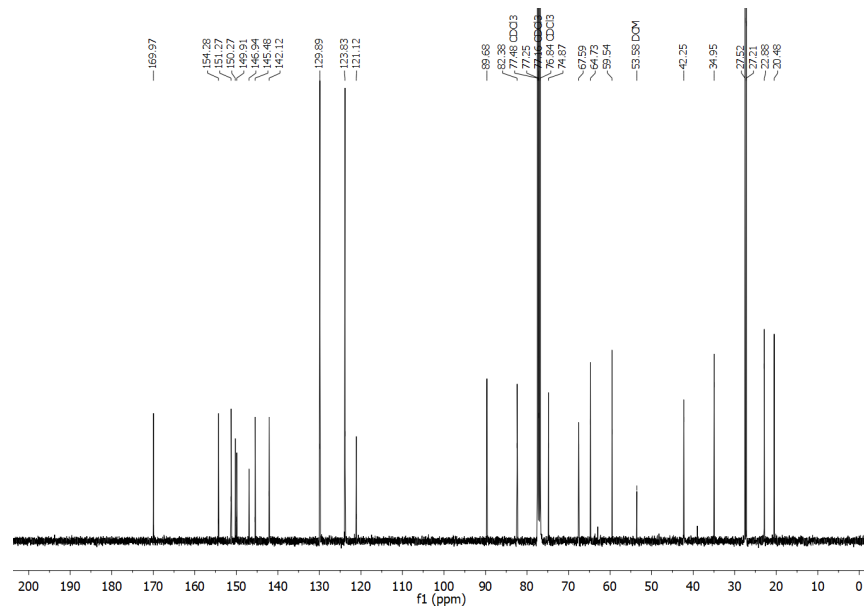
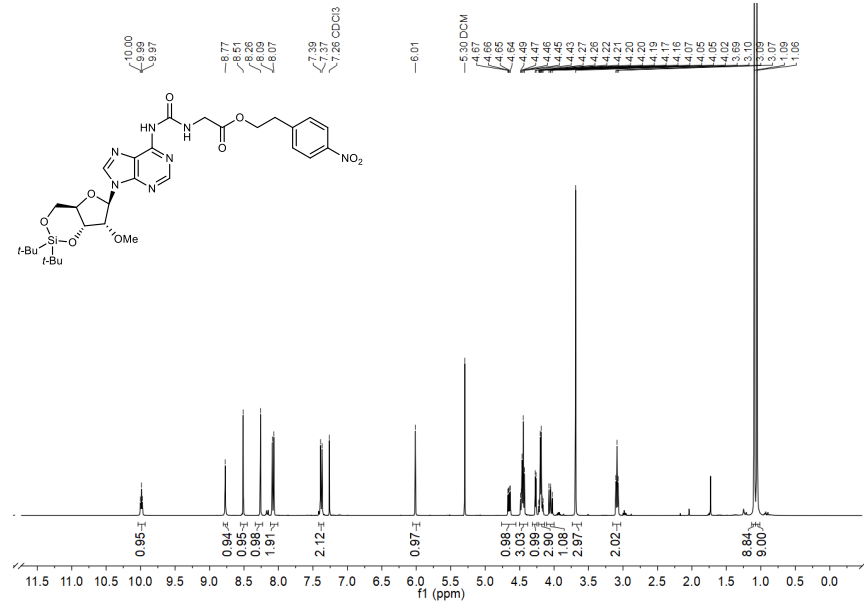


¹H and ¹³C(¹H) NMR spectra of compound 28



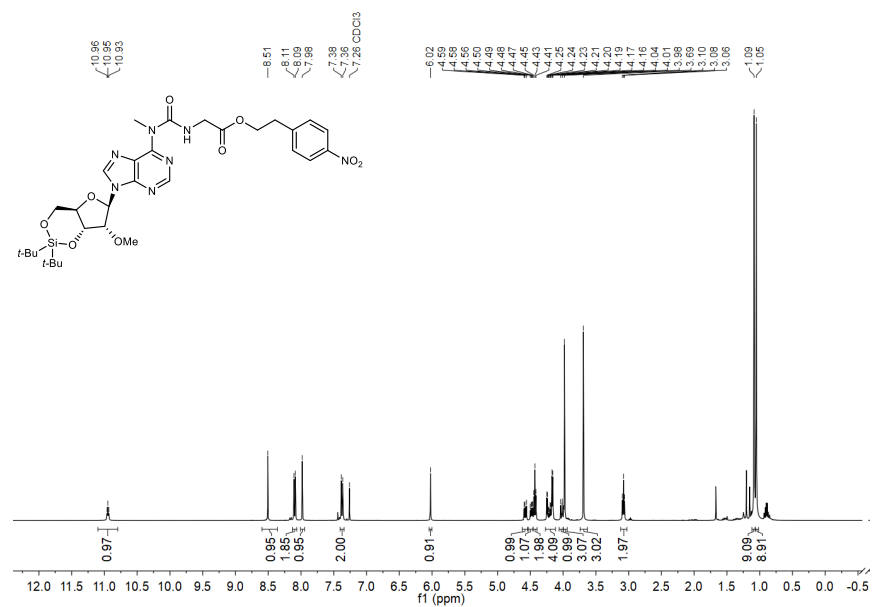
S150

¹H and ¹³C(¹H) NMR spectra of compound 29



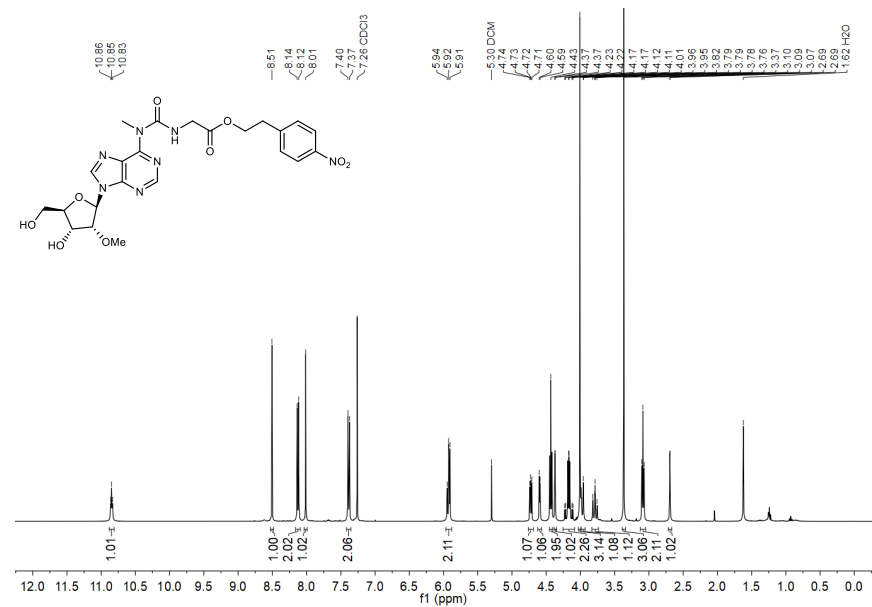
S151

¹H and ¹³C(¹H) NMR spectra of compound 30



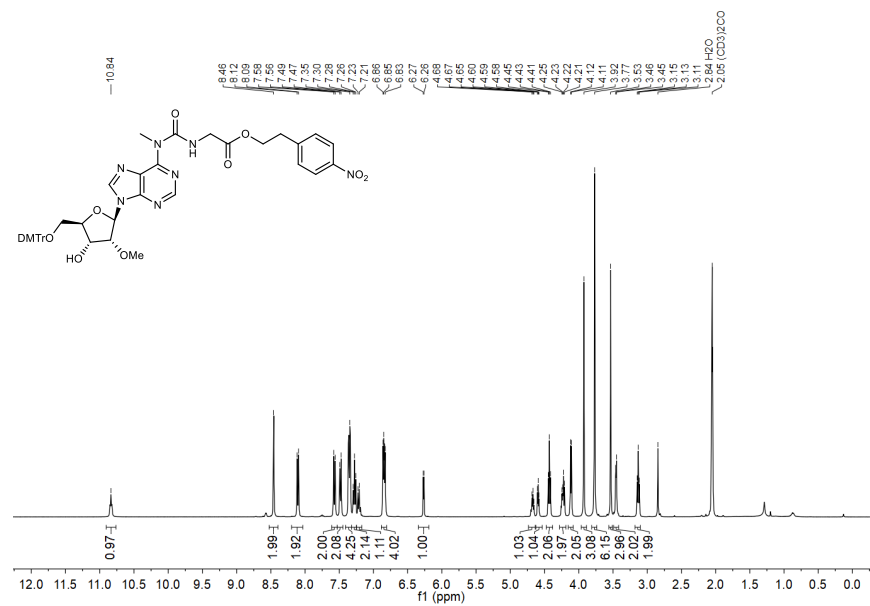
S152

¹H and ¹³C(¹H) NMR spectra of compound 31



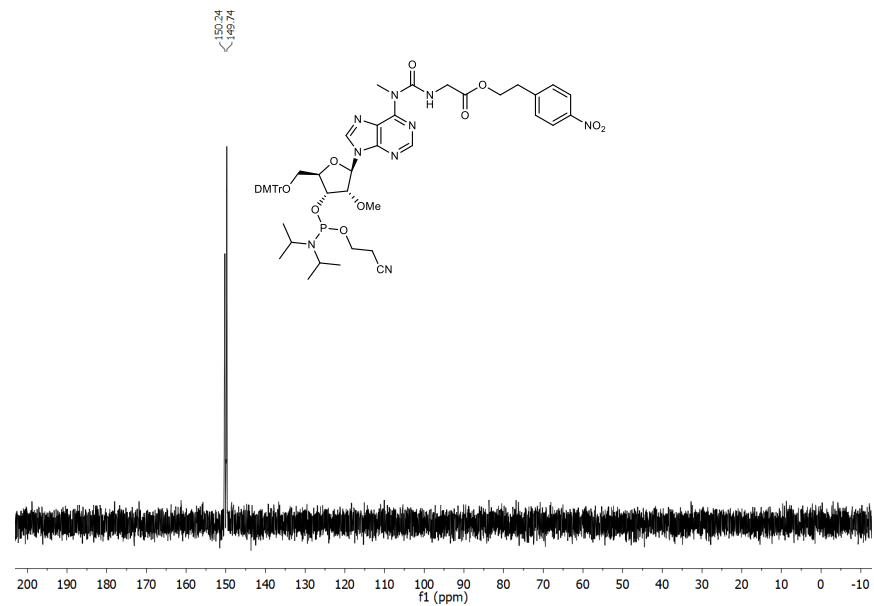
S153

¹H and ¹³C{¹H} NMR spectra of compound 32



S154

³¹P{¹H} NMR spectrum of compound 33



S155

18. References

1. Fulmer, G. R. *et al.* NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **29**, 2176-2179 (2010).
2. Tanpure, A. A. & Balasubramanian, S. Synthesis and Multiple Incorporations of 2'-O-Methyl-5-hydroxymethylcytidine, 5-Hydroxymethylcytidine and 5-Formylcytidine Monomers into RNA Oligonucleotides. *ChemBioChem* **18**, 2236-2241 (2017).
3. Shute, R. E. & Rich, D. H. Synthesis and Evaluation of Novel Activated Mixed Carbonate Reagents for the Introduction of the 2-(Trimethylsilyl)ethoxycarbonyl(Teoc)-Protecting Group. *Synthesis* **1987**, 346-349 (1987).
4. Nainyté, M. *et al.* Amino Acid Modified RNA Bases as Building Blocks of an Early Earth RNA-Peptide World. *Chem. Eur. J.* **26**, 14856–14860 (2020).
5. Serebryany, V. & Beigelman, L. An efficient preparation of protected ribonucleosides for phosphoramidite RNA synthesis. *Tetrahedron Lett.* **43**, 1983-1985 (2002).
6. Sundaram, M., Crain, P. F. & Davis, D. R. Synthesis and Characterization of the Native Anticodon Domain of *E. coli* tRNA^{Lys}: Simultaneous Incorporation of Modified Nucleosides mmm⁵s²U, t⁶A, and Pseudouridine Using Phosphoramidite Chemistry. *J. Org. Chem.* **65**, 5609-5614 (2000).
7. Matuszewski, M. & Sochacka, E. Stability studies on the newly discovered cyclic form of tRNA N⁶-threonylcarbamoyladenine (ct⁶A). *Bioorg. Med. Chem. Lett.* **24**, 2703-2706 (2014).
8. Schneider, C. *et al.* Noncanonical RNA Nucleosides as Molecular Fossils of an Early Earth—Generation by Prebiotic Methylations and Carbamoylations. *Angew. Chem. Int. Ed.* **57**, 5943-5946 (2018).
9. Himmelsbach, F., Schulz, B. S., Trichtinger, T., Charubala, R. & Pfeleiderer, W. The *p*-Nitrophenylethyl (NPE) Group: A Versatile New Blocking Group for Phosphate and Aglycone Protection in Nucleosides and Nucleotides. *Tetrahedron* **40**, 59-72 (1984).
10. Ferreira, F. & Morvan, F. Silyl Protecting Groups for Oligonucleotide Synthesis Removed by a ZnBr₂ Treatment. *Nucleosides, Nucleotides, and Nucleic Acids* **24**, 1009-1013 (2005).
11. Usanov, D. L., Chan, A. I., Maianti, J. P. & Liu, D. R. Second-generation DNA-templated macrocycle libraries for the discovery of bioactive small molecules. *Nat. Chem.* **10**, 704-714 (2018).
12. Hoops, S. *et al.* COPASI—a COmplex PAthway Simulator. *Bioinformatics* **22**, 3067-3074 (2006).
13. Jash, B., Tremmel, P., Jovanovic, D. & Richert, C. Single nucleotide translation without ribosomes. *Nat. Chem.* **13**, 751-757 (2021).
14. Mochrie, S. G. J. The Boltzmann factor, DNA melting, and Brownian ratchets: Topics in an introductory physics sequence for biology and premedical students. *Am. J. Phys* **79**, 1121-1126 (2011).
15. Senior, M. M., Jones, R. A. & Breslauer, K. J. Influence of loop residues on the relative stabilities of DNA hairpin structures. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6242-6246 (1988).
16. Xodo, L. E., Manzini, G., Quadrioglio, F., Marel, G. v. d. & van Boom, J. H. Hairpin structures in synthetic oligodeoxynucleotides: sequence effects on the duplex-to-hairpin transition. *Biochimie* **71**, 793-803 (1989).

Appendix II

Supporting Information

Loading of Amino Acids onto RNA in a Putative RNA-Peptide World

J. N. Singer, F. M. Müller, E. Węgrzyn, C. Hölzl, H. Hurmiz, C. Liu, L. Escobar, T. Carell**

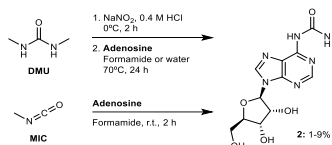
Table of Contents

1.	General information and instruments for nucleosides and phosphoramidites	S2
2.	Prebiotic synthesis of <i>N</i> ⁶ -methylurea adenosine	S2
3.	Prebiotic synthesis of amino acid-modified carbamoyl nucleosides	S4
3.1	Prebiotic synthesis of a series of amino acid-modified <i>N</i> ⁶ -carbamoyl adenosine nucleosides	S4
3.2	Prebiotic synthesis of amino acid-modified <i>N</i> ⁶ -carbamoyl guanosine nucleoside	S8
3.3	Prebiotic synthesis of amino acid-modified <i>N</i> ⁶ -carbamoyl cytidine nucleoside	S9
3.4	Prebiotic synthesis of amino acid-modified <i>N</i> ⁶ -methyl <i>N</i> ⁶ -carbamoyl adenosine nucleoside	S9
3.5	Control experiment of guanosine under the nitrosation conditions	S10
4.	Calibration curves of amino acid-modified carbamoyl nucleosides.....	S11
5.	General information and instruments for oligonucleotides.....	S12
6.	Prebiotic synthesis of amino acid-modified carbamoyl oligonucleotides	S14
6.1	Prebiotic synthesis of amino acid-modified carbamoyl oligonucleotides containing the four canonical bases	S14
6.2	Prebiotic synthesis of an oligonucleotide containing two amino acid-modified carbamoyl nucleotides at the terminal and internal positions	S15
6.3	Prebiotic synthesis of a series of amino acid-modified <i>N</i> ⁶ -carbamoyl adenosine oligonucleotides	S16
7.	Consecutive reactions using <i>N</i> ⁶ -methylurea adenosine oligonucleotide	S18
8.	Calibration curve of <i>N</i> ⁶ -methylurea adenosine oligonucleotide.....	S20
9.	Melting curve of double strand.....	S20
10.	Synthesis of methylurea and amino acid-modified carbamoyl nucleosides used as reference	S21
10.1	Synthesis of <i>N</i> ⁶ -methylurea adenosine	S21
10.2	Synthesis of amino acid-modified <i>N</i> ⁶ -carbamoyl adenosine nucleosides	S22
10.3	Synthesis of <i>N</i> ⁶ -methyl <i>N</i> ⁶ -methylurea adenosine	S25
10.4	Synthesis of amino acid-modified <i>N</i> ⁶ -methyl <i>N</i> ⁶ -carbamoyl adenosine nucleoside	S26
10.5	Synthesis of <i>N</i> ⁶ -methylurea guanosine and amino acid-modified <i>N</i> ⁶ -carbamoyl guanosine nucleoside used as reference.....	S26
10.6	Synthesis of <i>N</i> ⁶ -methylurea cytidine and amino acid-modified <i>N</i> ⁶ -carbamoyl cytidine nucleoside used as reference	S28
11.	Synthesis of methylurea nucleoside phosphoramidites	S29
11.1	Synthesis of <i>N</i> ⁶ -methylurea adenosine phosphoramidite.....	S29
11.2	Synthesis of <i>N</i> ⁶ -methylurea guanosine phosphoramidite.....	S31
11.3	Synthesis of <i>N</i> ⁶ -methylurea cytidine phosphoramidite.....	S32
12.	Synthesized oligonucleotides using a DNA/RNA automated synthesizer	S33
13.	NMR spectra of synthesized compounds	S34
14.	References.....	S72

1. General information and instruments for nucleosides and phosphoramidites

Reagents were purchased from commercial suppliers and used without further purification unless otherwise stated. Anhydrous solvents, stored under inert atmosphere, were also purchased. All reactions involving air/moisture sensitive reagents/intermediates were performed under inert atmosphere using oven-dried glassware. Routine ^1H NMR, $^{13}\text{C}\{^1\text{H}\}$ NMR and $^{31}\text{P}\{^1\text{H}\}$ NMR spectra were recorded on a Bruker Ascend 400 spectrometer (400 MHz for ^1H NMR, 100 MHz for ^{13}C NMR and 162 MHz for ^{31}P NMR), Bruker Ascend 500 spectrometer (500 MHz for ^1H NMR, 125 MHz for ^{13}C NMR and 202 MHz for ^{31}P NMR) or Bruker ARX 600 spectrometer (600 MHz for ^1H NMR, 150 MHz for ^{13}C NMR and 243 MHz for ^{31}P NMR). Deuterated solvents used are indicated in the characterization and chemical shifts (δ) are reported in ppm. Residual solvent peaks were used as reference.¹ All NMR J values are given in Hz. COSY, HSQC and HMBC experiments were recorded to help with the assignment of ^1H and ^{13}C signals. NMR spectra were analyzed using MestReNova software version 10.0. High Resolution Mass Spectra (HRMS) were measured on a Thermo Finnigan LTQ-FT with ESI as ionization mode. IR spectra were recorded on a Perkin-Elmer Spectrum BX II FT-IR instrument or Shimadzu IRSpirit FT-IR instrument. Both equipped with an ATR accessory. Column chromatography was performed with silica gel technical grade, 40-63 μm particle size. Reaction progress was monitored by Thin Layer Chromatography (TLC) analysis on silica gel 60 F254 and stained with *para*-anisaldehyde, potassium permanganate or cerium ammonium molybdate solution.

2. Prebiotic synthesis of *N*⁶-methylurea adenosine



Scheme S1. Prebiotic synthesis of *N*⁶-methylurea adenosine **2**.

Procedure for the prebiotic synthesis of **2** using DMU:

Step 1: The nitrosation reaction of 1,3-dimethylurea (**DMU**) was carried out following a procedure previously reported in the literature.² Step 2: Adenosine (2.67 mg, 10 μmol , 1 equiv.) and 1,3-dimethyl-1-nitrosourea (58.6 mg, 0.5 mmol, 50 equiv.) were dissolved in formamide (1 mL) or water (40 μL). The reaction was stirred at 70°C for 24 h. For the reaction in water solution, the crude was diluted with water (up to 1.0 mL) after 24 h. An aliquot (10 μL) of the crude reaction mixture was taken, diluted with water (up to 1.0 mL), filtered and analyzed by LC-MS (buffer A: 2 mM HCOONH₄ pH 5.5 in H₂O and buffer B: 2 mM HCOONH₄ pH 5.5 in 20:80 H₂O/MeCN; Gradient with B: 0-15% from 0 to 15 min and 15-20% from 15 to 30 min; Temperature = 40°C; Flow rate = 0.15 mL·min⁻¹ and Injection volume = 5 μL).

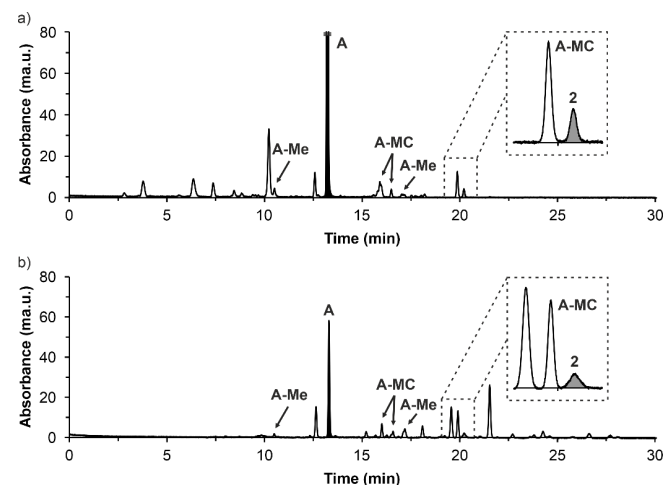


Figure S1. HPLC chromatograms of the crude reaction mixtures for the prebiotic synthesis of **2** using 1,3-dimethylurea (**DMU**) in: a) formamide and b) water. The chromatographic peaks assigned as **A-Me** corresponded to methylated Adenosine derivatives. In turn, the chromatographic peaks assigned as **A-MC** corresponded to Adenosine derivatives bearing a *N*-methylcarbamoyl substituent at the OH groups of the ribose. Structural assignments for **A-Me** and **A-MC** were not performed.

Procedure for the prebiotic synthesis of **2** using MIC:

Adenosine (10.0 mg, 37.4 μmol , 1 equiv.) and methylisocyanate (**MIC**, 12.2 μL , 206 μmol , 5.5 equiv.) were dissolved in formamide (1 mL). The reaction was stirred at r.t. for 2 h. An aliquot (25 μL) of the crude reaction mixture was taken, diluted with water (up to 1.0 mL), filtered and analyzed by LC-MS (buffer A: 2 mM HCOONH₄ pH 5.5 in H₂O and buffer B: 2 mM HCOONH₄ pH 5.5 in 20:80 H₂O/MeCN; Gradient with B: 0-25% from 0 to 45 min; Temperature = 40°C; Flow rate = 0.15 mL·min⁻¹ and Injection volume = 5 μL).

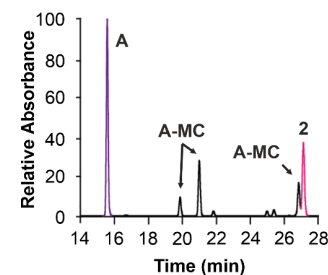


Figure S2. HPLC chromatogram of the crude reaction mixture for the prebiotic synthesis of **2** using methyl isocyanate (**MIC**) in formamide. The chromatographic peaks assigned as **A-MC** corresponded to Adenosine derivatives bearing a *N*-methylcarbamoyl substituent at the OH groups of the ribose. Structural assignments for **A-MC** were not performed.

Procedure for the hydrolysis reaction of the crude reaction mixture obtained from the prebiotic synthesis of **2** using DMU:

The crude reaction mixture in formamide solution (100 μL) was diluted with 50 mM borate buffer pH 9.5 (up to 1.0 mL). The reaction was stirred at 70°C for 24 h. An aliquot (100 μL) of the crude reaction mixture was taken, diluted with water (up to 1.0 mL), filtered and analyzed by LC-MS (buffer A: 2 mM HCOONH₄ pH 5.5 in H₂O and buffer B: 2 mM HCOONH₄ pH 5.5 in 20:80 H₂O/MeCN; Gradient with B: 0-15% from 0 to 15 min and 15-20% from 15 to 30 min; Temperature = 40°C; Flow rate = 0.15 mL·min⁻¹ and Injection volume = 5 μL).

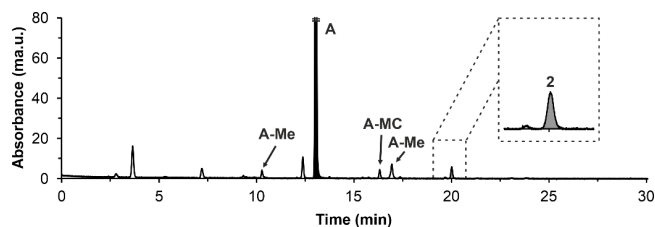
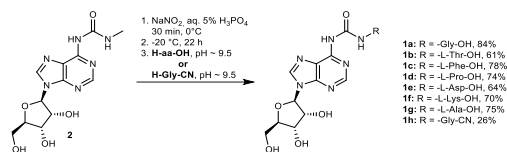


Figure S3. HPLC chromatogram of the crude reaction mixture after the hydrolysis reaction with 50 mM borate buffer pH 9.5 at 70°C for 24 h. The chromatographic peaks assigned as **A-Me** corresponded to methylated Adenosine derivatives. In turn, the chromatographic peak assigned as **A-MC** corresponded to an Adenosine derivative bearing a *N*-methylcarbamoyl substituent at one of the OH groups of the ribose. Structural assignments for **A-Me** and **A-MC** were not performed.

3. Prebiotic synthesis of amino acid-modified carbamoyl nucleosides

3.1 Prebiotic synthesis of a series of amino acid-modified *N*⁶-carbamoyl adenosine nucleosides



Scheme S2. Prebiotic synthesis of amino acid-modified *N*⁶-carbamoyl adenosine nucleosides **1a-h** from *N*⁶-methylurea adenosine **2** (optimized reaction conditions are shown).

Optimization of the reaction conditions using **2** and **H-Gly-OH**:

Step 1: *N*⁶-methylurea adenosine **2** (1.00 mg, 3.08 μmol, 1.0 equiv.) and NaNO₂ (2.66 mg, 38.54 μmol, 12.5 equiv.) were dissolved in an acidic aqueous solution (150 μL, see Table S1 for acids used). The reaction was stirred at 0°C (see Table S2 for time). Step 2: After that, the solution was kept in a freezer at -20°C (see Table S3 for time). Step 3: The thawed adenosine's solution was transferred to a 30 mM borate-buffered solution (3.00 mL) containing the amino acid **H-Gly-OH** (2.31 mg, 30.84 μmol, 10.0 equiv.). The pH was adjusted using a 4 M aqueous NaOH solution (see Table S4 for pH). The reaction was stirred at r.t. for 1 h. Finally, an aliquot (25 μL) of the crude reaction mixture was taken, diluted with water (up to 1.0 mL), filtered and analyzed by LC-MS (buffer A: 2 mM HCOONH₄ pH 5.5 in H₂O and buffer B: 2 mM HCOONH₄ pH 5.5 in 20:80 H₂O/MeCN; Gradient: 0-20% of B in 25 min; Flow rate = 0.15 mL·min⁻¹ and Injection volume = 5 μL).

Table S1. Screening of acidic aqueous solutions in step 1. Reaction conditions in other steps were kept constant: step 1) 30 min at 0°C; step 2) 22 h at -20°C and step 3) pH ~ 9.5.

Acidic aqueous solution in step 1	Yield (%) of 1a
1% H ₃ PO ₄	3
5% H₃PO₄	84
1 M HCl	78
1 M H ₂ SO ₄	64
10% acetic acid (AcOH)	Not detected
neat acetic acid (AcOH)	Not detected
50% formic acid (FA)	3
neat formic acid (FA)	18

Table S2. Screening of reaction time in step 1. Reaction conditions in other steps were kept constant: step 1) 5% H₃PO₄ at 0°C; step 2) 22 h at -20°C and step 3) pH ~ 9.5.

Time (min) in step 1	Yield (%) of 1a
30	84
60	76
120	57

Table S3. Screening of reaction time in step 2. Reaction conditions in other steps were kept constant: step 1) 5% H₃PO₄ at 0°C for 30 min; step 2) -20°C and step 3) pH ~ 9.5.

Time (h) in step 2	Yield (%) of 1a
0	4
1	16
3	42
22	84

Table S4. Screening of pH values in step 3. Reaction conditions in other steps were kept constant: step 1) 5% H₃PO₄ at 0°C for 30 min and step 2) 22 h at -20°C.

pH of aqueous solution in step 3	Yield (%) of 1a
-6.2	5
-6.9	24
-7.4	34
-8.6	60
-9.5	84

General procedure for the prebiotic synthesis of **1a-h** under the optimized reaction conditions:

In the general procedure, the amino acid was added in step 3.

Step 1: *N*⁶-methylurea adenosine **2** (1.00 mg, 3.08 μmol, 1.0 equiv.) and NaNO₂ (2.66 mg, 38.54 μmol, 12.5 equiv.) were dissolved in 5% aqueous H₃PO₄ solution (150 μL). The reaction was stirred at 0°C for 30 min. Step 2: After that, the solution was kept in the freezer at -20°C for 22 h. Step 3: The thawed adenosine's solution was transferred to a 30 mM borate-buffered solution (3.00 mL) containing the amino acid **H-aa-OH** or the amino nitrile **H-Gly-CN** (30.84 μmol, 10.0 equiv.). The pH was adjusted to ca. 9.5 using a 4 M aqueous NaOH solution (60 μL). The reaction was stirred at r.t. for 1 h. For **1a-g**, an aliquot (25 μL) of the crude reaction mixture was taken, diluted with water (up to 1.0 mL), filtered and analyzed by LC-MS (buffer A: 2 mM HCOONH₄ pH 5.5 in H₂O and buffer B: 2 mM HCOONH₄ pH 5.5 in 20:80 H₂O/MeCN; Gradient: 0-20% of B in 25 min; Flow rate = 0.15 mL·min⁻¹ and Injection volume = 5 μL). For **1h**, an aliquot (10 μL) of the crude reaction mixture was taken, diluted with water (up to 40 μL) and analyzed by HPLC (A: H₂O and B: 20:80 H₂O/MeCN; Gradient: 0-30% of B in 30 min; Flow rate = 1 mL·min⁻¹ and Injection volume = 20 μL).

Modified procedure for the prebiotic synthesis of **1c** under the optimized reaction conditions:

In the modified procedure, the amino acid was added in step 1 (nitrosation reaction).

Step 1: *N*⁶-methylurea adenosine **2** (1.00 mg, 3.08 μmol, 1.0 equiv.), the amino acid **H-Phe-OH** (30.84 μmol, 10.0 equiv.) and NaNO₂ (2.66 mg, 38.54 μmol, 12.5 equiv.) were dissolved in 5% aqueous H₃PO₄ solution (150 μL). The reaction was stirred at 0°C for 30 min. Step 2: After that, the solution was kept in the freezer at -20°C for 22 h. Step 3: The thawed adenosine's solution was transferred to a 30 mM borate-buffered solution (3.00 mL) and the pH was adjusted to ca. 9.5 using a 4 M aqueous NaOH solution (60 μL). The reaction was stirred at r.t. for 1 h. After that, an aliquot (10 μL) of the crude reaction mixture was taken, diluted with water (up to 40 μL) and analyzed by HPLC (A: H₂O and B: 20:80 H₂O/MeCN; Gradient: 0-30% of B in 50 min; Flow rate = 1 mL·min⁻¹ and Injection volume = 20 μL).

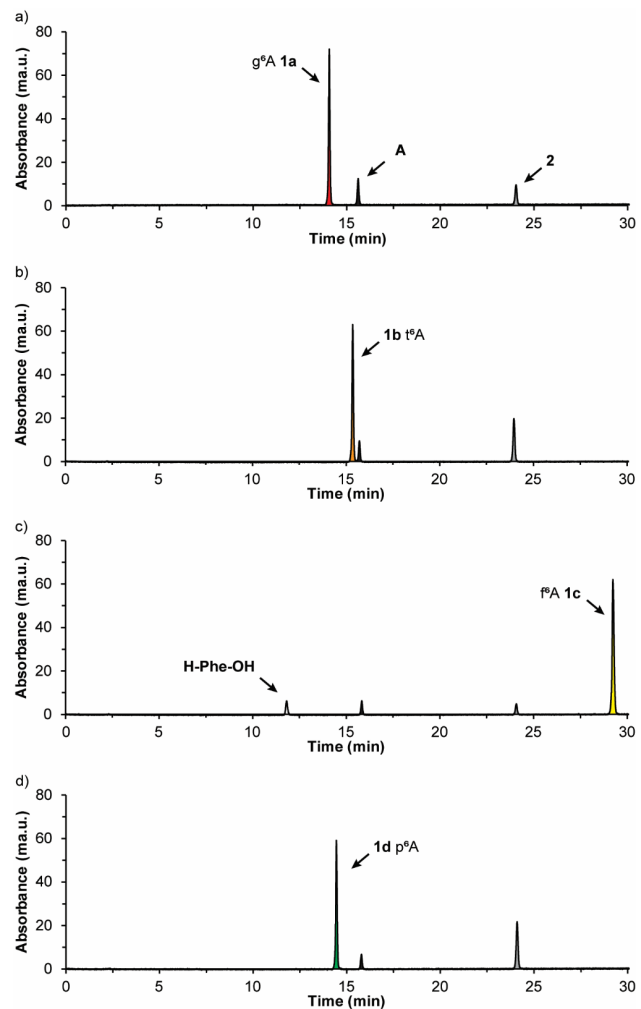


Figure S4. HPLC chromatograms of the crude reaction mixtures for the prebiotic synthesis of: a) g^A **1a**; b) t^A **1b**; c) f^A **1c** and d) p^A **1d**. **A** = Adenosine.

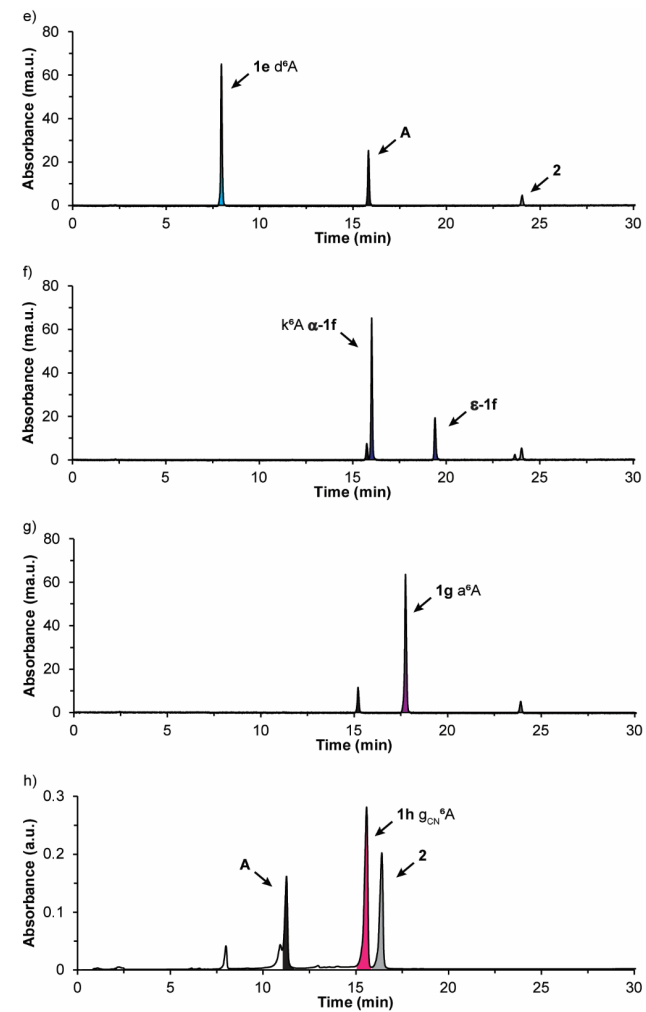


Figure S5. HPLC chromatograms of the crude reaction mixtures for the prebiotic synthesis of: e) d^A **1e**; f) k^A α -**1f** (ϵ -**1f** is the product of the reaction between the amino group adjacent to the carbon at the ϵ -position of **H-Lys-OH** with the nitrosated derivative of **1**); g) a^A **1g** and h) g_{CN}^A **1h**. **A** = Adenosine.

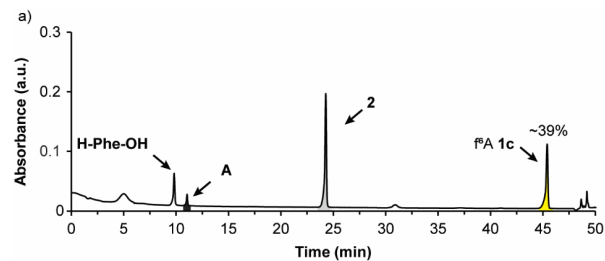


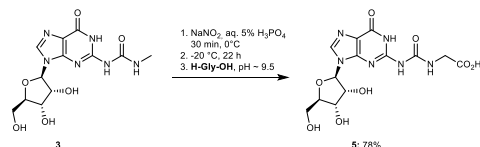
Figure S6. HPLC chromatogram of the crude reaction mixture for the prebiotic synthesis of *f*⁶A **1c** following the modified procedure, *i.e.* addition of the amino acid in step 1 (nitrosation reaction).

Table S5. Results obtained for the synthesis of **1a-h** under prebiotically plausible reaction conditions.

Compound	Yield (%)
1a ; g ⁶ A	84
1b ; t ⁶ A	61
1c ; f ⁶ A	78 (39) ^a
1d ; p ⁶ A	74
1e ; d ⁶ A	64
α-1f ; k ⁶ A	70
ε-1f	16
1g ; a ⁶ A	75
1h ; g ⁶ A	26

^a Result obtained following the modified procedure, *i.e.* addition of the amino acid in step 1 (nitrosation reaction).

3.2 Prebiotic synthesis of amino acid-modified *N*⁶-carbamoyl guanosine nucleoside



Scheme S3. Prebiotic synthesis of amino acid-modified *N*⁶-carbamoyl guanosine nucleoside **5** from *N*⁶-methylurea guanosine **3** under the optimized reaction conditions.

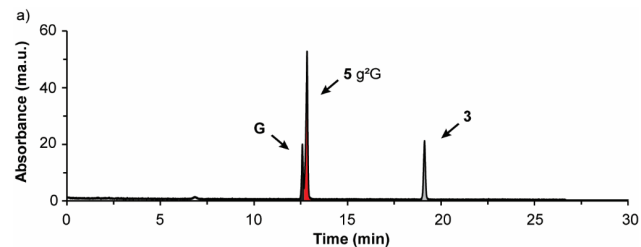
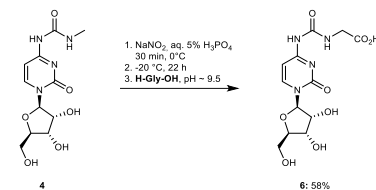


Figure S7. HPLC chromatogram of the crude reaction mixture for the prebiotic synthesis of g²G **5**. **G** = guanosine.

3.3 Prebiotic synthesis of amino acid-modified *N*⁶-carbamoyl cytidine nucleoside



Scheme S4. Prebiotic synthesis of amino acid-modified *N*⁶-carbamoyl cytidine nucleoside **6** from *N*⁶-methylurea cytidine **4** under the optimized reaction conditions.

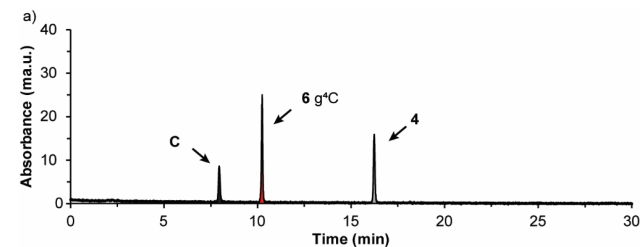
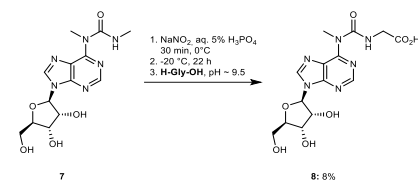


Figure S8. HPLC chromatogram of the crude reaction mixture for the prebiotic synthesis of g³C **6**. **C** = cytidine.

3.4 Prebiotic synthesis of amino acid-modified *N*⁶-methyl *N*⁶-carbamoyl adenosine nucleoside



Scheme S5. Prebiotic synthesis of amino acid-modified *N*⁶-methyl *N*⁶-carbamoyl adenosine nucleoside **8** from *N*⁶-methyl *N*⁶-methylurea adenosine **7** under the optimized reaction conditions.

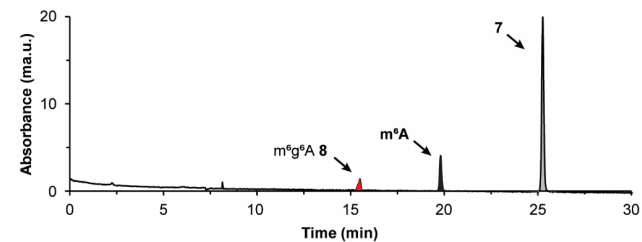
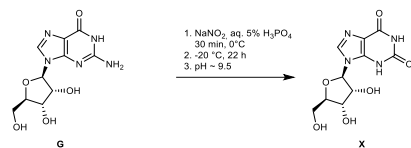


Figure S9. HPLC chromatogram of the crude reaction mixture for the prebiotic synthesis of m⁶g⁶A **8**. m⁶A = *N*⁶-methyl adenosine.

3.5 Control experiment of guanosine under the nitrosation conditions



Scheme S6. Control experiment of guanosine **G** under the nitrosation conditions. **X** = xanthosine.

The control experiment indicated that **G** gives **X** under the nitrosation conditions used in Section 3.1.

4. Calibration curves of amino acid-modified carbamoyl nucleosides

Amino acid-modified carbamoyl nucleosides **1a-h**, **5** and **6** were used for the development of calibration curves. Separate stock solutions of the modified nucleosides were prepared in water (100 μM). Dilute standard solutions of the modified nucleosides (1; 2; 4; 5; 6; 8 μM) were prepared in a final volume of 1 mL. The standard solutions were injected in an analytical UHPLC equipped with a C18 column (buffer A: 2 mM HCOONH₄ pH 5.5 in H₂O and buffer B: 2 mM HCOONH₄ pH 5.5 in 20:80 H₂O/MeCN; Gradient: 0-20% of B in 25 min; Flow rate = 0.15 mL·min⁻¹ and Injection volume = 20 μL). For **1h**, a stock solution was prepared in water (1 mM). Dilute standard solutions of **1h** (50; 100; 200; 300; 400; 500 μM) were prepared in a final volume of 20 μL. The standard solutions were injected in an analytical HPLC equipped with a C18 column (A: H₂O and B: 20:80 H₂O/MeCN; Gradient: 0-30% of B in 30 min; Flow rate = 1 mL·min⁻¹ and Injection volume = 20 μL). The absorbance was monitored at 260 nm and the areas of the chromatographic peaks were determined by integration. The plot of the chromatographic area (a.u.) versus the amount of each nucleoside followed a linear relationship.

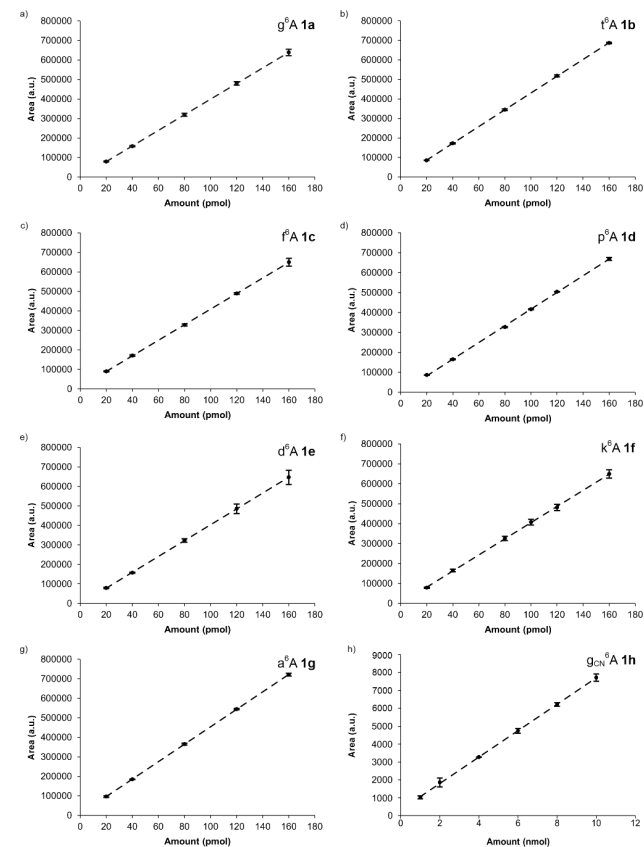


Figure S10. Chromatographic area (a.u.) vs. amount (pmol or nmol) of: a) **g⁰A 1a**; b) **t⁰A 1b**; c) **f⁰A 1c**; d) **p⁰A 1d**; e) **d⁰A 1e**; f) **k⁰A 1f**; g) **a⁰A 1g** and h) **g_{cn}⁰A 1h**. Lines show the fit of the data to a linear regression equation. Error bars are standard deviations from two independent experiments.

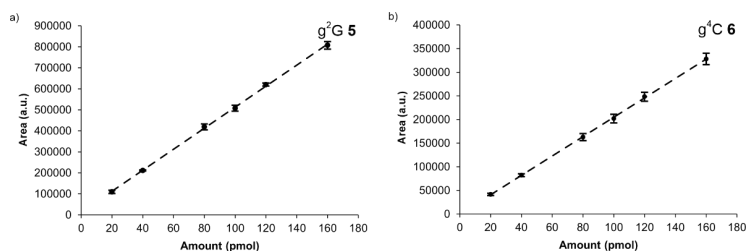


Figure S11. Chromatographic area (a.u.) vs. amount (pmol) of: a) g²G 5 and b) g⁴C 6. Lines show the fit of the data to a linear regression equation. Error bars are standard deviations from two independent experiments.

Table S6. Calibration curves ($y = mx + n$) obtained by analysis of the chromatographic peaks of **1a-h**, **5** and **6**.

Compound	Slope, m (pmol ⁻¹)	Intercept, n	r ²
g ⁴ A 1a	3999.8	-1195.1	0.99
t ⁶ A 1b	4300.7	-137.4	0.99
f ⁶ A 1c	3990.0	10177.5	0.99
p ⁶ A 1d	4183.8	-1542.8	0.99
d ⁶ A 1e	4062.7	-3010.2	0.99
k ⁶ A α-1f	4046.2	552.5	0.99
a ⁶ A 1g	4472.7	6578.1	0.99
g ⁶ A 1h ^a	738.7	318.5	0.99
g ² G 5	5000.9	12316.0	0.99
g ⁴ C 6	2050.4	-321.0	0.99

^a Slope (m) in nmol⁻¹.

5. General information and instruments for oligonucleotides

Synthesis and purification of oligonucleotides

Phosphoramidites of canonical ribonucleosides (Bz-A-CE, Dmf-G-CE, Ac-C-CE and U-CE) were purchased from LinkTech and Sigma-Aldrich. Oligonucleotides (ONs) were synthesized on a 1 μmol scale using RNA SynBase™ CPG 1000/110 and High Load Glen UnySupport™ as solid supports using an RNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with a standard phosphoramidite chemistry. ONs were synthesized in DMT-OFF mode using DCA as deblocking agent in CH₂Cl₂, BTT or Activator 42® as activator in MeCN, Ac₂O as capping reagent in pyridine/THF and I₂ as oxidizer in pyridine/H₂O.

Deprotection of npe and teoc groups

For the deprotection of the *para*-nitrophenylethyl (npe) group in ONs containing amino acid-modified carbamoyl adenosine nucleosides, the solid support beads were suspended in a 9:1 THF/DBU solution mixture (1 mL) and incubated at r.t. for 2 h.³ After that, the supernatant was removed, and the beads were washed with THF (3x1 mL).

For the deprotection of the 2-(trimethylsilyl)ethoxycarbonyl (teoc) group in ONs containing 5-methylaminomethyl uridine nucleosides, the solid support beads were suspended in a saturated solution of ZnBr₂ in 1:1 MeNO₂/IPA (1 mL) and incubated at r.t. overnight.⁴ After that, the supernatant was removed, and the beads were washed with 0.1 M EDTA in water (1 mL) and water (1 mL).

Cleavage from beads, deprotection of TBS groups and precipitation of the synthesized ON

The solid support beads were suspended in a 1:1 aqueous solution mixture (0.6 mL) of 30% NH₄OH and 40% MeNH₂. The suspension was heated at 65°C (8 min for SynBase™ CPG 1000/110 and 60 min for High Load Glen

UnySupport™). Subsequently, the supernatant was collected, and the beads were washed with water (2x0.3 mL). The combined aqueous solutions were concentrated under reduced pressure using a SpeedVac concentrator. After that, the crude was dissolved in DMSO (100 μL) and triethylamine trihydrofluoride (125 μL) was added. The solution was heated at 65°C for 1.5 h. Finally, the ON was precipitated by adding 3 M NaOAc in water (25 μL) and *n*-butanol (1 mL). The mixture was kept at -80°C for 2 h and centrifuged at 4°C for 1 h. The supernatant was removed and the white precipitate was lyophilized.

Purification of the synthesized ON by HPLC and desalting

The crude was purified by semi-preparative HPLC (1260 Infinity II Manual Preparative LC System from Agilent equipped with a G7114A detector) using a reverse-phase (RP) VP 250/10 Nucleodur 100-5 C18ec column from Macherey-Nagel (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-25% of B in 45 min; Flow rate = 5 mL·min⁻¹). The purified ON was analyzed by RP-HPLC (1260 Infinity II LC System from Agilent equipped with a G7165A detector) using an EC 250/4 Nucleodur 100-3 C18ec from Macherey-Nagel (Gradient: 0-30% of B in 45 min; Flow rate = 1 mL·min⁻¹). Finally, the purified ON was desalted using a C18 RP-cartridge from Waters.

Determination of the concentration and the mass of the synthesized ON

The absorbance of the synthesized ON in H₂O solution was measured using an IMPLen NanoPhotometer® N60/N50 at 260 nm. The extinction coefficient of the ON was calculated using the OligoAnalyzer Version 3.0 from Integrated DNA Technologies. For ONs incorporating non-canonical bases, the extinction coefficients were assumed to be identical to those containing only canonical counterparts.

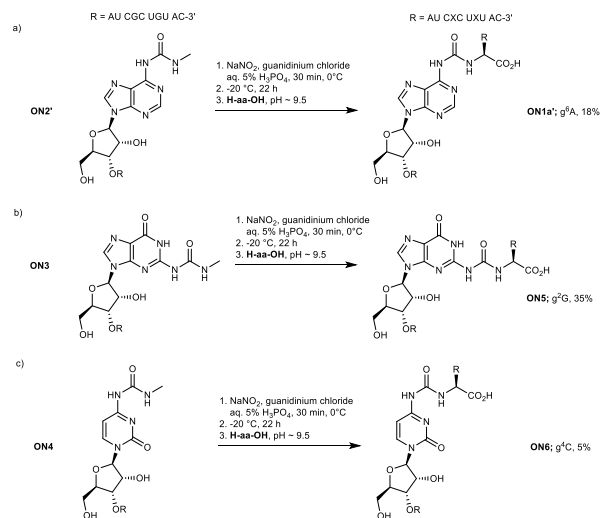
The synthesized ON (2-3 μL) was desalted on a 0.025 μm VSWP filter (Millipore), co-crystallized in a 3-hydroxypicolinic acid matrix (HPA, 1 μL) and analyzed by MALDI-TOF mass spectrometry (negative mode).

Enzymatic digestion of ONs into nucleosides

The ON (~100 pmol) was diluted with the nucleoside digestion mix reaction buffer (10X, 2 μL) and water (up to 19 μL). The nucleoside digestion mix enzyme (1 μL) was added to the ON's solution. The reaction was incubated at 37°C for 2 h. Finally, the crude reaction mixture was analyzed by LC-MS (buffer A: 2 mM HCOONH₄ pH 5.5 in H₂O and buffer B: 2 mM HCOONH₄ pH 5.5 in 20:80 H₂O/MeCN; Gradient: 0-20% of B in 25 min; Flow rate = 0.15 mL·min⁻¹ and Injection volume = 20 μL).

6. Prebiotic synthesis of amino acid-modified carbamoyl oligonucleotides

6.1 Prebiotic synthesis of amino acid-modified carbamoyl oligonucleotides containing the four canonical bases



Scheme S7. Prebiotic synthesis of amino acid-modified carbamoyl oligonucleotides from those containing: a) *N*⁶-methylurea adenosine $\text{ON2}'$; b) *N*⁶-methylurea guanosine ON3 and c) *N*⁶-methylurea cytidine ON4 .

General procedure for the nitrosation reaction with oligonucleotides

Step 1: The oligonucleotide ON (3 nmol, 1.0 equiv.) and NaNO_2 (1.5 μmol , 500 equiv.) were dissolved in 5% aqueous H_3PO_4 solution (60 μL) containing 100 mM guanidinium chloride (GdmCl). The reaction was stirred at 0°C for 30 min. Step 2: After that, the solution was kept in the freezer at -20°C for 22 h. Step 3: H-aa-OH (3 μmol , 1000 equiv.) in 30 mM borate-buffered solution (140 μL) was added to the thawed oligonucleotide's solution. The pH was adjusted to ca. 9.5 using a 4 M aqueous NaOH solution (ca. 19 μL). The reaction was stirred at r.t. for 1 h. Finally, the reaction was quenched with 1 M aqueous HCl solution (5 μL). An aliquot (75 μL) of the crude reaction mixture was taken, diluted with water (up to 100 μL) and analyzed by HPLC (buffer A: 0.1 M AcOH/ Et_3N pH 7 in H_2O and buffer B: 0.1 M AcOH/ Et_3N pH 7 in 20:80 $\text{H}_2\text{O}/\text{MeCN}$; Gradient: 0-30% of B in 45 min; Flow rate = 1 $\text{mL}\cdot\text{min}^{-1}$; Injection volume = 100 μL).

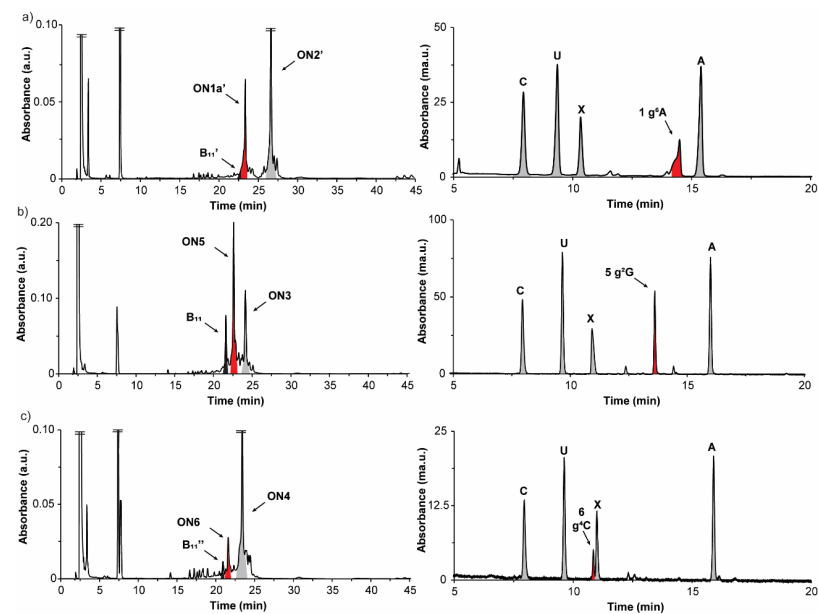
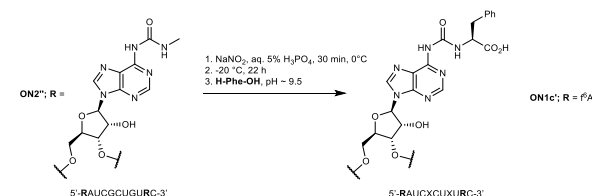


Figure S12. HPLC chromatograms of: left) the crude reaction mixtures for the prebiotic synthesis of a) $\text{ON1a}'$; b) ON5 and c) ON6 , and right) the enzymatic digestions of the corresponding products. B_{11} , B_{11}' and B_{11}'' are the oligonucleotides that do not contain the *N*-methylcarbamoyl substituent at the terminal nucleotide and have xanthosine instead of guanosine.

Table S7. HPLC retention times (0-30% of B in 45 min) and reaction yields of $\text{ON1a}'$, ON5 and ON6 .

Strand	t_r (min)	Yield%
$\text{ON1a}'$; g ³ A	23.4	18
ON5 ; g ² G	22.6	35
ON6 ; g ² C	21.5	5

6.2 Prebiotic synthesis of an oligonucleotide containing two amino acid-modified carbamoyl nucleotides at the terminal and internal positions



Scheme S8. Prebiotic synthesis of $\text{ON1c}'$ from $\text{ON2}''$ containing two *N*⁶-methylurea adenosine nucleotides. In this case, guanidinium chloride (GdmCl) was not added.

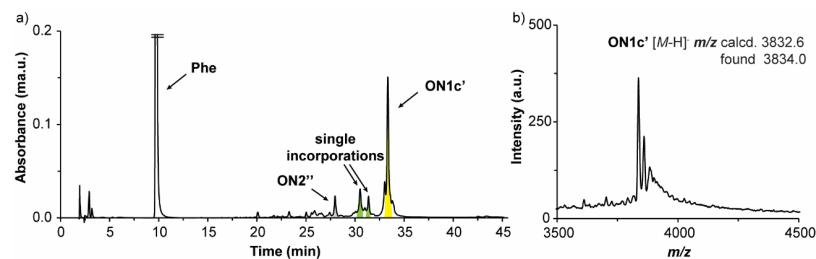
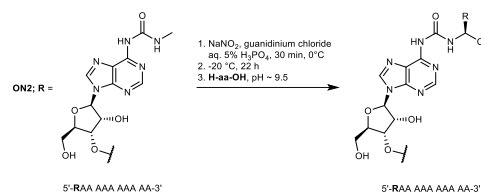


Figure S13. a) HPLC chromatogram of the crude reaction mixture for the prebiotic synthesis of 1c'. b) MALDI-TOF mass spectrum of the isolated product.

6.3 Prebiotic synthesis of a series of amino acid-modified *N*⁶-carbamoyl adenosine oligonucleotides



ON1a; R = *g*⁶A, 18%
 ON1b; R = *t*⁶A, 18%
 ON1c; R = *f*⁶A, 23%
 ON1d; R = *p*⁶A, 15%
 ON1e; R = *d*⁶A, 10%
 α-ON1f; R = *k*⁶A, 10%
 ON1g; R = *a*⁶A, 16%

Scheme S9. Prebiotic synthesis of amino acid-modified *N*⁶-carbamoyl adenosine oligonucleotides ON1a-g from *N*⁶-methylurea adenosine oligonucleotide ON2 (optimized reaction conditions are shown).

Effect of the addition of salts in the prebiotic synthesis of ON1g:

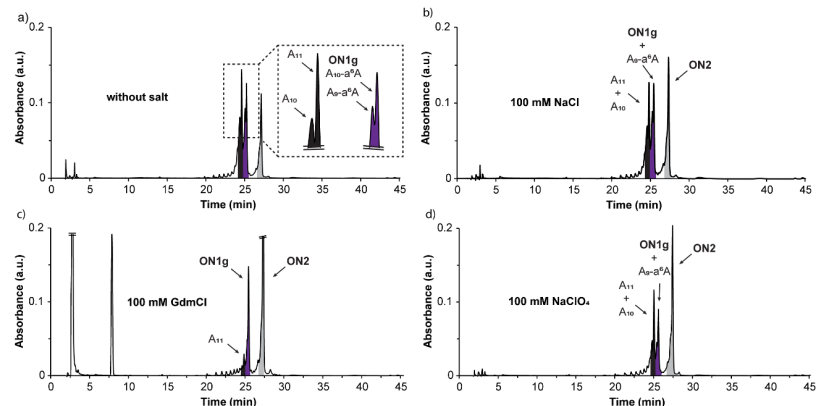


Figure S14. HPLC chromatograms of the crude reaction mixtures for the prebiotic synthesis of ON1g: a) without salt and with b) 100 mM NaCl; c) 100 mM guanidinium chloride (GdmCl) and d) 100 mM NaClO₄.

Table S8. Results obtained for the addition of salts in the prebiotic synthesis of ON1g.

Salt	A ₁₁ (%)	ON1g (%)	ON2 (%)	Degradation (%)
-	18	15	15	-52
100 mM NaCl	19	15	28	-38
100 mM GdmCl	4	15	53	-28
100 mM NaClO ₄	15	11	25	-49

Prebiotic synthesis of ON1a-g under the nitrosation conditions using guanidinium chloride:

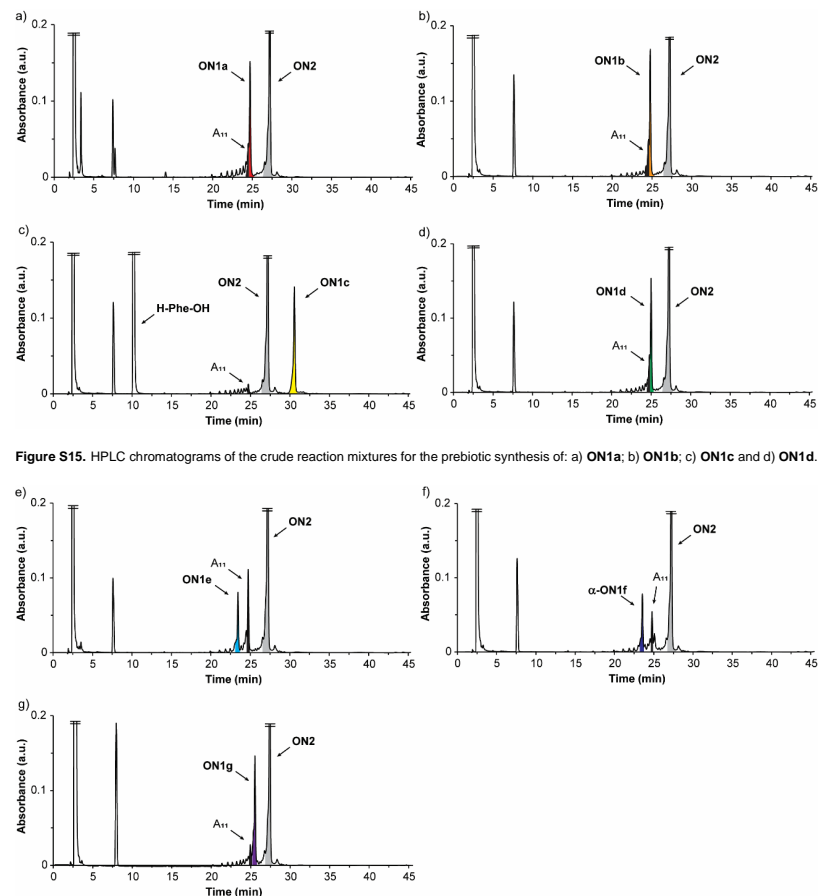


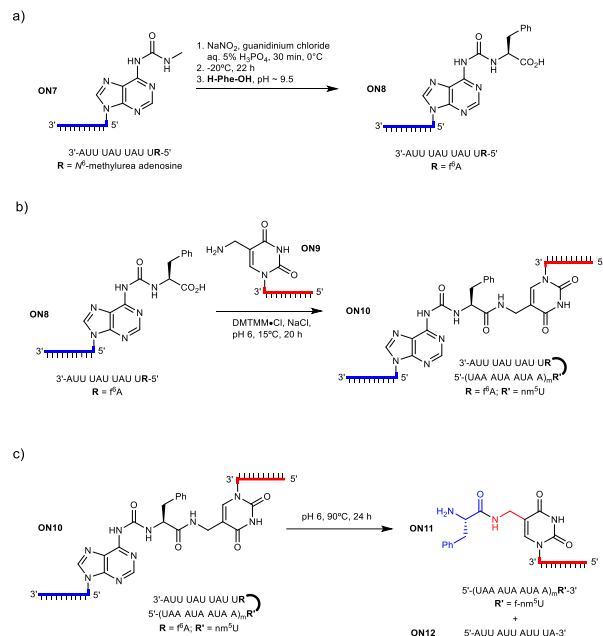
Figure S15. HPLC chromatograms of the crude reaction mixtures for the prebiotic synthesis of: a) ON1a; b) ON1b; c) ON1c and d) ON1d.

Figure S16. HPLC chromatograms of the crude reaction mixtures for the prebiotic synthesis of: e) ON1e; f) α-ON1f and g) ON1g.

Table S9. HPLC retention times (0-30% of B in 45 min), MALDI-TOF mass spectrometric analysis (negative mode) and reaction yields of ON1a-g.

Strand	t _R (min)	m/z calcd. for [M-H] ⁻	found	Yield%
ON1a; R = <i>g</i> ⁶ A	24.7	3657.6	3660.3	18
ON1b; R = <i>t</i> ⁶ A	24.8	3701.7	3702.7	18
ON1c; R = <i>f</i> ⁶ A	30.6	3747.7	3747.0	23
ON1d; R = <i>p</i> ⁶ A	25.0	3697.7	3699.7	15
ON1e; R = <i>d</i> ⁶ A	23.4	3715.6	3717.2	10
α-ON1f; R = <i>k</i> ⁶ A	23.6	3728.7	3730.3	10
ON1g; R = <i>a</i> ⁶ A	25.4	3671.6	3673.6	16

7. Consecutive reactions using *N*⁶-methylurea adenosine oligonucleotide



Scheme S10. Consecutive reactions: a) prebiotic synthesis of amino acid-modified carbamoyl oligonucleotide **ON8**; b) peptide coupling reaction between **ON8** and **ON9** and c) urea cleavage reaction of **ON10**.

Procedures for the stepwise reactions:

a) Prebiotic synthesis of amino acid-modified carbamoyl oligonucleotide **ON8**:

Step 1: The oligonucleotide **ON7** (20 nmol, 1.0 equiv.) and NaNO₂ (10 μmol, 500 equiv.) were dissolved in 5% aqueous H₃PO₄ solution (200 μL) containing 100 mM guanidinium chloride (GdmCl). The reaction was stirred at 0°C for 30 min. Step 2: After that, the solution was kept in the freezer at -20°C for 22 h. Step 3: **H-Phe-OH** (20 μmol, 1000 equiv.) in 30 mM borate-buffered solution (150 μL) was added to the thawed oligonucleotide's solution. The pH was adjusted to ca. 9.5 using a 4 M aqueous NaOH solution (ca. 50 μL). The reaction was stirred at r.t. for 1 h. Finally, the reaction was quenched with 1 M aqueous HCl solution (50 μL). An aliquot (22.5 μL) of the crude reaction mixture was taken, diluted with water (up to 100 μL) and analyzed by HPLC (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-30% of B in 45 min; Flow rate = 1 mL·min⁻¹; Injection volume = 100 μL) and MALDI-TOF mass spectrometry. The remaining crude was purified by semi-preparative HPLC. The purified **ON** was lyophilized, desalted, and redissolved in water. Subsequently, the concentration of **ON8** was determined.

b) Peptide coupling reaction between **ON8** and **ON9**:

An equimolar solution mixture of **ON8** (50 μM) and **ON9** (50 μM) containing MES buffer pH 6 (100 mM), NaCl (1 M) and DMTMM·Cl (50 mM) was incubated at 15°C for 20 h. After that, an aliquot (20 μL) of the crude reaction mixture was taken and analyzed by HPLC (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-30% of B in 45 min; Flow rate = 1 mL·min⁻¹; Injection volume = 20 μL) and MALDI-TOF mass spectrometry. The remaining crude was purified by semi-preparative HPLC. The purified **ON** was lyophilized, desalted, and redissolved in water. Subsequently, the concentration of **ON10** was determined.

c) Urea cleavage reaction of **ON10**:

The oligonucleotide **ON10** (20 μM) was diluted with MES buffer pH 6 (100 mM) containing NaCl (100 mM). The reaction mixture was heated at 90°C for 24 h. After that, an aliquot (100 μL) of the crude reaction mixture was taken and analyzed by HPLC (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-30% of B in 45 min; Flow rate = 1 mL·min⁻¹; Injection volume = 20 μL) and MALDI-TOF mass spectrometry.

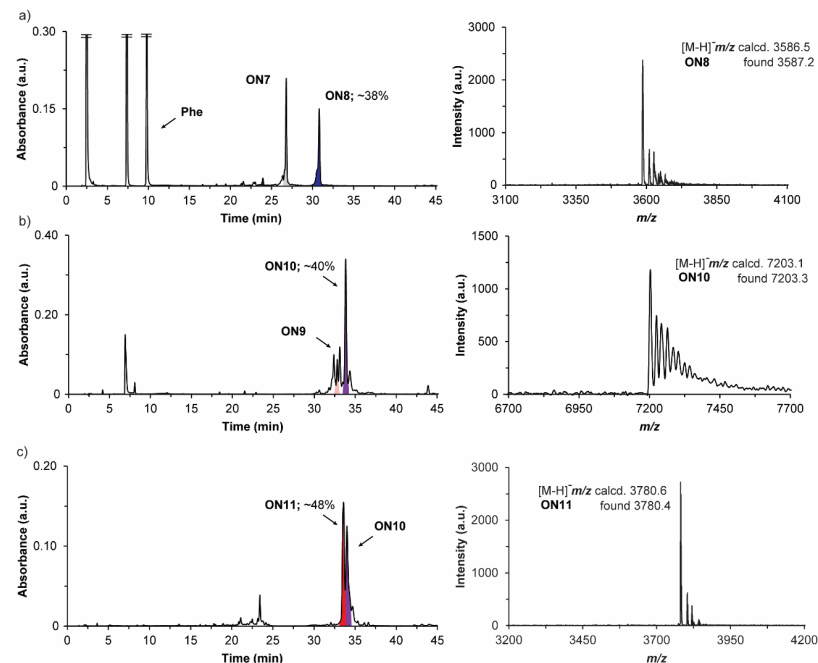


Figure S17. left) HPLC chromatograms of the crude reaction mixtures and right) MALDI-TOF mass spectra of the isolated oligonucleotides in the stepwise reactions: a) prebiotic synthesis of amino acid-modified carbamoyl oligonucleotide **ON8**; b) peptide coupling reaction between **ON8** and **ON9** and c) urea cleavage reaction of **ON10**. Yields of the oligonucleotide products were estimated using the areas of the chromatographic peaks in relation to that determined for reference compounds.

Procedure for the consecutive reactions (a → c):

The reactions a, b and c were performed using the conditions indicated above, but without performing a chromatographic purification after each reaction step. Under these conditions, only a filtration, using an Amicon® ultra centrifugal filter (0.5 mL, 3 kDa), was required after reactions a and b to remove the excess of salts and condensation reagents. For this experiment, 20 nmol of **ON7** were used as starting material and 10 nmol of **ON9** were added for reaction b.

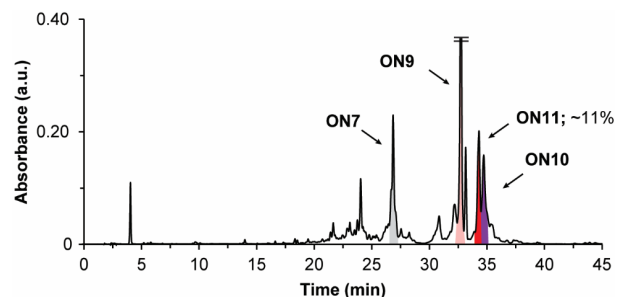


Figure S18: HPLC chromatogram of the crude reaction mixture after the consecutive reactions (a → c). Yield of the oligonucleotide product was estimated using the area of the chromatographic peak in relation to that determined for a reference compound.

8. Calibration curve of *N*⁶-methylurea adenosine oligonucleotide

*N*⁶-methylurea adenosine oligonucleotide **ON2** was used for the development of a HPLC calibration curve. A stock solution of **ON2** was prepared in water (100 μM). Separate standard solutions containing 1.2; 1.0; 0.8; 0.6; 0.4; 0.2 and 0.1 nmol of **ON2** were prepared in a final volume of 20 μL. The standard solutions were injected in an analytical HPLC equipped with a C18 column (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-30% of B in 45 min; Flow rate = 1 mL·min⁻¹). The absorbance was monitored at 260 nm and the areas of the chromatographic peaks were determined by integration of the HPLC chromatograms. The plot of the chromatographic area (a.u.) versus the amount (nmol) of the oligonucleotide followed a linear relationship.

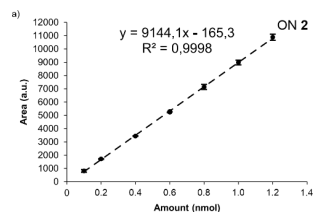


Figure S19. Chromatographic area (a.u.) vs. amount (nmol) of **ON2**. Line shows the fit of the data to a linear regression equation. Error bars are standard deviations from three independent experiments.

Table S10. Calibration curve ($y = mx + n$) obtained by analysis of the chromatographic peaks of **ON2**.

Strand	Slope, m (nmol ⁻¹)	Intercept, n	r ²
ON2	9144.1	-165.3	0.99

9. Melting curve of double strand

The UV melting curves were measured on a JASCO V-650 spectrometer at 260 nm using 10 mm QS cuvettes with a scanning rate of 1°C·min⁻¹. The obtained UV spectroscopic data were fit to a two-state melting model, *i.e.* double strand – random coil equilibrium, using a mono-sigmoidal Boltzmann function.⁵ The fit of the data returned the melting temperature.

For the experiments, we prepared aqueous solutions containing equimolar amounts of the oligonucleotides (2 μM), MES buffer pH 6 (100 mM) and NaCl (100 mM or 1 M). The oligonucleotides were annealed by heating to 95°C for 4 min and, subsequently, by cooling down slowly to 5°C before the variable-temperature UV spectroscopic experiments.

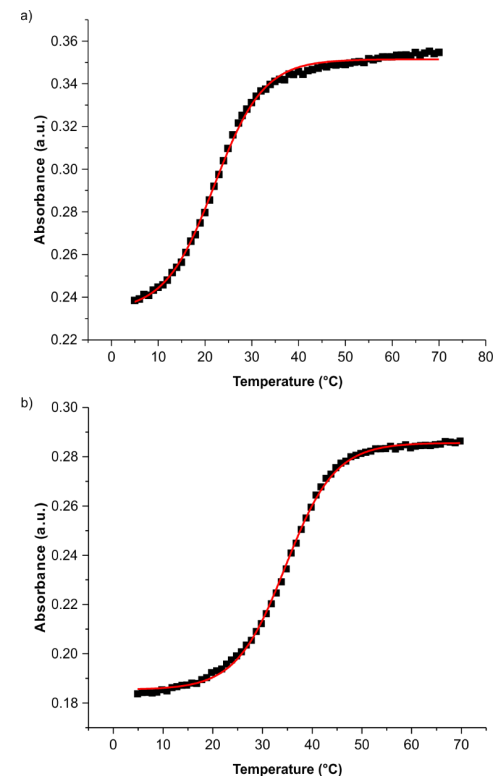
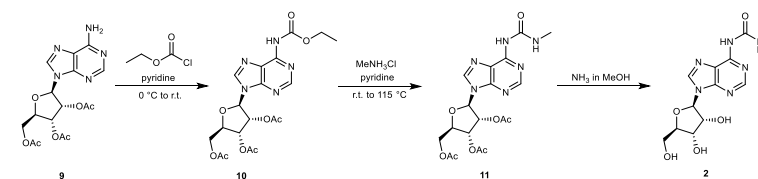


Figure S20. Melting curves of the RNA donor strand **ON8** with the RNA acceptor strand **ON9** containing: a) 100 mM NaCl and b) 1 M NaCl. The fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function returned the melting temperatures for the double strand: a) $T_m = 22.0^\circ\text{C}$ and b) $T_m = 34.6^\circ\text{C}$.

10. Synthesis of methylurea and amino acid-modified carbamoyl nucleosides used as reference

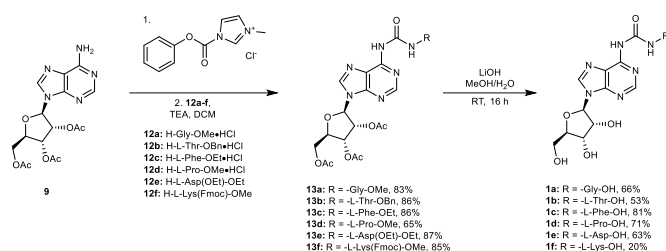
10.1 Synthesis of *N*⁶-methylurea adenosine



Scheme S11. Synthesis of *N*⁶-methylurea adenosine **2**.

Compounds **2**, **10** and **11** were synthesized following procedures previously reported in the literature.^{6,7}

10.2 Synthesis of amino acid-modified *N*⁶-carbamoyl adenosine nucleosides



Scheme S12. Synthesis of amino acid-modified *N*⁶-carbamoyl adenosine nucleosides **1a-f** used as reference.

General procedure for the synthesis of **13a-f**

Step 1: Acetyl-protected adenosine **9** (1.0 equiv.) and 1-*N*-methyl-3-phenoxycarbonyl-imidazolium chloride (2.0 equiv.) were dissolved in dry DCM. The reaction was stirred at r.t. for 14 h. Step 2: A suspension, in dry DCM, of the protected amino acid **12a-f** (2.0 equiv.), containing TEA (2 equiv.) in the case of **12a-d** (hydrochloride salts), was transferred to the adenosine's mixture. The reaction was stirred at r.t. for 16 h. After that, the crude was washed with sat. aq. NaHCO₃ solution and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (Na₂SO₄ or MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (*i*-Hex/EtOAc or DCM/IPA) affording **13a-f** as a white solid.

13a: Yield = 83%. *R*_f = 0.24 (EtOAc). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3240; 1741; 1698; 1611; 1589; 1538; 1532; 1470; 1366; 1206; 1089; 1043; 1019; 900; 798. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 9.95 (t, *J* = 5.5 Hz, 1H); 8.59 (s, 1H); 8.34 (s, 1H); 8.22 (s, 1H); 6.21 (d, *J* = 5.5 Hz, 1H); 5.96 (t, *J* = 5.5 Hz, 1H); 5.66 (dd, *J* = 5.5, 4.2 Hz, 1H); 4.49-4.35 (m, 3H); 4.24 (d, *J* = 5.5 Hz, 2H); 3.80 (s, 3H); 2.16 (s, 3H); 2.13 (s, 3H); 2.08 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 170.6; 170.5; 169.7; 169.5; 154.0; 151.7; 150.4; 150.4; 141.5; 121.1; 86.6; 80.6; 73.2; 70.8; 63.2; 52.5; 42.2; 20.9; 20.7; 20.6. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₀H₂₅N₆O₁₀ 509.1627; Found 509.1624.

13b: Yield = 86%. *R*_f = 0.50 (EtOAc). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3237; 1742; 1694; 1614; 1589; 1551; 1537; 1468; 1366; 1212; 1089; 1044; 1018; 923; 903; 798; 752; 735; 687. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.23 (d, *J* = 8.5 Hz, 1H); 8.70 (s, 1H); 8.54 (s, 1H); 8.28 (s, 1H); 7.40-7.33 (m, 5H); 6.22 (d, *J* = 5.5 Hz, 1H); 5.95 (dd, *J* = 5.5, 5.5 Hz, 1H); 5.64 (dd, *J* = 5.5, 4.2 Hz, 1H); 5.24 (d, *J* = 1.9 Hz, 2H); 4.74 (dd, *J* = 8.5, 2.9 Hz, 1H); 4.50-4.32 (m, 4H); 2.16 (s, 3H); 2.12 (s, 3H); 2.07 (s, 3H); 1.31 (d, *J* = 6.4 Hz, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 171.0; 170.6; 169.8; 169.5; 154.4; 151.8; 150.5; 150.4; 141.6; 135.5; 128.8; 128.6; 128.3; 121.0; 86.4; 80.6; 73.2; 70.8; 68.5; 67.4; 63.3; 59.0; 20.9; 20.7; 20.5; 20.1. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₈H₃₃N₆O₁₁ 629.2202; Found 629.2200.

13c: Yield = 86%. *R*_f = 0.45 (95:5 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3231; 1742; 1698; 1614; 1588; 1524; 1468; 1368; 1214; 1131; 1084; 1044; 921; 903; 798; 740; 700. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 9.99 (d, *J* = 7.7 Hz, 1H); 8.52 (s, 1H); 8.42 (s, 1H); 8.26 (s, 1H); 7.34-7.20 (m, 5H); 6.19 (d, *J* = 5.3 Hz, 1H); 5.98 (dd, *J* = 5.3, 5.3 Hz, 1H); 5.66 (dd, *J* = 5.3, 4.4 Hz, 1H); 4.91 (dt, *J* = 7.7, 6.1 Hz, 1H); 4.48-4.34 (m, 3H); 4.20 (q, *J* = 7.1 Hz, 2H); 3.24 (d, *J* = 6.1 Hz, 2H); 2.15 (s, 3H); 2.12 (s, 3H); 2.07 (s, 3H); 1.25 (t, *J* = 7.1 Hz, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 171.7; 170.5; 169.7; 169.5; 153.5; 151.4; 150.4; 150.3; 141.8; 136.4; 129.7; 128.6; 127.2; 121.0; 86.6; 80.5; 73.1; 70.7; 63.2; 61.5; 54.8; 38.3; 20.9; 20.7; 20.5; 14.3. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₈H₃₃N₆O₁₀ 613.2253; Found 613.2253.

13d: Yield = 65%. *R*_f = 0.13 (96:4 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 2956; 1740; 1690; 1647; 1607; 1403; 1367; 1213; 1043; 898; 727; 643. For major rotamer: ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 9.77 (s, 1H); 9.58 (s, 2H); 6.30 (d, *J* = 5.6 Hz, 1H); 6.07 (dd, *J* = 5.6, 5.6 Hz, 1H); 5.64 (dd, *J* = 5.6, 5.6 Hz, 1H); 4.43-4.38 (m, 3H); 4.27-4.24 (m, 1H); 3.72 (br s, 1H); 3.62 (s, 4H); 3.62 (s, 4H); 2.28-2.21 (m, 1H); 2.13 (s, 3H); 2.04 (s, 3H); 1.92 (br s, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 172.7; 170.1; 169.5; 169.4; 152.2; 151.9; 151.8; 151.1; 142.5; 123.8; 85.6; 79.6; 71.8; 70.1; 62.8; 58.9; 51.9; 46.9; 29.2; 24.5; 20.6; 20.4; 20.2. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₃H₂₉N₆O₁₀ 549.1939; Found 549.1942.

13e: Yield = 87%. *R*_f = 0.50 (97:3 DCM/IPA). ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.40 (d, *J* = 7.8 Hz, 1H); 8.95 (s, 1H); 8.58 (s, 1H); 8.40 (s, 1H); 6.22 (d, *J* = 5.5 Hz, 1H); 5.97 (dd, *J* = 5.5, 5.5 Hz, 1H); 5.65 (dd, *J* = 5.5, 4.2 Hz, 1H); 4.93 (dt, *J* = 7.8, 4.8 Hz, 1H); 4.44 (dq, *J* = 8.0, 3.5 Hz, 2H); 4.36 (dd, *J* = 12.8, 5.4 Hz, 1H);

4.25 (q, *J* = 7.1 Hz, 2H); 4.19-4.12 (m, 2H); 3.18-2.94 (m, 2H); 2.14 (s, 3H); 2.10 (s, 3H); 2.06 (s, 3H); 1.31-1.22 (m, 6H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 170.9; 170.7; 170.5; 169.7; 169.4; 153.8; 151.6; 150.5; 150.4; 142.1; 121.0; 86.5; 80.6; 73.1; 70.8; 63.2; 61.9; 61.1; 50.0; 37.1; 20.9; 20.7; 20.5; 14.3; 14.2. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₅H₃₃N₆O₁₂ 609.2151; Found 609.2143.

13f: Yield = 85%. *R*_f = 0.37 (95:5 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 2949; 1743; 1697; 1612; 1532; 1365; 1214; 1043; 741. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.06 (d, *J* = 7.8 Hz, 1H); 8.64 (s, 1H); 8.60 (s, 1H); 8.30 (s, 1H); 7.74 (d, *J* = 7.6 Hz, 2H); 7.57 (d, *J* = 7.4 Hz, 2H); 7.37 (t, *J* = 7.4 Hz, 2H); 7.28 (td, *J* = 7.5, 0.8 Hz, 2H); 6.22 (d, *J* = 5.5 Hz, 1H); 5.96 (t, *J* = 5.5 Hz, 1H); 5.64 (t, *J* = 4.6 Hz, 1H); 4.96 (t, *J* = 6.0 Hz, 1H); 4.70 (td, *J* = 7.7, 5.3 Hz, 1H); 4.47-4.36 (m, 5H); 4.20 (t, *J* = 7.0 Hz, 1H); 3.77 (s, 3H); 3.21 (hept, *J* = 6.8 Hz, 2H); 2.15 (s, 3H); 2.11 (s, 3H); 2.06 (s, 3H); 1.21 (s, 2H); 1.20 (s, 2H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 173.0; 170.5; 169.7; 169.5; 156.5; 153.7; 151.6; 150.5; 150.4; 144.1; 141.7; 141.4; 127.7; 127.1; 125.2; 121.1; 120.0; 86.5; 80.6; 73.1; 70.8; 66.6; 63.2; 53.2; 52.6; 47.4; 40.9; 32.4; 29.5; 22.8; 20.9; 20.7; 20.5. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₀H₄₄N₇O₁₂ 802.3042; Found 802.3046.

General procedure for the synthesis of **1a-f**

Acetyl-protected adenosine derivative **13a-f** (1.0 equiv.) was dissolved in 3:1 MeOH/H₂O and the solution was cooled to 0°C. LiOH·H₂O (10.0 equiv.) was added to the adenosine's solution and the reaction was stirred at r.t. for 2 h. After that, the solution was neutralized with 1 M aqueous HCl solution. The organic solvent was removed under reduced pressure and the remaining aqueous solution mixture was lyophilized. The crude was purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **1a-f** as a white solid. Note that for **1f**, the Fmoc protecting group was removed with 20% piperidine in DMF before the hydrolysis reaction.

1a: Yield = 66%. *t*_r = 14.0 min (0-30% of B in 45 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3499; 3377; 1720; 1656; 1610; 1557; 1254; 1227; 1100; 1056; 867; 803; 796; 745; 716. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.70 (br s, 1H); 9.90 (s, 1H); 9.65 (t, *J* = 5.6 Hz, 1H); 8.67 (s, 1H); 8.56 (s, 1H); 5.99 (d, *J* = 5.6 Hz, 1H); 5.55 (br s, 1H); 5.25 (br s, 1H); 5.15 (br s, 1H); 4.59 (t, *J* = 5.3 Hz, 1H); 4.17 (dd, *J* = 5.0, 3.6 Hz, 1H); 4.01-3.95 (m, 3H); 3.71-3.55 (m, 2H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 171.5; 153.7; 150.8; 150.4; 150.2; 142.3; 120.4; 87.7; 85.7; 73.8; 70.3; 61.3; 41.7. HRMS (ESI) *m/z*: [M-H]⁻ Calcd. for C₁₈H₁₅N₆O₇ 367.1008; Found 367.1009.

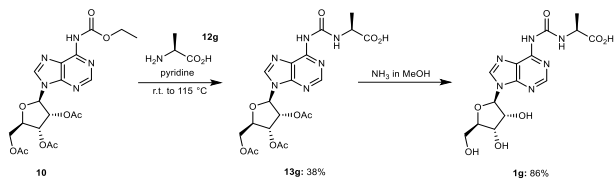
1b: Yield = 53%. *t*_r = 22.2 min (0-30% of B in 45 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3231; 1684; 1617; 1591; 1553; 1469; 1403; 1364; 1294; 1256; 1234; 1218; 1086; 1055; 896; 868; 848; 797; 753; 718; 695. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.62 (br s, 1H); 9.83 (s, 1H); 9.73 (d, *J* = 8.3 Hz, 1H); 8.68 (s, 1H); 8.55 (s, 1H); 5.99 (d, *J* = 5.6 Hz, 1H); 5.55 (d, *J* = 5.9 Hz, 1H); 5.29-5.17 (m, 1H); 5.14 (t, *J* = 5.6 Hz, 1H); 4.60 (q, *J* = 5.0 Hz, 1H); 4.30-4.22 (m, 2H); 4.18 (q, *J* = 4.1, 3.5 Hz, 1H); 3.97 (q, *J* = 3.9 Hz, 1H); 3.72-3.54 (m, 2H); 1.14 (d, *J* = 6.3 Hz, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 172.5; 153.8; 150.9; 150.4; 150.3; 142.3; 120.5; 87.7; 85.7; 73.8; 70.3; 66.2; 61.3; 58.7; 20.9. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₈H₂₁N₆O₈ 413.1415; Found 413.1415.

1c: Yield = 81%. *t*_r = 23.0 min (0-65% of B in 45 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3228; 1714; 1652; 1613; 1590; 1546; 1339; 1246; 1229; 1125; 1082; 986; 896; 868; 822; 797; 741; 699. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.92 (br s, 1H); 9.90 (s, 1H); 9.76 (d, *J* = 7.3 Hz, 1H); 8.66 (s, 1H); 8.46 (s, 1H); 7.36-7.20 (m, 5H); 5.98 (d, *J* = 5.6 Hz, 1H); 5.54 (s, 1H); 5.25 (s, 1H); 5.14 (s, 1H); 4.67-4.50 (m, 2H); 4.21-4.12 (m, 1H); 3.97 (q, *J* = 3.9 Hz, 1H); 3.74-3.51 (m, 2H); 3.21-3.04 (m, 2H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 172.9; 153.2; 150.6; 150.4; 150.1; 142.3; 137.0; 129.5; 128.4; 126.8; 120.3; 87.7; 85.7; 73.8; 70.3; 61.3; 54.4; 37.1. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₀H₂₃N₆O₇ 459.1623; Found 459.1624.

1d: Yield = 71%. *t*_r = 21.8 min (0-20% of B in 30 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3272; 2935; 1654; 1609; 1459; 1401; 1354; 1220; 1078; 1054; 892; 863; 640. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.55 (s, 1H); 9.64 (s, 1H); 8.58 (s, 2H); 5.97 (d, *J* = 5.3 Hz, 1H); 5.54 (d, *J* = 6.1 Hz, 1H); 5.24 (d, *J* = 6.1 Hz, 1H); 5.18 (s, 1H); 4.61 (s, 1H); 4.29 (s, 1H); 4.17-4.15 (m, 1H); 3.97-3.95 (m, 1H); 3.70-3.66 (m, 2H); 3.58-3.54 (m, 2H); 2.25-2.18 (s, 1H); 1.90 (s, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 173.8; 152.3; 151.6; 151.3; 142.0; 123.5; 87.6; 85.7; 73.6; 70.4; 61.4; 59.0; 46.9; 29.3; 24.4 (one carbon signal appeared too broad for an unequivocal assignment). HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₆H₁₈N₆O₇ 409.1466; Found 409.1465.

1e: Yield = 63%. *t*_r = 24.7 min (0-20% of B in 45 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3207; 2937; 1682; 1614; 1592; 1538; 1471; 1399; 1334; 1293; 1219; 1122; 1080; 1053; 985; 895; 866; 797; 704. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.79 (br s, 2H); 9.96 (d, *J* = 7.9 Hz, 1H); 9.92 (s, 1H); 8.67 (s, 1H); 8.51 (s, 1H); 5.99 (d, *J* = 5.6 Hz, 1H); 5.55 (d, *J* = 5.9 Hz, 1H); 5.24 (d, *J* = 5.0 Hz, 1H); 5.14 (s, 1H); 4.66-4.56 (m, 2H); 4.22-4.11 (m, 1H); 3.97 (q, *J* = 3.9 Hz, 1H); 3.75-3.53 (m, 2H); 2.93-2.73 (m, 2H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 172.4; 172.1; 153.2; 150.7; 150.5; 150.2; 142.3; 120.4; 87.7; 85.7; 73.8; 70.3; 61.3; 49.4; 36.6. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₅H₁₉N₆O₉ 427.1208; Found 427.1207.

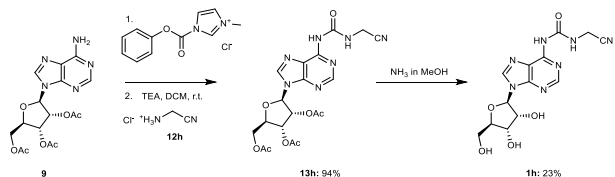
1f: Yield = 20%. t_R = 18.3 min (0-20% of B in 30 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3123; 2934; 1683; 1588; 1526; 1469; 1398; 1252; 1053; 796; 641. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 9.62 (d, *J* = 6.1 Hz, 1H); 8.65 (s, 1H); 8.51 (s, 1H); 5.97 (d, *J* = 5.5 Hz, 1H); 4.59 (t, *J* = 5.4 Hz, 1H); 4.17 (t, *J* = 4.2 Hz, 1H); 4.00-3.95 (m, 2H); 3.68 (dd, *J* = 12.0, 4.0 Hz, 1H); 3.56 (dd, *J* = 12.0, 4.0 Hz, 1H); 2.74 (t, *J* = 6.7 Hz, 2H); 1.81-1.67 (m, 2H); 1.56 (p, *J* = 7.6 Hz, 2H); 1.43-1.26 (m, 2H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 174.0; 152.2; 151.0; 150.4; 150.1; 142.1; 120.3; 87.7; 73.8; 70.3; 61.3; 55.1; 38.8; 32.3; 27.2; 21.9. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₇H₂₆N₇O₇ 440.1888; Found 440.1888.



Scheme S13. Synthesis of alanine-modified *N*⁶-carbamoyl adenosine nucleoside **1g** used as reference.

Acetyl-protected *N*⁶-alanine modified adenosine 13g: Carbamate **10** (501 mg, 1.08 mmol, 1.0 equiv.) was dissolved in dry pyridine. Subsequently, H-L-Ala-OH **12g** (192 mg, 2.15 mmol, 2.0 equiv.) was added to the carbamate's solution at r.t. The reaction was heated under reflux conditions for 16 h. After that, the reaction mixture was cooled to r.t. The suspension was filtered, and the residue was washed with EtOAc. The filtrate was concentrated under reduced pressure and co-evaporated with toluene. The crude was purified by recrystallisation from EtOH affording **13g** as a white solid (210 mg, 0.41 mmol, 38% yield). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3270; 2456; 1744; 1700; 1652; 1607; 1592; 1543; 1485; 1295; 1246; 1236; 1221; 1112; 1102; 1044; 1025; 994; 916; 802; 755. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.83 (br s, 1H); 9.95 (s, 1H); 9.69 (d, *J* = 7.0 Hz, 1H); 8.65 (s, 1H); 8.60 (s, 1H); 6.30 (d, *J* = 5.4 Hz, 1H); 6.03 (dd, *J* = 5.4, 5.4 Hz, 1H); 5.62 (dd, *J* = 5.4, 4.5 Hz, 1H); 4.46-4.23 (m, 4H); 2.12 (s, 3H); 2.03 (s, 3H); 2.02 (s, 3H); 1.42 (d, *J* = 7.2 Hz, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 174.2; 170.1; 169.5; 169.3; 152.8; 151.2; 150.4; 150.2; 142.7; 120.5; 85.7; 79.7; 72.0; 70.0; 62.8; 48.6; 20.6; 20.4; 20.2; 18.1. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₀H₂₅N₆O₁₀ 509.1627; Found 509.1627.

***N*⁶-alanine modified adenosine 1g**: Protected *N*⁶-alanine modified adenosine **13g** (200 mg, 0.39 mmol, 1.0 equiv.) was dissolved in 7 N NH₃ in MeOH. The reaction was stirred at r.t. overnight. After that, the crude was concentrated and purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **1g** as a white solid (130 mg, 0.34 mmol, 86% yield). t_R = 20.7 min (0-40% of B in 45 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3234; 2926; 1684; 1612; 1603; 1540; 1476; 1343; 1313; 1267; 1250; 1226; 1122; 1080; 861; 835; 782; 764; 746; 680. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.81 (br s, 1H); 9.87 (s, 1H); 9.74 (d, *J* = 7.0 Hz, 1H); 8.67 (s, 1H); 8.58 (s, 1H); 5.99 (d, *J* = 5.6 Hz, 1H); 5.55 (d, *J* = 5.6 Hz, 1H); 5.25 (s, 1H); 5.14 (t, *J* = 5.6 Hz, 1H); 4.59 (q, *J* = 4.6 Hz, 1H); 4.33 (p, *J* = 7.1 Hz, 1H); 4.17 (t, *J* = 4.5 Hz, 1H); 3.97 (q, *J* = 3.9 Hz, 1H); 3.70-3.58 (m, 2H); 1.42 (d, *J* = 7.2 Hz, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 174.2; 152.9; 150.9; 150.4; 150.2; 142.2, 120.4; 87.6; 85.7; 73.8; 70.3; 61.3; 48.6; 18.1. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₄H₁₉N₆O₇ 383.1310; Found 383.1311.



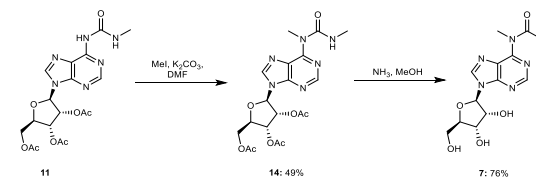
Scheme S14. Synthesis of amino nitrile-modified *N*⁶-carbamoyl adenosine nucleoside **1h** used as reference.

Acetyl-protected *N*⁶-amino nitrile modified adenosine 13h: Step 1: Acetyl-protected adenosine **9** (500 mg, 1.27 mmol, 1.0 equiv.) and 1-*N*-methyl-3-phenoxycarbonyl-imidazolium chloride (**607** mg, 2.54 mmol, 2.0 equiv.) were dissolved in dry DCM. The reaction was stirred at r.t. for 14 h. Step 2: A suspension, in dry DCM, of the amino nitrile hydrochloride **12h** (235 mg, 2.54 mmol, 2.0 equiv.), containing TEA (2 equiv.), was transferred to the adenosine's mixture. The reaction was stirred at r.t. for 16 h. After that, the reaction was quenched with sat. aq. NaHCO₃ solution and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column

chromatography on silica gel (95:5 DCM/IPA) affording **13h** as a white solid (570 mg, 1.20 mmol, 94% yield). R_f = 0.50 (95:5 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 1742; 1701; 1612; 1589; 1517; 1470; 1365; 1211; 1043; 904; 798; 642; 579; 507; 407. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.19 (t, *J* = 5.7 Hz, 1H); 9.17 (s, 1H); 8.58 (s, 1H); 8.48 (s, 1H); 6.23 (d, *J* = 5.2 Hz, 1H); 5.99 (dd, *J* = 5.2, 5.2 Hz, 1H); 5.74-5.64 (m, 1H); 4.51-4.45 (m, 2H); 4.42-4.35 (m, 3H); 2.16 (s, 3H); 2.12 (s, 3H); 2.09 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 170.6; 169.8; 169.5; 154.1; 151.2; 150.6; 149.9; 142.8; 121.1; 116.4; 86.8; 80.6; 73.2; 70.8; 63.3; 28.4; 21.0; 20.7; 20.6. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₉H₂₂N₇O₈ 476.1524; Found 476.1524.

***N*⁶-amino nitrile modified adenosine 1h**: Protected *N*⁶-amino nitrile modified adenosine **13h** (100 mg, 0.21 mmol, 1.0 equiv.) was dissolved in 7 N NH₃ in MeOH. The reaction was stirred at r.t. for 3 h. After that, the crude was concentrated and purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN) affording **1h** as a white solid (17 mg, 0.05 mmol, 23% yield). t_R = 14.1 min (0-30% of B in 30 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 1614; 1589; 1469; 1401; 1297; 1253; 1079; 1049; 1016; 984; 900; 865; 796; 644; 578; 511; 457; 444; 433; 414; 407. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 10.31 (s, 1H); 9.81 (t, *J* = 5.7 Hz, 1H); 8.69 (s, 1H); 8.57 (s, 1H); 5.99 (d, *J* = 5.7 Hz, 1H); 5.55 (d, *J* = 6.0 Hz, 1H); 5.26 (d, *J* = 5.0 Hz, 1H); 5.14 (t, *J* = 5.6 Hz, 1H); 4.59 (dd, *J* = 5.6, 5.6 Hz, 1H); 4.34 (d, *J* = 5.7 Hz, 2H); 4.17 (dd, *J* = 4.8 Hz, 1H); 3.97 (dd, *J* = 3.8, 3.8 Hz, 1H); 3.69 (ddd, *J* = 12.0, 5.2, 4.0 Hz, 1H); 3.57 (ddd, *J* = 12.0, 6.1, 3.8 Hz, 1H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 153.7; 150.7; 150.6; 149.9; 142.4; 120.5; 118.1; 87.7; 85.7; 73.8; 70.3; 61.3; 28.4. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₃H₁₆N₇O₅ 350.1207; Found 350.1208.

10.3 Synthesis of *N*⁶-methyl *N*⁶-methyleurea adenosine

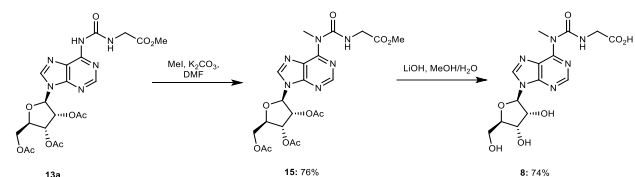


Scheme S15. Synthesis of *N*⁶-methyl *N*⁶-methyleurea adenosine **7**.

Acetyl-protected *N*⁶-methyl *N*⁶-methyleurea adenosine 14: Acetyl-protected *N*⁶-methyleurea adenosine **11** (300 mg, 0.67 mmol, 1.0 equiv.) was dissolved in dry DMF. The solution was cooled to 0°C and K₂CO₃ (276 mg, 2.00 mmol, 3.0 equiv.) was added. Finally, methyl iodide (82.9 μ L, 1.33 mmol, 2.0 equiv.) was added dropwise to the mixture. The reaction was stirred at r.t. for 16 h. After that, the crude reaction mixture was diluted with Et₂O and washed with sat. aq. NH₄Cl solution and water. The aqueous phase was re-extracted with Et₂O and the combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (98:2 DCM/IPA) affording **14** as a white foam (151 mg, 0.32 mmol, 49% yield). R_f = 0.55 (9:1 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3483; 3198; 2950; 1742; 1678; 1583; 1566; 1532; 1467; 1423; 1367; 1327; 1211; 1093; 1042; 1000; 920; 901; 795. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.31 (q, *J* = 4.9 Hz, 1H); 8.52 (s, 1H); 8.08 (s, 1H); 6.23 (d, *J* = 5.3 Hz, 1H); 5.93 (dd, *J* = 5.3, 5.3 Hz, 1H); 5.64 (dd, *J* = 5.3, 4.6 Hz, 1H); 4.54-4.32 (m, 3H); 3.99 (s, 3H); 2.99 (d, *J* = 4.9 Hz, 3H); 2.15 (s, 3H); 2.13 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 170.5; 169.8; 169.5; 156.6; 153.6; 152.0; 150.5; 139.2; 122.6; 86.5; 80.5; 73.2; 70.7; 63.2; 34.8; 27.5; 21.0; 20.7; 20.6. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₉H₂₅N₆O₈ 465.1728; Found 465.1726.

***N*⁶-methyl *N*⁶-methyleurea adenosine 7**: Acetyl-protected adenosine **14** (119 mg, 0.26 mmol, 1.0 equiv.) was dissolved in 7 N NH₃ in MeOH. The reaction was stirred at r.t. overnight. After that, the crude was concentrated and purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **7** as a white solid (66 mg, 0.20 mmol, 76% yield). t_R = 26.9 min (0-40% of B in 45 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3362; 3225; 3146; 1689; 1586; 1561; 1514; 1463; 1258; 1129; 1102; 1055; 1039; 983; 817; 793; 768, 740; 696. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 9.72 (q, *J* = 4.5 Hz, 1H); 8.69 (s, 1H); 8.57 (s, 1H); 6.01 (d, *J* = 5.6 Hz, 1H); 5.54 (d, *J* = 5.6 Hz, 1H); 5.26 (d, *J* = 4.9 Hz, 1H); 5.22-5.10 (m, 1H); 4.57 (q, *J* = 5.3 Hz, 1H); 4.17 (q, *J* = 4.5 Hz, 1H); 3.97 (q, *J* = 3.9 Hz, 1H); 3.78 (s, 3H); 3.67-3.59 (m, 2H); 2.81 (d, *J* = 4.5 Hz, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 155.8; 152.7; 151.8; 150.3; 141.3; 121.9; 87.6; 85.7; 73.8; 70.3; 61.2; 34.2; 27.2. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₃H₁₈N₆O₅ 339.1411; Found 339.1413.

10.4 Synthesis of amino acid-modified *N*⁶-methyl *N*⁶-carbamoyl adenosine nucleoside

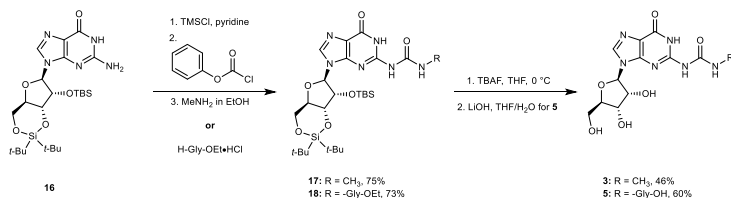


Scheme S16. Synthesis of glycine-modified *N*⁶-methyl *N*⁶-carbamoyl adenosine **8** used as reference.

Acetyl-protected *N*⁶-glycine modified *N*⁶-methyl adenosine **15:** Acetyl-protected *N*⁶-methylurea adenosine derivative **13a** (0.98 g, 1.94 mmol, 1.0 equiv.) was dissolved in dry DMF. The solution was cooled to 0°C and K₂CO₃ (0.80 g, 5.81 mmol, 3.0 equiv.) was added. Finally, methyl iodide (0.30 mL, 4.80 mmol, 2.5 equiv.) was added dropwise to the mixture. The reaction was stirred at r.t. for 16 h. After that, the crude reaction mixture was diluted with Et₂O, and washed with sat. aq. NH₄Cl solution and water. The aqueous phase was re-extracted with Et₂O and the combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (98:2 DCM/IPA) affording **15** as a white foam (0.77 g, 1.47 mmol, 76% yield). *R*_f = 0.69 (95:5 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3196; 2955; 1740; 1681; 1565; 1518; 1463; 1370; 1223; 1206; 1029; 922; 796. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.97 (t, *J* = 5.2 Hz, 1H); 8.57 (s, 1H); 8.10 (s, 1H); 6.25 (d, *J* = 5.4 Hz, 1H); 5.92 (dd, *J* = 5.5, 5.5 Hz, 1H); 5.63 (dd, *J* = 5.5, 4.5 Hz, 1H); 4.49-4.34 (m, 3H); 4.21 (d, *J* = 5.4 Hz, 2H); 4.00 (s, 3H); 3.78 (s, 3H); 2.15 (s, 3H); 2.14 (s, 3H); 2.08 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 171.0; 170.4; 169.8; 169.5; 156.0; 153.4; 152.2; 150.5; 139.4; 122.7; 86.4; 80.5; 73.2; 70.8; 63.2; 52.4; 43.0; 34.9; 21.0; 20.7; 20.5. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₁H₂₇N₆O₁₀ 523.1782; Found 523.1784.

***N*⁶-glycine modified *N*⁶-methyl adenosine **8**:** Acetyl-protected adenosine derivative **15** (103 mg, 0.20 mmol, 1.0 equiv.) was dissolved in 3:1 MeOH/H₂O. The solution was cooled to 0°C. LiOH·H₂O (83 mg, 2.0 mmol, 10.0 equiv.) was added to the solution and the reaction was stirred at r.t. for 2 h. After that, the solution was neutralized with 1 M aqueous HCl solution. The organic solvent was removed under reduced pressure and the remaining aqueous solution mixture was lyophilized. The crude was purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **8** as a white solid (56 mg, 0.15 mmol, 74% yield). *t*_R = 27.1 min (0-30% of B in 45 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3389; 3319; 3220; 3122; 2946; 1718; 1668; 1566; 1538; 1468; 1231; 1209; 1145; 1122; 1086; 1073; 1024; 815; 749. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.68 (s, 1H); 10.36 (t, *J* = 5.5 Hz, 1H); 8.74 (s, 1H); 8.60 (s, 1H); 6.03 (d, *J* = 5.3 Hz, 1H); 5.56 (br s, 1H); 5.25 (br s, 1H); 5.14 (br s, 1H); 4.57 (dd, *J* = 5.3, 5.3 Hz, 1H); 4.18 (dd, *J* = 4.3, 4.3 Hz, 1H); 4.00-3.90 (m, 3H); 3.82 (s, 3H); 3.74-3.51 (m, 2H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 171.5; 155.3; 152.4; 152.1; 150.1; 141.6; 121.9; 87.6; 85.6; 73.9; 70.2; 61.2; 42.6; 34.3. HRMS (ESI) *m/z*: [M-H]⁻ Calcd. for C₁₉H₁₇N₆O₈ 381.1164; Found 381.1165.

10.5 Synthesis of *N*⁶-methyleurea guanosine and amino acid-modified *N*⁶-carbamoyl guanosine nucleoside used as reference



Scheme S17. Synthesis of *N*⁶-methyleurea guanosine **3** and amino acid-modified *N*⁶-carbamoyl guanosine **5** used as reference.

Silyl-protected guanosine **16** was synthesized following a procedure previously reported in the literature.⁸

Silyl-protected *N*⁶-methyleurea guanosine **17:** Step 1: Silyl-protected guanosine **16** (4.00 g, 7.44 mmol, 1.0 equiv.) was suspended in dry pyridine and TMSCl (1.50 mL, 11.9 mmol, 1.6 equiv.) was added. The reaction was stirred at r.t. for 1 h. After that, phenyl chloroformate (1.50 mL, 11.9 mmol, 1.6 equiv.) was added and the reaction was stirred at r.t. for 5 h. Step 2: 33% MeNH₂ in EtOH (4.70 mL, 37.8 mmol, 5.1 equiv.) was added dropwise and the

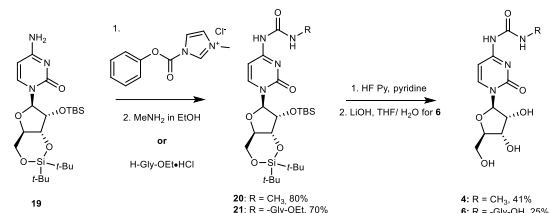
reaction was stirred at r.t. for 16 h. After that, the reaction was quenched with sat. aq. NaHCO₃ solution and the aqueous phase was extracted with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (95:5 DCM/MeOH) affording **17** as a white solid (3.32 g, 7.44 mmol, 75% yield). *R*_f = 0.47 (9:1 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3204; 2900; 2859; 1662; 1602; 1550; 1475; 1399; 1345; 1256; 1170; 1120; 1090; 1058; 1020; 970; 840; 781; 755; 733. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.01 (s, 1H); 9.84 (s, 1H); 8.15 (s, 1H); 7.00 (q, *J* = 4.5 Hz, 1H); 5.80 (s, 1H); 4.59 (d, *J* = 5.2 Hz, 1H); 4.35 (dd, *J* = 8.4, 4.3 Hz, 1H); 4.25 (dd, *J* = 9.1, 5.2 Hz, 1H); 4.08-3.92 (m, 2H); 2.72 (d, *J* = 4.5 Hz, 3H); 1.06 (s, 9H); 1.00 (s, 9H); 0.86 (s, 9H); 0.08 (s, 3H); 0.07 (s, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 155.4; 155.1; 149.0; 148.7; 137.2; 118.9; 90.0; 75.7; 74.9; 74.0; 66.9; 27.3; 26.9; 26.2; 25.7; 22.2; 20.0; 18.0; -4.6; -5.2. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₆H₄₇O₆N₆Si₂ 595.3090; Found 595.3088.

Silyl-protected *N*⁶-glycine modified guanosine **18:** Silyl-protected guanosine **16** (1.00 g, 1.86 mmol, 1.0 equiv.) was suspended in dry pyridine and TMSCl (0.40 mL, 3.15 mmol, 1.7 equiv.) was added. The reaction was stirred at r.t. for 2 h. After that, phenyl chloroformate (0.40 mL, 3.17 mmol, 1.7 equiv.) was added and the mixture was stirred at r.t. for 5 h. **H-Gly-OEt·HCl** (0.52 g, 3.72 mmol, 2.0 equiv.) was added to the solution and the mixture was stirred at r.t. for 16 h. After that, the reaction was quenched with sat. aq. NaHCO₃ solution and the aqueous phase was extracted with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (DCM/MeOH) affording **18** as a white solid (0.90 g, 1.35 mmol, 73% yield). *R*_f = 0.62 (9:1 DCM/MeOH). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 2932; 2859; 1827; 1735; 1661; 1618; 1550; 1472; 1397; 1366; 1251; 1142; 1054; 999; 940; 894; 832; 781; 747; 744; 687. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 11.83 (br s, 1H); 10.09 (br s, 1H); 8.17 (s, 1H); 7.53 (s, 1H); 5.83 (s, 1H); 4.59 (d, *J* = 5.1 Hz, 1H); 4.36 (dd, *J* = 8.3, 4.2 Hz, 1H); 4.26 (dd, *J* = 9.0, 5.1 Hz, 1H); 4.14 (q, *J* = 7.1 Hz, 2H); 4.07-3.92 (m, 4H); 1.21 (t, *J* = 7.1 Hz, 3H); 1.07 (s, 9H); 1.01 (s, 9H); 0.86 (s, 9H); 0.08 (s, 3H); 0.07 (s, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 169.7; 155.0; 148.8; 148.7; 148.5; 137.4; 119.1; 90.1; 75.7; 74.9; 74.1; 66.9; 60.8; 41.4; 27.3; 26.9; 25.7; 22.2; 20.0; 18.0; 14.1; -4.6; -5.1. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₉H₅₁O₈N₆Si₂ 667.3301; Found 667.3295.

***N*⁶-methyleurea guanosine **3**:** Silyl-protected cytidine derivative **17** (0.40 g, 0.67 mmol, 1.0 equiv.) was suspended in THF. 1 M tetrabutylammonium fluoride (TBAF) in THF (4.0 mL, 4.0 mmol, 6.0 equiv.) was added and the reaction was stirred at r.t. for 4 h. After that, the reaction was quenched by addition of methoxytrimethylsilane (0.56 mL, 4.0 mmol, 6.0 equiv.). The crude was concentrated under reduced pressure and purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **3** as a white solid (106 mg, 0.31 mmol, 46% yield). *t*_R = 12.1 min (0-30% of B in 30 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3257; 1703; 1659; 1636; 1554; 1497; 1475; 1406; 1354; 1267; 1233; 1170; 1125; 1090; 1058; 1042; 1020; 972; 862; 821; 781; 750; 720. ¹H NMR (500 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 11.97 (br s, 1H); 10.18 (br s, 1H); 8.16 (s, 1H); 6.82 (q, *J* = 4.7 Hz, 1H); 5.75 (d, *J* = 5.5 Hz, 1H); 5.47 (d, *J* = 5.9 Hz, 1H); 5.17 (d, *J* = 4.9 Hz, 1H); 5.04 (t, *J* = 5.4 Hz, 1H); 4.41 (q, *J* = 5.5 Hz, 1H); 4.11 (q, *J* = 4.6 Hz, 1H); 3.90 (q, *J* = 4.0 Hz, 1H); 3.64 (ddd, *J* = 12.0, 5.5, 4.0 Hz, 1H); 3.54 (ddd, *J* = 11.2, 5.5, 4.1 Hz, 1H); 2.72 (d, *J* = 4.7 Hz, 3H). ¹³C{¹H} NMR (126 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 155.6; 155.1; 149.2; 148.9; 137.1; 119.1; 86.9; 85.3; 74.0; 70.2; 61.1; 26.2. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₉H₁₇N₆O₈ 341.1204; Found 341.1203.

***N*⁶-glycine modified guanosine **5**:** Step 1: Silyl-protected guanosine derivative **18** (20 mg, 30 μmol, 1.0 equiv.) was suspended in THF. 1 M tetrabutylammonium fluoride (TBAF) in THF (0.18 mL, 0.18 mmol, 6.0 equiv.) was added and the reaction was stirred at r.t. for 4 h. The reaction was quenched by addition of methoxytrimethylsilane (25 μL, 0.18 mmol, 6.0 equiv.). The crude was concentrated under reduced pressure. Step 2: The residue was redissolved in 1:1 THF/H₂O and LiOH (13 mg, 0.30 mmol, 10.0 equiv.) was added. The reaction was stirred at 0°C for 1 h. Subsequently, the reaction was quenched by addition of aqueous 1 M HCl solution. The organic solvent was removed under reduced pressure and the remaining aqueous solution was lyophilized. The crude was purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **5** as a white solid (7.0 mg, 30 μmol, 60% yield). *t*_R = 11.8 min (0-30% of B in 30 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3370; 1693; 1611; 1532; 1460; 1362; 1209; 1051; 774; 665. ¹H NMR (500 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 8.19 (s, 1H); 7.61 (br s, 1H); 5.80 (d, *J* = 5.0 Hz, 1H); 5.06 (br s, 1H); 4.37 (t, *J* = 5.0 Hz, 1H); 4.11 (t, *J* = 4.6 Hz, 1H); 3.91 (q, *J* = 4.0 Hz, 1H); 3.80-3.72 (m, 2H); 3.66 (dd, *J* = 12.0, 4.0 Hz, 1H); 3.56 (dd, *J* = 12.0, 4.0 Hz, 1H). ¹³C{¹H} NMR (126 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 170.8; 155.2; 154.6; 148.9; 148.6; 137.2; 119.2; 87.2; 85.1; 74.1; 69.9; 61.0; 42.3. HRMS (ESI) *m/z*: [M-H]⁻ Calcd. for C₁₉H₁₅N₆O₈ 383.0956; Found 383.0964.

10.6 Synthesis of *N*⁶-methylurea cytidine and amino acid-modified *N*⁶-carbamoyl cytidine nucleoside used as reference



Scheme S18. Synthesis of *N*⁶-methylurea cytidine **4** and amino acid-modified *N*⁶-carbamoyl cytidine **6** used as reference.

Silyl-protected cytidine **19** was synthesized following a procedure previously reported in the literature.⁸

Silyl-protected *N*⁶-methyurea cytidine **20:** Silyl-protected cytidine **19** (3.00 g, 6.03 mmol, 1.0 equiv.) and 1-*N*-methyl-3-phenoxycarbonyl-imidazolium chloride (2.88 g, 12.1 mmol, 2.0 equiv.) were dissolved in dry DCM. The reaction was stirred at r.t. for 16 h. 2 M MeNH₂ in THF (15.1 mL, 30.1 mmol, 5.0 equiv.) was added to the solution and the mixture was stirred at r.t. for 16 h. After that, the reaction was quenched with sat. aq. NaHCO₃ solution and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (1:1 DCM/EtOAc) affording **20** as a white solid (2.67 g, 4.81 mmol, 80% yield). *R*_f = 0.21 (1:1 DCM/EtOAc). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 2933; 1720; 1641; 1562; 1509; 1472; 1416; 1382; 1326; 1204; 1166; 1145; 1129; 1115; 1079; 1053; 1023; 1000; 940; 902; 829; 783; 751; 712; 687; 652; 493; 441; 408. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.78 (s, 1H); 9.11 (s, 1H); 7.67 (d, *J* = 7.8 Hz, 1H); 7.60 (d, *J* = 7.8 Hz, 1H); 5.69 (s, 1H); 4.55 (dd, *J* = 9.3, 5.2 Hz, 1H); 4.35-4.24 (m, 2H); 4.02 (t, *J* = 10.0 Hz, 1H); 3.80 (dd, *J* = 9.7, 4.2 Hz, 1H); 2.82 (d, *J* = 4.3 Hz, 3H); 1.03 (s, 9H); 1.02 (s, 9H); 0.95 (s, 9H); 0.24 (s, 3H); 0.17 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 165.1; 156.4; 154.8; 141.3; 97.5; 94.2; 75.8; 75.4; 74.9; 67.9; 27.6; 27.1; 26.7; 26.0; 23.0; 20.5; 18.3; -4.2; -4.6. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₈H₄₇N₄O₈Si₂ 555.3028; Found 555.3029.

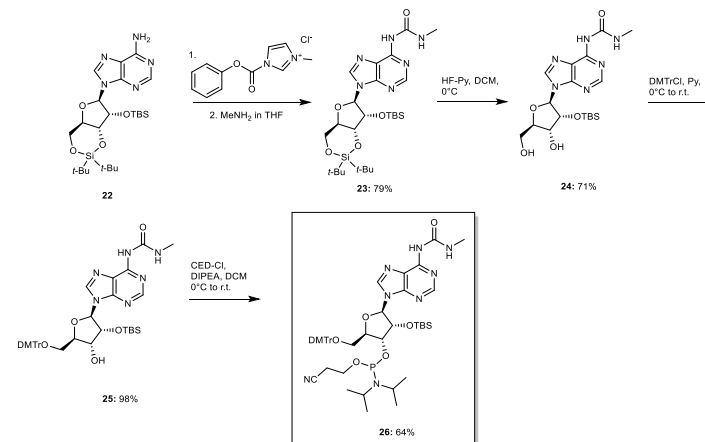
Silyl-protected *N*⁶-glycine modified cytidine **21:** Silyl-protected cytidine **19** (1.00 g, 2.01 mmol, 1.0 equiv.) and 1-*N*-methyl-3-phenoxycarbonyl-imidazolium chloride (0.90 g, 4.02 mmol, 2.0 equiv.) were dissolved in dry DCM. The reaction was stirred at r.t. for 14 h. Subsequently, a suspension in dry DCM of **H-Gly-OEt·HCl** (0.56 g, 4.02 mmol, 2.0 equiv.), containing TEA (0.84 mL, 6.03 mmol, 3.0 equiv.) was transferred to the reaction mixture. The reaction was stirred at r.t. for 16 h. After that, the reaction was quenched with sat. aq. NaHCO₃ solution and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (0 to 50% *i*-Hex/EtOAc) affording **21** as a white solid (0.88 g, 1.40 mmol, 70% yield). *R*_f = 0.33 (1:1 *i*-Hex/EtOAc); IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3229; 3063; 2932; 2857; 1718; 1643; 1570; 1503; 1471; 1434; 1373; 1263; 1187; 1165; 1126; 1053; 1026; 1010; 996; 937; 897; 826; 783; 752; 713; 685. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 10.88 (s, 1H); 9.68 (s, 1H); 7.70-7.55 (m, 2H); 5.64 (s, 1H); 4.52 (dd, *J* = 9.2, 5.2 Hz, 1H); 4.29-4.25 (m, 1H); 4.26-4.21 (m, 1H); 4.17 (q, *J* = 7.2 Hz, 2H); 4.02-3.96 (m, 1H); 3.96-3.90 (m, 2H); 3.75 (dd, *J* = 9.7, 4.1 Hz, 1H); 1.24 (t, *J* = 7.2 Hz, 3H); 1.00 (s, 9H); 0.99 (s, 9H); 0.91 (s, 9H); 0.16 (s, 3H); 0.12 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 170.1; 165.0; 156.3; 154.7; 141.3; 97.6; 94.0; 75.6; 75.3; 74.8; 67.8; 61.1; 41.7; 29.0; 25.9; 22.7; 20.4; 18.2; 14.3; -4.3; -4.8. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₈H₅₁N₄O₈Si₂ 627.3240; Found 627.3243.

***N*⁶-methylurea cytidine **4**:** Silyl-protected cytidine derivative **20** (0.37 g, 0.67 mmol, 1.0 equiv.) was dissolved in dry pyridine. 70% HF-pyridine (0.54 mL, 3.33 mmol, 5 equiv.) was added to the solution and the reaction was stirred at r.t. for 5 h. After that, the reaction was quenched by addition of methoxytrimethylsilane (1.38 mL, 4.79 mmol, 15 equiv.). The crude was concentrated under reduced pressure and purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **4** as a white solid (82 mg, 0.67 mmol, 41% yield). *t*_R = 13.2 min (0-15% of B in 15 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3257; 1712; 1693; 1651; 1602; 1561; 1538; 1483; 1374; 1334; 1308; 1229; 1154; 1138; 1098; 1061; 1032; 983; 944; 871; 851; 811; 792; 727; 704; 650; 601; 586; 438. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 9.92 (s, 1H); 8.24 (d, *J* = 7.4 Hz, 1H); 6.22 (br s, 1H); 5.76 (d, *J* = 3.5 Hz, 1H); 5.43 (br s, 1H); 5.12 (br s, 1H); 5.04 (br s, 1H); 4.00-3.91 (m, 2H); 3.87 (dt, *J* = 5.9, 3.0 Hz, 1H); 3.70 (dd, *J* = 12.2, 2.8 Hz, 1H); 3.57 (dd, *J* = 12.2, 3.1 Hz, 1H); 2.75 (d, *J* = 4.7 Hz, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 162.2; 154.2; 153.7; 143.7; 94.6; 89.8; 84.2; 74.3; 68.9; 60.1; 26.0. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₁H₁₇N₄O₆ 301.1142; Found 301.1144.

***N*⁶-glycine modified cytidine **6**:** Step 1: Silyl-protected cytidine derivative **21** (200 mg, 0.32 mmol, 1.0 equiv.) was dissolved in dry pyridine. 70% HF-pyridine (0.26 mL, 1.60 mmol, 5 equiv.) was added to the solution and the reaction was stirred at r.t. for 5 h. After that, the reaction was quenched by addition of methoxytrimethylsilane (0.66 mL, 4.79 mmol, 15 equiv.). The crude was concentrated under reduced pressure. Step 2: The obtained residue was redissolved in 1:1 THF/H₂O and LiOH (67 mg, 1.6 mmol, 5.0 equiv.) was added to the suspension. The reaction was stirred at r.t. for 1 h. Subsequently, the crude reaction mixture was neutralised by addition of Dowex ion-exchange resin. The resin was filtered off and the organic solvent was removed under reduced pressure. The remaining aqueous solution was lyophilized and the crude was purified by semi-preparative HPLC (A: H₂O and B: 20:80 H₂O/MeCN; both containing 0.1% of formic acid) affording **6** as a white solid (27 mg, 0.08 mmol, 25% yield). *t*_R = 12.1 min (0-15% of B in 15 min). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3339; 3191; 3096; 2942; 2544; 1703; 1655; 1608; 1572; 1530; 1468; 1446; 1371; 1361; 1307; 1252; 1219; 1171; 1135; 1107; 1085; 1059; 1035; 994; 983; 964; 915; 905; 849; 840; 805; 791; 755; 722; 696; 667. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 12.67 (br s, 1H); 10.04 (br s, 1H); 9.10 (br s, 1H); 8.28 (d, *J* = 7.4 Hz, 1H); 6.26 (s, 1H); 5.77 (d, *J* = 3.0 Hz, 1H); 5.46 (d, *J* = 4.7 Hz, 1H); 5.15 (t, *J* = 5.1 Hz, 1H); 5.07 (d, *J* = 5.3 Hz, 1H); 4.02-3.90 (m, 4H); 3.90-3.85 (m, 1H); 3.71 (dd, *J* = 12.1, 2.9 Hz, 1H); 3.58 (dd, *J* = 12.1, 3.1 Hz, 1H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 171.3; 162.2; 153.9; 153.7; 143.9; 94.6; 89.9; 84.2; 74.4; 68.8; 60.0; 41.4. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₂H₁₇N₄O₈ 345.1040; Found 345.1043.

11. Synthesis of methylurea nucleoside phosphoramidites

11.1 Synthesis of *N*⁶-methylurea adenosine phosphoramidite



Scheme S19. Synthesis of *N*⁶-methylurea adenosine phosphoramidite **26**.

Silyl-protected adenosine **22** was synthesized following a procedure previously reported in the literature.⁸

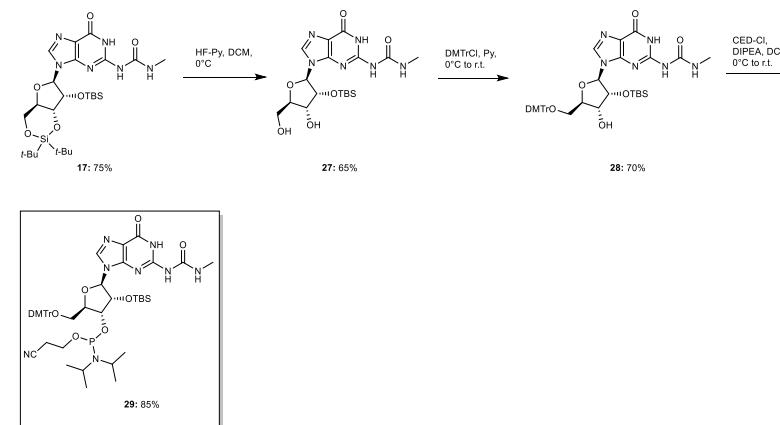
Silyl-protected *N*⁶-methylurea adenosine **23:** Silyl-protected adenosine **22** (4.00 g, 7.67 mmol, 1.0 equiv.) and 1-*N*-methyl-3-phenoxycarbonyl-imidazolium chloride (3.66 g, 15.3 mmol, 2.0 equiv.) were dissolved in dry DCM. The reaction was stirred at r.t. for 16 h. After that, the reaction was quenched with sat. aq. NaHCO₃ solution and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by column chromatography on silica gel (95:5 DCM/IPA) affording **23** as a white foam (3.52 g, 6.08 mmol, 79% yield). *R*_f = 0.58 (95:5 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3248; 2932; 2858; 1706; 1615; 1586; 1558; 1471; 1257; 1138; 1107; 1066; 1011; 1001; 893; 825; 783; 754. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 9.33 (br s, 1H); 8.50 (s, 1H); 8.12 (s, 1H); 5.98 (br s, 1H); 4.57 (d, *J* = 4.5 Hz, 1H); 4.50 (dd, *J* = 9.2, 5.1 Hz, 1H); 4.48-4.42 (m, 1H); 4.24 (td, *J* = 10.0, 5.1 Hz, 1H); 4.06 (dd, *J* = 10.5, 9.2 Hz, 1H); 3.01 (d, *J* = 4.7 Hz, 3H); 1.08 (s, 9H); 1.04 (s, 9H); 0.93 (s, 9H); 0.16 (s, 3H); 0.15 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 154.8; 149.7; 141.3; 121.0; 92.4; 76.0; 75.8; 74.9; 67.9; 27.6; 27.1; 26.8; 26.0; 22.9; 20.5; 18.5; -4.1; -4.9. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₂₆H₄₇N₆O₅Si₂ 579.3141; Found 579.3142.

5',3'-deprotected *N*⁶-methylurea adenosine 24: Modified adenosine **23** (1.68 g, 2.90 mmol, 1.0 equiv.) was dissolved in 9:1 DCM/pyridine in a plastic flask and the solution was cooled to 0°C. Subsequently, 70% HF-pyridine (5.0 equiv.) was added slowly to the adenosine's solution and the reaction was stirred at 0°C for 2 h. After that, the crude reaction mixture was diluted with sat. aq. NaHCO₃ solution and the crude was extracted three times with DCM. The combined organic layers were washed with water, dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by silica gel column chromatography (96:4 DCM/IPA) affording **24** as a white solid (0.90 g, 2.04 mmol, 71% yield). *R*_f = 0.48 (95:5 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3244; 2928; 2857; 1694; 1611; 1590; 1548; 1470; 1360; 1331; 1297; 1252; 1220; 1131; 1085; 1062; 1024; 865; 835; 779. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 9.26 (q, *J* = 4.7 Hz, 1H); 8.51 (s, 1H); 8.41 (s, 1H); 8.13 (s, 1H); 5.97 (dd, *J* = 11.9, 2.2 Hz, 1H); 5.84 (d, *J* = 7.3 Hz, 1H); 5.08 (dd, *J* = 7.3, 4.8 Hz, 1H); 4.40-4.30 (m, 2H); 3.97 (dt, *J* = 12.9, 2.0 Hz, 1H); 3.77 (ddd, *J* = 13.2, 11.9, 1.6 Hz, 1H); 3.02 (d, *J* = 4.7 Hz, 3H); 2.88 (d, *J* = 0.7 Hz, 1H); 0.79 (s, 9H); -0.19 (s, 3H); -0.39 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 154.4; 151.1; 151.0; 149.2; 143.1; 122.0; 91.2; 87.7; 74.7; 72.9; 63.4; 26.9; 25.6; 18.0; -5.2; -5.3. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₈H₃₁N₆O₅Si₂ 439.2120; Found 439.2121.

DMTr-protected *N*⁶-methylurea adenosine 25: 3',5'-Deprotected adenosine **24** (0.80 g, 1.82 mmol, 1.0 equiv.) was dissolved in dry pyridine. DMTrCl (0.87 g, 2.55 mmol, 1.4 equiv.) was added. The reaction was stirred at r.t. for 16 h. After that, the crude was concentrated under reduced pressure and purified by silica gel column chromatography (97:3 DCM/IPA, containing 0.1% pyridine). The DMTr-protected compound **25** was isolated as a white foam (1.34 g, 1.80 mmol, 98% yield). *R*_f = 0.70 (95:5 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 3553; 2927; 1702; 1608; 1589; 1548; 1507; 1468; 1299; 1248; 1174; 1066; 1031; 833; 780; 700. ¹H NMR (400 MHz, CDCl₃, 298 K): δ (ppm) = 9.35 (s, 1H); 8.72-8.61 (m, 1H); 8.54 (s, 1H); 8.45 (s, 1H); 7.50 (d, *J* = 7.5 Hz, 2H); 7.39-7.34 (m, 4H); 7.29-7.25 (m, 2H); 7.20 (t, *J* = 7.3 Hz, 1H); 6.86-6.82 (m, 4H); 6.15 (d, *J* = 4.7 Hz, 1H); 5.10 (t, *J* = 4.7 Hz, 1H); 4.51 (q, *J* = 5.0 Hz, 1H); 4.29 (q, *J* = 3.7, 3.3 Hz, 1H); 4.02 (d, *J* = 5.6 Hz, 1H); 3.76 (s, 3H); 3.76 (s, 3H); 3.49-3.42 (m, 2H); 2.91 (d, *J* = 4.6 Hz, 3H); 0.84 (s, 9H); 0.04 (s, 3H); -0.07 (s, 3H). ¹³C{¹H} NMR (100 MHz, CDCl₃, 298 K): δ (ppm) = 159.6; 154.8; 151.8; 151.4; 151.2; 146.1; 143.3; 136.7; 131.0; 130.9; 129.0; 128.6; 127.5; 121.5; 113.8; 89.8; 87.1; 84.8; 76.4; 72.0; 64.4; 55.4; 26.6; 26.1; 18.6; -4.7; -4.9. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₉H₄₉N₆O₇Si 741.3427; Found 741.3428.

***N*⁶-methylurea adenosine phosphoramidite 26:** 5'-DMTr-protected adenosine **25** (300 mg, 0.41 mmol, 1.0 equiv.) was dissolved in dry DCM and *N,N*-diisopropylethylamine (DIPEA) (0.28 mL, 1.62 mmol, 4.0 equiv.) was added to the solution. The mixture was cooled to 0°C, followed by the addition of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (CED-Cl) (0.23 mL, 1.01 mmol, 2.5 equiv.). The reaction was stirred at r.t. for 5 h. After that, sat. aq. NaHCO₃ solution was added to the crude reaction mixture and the aqueous phase was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by silica gel column chromatography (1:1 to 4:6 Hex/EtOAc, containing 0.1% pyridine). Finally, the product was lyophilized from benzene affording **26** as a mixture of diastereoisomers (245 mg, 0.26 mmol, 64% yield). *R*_f = 0.17 (1:1 Hex/EtOAc). ³¹P{¹H} NMR (162 MHz, acetone-*d*₆, 298 K): δ (ppm) = 150.4; 149.1. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₄₈H₆₆N₆O₅PSi 941.4505; Found 941.4488.

11.2 Synthesis of *N*⁶-methylurea guanosine phosphoramidite



Scheme S20. Synthesis of *N*⁶-methylurea guanosine phosphoramidite **29**.

The synthetic procedure of silyl-protected *N*⁶-methylurea guanosine **18** is described in Section 10.5.

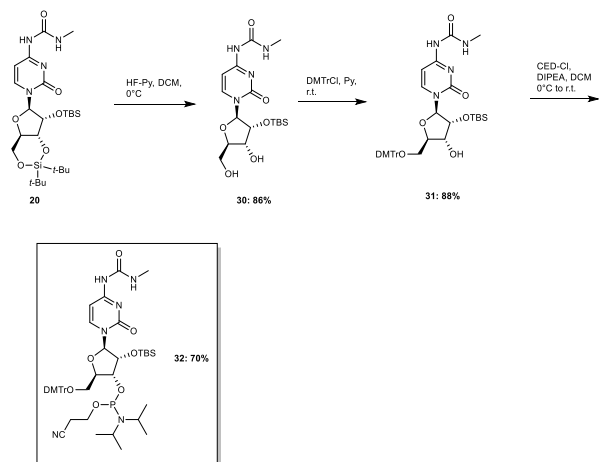
5',3'-deprotected *N*⁶-methylurea guanosine 27: Protected *N*⁶-methylurea guanosine derivative **17** (1.6 g, 2.7 mmol, 1.0 equiv.) was dissolved in 9:1 DCM/pyridine and the solution was cooled to 0°C. Subsequently, 70% HF-pyridine (0.35 mL, 13 mmol, 5.0 equiv.) was added slowly to the solution and the reaction was stirred at 0°C for 2 h. After that, the crude reaction mixture was diluted with sat. aq. NaHCO₃ solution and the crude was extracted with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by silica gel column chromatography (94:6 DCM/IPA) affording **27** as a white solid (0.80 g, 2.7 mmol, 65% yield). *R*_f = 0.4 (9:1 DCM/MeOH). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 2857; 1666; 1606; 1534; 1403; 1362; 1321; 1145; 1089; 1058; 1020; 908; 835; 779; 750; 641. ¹H NMR (400 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 11.98 (br s, 1H); 10.06 (br s, 1H); 8.18 (s, 1H); 6.86 (br s, 1H); 5.79 (d, *J* = 6.2 Hz, 1H); 5.10 (t, *J* = 5.4 Hz, 1H); 5.04 (d, *J* = 5.1 Hz, 1H); 4.51 (dd, *J* = 6.2, 4.9 Hz, 1H); 4.09 (td, *J* = 5.0, 2.9 Hz, 1H); 3.95 (q, *J* = 3.7 Hz, 1H); 3.66 (ddd, *J* = 12.0, 5.4, 4.1 Hz, 1H); 3.57 (ddd, *J* = 12.0, 5.4, 3.8 Hz, 1H); 2.71 (d, *J* = 4.6 Hz, 3H); 0.74 (s, 9H); -0.06 (s, 3H); -0.17 (s, 3H). ¹³C{¹H} NMR (100 MHz, DMSO-*d*₆, 298 K): δ (ppm) = 155.5; 149.3; 137.1; 119.1; 86.5; 85.9; 76.2; 70.5; 61.3; 26.2; 25.5; 17.8; -4.9; -5.4. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₁₈H₃₁O₆N₆Si 455.2068; Found 455.2069.

DMTr-protected *N*⁶-methylurea guanosine 28: 3',5'-Deprotected *N*⁶-methylurea guanosine derivative **27** (0.50 g, 1.1 mmol, 1.0 equiv.) was dissolved in dry pyridine and 4,4'-dimethoxytrityl chloride (DMTrCl) (0.56 g, 1.65 mmol, 1.5 equiv.) was added to the solution. The reaction was stirred at r.t. for 16 h. After that, the reaction mixture was concentrated under reduced pressure and the crude was purified by silica gel column chromatography (96:4 DCM/MeOH, containing 0.1% of pyridine). The DMTr-protected compound **28** was isolated as a white foam (0.59 g, 0.78 mmol, 70% yield). *R*_f = 0.50 (9:1 DCM/MeOH). IR (ATR): $\tilde{\nu}$ (cm⁻¹) = 2928; 1666; 1607; 1548; 1508; 1484; 1410; 1361; 1298; 1249; 1174; 1128; 1074; 1035; 969; 914; 833; 782; 752; 726; 699. ¹H NMR (400 MHz, acetone-*d*₆, 298 K): δ (ppm) = 12.21 (br s, 1H); 9.61 (br s, 1H); 8.00 (s, 1H); 7.49-7.44 (m, 2H); 7.38-7.17 (m, 8H); 6.88-6.80 (m, 4H); 6.79-6.70 (m, 1H); 5.94 (d, *J* = 5.1 Hz, 1H); 4.83 (t, *J* = 5.1 Hz, 1H); 4.50-4.42 (m, 1H); 4.22-4.15 (m, 1H); 3.89 (br s, 1H); 3.77 (s, 6H); 3.50-3.31 (m, 2H); 2.83 (d, *J* = 4.4 Hz, 3H); 0.84 (s, 9H); 0.05 (s, 3H); -0.07 (s, 3H). ¹³C{¹H} NMR (100 MHz, acetone-*d*₆, 298 K): δ (ppm) = 159.6; 156.6; 150.3; 146.0; 137.5; 136.6; 130.9; 128.9; 128.6; 127.6; 121.0; 113.9; 88.7; 87.1; 84.9; 77.3; 72.2; 64.8; 55.5; 26.7; 26.1; 18.7; -4.7; -4.9. HRMS (ESI) *m/z*: [M+H]⁺ Calcd. for C₃₉H₄₉O₉N₆Si 757.3375; Found 757.3375.

***N*⁶-methylurea guanosine phosphoramidite 29:** A solution of 5'-DMTr-protected guanosine derivative **28** (250 mg, 0.33 mmol, 1.0 equiv.) and DIPEA (0.23 mL, 1.3 mmol, 4.0 equiv.) in dry DCM was cooled to 0°C. 2-Cyanoethyl *N,N*-diisopropylchlorophosphoramidite (CED-Cl) (0.18 mL, 0.83 mmol, 2.5 equiv.) was added to the solution and the reaction mixture was stirred at r.t. for 5 h. After that, the reaction was quenched by addition of sat. aq. NaHCO₃ solution and the crude was extracted three times with DCM. The combined organic layers were dried (MgSO₄), filtered and concentrated under reduced pressure. The crude was purified by silica gel column chromatography (2:1 *n*-Hex/EtOAc, containing 0.1% of pyridine). Finally, the product was lyophilized from benzene affording **29** as

a mixture of diastereoisomers (270 mg, 0.28 mmol, 85% yield). $R_f = 0.36$ (9:1 DCM/MeOH). $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K): δ (ppm) = 150.7; 148.8. HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{48}\text{H}_{66}\text{O}_9\text{N}_8\text{PSi}$ 957.4453; Found 957.4436.

11.3 Synthesis of N^6 -methylurea cytidine phosphoramidite



Scheme S21. Synthesis of N^6 -methylurea cytidine phosphoramidite **32**.

The synthetic procedure of silyl-protected N^6 -methylurea cytidine **20** is described in Section 10.6.

5',3'-deprotected N^6 -methylurea cytidine 30: Protected cytidine **20** (250 mg, 0.45 mmol, 1.0 equiv.) was dissolved in 9:1 DCM/pyridine in a plastic flask and the solution was cooled to 0°C. Subsequently, 70% HF-pyridine (0.06 mL, 2.25 mmol, 5.0 equiv.) was added slowly to the solution and the reaction was stirred at 0°C for 2 h. After that, the crude reaction mixture was diluted with sat. aq. NaHCO_3 solution and the crude was extracted three times with DCM. The combined organic layers were washed with water, dried (MgSO_4), filtered and concentrated under reduced pressure. The crude was purified by silica gel column chromatography (97:3 DCM/IPA) affording **30** as a white solid (160 mg, 0.39 mmol, 86% yield). $R_f = 0.27$ (9:1 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm^{-1}) = 2928; 1704; 1641; 1567; 1504; 1462; 1385; 1362; 1275; 1251; 1112; 1065; 999; 960; 918; 839; 812; 789; 738; 691; 668; 637; 598; 418; 405. ^1H NMR (400 MHz, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 9.94 (s, 1H); 8.31 (d, $J = 7.5$ Hz, 1H); 6.21 (br s, 1H); 5.69 (d, $J = 2.6$ Hz, 1H); 5.18 (t, $J = 5.1$ Hz, 1H); 4.99 (d, $J = 5.3$ Hz, 1H); 4.07 (dd, $J = 4.0, 2.6$ Hz, 1H); 3.99-3.86 (m, 2H); 3.76 (ddd, $J = 12.1, 5.1, 2.4$ Hz, 1H); 3.60 (ddd, $J = 12.1, 5.1, 2.6$ Hz, 1H); 2.75 (d, $J = 4.6$ Hz, 3H); 0.86 (s, 9H); 0.07 (s, 3H), 0.05 (s, 3H). $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 162.2; 154.2; 153.6; 143.3; 94.4; 90.0; 83.8; 76.4; 68.1; 59.5; 26.0; 25.8; 18.0; -4.8; -4.9. HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{17}\text{H}_{31}\text{N}_4\text{O}_6\text{Si}$ 415.2007; Found 415.2010.

DMTr-protected N^6 -methylurea cytidine 31: 3',5'-Deprotected cytidine derivative **30** (100 mg, 0.24 mmol, 1.0 equiv.) was dissolved in dry pyridine and 4,4'-dimethoxytrityl chloride (DMTrCl) (123 mg, 0.36 mmol, 1.5 equiv.) was added to the solution. The reaction was stirred at r.t. for 16 h. After that, the reaction mixture was concentrated under reduced pressure and purified by silica gel column chromatography (96:4 DCM/IPA, containing 0.1% of pyridine) affording the DMTr-protected compound **31** as a white foam (152 mg, 0.24 mmol, 88% yield). $R_f = 0.40$ (9:1 DCM/IPA). IR (ATR): $\tilde{\nu}$ (cm^{-1}) = 3081; 1717; 1645; 1610; 1568; 1506; 1446; 1416; 1385; 1249; 1176; 1114; 1062; 1035; 1005; 828; 786; 754; 702; 585; 407. ^1H NMR (400 MHz, acetone- d_6 , 298 K): δ (ppm) = 9.22 (br s, 1H); 8.48 (br s, 1H); 7.53-7.48 (m, 2H); 7.41-7.33 (m, 6H); 7.27 (t, $J = 7.7$ Hz, 1H); 6.97-6.88 (m, 4H); 5.85 (d, $J = 1.1$ Hz, 1H); 4.52 (td, $J = 8.1, 4.5$ Hz, 1H); 4.42 (d, $J = 4.0$ Hz, 1H); 4.22 (dt, $J = 8.6, 2.7$ Hz, 1H); 3.87 (d, $J = 7.7$ Hz, 1H); 3.82 (s, 6H); 3.54 (d, $J = 2.4$ Hz, 2H); 2.79 (d, $J = 4.2$ Hz, 3H); 0.97 (s, 9H); 0.29 (s, 3H); 0.21 (s, 3H). $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, acetone- d_6 , 298 K): δ (ppm) = 159.7; 154.9; 145.4; 144.1; 136.8; 136.5; 131.0; 130.8; 129.1; 128.8; 127.8; 114.1; 92.3; 87.6; 83.1; 77.6; 69.7; 62.4; 55.5; 26.6; 26.3; 18.8; -4.2; -4.6. HRMS (ESI) m/z $[\text{M}+\text{H}]^+$ Calcd. for $\text{C}_{38}\text{H}_{49}\text{N}_4\text{O}_6\text{Si}$ 717.3314; Found 717.3317.

N^6 -methylurea cytidine phosphoramidite 32: A solution of 5'-DMTr-protected cytidine derivative **31** (90 mg, 0.13 mmol, 1.0 equiv.) and DIPEA (0.09 mL, 0.50 mmol, 4.0 equiv.) in dry DCM was cooled to 0°C. 2-Cyanoethyl N,N -diisopropylchlorophosphoramidite (CED-Cl) (0.07 mL, 0.31 mmol, 2.5 equiv.) was added to the solution and the reaction was stirred at r.t. for 5 h. After that, the reaction was quenched by addition of aq. sat. NaHCO_3 and the crude was extracted three times with DCM. The combined organic layers were dried (MgSO_4), filtered and concentrated under reduced pressure. The crude was purified by silica gel column chromatography (2:1 n -Hex/EOAc, containing 0.1% pyridine). Finally, the product was lyophilized from benzene affording **32** as a mixture of diastereoisomers (82 mg, 0.13 mmol, 70% yield). $R_f = 0.20$ (2:1 n -Hex/EOAc). $^{31}\text{P}\{^1\text{H}\}$ NMR (162 MHz, acetone- d_6 , 298 K): δ (ppm) = 150.4, 148.1. HRMS (ESI) m/z $[\text{M}-\text{H}]^-$ Calcd. for $\text{C}_{47}\text{H}_{64}\text{N}_6\text{O}_9\text{PSi}$ 915.4247; Found 915.4263.

12. Synthesized oligonucleotides using a DNA/RNA automated synthesizer

ON2: 5'-RAA AAA AAA AA-3'; **R** = N^6 -methylurea adenosine

ON2': 5'-RAU CGC UGU AC-3'; **R** = N^6 -methylurea adenosine

ON2'': 5'-RAU CGC UGU RC-3'; **R** = N^6 -methylurea adenosine

ON3: 5'-RAU CGC UGU AC-3'; **R** = N^6 -methylurea guanosine

ON4: 5'-RAU CGC UGU AC-3'; **R** = N^6 -methylurea cytidine

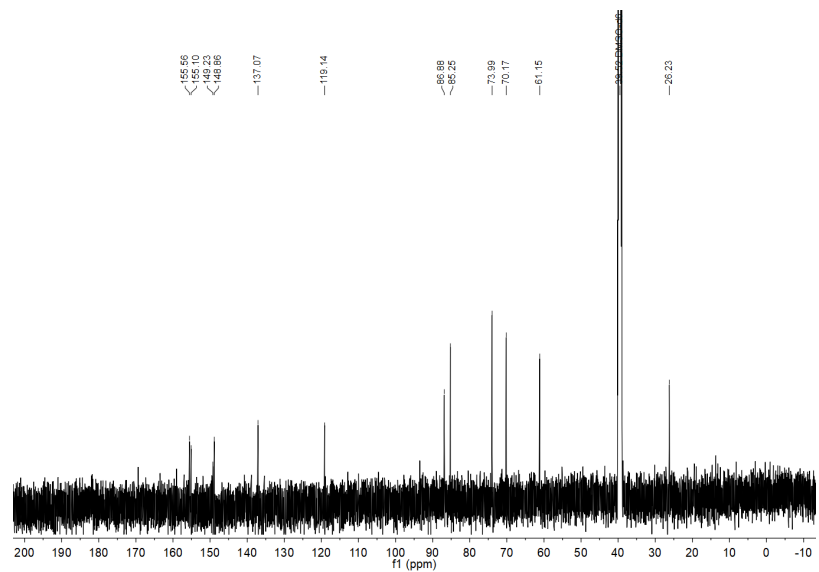
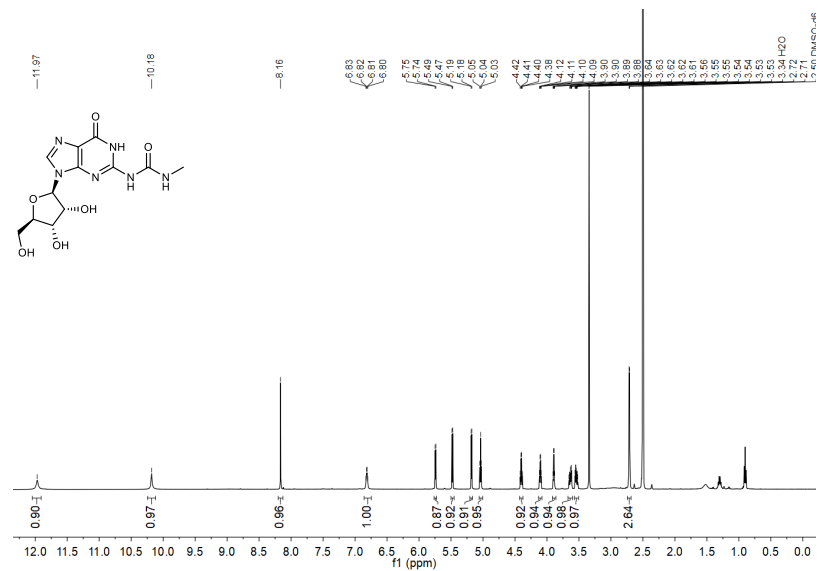
ON7: 5'-RUU AUU AUU UA-3'; **R** = N^6 -methylurea adenosine

ON9: 5'-(UAA AUA AUA A) $_m$ R'-3'; **R'** = nm⁵U

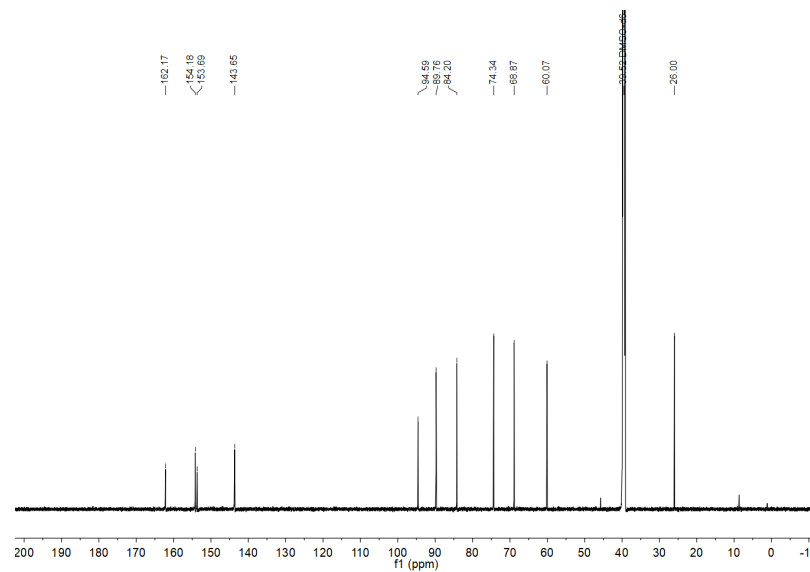
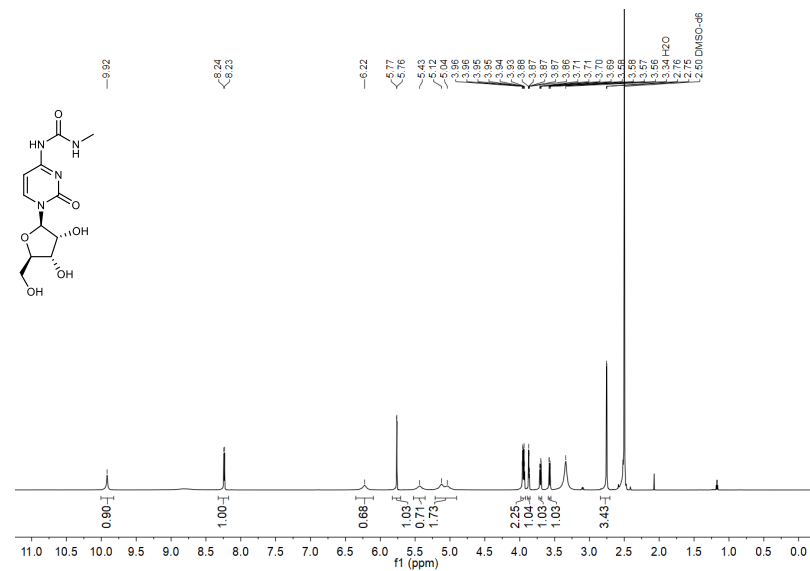
Table S11. HPLC retention times (0-30% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of **ON2-4**, **ON7** and **ON9**.

Strand	t_R (min)	m/z calcd. for $[\text{M}-\text{H}]^-$	Found
ON2	27.1	3613.6	3614.0
ON2'	26.7	3505.5	3504.4
ON2''	33.3	3832.6	3833.7
ON3	24.1	3520.5	3520.6
ON4	23.3	3480.5	3481.3
ON7	26.8	3452.4	3452.6
ON9	32.7	3633.5	3634.2

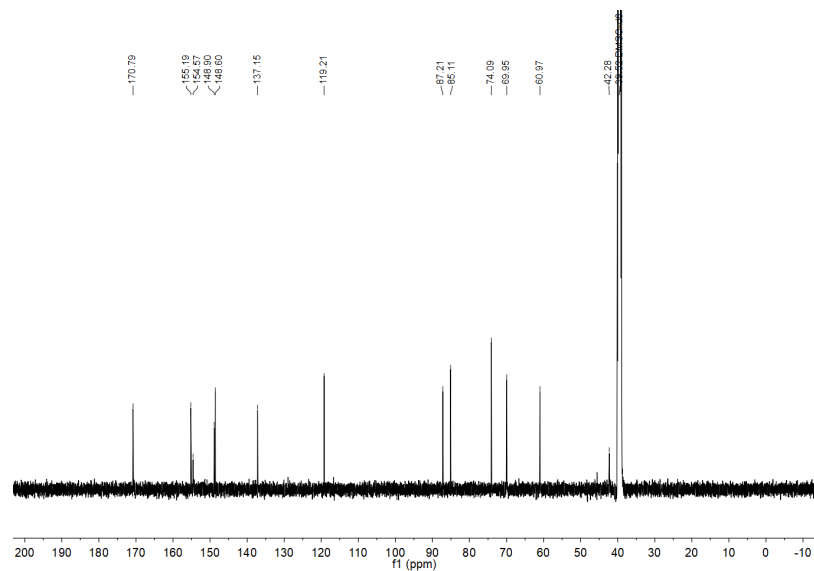
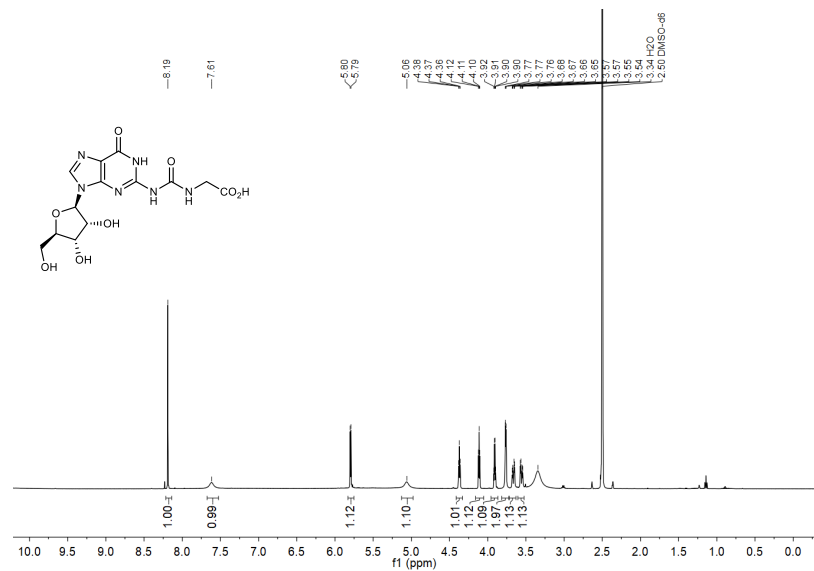
¹H and ¹³C(¹H) NMR spectra of compound 3



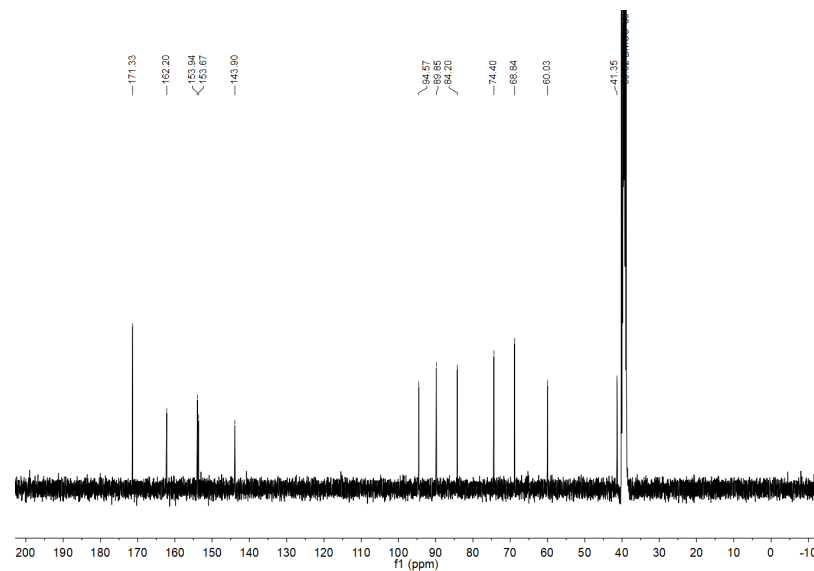
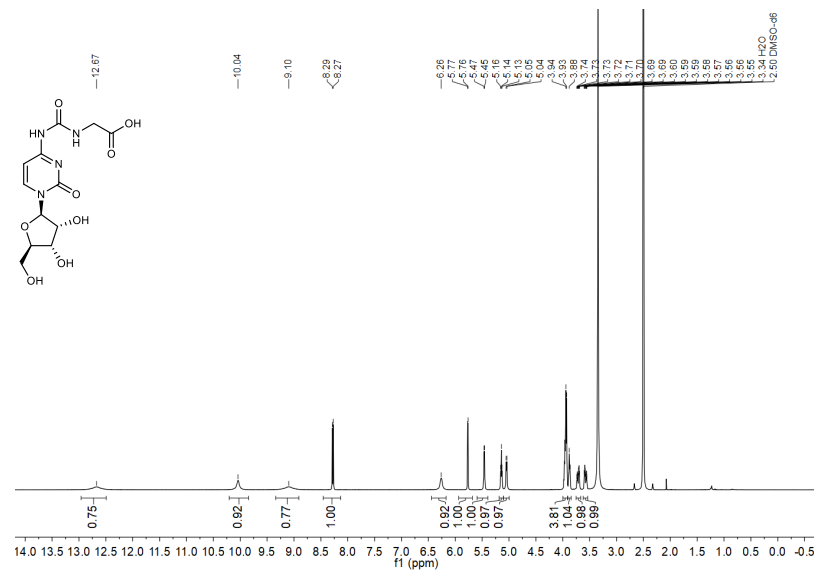
¹H and ¹³C(¹H) NMR spectra of compound 4



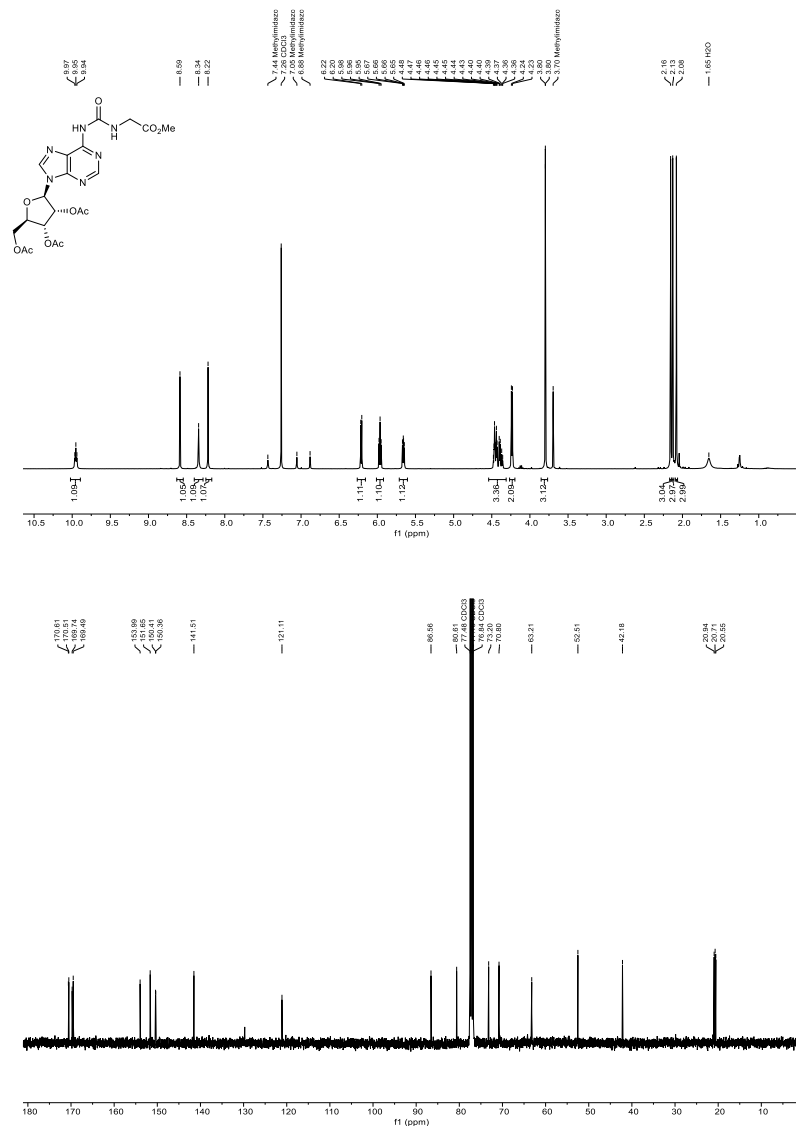
¹H and ¹³C(¹H) NMR spectra of compound 5



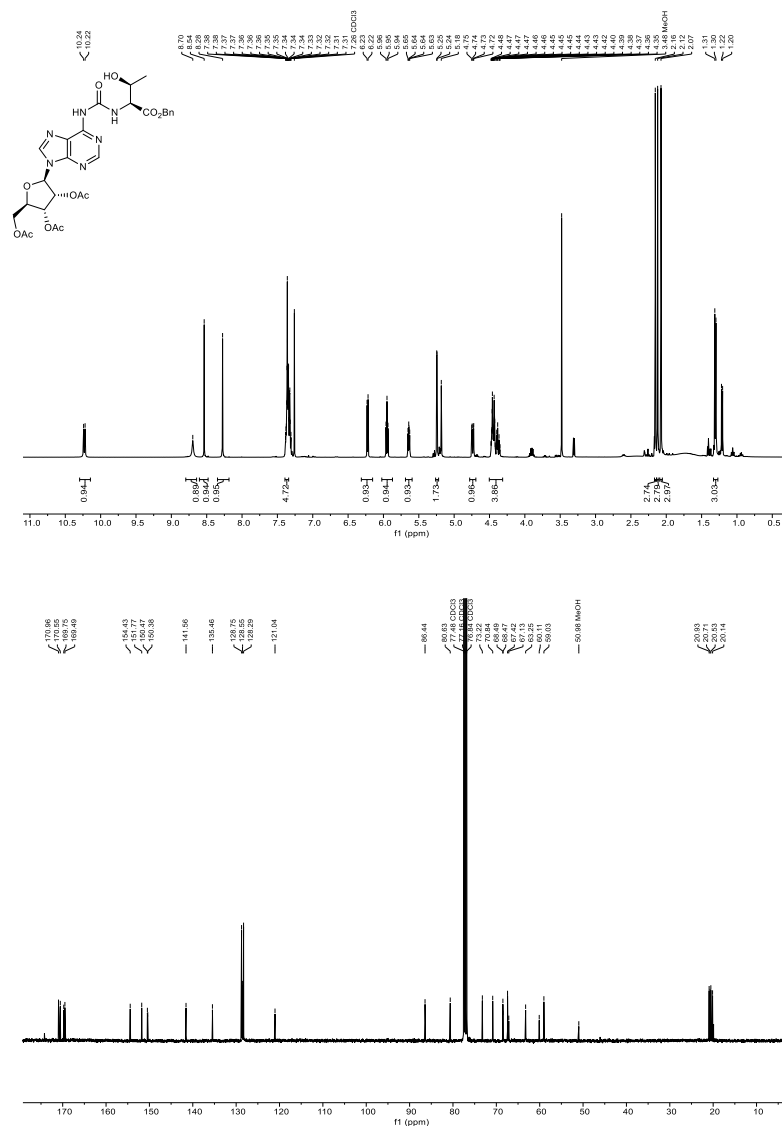
¹H and ¹³C(¹H) NMR spectra of compound 6



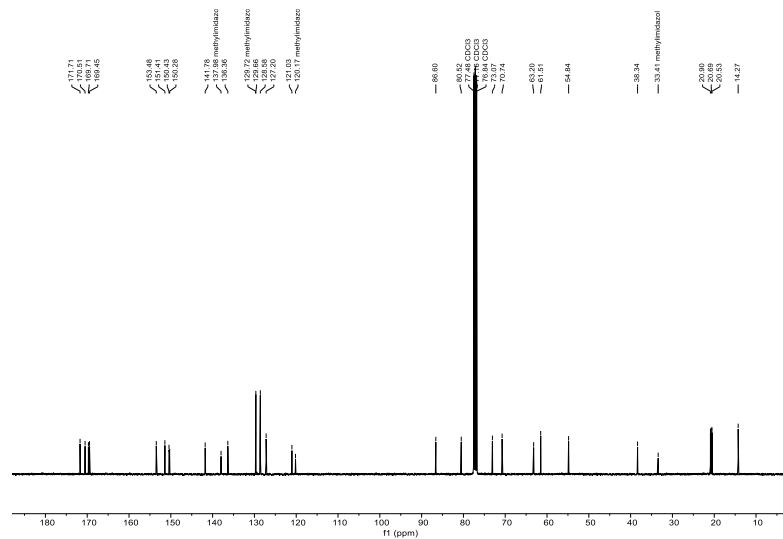
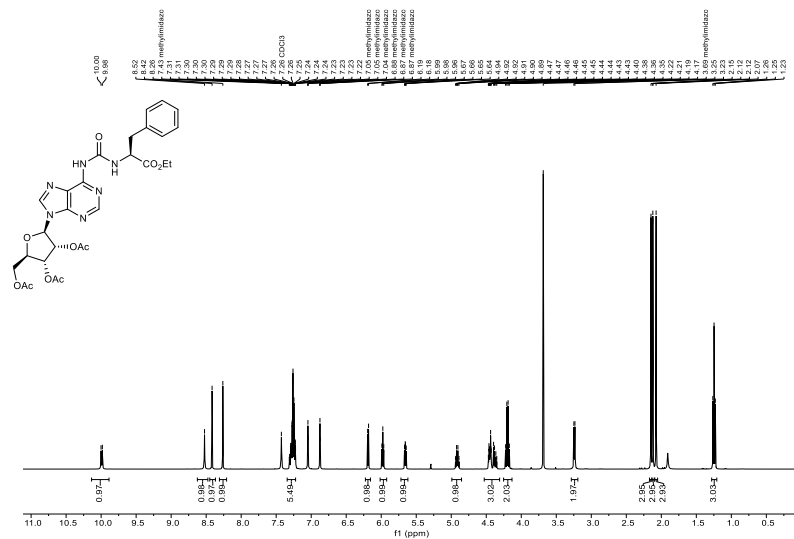
¹H and ¹³C{¹H} NMR spectra of compound 13a



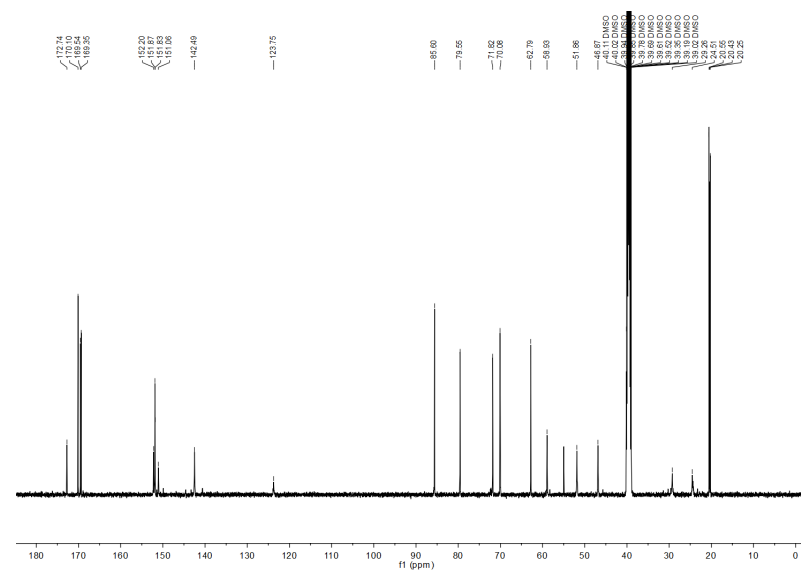
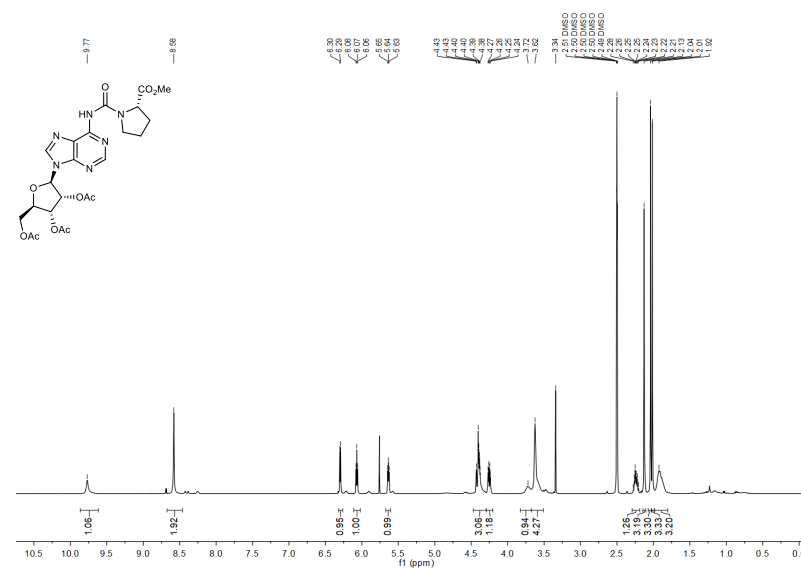
¹H and ¹³C{¹H} NMR spectra of compound 13b



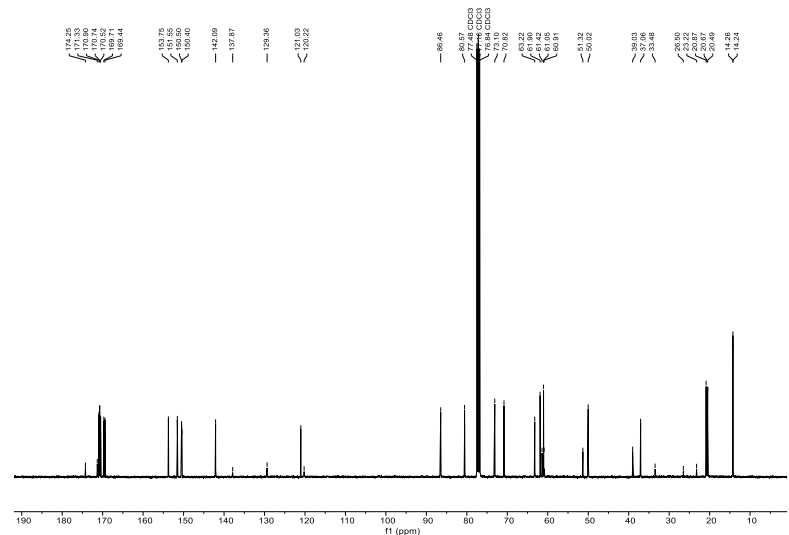
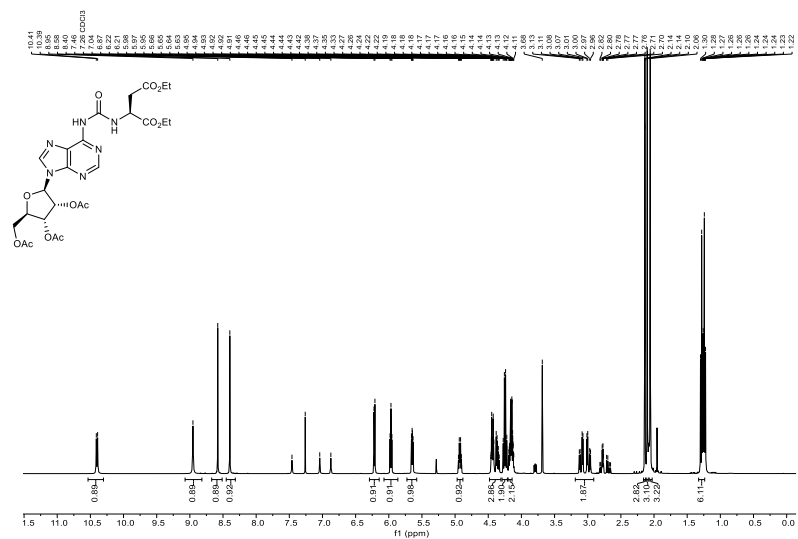
¹H and ¹³C{¹H} NMR spectra of compound 13c



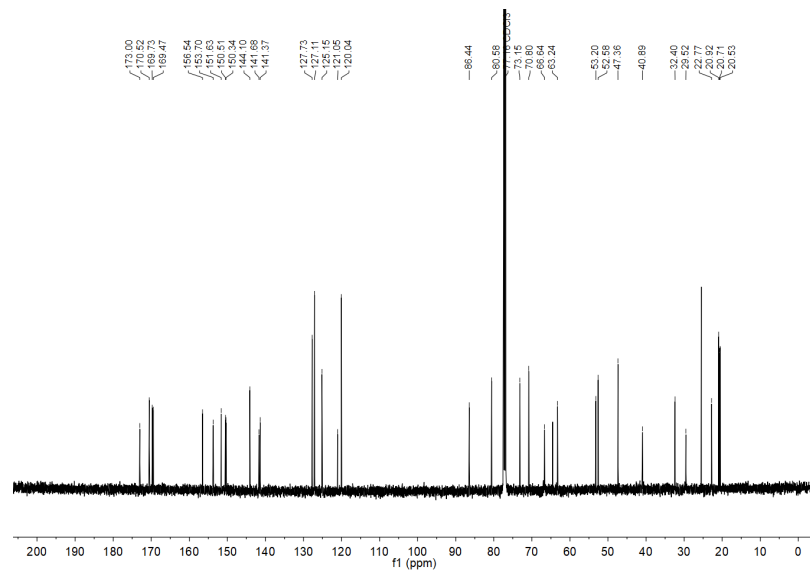
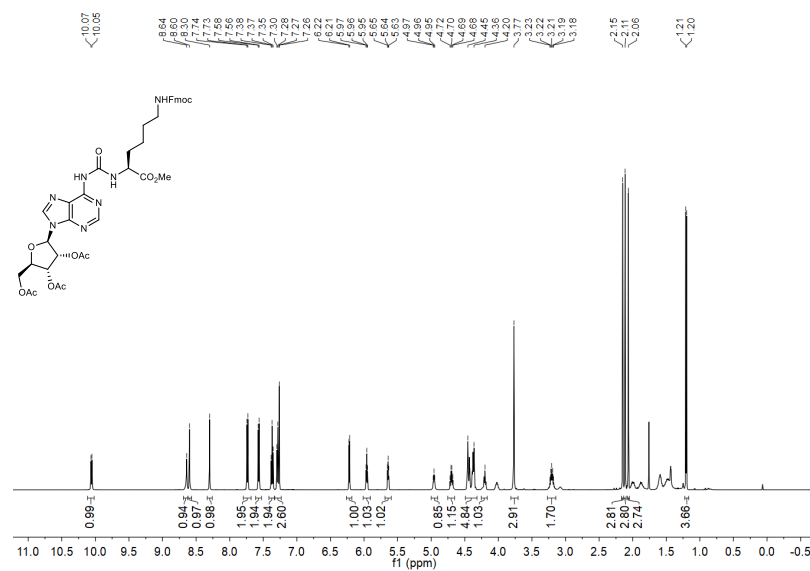
¹H and ¹³C{¹H} NMR spectra of compound 13d



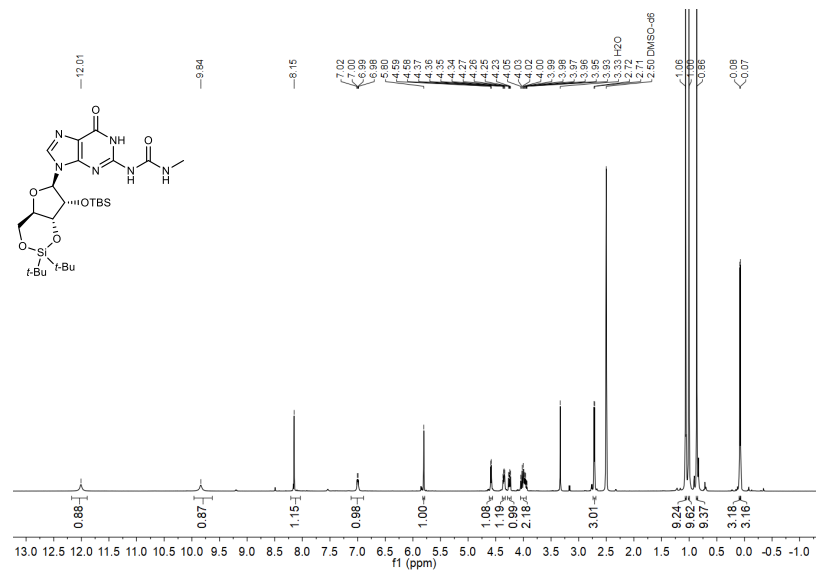
¹H and ¹³C(¹H) NMR spectra of compound 13e



¹H and ¹³C(¹H) NMR spectra of compound 13f

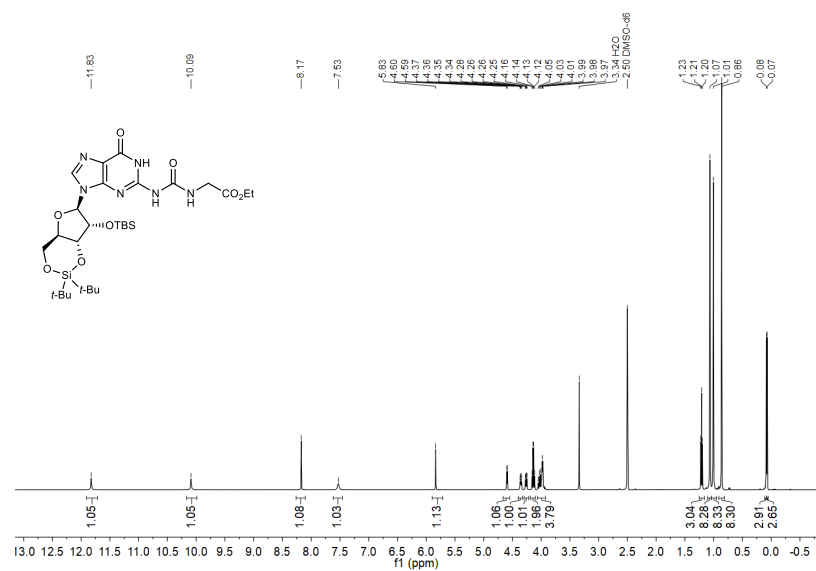


¹H and ¹³C(¹H) NMR spectra of compound 17



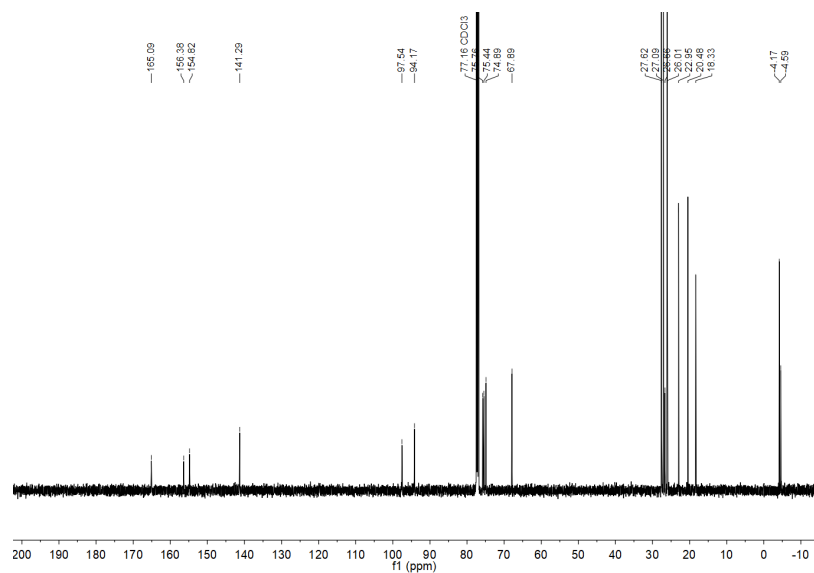
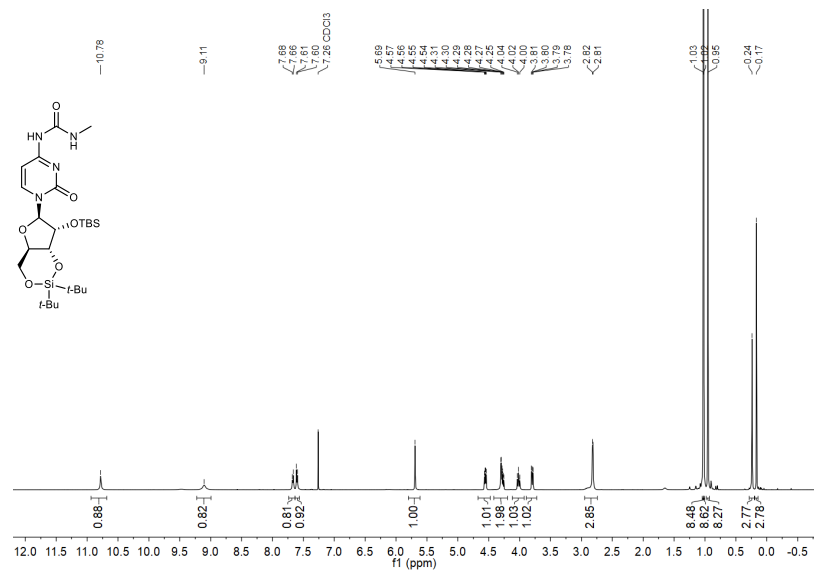
S58

¹H and ¹³C(¹H) NMR spectra of compound 18

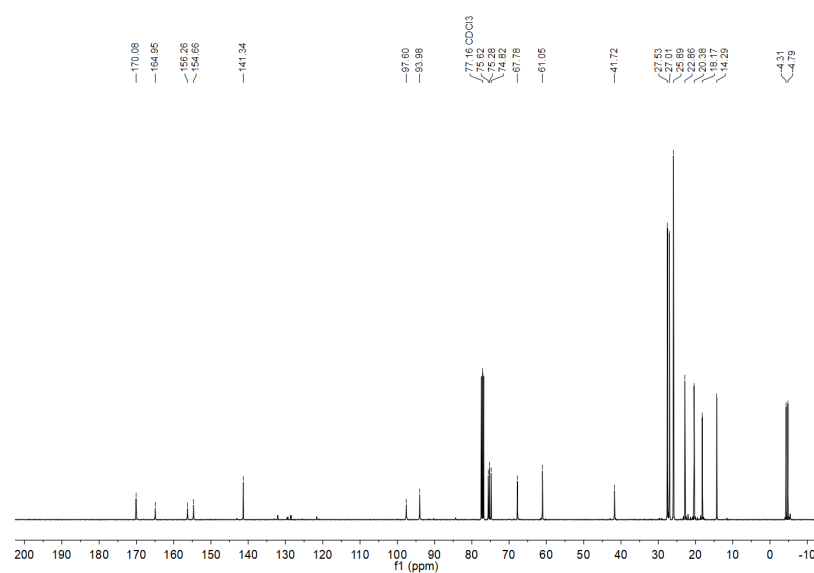
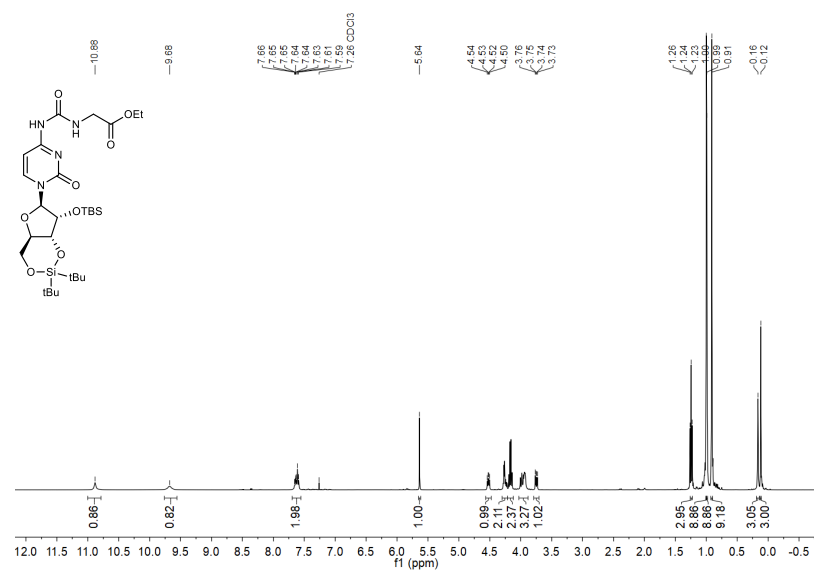


S59

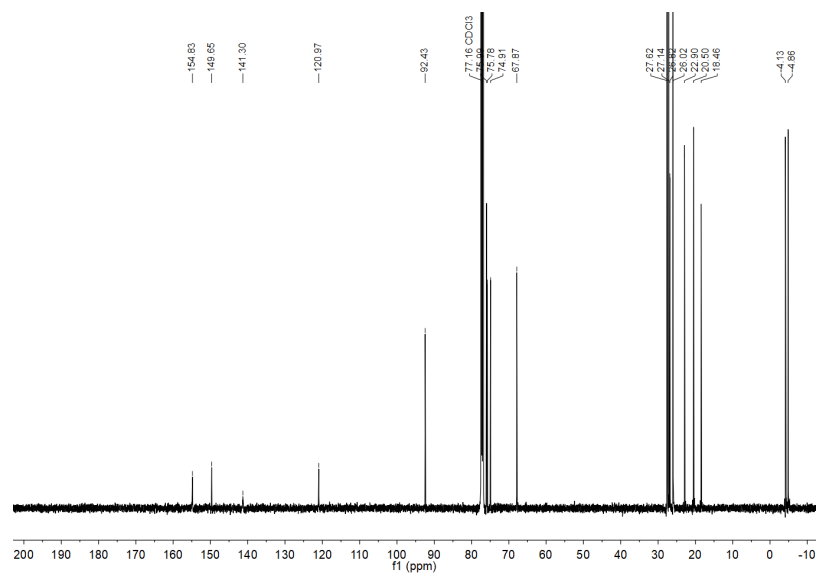
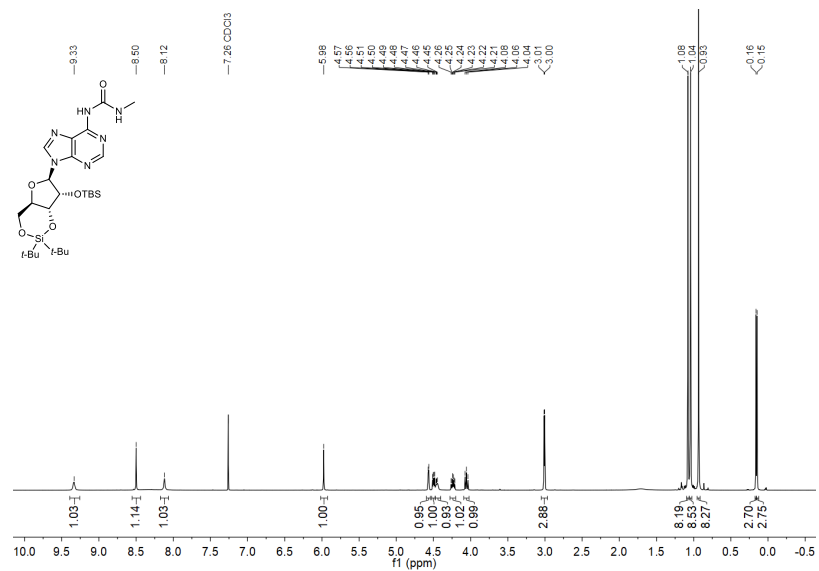
¹H and ¹³C(¹H) NMR spectra of compound 20



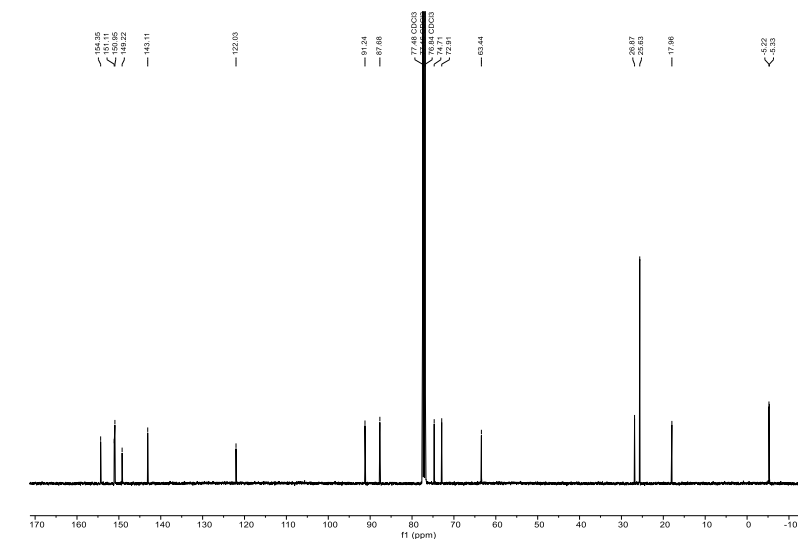
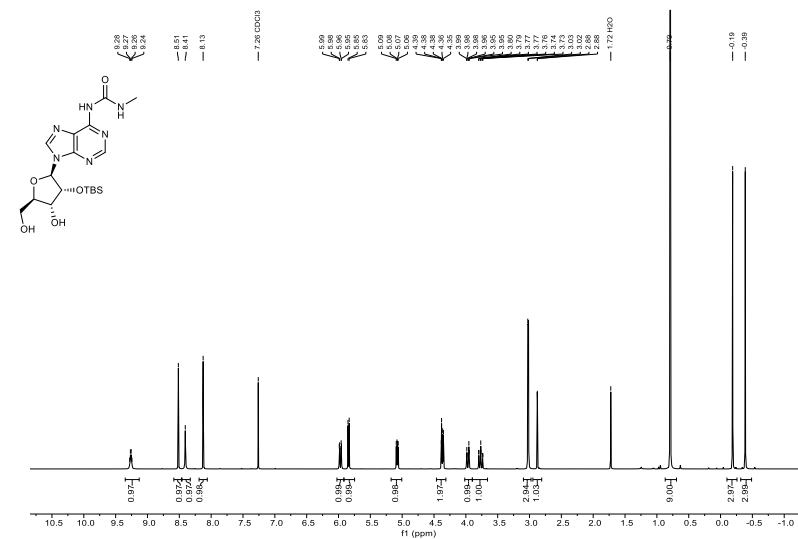
¹H and ¹³C(¹H) NMR spectra of compound 21



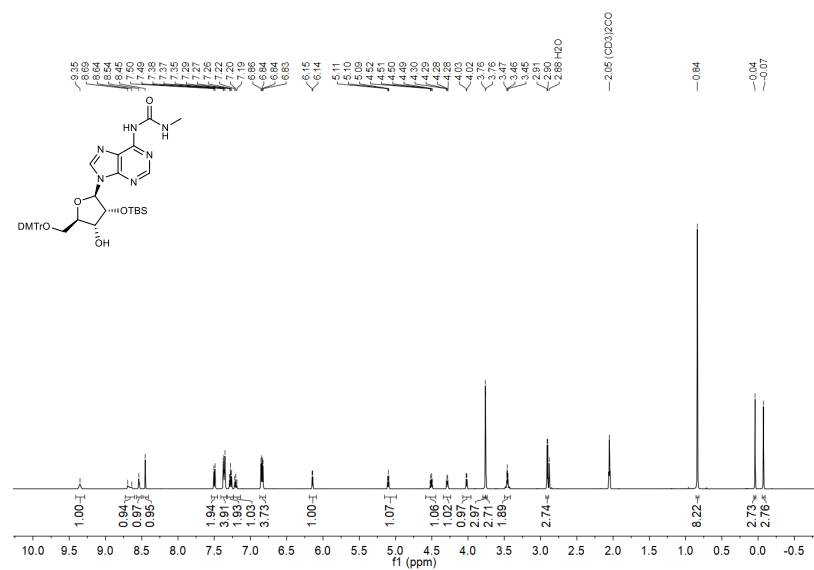
¹H and ¹³C(¹H) NMR spectra of compound 23



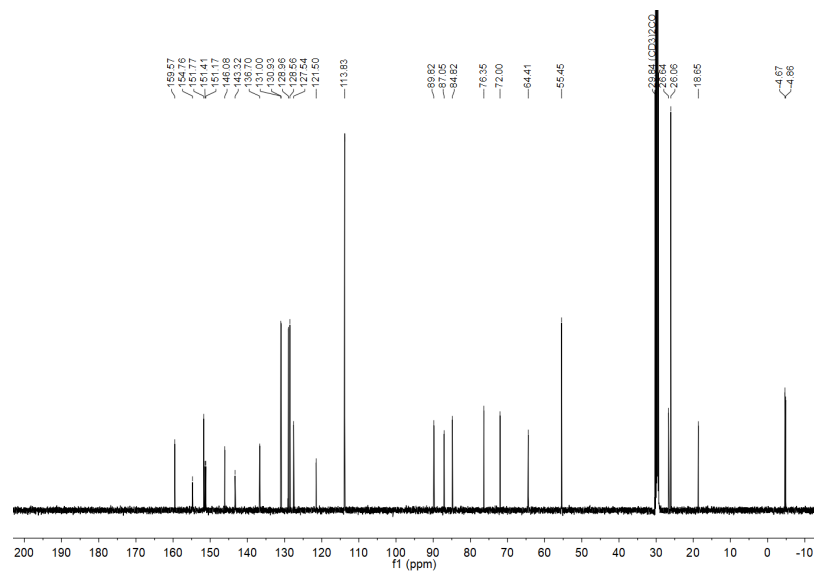
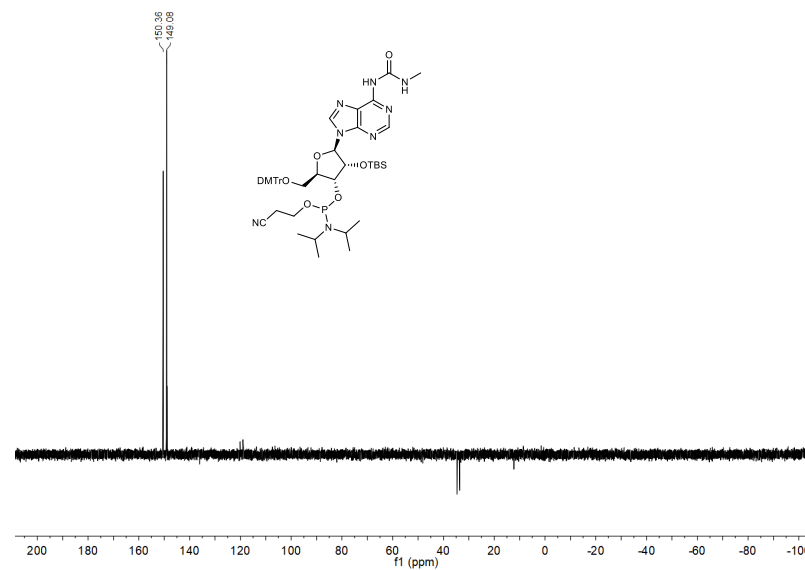
¹H and ¹³C(¹H) NMR spectra of compound 24



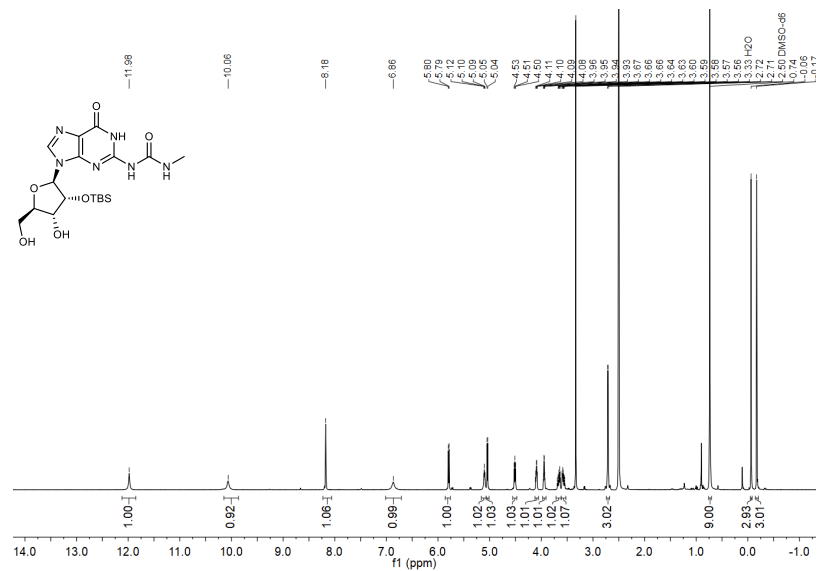
¹H and ¹³C(¹H) NMR spectra of compound 25



³¹P{¹H} NMR spectra of compound 26

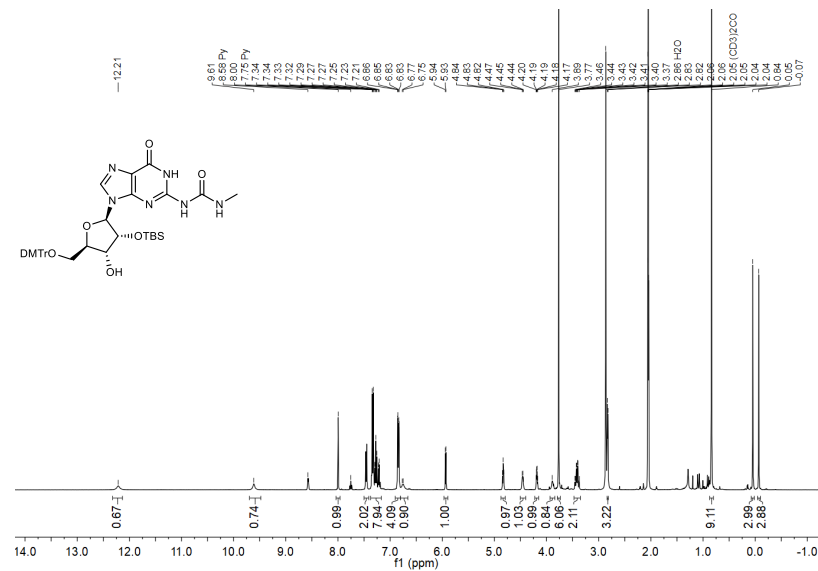


¹H and ¹³C{¹H} NMR spectra of compound 27



S66

¹H and ¹³C{¹H} NMR spectra of compound 28



S67

14. References

- ¹ G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176-2179.
- ² G. W. Breton, M. Turlington, *Tetrahedron Lett.* **2014**, *55*, 4661-4663.
- ³ F. Himmelsbach, B. S. Schulz, T. Trichtinger, R. Charubala, W. Pfeleiderer, *Tetrahedron* **1984**, *40*, 59-72.
- ⁴ F. Ferreira, F. Morvan, *Nucleosides, Nucleotides, and Nucleic Acids* **2005**, *24*, 1009-1013.
- ⁵ S. G. J. Mochrie, *Am. J. Phys.* **2011**, *79*, 1121-1126.
- ⁶ C. Schneider, S. Becker, H. Okamura, A. Crisp, T. Amatov, M. Stadlmeier, T. Carell, *Angew. Chem. Int. Ed.* **2018**, *57*, 5943-5946.
- ⁷ F. Müller, L. Escobar, F. Xu, E. Węgrzyn, M. Nainytė, T. Amatov, C. Y. Chan, A. Pichler, T. Carell, *Nature* **2022**, *605*, 279-284.
- ⁸ V. Serebryany, L. Beigelman, *Tetrahedron Lett.* **2002**, *43*, 1983-1985.

Appendix III

Supporting Information

RNA-Templated Peptide Bond Formation Promotes L-Homochirality

E. Węgrzyn, I. Mejdrová, F. M. Müller, M. Nainytė, L. Escobar, T. Carell**

RNA-Templated Peptide Bond Formation Promotes L-Homochirality

Ewa Węgrzyn,^{*,[a]} Ivana Mejdrová,^{*,[a]} Felix M. Müller,^[a] Milda Nainytė,^[a] Luis Escobar,^{*,[a]} and Thomas Carell^{*,[a]}

^[a] Department of Chemistry, Institute for Chemical Epigenetics (ICE-M), Ludwig-Maximilians-Universität (LMU) München, Butenandtstrasse 5-13, 81377 Munich (Germany)

* These authors contributed equally.

E-Mail: luisescobar1992@hotmail.es and thomas.carell@lmu.de

Supporting Information

Table of Contents

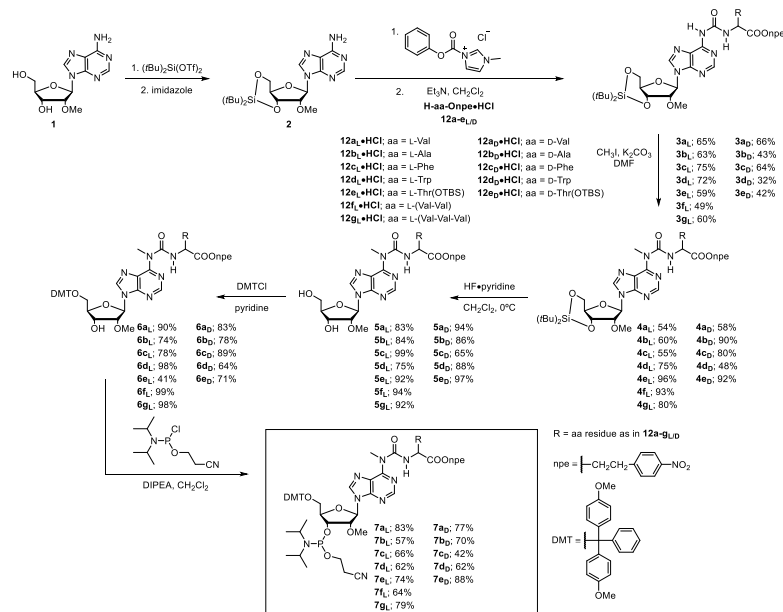
1. General information and instruments for nucleosides and phosphoramidites	S2
2. Synthesis and characterization data of phosphoramidites	S2
2.1 Amino acid-modified methyl <i>N</i> ⁶ -carbamoyl adenosine phosphoramidites.....	S2
2.2 Npe-protected amino acids	S57
2.3 5-Methylaminomethyl uridine phosphoramidite	S75
3. General information and instruments for oligonucleotides	S75
4. Synthesized oligonucleotides using a DNA/RNA automated synthesizer	S76
4.1 Donor strands containing an amino acid-modified methyl <i>N</i> ⁶ -carbamoyl adenosine at the 5'-end ...	S76
4.2 Acceptor strands containing a 5-methylaminomethyl uridine at the 3'-end	S77
5. Peptide coupling reactions between donor and acceptor oligonucleotides	S77
6. Calibration curve	S79
7. Pair-wise competitive peptide coupling reactions between donor and acceptor oligonucleotides	S79
8. Pair-wise competitive peptide coupling reactions between donor and acceptor oligonucleotides containing peptides.....	S83
9. Selected pair-wise competitive peptide coupling reactions at low temperature.....	S85
10. Kinetic experiments with different activators.....	S86
11. Melting curves	S88
12. Loading reaction of a selected amino acid onto an RNA strand	S88
13. References.....	S89

1. General information and instruments for nucleosides and phosphoramidites

Reagents were purchased from commercial suppliers and used without further purification unless otherwise stated. Anhydrous solvents, stored under inert atmosphere, were also purchased. All reactions involving air/moisture sensitive reagents/intermediates were performed under inert atmosphere using oven-dried glassware. Routine ^1H NMR, $^{13}\text{C}\{^1\text{H}\}$ NMR and $^{31}\text{P}\{^1\text{H}\}$ NMR spectra were recorded on a Bruker Ascend 400 spectrometer (400 MHz for ^1H NMR, 100 MHz for ^{13}C NMR and 162 MHz for ^{31}P NMR), Bruker Ascend 500 spectrometer (500 MHz for ^1H NMR, 125 MHz for ^{13}C NMR and 202 MHz for ^{31}P NMR) or Bruker ARX 600 spectrometer (600 MHz for ^1H NMR, 150 MHz for ^{13}C NMR and 243 MHz for ^{31}P NMR). Deuterated solvents used are indicated in the characterization and chemical shifts (δ) are reported in ppm. Residual solvent peaks were used as reference.¹ All NMR J values are given in Hz. COSY, HMQC and HMBC NMR experiments were recorded to help with the assignment of ^1H and ^{13}C signals. NMR spectra were analyzed using MestReNova software version 10.0. High Resolution Mass Spectra (HRMS) were measured on a Thermo Finnigan LTQ-FT with ESI as ionization mode. IR spectra were recorded on a Perkin-Elmer Spectrum BX II FT-IR instrument or Shimadzu IRSpirit FT-IR instrument. Both equipped with an ATR accessory. Column chromatography was performed with technical grade silica gel, 40-63 μm particle size. Reaction progress was monitored by Thin Layer Chromatography (TLC) analysis on silica gel 60 F254 and stained with *para*-anisaldehyde, potassium permanganate or cerium ammonium molybdate solution.

2. Synthesis and characterization data of phosphoramidites

2.1 Amino acid-modified methyl *N*⁶-carbamoyl adenosine phosphoramidites

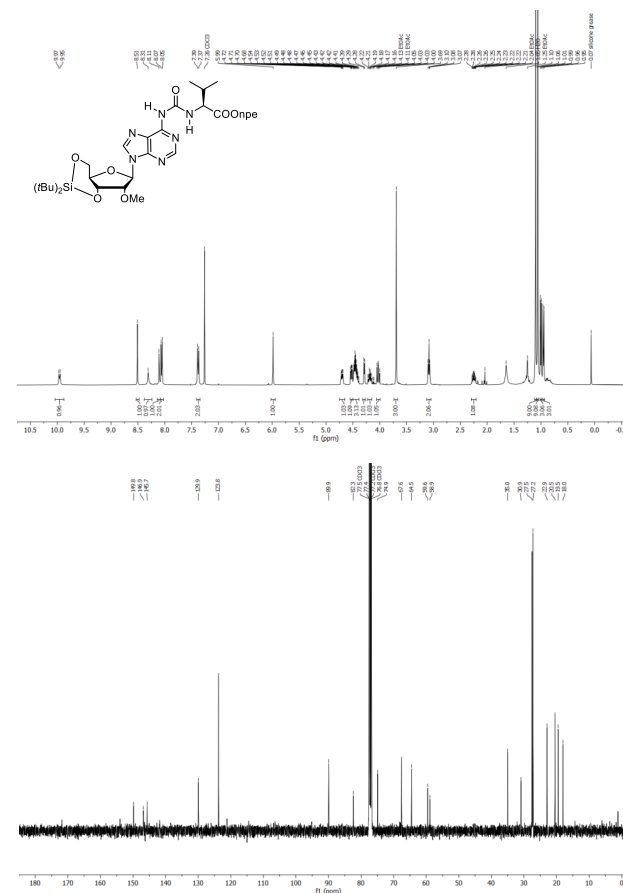


Scheme S1. Synthesis of amino acid-modified methyl *N*⁶-carbamoyl adenosine phosphoramidites.

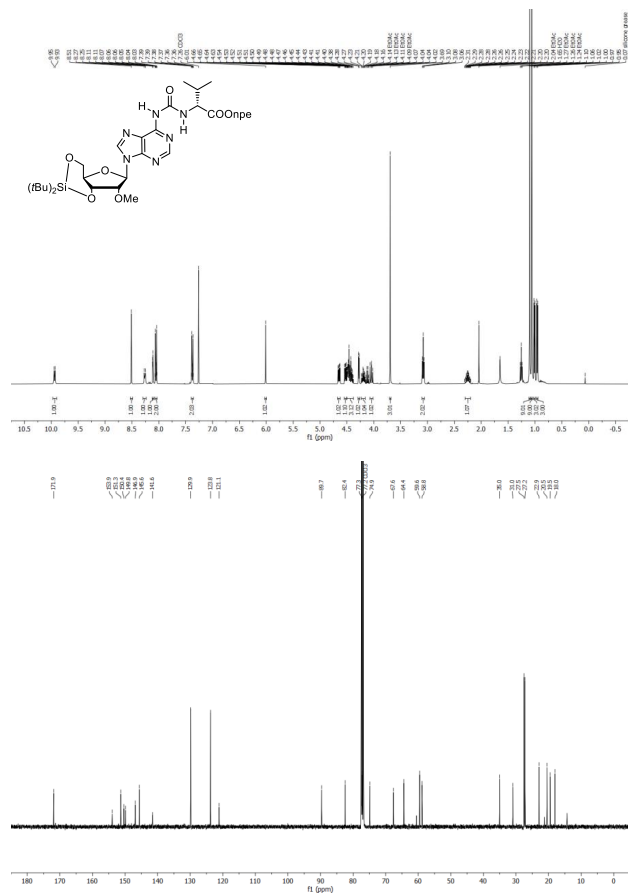
Compound **2** was synthesized following a procedure previously reported in the literature.²

General procedure for the synthesis of compound 3: Step 1. Compound **2** (1 equiv.) and *N*-methyl-3-phenoxy-carbonyl-imidazolium chloride (2 equiv.) were added to an oven dried round-bottom flask and kept under high-vacuum for 15 min. After that, dry CH_2Cl_2 was added under nitrogen atmosphere and the reaction was stirred at r.t. for 5 h. Step 2. Onpe-protected amino acid **12-HCl** (2 equiv.) was added to an oven dried round-bottom flask and suspended in dry CH_2Cl_2 , followed by the addition of Et_3N (2 equiv.). The suspension was added dropwise to the reaction mixture. The reaction was stirred at r.t. under nitrogen atmosphere for two days. After that, the reaction was quenched with aqueous saturated NaHCO_3 . The organic layer was separated and the crude was further extracted with CH_2Cl_2 . The combined organic layers were dried (Na_2SO_4), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

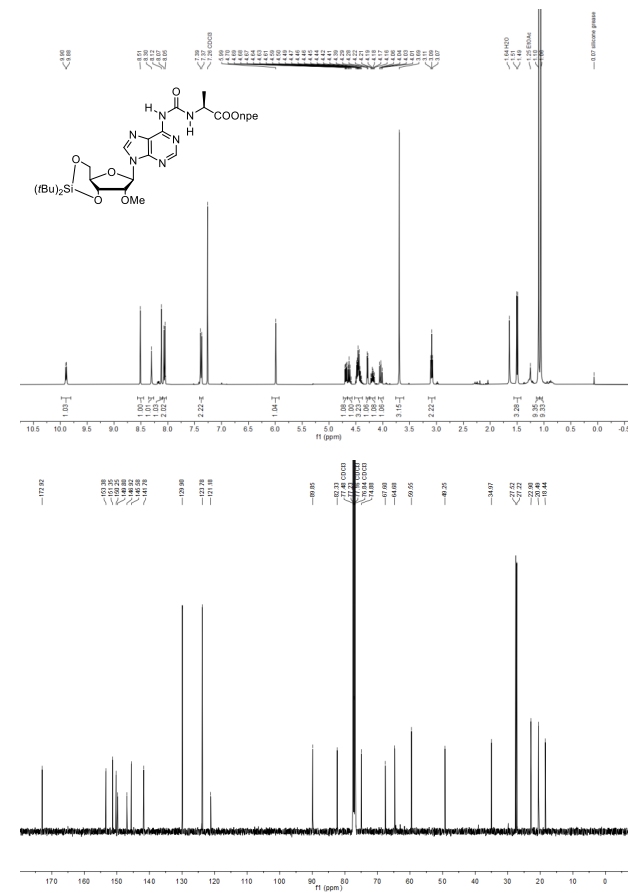
3a_L: Yield = 65%. R_f = 0.25 (2:8 iHex/EtOAc). ^1H NMR (500 MHz with cryoprobe, CDCl_3 , 298 K): δ (ppm) = 9.96 (d, J = 8.3 Hz, 1H); 8.51 (s, 1H); 8.31 (s, 1H); 8.11 (s, 1H); 8.06 (d, J = 8.4 Hz, 2H); 7.38 (d, J = 8.3 Hz, 2H); 5.99 (s, 1H); 4.70 (dd, J = 9.6, J = 4.6 Hz, 1H); 4.53 (dd, J = 8.3, J = 4.8 Hz, 1H); 4.49 – 4.38 (m, 3H); 4.29 (d, J = 4.6 Hz, 1H); 4.19 (td, J = 10.1, J = 4.9 Hz, 1H); 4.03 (dd, J = 10.5, J = 9.1 Hz, 1H); 3.69 (s, 3H); 3.08 (t, J = 6.5 Hz, 2H); 2.31 – 2.20 (m, 1H); 1.10 (s, 9H); 1.06 (s, 9H); 1.06 (d, J = 6.8 Hz, 3H); 0.95 (d, J = 6.8 Hz, 3H). $^{13}\text{C}\{^1\text{H}\}$ NMR (125 MHz with cryoprobe, CDCl_3 , 298 K): δ (ppm) = 149.8; 146.9; 145.7; 129.9; 123.8; 89.9; 82.3; 77.4; 74.9; 67.6; 64.5; 59.6; 58.9; 35.0; 30.9; 27.5; 27.2; 22.9; 20.5; 19.5; 18.0. FTIR ν_{max} (cm^{-1}): 3230 (w); 1739 (m); 1698 (s); 1519 (s); 1467 (m); 1344 (s); 1250 (m); 1138 (s); 1063 (s); 826 (s). HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{33}\text{H}_{46}\text{O}_9\text{N}_7\text{Si}$ 714.3277; Found 714.3278.



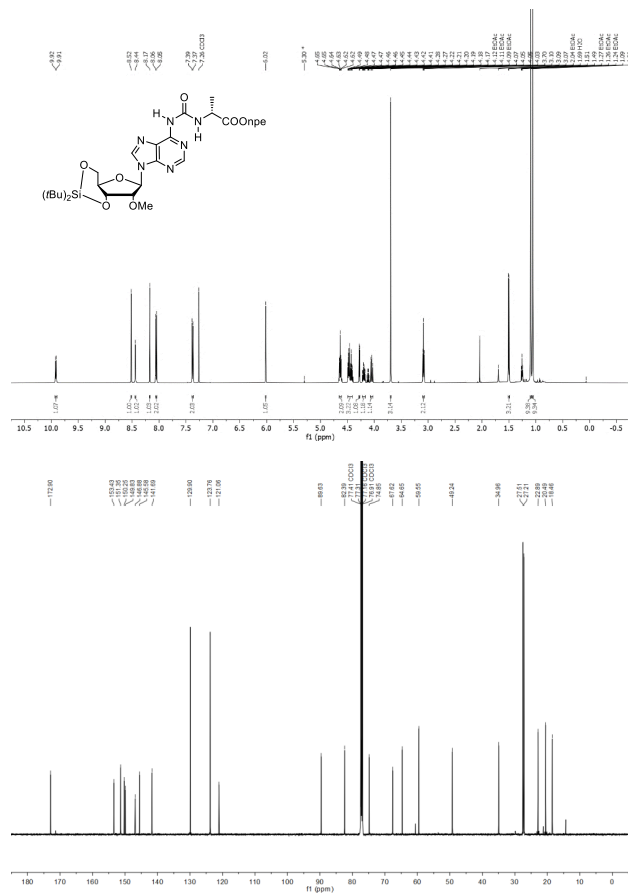
3a_b: Yield = 66%. Rf = 0.38 (2:3 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.94 (d, *J* = 8.9 Hz, 1H); 8.51 (s, 1H); 8.26 (d, *J* = 7.8 Hz, 1H); 8.11 (d, *J* = 2.5 Hz, 1H); 8.09 – 8.01 (m, 2H); 7.41 – 7.34 (m, 2H); 6.01 (s, 1H); 4.65 (dd, *J* = 9.6 Hz, *J* = 4.6 Hz, 1H); 4.53 (dd, *J* = 8.4 Hz, *J* = 4.9 Hz, 1H); 4.52 – 4.36 (m, 3H); 4.28 (d, *J* = 4.6 Hz, 1H); 4.20 (td, *J* = 10.1 Hz, *J* = 4.9 Hz, 1H); 4.04 (dd, *J* = 10.5 Hz, *J* = 9.1 Hz, 1H); 3.69 (s, 3H); 3.08 (t, *J* = 6.5 Hz, 2H); 2.31 – 2.20 (m, 1H); 1.10 (s, 9H); 1.06 (s, 9H); 1.01 (d, *J* = 6.8 Hz, 3H); 0.96 (d, *J* = 6.8 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 153.9; 151.3; 150.4; 149.8; 146.9; 145.6; 141.6; 129.9; 123.8; 121.1; 89.7; 82.4; 77.3; 74.9; 67.6; 64.4; 59.6; 58.8; 35.0; 31.0; 27.5; 27.2; 22.9; 20.5; 19.5; 18.0. FTIR *v*_{max} (cm⁻¹): 3242 (w); 1746 (m); 1700 (s); 1516 (s); 1467 (m); 1345 (s); 1256 (m); 1142 (s); 1062 (s); 826 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₃H₄₈O₉N₇Si 714.3277; Found 714.3273.



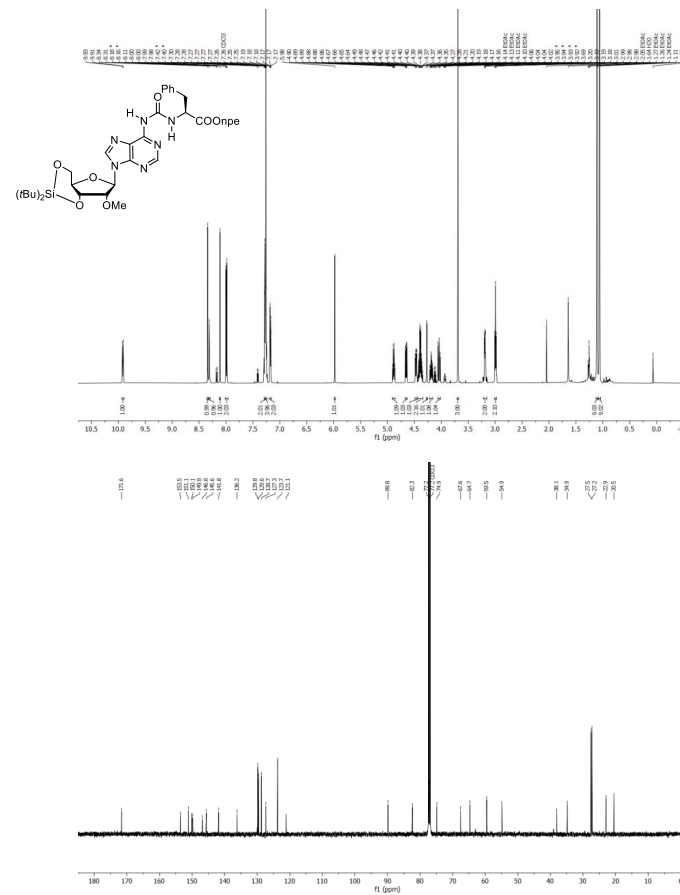
3b_c: Yield = 63%. Rf = 0.20 (85:15 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.89 (d, *J* = 7.3 Hz, 1H); 8.51 (s, 1H); 8.30 (s, 1H); 8.12 (s, 1H); 8.07 – 8.05 (m, 2H); 7.39 – 7.37 (m, 2H); 5.99 (s, 1H); 4.68 (dd, *J* = 9.6 Hz, *J* = 4.6 Hz, 1H); 4.66 – 4.59 (m, 1H); 4.48 – 4.41 (m, 3H); 4.28 (d, *J* = 4.6 Hz, 1H); 4.19 (ddd, *J* = 9.6 Hz, *J* = 9.6 Hz, *J* = 4.6 Hz, 1H); 4.03 (dd, *J* = 9.6 Hz, *J* = 9.6 Hz, 1H); 3.69 (s, 3H); 3.09 (t, *J* = 6.5 Hz, 2H); 1.50 (d, *J* = 7.3 Hz, 3H); 1.10 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.9; 153.4; 151.3; 150.2; 149.8; 146.9; 145.6; 141.8; 129.9; 123.8; 121.2; 89.9; 82.3; 77.2; 74.9; 67.6; 64.7; 59.5; 49.2; 35.0; 27.5; 27.2; 22.9; 20.5; 18.4. FTIR *v*_{max} (cm⁻¹): 2931 (w); 1742 (m); 1699 (s); 1519 (s); 1467 (m); 1344 (s); 1253 (m); 1137 (s); 1063 (s); 826 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₁H₄₄O₉N₇Si 686.2964; Found 686.2964.



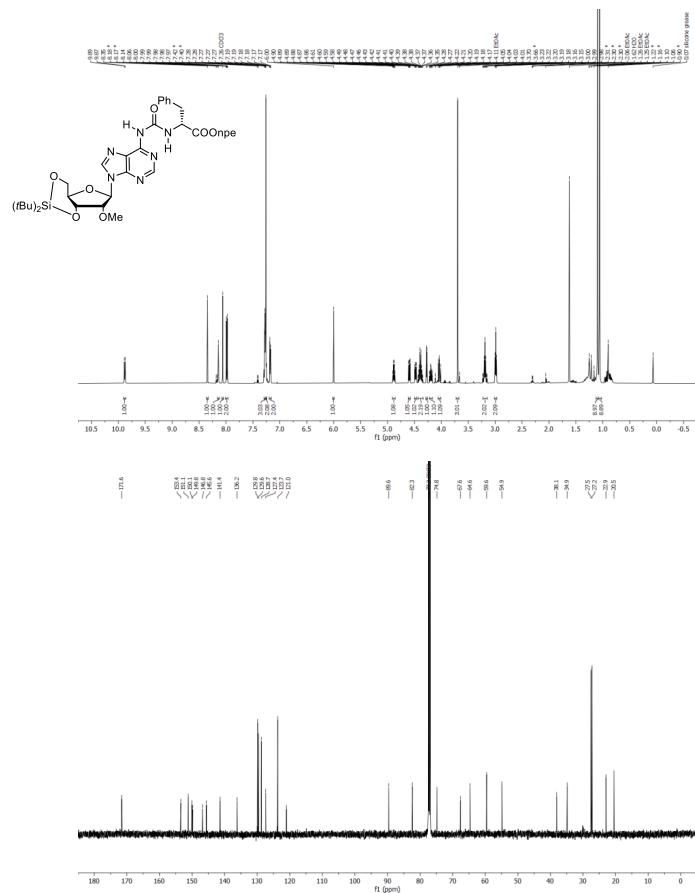
3b_o: Yield = 43% yield. R_f = 0.32 (1:1 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.92 (d, *J* = 7.3 Hz, 1H); 8.52 (s, 1H); 8.44 (s, 1H); 8.17 (s, 1H); 8.06 – 8.05 (m, 2H); 7.39 – 7.37 (m, 2H); 6.02 (s, 1H); 4.65 – 4.62 (m, 2H); 4.50 – 4.40 (m, 3H); 4.28 (d, *J* = 4.6 Hz, 1H); 4.20 (ddd, *J* = 10.3 Hz, *J* = 4.9 Hz, 1H); 4.05 (dd, *J* = 10.3 Hz, *J* = 10.3 Hz, 1H); 3.70 (s, 3H); 3.09 (t, *J* = 6.5 Hz, 2H); 1.50 (d, *J* = 7.2 Hz, 3H); 1.10 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.9; 153.4; 151.4; 150.2; 149.8; 146.9; 145.6; 141.7; 129.9; 123.8; 121.1; 89.6; 82.4; 77.3; 74.8; 67.6; 64.6; 59.6; 49.2; 35.0; 27.5; 27.2; 22.9; 20.5; 18.5. FTIR ν_{max} (cm⁻¹): 2936 (w); 2860 (w); 1743 (m); 1699 (s); 1612 (m); 1519 (s); 1469 (m); 1344 (s); 1259 (m); 1059 (s); 828 (s); 652 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₇H₄₄O₉N₇Si 686.2964; Found 686.2971.



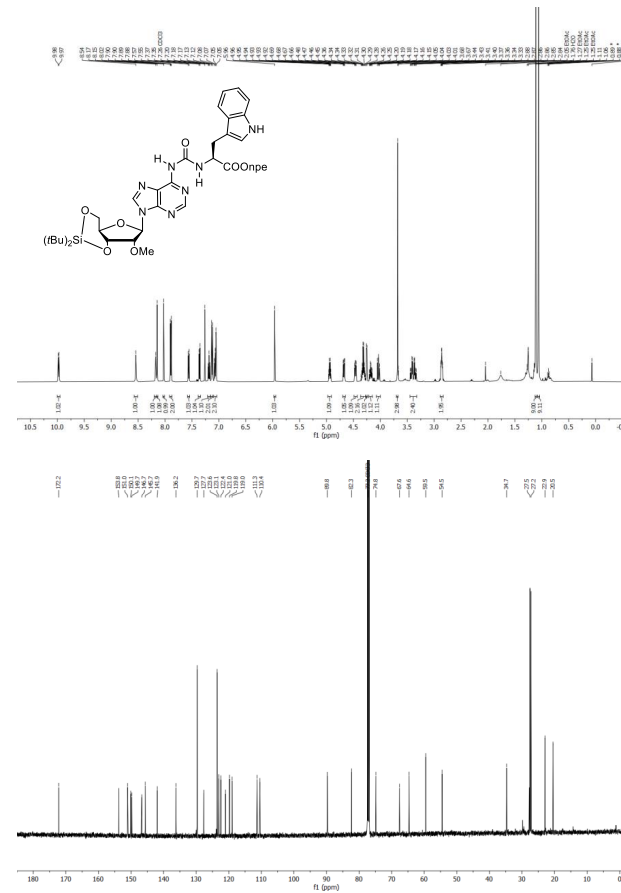
3c_l: Yield = 75%. R_f = 0.31 (2:8 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.92 (d, *J* = 7.6 Hz, 1H); 8.34 (s, 1H); 8.31 (s, 1H); 8.11 (s, 1H); 8.01 – 7.95 (m, 2H); 7.28 (d, *J* = 2.2 Hz, 2H); 7.26 (d, *J* = 1.2 Hz, 3H); 7.20 – 7.16 (m, 2H); 5.98 (s, 1H); 4.88 (dt, *J* = 7.7 Hz, *J* = 6.2 Hz, 1H); 4.65 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.47 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 1H); 4.43 – 4.35 (m, 2H); 4.27 (d, *J* = 4.6 Hz, 1H); 4.19 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.04 (dd, *J* = 10.5 Hz, *J* = 9.2 Hz, 1H); 3.69 (s, 3H); 3.19 (dd, *J* = 6.2 Hz, *J* = 3.2 Hz, 2H); 2.99 (t, *J* = 6.5 Hz, 2H); 1.11 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.6; 153.5; 151.1; 150.1; 149.8; 146.8; 145.6; 141.8; 136.2; 129.8; 129.6; 128.7; 127.3; 123.7; 121.1; 89.8; 82.3; 77.2; 74.9; 67.6; 64.7; 59.5; 54.9; 38.1; 34.9; 27.5; 27.2; 22.9; 20.5. FTIR ν_{max} (cm⁻¹): 2934 (w); 1740 (m); 1699 (s); 1519 (s); 1469 (m); 1344 (s); 1259 (m); 1138 (s); 1062 (s); 828 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₇H₄₆O₉N₇Si 762.3277; Found 762.3280.



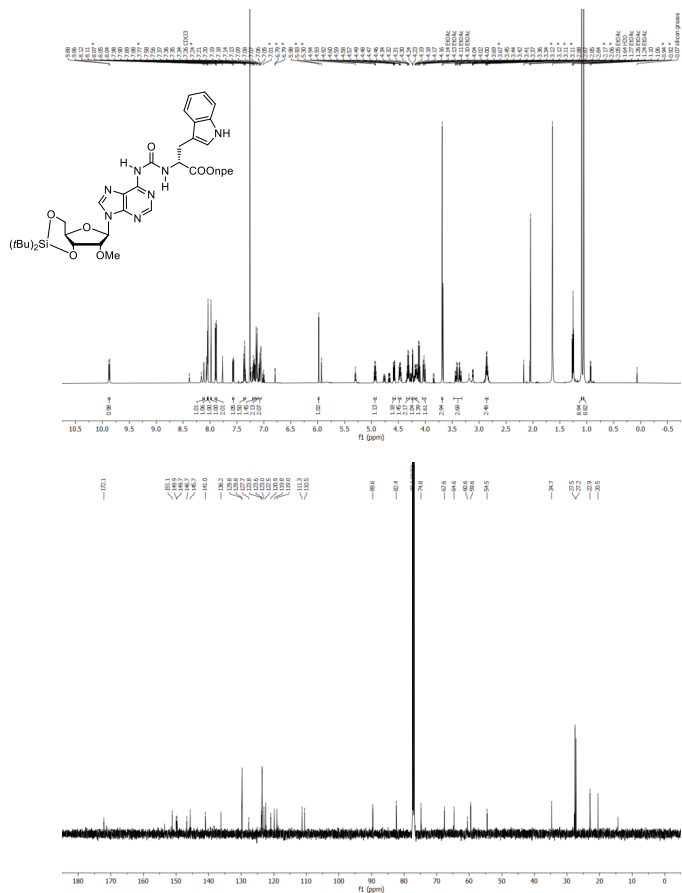
3c_b: Yield = 64%. R_f = 0.33 (3:7 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.88 (d, J = 7.7 Hz, 1H); 8.35 (s, 1H); 8.14 (s, 1H); 8.06 (s, 1H); 8.02 – 7.95 (m, 2H); 7.31 – 7.23 (m, 5H); 7.21 – 7.15 (m, 2H); 6.00 (s, 1H); 4.88 (dt, J = 7.7 Hz, J = 6.2 Hz, 1H); 4.59 (dd, J = 9.6 Hz, J = 4.6 Hz, 1H); 4.48 (dd, J = 9.2 Hz, J = 5.0 Hz, 1H); 4.45 – 4.33 (m, 2H); 4.27 (d, J = 4.6 Hz, 1H); 4.19 (td, J = 10.1 Hz, J = 5.0 Hz, 1H); 4.03 (dd, J = 10.5 Hz, J = 9.3 Hz, 1H); 3.70 (s, 3H); 3.25 – 3.13 (m, 2H); 2.99 (t, J = 6.5 Hz, 2H); 1.10 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.6; 153.4; 151.1; 150.1; 149.8; 146.8; 145.6; 141.4; 136.2; 129.8; 129.6; 128.7; 127.4; 123.7; 121.0; 89.6; 82.3; 74.8; 67.6; 64.6; 59.6; 54.9; 38.1; 34.9; 27.5; 27.2; 22.9; 20.5. FTIR ν_{max} (cm⁻¹): 2934 (w); 1740 (m); 1700 (s); 1519 (s); 1469 (m); 1345 (s); 1259 (m); 1140 (s); 1063 (s); 828 (s); 652 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₇H₄₈O₉N₇Si 762.3277; Found 762.3285.



3d_l: Yield = 72%. R_f = 0.45 (EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.98 (d, J = 7.6 Hz, 1H); 8.54 (s, 1H); 8.17 (s, 1H); 8.15 (s, 1H); 8.02 (s, 1H); 7.91 – 7.87 (m, 2H); 7.56 (d, J = 7.9 Hz, 1H); 7.36 (d, J = 8.1 Hz, 1H); 7.18 (t, J = 7.0 Hz, 1H); 7.13 (d, J = 8.7 Hz, 2H); 7.10 – 7.03 (m, 2H); 5.96 (s, 1H); 4.96 – 4.92 (m, 1H); 4.68 (dd, J = 9.7 Hz, J = 4.6 Hz, 1H); 4.46 (dd, J = 9.2 Hz, J = 5.0 Hz, 1H); 4.31 (tt, J = 11.2 Hz, J = 5.5 Hz, 2H); 4.25 (d, J = 4.6 Hz, 1H); 4.18 (td, J = 10.1 Hz, J = 5.0 Hz, 1H); 4.03 (dd, J = 10.6 Hz, J = 9.1 Hz, 1H); 3.68 (s, 3H); 3.45 – 3.32 (m, 2H); 2.86 (td, J = 6.4 Hz, J = 2.7 Hz, 2H); 1.11 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.2; 153.8; 151.0; 150.1; 149.7; 146.7; 145.7; 141.9; 136.2; 129.7; 127.7; 123.6; 123.1; 121.0; 119.8; 119.0; 111.3; 110.4; 89.8; 82.3; 74.8; 67.6; 64.6; 59.5; 54.5; 34.7; 27.5; 27.2; 22.9; 20.5. FTIR ν_{max} (cm⁻¹): 2934 (w); 1735 (m); 1697 (s); 1518 (s); 1472 (m); 1344 (s); 1197 (m); 1138 (s); 1062 (s); 828 (s); 739 (s); 652 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₉H₄₈O₉N₈Si 801.3386; Found 801.3396.

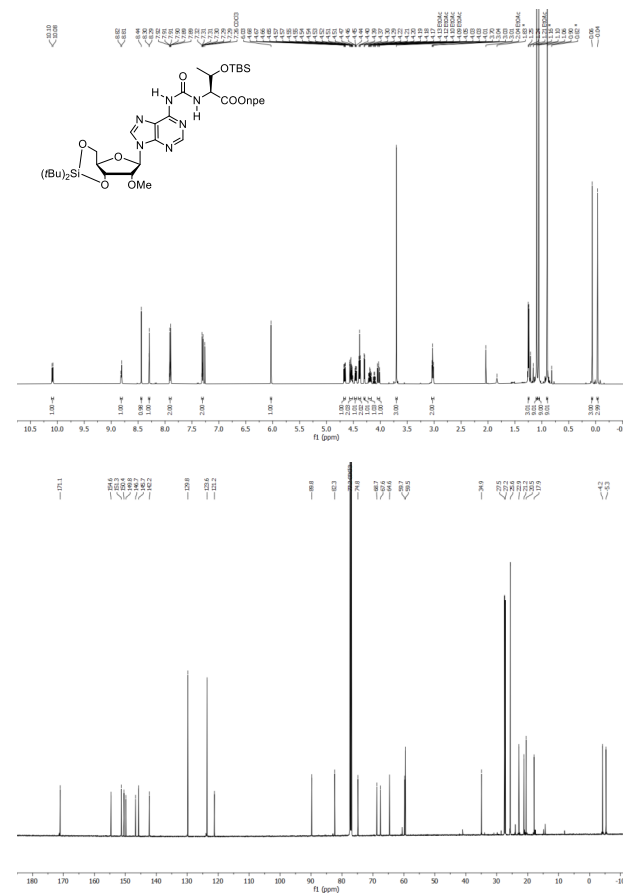


3d_b: Yield = 32%. R_f = 0.63 (EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.87 (d, *J* = 7.6 Hz, 1H); 8.12 (d, *J* = 2.8 Hz, 1H); 8.05 (s, 1H); 8.04 (s, 1H); 7.98 (s, 1H); 7.92 – 7.86 (m, 2H); 7.57 (d, *J* = 8.0 Hz, 1H); 7.39 – 7.32 (m, 1H); 7.19 (dd, *J* = 7.6 Hz, *J* = 6.5 Hz, 1H); 7.13 (d, *J* = 8.6 Hz, 2H); 7.11 – 7.03 (m, 2H); 5.98 (s, 1H); 4.97 – 4.89 (m, 1H); 4.58 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.51 – 4.43 (m, 1H); 4.36 – 4.27 (m, 2H); 4.24 (d, *J* = 4.6 Hz, 1H); 4.18 (dd, *J* = 10.0 Hz, *J* = 5.0 Hz, 1H); 4.06 – 3.98 (m, 1H); 3.69 (s, 3H); 3.46 – 3.32 (m, 2H); 2.86 (q, *J* = 6.1 Hz, 2H); 1.10 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.1; 151.1; 149.9; 149.7; 146.7; 145.7; 141.0; 136.2; 129.8; 129.8; 127.7; 123.8; 123.6; 123.0; 122.5; 120.9; 119.8; 119.0; 111.3; 110.5; 89.6; 82.4; 74.8; 67.6; 64.6; 60.6; 59.6; 54.5; 34.7; 27.5; 27.2; 22.9; 20.5; 14.3. FTIR *V*_{max} (cm⁻¹): 2934 (w); 1736 (m); 1696 (m); 1611 (m); 1518 (s); 1472 (m); 1344 (s); 1191 (m); 1138 (s); 1062 (s); 828 (s); 737 (s); 652 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₉H₄₉O₉N₆Si 801.3386; Found 801.3396.



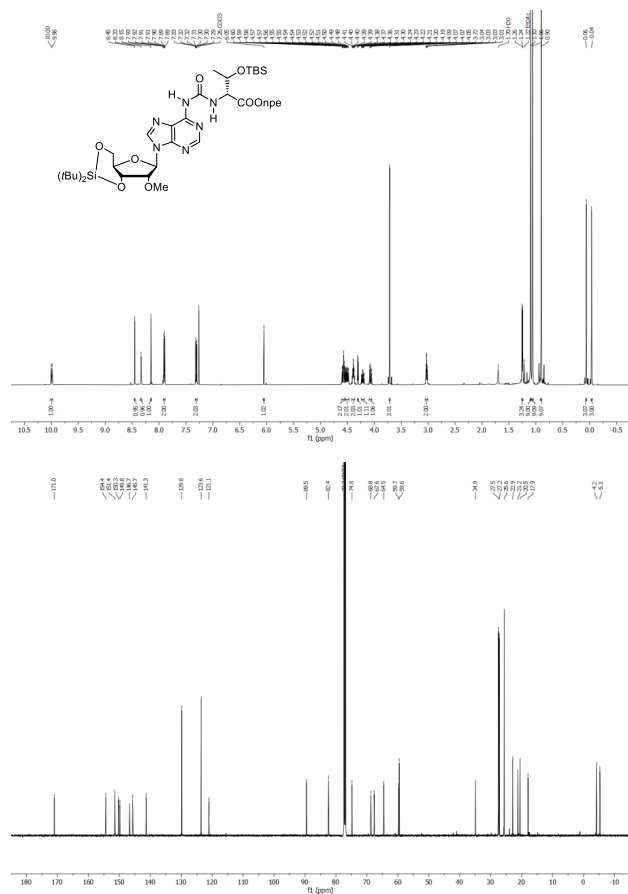
S10

3e_a: Yield = 59%. R_f = 0.38 (3:7 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.09 (d, *J* = 9.0 Hz, 1H); 8.81 (d, *J* = 6.3 Hz, 1H); 8.44 (s, 1H); 8.29 (d, *J* = 1.4 Hz, 1H); 7.94 – 7.87 (m, 2H); 7.32 – 7.28 (m, 2H); 6.03 (s, 1H); 4.66 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.60 – 4.49 (m, 2H); 4.46 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 1H); 4.39 (t, *J* = 6.4 Hz, 2H); 4.30 (d, *J* = 4.6 Hz, 1H); 4.19 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.03 (dd, *J* = 10.5 Hz, *J* = 9.2 Hz, 1H); 3.70 (s, 3H); 3.03 (t, *J* = 6.4 Hz, 2H); 1.25 (d, *J* = 6.4 Hz, 3H); 1.10 (s, 9H); 1.06 (s, 9H); 0.90 (s, 9H); 0.06 (s, 3H); -0.04 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.1; 154.6; 151.3; 150.4; 149.8; 146.7; 145.7; 142.2; 129.8; 123.6; 121.2; 89.8; 82.3; 74.8; 68.7; 67.6; 64.6; 59.7; 59.5; 34.9; 27.5; 27.2; 25.6; 22.9; 21.2; 20.5; 17.9; -4.2; -5.3. FTIR *V*_{max} (cm⁻¹): 2934 (w); 2858 (w); 1736 (m); 1701 (s); 1611 (m); 1520 (s); 1467 (m); 1345 (s); 1250 (m); 1137 (m); 1063 (s); 1030 (m); 827 (s); 652 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₈H₆₀O₁₀N₇Si₂ 830.3935; Found 830.3960.

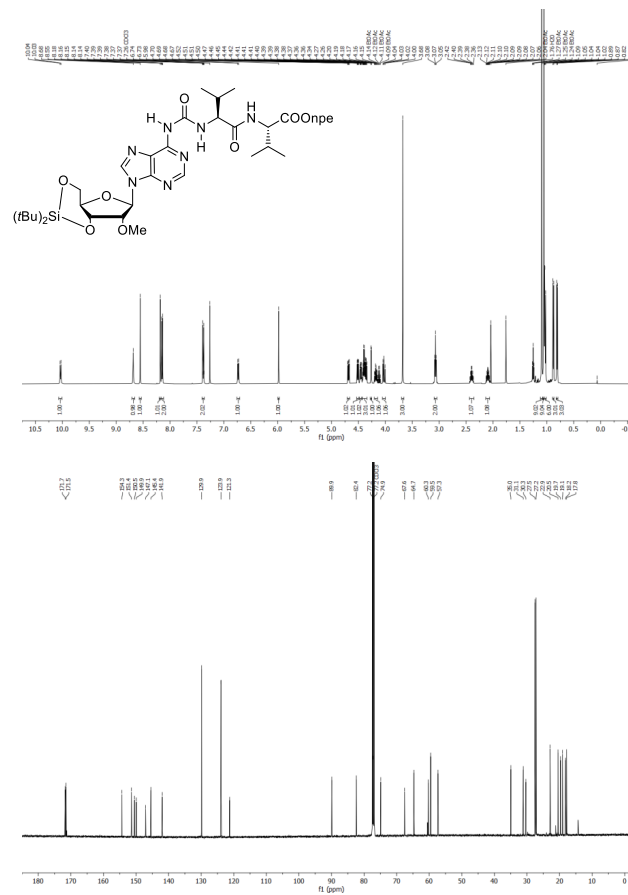


S11

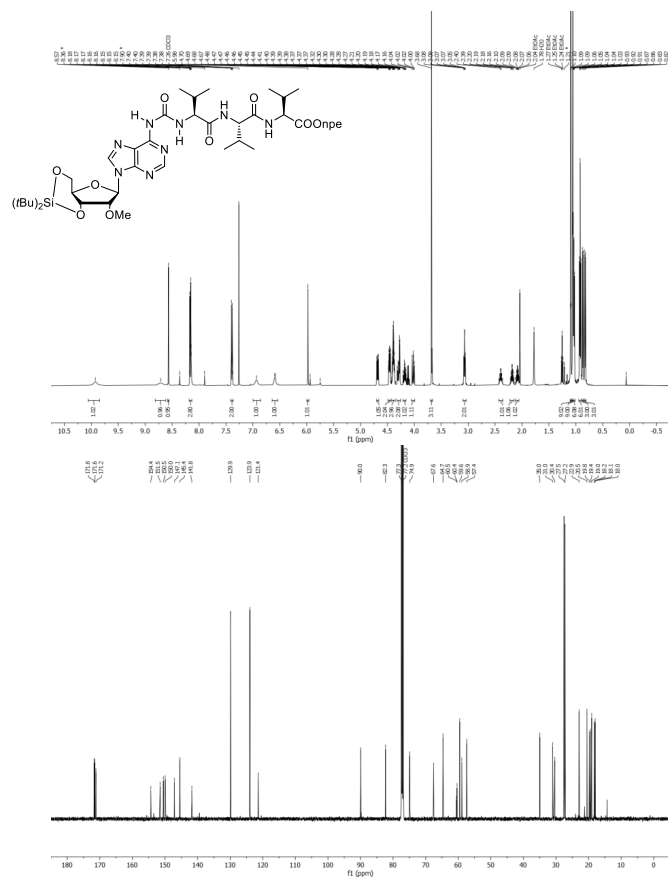
3e: Yield = 42%. Rf = 0.38 (2:3 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.99 (d, *J* = 9.1 Hz, 1H); 8.45 (s, 1H); 8.33 (s, 1H); 8.15 (s, 1H); 7.93 – 7.87 (m, 2H); 7.34 – 7.28 (m, 2H); 6.05 (s, 1H); 4.61 – 4.54 (m, 2H); 4.55 – 4.45 (m, 2H); 4.39 (td, *J* = 6.5 Hz, *J* = 3.2 Hz, 2H); 4.30 (d, *J* = 4.6 Hz, 1H); 4.22 (td, *J* = 10.2 Hz, *J* = 5.0 Hz, 1H); 4.07 (dd, *J* = 10.6 Hz, *J* = 9.2 Hz, 1H); 3.72 (s, 3H); 3.06 – 2.98 (m, 2H); 1.25 (d, *J* = 6.2 Hz, 3H); 1.10 (s, 9H); 1.06 (s, 9H); 0.90 (s, 9H); 0.06 (s, 3H); -0.04 (s, 3H). ¹³C NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.0; 154.4; 151.4; 150.3; 149.8; 146.7; 145.7; 141.3; 129.8; 123.6; 121.1; 89.5; 82.4; 74.8; 68.8; 67.6; 64.5; 59.7; 59.6; 34.9; 27.5; 27.2; 25.6; 22.9; 21.2; 20.5; 17.9; -4.2; -5.3. FTIR *v*_{max} (cm⁻¹): 2934 (w); 2858 (w); 1700 (s); 1611 (m); 1520 (s); 1467 (m); 1345 (s); 1252 (m); 1138 (m); 1063 (s); 828 (s); 654 (m); 443 (m). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₈H₆₀O₁₀N₇Si₂ 830.3935; Found 830.3966.



3f: Yield = 49%. Rf = 0.29 (1:4 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.03 (d, *J* = 8.2 Hz, 1H); 8.68 (s, 1H); 8.55 (s, 1H); 8.18 (s, 1H); 8.16 – 8.13 (m, 2H); 7.41 – 7.36 (m, 2H); 6.73 (d, *J* = 8.5 Hz, 1H); 5.98 (s, 1H); 4.68 (dd, *J* = 9.7 Hz, *J* = 4.7 Hz, 1H); 4.51 (dd, *J* = 8.5 Hz, *J* = 5.0 Hz, 1H); 4.46 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 1H); 4.43 – 4.33 (m, 3H); 4.26 (d, *J* = 4.6 Hz, 1H); 4.18 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H), 4.02 (dd, *J* = 10.5, 9.2 Hz, 1H), 3.68 (s, 3H), 3.07 (t, *J* = 6.7 Hz, 2H), 2.40 (h, *J* = 6.7 Hz, 1H), 2.10 (pd, *J* = 6.9, 5.0 Hz, 1H); 1.09 (s, 9H); 1.05 (s, 9H); 1.05 – 1.01 (m, 6H); 0.88 (d, *J* = 6.8 Hz, 3H); 0.81 (d, *J* = 6.9 Hz, 3H). ¹³C NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.7; 171.5; 154.3; 151.4; 150.5; 149.9; 147.1; 145.4; 141.9; 129.9; 123.9; 121.3; 89.9; 82.4; 77.2; 74.9; 67.6; 64.7; 60.3; 59.5; 57.3; 35.0; 31.1; 30.3; 27.5; 27.2; 22.9; 20.5; 19.7; 19.1; 18.2; 17.8. FTIR *v*_{max} (cm⁻¹): 2963 (w); 2933 (w); 1676 (m); 1519 (s); 1469 (m); 1345 (s); 1140 (s); 1063 (s); 828 (s); 652 (s). HRMS (ESI) *m/z* [M+Na]⁺ Calcd for C₃₈H₅₆O₁₀N₆SiNa 835.3780; Found 835.3786.



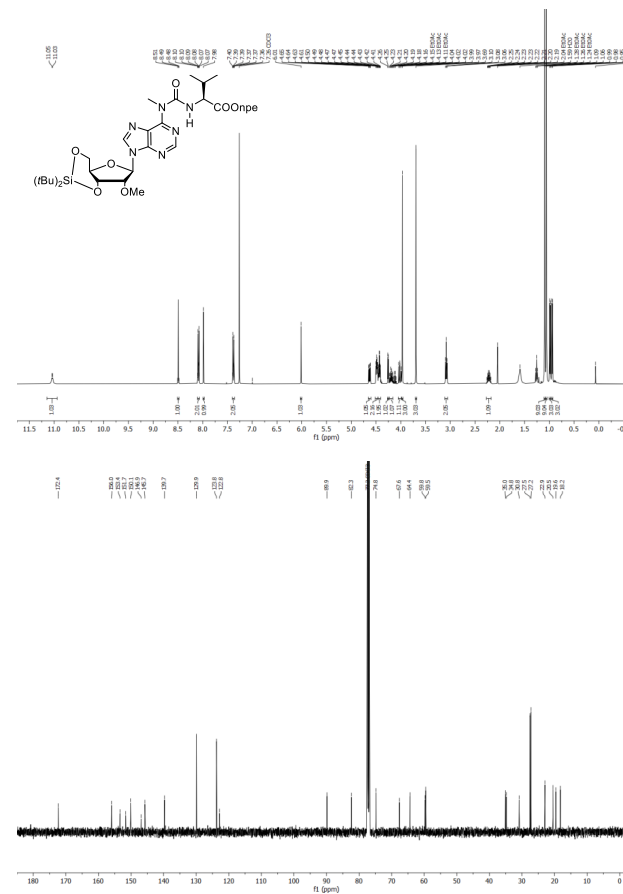
3g_L: Yield = 60%. R_f = 0.28 (1:4 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 9.93 (s, 1H); 8.71 (s, 1H); 8.57 (s, 1H); 8.19 – 8.13 (m, 3H); 7.43 – 7.36 (m, 2H); 6.93 (s, 1H); 6.59 (d, J = 8.1 Hz, 1H); 5.98 (s, 1H); 4.68 (dd, J = 9.7, J = 4.7 Hz, 1H); 4.46 (ddd, J = 8.6 Hz, J = 5.1 Hz, J = 2.3 Hz, 2H); 4.39 (tt, J = 5.5 Hz, J = 2.6 Hz, 3H); 4.33 – 4.26 (m, 2H); 4.18 (td, J = 10.1 Hz, J = 5.0 Hz, 1H); 4.02 (dd, J = 10.5 Hz, J = 9.2 Hz, 1H); 3.68 (s, 3H); 3.09 – 3.04 (m, 2H); 2.44 – 2.35 (m, 1H); 2.23 – 2.14 (m, 1H); 2.12 – 2.05 (m, 1H); 1.09 (s, 9H); 1.06 (s, 9H); 1.04 (dd, J = 7.0 Hz, J = 5.2 Hz, 6H); 0.92 (t, J = 6.5 Hz, 6H); 0.87 (d, J = 6.8 Hz, 3H), 0.83 (d, J = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.8; 171.6; 171.2; 154.4; 151.5; 150.5; 150.0; 147.1; 145.4; 141.8; 129.9; 123.9; 121.4; 90.0; 82.3; 77.3; 74.9; 67.6; 64.7; 60.5; 60.4; 59.6; 58.9; 57.4; 35.0; 31.0; 30.4; 27.5; 27.2; 22.9; 20.5; 19.8; 19.4; 19.0; 18.2; 18.1; 18.0. FTIR ν_{max} (cm⁻¹): 3291 (w); 2963 (w); 2934 (w); 1645 (m); 1520 (s); 1467 (m); 1345 (s); 1204 (m); 1138 (s); 1065 (s); 828 (s); 737 (m); 652 (s). HRMS (ESI) m/z: [M+Na]⁺ Calcd for C₄₃H₆₅O₁₁N₅SiNa 934.4465; Found 934.4465.



S14

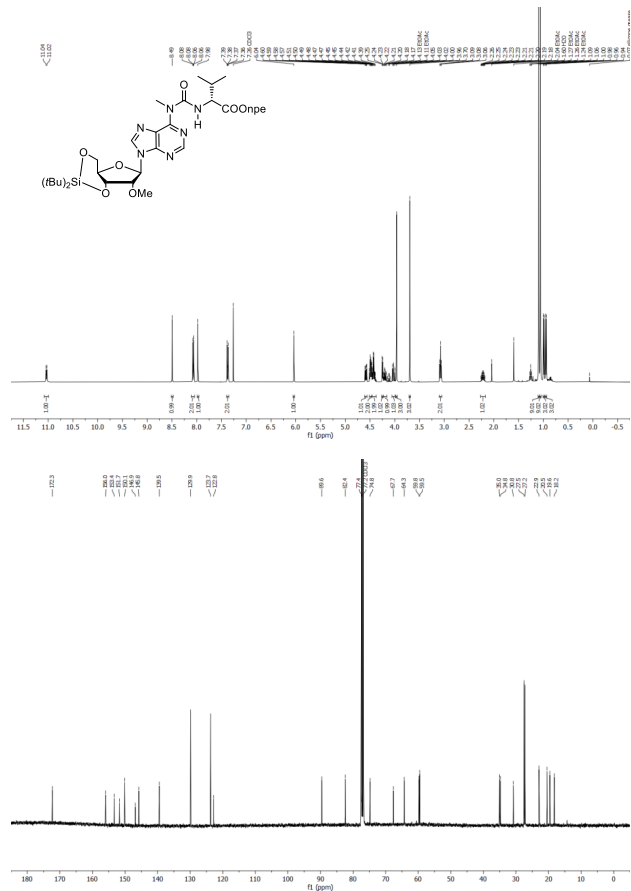
General procedure for the synthesis of compound 4: Compound **3** (1 equiv.) and K₂CO₃ (3 equiv.) were suspended in dry DMF. The suspension was stirred for 30 min at 0°C under nitrogen atmosphere. After that, CH₃I (2 equiv.) was added dropwise and the reaction was stirred at r.t. overnight. After that, the reaction mixture was diluted with CH₂Cl₂ and washed with aqueous saturated NH₄Cl and water. The organic layer was dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

4a_L: Yield = 54%. R_f = 0.33 (7:3 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 11.04 (d, J = 7.6 Hz, 1H); 8.49 (s, 1H); 8.13 – 8.05 (m, 2H); 7.98 (s, 1H); 7.42 – 7.34 (m, 2H); 6.01 (s, 1H); 4.63 (dd, J = 9.7 Hz, J = 4.7 Hz, 1H); 4.48 (dd, J = 9.1 Hz, J = 5.1 Hz, 2H); 4.43 (td, J = 6.6 Hz, J = 3.0 Hz, 2H); 4.26 (d, J = 4.7 Hz, 1H); 4.19 (td, J = 10.1 Hz, J = 5.0 Hz, 1H); 4.02 (dd, J = 10.5 Hz, J = 9.2 Hz, 1H); 3.97 (s, 3H); 3.69 (s, 3H); 3.08 (t, J = 6.6 Hz, 2H); 2.28 – 2.18 (m, 1H); 1.09 (s, 9H); 1.06 (s, 9H); 0.99 (d, J = 6.8 Hz, 3H); 0.94 (d, J = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.4; 156.0; 153.4; 151.7; 150.1; 146.9; 145.7; 139.7; 129.9; 123.8; 122.8; 89.9; 82.3; 74.8; 67.6; 64.4; 59.8; 59.5; 35.0; 34.8; 30.8; 27.5; 27.2; 22.9; 20.5; 19.6; 18.2. FTIR ν_{max} (cm⁻¹): 2932 (w); 1738 (w); 1687 (m); 1568 (m); 1519 (s); 1470 (m); 1344 (s); 1264 (m); 1135 (s); 1064 (s); 1012 (s); 827 (s). HRMS (ESI) m/z: [M+H]⁺ Calcd for C₃₄H₅₀O₉N₇Si 728.3434; Found 728.3420.

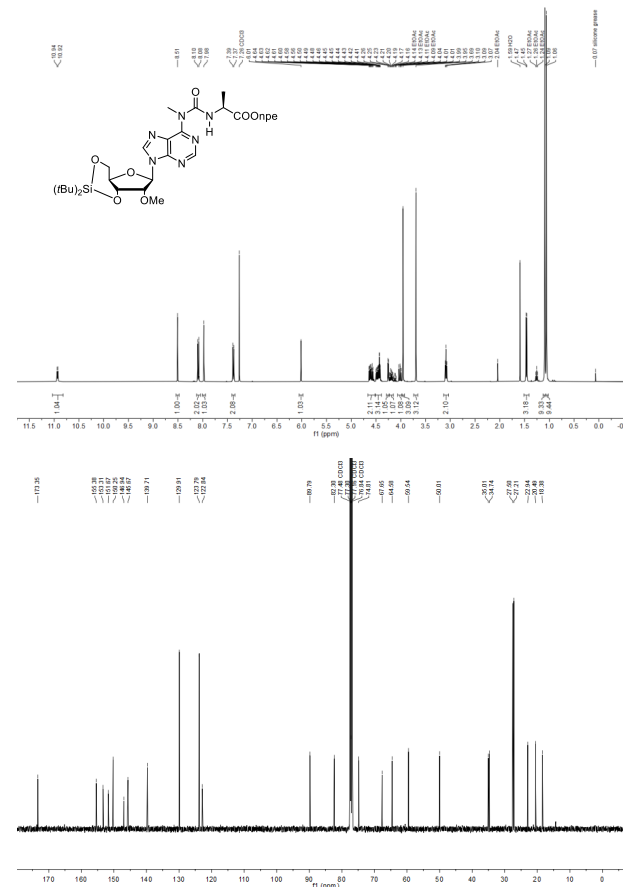


S15

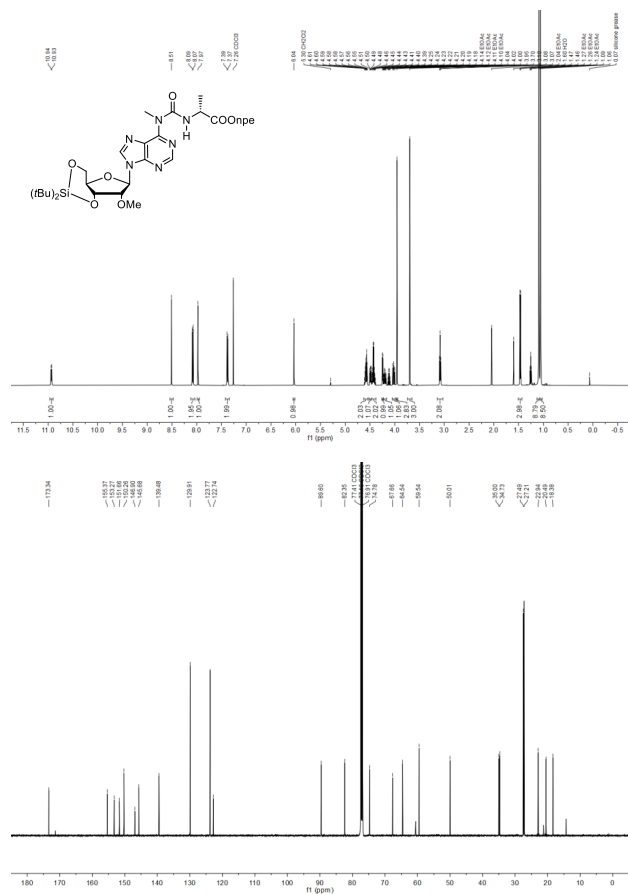
4a_b: Yield = 58%. R_f = 0.50 (3:2 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 11.03 (d, *J* = 7.6 Hz, 1H); 8.49 (s, 1H); 8.10 – 8.04 (m, 2H); 7.98 (s, 1H); 7.41 – 7.34 (m, 2H); 6.04 (s, 1H); 4.58 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.52 – 4.46 (m, 2H); 4.46 – 4.39 (m, 2H); 4.25 (d, *J* = 4.6 Hz, 1H); 4.20 (dt, *J* = 10.0 Hz, *J* = 5.1 Hz, 1H); 4.02 (dd, *J* = 10.5 Hz, *J* = 9.2 Hz, 1H); 3.96 (s, 3H); 3.70 (s, 3H); 3.08 (t, *J* = 6.5 Hz, 2H); 2.28 – 2.16 (m, 1H); 1.09 (s, 9H); 1.06 (s, 9H); 0.99 (d, *J* = 6.9 Hz, 3H); 0.95 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.3; 156.0; 153.4; 151.7; 150.1; 146.9; 145.8; 139.5; 129.9; 123.7; 122.8; 89.6; 82.4; 77.4; 74.8; 67.7; 64.3; 59.8; 59.5; 35.0; 34.8; 30.8; 27.5; 27.2; 22.9; 20.5; 19.6; 18.2. FTIR ν_{max} (cm⁻¹): 2935 (w); 1739 (w); 1683 (m); 1570 (m); 1520 (s); 1469 (m); 1346 (s); 1264 (m); 1136 (s); 1064 (s); 1015 (s); 907 (s); 728 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₄H₅₀O₉N₇Si 728.3434; Found 728.3432.



4b_c: Yield = 60%. R_f = 0.20 (9:1 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.93 (d, *J* = 6.7 Hz, 1H); 8.51 (s, 1H); 8.10 – 8.08 (m, 2H); 7.98 (s, 1H); 7.39 – 7.37 (m, 2H); 6.01 (s, 1H); 4.64 – 4.56 (m, 2H); 4.50 – 4.41 (m, 3H); 4.25 (d, *J* = 4.7 Hz, 1H); 4.19 (ddd, *J* = 10.3 Hz, *J* = 10.3 Hz, *J* = 4.7 Hz, 1H); 4.01 (dd, *J* = 10.3 Hz, *J* = 10.3 Hz, 1H); 3.95 (s, 3H); 3.69 (s, 3H); 3.09 (t, *J* = 6.5 Hz, 2H); 1.46 (d, *J* = 7.2 Hz, 3H); 1.09 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 173.3; 155.4; 153.3; 151.7; 150.2; 146.9; 145.7; 139.7; 129.9; 123.8; 122.8; 89.8; 82.3; 77.3; 74.8; 67.6; 64.6; 59.5; 50.0; 35.0; 34.7; 27.5; 27.2; 22.9; 20.5; 18.4. FTIR ν_{max} (cm⁻¹): 2933 (w); 1740 (w); 1683 (m); 1568 (m); 1519 (s); 1344 (s); 1265 (m); 1135 (m); 1016 (m); 827 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₂H₄₆O₉N₇Si 700.3120; Found 700.3121.

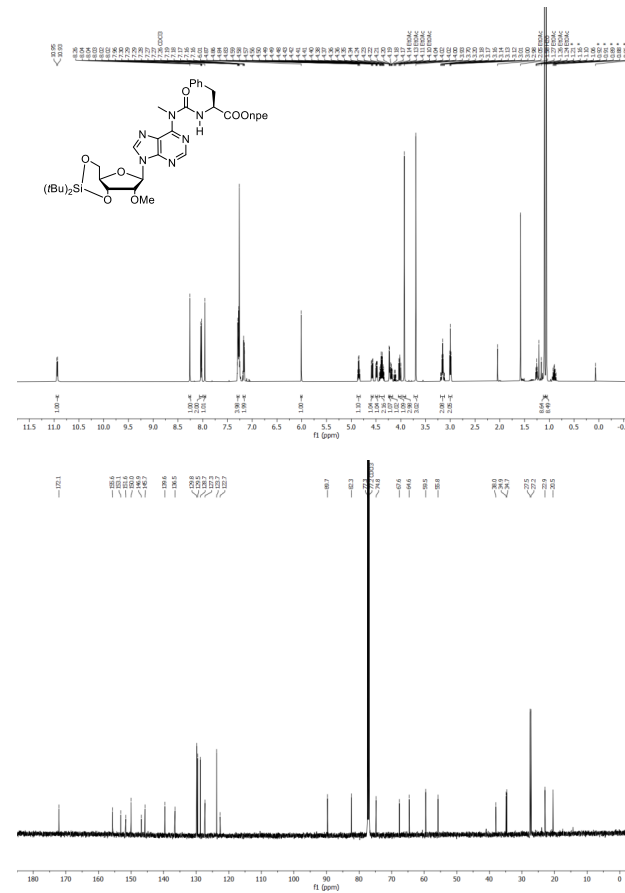


4b_o: Yield = 90% yield. Rf = 0.40 (8:2 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.93 (d, *J* = 6.6 Hz, 1H); 8.51 (s, 1H); 8.09 – 8.07 (m, 2H); 7.97 (s, 1H); 7.39 – 7.37 (m, 2H); 6.04 (s, 1H); 4.61 – 4.55 (m, 2H); 4.49 (dd, *J* = 9.3 Hz, *J* = 5.0 Hz, 1H); 4.46 – 4.39 (m, 2H); 4.24 (d, *J* = 4.6 Hz, 1H); 4.20 (ddd, *J* = 9.3 Hz, *J* = 9.3 Hz, *J* = 5.0 Hz, 1H); 4.02 (dd, *J* = 9.3 Hz, *J* = 9.3 Hz, 1H); 3.95 (s, 3H); 3.70 (s, 3H); 3.08 (t, *J* = 6.5 Hz, 2H); 1.47 (d, *J* = 7.3 Hz, 3H); 1.09 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 173.3; 155.4; 153.3; 151.7; 150.3; 146.9; 145.7; 139.5; 129.9; 123.8; 122.7; 89.6; 82.3; 77.4; 74.8; 67.7; 64.5; 59.5; 50.0; 35.0; 34.7; 27.5; 27.2; 22.9; 20.5; 18.4. FTIR *v*_{max} (cm⁻¹): 2860 (w); 1740 (m); 1684 (m); 1570 (m); 1518 (s); 1469 (m); 1345 (s); 1265 (m); 1137 (s); 1063 (s); 1016 (s); 828 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₂H₄₆O₉N₇Si 700.3120; Found 700.3126.



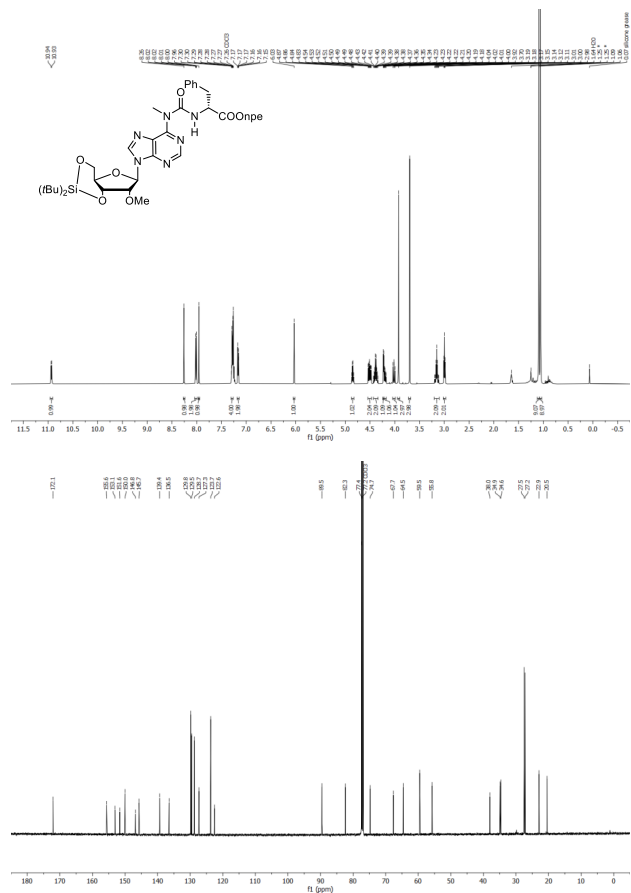
S18

4c_o: Yield = 55% yield. Rf = 0.48 (1:1 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.94 (d, *J* = 6.8 Hz, 1H); 8.26 (s, 1H); 8.04 – 8.02 (m, 2H); 7.96 (s, 1H); 7.30 – 7.26 (m, 4H); 7.19 – 7.16 (m, 2H); 6.01 (s, 1H); 4.85 (q, *J* = 6.5 Hz, 1H); 4.58 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.49 (dd, *J* = 9.3 Hz, *J* = 5.0 Hz, 1H); 4.44 – 4.33 (m, 2H); 4.23 (d, *J* = 4.6 Hz, 1H); 4.19 (dt, *J* = 10.1 Hz, *J* = 5.1 Hz, 1H); 4.02 (dd, *J* = 10.5 Hz, *J* = 9.3 Hz, 1H); 3.93 (s, 3H); 3.70 (s, 3H); 3.21 – 3.11 (m, 2H); 3.00 (t, *J* = 6.5 Hz, 2H); 1.10 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.1; 155.6; 153.1; 151.6; 150.0; 146.9; 145.7; 139.6; 136.5; 129.8; 129.5; 128.7; 127.3; 123.7; 122.7; 89.7; 82.3; 77.3; 74.8; 67.6; 64.6; 59.5; 55.8; 38.0; 34.9; 34.7; 27.5; 27.2; 22.9; 20.5. FTIR *v*_{max} (cm⁻¹): 2935 (w); 2860 (w); 1740 (m); 1684 (m); 1571 (m); 1518 (s); 1467 (m); 1345 (s); 1266 (m); 1135 (s); 1063 (s); 1015 (s); 828 (s); 652 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₈H₅₀O₉N₇Si 776.3433; Found 776.3443.

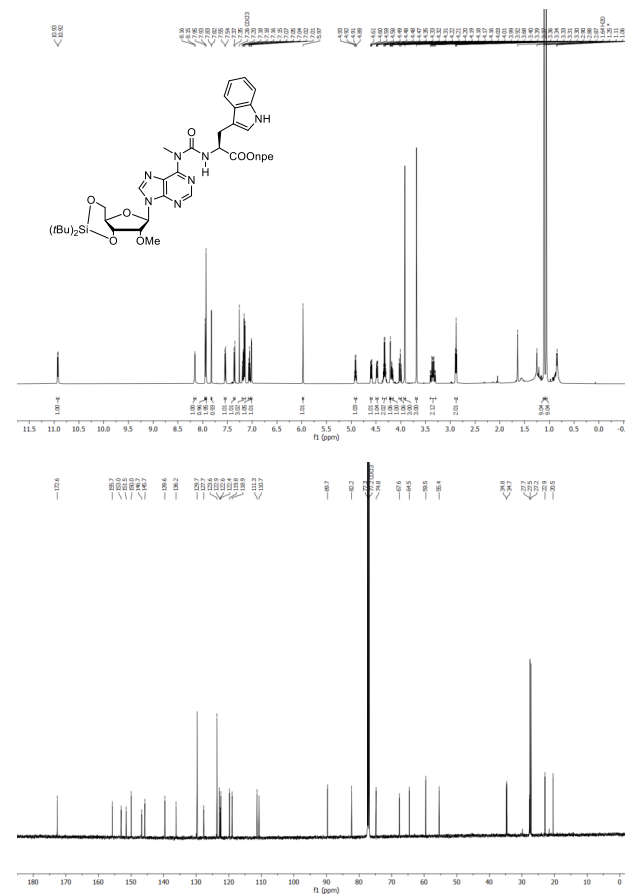


S19

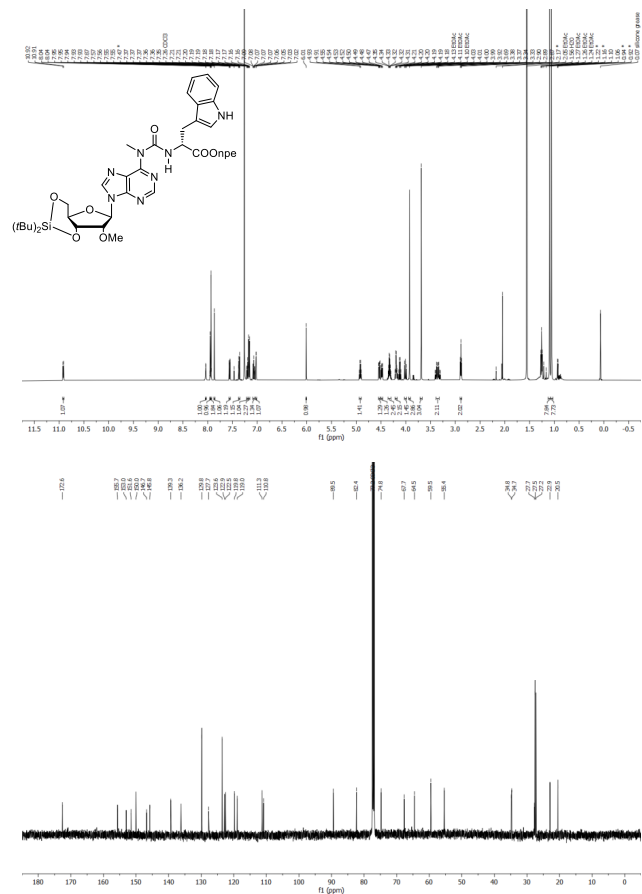
4c_b: Yield = 80% yield. R_f = 0.42 (1:1 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.93 (d, *J* = 6.9 Hz, 1H); 8.26 (s, 1H); 8.03 – 8.00 (m, 2H); 7.96 (s, 1H); 7.30 – 7.27 (m, 4H); 7.19 – 7.14 (m, 2H); 6.03 (s, 1H); 4.85 (q, *J* = 6.4 Hz, 1H); 4.54 – 4.48 (m, 2H); 4.38 (qt, *J* = 11.0 Hz, *J* = 6.5 Hz, 2H); 4.23 (d, *J* = 4.6 Hz, 1H); 4.19 (dd, *J* = 10.0 Hz, *J* = 5.0 Hz, 1H); 4.02 (dd, *J* = 10.6 Hz, *J* = 9.2 Hz, 1H); 3.92 (s, 3H); 3.70 (s, 3H); 3.20 – 3.10 (m, 2H); 3.00 (t, *J* = 6.5 Hz, 2H); 1.09 (s, 9H); 1.06 (s, 9H). ¹³C NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.1; 155.6; 153.1; 151.6; 150.0; 146.8; 145.7; 139.4; 136.5; 129.8; 129.5; 128.7; 127.3; 123.7; 122.6; 89.5; 82.3; 77.4; 74.7; 67.7; 64.5; 59.5; 55.8; 38.0; 34.9; 34.6; 27.5; 27.2; 22.9; 20.5. FTIR ν_{max} (cm⁻¹): 2934 (w); 2860 (w); 1742 (m); 1686 (m); 1571 (m); 1518 (s); 1467 (m); 1345 (s); 1266 (m); 1137 (s); 1063 (s); 1015 (s); 828 (s); 652 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₈H₅₀O₉N₇Si 776.3433; Found 776.3442.



4d_l: Yield = 75%. R_f = 0.48 (1:1 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.93 (d, *J* = 6.8 Hz, 1H); 8.16 (s, 1H); 7.95 (s, 1H); 7.93 (s, 2H); 7.83 (s, 1H); 7.54 (d, *J* = 7.9 Hz, 1H); 7.36 (d, *J* = 8.2 Hz, 1H); 7.20 – 7.15 (m, 3H); 7.06 (t, *J* = 7.5 Hz, 1H); 7.02 (d, *J* = 2.3 Hz, 1H); 5.97 (s, 1H); 4.91 (q, *J* = 6.3 Hz, 1H); 4.60 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.48 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 1H); 4.33 (q, *J* = 6.3 Hz, 2H); 4.21 (d, *J* = 4.7 Hz, 1H); 4.18 (dt, *J* = 10.0 Hz, *J* = 5.1 Hz, 1H); 4.01 (t, *J* = 9.9 Hz, 1H); 3.92 (s, 3H); 3.68 (s, 3H); 3.35 (qd, *J* = 14.8 Hz, *J* = 6.0 Hz, 2H); 2.88 (t, *J* = 6.4 Hz, 2H); 1.11 (s, 9H); 1.06 (s, 9H). ¹³C NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.6; 155.7; 153.0; 151.5; 150.0; 146.7; 145.7; 139.6; 136.2; 129.7; 127.7; 123.6; 122.9; 122.6; 122.4; 119.8; 118.9; 111.3; 110.7; 89.7; 82.2; 77.3; 74.8; 67.6; 64.5; 59.5; 55.4; 34.8; 34.7; 27.7; 27.5; 27.2; 22.9; 20.5. FTIR ν_{max} (cm⁻¹): 2934 (w); 2860 (w); 1735 (m); 1654 (m); 1571 (m); 1516 (s); 1459 (m); 1344 (s); 1267 (m); 1135 (s); 1063 (s); 1013 (s); 828 (s); 739 (s); 652 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₄₀H₅₁O₉NaSi 815.3542; Found 815.3555.

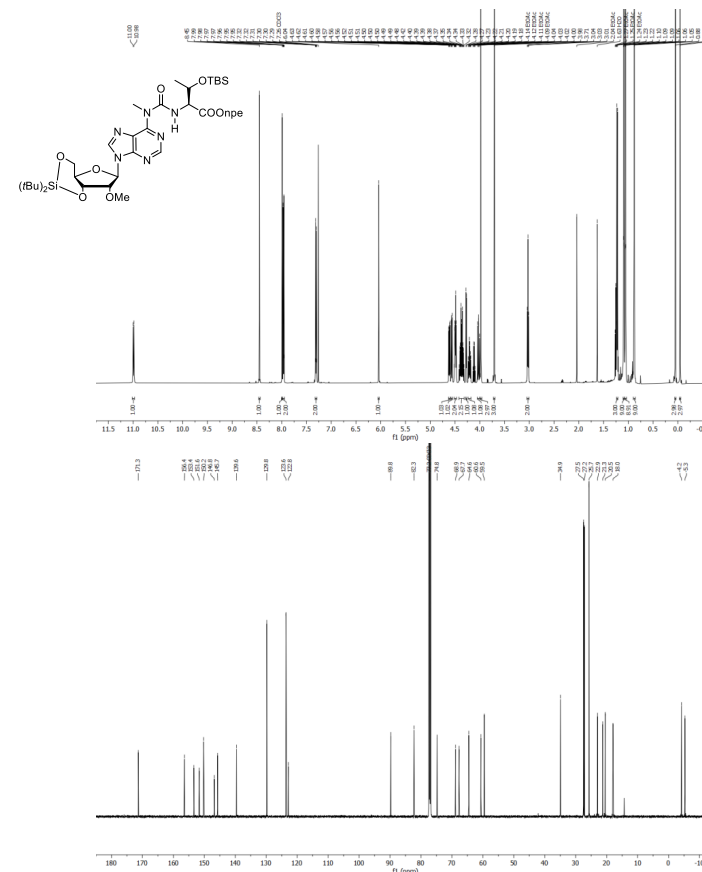


4d_b: Yield = 48%. Rf = 0.45 (1:1 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.92 (d, *J* = 6.9 Hz, 1H); 8.04 (s, 1H); 7.93 (s, 1H); 7.93 (s, 2H); 7.87 (s, 1H); 7.56 (dd, *J* = 8.0 Hz, *J* = 1.0 Hz, 1H); 7.36 (dt, *J* = 8.1 Hz, *J* = 0.9 Hz, 1H); 7.22 – 7.19 (m, 1H); 7.18 – 7.13 (m, 2H); 7.09 – 7.05 (m, 1H); 7.02 (d, *J* = 2.4 Hz, 1H); 6.01 (s, 1H); 4.92 (q, *J* = 6.2 Hz, 1H); 4.53 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.48 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 1H); 4.33 (ddt, *J* = 10.9 Hz, *J* = 6.6 Hz, *J* = 4.7 Hz, 2H); 4.22 – 4.17 (m, 2H); 4.01 (dd, *J* = 10.5 Hz, *J* = 9.3 Hz, 1H); 3.92 (s, 3H); 3.69 (s, 3H); 3.36 (qd, *J* = 14.8 Hz, *J* = 6.0 Hz, 2H); 2.89 (t, *J* = 6.4 Hz, 2H); 1.10 (s, 9H); 1.06 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.6; 155.7; 153.0; 151.6; 150.0; 146.7; 145.8; 139.3; 136.2; 129.8; 127.7; 123.6; 122.9; 122.5; 119.8; 119.0; 111.3; 110.8; 89.5; 82.4; 74.8; 67.7; 64.5; 59.5; 55.4; 34.8; 34.7; 27.7; 27.5; 27.2; 22.9; 20.5. FTIR *v*_{max} (cm⁻¹): 2934 (w); 2860 (w); 1736 (m); 1671 (m); 1571 (m); 1516 (s); 1466 (m); 1344 (s); 1267 (m); 1135 (s); 1063 (s); 1012 (s); 828 (s); 737 (s); 652 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₄₀H₅₁O₉N₅Si 815.3542; Found 815.3558.



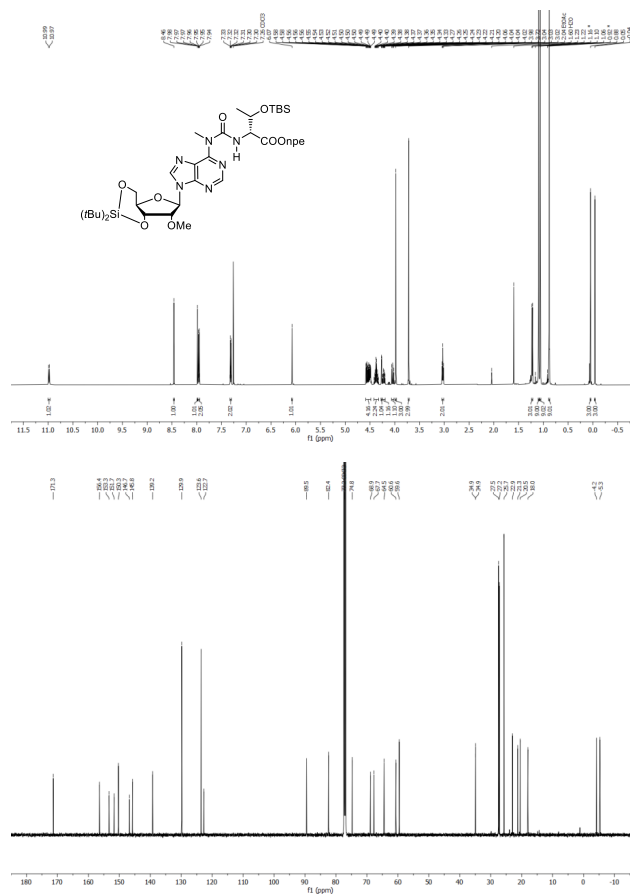
S22

4e_a: Yield = 96%. Rf = 0.59 (1:1 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.99 (d, *J* = 8.7 Hz, 1H); 8.45 (s, 1H); 7.99 (s, 1H); 7.98 – 7.94 (m, 2H); 7.33 – 7.28 (m, 2H); 6.04 (s, 1H); 4.61 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.57 (dd, *J* = 8.7 Hz, *J* = 1.7 Hz, 1H); 4.50 (ddd, *J* = 9.3 Hz, *J* = 6.4 Hz, *J* = 4.9 Hz, 2H); 4.42 – 4.31 (m, 2H); 4.27 (d, *J* = 4.6 Hz, 1H); 4.21 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.02 (dd, *J* = 10.5 Hz, *J* = 9.2 Hz, 1H); 3.98 (s, 3H); 3.71 (s, 3H); 3.03 (t, *J* = 6.5 Hz, 2H); 1.23 (d, *J* = 6.3 Hz, 3H); 1.09 (s, 9H); 1.06 (s, 9H); 0.88 (s, 9H); 0.05 (s, 3H); -0.04 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.3; 156.4; 153.4; 151.6; 150.2; 146.8; 145.7; 139.6; 129.8; 123.6; 122.8; 89.8; 82.3; 74.8; 68.9; 67.7; 64.6; 60.6; 59.5; 34.9; 27.5; 27.2; 25.7; 22.9; 21.3; 20.5; 18.0; -4.2; -5.3. FTIR *v*_{max} (cm⁻¹): 2934 (w); 2858 (w); 1736 (m); 1687 (m); 1569 (m); 1519 (s); 1466 (m); 1345 (s); 1253 (m); 1134 (s); 1065 (s); 1022 (s); 828 (s); 652 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₉H₆₂O₁₀N₇Si₂ 844.4091; Found 844.4122.

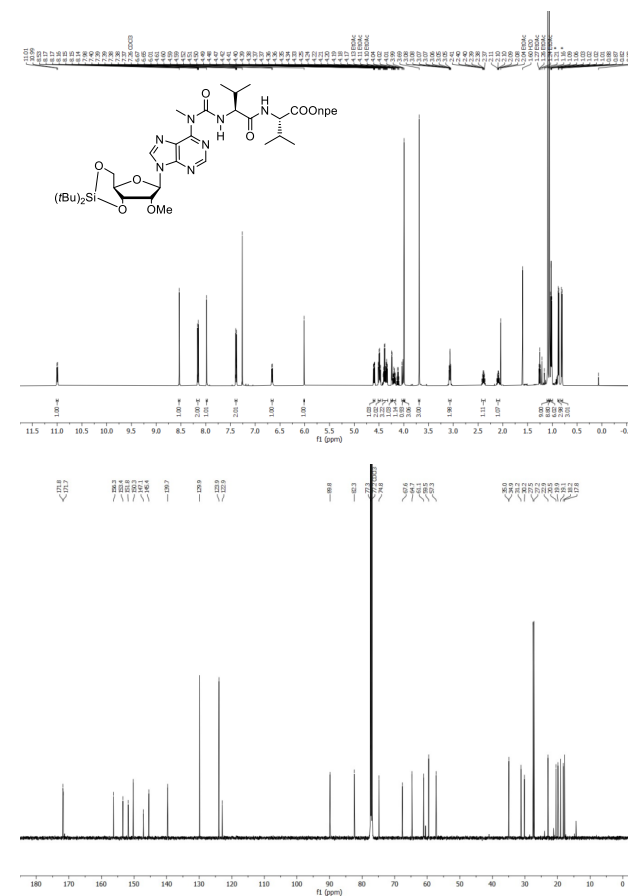


S23

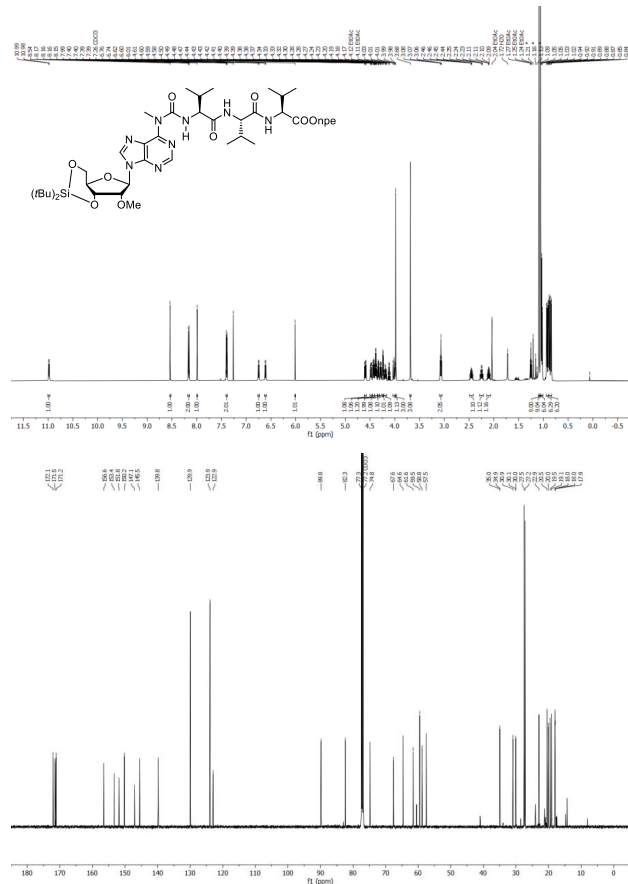
4e_b: Yield = 92%. R_f = 0.33 (3:2 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.98 (d, *J* = 8.7 Hz, 1H); 8.46 (s, 1H); 7.98 (s, 1H); 7.98 – 7.94 (m, 2H); 7.34 – 7.29 (m, 2H); 6.07 (s, 1H); 4.59 – 4.48 (m, 4H); 4.42 – 4.33 (m, 2H); 4.27 (d, *J* = 4.5 Hz, 1H); 4.22 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.04 (dd, *J* = 10.6 Hz, *J* = 9.3 Hz, 1H); 3.98 (s, 3H); 3.72 (s, 3H); 3.03 (t, *J* = 6.5 Hz, 2H); 1.22 (d, *J* = 6.3 Hz, 3H); 1.10 (s, 9H); 1.06 (s, 9H); 0.88 (s, 9H); 0.05 (s, 3H); -0.04 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.3; 156.4; 153.3; 151.7; 150.3; 146.7; 145.8; 139.2; 129.9; 123.6; 122.7; 89.5; 82.4; 74.8; 68.9; 67.7; 64.5; 60.6; 59.6; 34.9; 34.9; 27.5; 27.2; 25.7; 22.9; 21.3; 20.5; 18.0; -4.2; -5.3. FTIR *v*_{max} (cm⁻¹): 2934 (w); 2858 (w); 1736 (w); 1687 (m); 1568 (m); 1519 (s); 1464 (m); 1345 (s); 1253 (m); 1135 (m); 1065 (m); 1023 (s); 828 (s); 652 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₉H₆₂O₁₀N₇Si₂ 844.4091; Found 844.4128.



4f_L: Yield = 93%. R_f = 0.36 (1:1 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 11.00 (d, *J* = 7.7 Hz, 1H); 8.53 (s, 1H); 8.18 – 8.14 (m, 2H); 7.98 (s, 1H); 7.40 – 7.37 (m, 2H); 6.66 (d, *J* = 8.6 Hz, 1H); 6.01 (s, 1H); 4.60 (dd, *J* = 9.7 Hz, *J* = 4.7 Hz, 1H); 4.49 (td, *J* = 9.4 Hz, *J* = 9.0 Hz, *J* = 5.0 Hz, 2H); 4.43 – 4.32 (m, 3H); 4.24 (d, *J* = 4.6 Hz, 1H); 4.19 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.05 – 4.00 (m, 1H); 3.99 (s, 3H); 3.69 (s, 3H); 3.07 (td, *J* = 6.7 Hz, *J* = 1.3 Hz, 2H); 2.43 – 2.35 (m, 1H); 2.09 (pd, *J* = 6.8 Hz, *J* = 4.9 Hz, 1H); 1.09 (s, 9H); 1.06 (s, 9H); 1.02 (dd, *J* = 6.8 Hz, *J* = 5.7 Hz, 6H); 0.88 (d, *J* = 6.8 Hz, 3H); 0.81 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.8; 171.7; 156.3; 153.4; 151.8; 150.3; 147.1; 145.4; 139.7; 129.9; 123.9; 122.9; 89.8; 82.3; 77.3; 74.8; 67.6; 64.7; 61.1; 59.5; 57.3; 35.0; 34.9; 31.2; 30.2; 27.5; 27.2; 22.9; 20.5; 19.9; 19.1; 18.2; 17.8. FTIR *v*_{max} (cm⁻¹): 2963 (w); 2933 (w); 2863 (w); 1664 (m); 1575 (m); 1519 (s); 1467 (m); 1345 (m); 1137 (s); 1063 (m); 1013 (m); 828 (s); 652 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₉H₅₈O₁₀N₈SiNa 849.3934; Found 849.3943.



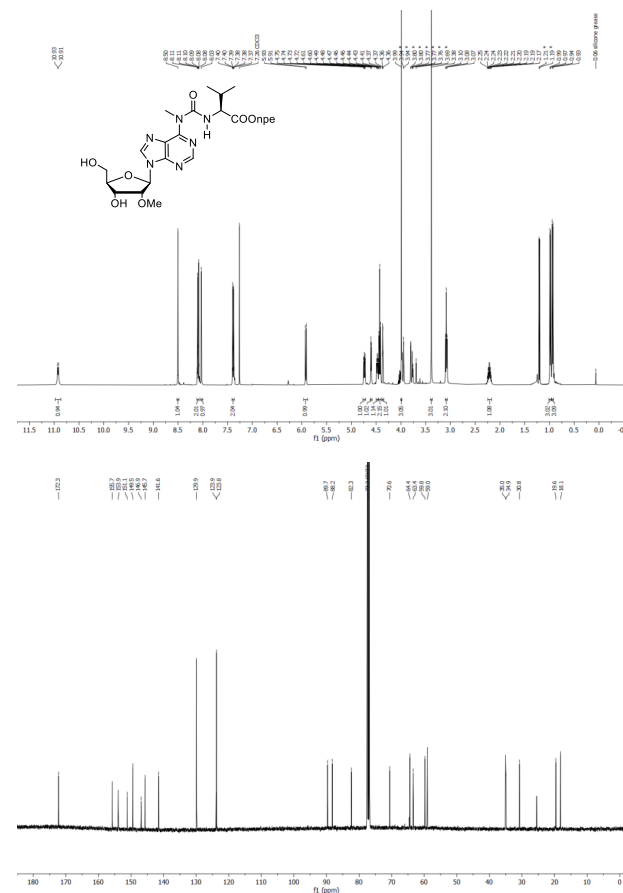
4g: Yield = 80%. Rf = 0.27 (1:1 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.98 (d, *J* = 6.9 Hz, 1H); 8.54 (s, 1H); 8.18 – 8.14 (m, 2H); 7.99 (s, 1H); 7.42 – 7.37 (m, 2H); 6.75 (d, *J* = 8.5 Hz, 1H); 6.61 (d, *J* = 8.5 Hz, 1H); 6.01 (s, 1H); 4.59 (dd, *J* = 9.7 Hz, *J* = 4.7 Hz, 1H); 4.48 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 1H); 4.43 (dd, *J* = 8.6 Hz, *J* = 5.3 Hz, 1H); 4.38 (td, *J* = 6.7 Hz, *J* = 5.0 Hz, 2H); 4.33 (dd, *J* = 7.0 Hz, *J* = 5.2 Hz, 1H); 4.28 (dd, *J* = 8.5 Hz, *J* = 6.2 Hz, 1H); 4.24 (d, *J* = 4.6 Hz, 1H); 4.19 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.01 (dd, *J* = 10.5 Hz, *J* = 9.3 Hz, 1H); 3.98 (s, 3H); 3.68 (s, 3H); 3.07 (t, *J* = 6.7 Hz, 2H); 2.45 (pd, *J* = 5.2 Hz, 1H); 2.25 (h, *J* = 6.8 Hz, 1H); 2.10 (pd, *J* = 6.9 Hz, *J* = 5.3 Hz, 1H); 1.09 (s, 9H); 1.05 (s, 9H); 1.03 (t, *J* = 6.6 Hz, 6H); 0.91 (dd, *J* = 15.0 Hz, *J* = 6.8 Hz, 6H); 0.86 (dd, *J* = 16.9 Hz, *J* = 6.9 Hz, 6H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.1; 171.5; 171.2; 156.6; 153.4; 151.8; 150.2; 147.1; 145.5; 139.8; 129.9; 123.9; 89.8; 82.3; 77.3; 74.8; 67.6; 64.6; 61.6; 59.5; 58.8; 57.5; 35.0; 34.9; 30.9; 30.1; 30.0; 27.5; 27.2; 22.9; 20.5; 20.0; 19.5; 19.1; 18.0; 18.0; 17.9. FTIR *v*_{max} (cm⁻¹): 3304 (w); 2963 (w); 1651 (m); 1520 (s); 1345 (m); 1136 (m); 1065 (m); 1013 (m); 828 (m); 652 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₄₄H₆₇O₇N₇SiNa 948.4621; Found 948.4617.



S26

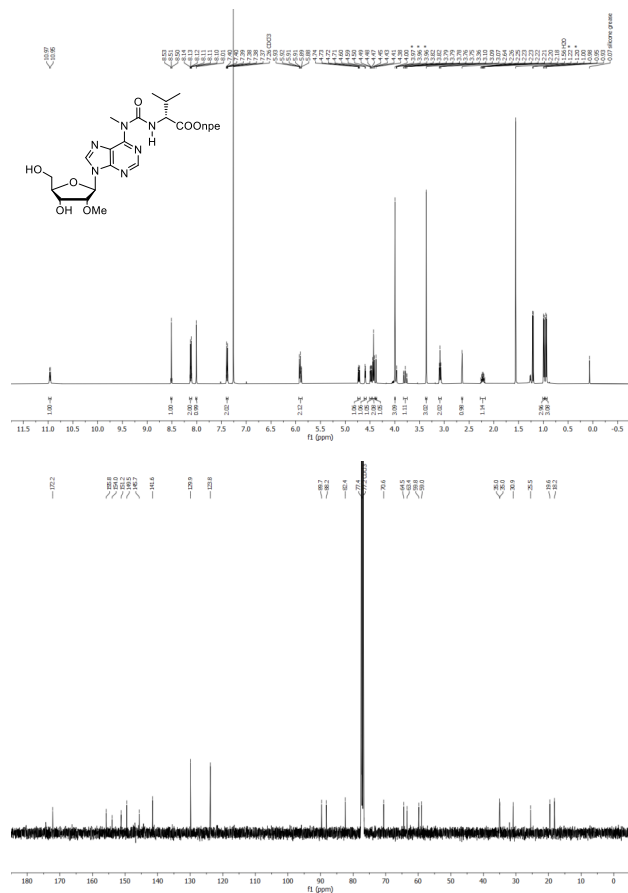
General procedure for the synthesis of compound 5: Compound **4** (1 equiv.) was added to a plastic flask and dissolved in dry 9:1 CH₂Cl₂/pyridine. The solution was stirred at 0°C. Finally, HF·pyridine (from a commercial solution containing 70% HF and 30% pyridine) was added and the reaction was stirred at 0°C for 2 h. After that, the reaction was quenched with aqueous saturated NaHCO₃ and CH₂Cl₂ was added. The organic layer was separated and the crude was further extracted with CH₂Cl₂. The combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

5a₁: Yield = 83%. Rf = 0.35 (94:6 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.92 (d, *J* = 7.5 Hz, 1H); 8.50 (s, 1H); 8.11 – 8.08 (m, 2H); 8.03 (s, 1H); 7.40 – 7.37 (m, 2H); 5.92 (d, *J* = 7.3 Hz, 1H); 4.74 (dd, *J* = 7.4, *J* = 4.7 Hz, 1H); 4.60 (d, *J* = 4.6 Hz, 1H); 4.48 (dd, *J* = 7.4, *J* = 4.8 Hz, 1H); 4.43 (t, *J* = 6.6 Hz, 2H); 4.36 (s, 1H); 3.99 (s, 3H); 3.38 (s, 3H); 3.08 (t, *J* = 6.6 Hz, 2H); 2.21 (pd, *J* = 6.8, *J* = 4.9 Hz, 1H); 0.98 (d, *J* = 6.8 Hz, 3H); 0.94 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.3; 155.7; 153.9; 151.1; 149.5; 146.9; 145.7; 141.6; 129.9; 123.9; 123.8; 89.7; 88.2; 82.3; 70.6; 64.4; 63.4; 59.8; 59.0; 35.0; 34.9; 30.8; 19.6; 18.1. FTIR *v*_{max} (cm⁻¹): 2963 (w); 1738 (m); 1682 (m); 1571 (s); 1517 (s); 1465 (m); 1344 (s); 1264 (m); 1186 (m); 1017 (s); 856 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₆H₃₄O₉N₇ 588.2412; Found 588.2407.

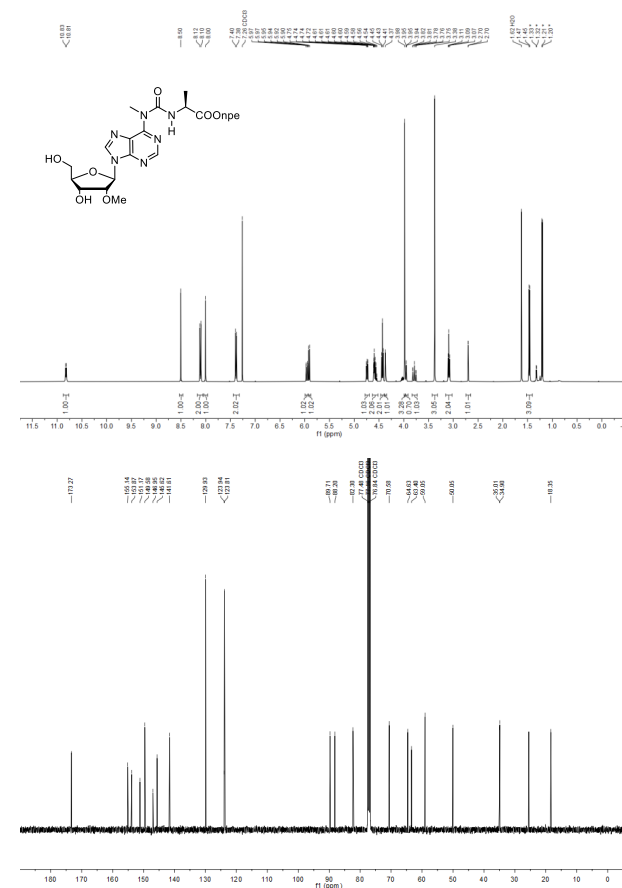


S27

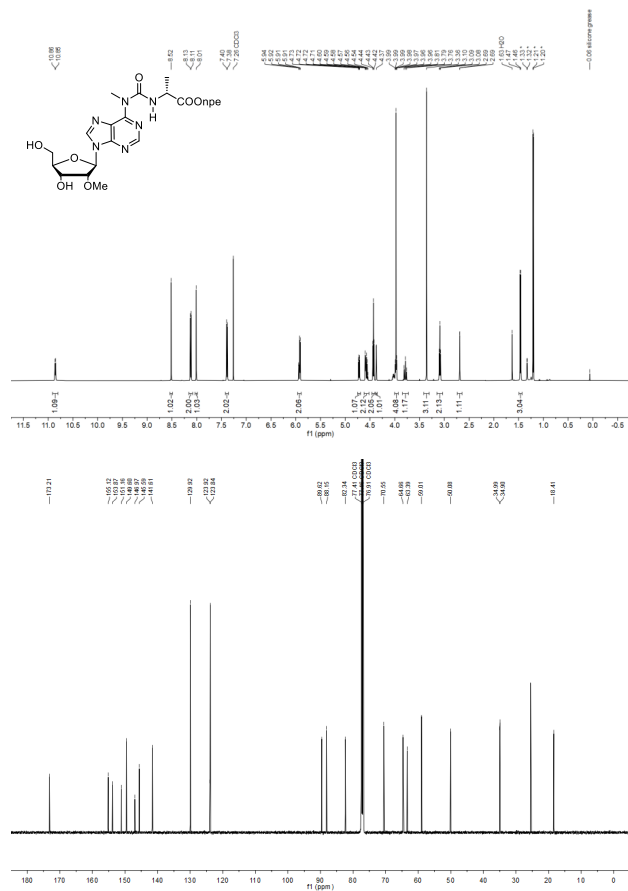
5a_b: Yield = 94%. R_f = 0.39 (95:5 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.96 (d, *J* = 7.5 Hz, 1H); 8.51 (s, 1H); 8.15 – 8.08 (m, 2H); 8.01 (s, 1H); 7.43 – 7.35 (m, 2H); 5.95 – 5.86 (m, 2H); 4.73 (dd, *J* = 7.4 Hz, *J* = 4.7 Hz, 1H); 4.60 (d, *J* = 4.7 Hz, 1H); 4.48 (dd, *J* = 7.6 Hz, *J* = 4.7 Hz, 1H); 4.43 (t, *J* = 6.6 Hz, 2H); 4.38 (s, 1H); 4.00 (s, 3H); 3.82 – 3.75 (m, 1H); 3.36 (s, 3H); 3.09 (t, *J* = 6.6 Hz, 2H); 2.64 (d, *J* = 1.5 Hz, 1H); 2.28 – 2.16 (m, 1H); 0.99 (d, *J* = 6.8 Hz, 3H); 0.94 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.2; 155.8; 154.0; 151.2; 149.5; 145.7; 141.6; 129.9; 123.8; 89.7; 88.2; 82.4; 77.4; 70.6; 64.5; 63.4; 59.8; 59.0; 35.0; 30.9; 25.5; 19.6; 18.2. FTIR ν_{max} (cm⁻¹): 2961 (w); 1735 (m); 1684 (m); 1570 (s); 1518 (s); 1465 (m); 1344 (s); 1263 (m); 1185 (m); 1017 (s); 856 (m). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₂₆H₃₄O₉N₇; 588.2412; Found 588.2411.



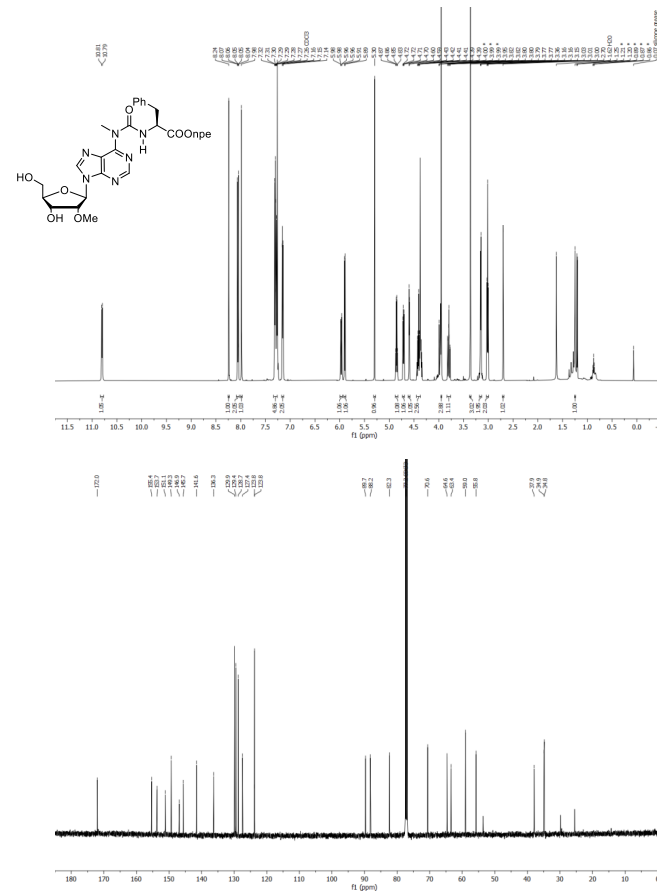
5b_L: Yield = 84%. R_f = 0.20 (95:5 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.82 (d, *J* = 6.6 Hz, 1H); 8.50 (s, 1H); 8.12 – 8.10 (m, 2H); 8.00 (s, 1H); 7.40 – 7.38 (m, 2H); 5.96 (dd, *J* = 11.7 Hz, *J* = 1.5 Hz, 1H); 5.91 (d, *J* = 7.4 Hz, 1H); 4.74 (dd, *J* = 7.4 Hz, *J* = 4.7 Hz, 1H); 4.61 – 4.54 (m, 2H); 4.43 (t, *J* = 6.6 Hz, 2H); 4.37 (s, 1H); 3.98 (s, 3H); 3.98 – 3.95 (m, 1H); 3.82 – 3.75 (m, 1H); 3.38 (s, 3H); 3.09 (t, *J* = 6.6 Hz, 2H); 2.70 (d, *J* = 1.5 Hz, 1H); 1.46 (d, *J* = 7.3 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 173.3; 155.1; 153.9; 151.2; 149.6; 146.9; 145.6; 141.6; 129.9; 123.9; 123.8; 89.7; 88.2; 82.3; 70.6; 64.6; 63.4; 59.0; 50.0; 35.0; 34.9; 18.3. FTIR ν_{max} (cm⁻¹): 2925 (w); 1738 (m); 1674 (m); 1568 (s); 1514 (s); 1343 (s); 1265 (m); 1176 (m); 1020 (m); 855 (m). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₂₄H₃₀O₉N₇; 560.2099; Found 560.2099.



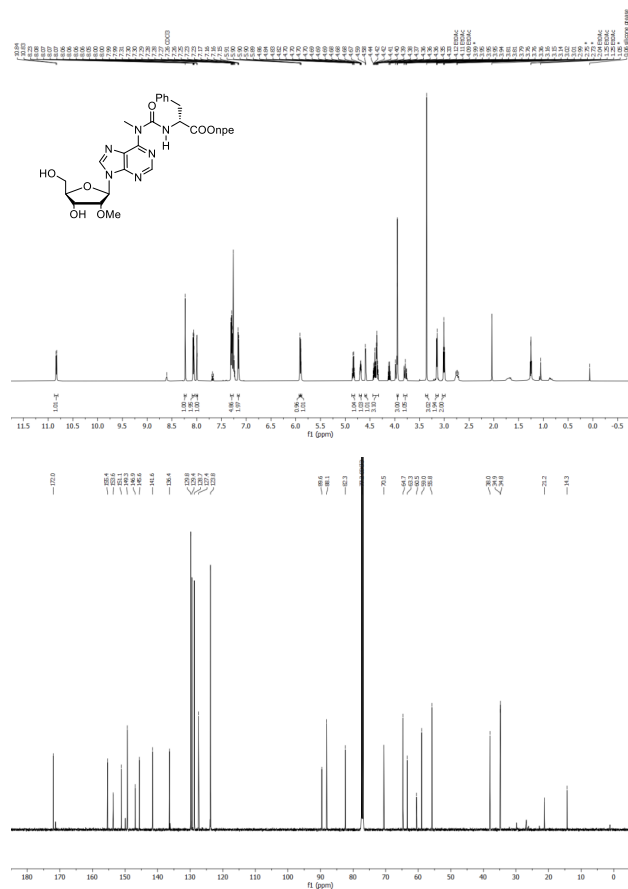
5b_o: Yield = 86%. R_f = 0.14 (95:5 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.86 (d, *J* = 6.6 Hz, 1H); 8.52 (s, 1H); 8.13 – 8.11 (m, 2H); 8.01 (s, 1H); 7.40 – 7.38 (m, 2H); 5.94 – 5.91 (m, 2H); 4.72 (dd, *J* = 7.4 Hz, *J* = 4.7 Hz, 1H); 4.60 – 4.54 (m, 2H); 4.43 (t, *J* = 6.6 Hz, 2H); 4.37 (s, 1H); 3.99 – 3.96 (m, 4H); 3.81 – 3.76 (m, 1H); 3.36 (s, 3H); 3.09 (t, *J* = 6.6 Hz, 2H); 2.69 (d, *J* = 1.5 Hz, 1H); 1.46 (d, *J* = 7.2 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 173.2; 155.1; 153.9; 151.2; 149.6; 147.0; 145.6; 141.6; 130.0; 123.9; 123.8; 89.6; 88.2; 82.3; 70.6; 64.7; 63.4; 59.0; 50.0; 35.0; 34.9; 18.4. FTIR ν_{max} (cm⁻¹): 2934 (w); 1740 (w); 1676 (m); 1571 (m); 1516 (s); 1466 (m); 1344 (s); 1266 (m); 1022 (m); 856 (m); 746 (m); 647 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₄H₃₀O₉N₇; 560.2099; Found 560.2099.



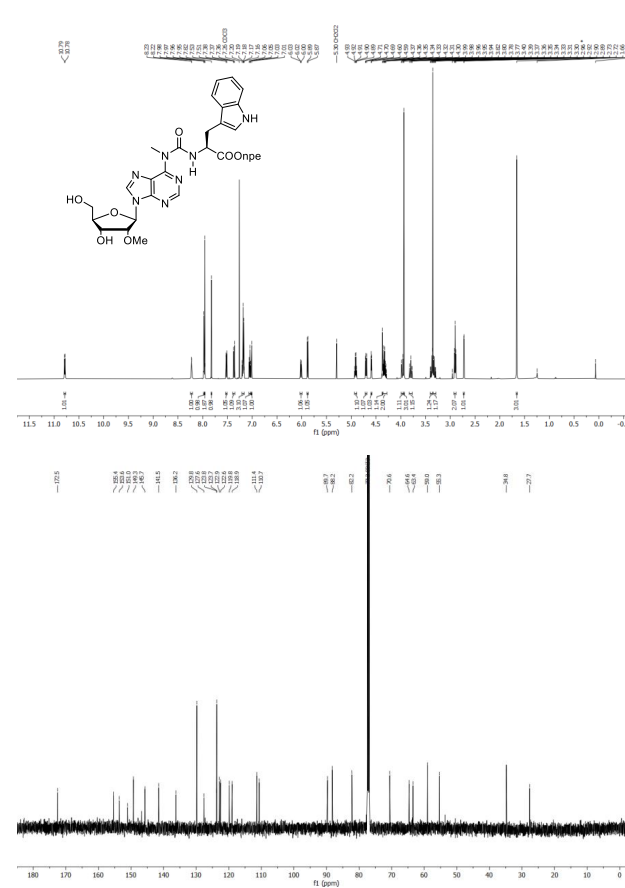
5c_l: Yield = 99%. R_f = 0.36 (95:5 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.80 (d, *J* = 6.8 Hz, 1H); 8.24 (s, 1H); 8.10 – 8.03 (m, 2H); 7.98 (s, 1H); 7.33 – 7.27 (m, 5H); 7.17 – 7.13 (m, 2H); 5.97 (dd, *J* = 11.7 Hz, *J* = 2.1 Hz, 1H); 5.90 (d, *J* = 7.4 Hz, 1H); 5.30 (s, 1H); 4.85 (q, *J* = 6.4 Hz, 1H); 4.71 (dd, *J* = 7.4 Hz, *J* = 4.6 Hz, 1H); 4.60 (d, *J* = 4.6 Hz, 1H); 4.47 – 4.38 (m, 1H); 3.95 (s, 3H); 3.80 (ddd, *J* = 13.1 Hz, *J* = 11.7 Hz, *J* = 1.6 Hz, 1H); 3.36 (s, 3H); 3.19 – 3.10 (m, 2H); 3.01 (t, *J* = 6.5 Hz, 2H); 2.70 (d, *J* = 1.6 Hz, 1H); 1.25 (s, 1H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.0; 155.4; 153.7; 151.1; 149.3; 146.9; 145.7; 141.6; 136.3; 129.9; 129.4; 128.7; 127.4; 123.8; 123.8; 89.7; 88.2; 82.3; 70.6; 64.6; 63.4; 59.0; 55.8; 37.9; 34.9; 34.8. FTIR ν_{max} (cm⁻¹): 3108 (w); 2947 (w); 1736 (m); 1686 (m); 1569 (s); 1515 (vs); 1464 (s); 1341 (s); 1272 (m); 1186 (s); 1022 (s); 698 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₀H₃₄O₉N₇; 636.2412; Found 636.2414.



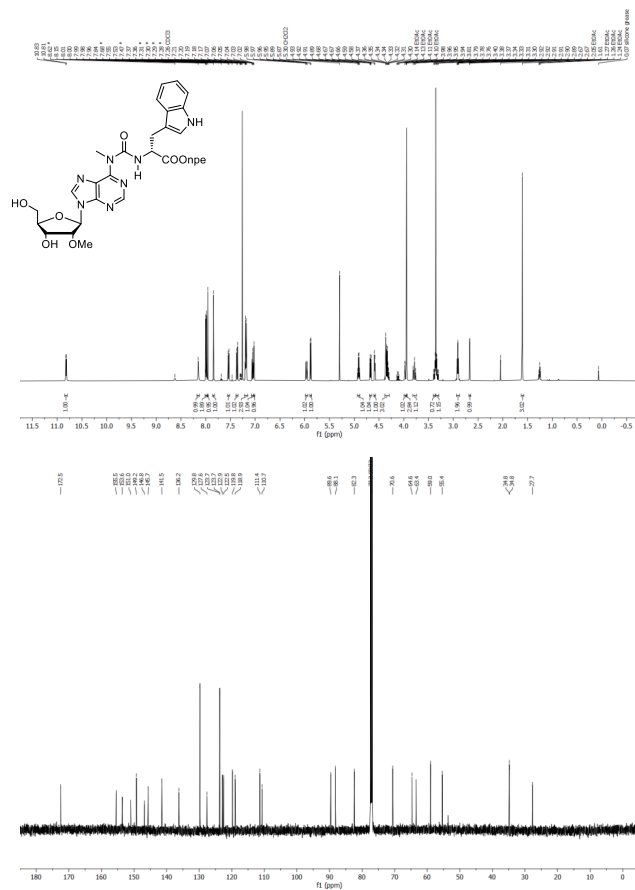
5c_b: Yield = 65%. R_f = 0.28 (95:5 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.83 (d, *J* = 6.7 Hz, 1H); 8.23 (s, 1H); 8.07 (ddd, *J* = 7.2 Hz, *J* = 1.9 Hz, *J* = 1.0 Hz, 2H); 8.00 (dd, *J* = 2.3 Hz, *J* = 1.3 Hz, 1H); 7.33 – 7.22 (m, 5H); 7.18 – 7.13 (m, 2H); 5.91 (s, 1H); 5.91 – 5.88 (m, 1H); 4.84 (q, *J* = 6.4 Hz, 1H); 4.71 – 4.66 (m, 1H); 4.59 (d, *J* = 4.6 Hz, 1H); 4.45 – 4.32 (m, 3H); 3.95 (q, *J* = 0.8 Hz, 3H); 3.84 – 3.75 (m, 1H); 3.36 (s, 3H); 3.18 – 3.12 (m, 2H); 3.01 (t, *J* = 6.6 Hz, 2H). ¹³C NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.0; 155.4; 153.6; 151.1; 149.3; 146.9; 145.6; 141.6; 129.8; 129.4; 128.7; 127.4; 123.8; 89.6; 88.1; 82.3; 70.5; 64.7; 63.3; 60.5; 59.0; 55.8; 38.0; 34.9; 34.8; 21.2; 14.3. FTIR ν_{max} (cm⁻¹): 3381 (w); 2929 (w); 1736 (m); 1677 (m); 1571 (s); 1515 (s); 1466 (s); 1344 (s); 1269 (m); 1178 (m); 1020 (s); 700 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₀H₃₄O₉N₇ 636.2412; Found 636.2419.



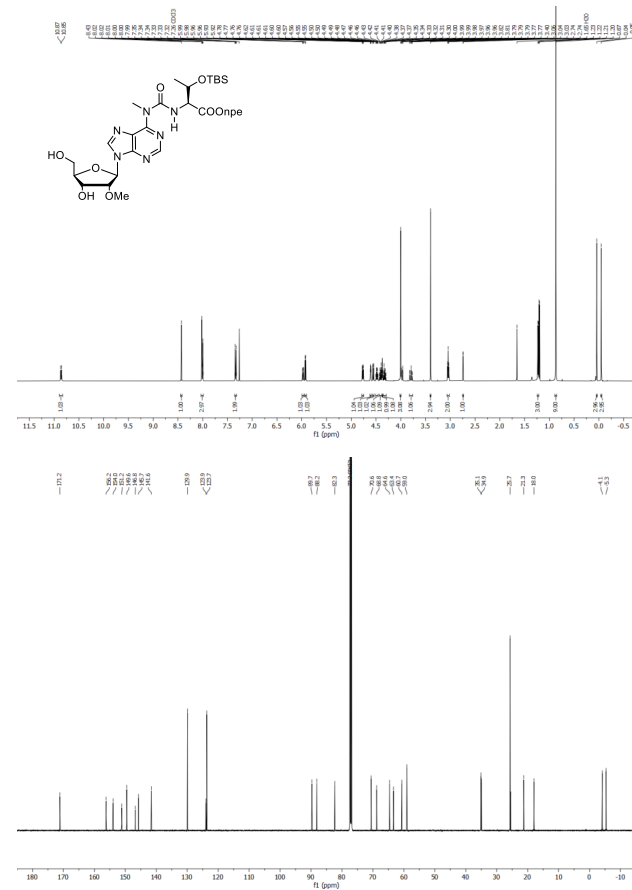
5d_L: Yield = 75%. R_f = 0.34 (95:5 CH₂Cl₂/CH₃OH). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.79 (d, *J* = 6.9 Hz, 1H); 8.25 – 8.20 (m, 1H); 7.97 (d, *J* = 1.9 Hz, 1H); 7.96 (s, 2H); 7.82 (s, 1H); 7.52 (d, *J* = 8.9 Hz, 1H); 7.41 – 7.33 (m, 1H); 7.19 (t, *J* = 8.4 Hz, 3H); 7.05 (t, *J* = 7.5 Hz, 1H); 7.01 (s, 1H); 6.06 – 5.98 (m, 1H); 5.88 (d, *J* = 7.4 Hz, 1H); 4.95 – 4.88 (m, 1H); 4.74 – 4.67 (m, 1H); 4.59 (d, *J* = 4.7 Hz, 1H); 4.37 (d, *J* = 1.9 Hz, 1H); 4.36 – 4.26 (m, 2H); 4.01 – 3.94 (m, 1H); 3.94 (s, 3H); 3.84 – 3.73 (m, 1H); 3.42 – 3.26 (m, 2H); 2.90 (t, *J* = 6.5 Hz, 2H); 2.72 (d, *J* = 1.5 Hz, 1H); 1.66 (s, 3H). ¹³C NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.5; 155.4; 153.6; 151.0; 149.3; 145.7; 141.5; 136.2; 129.8; 127.6; 123.7; 122.9; 122.6; 119.8; 118.9; 111.4; 110.7; 89.7; 88.2; 82.2; 70.6; 64.6; 63.4; 59.0; 55.3; 34.8; 27.7. FTIR ν_{max} (cm⁻¹): 3343 (w); 2934 (w); 1735 (m); 1670 (w); 1571 (m); 1516 (s); 1459 (m); 1342 (s); 1270 (m); 1217 (m); 1181 (m); 1022 (m); 743 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₃₂H₃₆O₉N₈ 675.2521; Found 675.2529.



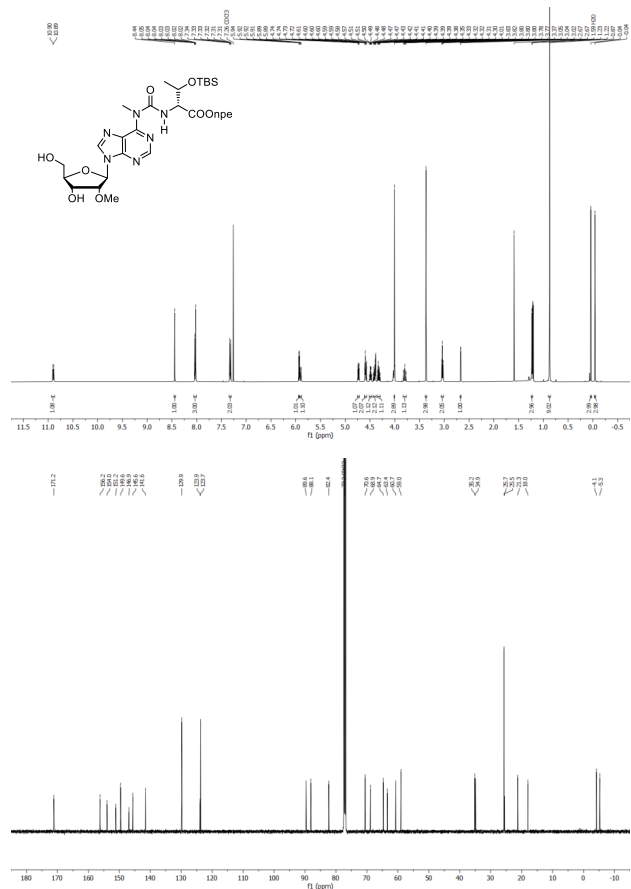
5d_b: Yield = 88%. R_f = 0.32 (95:5 CH₂Cl₂/CH₃OH). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.82 (d, *J* = 6.8 Hz, 1H); 8.16 – 8.13 (m, 1H); 8.01 – 7.98 (m, 2H); 7.96 (s, 1H); 7.84 (s, 1H); 7.54 (d, *J* = 7.9 Hz, 1H); 7.37 (d, *J* = 8.1 Hz, 1H); 7.21 – 7.15 (m, 3H); 7.05 (ddd, *J* = 8.1 Hz, *J* = 7.0 Hz, *J* = 1.0 Hz, 1H); 7.02 (d, *J* = 2.4 Hz, 1H); 5.96 (dd, *J* = 11.6 Hz, *J* = 2.1 Hz, 1H); 5.88 (d, *J* = 7.4 Hz, 1H); 4.91 (q, *J* = 6.2 Hz, 1H); 4.67 (dd, *J* = 7.4 Hz, *J* = 4.6 Hz, 1H); 4.59 (d, *J* = 4.6 Hz, 1H); 4.38 – 4.29 (m, 3H); 4.00 – 3.95 (m, 1H); 3.94 (s, 3H); 3.78 (ddd, *J* = 13.1 Hz, *J* = 11.7 Hz, *J* = 1.7 Hz, 1H); 3.40 – 3.30 (m, 2H); 2.91 (td, *J* = 6.5 Hz, *J* = 1.6 Hz, 2H); 2.67 (d, *J* = 1.5 Hz, 1H); 1.61 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.5; 155.5; 153.6; 151.0; 149.2; 146.8; 145.7; 141.5; 136.2; 129.8; 127.6; 123.7; 122.5; 119.8; 118.9; 111.4; 110.7; 89.6; 88.1; 82.3; 70.6; 64.6; 63.4; 59.0; 55.4; 34.8; 34.8; 27.7. FTIR *V*_{max} (cm⁻¹): 3310 (w); 2932 (w); 1735 (w); 1664 (m); 1571 (m); 1515 (s); 1460 (m); 1344 (s); 1269 (m); 1180 (m); 1089 (m); 1023 (m); 742 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₂H₃₅O₉N₈ 675.2521; Found 675.2533.



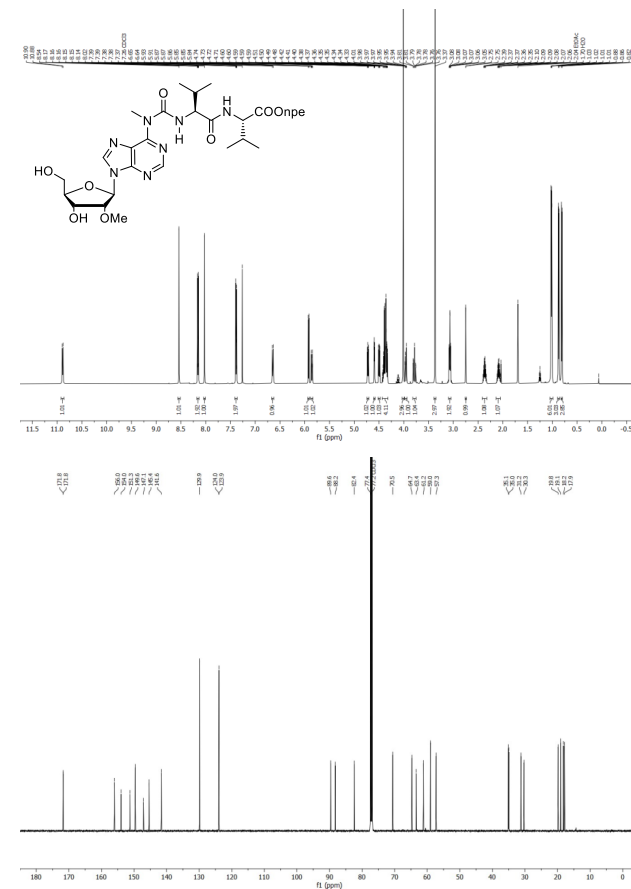
5e_a: Yield = 92%. R_f = 0.20 (96:4 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.86 (d, *J* = 8.6 Hz, 1H); 8.43 (s, 1H); 8.04 – 7.99 (m, 3H); 7.36 – 7.32 (m, 2H); 5.97 (dd, *J* = 11.7 Hz, *J* = 2.1 Hz, 1H); 5.93 (d, *J* = 7.3 Hz, 1H); 4.77 (dd, *J* = 7.3 Hz, *J* = 4.7 Hz, 1H); 4.61 (dt, *J* = 4.7 Hz, *J* = 1.3 Hz, 1H); 4.56 (dd, *J* = 8.6 Hz, *J* = 1.8 Hz, 1H); 4.48 (qd, *J* = 6.3 Hz, *J* = 1.8 Hz, 1H); 4.41 (dt, *J* = 11.1 Hz, *J* = 6.5 Hz, 1H); 4.37 (d, *J* = 1.2 Hz, 1H); 4.32 (dt, *J* = 11.0 Hz, *J* = 6.6 Hz, 1H); 4.00 (s, 3H); 3.79 (ddd, *J* = 13.1 Hz, *J* = 11.7 Hz, *J* = 1.7 Hz, 1H); 3.40 (s, 3H); 3.04 (t, *J* = 6.6 Hz, 2H); 2.74 (d, *J* = 1.6 Hz, 1H); 1.65 (s, 1H); 1.23 (d, *J* = 6.3 Hz, 3H); 0.87 (s, 9H); 0.04 (s, 3H); -0.05 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.2; 156.2; 154.0; 151.2; 149.6; 146.8; 145.7; 141.6; 129.9; 123.9; 123.7; 89.7; 88.2; 82.3; 70.6; 68.8; 64.6; 63.4; 60.7; 59.0; 35.1; 34.9; 25.7; 21.3; 18.0; -4.1; -5.3. FTIR *V*_{max} (cm⁻¹): 2933 (w); 1736 (w); 1678 (w); 1571 (m); 1518 (s); 1463 (m); 1345 (s); 1252 (m); 1088 (s); 1025 (m); 779 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₁H₄₆O₁₀N₇Si 704.3069; Found 704.3105.



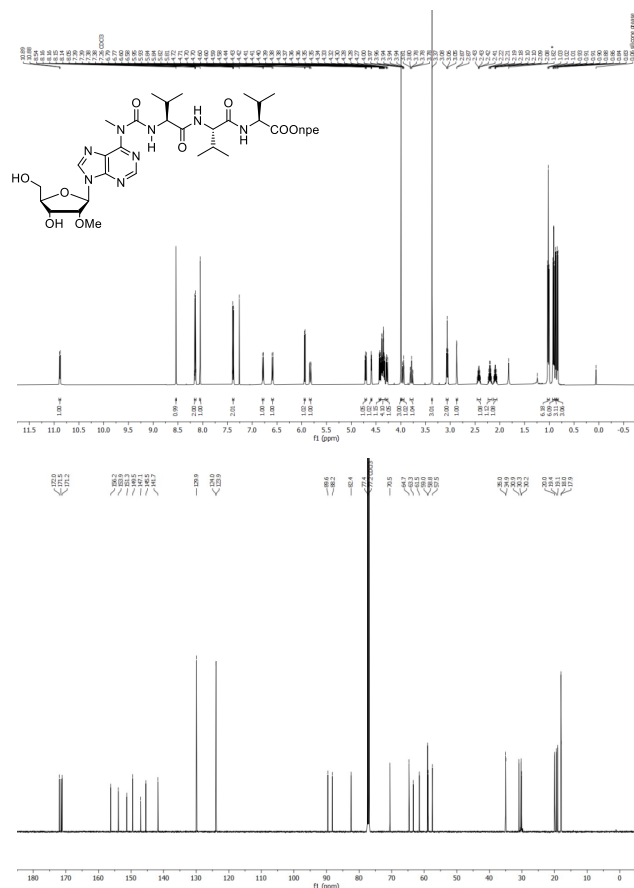
5e: Yield = 97%. Rf = 0.25 (95:5 CH₂Cl₂/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.90 (d, *J* = 8.8 Hz, 1H); 8.44 (s, 1H); 8.06 – 8.00 (m, 3H); 7.36 – 7.30 (m, 2H); 5.93 (d, *J* = 7.3 Hz, 1H); 5.90 (dd, *J* = 11.7 Hz, *J* = 2.1 Hz, 1H); 4.73 (dd, *J* = 7.3 Hz, *J* = 4.6 Hz, 1H); 4.62 – 4.55 (m, 2H); 4.49 (qd, *J* = 6.3 Hz, *J* = 1.8 Hz, 1H); 4.45 – 4.38 (m, 2H); 4.32 (dt, *J* = 11.1 Hz, *J* = 6.7 Hz, 1H); 4.01 (s, 3H); 3.80 (ddd, *J* = 13.1 Hz, *J* = 11.6 Hz, *J* = 1.7 Hz, 1H); 3.37 (s, 3H); 3.04 (t, *J* = 6.6 Hz, 2H); 2.67 (d, *J* = 1.6 Hz, 1H); 1.23 (d, *J* = 6.3 Hz, 3H); 0.87 (s, 9H); 0.04 (s, 3H); -0.04 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.2; 156.2; 154.0; 151.2; 149.6; 146.9; 145.6; 141.6; 129.9; 123.9; 123.7; 89.6; 88.1; 82.4; 70.6; 68.9; 64.7; 63.4; 60.7; 59.0; 35.2; 34.9; 25.7; 25.5; 21.3; 18.0; -4.1; -5.3. FTIR ν_{max} (cm⁻¹): 2932 (w); 1736 (w); 1681 (w); 1571 (m); 1518 (s); 1463 (m); 1345 (s); 1253 (m); 1214 (m); 1181 (m); 1088 (s); 1025 (m); 838 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₃₁H₄₆O₁₀N₇Si 704.3069; Found 704.3105.



5f: Yield = 94%. Rf = 0.46 (95:5 EtOAc/MeOH). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.89 (d, *J* = 7.7 Hz, 1H); 8.54 (s, 1H); 8.18 – 8.14 (m, 2H); 8.02 (s, 1H); 7.41 – 7.36 (m, 2H); 6.65 (d, *J* = 8.5 Hz, 1H); 5.92 (d, *J* = 7.3 Hz, 1H); 5.86 (dt, *J* = 11.5 Hz, *J* = 1.9 Hz, 1H); 4.72 (dd, *J* = 7.3 Hz, *J* = 4.6 Hz, 1H); 4.59 (dd, *J* = 4.6 Hz, *J* = 1.4 Hz, 1H); 4.50 (dd, *J* = 8.5 Hz, *J* = 4.9 Hz, 1H); 4.44 – 4.32 (m, 4H); 4.01 (s, 3H); 3.96 (dt, *J* = 13.0 Hz, *J* = 2.0 Hz, 1H); 3.78 (ddd, *J* = 13.1 Hz, *J* = 11.6 Hz, *J* = 1.7 Hz, 1H); 3.37 (s, 3H); 3.07 (td, *J* = 6.7 Hz, *J* = 1.5 Hz, 2H); 2.75 (d, *J* = 1.6 Hz, 1H); 2.36 (dq, *J* = 13.3 Hz, *J* = 6.4 Hz, 1H); 2.08 (pd, *J* = 6.9 Hz, *J* = 5.0 Hz, 1H); 1.02 (dd, *J* = 6.8 Hz, *J* = 2.4 Hz, 6H); 0.87 (d, *J* = 6.9 Hz, 3H); 0.81 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.8; 171.8; 156.0; 154.0; 151.3; 149.6; 147.1; 145.4; 141.6; 129.9; 124.0; 123.9; 89.6; 88.2; 82.4; 77.4; 70.5; 64.7; 63.4; 61.2; 59.0; 57.3; 35.1; 35.0; 31.2; 30.3; 19.8; 19.1; 18.2; 17.9. FTIR ν_{max} (cm⁻¹): 3285 (w); 3111 (w); 2963 (w); 1743 (m); 1696 (m); 1656 (s); 1568 (m); 1457 (m); 1347 (s); 1112 (s); 1066 (m); 1022 (m); 645 (s). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₁H₄₂O₁₀N₆Na 709.2916; Found 709.2904.



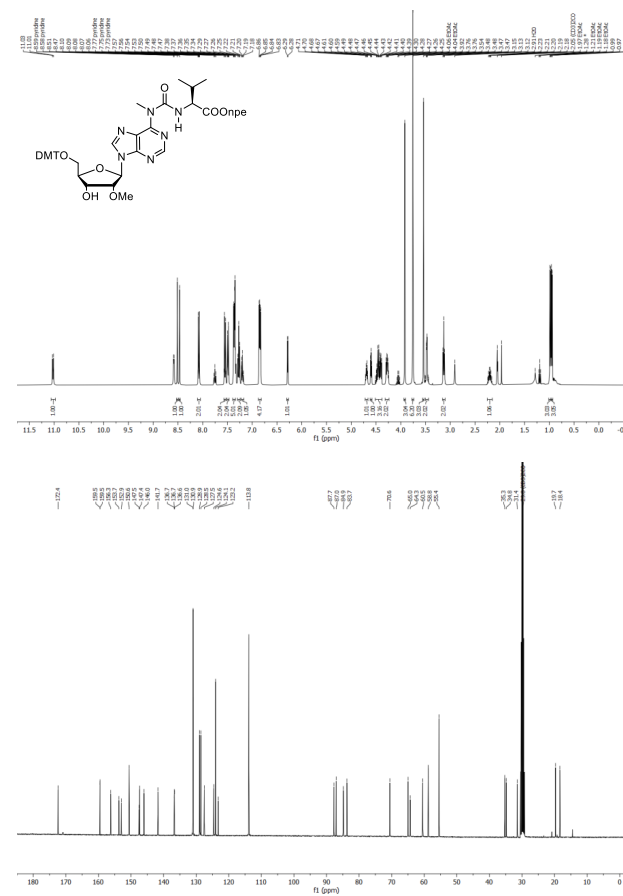
5g₁: Yield = 92%. R_f = 0.43 (9:1 DCM/IPA). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.88 (d, J = 7.1 Hz, 1H); 8.54 (s, 1H); 8.18 – 8.14 (m, 2H); 8.05 (s, 1H); 7.40 – 7.37 (m, 2H); 6.78 (d, J = 8.6 Hz, 1H); 6.59 (d, J = 8.6 Hz, 1H); 5.94 (d, J = 7.3 Hz, 1H); 5.83 (dd, J = 11.5 Hz, J = 2.2 Hz, 1H); 4.71 (dd, J = 7.3 Hz, J = 4.7 Hz, 1H); 4.60 – 4.58 (m, 1H); 4.42 (dd, J = 8.5 Hz, J = 5.2 Hz, 1H); 4.40 – 4.32 (m, 4H); 4.28 (dd, J = 8.6 Hz, J = 6.5 Hz, 1H); 4.00 (s, 3H); 3.95 (dt, J = 12.9 Hz, J = 2.0 Hz, 1H); 3.78 (ddd, J = 13.0 Hz, J = 11.5 Hz, J = 1.7 Hz, 1H); 3.37 (s, 3H); 3.06 (t, J = 6.7 Hz, 2H); 2.87 (d, J = 1.9 Hz, 1H); 2.42 (pd, J = 6.8 Hz, J = 5.2 Hz, 1H); 2.20 (h, J = 6.7 Hz, 1H); 2.09 (pd, J = 6.9 Hz, J = 5.2 Hz, 1H); 1.02 (t, J = 7.1 Hz, 6H); 0.91 (dd, J = 8.1 Hz, J = 6.8 Hz, 6H); 0.87 (d, J = 6.9 Hz, 3H); 0.83 (d, J = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.0; 171.5; 171.2; 156.2; 153.9; 151.3; 149.5; 147.1; 145.5; 141.7; 129.9; 124.0; 123.9; 89.6; 88.2; 82.4; 77.4; 70.5; 64.7; 63.3; 61.5; 59.0; 58.8; 57.5; 35.0; 34.9; 30.9; 30.3; 30.2; 20.0; 19.4; 19.1; 18.0; 17.9. FTIR *v*_{max} (cm⁻¹): 3274 (w); 2962 (w); 1644 (s); 1516 (s); 1344 (m); 1114 (m); 1022 (m); 645 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₆H₅₁O₁₁N₉Na 808.3600; Found 808.3582.



S38

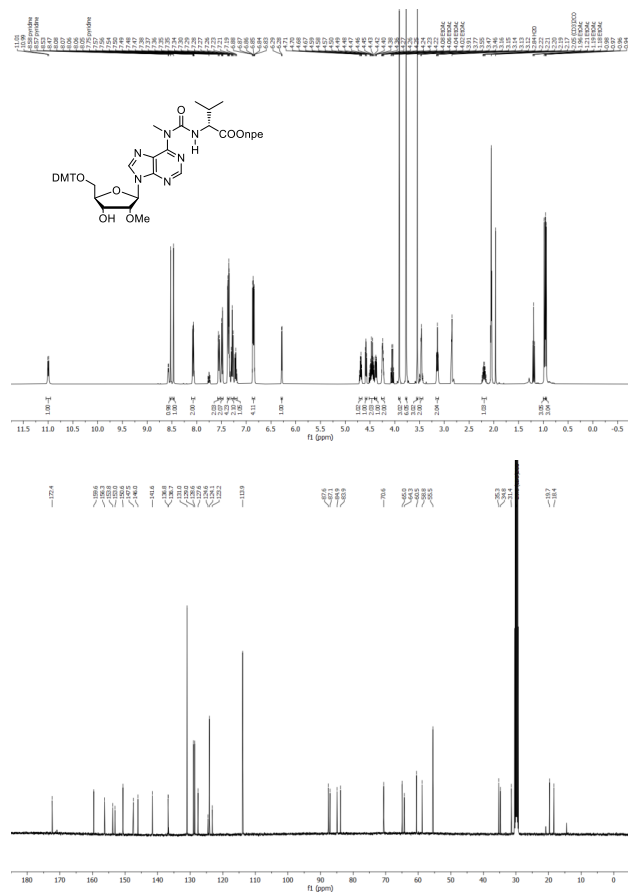
General procedure for the synthesis of compound 6: Compound **5** (1 equiv.) was dissolved in dry pyridine and stirred under nitrogen atmosphere at r.t. 4,4-Dimethoxytrityl chloride (1.5 equiv.) was added in two portions and the reaction was stirred at r.t. overnight. After that, the crude was concentrated and purified by silica gel column chromatography (eluent containing 0.1% pyridine) affording the product as a white foam.

6a₁: Yield = 90%. R_f = 0.25 (8:2 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 11.02 (d, J = 7.7 Hz, 1H); 8.51 (s, 1H); 8.47 (s, 1H); 8.11 – 8.05 (m, 2H); 7.58 – 7.52 (m, 2H); 7.52 – 7.46 (m, 2H); 7.36 (dt, J = 9.1, J = 3.3 Hz, 4H); 7.27 (dd, J = 8.4 Hz, J = 6.7 Hz, 2H); 7.23 – 7.15 (m, 1H); 6.85 (dd, J = 8.9 Hz, J = 3.7 Hz, 4H); 6.28 (d, J = 3.9 Hz, 1H); 4.69 (q, J = 5.7 Hz, 1H); 4.60 (t, J = 4.5 Hz, 1H); 4.53 – 4.36 (m, 3H); 4.32 – 4.24 (m, 2H); 3.92 (s, 3H); 3.76 (d, J = 1.1 Hz, 6H); 3.54 (s, 3H); 3.51 – 3.44 (m, 2H); 3.13 (t, J = 6.4 Hz, 2H); 2.20 (pd, J = 6.9 Hz, J = 4.8 Hz, 1H); 0.98 (d, J = 6.9 Hz, 3H); 0.95 (d, J = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.4; 159.5; 156.3; 153.7; 152.9; 150.6; 147.5; 147.4; 146.0; 141.7; 136.7; 136.6; 131.0; 130.9; 128.9; 128.5; 127.5; 124.6; 124.1; 123.2; 113.8; 87.7; 87.0; 84.9; 83.7; 70.6; 65.0; 64.3; 60.5; 58.8; 55.4; 35.3; 34.8; 31.4; 19.7; 18.4. FTIR *v*_{max} (cm⁻¹): 2931 (w); 1738 (w); 1684 (m); 1571 (m); 1509 (s); 1464 (m); 1345 (s); 1249 (s); 1176 (s); 1018 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₄₇H₅₂O₁₁N₇ 890.3719; Found 890.3696.

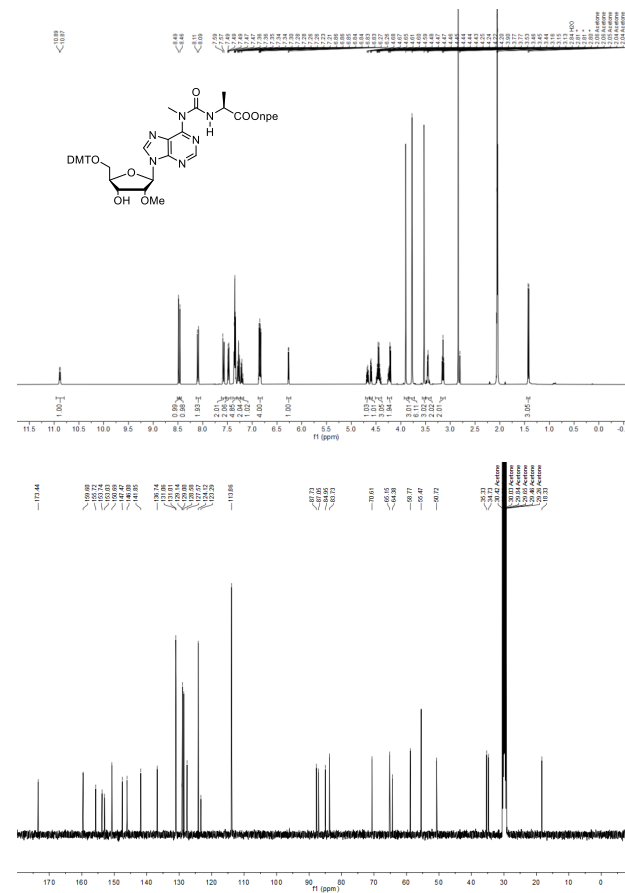


S39

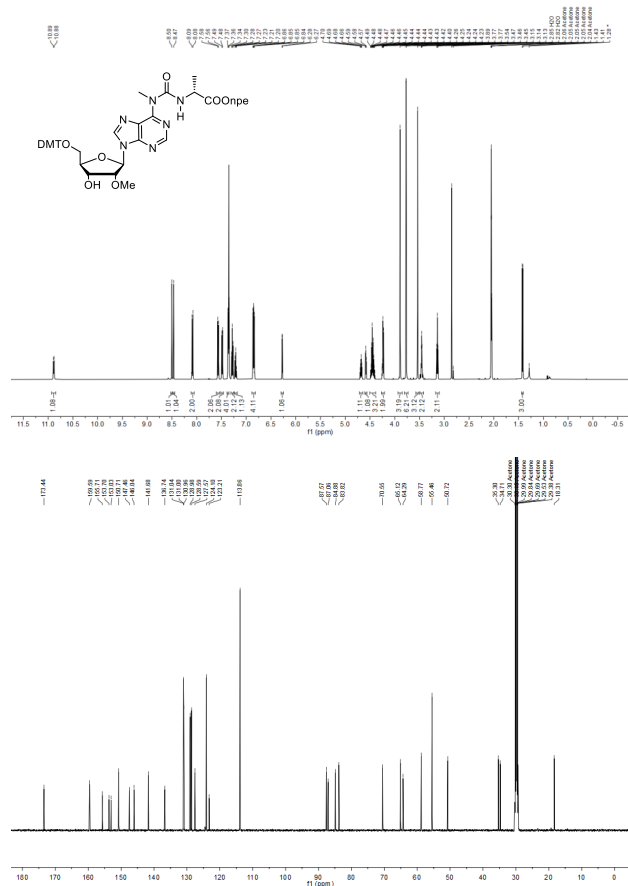
6a_b: Yield = 83%. R_f = 0.20 (8:2 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 11.00 (d, *J* = 7.7 Hz, 1H); 8.53 (s, 1H); 8.47 (s, 1H); 8.11 – 8.03 (m, 2H); 7.59 – 7.52 (m, 2H); 7.51 – 7.47 (m, 2H); 7.39 – 7.33 (m, 4H); 7.28 (dd, *J* = 8.4 Hz, *J* = 6.6 Hz, 2H); 7.24 – 7.18 (m, 1H); 6.89 – 6.80 (m, 4H); 6.28 (d, *J* = 3.9 Hz, 1H); 4.69 (q, *J* = 5.7 Hz, 1H); 4.58 (dd, *J* = 5.0 Hz, *J* = 4.0 Hz, 1H); 4.47 (ddt, *J* = 14.7 Hz, *J* = 11.1 Hz, *J* = 5.7 Hz, 2H); 4.38 (dd, *J* = 7.7 Hz, *J* = 4.9 Hz, 1H); 4.25 (dq, *J* = 6.8 Hz, *J* = 3.6 Hz, *J* = 2.5 Hz, 2H); 3.91 (s, 3H); 3.77 (d, *J* = 1.2 Hz, 6H); 3.55 (s, 3H); 3.50 – 3.43 (m, 2H); 3.18 – 3.11 (m, 2H); 2.20 (pd, *J* = 6.9 Hz, *J* = 4.9 Hz, 1H); 0.97 (d, *J* = 6.9 Hz, 3H); 0.95 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.4; 159.6; 156.3; 153.8; 153.0; 150.6; 147.5; 146.0; 141.6; 136.8; 136.7; 131.0; 129.0; 128.6; 127.6; 124.6; 124.1; 123.2; 113.9; 87.6; 87.1; 84.9; 83.9; 70.6; 65.0; 64.3; 60.5; 58.8; 55.5; 35.3; 34.8; 31.4; 19.7; 18.4. FTIR ν_{max} (cm⁻¹): 2960 (w); 1737 (w); 1684 (m); 1570 (m); 1509 (s); 1464 (m); 1344 (s); 1248 (s); 1176 (s); 1018 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₄₇H₅₂O₁₁N₇ 890.3719; Found 890.3694.



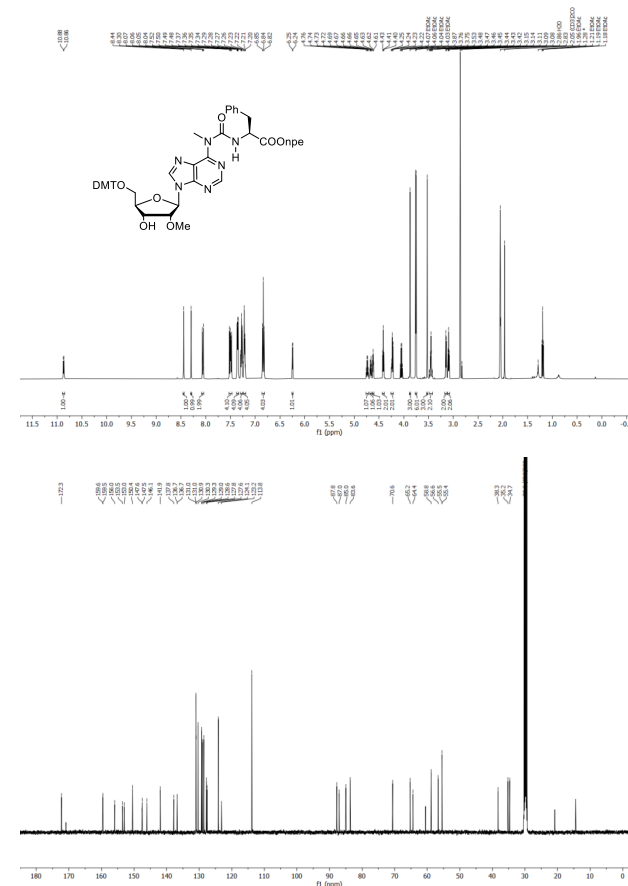
6b₁: Yield = 74%. R_f = 0.20 (8:2 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.88 (d, *J* = 6.7 Hz, 1H); 8.49 (s, 1H); 8.46 (s, 1H); 8.11 – 8.09 (m, 2H); 7.59 – 7.57 (m, 2H); 7.49 (d, *J* = 7.4 Hz, 2H); 7.36 – 7.34 (m, 4H); 7.28 (dd, *J* = 7.4 Hz, *J* = 7.4 Hz, 2H); 7.21 (t, *J* = 7.4 Hz, 1H); 6.86 – 6.83 (m, 4H); 6.27 (d, *J* = 4.1 Hz, 1H); 4.69 – 4.65 (m, 1H); 4.60 (dd, *J* = 4.1 Hz, *J* = 4.1 Hz, 1H); 4.49 – 4.40 (m, 3H); 4.26 – 4.20 (m, 1H); 4.21 (d, *J* = 6.7 Hz, 1H); 3.90 (s, 3H); 3.77 (s, 3H); 3.76 (s, 3H); 3.53 (s, 3H); 3.46 – 3.44 (m, 2H); 3.15 (t, *J* = 7.2 Hz, 2H); 1.42 (d, *J* = 7.2 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 173.4; 159.6; 155.7; 153.7; 153.0; 150.7; 147.5; 146.1; 141.9; 136.7; 131.1; 131.0; 129.1; 129.0; 128.6; 127.6; 124.1; 123.3; 113.9; 87.7; 87.0; 84.9; 83.7; 70.6; 65.2; 64.4; 58.8; 55.5; 50.7; 35.3; 34.7; 18.3. FTIR ν_{max} (cm⁻¹): 2933 (w); 1739 (w); 1682 (m); 1568 (m); 1509 (s); 1463 (m); 1345 (m); 1249 (m); 1176 (m); 1022 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₄₅H₄₆O₁₁N₇ 862.3406; Found 862.3388.



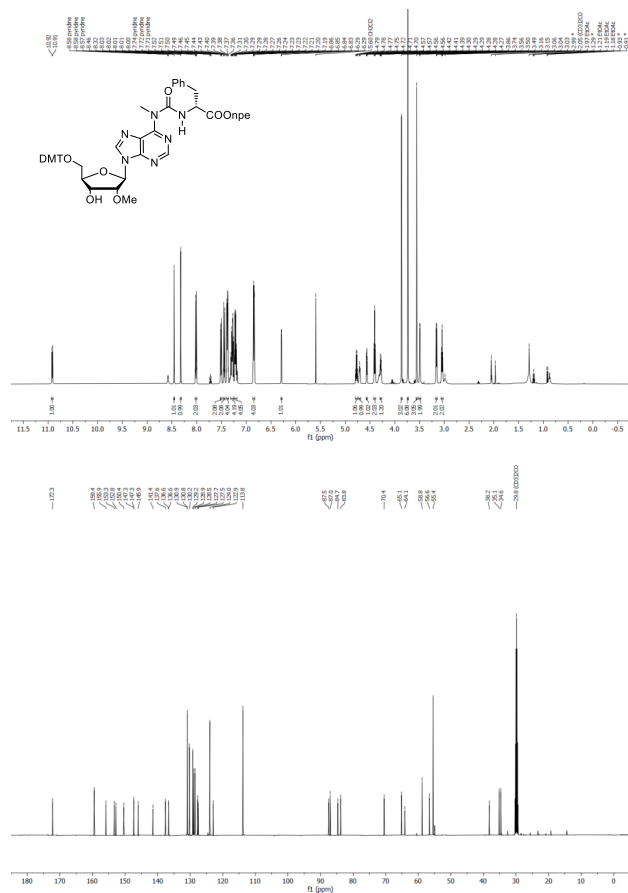
6b_o: Yield = 78% yield. R_f = 0.10 (8:2 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.89 (d, *J* = 6.7 Hz, 1H); 8.50 (s, 1H); 8.47 (s, 1H); 8.09 – 8.08 (m, 2H); 7.58 – 7.56 (m, 2H); 7.49 (d, *J* = 7.7 Hz, 2H); 7.37 – 7.34 (m, 4H); 7.28 (dd, *J* = 7.7 Hz, *J* = 7.7 Hz, 2H); 7.21 (t, *J* = 7.7 Hz, 1H); 6.86 – 6.84 (m, 4H); 6.27 (d, *J* = 4.0 Hz, 1H); 4.70 – 4.66 (m, 1H); 4.58 (dd, *J* = 4.4 Hz, *J* = 4.4 Hz, 1H); 4.49 – 4.40 (m, 3H); 4.26 – 4.23 (m, 2H); 3.89 (s, 3H); 3.77 (s, 3H); 3.76 (s, 3H); 3.54 (s, 3H); 3.47 – 3.45 (m, 2H); 3.14 (t, *J* = 6.3 Hz, 2H); 1.42 (d, *J* = 7.2 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 173.4; 159.6; 155.7; 153.7; 153.0; 150.7; 147.5; 146.0; 141.7; 136.7; 131.1; 131.0; 130.9; 129.0; 128.6; 127.6; 124.1; 123.2; 113.9; 87.6; 87.1; 84.9; 83.8; 70.6; 65.1; 64.3; 58.8; 55.5; 50.7; 35.3; 34.7; 18.3. FTIR ν_{max} (cm⁻¹): 2934 (w); 1742 (w); 1683 (m); 1571 (m); 1509 (s); 1464 (m); 1345 (m); 1250 (s); 1175 (m); 1023 (s); 828 (m); 583 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₄₈H₄₈O₁₁N₇: 862.3406; Found 862.3411.



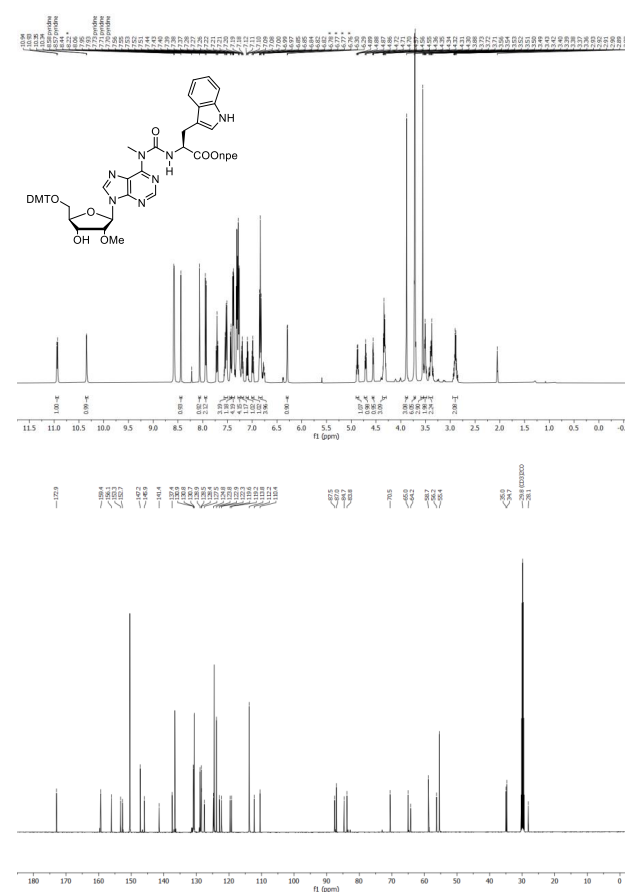
6c_o: Yield = 78%. R_f = 0.50 (1:1 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.87 (d, *J* = 7.0 Hz, 1H); 8.44 (s, 1H); 8.30 (s, 1H); 8.09 – 8.03 (m, 2H); 7.54 – 7.45 (m, 4H); 7.39 – 7.32 (m, 4H); 7.30 – 7.25 (m, 4H); 7.24 – 7.20 (m, 4H); 6.84 (t, *J* = 8.8 Hz, 4H); 6.25 (d, *J* = 4.0 Hz, 1H); 4.74 (q, *J* = 6.5 Hz, 1H); 4.69 – 4.65 (m, 1H); 4.62 (t, *J* = 4.5 Hz, 1H); 4.44 – 4.38 (m, 2H); 4.26 – 4.20 (m, 2H); 3.87 (s, 3H); 3.76 (d, *J* = 4.2 Hz, 6H); 3.53 (s, 3H); 3.49 – 3.42 (m, 2H); 3.15 (d, *J* = 6.2 Hz, 2H); 3.09 (t, *J* = 6.3 Hz, 2H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.3; 159.6; 159.5; 156.0; 153.5; 153.0; 150.4; 147.5; 146.1; 141.9; 137.8; 136.7; 131.0; 130.9; 130.3; 129.3; 129.0; 128.6; 127.8; 127.6; 124.1; 123.2; 113.8; 87.8; 87.0; 85.0; 83.6; 70.6; 65.2; 64.4; 58.8; 56.6; 55.5; 38.3; 35.2; 34.7. FTIR ν_{max} (cm⁻¹): 2932 (w); 1735 (w); 1683 (m); 1571 (m); 1509 (s); 1464 (m); 1345 (m); 1247 (s); 1176 (s); 1029 (s); 699 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₁H₅₂O₁₁N₇: 938.3719; Found 938.3724.



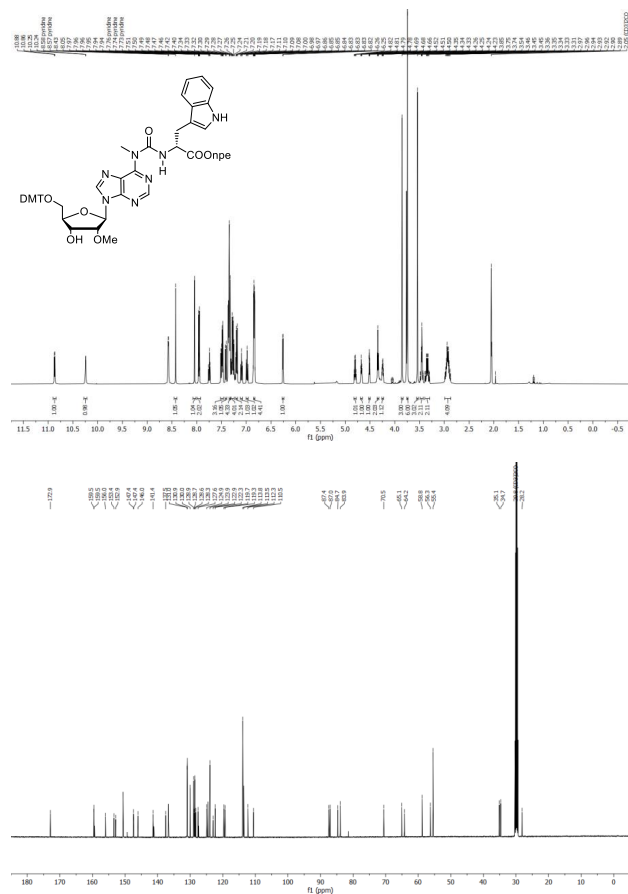
6c: Yield = 89%. Rf = 0.60 (7:3 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 10.92 (d, J = 6.9 Hz, 1H); 8.46 (s, 1H); 8.32 (s, 1H); 8.05 – 7.97 (m, 2H); 7.53 – 7.49 (m, 2H); 7.47 – 7.42 (m, 2H); 7.41 – 7.35 (m, 4H); 7.32 – 7.25 (m, 4H); 7.24 – 7.17 (m, 4H); 6.88 – 6.83 (m, 4H); 6.29 (d, J = 3.8 Hz, 1H); 4.77 (q, J = 6.4 Hz, 1H); 4.70 (d, J = 5.6 Hz, 1H); 4.57 (dd, J = 5.0 Hz, J = 3.8 Hz, 1H); 4.41 (t, J = 6.3 Hz, 2H); 4.28 (dt, J = 5.9 Hz, J = 4.1 Hz, 1H); 3.86 (s, 3H); 3.74 (s, 6H); 3.56 (s, 3H); 3.49 (d, J = 4.1 Hz, 2H); 3.16 (d, J = 6.2 Hz, 2H); 3.04 (t, J = 6.4 Hz, 2H). ¹³C NMR (125 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 172.3; 159.4; 155.9; 153.3; 152.8; 150.4; 147.3; 145.9; 141.4; 137.6; 136.6; 130.9; 130.8; 130.2; 129.2; 128.9; 128.5; 127.7; 127.5; 124.0; 122.9; 113.8; 87.5; 87.0; 84.7; 83.8; 70.4; 65.1; 64.1; 58.8; 56.6; 55.4; 38.2; 35.1; 34.6. FTIR *V*_{max} (cm⁻¹): 2929 (w); 1740 (w); 1683 (w); 1571 (m); 1509 (s); 1464 (m); 1345 (m); 1249 (m); 1176 (m); 1029 (m); 700 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₁H₅₂O₁₁N₇ 938.3719; Found 938.3728.



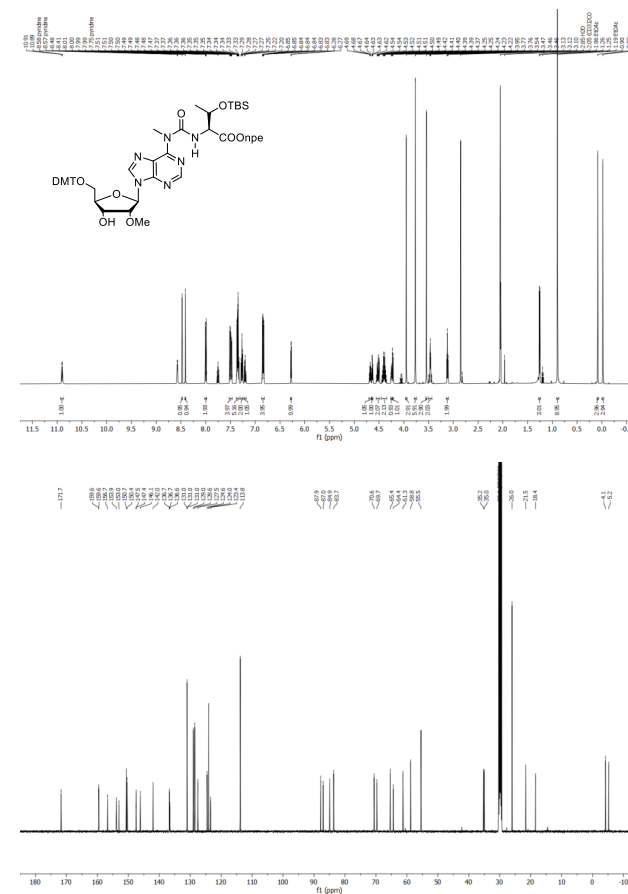
6d: Yield = 98%. Rf = 0.31 (95:5 CH₂Cl₂/CH₃OH). ¹H NMR (500 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 10.94 (d, J = 6.7 Hz, 1H); 10.34 (d, J = 2.6 Hz, 1H); 8.44 (s, 1H); 8.06 (s, 1H); 7.94 (d, J = 8.6 Hz, 2H); 7.58 – 7.48 (m, 3H); 7.43 (d, J = 8.2 Hz, 1H); 7.38 (dd, J = 8.9 Hz, J = 3.1 Hz, 4H); 7.29 – 7.24 (m, 4H); 7.23 – 7.17 (m, 1H); 7.12 – 7.07 (m, 1H); 6.99 (t, J = 7.4 Hz, 1H); 6.84 (t, J = 9.0 Hz, 4H); 6.29 (d, J = 3.7 Hz, 1H); 4.87 (q, J = 6.1 Hz, 1H); 4.71 (t, J = 5.4 Hz, 1H); 4.56 (t, J = 4.3 Hz, 1H); 4.33 (h, J = 4.9 Hz, 3H); 3.88 (s, 3H); 3.72 (d, J = 4.4 Hz, 6H); 3.56 (s, 3H); 3.54 – 3.46 (m, 2H); 3.45 – 3.33 (m, 2H); 2.96 – 2.81 (m, 2H). ¹³C NMR (125 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 172.9; 159.4; 156.1; 153.3; 152.7; 147.2; 145.9; 141.4; 137.4; 130.9; 130.8; 130.7; 128.9; 128.5; 128.4; 127.5; 124.8; 123.8; 122.9; 122.3; 119.6; 119.2; 113.8; 112.2; 110.4; 87.5; 87.0; 84.7; 83.8; 70.5; 65.0; 64.2; 58.7; 56.2; 55.4; 35.0; 34.7; 28.1. FTIR *V*_{max} (cm⁻¹): 2933 (w); 1736 (w); 1678 (w); 1572 (m); 1509 (s); 1344 (m); 1249 (m); 1176 (m); 1030 (s); 744 (s); 701 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₃H₅₃O₁₁N₈ 977.3828; Found 977.3848.



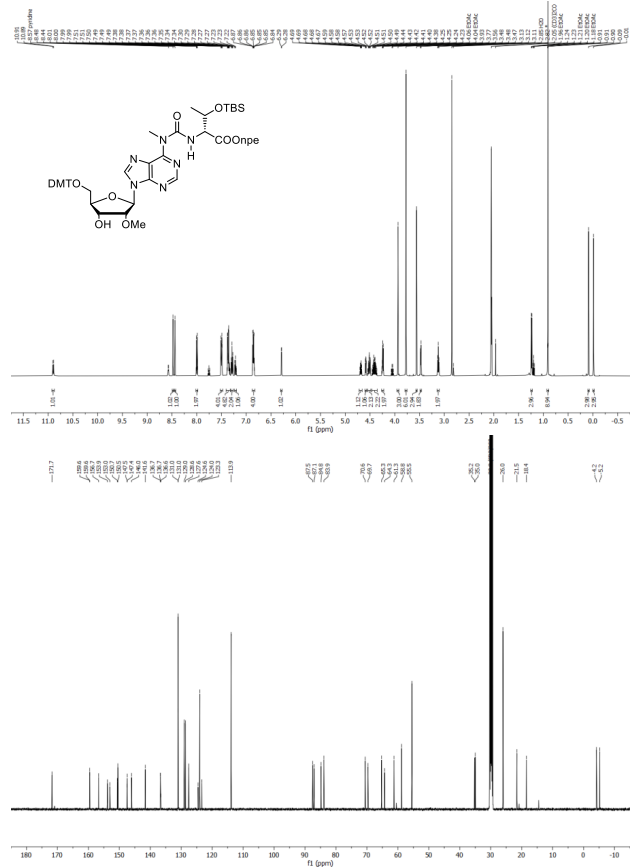
6d_b: Yield = 64%. R_f = 0.43 (95:5 CH₂Cl₂/CH₃OH). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.87 (d, *J* = 6.9 Hz, 1H); 10.24 (s, 1H); 8.43 (s, 1H); 8.05 (s, 1H); 7.99 – 7.91 (m, 2H); 7.53 – 7.45 (m, 3H); 7.41 (d, *J* = 8.1 Hz, 1H); 7.35 – 7.32 (m, 4H); 7.30 – 7.23 (m, 4H); 7.22 – 7.17 (m, 2H); 7.10 (ddd, *J* = 8.2 Hz, *J* = 6.9 Hz, *J* = 1.2 Hz, 1H); 6.98 (ddd, *J* = 7.9 Hz, *J* = 6.9 Hz, *J* = 1.0 Hz, 1H); 6.89 – 6.79 (m, 4H); 6.26 (d, *J* = 3.8 Hz, 1H); 4.83 – 4.78 (m, 1H); 4.68 (t, *J* = 5.5 Hz, 1H); 4.51 (dd, *J* = 5.0 Hz, *J* = 3.8 Hz, 1H); 4.34 (td, *J* = 6.1 Hz, *J* = 1.5 Hz, 2H); 4.24 (td, *J* = 5.1 Hz, *J* = 3.4 Hz, 1H); 3.85 (s, 3H); 3.75 (d, *J* = 1.2 Hz, 6H); 3.54 (s, 3H); 3.50 – 3.41 (m, 2H); 3.40 – 3.29 (m, 2H); 3.00 – 2.85 (m, 4H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.9; 159.5; 156.0; 153.4; 152.9; 147.4; 146.0; 141.4; 137.5; 131.0; 130.9; 130.0; 128.9; 128.7; 128.6; 128.3; 127.6; 124.9; 123.9; 122.9; 122.3; 119.7; 119.3; 113.8; 113.5; 112.3; 110.5; 87.4; 87.0; 84.7; 83.9; 70.5; 65.1; 64.2; 58.8; 56.3; 55.4; 35.1; 34.7; 28.2. FTIR *v*_{max} (cm⁻¹): 2934 (w); 1736 (w); 1673 (w); 1572 (m); 1509 (s); 1460 (m); 1344 (m); 1249 (m); 1176 (m); 1029 (m); 701 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₅₃H₅₃O₁₁N₈ 977.3828; Found 977.3850.



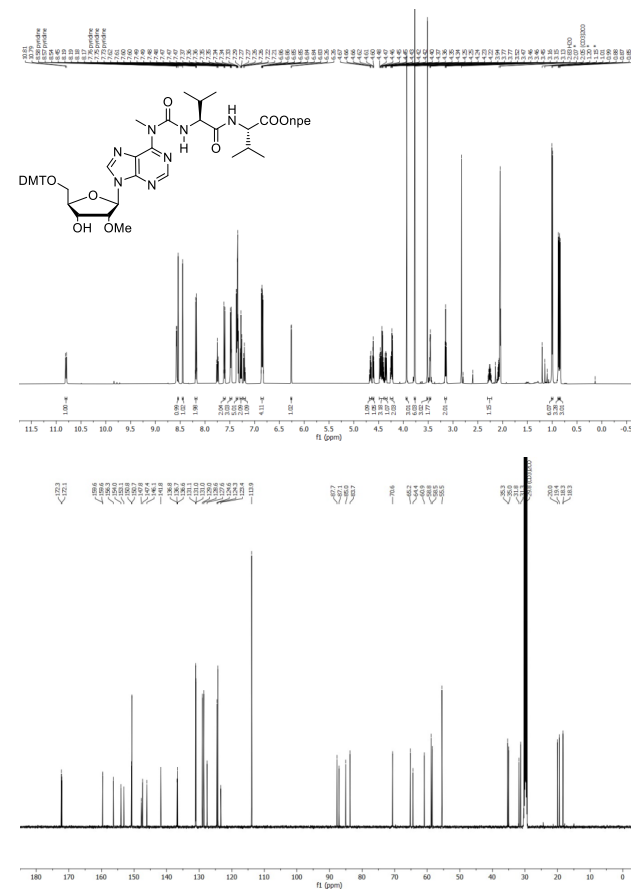
6e_a: Yield = 41%. R_f = 0.22 (9:1 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.90 (d, *J* = 8.7 Hz, 1H); 8.48 (s, 1H); 8.41 (s, 1H); 8.02 – 7.98 (m, 2H); 7.54 – 7.44 (m, 4H); 7.40 – 7.32 (m, 5H); 7.30 – 7.24 (m, 2H); 7.23 – 7.17 (m, 1H); 6.88 – 6.80 (m, 4H); 6.28 (d, *J* = 3.9 Hz, 1H); 4.68 (dt, *J* = 6.6 Hz, *J* = 5.3 Hz, 1H); 4.63 (dd, *J* = 5.1 Hz, *J* = 3.9 Hz, 1H); 4.57 – 4.48 (m, 2H); 4.40 (qt, *J* = 11.1 Hz, *J* = 6.2 Hz, 2H); 4.25 (dd, *J* = 5.3 Hz, *J* = 3.7 Hz, 1H); 4.22 (d, *J* = 6.8 Hz, 1H); 3.95 (s, 3H); 3.77 (d, *J* = 1.2 Hz, 6H); 3.54 (s, 3H); 3.50 – 3.43 (m, 2H); 3.12 (t, *J* = 6.2 Hz, 2H); 1.26 (d, *J* = 6.2 Hz, 2H); 0.90 (s, 9H); 0.08 (s, 3H); -0.02 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 171.7; 159.6; 159.6; 156.7; 153.9; 153.0; 150.7; 150.4; 147.5; 147.4; 146.1; 142.0; 136.7; 136.7; 136.6; 131.0; 131.0; 131.0; 129.0; 128.6; 127.5; 124.6; 124.0; 123.4; 113.8; 87.9; 87.0; 84.9; 83.7; 70.6; 69.7; 65.4; 64.4; 61.3; 58.8; 55.5; 35.2; 35.0; 26.0; 21.5; 18.4; -4.1; -5.2. FTIR *v*_{max} (cm⁻¹): 2932 (w); 2857 (w); 1735 (w); 1684 (m); 1571 (m); 1509 (s); 1463 (m); 1345 (m); 1249 (s); 1176 (m); 1029 (s); 826 (s); 701 (s). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₅₂H₆₃O₁₂N₇Si 1006.4376; Found 1006.4434.



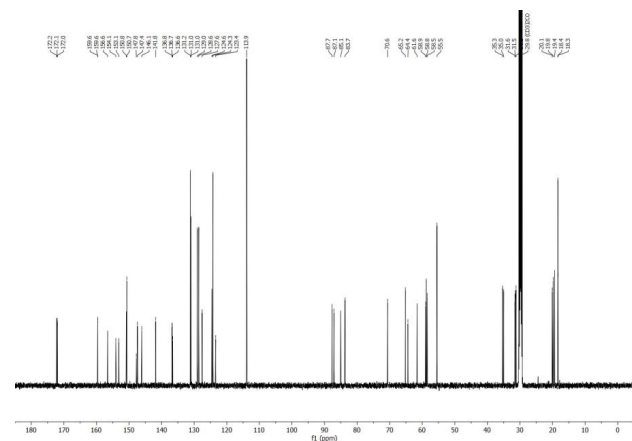
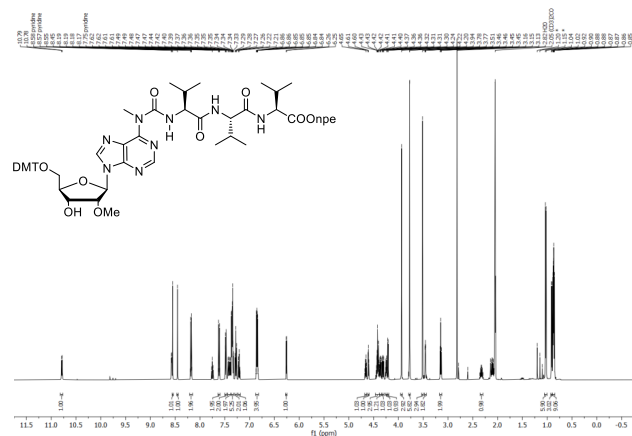
6e: Yield = 71%. Rf = 0.27 (4:1 CH₂Cl₂/EtOAc). ¹H NMR (500 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 10.90 (d, *J* = 8.7 Hz, 1H); 8.48 (s, 1H); 8.44 (s, 1H); 8.02 – 7.97 (m, 2H); 7.55 – 7.45 (m, 4H); 7.41 – 7.33 (m, 5H); 7.29 (t, *J* = 7.6 Hz, 2H); 7.24 – 7.19 (m, 1H); 6.88 – 6.82 (m, 4H); 6.29 (d, *J* = 3.9 Hz, 1H); 4.69 (dt, *J* = 6.8 Hz, *J* = 5.3 Hz, 1H); 4.58 (dd, *J* = 5.0 Hz, *J* = 3.9 Hz, 1H); 4.56 – 4.48 (m, 2H); 4.48 – 4.35 (m, 2H); 4.24 (dd, *J* = 6.3 Hz, *J* = 3.0 Hz, 2H); 3.93 (s, 3H); 3.77 (s, 6H); 3.56 (s, 3H); 3.52 – 3.45 (m, 2H); 3.12 (t, *J* = 6.3 Hz, 2H); 1.24 (d, *J* = 6.2 Hz, 3H); 0.91 (s, 9H); 0.09 (s, 3H); -0.01 (s, 3H). ¹³C NMR (125 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 171.7; 159.6; 159.6; 156.7; 153.9; 153.0; 150.7; 150.5; 147.5; 147.4; 146.0; 141.6; 136.7; 136.7; 136.6; 131.0; 131.0; 129.0; 128.6; 127.6; 124.6; 124.0; 123.3; 113.9; 87.5; 87.1; 84.8; 83.9; 70.6; 69.7; 65.3; 64.3; 61.3; 58.8; 55.5; 35.2; 35.0; 26.0; 21.5; 18.4; -4.2; -5.2. FTIR *v*_{max} (cm⁻¹): 2932 (w); 1736 (w); 1686 (w); 1571 (m); 1509 (s); 1463 (m); 1345 (m); 1249 (s); 1176 (m); 1086 (m); 1028 (s); 826 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₂H₃₃O₁₂N₇Si 1006.4376; Found 1006.4418.



6f: Yield = 99%. Rf = 0.60 (1:1 iHex/Acetone). ¹H NMR (500 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 10.81 (d, *J* = 8.0 Hz, 1H); 8.55 (s, 1H); 8.46 (s, 1H); 8.21 – 8.17 (m, 2H); 7.64 – 7.60 (m, 2H); 7.49 (dt, *J* = 8.4 Hz, *J* = 1.8 Hz, 3H); 7.39 – 7.34 (m, 5H); 7.28 (dd, *J* = 8.5 Hz, *J* = 6.8 Hz, 2H); 7.24 – 7.19 (m, 1H); 6.88 – 6.83 (m, 4H); 6.27 (d, *J* = 4.2 Hz, 1H); 4.67 (dt, *J* = 6.4 Hz, *J* = 5.1 Hz, 1H); 4.62 (t, *J* = 4.6 Hz, 1H); 4.50 – 4.40 (m, 3H); 4.36 (dd, *J* = 8.2 Hz, *J* = 5.7 Hz, 1H); 4.28 – 4.22 (m, 2H); 3.95 (s, 3H); 3.78 (d, *J* = 1.5 Hz, 6H); 3.53 (s, 3H); 3.47 (dd, *J* = 4.3 Hz, *J* = 2.4 Hz, 2H); 3.16 (t, *J* = 6.5 Hz, 2H); 2.27 (pd, *J* = 6.9 Hz, *J* = 5.2 Hz, 1H); 1.01 (d, *J* = 6.8 Hz, 6H); 0.88 (d, *J* = 6.8 Hz, 3H); 0.86 (d, *J* = 6.9 Hz, 3H). ¹³C NMR (125 MHz with cryoprobe, acetone-d₆, 298 K): δ (ppm) = 172.3; 172.1; 159.6; 159.6; 156.3; 154.0; 153.1; 150.8; 150.7; 147.8; 147.4; 146.1; 141.8; 136.8; 136.7; 136.6; 131.1; 131.0; 129.0; 128.6; 127.6; 124.6; 124.3; 123.4; 113.9; 87.7; 87.1; 85.0; 83.7; 70.6; 65.2; 64.4; 60.9; 58.8; 58.5; 55.5; 35.3; 35.0; 31.8; 31.3; 20.0; 19.4; 18.3; 18.3. FTIR *v*_{max} (cm⁻¹): 2963 (w); 1662 (m); 1577 (m); 1509 (s); 1345 (m); 1250 (s); 1177 (m); 1032 (m); 828 (m); 701 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₂₂H₃₀O₁₂N₆Na 1011.4222; Found 1011.4209.

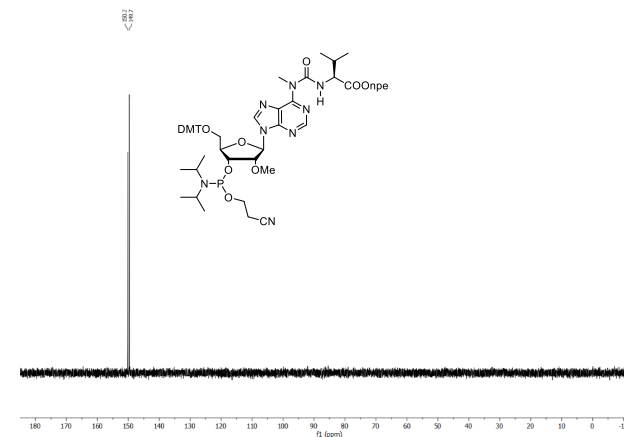


6g_l: Yield = 98%. R_f = 0.27 (3:2 iHex/Acetone). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.78 (d, *J* = 7.5 Hz, 1H); 8.55 (s, 1H); 8.45 (s, 1H); 8.20 – 8.16 (m, 2H); 7.63 – 7.58 (m, 2H); 7.51 – 7.45 (m, 2H); 7.41 (dd, *J* = 16.9 Hz, *J* = 8.5 Hz, 2H); 7.35 (qt, *J* = 5.6 Hz, *J* = 2.4 Hz, 5H); 7.30 – 7.26 (m, 2H); 7.23 – 7.19 (m, 1H); 6.88 – 6.82 (m, 4H); 6.26 (d, *J* = 4.2 Hz, 1H); 4.66 (dt, *J* = 6.5 Hz, *J* = 5.1 Hz, 1H); 4.60 (t, *J* = 4.6 Hz, 1H); 4.46 – 4.39 (m, 3H); 4.39 – 4.34 (m, 1H); 4.31 (dd, *J* = 8.2 Hz, *J* = 5.7 Hz, 1H); 4.26 – 4.23 (m, 1H); 4.21 (d, *J* = 6.5 Hz, 1H); 3.94 (s, 3H); 3.77 (d, *J* = 1.3 Hz, 6H); 3.51 (s, 3H); 3.46 (dd, *J* = 4.3 Hz, *J* = 2.4 Hz, 2H); 3.15 (t, *J* = 6.4 Hz, 2H); 2.33 (pd, *J* = 6.9 Hz, *J* = 5.0 Hz, 1H); 1.03 (d, *J* = 6.8 Hz, 6H); 0.91 (d, *J* = 6.8 Hz, 3H); 0.89 – 0.85 (m, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.2; 172.1; 172.0; 159.6; 159.6; 156.6; 154.1; 153.1; 150.8; 150.7; 147.8; 147.4; 146.1; 141.8; 136.8; 136.7; 136.6; 131.2; 131.0; 131.0; 129.0; 128.6; 127.6; 124.6; 124.3; 123.4; 123.4; 113.9; 87.7; 87.1; 85.1; 83.7; 70.6; 65.2; 64.4; 61.6; 58.9; 58.8; 58.5; 55.5; 35.3; 35.0; 31.6; 31.5; 31.2; 20.1; 19.8; 19.4; 18.4; 18.3. FTIR *v*_{max} (cm⁻¹): 2963 (w); 2361 (w); 1654 (m); 1509 (s); 1345 (m); 1250 (m); 1177 (m); 1032 (m); 703 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₅₇H₆₉O₁₃N₉Na 1110.4906; Found 1110.4897.

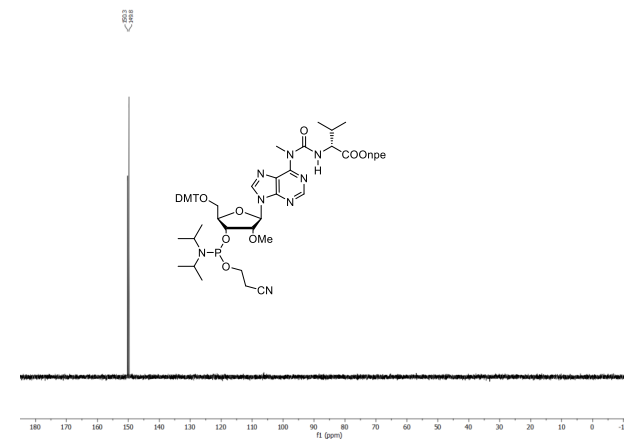


General procedure for the synthesis of compound 7: Compound **6** (1 equiv.) was added to a dry-oven round-bottom flask and dissolved in dry CH₂Cl₂. The solution was stirred under Argon atmosphere at 0°C. DIPEA (4 equiv.) was added dropwise. Finally, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (2.5 equiv.) was added dropwise. The reaction was stirred at r.t. for 3 h. After that, the reaction was stopped and diluted with CH₂Cl₂. The crude was washed with aqueous saturated NaHCO₃ and the organic layer was separated. The crude was further extracted with CH₂Cl₂. The combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography (eluent containing 0.1% pyridine). The products were isolated as a mixture of diastereoisomers as a white foam. Finally, the product was lyophilized from benzene.

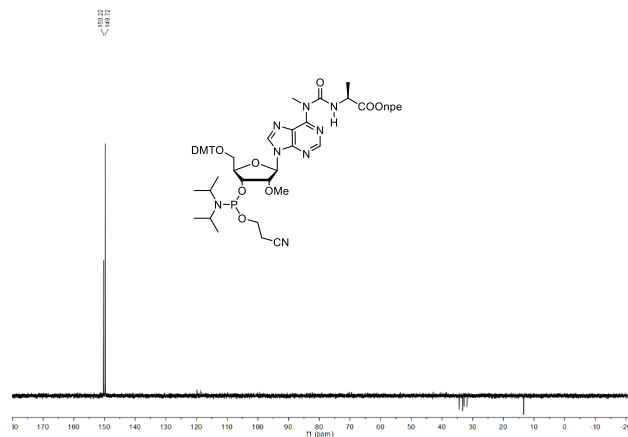
7a_l: Yield = 83%. R_f = 0.70 (3:7 iHex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.7. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₆H₆₈N₉O₁₂P 1090.4797; Found 1090.4759.



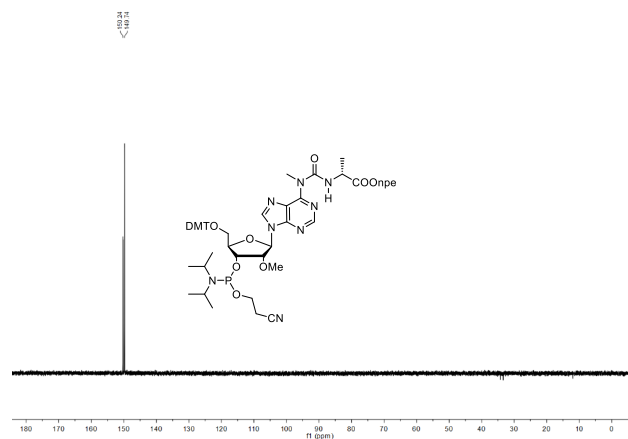
7a_b: Yield = 77%. R_f = 0.50; 0.30 (1:1 iHex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.3; 149.8. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₆H₆₈N₉O₁₂P 1090.4797; Found 1090.4762.



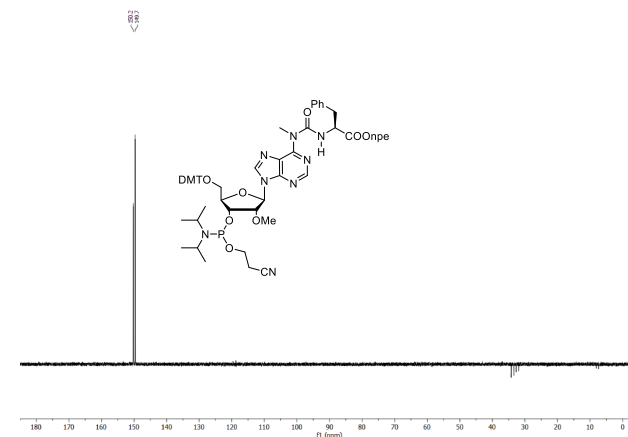
7b_L: Yield = 57%. Rf = 0.40 (8:2 CH₂Cl₂/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.7. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₄H₆₈N₉O₁₂P 1062.4484; Found 1062.4466.



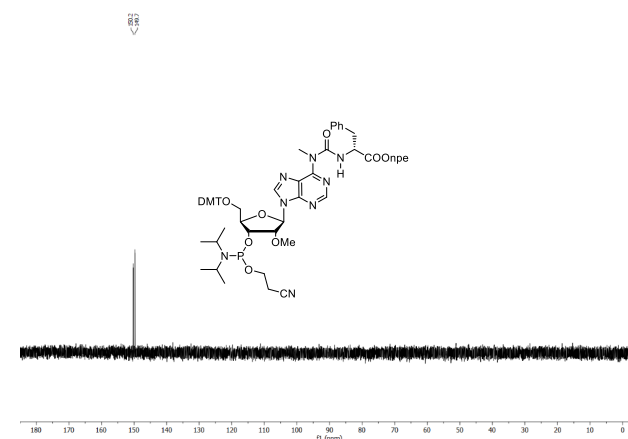
7b_B: Yield = 70%. Rf = 0.26 (8:2 CH₂Cl₂/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.7. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₄H₆₈N₉O₁₂P 1062.4484; Found 1062.4496.



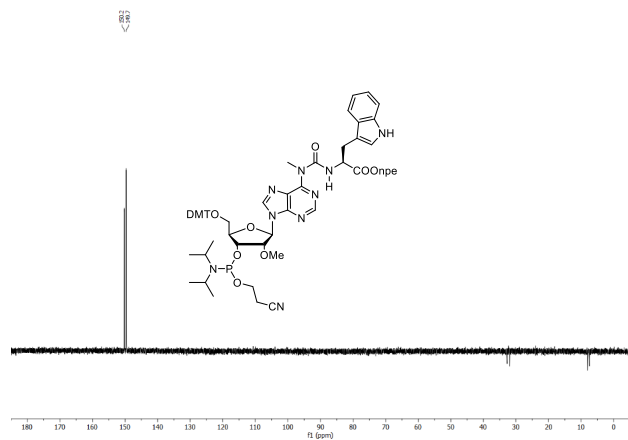
7c_L: Yield = 66%. Rf = 0.30; 0.27 (1:1 Hex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.7. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₆₀H₆₈N₉O₁₂P 1138.4797; Found 1138.4831.



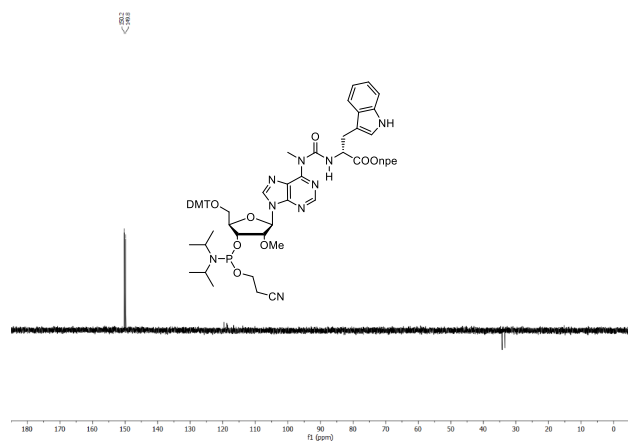
7c_B: Yield = 42%. Rf = 0.31; 0.12 (1:1 iHex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.7. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₆₀H₆₈N₉O₁₂P 1138.4797; Found 1138.4818.



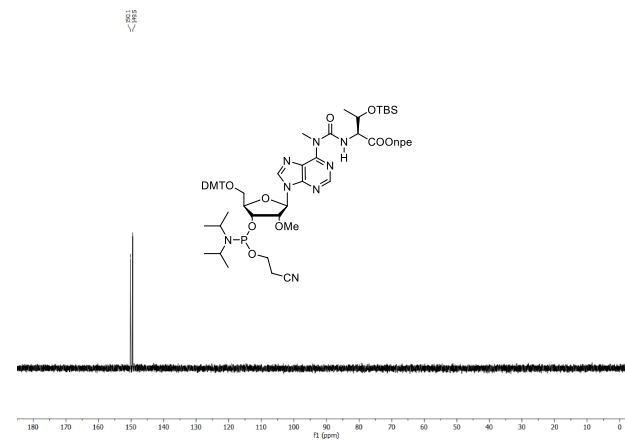
7d_L: Yield = 62%. Rf = 0.27 (1:1 Hex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.7. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₆₂H₇₀N₁₀O₁₂P 1177.4906; Found 1177.4940.



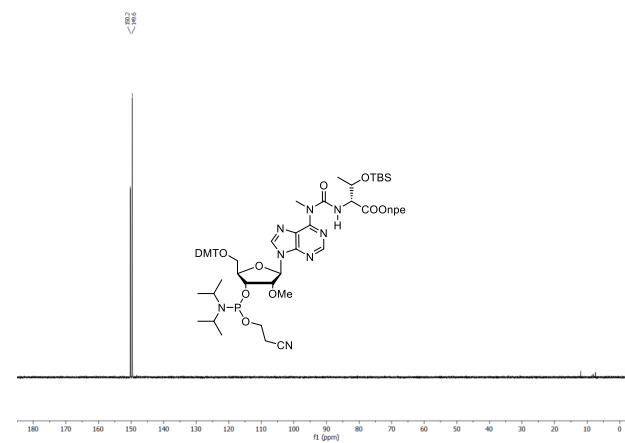
7d_B: Yield = 62%. Rf = 0.30 (1:1 Hex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.8. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₆₂H₇₀N₁₀O₁₂P 1177.4906; Found 1177.4943.



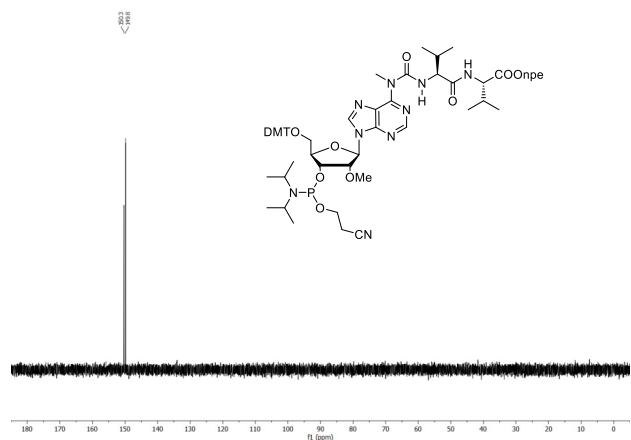
7e_L: Yield = 74%. Rf = 0.35; 0.44 (1:1 Hex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.1; 149.5. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₆₁H₆₇N₉O₁₃SiP 1206.5454; Found 1206.5542.



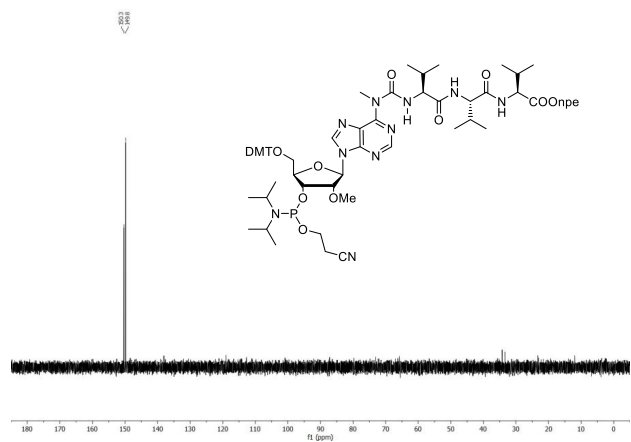
7e_B: Yield = 88%. Rf = 0.34; 0.41 (1:1 Hex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.2; 149.6. HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₆₁H₆₇N₉O₁₃SiP 1206.5454; Found 1206.5512.



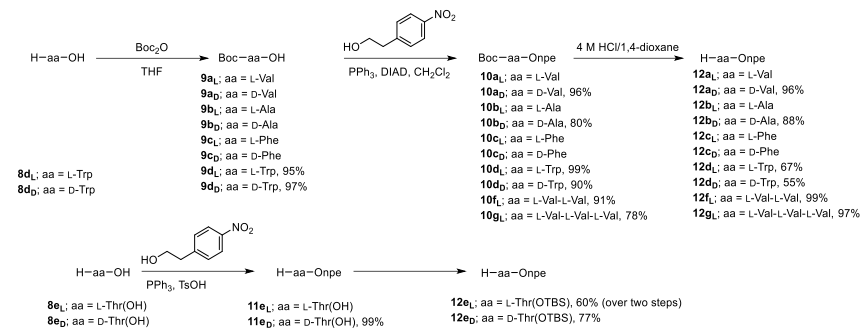
7f_L: Yield = 64%. R_f = 0.35; 0.45 (1:1 iHex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.3; 149.8. HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₆₁H₇₇O₁₃N₁₀PNa 1211.5301; Found 1211.5274.



7g_L: Yield = 79%. R_f = 0.60 (1:4 iHex/EtOAc). ³¹P{¹H} NMR (202 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 150.3; 149.8. HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₆₆H₈₆O₁₄N₁₁PNa 1310.5985; Found 1310.5954.



2.2 Npe-protected amino acids



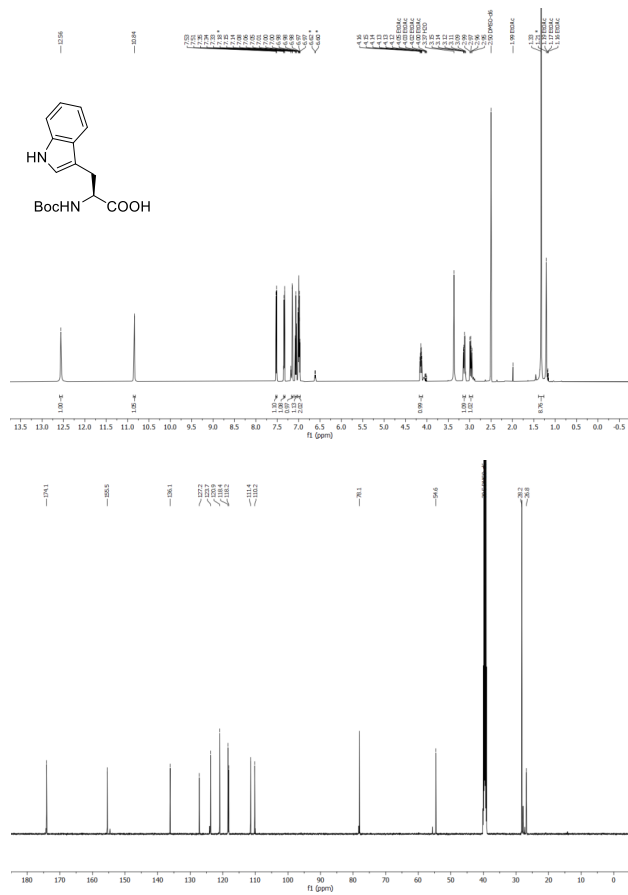
Scheme S2. Synthesis of Onpe-protected amino acids **12·HCl**.

Compounds **9a_L**, **9a_D**, **9b_L**, **9b_D**, **9c_L** and **9c_D** are commercially available.

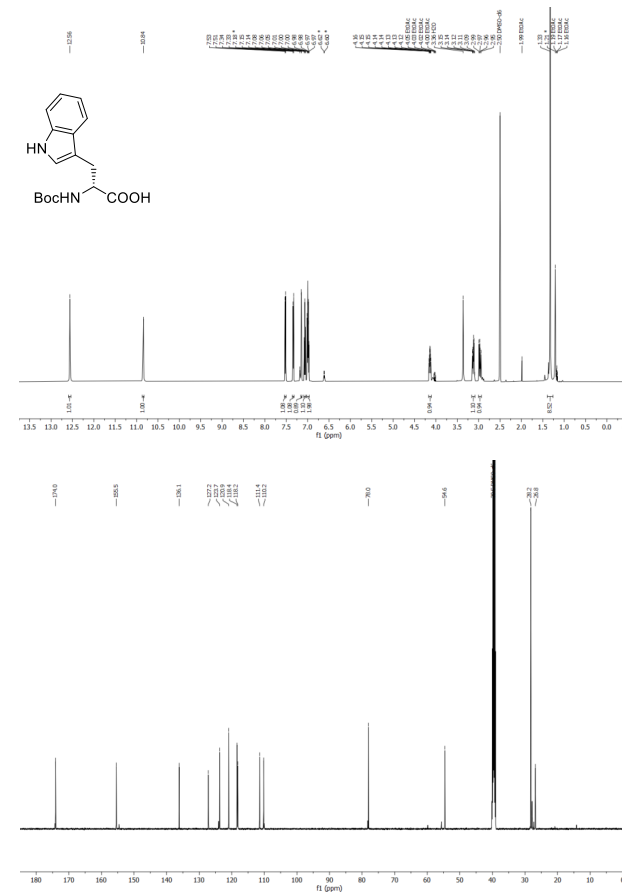
The Boc- and Onpe-protected amino acids **10a_L**, **10b_L**, **10c_L**, **10c_D**, **11e_L** and their corresponding Boc-deprotected derivatives **12a_L·HCl**, **12b_L·HCl**, **12c_L·HCl** and **12c_D·HCl** were synthesized following a procedure previously reported in the literature.^{2,3}

General procedure for the synthesis of compound 9: Compound **8** (1 equiv.) was dissolved in 1 M NaOH in water. Boc₂O (1.2 equiv.) was dissolved in THF and, subsequently, it was added dropwise to the aqueous solution containing **8**. The reaction was stirred at r.t. overnight. After that, the organic solvent was removed under reduced pressure and the pH of the residual aqueous solution was adjusted to 4 with 1 M HCl in water. The crude was extracted with EtOAc and the organic layer was washed with brine, dried (Na₂SO₄), filtered and concentrated to afford the product as a white solid.

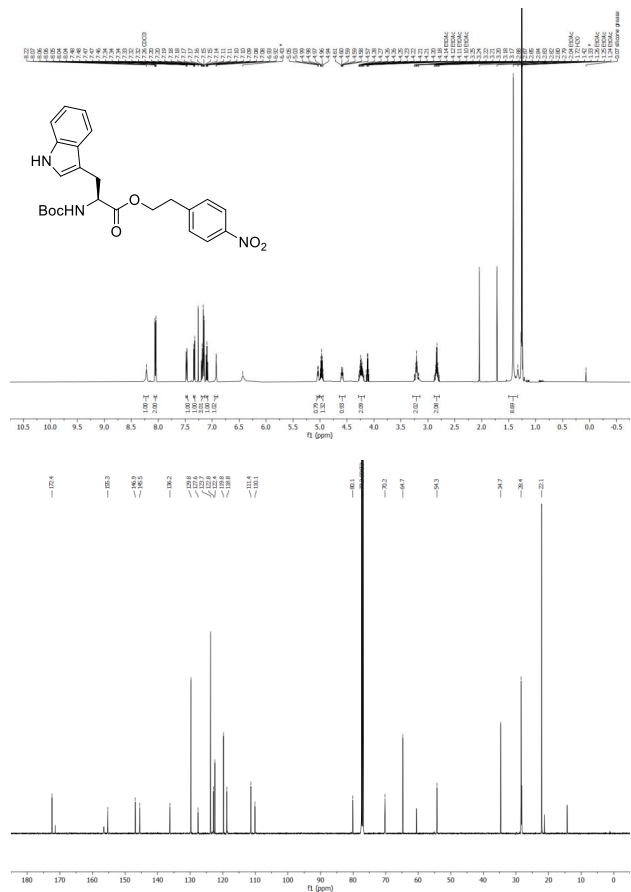
9d_l: Yield = 95%. ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 12.56 (s, 1H); 10.84 (s, 1H); 7.52 (d, *J* = 7.9 Hz, 1H); 7.33 (d, *J* = 8.0 Hz, 1H); 7.15 (d, *J* = 2.3 Hz, 1H); 7.06 (t, *J* = 6.9 Hz, 1H); 7.03 – 6.95 (m, 2H); 4.18 – 4.10 (m, 1H); 3.17 – 3.09 (m, 1H); 3.01 – 2.93 (m, 1H); 1.33 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 174.1; 155.5; 136.1; 127.2; 123.7; 120.9; 118.4; 118.2; 111.4; 110.2; 78.1; 54.6; 28.2; 26.8. FTIR *v*_{max} (cm⁻¹): 3371 (w); 1717 (s); 1643 (m), 1437 (m); 1397 (s); 1370 (m); 1144 (m); 740 (s). HRMS (ESI) *m/z*: [M-H]⁻ Calcd for C₁₆H₁₉N₂O₄ 303.1350; Found 303.1350. The analytical data is in agreement with that reported in the literature.⁴



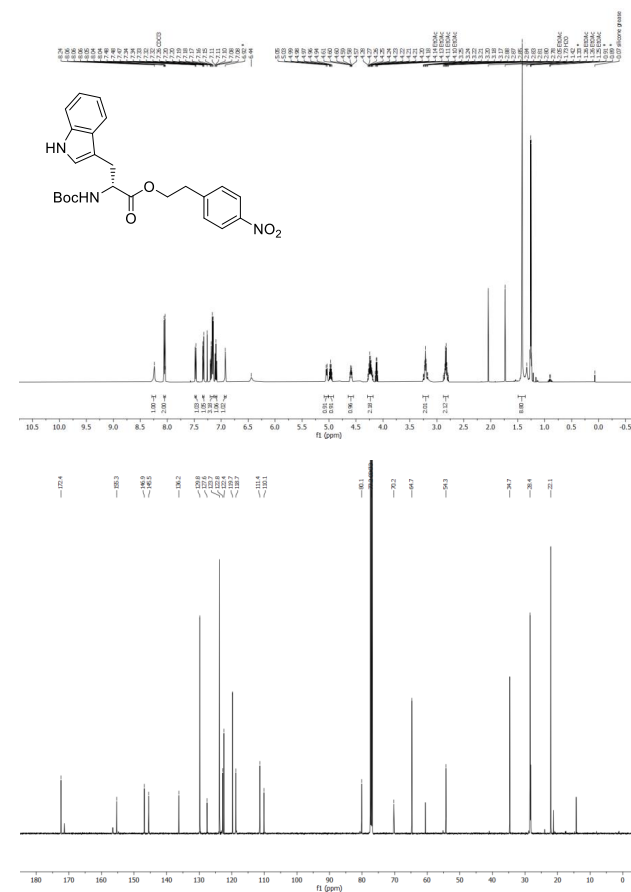
9d_b: Yield = 97%. ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 12.56 (s, 1H); 10.84 (s, 1H); 7.52 (d, *J* = 7.9 Hz, 1H); 7.33 (d, *J* = 8.1 Hz, 1H); 7.15 (d, *J* = 2.3 Hz, 1H); 7.06 (t, *J* = 7.5 Hz, 1H); 7.03 – 6.95 (m, 2H); 4.14 (ddd, *J* = 9.4 Hz, *J* = 8.0 Hz, *J* = 4.6 Hz, 1H); 3.13 (dd, *J* = 14.6 Hz, *J* = 4.7 Hz, 1H); 2.97 (dd, *J* = 14.6 Hz, *J* = 9.5 Hz, 1H); 1.33 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 174.0; 155.5; 136.1; 127.2; 123.7; 120.9; 118.4; 118.2; 111.4; 110.2; 78.0; 54.6; 28.2; 26.8. FTIR *v*_{max} (cm⁻¹): 3373 (w); 1717 (s); 1648 (m), 1437 (m); 1397 (s); 1370 (m); 1144 (m); 742 (s). HRMS (ESI) *m/z*: [M-H]⁻ Calcd for C₁₆H₁₉N₂O₄ 303.1350; Found 303.1350.



10d_L: Yield = 99%. R_f = 0.25 (7:3 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.22 (s, 1H); 8.08 – 8.02 (m, 2H); 7.47 (dd, *J* = 7.9 Hz, *J* = 1.0 Hz, 1H); 7.33 (dt, *J* = 8.1 Hz, *J* = 0.9 Hz, 1H); 7.21 – 7.13 (m, 3H); 7.10 (ddd, *J* = 8.0 Hz, *J* = 7.0 Hz, *J* = 1.0 Hz, 1H); 6.92 (d, *J* = 2.4 Hz, 1H); 5.04 (d, *J* = 8.2 Hz, 1H); 4.97 (p, *J* = 6.3 Hz, 1H); 4.59 (dt, *J* = 8.3 Hz, *J* = 5.9 Hz, 1H); 4.30 – 4.17 (m, 2H); 3.27 – 3.16 (m, 2H); 2.83 (hept, *J* = 7.1 Hz, *J* = 6.6 Hz, 2H); 1.42 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.4, 155.3, 146.9, 145.5, 136.2, 129.8, 127.6, 123.7, 122.8, 122.4, 119.8, 118.8, 111.4, 110.1, 80.1, 70.2, 64.7, 54.3, 34.7, 28.4, 22.1. FTIR ν_{max} (cm⁻¹): 3281 (w); 2980 (w); 1735 (s); 1696 (s); 1516 (s); 1345 (s); 1232 (s); 1180 (s); 1107 (s); 1051 (s); 749 (m). HRMS (ESI) *m/z* [M-H]⁻ Calcd for C₂₄H₂₆N₃O₆ 452.1827; Found 452.1825.

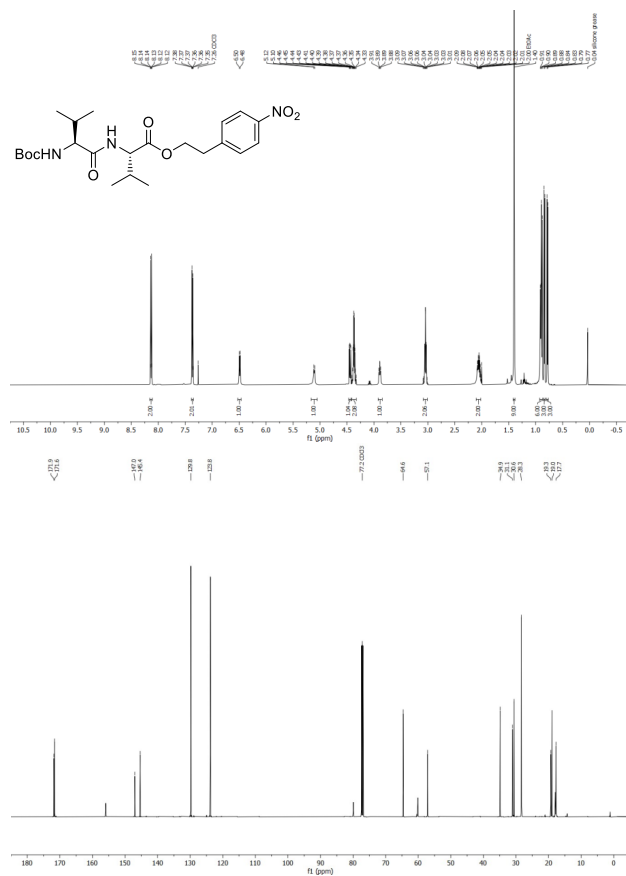


10d_B: Yield = 90%. R_f = 0.43 (6:4 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.24 (s, 1H); 8.08 – 8.02 (m, 2H); 7.51 – 7.45 (m, 1H); 7.36 – 7.31 (m, 1H); 7.22 – 7.13 (m, 3H); 7.10 (ddd, *J* = 8.1 Hz, *J* = 7.0 Hz, *J* = 1.0 Hz, 1H); 6.92 (d, *J* = 2.4 Hz, 1H); 5.04 (d, *J* = 8.2 Hz, 1H); 4.97 (p, *J* = 6.2 Hz, 1H); 4.59 (dt, *J* = 8.2 Hz, *J* = 5.9 Hz, 1H); 4.24 (tq, *J* = 11.1 Hz, *J* = 5.5 Hz, *J* = 4.4 Hz, 2H); 3.27 – 3.16 (m, 2H); 2.83 (hept, *J* = 7.1 Hz, *J* = 6.6 Hz, 2H); 1.42 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.4; 155.3; 146.9; 145.5; 136.2; 129.8; 127.6; 123.7; 122.8; 122.4; 119.7; 118.7; 111.4; 110.1; 80.1; 70.2; 64.7; 54.3; 34.7; 28.4; 22.1. FTIR ν_{max} (cm⁻¹): 3284 (w); 2980 (w); 1690 (s); 1519 (s); 1345 (s); 1246 (s); 1177 (s); 1107 (s); 1049 (m); 744 (m). HRMS (ESI) *m/z* [M-H]⁻ Calcd for C₂₄H₂₆N₃O₆ 452.1827; Found 452.1825.

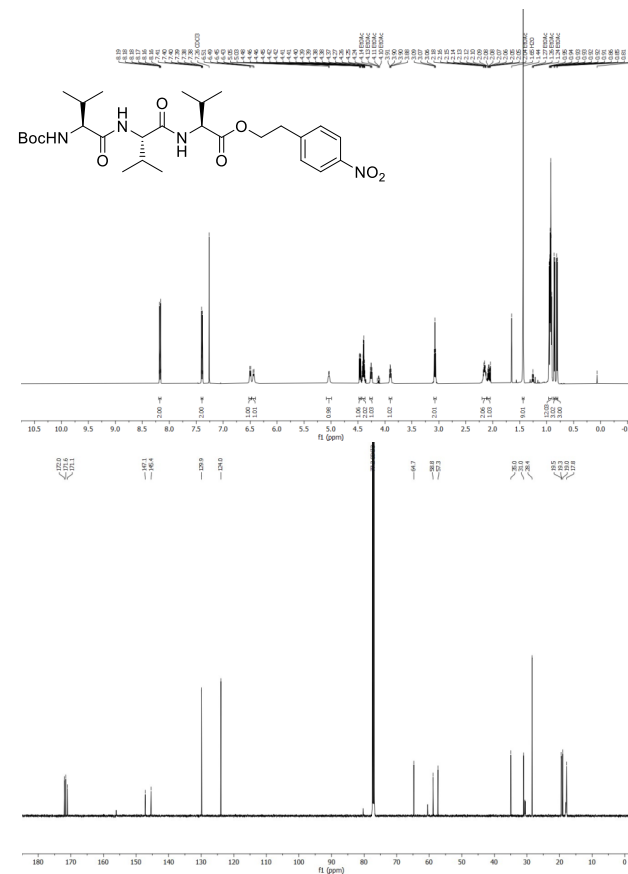


General procedure for the synthesis of peptides: Compound **12a_L** or **12f_L** were dissolved in dry CH₂Cl₂ and compound **9a_L** was added. The mixture was stirred at 0°C in an ice bath. 2-(1*H*-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium-hexafluorophosphat (HBTU) and diisopropylethylamine (DIPEA) were added and the reaction was stirred at r.t. for another 2 h. The reaction was quenched with aqueous saturated NH₄Cl and extracted with CH₂Cl₂. The combined organic layers were dried (MgSO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white solid.

10f_L: Synthesized by peptide coupling reaction of **12a_L** with commercially available **9a_L**. Yield = 91%. Rf = 0.38 (2:3 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.16 – 8.11 (m, 2H); 7.39 – 7.34 (m, 2H); 6.49 (d, *J* = 8.6 Hz, 1H); 5.11 (d, *J* = 8.9 Hz, 1H); 4.44 (dd, *J* = 8.6 Hz, *J* = 5.1 Hz, 1H); 4.41 – 4.32 (m, 2H); 3.92 – 3.85 (m, 1H); 3.04 (td, *J* = 6.7 Hz, *J* = 1.4 Hz, 2H); 2.11 – 2.01 (m, 2H); 1.40 (s, 9H); 0.90 (dd, *J* = 11.0 Hz, *J* = 6.8 Hz, 6H); 0.84 (d, *J* = 6.9 Hz, 3H); 0.78 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 171.6; 147.0; 145.4; 129.8; 123.8; 64.6; 57.1; 34.9; 31.1; 30.6; 28.3; 19.3; 19.0; 17.7. FTIR *v*_{max} (cm⁻¹): 2968 (w); 1678 (m); 1522 (s); 1348 (s); 1160 (m); 913 (m). HRMS (ESI) *m/z* [M+Na]⁺ Calcd for C₂₃H₃₅N₃O₇Na 488.2367; Found 488.2371.

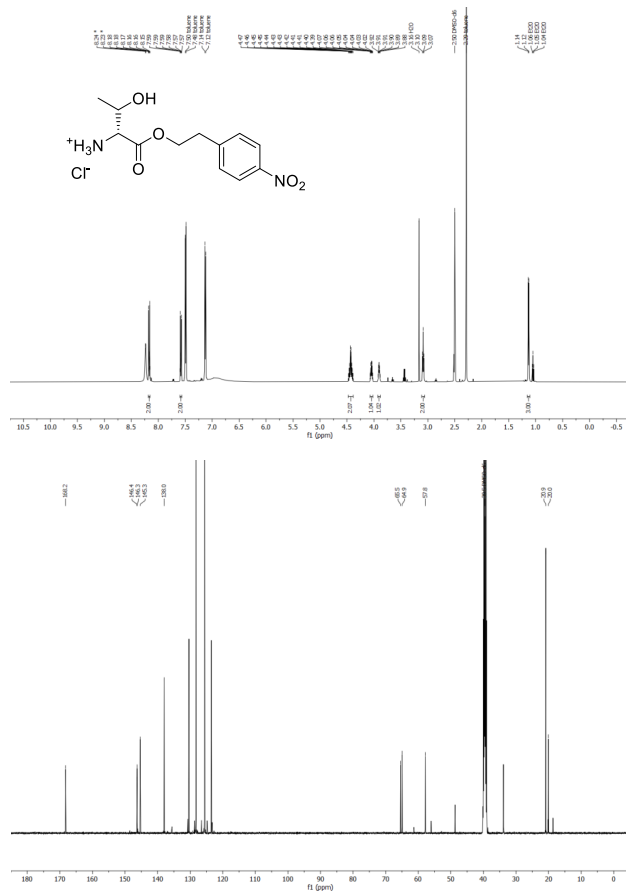


10g_L: Synthesized by peptide coupling reaction of **12f_L** with commercially available **9a_L**. Yield = 78%. Rf = 0.56 (7:3 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.19 – 8.15 (m, 2H); 7.41 – 7.38 (m, 2H); 6.50 (d, *J* = 8.5 Hz, 1H); 6.44 (d, *J* = 8.7 Hz, 1H); 5.04 (d, *J* = 8.1 Hz, 1H); 4.46 (dd, *J* = 8.6 Hz, *J* = 5.1 Hz, 1H); 4.40 (dd, *J* = 11.0 Hz, *J* = 6.8 Hz, *J* = 4.3 Hz, 2H); 4.25 (dd, *J* = 8.5 Hz, *J* = 6.7 Hz, 1H); 3.90 (dd, *J* = 8.3 Hz, *J* = 6.3 Hz, 1H); 3.07 (t, *J* = 6.7 Hz, 2H); 2.15 (p, *J* = 6.1 Hz, *J* = 5.5 Hz, 2H); 2.10 – 2.05 (m, 1H); 1.44 (s, 9H); 0.96 – 0.89 (m, 12H); 0.86 (d, *J* = 6.8 Hz, 3H); 0.81 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 172.0; 171.6; 171.1; 147.1; 145.4; 129.9; 124.0; 64.7; 58.8; 57.3; 35.0; 31.0; 28.4; 19.5; 19.3; 19.0; 17.8. FTIR *v*_{max} (cm⁻¹): 2161 (w); 1522 (s); 1348 (s); 1158 (m); 908 (s). HRMS (ESI) *m/z* [M+Na]⁺ Calcd for C₂₈H₄₄N₄O₈Na 587.3051; Found 587.3042.



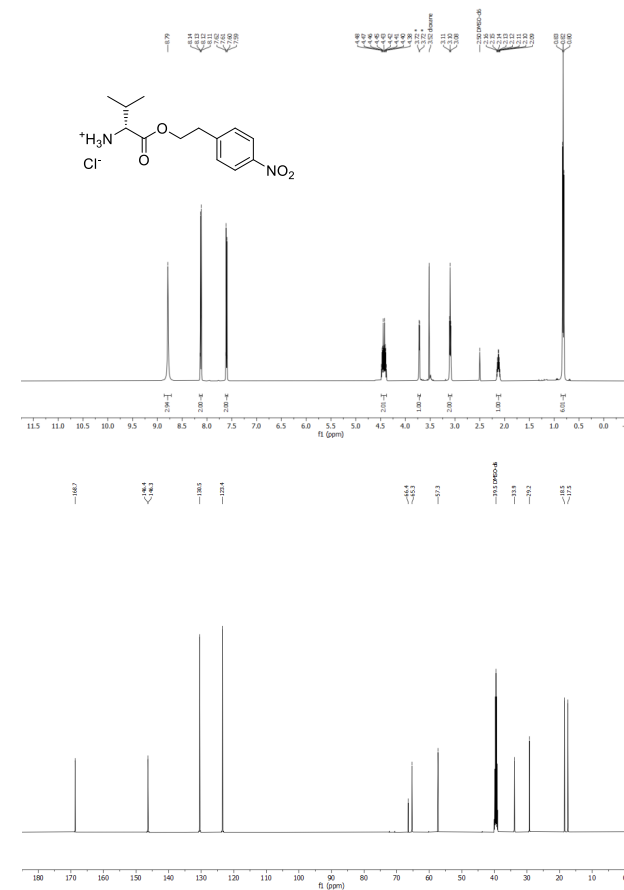
General procedure for the synthesis of compound 11: Compound **8** (1 equiv.), 2-(4-nitrophenyl)-ethanol (3 equiv.) and TsOH (2 equiv.) were dissolved in dry toluene. The reaction mixture was refluxed in a Dean-Stark-apparatus for 20 h. After that, the reaction mixture was cooled down to r.t. and Et₂O was added. The crude was concentrated under reduced pressure and purified by silica gel column chromatography affording the product as a brown oil.

11e_b: Yield = 99%. R_f = 0.30 (93:7 CH₂Cl₂/CH₃OH). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 8.19 – 8.15 (m, 2H); 7.60 – 7.56 (m, 2H); 4.48 – 4.38 (m, 2H); 4.05 (qd, *J* = 6.5 Hz, *J* = 3.8 Hz, 1H), 3.90 (dt, *J* = 9.3 Hz, *J* = 4.6 Hz, 1H); 3.09 (t, *J* = 6.4 Hz, 2H); 1.13 (d, *J* = 6.5 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 168.2; 146.4; 146.3; 145.3; 138.0; 65.5; 64.9; 57.8; 20.9; 20.0. FTIR *v*_{max} (cm⁻¹): 3033 (w); 2947 (w); 1739 (m); 1510 (m); 1352 (m); 1119 (s); 1032 (s); 1006 (s); 816 (m); 681 (s); 559 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₁₂H₁₇N₂O₅ 269.1132; Found 269.1132.

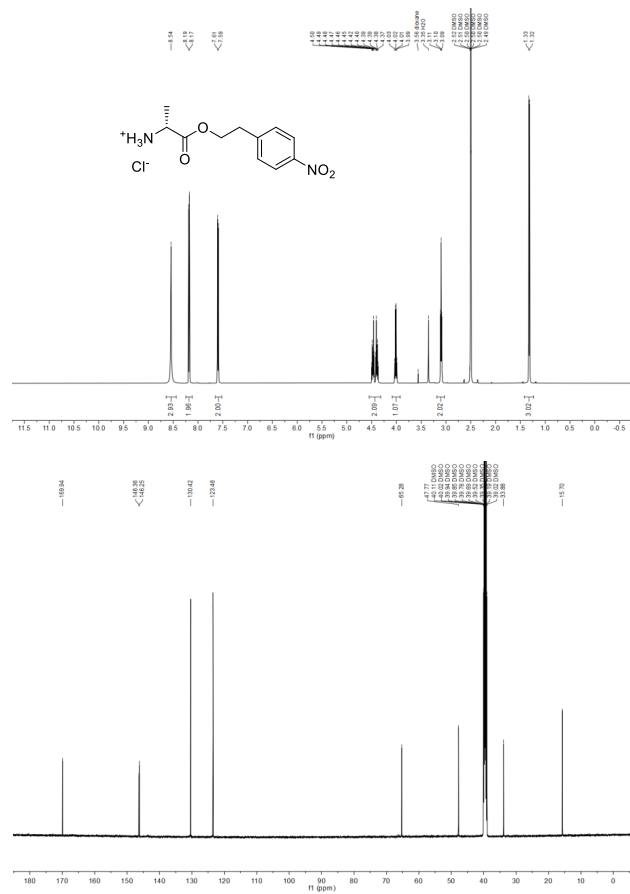


General procedure for the synthesis of compound 12a-d,f,g: Compound **10** (1 equiv.) was dissolved in 4 M HCl/1,4-dioxane at 0°C. The reaction was stirred at 0°C for 5 min and at r.t. for 1 h. After that, the crude was concentrated under reduced pressure. The crude was triturated with Et₂O and filtered. The white precipitate was washed with additional Et₂O. Finally, the product was dried under high vacuum.

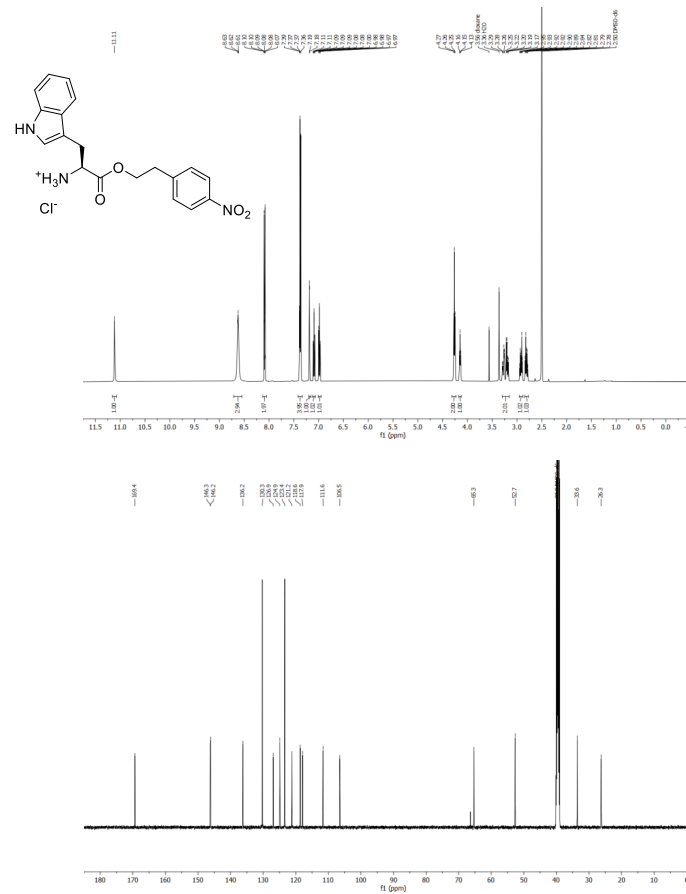
12a_b: Yield = 96%. ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 8.79 (s, 3H); 8.17 – 8.09 (m, 2H); 7.64 – 7.56 (m, 2H); 4.50 – 4.35 (m, 2H); 3.72 (d, *J* = 4.6 Hz, 1H); 3.10 (t, *J* = 6.4 Hz, 2H); 2.13 (pd, *J* = 7.0, *J* = 4.6 Hz, 1H); 0.82 (t, *J* = 7.4 Hz, 6H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 168.7; 146.4; 146.3; 130.5; 123.4; 66.4; 65.3; 57.3; 33.9; 29.2; 18.5; 17.5. FTIR *v*_{max} (cm⁻¹): 2873 (w); 1735 (s); 1590 (w); 1508 (s); 1466 (m); 1343 (s); 1232 (m); 1207 (m); 857 (m); 736 (m). HRMS (ESI) *m/z*: [M-Cl]⁺ Calcd for C₁₃H₁₉N₂O₄ 267.1339; Found 267.1340.



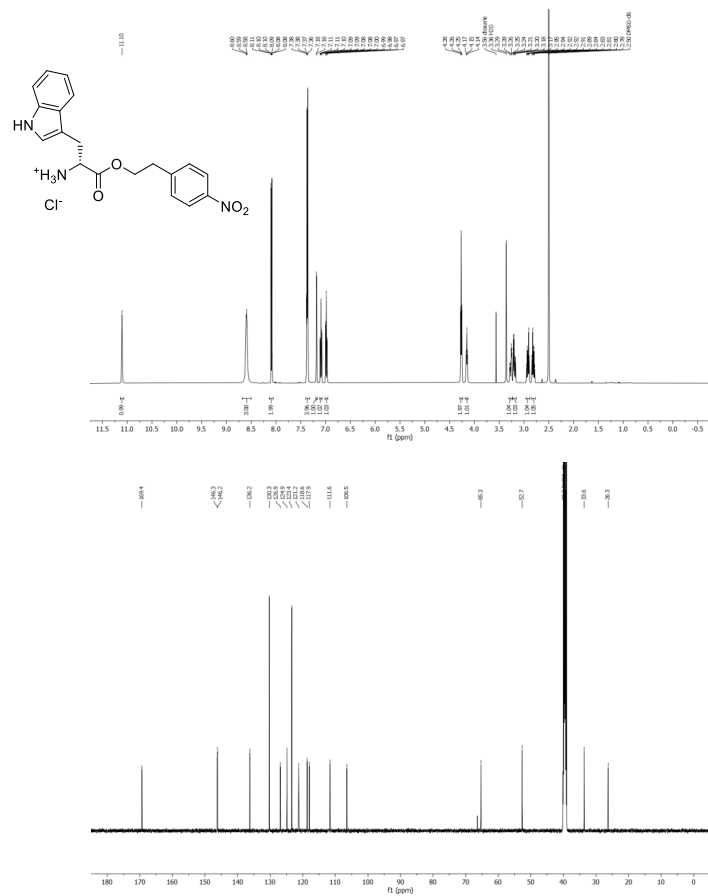
12b_b: Yield = 88%. ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 8.54 (s, 3H); 8.19 – 8.17 (m, 2H); 7.61 – 7.59 (m, 2H); 4.50 – 4.37 (m, 2H); 4.01 (q, *J* = 7.3 Hz, 1H); 3.10 (t, *J* = 6.3 Hz, 2H); 1.32 (d, *J* = 7.3 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 169.9; 146.4; 146.2; 130.4; 123.5; 62.3; 47.8; 33.9; 15.7. FTIR *v*_{max} (cm⁻¹): 2880 (m); 1737 (s); 1599 (m); 1510 (s); 1483 (m); 1345 (s); 1236 (m); 1197 (s); 1111 (s); 821 (m). HRMS (ESI) *m/z*: [M-Cl]⁺ Calcd for C₁₁H₁₅N₂O₄ 239.1026; Found 239.1025.



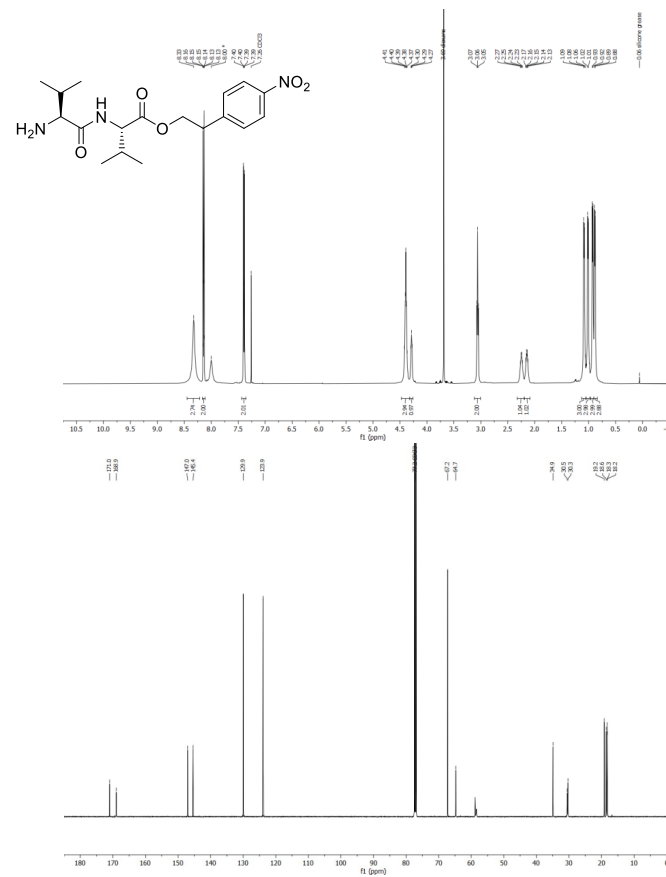
12d_L: Yield = 67%. ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 11.11 (s, 1H); 8.65 – 8.60 (m, 3H); 8.12 – 8.06 (m, 2H); 7.37 (dd, *J* = 8.0 Hz, *J* = 6.1 Hz, 4H); 7.18 (d, *J* = 2.4 Hz, 1H); 7.12 – 7.08 (m, 1H); 7.02 – 6.95 (m, 1H); 4.26 (t, *J* = 6.3 Hz, 2H); 4.15 (t, *J* = 6.5 Hz, 1H); 3.31 – 3.16 (m, 2H); 2.92 (dt, *J* = 14.2 Hz, *J* = 6.1 Hz, 1H); 2.81 (dt, *J* = 13.8 Hz, *J* = 6.4 Hz, 1H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 169.4; 146.3; 146.2; 136.2; 130.3; 126.9; 124.9; 123.4; 121.2; 118.6; 117.9; 111.6; 106.5; 65.3; 52.7; 33.6; 26.3. FTIR *v*_{max} (cm⁻¹): 3282 (w); 1745 (s); 1516 (s); 1497 (m); 1349 (s); 1230 (m); 1107 (m); 855 (m); 743 (s). HRMS (ESI) *m/z*: [M-Cl]⁺ Calcd for C₁₉H₂₀N₂O₄ 354.1448; Found 354.1449.



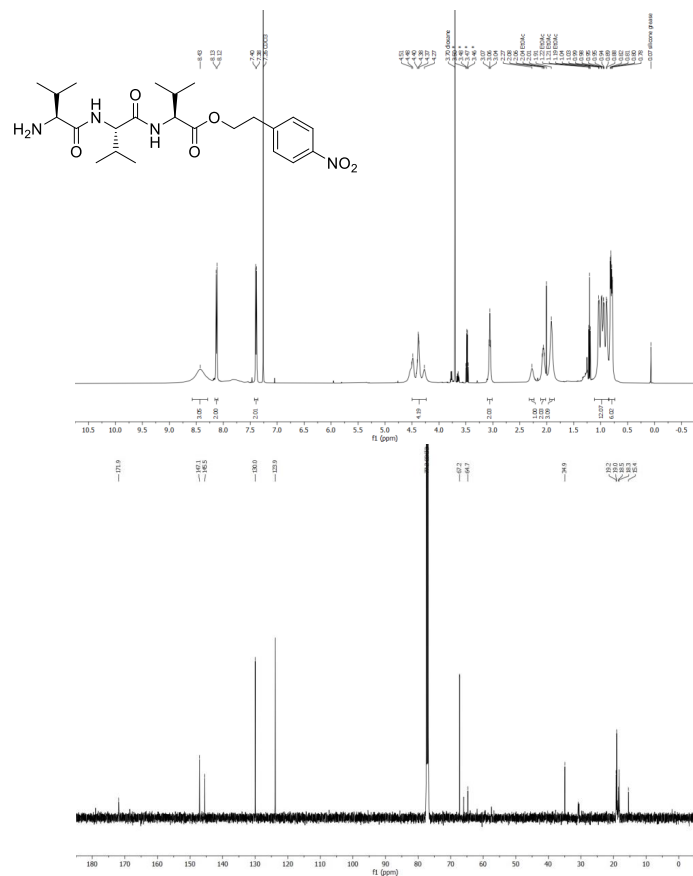
12db: Yield = 55%. ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 11.10 (s, 1H); 8.62 – 8.57 (m, 3H); 8.12 – 8.06 (m, 2H); 7.37 (dd, *J* = 8.1 Hz, *J* = 4.5 Hz, 4H); 7.18 (d, *J* = 2.4 Hz, 1H); 7.13 – 7.06 (m, 1H); 7.02 – 6.95 (m, 1H); 4.26 (t, *J* = 6.3 Hz, 2H); 4.15 (t, *J* = 6.5 Hz, 1H); 3.30 – 3.23 (m, 1H); 3.19 (dd, *J* = 14.8 Hz, *J* = 7.3 Hz, 1H); 2.92 (dt, *J* = 14.2 Hz, *J* = 6.1 Hz, 1H); 2.81 (dt, *J* = 13.8 Hz, *J* = 6.5 Hz, 1H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 169.4; 146.3; 146.2; 136.2; 130.3; 126.9; 124.9; 123.4; 121.2; 118.6; 117.9; 111.6; 106.5; 65.3; 52.7; 33.6; 26.3. FTIR *v*_{max} (cm⁻¹): 3281 (w); 1745 (s); 1516 (s); 1496 (m); 1349 (s); 1230 (m); 1105 (m); 855 (m); 743 (s). HRMS (ESI) *m/z*: [M-Cl]⁺ Calcd for C₁₉H₂₀N₃O₄ 354.1448; Found 354.1453.



12f: Yield = 99%. ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.33 (s, 3H); 8.17 – 8.11 (m, 2H); 7.42 – 7.36 (m, 2H); 4.38 (q, *J* = 5.5 Hz, 3H); 4.29 (t, *J* = 5.9 Hz, 1H); 3.06 (t, *J* = 6.8 Hz, 2H); 2.25 (q, *J* = 6.6 Hz, 1H); 2.15 (q, *J* = 6.4 Hz, 1H); 1.09 (d, *J* = 6.4 Hz, 3H); 1.01 (d, *J* = 6.5 Hz, 3H); 0.93 (d, *J* = 6.5 Hz, 3H); 0.89 (d, *J* = 6.6 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.0; 168.9; 147.0; 145.4; 129.9; 123.9; 67.2; 64.7; 34.9; 30.5; 30.3; 19.2; 18.6; 18.3; 18.2. FTIR *v*_{max} (cm⁻¹): 2966 (w); 1739 (m); 1673 (m); 1516 (s); 1344 (s); 1188 (m); 1151 (m); 856 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₁₈H₂₀N₃O₅ 366.2023; Found 366.2026.



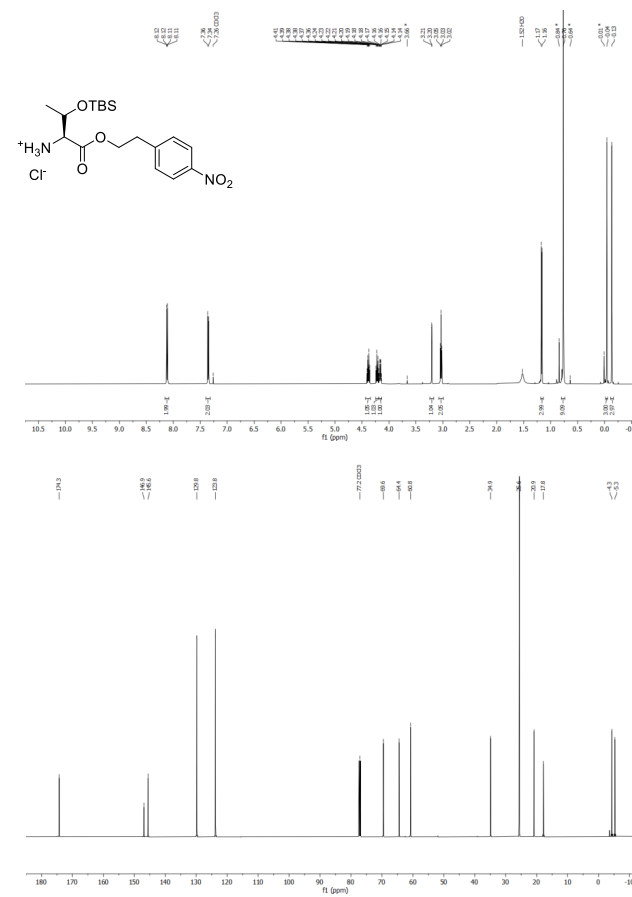
12g_L: Yield = 97%. ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.43 (s, 3H); 8.12 (d, *J* = 8.4 Hz, 2H); 7.39 (d, *J* = 8.4 Hz, 2H); 4.52 – 4.25 (m, 4H); 3.09 – 3.02 (m, 2H); 2.27 (s, 1H); 2.07 (d, *J* = 10.7 Hz, 2H); 1.91 (s, 3H); 1.08 – 0.84 (m, 12H); 0.80 (dd, *J* = 13.6 Hz, *J* = 6.5 Hz, 6H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 147.1; 145.5; 130.0; 123.9; 67.2; 64.7; 34.9; 19.2; 19.0; 18.5; 18.3; 15.4. FTIR ν_{max} (cm⁻¹): 2966(w); 1739 (m); 1651 (s); 1516 (s); 1344 (s); 1188 (m); 1154 (m); 856 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₃H₃₇N₄O₆ 465.2707; Found 465.2705.



S72

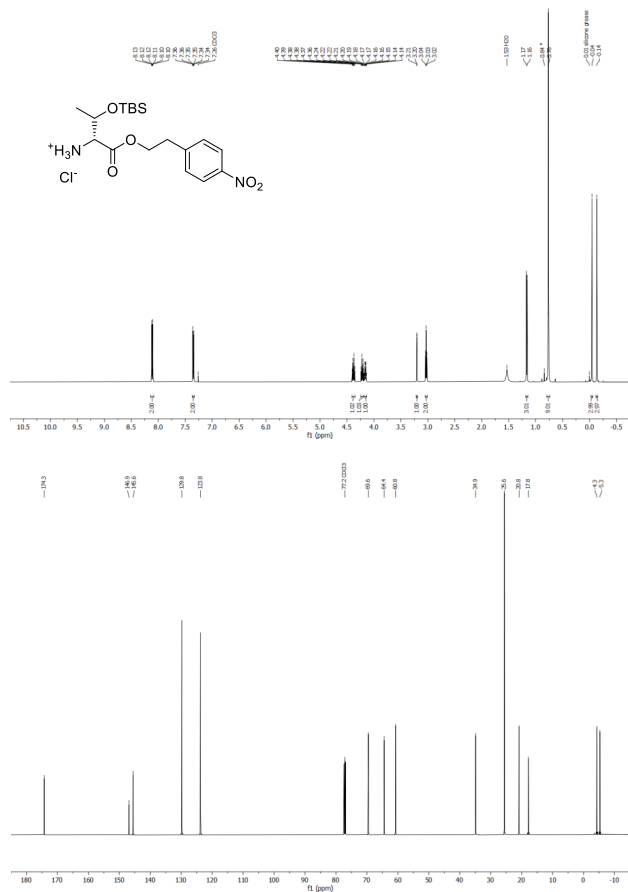
General procedure for the synthesis of compound 12e: Compound **11e** was dissolved in dry pyridine, then *tert*-butyldimethylsilyl chloride (TBDMS-Cl) (1.5 equiv.) and imidazole (3 equiv.) was added. After 10 min, another portion of TBDMS-Cl (1.5 equiv.) was added and the reaction was stirred at r.t. for 23 h. The reaction mixture was diluted with CH₂Cl₂ and washed with aq. sat. NaHCO₃ and water. The organic layer was dried (MgSO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a brown oil.

12e_L: Yield = 60% (over two steps). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.11 (d, *J* = 8.7 Hz, 2H); 7.35 (s, *J* = 8.7 Hz, 2H); 4.38 (dt, *J* = 11.0 Hz, *J* = 6.8 Hz, 1H); 4.22 (dt, *J* = 11.0 Hz, *J* = 6.8 Hz, 1H); 4.16 (qd, *J* = 6.3 Hz, *J* = 2.8 Hz, 1H); 3.21 (d, *J* = 2.8 Hz, 1H); 3.03 (t, *J* = 6.8 Hz, 2H); 1.17 (d, *J* = 6.3 Hz, 3H); 0.76 (s, 9H); -0.04 (s, 3H); -0.13 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 174.3; 146.9; 145.6; 129.8; 123.8; 69.6; 64.4; 60.8; 34.9; 25.6; 20.9; 17.8; -4.3; -5.3. FTIR ν_{max} (cm⁻¹): 2930 (w); 2857 (w); 1739 (m); 1601 (w); 1519 (s); 1473 (w); 1345 (vs); 1252 (m); 1154 (m); 1075 (m); 969 (m); 835 (s); 775 (vs). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₁₈H₃₁N₂O₅Si 383.1996; Found 383.1996.



S73

12ep: Yield = 77%. ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 8.14 – 8.08 (m, 2H); 7.39 – 7.31 (m, 2H); 4.38 (dt, *J* = 11.0 Hz, *J* = 6.7 Hz, 1H); 4.21 (dt, *J* = 11.1 Hz, *J* = 6.9 Hz, 1H); 4.16 (qd, *J* = 6.3 Hz, *J* = 2.8 Hz, 1H); 3.20 (d, *J* = 2.9 Hz, 1H); 3.03 (t, *J* = 6.8 Hz, 2H); 1.16 (d, *J* = 6.3 Hz, 3H); 0.76 (s, 9H); -0.04 (s, 3H); -0.14 (s, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 174.3; 146.9; 145.6; 129.8; 123.8; 69.6; 64.4; 60.8; 34.9; 25.6; 20.8; 17.8; -4.3; -5.3. FTIR ν_{max} (cm⁻¹): 2930 (w); 2857 (w); 1739 (m); 1601 (w); 1519 (s); 1345 (vs); 1252 (m); 1154 (m); 1075 (m); 969 (m); 835 (s); 775 (vs). HRMS (ESI) *m/z* [M+H]⁺ Calcd for C₁₈H₃₁N₂O₅Si 383.1996; Found 383.1996.



2.3 5-Methylaminomethyl uridine phosphoramidite

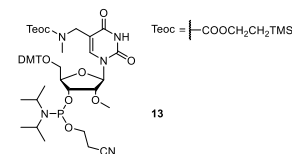


Figure S1. 5-methylaminomethyl uridine phosphoramidite **13**.

Compound **13** was synthesized following a procedure previously reported in the literature.²

3. General information and instruments for oligonucleotides

Synthesis and purification of oligonucleotides

Phosphoramidites of 2'-O-Me ribonucleosides (2'-OMe-Bz-A-CE, 2'-OMe-Dmf-G-CE, 2'-OMe-Ac-C-CE and 2'-OMe-U-CE) were purchased from LinkTech and Sigma-Aldrich. Oligonucleotides (ONs) were synthesized on a 1 μmol scale using RNA SynBase™ CPG 1000/110 and High Load Glen UnySupport™ as solid supports using an RNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with a standard phosphoramidite chemistry. ONs were synthesized in DMT-OFF mode using DCA as a deblocking agent in CH₂Cl₂, BTT or Activator 42® as activator in MeCN, Ac₂O as capping reagent in pyridine/THF and I₂ as oxidizer in pyridine/H₂O.

Deprotection of Onpe and teoc groups

For the deprotection of the *para*-nitrophenylethyl (Onpe) group in ONs containing amino acid-modified carbamoyl adenosine nucleosides, the solid support beads were suspended in a 9:1 THF/DBU solution mixture (1 mL) and incubated at r.t. for 2 h.⁵ After that, the supernatant was removed, and the beads were washed with THF (3x1 mL).

For the deprotection of the 2-(trimethylsilyl)ethoxycarbonyl (teoc) group in ONs containing 5-methylaminomethyl uridine nucleosides, the solid support beads were suspended in a saturated solution of ZnBr₂ in 1:1 MeNO₂/IPA (1 mL) and incubated at r.t. overnight.⁶ After that, the supernatant was removed, and the beads were washed with 0.1 M EDTA in water (1 mL) and water (1 mL).

Coupling of amino acids to ONs anchored to the solid support beads

The solid support beads (1 μmol) in an Eppendorf tube were washed with dry DMF (0.3 mL). In a separate Eppendorf tube, Boc-protected L- or D-Valine, DMTMM•BF₄ (100 μmol) as activator and dry DIPEA (200 μmol) were dissolved in dry DMF (0.6 mL). Subsequently, the amino acid solution was added to the solid support beads and the reaction was incubated in an orbital shaker at r.t. for 1 h. The suspension was centrifuged and the supernatant was removed. The solid support beads were washed with dry DMF (2x0.3 mL) and dry MeCN (2x0.3 mL). Finally, the beads were dried using a SpeedVac concentrator.

For the deprotection of the *tert*-butyloxycarbonyl (Boc) group in ONs after the coupling of a Boc-protected amino acid or peptide, the solid support beads were suspended in a 1:1 TFA/CH₂Cl₂ solution mixture (0.5 mL) and incubated for 5 min at r.t.⁷ After that, the supernatant was removed and the solid support beads were washed with CH₂Cl₂ (2x0.5 mL).

Cleavage from beads and precipitation of the synthesized ON

The solid support beads were suspended in a 1:1 aqueous solution mixture (0.6 mL) of 30% NH₄OH and 40% MeNH₂. The suspension was heated at 65°C (8 min for SynBase™ CPG 1000/110 and 60 min for High Load Glen UnySupport™). Subsequently, the supernatant was collected, and the beads were washed with water (2x0.3 mL). The combined aqueous solutions were concentrated under reduced pressure using a SpeedVac concentrator. After that, the crude was dissolved in DMSO (100 μL) and the ON was precipitated by adding 3 M NaOAc in water (25 μL) and *n*-butanol (1 mL). The mixture was kept at -80°C for 2 h and centrifuged at 4°C for 1 h. The supernatant was removed, and the white precipitate was lyophilized. For **ON1e₁₀**, the TBS-protecting group of the amino acid was cleaved by incubation of the ON in DMSO (100 μL) and HF•NEt₃ (100 μL) at 65°C for 1.5 h, followed by precipitation with 3 M NaOAc as described above.

Purification of the synthesized ON by HPLC and desalting

The crude was purified by semi-preparative HPLC (1260 Infinity II Manual Preparative LC System from Agilent equipped with a G7114A detector) using a reverse-phase (RP) VP 250/10 Nucleodur 100-5 C18ec column from Macherey-Nagel (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-25% of B in 45 min; Flow rate = 5 mL·min⁻¹). The purified ON was analyzed by RP-HPLC (1260 Infinity II LC System from Agilent equipped with a G7165A detector) using an EC 250/4 Nucleodur 100-3 C18ec from Macherey-Nagel (Gradient: 0-30% of B in 45 min; Flow rate = 1 mL·min⁻¹). Finally, the purified ON was desalted using a C18 RP-cartridge from Waters.

Determination of the concentration and the mass of the synthesized ON

The absorbance of the synthesized ON in H₂O solution was measured using an IMPLEN NanoPhotometer® N60/N50 at 260 nm. The extinction coefficient of the ON was calculated using the OligoAnalyzer Version 3.0 from Integrated DNA Technologies. For ONs incorporating non-canonical bases, the extinction coefficients were assumed to be identical to those containing only canonical counterparts.

The synthesized ON (2-3 μL) was desalted on a 0.025 μm VSWP filter (Millipore), co-crystallized in a 3-hydroxypicolinic acid matrix (HPA, 1 μL) and analyzed by matrix-assisted laser desorption/ionization – time-of-flight (MALDI-TOF) mass spectrometry (negative mode).

4. Synthesized oligonucleotides using a DNA/RNA automated synthesizer

4.1 Donor strands containing an amino acid-modified methyl N⁶-carbamoyl adenosine at the 5'-end

Table S1. HPLC retention times (0-40% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of donor strands **ON1** containing an amino acid-modified N⁶-carbamoyl adenosine at the 5'-end.

Sequence	Donor strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
	ON1a_L ; X = m ⁶ V _L ⁶ A _m	27.8	2417.5	2417.2
	ON1a_D ; X = m ⁶ V _D ⁶ A _m	28.9	2417.5	2417.0
	ON1b_L ; X = m ⁶ a _L ⁶ A _m	24.7	2389.5	2388.0
	ON1b_D ; X = m ⁶ a _D ⁶ A _m	25.7	2389.5	2389.5
	ON1c_L ; X = m ⁶ f _L ⁶ A _m	29.7	2465.5	2465.4
	ON1c_D ; X = m ⁶ f _D ⁶ A _m	31.4	2465.5	2465.4
5'-X(AUCGCU) _m -3'	ON1d_L ; X = m ⁶ W _L ⁶ A _m	29.2	2504.5	2504.7
	ON1d_D ; X = m ⁶ W _D ⁶ A _m	30.3	2504.5	2504.1
	ON1e_L ; X = m ⁶ t _L ⁶ A _m	24.2	2419.4	2420.0
	ON1e_D ; X = m ⁶ t _D ⁶ A _m	25.1	2419.4	2420.0
	ON1f_L ; X = m ⁶ V _L V _L ⁶ A _m	30.1	2516.5	2516.1
	ON1g_L ; X = m ⁶ V _L V _L V _L ⁶ A _m	31.7	2615.5	2614.3
	ON1h ; X = m ⁶ g ⁶ A _m	24.0	2375.4	2374.8

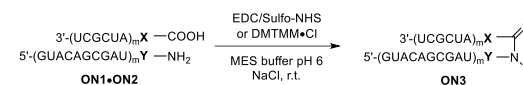
4.2 Acceptor strands containing a 5-methylaminomethyl uridine at the 3'-end

Table S2. HPLC retention times (0-40% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of acceptor strands **ON2** containing a methylaminomethyl uridine at the 3'-end.

Sequence	Acceptor strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
	ON2a_L ; Y = v _L -mnm ⁶ U _m	24.8	3785.4	3786.0
5'-(GUACAGCGAU) _m -3'	ON2a_D ; Y = v _D -mnm ⁶ U _m	24.7	3785.4	3788.7
	ON2b ; Y = mnm ⁶ U _m	24.3	3686.2	3686.3

5. Peptide coupling reactions between donor and acceptor oligonucleotides

Stock solutions of MES buffer pH 6 (400 mM), NaCl (1 M) and activator (500 mM) were prepared in water. Subsequently, equimolar amounts of **ON1** and **ON2** (3-5 nmol) were annealed at 95°C for 4 min in water containing NaCl (half of the volume required for the reaction). Finally, buffer, NaCl, activator solutions and water were added to the ON solution and the reaction was incubated in a ThermoMixer at 25°C for 2-6 h. The concentration of the components in the reaction mixture was: 50 μM of **ON1**, 50 μM of **ON2**, 100 mM of buffer, 100 mM of NaCl and 50 mM of activator.



Scheme S3. Peptide coupling reactions between oligonucleotides **ON1** and **ON2** to give hairpin **ON3**.

The peptide coupling reactions were only performed to isolate and characterize the products by HPLC (t_R) and mass spectrometry (m/z). The isolated products were used as references for the analyses of the pair-wise competitive reactions. In this respect, the reaction yields are not given.

Table S3. HPLC retention times (0-40% of B in 45 min) and MALDI-TOF mass spectrometric analysis (negative mode) of isolated hairpin products **ON3**.

Donor strand	Acceptor strand	Hairpin strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
ON1a_L ; X = m ⁶ v _L ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3a_{LL}	32.2	6186.9	6186.1
ON1a_D ; X = m ⁶ v _D ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3a_{DL}	33.6	6186.9	6187.3
ON1a_L ; X = m ⁶ v _L ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3a_{LD}	32.8	6186.9	6187.6
ON1a_D ; X = m ⁶ v _D ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3a_{DD}	32.1	6186.9	6187.6
ON1b_L ; X = m ⁶ a _L ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3b_{LL}	31.1	6158.8	6159.4
ON1b_D ; X = m ⁶ a _D ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3b_{DL}	31.4	6158.8	6159.4
ON1b_L ; X = m ⁶ a _L ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3b_{LD}	30.4	6158.8	6159.0
ON1b_D ; X = m ⁶ a _D ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3b_{DD}	29.8	6158.8	6158.6
ON1c_L ; X = m ⁶ f _L ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3c_{LL}	34.8	6234.9	6235.3
ON1c_D ; X = m ⁶ f _D ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3c_{DL}	35.7	6234.9	6234.5
ON1c_L ; X = m ⁶ f _L ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3c_{LD}	34.7	6234.9	6235.3
ON1c_D ; X = m ⁶ f _D ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3c_{DD}	34.0	6234.9	6234.9
ON1d_L ; X = m ⁶ w _L ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3d_{LL}	34.2	6273.9	6272.4
ON1d_D ; X = m ⁶ w _D ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3d_{DL}	34.3	6273.9	6272.8
ON1d_L ; X = m ⁶ w _L ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3d_{LD}	33.0	6273.9	6272.4
ON1d_D ; X = m ⁶ w _D ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3d_{DD}	33.5	6273.9	6274.3
ON1e_L ; X = m ⁶ t _L ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3e_{LL}	29.9	6188.8	6188.4
ON1e_D ; X = m ⁶ t _D ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3e_{DL}	29.1	6188.8	6188.8
ON1e_L ; X = m ⁶ t _L ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3e_{LD}	29.3	6188.8	6188.4
ON1e_D ; X = m ⁶ t _D ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3e_{DD}	29.1	6188.8	6188.0
ON1f_L ; X = m ⁶ v _L v _L ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3f_{LL}	34.0	6285.9	6286.7
ON1f_D ; X = m ⁶ v _D v _D ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3f_{LD}	36.1	6285.9	6285.6
ON1g_L ; X = m ⁶ v _L v _L v _L ⁶ A _m	ON2a_L ; Y = v _L -mnm ⁶ U _m	ON3g_{LL}	35.5	6384.9	6385.0
ON1g_D ; X = m ⁶ v _D v _D v _D ⁶ A _m	ON2a_D ; Y = v _D -mnm ⁶ U _m	ON3g_{LD}	36.3	6384.9	6385.0
ON1h ; X = m ⁶ g ⁶ A _m	ON2b ; Y = mnm ⁶ U _m	ON3h	26.8	6045.7	6044.8

6. Calibration curve

Hairpin oligonucleotide **ON3h** was used for the development of a HPLC calibration curve. A stock solution of **ON3h** was prepared in water (100 μM). Separate standard solutions containing 1.2; 1.0; 0.8; 0.6; 0.4; 0.2 and 0.1 nmol of **ON3h** were prepared in a final volume of 20 μL. The standard solutions were injected in an analytical HPLC equipped with a C18 column (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-40% of B in 45 min; Flow rate = 1 mL·min⁻¹). The absorbance was monitored at 260 nm and the areas of the chromatographic peaks were determined by integration of the HPL-chromatograms. The plot of the chromatographic area (a.u.) versus the amount (nmol) of the oligonucleotide followed a linear relationship.

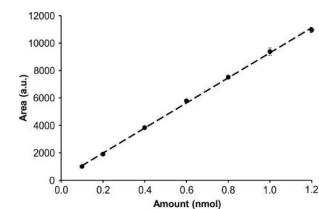


Figure S2. Chromatographic area (a.u.) vs. amount (nmol) of **ON3h**. Line shows the fit of the data to a linear regression equation. Error bars are standard deviations from two independent experiments.

Table S4. Calibration curve ($y = mx + n$) obtained by HPLC analysis of **ON3h** and calculated extinction coefficients of **ON1h**, **ON2b** and **ON3h**.

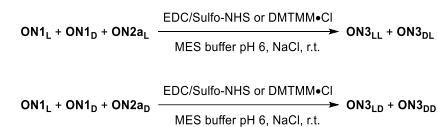
Strand	Slope, m (nmol ⁻¹)	Intercept, n	r ²	ε (M ⁻¹ ·cm ⁻¹)
ON1h ^a	-	-	-	68800
ON2b ^b	-	-	-	113400
ON3h ^b	9142.5	149.4	0.99	159182

The extinction coefficients were calculated using: ^a the OligoAnalyzer Version 3.0 from Integrated DNA Technologies and ^b an hypochromicity value of $h = 0.827$ for the section of the hairpin forming base pairs as reported in the literature.⁸

7. Pair-wise competitive peptide coupling reactions between donor and acceptor oligonucleotides

EDC/Sulfo-NHS and DMTMM·Cl as activators

The peptide coupling reactions were carried out under identical conditions to those described in Section 5 using EDC/Sulfo-NHS or DMTMM·Cl as activator. An equimolar solution of **ON1_L** and **ON1_D** was prepared in water and analyzed by HPLC. The 1:1 solution mixture of donor strands, **ON1_L** and **ON1_D**, was used to perform the coupling reactions with 1 equiv. of acceptor strand, **ON2a_L** or **ON2a_D**.



Scheme S4. Pair-wise competitive peptide coupling reactions between **ON1_L** and **ON1_D** with **ON2a_L** or **ON2a_D**.

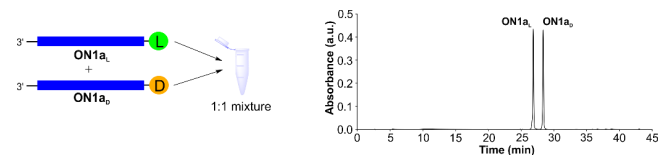


Figure S3. Example of an equimolar mixture of **ON1a_L** and **ON1a_D** analyzed by HPLC. The integration of the chromatographic peaks confirmed the presence of the two donor strands in equal amounts.

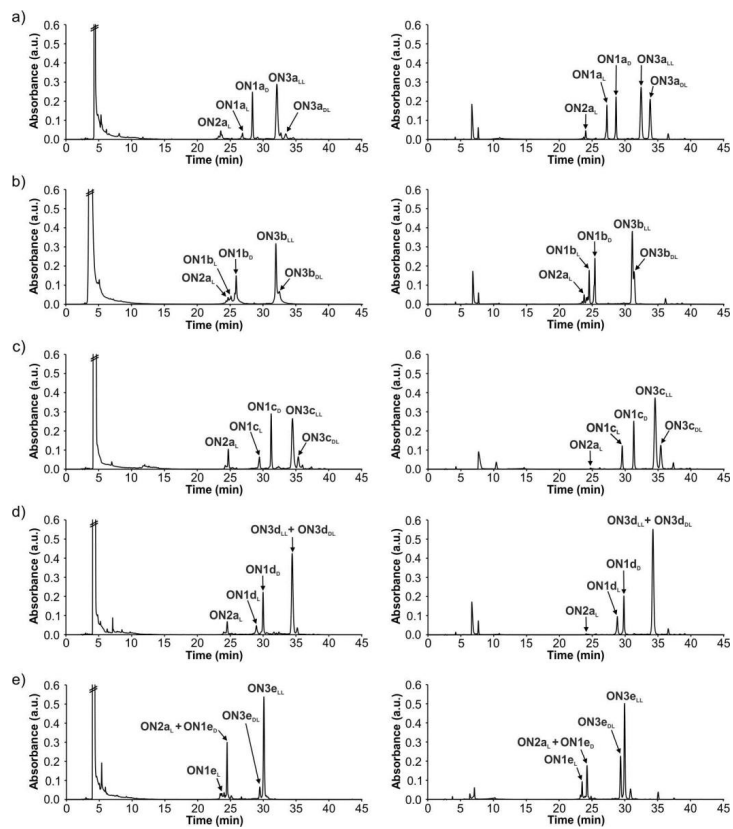


Figure S4. HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of **ON2a_L** with equimolar amounts of: a) **ON1a_L** and **ON1a_D**; b) **ON1b_L** and **ON1b_D**; c) **ON1c_L** and **ON1c_D**; d) **ON1d_L** and **ON1d_D** and e) **ON1e_L** and **ON1e_D**. Reactions performed with: left) EDC/Sulfo-NHS (2 h) and right) DMTMM-Cl (6 h).

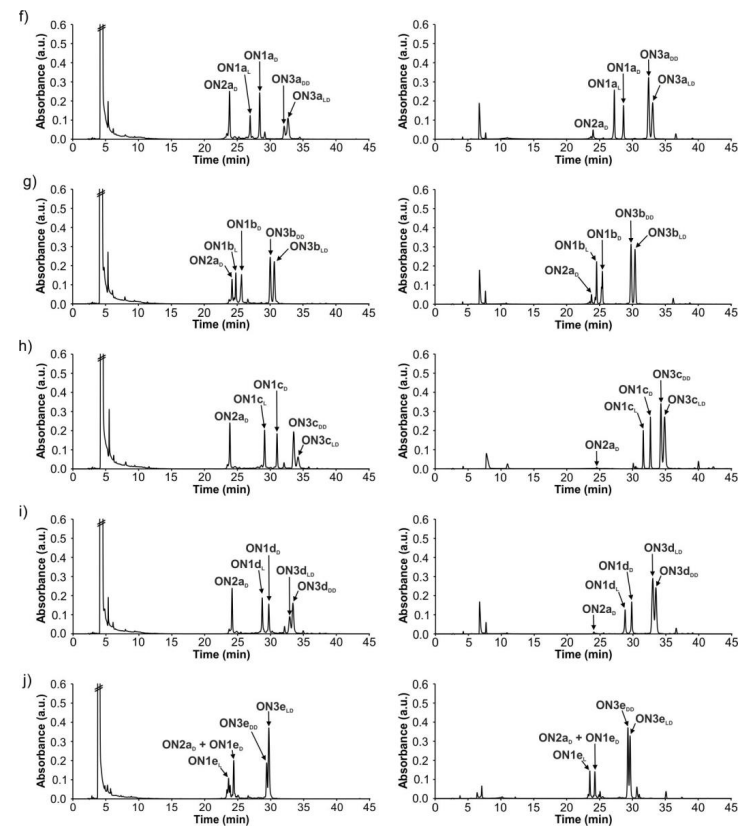


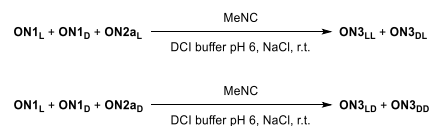
Figure S5. HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of **ON2a_D** with equimolar amounts of: f) **ON1a_D** and **ON1a_L**; g) **ON1b_D** and **ON1b_L**; h) **ON1c_D** and **ON1c_L**; i) **ON1d_D** and **ON1d_L** and j) **ON1e_D** and **ON1e_L**. Reactions performed with: left) EDC/Sulfo-NHS (2 h) and right) DMTMM-Cl (6 h).

Table S5. Results obtained in the pair-wise competitive peptide coupling reactions of **ON2a_L** and **ON2a_D** with equimolar amounts of **ON1_L** and **ON1_D** using EDC/Sulfo-NHS (or DMTMM-Cl) as activator (average of two experiments). Errors were determined to be lower than 10%.

Acceptor strand	Donor strands	Overall yield (%)	ON3 _{LL} /ON3 _{DL} ratio
ON2a_L	ON1a_L + ON1a_D	69 (90)	93:7 (58:42)
	ON1b_L + ON1b_D	72 (85)	80:20 (70:30)
	ON1c_L + ON1c_D	61 (99)	86:14 (77:23)
	ON1d_L + ON1d_D	74 (99)	n.d.
ON2a_D	ON1e_L + ON1e_D	73 (84)	91:9 (70:30)
	ON1a_L + ON1a_D	36 (90)	66:34 (41:59)
	ON1b_L + ON1b_D	55 (88)	52:48 (50:50)
	ON1c_L + ON1c_D	40 (99)	29:71 (52:48)
ON2a_D	ON1d_L + ON1d_D	48 (99)	38:62 (50:50)
	ON1e_L + ON1e_D	65 (84)	71:29 (47:53)

Methyl isonitrile (MeNC) in 4,5-dicyanoimidazole as activator

Stock solutions of 4,5-dicyanoimidazole (DCI) buffer pH 6 (100 mM), NaCl (4 M) and methyl isonitrile (MeNC) (500 mM) were prepared in water. Subsequently, equimolar amounts of **ON1** and **ON2** (1.5 nmol) were annealed at 95°C for 4 min in water containing NaCl (half of the volume required for the reaction). Finally, buffer, NaCl, activator solutions and water were added to the ON solution and the reaction was incubated in a ThermoMixer at 25°C for 24 h. The concentration of the components in the reaction mixture was: 50 μM of **ON1**, 50 μM of **ON2**, 50 mM of DCI, 100 mM of NaCl and 50 mM of MeNC.



Scheme S5. Pair-wise competitive peptide coupling reactions between **ON1_L** and **ON1_D** with **ON2a_L** or **ON2a_D**.

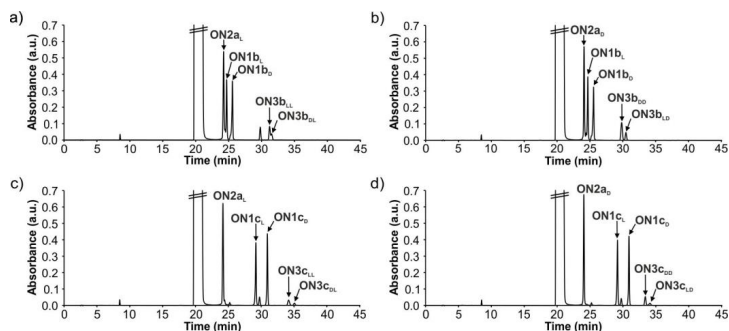


Figure S6. HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of equimolar amounts of **ON1b_L** and **ON1b_D** with a) **ON2a_L** and b) **ON2a_D** or equimolar amounts of **ON1c_L** and **ON1c_D** with c) **ON2a_L** and d) **ON2a_D**. Chromatographic peak at ca. 30 min is, probably, an adduct formed between MeNC and DCI.

Table S6. Results obtained in the pair-wise competitive peptide coupling reactions of **ON2a_L** and **ON2a_D** with equimolar amounts of **ON1_L** and **ON1_D** using MeNC as activator (average of two experiments). Errors were determined to be lower than 10%.

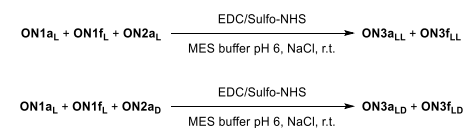
Acceptor strand	Donor strands	Overall yield (%)	ON3 _{LL} /ON3 _{DL} ratio
ON2a_L	ON1b_L + ON1b_D	10	75:25
	ON1c_L + ON1c_D	5	80:20
Acceptor strand	Donor strands	Overall yield (%)	ON3 _{LD} /ON3 _{DD} ratio
ON2a_D	ON1b_L + ON1b_D	16	28:72 ^a
	ON1c_L + ON1c_D	7	17:83

^a Maximum estimate due to possible signal overlap.

8. Pair-wise competitive peptide coupling reactions between donor and acceptor oligonucleotides containing peptides

Competitive reactions between two donor strands with one acceptor strand

The peptide coupling reactions were carried out under identical conditions to those described in Section 5 using EDC/Sulfo-NHS as activator. Equimolar solutions of **ON1a_L** and **ON1f_L** or **ON1a_L** and **ON1g_L** were prepared in water and analyzed by HPLC. The 1:1 solution mixture of donor strands was used to perform the coupling reactions with 1 equiv. of acceptor strand, **ON2a_L** or **ON2a_D**.



Scheme S6. Pair-wise competitive peptide coupling reactions between **ON1a_L** and **ON1f_L** with **ON2a_L** or **ON2a_D**. The same competitive reactions were performed by replacing **ON1f_L** by **ON1g_L**.

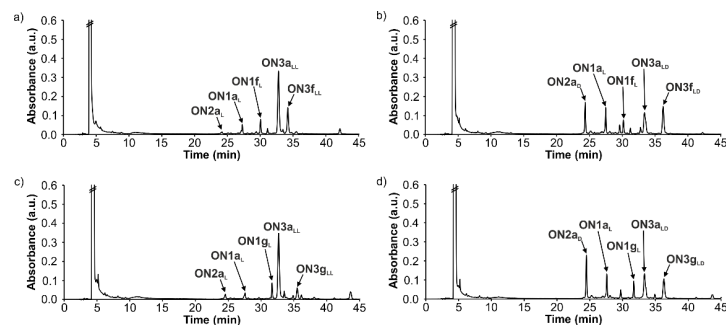


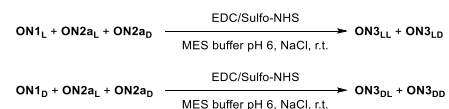
Figure S7. HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of equimolar amounts of **ON1a_L** and **ON1f_L** with a) **ON2a_L** and b) **ON2a_D**, or equimolar amounts of **ON1a_L** and **ON1g_L** with c) **ON2a_L** and d) **ON2a_D**. Reactions performed with EDC/Sulfo-NHS (2 h).

Table S7. Results obtained in the pair-wise competitive peptide coupling reactions of **ON2a_L** and **ON2a_D** with equimolar amounts of **ON1a_L** and **ON1f_L**, or equimolar amounts of **ON1a_L** and **ON1g_L**, using EDC/Sulfo-NHS as activator (average of two experiments). Errors were determined to be lower than 10%.

Acceptor strand	Donor strands	Overall yield (%)	ON3a _{LL} /ON3(f/g) _{LL} ratio
ON2a_L	ON1a_L + ON1f_L	79	73:27
	ON1a_L + ON1g_L	71	89:11
Acceptor strand	Donor strands	Overall yield (%)	ON3a _{LD} /ON3(f/g) _{LD} ratio
ON2a_D	ON1a_L + ON1f_L	49	51:49
	ON1a_L + ON1g_L	33	58:42

Competitive reactions between two acceptor strands with one donor strand

The peptide coupling reactions were carried out under identical conditions to those described in Section 5 using EDC/Sulfo-NHS as activator. An equimolar solution of **ON2a_L** and **ON2a_D** was prepared in water and analyzed by HPLC. The 1:1 solution mixture of acceptor strands, **ON2a_L** and **ON2a_D**, was used to perform the coupling reactions with 1 equiv. of donor strand, **ON1a_L**, **ON1a_D**, **ON1f_L** or **ON1g_L**.



Scheme S7. Pair-wise competitive peptide coupling reactions between **ON2a_L** and **ON2a_D** with **ON1_L** or **ON1_D**.

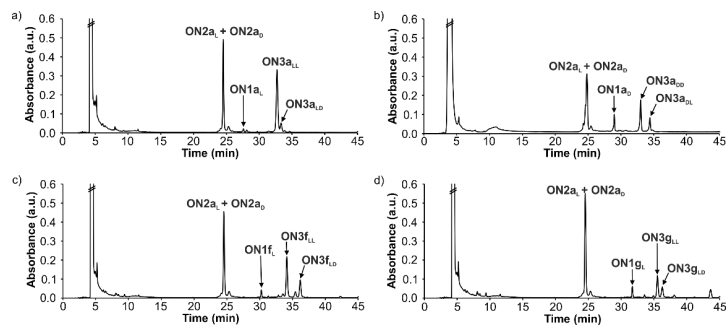


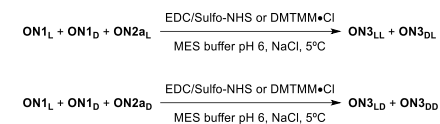
Figure S8. HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of equimolar amounts of **ON2a_L** and **ON2a_D** with a) **ON1a_L**; b) **ON1a_D**; c) **ON1f_L**, and d) **ON1g_L**. Reactions performed with EDC/Sulfo-NHS (2 h).

Table S8. Results obtained in the pair-wise competitive peptide coupling reactions of **ON2a_L** and **ON2a_D** with **ON1a_L**, **ON1a_D**, **ON1f_L** or **ON1g_L**, using EDC/Sulfo-NHS as activator (average of two experiments). Errors were determined to be lower than 10%.

Acceptor strands	Donor strand	Overall yield (%)	ON3 _{LL} /ON3 _{LD} ratio
ON2a_L + ON2a_D	ON1a_L	73	87:13
	ON1f_L	54	66:34
	ON1g_L	32	67:33
Acceptor strands	Donor strand	Overall yield (%)	ON3 _{DL} /ON3 _{DD} ratio
ON2a_L + ON2a_D	ON1a_D	35	31:69

9. Selected pair-wise competitive peptide coupling reactions at low temperature

The peptide coupling reactions were carried out at 5°C using identical conditions to those described in Section 5 and Section 7 using EDC/Sulfo-NHS or DMTMM-Cl as activator.



Scheme S8. Pair-wise competitive peptide coupling reactions between **ON1_L** and **ON1_D** with **ON2a_L** or **ON2a_D** at 5°C.

Table S9. Results obtained in the pair-wise competitive peptide coupling reactions of **ON2a_L** and **ON2a_D** with equimolar amounts of **ON1_L** and **ON1_D** using EDC/Sulfo-NHS (or DMTMM-Cl) as activator (average of two experiments) at 5°C. Errors were determined to be lower than 10%.

Acceptor strand	Donor strands	Overall yield (%)	ON3 _{LL} /ON3 _{DL} ratio
ON2a_L	ON1a_L + ON1a_D	65 (90)	98:2 (58:42)
	ON1c_L + ON1c_D	43 (96)	88:12 (83:17)
Acceptor strand	Donor strands	Overall yield (%)	ON3 _{LD} /ON3 _{DD} ratio
ON2a_D	ON1a_L + ON1a_D	26 (94)	58:42 (47:53)
	ON1c_L + ON1c_D	27 (91)	23:77 (64:36)

10. Kinetic experiments with different activators

The peptide coupling reactions were carried out under identical conditions to those described in Section 5 using EDC/Sulfo-NHS or DMTMM-Cl as activator. The coupling data (concentration of product vs. time) was fit to the corresponding theoretical kinetic model using the Parameter Estimation Module of COPASI software Version 4.29.⁹ We introduced the theoretical kinetic model shown below:

Double strand \rightarrow Hairpin strand; k_{app}

The initial concentration of the double strand was refined as variable but constrained between 35 and 50×10^{-6} M. The fit of the data returned the rate constant value (k_{app}).

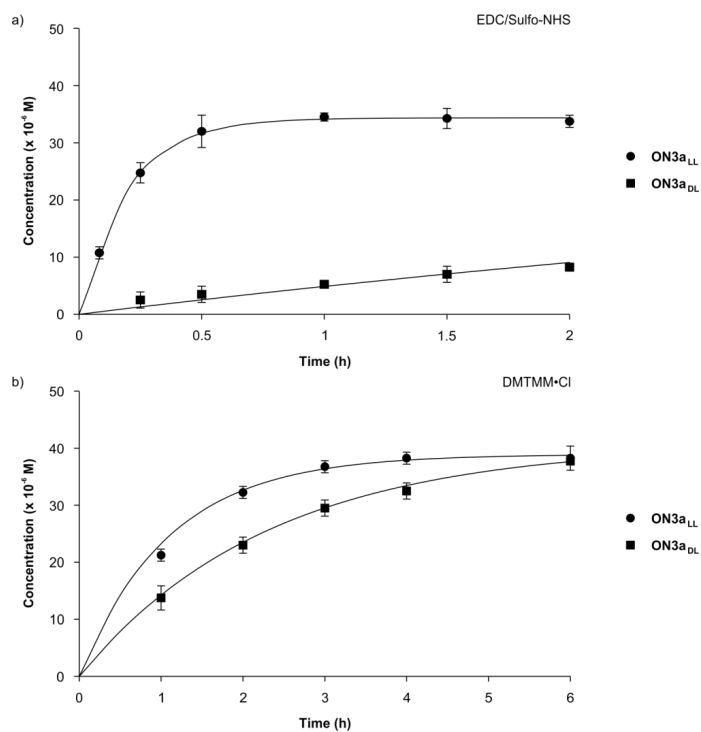


Figure S9. Concentration of the product (M) vs. time (h) for the peptide coupling reactions of **ON1a_L** and **ON1a_D** with **ON2a_L** using: a) EDC/Sulfo-NHS and b) DMTMM-Cl. Lines show fit of the data to the theoretical kinetic model. Error bars are the standard deviations.

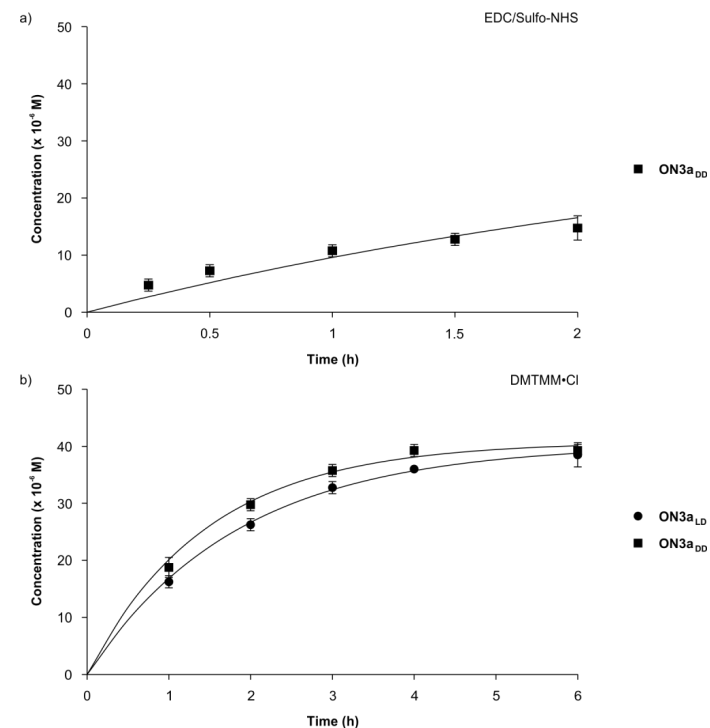


Figure S10. Concentration of the product (M) vs. time (h) for the peptide coupling reactions of **ON1a_L** and **ON1a_D** with **ON2a_D** using: a) EDC/Sulfo-NHS and b) DMTMM-Cl. Lines show fit of the data to the theoretical kinetic model. Error bars are the standard deviations.

Table S10. Calculated rate constant values (k_{app}) for the peptide coupling reactions of **ON1a_L** or **ON1a_D** with **ON2a_L** or **ON2a_D** (average of two experiments). Errors are indicated as standard deviations.

Activator	Donor strand	Acceptor strand	Hairpin strand	k_{app} (h ⁻¹)	Calcd. ratio ^a
EDC/Sulfo-NHS	ON1a_L ; $X = m^{\delta}v_L^{\delta}A_m$	ON2a_L ; $Y = v_L - mnm^{\delta}U_m$	ON3a_{LL}	5.04 ± 0.12	87
	ON1a_D ; $X = m^{\delta}v_D^{\delta}A_m$		ON3a_{DL}	0.15 ± 0.01	13
	ON1a_L ; $X = m^{\delta}v_L^{\delta}A_m$	ON2a_D ; $Y = v_D - mnm^{\delta}U_m$	ON3a_{LD}	n.d.	n.d.
	ON1a_D ; $X = m^{\delta}v_D^{\delta}A_m$		ON3a_{DD}	0.32 ± 0.01	n.d.
DMTMM-Cl	ON1a_L ; $X = m^{\delta}v_L^{\delta}A_m$	ON2a_L ; $Y = v_L - mnm^{\delta}U_m$	ON3a_{LL}	0.91 ± 0.07	59
	ON1a_D ; $X = m^{\delta}v_D^{\delta}A_m$		ON3a_{DL}	0.43 ± 0.01	41
	ON1a_L ; $X = m^{\delta}v_L^{\delta}A_m$	ON2a_D ; $Y = v_D - mnm^{\delta}U_m$	ON3a_{LD}	0.54 ± 0.01	47
	ON1a_D ; $X = m^{\delta}v_D^{\delta}A_m$		ON3a_{DD}	0.68 ± 0.03	53

^a Product ratio calculated from experimental rate constants using the Time Course Module of COPASI software Version 4.29.⁹ The calculated ratios are similar to those determined in the pair-wise competition experiments. n.d. = not determined due to signal overlap with activated species.

11. Melting curves

The UV melting curves were measured on a JASCO V-650 spectrometer at 260 nm using 10 mm QS cuvettes with a scanning rate of 1°C·min⁻¹. The obtained UV spectroscopic data were fit to a two-state melting model, *i.e.* double strand – random coil equilibrium, using a mono-sigmoidal Boltzmann function.¹⁰ The fit of the data returned the melting temperature.

For the experiments, we prepared aqueous solutions containing equimolar amounts of the oligonucleotides (ca. 5 μM), 10 mM phosphate buffer at pH 7 and 150 mM NaCl. The oligonucleotides were annealed by heating to 95°C for 4 min and, subsequently, by cooling down slowly to 5°C before the variable-temperature UV spectroscopic experiments.

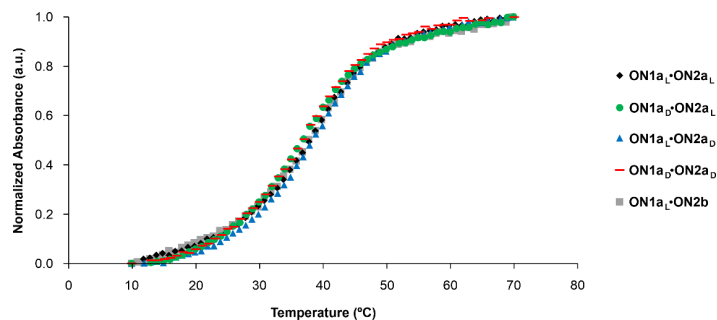
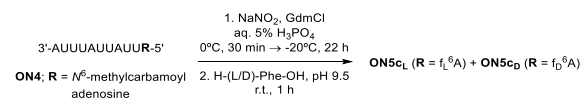


Figure S11. Melting curves of the RNA donor strands **ON1a_L** and **ON1a_P** with the RNA acceptor strands **ON2a_L**, **ON2a_D** and **ON2b**. The fit of the data to a two-state melting model using a mono-sigmoidal Boltzmann function returned very similar melting temperatures for all double strands, $T_m = 37.1 \pm 0.8^\circ\text{C}$.

12. Loading reaction of a selected amino acid onto an RNA strand

We synthesized the RNA strand **ON4**, bearing an *N*⁶-methylcarbamoyl adenosine at the 5'-end [3'-AUUUUUUUUUR-5' (R = *N*⁶-methylcarbamoyl adenosine)], following a procedure previously reported in the literature.¹¹ Next, we performed the loading reaction according to the following protocol: Step 1. **ON4** (1.0 equiv.) and NaNO₂ (500 equiv.) were dissolved in 5% aq. H₃PO₄ solution (20 μL/nmol) containing 100 mM guanidinium chloride (GdmCl). The reaction was stirred at 0°C for 30 min and kept in a freezer at -22°C for 22 h. Step 2. A racemic mixture of **H-Phe-OH** (1000 equiv.) in 30 mM borate-buffered solution (150 μL) was added to the thawed oligonucleotide's solution. The pH was adjusted to ca. 9.5 using a 4 M aq. NaOH solution. The reaction was stirred at r.t. for 1 h. Finally, the reaction was neutralised with 1 M aq. HCl solution. An aliquot (22.5 μL) of the crude reaction mixture was taken, diluted with water (up to 100 μL) and analyzed by HPLC (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-40% of B in 45 min; Flow rate = 1 mL·min⁻¹; Injection volume = 100 μL).



Scheme S9. Loading reaction of a racemic mixture of L- and D-Phe onto **ON4**.

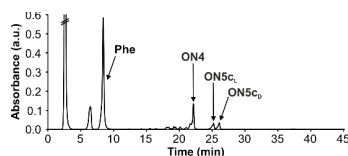


Figure S12. HPLC-chromatogram of the crude reaction mixture for the loading reaction of a racemic mixture of L- and D-Phe onto **ON4**. Integration of the chromatographic peaks indicated the formation of **ON5c_L** and **ON5c_D** in equal amounts.

Table S11. Results obtained in the loading reaction of a racemic mixture of L- and D-Phe onto **ON4** (average of two experiments). Errors were determined to be lower than 10%.

RNA strand	Amino acids	Overall yield (%) ^a	ON5c _L /ON5c _D ratio ^b
ON4	L-Phe + D-Phe	35	1:1 (765:765)

^a Calculated using the areas of the chromatographic peaks in relation to that determined for a reference compound. ^b Areas (a.u.) of the chromatographic peaks in parenthesis.

13. References

- G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176-2179.
- F. Müller, L. Escobar, F. Xu, E. Węgrzyn, M. Nainytė, T. Amatov, C. Y. Chan, A. Pichler, T. Carell, *Nature* **2022**, *605*, 279-284.
- M. Nainytė, F. Müller, G. Ganazzoli, C.-Y. Chan, A. Crisp, D. Globisch, T. Carell, *Chem. Eur. J.* **2020**, *26*, 14856-14860.
- X. Li, J. Wu, X. Li, W. Mu, X. Liu, Y. Jin, W. Xu, Y. Zhang, *Biorg. Med. Chem.* **2015**, *23*, 6258-6270.
- F. Himmelsbach, B. S. Schulz, T. Trichtinger, R. Charubala, W. Pfeleiderer, *Tetrahedron* **1984**, *40*, 59-72.
- F. Ferreira, F. Morvan, *Nucleosides, Nucleotides, and Nucleic Acids* **2005**, *24*, 1009-1013.
- D. L. Usanov, A. I. Chan, J. P. Maianti, D. R. Liu, *Nature Chem.* **2018**, *10*, 704-714.
- A. V. Tataurov, Y. You, R. Owczarzy, *Biophys. Chem.* **2008**, *133*, 66-70.
- S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, U. Kummer, *Bioinformatics*, **2006**, *22*, 3067-3074.
- S. G. J. Mochrie, *Am. J. Phys.* **2011**, *79*, 1121-1126.
- J. N. Singer, F. M. Müller, E. Węgrzyn, C. Hölzl, H. Hurmiz, C. Liu, L. Escobar, T. Carell, *Angew. Chem. Int. Ed.* **2023**, *62*, e202302360.

Appendix IV

Gradual evolution of a homo-L-peptide world on homo-D-configured RNA and DNA

Ewa Węgrzyn,^{*a} Ivana Mejdrová,^{*a} and Thomas Carell^a

^a Department of Chemistry, Center for Nucleic Acids Therapies at the Institute for Chemical Epigenetics (ICE-M), Ludwig-Maximilians-Universität (LMU) München, Butenandtstrasse 5-13, 81377 Munich (Germany)

* These authors contributed equally.

E-Mail: thomas.carell@lmu.de

Supporting Information

Table of Contents

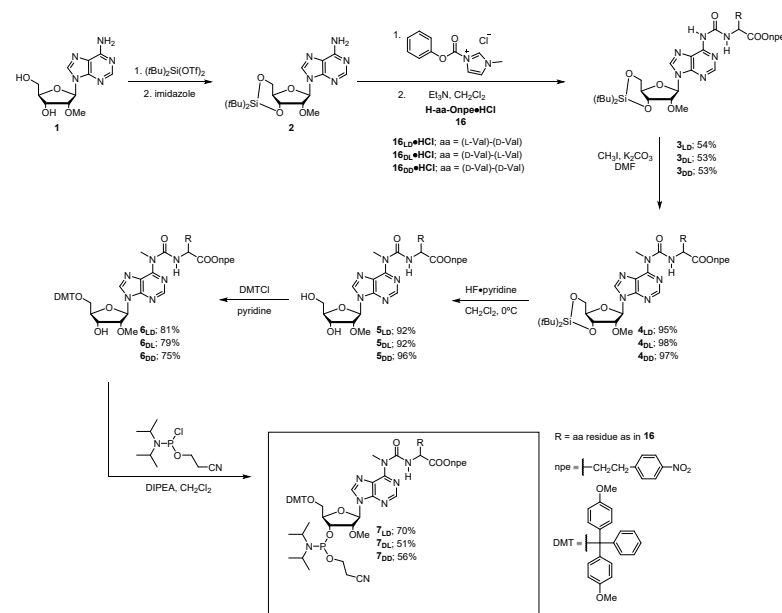
1. General information and instruments for nucleosides and phosphoramidites	2
2. Synthesis and characterization data of phosphoramidites.....	2
2.1 Amino acid-modified methyl <i>N</i> ⁶ -carbamoyl adenosine phosphoramidites.....	2
2.2 Amino acid-modified methyl <i>N</i> ⁶ -carbamoyl deoxy adenosine phosphoramidites	17
2.3 Npe-protected amino acids.....	28
2.4 5-Methylaminomethyl uridine phosphoramidite	35
2.5 5-Aminomethyl thymidine phosphoramidite.....	35
3. General information and instruments for oligonucleotides.....	42
4. Synthesized oligonucleotides using a DNA/RNA automated synthesizer.....	43
4.1 Donor strands containing amino acid-modified methyl <i>N</i> ⁶ -carbamoyl adenosine at the 5'-end.....	43
4.2 Acceptor strands containing 5-methylaminomethyl uridine at the 3'-end	43
4.3 Donor strands containing amino acid-modified methyl <i>N</i> ⁶ -carbamoyl deoxy adenosine at the 5'-end.....	43
4.4 Acceptor strands containing 5-aminomethyl thymidine at the 3'-end	43
5. Peptide coupling reactions between donor and acceptor oligonucleotides	44
6. Calibration curve hairpin	45
7. Cleavage reaction of hairpin products	46
8. Competitive coupling and cleavage reactions (one pot)	47
9. Pair-wise competitive peptide coupling reactions between donor and acceptor oligonucleotides containing dipeptides.....	48
10. Pair-wise competitive peptide coupling reactions between donor and acceptor deoxy oligonucleotides	51
11. Circular dichroism measurements	52
12. References.....	52

1. General information and instruments for nucleosides and phosphoramidites

Reagents were purchased from commercial suppliers and used without further purification unless otherwise stated. Anhydrous solvents, stored under inert atmosphere, were also purchased. All reactions involving air/moisture sensitive reagents/intermediates were performed under inert atmosphere using oven-dried glassware. Routine ¹H NMR, ¹³C NMR and ³¹P NMR spectra were recorded on a Bruker Ascend 500 spectrometer (500 MHz for ¹H NMR, 125 MHz for ¹³C NMR and 202 MHz for ³¹P NMR). Deuterated solvents used are indicated in the characterization and chemical shifts (δ) are reported in ppm. Residual solvent peaks were used as reference.¹ All NMR *J* values are given in Hz. COSY, HMQC and HMBC NMR experiments were recorded to help with the assignment of ¹H and ¹³C signals. NMR spectra were analysed using MestReNova software version 10.0. High Resolution Mass Spectra (HRMS) were measured on a Thermo Finnigan LTQ-FT with ESI as ionization mode. IR spectra were recorded on a Perkin-Elmer Spectrum BX II FT-IR instrument or Shimadzu IRSpirit FT-IR instrument. Both equipped with an ATR accessory. Column chromatography was performed with technical grade silica gel, 40-63 μm particle size. Reaction progress was monitored by Thin Layer Chromatography (TLC) analysis on silica gel 60 F254 and stained with *para*-anisaldehyde, potassium permanganate or ninhydrin solution.

2. Synthesis and characterization data of phosphoramidites

2.1 Amino acid-modified methyl *N*⁶-carbamoyl adenosine phosphoramidites

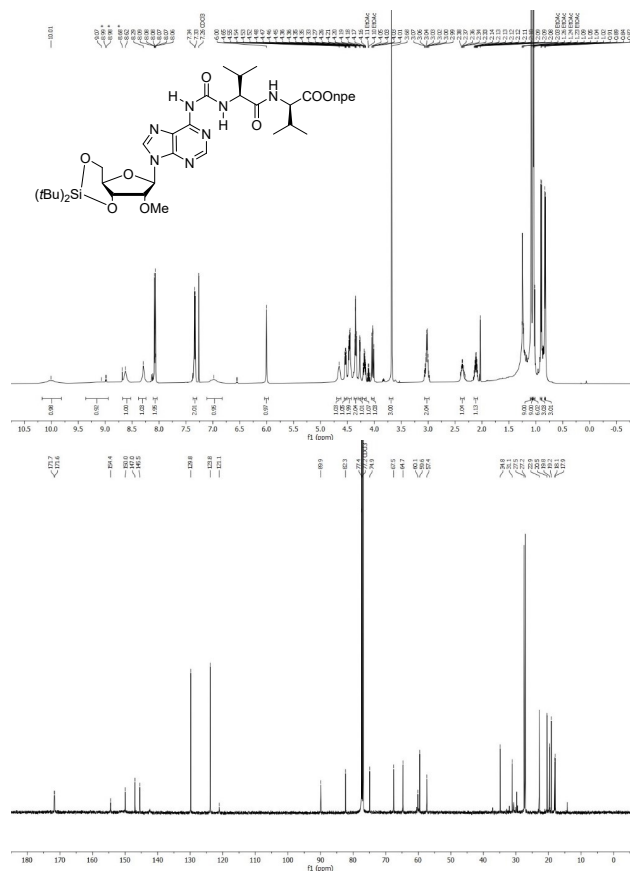


Scheme S1. Synthesis of amino acid-modified methyl *N*⁶-carbamoyl adenosine phosphoramidites 7.

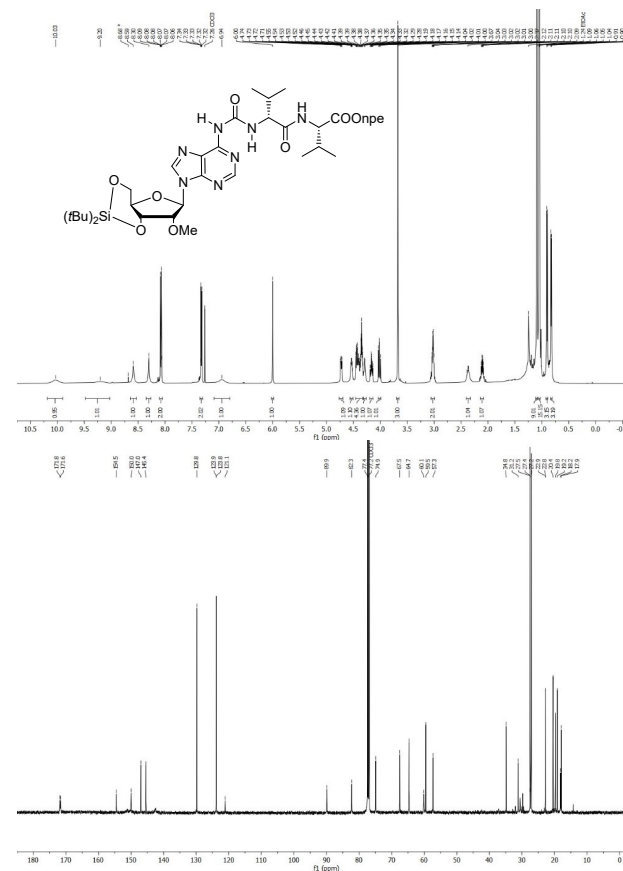
Compound **2** was synthesized following a procedure previously reported in the literature.²

General procedure for the synthesis of compound 3: Step 1. Compound **2** (1 equiv.) and 1-*N*-methyl-3-phenoxy-carbonyl-imidazolium chloride (2 equiv.) were added to an oven dried round-bottom flask and kept under high-vacuum for 15 min. After that, dry CH₂Cl₂ was added under nitrogen atmosphere and the reaction was stirred at r.t. for 5 h. Step 2. Onpe-protected amino acid **16**•HCl (2 equiv.) was added to an oven-dried round-bottom flask and suspended in dry CH₂Cl₂, followed by the addition of Et₃N (2 equiv.). The suspension was added dropwise to the reaction mixture. The reaction was stirred at r.t. under nitrogen atmosphere for 20 h. After that, the reaction was quenched with aqueous saturated NaHCO₃. The organic layer was separated and the crude was further extracted with CH₂Cl₂. The combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

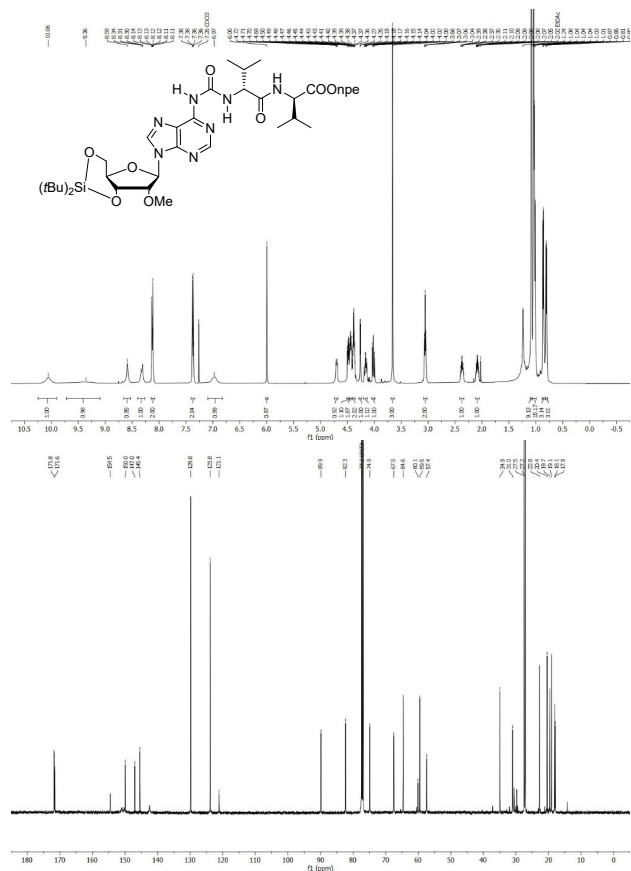
3_{1b}: Yield = 54%. Rf = 0.30 (EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.01 (s, 1H); 9.07 (s, 1H); 8.62 (s, 1H); 8.29 (s, 1H); 8.11 – 8.04 (m, 2H); 7.34 (d, *J* = 8.4 Hz, 2H); 6.98 (s, 1H); 6.00 (s, 1H); 4.65 (s, 1H); 4.53 (dd, *J* = 8.7 Hz, *J* = 5.0 Hz, 1H); 4.46 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 2H); 4.38 – 4.32 (m, 2H); 4.27 (d, *J* = 4.6 Hz, 1H); 4.18 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.03 (dd, *J* = 10.5 Hz, *J* = 9.2 Hz, 1H); 3.68 (s, 3H); 3.03 (hept, *J* = 8.1 Hz, *J* = 7.4 Hz, 2H); 2.36 (h, *J* = 6.7 Hz, 1H); 2.11 (pd, *J* = 6.9 Hz, *J* = 5.2 Hz, 1H); 1.09 (s, 9H); 1.05 (s, 9H); 1.03 (d, *J* = 9.2 Hz, 6H); 0.90 (d, *J* = 6.8 Hz, 3H); 0.83 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.7; 171.6; 154.4; 150.0; 147.0; 145.5; 129.8; 123.8; 121.1; 89.9; 82.3; 77.4; 74.9; 67.5; 64.7; 60.1; 59.6; 57.4; 34.8; 31.1; 27.5; 27.2; 22.9; 20.5; 19.8; 19.2; 18.1; 17.9. FTIR ν_{max} (cm⁻¹): 2933 (w); 2860 (w); 1677 (m); 1520 (s); 1467 (m); 1345 (s); 1257 (m); 1138 (s); 1063 (s); 828 (s). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₈H₅₆O₁₀N₈SiNa 835.3780; Found 835.3803.



3_{1c}: Yield = 53%. Rf = 0.36 (EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.03 (s, 1H); 9.20 (s, 1H); 8.59 (s, 1H); 8.30 (s, 1H); 8.10 – 8.05 (m, 2H); 7.35 – 7.31 (m, 2H); 6.94 (s, 1H); 6.00 (s, 1H); 4.73 (dd, *J* = 9.6 Hz, *J* = 4.7 Hz, 1H); 4.56 – 4.51 (m, 1H); 4.47 – 4.31 (m, 4H); 4.31 – 4.27 (m, 1H); 4.16 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.02 (dd, *J* = 10.5 Hz, *J* = 9.2 Hz, 1H); 3.67 (s, 3H); 3.02 (td, *J* = 6.7 Hz, *J* = 3.8 Hz, 2H); 2.42 – 2.32 (m, 1H); 2.15 – 2.07 (m, 1H); 1.09 (s, 9H); 1.06 – 1.03 (m, 15H); 0.90 (d, *J* = 6.8 Hz, 3H); 0.83 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.8; 171.6; 154.5; 150.0; 147.0; 145.4; 129.8; 123.9; 123.8; 121.1; 89.9; 82.3; 77.4; 74.9; 67.5; 64.7; 60.1; 59.5; 57.3; 34.8; 31.2; 27.5; 27.4; 27.2; 22.9; 22.8; 20.4; 19.8; 19.2; 18.2; 17.9. FTIR ν_{max} (cm⁻¹): 2934 (w); 2860 (w); 1677 (m); 1520 (s); 1466 (m); 1345 (s); 1134 (m); 1063 (m); 828 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₈H₅₆O₁₀N₈SiNa 835.3780; Found 835.3801.



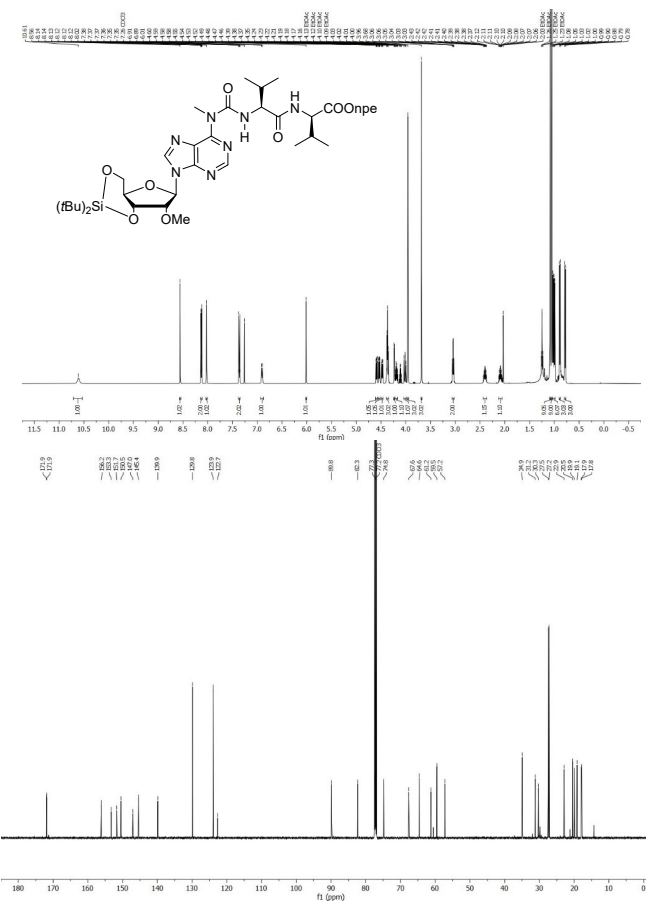
3_{0D}: Yield = 53%. R_f = 0.38 (EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.06 (s, 1H); 9.36 (s, 1H); 8.59 (s, 1H); 8.32 (t, J = 8.1 Hz, 1H); 8.12 (td, J = 5.3 Hz, J = 2.6 Hz, 2H); 7.39 – 7.35 (m, 2H); 6.97 (s, 1H); 6.00 (s, 1H); 4.70 (dd, J = 10.0 Hz, J = 4.6 Hz, 1H); 4.49 (dd, J = 8.4 Hz, J = 5.1 Hz, 1H); 4.44 (dt, J = 6.8 Hz, J = 3.3 Hz, 2H); 4.38 (td, J = 6.7 Hz, J = 3.9 Hz, 2H); 4.26 (d, J = 4.6 Hz, 1H); 4.16 (td, J = 10.0 Hz, J = 4.9 Hz, 1H); 4.02 (dd, J = 10.5 Hz, J = 9.2 Hz, 1H); 3.66 (s, 3H); 3.06 (t, J = 6.7 Hz, 2H); 2.37 (h, J = 6.8 Hz, 1H); 2.12 – 2.05 (m, 1H); 1.08 (s, 9H); 1.06 – 1.00 (m, 15H); 0.86 (d, J = 6.8 Hz, 3H); 0.81 (d, J = 6.7 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.8; 171.6; 154.5; 150.0; 147.0; 145.4; 129.8; 123.8; 121.1; 89.9; 82.3; 74.9; 67.5; 64.6; 60.1; 59.5; 57.4; 34.9; 31.0; 27.5; 27.2; 22.8; 20.4; 19.7; 19.1; 18.1; 17.9. FTIR ν_{max} (cm⁻¹): 2934 (w); 2860 (w); 1676 (m); 1520 (s); 1467 (m); 1345 (s); 1138 (m); 1063 (m); 828 (m). HRMS (ESI) m/z: [M+Na]⁺ Calcd for C₃₈H₅₆O₁₀N₈SiNa 835.3780; Found 835.3802.



S5

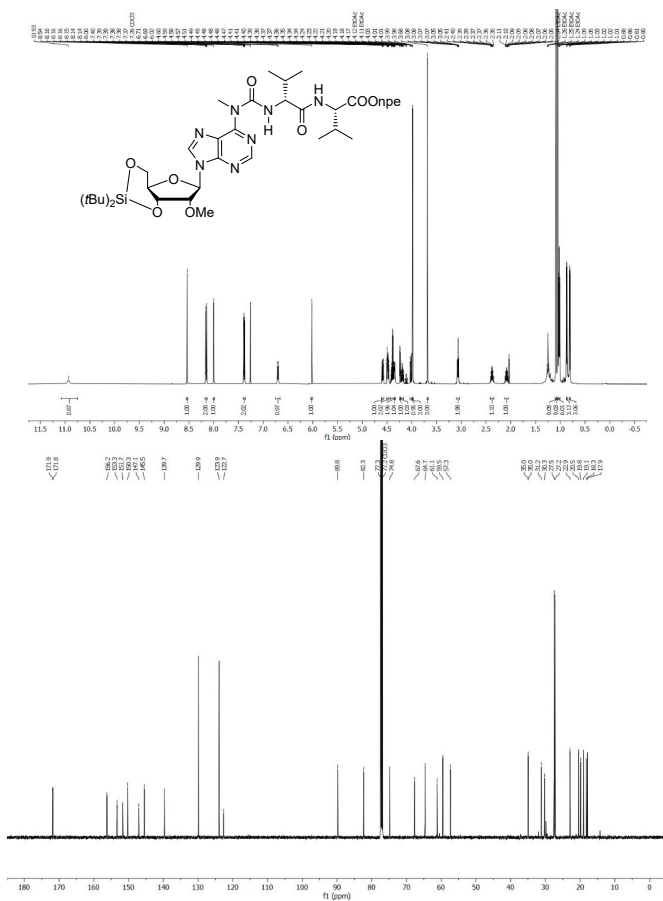
General procedure for the synthesis of compound 4: Compound **3** (1 equiv.) and K₂CO₃ (3 equiv.) were suspended in dry DMF. The suspension was stirred for 30 min at 0°C under nitrogen atmosphere. After that, CH₃I (2 equiv.) was added dropwise and the reaction was stirred at r.t. overnight. After that, the reaction mixture was diluted with CH₂Cl₂ and washed with aqueous saturated NH₄Cl and water. The organic layer was dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

4₀: Yield = 95%. R_f = 0.57 (3:7 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.61 (s, 1H); 8.56 (s, 1H); 8.15 – 8.11 (m, 2H); 8.02 (s, 1H); 7.38 – 7.34 (m, 2H); 6.90 (d, J = 8.7 Hz, 1H); 6.01 (s, 1H); 4.59 (dd, J = 9.7 Hz, J = 4.6 Hz, 1H); 4.54 (dd, J = 8.7 Hz, J = 5.0 Hz, 1H); 4.48 (dd, J = 9.2 Hz, J = 5.0 Hz, 1H); 4.37 (t, J = 6.7 Hz, 3H); 4.23 (d, J = 4.6 Hz, 1H); 4.19 (td, J = 10.1 Hz, J = 5.0 Hz, 1H); 4.01 (dd, J = 10.5 Hz, J = 9.2 Hz, 1H); 3.96 (s, 3H); 3.68 (s, 3H); 3.04 (td, J = 6.7 Hz, J = 1.6 Hz, 2H); 2.40 (pd, J = 6.9 Hz, J = 5.1 Hz, 1H); 2.09 (pd, J = 6.9 Hz, J = 5.0 Hz, 1H); 1.08 (s, 9H); 1.05 (s, 9H); 1.01 (dd, J = 14.0 Hz, J = 6.8 Hz, 6H); 0.89 (d, J = 6.8 Hz, 3H); 0.78 (d, J = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 171.9; 156.2; 153.3; 151.7; 150.5; 147.0; 145.4; 139.9; 129.8; 123.9; 122.7; 89.8; 82.3; 77.3; 74.8; 67.6; 64.6; 61.2; 59.5; 57.2; 34.9; 31.2; 30.3; 27.5; 27.2; 22.9; 20.5; 19.9; 19.1; 17.9; 17.8. FTIR ν_{max} (cm⁻¹): 2934 (w); 2860 (w); 1670 (m); 1578 (m); 1520 (s); 1466 (m); 1345 (m); 1137 (m); 1065 (m); 1012 (m); 828 (m). HRMS (ESI) m/z: [M+Na]⁺ Calcd for C₃₉H₅₈O₁₀N₈SiNa 849.3934; Found 849.3958.

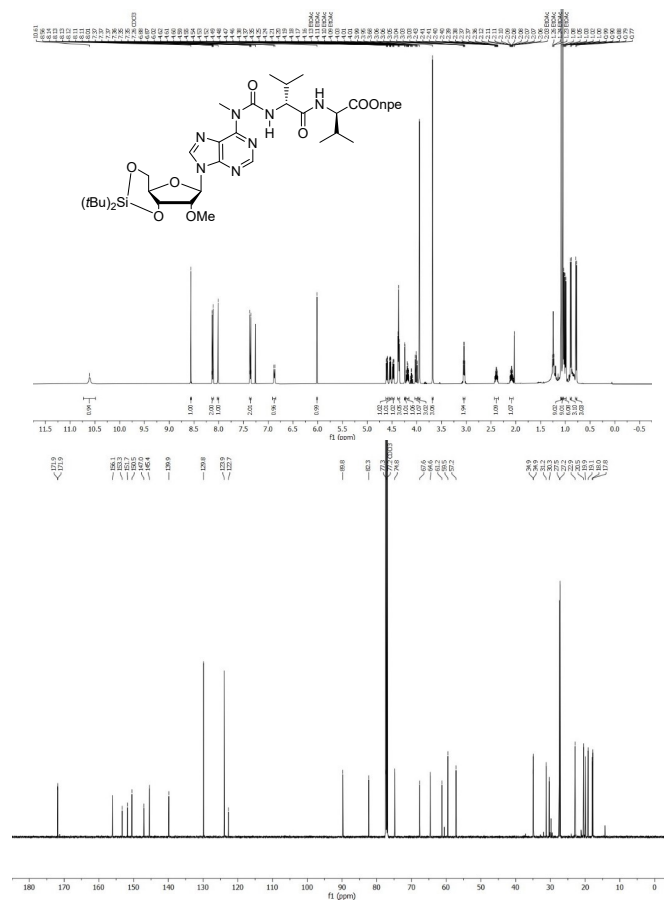


S6

4_{6i}: Yield = 98%. R_f = 0.64 (3:7 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.93 (s, 1H); 8.54 (s, 1H); 8.17 – 8.13 (m, 2H); 8.00 (s, 1H); 7.40 – 7.37 (m, 2H); 6.70 (d, *J* = 8.5 Hz, 1H); 6.02 (s, 1H); 4.59 (dd, *J* = 9.7 Hz, *J* = 4.6 Hz, 1H); 4.51 – 4.46 (m, 2H); 4.39 (q, *J* = 6.8 Hz, 2H); 4.35 (dd, *J* = 6.3 Hz, *J* = 2.8 Hz, 1H); 4.24 (d, *J* = 4.6 Hz, 1H); 4.19 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.01 (dd, *J* = 10.6 Hz, *J* = 9.4 Hz, 1H); 3.98 (s, 3H); 3.68 (s, 3H); 3.07 (td, *J* = 6.6 Hz, *J* = 1.4 Hz, 2H); 2.42 – 2.34 (m, 1H); 2.09 (pd, *J* = 6.9 Hz, *J* = 5.0 Hz, 1H); 1.09 (s, 9H); 1.05 (s, 9H); 1.02 (dd, *J* = 6.8 Hz, *J* = 5.4 Hz, 6H); 0.87 (d, *J* = 6.9 Hz, 3H); 0.81 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 171.8; 156.2; 153.3; 151.7; 150.3; 147.1; 145.5; 139.7; 129.9; 123.9; 122.7; 89.8; 82.3; 77.3; 74.8; 67.6; 64.7; 61.1; 59.5; 57.3; 35.0; 35.0; 31.2; 30.3; 27.5; 27.2; 22.9; 20.5; 19.8; 19.1; 18.3; 17.9. FTIR ν_{max} (cm⁻¹): 2934 (w); 2860 (w); 1670 (m); 1578 (m); 1520 (s); 1466 (m); 1345 (m); 1137 (m); 1065 (m); 1013 (m); 828 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₉H₅₈O₁₀N₈SiNa 849.3934; Found 849.3957.

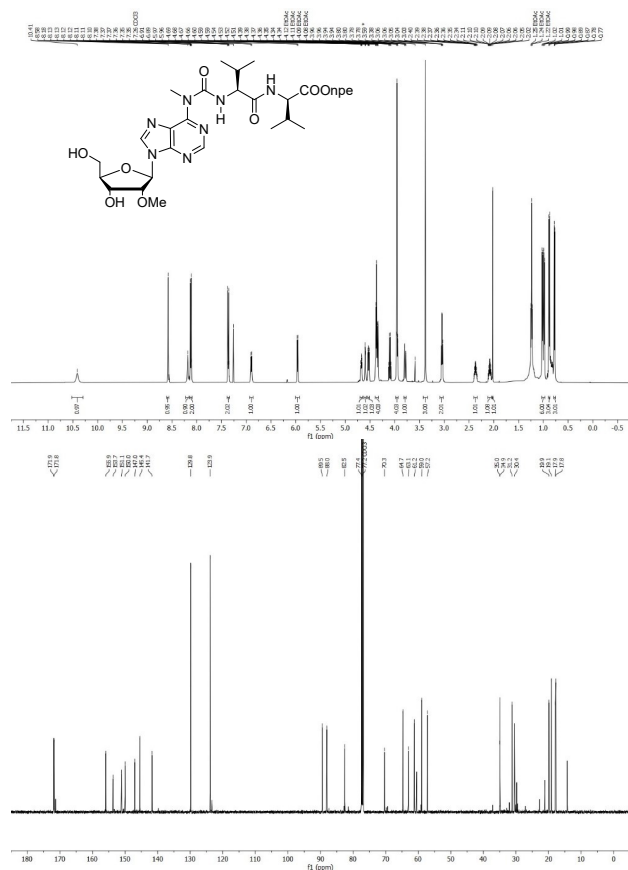


4_{6o}: Yield = 97%. R_f = 0.64 (3:7 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.61 (s, 1H); 8.56 (s, 1H); 8.14 – 8.10 (m, 2H); 8.01 (s, 1H); 7.38 – 7.35 (m, 2H); 6.88 (d, *J* = 8.8 Hz, 1H); 6.02 (s, 1H); 4.60 (dd, *J* = 9.7 Hz, *J* = 4.7 Hz, 1H); 4.54 (dd, *J* = 8.8 Hz, *J* = 4.9 Hz, 1H); 4.48 (dd, *J* = 9.2 Hz, *J* = 5.0 Hz, 1H); 4.37 (t, *J* = 6.6 Hz, 3H); 4.24 (d, *J* = 4.6 Hz, 1H); 4.19 (td, *J* = 10.1 Hz, *J* = 5.0 Hz, 1H); 4.01 (dd, *J* = 10.6 Hz, *J* = 9.2 Hz, 1H); 3.95 (s, 3H); 3.68 (s, 3H); 3.05 (td, *J* = 6.7 Hz, *J* = 2.1 Hz, 2H); 2.39 (pd, *J* = 6.9 Hz, *J* = 5.1 Hz, 1H); 2.09 (pd, *J* = 6.9 Hz, *J* = 4.9 Hz, 1H); 1.08 (s, 9H); 1.05 (s, 9H); 1.01 (dd, *J* = 15.6 Hz, *J* = 6.8 Hz, 6H); 0.89 (d, *J* = 6.8 Hz, 3H); 0.78 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 171.9; 156.1; 153.3; 151.7; 150.5; 147.0; 145.4; 139.9; 129.8; 123.9; 122.7; 89.8; 82.3; 77.3; 74.8; 67.6; 64.6; 61.2; 59.5; 57.2; 34.9; 34.9; 31.2; 30.3; 27.5; 27.2; 22.9; 20.5; 19.9; 19.1; 18.0; 17.8. FTIR ν_{max} (cm⁻¹): 2934 (w); 2860 (w); 1663 (m); 1559 (m); 1520 (s); 1466 (m); 1345 (m); 1137 (m); 1065 (m); 1013 (m); 828 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₉H₅₈O₁₀N₈SiNa 849.3934; Found 849.3960.



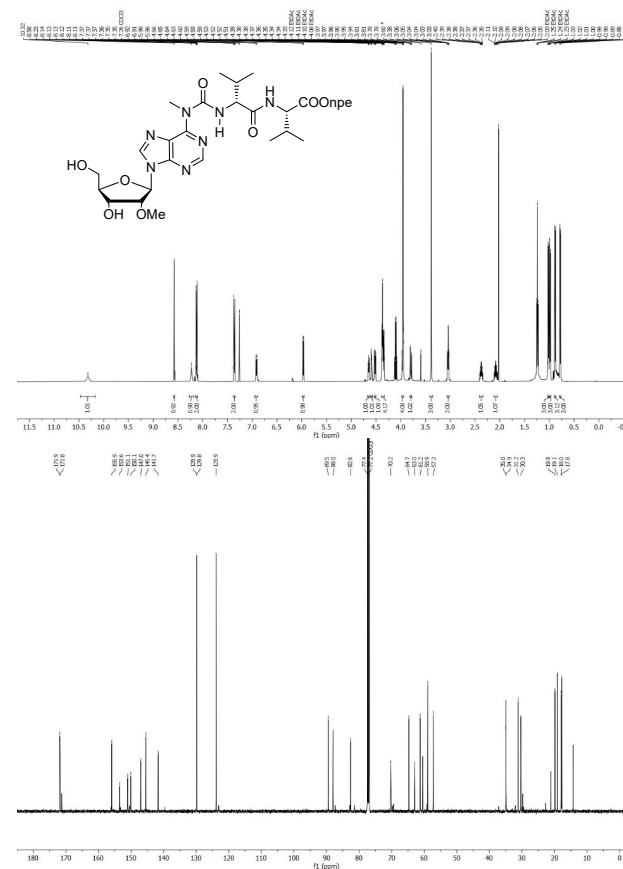
General procedure for the synthesis of compound 5: Compound **4** (1 equiv.) was added to a plastic flask and dissolved in dry 9:1 CH₂Cl₂/pyridine. The solution was stirred at 0°C. Finally, HF·pyridine (from a commercial solution containing 70% HF and 30% pyridine) was added and the reaction was stirred at 0°C for 2 h. After that, the reaction was quenched with aqueous saturated NaHCO₃ and CH₂Cl₂ was added. The organic layer was separated and the crude was further extracted with CH₂Cl₂. The combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

5_{1b}: Yield = 92%. R_f = 0.38 (95:5 CH₂Cl₂:MeOH). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.41 (s, 1H); 8.58 (s, 1H); 8.18 (s, 1H); 8.14 – 8.09 (m, 2H); 7.39 – 7.34 (m, 2H); 6.90 (d, *J* = 8.7 Hz, 1H); 5.96 (d, *J* = 6.9 Hz, 1H); 4.67 (dd, *J* = 6.9 Hz, *J* = 4.6 Hz, 1H); 4.60 (d, *J* = 4.2 Hz, 1H); 4.53 (dd, *J* = 8.8 Hz, *J* = 4.9 Hz, 1H); 4.41 – 4.33 (m, 4H); 3.98 – 3.92 (m, 4H); 3.79 (dd, *J* = 13.0 Hz, *J* = 1.7 Hz, 1H); 3.38 (s, 3H); 3.07 – 3.02 (m, 2H); 2.41 – 2.33 (m, 1H); 2.08 (pd, *J* = 6.9 Hz, *J* = 4.9 Hz, 1H); 2.02 (s, 1H); 1.00 (dd, *J* = 17.5 Hz, *J* = 6.8 Hz, 6H); 0.88 (d, *J* = 6.9 Hz, 3H); 0.77 (d, *J* = 6.8 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 171.8; 155.9; 153.7; 151.1; 150.0; 147.0; 145.4; 141.7; 129.8; 123.9; 89.5; 88.0; 82.5; 77.4; 70.3; 64.7; 63.1; 59.0; 57.2; 35.0; 34.9; 31.2; 30.4; 19.9; 19.1; 17.9; 17.8. FTIR ν_{max} (cm⁻¹): 2926 (w); 1663 (m); 1518 (s); 1464 (m); 1345 (m); 1266 (m); 1016 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₁H₄₂O₁₀N₈Na 709.2916; Found 709.2918.



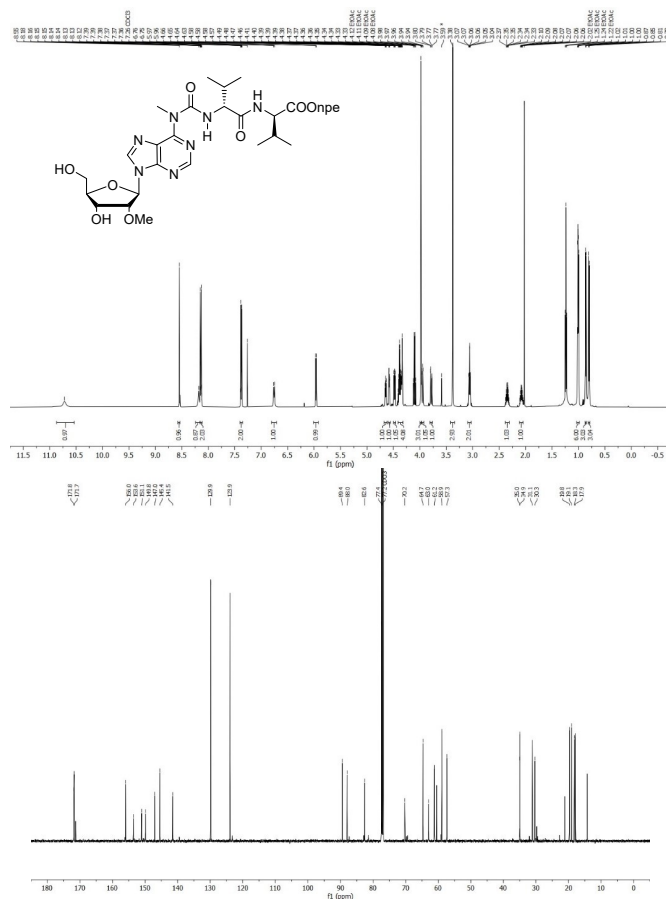
S9

5_{0b}: Yield = 92%. R_f = 0.36 (95:5 CH₂Cl₂:MeOH). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.32 (s, 1H); 8.58 (s, 1H); 8.23 (s, 1H); 8.14 – 8.11 (m, 2H); 7.38 – 7.34 (m, 2H); 6.92 (d, *J* = 8.8 Hz, 1H); 5.97 (d, *J* = 6.7 Hz, 1H); 4.65 (dd, *J* = 6.8 Hz, *J* = 4.6 Hz, 1H); 4.59 (dd, *J* = 4.7 Hz, *J* = 1.6 Hz, 1H); 4.52 (dd, *J* = 8.8 Hz, *J* = 4.9 Hz, 1H); 4.39 – 4.33 (m, 4H); 3.98 – 3.94 (m, 4H); 3.79 (dd, *J* = 13.0 Hz, *J* = 1.7 Hz, 1H); 3.38 (s, 3H); 3.04 (td, *J* = 6.7 Hz, *J* = 1.5 Hz, 2H); 2.37 (pd, *J* = 6.8 Hz, *J* = 5.2 Hz, 1H); 2.08 (pd, *J* = 6.9 Hz, *J* = 5.0 Hz, 1H); 1.02 (d, *J* = 6.8 Hz, 3H); 0.99 (d, *J* = 6.8 Hz, 3H); 0.88 (d, *J* = 6.9 Hz, 3H); 0.78 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.9; 171.8; 155.9; 153.6; 151.1; 150.1; 147.0; 145.4; 141.7; 129.9; 129.8; 123.9; 89.5; 88.0; 82.6; 77.4; 70.2; 64.7; 63.0; 61.2; 58.9; 57.2; 35.0; 34.9; 31.2; 30.3; 19.9; 19.1; 18.0; 17.8. FTIR ν_{max} (cm⁻¹): 2929 (w); 1663 (m); 1518 (s); 1464 (m); 1345 (m); 1016 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₁H₄₂O₁₀N₈Na 709.2916; Found 709.2918.



S10

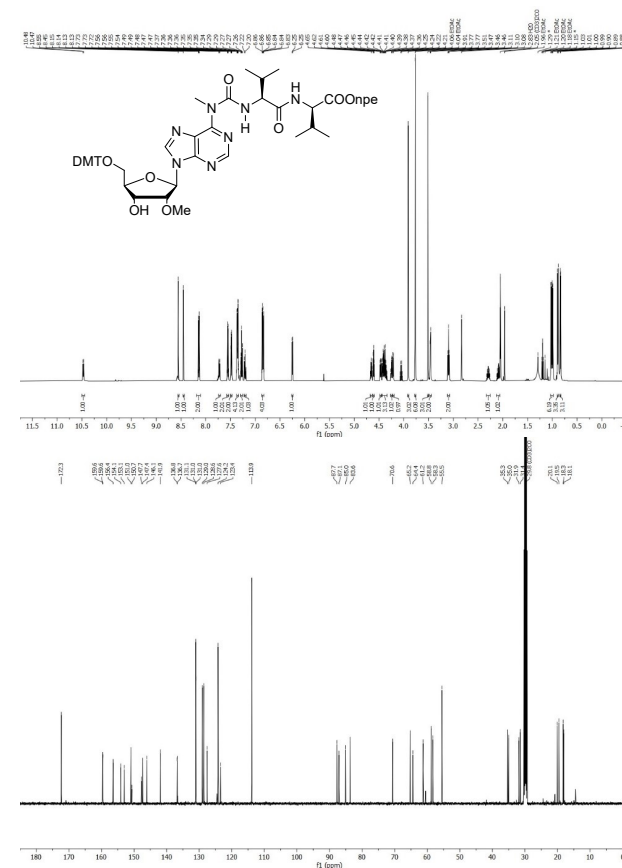
5_{0b}: Yield = 96%. R_f = 0.37 (95:5 CH₂Cl₂:MeOH). ¹H NMR (500 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 10.72 (s, 1H); 8.55 (s, 1H); 8.18 (s, 1H); 8.16 – 8.12 (m, 2H); 7.40 – 7.36 (m, 2H); 6.76 (d, *J* = 8.4 Hz, 1H); 5.96 (d, *J* = 6.8 Hz, 1H); 4.64 (dd, *J* = 6.9 Hz, *J* = 4.6 Hz, 1H); 4.58 (dd, *J* = 4.7 Hz, *J* = 1.6 Hz, 1H); 4.48 (dd, *J* = 8.5 Hz, *J* = 5.0 Hz, 1H); 4.41 – 4.32 (m, 4H); 3.98 (s, 3H); 3.95 (dd, *J* = 13.0 Hz, *J* = 1.9 Hz, 1H); 3.78 (dd, *J* = 13.0 Hz, *J* = 1.7 Hz, 1H); 3.38 (s, 3H); 3.06 (td, *J* = 6.7 Hz, *J* = 1.9 Hz, 2H); 2.39 – 2.31 (m, 1H); 2.08 (pd, *J* = 6.9 Hz, *J* = 5.0 Hz, 1H); 1.01 (dd, *J* = 6.8 Hz, *J* = 3.8 Hz, 6H); 0.86 (d, *J* = 6.8 Hz, 3H); 0.80 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, CDCl₃, 298 K): δ (ppm) = 171.8; 171.7; 156.0; 153.6; 151.1; 149.8; 147.0; 145.4; 141.5; 129.9; 123.9; 89.4; 88.0; 82.6; 77.4; 70.2; 64.7; 63.0; 61.2; 58.9; 57.3; 35.0; 34.9; 31.1; 30.3; 19.8; 19.1; 18.3; 17.9. FTIR ν_{max} (cm⁻¹): 2927 (w); 1654 (m); 1518 (s); 1464 (m); 1345 (m); 1016 (m). HRMS (ESI) *m/z*: [M+Na]⁺ Calcd for C₃₁H₄₂O₁₀N₈Na 709.2916; Found 709.2917.



S11

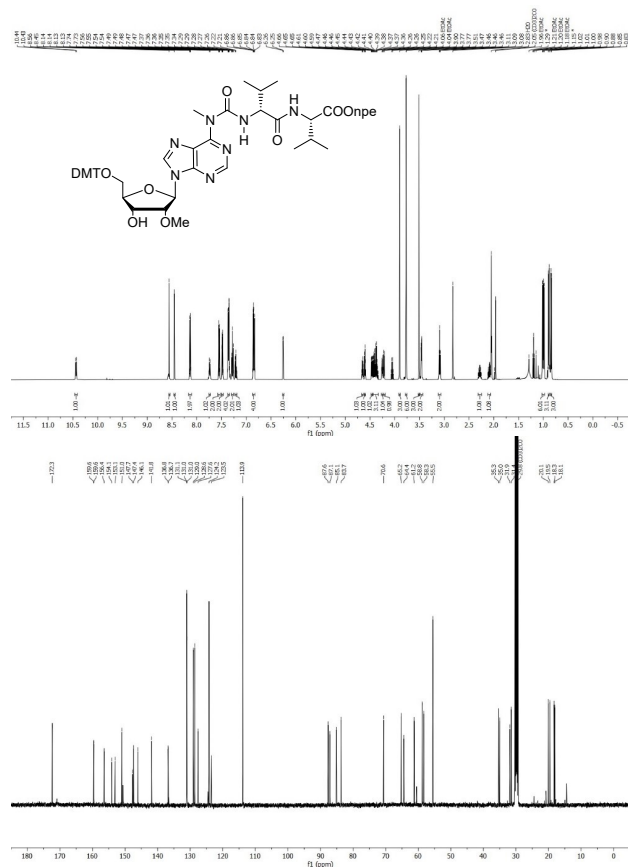
General procedure for the synthesis of compound 6: Compound **5** (1 equiv.) was dissolved in dry pyridine and stirred under nitrogen atmosphere at r.t. 4,4-Dimethoxytrityl chloride (1.5 equiv.) was added in two portions and the reaction was stirred at r.t. overnight. After that, the crude was concentrated and purified by silica gel column chromatography (eluent containing 0.1% pyridine) affording the product as a white foam.

6_{1b}: Yield = 81%. R_f = 0.53 (1:1 iHex/Acetone). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.47 (d, *J* = 7.8 Hz, 1H); 8.55 (s, 1H); 8.45 (s, 1H); 8.17 – 8.11 (m, 2H); 7.74 – 7.70 (m, 1H); 7.57 – 7.53 (m, 2H); 7.51 – 7.45 (m, 2H); 7.38 – 7.33 (m, 4H); 7.30 – 7.25 (m, 2H); 7.23 – 7.18 (m, 1H); 6.87 – 6.82 (m, 4H); 6.25 (d, *J* = 4.2 Hz, 1H); 4.66 (dt, *J* = 6.4 Hz, *J* = 5.1 Hz, 1H); 4.61 (t, *J* = 4.6 Hz, 1H); 4.47 (dd, *J* = 7.8 Hz, *J* = 4.9 Hz, 1H); 4.44 – 4.33 (m, 3H); 4.25 (q, *J* = 4.6 Hz, 1H); 4.22 (d, *J* = 6.4 Hz, 1H); 3.91 (s, 3H); 3.77 (d, *J* = 1.5 Hz, 6H); 3.51 (s, 3H); 3.48 – 3.45 (m, 2H); 3.10 (t, *J* = 6.5 Hz, 2H); 2.29 (pd, *J* = 6.9 Hz, *J* = 4.9 Hz, 1H); 2.13 – 2.06 (m, 1H); 1.01 (dd, *J* = 12.5 Hz, *J* = 6.8 Hz, 6H); 0.89 (d, *J* = 6.8 Hz, 3H); 0.84 (d, *J* = 6.8 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.3; 159.6; 159.6; 156.4; 154.1; 153.1; 151.0; 150.7; 147.7; 147.4; 146.1; 141.9; 136.8; 136.7; 131.1; 131.0; 131.0; 129.0; 128.6; 127.6; 124.2; 123.4; 113.9; 87.7; 87.1; 85.0; 83.6; 70.6; 65.2; 64.4; 61.2; 58.8; 58.3; 55.5; 35.3; 35.0; 31.9; 31.4; 20.1; 19.5; 18.3; 18.1. FTIR ν_{max} (cm⁻¹): 2960 (w); 2930 (w); 1670 (m); 1576 (m); 1507 (s); 1464 (m); 1345 (m); 1249 (m); 1177 (m); 1033 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₂H₆₁O₁₂N₈ 989.4403; Found 989.4419.

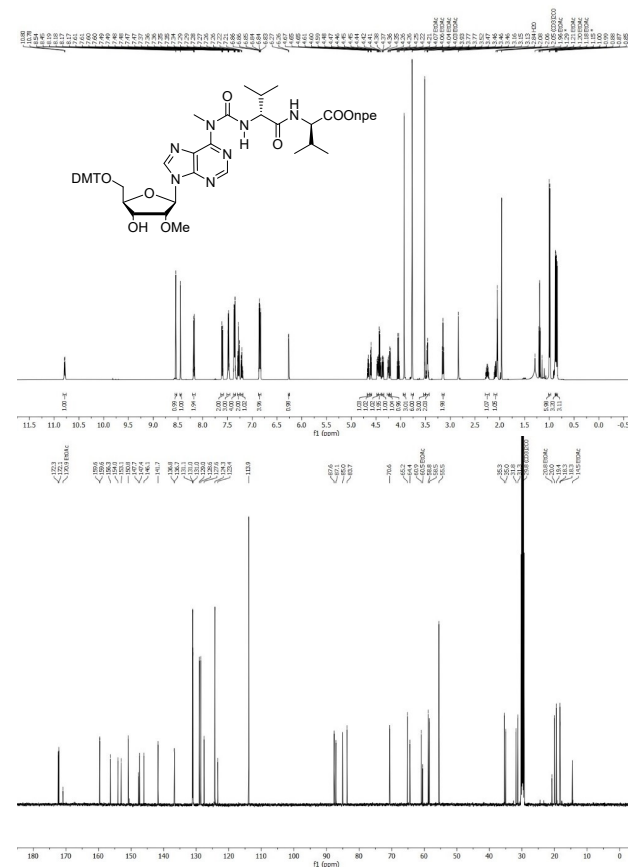


S12

6₀₁: Yield = 79%. Rf = 0.53 (1:1 iHex/Acetone). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.44 (d, *J* = 7.7 Hz, 1H); 8.56 (s, 1H); 8.45 (s, 1H); 8.16 – 8.11 (m, 2H); 7.76 – 7.72 (m, 1H); 7.57 – 7.53 (m, 2H); 7.50 – 7.46 (m, 2H); 7.36 (dq, *J* = 8.5 Hz, *J* = 3.2 Hz, 4H); 7.30 – 7.25 (m, 2H); 7.23 – 7.19 (m, 1H); 6.87 – 6.82 (m, 4H); 6.26 (d, *J* = 4.3 Hz, 1H); 4.65 (dt, *J* = 6.4 Hz, *J* = 5.1 Hz, 1H); 4.60 (t, *J* = 4.7 Hz, 1H); 4.46 (dd, *J* = 7.8 Hz, *J* = 4.9 Hz, 1H); 4.44 – 4.33 (m, 3H); 4.27 – 4.23 (m, 1H); 4.22 (d, *J* = 6.4 Hz, 1H); 3.90 (s, 3H); 3.77 (d, *J* = 1.9 Hz, 6H); 3.51 (s, 3H); 3.46 (dd, *J* = 4.4 Hz, *J* = 1.5 Hz, 2H); 3.09 (t, *J* = 6.4 Hz, 2H); 2.28 (pd, *J* = 6.9 Hz, *J* = 4.8 Hz, 1H); 2.10 (qd, *J* = 7.0 Hz, *J* = 5.9 Hz, 1H); 1.00 (dd, *J* = 11.5 Hz, *J* = 6.9 Hz, 6H); 0.89 (d, *J* = 6.9 Hz, 3H); 0.84 (d, *J* = 6.8 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.3; 159.6; 159.6; 156.4; 154.1; 153.1; 151.0; 147.7; 147.4; 146.1; 141.8; 136.8; 136.7; 131.1; 131.0; 131.0; 129.0; 128.6; 127.6; 124.2; 123.5; 113.9; 87.6; 87.1; 85.1; 83.7; 70.6; 65.2; 64.4; 61.2; 58.8; 58.3; 55.5; 35.3; 35.0; 31.9; 31.4; 20.1; 19.5; 18.3; 18.1. FTIR ν_{max} (cm⁻¹): 2960 (w); 2930 (w); 1670 (m); 1576 (m); 1507 (s); 1464 (m); 1345 (m); 1250 (m); 1177 (m); 1033 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₂H₆₁O₁₂N₈ 989.4403; Found 989.4421.

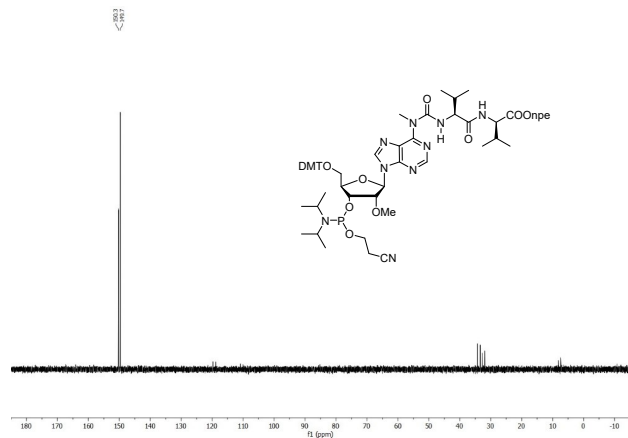


6₀₀: Yield = 75%. Rf = 0.53 (1:1 iHex/Acetone). ¹H NMR (500 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 10.79 (d, *J* = 8.0 Hz, 1H); 8.54 (s, 1H); 8.45 (s, 1H); 8.21 – 8.15 (m, 2H); 7.63 – 7.57 (m, 2H); 7.51 – 7.45 (m, 3H); 7.36 (dq, *J* = 8.6 Hz, *J* = 3.2 Hz, 4H); 7.31 – 7.24 (m, 2H); 7.24 – 7.17 (m, 1H); 6.89 – 6.81 (m, 4H); 6.26 (d, *J* = 4.2 Hz, 1H); 4.66 (dt, *J* = 6.4 Hz, *J* = 5.2 Hz, 1H); 4.60 (t, *J* = 4.6 Hz, 1H); 4.50 – 4.44 (m, 1H); 4.43 (q, *J* = 6.6 Hz, 2H); 4.36 (dd, *J* = 8.3 Hz, *J* = 5.8 Hz, 1H); 4.28 – 4.22 (m, 1H); 4.22 (d, *J* = 6.5 Hz, 1H); 3.93 (s, 3H); 3.77 (d, *J* = 1.4 Hz, 6H); 3.52 (s, 3H); 3.51 – 3.42 (m, 2H); 3.15 (t, *J* = 6.5 Hz, 2H); 2.25 (pd, *J* = 6.9 Hz, *J* = 5.2 Hz, 1H); 2.13 – 2.06 (m, 1H); 1.00 (d, *J* = 6.9 Hz, 6H); 0.87 (d, *J* = 6.9 Hz, 3H); 0.85 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, acetone-*d*₆, 298 K): δ (ppm) = 172.3; 172.1; 170.9; 159.6; 159.6; 156.3; 154.0; 153.1; 150.8; 147.7; 147.4; 146.1; 141.7; 136.8; 136.7; 131.1; 131.0; 131.0; 129.0; 128.6; 127.6; 124.3; 123.4; 113.9; 87.6; 87.1; 85.0; 83.7; 70.6; 65.2; 64.4; 60.9; 60.5; 58.8; 58.5; 55.5; 35.3; 35.0; 31.8; 31.3; 20.8; 20.0; 19.4; 18.3; 18.3; 14.5. FTIR ν_{max} (cm⁻¹): 2962 (w); 2932 (w); 1670 (m); 1576 (m); 1509 (s); 1464 (m); 1345 (m); 1250 (m); 1177 (m); 1033 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₅₂H₆₁O₁₂N₈ 989.4403; Found 989.4426.

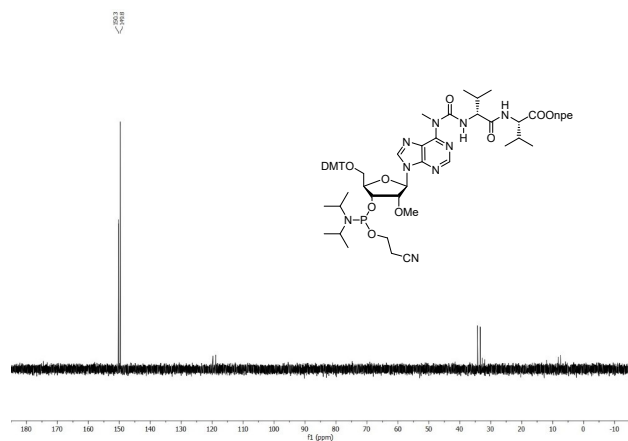


General procedure for the synthesis of compound 7: Compound **6** (1 equiv.) was added to a dry-oven round-bottom flask and dissolved in dry CH_2Cl_2 . The solution was stirred under Argon atmosphere at 0°C . DIPEA (4 equiv.) was added dropwise. Finally, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (2.5 equiv.) was added dropwise. The reaction was stirred at r.t. for 3 h. After that, the reaction was stopped and diluted with CH_2Cl_2 . The crude was washed with aqueous saturated NaHCO_3 and the organic layer was separated. The crude was further extracted with CH_2Cl_2 . The combined organic layers were dried (Na_2SO_4), filtered and concentrated. The crude was purified by silica gel column chromatography (eluent containing 0.1% pyridine). The products were isolated as a mixture of diastereoisomers as a white foam. Finally, the product was lyophilized from benzene.

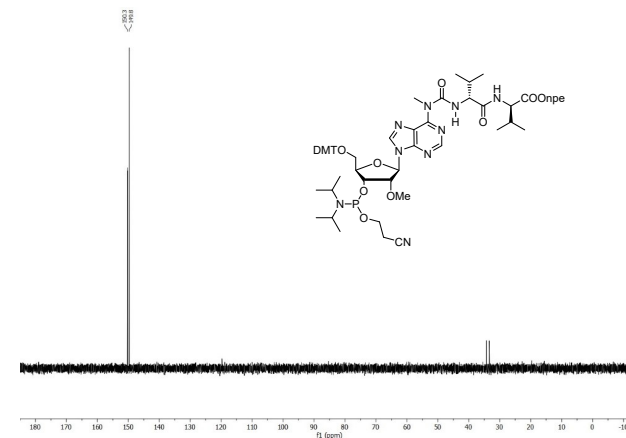
7_{LD}: Yield = 70%. Rf = 0.34; 0.25 (2:3 iHex/EtOAc). $^{31}\text{P}\{^1\text{H}\}$ NMR (202 MHz with cryoprobe, acetone- d_6 , 298 K): δ (ppm) = 150.3; 149.7. HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{61}\text{H}_{78}\text{O}_{13}\text{N}_{10}\text{P}$ 1189.5481; Found 1189.5517.



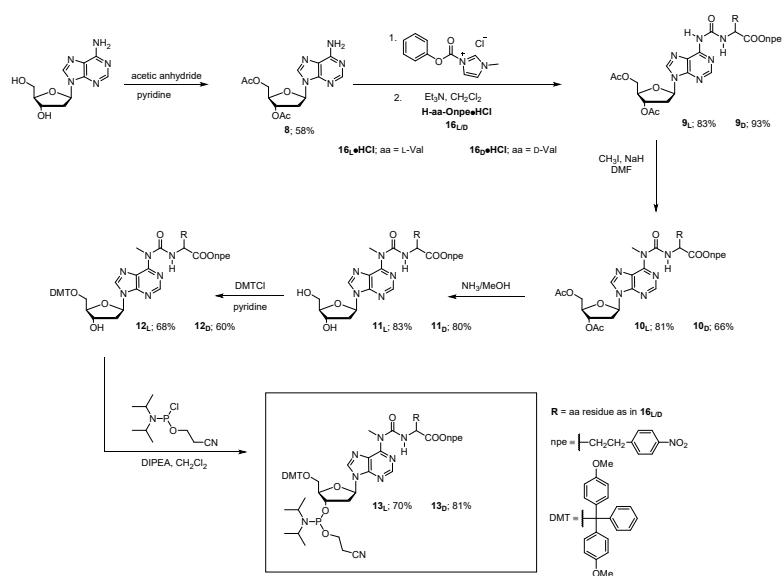
7_{DD}: Yield = 51%. Rf = 0.40; 0.30 (2:3 iHex/EtOAc). $^{31}\text{P}\{^1\text{H}\}$ NMR (202 MHz with cryoprobe, acetone- d_6 , 298 K): δ (ppm) = 150.3; 149.8. HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{61}\text{H}_{78}\text{O}_{13}\text{N}_{10}\text{P}$ 1189.5481; Found 1189.5523.



7_{DD}: Yield = 56%. Rf = 0.38; 0.28 (2:3 iHex/EtOAc). $^{31}\text{P}\{^1\text{H}\}$ NMR (202 MHz with cryoprobe, acetone- d_6 , 298 K): δ (ppm) = 150.3; 149.8. HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{61}\text{H}_{78}\text{O}_{13}\text{N}_{10}\text{P}$ 1189.5481; Found 1189.5506.



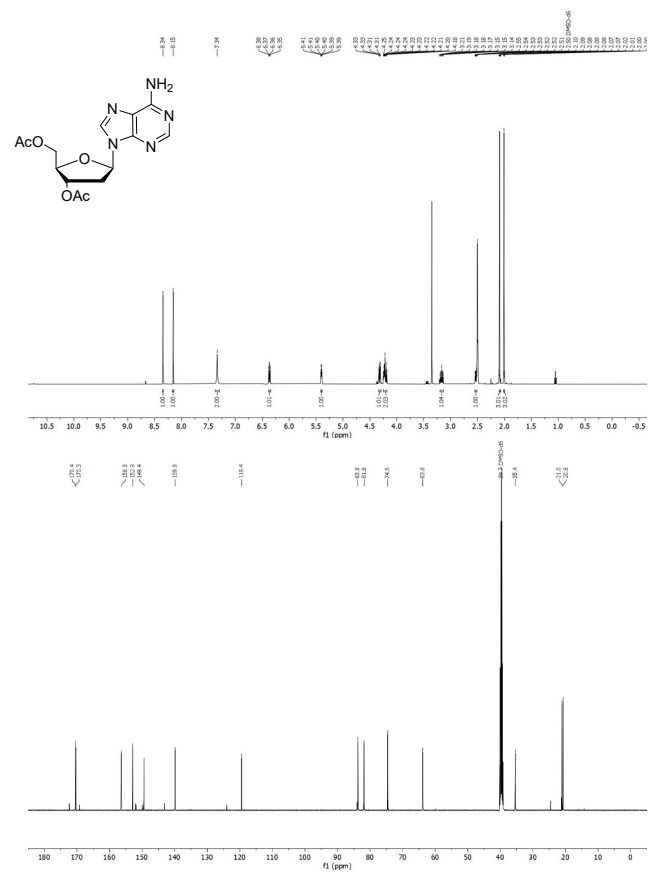
2.2 Amino acid-modified methyl *N*⁶-carbamoyl deoxy adenosine phosphoramidites



Scheme 52. Synthesis of amino acid-modified methyl *N*⁶-carbamoyl deoxy adenosine phosphoramidites.

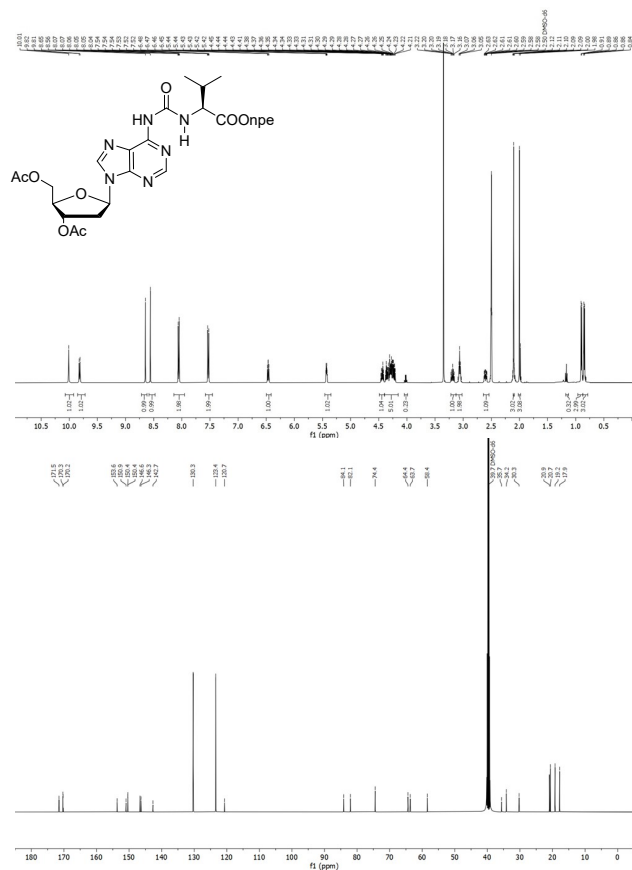
General procedure for the synthesis of compound 8: 2'-Deoxyadenosine (1 equiv.) was dissolved in dry pyridine and acetic anhydride (5 equiv.) was added dropwise. The reaction mixture was stirred for 2 h, then cooled down to 0°C and quenched with water. The volatiles were evaporated, and the residue was re-dissolved in CH₂Cl₂ and washed with aqueous saturated NaHCO₃. The aqueous phase was extracted twice with CH₂Cl₂, and the combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography, followed by recrystallization from EtOH.

8: Yield = 58%. R_f = 0.38 (9:1 EtOAc/MeOH). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 8.34 (s, 1H); 8.15 (s, 1H); 7.34 (s, 2H); 6.37 (dd, *J* = 8.1 Hz, *J* = 6.2 Hz, 1H); 5.40 (dt, *J* = 6.4 Hz, *J* = 2.5 Hz, 1H); 4.32 (dd, *J* = 10.8 Hz, *J* = 3.8 Hz, 1H); 4.26–4.18 (m, 2H); 3.17 (ddd, *J* = 14.4 Hz, *J* = 8.2 Hz, *J* = 6.5 Hz, 1H); 2.56–2.51 (m, 1H); 2.09 (s, 3H); 2.01 (s, 3H). ¹³C NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 170.4; 170.3; 156.3; 152.9; 149.4; 139.8; 119.5; 83.8; 81.8; 74.6; 63.8; 35.4; 21.0; 20.8. FTIR *v*_{max} (cm⁻¹): 3306 (w); 3155 (w); 1733 (s); 1668 (s); 1602 (m); 1573 (m); 1507 (w); 1470 (m); 1362 (m); 1260 (m); 1156 (m); 1070 (m); 1041 (s); 975 (m); 939 (m); 906 (m); 857 (m); 837 (m); 799 (w); 748 (w); 691 (m); 649 (m); 638 (s). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₁₄H₁₈O₅N₅ 336.1302; Found 336.1303.

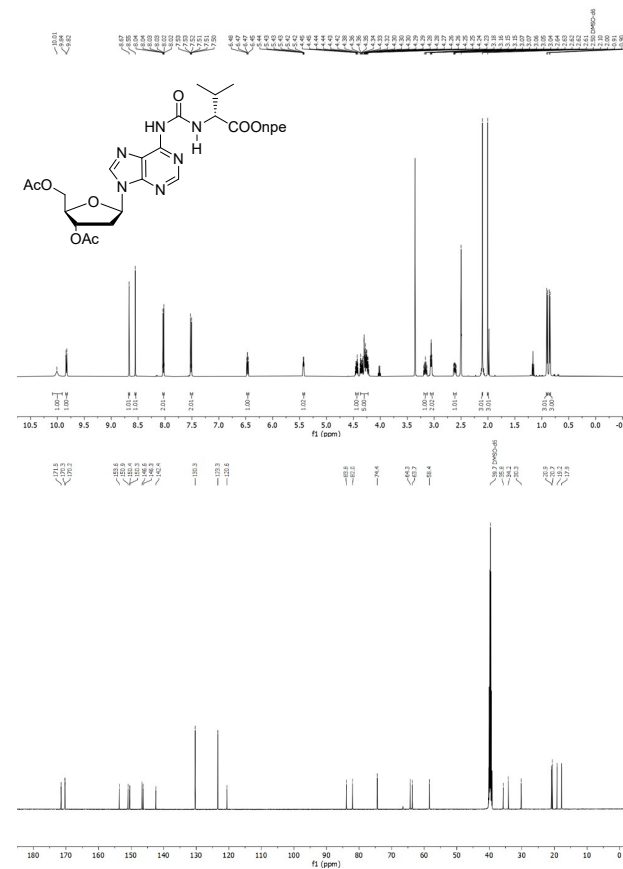


General procedure for the synthesis of compound 9: Step 1. Compound **8** (1 equiv.) and 1-*N*-methyl-3-phenoxy-carbonyl-imidazolium chloride (2 equiv.) were added to an oven dried round-bottom flask and kept under vacuum for 15 min. After that, dry CH_2Cl_2 was added under nitrogen atmosphere and the reaction was stirred at r.t. for 5 h. Step 2. Onpe-protected amino acid **16•HCl** (2 equiv.) was added to an oven dried flask and suspended in dry CH_2Cl_2 , followed by the addition of Et_3N (2 equiv.). The suspension was added dropwise to the reaction mixture. The reaction was stirred at r.t. under nitrogen atmosphere for 20 h. After that, the reaction was quenched with aqueous saturated NaHCO_3 . The organic layer was separated and the crude was further extracted with CH_2Cl_2 . The combined organic layers were dried (Na_2SO_4), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

9_c: Yield = 83%. Rf = 0.45 (10:0.5 EtOAc/MeOH). ^1H NMR (500 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 10.01 (s, 1H); 9.82 (d, J = 8.0 Hz, 1H); 8.65 (s, 1H); 8.56 (s, 1H); 8.18 – 7.92 (m, 2H); 7.64 – 7.44 (m, 2H); 6.47 (dd, J = 7.8 Hz, J = 6.3 Hz, 1H); 5.43 (dt, J = 6.5 Hz, J = 2.7 Hz, 1H); 4.44 (dt, J = 11.1 Hz, J = 6.3 Hz, 1H); 4.39 – 4.19 (m, 4H); 3.19 (ddd, J = 14.3 Hz, J = 8.0 Hz, J = 6.5 Hz, 1H); 3.06 (t, J = 6.4 Hz, 2H); 2.60 (ddd, J = 14.2 Hz, J = 6.3 Hz, J = 2.7 Hz, 1H); 2.10 (s, 3H); 2.00 (s, 3H); 0.90 (d, J = 6.8 Hz, 3H); 0.85 (d, J = 6.8 Hz, 3H). ^{13}C NMR (125 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 171.5; 170.3; 170.2; 153.6; 150.9; 150.4; 146.6; 146.3; 142.7; 130.3; 123.4; 120.7; 84.1; 82.1; 74.4; 64.4; 63.7; 58.4; 35.7; 34.2; 30.3; 20.9; 20.7; 19.2; 17.9. FTIR ν_{max} (cm^{-1}): 2964 (w); 1737 (m); 1697 (m); 1609 (w); 1585 (m); 1517 (s); 1467 (m); 1345 (s); 1222 (s); 1107 (m); 1018 (m); 941 (m); 856 (m); 797 (w); 747 (w); 696 (m); 646 (m). HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{28}\text{H}_{34}\text{O}_{10}\text{N}_2$, 628.2362; Found 628.2357.

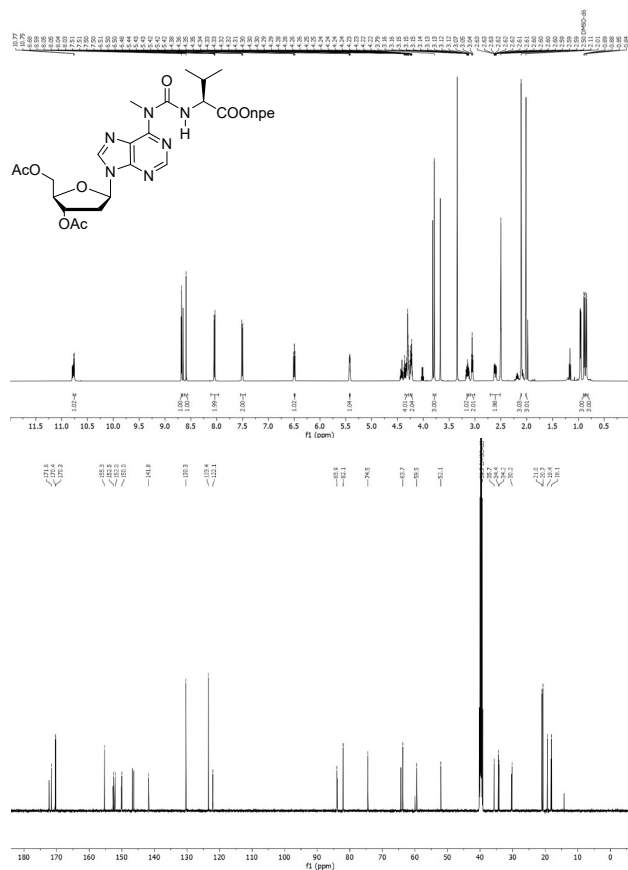


9_b: Yield = 93%. Rf = 0.45 (10:0.5 EtOAc/MeOH). ^1H NMR (500 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 10.01 (s, 1H); 9.83 (d, J = 8.1 Hz, 1H); 8.67 (s, 1H); 8.55 (s, 1H); 8.07 – 7.94 (m, 2H); 7.63 – 7.41 (m, 2H); 6.47 (dd, J = 7.9 Hz, J = 6.3 Hz, 1H); 5.43 (dt, J = 6.5 Hz, J = 2.5 Hz, 1H); 4.44 (dt, J = 11.0 Hz, J = 6.2 Hz, 1H); 4.40 – 4.20 (m, 5H); 3.16 (ddd, J = 14.3 Hz, J = 8.0 Hz, J = 6.5 Hz, 1H); 3.09 – 3.03 (m, 2H); 2.61 (ddd, J = 14.2 Hz, J = 6.3 Hz, J = 2.7 Hz, 1H); 2.10 (s, 3H); 2.00 (s, 3H); 0.90 (d, J = 6.8 Hz, 3H); 0.87 (s, 3H). ^{13}C NMR (125 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 171.5; 170.3; 170.2; 153.6; 150.9; 150.4; 150.3; 146.6; 146.3; 142.4; 130.3; 123.3; 120.6; 83.8; 82.0; 74.4; 64.3; 63.7; 58.4; 35.8; 34.2; 30.3; 20.9; 20.7; 19.2; 17.9. FTIR ν_{max} (cm^{-1}): 2964 (w); 1737 (m); 1697 (m); 1609 (w); 1585 (m); 1517 (s); 1467 (m); 1345 (s); 1222 (s); 1107 (m); 1018 (m); 941 (m); 856 (m); 797 (w); 747 (w); 696 (m); 646 (m). HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{28}\text{H}_{34}\text{O}_{10}\text{N}_2$, 628.2362; Found 628.2358.

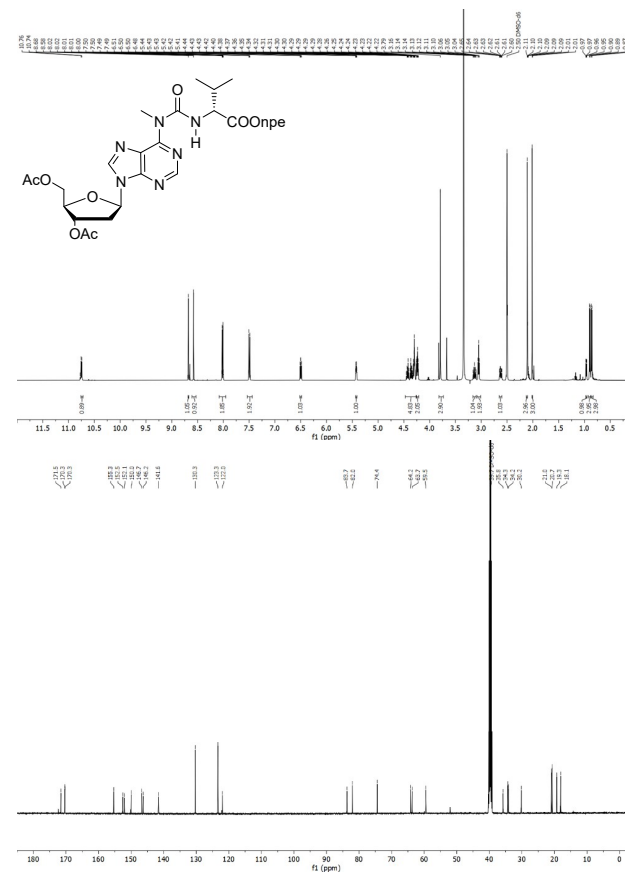


General procedure for the synthesis of compound 10: Compound 9 (1 equiv.) was dissolved in dry DMF under N₂ atmosphere. The solution was cooled to 0°C and NaH (1.05 equiv.) was added. The mixture was stirred at r.t. for 1 h. Then, MeI (1.05 equiv.) was added dropwise and the reaction was stirred for 1 h. The reaction mixture was diluted with EtOAc and washed with water. The combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography affording the product as a white foam.

10_c: Yield = 81%. Rf = 0.50 (5:1 EtOAc/iHex). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 10.76 (d, *J* = 7.5 Hz, 1H); 8.68 (s, 1H); 8.59 (s, 1H); 8.08 – 8.01 (m, 2H); 7.53 – 7.48 (m, 2H); 6.50 (dd, *J* = 7.9 Hz, *J* = 6.3 Hz, 1H); 5.42 (dq, *J* = 7.0 Hz, *J* = 2.3 Hz, 1H); 4.38 – 4.27 (m, 4H); 4.26 – 4.20 (m, 2H); 3.79 (s, 3H); 3.18 – 3.10 (m, 1H); 3.05 (t, *J* = 6.3 Hz, 2H); 2.61 (dddd, *J* = 14.2 Hz, *J* = 6.4 Hz, *J* = 2.8 Hz, *J* = 1.3 Hz, 2H); 2.11 (s, 3H); 2.01 (s, 3H); 0.89 (d, *J* = 6.8 Hz, 3H); 0.85 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 171.6; 170.4; 170.3; 155.3; 152.5; 152.0; 150.0; 141.7; 130.3; 123.4; 122.1; 83.9; 82.1; 74.5; 63.8; 59.5; 52.1; 35.7; 34.4; 34.2; 30.2; 21.0; 20.7; 19.4; 18.1. FTIR ν_{max} (cm⁻¹): 1735 (m); 1682 (m); 1572 (m); 1517 (s); 1465 (m); 1345 (s); 1218 (s); 1182 (s); 1106 (m); 1049 (m); 1018 (m); 942 (m); 856 (m); 797 (m); 747 (m); 698 (m); 646 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₉H₃₆O₁₀N₇; 642.2518; Found 642.2515.

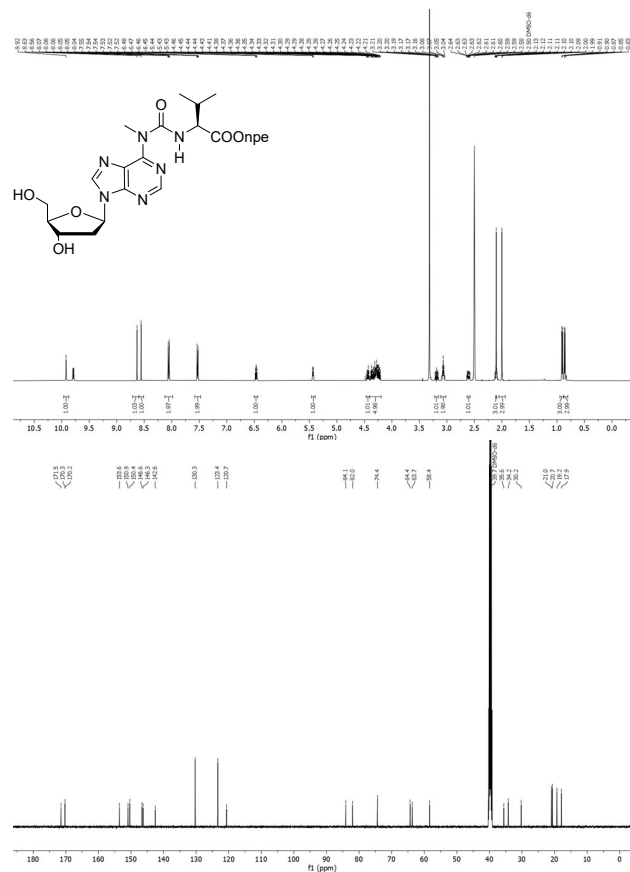


10_b: Yield = 66%. Rf = 0.53 (5:1 EtOAc/iHex). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 10.75 (d, *J* = 7.7 Hz, 1H); 8.68 (s, 1H); 8.58 (s, 1H); 8.06 – 7.98 (m, 2H); 7.53 – 7.45 (m, 2H); 6.50 (dd, *J* = 7.9 Hz, *J* = 6.2 Hz, 1H); 5.42 (dt, *J* = 6.5 Hz, *J* = 2.6 Hz, 1H); 4.46 – 4.27 (m, 5H); 4.27 – 4.20 (m, 2H); 3.79 (s, 3H); 3.13 (ddd, *J* = 14.4 Hz, *J* = 8.1 Hz, *J* = 6.5 Hz, 1H); 3.05 (t, *J* = 6.3 Hz, 2H); 2.62 (ddd, *J* = 14.2 Hz, *J* = 6.3 Hz, *J* = 2.7 Hz, 1H); 2.11 (s, 3H); 2.01 (s, 3H); 0.90 (d, *J* = 6.8 Hz, 3H); 0.86 (d, *J* = 6.8 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 171.5; 170.3; 170.3; 155.3; 152.5; 152.1; 150.0; 146.7; 146.3; 141.6; 130.3; 123.3; 122.0; 83.7; 82.0; 74.4; 64.2; 63.7; 59.5; 35.8; 34.4; 34.2; 30.2; 20.9; 20.7; 19.3; 18.1. FTIR ν_{max} (cm⁻¹): 2964 (w); 1737 (m); 1685 (m); 1517 (s); 1467 (m); 1345 (s); 1221 (s); 1184 (m); 1107 (m); 1018 (m); 941 (m); 856 (m); 797 (w); 747 (m); 696 (m); 646 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₉H₃₆O₁₀N₇; 642.2518; Found 642.2515.

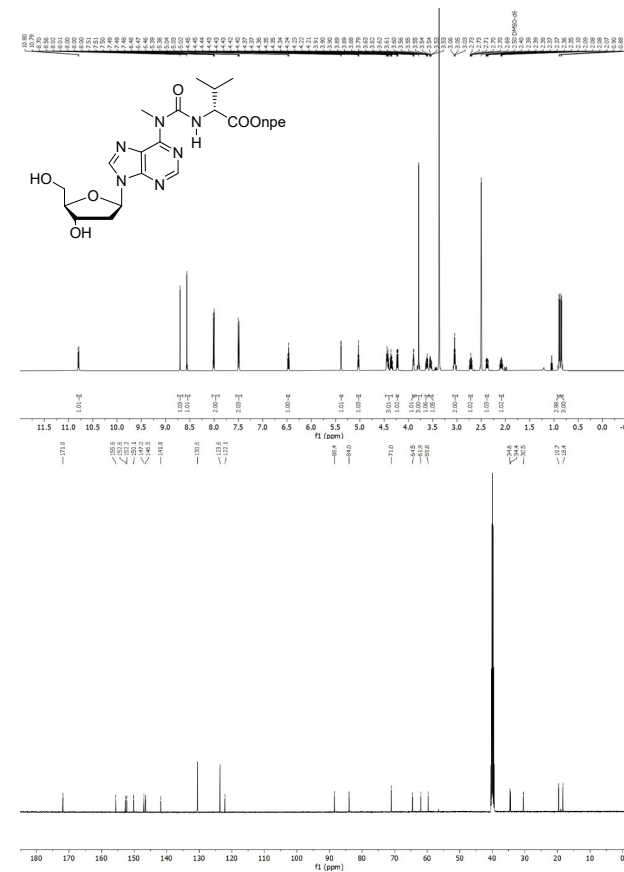


General procedure for the synthesis of compound 11: Compound **10** (1 equiv.) was diluted with 7 N NH₃/MeOH and stirred at r.t. for 2 h. Then, the reaction mixture was concentrated and the crude was purified by silica gel column chromatography.

11: Yield = 83%. Rf = 0.46 (10:0.75 CH₂Cl₂/MeOH). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 9.92 (s, 1H); 8.63 (s, 1H); 8.56 (s, 1H); 8.11 – 8.01 (m, 2H); 7.56 – 7.45 (m, 2H); 6.47 (dd, *J* = 7.8 Hz, *J* = 6.3 Hz, 1H); 5.49 – 5.39 (m, 1H); 4.44 (dt, *J* = 11.0 Hz, *J* = 6.3 Hz, 1H); 4.39 – 4.20 (m, 5H); 3.19 (ddd, *J* = 14.3 Hz, *J* = 7.8 Hz, *J* = 6.5 Hz, 1H); 3.07 (t, *J* = 6.4 Hz, 2H); 2.65 – 2.56 (m, 1H); 2.10 (s, 4H); 2.00 (s, 3H); 0.90 (d, *J* = 6.8 Hz, 3H); 0.86 (d, *J* = 6.8 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 171.5; 170.3; 170.2; 153.6; 150.9; 150.4; 146.6; 146.3; 142.6; 130.3; 123.4; 120.7; 84.1; 82.0; 74.4; 64.4; 63.7; 58.4; 35.6; 34.2; 30.2; 20.9; 20.7; 19.3; 17.9. FTIR *v*_{max} (cm⁻¹): 1735 (w); 1678 (w); 1572 (m); 1516 (s); 1465 (m); 1343 (s); 1261 (m); 1182 (m); 1050 (m); 1017 (m); 941 (m); 856 (m); 796 (m); 747 (m); 696 (m); 646 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₅H₃₂O₈N₄, 558.2307; Found 558.2321.

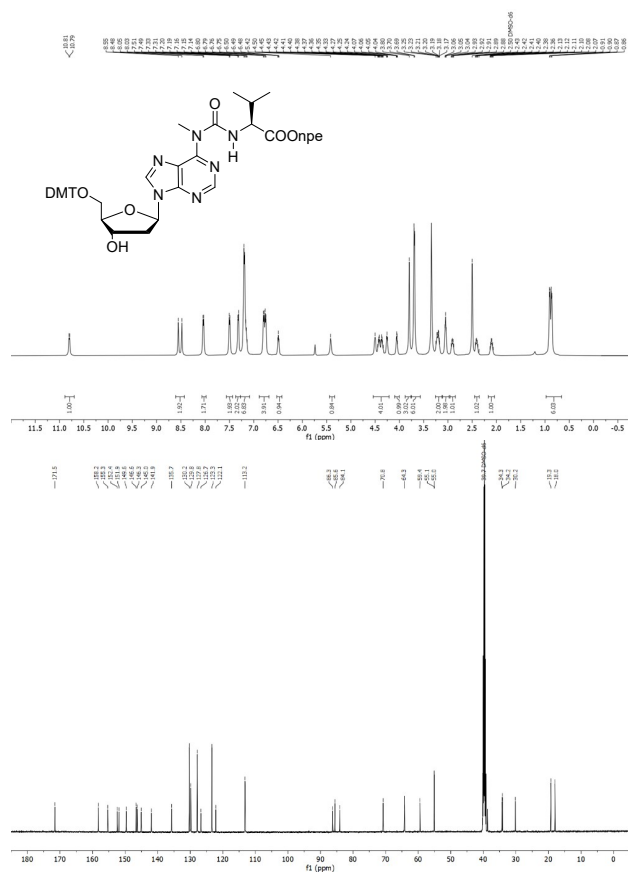


11b: Yield = 80%. Rf = 0.47 (10:0.75 CH₂Cl₂/MeOH). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 10.80 (d, *J* = 7.5 Hz, 1H); 8.70 (s, 1H); 8.56 (s, 1H); 8.04 – 7.96 (m, 2H); 7.54 – 7.45 (m, 2H); 6.47 (t, *J* = 6.7 Hz, 1H); 5.39 (d, *J* = 4.2 Hz, 1H); 5.03 (t, *J* = 5.5 Hz, 1H); 4.47 – 4.30 (m, 3H); 4.23 (dd, *J* = 7.6 Hz, *J* = 4.9 Hz, 1H); 3.90 (td, *J* = 4.4 Hz, *J* = 3.0 Hz, 1H); 3.79 (s, 3H); 3.62 (dt, *J* = 11.7 Hz, *J* = 5.0 Hz, 1H); 3.54 (ddd, *J* = 11.7 Hz, *J* = 5.6 Hz, *J* = 4.4 Hz, 1H); 3.05 (t, *J* = 6.3 Hz, 2H); 2.71 (ddd, *J* = 13.1 Hz, *J* = 7.2 Hz, *J* = 5.9 Hz, 1H); 2.38 (ddd, *J* = 13.3 Hz, *J* = 6.3 Hz, *J* = 3.6 Hz, 1H); 2.09 (pd, *J* = 6.8 Hz, *J* = 5.0 Hz, 1H); 0.89 (d, *J* = 6.9 Hz, 3H); 0.85 (d, *J* = 6.9 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 171.8; 155.4; 152.6; 150.1; 147.0; 146.5; 141.8; 130.6; 123.6; 122.1; 88.4; 84.0; 71.0; 64.5; 61.9; 59.8; 34.6; 34.4; 30.5; 19.7; 18.4. FTIR *v*_{max} (cm⁻¹): 1735 (w); 1678 (w); 1572 (m); 1516 (s); 1465 (m); 1421 (m); 1343 (s); 1263 (m); 1184 (m); 1050 (m); 1017 (m); 941 (m); 856 (m); 796 (m); 747 (m); 696 (m); 646 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₂₅H₃₂O₈N₄, 558.2307; Found 558.2321.

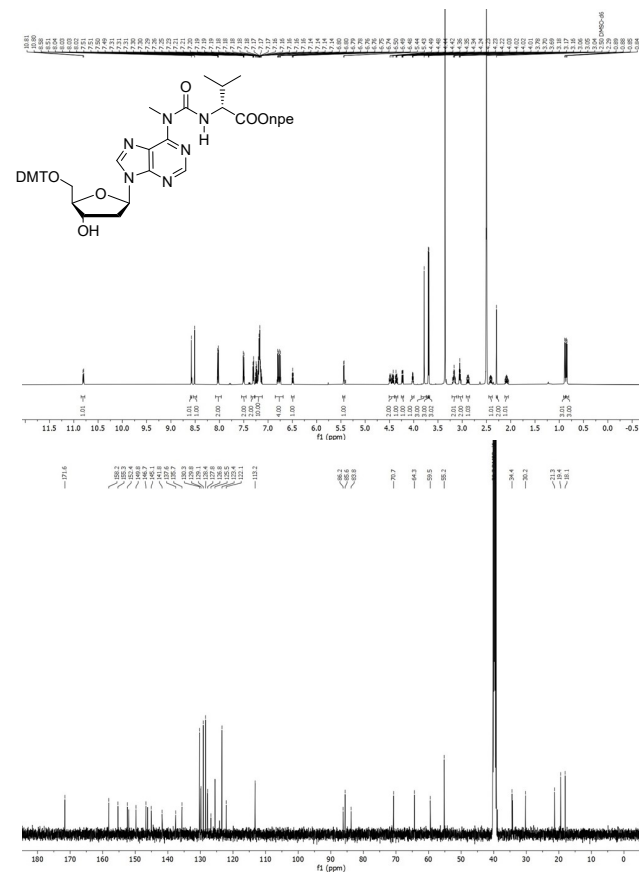


General procedure for the synthesis of compound 12: Compound 11 (1 equiv.) was dissolved in dry pyridine and stirred under nitrogen atmosphere at r.t. 4,4-Dimethoxytrityl chloride (1.5 equiv.) was added in two portions and the reaction was stirred at r.t. overnight. After that, the crude was concentrated and purified by silica gel column chromatography (eluent containing 0.1% pyridine) affording the product as a colourless foam.

12: Yield = 68%. Rf = 0.39 (10:0.5 CH₂Cl₂/MeOH). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 10.80 (d, *J* = 7.6 Hz, 1H); 8.52 (d, *J* = 36.5 Hz, 2H); 8.04 (d, *J* = 8.3 Hz, 2H); 7.50 (d, *J* = 8.2 Hz, 2H); 7.32 (d, *J* = 7.6 Hz, 2H); 7.26 – 7.12 (m, 7H); 6.78 (dd, *J* = 19.0 Hz, *J* = 8.4 Hz, 4H); 6.49 (t, *J* = 6.4 Hz, 1H); 5.42 (s, 1H); 4.56 – 4.20 (m, 4H); 4.06 (t, *J* = 4.6 Hz, 1H); 3.80 (s, 3H); 3.69 (d, *J* = 6.6 Hz, 6H); 3.26 – 3.14 (m, 2H); 3.05 (t, *J* = 6.4 Hz, 2H); 2.91 (dt, *J* = 13.0 Hz, *J* = 6.1 Hz, 1H); 2.41 (dt, *J* = 12.7 Hz, *J* = 5.9 Hz, 1H); 2.10 (dq, *J* = 13.6 Hz, *J* = 6.6 Hz, 1H); 0.90 (d, *J* = 6.8 Hz, 3H); 0.86 (d, *J* = 7.2 Hz, 3H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 171.5; 158.2; 155.3; 152.4; 151.9; 149.6; 146.6; 146.3; 145.1; 141.9; 135.7; 130.2; 129.8; 127.8; 126.7; 123.3; 122.1; 113.2; 86.3; 85.6; 84.1; 70.8; 64.3; 59.4; 55.1; 55.0; 34.3; 34.2; 30.2; 19.3; 18.0. FTIR ν_{max} (cm⁻¹): 2966 (w); 1731 (w); 1685 (w); 1606 (w); 1575 (s); 1518 (s); 1510 (s); 1464 (m); 1421 (w); 1391 (w); 1345 (s); 1299 (m); 1250 (s); 1218 (m); 1178 (s); 1151 (m); 1109 (m); 1073 (m); 1060 (m); 1033 (s); 1005 (m); 977 (m); 952 (m); 902 (w); 852 (w); 827 (m); 797 (m); 768 (m); 748 (m); 728 (m); 699 (m); 649 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₄₆H₅₀O₁₀N₄, 860.3614; Found 860.3643.

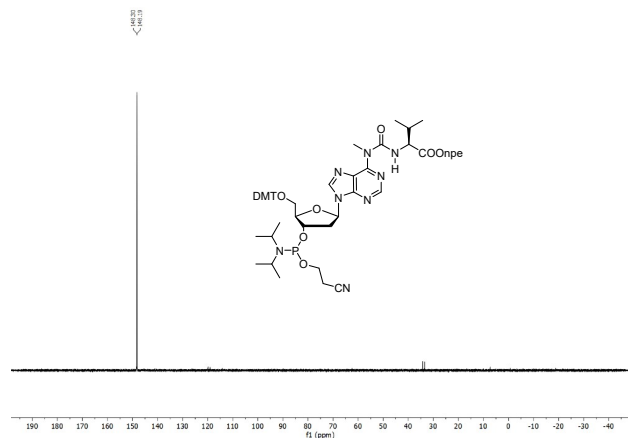


12_b: Yield = 60%. Rf = 0.46 (10:0.75 CH₂Cl₂/MeOH). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 10.80 (d, *J* = 7.5 Hz, 1H); 8.58 (s, 1H); 8.51 (s, 1H); 8.07 – 8.00 (m, 2H); 7.54 – 7.46 (m, 2H); 7.34 – 7.27 (m, 2H); 7.27 – 7.11 (m, 11H); 6.83 – 6.71 (m, 4H); 6.49 (t, *J* = 6.2 Hz, 1H); 5.44 (d, *J* = 4.6 Hz, 1H); 4.52 – 4.39 (m, 2H); 4.35 (dt, *J* = 11.0 Hz, *J* = 6.1 Hz, 1H); 4.23 (dd, *J* = 7.6 Hz, *J* = 4.9 Hz, 1H); 4.02 (dt, *J* = 6.0 Hz, *J* = 4.1 Hz, 1H); 3.78 (s, 3H); 3.69 (d, *J* = 5.9 Hz, 6H); 3.17 (h, *J* = 6.1 Hz, 2H); 3.05 (t, *J* = 6.3 Hz, 2H); 2.92 – 2.83 (m, 1H); 2.41 (ddd, *J* = 13.4 Hz, *J* = 6.9 Hz, *J* = 5.1 Hz, 1H); 2.29 (s, 2H); 2.09 (pd, *J* = 6.9 Hz, *J* = 4.9 Hz, 1H); 0.87 (dd, *J* = 17.3 Hz, *J* = 6.8 Hz, 6H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 171.6; 158.2; 155.3; 152.4; 149.8; 146.8; 145.11; 141.8; 137.6; 135.7; 130.3; 129.8; 129.1; 128.4; 127.8; 126.8; 125.6; 123.4; 122.1; 113.2; 86.2; 85.6; 83.8; 70.7; 64.3; 59.5; 55.2; 34.4; 30.2; 21.3; 19.4; 18.1. FTIR ν_{max} (cm⁻¹): 1735 (m); 1606 (m); 1465 (m); 1345 (s); 1300 (m); 1248 (s); 1217 (s); 1175 (s); 1028 (s); 1028 (m); 944 (m); 856 (m); 827 (m); 797 (m); 748 (m); 698 (m); 648 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₄₆H₅₀O₁₀N₄, 860.3614; Found 860.3637.

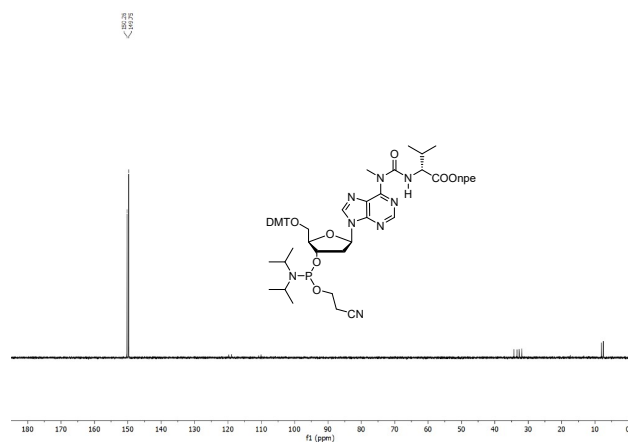


General procedure for the synthesis of compound 13: Compound **12** (1 equiv.) was added to a dry-oven flask and dissolved in dry CH_2Cl_2 . The solution was stirred under Argon atmosphere at 0°C . DIPEA (4 equiv.) was added dropwise. Finally, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (2.5 equiv.) was added dropwise. The reaction was stirred at r.t. for 3 h. After that, the reaction was stopped and diluted with CH_2Cl_2 . The crude was washed with aqueous saturated NaHCO_3 and the organic layer was separated. The crude was further extracted with CH_2Cl_2 . The combined organic layers were dried (Na_2SO_4), filtered and concentrated. The crude was purified by silica gel column chromatography (eluent containing 0.1% pyridine). The products were isolated as a mixture of diastereoisomers as a white foam. Finally, the product was lyophilized from benzene.

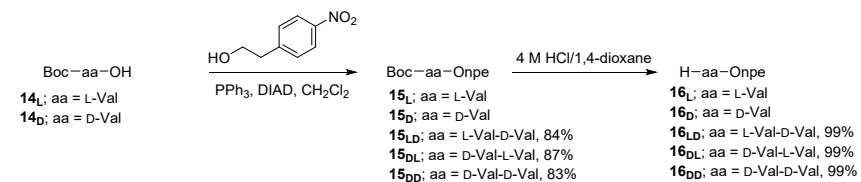
13_L: Yield = 70%. R_f = 0.70 (1:5 iHex/EtOAc). $^{31}\text{P}\{^1\text{H}\}$ NMR (202 MHz with cryoprobe, acetone- d_6 , 298 K): δ (ppm) = 148.3; 148.2. HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{55}\text{H}_{67}\text{N}_9\text{O}_{11}\text{P}$ 1060.4692; Found 1060.4742.



13_D: Yield = 81%. R_f = 0.70 (1:5 iHex/EtOAc). $^{31}\text{P}\{^1\text{H}\}$ NMR (202 MHz with cryoprobe, acetone- d_6 , 298 K): δ (ppm) = 150.3; 149.8. HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{55}\text{H}_{67}\text{N}_9\text{O}_{11}\text{P}$ 1060.4692; Found 1060.4751.



2.3 Npe-protected amino acids

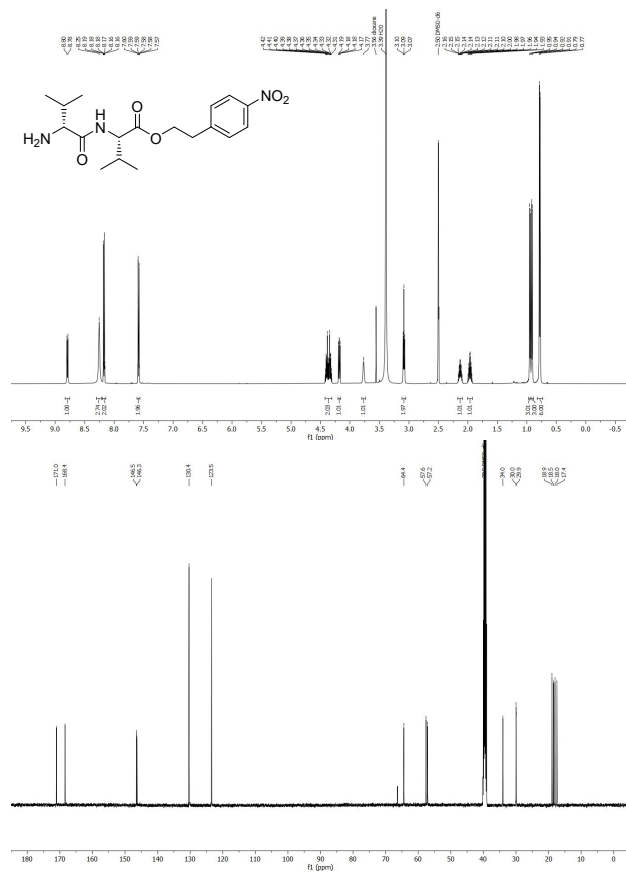


Scheme S3. Synthesis of Onpe-protected amino acids **16•HCl**.

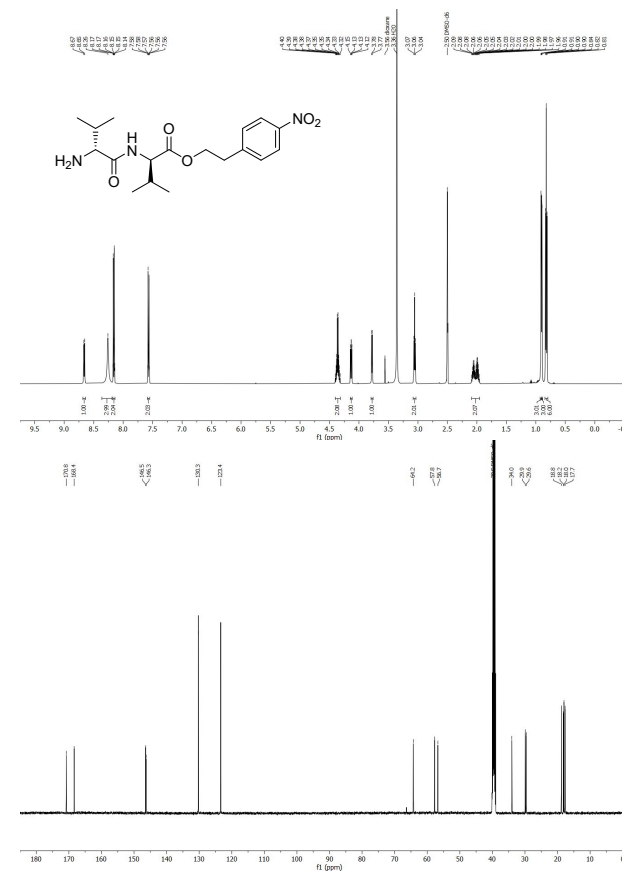
Compounds **14_L**, **14_D** are commercially available.

The Boc- and Onpe-protected amino acids **15_L** and **15_D** and their corresponding Onpe-deprotected derivatives **16_L•HCl** and **16_D•HCl** were synthesized following a procedure previously reported in the literature.^{2,3}

16_{0L}: Yield = 99%. ¹H NMR (500 MHz with cryoprobe, DMSO-d₆, 298 K): δ (ppm) = 8.79 (d, *J* = 8.3 Hz, 1H); 8.25 (s, 3H); 8.20 – 8.15 (m, 2H); 7.61 – 7.56 (m, 2H); 4.43 – 4.30 (m, 2H); 4.18 (dd, *J* = 8.2 Hz, *J* = 6.1 Hz, 1H); 3.77 (s, 1H); 3.09 (t, *J* = 6.4 Hz, 2H); 2.13 (pd, *J* = 6.9 Hz, *J* = 5.1 Hz, 1H); 1.96 (dq, *J* = 13.5 Hz, *J* = 6.8 Hz, 1H); 0.95 (d, *J* = 6.9 Hz, 3H); 0.91 (d, *J* = 6.9 Hz, 3H); 0.78 (d, *J* = 6.8 Hz, 6H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-d₆, 298 K): δ (ppm) = 171.0; 168.4; 146.5; 146.3; 130.4; 123.5; 64.4; 57.6; 57.2; 34.0; 30.0; 29.9; 18.9; 18.5; 18.0; 17.4. FTIR *v*_{max} (cm⁻¹): 3206 (w); 2966 (w); 1733 (m); 1677 (m); 1518 (s); 1464 (m); 1344 (s); 1275 (m); 856 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₁₈H₂₈O₅N₃ 366.2023; Found 366.2029.



16_{0P}: Yield = 99%. ¹H NMR (500 MHz with cryoprobe, DMSO-d₆, 298 K): δ (ppm) = 8.66 (d, *J* = 7.4 Hz, 1H); 8.26 (s, 3H); 8.18 – 8.14 (m, 2H); 7.59 – 7.55 (m, 2H); 4.41 – 4.31 (m, 2H); 4.13 (dd, *J* = 7.4 Hz, *J* = 5.7 Hz, 1H); 3.78 (d, *J* = 5.6 Hz, 1H); 3.06 (t, *J* = 6.4 Hz, 2H); 2.09 – 1.95 (m, 2H); 0.91 (d, *J* = 2.0 Hz, 3H); 0.90 (d, *J* = 2.0 Hz, 3H); 0.82 (t, *J* = 6.5 Hz, 6H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-d₆, 298 K): δ (ppm) = 170.8; 168.4; 146.5; 146.3; 130.3; 123.4; 64.2; 57.8; 56.7; 34.0; 29.9; 29.6; 18.8; 18.2; 18.0; 17.7. FTIR *v*_{max} (cm⁻¹): 3185 (w); 2966 (w); 1746 (m); 1660 (m); 1520 (s); 1447 (m); 1345 (s); 1187 (m); 1142 (m); 852 (m). HRMS (ESI) *m/z*: [M+H]⁺ Calcd for C₁₈H₂₈O₅N₃ 366.2023; Found 366.2033.



2.4 5-Methylaminomethyl uridine phosphoramidite

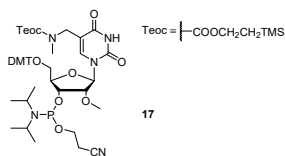
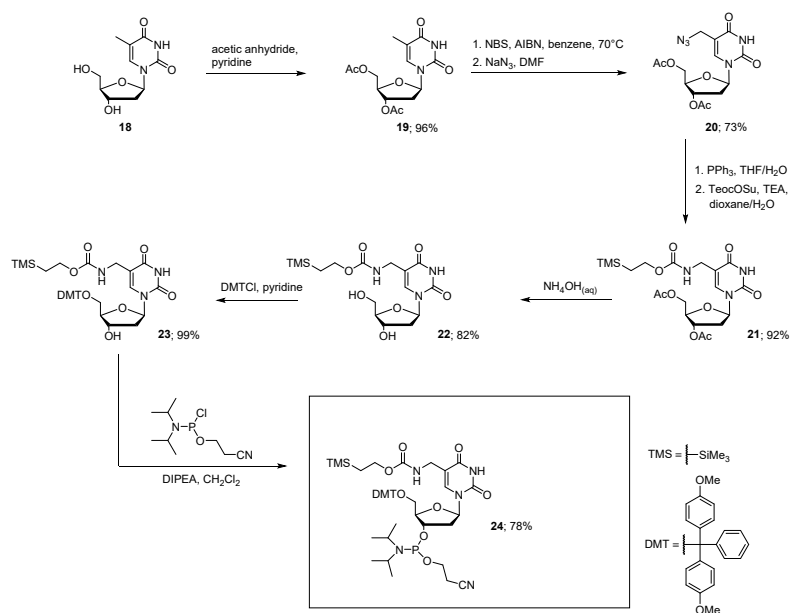


Figure S1. 5-methylaminomethyl uridine phosphoramidite 17.

Compound 17 was synthesized following a procedure previously reported in the literature.²

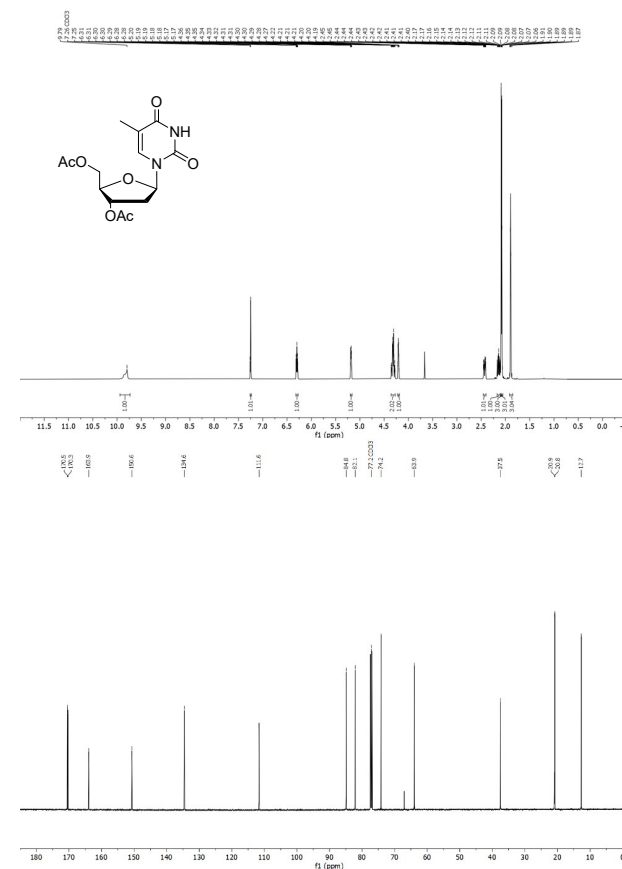
2.5 5-Aminomethyl thymidine phosphoramidite



Scheme S4. Synthesis of 5-aminomethyl thymidine phosphoramidite.

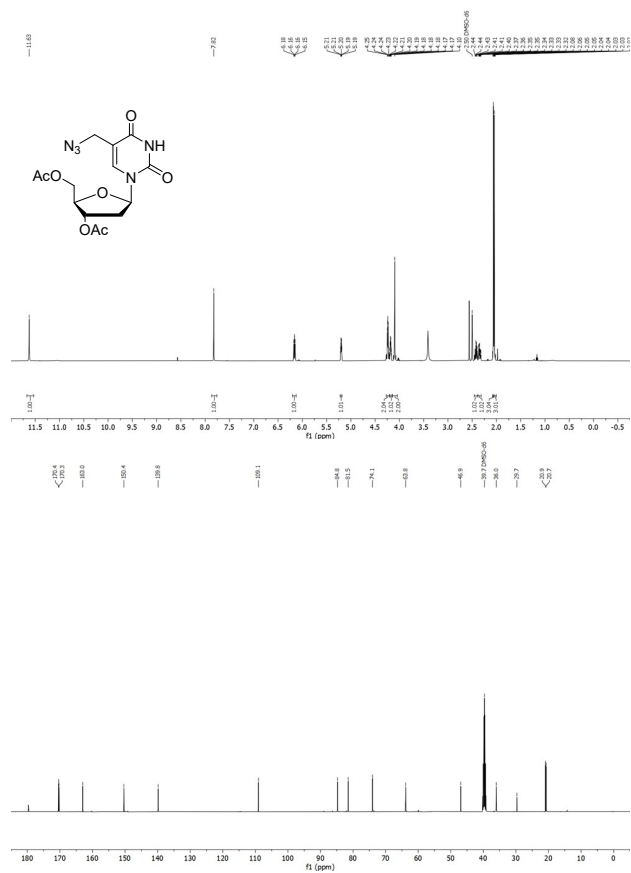
General procedure for the synthesis of compound 19: Thymidine 18 (1 equiv.) was dissolved in dry pyridine, then acetic anhydride (3 equiv.) was added slowly. The mixture was stirred at r.t. for 2 h, then quenched with CH_2Cl_2 , diluted with EtOAc and washed with water. The aqueous phase was extracted twice with EtOAc, then the combined organic layers were dried (Na_2SO_4), filtered and concentrated. The crude was purified by silica gel column chromatography.

19: Yield = 96%. Rf = 0.70 (10:0.5 $\text{CH}_2\text{Cl}_2/\text{MeOH}$). ^1H NMR (500 MHz with cryoprobe, CDCl_3 , 298 K): δ (ppm) = 9.83 (d, J = 30.5 Hz, 1H); 7.25 (d, J = 1.6 Hz, 1H); 6.30 (ddd, J = 8.1 Hz, J = 5.6 Hz, J = 2.1 Hz, 1H); 5.18 (dq, J = 6.8 Hz, J = 2.2 Hz, 1H); 4.36 – 4.27 (m, 2H); 4.21 (dt, J = 4.0 Hz, J = 2.8 Hz, 1H); 2.46 – 2.40 (m, 1H); 2.18 – 2.10 (m, 1H); 2.09 (d, J = 2.8 Hz, 3H); 2.07 (d, J = 2.9 Hz, 3H); 1.92 – 1.87 (m, 3H). ^{13}C NMR (125 MHz with cryoprobe, CDCl_3 , 298 K): δ (ppm) = 170.5; 170.3; 163.9; 150.6; 134.6; 111.6; 84.8; 82.1; 74.2; 63.9; 37.5; 20.9; 20.9; 12.7. FTIR ν_{max} (cm^{-1}): 1747 (m); 1731 (m); 1700 (m); 1662 (s); 1474 (w); 1376 (m); 1243 (s); 1120 (s); 1093 (m); 1060 (m); 1027 (s); 951 (w); 883 (w); 865 (m); 763 (w); 629 (w). HRMS (ESI) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_7$, 327.1187; Found 327.1191.



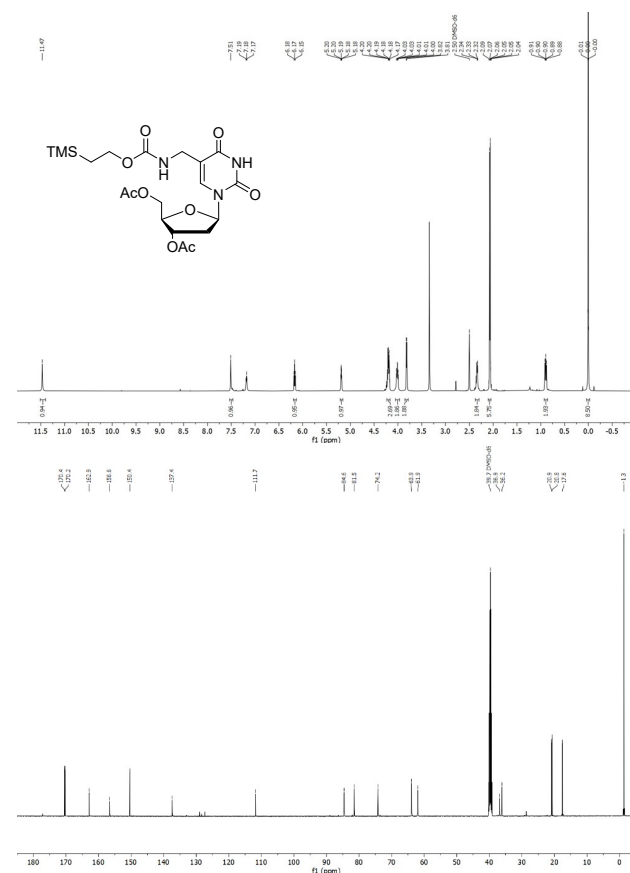
General procedure for the synthesis of compound 20: Compound **19** (1 equiv.) was suspended in dry benzene and degassed. *N*-Bromosuccinimide (1.2 equiv.) was added, the mixture was degassed again and heated to 70°C. Then, AIBN (0.5 equiv.) was added. The reaction mixture was stirred for 2 h, then another portion of AIBN (0.25 equiv.) was added. The reaction mixture was concentrated, the crude was dissolved in dry DMF under N₂ and NaN₃ (1 equiv.) was added. The reaction was stirred at r.t. for 20 h, then the reaction mixture was diluted with water. The aqueous phase was extracted with EtOAc, then the combined organic layers were dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography.

20: Yield = 73%. R_f = 0.35 (1:3 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 11.63 (s, 1H); 7.82 (s, 1H); 6.16 (dd, *J* = 8.0 Hz, *J* = 6.2 Hz, 1H); 5.20 (dt, *J* = 6.4 Hz, *J* = 2.9 Hz, 1H); 4.24 (dd, *J* = 4.8 Hz, *J* = 3.4 Hz, 2H); 4.18 (td, *J* = 4.8 Hz, *J* = 2.9 Hz, 1H); 4.10 (s, 2H); 2.43 (ddd, *J* = 14.6 Hz, *J* = 8.1 Hz, *J* = 6.8 Hz, 1H); 2.34 (ddd, *J* = 14.4 Hz, *J* = 6.3 Hz, *J* = 2.8 Hz, 1H); 2.05 (d, *J* = 9.1 Hz, 6H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 170.4; 170.3; 163.0; 150.4; 139.8; 109.1; 84.8; 81.5; 74.1; 63.9; 46.9; 36.0; 20.9; 20.7. FTIR ν_{max} (cm⁻¹): 2106 (w); 1681 (s); 1602 (w); 1464 (m); 1366 (m); 1224 (s); 1100 (m); 685 (w); 603 (w). HRMS (ESI) *m/z*: [M-H]⁻ Calcd for C₁₄H₁₆N₃O₇ 366.1055; Found 366.1056.



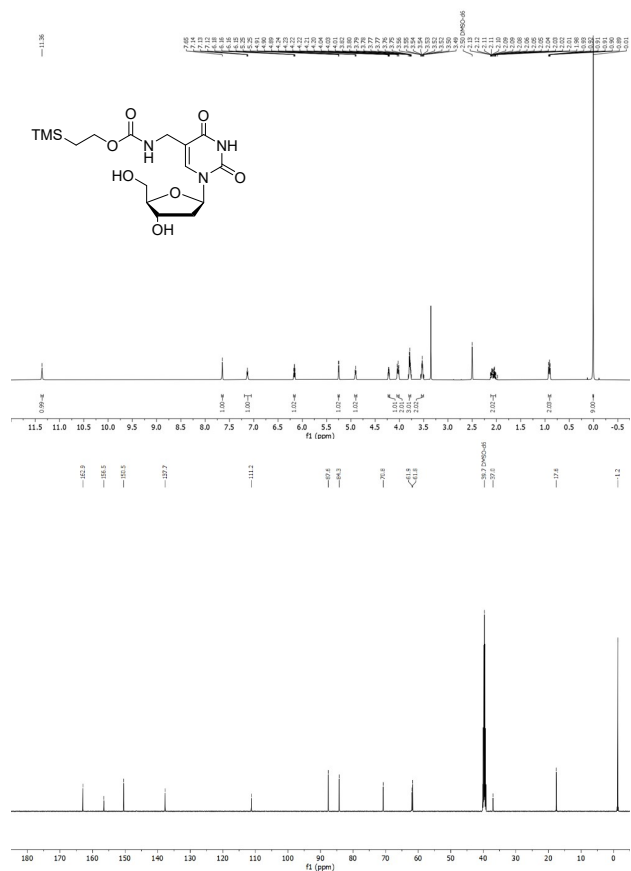
General procedure for the synthesis of compound 21: Step 1. Compound **20** (1 equiv.) was dissolved in THF/H₂O mixture (10:1) and PPh₃ (1.5 equiv.) was added. The reaction was stirred for 3 h, then the volatiles were evaporated and the crude was used for the next step without further purification. Step 2. The crude (1 equiv.) was dissolved in dioxane/water mixture (1:1) and TeocOsU (1.1 equiv.) was added, followed by TEA (1.5 equiv.). The reaction mixture was stirred at r.t. for 20 h. After that, the crude was diluted with EtOAc, washed with water, dried (Na₂SO₄), filtered and concentrated. The crude was purified by silica gel column chromatography.

21: Yield = 92%. R_f = 0.25 (3:7 iHex/EtOAc). ¹H NMR (500 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 11.47 (s, 1H); 7.51 (s, 1H); 7.28 – 7.09 (m, 1H); 6.17 (t, *J* = 7.1 Hz, 1H); 5.19 (dt, *J* = 5.9 Hz, *J* = 2.9 Hz, 1H); 4.25 – 4.15 (m, 3H); 4.02 (dt, *J* = 9.0 Hz, *J* = 3.8 Hz, 2H); 3.82 (d, *J* = 5.6 Hz, 2H); 2.37 – 2.28 (m, 2H); 2.07 (d, *J* = 5.2 Hz, 6H); 0.94 – 0.85 (m, 2H); 0.00 (s, 9H). ¹³C{¹H} NMR (125 MHz with cryoprobe, DMSO-*d*₆, 298 K): δ (ppm) = 170.4; 170.2; 162.9; 156.6; 150.4; 137.3; 111.7; 84.6; 81.5; 74.2; 63.9; 61.9; 36.9; 36.2; 20.9; 20.8; 17.6; -1.3. FTIR ν_{max} (cm⁻¹): 1672 (s); 1467 (w); 1376 (w); 1221 (s); 1135 (w); 1100 (m); 1059 (m); 1026 (m); 951 (w); 860 (m); 834 (m); 763 (w); 694 (w); 603 (w). HRMS (ESI) *m/z*: [M-H]⁻ Calcd for C₂₀H₃₀N₃O₅Si 484.1757; Found 484.1760.



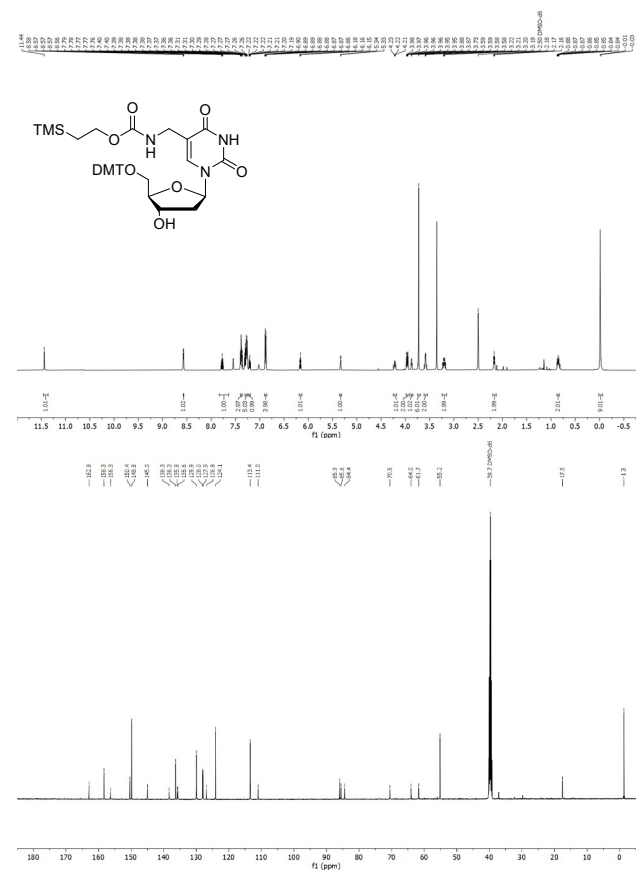
General procedure for the synthesis of compound 22: Compound **21** was dissolved in aq. NH_4OH and stirred at r.t. for 3 h. The reaction mixture was extracted with EtOAc, the organic layers were washed with water, then with aq. NH_4Cl solution, then dried (Na_2SO_4), filtered and concentrated. The crude was purified by silica gel column chromatography to afford the product as a white solid.

22: Yield = 82%. Rf = 0.38 (10:1 $\text{CH}_2\text{Cl}_2/\text{MeOH}$). ^1H NMR (500 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 11.36 (s, 1H); 7.65 (s, 1H); 7.13 (t, $J = 5.6$ Hz, 1H); 6.16 (dd, $J = 7.5$ Hz, $J = 6.1$ Hz, 1H); 5.25 (d, $J = 4.2$ Hz, 1H); 4.90 (t, $J = 5.3$ Hz, 1H); 4.22 (dq, $J = 6.5$ Hz, $J = 3.1$ Hz, 1H); 4.09 – 3.95 (m, 2H); 3.78 (dd, $J = 8.2$ Hz, $J = 4.4$ Hz, 3H); 3.57 – 3.49 (m, 2H); 2.14 – 1.96 (m, 2H); 0.91 (dd, $J = 9.4$ Hz, $J = 7.4$ Hz, 2H); 0.01 (s, 9H). $^{13}\text{C}\{^1\text{H}\}$ NMR (125 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 162.9; 156.5; 150.5; 137.7; 111.2; 87.6; 84.3; 70.8; 61.9; 61.8; 37.0; 17.6; -1.2. FTIR ν_{max} (cm^{-1}): 2953 (w); 1673 (s); 1468 (m); 1247 (s); 1092 (m); 1049 (m); 941 (m); 833 (s); 761 (m); 694 (m). HRMS (ESI) m/z : [M-H]⁻ Calcd for $\text{C}_{16}\text{H}_{26}\text{N}_3\text{O}_7\text{Si}$ 400.1546; Found 400.1548.



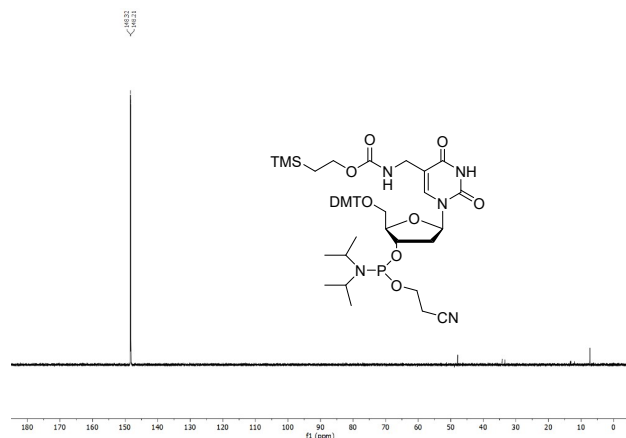
Compound **23** was synthesized following a procedure described above for compound **6** (Section 2.1).

23: Yield = 99%. Rf = 0.44 (1:2.5 iHex/EtOAc). ^1H NMR (500 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 11.44 (s, 1H); 8.58 – 8.56 (m, 2H); 7.77 (tt, $J = 7.6$ Hz, $J = 1.8$ Hz, 1H); 7.42 – 7.34 (m, 4H); 7.33 – 7.24 (m, 6H); 7.24 – 7.18 (m, 1H); 6.91 – 6.85 (m, 4H); 6.16 (t, $J = 6.6$ Hz, 1H); 5.33 (d, $J = 4.7$ Hz, 1H); 4.22 (p, $J = 5.1$ Hz, 1H); 3.99 – 3.93 (m, 2H); 3.87 (q, $J = 4.3$ Hz, 1H); 3.73 (s, 6H); 3.59 (dd, $J = 5.2$ Hz, $J = 2.5$ Hz, 2H); 3.20 (qd, $J = 10.4$ Hz, $J = 4.5$ Hz, 2H); 2.17 (t, $J = 6.0$ Hz, 2H); 0.86 (td, $J = 8.0$ Hz, $J = 3.4$ Hz, 2H); -0.01 (s, 9H). $^{13}\text{C}\{^1\text{H}\}$ NMR (125 MHz with cryoprobe, $\text{DMSO}-d_6$, 298 K): δ (ppm) = 162.9; 158.3; 156.3; 150.4; 149.8; 145.0; 138.3; 136.3; 135.8; 135.6; 129.9; 128.0; 127.9; 126.9; 124.1; 113.4; 110.9; 85.9; 85.6; 84.4; 70.5; 64.0; 61.7; 55.2; 17.6; -1.3. FTIR ν_{max} (cm^{-1}): 2951 (w); 1649 (s); 1608 (w); 1507 (m); 1464 (m); 1247 (s); 1175 (m); 1090 (m); 1031 (s); 967 (w); 827 (s); 791 (m); 754 (m); 727 (w); 701 (s). HRMS (ESI) m/z : [M-H]⁻ Calcd for $\text{C}_{37}\text{H}_{44}\text{N}_3\text{O}_9\text{Si}$ 702.2852; Found 702.2860.



Compound **24** was synthesized following a procedure describe above for compound **7** (Section 2.1).

24: Yield = 78%. Rf = 0.53 (1:1 iHex/EtOAc). $^{31}\text{P}\{^1\text{H}\}$ NMR (202 MHz with cryoprobe, acetone- d_6 , 298 K): δ (ppm) = 148.3; 148.2. HRMS (ESI) m/z : [M+Cl] Calcd for $\text{C}_{46}\text{H}_{62}\text{N}_5\text{O}_{10}\text{PSiCl}$ 938.3698; Found 938.3699.



3. General information and instruments for oligonucleotides

Synthesis and purification of oligonucleotides

Phosphoramidites of 2'-O-Me ribonucleosides (2'-OMe-Bz-A-CE, 2'-OMe-Dmf-G-CE, 2'-OMe-Ac-C-CE and 2'-OMe-U-CE) and deoxyribonucleosides (Bz-dA-CE, Dmf-dG-CE, Ac-dC-CE, T-CE) were purchased from LinkTech and Sigma-Aldrich. Oligonucleotides (ONs) were synthesized on a 1 μmol scale using RNA SynBase™ CPG 1000/110 and High Load Glen UnySupport™ as solid supports using an RNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with a standard phosphoramidite chemistry. ONs were synthesized in DMT-OFF mode using DCA as a deblocking agent in CH_2Cl_2 , BTT or Activator 42[®] as activator in MeCN, Ac_2O as capping reagent in pyridine/THF and I_2 as oxidizer in pyridine/ H_2O .

Deprotection of Onpe and teoc groups

For the deprotection of the *para*-nitrophenylethyl (Onpe) group in ONs containing amino acid-modified carbamoyl adenosine nucleosides, the solid support beads were suspended in a 9:1 THF/DBU solution mixture (1 mL) and incubated at r.t. for 2 h.⁴ After that, the supernatant was removed, and the beads were washed with THF (3 \times 1 mL).

For the deprotection of the 2-(trimethylsilyl)ethoxycarbonyl (teoc) group in ONs containing 5-methylaminomethyl uridine nucleosides, the solid support beads were suspended in a saturated solution of ZnBr_2 in 1:1 MeNO₂/IPA (1 mL) and incubated at r.t. overnight.⁵ After that, the supernatant was removed and the beads were washed with 0.1 M EDTA in water (1 mL) and water (1 mL).

Coupling of amino acids to ONs anchored to the solid support beads

The solid support beads (1 μmol) in an Eppendorf tube were washed with dry DMF (0.3 mL). In a separate Eppendorf tube, Boc-protected L- or D-Valine, DMTMM $\cdot\text{BF}_4$ (100 μmol) as activator and dry DIPEA (200 μmol) were dissolved in dry DMF (0.6 mL). Subsequently, the amino acid solution was added to the solid support beads and the reaction was incubated in an orbital shaker at r.t. for 1 h. The suspension was centrifuged and the supernatant was removed. The solid support beads were washed with dry DMF (2 \times 0.3 mL) and dry MeCN (2 \times 0.3 mL). Finally, the beads were dried using a SpeedVac concentrator.

For the deprotection of the *tert*-butyloxycarbonyl (Boc) group in ONs after the coupling of a Boc-protected amino acid or peptide, the solid support beads were suspended in a 1:1 TFA/ CH_2Cl_2 solution mixture (0.5 mL) and incubated for 5 min at r.t.⁶ After that, the supernatant was removed and the solid support beads were washed with CH_2Cl_2 (2 \times 0.5 mL).

Cleavage from beads and precipitation of the synthesized ON

The solid support beads were suspended in a 1:1 aqueous solution mixture (0.6 mL) of 30% NH_4OH and 40% MeNH₂. The suspension was heated at 65°C (8 min for SynBase™ CPG 1000/110 and 60 min for High Load Glen UnySupport™). The ONs containing dipeptide-modified carbamoyl adenosine derivatives were cleaved from the solid support beads using a 30% NH_4OH aqueous solution (0.6 mL) at r.t. overnight. Subsequently, the supernatant was collected, and the beads were washed with water (2 \times 0.3 mL). The combined aqueous solutions were concentrated under reduced pressure using a SpeedVac concentrator. After that, the crude was dissolved in DMSO (100 μL) and the ON was precipitated by adding 3 M NaOAc in water (25 μL) and *n*-butanol (1 mL). The mixture was kept at -80°C for 2 h and centrifuged at 4°C for 1 h. The supernatant was removed, and the white precipitate was lyophilized.

Purification of the synthesized ON by HPLC and desalting

The crude was purified by semi-preparative HPLC (1260 Infinity II Manual Preparative LC System from Agilent equipped with a G7114A detector) using a reverse-phase (RP) VP 250/10 Nucleodur 100-5 C18ec column from Macherey-Nagel (buffer A: 0.1 M AcOH/ Et_3N pH 7 in H_2O and buffer B: 0.1 M AcOH/ Et_3N pH 7 in 20:80 H_2O /MeCN; Gradient: 0-25% of B in 45 min; Flow rate = 5 mL $\cdot\text{min}^{-1}$). The purified ON was analyzed by RP-HPLC (1260 Infinity II LC System from Agilent equipped with a G7165A detector) using an EC 250/4 Nucleodur 100-3 C18ec from Macherey-Nagel (Gradient: 0-30% of B in 45 min; Flow rate = 1 mL $\cdot\text{min}^{-1}$). Finally, the purified ON was desalted using a C18 RP-cartridge from Waters.

Determination of the concentration and the mass of the synthesized ON

The absorbance of the synthesized ON in H_2O solution was measured using an IMPLEN NanoPhotometer[®] N60/N50 at 260 nm. The extinction coefficient of the ON was calculated using the OligoAnalyzer Version 3.0 from Integrated DNA Technologies. For ONs incorporating non-canonical bases, the extinction coefficients were assumed to be identical to those containing only canonical counterparts.

The synthesized ON (2-3 μL) was desalted on a 0.025 μm VSWP filter (Millipore), co-crystallized in a 3-hydroxyisobutyric acid matrix (HPA, 1 μL) and analyzed by matrix-assisted laser desorption/ionization – time-of-flight (MALDI-ToF) mass spectrometry (negative mode).

4. Synthesized oligonucleotides using a DNA/RNA automated synthesizer

4.1 Donor strands containing amino acid-modified methyl *N*⁶-carbamoyl adenosine at the 5'-end

Table S1. HPLC retention times (0-40% of B in 45 min) and MALDI-ToF mass spectrometric analysis (negative mode) of donor strands **ON1_{R1}** containing a Val- or Val-Val-modified *N*⁶-carbamoyl adenosine at the 5'-end. Subscript _L stands for the L-Val enantiomer and subscript _D stands for the D-Val enantiomer.

Sequence	Donor strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
5'-R ¹ (AUCGCU) _m -3'	ON1_L ; R ¹ = m ⁶ V _L ⁶ A _m	27.8	2417.5	2417.2
	ON1_D ; R ¹ = m ⁶ V _D ⁶ A _m	28.9	2417.5	2417.0
	ON1_{LD} ; R ¹ = m ⁶ V _L V _D ⁶ A _m	30.1	2516.5	2516.5
	ON1_{DL} ; R ¹ = m ⁶ V _D V _L ⁶ A _m	31.0	2516.5	2516.3
	ON1_{DD} ; R ¹ = m ⁶ V _D V _D ⁶ A _m	33.9	2516.5	2516.4

4.2 Acceptor strands containing 5-methylaminomethyl uridine at the 3'-end

Table S2. HPLC retention times (0-40% of B in 45 min) and MALDI-ToF mass spectrometric analysis (negative mode) of acceptor strands **ON2_{R2}** containing a methylaminomethyl uridine at the 3'-end.

Sequence	Acceptor strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
5'-(GUACAGCGAU) _m R ² -3'	ON2_L ; R ² = v _L -mnm ⁵ U _m	24.8	3785.4	3786.0
	ON2_D ; R ² = v _D -mnm ⁵ U _m	24.7	3785.4	3788.7

4.3 Donor strands containing amino acid-modified methyl *N*⁶-carbamoyl deoxy adenosine at the 5'-end

Table S3. HPLC retention times (0-40% of B in 45 min) and MALDI-ToF mass spectrometric analysis (negative mode) of donor strands **dON1_{R1}** containing an L- or D-Val-modified *N*⁶-carbamoyl deoxy adenosine at the 5'-end.

Sequence	Donor strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
5'-R ¹ d(ATCGCT)-3'	dON1_L ; R ¹ = m ⁶ V _L ⁶ dA	24.5	2236.6	2236.3
	dON1_D ; R ¹ = m ⁶ V _D ⁶ dA	23.2	2236.6	2236.2

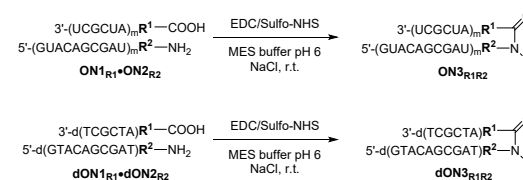
4.4 Acceptor strands containing 5-aminomethyl thymidine at the 3'-end

Table S4: HPLC retention times (0-40% of B in 45 min) and MALDI-ToF mass spectrometric analysis (negative mode) of acceptor strands **dON2_{R2}** containing an aminomethyl thymidine at the 3'-end.

Sequence	Acceptor strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
5'-d(GTACAGCGAT)R ² -3'	dON2_L ; R ² = v _L -nm ⁵ dT	19.3	3470.4	3469.1
	dON2_D ; R ² = v _D -nm ⁵ dT	19.2	3470.4	3469.8

5. Peptide coupling reactions between donor and acceptor oligonucleotides

Stock solutions of MES buffer pH 6 (400 mM), NaCl (1 M) and activator (500 mM) were prepared in water. Subsequently, equimolar amounts of (**dON1_{R1}**) and (**dON2_{R2}**) (3-5 nmol) were annealed at 95°C for 4 min in water containing NaCl (half of the volume required for the reaction). Finally, buffer, NaCl, activator solutions and water were added to the ON solution and the reaction was incubated in a ThermoMixer at 25°C for 2 h. The concentration of the components in the reaction mixture was: 50 μM of (**dON1_{R1}**), 50 μM of (**dON2_{R2}**), 100 mM of buffer, 100 mM of NaCl and 50 mM of EDC/Sulfo-NHS.



Scheme S5. Peptide coupling reactions between oligonucleotides (**dON1_{R1}**) and (**dON2_{R2}**) to give hairpin (**dON3_{R1R2}**).

The peptide coupling reactions were only performed to isolate and characterize the products by HPLC (t_R) and mass spectrometry (m/z). The isolated products were used as references for the analyses of the pair-wise competitive reactions. In this respect, the reaction yields are not given.

Table S5. HPLC retention times (0-40% of B in 45 min) and MALDI-ToF mass spectrometric analysis (negative mode) of isolated hairpin products (**dON3_{R1R2}**).

Donor strand	Acceptor strand	Hairpin strand	t _R (min)	m/z calcd. for [M-H] ⁻	found
ON1_L ; R ¹ = m ⁶ V _L ⁶ A _m	ON2_L ; R ² = v _L -mnm ⁵ U _m	ON3_{LL}	32.2	6186.9	6186.1
ON1_D ; R ¹ = m ⁶ V _D ⁶ A _m	ON2_L ; R ² = v _L -mnm ⁵ U _m	ON3_{DL}	33.6	6186.9	6187.3
ON1_{LD} ; R ¹ = m ⁶ V _L V _D ⁶ A _m	ON2_L ; R ² = v _L -mnm ⁵ U _m	ON3_{LLD}	34.0	6285.9	6286.7
ON1_{LD} ; R ¹ = m ⁶ V _L V _L ⁶ A _m	ON2_D ; R ² = v _D -mnm ⁵ U _m	ON3_{LLD}	36.1	6285.9	6285.6
ON1_{LD} ; R ¹ = m ⁶ V _L V _D ⁶ A _m	ON2_L ; R ² = v _L -mnm ⁵ U _m	ON3_{LDL}	35.5	6285.9	6285.9
ON1_{LD} ; R ¹ = m ⁶ V _L V _D ⁶ A _m	ON2_D ; R ² = v _D -mnm ⁵ U _m	ON3_{LDD}	33.5	6285.9	6286.2
ON1_{DL} ; R ¹ = m ⁶ V _D V _L ⁶ A _m	ON2_L ; R ² = v _L -mnm ⁵ U _m	ON3_{DLL}	34.6	6285.9	6286.9
ON1_{DL} ; R ¹ = m ⁶ V _D V _D ⁶ A _m	ON2_D ; R ² = v _D -mnm ⁵ U _m	ON3_{DL}	35.5	6285.9	6285.7
ON1_{DD} ; R ¹ = m ⁶ V _D V _D ⁶ A _m	ON2_L ; R ² = v _L -mnm ⁵ U _m	ON3_{DDL}	34.5	6285.9	6287.0
ON1_{DD} ; R ¹ = m ⁶ V _D V _D ⁶ A _m	ON2_D ; R ² = v _D -mnm ⁵ U _m	ON3_{DDD}	35.3	6285.9	6286.4
dON1_L ; R ¹ = m ⁶ V _L ⁶ dA	dON2_L ; R ² = v _L -nm ⁵ dT	dON3_{LL}	22.1	5689.0	5689.7
dON1_D ; R ¹ = m ⁶ V _D ⁶ dA	dON2_L ; R ² = v _L -nm ⁵ dT	dON3_{DL}	25.1	5689.0	5689.2
dON1_{LD} ; R ¹ = m ⁶ V _L ⁶ dA	dON2_D ; R ² = v _D -nm ⁵ dT	dON3_{LDD}	24.3	5689.0	5689.0
dON1_{LD} ; R ¹ = m ⁶ V _L ⁶ dA	dON2_L ; R ² = v _L -nm ⁵ dT	dON3_{LDL}	23.6	5689.0	5688.8
ON1_{cal} ; R ¹ = m ⁶ G ⁶ A _m	ON2_{cal} ; R ² = mnm ⁵ U _m	ON3_{cal}	26.8	6045.7	6044.8

6. Calibration curve hairpin

Hairpin oligonucleotide **ON3_{cal}** or acceptor oligonucleotide **ON2_L** was used for the development of a HPLC calibration curve. A stock solution of **ON3_{cal}** or **ON2_L** was prepared in water (100 μM). Separate standard solutions containing 1.2; 1.0; 0.8; 0.6; 0.4; 0.2 and 0.1 nmol of **ON3_{cal}** or **ON2_L** were prepared in a final volume of 20 μL. The standard solutions were injected in an analytical HPLC equipped with a C18 column (buffer A: 0.1 M AcOH/Et₃N pH 7 in H₂O and buffer B: 0.1 M AcOH/Et₃N pH 7 in 20:80 H₂O/MeCN; Gradient: 0-40% of B in 45 min; Flow rate = 1 mL·min⁻¹). The absorbance was monitored at 260 nm and the areas of the chromatographic peaks were determined by integration of the HPLC-chromatograms. The plot of the chromatographic area (a.u.) versus the amount (nmol) of the oligonucleotide followed a linear relationship.

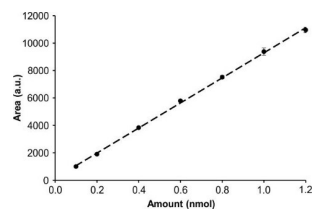


Figure S2. Chromatographic area (a.u.) vs. amount (nmol) of **ON3_{cal}**. Line shows the fit of the data to a linear regression equation. Error bars are standard deviations from two independent experiments.

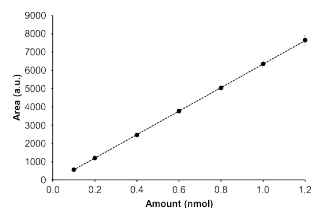


Figure S3. Chromatographic area (a.u.) vs. amount (nmol) of **ON2_L**. Line shows the fit of the data to a linear regression equation. Error bars are standard deviations from two independent experiments.

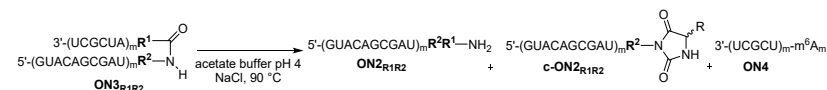
Table S6. Calibration curve ($y = mx + n$) obtained by HPLC analysis of **ON3_{cal}** and **ON2_L** and calculated extinction coefficients of **ON1_{cal}**, **ON2_{cal}**, **ON3_{cal}** and **ON2_L**.

Strand	Slope, m (nmol ⁻¹)	Intercept, n	r ²	ϵ (M ⁻¹ ·cm ⁻¹)
ON1_{cal} ^a	-	-	-	68800
ON2_{cal} ^a	-	-	-	113400
ON3_{cal} ^b	9142.5	149.4	0.99	159182
ON2_L	6428.6	-89.8	1.0	113400

The extinction coefficients were calculated using: ^a the OligoAnalyzer Version 3.0 from Integrated DNA Technologies and ^b an hypochromicity value of $h = 0.827$ for the section of the hairpin forming base pairs as reported in the literature.⁷

7. Cleavage reaction of hairpin products

A stock solution of acetate buffer pH 4 (400 mM) and NaCl (1 M) were prepared in water. Subsequently, **ON3_{R1R2}** was mixed with buffer, NaCl, and water and the reaction mixture was incubated at 90°C for 20 h. The concentration of the components in the reaction mixture was: 50 μM of **ON3_{R1R2}**, 100 mM of buffer, 100 mM of NaCl.



Scheme S6: Cleavage reactions of hairpin products **ON3_{R1R2}** to give **ON2_{R1R2}**, **c-ON2_{R1R2}** and **ON4**.

The hairpin cleavage reactions were only performed to isolate and characterize the products by HPLC (t_R) and mass spectrometry (m/z). The isolated products were used as references for the analyses of the pair-wise competitive reactions. In this respect, the reaction yields are not given.

Table S7. HPLC retention times (0-40% of B in 45 min) and MALDI-ToF mass spectrometric analysis (negative mode) of isolated cleavage products **ON2_{R1R2}**, **c-ON2_{R1R2}** and **ON4**.

Acceptor strand	t_R (min)	m/z calcd. for [M-H]	found
ON2_{LL}	25.9	3884.4	3885.3
c-ON2_{LL}	30.0	3910.3	3911.3
ON2_{DL}	28.0	3884.4	3885.6
c-ON2_{DL}	30.0	3910.3	3910.6
ON2_{LLL}	27.6	3983.4	3984.5
c-ON2_{LLL}	34.0	4009.4	4011.5
ON2_{LDL}	31.4	3983.4	3983.9
c-ON2_{LDL}	32.8	4009.4	4012.3
ON2_{DLL}	28.7	3983.4	3985.0
c-ON2_{DLL}	34.3	4009.4	4011.0
ON2_{ODL}	28.7	3983.4	3984.3
c-ON2_{ODL}	34.3	4009.4	4010.4
ON4	26.0	2274.3	2272.6

8. Competitive coupling and cleavage reactions (one-pot)

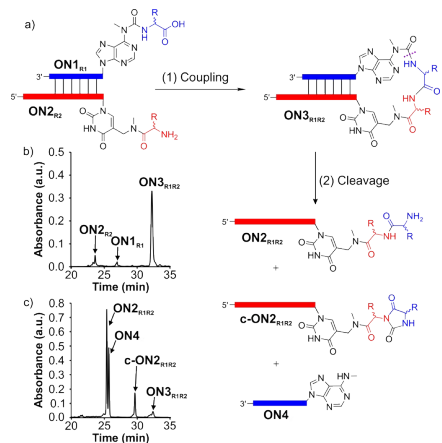
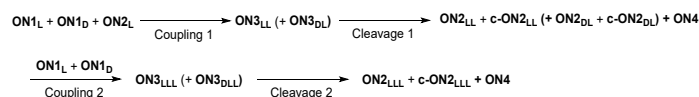


Figure S4. a) Coupling and cleavage reactions between acceptor and donor strands. Reaction conditions: (1) Coupling: EDC/Sulfo-NHS, MES buffer pH 6, NaCl, r.t., 2 h; (2) Cleavage: acetate buffer pH 4, NaCl, 90 °C, 48 h. HPL-chromatograms of the b) coupling reaction of **ON1_{R1}** and **ON2_{R2}** to yield **ON3_{R1R2}** and of the c) cleavage reaction of **ON3_{R1R2}** to yield **ON2_{R1R2}**, **c-ON2_{R1R2}** and **ON4**.



Scheme S7. Schematic representation of the one-pot coupling/cleavage cycle.

The one-pot reactions were carried out using 20 nmol of **ON2_L** as starting material. An equimolar solution of 20 nmol **ON1_L** and **ON1_D** was prepared and the coupling reaction was carried out as described in Section 5. After each coupling step, the crude mixture was filtered using an Amicon® ultra centrifugal filter (3 kDa nominal molecular weight cut-off) to remove the unreacted activator and to exchange the buffer solution. Then, the cleavage reaction was performed as described in Section 7, although the reaction time was extended to 48 h to optimize the yield. After each step, 20 µL of the reaction mixture was analysed by HPLC.

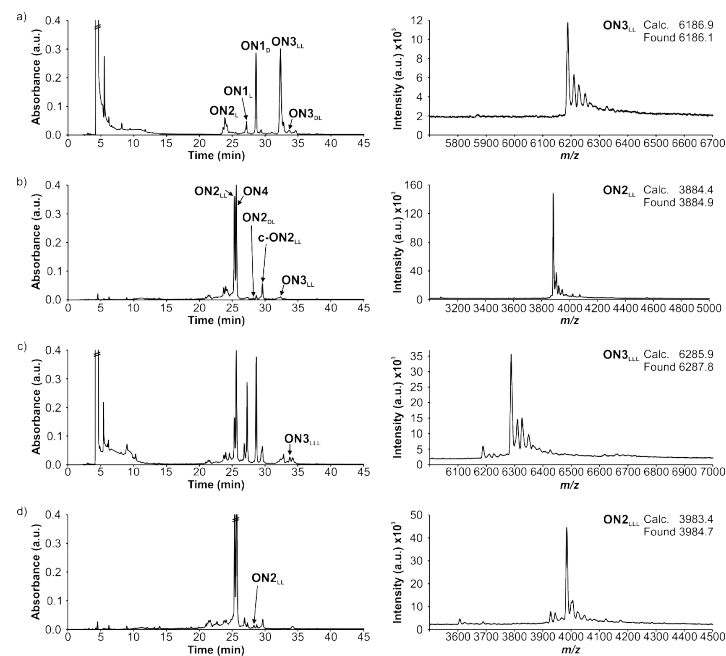


Figure S5: HPL-chromatograms of the crude reaction mixtures of the one-pot reaction and MALDI-ToF spectra of the isolated products; a) first coupling; b) first cleavage; c) second coupling and d) second cleavage.

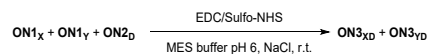
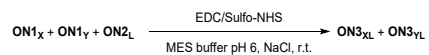
Table S8: Results obtained in one-pot coupling and cleavage cycle.

Step	Activator	pH	T (°C)	Time (h)	Yield (%)	LL vs. DL ratio
Coupling 1	EDC/Sulfo-NHS	6	25	2	58%	94:6
Cleavage 1	-	4	90	48	42% (over two steps)	98:2
Coupling 2	EDC/Sulfo-NHS	6	25	2	3% (over three steps)	n.d.
Cleavage 2	-	4	90	48	1% (over four steps)	n.d.

9. Pair-wise competitive peptide coupling reactions between donor and acceptor oligonucleotides containing dipeptides

EDC/Sulfo-NHS as activator

The peptide coupling reactions were carried out under identical conditions to those described in Section 5 using EDC/Sulfo-NHS as activator. An equimolar solution of **ON1_x** and **ON1_y** was prepared in water and analyzed by HPLC. The 1:1 solution mixture of donor strands, **ON1_x** and **ON1_y**, was used to perform the coupling reactions with 1 equiv. of acceptor strand, **ON2_L** or **ON2_D**.



x and y can be either LL, LD, DL or DD, but x ≠ y

Scheme S8. Pair-wise competitive peptide coupling reactions of ON1_x and ON1_y with ON2_L or ON2_D.

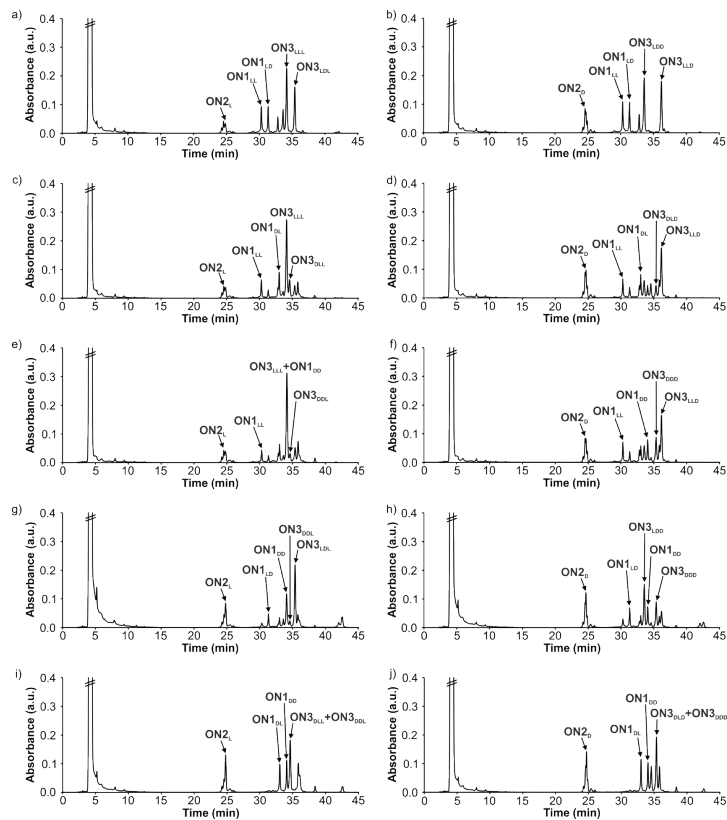


Figure S6. HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of ON2_L with equimolar amounts of: a) ON1_{LL} and ON1_{LD}; c) ON1_{LL} and ON1_{DL}; e) ON1_{LD} and ON1_{DD}; g) ON1_{LD} and ON1_{DL} and i) ON1_{DL} and ON1_{DD} with equimolar amounts of: b) ON1_{LL} and ON1_{LD}; d) ON1_{LL} and ON1_{DL}; f) ON1_{LD} and ON1_{DD}; h) ON1_{LD} and ON1_{DL} and j) ON1_{DL} and ON1_{DD}. Reactions were performed with EDC/Sulfo-NHS (2 h).

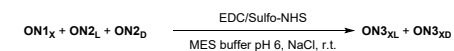
Table S9. Results obtained in the pair-wise competitive peptide coupling reactions of ON2_L and ON2_D with equimolar amounts of ON1_x and ON1_y (x and y can be either LL, LD, DL or DD, but x ≠ y) using EDC/Sulfo-NHS as activator (average of two experiments). Errors were determined to be lower than 10%.

Acceptor strand	Donor strands	Overall yield (%)	ON3 _{xL} /ON3 _{yL} ratio
ON2 _L	ON1 _{LL} + ON1 _{LD}	56	59:41
	ON1 _{LL} + ON1 _{DL}	48	81:19
	ON1 _{LD} + ON1 _{DD}	50	n.d.
	ON1 _{LD} + ON1 _{DL}	48	62:38
ON2 _D	ON1 _{LL} + ON1 _{LD}	50	50:50
	ON1 _{LL} + ON1 _{DL}	30	80:20
	ON1 _{LD} + ON1 _{DD}	37	66:34
	ON1 _{LD} + ON1 _{DL}	34	64:36
	ON1 _{DL} + ON1 _{DD}	24	n.d.

n.d. = not determined (due to peak overlap)

Competitive reactions between two acceptor strands with one donor strand

The peptide coupling reactions were carried out under identical conditions to those described in Section 5 using EDC/Sulfo-NHS as activator. An equimolar solution of ON2_L and ON2_D was prepared in water and analyzed by HPLC. The 1:1 solution mixture of acceptor strands, ON2_L and ON2_D, was used to perform the coupling reactions with 1 equiv. of donor strand, ON1_{LL}, ON1_{LD}, ON1_{DL} or ON1_{DD}.



x can be either LL, LD, DL or DD

Scheme S9. Pair-wise competitive peptide coupling reactions between ON2_L and ON2_D with ON1_{LL}, ON1_{LD}, ON1_{DL} or ON1_{DD}.

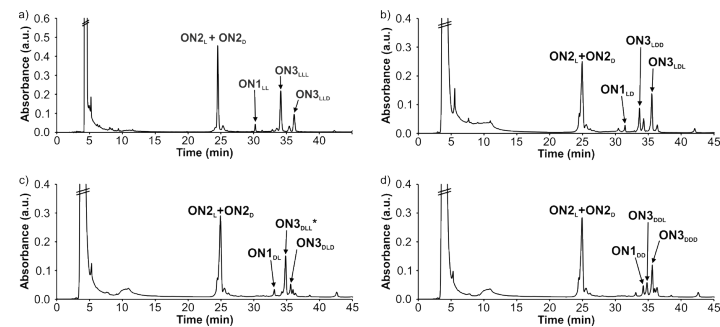


Figure S7. HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of equimolar amounts of ON2_L and ON2_D with a) ON1_{LL}; b) ON1_{LD}; c) ON1_{DL} and d) ON1_{DD}. Reactions performed with EDC/Sulfo-NHS (2 h).

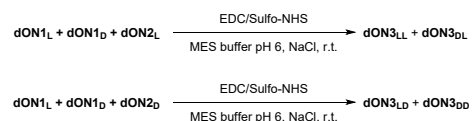
Table S10: Results obtained in the pair-wise competitive peptide coupling reactions of equimolar amounts of **ON2_L** and **ON2_D** with **ON1_L**, **ON1_D**, **ON1_D** or **ON1_D** using EDC/Sulfo-NHS as activator (average of two experiments). Errors were determined to be lower than 10%.

Acceptor strands	Donor strand	Overall yield (%)	ON3 _L /ON3 _D ratio
ON2_L + ON2_D	ON1_L	54	66:34
	ON1_D	40	62:38
	ON1_D	n.d.	n.d.
	ON1_D	27	29:71

n.d. = not determined (due to overlap with impurity)

10. Pair-wise competitive peptide coupling reactions between donor and acceptor deoxy oligonucleotides

The competitive peptide coupling reactions of an equimolar mixture of **dON1_L** and **dON1_D** with **dON2_L** or **dON2_D** were carried out under identical conditions to those described in Section 5 using EDC/Sulfo-NHS as activator.



Scheme S10: Pair-wise competitive peptide coupling reactions between **dON1_L** and **dON1_D** with **dON2_L** or **dON2_D**.

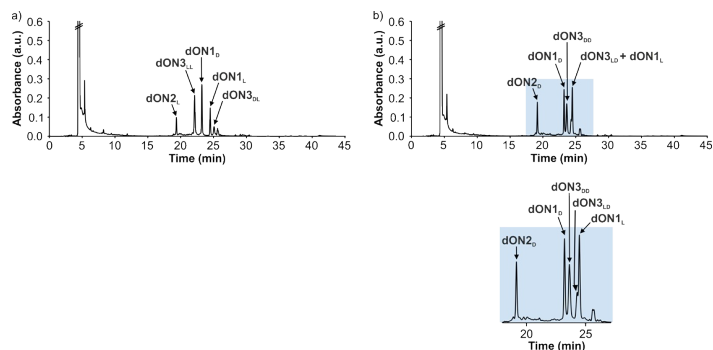


Figure S8: HPL-chromatograms of the crude reaction mixtures for the pair-wise competitive peptide coupling reactions of a) **dON2_L** and b) **dON2_D** with equimolar amounts of **dON1_L** and **dON1_D**. Reactions performed with EDC/Sulfo-NHS (2 h).

Table S11: Results obtained in the pair-wise competitive peptide coupling reactions of **dON2_L** and **dON2_D** with equimolar amounts of **dON1_L** and **dON1_D** using EDC/Sulfo-NHS as activator (average of two experiments). Errors were determined to be lower than 10%.

Acceptor strand	Donor strands	Overall yield (%)	dON3 _{LL} /dON3 _{DL} ratio
dON2_L	dON1_L + dON1_D	24	85:15
Acceptor strand	Donor strands	Overall yield (%)	dON3 _{LD} /dON3 _{DD} ratio
dON2_D	dON1_L + dON1_D	23	31:69

11. Circular dichroism measurements

The circular dichroism (CD) spectra were measured on a Jasco J-810 spectropolarimeter using 5 mm cuvettes in the spectral range from 220 nm to 310 nm and are the average of three measurements. For the experiments we prepared aqueous solutions with the following concentrations: [ON] = 5 μM, [NaCl] = 150 mM, [MES buffer] = 10 mM. The oligonucleotides were annealed by heating to 95°C for 4 min and, subsequently, by cooling down slowly to 5°C before the circular dichroism measurement.

The experiments were conducted with a) an annealed duplex of **ON1_L** and **ON2_L**, b) with the annealed hairpin **ON3_{LL}** and with c) an annealed duplex of canonical strands **ON5** and **ON6**.

ON5: 5'-(AAUCGCU)_m-3' and **ON6** 3'-(UUAGCGACAUG)_m-5' (with 2'OMe nucleotides)

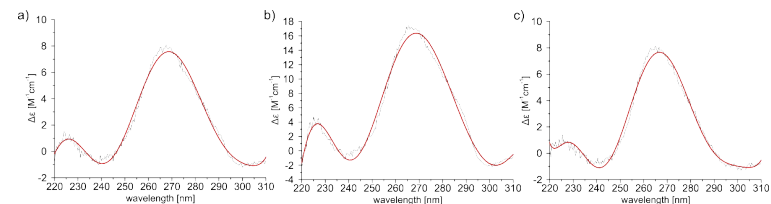


Figure S9: Circular dichroism (CD) spectra of a) an annealed duplex of **ON1_L** and **ON2_L**, b) annealed hairpin of **ON3_{LL}** and c) annealed duplex of canonical strands **ON5** and **ON6**.

12. References

- G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176-2179.
- F. Müller, L. Escobar, F. Xu, E. Węgrzyn, M. Nainytė, T. Amatov, C. Y. Chan, A. Pichler and T. Carell, *Nature*, 2022, **605**, 279-284.
- M. Nainytė, F. Müller, G. Ganazzoli, C.-Y. Chan, A. Crisp, D. Globisch and T. Carell, *Chem. Eur. J.*, 2020, **26**, 14856-14860.
- F. Himmelsbach, B. S. Schulz, T. Trichtinger, R. Charubala and W. Pfeleiderer, *Tetrahedron*, 1984, **40**, 59-72.
- F. Ferreira and F. Morvan, *Nucleosides, Nucleotides Nucleic Acids*, 2005, **24**, 1009-1013.
- A. V. Tataurov, Y. You and R. Owczarzy, *Biophys. Chem.*, 2008, **133**, 66-70.