Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN "SCIENZE CHIMICHE"

CICLO XXXI

COORDINATORE Prof. Bignozzi Carlo Alberto

Design, synthesis and biological evaluation of novel nucleoside and oligonucleotide conjugates of bio-medical interest.

Settore Scientifico Disciplinare CHIM/06

<table>
<tr><td>Dottorando
Dott. Chinaglia Nicola</td><td>Tutore
Prof.ssa Perrone Daniela</td></tr>
</table>

Anni 2015/2018

"Cercai di spiegargli che la nobiltà dell'Uomo, acquisita in cento secoli di prove e di errori, era consistita nel farsi signore della materia, e che io mi ero iscritto a Chimica perché a questa nobiltà mi volevo mantenere fedele. Che vincere la materia è comprenderla, e comprendere la materia è necessario per comprendere l'universo e noi stessi: e che quindi il Sistema Periodico di Mendeleev era una poesia, più alta e più solenne di tutte le poesie digerite in liceo. A pensarci bene, aveva persino le rime."

Primo Levi

"Le incredibili scoperte della chimica esprimono con forza la magia della natura."

Johann Wolfgang von Goethe

Dedicato alla mia famiglia.

Summary

Keywords

Bioconjugates, click chemistry, nucleosides, nucleotides, bile acids, oligonucleotides, phosphoramidite, exon skipping, Duchenne Muscular Dystrophy, glycosyltransferases.

First part: Introduction to bioconjugate chemistry

Bioconjugation is the process of linking or connecting a biological molecule with another moiety. These moieties may include other biomolecules, synthetic polymers or small molecules such as ligands, drugs or fluorescent dyes, among a multitude of other possibilities [1,2].

The linking of two biomolecules forms a hybrid, the bioconjugate, in which the properties of the parent molecules are integrated, yielding a single entity with two different functions (Scheme I). Chemistry of any suitable bioconjugate must be compatible with such an environment, while at the same time preserving the function of the biomolecules. Conjugates are generally formed through the addition of separate but reactively complementary functional groups to each of the two biomolecules. A desired bioconjugate may be obtained by combining two modified biomolecules together. Such bioconjugate can be explored in various fields of research with myriad applications [3,4,5,6].
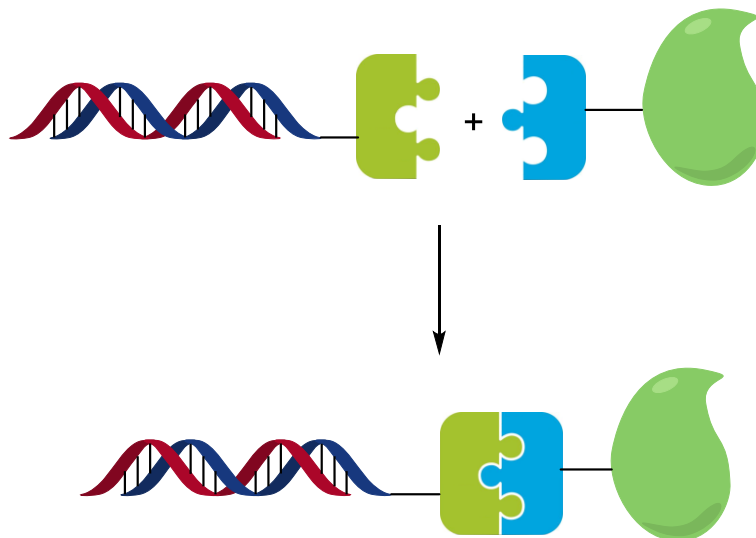
Scheme I Linking of two biomolecules to form a bioconjugate.

1 Traditional bioconjugate reactions

1.1 Introduction

In bioconjugate chemistry, the term "traditional" (although there is not a formal definition) means reactions that satisfy two criteria:

- ➢ reactions with a native functional group in a biomolecule under mild aqueous conditions;
- ➢ used chemistry by many researchers over many years with continued application today.

The native functional groups in target biomolecules are the primary sites for traditional bioconjugate reactions. These functional groups are generally nucleophiles or electrophiles in and of themselves, such that the reactions of a certain functional group in a protein, for example, will be the same reactions that can be used with that functional group in a nucleic acid, lipid, or carbohydrate. Anyway, optimization of those reactions and the scope of their applicability can vary from a biomolecule to another. Each traditional bioconjugate reaction is modular: provided that the requisite functional groups are present, a given reaction can be used to conjugate biomolecules with one another, label a biomolecule with a small molecule (such as a fluorescent dye, contrast agent or drug), or immobilize a biomolecule on a surface or within a matrix [1].

Some typical reactions of the most used functional groups in the traditional bioconjugate chemistry will be illustrated below.

1.1.1 Typical traditional bioconjugate reactions of amines

Primary amines (R-NH$_2$) are one of the most common nucleophiles in biomolecules. Carbonyls, active esters, and isothiocyanates are typically reacted with amine groups for purposes of bioconjugation, and many reagents are available with these functional groups. Conversely, carbonyl groups and active esters (from carboxyl groups) are either available or can be introduced to many biomolecules, and will react with amines [3,4,5,6]. The most common bioconjugate reactions of amines are those with an activated carboxylic acid. Scheme 1.1a illustrates the activation of a carboxylic acid with a carbodiimide to form an *O*-acylisourea intermediate that reacts with primary amines to form a very stable amide linkage [7]. Water-soluble *N*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide (EDC) is the carbodiimide reagent of choice for most bioconjugate reactions. Unfortunately, both the EDC and the *O*-acylisourea intermediate are unstable, and hydrolysis is a major competing reaction [7,8], generally necessitating excesses of carbodiimide. A way of improving the efficiency of carbodiimide conjugation reactions is to convert the *O*-acylisourea intermediate into a more stable succinimidyl ester [9], as shown in Scheme 1.1b,c. This procedure can be as simple as adding *N*-hydroxysuccinimide (NHS) to a reaction mixture with EDC, forming the succinimidyl ester *in situ* for reaction with the amine reagent.
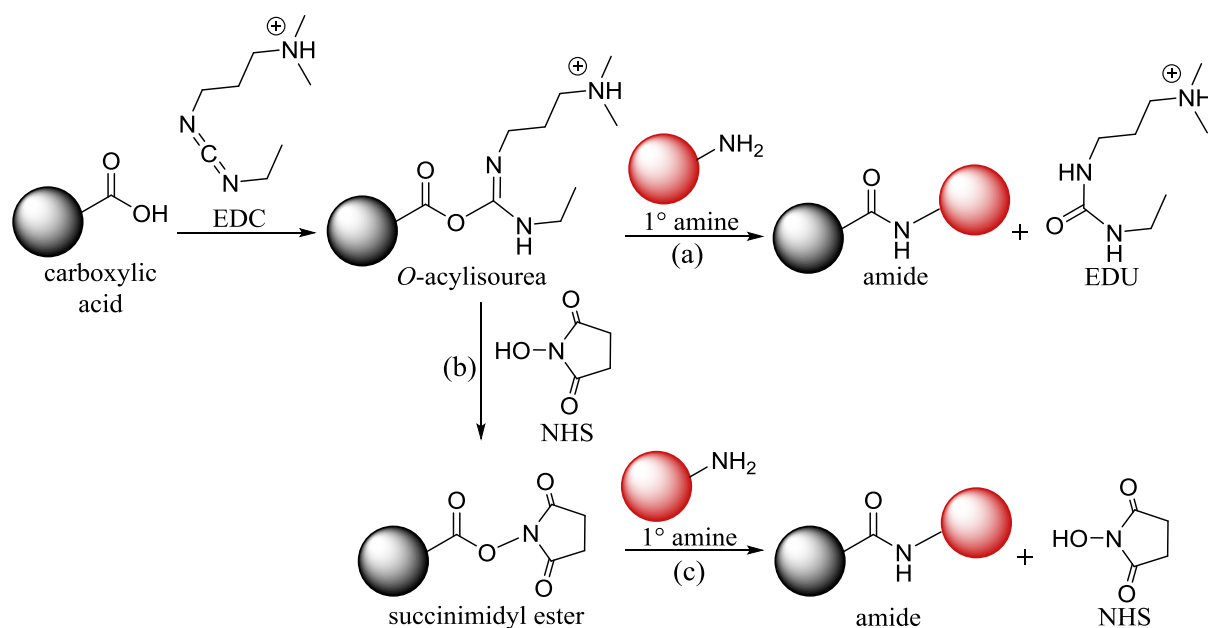
Scheme 1.1 *N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide (EDC) activates a carboxylic acid to an O-acylisourea intermediate that can react with (a) a primary amine to yield an amide or (b) N-hydroxysuccinimide (NHS) to yield a more stable but still reactive succinimidyl ester. (c) The succinimidyl ester reacts with an amine to yield an amide.*

Alternatively, two-step conjugation procedures are sometimes utilized, where EDC and NHS are first added to the carboxylic acid reagent and the amine reagent is added in the second step. Although succinimidyl esters are more stable toward hydrolysis than *O*-acylisoureas, hydrolysis remains a competing reaction [2,10].

Primary amines react spontaneously with aldehydes and ketons to form imines (or Schiff bases) [3,10,11]. These reactions proceed in aqueous media, but are reversible and the equilibrium can shifts to the unconjugated amine and carbonyl groups [3,5] (Scheme 1.2a). To overcome this problem, is then performed a reduction that produces a stable secondary amine [12,13] (Scheme 1.2b).
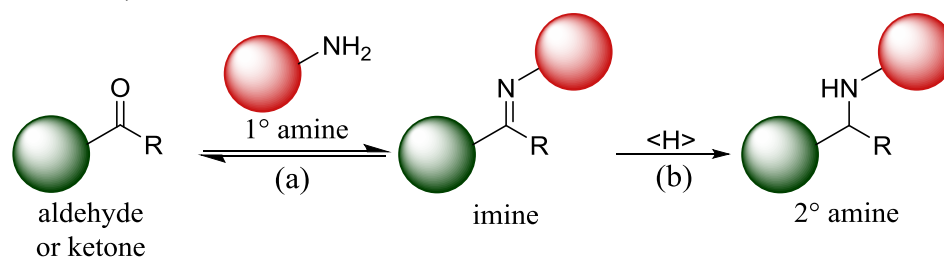
Scheme 1.2 *(a) Reaction between a carbonyl (aldehyde or ketone) and a primary amine to form an unstable imine, followed by (b) reduction to a stable secondary amine with sodium cyanoborohydride.*

1.1.2 Typical traditional bioconjugate reactions of thiols

Thiols (R-SH), and the thiolate anion (R-S⁻) in particular, are the strongest biological nucleophiles [14,15]. Thiols react with many of the same functional groups as amines. For example, thiols will react with active esters to form unstable thioesters [11]. The most common reagents for modifying selectively thiol groups are maleimides, with which happens a Michael addition to form a stable thioether linkage [1,3,4,5] (Scheme 1.3).
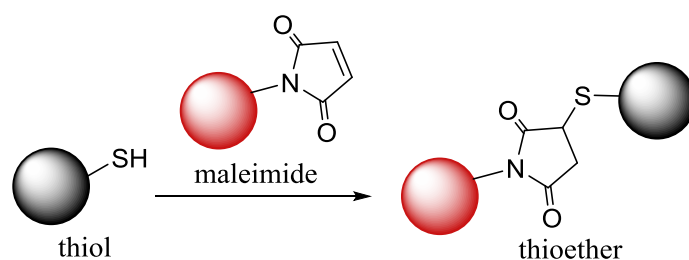
Scheme 1.3 *Michael addition between a thiol and a maleimide to form a stable thioether linkage.*

Another important reaction of thiols is the thiol-disulfide exchange. This reaction, which involves an activated disulfide and a thiol, is very selective for thiols and not subject to competing hydrolysis. Pyridyl disulfide derivatives are the most common reagents for thiol–disulfide exchange, which form pyridine-2-thione as a by-product of the reaction [3] (Scheme 1.4a). The so activated thiol as pyridyl disulfide derivative can react with another thiol to form a disulfide (Scheme 1.4b). The resulting S-S bond is sensitive to reduction: this can be a drawback or an advantage, depending on the application [1].
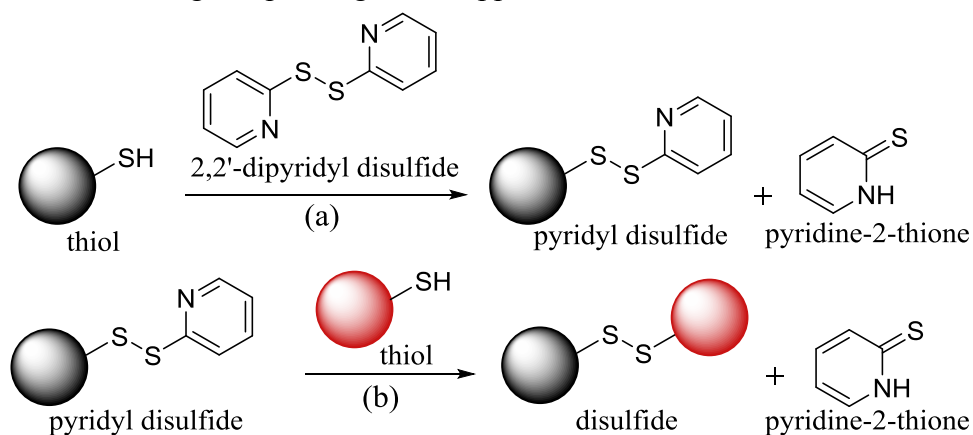
Scheme 1.4 *(a) Activation of a thiol group with 2,2′-dipyridyl disulfide. (b) Thiol–disulfide exchange reaction between a pyridyl disulfide derivative and a thiol to form a new disulfide linkage. Pyridine-2-thione is formed as by-product in both steps.*

1.2 Cross-linking strategies

Below will be discussed strategies for using traditional bioconjugate reaction for some purposes.

1.2.1 Traceless ligations

Traceless ligation (or zero-length cross-linking) refers to the direct formation of new covalent bonds between two biomolecules through an activating agent or reactive group that is not incorporated into the final conjugate, leaving no residual atoms. The most common example of traceless ligation is amide coupling through carbodiimide activation of carboxyl groups or *via* a succinimidyl ester.

Advantages of traceless ligation include minimal (if any) non-native structure in the final bioconjugate and minimization of the final conjugate size. For traceless ligation to be effective, the reactive functional groups must be mutually accessible. Functional groups that are submerged within biomolecular structures will not be able to react if steric hindrance prevents the approach of the complementary functional group. Although there are many activating agents that are potentially capable of traceless ligation, only a small subgroup of these agents are suitably mild and stable for bioconjugate reactions [1].

1.2.2 Homo- and heterobifunctional linkers

Many bioconjugated methods are based on cross-linking agents with reactive functional groups at opposite ends of an alkyl or polyethylene glycol (PEG) spacer [3,4,5,10,16]. These reagents can be homofunctional, if functional groups have the same reactivity (for example bis-NHS esters and bis-maleimido-PEG$_3$, Scheme 1.5ab), or heterobifunctional, if functional groups have a different reactivity (for example NHS and maleimide groups separated by spacer, Scheme 1.5c). The most common uses of these reagents are to connect biomolecules together or to attach to supports and surfaces. A great advantage in use of these reagents is the presence of a linker of discrete length which can diminish biomolecular steric effects [16]. Heterobifunctional cross-linkers often exploit different reactivity of amine and thiols, as shown in Scheme 1.5b.
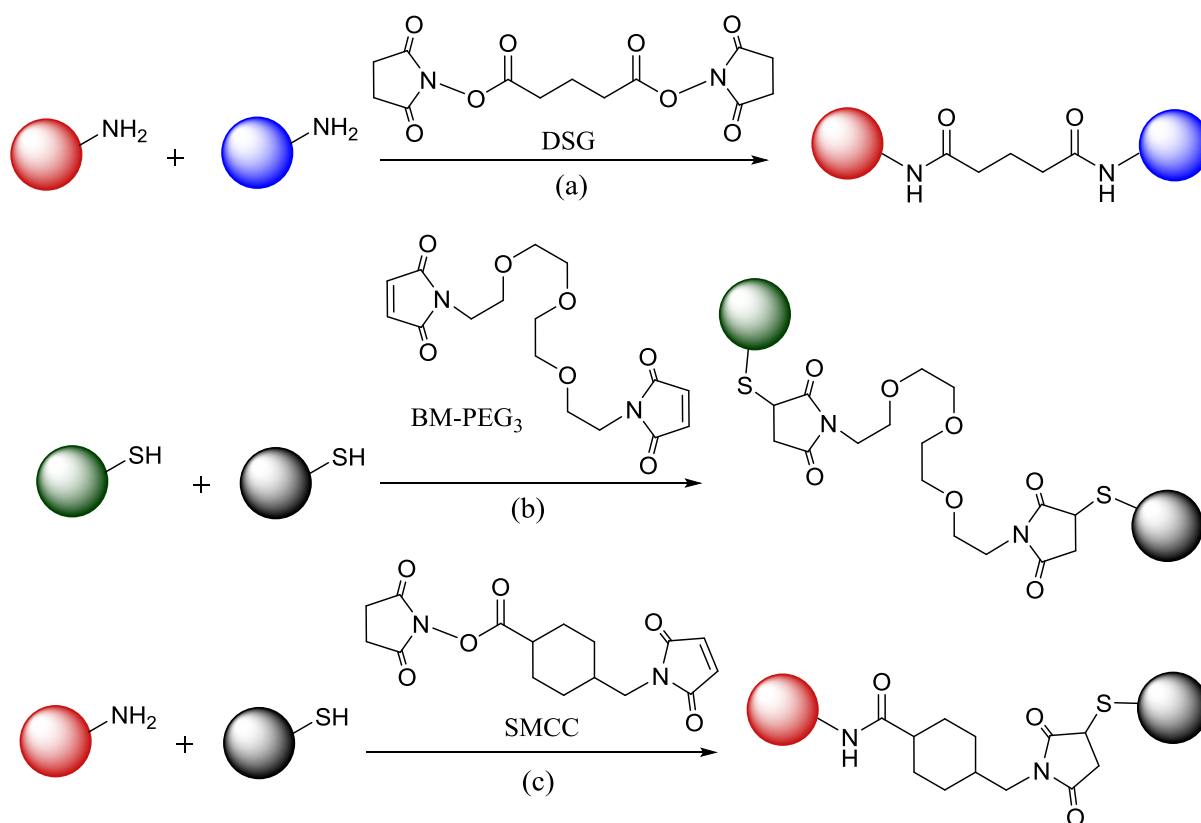
Scheme 1.5 *Examples of reaction with (a,b) homobifunctional and (c) heterobifunctional cross-linkers. DSG = disuccinimidyl glutarate, BM-PEG$_3$= bismaleimidotriethyleneglycol, SMCC = succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate.*

In the selection of appropriate cross-linker, it is fundamental to consider the desired stability of the final linkage: it is possible to have long-term stability conjugations over a wide range of conditions or to create reversible conjugations. In this last cases, the selected cross-linker can include a breakable spacer [3,17,18], which incorporate groups like disulfides, diols or esters that can be cleaved by reduction, oxidation, or a strong nucleophile attack, for example from hydroxylamine. Other reversible cross-linking strategies can exploit the lability due to the photoreactivity or pH of certain functional groups, competitive binding reactions, or even include a structure that is recognized as a substrate by hydrolytic enzymes [1,17,18,19,20].

1.2.3 Functional group conversion

Limited functional groups variety in biomolecules, reagent availability, efficiency or chemoselectivity of one reaction against another in a given set of conditions and successive reaction phases may make it advantageous to introduce a new functional group in a biomolecule compared to an already existing one.

Bisamines (for example ethylenediamine or hexamethylenediamine) are commonly used for converting electrophilic groups into amines [3]. Ethylenediamine can be used to convert an activated carboxylic group into amine with formation of an amidic link (Scheme 1.6a); the same bisamine can react with aldehydes or ketones to form an imine, which after reduction yield the desired amine-derivative with formation of an amino link (Scheme 1.6b).
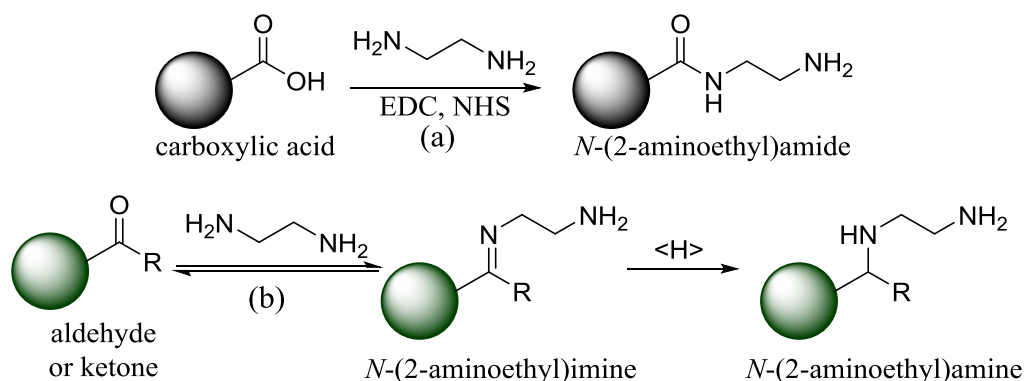
Scheme 1.6 *Conversion in amine of (a) carboxylic acids and (b) aldehydes or ketones using ethylenediamine.*

Undesired intra- or intermolecular cross-linking can occur if both amino groups of the same molecule of bisamine react. To minimize this problem, it is opportune to utilize a sufficiently large excess of bisamines, preferably with short spacers between amino functionality.

Using cistamine it is possible to convert carboxylic groups into thiols [3]. Activated carboxylic acids react with amino groups of cistamine forming an amidic link without interference from the disulfide, then the reduction breaks the S-S bond yielding the thiol (Scheme 1.7).
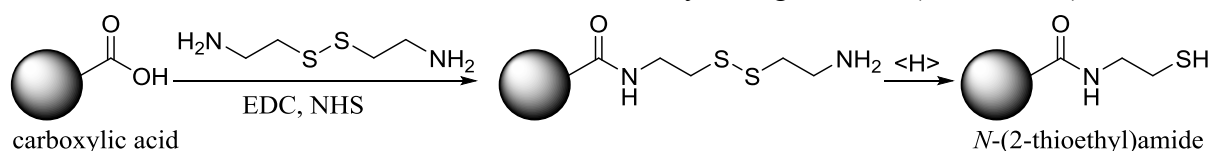
Scheme 1.7 *Conversion of a carboxylic acid into a thiol.*

In this case undesired cross-linking are not a problem, because if both NH_2 functionality of the same molecule of cistamine react, the subsequent breaking of S-S bond forms two molecules of the desidered thiol.

It's also possible to convert amines into carboxylic acids or thiols. To obtain carboxylic acids, it can be used succinic anhydride (Scheme 1.8a) or homobifunctional succimidyl esters (Scheme 1.8b); in this last case, where one succimidyl ester group reacts to form an amidic link and the other is hydrolyzed to form the carboxylic acid, it is opportune to utilize a sufficiently large excess of linker, preferably with short length, to minimize undesired cross-linking [1].
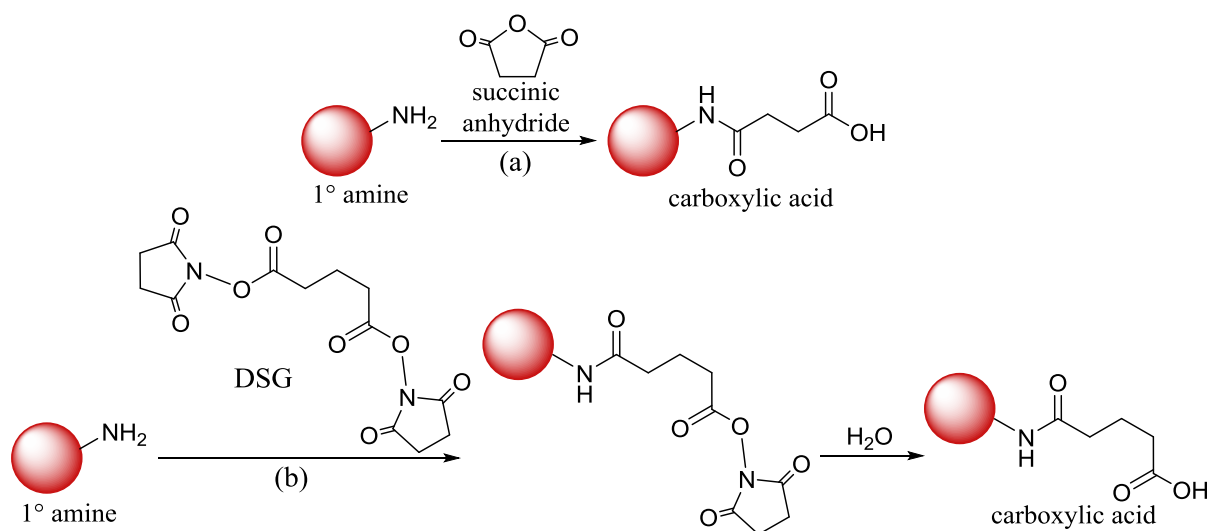
Scheme 1.8 *Conversion of an amino group to a carboxylic acid group using (a) succinic anhydride and (b) disuccinimidyl glutarate (DSG).*

To obtain a thiol from an amino group can be used 2-iminothiolane (Traut's reagent, Scheme 1.9) [3,5,21].
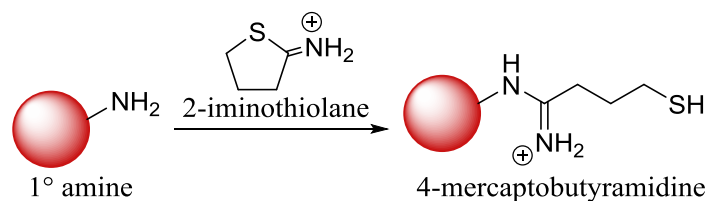
Scheme 1.9 *Conversion of an amine to a thiol using 2-iminothiolane.*

1.3 Challenges associated with traditional bioconjugate reactions

The characteristics of an ideal bioconjugate reaction are:

- ➢ reaction would proceed rapidly with stoichiometric efficiency at sub-micromolar concentrations of reagents and across physiologically relevant ranges of pH and temperature;
- ➢ conjugation would occur exclusively at the targeted site(s) with no side reactions, without interfering with other bioconjugate reactions, and without interference from complex biological matrices and media;
- ➢ reagents and derivatives for the conjugate reaction should also be widely available or easily prepared, stable, and highly soluble in water, as well as nontoxic.

There are very few, if any, traditional bioconjugate reactions that meet all of these ideal criteria.

An important factor which limits the efficiency of many traditional bioconjugate reactions is the competing hydrolysis. Large excesses of active agents (for example succinimidyl esters) or activating agents (for example carbodiimides) are often required to compensate for competing hydrolysis. For some reagents, limited aqueous solubility may require the use of a cosolvent (DMF, DMSO, MeOH, CH$_3$CN, …) to achieve sufficiently high concentrations. Good control of pH, temperature and reaction time can limit undesired side reactions [10]. Empirical optimization of conditions is anyway required for each new bioconjugate reaction: reproducibility may be then limited.

Hydrolysis is not the only problem of limited chemoselectivity. For example, the potent nucleophilicity of thiols may require that them be blocked or reacted prior to any amines because the thiols will react with most amine-reactive agents. Changing the solvent from aqueous to organic in systems that can tolerate it, the effects of competing hydrolysis can be reduced, but other issues of chemoselectivity can be aggravate.

The limited diversity of reactive functional groups available in biomolecules represents another important challenge for traditional bioconjugate chemistry. Most biomolecules are polymers that comprise a limited set of monomers that are repeated many times in a sequence (for example amino acids in proteins, nucleotides in nucleic acids and saccharides in carbohydrates): targeting bioconjugation to specific sites, or targeting specific molecules in a biological system, can be consequently difficult or impossible [10].

Thiol-maleimide coupling and thiol-disulfide exchange are the closest reactions to an ideal bioconjugate reaction. Positive characteristics of these reactions are:

➢ the scarcity of thiols in biomolecules, then the possibility of reacting only specific sites;

➢ selectivity near neutral pH, then limited undesired reactions;

➢ slow hydrolysis rates (compared to other traditional bioconjugate reactions);

➢ large excesses of reagents are often not required;

➢ mild reaction conditions, micromolar or lower concentrations allowed, short reaction times.

Unfortunately, also this reactions show some limitations:

➢ maleimides and their products can hydrolyze at certain pH ranges or if still react for too long time;

➢ used reactives for reduction of disulfides to thiols can interfere with the conjugation reaction, then a two-step process with an intermediate purification may be necessary;

➢ lability of disulfide bonds in cellular reducing environments restricts the use of thiol-disulfide exchange to *in vitro* applications;

➢ maleimide and pyridyl disulfide reagents are less widely available than succinimidyl ester reagents.

Despite these limitations, thiol-maleimide coupling and thiol-disulphide exchange reactions are very useful and represent an important reference for the development of the new chemoselective bioconjugation methods [1].

1.4 Conclusion

Above discussed challenges and limitations should not be considered as prohibitions: in fact there are many applications where traditional bioconjugate reactions are convenient and feasible and provide satisfactory efficiency and levels of control. However, new and more sophisticated applications of bioconjugates require new levels of efficiency and new levels of control, leading to the development of new methodologies that overcome many of the challenges and limitations associated with traditional bioconjugate reactions [1].

Traditional bioconjugate reactions, in particular bioconjugations between activated carboxylic acids and amines in presence of carbodiimides, will be used several times in this thesis.

1.5 Bibliography

1. Algar, W. R.; Dawson, P. E.; Medintz, I. L. *Chemoselective and Bioorthogonal Ligation Reactions: Concepts and Applications,* I ed.; Wiley-VCH Verlag GmbH & Co. KGaA., **2017**.

2. Meares, C. F. Introduction to bioconjugate chemistry. *Perspectives in Bioconjugate Chemistry* **1993,** 1-8.

3. Hermanson, G. T. *Bioconjugate Techniques,* III ed.; Academic Press, Elsevier: New York, **2013**.

4. Wong, S. S.; Jameson, D. M. *Chemistry of Protein and Nucleic Acid Cross-Linking and Conjugation,* II ed.; CRC Press, Taylor & Francis Group: Boca Raton, **2014**.

5. Narain, R. *Chemistry of Bioconjugates: Synthesis, Characterization and Biomedical Applications;* John Wiley & Sons, Inc.: Hoboken, **2014**.

6. Lundblad, R. L. *Chemical Reagents for Protein Modification,* IV ed.; CRC Press, Taylor & Francis Group: Boca Raton, **2014**.

7. Nakajima, N.; Ikada, Y. Mechanism of amide formation by carbodiimide for bioconjugation in aqueous media. *Bioconj. Chem.* **1995,** *6,* 123-130.

8. Gilles, M. A.; Hudson, A. Q.; Borders, C. L. J. Stability of water-soluble carbodiimides in aqueous solution. *Analytical Biochemistry* **1990,** *184,* 244-248.

9. Staros, J. V.; Wright, R. W.; Swingle, D. M. Enhancement by N-hydroxysulfosuccinimide of water-soluble carbodiimide-mediated coupling reactions. *Analytical Biochemistry* **1986,** *156,* 220-222.

10. Brinkley, M. A brief survey of methods for preparing protein conjugates with dyes, haptens, and cross-linking reagents. *Bioconj. Chem.* **1992,** *3,* 2-13.

11. Wilbur, D. S. Radiohalogenation of proteins: an overview of radionuclides, labeling methods, and reagent for conjugate labeling. *Bioconj. Chem.* **1992,** *3,* 433-470.

12. Jentoft, N.; Dearborn, D. G. Labeling of proteins by reductive methylation using sodium cyanoborohydride. *Journal of Biological Chemistry* **1979,** *30,* 4359-4365.

13. Baslé, E.; Joubert, N.; Pucheault, M. Protein chemical modification on endogenous amino acids. *Chemical Biology* **2010,** *17,* 213-227.

14. DeGraaf, A. J.; Kooijman, M.; Hennik, W. E.; Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconj. Chem.* **2009,** *20,* 1281-1295.

15. Kalia, J.; Raines, R. T. Advances in bioconjugation. *Current Organic Chemistry* **2010,** *14,* 138-147.

16. Mattson, G.; Conklin, E.; Desai, S.; Nielander, G.; Savage, M. D.; Morgensen, S. A practical approach to crosslinking. *Molecular Biology Reports* **1993,** *17,* 167-183.

17. Petrotchenko, E. V.; Borchers, C. H. Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrometry Reviews* **2010,** *29,* 862-876.

18. Leriche, G.; Chisholm, L.; Wagner, A. Cleavable linkers in chemical biology. *Bioorg. Med. Chem.* **2012,** *20,* 571-582.

19. Siegel, D. Applications of reversible covalent chemistry in analytical sample preparation. *Analyst* **2012,** *137,* 5457-5482.

20. Christie, R. J.; Anderson, D. J.; Grainger, D. W. Comparison of hydrazone heterobifunctional cross-linking agents for reversible conjugation of thiol containing chemistry. *Bioconjugate Chemistry* **2010,** *21,* 1779-1787.

21. Jue, R.; Lambert, J. M.; Pierce, L. R.; Traut, R. R. Addition of sulfhydryl groups of Escherichia coli ribosomes by protein modification with 2-iminothiolane (methyl 4-mercaptobutyrimidate). *Biochemistry* **1978,** *236,* 5399-5406.

2 [3+2]-Dipolar cycloadditions in bioconjugation

2.1 Introduction

Several applications of bioconjugates require high chemoselectivity in biological systems [1,2,3,4,5], with formation of covalent bonds only between the desired coupling partners while ignoring the multitude of other functional groups present in biological environment [6,7,8,9].

In this class of reactions, [3+2]-dipolar cycloadditions (or 1,3-dipolar cycloadditions) have been very successful for various reasons [10,11,12]. [3+2]-dipolar cycloadditions are processes in which two reagents, a 1,3-dipole and a dipolarophile, bind to form a huge variety of pentatomic heterocycles [13,14,15,16,17,18] (Scheme 2.1). Like the Diels-Alder reactions, they are concerted cycloadditions $[4\pi + 2\pi]$.
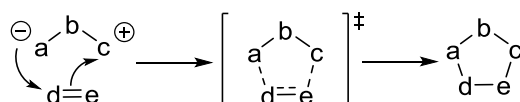
Scheme 2.1 *General pericyclic mechanism for 1,3-dipolar cycloadditions.*

1,3-dipolar molecules are isoelectronic with the allyl anion and possess 4 π electrons; there have at least one resonant structure with separate charges in positions 1 and 3. Some examples of 1,3-dipolar molecules are shown in Figure 2.1.
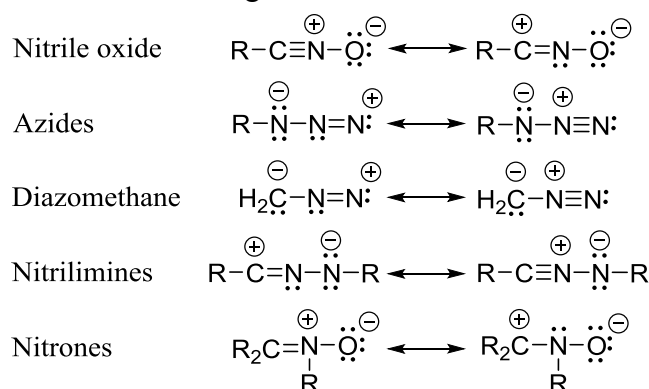
Figure 2.1 *Examples of 1,3-dipoles.*

The dipolarophile is usually an alkene or an alkyne, but also other functional groups with multiple bonds (imines, azo and nitro groups, etc.) can act as dipolarophiles. The reactivity of the dipolarophile depends on the substituents present on the π bond and on the nature of the 1,3-dipole involved in the reaction.

Despite the large number of functional groups that could be theoretically used, the majority of bioconjugation applications have centered on a subset of dipoles and dipolarophiles, with a specific focus on the azide and alkyne functional groups [4,19,20]. These functional groups have the characteristics of excellent stability equilibrium in physiological conditions and are very rare in cellular environment [6,21]: azides and alkynes represent then excellent candidates for synthesis of bioconjugates.

Although highly complementary, both azide and alkyne are kinetically enough stable: this feature is essential to ensure its biocompatibility, but it is also a challenge, because it creates a need for effective modes of activation to stimulate the desired cycloaddition. To this end, two general strategies for alkyne activation have been developed: use of a copper catalyst and use of cyclic alkynes with high ring tension [1,8,22,23,24]. From copper-catalyzed strategies, the

Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC), which will be used in many cases in this thesis, will be illustrated in detail below.

2.2 Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC)

It is known that the reactions between azides (1,3-dipoles) and acetylenes (dipolarophiles), conducted at high temperatures (azide-alkyne cycloaddition AAC, or Huisgen's reaction) [25,26,27,28], form mixtures of 1,4 and 1,5-disubstituted triazoles (Scheme 2.2a): therefore it is essential to identify reaction conditions that control regioselectivity. The discovery that catalytic quantities of Cu(I) increase the reaction rate and control the regiochemistry in a highly specific manner for 1,4-disubstituted triazoles (Scheme 2.2b) was accomplished by Sharpless and Meldal in 2002 [29,30]. The reaction mechanism is no longer of a concerted type, but proceeds according to a sequence with successive stages, involving the formation of copper-alkyne and copper-azide complexes. The improved version of Huisgen's reaction has thus earned the epithet of "click", a term coined by Sharpless itself [31], as it is stereospecific, modular, with a wide range of applicability, proceeds with high yields and generates easily separable by-products through the common purification techniques. Furthermore, the click reaction requires simple operating conditions and can take place in any solvent, including water and other green solvents.
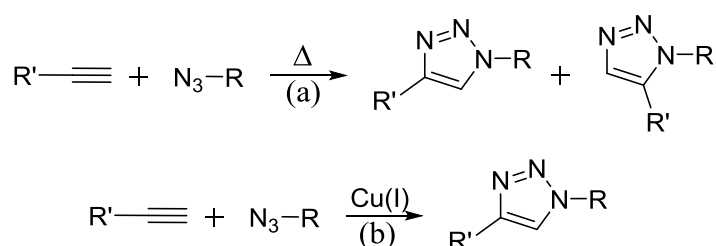
Scheme 2.2 *Thermal (a) and Cu(I)-catalyzed (b) azide-alkyne cycloaddition.*

Since being discovered, CuAAC's exceptionally properties have established it as one of the most versatile reactions in the modern synthetic chemist's toolbox, allowing it to be utilized in an ever-expanding number of applications [32,33,34]. As previously anticipated, the high affinity of the azide group towards the triple bond (chemoselectivity) and, in contrast, the inertia of both functional groups towards most of the substituents connected to the basic structure of many biomolecules (bioortogonality), as well as the stability of the triazole towards chemical and enzymatic degradation, are peculiarities that contribute to use the click reaction to covalently attach molecular bioactive entities. In this operation the use of mild and neutral reaction conditions are essential prerequisites.

However, the CuAAC reaction is not without complications: the greatest limitation is the high concentrations of copper salts often required to achieve a synthetically useful coupling rate. The consequence of this strategy is an increase in the generation of reactive oxygen species (ROS) deriving from the oxidation / reduction chemistry of the copper complexes formed during the non-specific binding to biomolecules [1,35,36].

2.2.1 Mechanism

The first proposed mechanism for the Cu(I)-catalyzed azide-alkyne cycloaddition by Fokin and coworkers is shown in Scheme 2.3: after the formation of σ-Cu(I) acetylide **3**, the azide **4** is coordinated through the negatively charged nitrogen to give the coordinated complex **5** [29]; at this point the complex undergoes a gradual cycloaddition, first giving a 6-term cyclic

intermediate containing the copper (**6**) through the formation of the C-N bond, followed by the rearrangement to generate the acetylide of Cu(I) **7**; the final protonation regenerates the catalyst (**1**) and releases the 1,4-disubstituted triazole **8**.
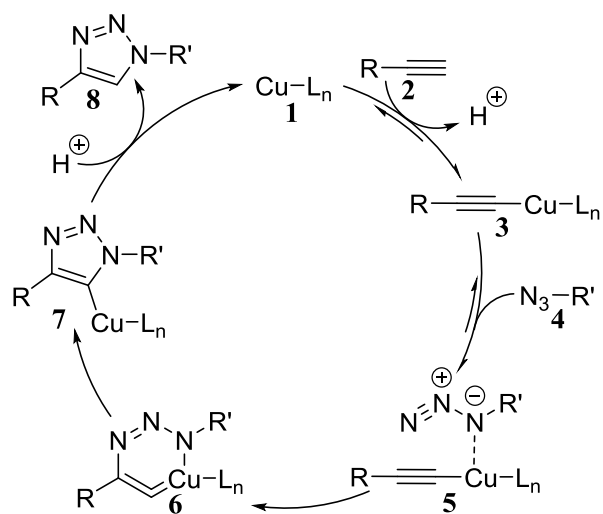
Scheme 2.3 *The first proposed mechanism for CuAAC.*

This catalytic cycle was supported by some evidences:

1) the link between the azide and the copper center (**5**) through internal nitrogen is confirmed by computational investigations (Density Functional Theory, DFT) [37];

2) the current understanding of the relative dipole additions to σ-Cu(I) acetylates, including nitrones and nitrile oxides, confirms stepwise nature of cycloaddition.

Other factors however discredit the first mechanism proposed by Fokin and coworkers:

➢ the intermediate **6** is very high in energy;

➢ kinetic studies have shown that the calculated energy for the transition step between intermediates **5** and **6**, which was predicted to be the rate-limiting step, is too high to justify the observed reaction rates.

An alternate pathway involving a dinuclear copper intermediate was suggested by Finn and coworkers [38], observing a second-order dependence on copper for certain concentrations of azide and alkyne. Successively Straub and Fokin confirmed this hypothesis through computational investigations [39,40,41]. Later Fokin proved experimentally the involvement of dinuclear copper intermediates [42] and proposed a new mechanism (Scheme 2.4).

The biggest difference compared to the originally proposed mechanism is the intermediacy of a σ-Cu(I) acetylide bearing a π-coordinated copper (**10**). After coordination of internal nitrogen of azide **4** to the σ-coordinated copper (**11**), the cycle containing an endo- and an exocyclic copper centers (**12**) is formed. Reductive elimination gives the key intermediate (**7**) and the final protonation regenerates the catalyst (**1**) and releases the 1,4-disubstituted triazole (**8**).
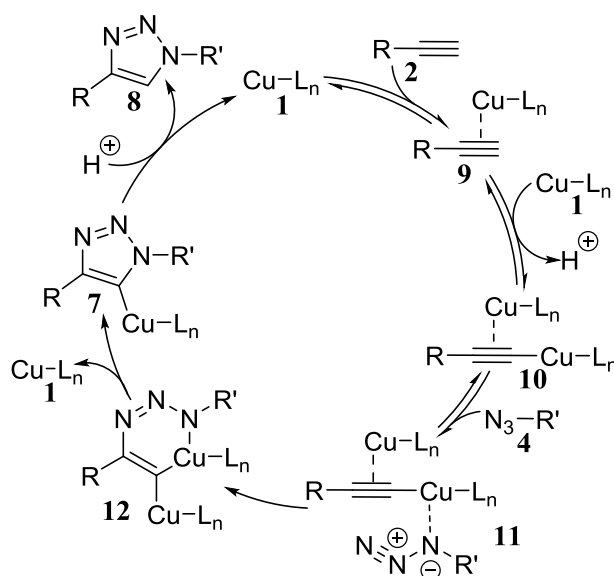
Scheme 2.4 Dinuclear mechanism for CuAAC.

More recently, Bertrand demonstrated that the nature of CuAAC is dynamic [43]: in facts, the mechanism can be considered as a spectrum of possible intermediates, with relative importance of each one highly dependent on the reaction conditions. It follows that both mononuclear and dinuclear σ-acetylides can be productive intermediates in the proper reaction conditions [44]. The proposed mechanism with competing catalytic cycles involving mono- (**3**) and dinuclear (**10**) acetylides [1] is shown in Scheme 2.5.
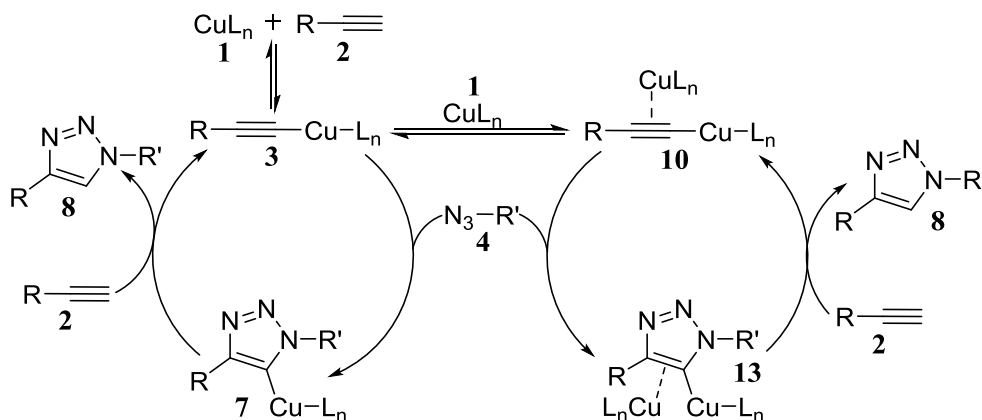
Scheme 2.5 Proposed mechanism with competing catalytic cycles involving mono- and dinuclear acetylides.

2.2.2 Benefits, limitations and applications

Despite the high biocompatibility of both azides and alkynes, some minor problems may affect these functional groups; these complications, illustrated below, concern particular compounds and can be avoided by designing alternative molecular tags:

➢ azides are susceptible to being attacked and reduced by the soft nucleophiles, such as phosphines [45];

➢ highly electron-poor aryl-azides can form nitrenes by photochemical reactions [46];

➢ alkynes can react and give coupling with free thiols [47];

➢ alkynes with adjacent carbonyl function, such as propiolate derivatives, are highly reactive Michael acceptors [48].

One of the major advantages of CuAAC is its exceptionally wide chemical scope, difficult to reach by other bioortogonal synthetic reaction strategies [7,36].

The greatest limitation of CuAAC is the need to use Cu (I) as a catalyst, indispensable for the activation of the alkino as copper acetylide. To activate the kinetically stable alkyne group, copper performs the double action of weakening the electronic concentration of the π-system and polarizing the C-C bond [49]. A significant challenge is represented by the ability to supply and maintain copper at a significantly high concentration and only in the catalytically active state, i.e. Cu(I).

The use of copper is also problematic because of its propensity to form ROS, which damage the biomolecules and lead to cell toxicity [35,50]. The oxidative damage resulting from the incorporation of copper complexes is due to the generation of hydroxyl and alkoxyl radicals through the Fenton reaction [51], which involves multiple single-electron transfer events between the metal center and the reactive oxygen partners (Scheme 2.6).
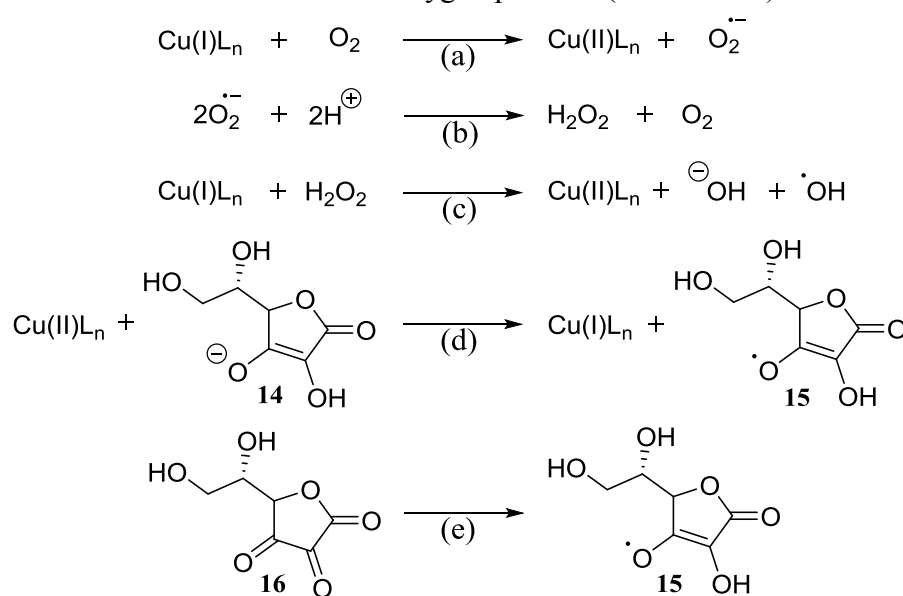
Scheme 2.6 *Single-electron processes involved in Cu(I)/Cu(II) oxidative cycling.*

In addition to the production of hydroxyl radicals due to the oxidation of Cu(I) to Cu(II) [52] (Scheme 2.6abc), an alkyl radical (**15**) can be produced if sodium ascorbate (**14**) is used to reduce Cu(II) to Cu(I) *in situ* [24,36] (Scheme 2.6d). This radical **15** is also a by-product deriving from the final product of the reduction of copper *in situ*, i.e. dehydroascorbate (**16**, Scheme 2.6e).

A strategy to help improve the ability to localize the catalyst with the azide and alkyne coupling partners is to add a chelation group adjacent to the reactive center. Because of the conserved linear geometry of copper acetylide, this strategy is successful only when it is used with the azide component. After this strategy had been tested by Zhu on pyridyl azides [53], Ting extended it to bioconjugates using benzyl azides (**17**) and picolyl azides (**18**) with low copper concentrations under *in vitro* protein labeling conditions [54] (Scheme 2.7). It has been subsequently shown that the picolyl azide allows CuAAC to occur faster and with reduced Cu concentrations [55].
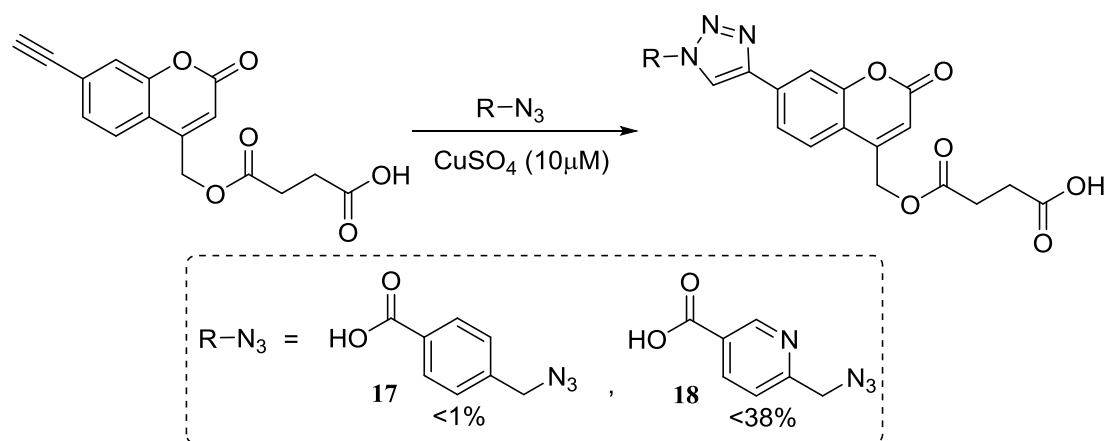
Scheme 2.7 *CuAAC reactions performed by Ting with benzyl azide and picolyl azide and related conversions reported after 30 minutes* [54].

Another strategy to accelerate the rate of productive coupling by limiting the degree of oxidative stress on a living system implies the incorporation of supporting ligands (Figure 2.2). The first support ligand that showed to accelerate CuAAc for both synthetic applications and bioconjugations was the tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA). The poor water solubility of this compound subsequently led to the development of other tris((triazolyl)methyl)amine support ligands with polar groups, such as THPTA [36], BTTAA [56,57] and TABTA [55]. A support ligand with an alternative structure is represented by the bathophenanthrolinedisulfonate disodium salt (BPS), which however requires a large excess of copper to give an optimal coupling [58].

TBTA: R=R'=Bn
THPTA: R=R'=CH$_2$CH$_2$CH$_2$OH
BTTAA: R=*t*-Bu, R'=CH$_2$COOH
TABTA: R=*t*-Bu, R'=CH$_2$CH$_2$N(CH$_3$)$_3$
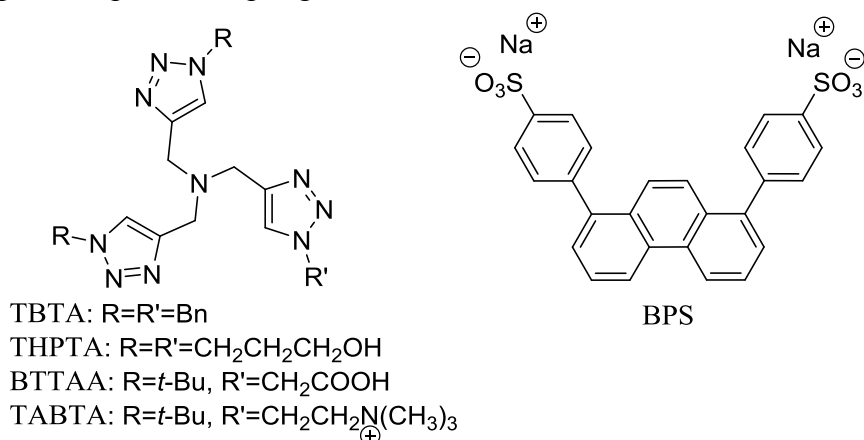
Figure 2.2 *Accelerating ligands for CuAAC applied to bioconjugation.*

These ligands accelerate CuAAC by stabilizing the oxidation state Cu(I), thus allowing a greater amount of metal to remain catalytically active [59]. They also closely bond copper to prevent Cu (II) oxidation, thus reducing the propensity to generate ROS. However, it has been demonstrated by Pezacki that various copper / ligand complexes, such as those formed with BPS, TBTA and THPTA, show a high toxicity, higher than the levels observed with non-combined Cu(II) [50]. This study clearly demonstrates that both the native toxicity of the copper / ligand complex and the bioconjugation reaction kinetics must be considered for the future design of ligands.

In addition to limiting oxidative stress, it is crucial to preserve the function of a target biomolecule when planning to label live cells. In fact, in this context non-specific damages to proteins and splitting of the strand in high molecular weight DNA are common issues [60]. To

further complicate matters, there is the redox activity of the copper catalyst, which can interfere on systems with pendant-free thiol functional groups. It was demonstrated by Bode that oligopeptides containing a cysteine residue (*i.e.* with a SH group) were highly susceptible to oxidative dimerization by disulfide formation [61] (Scheme 2.8); this collateral reaction complicates the isolation of the desired protein conjugate, even if it does not stop the desired triazole formation. The removal of the cysteine eliminated this challenge.

Scheme 2.8 *Unwanted disulfide formation during peptide functionalization.*
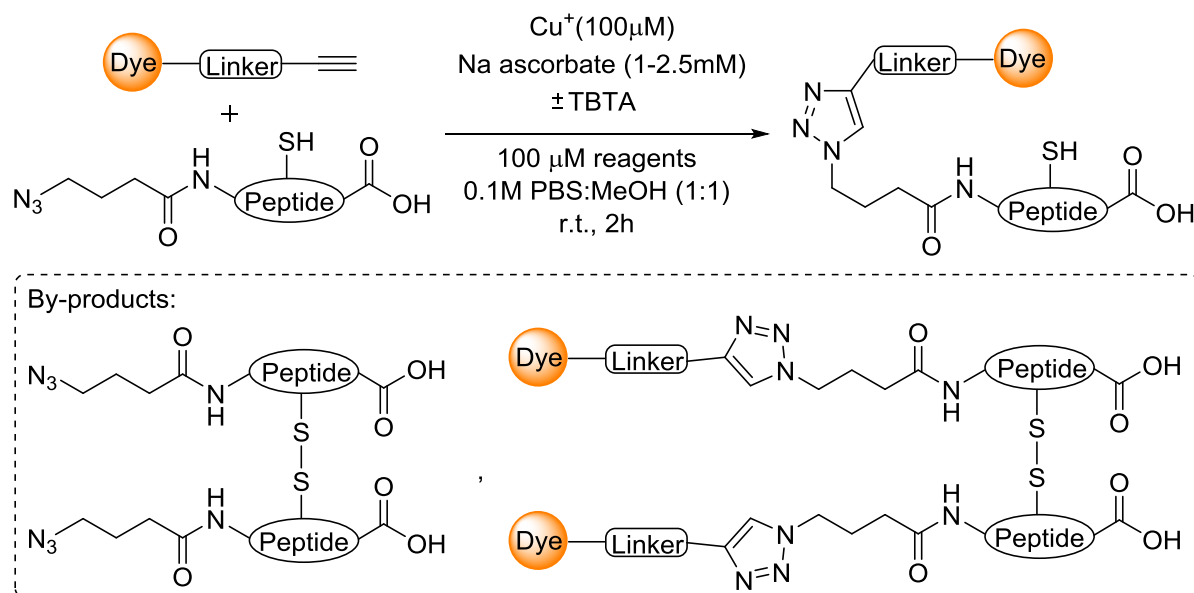
The addition of exogenous reducing agents, such as TCEP or Trolox (Figure 2.3), to the use of supportive ligands can further help to suppress oxidative damage [62,63]. Furthermore, it has been observed that the use of DMSO as a cosolvent can radically suppress the degree of long-stranded DNA cleavage [64]; however, the intrinsic toxicity of DMSO makes it difficult to use it *in vivo* [1].
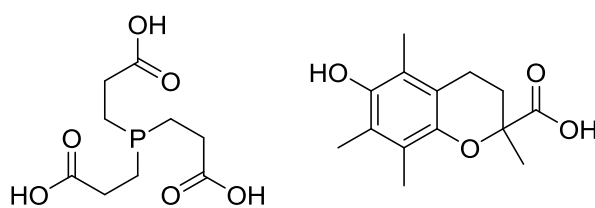
Figure 2.3 *Reducing agents TCEP and Trolox.*

2.2.3 Conclusion

The highest quality of CuAAC, which makes it a powerful tool with proven efficacy for both *in vitro* and *in vivo* applications, is the multitude of usable substrates which allows easy incorporation of many complementary tags. However, cellular toxicity of copper catalysts has limited *in vivo* applications, making it a primary tool for the labeling of biomolecules *in vitro*. With the development of new ligands and reaction protocols, it will soon be possible to expand the field of *in vivo* applications of CuAAC. Although every particular application requires the tuning of the catalyst mix, the following general conditions can be identified [1]:

➢ copper and probe are usually introduced in a 1:1 stoichiometry, at approximately 30–75 μM;

➢ the catalyst is introduced in the Cu(II) oxidation state (CuSO$_4$ is the most common copper source due to its solubility and availability);

> sodium ascorbate is added to initiate the reaction (typically in a ratio of 10:1–20:1 relative to copper);

> ligand (if used) is usually introduced in excess (typical ratio of 1:1–4:1 relative to copper). Too much additional ligand may suppress the rate of reaction by sequestering copper;

> supportive solvents and buffers are tolerated and can be applied based on the requirements of the substrate and probe.

CuAAC will be used many times to the synthesis of bioconjugates in this thesis.

2.3 Bibliography

1. Algar, W. R.; Dawson, P. E.; Medintz, I. L. *Chemoselective and Bioorthogonal Ligation Reactions: Concepts and Applications,* I ed.; Wiley-VCH Verlag GmbH & Co. KGaA., **2017**.

2. Prescher, J. A.; Bertozzi, C. R. *Nature Chemical Biology* **2005,** *1* (1), 13-21.

3. Sletten, E. M.; Bertozzi, C. R. *Angewandte Chemie International Edition* **2009,** *48* (38), 6974-6998.

4. Patterson, D. M.; Nazarova, L. A.; Prescher, J. A. *ACS Chemical Biology* **2014,** *9* (3), 592-605.

5. Stephanopoulos, N.; Francis, M. B. *Nature Chemical Biology* **2011,** *7* (12), 876-884.

6. Shih, H.-W.; Kamber, D. N.; Prescher, J. A. *Current Opinion in Chemical Biology* **2014,** *21,* 103-111.

7. Mckay, C. S.; Finn, M. G. *Chemistry and Biology* **2010,** *39* (4), 1075-1101.

8. Jewett, J. C.; Bertozzi, C. R. *Chemical Society Reviews* **2010,** *39* (4), 1272-1279.

9. Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006,** *312* (5771), 217-224.

10. Hashimoto, T.; Maruoka, K. *Chemical Reviews* **2015,** *115* (11), 5366-5412.

11. Coldham, I.; Hufton, R. *Chemical Reviews* **2005,** *105* (7), 2765-2810.

12. Pearson, W. H. *Pure and Applied Chemistry* **2002,** *74* (8), 1339-1347.

13. Huisgen, R. *Angew. Chem.* **1963,** *75,* 604.

14. Huisgen, R. *Angew. Chem. Int. Ed. Engl.* **1963,** *2,* 565.

15. Huisgen, R. *Angew. Chem.* **1963,** *75,* 741.

16. Huisgen, R. *Angew. Chem Int. Ed. Engl.* **1963,** *2,* 633.

17. Huisgen, R. *1,3-Dipolar Cycloaddition Chemistry;* Wiley: New York, **1984**; pp 1-176.

18. Padwa, A. *Comprehensive Organic Synthesis;* Pergamon: Oxford, **1991**; Vol. 4, pp 1069-1109.

19. Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chemical Biology* **2006,** *1* (10), 644-648.

20. Zhang, X.; Zhang, Y. *Molecules* **2013,** *18* (6), 7145-7159.

21. Bianco, A.; Townsley, F. M.; Greiss, S.; Lang, K.; Chin, J. W. *Nature Chemical Biology* **2012,** *8* (9), 748-750.

22. Yang, M.; Yang, Y.; Chen, P. R. *Topics in Current Chemistry* **2015,** 1-29.

23. Best, M. D. *Biochemistry* **2009,** *48* (28), 6571-6584.

24. Lallana, E.; Riguera, R.; Fernandez-Megia, E. *Angew. Chem. Int. Ed.* **2011,** *50* (38), 8794-8804.

25. Huisgen, R. *Pure Appl. Chem.* **1989,** *61,* 613.

26. Huisgen, R.; Szeimies, G.; Moebius, L. *Chem. Ber.* **1967,** *100,* 2494.

27. Lwowski, W. 1,3-Dipolar Cycloaddition Chemistry. Padwa, A., Ed.; Wiley: New York, **1984**; Chapter 5, Vol. 1, pp 559-645.

28. Wamhoff, H. *Comprehensive Heterocyclic Chemistry;* Pergamon: Oxford, **1984**; Vol. 5, pp 669-732.

29. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2002,** *41* (14), 2596-2599.

30. Tornoe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002,** *67,* 3057.

31. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001,** *40* (11), 2004-2021.

32. Hein, J. E.; Fokin, V. V. *Chemical Society Reviews* **2010,** *39* (4), 1302-1315.

33. Moses, J. E.; Moorhouse, A. D. *Chemical Society Reviews* **2007,** *36* (8), 1249.

34. Meldal, M.; Tornoe, C. W. *Chemical Reviews* **2008,** *108* (8), 2952-3015.

35. Brewer, G. J. *Chemical Research in Toxicology* **2009,** *23* (2), 319-326.

36. Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G. *Bioconjugate Chemistry* **2010,** *21* (10), 1912-1916.

37. Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. *Journal of the American Chemical Society* **2005,** *127* (1), 210-216.

38. Rodionov, V. O.; Fokin, V. V.; Finn, M. G. *Angew. Chem. Int. Ed.* **2005,** *44* (15), 2210-2215.

39. Straub, B. F. *Chemical Communications* **2007,** *37,* 3868.

40. Ahlquist, M.; Fokin, V. V. *Organometallics* **2007,** *26* (18), 4389-4391.

41. Makarem, A.; Berg, R.; Rominger, F.; Straub, B. F. *Angew. Chem. Int. Ed.* **2015,** *54* (25), 7431-7435.

42. Worrell, B. T.; Malik, J. A.; Fokin, V. V. *Science* **2013,** *340* (6131), 457-460.

43. Bidal, Y. D.; Lesieur, M.; Melaimi, M.; Nahra, F.; Cordes, D. B.; Athukorala Arachchige, K. S.; Slawin, A. M. Z.; Bertrand, G.; Cazin, C. S. J. *Advanced Synthesis and Catalysis* **2015,** *357* (14-15), 3155-3161.

44. Jin, L.; Tolentino, D. R.; Melaimi, M.; Bertrand, G. *Science Advances* **2015,** *1* (5), e1500304.

45. Debets, M. F.; van der Doelen, C. W. J.; Rutjes, F. P. J. T.; van Delft, F. L. *ChemBioChem* **2010,** *11* (9), 1168-1184.

46. Norberg, O.; Deng, L.; Yan, M.; Ramstrom, O. *Bioconjugate Chemistry* **2009,** *20* (12), 2364-2370.

47. Hoyle, C. E.; Bowman, C. N. *Angew. Chem. Int. Ed.* **2010,** *49* (9), 1540-1573.

48. Declerck, V.; Martinez, J.; Lamaty, F. *Chemical Reviews* **2009,** *109* (1), 1-48.

49. Siemsen, P.; Livingston, R. C.; Diederich, F. *Angew. Chem. Int. Ed.* **2000,** *39* (15), 2632-2657.

50. Kennedy, D. C.; Mckay, C. S.; Legault, M. C. B.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P. *Journal of the American Chemical Society* **2011,** *133* (44), 17993-18001.

51. Tabbì, G.; Fry, S. C.; Bonomo, R. P. *Journal of Inorganic Biochemistry* **2001,** *84* (3-4), 179-187.

52. Biaglow, J. E.; Manevich, Y.; Uckun, F.; Held, K. D. *Free Radical Biology and Medicine* **1997,** *22* (7), 1129-1138.

53. Kuang, G.-C.; Guha, P. M.; Brotherton, W. S.; Simmons, J. T.; Stankee, L. A.; Nguyen, B. T.; Clark, R. J.; Zhu, L. *Journal of the American Chemical Society* **2011,** 110816125108031.

54. Uttamapinant, C.; Tangpeerachaikul, A.; Grecian, S.; Clarke, S.; Singh, U.; Slade, P.; Gee, K. R.; Ting, A. Y. *Angew. Chem. Int. Ed.* **2012,** *51* (24), 5852-5856.

55. Rudolf, G. C.; Sieber, S. A. *ChemBioChem* **2013,** *14* (18), 2447-2455.

56. Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; Soriano del Amo, D.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. *Angew. Chem. Int. Ed.* **2011,** *50* (35), 8051-8056.

57. Soriano del Amo, D.; Wang, W.; Jiang, H.; Besanceney, C.; Yan, A. C.; Levy, M.; Liu, Y.; Marlow, F. L.; Wu, P. *Journal of the American Chemical Society* **2010,** *132* (47), 16893-16899.

58. Lewis, W. G.; Magallon, F. G.; Fokin, V. V.; Finn, M. G. *Journal of the American Chemical Society* **2004,** *126* (30), 9152-9153.

59. Chan, T. R.; Hilgraf, R.; Sharpless, K. B.; Fokin, V. V. *Organic Letters* **2004,** *6* (17), 2853-2855.

60. Eiberger, W.; Volkmer, B.; Amouroux, R.; Dherin, C.; Radicella, J. P.; Epe, B. *DNA Repair* **2008,** *7* (6), 912-921.

61. Saito, F.; Noda, H.; Bode, J. W. *ACS Chemical Biology* **2015,** *10* (4), 1026-1033.

62. Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. A. G. *Angew. Chem. Int. Ed.* **2009,** *48* (52), 9879-9883.

63. Rasnik, I.; McKinney, S. A.; Ha, T. *Nature Methods* **2006,** *3* (11), 891-893.

64. Abel, G. R.; Calabrese, Z. A.; Ayco, J.; Hein, J. E.; Ye, T. *Bioconjugate Chemistry* **2016,** *27* (3), 698-704.

Second part: Bile acids bioconjugates

Bile acids, thanks to their amphiphilic properties (in fact they have a hydrophobic and a hydrophilic side, as shown in Figure II), represent interesting carriers of poorly lipophilic molecules: hence the idea of synthesizing bioconjugates that exploit these molecules as transporters.
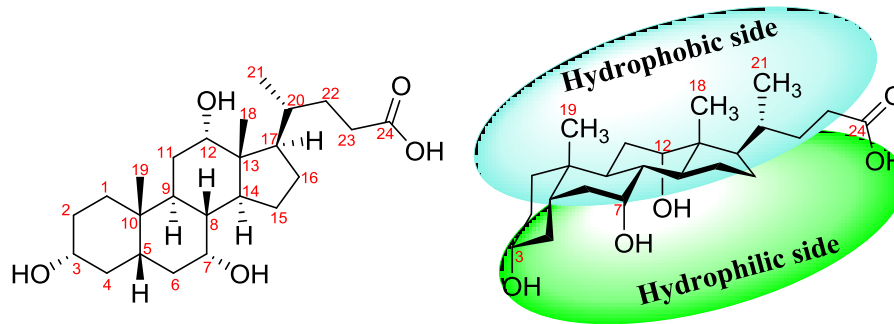
Figure II *Structure of cholic acid with its hydrophobic and hydrophilic sides.*

3   Introduction to bile acids chemistry

3.1   Origin of bile acids

Bile acids are the final products of cholesterol metabolism (Scheme 3.1) and represent one of the main ways of eliminating it from the body. At physiological pH they are in the form of anions and for this reason they are also called bile salts [1].
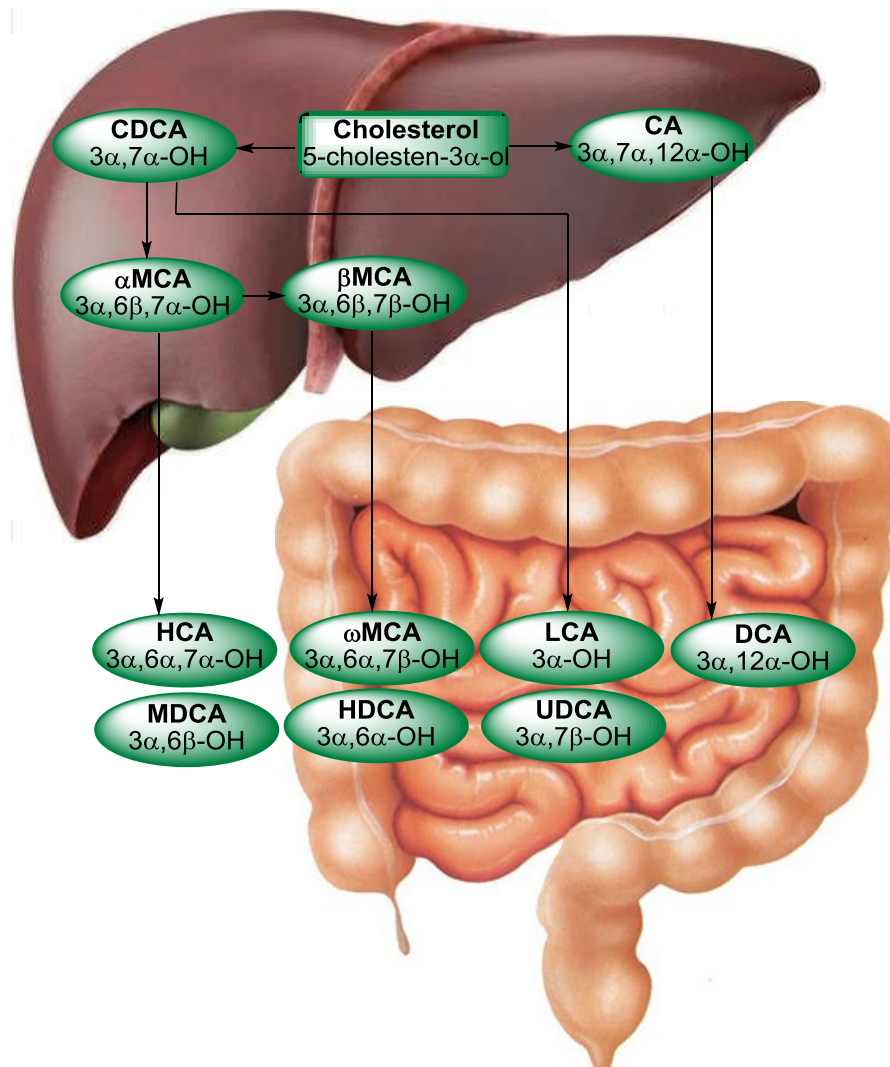
Scheme 3.1 *Conversion of cholesterol into primary (in the liver) and secondary (in the intestine) bile acids.*

Bile acids can be divided into 3 categories: primary, secondary and tertiary. In the liver, the cholesterol undergoes a series of catabolic reactions leading to the synthesis of primary bile acids: cholic acid (CA) and chenodeoxycholic acid (CDCA). These are stored in the bile inside the gallbladder after being conjugated with the amino acids taurine and glycine, by the microsomal enzyme cholil-CoA glycine/taurine aminotransferase, respectively giving taurocholic and taurochenodeoxycholic acid and to glycolic and glycochenodeoxycholic acid, which are more hydrophilic than the starting compounds. The conjugation between the amino group of the two amino acids and the carboxyl group of bile acids forms an amide bond. The bile is then poured into the small intestine where part of the primary bile acids undergoes the deconjugation and dehydroxylation reactions in C-7 by the bacterial flora. In this way are synthesized the secondary bile acids, deoxycholic acid (DCA) and lithocholic acid (LCA), which derive respectively from the cholic acid and chenodeoxycholic acid. Subsequently, the

secondary bile acids are conveyed into the liver through the portal vein and, after being conjugated with glycine and taurine, are excreted into the bile again [2].

Other transformations occurring in the intestinal lumen are the hydrolysis of taurine and glycine and the epimerization in C-7 (transformation of the substituent from α to β) of the chenodeoxycholic acid which becomes ursodeoxycholic acid (UDCA). Furthermore, in the liver, the lithocholic acid, which is hepatotoxic, is sulphated to increase its water-solubility and consequently decrease its reabsorption: this results in the sulfolitocholic acid, a tertiary bile acid [3,4]. The chemical structure of the most common bile acids is shown in Scheme 3.2.
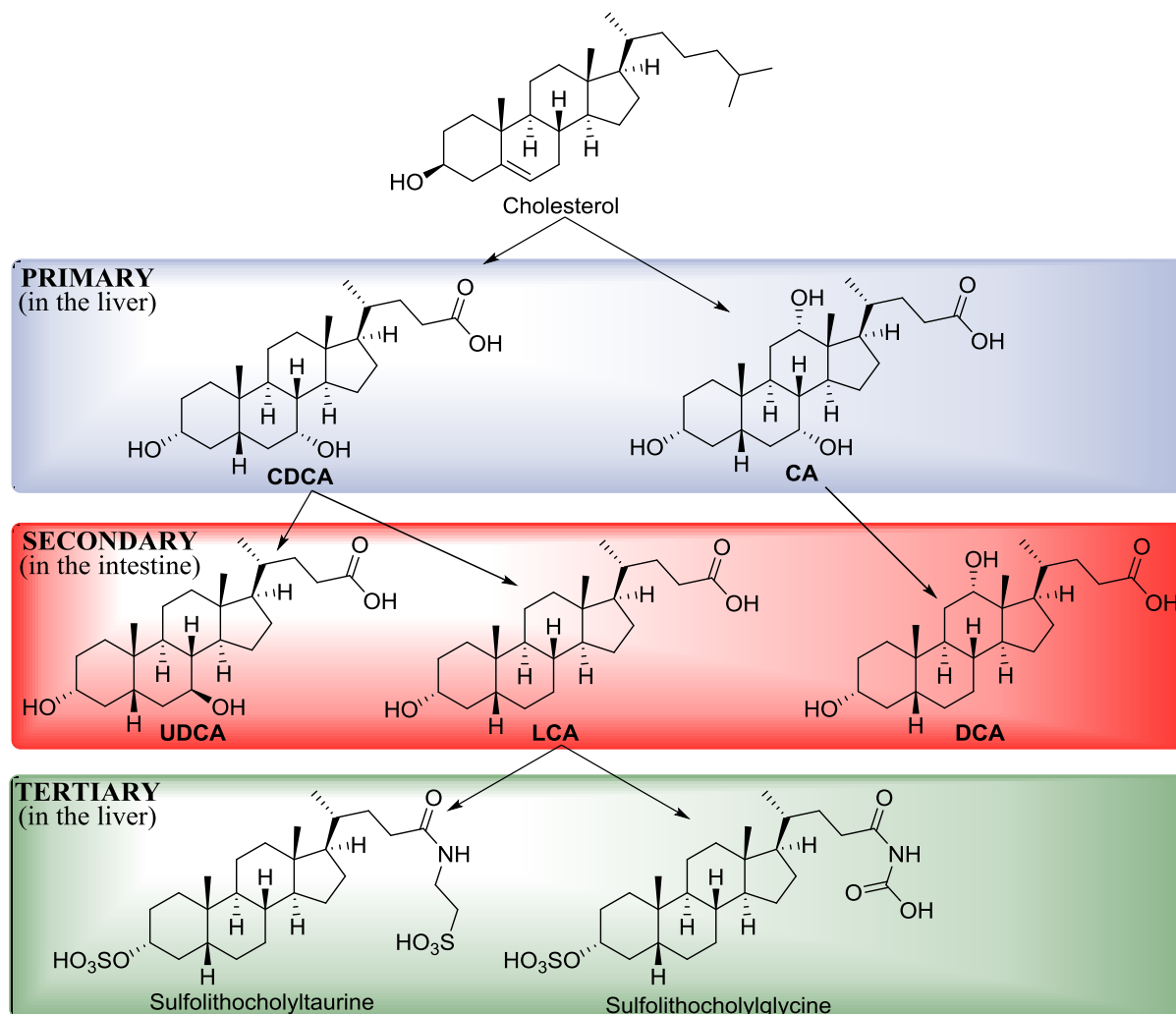
Scheme 3.2 Chemical structure and transformation of the most common human bile acids.

A curiosity derives from the fact that many bile acids take their name from the animal in which they are most abundant or from which they were isolated for the first time: for example, chenodeoxycholic acid was isolated for the first time from the bile of the domestic goose (in Greek χήνα), while the ursodeoxycholic acid results to be most abundant in the bear (in Latin ursus) [5,6,7,8,9].

The main role of bile acids is the digestion of lipids and fat-soluble vitamins, but their antibacterial and antiviral properties are also known [10].

3.2 Structure and activity

Bile acids consist of 24 carbon atoms and structurally they are characterized by the cyclopentanoperhydrophenanthrene ring [8], more commonly known as steroid nucleus. They also present methyl groups in positions C-10 and C-13; a side chain in C-17 with a terminal

carboxylic group; one or more hydroxyl groups in position C-3, C-6, C-7, C-12 which modulate their properties, such as solubility, effect of pH on their solubility, ionization, and lipophilia [5,6,7,8,9]. As can be seen in Figure 3.1, the steroid nucleus consists of three rings with 6 carbon atoms, called A, B and C, which have a chair configuration, and a ring with 5 carbon atoms called D. The joints between the rings B/C and C/D are in *trans*, while between the rings A and B the junction is in *cis*. This feature has as primary consequence the acquisition of a β configuration of the hydrogen in position 5: bile acids are therefore steroids of the 5β series, different from other natural steroids of the 5α series with *trans* junction between the rings A and B.
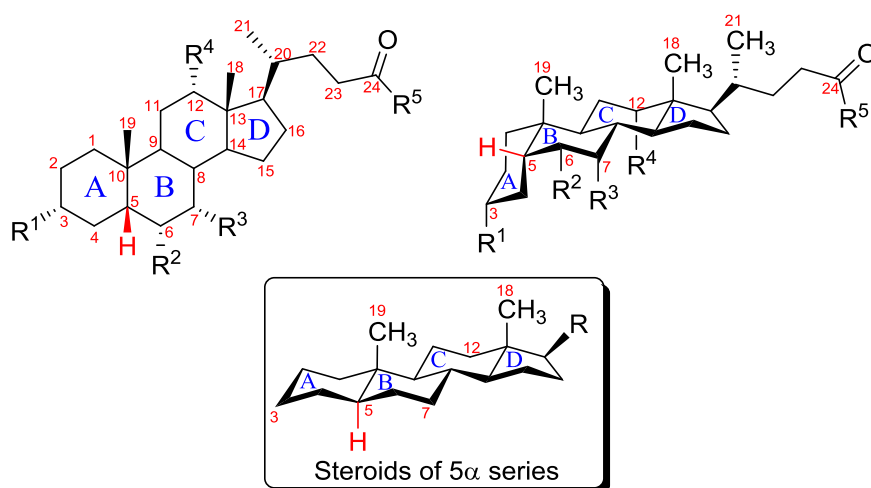
Figure 3.1 *Structure of bile acids, belonging to the 5β series, and steroids of the 5α series.*

Thanks to this particular conformation all the hydroxyls that may be present in C-3, C-6, C-7 and C-12 are in position α (downwards) respect to the ring plane (with the exception of ursodeoxycholic acid, which has the OH in C-7 in position β), while the methyl groups in C-10 and C-13 are in position β (upwards). Consequently the upper part of the bile acids is apolar while the inferior part is polar: for this reason they are amphiphilic molecules able to interface both with polar and apolar molecules.

This feature gives them properties such as detergency, or rather the ability to form micelles in aqueous solution [11]. The micelles are small clusters of molecules in which the polar groups that are outside, facing the aqueous environment, shield the apolar groups that are grouped inside. Thanks to this characteristic, bile acids are able to emulsify the lipids present in the intestine into tiny droplets and to promote their absorption. A similar function is performed in the gallbladder where, by emulsifying cholesterol, they prevent its precipitation and favor its elimination through the faeces.

A possible alteration of the biliary function is the formation of bile stones given by an excessive concentration of cholesterol inside the gallbladder. To oppose the formation of cholesterol stones or to create suitable conditions for their dissolution, ursodeoxycholic acid is used as a drug. The ursodeoxycholic acid is considered of greater pharmacological importance than the chenodeoxycholic acid in the therapy against biliary calculosis, because it possesses high efficacy and total absence of side effects.

Moreover, UDCA and its taurinated derivative TUDCA, thanks to their higher hydrophilicity compared to other bile acids, have shown anti-apoptotic effects on the proliferation of some cells, playing the role of neuroprotectors in some neurodegenerative diseases such as Alzheimer's and Parkinson's [12].

The ursodeoxycholic acid can be obtained from bear bile (a practice used in traditional Chinese medicine) or from cholic and chenodeoxycholic acids of bovine and swine origin; in the latter case, the starting bile acids can be transformed through chemical reactions (characterized by high costs and high environmental impact [13,14]) or, more conveniently, through microbiological and enzymatic reactions [3].

3.3 Bile acid derivatives

Bile acids are good carriers for poorly lipophilic molecules due to their amphiphilic character. They are able to interface with these molecules through their hydrophilic component and, at the same time, the lipophilic component is necessary to convey them through the cytoplasmic membranes, consisting of phospholipids. From this particular characteristic, combined with properties such as rigidity, chirality, modularity, good availability and low costs, the idea of synthesizing bioconjugates was born. The recommended positions for the synthesis of bioconjugates are C-3 and C-24 [15].

It is thought that the lipophilicity of these molecules is correlated with the induction of cytotoxicity and apoptosis in human tumor cells. Synthetic bile acid derivatives, especially chenodeoxycholic and ursodeoxycholic acids conjugates, significantly inhibit cell growth and induce apoptosis in various cancerous cell lines, including breast cancer, T-cell leukemia, hepatocellular carcinoma. A vast amount of bile acid conjugates with improved pharmacological profiles in terms of bioavailability and biostability is also described in the literature, thus demonstrating further potential from precisely modified bile acids, in order to act as hybrid molecules and bases for discovery and development of new drugs [16].

An example of conjugation is proposed in a study aimed at increasing the oral absorption of cytarabine through conjugation with CA, CDCA, HDCA (hyodeoxycholic acid) and UDCA performed by Zhang [17] (Scheme 3.3). Cytarabine is a nucleoside analog of cytidine, widely used in the treatment of acute and chronic myeloblastic leukemia [18]. The cytotoxic effect of cytarabine is given by its triphosphated form (ara-CTP) which interferes with the synthesis of DNA being incorporated into the DNA chain and thus inhibiting the DNA polymerase, which fails to recognize this modified nucleotide (instead of sugar deoxyribose is present arabinose). The conjugation is given by an amide bond formed between the amine function in C-4 of the cytarabine and the carboxylic group in C-24 of the bile acid.

CA: R^1=H, R^2=α−OH, R^3=α−OH
CDCA: R^1=H, R^2=α−OH, R^3=H
HDCA: R^1=α−OH, R^2=H, R^3=H
UDCA: R^1=H, R^2=β−OH, R^3=H
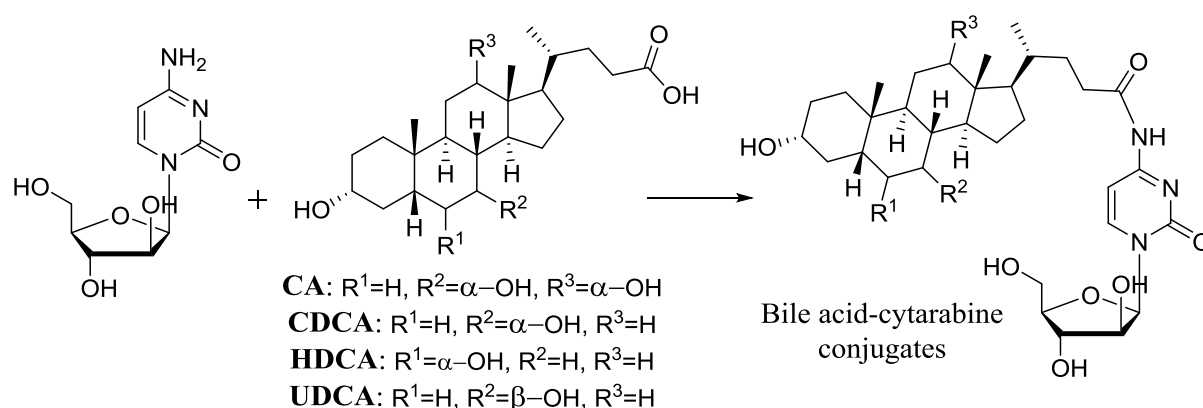
Scheme 3.3 Synthesis of bile acid-cytarabine conjugates performed by Zhang.

The pharmacokinetic study in rats found that the UDCA-cytarabine conjugate has twice the oral bioavailability compared to the cytarabine itself. This is probably due to the increased lipophilicity of the compound which favors cellular uptake through a passive transport and to

the active transport of the conjugate into the enterocytes through the ASBT (apical Na$^+$ dependent bile acid transporter) [17].

A second example is represented by a bioconjugate of Zidovudine (AZT), a nucleoside analogue-based drug, with UDCA (Figure 3.2), developed by Dalpiaz *et alter*, in order to remedy the low ability of AZT to cross the Central Nervous System or the intracellular compartments [19,20,21].
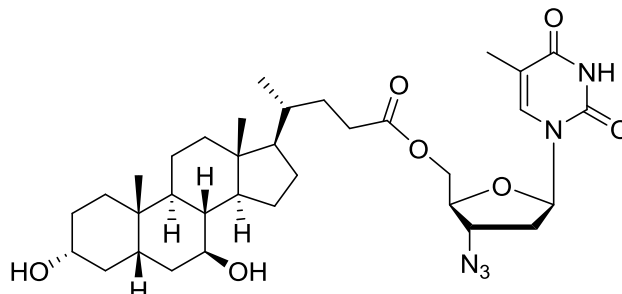
Figure 3.2 UDCA-AZT conjugate developed by Dalpiaz and co-workers.

3.4 Bibliography

1. Kuipers, F.; Bloks, V. W.; Groen, A. K. Beyond intestinal soap-bile acids in metabolic control. *Nat. Rev. Endocrinol.* **2014,** *10* (8), 488-498.

2. Carey, M. C. Physical-chemical properties of bile acids and their salts. In *Sterols and bile acids.;* Danielsson, H., Sjovall, J., Eds.; Elsevier, **1985**; Chapter 13, Vol. 12, pp 345-403.

3. Maldonado Rodriguez, M. E. *Biotrasformazioni di acidi biliari;* Ferrara, **2012**.

4. Hylemon, P. B.; Mallonee, D. H. Molecular biology of intestinal bile salt. *Chimica Oggi* **1991,** No. April, 25 – 29.

5. Lillienau, J.; Hagey, L. R.; Borgstrom, B. Hepatic and ileal transport and effect on biliary secretion of norursocholic acid and its conjugates in rats. *Am. J. Physiol.* **1991,** *261* (6 Pt 1), G1057-G1064.

6. Sturman, J. A.; Gaull, G. E. Taurine in the brain and liver of the developing human and monkey. *J. Neurochem.* **1975,** *25* (6), 831-835.

7. Hardison, W. G. Hepatic taurine concentration and dietary taurine as regulators of bile acid conjugation with taurine. *Gastroenterology* **1978,** *75* (1), 71-75.

8. Cacchi, S.; Caglioti, L.; Rosini, G. *Chimica delle sostanze naturali. Complementi di biochimica.;* Ambrosiana: Milano, **1971**.

9. Nair, P. P.; Kritchevsky, D. *The Bile Acids Chemistry, Physiology, and Metabolism.;* Plenum, **1971**; Vol. 1.

10. Hiss, D. C.; Gabriels, G. A.; Folb, P. I. Combination of tunicamycin with anticancer drugs synergistically enhances their toxicity in multidrug-resistant human ovarian cystadenocarcinoma cells. *Cancer Cell Int.* **2007,** *7:5*.

11. Eskandar, M. Absorption-Enhancing Effects of Bile Salts. *Molecules* **2015,** *20,* 14451-14473.

12. Ackerman, H. D.; Gerhard, G. S. Bile Acids in Neurodegenerative Disorders. *Frontiers in Aging Neuroscience* **2016,** *8,* 263.

13. Fieser, L. F.; Rajacopalan, S. *J. Am. Chem.* **1950,** *72,* 5530.

14. Hofmann, A. F. *Acta. Chem. Scand.* **1963,** *14,* 17-20.

15. Rais, R.; Acharya, C.; Mackerell, A.; Polli, J. E. Structural determinants for transport across the intestinal bile acid transporter using C-24 bile acid conjugates. *Mol. Pharm.* **2010,** *7,* 2240-2254.

16. Perrone, D.; Bortolini, O.; Fogagnolo, M.; Marchesi, E.; Mari, L.; Massarenti, C.; Navacchia, M. L.; Sforza, F.; Varani, K.; Capobianco, M. L. *New J. Chem.* **2013,** *37,* 3559-3567.

17. Zhang, D. Transporter-targeted cholic acid-cytarabine conjugates for improved. *International Journal of Pharmaceutics* **2016,** *511,* 161-169.

18. Rustum, Y.; Raymakers, R. 1-Beta-arabinofuranosylcytosine in therapy of leukemia: preclinical and clinical overview. *Pharmacol. Ther.* **1992,** *56,* 307-321.

19. Dalpiaz, A.; Paganetto, G.; Pavan, B.; Fogagnolo, M.; Medici, A.; Beggiato, S.; Perrone, D. Zidovudine and Ursodeoxycholic Acid Conjugation: Design of a New Prodrug Potentially Able To Bypass the Active Efflux Transport Systems of the Central Nervous System. *Molecular Pharmaceutics* **2012,** *9* (4), 957-968.

20. Dalpiaz, A.; Ferraro, L.; Perrone, D.; Leo, E.; Iannuccelli, V.; Pavan, B.; Paganetto, G.; Beggiato, S.; Scalia, S. Brain Uptake of a Zidovudine Prodrug after Nasal Administration of Solid Lipid Microparticles. *Molecular Pharmaceutics* **2014,** *11* (5), 1550-1561.

21. Dalpiaz, A.; Contado, C.; Mari, L.; Perrone, D.; Pavan, B.; Paganetto, G.; Hanuskova, M.; Vighi, E.; Leo, E. Development and characterization of PLGA nanoparticles as delivery systems of a prodrug of zidovudine obtained by its conjugation with ursodeoxycholic acid. *Drug Delivery* **2014,** *21* (3), 221-232.

4 NO photoreleaser-deoxyadenosine and -bile acid derivative bioconjugates as novel potential photochemotherapeutics

4.1 Introduction

The biological relevance of nitric oxide (NO) in many bioregulatory systems such as neurotransmission, vasodilatation, hormone secretion, and immune stimulation is nowadays well-established [1,2]. Recent studies demonstrated that NO is a very efficient antimicrobial and antioxidant agent that can be employed to tackle bacterial infections and cardiovascular diseases [3,4]. Moreover NO can be used as an anticancer agent or as a chemosensitizer enhancing the effect of traditional cancer therapies [3,4,5,6,7]. In this view, some conventional drugs may be in principle modified with a NO releasing moiety in order to produce synergistic effects due to the combined active species. Although several therapies are based on the regulation of NO synthase activity responsible for the endogenous NO concentration [7,8,9], there is an increasing interest in innovative therapies based on exogenous NO releasing agents [10]. The effect of a NO releasing drug is strongly related to the concentration and dose of NO delivered. For instance, NO concentrations in the micromolar range inhibit the growth of tumor cells, whereas picomolar NO concentrations encourage cell proliferation [11,12,13].

Strategies to control the NO release can be based on the use of an external trigger that allows the NO release under a specific stimulus (luminous, chemical, or electrochemical) [14,15]. Among these, light is the most appealing due to its peculiar advantages of not perturbing the physiological values of parameters such as temperature, pH, and ionic strength, fundamental prerequisites for biomedical applications [16]. These unique features make photoactivated NO donors (photocages, Scheme 4.1) a powerful arsenal in the burgeoning field of nanomedicine with intriguing potential to tackle cancer diseases in a noninvasive way [17,18,19,20].
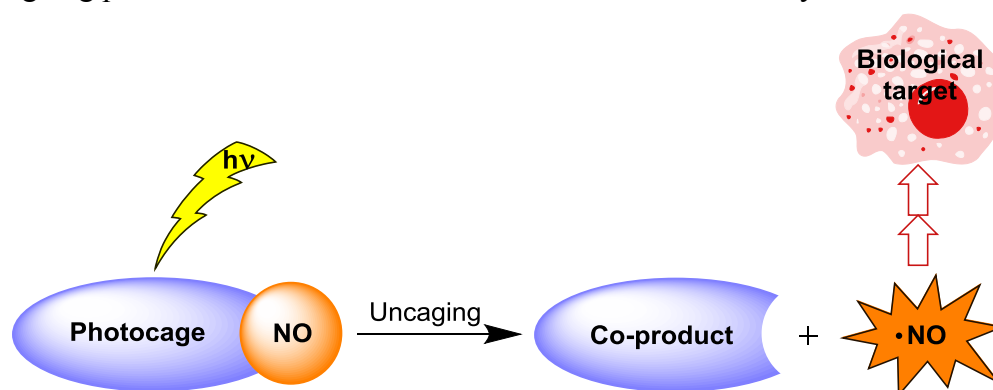
Scheme 4.1 Uncaging of NO through light excitation.

In the last years Sortino and co-workers have developed a number of molecular and supramolecular systems with anticancer activity based on the 4-nitro-3-(trifluoromethyl)aniline as suitable photocage releasing NO upon irradiation with visible light [16,17,21,22,23]. The mechanism of NO release involves a nitro-to-nitrite photorearrangement followed by the rupture of the O−NO bond with formation of NO and a phenoxy radical intermediate [21] (Scheme 4.2).
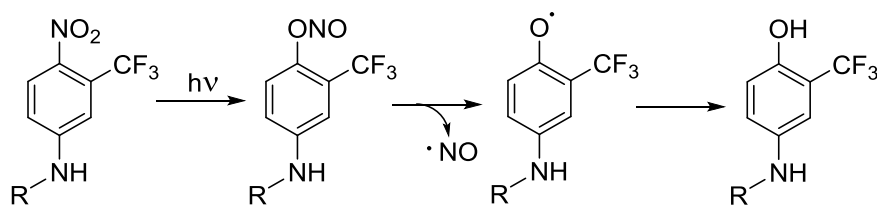
Scheme 4.2 *Mechanism of NO release of 4-nitro-3-(trifluoromethyl)aniline derivatives.*

Our interest in the study of bioconjugate functional compounds for biosensing and pharmacological applications prompted us to scout new molecules integrating a photocaged NO radical with 2′-deoxyadenosine and bile acid derivatives in the same molecular skeleton. This was motivated by the already assessed cytotoxic activity of 2′-deoxyadenosine and bile acid based compounds as well as on their chemical properties. It is well established that modified 2′-deoxyadenosines are important cytotoxic agents against lymphoid and myeloid tumors [24,25]. Bile acid derivatives, with particular regard to chenodeoxycholic (CDC) and ursodeoxycholic (UDC) bile acids, are well-known to inhibit the growth cells and to induce apoptosis in many human cancer cell lines [26,27,28,29]. From the chemical point of view, 2′-deoxyadenosine can be modified at C-8 position keeping unchanged its intrinsic characteristic of recognition of natural nucleic acids through specific hydrogen bond patterns (Watson−Crick and Hoogsteen). However, the intrinsic properties of bile acids such as rigidity, chirality, amphiphilicity, etc., can be fine-tuned. Moreover, 2′-deossiadenosine/bile acid conjugates embodying the 1,2,3-triazole ring as well as a S-alkyl unit were found to exhibit a good antiproliferative activity selectively toward both leukemic T Jurkat and K562 cells [30]. Along these lines, we adopted similar molecular skeletons but we changed one by one the 2′-deoxyadenosine and bile acid derivatives with a photocage unit.

Therefore, we designed the synthesis of some novel bioconjugates of 4-nitro-3-(trifluoromethyl)-aniline with 2′-deoxyadenosine, UDC and CDC bile acid moieties (namely, **photocage-SdAdo**, **photocage-UDC** and **photocage-CDC**) that can be considered as progenitors of large variety of new bioconjugates and, combining the intrinsic characteristics of the biological components with that of NO, could be powerful tools in cancer therapy. Photochemical studies on the effective release of NO have been performed. A biological screening for the cytotoxic effects toward a selection of two cancer cell lines, leukemic K562 and colon carcinoma HCT116, as well as the normal fibroblast skin cells, and subsequently a photobiological study on **photocage-SdAdo** toward the HCT116 cancer cell line were also conducted.

4.2   Results and discussion

4.2.1   Synthesis of photocage conjugates

Recently was reported a new methodology for the preparation of 2′-deoxyadenosine modified at C-8 position enabling the conjugation with a variety of functional molecules such as chromophores, fluorophores, intercalators, alkylating agents, and functional groups for the click chemistry [31,32]. The 2′-deoxyadenosine derivatives synthesized have been demonstrated to be quite powerful derivatives for further applications [30,31,33]. Following the procedure previously described [31,32], the conjugate **photocage-SdAdo** was prepared by reaction between the commercial 8-bromo-2′-deoxyadenosine **1** and 5-(4-nitro-3-trifluoromethyl)phenylamino-pentane-1-thiol **2** through a reaction in water mediated by

triethylamine (TEA) (Scheme 4.3). The reaction was performed in a sealed tube at 100 °C for 2 h. A simple and very efficient work up based on the extraction of the warm reaction mixture with ethyl acetate led to the pure target bioconjugate photocage-SdAdo in 20% yield.
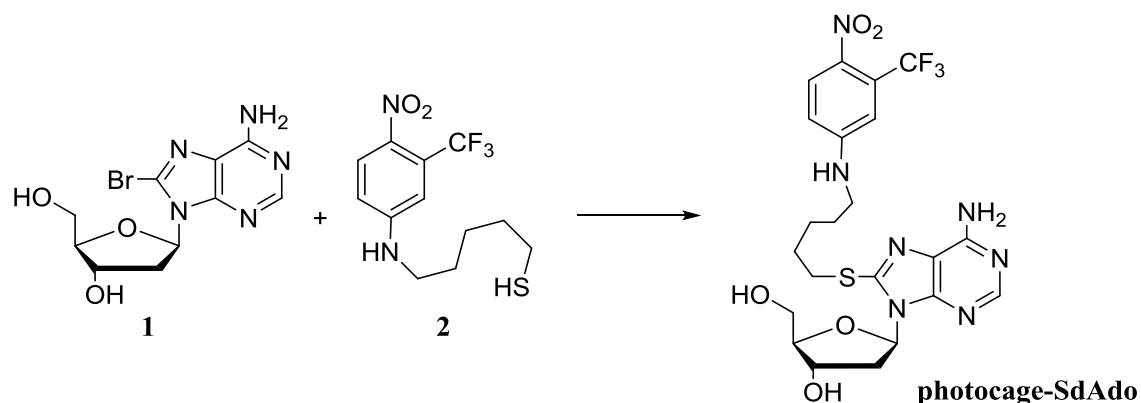
Scheme 4.3 Synthesis of photocage-SdAdo. Reagents and conditions: H_2O, TEA, 100 °C, 2 h (20% yield).

The synthesis of **photocage-CDC** and **photocage-UDC** was performed *via* a Cu(I) mediated 1,3-dipolar cycloaddition reaction starting from the appropriate 3-azido bile acid **3a,b** with the alkyne **4**[30] (Scheme 4.4). In both cases the reaction performed by reacting under microwave irradiation at 80 °C for 30 min, a 1:1.5:0.4:2 molar ratio of the alkyne **4**, the appropriate azide **3a,b**, CuSO$_4$·5H$_2$O, and sodium ascorbate, respectively, in THF/tBuOH/H$_2$O mixture (1.5:1:1, *v/v*) provided the target compound in 70−80% yield after chromatographic purification.

3a N$_3$-CDCA (7α-OH)
3b N$_3$-UDCA (7β-OH)

4

photocage-CDC (7α-OH)
photocage-UDC (7β-OH)
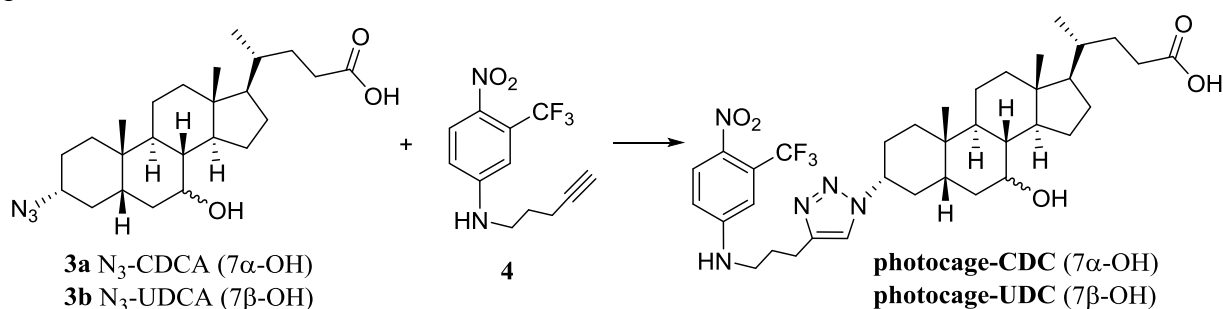
Scheme 4.4 Synthesis of photocage-CDC and –UDC. Reagents and conditions: CuSO$_4$·5H$_2$O, sodium ascorbate, THF/tBuOH/H$_2$O (1.5:1:1), microwaves: 80 °C, 30 min (70−80% yield).

The syntheses of the compounds **2** and **4** are shown in Scheme 4.5. The thiol **2** was prepared through a multistep synthesis, starting from commercial 4-chloro-1-nitro-2-trifluoromethyl-benzene **5** in 34% overall yield. Starting from the same nitrobenzene derivative **5** the terminal alkyne **4** was obtained in good yield.
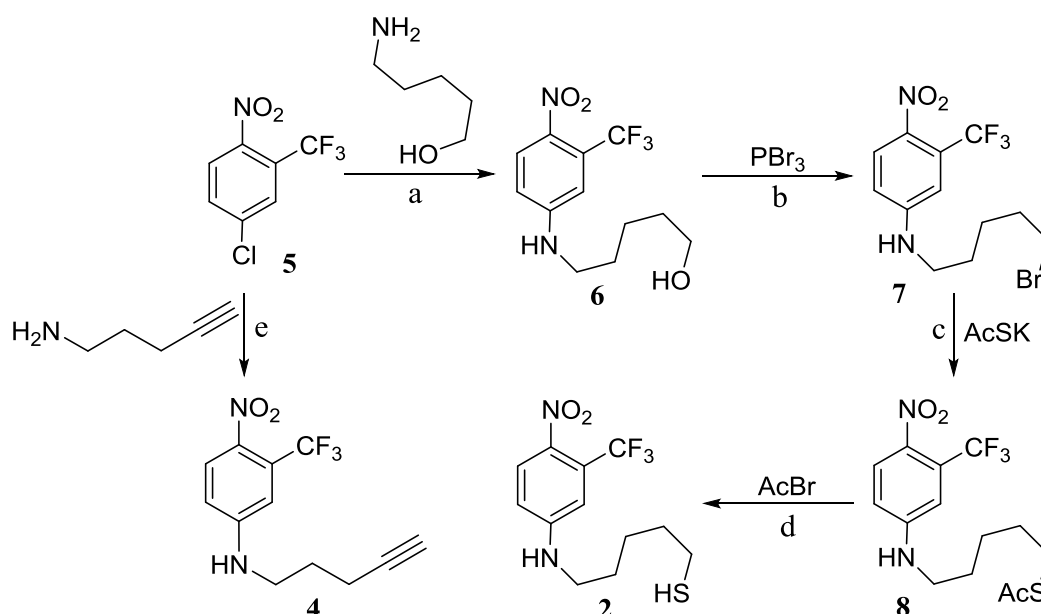
Scheme 4.5 Syntheses of thiol **2** and alkine **4**. Reagents and conditions: (a) DMSO, K_2CO_3, 85°C, 24 h (90% yield); (b) DCM, 20°C, 16 h (50% yield); (c) DMSO, 50°C, 16 h (80% yield); (d) MeOH, −78°C, 4 h (95% yield); (e) DMSO, K_2CO_3, 85°C, 24 h (65% yield).

4.2.2 Biological assays

All the biological trials were performed first of all in the dark. All bioconjugates were tested *in vitro* against a selection of two human cancer cell lines such as K562 leukemia cells and the colon cancer cell line HCT116 and normal human skin fibroblast cells as a control. Cisplatin was chosen as a reference compound. The cytotoxicity was evaluated using MTT assay.

Cell growth inhibition was determined at concentrations of 10, 25, and 50 μM for the cancer cells and up to 100 μM for fibroblast cells after 72 h. No cytotoxicity was observed with both **photocage-SdAdo** and **photocage-UDC** bioconjugates on the fibroblast cells since the percent cell growth inhibition was found under the detection limit up to 100 μM (Table 4.1). In contrast, a significant concentration-dependent antiproliferative effect was found with both conjugates against the K562 and HCT116 cancer cells lines (Table 4.1) with **photocage-UDC** exhibiting a higher cytotoxic effect than **photocage-SdAdo** and some cytoselectivity. In the case of **photocage-CDC**, IC_{50} values similar to **photocage-UDC** against K562 and HTC116 cancer cells were found. However, the IC_{50} value of 32.0 μM against fibroblast cells prompted us to exclude this compound from further studies. Based on these results we evaluated the **photocage-SdAdo** conjugate, with lower, but still good, cytotoxic activity with respect to **photocage-UDC**, the best candidate to highlight the cytotoxic effect of the light, resulting in the production of NO, toward the HTC116 cancer cells.

	compound	IC_{50} (μM)		
		K562	HCT116	fibroblast
dark	photocage-CDC	24.0 ± 1.2	21.0 ± 0.8	32.0 ± 1.0
	photocage-UDC	31.6 ± 1.6	14.0 ± 1.0	≫100
	photocage-SdAdo	68.1 ± 1.3	66.2 ± 3.3	≫100
light	photocage-SdAdo		31.0 ± 2.5	
	cisplatin	5.40 ± 1.0	8.5 ± 1.2	25.4 ± 3.5

Table 4.1 IC_{50} values determined from the dose−response curves using MTT assay after 72 h incubation time in the dark and also upon irradiation in the case of photocage-SdAdo. Results are expressed as the mean of three independent experiments ±SD. Cisplatin was used as a reference compound.

Figure 4.1 shows the cytotoxicity observed for different concentrations of the conjugate **photogage-SdAdo** at 72 h in the case of samples irradiated for 40 min with visible light and, for comparison, kept in the dark. The data obtained show an increase of the growth inhibition at all concentrations. In particular, the percent of growth inhibition observed for 50 μM solution of **photocage-SdAdo** upon irradiation was around double that observed in the dark.
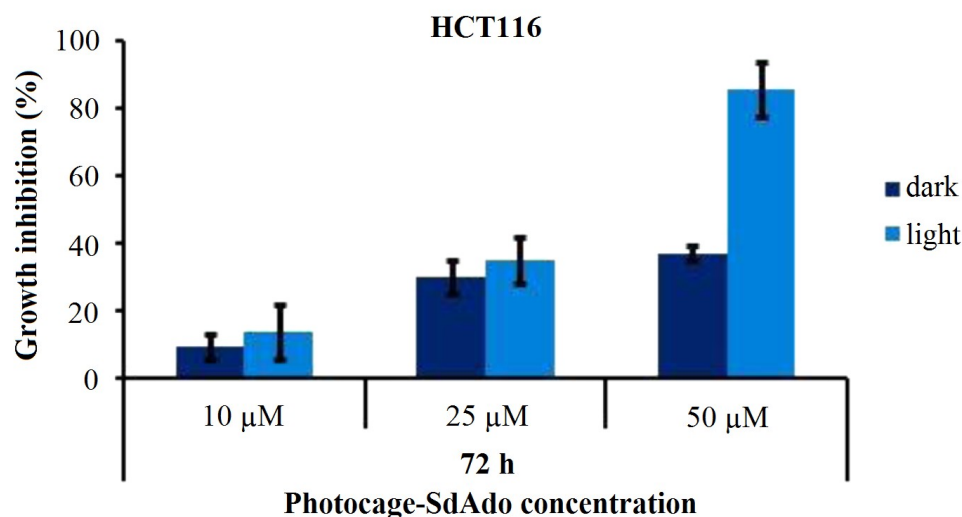
Figure 4.1 *Antiproliferative effect of photocage-SdAdo on HCT116 cell line observed at 72 h after 40 min light irradiation (λ = 420 nm, 7 mW cm⁻²) and, for comparison, in the dark* [34].

4.2.3 Photochemical experiments

The photochemical experiments were focused on **photocage-SdAdo** and **photocage-UDC**. **Photocage-CDC** was not studied in consideration of its high cytotoxicity against normal fibroblast cells at low micromolar concentrations (see Table 4.1). Figure 4.2 shows the absorption spectra of **photocage-SdAdo** and **photocage-UDC**. As expected, the nitroaniline chromogenic unit strongly dominates the absorption in the visible range, whereas the UV region is characterized by the typical absorption band of the deoxyadenosine fragment in the case of **photocage-SdAdo**.
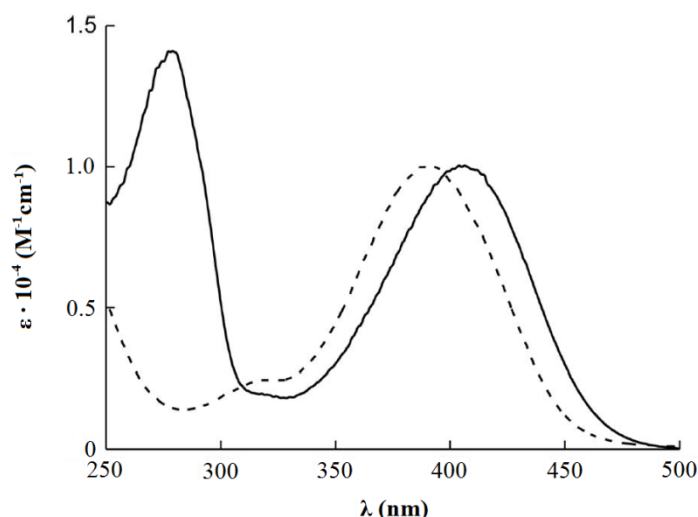
Figure 4.2 *Absorption spectra of photocage-SdAdo (solid) and photocage-UDC (dotted) in phosphate buffer (10 mM, pH = 7.4) with DMSO 0.5%* [34].

The molar absorbivity of the visible band for both conjugates is ca. 10.000 M^{-1} cm^{-1}. This value is basically the same to that already reported for the NO photocage alone [21,35], ruling out any significant interaction, in the ground state, between the nitroaniline chromophores and

the other functional molecular components in both **photocage-SdAdo** and **photocage-UDC**. To exploit the cytotoxicity effects of NO along with those of 2′-deoxyadenosines and bile acids, the NO photoreleasing capability of the NO photocage needs to be preserved after its covalent conjugation. Note that, in contrast to nonphotoresponsive compounds, the preservation of the photobehavior of a photoactive center after its covalent linking with other molecular components is not a "trivial result". In some cases, the response to light of the photoactive unit can be considerably influenced, in both nature and efficiency, by the occurrence of competitive photoprocesses (i.e., photoinduced energy and/or electron transfer, nonradiative deactivation, etc.), occurring upon light absorption. The most convenient methodology to demonstrate the NO generation from the molecular conjugates under visible light stimuli is the direct and in real-time monitoring of this radical species. To this end, an ultrasensitive NO electrode was used, which directly detects NO, with nanomolar concentration sensitivity, by an amperometric technique [36].

The results illustrated in Figure 4.3 provide unambiguous evidence that both **photocage-SdAdo** and **photocage-UDC** are stable in the dark but supply NO upon illumination with visible light.
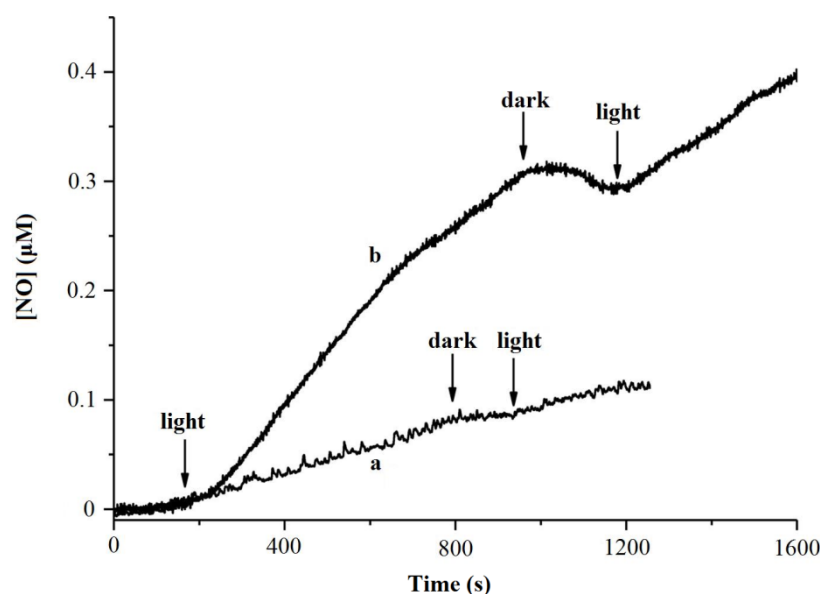
Figure 4.3 *NO release profiles observed upon visible light irradiation (λ = 420 nm, 1.5 mW cm⁻²) of solutions of photocage-SdAdo (a) and photocage-UDC (b) (50 μM) in phosphate buffer (10 mM, pH = 7.4) with DMSO 0.5% [34].*

The release process is strictly regulated by the external light inputs, as confirmed by the linear NO generation, which promptly stops as the light turns off and restarts as the illumination turns on again. From the slopes of the photoamperograms obtained, one can note that the NO photorelease by **photocage-UDC** is almost 30% larger than that of **photocage-SdAdo**.

In principle, one can tentatively attribute this different photoreactivity to side reactions occurring in the case of **photocage-SdAdo**, competitive with NO photorelease. However, this is not the case. In fact, the absorption spectral changes of **photocage-SdAdo** as a function of the irradiation time show a photobleaching of the visible band without any significant shift in the absorption maximum (Figure 4.4).
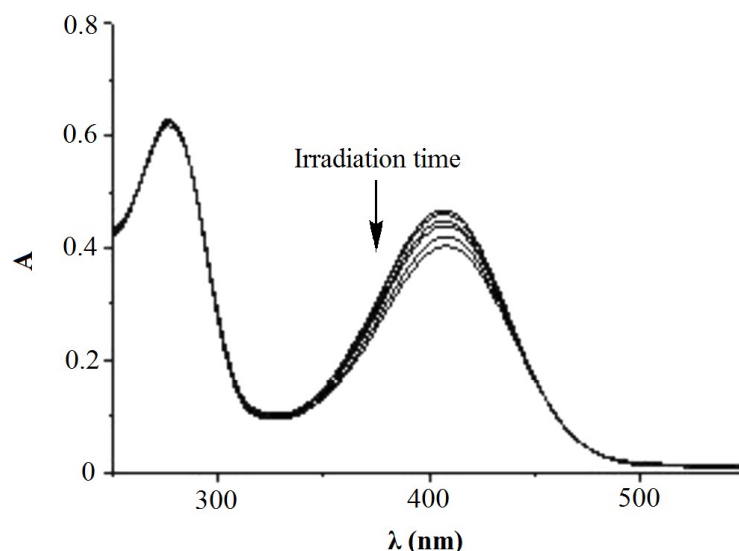
Figure 4.4 Absorption spectral changes observed upon visible light irradiation (λ = 420 nm, 7 mW cm⁻²) of an aqueous solution of photocage-SdAdo (50 μM) in phosphate buffer (10 mM, pH = 7.4) with DMSO 0.5% from 0 to 70 min [34].

This spectral behavior is in excellent agreement with the photochemical pathway leading to the NO release previously proposed in the case of the single NO photodonor unit [21]. Comparative photolysis experiments carried out with an optically matched solution of **photocage-UDC** (Figure 4.5) show that the kinetics of the photobleaching in the case of **photocage-UDC** is faster than that observed for **photocage-SdAdo**, in good agreement with the NO photoreleasing behavior. On the basis of these results, the higher photochemical reactivity of **photocage-UDC** could be probably due to an active role of the bile acid substituent in the pathways following the NO release and leading to the steady state coproduct *via* the intermediate phenoxy radical [21] (*i.e.*, intramolecular H-abstraction).
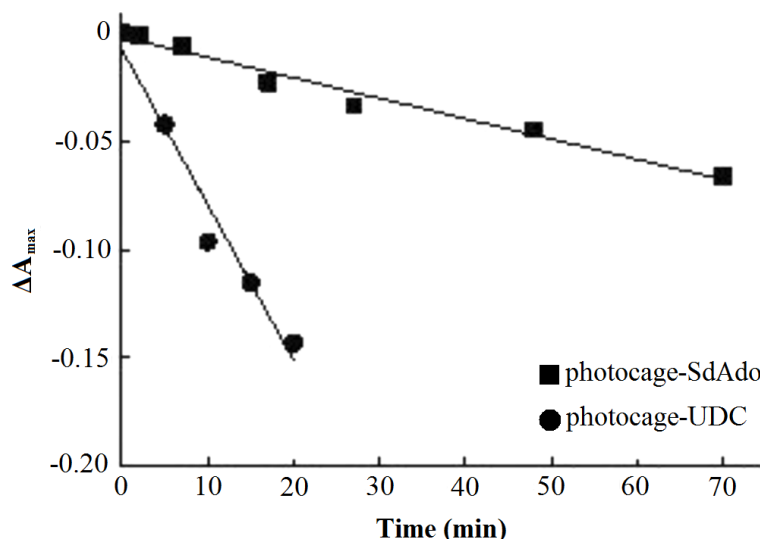
Figure 4.5 Kinetic profiles, monitored at λ_{max}, and the related linear fittings, observed for 50 μM solutions of photocage-SdAdo and photocage-UDC [34].

4.3   Conclusion

We have demonstrated that the covalent linking of a NO photoreleaser with modified 2′-deoxyadenosine as well as CDC and UDC bile acid derivatives leads to molecular conjugates displaying an interesting cytotoxicity toward both leukemia K562 and colon carcinoma

HCT116 cancer cells. In particular, **photocage-SdAdo** and **photocage-UDC**, with no significant cytotoxicity to the human fibroblast skin cells and good antiproliferative effect against both K562 and HCT116 cells, were found to preserve the NO photoreleasing capability of the NO photodonor after its covalent conjugation. In fact, although with different efficiency, both the synthesized compounds exhibited NO delivery exclusively regulated by visible light inputs. This feature offers the possibility to enhance the cytotoxic effects of the conjugates upon light irradiation, combining the chemotherapeutic action with the cytotoxic activity of the photogenerated NO. This was clearly demonstrated in the case of **photocage-SdAdo** for which, upon illumination, the IC_{50} value after 72 h of incubation was found more than halved, decreasing from 66.2 μM to 31.0 μM. This result, being the first evidence *in vitro* of combined chemo- and phototherapeutic activities in the same molecule, represents an intriguing starting point toward the development of a larger number of molecular hybrids [34].

4.4 Experimental

4.4.1 General information

The reactions for the synthesis of **photocage-CDC** and **photocage-UDC** were monitored by TLC on pre-coated Silica Gel plates (thickness 0.25 mm, Merck), and phosphomolybdicacid solution was used as the spray reagent to visualize the steroids. The reactions for the synthesis of intermediates **2**, **4**, **6**, **7**, **8** and of **photocage-SdAdo** were monitored by HPC-MS. Flash column chromatography was performed on silica gel 60 (230-400 mesh) or with a combiflash apparatus. The microwave (MW) irradiation was performed using a *Biotage Initiator* apparatus. Optimization experiments were performed in the 'single-run' mode, i.e., by manual filling of reaction vials and by specifying the irradiation time and maximum temperature. Melting points were determined using a capillary apparatus. HPLC-MS analyses were performed on a *Agilent 1260* with a diode array detector using a *Zorbax C8* column (4.6 × 150 mm, 5 μm) with a linear gradient water/acenitrile at a 0.5 mL/min flow rate, detection at λ = 260 nm and an *Esquire 3000 Plus Bruker* mass spectrometer. ESI-HRMS were acquired on an *Agilent Dual ESI Q TOF 6520* using methanol. NMR spectra were recorded with a *Varian Mercury 400 MHz* instrument. UV/vis absorption and fluorescence spectra were recorded with a *Jasco V 650* spectrophotometer. Irradiation experiments were performed in a quartz cell (1 cm path length, 3 mL capacity) with RPR lamps with emission centered at λ = 420 nm in a *Rayonet* photochemical reactor.

NO release was measured with a *World Precision Instrument*, ISO-NO meter, equipped with a data acquisition system, and based on direct amperometric detection of NO with short response time (< 5 s) and sensitivity range 1 nM–20 μM. The analog signal was digitalized with a four-channel recording system and transferred to a computer. The sensor was accurately calibrated by mixing standard solutions of $NaNO_2$ with 0.1 M H_2SO_4 and 0.1 M KI according to the reaction:

$$4H^+ + 2I^- + 2NO_2^- \rightarrow 2H_2O + 2NO + I_2$$
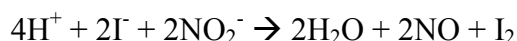
Irradiation was in a quartz cell (1 cm path length, 3 mL capacity); NO measurements were carried out under stirring with the electrode positioned outside the light path, to avoid NO signal artefacts due to photoelectric interference on the ISO-NO electrode.

The intermediates **2**, **4**, **6**, **7** and **8** were characterized by LC- MS, ^{1}H-NMR and ^{13}C-NMR. In the case of bioconjugates **photocage-SdAdo**, **photocage-CDC** and **photocage-UDC** also HRMS was performed.

4.4.2 Derivatization of photocage
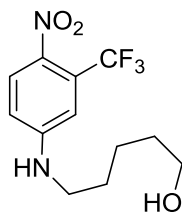
5-(4-Nitro-3-trifluoromethyl-phenylamino)pentan-1-olo (6)

5-Amino-pentan-1-olo (2 mmol, 206 mg) and K_2CO_3 (150 mg) were added to a solution of 4-chloro-1-nitro-2-trifluoromethyl benzene **5** (1 mmol, 225 mg, 0.15 mL) in dry DMSO (10 mL) in a sealed tube. The mixture was stirred at 85°C for 24 h. The reaction mixture was extracted with ethyl acetate (2 x 15 mL) and washed with water/brine 1:1 solution (5 x 25 mL). The organic phase was dried over sodium sulfate and evaporated to dryness under reduced pressure. The residue was purified on silica gel by gradient elution from petroleum ether 100% to petroleum ether/ethyl acetate 40/60 to obtain the target compound (0.6 mmol, 175 mg, 60% yield). ^{1}H-NMR (CDCl$_3$): δ= 7.98 (1H, d, J = 5.5 Hz), 6.87 (1H, d, J = 1.5 Hz), 6.64 (1H, dd, J$_1$ = 1.5 Hz, J$_2$ = 5.5 Hz), 3.67 (2H, J = 4.0 Hz), 3.22 (2H, t, J = 4.5 Hz), 1.70 (2H, m), 1.62 (2H, m), 1.50 (2H, m). ^{13}C-NMR (CDCl$_3$, selected data): δ = 152 (q), 129 (CH), 124 (q), 121 (q), 113 (CH), 112 (CH), 63 (CH$_2$), 44 (CH$_2$), 32 (CH$_2$), 29 (CH$_2$), 23 (CH$_2$). MS (ESI, ES+) *m/z*: calculated for C$_{12}$H$_{15}$F$_3$N$_2$O$_3$ 292.26; found 293 [M+H]$^+$, 315 [M+Na]$^+$, 331 [M+K]$^+$.
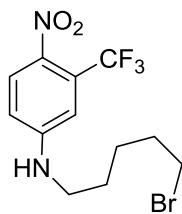
(5-Bromo-pentyl)-(4-nitro-3-trifluoromethyl-phenyl) amine (7)

Compound **6** (1 mmol, 290 mg) was dissolved in fluorobenzene (10 mL) and PBr$_3$ 1 M solution in dichloromethane was added dropwise (3 eq, 3 mL). The reaction mixture was left under stirring overnight and then evaporated to dryness under reduced pressure. The residue was suspended in water and NaOH 0.2 M was added up to pH = 12. The water solution was extracted with ethyl acetate (2 x 15 mL) and the organic layer washed with water/brine 1:1 solution (2 x 10 mL). The organic phase was dried over sodium sulfate and evaporated to dryness. The residue was purified on silica gel by gradient elution from petroleum ether 100% to petroleum ether/ethyl acetate 80/20 to obtain the target compound (0.5 mmol, 180 mg, 50% yield). ^{1}H-NMR (CDCl$_3$): δ= 7.98 (1H, d, J = 9.2 Hz), 6.87 (1H, d, J = 2.5 Hz), 6.68 (1H, dd, J$_1$ = 2.5 Hz, J$_2$ = 8.8 Hz), 3.43 (2H, J = 6.5 Hz), 3.25 (2H, t, J = 7.0 Hz), 1.92 (2H, m), 1.71 (2H, m), 1.60 (2H, m). ^{13}C-NMR (CDCl$_3$): δ = 152 (q), 129 (CH), 124 (q), 121 (q), 113 (CH), 112 (CH), 44 (CH$_2$), 33 (CH$_2$), 32 (CH$_2$), 28 (CH$_2$), 26 (CH$_2$). MS (ESI, ES+) *m/z*: calculated for C$_{12}$H$_{14}$BrF$_3$N$_2$O$_2$ 355.16; found 355.36 [M+H]$^+$.
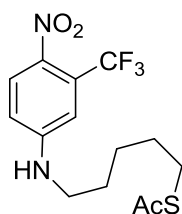
S-[5-(4-nitro-3-trifluoromethyl-phenylamine) penthyl] ester (8)

Compound **7** (1 mmol, 355 mg) was dissolved in dry DMF (10 mL) in a sealed tube. Potassium thioacetate (1.3 mmol, 150 mg) was added and the mixture was stirred at 50°C for 4 h. The reaction mixture was extracted with ethyl acetate (2 x 15 mL) and washed with water/brine 1:1 solution (5 x 25 mL). The organic phase was dried over sodium sulfate and evaporated to dryness under reduced pressure. The residue was purified on silica gel by gradient elution from petroleum ether 100% to petroleum ether/ethyl acetate 85/15 to obtain the target compound (0.8 mmol, 280 mg, 80% yield). ^{1}H-NMR (CDCl$_3$): δ= 7.98 (1H, d, J = 9.0 Hz), 6.88 (1H, d, J = 2.8 Hz), 6.65 (1H, dd, J$_1$ = 2.8 Hz, J$_2$ = 9.0 Hz), 3.21 (2H, t, J = 6.8 Hz), 2.89 (2H, t, J = 7.2 Hz), 2.34 (3H, s), 1.66 (4H, m), 1.71 (2H, m), 1.48 (2H, m). ^{13}C-NMR (CDCl$_3$): δ= 196 (q, C=O), 152 (q), 129 (CH), 124 (q), 121 (q), 113 (CH), 111 (CH), 44 (CH$_3$), 31 (CH$_2$), 30 (CH$_2$), 29 (CH$_2$),

28 (CH_2), 26 (CH_2). MS (ESI, ES+) *m/z*: calculated for $C_{14}H_{17}F_3N_2O_3S$ 350.36; found 351 $[M+H]^+$, 373 $[M+Na]^+$.
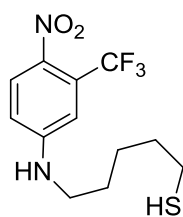
5-(4-nitro-3-trifluoromethyl-phenylamine)-pentane-1-thiol (2)

Compound **8** (1 mmol, 350 mg) was dissolved in CH_3OH (50 mL) and cooled down to -78°C. CH_3COBr (20 mmol, 2 mL) was added dropwise. The reaction mixture was warmed up at room temperature and left under stirring for 4 h. The solvent was evaporated to dryness under reduced pressure and the residue was solved in ethyl acetate and washed with water in order to reach pH = 7. The organic layer was dried over sodium sulfate and evaporated to dryness. The residue was used without further purification (0.95 mmol, 330 mg, 95% yield). ^{1}H-NMR ($CDCl_3$): δ= 8.00 (1H, d, J = 5.5 Hz), 6.88 (1H, d, J = 1.8 Hz), 6.65 (1H, dd, J_1 = 1.8 Hz, J_2 = 5.5 Hz), 3.23 (2H, J = 4.2 Hz), 2.56 (2H, q, J_1 = 4.5 Hz, J_2 = 5.0 Hz), 1.68 (4H, m), 1.53 (2H, m), 1.35 (1H, t, J = 5.0 Hz; disappeared upon D_2O shake; SH signal). ^{13}C-NMR ($CDCl_3$): δ= 152 (q), 129 (CH), 124 (q), 121 (q), 113 (CH), 111 (CH), 44 (CH_2), 33 (CH_2), 29 (CH_2), 26 (CH_2), 25 (CH_2). MS (ESI, ES+) *m/z*: calculated for $C_{12}H_{15}F_3N_2O_2S$ 308.32; found 309 $[M+H]^+$.
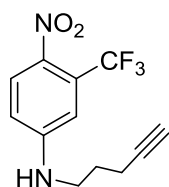
(4-nitro-3-trifluoromethyl-phenyl)-pent-4-inyl-amine (4)

4-Pentyn-1-amine (2 mmol, 206 mg) and K_2CO_3 (150 mg) were added to a solution of 4-chloro-1-nitro-2-trifluoromethyl benzene **5** (1 mmol, 225 mg, 0.15 mL) in dry DMSO (10 mL) in a sealed tube. The mixture was stirred at 85°C for 24 h. The reaction mixture was extracted with ethyl acetate (2 x 15 mL) and washed with water/brine 1:1 solution (5 x 25 mL). The organic phase was dried over sodium sulfate and evaporated to dryness under reduced pressure. The residue was purified on silica gel by gradient elution from petroleum ether 100% to petroleum ether/ethyl acetate 80/20 to obtain the target compound (0.65 mmol, 340 mg, 65% yield). ^{1}H-NMR ($CDCl_3$): δ = 8.00 (1H, d, J = 9.2 Hz), 6.90 (1H, d, J = 2.8 Hz), 6.68 (1H, dd, J_1 = 2.8 Hz, J_2 = 9.2 Hz), 3.40 (2H, t, J = 6.8 Hz; collapsing to s upon irradiation at δ = 1.87), 2.35 (2H, dt, J_d = 2.8 Hz, J_t = 6.8 Hz; collapsing to d J = 2.8 Hz upon irradiation at δ = 1.87), 2.05 (1H, t, J = 2.8 Hz), 1.86 (2H, m), 1.71 (2H, m), 1.48 (2H, m). ^{13}C-NMR ($CDCl_3$): δ = 152 (q), 129 (CH), 127 (q), 124 (q), 121 (q), 113 (CH), 83 (q), 70 (CH), 43 (CH_2), 27 (CH_2), 16 (CH_2). MS (ESI, ES+) *m/z*: calculated for $C_{12}H_{11}F_3N_2O_2$ 272.23; found 273 $[M+H]^+$.

4.4.3 Syntheses of conjugates
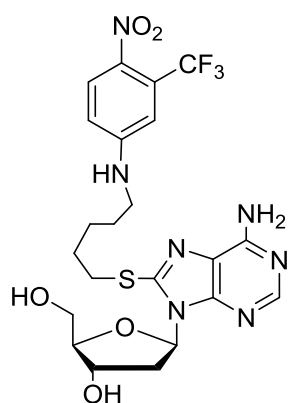
Photocage-SdAdo

5-(4-Nitro-3-trifluoromethyl)phenylamino-pentane-1-thiol **2** (3 mmol, 920 mg) and triethylamine (1.5 mL, 10 mmol) were added to a 1.6 mM suspension of commercial 8-bromo-2′-deoxyadenosine **1** (330 mg, 1.0 mmol) in water. The resulting solution was heated at 100°C for 2 h. The warm reaction mixture was extracted with ethylacetate (2 x 100 mL) and the solvent evaporated under reduced pressure. The target compound was obtained in 20% yield. ^{1}H-NMR (CD_3OD): δ= 1.59 (2H, m), 1.69 (2H, m), 1.84 (2H, m), 2.21 (1H,

ddd, J_1 = 1.6 Hz, J_2 = 5.6 Hz, J_3 = 13.0 Hz, collapsing to dd J = 5.6 Hz, J = 13 Hz upon irradiation at δ 4.60; collapsing to dd J = 1.6 Hz, J = 13.0 Hz upon irradiation at δ 6.38;), 2.99 (1H, ddd, J_1 = 5.6 Hz, J_2 = 9.0 Hz, J_3 = 13.0 Hz, collapsing to dd J = 5.6 Hz, J = 13.0 Hz upon irradiation at δ 4.60; collapsing to dd J = 9.0 Hz, J = 13.0 Hz upon irradiation at δ 6.38), 3.20 (2H, t, J = 6.8 Hz), 3.36 (2H, t, J = 8.4 Hz), 3.80 (2H, ABX system, J_{AB} = 13.0 Hz, J_{AX} = 3.0 Hz, collapsing to AB system upon irradiation at δ 4.08), 4.08 (1H, m, collapsing to t J = 3.0 Hz, upon irradiation at δ 4.60), 4.60 (1H, m, collapsing to dd, J_1 = 1.6 Hz, J_2 = 6.0 Hz, upon irradiation to δ 4.08), 6.38 (1H, dd, J_1 = J_2 = 5.6), 6.70 (1H, dd, J_1 = 3.0 Hz, J_2 = 9.0 HZ), 6.94 (1H, d, J = 3.0 Hz), 7.96 (1H, d, J = 9.0 Hz), 8.05 (1H, s). ^{13}C-NMR (CD$_3$OD): δ = 154 (q), 153 (q), 151 (CH), 150 (q), 121 (q), 120 (q), 151 (CH), 129 (CH), 89 (CH), 86 (CH), 72 (CH), 63 (CH$_2$), 42 (CH$_2$), 39 (CH$_2$), 32 (CH$_2$), 29 (CH$_2$), 28 (CH$_2$), 26 (CH$_2$). HRMS calculated for [C$_{22}$H$_{25}$F$_3$N$_7$O$_5$S + H]$^+$ 558.1668; found 558.1667.

4.4.3.1 General procedure for the "click" reaction

To a solution of the appropriate alkyne (0.03 mmol) in 1.4 mL of a 1:1:1.5 mixture of H$_2$O/*t*-BuOH/THF (v/v), sodium ascorbate (0.06 mmol), CuSO$_4$·5H$_2$O (0.012 mmol), and the proper azide **3a,b** (0.04 mmol; prepared according to the literature procedures [30]) was added. The resulting mixture was premixed for 30 s, then heated in a sealed glass tube in a *Biotage Initiator* microwave apparatus at 80°C for 30 min. After cooling at room temperature, solvents were removed *in vacuo* and the crude material was purified by flash chromatography on silica gel with cyclohexane/EtOAc 2:1 and AcOH 0.1%, as an eluent.
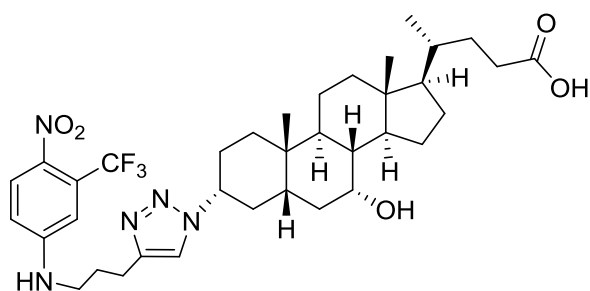
Photocage-CDC
Yellow amorphous solid, yield 68%. ^{1}H-NMR (CD$_3$OD): δ= 8.01 (1H, d, J = 9.1 Hz), 7.82 (1H, s), 6.96 (1H, d, J = 2.5 Hz), 6.72 (1H, d, J = 9.2 Hz), 4.41-4.29 (1H, m), 3.80 (1H, d, J = 2.5 Hz), 3.25 (2H, t, J = 6.9 Hz), 2.84-2-72 (3H, m), 2.39-2.27(1H, m), 2.24-2.14 (1H, m), 2.06-1.05 (32H, m), 1.00 (3H, s), 0.95 (3H, d, J = 6.5 Hz), 0.69 (3H, s). ^{13}C-NMR (CD$_3$OD): δ = 178.08, 154,32, 147.77, 136.13, 127.10 (q, CF$_3$) 126.13, 125.20, 122.20, 121.10, 112.67, 112.17, 68.87, 62.67, 57.34, 51.54, 43.57, 43.41, 43.17, 40.91, 40.68, 38.16, 36.85, 36.76, 36.33, 35.60, 34.09, 32.35, 32.07, 29.40, 29.24, 24.55, 23.72, 23.28, 21.78, 18.79, 12.19. HRMS calculated for [C$_{36}$H$_{51}$F$_3$N$_5$O$_5$ – H]$^-$ 688.3764; found 688.3767.
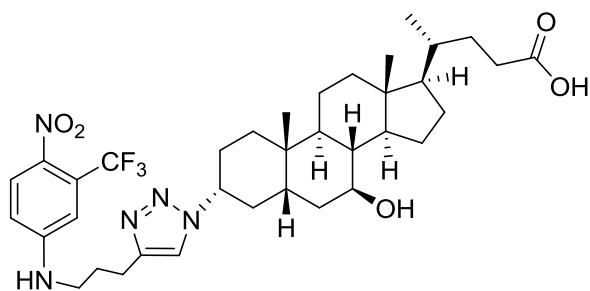
Photocage-UDC
Yellow amorphous solid, yield 66%. ^{1}H-NMR (DMSO-d_6): δ = 11.50 (1H, bs), 8.05 (1H, d, J = 9.5 Hz), 7.99 (1H, s, 5-H triazole), 7.63 (1H, t, J = 5.0 Hz), 7.03 (1H, bs), 6.79 (1H, dd, J = 9.5, 2.0 Hz), 4.43-4.32 (1H, m), 4.90 (1H, bs, D$_2$O ex), 3.40-3.27 (2H, m), 3.23 (2H, dt, J = 7.5, 5.0 Hz), 2.69 (2H, t, J = 7.5 Hz), 2.35-2.15 (1H, m), 2.13-2.01 (2H, m), 1.93-0.95 (24H, m), 0.92 (3H, s), 0.86 (3H, d, J = 6.5 Hz), 0,60 (3H, s). ^{13}C-NMR (DMSO-d_6): δ = 174.9, 153.2, 145.8, 133.4, 129.8, 126.6, 124.9 (q, CF$_3$), 123.9, 121.1, 119.9, 69.1, 59.5, 55.4, 54.7, 43.0, 42.5, 41.8, 39.5, 38.4, 37.3, 35.0, 34.8, 34.1, 33.8, 30.73, 30.72, 28.1, 27.8, 27.7,

26.6, 23.1, 22.6, 20.9, 18.3, 12.0. HRMS calculated for $[C_{36}H_{51}F_3N_5O_5 - H]^-$ 688.3764; found 688.3762.

4.4.4 Biological assay

Cell growth inhibition assays were carried out using the leukemia cell line K562 and colon carcinoma HCT116. Cell lines were obtained from ATCC (Manassas, VA) and maintained in RPMI 1640, supplemented with 10% fetal bovine serum (FBS), penicillin (100 Units mL^{-1}), streptomycin (100 µg mL^{-1}) and glutamine (2mM) (complete medium); the pH of the medium was 7.2 and the incubation was performed at 37°C in a 5% CO_2 atmosphere. Adherent cells were routinely used at 70% of confluence and passaged every 3 days by treatment with 0.05% Trypsin-EDTA (Lonza). K562 cells were routinely fed every 3 days. The antiproliferative activity of the compounds was tested with 3-(4,5-dimethylthiozol-2-yl)2,5-diphenyltetrazolium bromide solution (MTT) assay. K562 and HCT116 were seeded in triplicate in 96-well trays respectively at the density of $5 \cdot 10^3$ and 10^4 in 50 µL of complete medium. Stock solutions (50 mM) of each compound were made in DMSO and diluted in complete medium to give final concentrations of 50, 25 and 10 µM. Untreated cells were placed in every plate as a negative control. The cells were exposed to the compounds, in 100 µL total volume, for 72 hours.

The photocytotoxicity experiments were carried out by irradiating cells, incubated with photoactive components, with the irradiation source described above for 40 min. In this case, cell viability was measured after 72 h. After each incubation time 25 µL of a 12 mM solution of MTT were added. After two hours of incubation, 100 µL of lysing buffer (50% DMF + 20% sodium dodecyl sulfate (SDS), pH 4.7) were added to convert the MTT solution into a violet colored formazane. After additional 18 hours the solution absorbance, proportional to the number of live cells, was measured by spectrophotometer at 570 nm and converted into % of growth inhibition.

4.5 Bibliography

1. Wang, P. G.; Cai, T. B.; Taniguchi, N. Nitric Oxide Donors For Pharmaceutical and Biological Applications Preface. *Nitric Oxide Donors: For Pharmaceutical and Biological Applications* **2005,** Xv-Xvii.

2. Kerwin, J. F.; Lancaster, J. R.; Feldman, P. L. Nitric-Oxide - a New Paradigm for 2nd-Messengers. *J. Med. Chem.* **1995,** *38* (22), 4343−4362.

3. Carpenter, A. W.; Schoenfisch, M. H. Nitric oxide release: Part II. Therapeutic applications. *Chem. Soc. Rev.* **2012,** *41* (10), 3742−3752.

4. Halpenny, G. M.; Mascharak, P. K. Emerging Antimicrobial Applications of Nitric Oxide (NO) and NO-Releasing Materials. *Anti-Infect. Agents Med. Chem.* **2010,** *9* (4), 187−197.

5. Fang, F. C. *Nitric Oxide and Infections;* Kluwer Academic/Plenum Publishers: New York, **1999**.

6. Yasuda, H. Solid tumor physiology and hypoxia-induced chemo/radio-resistance: Novel strategy for cancer therapy: Nitric oxide donor as a therapeutic enhancer. *Nitric Oxide* **2008,** *19* (2), 205−216.

7. Ostrowski, A. D.; Ford, P. C. Metal complexes as photochemical nitric oxide precursors: Potential applications in the treatment of tumors. *Dalton T* **2009,** *48,* 10660−10669.

8. McCarthy, H. O.; Coulter, J. A.; Robson, T.; Hirst, D. G. Gene therapy via inducible nitric oxide synthase: a tool for the treatment of a diverse range of pathological conditions. *J. Pharm. Pharmacol.* **2008,** *60* (8), 999−1017.

9. Maddaford, S.; Annedi, S. C.; Ramnauth, J.; Rakhit, S. Advancements in the Development of Nitric Oxide Synthase Inhibitors. *Annu. Rep. Med. Chem.* **2009,** *44,* 27−50.

10. Riccio, D. A.; Schoenfisch, M. H. Nitric oxide release: Part I. Macromolecular scaffolds. *Chem. Soc. Rev.* **2012,** *41* (10), 3731−3741.

11. Fukumura, D.; Kashiwagi, S.; Jain, R. K. The role of nitric oxide in tumour progression. *Nat. Rev. Cancer* **2006,** *6* (7), 521−534.

12. Ridnour, L. A.; Thomas, D. D.; Donzelli, S.; Espey, M. G.; Roberts, D. D.; Wink, D. A.; Isenberg, J. S. The biphasic nature of nitric oxide responses in tumor biology. *Antioxid. Redox Signaling* **2006,** *8* (7-8), 1329−1337.

13. Chang, C. F.; Diers, A. R.; Hogg, N. Cancer cell metabolism and the modulating effects of nitric oxide. *Free Radical Biol. Med.* **2015,** *79,* 324−336.

14. Wimalawansa, S. J. Nitric oxide: new evidence for novel therapeutic indications. *Expert Opin. Pharmacother.* **2008,** *9* (11), 1935−1954.

15. Wang, P. G.; Xian, M.; Tang, X. P.; Wu, X. J.; Wen, Z.; Cai, T. W.; Janczuk, A. J. Nitric oxide donors: Chemical activities and biological applications. *Chem. Rev.* **2002,** *102* (4), 1091−1134.

16. Sortino, S. Photoactivated nanomaterials for biomedical release applications. *J. Mater. Chem.* **2012,** *22* (2), 301−318.

17. Sortino, S. Light-controlled nitric oxide delivering molecular assemblies. *Chem. Soc. Rev.* **2010,** *39* (8), 2903−2913.

18. Ford, P. C. Polychromophoric metal complexes for generating the bioregulatory agent nitric oxide by single- and two-photon excitation. *Acc. Chem. Res.* **2008,** *41* (2), 190−200.

19. Fry, N. L.; Mascharak, P. K. Photoactive Ruthenium Nitrosyls as NO Donors: How To Sensitize Them toward Visible Light. *Acc. Chem. Res.* **2011,** *44* (4), 289−298.

20. Ford, P. C. Photochemical delivery of nitric oxide. *Nitric Oxide* **2013,** *34,* 56−64.

21. Caruso, E. B.; Petralia, S.; Conoci, S.; Giuffrida, S.; Sortino, S. Photodelivery of nitric oxide from water-soluble platinum nanoparticles. *J. Am. Chem. Soc.* **2007,** *129* (3), 480-481.

22. Fraix, A.; Marino, N.; Sortino, S. Phototherapeutic Release of Nitric Oxide with Engineered Nanoconstructs. *Top. Curr. Chem.* **2016,** *370,* 225−257.

23. Fraix, A.; Sortino, S. Photoactivable Platforms for Nitric Oxide Delivery with Fluorescence Imaging. *Chem. - Asian J.* **2015,** *10* (5), 1116−1125.

24. Genini, D.; Adachi, S.; Chao, Q.; Rose, D. W.; Carrera, C. J.; Cottam, H. B.; Carson, D. A.; Leoni, M. L. Deoxyadenosine analogs induce programmed cell death in chronic lymphocytic leukemia cells by damaging the DNA and by directly affecting the mitochondria. *Blood* **2000,** *96* (10), 3537−3543.

25. Van Der Neste, E.; Cardoen, S.; Offner, F.; Bontemps, F. Old and new insights into the mechanisms of action of two nucleoside analogs active in lymphoid malignancies:

fludarabine and cladribine (Review). *Int. J. Oncol.* **2005,** *27* (4), 1113-1124.

26. Choi, Y. H.; Im, E. O.; Suh, H.; Jin, Y.; Lee, W. H.; Yoo, Y. H.; Kim, K. W.; Kim, N. D. Apoptotic activity of novel bile acid derivatives in human leukemic T cells through the activation of caspases. *Int. J. Oncol.* **2001,** *18* (5), 979-984.

27. Choi, Y. H.; Im, E. O.; Suh, H.; Jin, Y. G.; Yoo, Y. H.; Kim, N. D. Apoptosis and modulation of cell cycle control by synthetic derivatives of ursodeoxycholic acid and chenodeoxycholic acid in human prostate cancer cells. *Cancer Lett.* **2003,** *199* (2), 157-167.

28. Jeong, J. H.; Park, J. S.; Moon, B.; Kim, M. C.; Kim, J. K.; Lee, S.; Suh, H.; Kim, N. D.; Kim, J. M.; Park, Y. C.; Yoo, Y. H. Orphan nuclear receptor Nur77 traslocates to mitochondria in the early phase of apoptosis induced by synthetic chenodeoxycholic acid derivatives in human stomach cancer cell line SNU-1. *Ann. N. Y. Acad. Sci.* **2003,** *1010,* 171-177.

29. Park, S. E.; Choi, H. J.; Yee, S. B.; Chung, H. Y.; Suh, H.; Choi, Y. H.; Yoo, Y. H.; Kim, N. D. Synthetic bile acid derivatives inhibit cell proliferation and induce apoptosis in HT-29 human colon cancer cells. *Int. J. Oncol.* **2004,** *25* (1), 231-236.

30. Perrone, D.; Bortolini, O.; Fogagnolo, M.; Marchesi, E.; Mari, L.; Massarenti, C.; Navacchia, M. L.; Sforza, F.; Varani, K.; Capobianco, M. L. Synthesis and in vitro cytotoxicity of deoxyadenosine–bile acid conjugates linked with 1,2,3-triazole. *New. J. Chem.* **2013,** *37,* 3559–3567.

31. Capobianco, M. L.; Marchesi, E.; Perrone, D.; Navacchia, M. L. Labeling Deoxyadenosine for the Preparation of Functional Conjugated Oligonucleotides. *Bioconjugate Chem.* **2013,** *24* (8), 1398-1407.

32. Chatgilialoglu, C.; Navacchia, M. L.; Postigo, A. A facile one-pot synthesis of 8-oxo-7,8-dihydro-(2′-deoxy)adenosine in water. *Tetrahedron Lett.* **2006,** *47* (5), 711-714.

33. Bosi, V.; Sarti, E.; Navacchia, M. L.; Perrone, D.; Pasti, L.; Cavazzini, A.; Capobianco, M. L. Gold-nanoparticle extraction and reversed-electrode-polarity stacking mode combined to enhance capillary electrophoresis sensitivity for conjugated nucleosides and oligonucleotides containing thioether linkers. *Anal. Bioanal. Chem.* **2015,** *407* (18), 5405-5415.

34. Navacchia, M. L.; Fraix, A.; Chinaglia, N.; Gallerani, E.; Perrone, D.; Cardile, V.; Graziano, A. C. E.; Capobianco, M. L.; Sortino, S. NO Photoreleaser-Deoxyadenosine and -Bile Acid Derivative Bioconjugates as Novel Potential Photochemotherapeutics. *ACS Med. Chem. Lett.* **2016,** *7,* 939−943.

35. Callari, F. L.; Sortino, S. Amplified nitric oxide photorelease in DNA proximity. *Chem. Commun.* **2008,** No. 17, 1971−1973.

36. Coneski, P. N.; Schoenfisch, M. H. Nitric oxide release: Part III. Measurement and reporting. *Chem. Soc. Rev.* **2012,** *41* (10), 3753−3758.

5 Rational design of nucleoside–bile acid conjugates incorporating a triazole moiety for anticancer evaluation and SAR exploration

5.1 Introduction

Important traditional chemotherapeutic drugs or anticancer agents were mostly derived from natural sources through synthetic structural modifications. Successful examples of this approach are represented, among others, by the anthracyclines, taxanes and camptothecins (Figure 5.1) that are still considered a structural platform for discovering new anticancer drugs [1].

Idarubicin — Baccatin III — Camptothecin
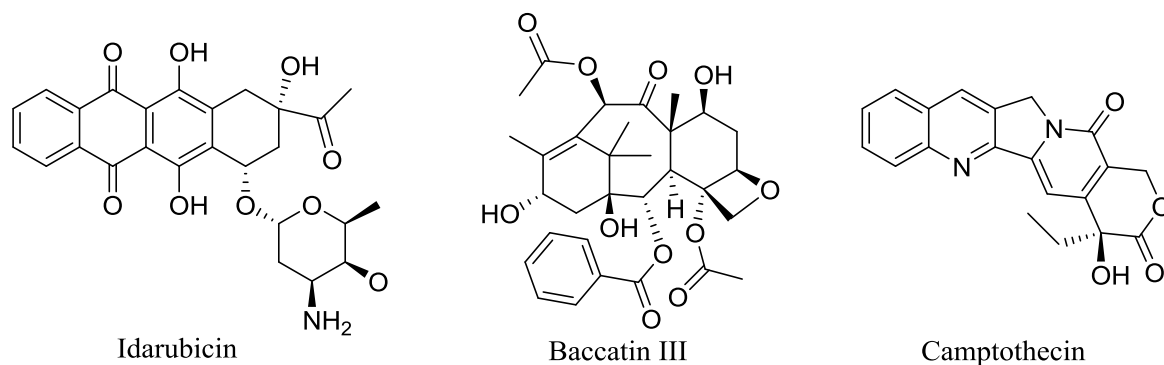
Figure 5.1 *Structures of idarubicin (an anthracycline), baccatin III (a taxane) and camptothecin.*

Nucleosides and nucleotides—endogenous small molecules that can be chemically fine-tuned leading to the corresponding analogues—can behave as antimetabolites and can inhibit the cellular division and viral replication by incorporation into DNA or RNA, resulting in potential therapeutic benefits. They can also act as inhibitors of essential enzymes such as DNA polymerases, kinases and so on. In such a way, they would operate by stopping the synthesis of pre-DNA molecule building blocks or by direct damage of the DNA in the nucleus of the cell or by effecting the synthesis or by breakdown of the mitotic spindles. Currently, several nucleoside and nucleotide analogues derived from 2′-deoxycitidine, 2′-deoxyadenosine and 2′-deoxyguanosine have been approved by the FDA as anti-cancer drugs or anti-viral agents [2].

Despite their therapeutic potential, the bioavailability of hydrophilic nucleoside-based drugs remains a critical negative feature since they do not readily cross the plasma membrane by passive diffusion, and accordingly, their clinical efficacy also depends on nucleoside delivery systems [3,4].

With the aim to discover new nucleoside analogues with anticancer activity we consider conjugation as a powerful approach. In principle, a targeted conjugation can be helpful to tune the cytotoxicity, for instance by coupling a 2′-deoxyadenosine derivative with a NO photodonor unit the intrinsic cytotoxicity of the bioconjugate combined upon light irradiation with that of the photogenerated NO leads to an interplay of anticancer mechanisms of action [5], as seen previously (see "NO photoreleaser-deoxyadenosine and -bile acid derivative bioconjugates as novel potential photochemotherapeutics"). Furthermore, 2′-deoxyadenosine derivatives conjugated with cheno- and urso-deoxycholic acids through a triazole or a thioalkyl unit tested on four cancer cell lines (K562, Jurkat, HCT116 and A2780) showed interesting antiproliferative activity selectively towards leukemic T-cells whereas no

cytotoxicity against the solid tumors HCT116 and A2780 was found [6]. In our studies, bile acids (BAs) were chosen as combination partners by virtue of their biological as well as physico-chemical properties (see "Introduction to bile acids chemistry"). For instance, the cytotoxic activity of certain BAs and BA-derivatives is well recognized, including the potential of several unconjugated BAs to induce cell death in a wide range of cells, through their non-specific ability to disrupt cell membranes (biological surfactant feature) or receptor-mediated interactions and DNA oxidative damage [7]. On the other hand, the conjugation with hydrophilic glycine and taurine can dramatically decreases BA cytotoxicity while enhancing the neuroprotective effects [8,9]. Moreover, taking advantage of their organotropism in the enterohepatic circulation mediated by the BA transport systems, the presence of BA units can be helpful in targeting a drug conjugate to the liver or to improve its metabolic stability [10]. It has been also reported that the conjugation of zidovuidine (AZT), a nucleoside analogue-based drug, with ursodesoxycholic acid increases the poor permeability of AZT through the intracellular departments [11]. Thanks to their intrinsic chemical features BAs can be fine tailored.

We present herein a study on the synthesis and biological evaluation of nucleoside-BA conjugates obtained by combining a selection of nucleoside analogues and bile acid derivatives (Figure 5.2). For this purpose, 2′-deoxyadenosine (dA), 2′-deoxyguanosine (dG), 2′-deoxyuridine (dU), as well as adenosine (A) and guanosine (G), modified at a suitable position with an alkynyl chain containing an acetylenic bond, were conjugated by means of the click reaction with cheno- (CDC), urso- (UDC), taurourso- (TUDCA), *nor*-cheno (*nor*-CDC) and *nor*-urso- (*nor*-UDC) deoxycholic acid derivatives equipped with the azido group at the head or the tail position. The new nucleoside-BA conjugates were characterized and tested *in vitro* against two types of cancer cell lines: leukemic K562, a hematological cancer, and the solid tumor HCT116 colon carcinoma, as well as on normal fibroblast cells.
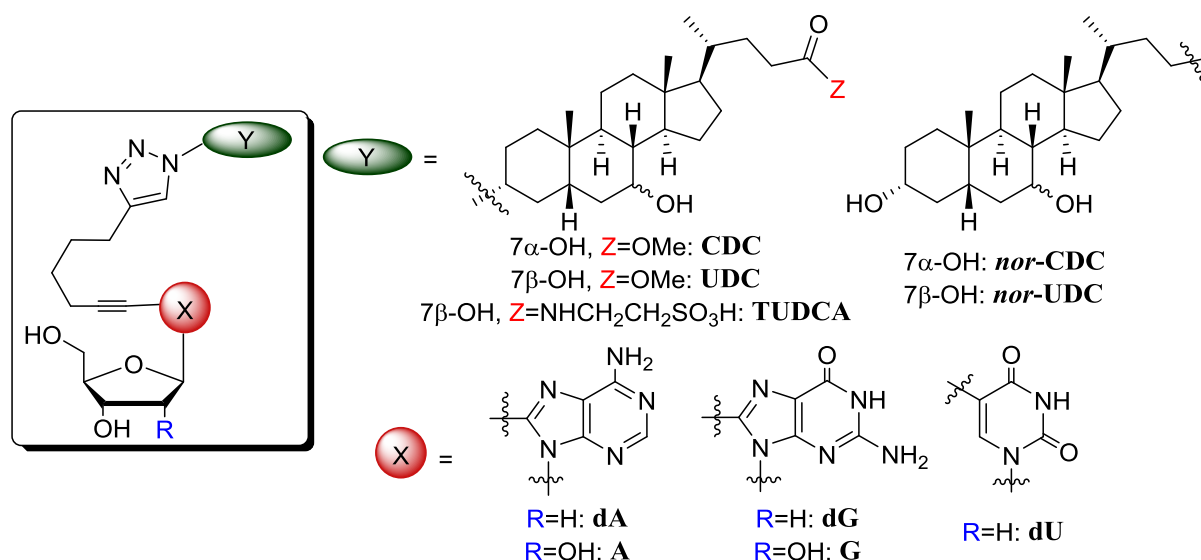

Figure 5.2 *Sketch of novel nucleoside-bile acid conjugates.*

5.2 Results and discussion

5.2.1 Synthesis of nucleoside-bile acid conjugates

The click chemistry approach, being a specific and high yield reaction, was considered a good synthetic approach for the preparation of our target nucleoside-BA hybrids. Moreover, the

triazole moiety resulting from the 1,3-cycloaddition is biologically relevant, being able to improve the biostability, bioavailability and also the anticancer activity of bioactive compounds [12,13,14,15].

Click chemistry requires the presence of a terminal alkyne moiety and an azido group. To provide those features the nucleoside units were modified with an alkynyl moiety at C-8 position in the case of the purine bases and at the C-6 position in the case of the pyrimidine one in order to keep unchanged their intrinsic characteristic of recognition of natural nucleic acids through specific hydrogen bond patterns (Watson-Crick and Hoogsteen). 8-(1,7-Octadynyl)-2′-deoxyadenosine (**ALK-dA**) was prepared as previously reported [6]. Similarly, the 8-(1,7-octadynyl)-derivative of A, dG and G, as well as 6-(1,7-octadynyl)-2′-deoxyuridine (namely **ALK-A**, **ALK-dG**, **ALK-G** and **ALK-dU**, respectively) were synthesized through a standard palladium catalyzed cross-coupling reaction starting from commercially available 8-bromo-A, 8-bromo-dG, 8-bromo-G and 6-iodo-dU [16] (Scheme 5.1).

R=H: **8-Br-dA**
R=OH: **8-Br-A**

R=H: **8-Br-dG**
R=OH: **8-Br-G**

6-I-dU

R=H: **ALK-dA**
R=OH: **ALK-A**

R=H: **ALK-dG**
R=OH: **ALK-G**

ALK-dU
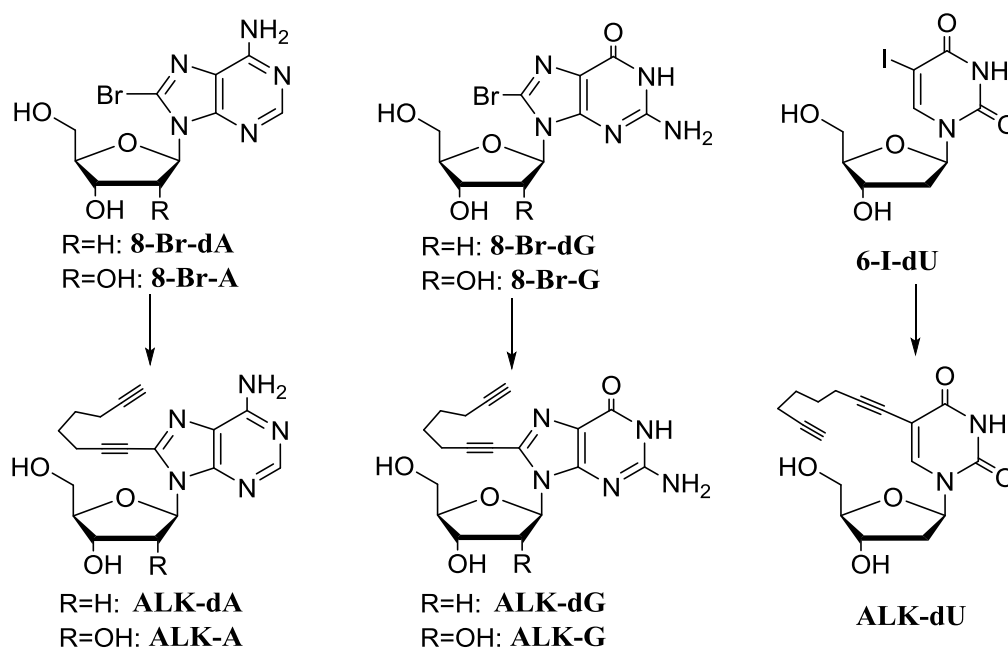
Scheme 5.1 *Syntheses of alkyne-nucleoside intermediates. Reagents and conditions: 1,7-octadiyne, (PPh₃)₂PdCl₂, CuI, DMF, TEA, 50°C, 2.30 h (70-80% yield).*

The azido-BA derivatives **3α-N₃-CDC** and **3α-N₃-UDC** were synthesized starting from commercially available BAs, using a synthetic approach described previously [6,17] (Scheme 5.2). The 3α-N₃ derivative of TUDCA was prepared in three steps in 75% overall yield from the corresponding **3α-N₃-UDC**: the methyl ester was hydrolyzed with a 1.5 M LiOH in MeOH to the corresponding acid, which in turn was coupled with the aminoethanesulfonic acid taurine after activation of the free acid with ethyl chloroformate (Scheme 5.2). Finally, the **23-N₃-*nor*-CDC** and **23-N₃-*nor*-UDC** were synthesized starting from the corresponding bile acids following a recently reported metal free iodo-decarboxylation method [18]. Accordingly, the C3 and C7 free hydroxyl groups of CDCA and UDCA were firstly protected as formyloxy derivatives by using formic acid at 55°C for 24 h, then concentrated *in vacuo* and the residues irradiated for 2 h in presence of 1,3-diiodo-5,5-dimethylhydantoin (DIH) as a sole reagent. After purification by flash chromatography the **23-I-*nor*-CDC** and **-UDC** derivatives (85–90% yield) were converted into the target compounds through a nucleophilic

substitution with sodium azide in DMF at room temperature, followed by hydrolysis of the formate esters with 25% NaOH which allowed the precipitation of 23-N$_3$-*nor* BAs derivatives as pure compounds in satisfactory yields (75–78% after two steps, Scheme 5.2).
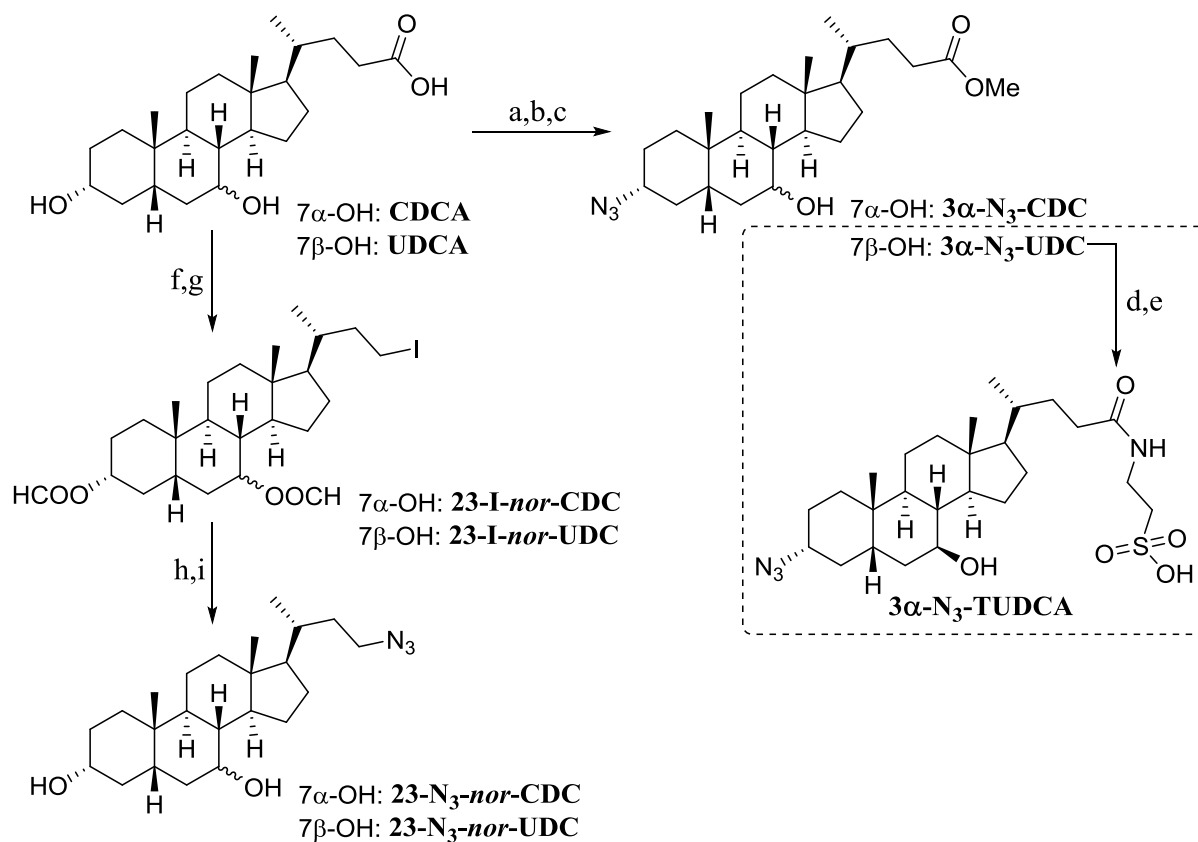
Scheme 5.2 *Syntheses of azido-bile acid intermediates. Reagents and conditions: (a) H$_2$SO$_4$, MeOH, 100°C, 2h; (b) I$_2$, Ph$_3$P, imidazole, 1,3-dioxolane, r.t., 30min (55-65% yield after two steps); (c) NaN$_3$, DMF, rt, 6h, (87-94% yield); (d) LiOH 1.5M, MeOH, r.t., 21h (94% yield); (e) TEA, ClCOOEt, THF, 0°C to r.t., 1h; then taurine, NaOH 10%, r.t., 12h (80% yield); (f) HCOOH, 55°C, 24h; (g) DIH, DCE, hv, reflux, 2h (85-90% yield after two steps); (h) NaN$_3$, DMF, r.t., 6h; (i) NaOH 25%, MeOH, r.t., 2h (75-78% yield after two steps).*

The click chemistry reaction was performed *via* a Cu(I)-mediated 1,3-dipolar cycloaddition under commonly used conditions: a (1:1:1.5) H$_2$O/*tert*-BuOH/THF (*v/v*) solution of the appropriate alkyne–nucleoside derivative and of the BA-azide in the presence of the CuSO$_4$ catalyst and sodium ascorbate was stirred at room temperature for 18 h. The target conjugates were obtained in yields ranging from 60% to 90% after purification. In Scheme 5.3 is depicted, for example, the synthesis of the conjugate compounds of adenosine with all the BA selected (namely **A-CDC**, **A-UDC**, **A-TUDCA**, **A-*nor*-CDC** and **A-*nor*-UDC**).
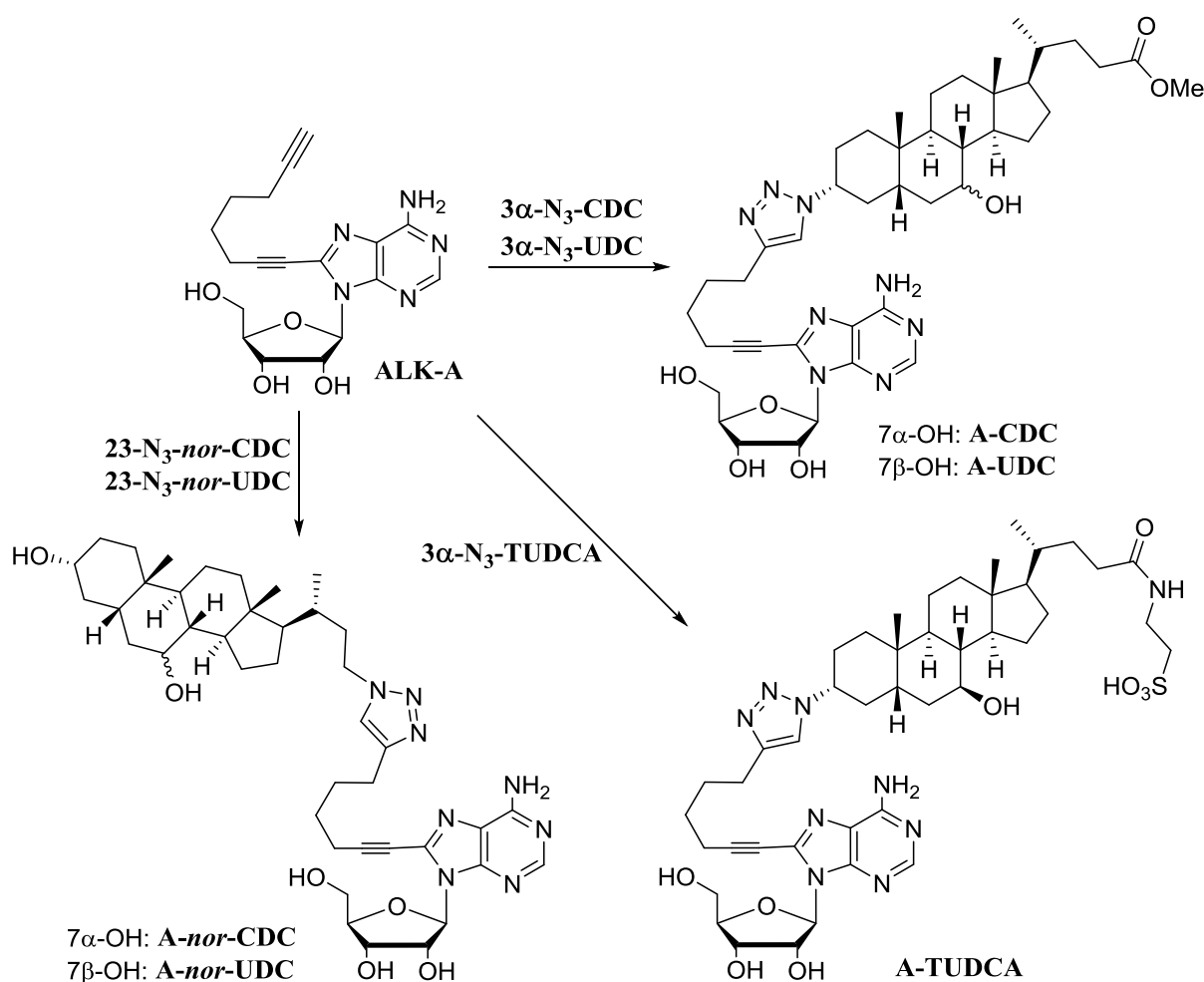
Scheme 5.3 Syntheses of A-CDC, A-UDC, A-nor-CDC, A-nor-UDC and A-TUDCA conjugates. Reagents and conditions: CuSO₄·5H₂O, sodium ascorbate, 1:1:1.5 H₂O/tert-BuOH/THF (v/v), 25°C, 18 h (63-80% yield).

5.2.2 Biological assays

All conjugated compounds prepared *via* click chemistry and listed in Table 5.1 were evaluated *in vitro* for their cytotoxic activity against K562 leukemia cells and the colon cancer cell line HCT116. Normal human skin fibroblast cells were chosen as a control and cisplatin served as a reference compound. The cytotoxicity was evaluated using the MTT assay (details are reported in the Experimental section). In all experiments the cell growth inhibition of K562 and HCT116 was determined for each compound at concentrations of 10, 25 and 50 μM after 72 h of treatment and up to 100 μM in the case of fibroblasts.

Structure	Compound	IC₅₀ (μM)		
		K562	**HCT116**	**fibroblast**
	7α-OH, R=H: **dA-CDC** [6]	8.5±4.0	>200	>200
	7β-OH, R=H: **dA-UDC** [6]	>100	>200	>200
	7α-OH, R=OH: **A-CDC**	43.5±1.3	23.1±1.7	>100
	7β-OH, R=OH: **A-UDC**	>50	>50	>100

Structure	Compound	IC$_{50}$ (μM)		
		K562	HCT116	fibroblast
	7α-OH, R=H: **dA-*nor*-CDC**	16.2±2.2	17.0±2.5	91.5±3.5
	7β-OH, R=H: **dA-*nor*-UDC**	>50	44.8±3.5	>100
	7α-OH, R=OH: **A-*nor*-CDC**	23.6±1.2	44.1±2.5	>100
	7β-OH, R=OH: **A-*nor*-UDC**	>50	>50	>100
	R=H: **dA-TUDCA**	>50	>50	>100
	R=OH: **A-TUDCA**	>50	>50	>100
	7α-OH, R=H: **dG-CDC**	>50	>50	>100
	7β-OH, R=H: **dG-UDC**	>50	>50	>100
	7α-OH, R=OH: **G-CDC**	>50	25.3±3.8	>100
	7β-OH, R=OH: **G-UDC**	>50	>50	>100
	7α-OH, R=H: **dG-*nor*-CDC**	>50	>50	>100
	7β-OH, R=H: **dG-*nor*-UDC**	>50	>50	>100
	7α-OH, R=OH: **G-*nor*-CDC**	>50	>50	>100
	7β-OH, R=OH: **G-*nor*-UDC**	>50	>50	>100
	R=H: **dG-TUDCA**	>50	>50	>100
	R=OH: **G-TUDCA**	>50	>50	>100
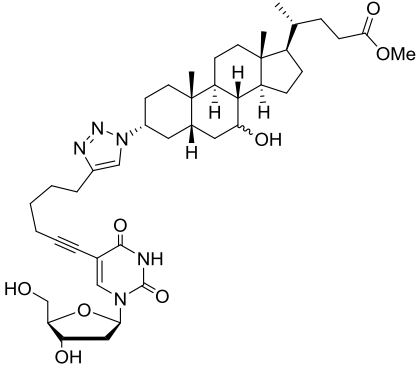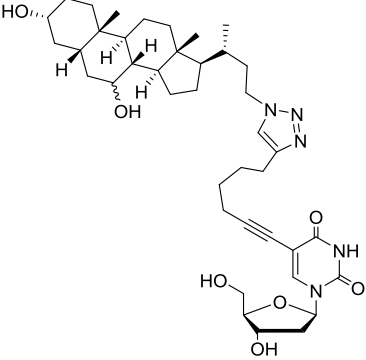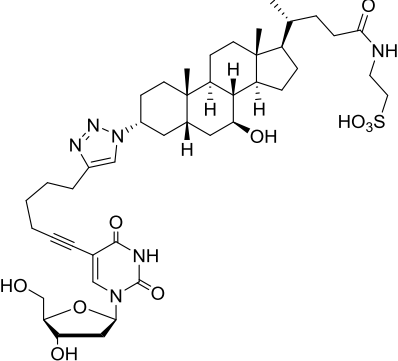
Structure	Compound	IC$_{50}$ (μM)		
		K562	HCT116	fibroblast
	7α-OH: **dU-CDC**	>50	>50	>100
	7β-OH: **dU-UDC**	21.5±2.0	23.5±1.6	91.5±4.5
	7α-OH: **dU-*nor*-CDC**	42.9±1.9	>50	>100
	7β-OH: **dU-*nor*-UDC**	24.8±1.5	43.0±3.5	>100
	dU-TUDCA	>50	>50	>100
	ALK-dA [6]	>50	>50	>100
	ALK-A	>50	>50	>100
	ALK-dG	>50	>50	>100
	ALK-G	>50	>50	>100
	ALK-dU	>50	>50	>100
	3α-N$_3$-CDC	23.0±2.0	31.0±2.3	>100
	3α-N$_3$-UDC	25.0±1.0	22.0±1.5	>100
	23-N$_3$-nor-CDC	21±1.2	25±2.0	79±3.1
	23-N$_3$-nor-UDC	15±1.0	22±1.4	81±2.0
	3α-N$_3$-TUDCA	>50	>50	>100
	cisplatin	5.4±1.0	8.5±1.2	23.6±3.5

Table 5.1 *IC$_{50}$ values determined from the dose–response curves using MTT assay after 72 h incubation time. Results are expressed as the mean of three independent experiments ±SD. Cisplatin was used as a reference compound. Where indicated IC$_{50}$>50 inhibition found was <20% at 50 μM; where indicated IC$_{50}$>100 inhibition found was <10% at 100 μM.*

Figure 5.3 shows the antiproliferative activity of the most active conjugates against K562 and HCT116 cancer cells whereas in Table 5.1 the IC$_{50}$ values are reported for all the compounds tested, including the alkyne-nucleoside and BA-azide building blocks.
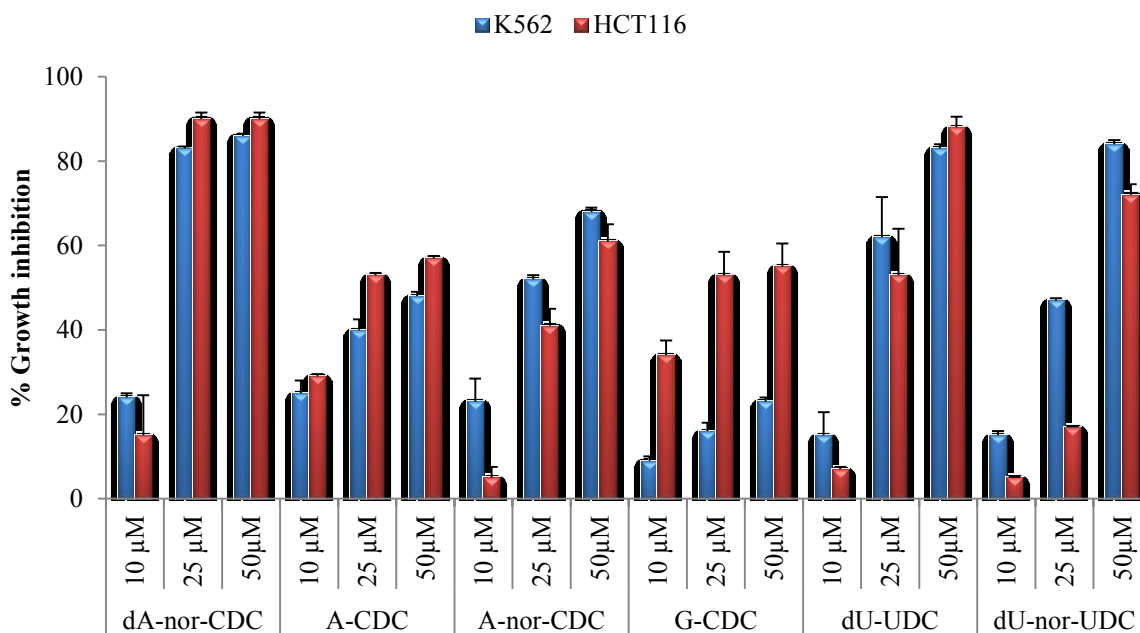
Figure 5.3 *Antiproliferative activity of the most active conjugates against K562 and HCT116 cancer cells at 10, 25 and 50 μM after 72h.*

With regard to the 2′-deoxyadenosine derivatives, **dA-*nor*-CDC** was found cytotoxic against both K562 and HCT116, with comparable IC_{50} values (16.2 and 17.0 μM respectively; Table 5.1). On the other hand, **dA-*nor*-UDC** showed a preferential cytotoxicity against HCT116 cancer cells (IC_{50} = 44.8 μM vs. 87.1 μM extrapolated value; Table 5.1).

In the case of the adenosine derivatives, both **A-CDC** and **A-*nor*-CDC** were found active against K562 and HCT116, but with an opposite cytoselectivity. Indeed, **A-CDC** showed a higher cytotoxicity against HCT116, whereas **A-*nor*-CDC** was found more toxic towards K562. None of the bioconjugates of 2′-deoxyguanosine and guanosine series showed any cytotoxicity, with the only exception of **G-CDC** which was found to be selectively cytotoxic against HCT116 cancer cells (IC_{50} = 25.3 μM; Table 5.1).

In the 2′-deoxyuridine series, **dU-UDC** was found cytotoxic towards both K562 and HCT116, with comparable IC_{50} values (21.5 and 23.5 μM respectively; Table 5.1), whereas **dU-*nor*-UDC** showed cytoselectivity towards K562, with IC_{50} = 24.8 μM vs. HCT116 IC_{50} = 43.0 μM (Figure 5.3, Table 5.1). No cytotoxic activity was found for the conjugates with TUDCA.

The nucleoside-alkyne derivatives (namely **ALK-A**, **ALK-G**, **ALK-dG** and **ALK-dU**) and the N_3-BA building blocks (namely **3α-N₃-CDC**, **3α-N₃-UDC**, **3α-N₃-TUDCA**, **23-N₃-*nor*-CDC** and **23-N₃-*nor*-UDC**) were also tested *in vitro*. The nucleoside-alkyne derivatives were found to consistently not be cytotoxic towards any of tested cell lines with the results previously reported for **ALK-dA** [6]. As far as for the 3α-N₃-BA is concerned, we found that **3α-N₃-CDC** and **3α-N₃-UDC** are active against both cancer cell lines indiscriminately and not active towards the fibroblast cells up to 100 μM, whereas **3α-N₃-TUDCA** was found to not be cytotoxic against any of the cell lines at the concentrations tested. On the other hand, the 23-N₃-*nor*-BA series showed cytotoxicity against the cancer cell lines, with IC_{50} values ≤ 25 μM, and, to a minor extent, also towards the fibroblasts (Table 5.1).

To determine whether the antiproliferative activity induced by **dA-*nor*-CDC** was related to apoptosis, as previously reported for the corresponding **dA-CDC** conjugate [6], K562 cells were treated with compound **dA-*nor*-CDC** (25 and 50 μM) for 24 h, then assayed by flow

cytometry analysis with Annexin V-FITC staining. The results of the cell apoptosis assay indicated that compound **dA-*nor*-CDC** induced apoptosis in a dose dependent manner (Figure 5.4).
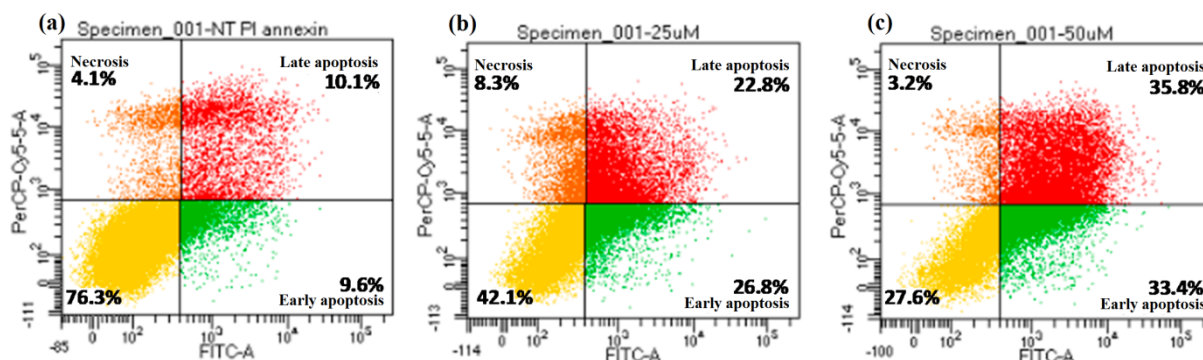

Figure 5.4 *dA-nor-CDC induced apoptosis in K562 cells after 24 h treatment: (a) control: untreated cells; (b) dA-nor-CDC 25 µM; (c) dA-nor-CDC 50 µM* [19].

5.2.3 Structure–activity relationship exploration

The reported *in vitro* screening highlighted some compounds with an interesting anticancer activity, with IC$_{50}$ values ≤ 25 µM, which are **dA-*nor*-CDC**, **A-*nor*-CDC**, **dU-UDC** and **dU-*nor*-UDC** with respect to K562 leukemia cells and **dA-*nor*-CDC**, **A-CDC**, **G-CDC** and **dU-UDC** with respect to HCT116 colon carcinoma (Table 5.1, Figure 5.3). Among them, only **dA-*nor*-CDC** and **dU-UDC** showed good anti-proliferative activity against both K562 and HCT116, with comparable IC$_{50}$ values. It is worth noting that these two compounds showed also a higher IC$_{50}$ value (around 100 µM) respect to the other conjugates towards the fibroblast cells therefore the lack of cytoselectivity among the selected cancer lines could be related to the greater activity of the compounds (Table 5.1).

In agreement with literature [6], the A/dA-based bioconjugates were confirmed to be potential active anticancer compounds. Besides, the present screening also evidenced G- and U-based conjugates with interesting cytotoxicity/cytoselectivity. Moreover, it can be observed that CDC/*nor*-CDC scaffolds conjugated with A/dA nucleosides showed in all cases a fair cytotoxic activity and cytoselectivity. Conversely, UDC/*nor*-UDC scaffolds showed cytotoxic activity only when coupled with 2′-deoxyuridine (Table 5.1).

In the previous paper on 2′-deoxyadenosine-BA conjugates [6], including **dA-CDC** and **dA-UDC** (those IC$_{50}$ values are also reported in Table 5.1 for comparison), it was been demonstrated that the conjugation of dA with CDC and UDC actually plays a crucial role in the cytotoxic/cytoselective process. Starting from this point we would like to discuss the possible structure-activity relationship in the light of the biological evaluation of the nucleoside-BA conjugates incorporating a triazole moiety herein reported.

A marked cytoselectivity trend among the selected cancer lines can be identified in the adenosine series. In fact, compound **A-CDC** is preferentially cytotoxic against HCT116 cells, whereas the corresponding **A-*nor*-CDC** derivative showed cytoselectivity towards K562 cells. Therefore, CDC bile acid seems to address the cytoselectivity to HCT116 whereas *nor*-CDC bile acid does the same for K562 cells. This seems to be supported by the data of both the G and dU series where **G-CDC** was found highly cytoselective against the HCT116 and **dU-*nor*-CDC** cytoselective against the K562 as expected in the light of the previous consideration. However, this hypothesis is in contrast with the cytoselectivity of **dA-CDC** that is selective against the K562 unless it is a CDC derivative. The overall data also indicate

that the CDC/*nor*-CDC derivatives are more active than the corresponding UDC/*nor*-UDC, except for the conjugates with 20-deoxyuridine (Table 5.1, Figure 5.3). Therefore, the conjugation with a pyrimidine nucleoside seems to improve the anticancer activity of the UDC/*nor*-UDC series. Looking more deeply through the biological data it can be seen that also the sugar nature, deoxy- or ribo-, seems to influence the cytoselectivity being the ribo form more active against the HCT116 in the adenine series (comparison between **dA-CDC** and **A-CDC**) and also in the corresponding guanine series (**G-CDC**) (Table 5.1). Finally, in the case of TUDCA-conjugates no cytotoxic activity was found, as for the corresponding **3-N$_3$-TUDCA** building block. The overall data herein debated seems to indicate that the cytoselectivity is mainly driven by the BA and can be fine-tuned by the nucleoside nature, *i.e.*, purine or pyrimidine, deoxy- or ribo-[19].

5.3 Conclusion

A conjugation approach by means of click chemistry was exploited in order to synthesize a library of novel fully bio-inspired conjugates combining dA, A, dG, G and dU with CDC, UDC, TUDCA, *nor*-CDC and *nor*-UDC bile acids derivatives. All the nucleoside-BA conjugates were tested for their *in vitro* anti-proliferative activity against two cancer cells lines and their cytotoxicity towards human fibroblast normal cells. In most of the cases negligible cytotoxicity toward fibroblast was found. Six compounds displayed an interesting anti-proliferative activity with IC$_{50}$ value ≤ 25 μM.

In particular, **A-*nor*-CDC** and **dU-*nor*-UDC** were found to be selectively cytotoxic against K562 leukemia cells; **A-CDC** and **G-CDC** were found to be selectively cytotoxic against HCT116; **dA-*nor*-CDC** and **dU-UDC** showed good anti-proliferative activity against both K562 and HCT116.

Furthermore, the mechanism of K562 cell death was investigated in the case of **dA-*nor*-CDC** which showed a high percentage of specific apoptosis.

A possible structure–activity relationship was also investigated. In the light of the present data we reason that the cytoselectivity is mainly driven by the nature of the BA but also influenced by the nature of the nucleobase and the sugar form of the nucleoside, *i.e.* deoxy- or ribo-. Therefore, the cytotoxicity could be considered as the result of an interplay of the fine details of the structure of the conjugates.

This study confirmed that the conjugation of nucleosides and BAs can actually pave the way to new compounds for anticancer therapy and that unless the structure–activity relationship is not self evident there is a common thread.

5.4 Experimental

5.4.1 General information

Reactions were monitored by TLC on pre-coated Silica Gel plates (thickness 0.25 mm, Merck), and phosphomolybdic acid solution was used as the spray reagent to visualize the steroids. Flash column chromatography was performed on silica gel 60 (230-400 mesh). HPLC-MS analyses were performed on an *Agilent 1100* HPLC system and an *Esquire 3000 Plus Bruker* mass spectrometer using a *Zorbax C8* column (4.6 x 150 mm, 5 μm) (linear gradient water/CH$_3$CN at a 0.5 mL/min flow rate, detection at λ = 260 nm). ESI-HRMS were acquired on an *Agilent Ual ESI Q TOF 6520*, in positive-ion mode, using methanol. NMR spectra were recorded for DMSO-*d$_6$* solution, unless otherwise specified, with a *Varian*

Mercury Plus 400 MHz instrument. IR spectra were recorded on a *Perkin–Elmer Spectrum 100 FT-IR* spectrometer. 8-Br-adenosine, 8-Br-2′-deoxy-adenosine, 8-Br-2′-deoxy-guanosine, 8-Br-guanosine, 5-I-2′-deoxyuridine, chenodeoxycholic acid and ursodeoxycholic acid are commercially available compounds used without further purification. The corresponding azides, *methyl 3α-azido-7α-hydroxy-5β-cholan-24-oate* (**3α-N₃-CDC**) and *methyl 3α-azido-7β-hydroxy-5β-cholan-24-oate* (**3α-N₃-UDC**), were prepared according to the literature procedures [6,17].

5.4.2 Derivatization of nucleosides

5.4.2.1 General procedure for the synthesis of alkynes

Alkynes were prepared following the procedure reported in the literature [16]. In all cases they were obtained in 70–80% yield and no chromatographic purification was necessary. An analytical sample for the characterization analyses was obtained after flash chromatography using CH_2Cl_2/MeOH 9:1 as eluent in all cases.
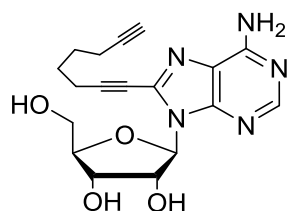
ALK-A

8-(1,7-Octadynyl)-adenosine. ^{1}H-NMR: δ = 8.18 (1H, s, H2), 7.56 (2H, br s; disappeared upon D_2O shake; NH₂), 5.90 (1H, d, J = 7.2 Hz; collapsing to s upon irradiation at δ 4.98; H1'), 5.54 (1H, m; disappeared upon D_2O shake; C5′-OH), 5.39 (1H, d, J = 6.4 Hz; collapsing to s upon irradiation at δ 4.98; disappeared upon D_2O shake; C2′-OH), 5.16 (1H, d, J = 3.6 Hz; collapsing to s upon irradiation at δ 4.17; disappeared upon D_2O shake; C3′-OH), 4.98 (1H, m; collapsing to dd, J_1 = 7.2 Hz and J_2 = 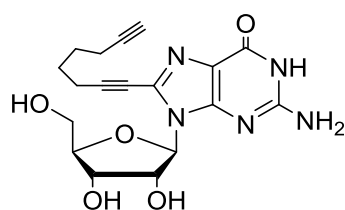6.4 Hz, upon irradiation at δ 4.17; H2'), 4.18 (1H, m; H3'), 3.96 (1H, m; H4'), 3.64 (1H, m; H5'), 3.50 (1H, m; H5''), 2.53 (3H, m), 2.37 (2H, m), 1.64 (4H, m). ^{13}C-NMR: δ = 153.7 (CH), 153.4 (q), 149.5 (q), 132.0 (q), 94.7 (q), 94.6 (q), 89.9 (CH), 87.2 (CH), 72.1 (CH), 71.6 (CH), 66.2 (q), 62.8 (CH), 55.5 (CH₂), 27.5 (CH₂), 27.3 (CH₂), 18.7 (CH₂), 18.5 (CH₂). MS (ESI, ES+) *m/z*: calculated for $C_{18}H_{21}N_5O_4$ 371.40; found 372 [M+H]⁺, 394 [M+Na]⁺.

ALK-G

8-(1,7-octadynyl)-guanosine. ^{1}H-NMR: δ = 5.75 (1H, d, J = 7.2 Hz; collapsing to s upon irradiation at δ 4.90; H1'), 4.98 (1H, m; collapsing to dd, J_1 = 7.2 Hz and J_2 = 6.8 Hz, upon irradiation at δ 4.08; H2'), 4.08-4.05 (1H, m; H3'), 3.94-3.91 (1H, m; H4'), 3.64-3.40 (2H, m; H5', H5''), 2.75 (1H, t, J = 2.8 Hz), 2.58-2.50 (2H, m), 2.23-2.18 (2H, m), 1.65-1.53 (4H, m). ^{13}C-NMR: δ = 163.4 (q), 137.1 (q), 134.6 (q), 131.4 (q), 118.9 (q), 94.5 (q), 89.5 (q), 87.1 (CH), 84.9 (CH), 72.4 (q), 72.1 (CH), 72.1 (CH), 71.8 (CH), 63.3 (CH₂), 27.8 (CH₂), 26.7 (CH₂), 18.8 (CH₂), 17.9 (CH₂). MS (ESI, ES+) *m/z*: calculated for $C_{18}H_{21}N_5O_5$ 387.40; found 388 [M+H]⁺, 410 [M+Na]⁺.
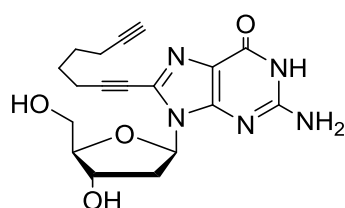
ALK-dG

8-(1,7-octadynyl)- 2′-deoxyguanosine. ^{1}H-NMR: δ = 6.23-6.18 (1H, m; H1'), 4.40-4.38 (1H, m; H3'), 3.90-3.83 (1H, m; H4'), 3.68-3.60 (1H, m; H5'), 3.58-3.42 (1H, m; H5''), 3.02-2.95 (1H,

m; H2''), 2.75 (1H, t, J = 2.8 Hz), 2.58-2.50 (2H, m), 2.25-2.19 (2H, m), 2.18-1.98 (1H, m; H2'), 1.70-1.52 (4H, m). ^{13}C-NMR: δ = 163.4 (q), 137.1 (q), 134.6 (q), 131.4 (q), 118.9 (q), 94.5 (q), 89.5 (q), 87.1 (CH), 84.9 (CH), 72.4 (q), 72.1 (CH), 71.8 (CH), 63.3 (CH$_2$), 40.2 (CH$_2$), 27.8 (CH$_2$), 26.7 (CH$_2$), 18.8 (CH$_2$), 17.9 (CH$_2$). MS (ESI, ES+) m/z: calculated for C$_{18}$H$_{21}$N$_5$O$_4$ 371.40; found 394 [M+Na]$^+$. (ESI, ES-) m/z: found 370 [M-1]$^-$.
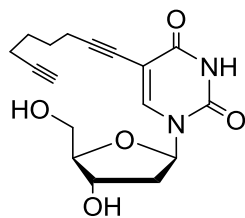

ALK-dU

5-(1,7-octadynyl)-2'-deoxyuridine. ^{1}H-NMR: δ = 8.10 (1H, s, H6), 6.08 (1H, dd, J$_1$ = J$_1$ = 6.4 Hz; collapsing to s upon irradiation at δ 2.08; H1'), 5.22 (1H, m; disappeared upon D$_2$O shake; C5′-OH), 5.07 (1H, m, disappeared upon D$_2$O shake; C3′-OH), 4.21-4.18 (1H, m; H3'), 3.78 (1H, m; collapsing to d, J = 2.8 Hz upon irradiation at δ 3.55; H4'), 3.58-3.53 (2H, m; H5'and H5''), 2.74 (1H, t, J = 2.4 Hz; collapsing to s upon irradiation at δ 2.18), 2.362.32 (2H, m), 2.20-2.17 (2H, m), 2.11-2.06 (2H, m, H2'and H2''), 1.60-1.53 (4H, m). ^{13}C-NMR: δ = 163.2 (q), 150.1 (q), 143.4 (CH), 99.7 (q), 93.7 (q), 88.1 (CH), 85.3 (CH), 85.1 (q), 72.0 (q), 70.8 (CH), 61.5 (CH$_2$), 46.4 (CH$_2$), 27.9 (CH$_2$), 27.7 (CH$_2$), 19.0 (CH$_2$), 17.9 (CH$_2$). MS (ESI, ES+) *m/z*: calculated for C$_{17}$H$_{20}$N$_2$O$_5$ 332.36; found 333 [M+H]$^+$, 355 [M+Na]$^+$, 665 [2M+H]$^+$, 687 [2M+Na]$^+$.

5.4.3 Derivatization of bile acids

5.4.3.1 Synthesis of Diformyloxy-5β-23-iodio-24-norcholane

3α,7β-Diformyloxy-5β-23-iodio-24-norcholane and 3α,7α-diformyloxy-5β-23-iodio-24-norcholane were prepared from the corresponding bile acid according to the literature procedure [18]. A mixture of bile acid (2.5 mmol) and formic acid (4 mL) was stirred at 55°C for 24 h and concentrated *in vacuo*. The residue was crystallized by adding water to warm EtOH solution and used in the next step. A solution of diformyloxy bile acid (0.5 mmol) and DIH (228 mg, 0.6 mmol) in DCE (4 mL) was irradiated for 2 h under reflux conditions. After chromatography on silica gel (eluent, 0:100 to 50:50 EtOAc/hexane) **23-I-*nor*-UDC** and **23-I-*nor*-CDC** were obtained.
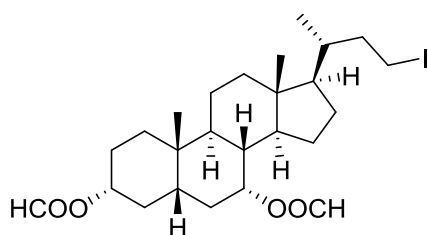
23-I-*nor*-CDC

3α,7α-Diformyloxy-5β-23-iodio-24-norcholane. Yield 90%. ^{1}H-NMR (CDCl$_3$): δ= 8.08 (s, 1H), 8.02 (s, 1H), 5.03 (br s, 1H), 4.78-4.65 (m, 1H), 3.38-3.23 (m, 1H), 3.17-3.04 (m, 1H), 2.18-1.03 (m, 24H), 0.98 (s, 3H), 0.96 (d, J= 6.2 Hz, 3H), 0.64 (s, 3H); ^{13}C-NMR (CDCl$_3$): δ =160.8 (q), 74.1 (CH), 71.4 (CH), 55.6 (CH), 50.1 (CH), 42.8 (q), 40.9 (CH), 40.2 (CH$_2$), 39.4 (CH$_2$), 37.9 (CH), 37.1 (CH), 34.8 (CH$_2$), 34.6 (CH), 34.0 (q), 33.8 (CH$_2$), 31.5 (CH$_2$), 28.0 (CH$_2$), 26.8 (CH$_2$), 23.5 (CH$_2$), 22.7 (CH$_3$), 20.6 (CH$_2$), 17.9 (CH$_3$), 11.8 (CH$_3$), 5.2 (CH$_2$). MS (ESI, ES+) m/z: calculated for C$_{25}$H$_{39}$IO$_4$ 530.49; found 553 [M+Na]$^+$.
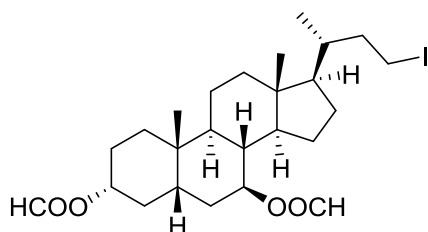
23-I-*nor*-UDC

3α,7β Diformyloxy-5β-23-iodio-24-norcholane. Yield 85%. ^{1}H-NMR (CDCl$_3$): δ= 7.99 (s, 1H), 7.97 (s, 1H), 4.94-4.75 (m, 2H), 3.33-3.23 (m, 1H), 3.12-3.02 (m, 1H),

2.05-1.12 (m, 24H), 0.98 (s, 3H), 0.91 (d, J= 6.15 Hz, 3H), 0.64 (s, 3H). ^{13}C-NMR (CDCl$_3$): δ = 161.0 (q), 160.6 (q), 73.5 (CH), 73.3 (CH), 55.2 (CH), 54.8 (CH), 43.7 (q), 42.0 (CH), 40.1 (CH$_2$), 39.8 (CH$_2$), 39.7 (CH), 39.4 (CH), 36.9 (CH), 34.3 (CH$_2$), 33.9 (q), 32.8 (CH$_2$), 32.7 (CH$_2$), 28.3 (CH$_2$), 26.3 (CH$_2$), 25.8 (CH$_2$), 23.2 (CH$_3$), 21.2 (CH$_2$), 17.86 (CH$_3$), 12.0 (CH$_3$), 5.3 (CH$_2$). MS (ESI, ES+) m/z: calculated for C$_{25}$H$_{39}$IO$_4$ 530.49; found 553 [M+Na]$^+$.

5.4.3.2 General procedure for the synthesis of nor-azides

The 23-iodo derivative (0.5 mmol) was dissolved in DMF (3 mL) and NaN$_3$ (4 mmol) was added. The reaction mixture was stirred at room temperature overnight and then poured into water (8 mL) and extracted twice with Et$_2$O (12 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered and concentrated *in vacuo* to give the diformyloxy azido-compound. The pale yellow solid was treated with 25% NaOH in MeOH at room temperature monitoring by TLC (AcOEt/cyclohexane 1:1) until disappearing of the starting material (2 h for UDC, 12 h for CDC). The corresponding dihydroxy azido derivatives **23-N$_3$-*nor*-UDC** and **23-N$_3$-*nor*-CDC** were precipitated by adding water to the solution.
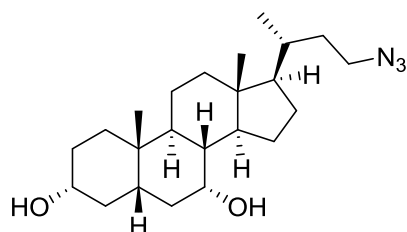
23-N$_3$-*nor*-CDC

3α,7α-Dihydroxy-5β-23-azido-24-norcholane. Amorphous white solid, yield 78%. IR: ν (cm^{-1}) 3439 (O–H), 2970-2864 (C-H), 2091 (N$_3$). ^{1}H-NMR: δ= 4.29 (d, J= 4.6 Hz, 1H), 4.11 (d, J= 3.2 Hz, 1H), 3.61 (br s, 1H), 3.44-3.38 (m, 1H), 3.29-3.08 (m, 2H), 2.22-2.09 (m, 1H), 1.95-1.58 (m, 8H), 1.48-0.96 (m, 15H), 0.89 (d, J= 6.5 Hz, 3H), 0.82 (s, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ =70.3 (CH), 66.1 (CH), 55.7 (CH), 50.0 (CH), 48.3 (CH$_2$), 42.0 (q), 41.4 (CH), 39.6 (CH$_2$), 39.3 (CH$_2$), 39.1 (CH), 35.3 (CH$_2$), 34.8 (CH$_2$), 34.7 (q), 34.2 (CH$_2$), 33.2 (CH), 32.3 (CH), 30.5 (CH$_2$), 27.9 (CH$_2$), 23.1 (CH$_2$), 22.7 (CH$_3$), 20.2 (CH$_2$), 18.2 (CH$_3$), 11.6 (CH$_3$). MS (ESI, ES+) m/z: calculated for C$_{23}$H$_{39}$N$_3$O$_2$ 389.58; found 1191 [3M+Na]$^+$.
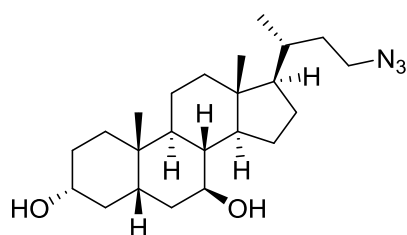
23-N$_3$-*nor*-UDC

3α,7β-Dihydroxy-5β-23-azido-24-norcholane. Amorphous white solid, yield 75%. IR: ν (cm^{-1}) 3593 (O–H), 3447 (O–H), 2970-2855 (C-H), 2086 (N$_3$). ^{1}H-NMR: δ= 4.43 (d, J= 4.5 Hz, 1H), 3.86 (d, J= 6.8 Hz, 1H), 3.42-3.31 (m, 1H), 3.29-3.18 (m, 3H), 1.98-1.58 (m, 6H), 1.51-0.91 (m, 18H), 0.89 (d, J= 6.4 Hz, 3H), 0.82 (s, 3H), 0.59 (s, 3H). ^{13}C-NMR: δ =69.7 (CH), 69.4 (CH), 55.8 (CH), 54.8 (CH), 48.3 (CH$_2$), 43.1 (q), 42.9 (CH), 42.1 (CH), 39.7 (CH$_2$), 38.7 (CH), 37.7 (CH$_2$), 37.2 (CH$_2$), 34.8 (CH$_2$), 34.3 (CH$_2$), 33.7 (q), 33.1 (CH), 30.2 (CH$_2$), 28.2 (CH$_2$), 26.7 (CH$_2$), 23.3 (CH$_3$), 20.8 (CH$_2$), 18.4 (CH$_3$), 11.9 (CH$_3$). MS (ESI, ES+) m/z: calculated for C$_{23}$H$_{39}$N$_3$O$_2$ 389.58; found 1191 [3M+Na]$^+$.
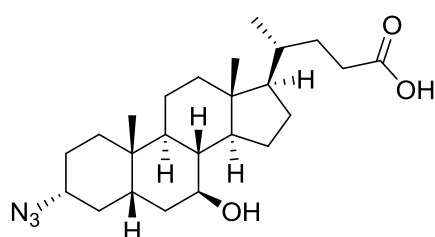
3α-N$_3$-UDCA

3α-Azido-7β-hydroxy-5β-cholic acid. LiOH 1.5 M (23 mmol; 15 mL) was added to a solution of **3α-N$_3$-UDC** (1.0 g; 2.3 mmol) in methanol (10 mL). The mixture was stirred at room temperature for 21 h. Then 2N HCl was

added until pH = 4–5 and the solution was extracted with ethyl acetate (2·20 mL). The combined organic phase was washed with water, dried over $MgSO_4$ and concentrated *in vacuo* to afford a white powder. Yield 94%. ^{1}H-NMR: δ = 11.93 (br s, 1H), 3.91 (br s, 1H), 3.41–3.22 (m, 2H), 2.25–0.82 (m, 32H), 0.60 (s, 3H); ^{13}C-NMR: δ = 174.8 (q), 69.1 (CH), 60.0 (CH), 55.4 (CH), 54.5 (CH), 42.9 (CH), 42.8 (q), 42.0 (CH), 39.4 (CH_2), 38.4 (CH), 37.2 (CH_2), 34.7 (CH), 34.5 (CH_2), 33.6 (q), 32.7 (CH_2), 30.6 (CH_2), 28.0 (CH_2), 26.6 (CH_2), 26.0 (CH_2), 23.1 (CH_3), 20.7 (CH_2), 18.2 (CH_3), 11.9 (CH_3). MS (ESI, ES+) m/z: calculated for $C_{24}H_{39}N_3O_3$ 417.59; found 440 $[M+Na]^+$.
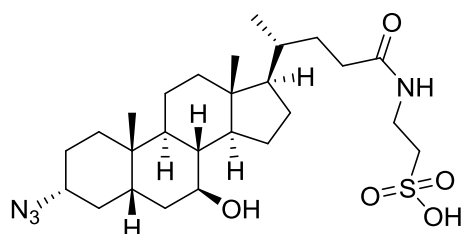
3α-N₃-TUDCA

(3α-Azido-7β-hydroxy-5β-cholan-24-oyl)-2-aminoethanesulfonic acid. To a solution of **3α-N₃-UDCA** (500 mg, 1.19 mmol) in anhydrous THF (5 mL) stirred at 0°C were added triethylamine (0.18 mL, 1.3 mmol) and ethyl chloroformate (0.13 mL, 1.3 mmol). After 2 h at room temperature a solution of taurine (136 mg, 1.3 mmol) in $NaOH/H_2O$ (1 mL, 1.43 mmol) was added. The reaction mixture was stirred at room temperature overnight and then acidified with 5% HCl to pH 1. After evaporation of THF, the mixture was diluted with water and washed with EtOAc. The aqueous phase was extracted with *n*-butanol and the organic layer dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure to give the title compound as an amorphous white solid. Yield 80%. IR: v (cm^{-1}) 3309 (O-H), 2931–2866 (C-H), 2090 (N3), 1648 (C=O). ^{1}H-NMR: δ = 7.72 (br s, 1H), 6.83–6.80 (m, 1H), 3.98–3.88 (m, 2H), 3.28–3.17 (m, 2H), 3.15–3.01 (m, 2H), 2.68–2.75 (m, 1H), 2.57–2.48 (m, 2H), 2.08–0.93(m, 23H), 0.87 (s, 3H), 0.85 (d, J = 6.4 Hz, 3H), 0.59 (s, 3H). ^{13}C-NMR: δ = 171.9 (q), 69.0 (CH), 60.0 (q), 55.4 (CH), 54.5 (CH), 50.5 (CH_2), 45.6 (CH_2), 42.9 (CH), 42.8 (q), 42.1 (CH), 40.0 (CH), 38.4 (CH), 37.2 (CH_2), 35.4 (CH_2), 34.8 (CH), 34.5 (CH_2), 33.6 (CH_2), 32.5 (CH_2), 31.4 (CH_2), 28.0 (CH_2), 26.6 (CH_2), 26.0 (CH_2), 23.1 (CH_3), 20.7 (CH_2), 18.4 (CH_3), 11.9 (CH_3). MS (ESI, ES+) m/z: calculated for $C_{26}H_{44}N_4O_5S$ 524.72; found 547 $[M+Na]^+$.

5.4.4 "Click" reactions

5.4.4.1 *General procedure for the "click" reaction*
To a solution of the appropriate alkyne **ALK-dA, ALK-A, ALK-G, ALK-dG, ALK-dU** (0.03 mmol) in 1.4 ml of a 1:1:1.5 mixture of H_2O/*tert*-BuOH /THF (*v/v*), sodium ascorbate (0.06 mmol) and copper(II) sulfate (0.012 mmol) were added. Then the appropriate azide **3α-N₃-CDC**, **3α-N₃-UDC**, **23-N₃-*nor*-CDC**, **23-N₃-*nor*-UDC** or **3α-N₃-TUDCA** (0.045 mmol) was added and the resulting solution was stirred at room temperature overnight.

5.4.4.2 *Purification of conjugates with 3α-N₃-CDC or 3α-N₃-UDC*
The mixture was concentrated under reduced pressure, added with water and extracted with dichloromethane. The organic layers was dried over Na_2SO_4, filtered and concentrated *in vacuo*. The resulting crude solid was washed three times with Et_2O.
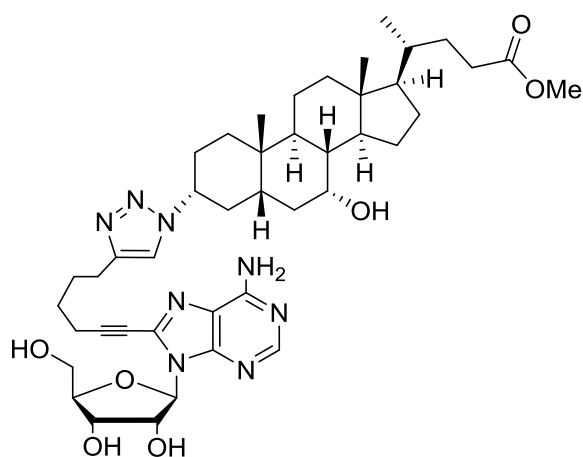
A-CDC

Colourless syrup, yield 80%. IR: v (cm^{-1}) 3418-3315 (O–H), 2928-2866 (C–H), 2241 (C≡C), 1693 (C=O), 1665-1524 (C=C,C=N). ^{1}H-NMR : δ = 8.18 (br s, 1H), 7.88 (s, 1H), 7.58 (br s, 2H), 5.98 (d, J= 6.83 Hz 1H), 5.61-5.58 (m, 1H), 5.42 (d, J= 6.25 Hz 1H), 5.20 (d, J= 4.30 Hz, 1H), 5.12-4.98 (m, 1H), 4.32-4.16 (m, 3H), 3.98 (m, 1H), 3.72-3.42 (s, 7H), 2.74-2.59 (m, 5H), 2.38-2.16 (m, 3H), 2.01-0.95 (m, 25H), 0.90 (s, 3H), 0.86 (d, J= 6.44 Hz, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ = 173.8 (q), 156.0 (q), 153.2 (CH), 148.3 (q), 146.1 (q), 134.0 (q), 120.0 (q), 119.9 (CH), 97.4 (q), 89.3 (CH), 86.6 (CH), 71.5 (CH), 71.08 (CH), 70.3 (q), 66.1 (CH), 62.2 (CH$_2$), 60.1 (CH), 55.4 (CH), 51.2 (CH$_3$), 49.9 (CH), 41.9 (CH), 41.7 (q), 40.1 (CH$_2$), 37.7 (CH$_2$), 35.4 (CH$_2$), 34.9 (q), 34.8 (CH), 34.4 (CH$_2$) , 32.2 (CH), 30.6 (CH$_2$), 30.4 (CH$_2$), 28.2 (CH$_2$), 27.7 (CH$_2$), 27.2 (CH$_2$), 27.1 (CH$_2$), 24.6 (CH$_2$), 23.1 (CH$_2$), 22.6 (CH$_3$), 20.3 (CH$_2$), 18.4 (CH$_2$), 18.1 (CH$_3$), 11.6 (CH$_3$). HRMS calculated for [C$_{43}$H$_{62}$N$_8$O$_7$+H]$^+$ 803.4814; found 803.4819.

A-UDC

Colourless syrup, yield 78%. IR: v (cm^{-1}) 3411 (O–H), 2926-2865 (C–H), 2240 (C≡C), 1693 (C=O), 1660-1524 (C=C,C=N). ^{1}H-NMR : δ = 8.18 (br s, 1H), 8.02 (s, 1H), 7.58 (br s, 2H), 5.98 (d, J= 6.83 Hz 1H), 5.61-5.58 (m, 1H), 5.42 (d, J= 6.23 Hz 1H), 5.20 (d, J= 4.31 Hz, 1H), 5.12-4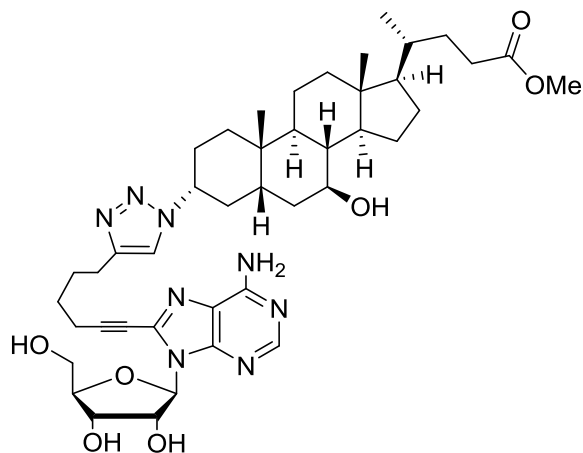.98 (m, 1H), 4.38 (br s, 1H), 4.18 (br, s, 1H), 3.98 (br, s, 1H), 3.91 (br s, 1H), 3.71-3.62 (m, 1H), 3.58-3.45 (m, 4H), 2.74-2.59 (m, 5H), 2.38-2.11 (m, 5H), 2.01-0.95 (m, 25H), 0.92 (s, 3H), 0.84 (d, J= 6.44 Hz, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ = 174.2 (q), 156.4 (q), 153.4 (CH), 148.7 (q), 142.8 (q), 133.9 (q), 120.4 (q), 120.2 (CH), 98.0 (q), 88.8 (CH), 85.6 (CH), 71.8 (q), 70.8 (CH), 66.5 (CH), 62.7 (CH$_2$), 60.5 (CH), 55.9 (CH), 51.6 (CH$_3$), 50.4 (CH), 42.4 (q), 42.1 (CH), 39.6 (CH), 38.0 (CH$_2$), 37.7 (CH$_2$), 35.9 (CH$_2$), 35.3 (CH), 35.3 (CH$_2$), 34.9 (q), 32.7 (CH$_2$), 31.1 (CH), 30.8 (CH$_2$), 28.7 (CH$_2$), 28.2 (CH$_2$), 27.6 (CH$_2$), 27.5 (CH$_2$), 25.0 (CH$_2$), 23.5 (CH$_2$), 23.1 (CH$_2$), 20.7 (CH$_3$), 18.8 (CH$_2$), 18.6 (CH$_3$), 12.1 (CH$_3$). HRMS calculated for [C$_{43}$H$_{62}$N$_8$O$_7$+H]$^+$ 803.4814; found 803.4829.

dG-CDC

Colourless syrup, yield 76%. IR: v (cm^{-1}) 3327 (O–H), 2933-2865 (C–H), 2243 (C≡C), 1692 (C=O), 1629-1568 (C=C,C=N). ^{1}H-NMR : δ = 10.82 (br s, 1H), 7.85 (s, 1H), 6.57 (br s, 2H), 6.25-6.19 (m, 1H), 5.22 (br s, 1H), 4.86 (br s, 1H), 4.4-4.16 (m, 3H), 3.8 (m, 1H), 3.6-3.58 (m, 2H), 3.59 (s, 3H), 3.57-3.42 (m, 2H), 3.15-2.99 (m, 1H), 2.71-2.62 (m, 2H), 2.53 (m, 2H), 2.38-2.01 (m, 5H), 1.92-0.99 (m, 25H), 0.89 (s, 3H), 0.86 (d, J= 6.40 Hz, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ = 173.6 (q), 156.5 (q), 153.6 (q), 150.5 (q), 146.3 (q), 130.1 (q), 121.6 (CH),
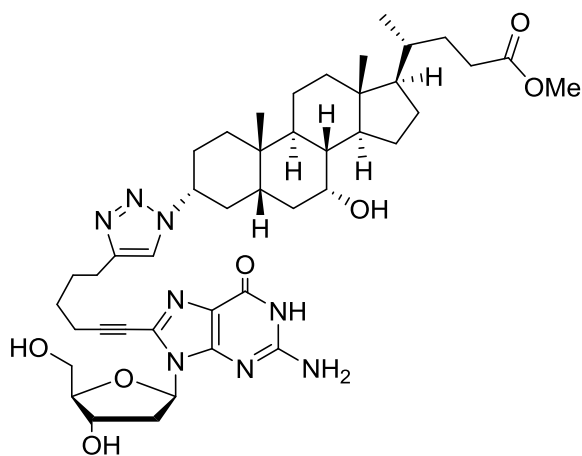
116.8 (q), 95.2 (q), 87.6 (CH), 84.6 (CH), 70.9 (q), 70.6 (CH), 70.5 (CH), 66.0 (CH), 62.0 (CH$_2$), 60.0 (CH), 55.4 (CH), 51.1 (CH$_3$), 49.8 (CH), 41.9 (CH), 41.6 (q), 40.0 (CH$_2$), 39.6 (CH), 38.8 (CH$_2$), 37.1 (CH$_2$), 35.4 (CH$_2$), 34.7 (q), 34.3 (CH$_2$), 32.1 (CH), 30.6 (CH$_2$), 30.3 (CH$_2$), 28.1 (CH$_2$), 27.7 (CH$_2$), 27.1 (CH$_2$), 24.5 (CH$_2$), 23.0 (CH$_2$), 22.6 (CH$_3$), 20.2 (CH$_2$), 18.3 (CH$_2$), 18.2 (CH$_3$), 11.6 (CH$_3$). HRMS calculated for [C$_{43}$H$_{62}$N$_8$O$_7$+H]$^+$ 803.4814; found 803.4815.
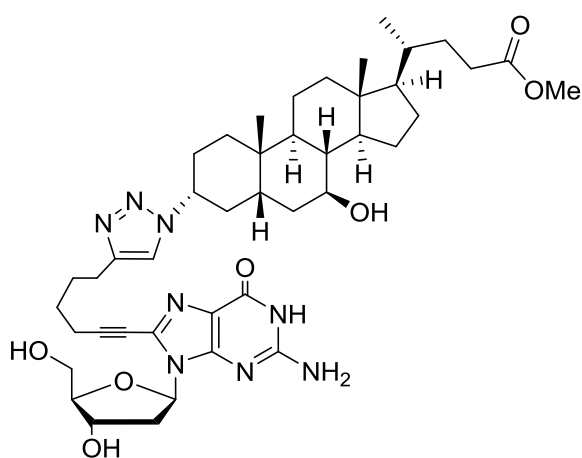
dG-UDC

Colourless syrup, yield 78%. IR: v (cm^{-1}) 3312 (O–H), 2929-2860 (C–H), 2240 (C≡C), 1686 (C=O), 1601-1565 (C=C,C=N). ^{1}H-NMR : δ = 10.78 (br s, 1H), 7.99 (s, 1H), 6.48 (br s, 2H), 6.28-6.18 (m, 1H), 5.22 (br s, 1H), 4.93-4.82 (m, 1H), 4,38 (br s, 2H), 3.91 (d, J= 6.45 Hz, 1H), 3.79-3.75 (m, 1H), 3.62-3.41 (m, 5H), 3.19-2.98 (m, 1H), 2.68-0.99 (m, 36H), 0.97 (s, 3H), 0.86 (d, J= 6.44 Hz, 3H), 0.59 (s, 3H). ^{13}C-NMR: δ = 173.7 (q), 158.3 (q), 155.9 (q), 153.6 (q), 150.5 (q), 149.9 (q), 129.5 (q), 119.8 (CH), 94.0 (q), 87.6 (CH), 83.6 (CH), 71.8 (q), 71.0 (CH), 68.9 (CH), 68.0 (CH), 62.0 (CH$_2$), 59.5 (CH), 55.2 (CH), 54.5 (CH), 51.1 (CH$_3$), 42.9 (CH), 42.4 (CH), 41.8 (q), 40.0 (CH$_2$), 37.2 (CH$_2$), 36.8 (CH$_2$), 34.9 (CH$_2$), 34.7 (CH), 34.0 (q), 33.7 (CH$_2$), 30.6 (CH$_2$), 30.2 (CH$_2$), 28.1 (CH$_2$), 27.5 (CH$_2$), 27.1 (CH$_2$), 26.5 (CH$_2$), 24.5 (CH$_2$), 23.8 (CH$_2$), 23.1 (CH$_3$), 20.8 (CH$_2$), 18.2 (CH$_2$), 18.1 (CH$_3$), 11.9 (CH$_3$). HRMS calculated for [C$_{43}$H$_{62}$N$_8$O$_7$+H]$^+$ 803.4814; found 803.4816.

G-CDC

Colourless syrup, yield 80%. IR: v (cm^{-1}) 3330 (O–H), 2918-2850 (C–H), 2238 (C≡C), 1736 (C=O), 1645-1580 (C=C,C=N). ^{1}H-NMR : δ = 10.78 (br s, 1H), 7.88 (s, 1H), 6.48 (br s, 2H), 5.78 (d, J= 6.44 Hz, 1H), 5.40 (d, J= 6.25 Hz, 1H), 5.11 (d, J= 4.88, 1H), 4.99-4.82 (m, 2H), 4,32-4.18 (m, 2H), 4.15 (br s, 1H), 3.82 (br s, 1H), 3.63-3.41 (m, 4H), 2.75-2.15 (m, 5H), 1.96-0.99 (m, 31 H), 0.88 (s, 3H), 0.84 (d, J= 6.44 Hz, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ = 173.7 (q), 157.4 (q), 155.9 (q), 153.7 (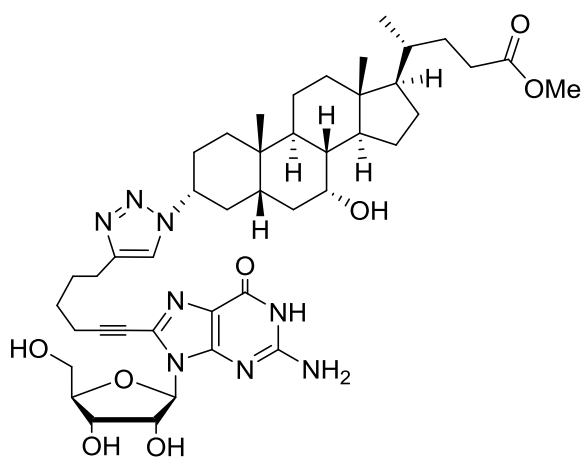q), 150.7 (q), 130.2 (q), 119.9 (CH), 116.8 (q), 95.2 (q), 88.3 (CH), 85.5 (CH), 70.9 (q), 70.6 (CH), 70.5 (CH), 66.0 (CH), 62.0

(CH$_2$), 60.0 (CH), 55.4 (CH), 51.1 (CH$_3$), 49.8 (CH), 41.9 (CH), 41.6 (q), 40.0 (CH$_2$), 39.6 (CH), 38.8 (CH$_2$), 37.1 (CH$_2$), 35.4 (CH$_2$), 34.8 (CH), 34.7 (q), 34.3 (CH$_2$), 32.1 (CH), 30.6 (CH$_2$), 30.3 (CH$_2$), 28.1 (CH$_2$), 27.7 (CH$_2$), 27.1 (CH$_2$), 24.5 (CH$_2$), 23.0 (CH$_2$), 22.5 (CH$_3$), 20.2 (CH$_2$), 18.3 (CH$_2$), 18.1 (CH$_3$), 11.6 (CH$_3$). HRMS calculated for [C$_{43}$H$_{62}$N$_8$O$_8$+H]$^+$ 819.4763; found 819.4768.
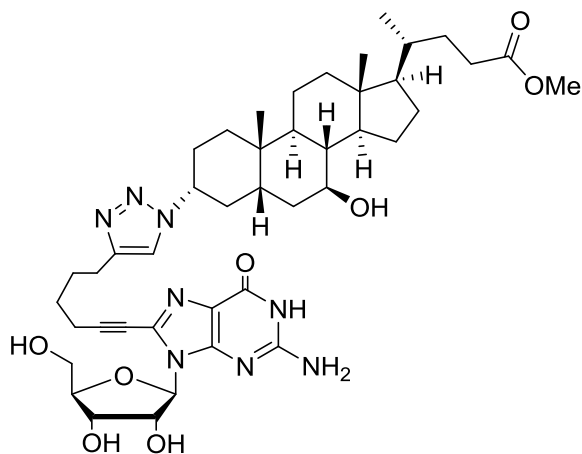

G-UDC

Light yellow syrup, yield 78%. IR: v (cm^{-1}) 3327 (O–H), 2930-2875 (C–H), 2238 (C≡C), 1735 (C=O), 1645-1572 (C=C,C=N). ^{1}H-NMR : δ = 10.78 (br s, 1H), 8.02 (s, 1H), 6.48 (br s, 2H), 5.78 (d, J= 6.43 Hz, 1H), 5.40 (d, J= 6.24 Hz, 1H), 5.11 (d, J= 4.88, 1H), 4.99-4.82 (m, 2H), 4.42-4.35(m, 1H), 4,18-4.12 (m, 1H), 3.91 (d, J= 6.64 Hz, 1H), 3.823.78 (m, 1H), 3.68-3.38 (m, 6H), 2.71-0.98 (m, 34H), 0.92 (s, 3H), 0.83 (d, J= 6.44 Hz, 3H), 0.58 (s, 3H). ^{13}C-NMR: δ= 173.5 (q), 156.0 (q), 153.7 (q), 150.7 (q), 146.3 (q), 130.2 (q), 121.6 (CH), 116.8 (q), 95.2 (q), 88.3 (CH), 85.5 (CH), 71.8 (q), 71.0 (CH), 68.9 (CH), 68.0 (CH), 62.0 (CH$_2$), 59.5 (CH), 55.2 (CH), 54.5 (CH), 51.1 (CH$_3$), 42.9 (CH), 42.4 (CH), 41.8 (q), 40.0 (CH$_2$), 37.2 (CH$_2$), 36.8 (CH$_2$), 34.9 (CH$_2$), 34.7 (CH), 34.0 (q), 33.7 (CH$_2$), 30.6 (CH$_2$), 30.2 (CH$_2$), 28.1 (CH$_2$), 27.5 (CH$_2$), 27.1 (CH$_2$), 26.5 (CH$_2$), 24.5 (CH$_2$), 23.8 (CH$_2$), 23.1 (CH$_3$), 20.8 (CH$_2$), 18.2 (CH$_2$), 18.1 (CH$_3$), 11.9 (CH$_3$). HRMS calculated for [C$_{43}$H$_{62}$N$_8$O$_7$+H]$^+$ 819.4763; found 819.4766.
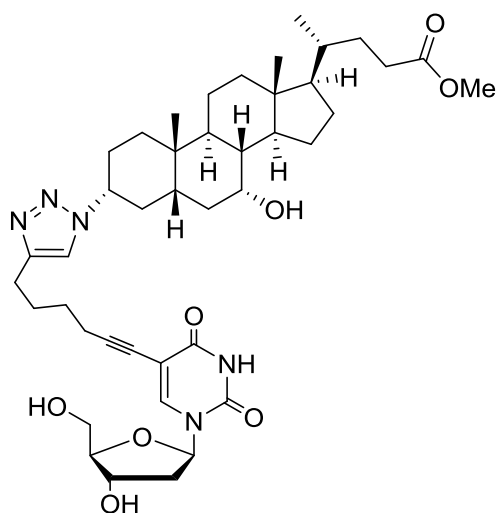
dU-CDC

Amorphous white solid, yield 90%. IR: v (cm^{-1}) 3440-3310 (O–H), 2930-2861 (C–H), 2244 (C≡C), 1693 (C=O), 1633-1565 (C=C,C=N). ^{1}H-NMR: δ = 8.62 (s, 1H), 7.81 (s, 1H), 6.42 (s, 1H), 6.17-6.12 (m, 1H), 5.28 (d, J= 4.3 Hz, 1H), 5.14-5.09 (m, 1H), 4.24-4.18 (m, 3H), 3.88 (q, J= 3.71 Hz, 1H), 3.67-3.58 (m, 3H), 3.57 (s, 3H), 2.78-2.58 (m, 5H), 2.40-2.24 (m, 2H), 2.23-2.13 (m, 1H), 2.08-1.98 (m, 1H), 1.92-0.95 (m, 27H), 0.89 (s, 3H), 0.86 (d, J= 6.44 Hz, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ = 173.8 (q), 171.2 (q), 158.1 (q), 153.8 (q), 146.1 (q), 136.8 (CH), 119.8 (CH), 106.4 (q) , 99.9 (q), 88.1 (CH), 87.4 (CH), 69.7 (CH), 66.1 (CH), 60.8 (CH), 60.1 (CH$_2$), 55.5 (CH), 51.2 (CH), 49.9 (CH$_3$), 42.0 (q), 41.7 (CH), 41.21 (CH), 37.3 (CH$_2$), 35.4 (CH$_2$), 34.9 (CH), 34.8 (CH$_2$), 34.4 (q), 32.2 (CH), 30.7 (CH$_2$), 30.4 (CH$_2$), 28.3 (CH$_2$), 27.8 (CH$_2$), 27.2 (CH$_2$), 27.1 (CH$_2$), 26.0 (CH$_2$), 24.7 (CH$_2$), 23.1 (CH$_2$), 22.6 (CH$_3$), 20.3 (CH$_2$), 18.2 (CH$_3$), 11.7 (CH$_3$). HRMS calculated for [C$_{42}$H$_{61}$N$_5$O$_8$+H]$^+$ 764.4592; found 764.4602.
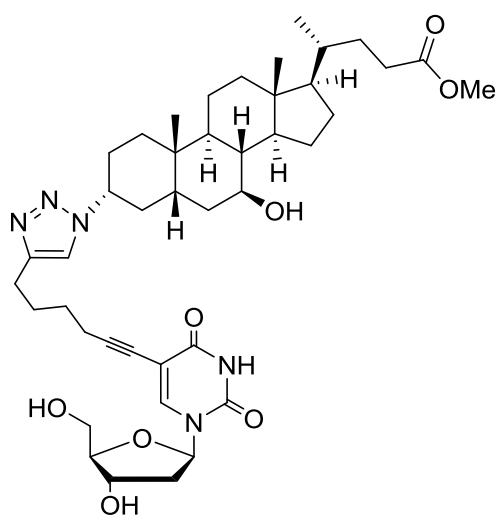
dU-UDC

Amorphous white solid, yield 70%. IR: v (cm^{-1}) 3447-3309 (O–H), 3016-2942 (C–H), 2254 (C≡C), 1690 (C=O), 1645-1522 (C=C, C=N). ^{1}H-NMR : δ = 11.58 (s, 1H), 8.18 (s, 1H), 7.98 (s, 1H), 6.12-6.07 (m, 1H), 5.28-5.21 (m, 1H), 5.18-5.15 (m, 1H), 4.42-4.28 (m, 1H), 4.21 (br s, 1H), 3.95 (d, J= 6.64 Hz, 1H), 3.78 (br s, 1H), 3.65-3.48 (m, 4H), 2.65-2.58 (m, 2H), 2.41-0.98 (m, 36H), 0.95 (s, 3H), 0.85 (d, J= 6.44 Hz, 3H), 0.62 (s, 3H). ^{13}C-NMR: δ = 175.4 (q), 171.8 (q), 162.2 (q), 149.9 (q), 146.7 (q), 143.2 (CH), 120.3 (CH), 99.4 (q), 88.0 (CH), 85.0 (CH), 80.7 (CH), 70.6 (CH), 69.5 (CH), 61.4 (CH$_2$), 60.0 (CH), 55.8 (CH), 51.6 (CH$_3$), 43.5 (CH), 43.0 (CH), 42.5 (q), 40.5 (CH$_2$), 40.0 (CH$_2$), 38.8 (CH), 37.74 (CH$_2$), 35.4 (CH$_2$), 35.2 (CH), 34.6 (CH$_2$), 34.3 (q), 31.2 (CH$_2$), 30.9 (CH$_2$), 29.5 (CH$_2$), 28.6 (CH$_2$), 28.2 (CH$_2$), 27.1 (CH$_2$), 25.1 (CH$_2$), 23.6 (CH$_3$), 21.4 (CH$_2$), 19.0 (CH$_2$), 18.7 (CH$_3$), 12.4 (CH$_3$). HRMS calculated for [C$_{42}$H$_{61}$N$_5$O$_8$+H]$^+$ 764.4592; found 764.4600.

5.4.4.3 *Purification of conjugates with 23-N$_3$-nor-CDC or 23-N$_3$-nor-UDC*

The mixture was concentrated *in vacuo* until the complete elimination of THF and *tert*-BuOH. The crude precipitated solid was filtered, washed with water, EtOH, EtOAc and finally dried with Et$_2$O.

dA-*nor*-CDC

Light yellow syrup, yield 68%. IR: v (cm^{-1}) 3440-3332 (O–H), 2926-2865 (C–H), 2242 (C≡C), 1650-1570 (C=C,C=N). ^{1}H-NMR: δ = 8.16 (br s, 1H), 7.89 (s, 1H), 7.58 (br s, 2H), 6.42-6.38 (m, 1H), 5.43-5.39 (m, 1H), 5.32 (d, J= 4.30 Hz, 1H), 4.45 (br s, 1H), 4.36-4.22 (m, 3H), 4.08 (d, J= 4.12 Hz, 1H), 3.90-3.85 (m, 1H), 3.72-3.42 (m, 3H), 3.20-3.03 (m, 2H), 2.71-2.57 (m, 4H), 2.19-2.09 (m, 2H), 1.97-0.99 (m, 27 H), 0.91 (d, J= 6.25 Hz, 3H), 0.75 (s, 3H), 0.48 (s,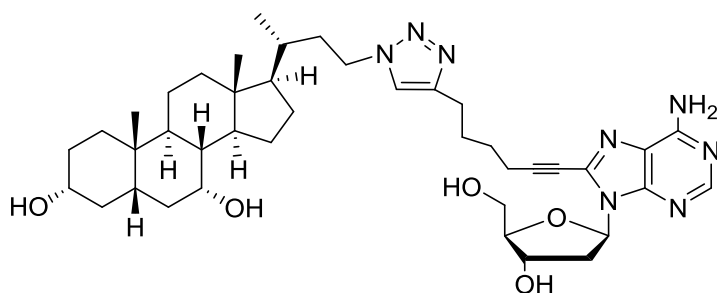 3H). ^{13}C-NMR: δ= 155.8 (q), 152.9 (CH), 148.2 (q), 146.3 (q), 133.3 (q), 121.6 (CH), 119.0 (q), 97.4 (q), 88.2 (CH), 85.1 (CH), 71.3 (CH), 70.2 (CH), 69.2 (q), 66.0 (CH), 62.1 (CH$_2$), 55.2 (CH), 50.1 (CH$_2$), 49.8 (CH), 46.8 (CH$_2$), 41.8 (q), 41.3 (CH), 40.3 (CH), 39.2 (CH$_2$), 37.5 (CH$_2$), 36.0 (CH$_2$), 35.2 (CH$_2$), 34.6 (q), 34.5 (CH$_2$), 32.9 (CH), 32.1 (CH), 30.4 (CH$_2$), 28.0 (CH$_2$), 27.7 (CH$_2$), 26.7 (CH$_2$), 24.2 (CH$_2$), 22.9 (CH$_2$), 22.5 (CH$_3$), 20.0 (CH$_2$), 18.1 (CH$_2$), 18.1 (CH$_3$), 11.3 (CH$_3$). HRMS calculated for [C$_{41}$H$_{60}$N$_8$O$_5$+H]$^+$ 745.4759; found 745.4767.
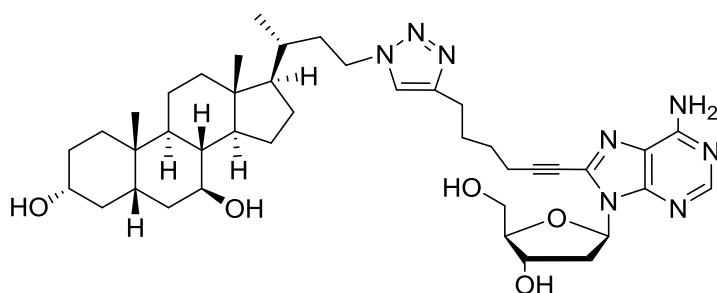
dA-*nor*-UDC

Light yellow syrup, 72%. IR: v (cm^{-1}) 3440-3317 (O–H), 2928-2861 (C–H), 2243 (C≡C), 1603-1569 (C=C,C=N). ^{1}H-NMR: δ = 8.18 (br s, 1H), 7.88 (s, 1H),

7.55 (br s, 2H), 6.42-6.39 (m, 1H), 5.42-5.39 (m, 1H), 5.37 (d, J= 4.30 Hz, 1H), 4.43 (br s, 2H), 4.38-4.20 (m, 2H), 3.92-3.81 (m, 2H), 3.70-3.61 (m, 1H), 3.53-3.42 (m, 1H), 3.38-3.18 (m, 3H), 3.17-3.04 (m, 1H), 2.72-2.57 (m, 4H), 2.20-2.11 (m, 1H), 1.97-0.99 (m, 27 H), 0.91 (d, J= 6.25 Hz, 3H), 0.81 (s, 3H), 0.52 (s, 3H). ^{13}C-NMR: δ = 156.4 (q), 153.3 (CH), 148.7 (q), 147.0 (q), 133.9 (q), 122.3 (CH), 120.0 (q), 97.9 (q), 88.8 (CH), 85.7 (CH), 71.8 (CH), 70.8 (q) 70.2 (CH), 69.8 (CH), 62.7 (CH_2), 56.2 (CH), 54.9 (CH), 48.8 (CH_2), 47.5 (CH_2), 43.5 (q), 43.4 (CH), 42.6 (CH), 40.0 (CH_2), 39.1 (CH), 38.1 (CH_2), 37.7 (CH_2), 36.7 (CH_2), 35.2 (CH_2), 34.7 (CH_2), 34.1 (q), 33.4 (CH), 30.7 ($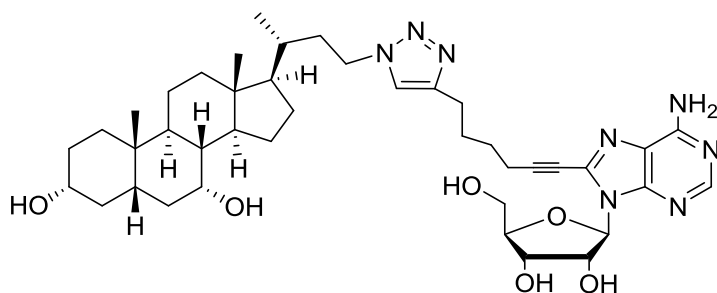CH_2$), 28.6 ($CH_2$), 27.3 ($CH_2$), 27.1 ($CH_2$), 24.8 ($CH_2$), 23.7 ($CH_3$), 21.2 ($CH_2$), 18.9 ($CH_3$), 18.8 ($CH_2$), 12.3 ($CH_3$). HRMS calculated for $[C_{41}H_{60}N_8O_5+H]^+$ 745.4759; found 745.4765.

A-*nor*-CDC

Amorphous white solid, yield 68%. IR: v (cm^{-1}) 3443-3318 (O–H), 2932-2863 (C–H), 2241 (C≡C), 1640-1570 (C=C,C=N). ^{1}H-NMR: δ = 8.12 (s, 1H), 7.87 (s, 1H), 7.58 (br s, 2H), 5.92 (d, J= 6.83 Hz, 1H), 5.60-5.55 (m, 1H), 5.41 (d, J= 6.25 Hz, 1H), 5.20 (d, J= 4.30 Hz, 1H), 5.00-4.96 (m, 1H), 4.32-4.28 (m, 2H), 4.19-4.12 (m, 1H), 4.08 (d, J= 4.12 Hz, 1H), 3.98-3.96 (m, 1H), 3.72-3.44 (m, 3H), 3.20-3.12 (m, 1H), 2.68-2.57 (m, 4H), 2.38-2.16 (m, 3H), 1.98-0.99 (m, 26H), 0.95 (d, J= 6.25 Hz, 3H), 0.79 (s, 3H), 0.52 (s, 3H). ^{13}C-NMR: δ = 155.8 (q), 152.9 (CH), 148.1 (q), 146.3 (q), 133.9 (q), 121.6 (CH), 119.0 (q), 97.2 (q), 89.2 (CH), 86.5 (CH), 71.4 (CH), 70.9 (CH), 70.2 (CH), 69.5 (q), 66.0 (CH), 62.1 (CH_2), 55.2 (CH), 49.9 (CH), 46.8 (CH_2), 41.8 (q), 41.2 (CH), 40.0 (CH), 36.1 (CH_2), 35.2 (CH_2), 34.4 (q), 34.6 (CH_2), 33.0 (CH), 32.1 (CH), 30.4 (CH_2), 28.9 (CH_2), 28.6 (CH_2), 28.0 (CH_2), 27.7 (CH_2), 26.8 (CH_2), 24.3 (CH_2), 23.0 (CH_2), 22.6 (CH_3), 20.1 (CH_2), 18.2 (CH_2), 18.2 (CH_3), 11.4 (CH_3). HRMS calculated for $[C_{41}H_{60}N_8O_6+H]^+$ 761.4708; found 761.4704.
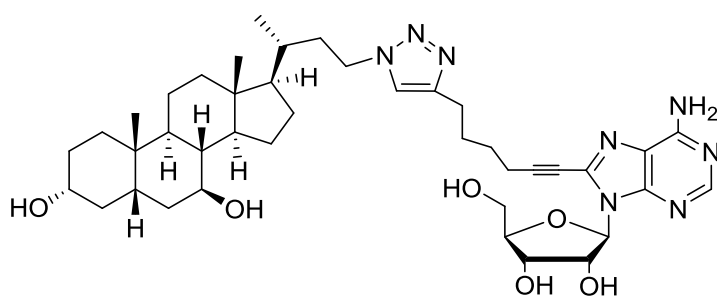
A-*nor*-UDC

Amorphous white solid, yield 71%. IR: v (cm^{-1}) 3440-3318 (O–H), 2930-2863 (C–H), 2247 (C≡C), 1640-1570 (C=C,C=N). ^{1}H-NMR: δ = 8.20 (br s, 1H), 7.89 (s, 1H), 7.58 (br s, 2H), 5.92 (d, J= 6.83 Hz, 1H), 5.68-5.59 (m, 1H), 5.41 (d, J= 6.25 Hz, 1H), 5.22 (d, J= 4.30 Hz, 1H), 5.00-4.96 (m, 1H), 4.52 (d, J= 4.30 Hz, 1H), 4.35-4.12 (m, 3H), 3.98-3.88 (m, 3H), 3.80-3.44 (m, 3H), 2.71-2.52 (m, 4H), 1.97-0.99 (m, 28H), 0.93 (d, J= 6.25 Hz, 3H), 0.82 (s, 3H), 0.53 (s, 3H). ^{13}C-NMR: δ = 156.4 (q), 153.5 (CH), 148.7 (q), 146.8 (q), 134.4 (q), 122.3 (CH), 119.5 (q), 97.8 (q), 89.7 (CH), 87.0 (CH), 72.0 (CH), 71.4 (CH), 70.7 (q), 70.1 (CH), 69.9 (CH), 62.6 (CH_2), 56.2 (CH), 54.9 (CH), 47.4 (CH_2), 43.5 (q), 43.4 (CH), 42.6 (CH), 39.1 (CH), 38.1 (CH_2), 37.7 (CH_2), 36.7 (CH_2), 35.3 (CH_2), 34.2 (q), 33.4 (CH), 30.6 (CH_2), 28.6 (CH_2), 27.3 (CH_2), 27.1 (CH_2), 24.8 (CH_2), 23.7 (CH_3), 21.2 (CH_2), 18.9 (CH_3), 18.8 (CH_2), 18.3 (CH_2), 12.3 (CH_3). HRMS calculated for $[C_{41}H_{60}N_8O_6+H]^+$ 761.4708; found 761.4715.
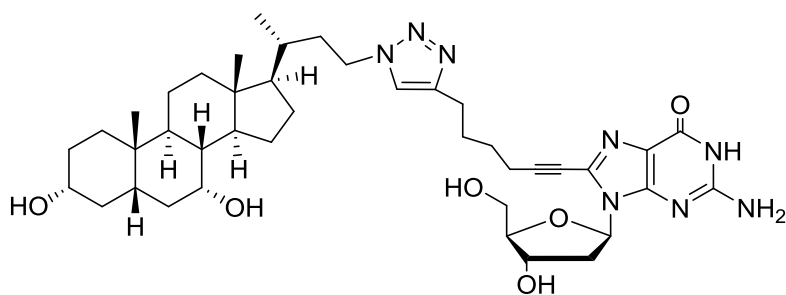
dG-*nor*-CDC

Amorphous white solid, yield 64%; IR: v (cm^{-1}) 3402-3320 (O–H), 2924-2850 (C–H), 2233 (C≡C), 1688 (C=O), 1609-1523 (C=C,C=N). ^{1}H-NMR : δ = 10.78 (br s, 1H), 7.88 (s, 1H), 6.49 (br s, 2H), 6.24-6.18 (m, 1H), 5.23 (br s, 1H), 4.91 (br s, 1H), 4.42-4.21 (m, 4H), 4.08 (br s, 1H), 3.82-3.75 (m, 1H), 3.64-3.41 (m, 3H), 3.11-2.98 (m, 2H), 2.69-2.59 (m, 2H), 2.19-0.98 (m, 31H), 0.98 (d, J= 6.25 Hz, 3H), 0.78 (s, 3H), 0.52 (s, 3H). ^{13}C-NMR: δ = 156.0 (q), 153.6 (q), 150.4 (q), 146.3 (q), 129.6 (q), 121.6 (CH), 116.8 (q), 95.1 (q), 87.6 (CH), 83.6 (CH), 71.4 (CH), 70.2 (CH), 69.9 (q), 66.0 (CH), 62.0 (CH$_2$), 55.2 (CH), 49.8 (CH), 46.8 (CH$_2$), 41.8 (q), 41.3 (CH), 40.3 (CH), 39.6 (CH$_2$), 39.4 (CH), 38.9 (CH$_2$), 36.8 (CH$_2$), 36.0 (CH$_2$), 35.2 (CH$_2$), 34.6 (CH$_2$), 32.9 (q), 32.1 (CH), 30.4 (CH$_2$), 28.0 (CH$_2$), 27.7 (CH$_2$), 26.8 (CH$_2$), 24.2 (CH$_2$), 22.9 (CH$_2$), 22.6 (CH$_3$), 20.1 (CH$_2$), 18.2 (CH$_2$), 18.1 (CH$_3$), 11.4 (CH$_3$). HRMS calculated for [C$_{41}$H$_{60}$N$_8$O$_6$+H]$^+$ 761.4708; found 761.4705.
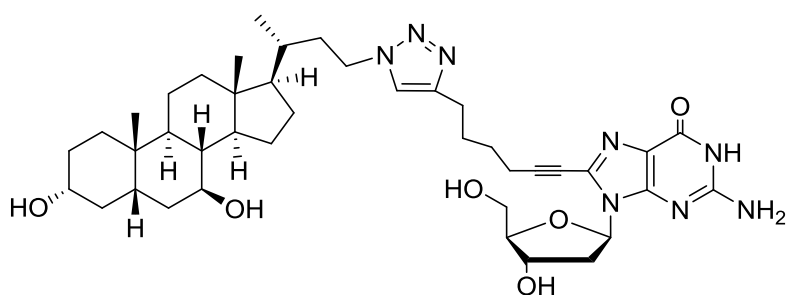
dG-*nor*-UDC

Amorphous white solid, yield 65%; IR: v (cm^{-1}) 3405-3322 (O–H), 2932-2852 (C–H), 2243 (C≡C), 1687 (C=O), 1645-1523 (C=C,C=N). ^{1}H-NMR : δ = 7.87 (s, 1H), 6.48 (br s, 2H), 6.25-6.19 (m, 1H), 5.22 (br s, 1H), 4.89 (br s, 1H), 4,48-4.22 (m, 4H), 3.90-3.75 (m, 3H), 3.62-3.41 (m, 2H), 3.09-2.98 (m, 1H), 2.75-2.62 (m, 2H), 2.32-0.99 (m, 33H), 0.98 (d, J= 6.25 Hz, 3H), 0.88 (s, 3H), 0.57 (s, 3H). ^{13}C-NMR: δ = 155.9 (q), 153.5 (q), 150.4 (q), 146.3 (q), 129.7 (q), 121.7 (CH), 116.7 (q), 95.2 (q), 87.6 (CH), 83.6 (CH), 71.0 (CH), 69.9 (q), 69.6 (CH), 69.3 (CH), 62.0 (CH$_2$), 55.7 (CH), 54.3 (CH), 46.9 (CH$_2$), 42.9 (q), 42.8 (CH), 42.0 (CH), 39.5 (CH$_2$), 38.7 (CH), 37.6 (CH$_2$), 37.1 (CH$_2$), 36.8 (CH$_2$), 36.1 (CH$_2$), 34.7 (CH$_2$), 33.6 (q), 32.8 (CH), 30.1 (CH$_2$), 28.9 (CH$_2$), 28.0 (CH$_2$), 26.8 (CH$_2$), 26.5 (CH$_2$), 24.2 (CH$_2$), 23.2 (CH$_3$), 20.7 (CH$_2$), 18.3 (CH$_3$), 18.2 (CH$_2$), 11.8 (CH$_3$). HRMS calculated for [C$_{41}$H$_{60}$N$_8$O$_6$+H]$^+$ 761.4708; found 761.4705.

G-*nor*-CDC

Light yellow syrup, yield 69%. IR: v (cm^{-1}) 3409-3310 (O–H), 2926-2857 (C–H), 2243 (C≡C), 1693 (C=O), 1640-1526 (C=C,C=N). ^{1}H-NMR: δ = 10.87 (br s, 1H), 7.89 (s, 1H), 6.52 (br s, 2H), 5.78 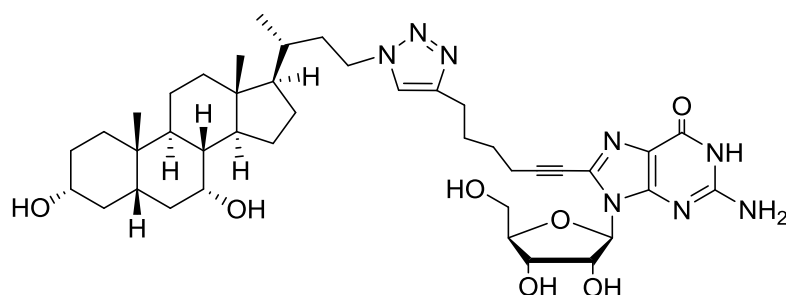(d, J= 6.44 Hz, 1H), 5.40 (d, J= 6.25 Hz, 1H), 5,05 (br s, 1H), 4.99-4.84 (m, 2H), 4.38-4.21 (m, 4H), 4.09 (br s, 2H), 3.83 (br s, 1H), 3.68-3.55 (m, 2H), 3.54-3.43 (m, 1H), 3.21-3.09 (m, 2H), 2.70-2.60 (m, 2H), 2.55-2.48 (m, 2H), 2.22-2.15 (m,

2H), 1.98-1.02 (m, 23H), 0.93 (d, J= 6.25 Hz, 3H), 0.79 (s, 3H), 0.52 (s, 3H). ^{13}C-NMR: δ = 156.0 (q), 153.6 (q), 150.7 (q), 146.4 (q), 130.3 (q), 121.6 (CH), 116.9 (q), 95.2 (q), 88.4 (CH), 85.5 (CH), 72.0 (CH), 70.2 (CH), 69.8 (q), 66.0 (CH), 62.1 (CH$_2$), 55.5 (CH), 49.9 (CH), 48.2 (CH$_2$), 41.9 (CH), 41.7 (q), 41.3 (CH), 41.0 (CH), 39.6 (CH$_2$), 36.1 (CH), 35.2 (CH$_2$), 34.6 (CH$_2$), 34.1 (CH$_2$), 33.1 (q), 32.1 (CH), 30.4 (CH$_2$), 28.0 (CH$_2$), 27.8 (CH$_2$), 27.0 (CH$_2$), 26.5 (CH$_2$), 24.3 (CH$_2$), 23.0 (CH$_2$), 22.6 (CH$_3$), 20.1 (CH$_2$), 18.1 (CH$_3$), 18.0 (CH$_2$), 11.5 (CH$_3$). HRMS calculated for [C$_{41}$H$_{60}$N$_8$O$_7$+H]$^+$ 777.4657; found 777.4657.
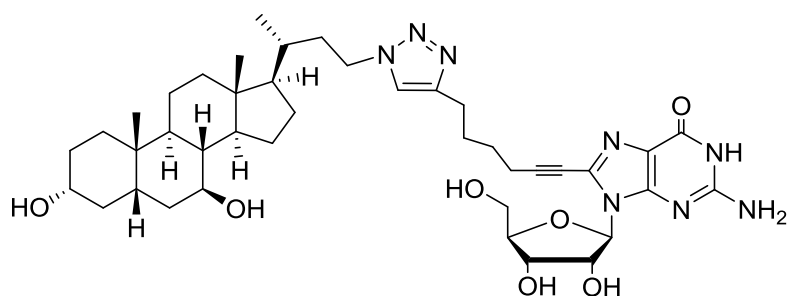
G-*nor*-UDC

Light yellow syrup, yield 68%. IR: v (cm^{-1}) 3413-3314 (O–H), 2930-2860 (C–H), 2238 (C≡C), 1693 (C=O), 1640-1526 (C=C,C=N). ^{1}H-NMR: δ = 10.79 (br s, 1H), 7.88 (s, 1H), 6.50 (br s, 2H), 5.78 (d, J= 6.44 Hz, 1H), 5.40 (d, J= 6.25 Hz, 1H), 5,05 (br s, 1H), 4.99-4.83 (m, 2H), 4.42 (br s, 1H), 4.39-4.18 (m, 2H), 4.04 (br s, 1H), 3.98-3.78 (m, 2H), 3.72-3.58 (m, 1H), 3.57-3.42 (m, 1H), 2.78-2.45 (m, 4H), 1.98-0.99 (m, 30H), 0.98 (d, J= 6.25 Hz, 3H), 0.82 (s, 3H), 0.58 (s, 3H). ^{13}C-NMR: δ = 156.0 (q), 153.8 (q), 150.7 (q), 146.4 (q), 130.3 (q), 121.8 (CH), 116.9 (q), 95.1 (q), 88.4 (CH), 85.6 (CH), 71.0 (q), 70.7 (CH), 70.6 (CH), 69.7 (CH), 69.4 (CH), 62.1 (CH$_2$), 55.8 (CH), 54.5 (CH), 47.0 (CH$_2$), 43.1 (q), 43.0 (CH), 42.2 (CH), 40.0 (CH$_2$), 38.7 (CH), 37.7 (CH$_2$), 37.3 (CH$_2$), 36.2 (CH$_2$), 34.8 (CH$_2$), 33.7 (q) , 33.0 (CH), 30.2 (CH$_2$), 28.2 (CH$_2$), 27.0 (CH$_2$), 26.7 (CH$_2$), 24.4 (CH$_2$), 23.3 (CH$_3$), 20.8 (CH$_2$), 18.4 (CH$_3$), 18.3 (CH$_2$), 11.9 (CH$_3$). HRMS calculated for [C$_{41}$H$_{60}$N$_8$O$_7$+H]$^+$ 776.4658; found 776.4668.
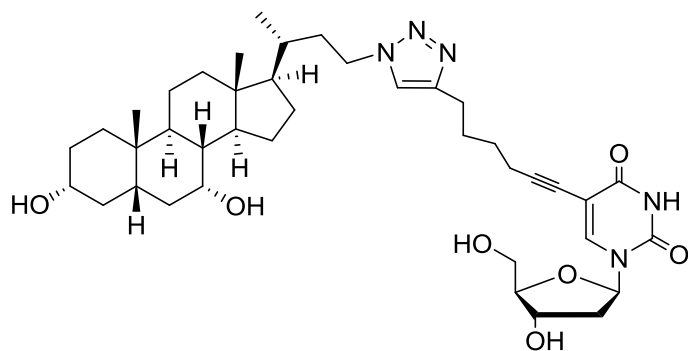
dU-*nor*-CDC

Light yellow syrup, yield 71%. IR: v (cm^{-1}) 3446 (O–H), 3100-2850 (C–H), 2255 (C≡C), 1665 (C=O), 1634-1565 (C=C,C=N). ^{1}H-NMR: δ = 8.65 (s, 1H), 7.82 (s, 1H), 6.39 (s, 1H), 6.19-6.09 (m, 1H), 5.23 (br s, 1H), 5.10 (br s, 1H), 4.38-4.17 (m, 3H), 4.05 (br s, 1H), 3.86 (br s, 1H), 3.69-3.52 (m, 3H), 3.21-3.04 (m, 1H), 2.71-2.58 (m, 3H), 2.39-2.25 (m, 1H), 2.21-2.03 (m, 1H), 2.03-0.99 (m, 30 H), 0.93 (d, J= 6.25 Hz, 3H), 0.79 (s, 3H), 0.44 (s, 3H). ^{13}C-NMR: δ = 158.0 (q), 153.6 (q), 146.3 (q), 136.4 (CH), 121.6 (CH), 106.2 (q), 99.7 (q), 88.0 (CH), 87.3 (CH), 70.2 (CH), 69.6 (CH), 69.5 (q), 66.0 (CH), 60.7 (CH$_2$), 55.2 (CH), 49.9 (CH), 46.8 (CH$_2$), 41.8 (q), 41.3 (CH), 41.1 (CH$_2$), 40.3 (CH), 39.2 (CH$_2$), 39.0 (CH$_2$), 36.0 (CH$_2$), 35.2 (CH$_2$), 34.7 (q), 34.6 (CH$_2$), 32.9 (CH), 32.1 (CH), 30.4 (CH$_2$), 28.1 (CH$_2$), 27.7 (CH$_2$), 27.0 (CH$_2$), 25.6 (CH$_2$), 24.4 (CH$_2$), 22.9 (CH$_2$), 22.6 (CH$_3$), 20.1 (CH$_2$), 18.1 (CH$_3$), 11.3 (CH$_3$). HRMS calculated for [C$_{40}$H$_{59}$N$_5$O$_7$+H]$^+$ 722.4487; found 722.4492.
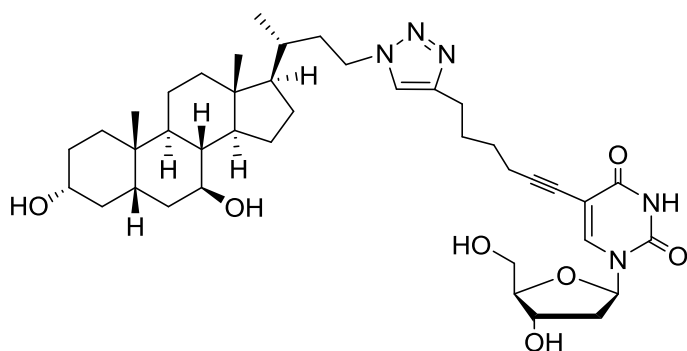
dU-*nor*-UDC

Light yellow syrup, yield 70%. IR: v (cm^{-1}) 3441 (O–H), 3016-2862 (C–H), 2255 (C≡C), 1691 (C=O), 1640-1572 (C=C,C=N). ^{1}H-NMR: δ = 11.52 (s, 1H), 8.12 (s, 1H), 7.85 (s, 1H), 6.12-6.09 (m, 1H), 5.23 (br s, 1H), 5.08 (br s, 1H), 4.42 (br s, 1H), 4.38-4.18 (m, 3H), 3.85 (br s, 1H), 3.79 (br s, 1H), 3.63-3.51 (m, 2H), 3.30-3.19 (m, 2H), 2.78-2.55 (m, 2H), 2.40-2.31 (m, 2H), 2.12-2.02 (m, 2H), 1.98-0.99 (m, 28 H), 0.92 (d, J= 6.25 Hz, 3H), 0.83 (s, 3H), 0.55 (s, 3H). ^{13}C-NMR: δ = 161.9 (q), 149.3 (q), 146.6 (q), 142.6 (CH), 121.6 (CH), 98.9 (q), 92.9 (q), 87.4 (CH), 84.4 (CH), 72.9 (q), 70.0 (CH), 69.6 (CH), 69.3 (CH), 60.8 (CH$_2$), 55.7 (CH), 54.4 (CH), 48.2 (CH$_2$), 46.9 (CH$_2$), 43.0 (q), 42.9 (CH), 42.0 (CH), 37.6 (CH$_2$), 37.1 (CH$_2$), 36.1 (CH$_2$), 34.7 (CH$_2$), 34.2 (CH$_2$), 33.6 (q), 33.0 (CH), 32.9 (CH), 30.1 (CH$_2$), 28.1 (CH$_2$), 28.0 (CH$_2$), 27.4 (CH$_2$), 26.5 (CH$_2$), 24.3 (CH$_2$), 23.2 (CH$_3$), 20.7 (CH$_2$), 18.4 (CH$_2$), 18.3 (CH$_3$), 11.8 (CH$_3$). HRMS calculated for [C$_{40}$H$_{59}$N$_5$O$_7$ +H]$^+$ 722.4487; found 722.4490.

5.4.4.4 Purification of conjugates with 3α-N$_3$-TUDCA

The mixture was concentrated under reduced pressure, added with water and extracted with *n*-butanol. The organic layers was dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo*. The crude white solid was washed twice with EtOH (10 mL) and dried with Et$_2$O.
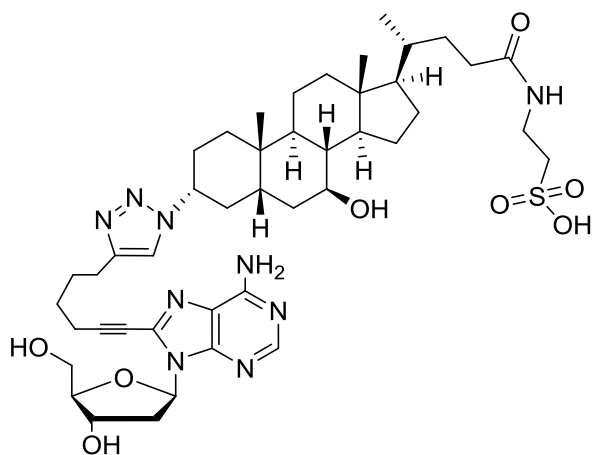
dA-TUDCA

Light yellow syrup, yield 68%. IR: v (cm^{-1}) 3340–3284 (O-H), 2918–2860 (C-H), 2240 (C≡C), 1651 (C=O), 1640–1514 (C=C, C=N). ^{1}H-NMR: δ = 8.16 (s, 1H), 8.02 (s, 1H), 7.70–7.61 (m, 1H), 7.52 (br, s, 2H), 6.43–6.38 (m, 1H), 5.42–5.25 (m, 2H), 4.48–4.12 (m, 5H), 3.95–3.82 (m, 2H), 3.72–3.59 (m, 1H), 3.55–3.41 (m, 1H), 3.40–3.21 (m, 5H), 3.19–3.12 (m, 2H), 2.85–2.71 (m, 2H), 2.69–2.58 (m, 3H), 2.28–0.99 (m, 27H), 0.93 (s, 3H), 0.88 (d, J = 6.41 Hz, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ = 155.8 (q), 152.9 (q), 148.2 (q), 146.3 (q), 133.3 (q), 121.7 (CH), 119.2 (CH), 97.4 (q), 88.2 (CH), 85.1 (CH), 72.0 (q), 71.2 (CH), 70.2 (q), 69.6 (CH), 69.3 (CH), 62.1 (CH$_2$), 55.6 (CH), 54.3 (CH), 50.1 (CH$_2$), 46.9 (CH$_2$), 42.9 (q), 42.8 (CH), 42.0 (CH), 40.3 (CH), 39.9 (CH$_2$), 39.3 (CH$_2$), 38.6 (CH), 37.6 (CH$_2$), 37.5 (CH$_2$), 37.1 (CH$_2$), 36.1 (CH$_2$), 34.7 (CH$_2$), 33.6 (q), 32.9 (CH$_2$), 30.1 (CH$_2$), 28.1 (CH$_2$), 26.7 (CH$_2$), 26.5 (CH$_2$), 24.2 (CH$_2$), 23.2 (CH$_3$), 20.7 (CH$_2$), 18.3 (CH$_3$), 18.2 (CH$_2$), 11.8 (CH$_3$). HRMS calculated for [C$_{44}$H$_{65}$N$_9$O$_8$S + H]$^+$ 880.4749; found 880.4751.
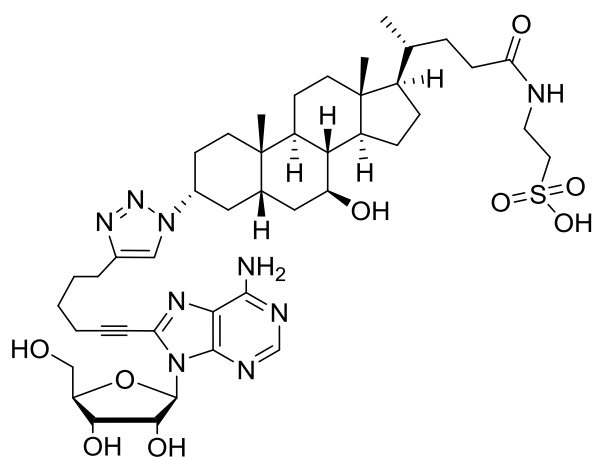
A-TUDCA

Light yellow syrup, yield 63%. IR: v (cm^{-1}) 3340–3311 (O-H), 2929–2870 (C-H), 2243 (C≡C), 1640 (C=O), 1650–1527 (C=C,C=N). ^{1}H-NMR: δ = 8.16 (s, 1H), 8.02 (s, 1H), 7.73–7.63 (m, 1H), 7.58 (br, s, 2H), 5.98 (d, J = 6.83 Hz, 1H), 5.60–5.55 (m, 1H), 5.43 (d, J = 6.25 Hz, 1H), 5.22 (d, J = 4.58 Hz, 1H), 5.02–4.97 (m, 1H), 4.48–4.31 (m, 1H), 4.21–4.12 (m, 1H), 3.98–3.87 (m, 2H), 3.72–3.61 (m, 1H), 3.58–3.45 (m, 1H), 3.33–3.23 (m, 3H), 2.69–2.55 (m, 3H), 2.15–0.99 (m, 35H), 0.92 (s, 3H), 0.88 (d, J = 6.41 Hz, 3H), 0.60 (s, 3H). ^{13}C-NMR: δ = 172.6 (q), 156.4 (q), 153.4 (CH), 148.7 (q), 146.6 (q), 134.4 (q), 120.1 (CH), 97.8 (q), 89.7 (CH), 87.0 (CH), 72.0 (q), 71.9 (CH), 71.5 (CH), 70.7 (q), 69.5 (CH), 62.7 (CH$_2$), 60.0 (CH), 55.8 (CH), 55.2 (CH), 51.0 (CH$_2$), 43.5 (CH), 43.0 (CH), 42.0 (q), 40.0 (CH$_2$), 38.8 (CH), 37.7 (CH$_2$), 35.9 (CH$_2$), 35.6 (CH$_2$), 35.4 (CH), 34.6 (CH$_2$), 34.2 (q), 33.0 (CH$_2$), 32.0 (CH$_2$), 28.6 (CH$_2$), 28.1 (CH$_2$), 27.5 (CH$_2$), 27.1 (CH$_2$), 25.1 (CH$_2$), 23.6 (CH$_3$), 21.4 (CH$_2$), 18.9 (CH$_3$), 18.8 (CH$_2$), 18.3 (CH$_2$), 12.5 (CH$_3$). HRMS calculated for [C$_{44}$H$_{65}$N$_9$O$_9$S + H]$^+$ 896.4698; found 896.4701.
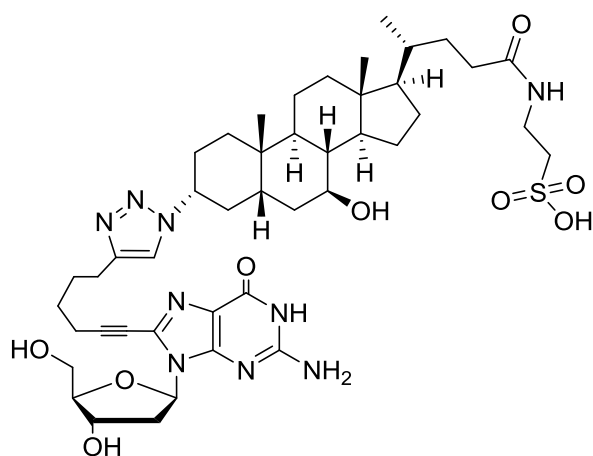
dG-TUDCA

Yield 65%. IR: v (cm^{-1}) 3397 (O-H), 2930–2863 (C-H), 2254 (C≡C), 1690, 1646 (C=O), 1638–1552 (C=C,C=N). ^{1}H-NMR: δ = 11.18 (s, 1H), 8.02 (s, 1H), 7.79–7.70 (m, 1H), 6.83 (br, s, 2H), 6.25–6.18 (m, 1H), 5.35 (d, J = 4.31 Hz, 1H), 4.99 (br, s, 1H), 4.38 (br, s, 1H), 3.92 (d, J = 6.42 Hz, 1H), 3.78 (br, s, 1H), 3.63–3.55 (m, 1H), 3.50–3.38 (m, 1H), 3.30–3.20 (m, 4H), 3.05–2.97 (m, 1H), 2.68–2.58 (m, 2H), 2.56–2.49 (m, 4H), 2.12–0.98 (m, 32 H), 0.91 (s, 3H), 0.84 (d, J = 6.41 Hz, 3H), 0.59 (s, 3H). ^{13}C-NMR: δ = 172.6 (q), 156.5 (q), 154.3 (q), 151.0 (q), 146.7 (q), 132.2 (q), 130.2 (q), 120.3 (CH), 117.2 (q), 95.7 (q), 88.1 (CH), 84.1 (CH), 71.6 (CH), 69.6 (q), 69.5 (CH), 62.5 (CH$_2$), 60.1 (CH), 55.7 (CH), 55.1 (CH), 51.1 (CH$_2$), 43.5 (CH), 43.03 (CH), 38.8 (CH), 37.7 (CH$_2$), 37.4 (CH$_2$), 35.9 (CH$_2$), 35.5 (CH$_2$), 35.3 (CH), 34.6 (CH$_2$), 34.2 (q), 33.0 (CH$_2$), 32.9 (CH$_2$), 32.0 (CH$_2$), 28.6 (CH$_2$), 28.1 (CH$_2$), 27.6 (CH$_2$), 27.1 (CH$_2$), 25.1 (CH$_2$), 23.6 (CH$_3$), 22.5 (CH$_2$), 21.4 (CH$_2$), 18.9 (CH$_3$), 18.8 (CH$_2$), 12.5 (CH$_3$). HRMS calculated for [C$_{44}$H$_{65}$N$_9$O$_9$S + H]+ 896.4698; found 896.4703.

G-TUDCA

Yield 60%. IR: v (cm^{-1}) 3329 (O-H), 2933–2867 (C-H), 2254 (C≡C), 1694, 1648 (C=O), 1640–1518 (C=C,C=N). ^{1}H-NMR: δ = 11.13 (s, 1H), 8.02 (s, 1H), 7.75 (br, s, 1H), 6.85 (s, 2H), 5.75 (d, J = 6.10 Hz, 1H), 5.43 (d, J = 6.11 Hz, 1H), 5,11 (d, J = 4.89 Hz, 1H), 4.92–4.97 (m, 1H), 4.88 (q, J = 5.80 Hz, 1H), 4.38 (br, s, 1H), 4.11 (br, s, 1H), 3.92 (d, J = 6.41 Hz, 1H), 3.82 (br, s, 1H), 3.65–3.58 (m, 1H), 3.52–3.38 (m, 1H), 3.33–3.23 (m, 4H), 2.66–2.63 (m, 2H),
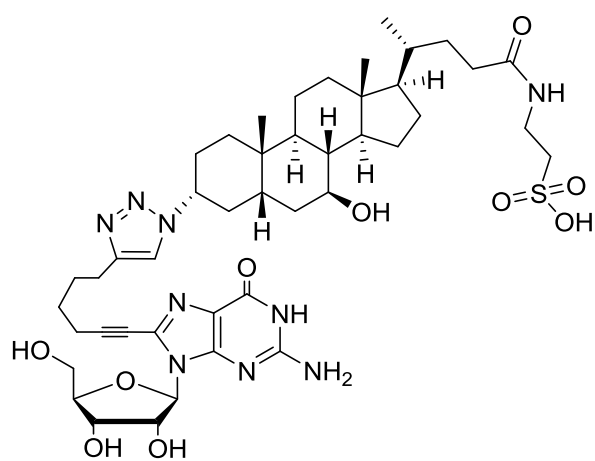
2.56–2.49 (m, 4H), 2.09–0.98 (m, 30H), 0.92 (s, 3H), 0.86 (d, J = 6.41 Hz, 3H), 0.59 (s, 3H). ^{13}C-NMR: δ = 172.2 (q), 155.9 (q), 154.1 (q), 150.7 (q), 146.2 (q), 130.1 (q), 129.5 (q) 119.9 (CH), 116.9 (q), 95.1 (q), 88.4 (CH), 85.5 (CH), 71.1 (q), 70.8 (CH), 70.6 (CH), 69.0 (CH), 62.0 (CH$_2$), 59.6 (CH), 55.3 (CH), 54.6 (CH), 50.6 (CH$_2$), 43.0 (CH), 42.6 (q), 38.3 (CH$_2$), 38.2 (CH), 37.3 (CH$_2$), 37.2 (CH), 35.42 (CH$_2$), 34.9 (CH$_2$), 34.1 (q), 33.8 (CH$_2$), 32.6 (CH$_2$) 32.5 (CH$_2$), 31.5 (CH$_2$), 28.2 (CH$_2$), 27.7 (CH$_2$), 27.2 (CH$_2$), 26.7 (CH$_2$), 24.6 (CH$_2$), 23.17 (CH$_3$), 20.9 (CH$_2$), 18.5 (CH$_3$), 18.4 (CH$_2$), 12.0 (CH$_3$). HRMS calculated for [C$_{44}$H$_{65}$N$_9$O$_{10}$S + H]$^+$ 912.4647; found 912.4650.

dU-TUDCA

Yield 65%. IR: ν (cm^{-1}) 3442 (O-H), 3015–2860 (C-H), 1690, 1668 (C=O), 1640–1520 (C=C,C=N). ^{1}H-NMR: δ = 8.12 (s, 1H), 8.01 (s, 1H), 7.69–7.61 (m, 2H), 6.32–6.18 (m, 1H), 5.23 (d, J = 4.32 Hz, 1H), 5.19–5.11 (m, 1H), 4.42–4.32 (m, 1H), 4.28–4.19 (m, 2H), 3.91 (d, J = 6.40 Hz, 1H), 3.78 (br, s, 1H), 3.63–3.55 (m, 2H), 3.50–3.38 (m, 2H), 3.30–3.20 (m, 4H), 3.11–3.02 (m, 2H), 2.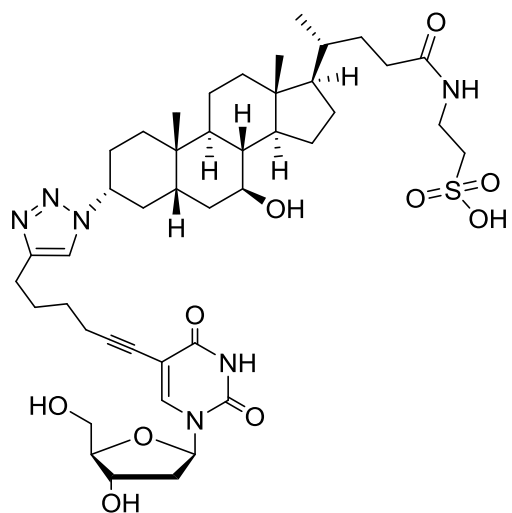65–2.54 (m, 2H), 2.42–2.34 (m, 2H), 2.12–1.08 (m, 29 H), 0.92 (s, 3H), 0.88 (d, J = 6.41 Hz, 3H), 0.61 (s, 3H). ^{13}C-NMR: δ = 175.4 (q), 172.60 (q), 161.9 (q), 149.3 (q), 146.6 (q), 143.2 (CH), 120.3 (CH), 99.4 (q), 93.5 (q), 88.0 (CH), 85.0 (CH), 80.7 (CH), 71.6 (CH), 69.6 (CH), 61.4 (CH$_2$), 60.1 (CH), 55.1 (CH), 51.1 (CH), 43.5 (CH), 43.0 (q), 42.6 (CH$_2$), 40.0 (CH$_2$), 38.8 (CH), 37.74 (CH$_2$), 35.5 (CH$_2$), 35.3 (CH), 34.6 (CH$_2$), 33.6 (q), 33.8 (CH$_2$), 33.0 (CH$_2$), 32.9 (CH$_2$), 28.6 (CH$_2$), 28.1 (CH$_2$), 27.7 (CH$_2$), 27.1 (CH$_2$), 26.7 (CH$_2$), 25.1 (CH$_3$), 24.6 (CH$_2$), 23.6 (CH$_2$), 21.4 (CH$_2$), 18.9 (CH$_3$), 18.8 (CH$_2$), 12.4 (CH$_3$). MS (ESI, ES+) m/z: calculated for [C$_{43}$H$_{64}$N$_6$O$_{10}$S + Na]$^+$ 880.07; found 880.56.

5.4.5 Biological assay

5.4.5.1 Cell Lines and Culture

Cell growth inhibition assays were carried out using the leukemia cell line K562 and colon carcinoma HCT116. Cell lines were obtained from ATCC and maintained in RPMI 1640, supplemented with 10% fetal bovine serum (FBS), penicillin (100 Units mL^{-1}), streptomycin (100 μg mL^{-1}) and glutamine (2 mM) (complete medium); the pH of the medium was 7.2 and the incubation was performed at 37°C in a 5% CO$_2$ atmosphere. Adherent cells were routinely used at 70% of confluence and passaged every 3 days by treatment with 0.05% Trypsin-EDTA (Lonza). K562 cells were routinely fed every 3 days.

5.4.5.2 Evaluation of Anti-Proliferative Activity (MTT Assay)

The antiproliferative activity of the compounds was tested using the 3-(4,5-dimethylthiozol-2-yl)2,5-diphenyltetrazolium bromide solution (MTT) assay. K562 and HCT116 were seeded in triplicate in 96-well trays respectively at the density of $5 \cdot 10^3$ and 10^4 in 50 µL of complete medium. Stock solutions (50 mM) of each compound were made in DMSO and diluted in complete medium to give final concentrations of 50, 25 and 10 µM. Untreated cells were placed in every plate as a negative control. The cells were exposed to the compounds, in 100 µL total volume, for 72 h.

5.4.5.3 Evaluation of Percentage of Apoptosis (Annexin V Staining)

The percentage of apoptotic cells was assessed using the Annexin V assay (Clontech Laboratories, Inc. A Takara Bio Company). Propidium iodide (PI) was used to avoid necrotic cell detection (Annexin-/PI+). The drug-induced apoptotic rate (Annexin+/PI- and Annexin V+/PI+) was compared with the apoptosis in the absence of the drugs used as control (spontaneous apoptosis). K562 cells were cultured in RPMI + 10% FBS in a 6-weels plate for 24 h in the presence of compounds **dA-*nor*-CDC** at concentration of 50 and 25 µM. Cells were washed once with saline buffer (PBS) and resuspended in 250 µL of Binding Buffer 1· containing 100 ng of FITC-labeled annexin V and in control sample, 500 ng of PI. Incubation with annexin V for 15 min on ice in dark was directly followed by flow cytometric analysis of the cells with a FACScan (Becton Dickinson) at 488 nm and quantified using the *Cell Quest Pro* software (Becton Dickinson).

5.5 Bibliography

1. Nobili, S.; Lippi, D.; Witort, E.; Donnini, M.; Bausi, L.; Mini, E.; Capaccioli, S. Natural compounds for cancer treatment and prevention. *Pharmacol. Res.* **2009,** *59,* 365-422.

2. Jordeim, L. P.; Durantel, D.; Zoulim, F.; Dumontet, C. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat. Rev. Drug Discov.* **2013,** *12,* 447-464.

3. Clarke, M. L.; Mackey, J. R.; Baldwin, S. A.; Young, J. D.; Cass, C. E. *Clinically relevant resistance in cancer chemotherapy.;* Springer: Cham, Switzerland, **2002**; Vol. 112, pp 27-47.

4. Johnson, Z. L.; Lee, J. H.; Lee, K.; Lee, M.; Kwon, D. Y.; Hong, J.; Lee, S. Y. Structural basis of nucleoside and nucleoside drug selectivity by concentrative nucleoside transporters. *eLife* **2014,** *3,* e03604.

5. Navacchia, M. L.; Fraix, A.; Chinaglia, N.; Gallerani, E.; Perrone, D.; Cardile, V.; Graziano, A. C. E.; Capobianco, M. L.; Sortino, S. NO Photoreleaser-Deoxyadenosine and -Bile Acid Derivative Bioconjugates as Novel Potential Photochemotherapeutics. *ACS Med. Chem. Lett.* **2016,** *7,* 939−943.

6. Perrone, D.; Bortolini, O.; Fogagnolo, M.; Marchesi, E.; Mari, L.; Massarenti, C.; Navacchia, M. L.; Sforza, F.; Varani, K.; Capobianco, M. L. Synthesis and in vitro cytotoxicity of deoxyadenosine–bile acid conjugates linked with 1,2,3-triazole. *New. J. Chem.* **2013,** *37,* 3559–3567.

7. Degirolamo, C.; Modica, S.; Palasciano, G.; Moschetta, A. Bile acids and colon cancer:

solving the puzzle with nuclear receptors. *Trends Mol. Med.* **2011,** *17,* 564-572.

8. Hofmann, A. F.; Hagey, L. R. Bile acids: chemistry, pathochemistry, biology, pathobiology, and therapeutics. *Cell. Mod. Life Sci.* **2008,** *65,* 2461-2483.

9. Kramer, W. Transporters, trojan horses and therapeutics: suitability of bile acid and peptide transporters for drug delivery. *Biol. Chem.* **2011,** *392,* 77-94.

10. Nurunnabi, M. D.; Khatun, Z.; Revuri, V.; Nafiujjaman, M. D.; Cha, S.; Cho, S.; Moo Huh, K.; Lee, Y. Synthesis and anti-proliferative activity evaluation of novel 1,4-naphthoquinones possessing pyrido [2,3-d]pyrimidine scaffolds. *RSC Adv.* **2016,** *6*.

11. Dalpiaz, A.; Paganetto, G.; Pavan, B.; Fogagnolo, M.; Medici, A.; Beggiato, S.; Perrone, D. Zidovudine and Ursodeoxycholic Acid Conjugation: Design of a New Prodrug Potentially Able To Bypass the Active Efflux Transport Systems of the Central Nervous System. *Molecular Pharmaceutics* **2012,** *9* (4), 957-968.

12. Nguyen, B. C.; Takahashi, H.; Uto, Y.; Shahinozzaman, M. D.; Tawata, S.; Maruta, H. 1,2,3-Triazolyl ester of Ketorolac: a "click chemistry"-based highly potent PAK1-blocking cancer-killer. *Eur. J. Med. Chem.* **2017,** *126,* 270-276.

13. Kolb, H. C.; Sharpless, K. B. The growing impact of click chemistry on drug discovery. *Drug Discov. Today* **2003,** *8,* 1128-1137.

14. Dalvie, D. K.; Kalgutkar, A. S.; Khojasteh-Bakht, S. C.; Obach, R. S.; O'Donnell, J. P. Biotransformation reactions of five-membered aromatic heterocycling rings. *Chem. Res. Toxicol.* **2002,** *15,* 269-299.

15. Horne, W. S.; Yadav, M. K.; Stout, C. D.; Ghadiri, M. R. Heterocyclic peptide backbone modifications in an α-helical coiled coil. *J. Am. Chem. Soc.* **2004,** *126,* 15366-15367.

16. Sàgi, G.; Otvos, L.; Ikeda, S.; Andrei, G.; Snoeck, R.; Clercqt, E. D. Synthesis and Antiviral Activities of 8-Alkynyl-, 8-Alkenyl-, and 8-Alkyl-20-deoxyadenosine Analogs. *J. Med. Chem.* **1994,** *37,* 1307–1311.

17. Massarenti, C.; Bortolini, O.; Fantin, G.; Cristofaro, D.; Ragno, D.; Perrone, D.; Marchesi, E.; Toniolo, G.; Massi, A. Fluorous-tag assisted synthesis of bile acid–bisphosphonate conjugates via orthogonal click reactions: An access to potential anti-resorption bone drugs. *Org. Biomol. Chem.* **2017,** *15,* 4907–4920.

18. Kulbitski, K.; Nisnevich, G.; Gandelman, M. Metal-Free Efficient, General and Facile Iododecarboxylation Method with Biodegradable Co-Products. *Adv. Synth. Catal.* **2011,** *353,* 1438–1442.

19. Navacchia, M. L.; Marchesi, E.; Mari, L.; Chinaglia, N.; Gallerani, E.; Gavioli, R.; Capobianco, M. L.; Perrone, D. Rational Design of Nucleoside–Bile Acid Conjugates Incorporating a Triazole Moiety for Anticancer Evaluation and SAR Exploration. *Molecules* **2017,** *22,* 1710.

6   Introduction to oligonucleotide chemistry

6.1   Structure of oligonucleotides

Oligonucleotides (ONs) are short sequences of nucleotides, the same monomers that make up the nucleic acids DNA and RNA. Generally they are composed of a few tens of nucleotides, but they can reach up to a few hundreds.

The nucleotides are constituted of three parts: a nitrogenous base, a pentose sugar and a phosphate group [1,2]. There are two families of nitrogenous bases: the purines and the pyrimidines. The pyrimidines are characterized by an hexatomic ring consisting of carbon and nitrogen atoms; cytosine (C), thymine (T) and uracil (U) belong to this category. The purines are adenine (A) and guanine (G); they are constituted by a pentatomic ring fused with a pyrimidinic ring. The various types of purines and pyrimidines differ for the functional groups linked to the rings. A clarification to make is that the thymine is found only in the DNA while the uracil is only in the RNA. Furthermore, in RNA, sugar is ribose, while in DNA it is deoxyribose. The only difference between the two sugars is that deoxyribose is devoid of the hydroxyl group in position 2'. The union between the sugar and a nitrogenous base, through a β-N-glycosidic bond, leads to the formation of a nucleoside; if a phosphate group also binds to the carbon atom in 5' of the sugar, the molecule is called nucleotide or nucleoside monophosphate. The nucleotides are bonded together by covalent bonds called phosphodiester bonds, present between the nucleoside 5′-phosphate and the 3′-hydroxyl group of the adjacent nucleotide. In this way a sugar-phosphate skeleton is formed, from which the nitrogenous bases protrude [3] (Figure 6.1).
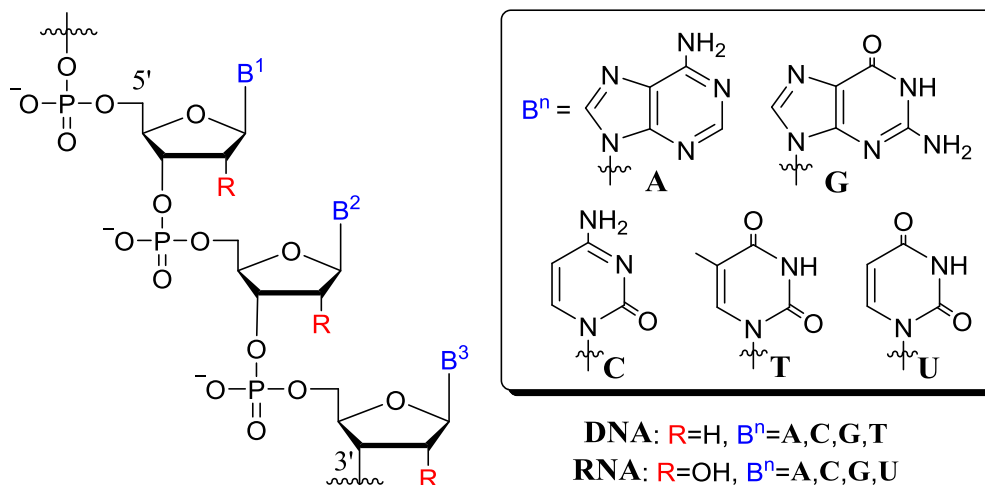

Figure 6.1 *General structure of oligonucleotides.*

6.2   Modulation of gene expression by oligonucleotides

According to the DNA double helix model proposed by Watson and Crick [4], the nitrogen bases of a filament are coupled with those of the complementary strand by hydrogen bonds. More precisely A can only be paired with T, while G only with C [5] (Figure 6.2).
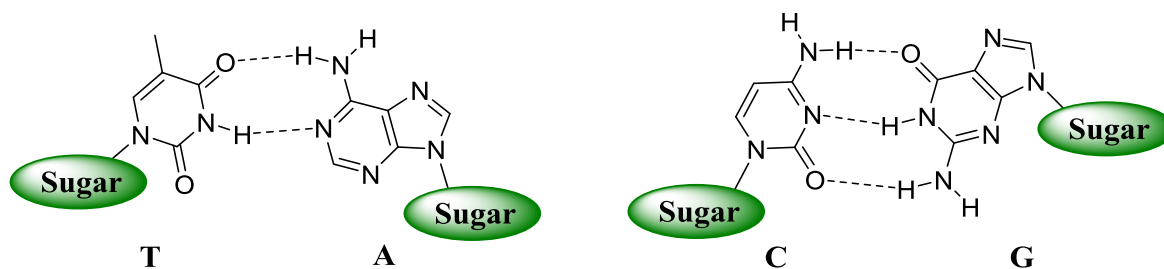
Figure 6.2 *Base pairing according to the model of Watson and Crick.*

The study of nucleic acids and their interactions with proteins is at the base of a branch of biology, called molecular biology. The dogma of this discipline is that starting from the DNA, following its transcription, it is synthesized the messenger RNA (mRNA) that before leaving the nucleus, in eukaryotes, is subjected to processes such as the attack of a cap at the 5′-end, of a poly-A tail at the 3′-end and, finally, the splicing. In the cytoplasm, from the mature mRNA, proteins are synthesized through the translation process [1,2] (Scheme 6.1A).

The relationship between genes and proteins that led to the enunciation of the dogma was proposed for the first time, at the beginning of the 1900s, by Garrod, to be then developed and confirmed by Beadle and Tatum in the forties of the twentieth century [6].

Most of the drugs used today go to act on the last stage of the process just described, or on the enzymes (proteins) that perform fundamental functions in our body. In recent years numerous studies have been developed with the aim of creating molecules that act on different targets than proteins: some of these molecules are precisely the oligonucleotides, which are able to bind to DNA and RNA in a specific way, and which act through different mechanisms of action, called "code blockers" [7,8,9,10].
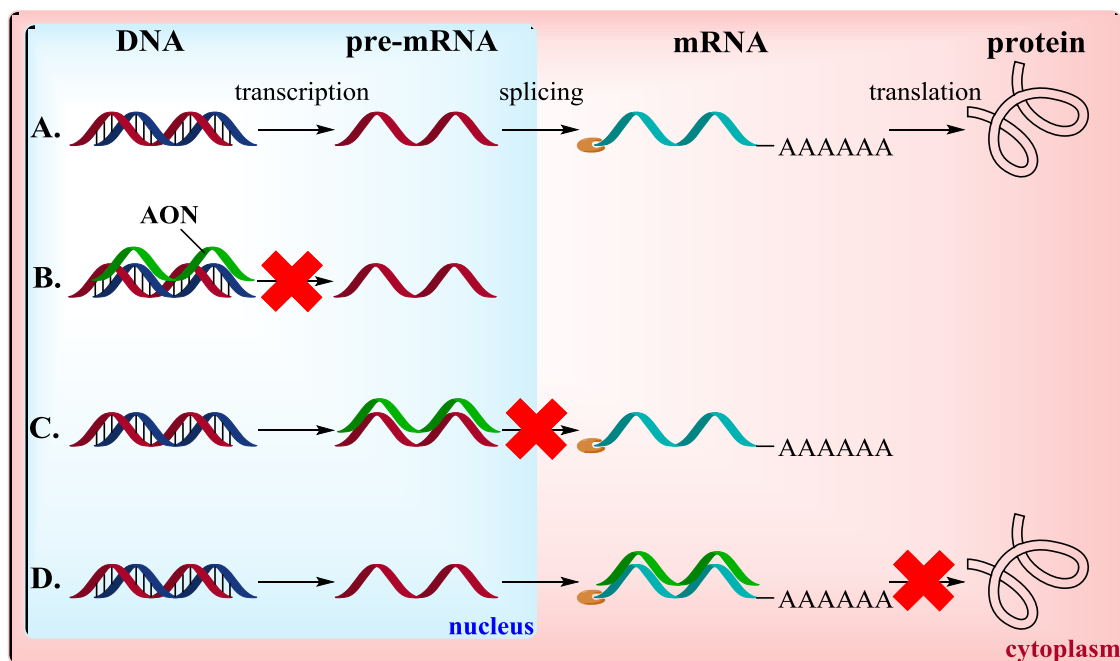
Scheme 6.1 *Normal process of protein synthesis from the DNA through transcription, splicing and translation (A) and mechanisms of action of AON with steric block (B, C and D).*

The first mechanism, developed in 1978, is antisense technology [11,12,13], which consists in the synthesis of single-stranded oligonucleotides able to bind to certain sequences of DNA or RNA complementary to them, which act as targets, through Watson-Crick bonds. The antisense oligonucleotides (abbreviated as AONs or ASOs) are able to inhibit the expression

of the target gene thanks to their ability to hybridize with these "sense" sequences, which in the absence of AON would normally be transcribed or translated.

There are four main mechanisms of action of the AONs: the steric block of the transcription, splicing or translation mechanisms [7], the activation of RNase H [8], the gene silencing induced by siRNA (small interfering RNA) [14] and the alternative splicing [3,15].

The target DNA or RNA sequences may belong to endogenous cells containing mutations and may therefore be harmful to the organism or belong to tumor cells.

6.2.1 Steric block of transcription, splicing or translation

The block due to a steric encumbrance occurs inside the nucleus when the AON hybridizes with the DNA and blocks the transcription, or when it hybridizes with the pre-mRNA thus preventing the splicing and the synthesis of the mature mRNA (Scheme 6.1B-C). In the cytoplasm, the AON binds to the mRNA site where ribosome binding occurs: in this way the ribosome is not able to bind to the mRNA and perform the translation, with the consequent lack of protein synthesis [3,8] (Scheme 6.1D).

6.2.2 Activation of RNase H

Completely different is the mechanism that causes the activation of RNase H: this enzyme is able to recognize the hybrid AON-RNA (when the AON is DNA based), both at the level of mature mRNA and pre-mRNA, and to degrade the molecule of RNA. The RNases H belong to the category of hydrolases: they in fact hydrolyze the phosphodiester bonds of RNA, releasing oligonucleotides and mononucleotides of the 5′-phosphate-3′-OH type [16]. Once the hydrolytic action is complete, the AON is released and can thus be bound to a new RNA strand; it therefore acts with a catalytic mechanism (Scheme 6.2).
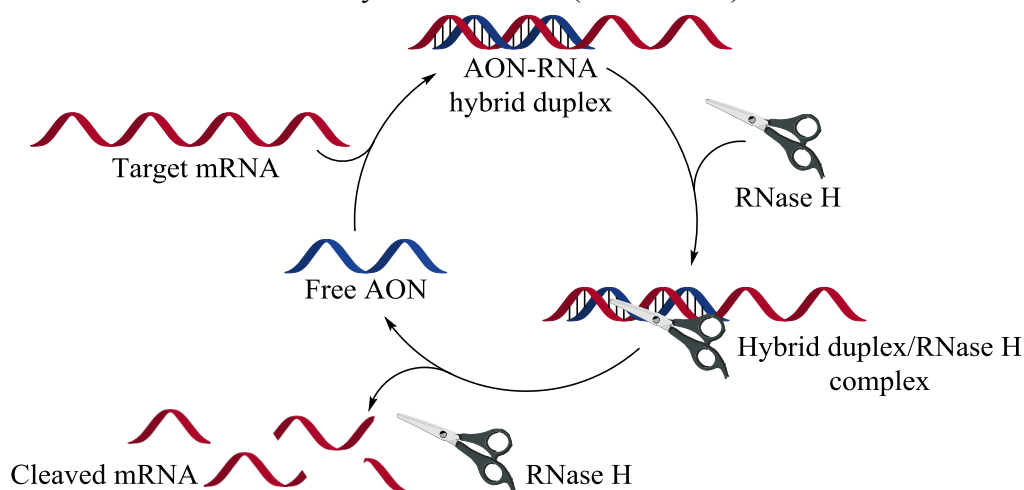
Scheme 6.2 *Mechanism of activation of RNase H.*

6.2.3 Gene silencing induced by siRNA

Another important mechanism is gene silencing induced by siRNA, small fragments of double-stranded RNA of the length of 21-23 nucleotides. They are able to inhibit the expression of specific targets to mRNA, involved in the pathogenesis of important diseases, through their degradation. As shown in Scheme 6.3, an ATP-dependent helicase recognizes these short double helixes and resolves the siRNA molecule in the two individual RNA strands, one called "sense" strand and the other "antisense" strand. The antisense strand is incorporated into the protein complex RISC (RNA-induced silencing complex), acting as a guide for the recognition of the complementary mRNA target sequence, and its subsequent

fragmentation. After the mRNA molecule has been destroyed, the RISC complex continues its enzymatic activity, recognizing and destroying other molecules of the same mRNA target in multiple cycles of catalytic activity [14].
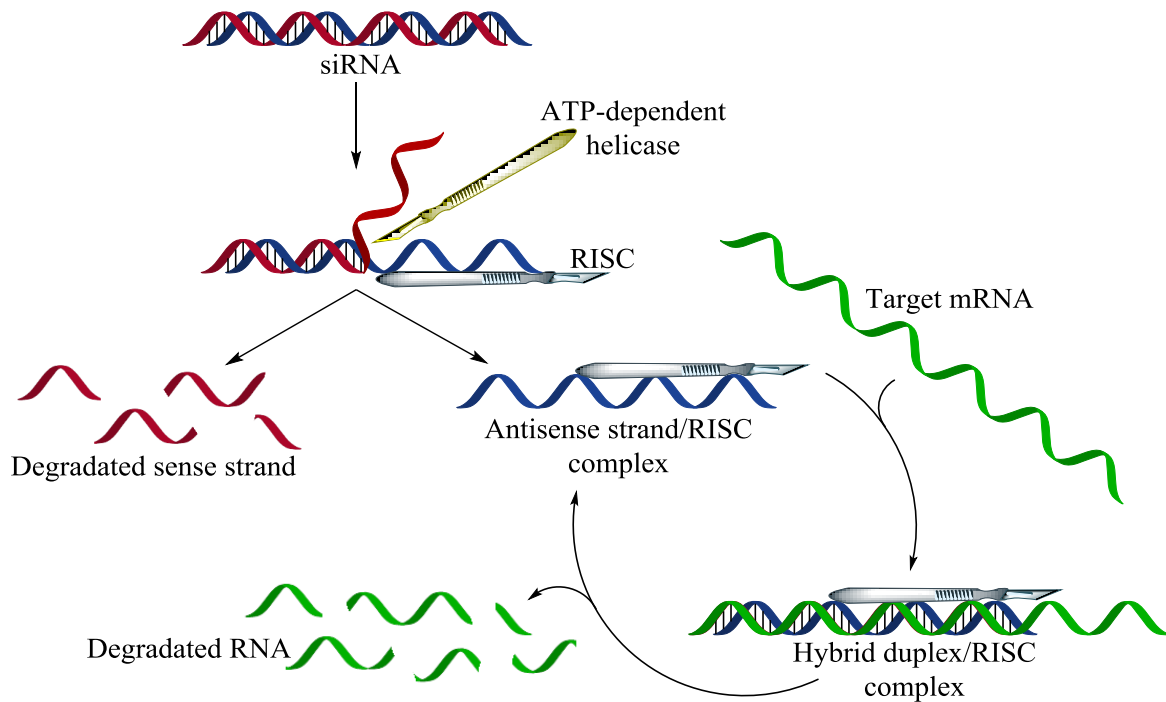
Scheme 6.3 *Mechanism of action of siRNA.*

6.2.4 Alternative splicing

To understand what the alternative splicing technique is, let's see what the splicing is. The pre-mRNA consists of exons, *i.e.* coding regions containing genetic information, interspersed with introns, which are non-coding sequences that have the task of modulation. Before leaving the nucleus, the pre-mRNA undergoes splicing by some enzymes that cut the introns of the molecule and join the exons forming a mature mRNA molecule with a continuous coding sequence; starting from this the translation is carried out which leads to the synthesis of the corresponding protein [6] (Scheme 6.4). Through the splicing process, starting from a single gene, different types of proteins can be created.
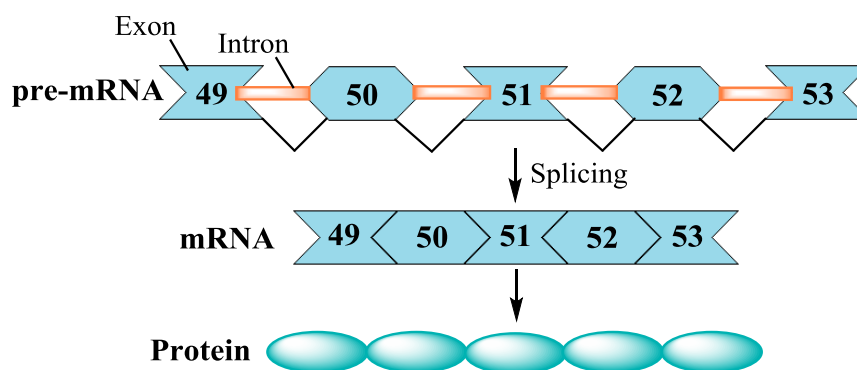
Scheme 6.4 *Process of splicing of pre-mRNA which leads to the synthesis of a protein.*

Some genetic diseases are caused by changes in DNA that lead to the synthesis of dysfunctional proteins or the lack of production of the same. For example, in the case of Duchenne Muscular Dystrophy (DMD), one of the exons has a mutation, and as a result of splicing, a mature mRNA without meaning is generated, which is not translated and from which no protein is synthesized (Scheme 6.5).
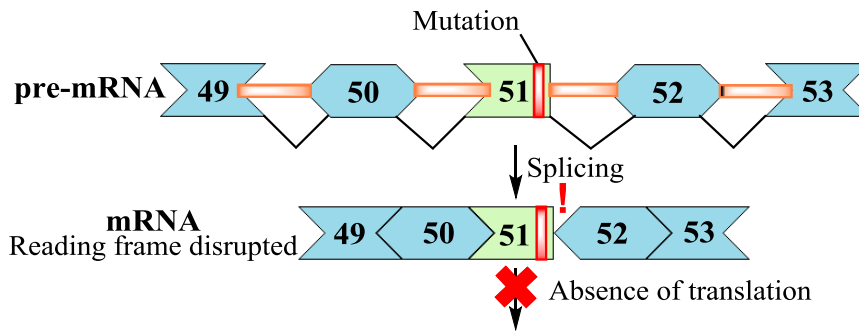
Scheme 6.5 *Protein synthesis absent due to a genetic defect.*

For this type of pathology, a possible therapy that employs AONs exploits a type of alternative splicing called exon skipping [17]. The exon skipping is a gene therapy approach that aims to eliminate molecular damage by acting on the pre-mRNA that codes for that particular protein. In particular, the AON binds to the exon presenting the mutation, thus causing it to be omitted from the mature mRNA. In this way the mutation is removed from the mRNA, that can be translated into the corresponding protein; it will be partly different from the original protein, as it is devoid of an exon, but will nevertheless be functioning [3] (Scheme 6.6).
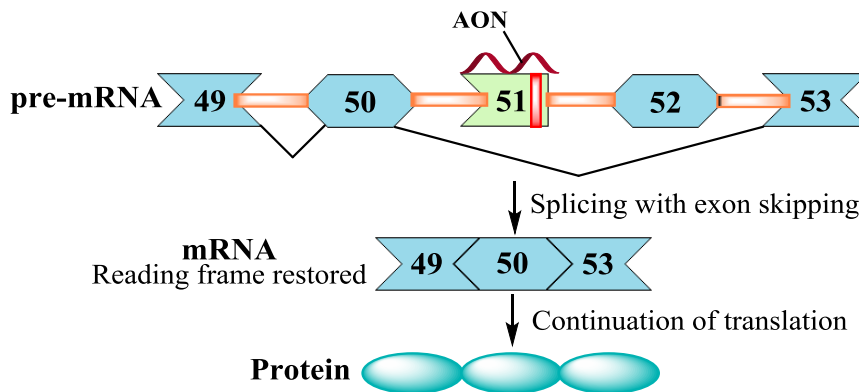
Scheme 6.6 *Modulation of splicing through exon skipping approach.*

6.2.5 AONs in gene therapy

In order for antisense oligonucleotides to be used in gene therapy it is necessary that they exhibit the following properties in the *in vitro* and *in vivo* tests: favorable delivery, non-toxicity and stability. Delivery is the ability of AONs to reach the site where the disease or genes that cause it are located, while non-toxicity is the lack of secondary toxic effects on endogenous cells. Furthermore, these molecules must be stable in biological fluids at least until they have reached their target. The DNA-AONs are in part more stable than the RNA-AONs, but still tend to be very unstable. Therefore, stability is the major limitation of unmodified antisense oligonucleotides; to increase it, over the years more stable analogues have been created that have some structural changes (see paragraph 6.3: Modified oligonucleotides).

In recent years, there has been a growing interest in these innovative molecules able to modulate gene expression and which we hope will be used in the therapeutic and diagnostic field. There are numerous preclinical and clinical studies on these hypothetical drugs for the treatment of illnesses such as cancer, cardiopathies, muscular dystrophies, infectious, inflammatory (viral and genetic) and metabolic diseases [18,19]. Some AONs that have reached the most advanced clinical stages are listed in Table 6.1 [20]. Recently *SPINRAZA*® has been

marked for the treatment of spinal muscular atrophy (SMA) in pediatric and adult patients. The marketing was approved in December 2016 by the FDA (Food and Drug Administration) and in June 2017 by the European Commission [20].

Therapeutic area	Drug	Target	Indication	Phase
Neurological	SPINRAZA® (nusinersen)	SMN2	Spinal Muscular Atrophy	Commercialized
	IONIS-HTT$_{Rx}$ (RG6042)	HTT	Huntington's Disease	Phase 2
	IONIS-SOD1$_{Rx}$ (BIIB067)	SOD1	Amyotrophic Lateral Sclerosis	Phase 2
	IONIS-MAPT$_{Rx}$ (BIIB080)	TAU	Alzheimer's Disease	Phase 2
	IONIS-MAPT$_{Rx}$ (BIIB080)	TAU	Frontotemporal Dementia	Phase 2
	IONIS-C9$_{Rx}$ (BIIB078)	C9orf72	Amyotrophic Lateral Sclerosis	Phase 2
Severe & Rare	TEGSEDI™ (inotersen)	TTR	hATTR	Commercialized
	KYNAMRO®	ApoB-100	Homozygous FH	Commercialized
	WAYLIVRA™ (volanesorsen)	ApoCIII	Familiar Chylomicronemia Syndrome	Registration
			Familiar Partial Lipodystrophy	Phase 3
	AKCEA-ANGPTL3-L$_{Rx}$	ANGPTL3	Rare Hyperlipidermias	Phase 2
	IONIS-GHR-L$_{Rx}$	GHr	Acromegaly	Phase 2
	IONIS-PKK$_{Rx}$	PKK	Hereditary Angioedema	Phase 1
	IONIS-PKK-L$_{Rx}$	PKK	Hereditary Angioedema	Phase 1
	IONIS-TMPRSS6-L$_{Rx}$	TMPRSS6	B-Thalassemia	Phase 1
Cardiometabolic & Renal	AKCEA-ANGPTL3-L$_{Rx}$	ANGPTL3	NAFLD/Metabolic Complications	Phase 2
	AKCEA-APO(a)-L$_{Rx}$	Apo(a)	CVD	Phase 2
	AKCEA-APOCIII-L$_{Rx}$	ApoCIII	CVD	Phase 2
	IONIS-GCGR$_{Rx}$	GCGR	Diabetes	Phase 2
	IONIS-FXI$_{Rx}$ (BAY 2306001)	Factor XI	Clotting Disorders	Phase 2
	IONIS-DGAT2$_{Rx}$	DGAT2	NASH	Phase 2
	IONIS-AGT-L$_{RX}$	AGT	Treatment-Resistant Hypertension	Phase 1
	IONIS-AZ4-2.5-L$_{Rx}$	Undisclosed	Cardiovascular Disease	Phase 1
	IONIS-FXI-L$_{Rx}$	Factor XI	Clotting Disease	Phase 1
Cancer	IONIS-AR-2.5$_{Rx}$	AR	Prostate Cancer	Phase 2
	Danvatirsen (IONIS-STAT3-2.5$_{Rx}$)	STAT3	Cancer	Phase 2
	IONIS-KRAS-2.5$_{Rx}$ (AZD4785)	KRAS	Cancer	Phase 2
Other	IONIS-HBV$_{Rx}$	HBV	Hepatitis B Virus Infection	Phase 2
	IONIS-HBV-L$_{Rx}$	HBV	Hepatitis B Virus Infection	Phase 2
	IONIS-FB-L$_{Rx}$	Complement Factor B	Complement-Mediated Diseases	Phase 2
	IONIS-JBI1-2.5$_{Rx}$	Undisclosed	Gl Autoimmune Disease	Phase 1

Table 6.1 *Examples of AONs undergoing advanced clinical studies* [20].

6.3 Modified oligonucleotides

The polyanionic character of the oligonucleotides, due to the presence of phosphate groups containing negative charges, makes it difficult for them to cross the cytoplasmic and nuclear membranes that is necessary to reach their site of action [21,22,23]. The AONs are also relatively unstable in the blood because they are picked up by the endothelial reticulum system, which has the function of eliminating the foreign substances from the organism, and, following the degradation by the nucleases, they are aggregated with the plasma proteins with the resulting renal excretion (their plasma half-life is a few minutes) [24]. Moreover, in some cases they show a poor affinity for the target RNA.

To increase the stability of AONs and their cellular uptake, two main approaches are reported in the literature: the introduction of structural chemical modifications and / or conjugation with lipids, peptides, polymeric molecules, nanoparticles and liposomes [25,26,27].

The potential advantages of these strategies are:

1) greater intrinsic stability;
2) greater stability towards nucleases;
3) improved delivery into appropriate cellular compartments, such as cytoplasm or nucleus, or into specific tissues that exhibit specific types of specific receptors;
4) faster passage from the blood to the tissues where they are widely distributed.

The chemical modifications on the structure of the nucleotides must allow to maintain their ability to recognize the complementary bases and therefore to be hybridized with the target molecule. The sites of the oligonucleotides on which the various modifications were made are the phosphate group (commonly called backbone), the sugar and the nitrogenous bases. Figure 6.3 shows schematically the various sites subject to modification, while some of the main modifications will be described below.
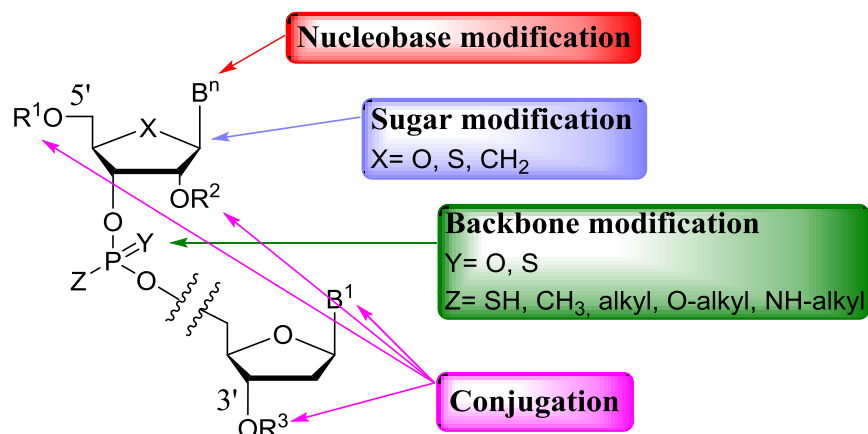
Figure 6.3 *Possible sites of modification on oligonucleotides.*

6.3.1 Backbone modifications

One of the most widespread modifications on the backbone is the replacement of a non-bridge oxygen of the phosphate group with a sulfur atom, leading to the formation of phosphorothioate oligonucleotides [28,29,30] (Figure 6.4). The main advantages of this type of modification are the increased stability towards nucleases [31,13], a greater uptake [32,33,34,35], a good hybridization capacity [36] and the ability to activate RNase H as a mechanism of action [13].

If both oxygens not bound to one of the two nucleoside units are replaced with a sulfur, the phosphorodithioates are obtained (Figure 6.4), which are resistant to nucleases but have little affinity towards the target: for this reason they are not very used [37]. In boranophosphate

oligonucleotides, on the other hand, oxygen is replaced by a borane group (BH_3^-, Figure 6.4); they are much more stable against nucleases than DNA and slightly more stable than phosphorothioates. It was also been seen that they are able to activate the RNase H of *Escherichia Coli*. Another type of modification on the backbone is given by the complete substitution of the phosphate group with other functional groups: for example in the literature [38] the substitution with an amide group associated with the insertion of a methoxyl group in position 2' is reported (Figure 6.4). In this case the affinity towards the complementary RNA is only partly increased, but the amide group has a high resistance against nucleases as it is different from their classical target (the phosphodiester group).
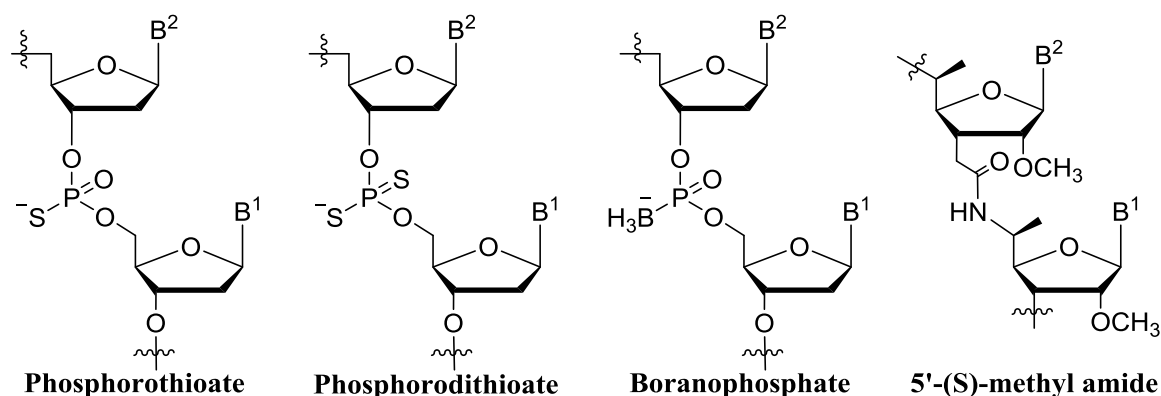

Phosphorothioate   **Phosphorodithioate**   **Boranophosphate**   **5'-(S)-methyl amide**

Figure 6.4 Examples of backbone modifications.

6.3.2 Sugar modifications

In the pentose, the position 2' is the one on which more modifications are made to increase the pharmacological properties of the oligonucleotides. The changes in this position have two advantages: greater affinity and less susceptibility to nucleases, given by the proximity of the substituent in 2' to the phosphate group in position 3'. At the same time, however, this modification greatly reduces or completely inhibits the ability to activate RNase H [39], thereby disrupting this type of anti-sense mechanism. The most common modifications in this position are given by the replacement of the hydroxyl group with a fluorine atom or with 2'-*O*-alkyl groups, such as *O*-methyl and *O*-methoxyethyl (MOE, Figure 6.5) [40]. The high electronegativity of the fluorine atom induces an increase in affinity with respect to RNA, which results to be lower in the presence of the 2'-*O*-alkyl groups. The 2'-*O*-methoxyethyl modification is currently the most advanced of these types of modifications and has undergone several clinical trials. The AONs with this modification show a further increase in resistance to nucleases and a possible decrease in toxicity [41,42,43,44].

$X= F$
OCH_3
OCH_2CH_3
$O(CH_2)_3CH_3$
$OCH_2CH=CH_2$
$O(CH_2)_2OCH_3$
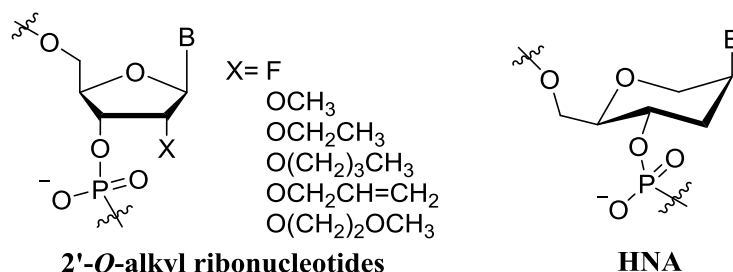
2'-*O*-alkyl ribonucleotides   **HNA**

Figure 6.5 Examples of sugar modifications.

Another type of modification concerns the furanosic scaffold, which can be replaced with a six-terms sugar, thus containing an extra carbon atom. In this category, of particular

importance is the hexitol nucleic acid (HNA, Figure 6.5), which is able to hybridize in a stable way with the RNA and which has a high resistance to nucleases [45].

6.3.3 Nucleobase modifications

Among the various modifications that have been made to the nitrogenous bases, we report the insertion in position 5 in the pyrimidine-based bases of an alkyl group, specifically of a propino (Figure 6.6). This group, being rich in π electrons, increases stacking interactions between the bases, thus increasing the stability of the duplex. The stacking interactions are aromatic interactions between adjacent and parallel nitrogenous bases which have the function of favoring the stability of the double helix. This type of modification also maintains Watson and Crick type recognition unchanged [38].
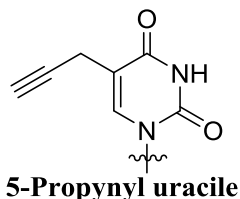
5-Propynyl uracile

Figure 6.6 *Example of nucleobase modification.*

6.3.4 Conjugations with neutral lipids

To increase cellular uptake, a commonly used method is the use of cationic lipids or polymers, which interact electrostatically with the oligonucleotides, leading to the formation of nanoplexes with sizes ranging from 10 to a few hundred nanometers. The advantages provided by this type of formulations are given by greater stability in biological fluids, cellular uptake and biodistribution. The main problem with this strategy is the toxicity given by the cationic character of lipids or polymers that does not make it suitable for use in the biomedical field [46,47].

Based on these considerations, it is interesting to use non-ionic lipids which, due to their natural origin, are non-toxic and biocompatible. Unfortunately the association of AONs through only physical interactions (for example the adsorption) with nanoparticles composed of neutral lipids appears to be poor compared to that with cationic lipids [26,48].

To solve this problem, a possible approach is to associate AONs with neutral lipids through a stable covalent bond. The conjugation with neutral lipids through a covalent bond can be carried out both in the 3' and 5' positions of the oligonucleotide. Both strategies have advantages and disadvantages:

> ➢ the conjugation in 3' has the advantage of the increase in stability against nucleases [49,50]; the main disadvantage is given by a greater technical difficulty in the synthesis;
> ➢ the advantages of conjugation in 5' are a greater ease in the technical execution and the ability to quickly build libraries of oligonucleotides by changing only the conjugate, which is inserted at the end of the whole synthesis; the disadvantages are the non-particular influence on stability against nucleases and the variability of uptake.

It is possible to insert a neutral molecule also within the oligonucleotide chain: to do this it is necessary to have units that compose the AON already conjugated with the neutral molecule or bearing an appropriate linker to which the lipid is bound in the post-synthetic phase.

Another evaluation to do is whether to perform the coupling in solid phase or in solution. To perform the coupling in solution (Scheme 6.7) it is necessary to remove the synthesized oligonucleotide from the support, purify it if necessary, combine it with the neutral lipid and purify it again. This technique turns out to be more laborious than solid-phase conjugation

due to the complexity of the purification method, which in solid phase is performed by simple washing of the reagents left in the reaction environment. Moreover, in the phase of coupling in solution it is necessary to use large excesses of the group to be conjugated and the chemical reactions used must be compatible with the nature of the AON.
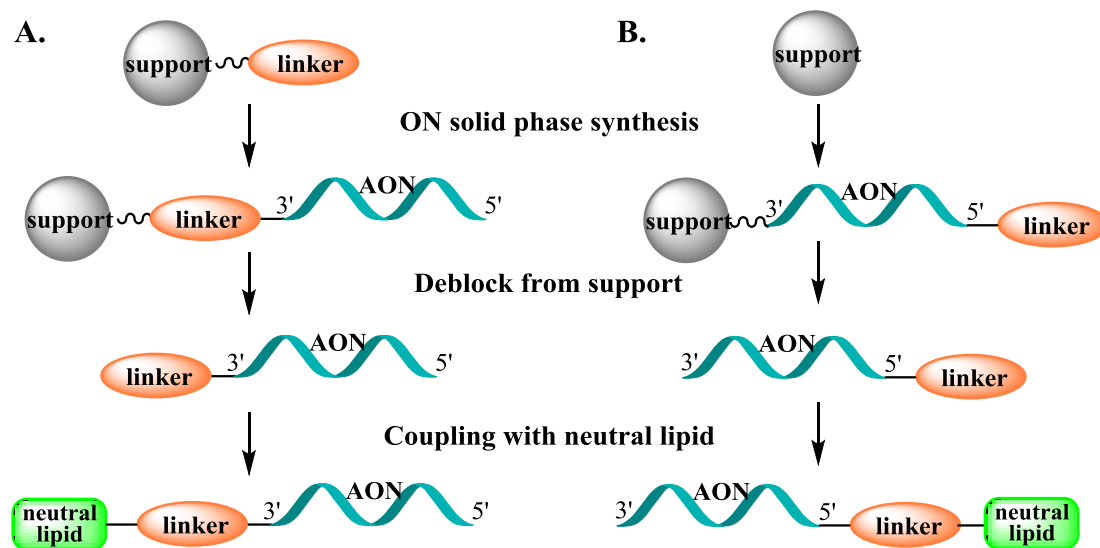
Scheme 6.7 *Syntheses of AON-3′-neutral lipid (A) and AON-5′-neutral lipid (B) conjugates with coupling in solution.*

The use of a solid support can avoid problems related to the poor solubility of one of the reagents. The covalent bonds most commonly used to conjugate a lipid molecule to the oligonucleotide are thioether, disulfide, phosphate and phosphoramidate. To avoid possible problems of steric encumbrance, a simple alkyl chain is often used as a linker between the oligonucleotide and the lipid molecule [51].

To perform the solid phase coupling in position 3' it is necessary to have a support already functionalized with the lipophilic molecule, starting from which the whole oligonucleotide will be synthesized; at the end of the synthesis the conjugated lipophilic molecule-oligonucleotide is released from the support by an ammonolysis reaction (Scheme 6.8A). In the coupling in position 5', instead, the lipophilic molecule, suitably functionalized, is conjugated to the oligonucleotide at the end of the synthetic process in solid phase (Scheme 6.8B).
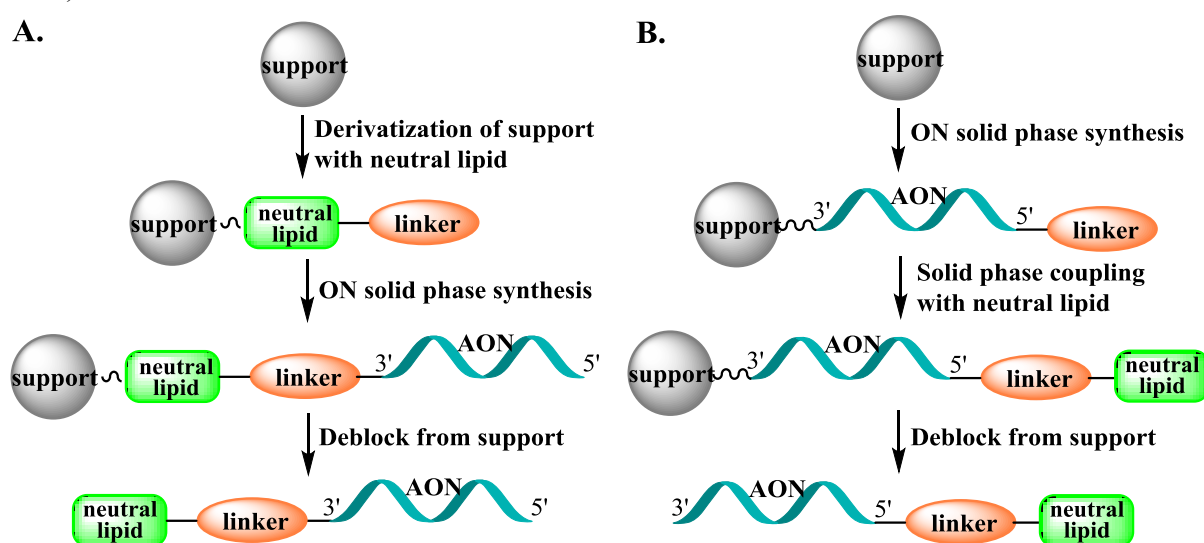
Scheme 6.8 *Syntheses of AON-3′-neutral lipid (A) and AON-5′-neutral lipid (B) conjugates with solid phase coupling.*

Some examples of neutral lipids used in conjugation to improve the pharmacokinetic and pharmacodynamic properties of AONs are cholesterol [52,53], fatty acids [54], vitamin E [55] and bile acids [56,57,58,59].

For example, in literature is reported the conjugation of an antisense oligonucleotide in position 5' with a derivative of a bile acid, the taurocholic acid [59] (Figure 6.7). The attack is made possible thanks to the synthesis of a phosphoroamidite containing taurocholic acid which was inserted as the last amidite of the series. The oligonucleotide in question is able to inhibit the translation of the HCV virus (*Hepatitis C Virus*) *in vitro*. The tests carried out showed that the oligonucleotide taurocholate has a greater cellular uptake, selective for hepatocytes, and therefore a greater inhibitory capacity against the gene expression of HCV compared to the non-conjugated analogue, both *in vitro* and *in vivo* [59].
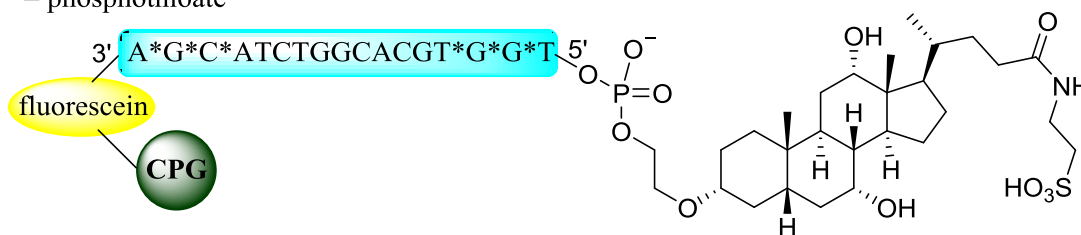

Figure 6.7 *Example of conjugation in position 5' of the antisense oligonucleotide with taurocholic acid.*

6.4 Solid phase synthesis of oligonucleotides

Solid phase synthesis is a technique used for the synthesis of biomacromolecules consisting of monomers. The first monomer of the series is bound to a solid support on which a solution containing the reagents is flowed and, through various steps, the designed compound is synthesized.

This method was developed by Robert Bruce Merrifield for peptide synthesis [60] and earned him the Nobel Prize for chemistry in 1984. Today it also has other applications, such as the synthesis of oligonucleotides and combinatorial chemistry.

Oligonucleotides can also be synthesized in solution, but solid phase synthesis has some advantages: the simplicity of purification operations after each step and the use of synthesizers that make the process automatic. Moreover, compared to the liquid phase synthesis, a smaller quantity of reagents is required, but at the same time it is possible, if necessary, to use an excess of reagents to push the reaction to completion, which will be easily eliminated [61].

6.4.1 Solid support

The solid support must possess a series of requirements:
- ➢ sufficient rigidity and pressure resistance;
- ➢ derivatizability with functional groups;
- ➢ insolubility and inertia towards reagents and solvents used;
- ➢ must not swell and must possess high porosity for facilitate access to reagents thanks to a larger surface area available.

There are several types of solid support, including:
- ➢ polystyrene resins: the first used, called "Merrifield resin", consists of inert polystyrene microspheres containing about 5% of phenyl groups having a chloromethyl group (-CH$_2$Cl) in *para* position [60]. This category also includes *Primer Support^{TM} 5G Amino* and *Custom Primer Support^{TM} C6 Amino*, solid supports which

will be discussed later, which consist of polystyrene bound to amino groups. The typical loading for this type of support is 350-400 μmol/g;

➤ glass supports, called "controlled pore glass", commonly abbreviated with CPG, which are made up of glass balls with controlled porosity. They are rigid, non-swollen, chemically inert and characterized by deep pores within which synthesis occurs. They have a variable loading from 10 to 50 μmol/g [62];

➤ polystyrene-PEG copolymers (*TentaGel*® resins), consisting of polystyrene microspheres on which polyethylene glycol (PEG) molecules are bound [63].

The main difference between polystyrene and CPG supports is that the first ones, in contact with the synthetic solvents, are subject to swelling: this can cause over pressure problems inside the synthesizer. The polystyrene supports, however, are characterized by a significantly higher loading value than the CPG, which allows working on larger synthesis scales, also possess greater resistance and chemical inertia, such as to avoid contamination of the synthesis product with by-products deriving from the decomposition of the support itself; moreover some solid phase syntheses are incompatible or ineffective if carried out on CPG. The polystyrene resins are therefore more suitable for large-scale synthesis of short oligonucleotides (up to 25 nucleotides) and for biomedical applications [3].

An important characteristic of a solid support is its loading, expressed in μmol per gram. It depends on the number of functional groups present on the microspheres. Starting from the loading value, the amount of support needed to obtain a given synthesis scale is calculated, *i.e.* the initial amount of monomer bound to the support, which is expressed in μmol. The equation that connects the two values is as follows:

$$Synthesis\ scale\ (\mu mol) = amount\ of\ support\ (g) \cdot loading\ of\ support\ (^{\mu mol}/_g)$$

Some types of supports available on the market present the first nucleoside of the series linked in position 3' on them, from which the entire oligonucleotide chain will be constructed. The nucleoside unit can be further functionalized with spacer molecules, protector groups and modifications on the nitrogenous bases [3].

Other types of support are functionalized with spacer or linker molecules on which the oligonucleotide will be constructed (Figure 6.8). Their purpose is to functionalize the oligonucleotide in position 3' with a molecule, the linker itself, which has the function of facilitating the subsequent attack of other molecules. At the end of solid phase synthesis, in fact, the link between support and linker is split, while the one between linker and oligonucleotide remains.
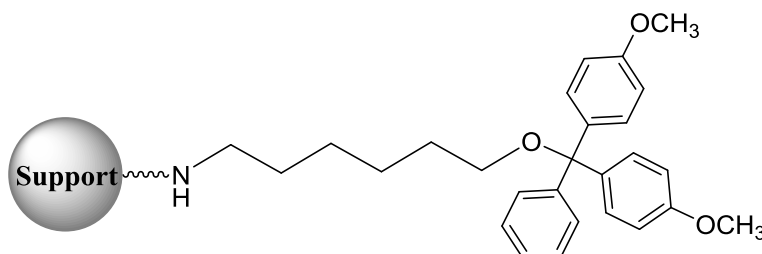

Figure 6.8 *Example of support functionalized with a linker: Custom Primer Support™ C6 Amino.*

The supports available on the market containing a linker have high costs; to reduce these expenses it is possible to attach a personalized linker to a cheaper support already present on the market. In this regard, examples of carboxylic, aminic, carboxamide and mercaptoalkylic

spacers are reported in the literature, which allow to obtain modified oligonucleotides in the 3' position [64]. They consist of a moderately susceptible ester bond to nucleophilic species which, after synthesis termination, is split (Figure 6.9). In this way the newly formed oligonucleotide detaches from the support and, at the same time, various types of functionality are introduced in position 3' [64].
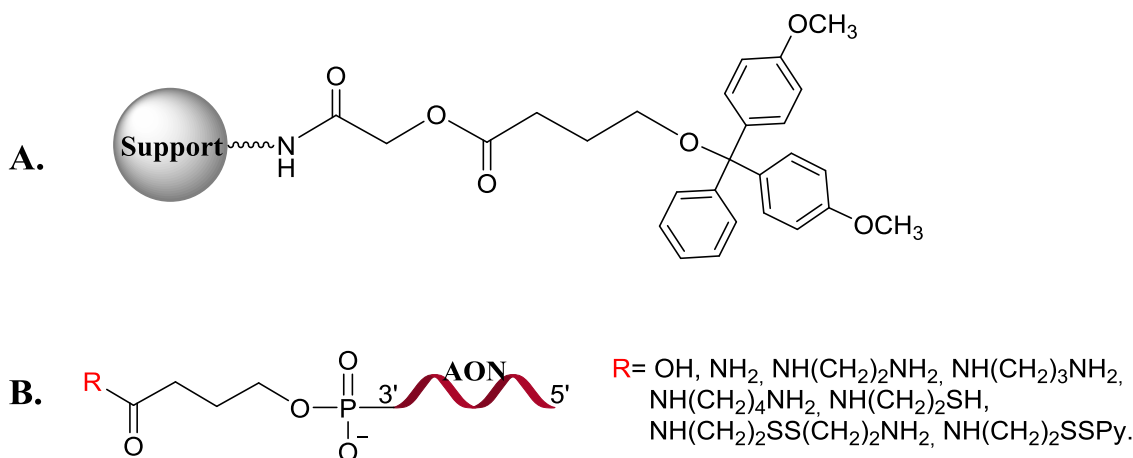

Figure 6.9 *Support modified with a spacer (A) and oligonucleotides functionalized in position 3' with different spacers (B) following the attack of the nucleophile R which causes the separation from the support.*

6.4.2 Conditions for effective solid phase synthesis

As seen previously, the building blocks that make up the oligonucleotides are the nucleotides, which bind to each other through a phosphodiestereal link between two functional groups: the 3′-OH and the 5′-OH. In solid phase synthesis, the growth of the oligonucleotide occurs in the 3'→5' direction, as opposed to what occurs in nature by DNA polymerase, which works in the 5'→3' direction. The direction of the synthesis is reversed with respect to the natural one in order to exploit the greater reactivity of the primary hydroxyl in 5' with respect to the secondary in 3' during the coupling reactions. In order that only the phosphoramidite group in position 3' is able to carry out the nucleophilic attack, it is necessary to protect the 5′-OH. The protector groups commonly used are the triphenylmethyl groups: trityl, 4-monomethoxytrityl (MMT) and 4,4'-dimethoxytrityl (DMT). The greater the number of methoxyl groups on the phenyl, the easier the removal of the protecting group is. Among those just mentioned, the dimethoxytrityl group is more easily removed, under mild conditions, in the presence of trichloroacetic or dichloroacetic acid, in non-aqueous solvents [65]. After the attack of each nucleotide, the end of the oligonucleotide chain must, in fact, be deprotected to allow the attack of the next nucleoside phosphoramidite. Since deprotection occurs after each coupling cycle, the protector groups described above are referred to as "temporary". The "semipermanent" protecting groups are instead removed only at the end of the synthesis of the entire oligonucleotide chain, and are those used to protect the amino groups present on the adenine, cytosine and guanine bases (Figure 6.10). In fact, they are good nucleophiles that can compete in the coupling reaction. It is important to underline how this protection/deprotection system of the different functional groups is fundamental for the solid phase synthesis of the desired oligonucleotide, minimizing the possibility of unwanted couplings [3].
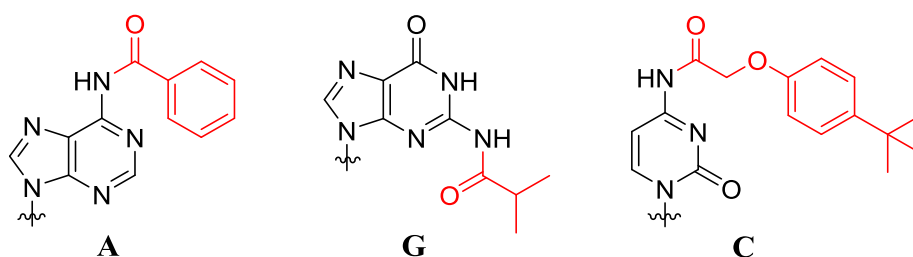
Figure 6.10 *Semipermanent protecting groups used for the protection of adenine, guanine and cytosine: respectively benzoyl, isobutyryl and tert-butylphenoxyacetyl. Uracil and thymine do not need protections instead.*

Another possible nucleophile that can compete with amides is the water present, even in small quantities, in the reagents or in the air. To avoid this drawback, reagents and solvents created specifically for the synthesis of oligonucleotides are used: they contain a concentration of water lower than 30ppm. Furthermore, to ensure the anhydrous conditions, it is necessary to keep the system inside the synthesizer constantly under pressure of anhydrous inert gas (N_2 or Ar) [3].

6.4.3   The automatic synthesizer

The chemical assembly of the oligonucleotide chain takes place within an automatic synthesizer. This one interfaces with a software (*Unikorn* in our case) that allows to program the synthesis and to monitor on-line the most important parameters related to the synthesis process, such as system pressure, reagent conductivity, their UV-Visible absorption and the efficiency of the coupling reaction. The most advanced synthesizers allow the synthesis of oligonucleotides both based on classical phosphodiester DNA/RNA and those modified.

The main components of the oligosynthesizers used by us, the "*ÄKTA oligopilot plus*" of *GE Heathcare* (Figure 6.11), are:

> ➢ Pump P-900, a high performance pump system;
> ➢ UV-900 monitor, with multiple wavelength, which allows simultaneous monitoring of up to three wavelengths between 190 and 700 nm;
> ➢ PH/C-900 monitor, for conductivity and pH monitoring.

Figure 6.11 *The oligosynthesizer "ÄKTA oligopilot plus".*

The system is equipped with a series of valves and tubes that allow the pumps to draw and distribute the amidites (arranged on special slides outside the synthesizer door) and the various reagents necessary for the synthesis. The instrument works in an inert atmosphere (anhydrous argon or nitrogen) to allow the absence of humidity during the reactions; the gas, coming from an external line, is maintained at constant pressure through a pressure regulator

and distributed to the various reactants through two manifolds. The presence of inert gas lines assists the pumps in creating a regular flow and therefore making synthesis more efficient. Outside the synthesizer is placed the housing for the reactor inside which the support is loaded and the synthesis takes place.

There are two types of reactors that differ in terms of material and size, based on the synthesis scale to be achieved: the columns in polymeric material (small cassettes), which have a fixed volume of 0.6 ml and are used for synthesis in range of 1-7 µmolar scale, and steel columns, which are available in various volumes and are used for synthesis on larger scales [3].

6.4.4 Synthesis cycle of oligonucleotides

Since the 1950s, various researchers have dedicated themselves to the search for an effective method for the synthesis of oligonucleotides; this has led to the development of a method, now well established, which is commonly used today: the phosphoramidite method. This method was conceived by Marvin Caruthers in the 80s and consists of the addition of a single nucleotide phosphoramidite for each cycle of synthesis [61,66,67]. A phosphoramidite consists of a nitrogenous base bound to a pentose sugar with the hydroxyl in 5' protected by a dimethoxytrityl group. The 3' hydroxyl is not bound to the phosphate group, as in natural nucleotides, because this group has limitations in the chemical synthesis of oligonucleotides. Instead it is bound to a phosphite group (P^{III}) containing also a β-cyanoethyl group, which is removed only at the end of the synthesis and which has the function of protecting the free hydroxyl, and an amino group (Figure 6.12). The compound, so structured, results to have a certain stability, and can be stored both in the solid state for several months, at a temperature of -20°C, and in solution in anhydrous solvents for short periods [3].
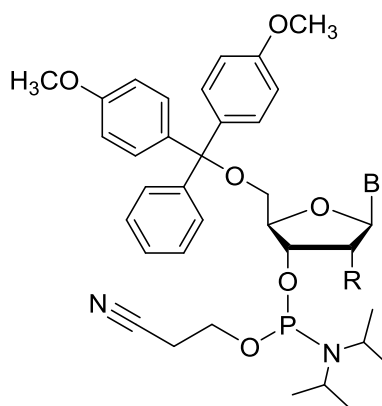
Figure 6.12 Structure of a nucleoside-phosphoramidite.

As stated previously, the oligonucleotide synthesis consists of a series of cycles, each of which corresponds to the addition of a single nucleoside-phosphoramidite. Specifically, we will talk about the synthesis of phosphorothioate oligonucleotides.

Each cycle, shown in Scheme 6.9, consists of 4 steps:

1. detritylation, which consists in the removal of the dimethoxytrityl protecting group from the support or from the first nucleoside bound to the support, in the first synthesis cycle, or from the last attached nucleotide, in the subsequent cycles;
2. coupling, that is the condensation with a nucleoside-phosphoramidite entering the activated form;
3. thio-oxidation of the phosphite ester to thiophosphate ester, a passage made necessary by the instability of the P^{III} in the phosphite;

4. capping, which consists in the protection of 5′-OH functionalities that have not reacted with phosphoramidite in the coupling phase (incomplete chain). Starting from this protected nucleotide the chain will no longer grow, thus generating the failed sequences.
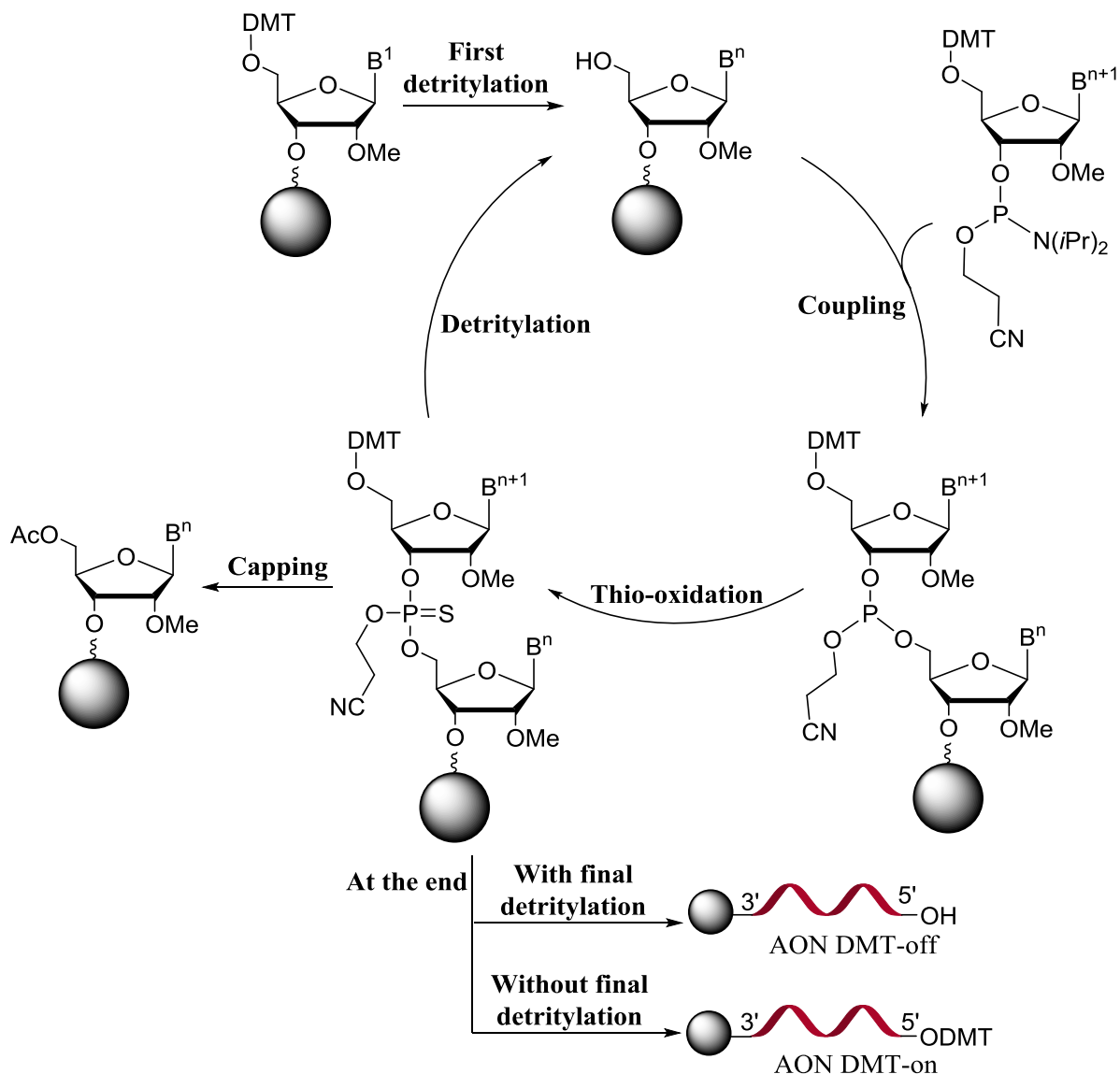
Scheme 6.9 *Synthesis cycle of oligonucleotides.*

The four steps of synthesis cycle of 2′-*O*-methyl phosphorothioate oligonucleotides used by us will be described in more detail below.

6.4.4.1 Detritylation

Removal of the dimethoxytrityl group occurs in the presence of a 3% dichloroacetic acid solution in toluene [68].

As an alternative to toluene, chlorinated solvents can be used, for example methylene chloride, but being more dangerous and having higher disposal costs, it is preferred to avoid its use. The deprotection releases the hydroxyl group in 5', which will be involved in the subsequent coupling step. Precisely for this reason the detritylation is a fundamental step of the synthesis, because in the sites where it does not happen the nucleotide chain will not grow, thus leading to a low yield of the reaction; unfortunately it is known that the detritylation reaction is very often incomplete [69,70]. A possible side reaction caused by the acidic

environment is the depurination, that is the splitting of the β-N-glycosidic bond that causes the detachment of the purine bases (adenine and guanine) [71,72].

The UV-VIS measurement of the efficiency of each individual detritylation reaction is useful to follow the trend of the synthesis in real time through the software. The dimethoxytrityl cation, released in an acidic environment, is in fact a chromophore group which absorbs at λ_{max}=498 nm; it is characterized by a red/orange color due to the total delocalization of the positive charge on the aromatic rings that make it up.

The detritylation reaction is performed before each single attack of a phosphoramidite; only the last nucleotide of the chain may not be detrityled, for reasons related to the purification process that will be explained later, giving rise to an AON called DMT-on.

6.4.4.2 Coupling

The coupling step, performed in CH_3CN, consists of a nucleophilic attack by the 5′-OH of a nucleotide on the P^{III} of the incoming phosphoramidite (Scheme 6.9). This coupling is the basis of the formation of the thiophosphate bond that characterizes this type of oligonucleotides. The outgoing group of this nucleophilic substitution is diisopropylamine. It is a poor leaving group because of its relatively high basicity: it is therefore necessary an activating reagent that neutralizes the diisopropylamine and makes it a better leaving group [73].

The first activating agent used in solid phase synthesis was 1H-tetrazole [74] (Figure 6.13), which has good activating capacities, but has some limitations such as: low solubility, especially at temperatures below 20°C (temperatures at which the synthesis usually takes place), which can lead to crystallization, and the possible removal of the DMT group due to its acidity (pK$_a$=4.89). To avoid these drawbacks, it is preferred to use other activating agents such as 5-benzylthio-1H-tetrazole (BTT) [75,76,77,78] and 5-(bis-3,5-trifluoromethylphenyl)-1H-tetrazole, more commonly called "*Activator 42*" [79,80,81] (Figure 6.13).
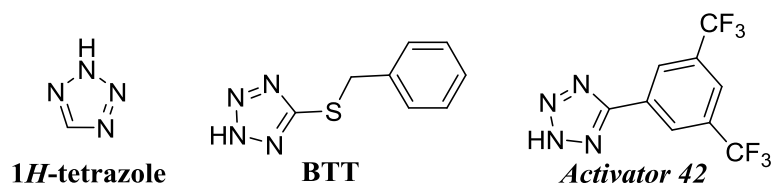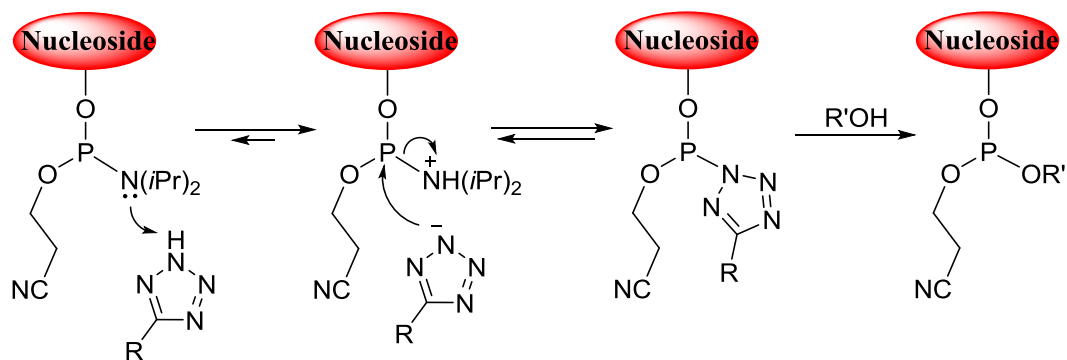

Figure 6.13 *Structures of the most common activating reagents.*

The activating agent acts as a donor of protons first and as nucleophile after [82], as shown in Scheme 6.10.

**Scheme 6.10** *Coupling mechanism activated by tetrazole reactants.*

The time in which each coupling reaction occurs with a certain amidite is called "recycle time", *i.e.* the time in which the solution containing the amidite is made to flow on the support before moving on to the next step. This is an important parameter of the synthesis that must be appropriately chosen to achieve the greatest possible reaction efficiency. In fact, if the recycling time is too short, the phosphoramidite does not have the time necessary to couple with all the available sites and nucleotide sequences shorter than the designed one are formed. If, on the other hand, the recycling time is too long, there is a risk that the phosphoramidite will re-couple on the site where it has just reacted (multiple couplings) leading to the formation of sequences longer than the one planned. The latter phenomenon is mainly found when the sequences to be synthesized are rich in nucleoside-phosphoramidites of the "g" type, which are more susceptible to tetrazolic acid reactivity, and therefore can also be coupled several times within the same coupling cycle.

Another parameter to consider is coupling efficiency, which is used to measure the efficiency of the synthesizer in adding new bases to the oligonucleotide chain being formed. If each available base on the chain has successfully reacted with the new input base, the coupling efficiency will be 100%. The coupling efficiency is not always so high. If, for example, the maximum value that can be obtained is 99%, it means that at each coupling step 1% of the available bases fails to react with the new input base.

6.4.4.3 Thio-oxidation

The thio-oxidation reaction transforms the phosphite triester (P^{III}) obtained with the coupling reaction into thiophosphate (P^V). In this way the reactive and unstable P^{III} becomes the more stable P^V.

The thiophosphate bond differs from the classic phosphate bond contained in DNA and RNA by the presence of a sulfur atom in place of an oxygen atom. This is a typical modification on the backbone that has already been discussed (see paragraph 6.3.1). From the ^{13}P-NMR analysis of the final product it is however often possible to note the presence of some phosphate groups (see paragraph 6.8): this is due to the poor efficiency of the sulfurization reaction or to the presence of water in the reaction environment. The thio-oxidating reagent most commonly used is phenylacetyl disulfide (PADS, Scheme 6.11) [83,84]. To improve the performance of this reagent, it is useful to add to the solution of the reagent in anhydrous acetonitrile organic bases, such as pyridine, which have the function of co-solvents [85].
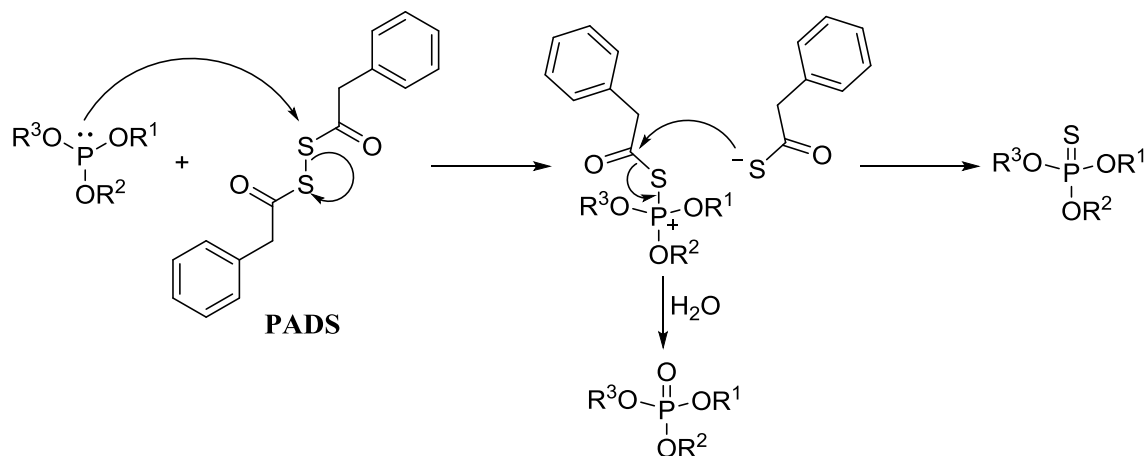
PADS

Scheme 6.11 Thio-oxidation mechanism by PADS. The presence of water in the reaction environment causes the synthesis of phosphate by-products.

Moreover, it has been seen that these solutions require a period of aging [86], equal to at least one day, before their use, but at the same time it is necessary to use them in a short period of time (about fifteen days) because they degrade quickly.

The hypothesized mechanism of PADS (Scheme 6.11) consists initially in the attack of the phosphite triester on the sulfur of the sulfurizing agent, leading to the formation of an intermediate phosphonium ion, from which the desired thio-phosphate triester derives [83]. Thioxidation is normally not stereoselective.

6.4.4.4 Capping

Since the coupling phase does not have 100% efficiency, it is necessary to block the non-reacted 5′-OH functions, to avoid that these failed sequences continue to grow and make purification more difficult. The block consists in protecting the hydroxyl groups at the 5' ends with an acetyl group. Note that only the failed sequences have free 5′-OH groups that can be acetylated, while the correct sequences have the 5′-ends protected as -ODMT. The acetylation reaction is carried out using *CappingA*, a 20% *N*-methylimidazole solution in CH_3CN, and *CappingB*, which corresponds to an 1:1 mixture of capping solutions *B1* (a 40% acetic anhydride solution in CH_3CN) and capping *B2* (a 60% solution of 2,6-lutidine in CH_3CN). Once the capping stage is complete, the cycle starts again with the detritylation of the newly attached nucleotide.

6.4.5 DEA treatment

Once the whole oligonucleotide chain is synthesized, it is treated with a 20% ($V:V$) diethylamine (DEA) solution in anhydrous acetonitrile, to avoid the formation of by-products in the subsequent deprotection phase of the oligonucleotide [87]. During this phase, the cyanoethyl protecting groups located on the thiophosphate bridges are removed, releasing the acrylonitrile (AN) group, which is washed away. The acrylonitrile otherwise could be formed when the AON is unblocked from the solid support in NH_4OH, and add with a reaction of Michael to nitrogen in position 3 present on the thymidine bases, leading to the formation of by-products [88] (Scheme 6.12).
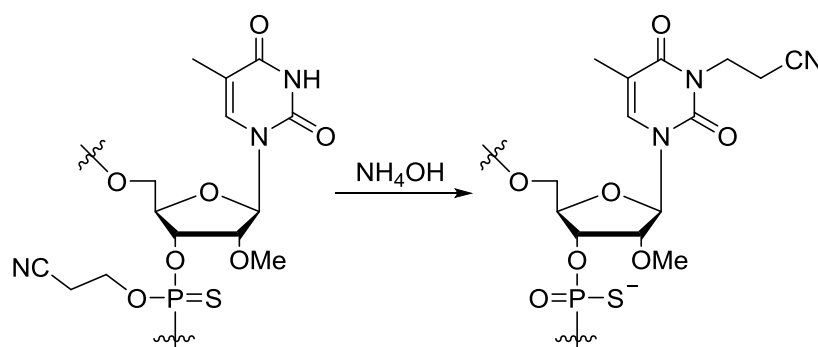
Scheme 6.12 Forming of by-products by alkylation in position 3 of thymidine residues during treatment with NH_4OH.

6.4.6 Deprotection and deblock of the oligonucleotide from the support

The process just described is the last one that runs inside the synthesizer. The obtained compound is then extracted from the reactor and treated with a concentrated solution of ammonium hydroxide at 33% at 60°C for at least 10 hours [89,90]. As an alternative to ammonium hydroxide, gaseous ammonia [91] or methylamine [92] can be used. In this way the oligonucleotide detaches from the support and the exocyclic amino functions present on the nitrogenous bases undergo a deprotection process.

6.5 Purification of the oligonucleotide by RP-HPLC

The analytical profile of the crude oligonucleotide shows the presence of the full lenght oligonucleotide and of various by-products due to a non-total efficiency in solid phase synthesis. The main impurities, as anticipated, may be shorter (short-mers) or longer (long-mers) oligonucleotide chains than the target sequence.

It is therefore necessary to subject the crude compound to a purification process. In general, the techniques most used in the purification of DNA- or RNA-based oligonucleotides are three: reverse phase high performance liquid chromatography (RP-HPLC) [93], ion-exchange high performance liquid chromatography (IEX-HPLC) [94,95,96] or a combination of both procedures.

RP-HPLC allows the separation of various products by exploiting their different polarity. For this reason, when this technique is chosen, during the solid phase synthesis the detritylation of the last amidite of the sequence is often not performed, thus generating a DMT-on oligonucleotide much more apolar than all the shorter chains on which it is not present the dimethoxytrityl group. This allows a good separation of the product from the impurities but makes a subsequent detritylation reaction in solution necessary.

IEX-HPLC, on the other hand, exploits the interaction between the ions present on the oligonucleotide and those of opposite charge present on the resin. In this case it is not necessary to work with DMT-on oligonucleotides and it is possible to avoid the detritylation reaction in solution described above. This technique, in addition to reducing purification steps, allows work in the absence of organic solvents. However, a major disadvantage of ion exchange is the need for high quantities of aqueous salt solutions to release the oligonucleotide from the chromatographic column, with the risk of not being able to completely recover it. Furthermore, a tedious desalination phase of the oligonucleotide is required with the use of very high amounts of aqueous solvent which extends the working time.

In the research work performed by us, RP-HPLC was used for the purification of the oligonucleotides. In fact, this technique proved to be the most efficient for our type of oligonucleotides, guaranteeing a good compromise regarding solvent consumption and purification times. As previously described, in this procedure the newly synthesized oligonucleotide often still has the dimethoxytrityl protection group at the 5' end (DMT-on mode), while the impurities, in particular the short-mers, have the free OH at the 5' end (the acetyl group inserted during the capping is hydrolyzed during the deprotection step by NH_4OH) (Figure 6.14).

The different hydrophilicity of these protecting groups is exploited to purify the oligonucleotide by reverse phase liquid chromatography. If the oligonucleotide is detrityled at the end of the synthesis process (DMT-off mode), its retention time is very close to that of the short-mers, because the full length oligonucleotide and the failed sequences have comparable hydrophilicity, and the separation of compounds by reverse phase chromatography does not give satisfactory results.

Commonly in RP-HPLC a salt buffer of triethylammonium acetate (TEAA) is used as eluent phase [97,98,99], to which a growing percentage of acetonitrile is added according to a defined concentration gradient. If this procedure is used, the oligonucleotide chain is then subjected to further purification steps which will be described below.
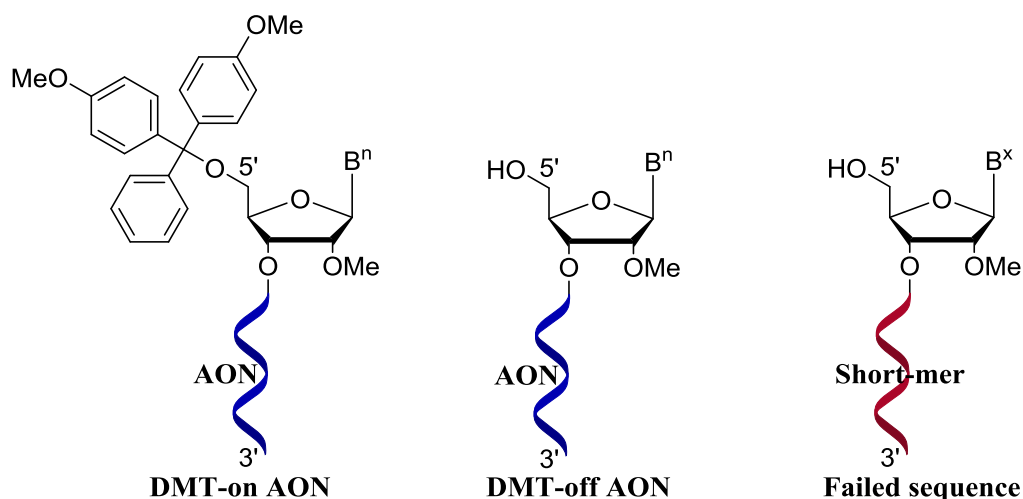
Figure 6.14 *Difference between DMT-on oligonucleotides, DMT-off oligonucleotides and failed sequences.*

6.6 Detritylation in solution

The compound obtained from the chromatography now contains only the full lenght oligonucleotide chain in which the hydroxyl group in 5' is still protected with the DMT group (DMT-on). After being lyophilized, the detritylation of the compound is performed in solution using saline buffer (sodium acetate at pH = 3). This reaction must take place in a short time (about an hour) to avoid the depurination caused by the acidic environment [71]. This reaction time allows a good detritylation efficiency only when working at certain concentrations [100]: it is therefore necessary to derive the AON concentration by spectrophotometric analysis. From the value of the optical density (OD) of the sample, it is possible to calculate the quantity of water in which the AON must be dissolved, considering that the reaction has a good kinetics when a concentration of 1000 OD per ml of H_2O is respected. An excessive dilution, in fact, would cause the detritylation reaction to slow down, with a greater probability of detecting phenomena of depurination [71,72].

An alternative method for detrityling in solution consists in dissolving the AON in H_2O and leaving it at r.t. for 18 hours. The moderate acidity of the water is in some cases sufficient to remove the dimethoxytrityl group, which precipitates as 4,4'-dimethoxytritanol (DMT-OH) and can be easily removed by centrifuging and decanting the mixture.

6.6.1 Purification of the DMT-off AON

The DMT-off compound is subjected to a further RP-HPLC purification to eliminate the dimethoxytrityl alcohol which was formed during the detritylation reaction in solution, the residues of sodium acetate salts and any DMT-on oligonucleotide residues that did not react. Following the course of the chromatography, it is possible to see a sudden increase in the conductivity at the sodium salts outlet. The compound is then again lyophilized.

6.7 Ion exchange

If the oligonucleotide is to be used for *in vitro* tests, it is necessary to exchange the triethylammonium ions (TEA^+), present as counterions of the thiophosphate groups, with Na^+ ions, because the first ones show cellular toxicity. This can be done by filtration on ion-exchange resin [3].

6.8 Qualitative analysis

The purity of the synthesized oligonucleotide can be verified by RP-HPLC analysis: the presence of a single peak in the chromatographic profile, corresponding to the AON, indicates a high degree of purity (Figure 6.15).
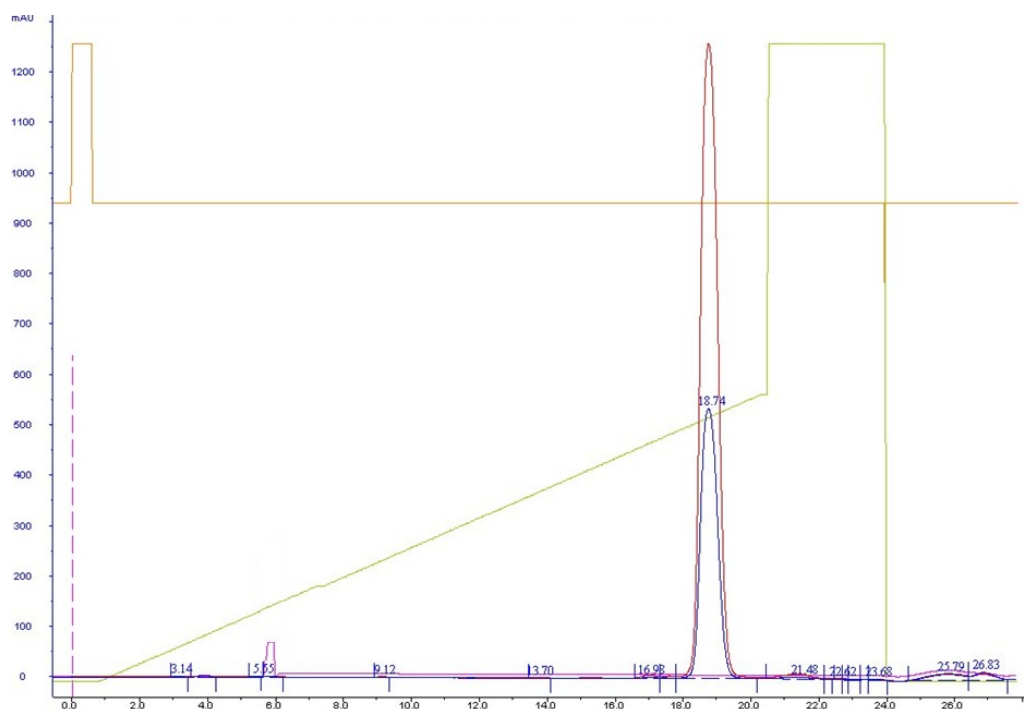

Figure 6.15 RP-HPLC profile of a pure oligonucleotide.

By using ^{31}P Nuclear Magnetic Resonance spectroscopy (^{31}P-NMR), it is also possible to estimate the eventual degree of impurity attributable to a part of the product that is not completely phosphorothioate (presence of partial mono, di, tri, etc. phosphodiesteres (PO)), deriving from an inefficient sulfurization reaction or generated during the ammonolysis step. The chemical shift of phosphorothioate groups is in the 60-50 ppm range, while that of phosphodiesteres is between 0 and -5 ppm (Figure 6.16).
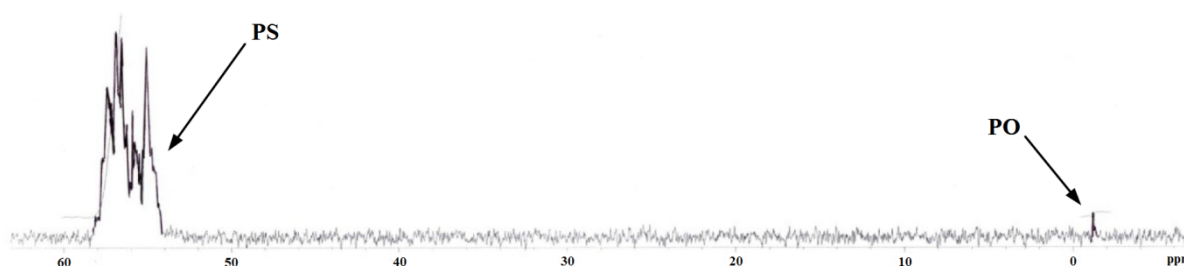

Figure 6.16 ^{31}P-NMR spectrum of a phosphorothioate oligonucleotide with presence of PO impurity.

Another fundamental technique for the characterization of oligonucleotides is mass spectroscopy. In the case of ESI-MS spectra, a pattern of peaks attributable to the various charges assumed by the oligonucleotide is often visible (Figure 6.17).
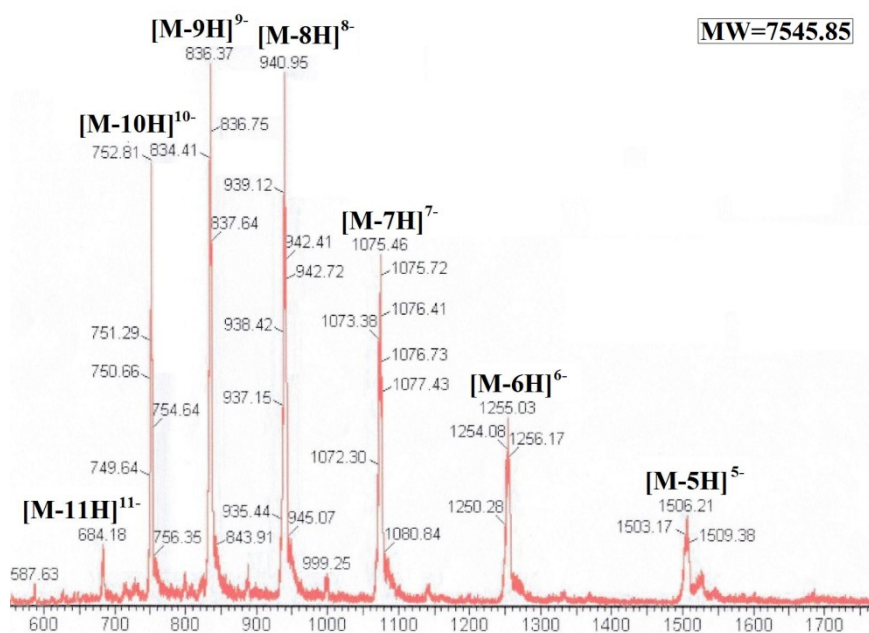
Figure 6.17 *Example of ESI-MS spectrum of an oligonucleotide.*

6.9 Bibliography

1. Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular Biology of the Cell,* 4th ed.; Garland: New York, **2002**.

2. Lewin, B. *Genes VII;* Oxford University: New York, **2000**.

3. Mari, L. *Sintesi e Purificazione di Oligonucleotidi e Nucleosidi 2'-funzionalizzati;* Ferrara, **2008**.

4. Watson, J. D.; Krick, F. H. *Nature* **1953,** *171,* 737.

5. Marmur, J.; et al. *Prog. Nuc. Acid. Res.* **1963,** *1,* 231.

6. Campbell, N. A. *Principi di Biologia;* Zanichelli, 2009.

7. Liebhaber, S. A.; Cash, F. E.; Shakin, S. H. *J. Biol. Chem.* **1984,** *259,* 15597-15602.

8. Walder, R. Y.; Walder, J. A. *Proc. Natl. Acad.* **1998,** *85,* 5011-5015.

9. Crooke, S. T. *Biochem. Biophys. Acta* **1999,** *31,* 1489.

10. Zamore, P. D.; Tuschl, T.; Sharp, P. A.; Bartel, D. P. *Cell.* **2000,** *101,* 25.

11. Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci.* **1978,** *75,* 280.

12. Wagner, R. W. *Nature* **1994,** *372,* 333.

13. Cook, P. D. In *Antisense Research and Application;* CRC, **1993**.

14. Foye, W. O.; Lemke, T. L.; Williams, D. A.; Roche, V. F.; Zito, S. W. *Principi di Chimica Farmaceutica,* 6th ed.; Piccin, **2014**.

15. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010,** *50,* 259-293.

16. Crouch, R.; Dirksen, M. Cold Spring Harbor Monograph Series, Plainview. *Nucleases* **1982,** *14,* 211-254.

17. Spitali, P.; Rimessi, P.; Fabris, M.; Perrone, D.; Falzarano, S.; Bovolenta, M.; Trabanelli, C.; Mari, L.; Bassi, E.; Tuffery, S.; Gualandi, F.; Maraldi, N. M.; Sabatelli-Giraud, P.; Medici, A.; Merlini, L.; Ferlini, A. *Hum. Mutat.* **2009,** *30,* 1527-1534.

18. Culliton, B. J. Clinical investigation: An endangered science. *Nature Medicine* **1995,** *1* (4), 281.

19. Burkhard, J.; Uwe, Z. *Lancet Oncology* **2002,** *3,* 672-683.

20. Ionis Pharmaceuticals. Pipeline - Ionis Pharmaceuticals. http://www.ionispharma.com/pipeline/ (accessed November 8, **2018**).

21. Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Biologia molecolare della cellula,* 6th ed.; Zanichelli, **2016**.

22. Miller, P. S.; McParland, K. B.; Jayaram, K. *Biochem.* **1981,** *20,* 1874.

23. Nolen, H. W.; Catz, P.; Friend, D. R. *Intl. J. Pharm.* **1994,** *107,* 169.

24. Shaller, H.; Weimann, G.; Lerch, B.; Khorana, H. G. *J. Am. Chem. Soc.* **1963,** *85,* 3821.

25. Gray, G. D.; Basu, S.; Wickstrom, E. *Biochem. Pharmacol.* **1997,** *15* (53), 1465-1476.

26. Rimessi, P.; Sabatelli, P.; Fabris, M.; Braghetta, P.; Bassi, E.; Spitali, P.; Vattemi, G.; Tomelleri, G.; Mari, L.; Perrone, D.; Medici, A.; Neri, M.; Bovolenta, M.; Martoni, E.; Maraldi, N. M.; Gualandi, F.; Merlini, L.; Ballestri, M.; Tondelli, L.; Sparnacci, K.; Bonaldo, P.; Caputo, A.; Laus, M.; Ferlini, A. *Mol. Ther.* **2009,** *17,* 820-827.

27. Shokrzadeh, N.; Winkler, A. M.; Mehrdad, D.; Winkler, J. *Bioorg. Med. Chem. Lett.* **2014,** *24,* 5758-5761.

28. Uhlman, E.; Peyman, A. *Chem. Rev.* **1990,** *90,* 543.

29. Cohen, J. S. In *Antisense Research and Application;* CRC, **1993**; p 205.

30. Stein, C. A.; Cheng, Y. C. *Science* **1993,** *261,* 1004.

31. Campbell, J. M.; Bacon, T. A.; Wickstrom, E. *J. Biochem. Biophys. Methods* **1990,** *20,* 259.

32. Loke, L. S.; Stein, C. A.; Zang, X. H.; Mori, K.; Nakanisci, M.; Subasinghe, C.; Cohen, J. S.; Nekers, L. M. *Proc. Natl. Acad. Sci. USA* **1989,** *86,* 3474.

33. Crooke, R. M. *Anti-Cancer Drug Design* **1991,** *6,* 609.

34. Jarosewski, J. W.; Cohen, J. S. *Adv. Drug. Deliv. Rev.* **1991,** *6,* 235.

35. Chin, D. J.; Green, G. A.; Zon, G.; Szoka, F. G.; Straubunger, R. M. *New Biologist* **1990,** *2,* 1091.

36. Stec, W. J.; Wilk, A. *Angew. Chem. Int.* **1994,** *33,* 709.

37. Ghosh, M. K.; Ghosh, K.; Dahl, O.; Cohen, J. S. *Nucleic Acids Res.* **1993,** *21,* 5761.

38. Crooke, S. T. *Antisense Drug Technology: Principles, Strategies and Applications;* CRC, **2007**.

39. Hinoue, H.; Hayase, Y.; Iwai, S.; Ohtsuka, E. *FEBS Lett.* **1987,** *215,* 327-330.

40. Prakash, T. P.; Balkrishen, B. *Curr. Top. Med. Chem.* **2007,** *7,* 641-649.

41. Sproat, B. S.; Lamond, A. I. In *Antisense Research and Application;* CRC, **1993**.

42. Morvan, F.; Porumb, H.; Degols, G.; Lefebvre, I.; Pompon, A.; Sproat, B. S.; Rayner, B.; Malvy, C.; Lebleu, B.; Imbach, J. L. *J. Med. Chem.* **1993,** *36,* 280.

43. Monia, B. P.; Lesnik, E. A.; Gonzales, C.; Lima, W. F.; McGee, D.; Guinosso, C. J.; Kawasaki, A. M.; Cook, P. D.; Freier, S. M. *J. Biol. Chem.* **1993,** *36,* 14514-14522.

44. Sproat, B. S.; Lamond, A. B.; Beijer, B.; Neuner, P.; Ryder, U. *Nucleic Acids Res.* **1989,** *17,* 3373.

45. De Bouvere, B.; Kerreinans, L.; Hendrix, C.; De Winter, H.; Schepers, G.; Van

Aerschot, A.; Herdewijn, P. Hexitol Nucleic Acids (HNA): Synthesys and Properties. *Nucleosides and Nucleotides* **1997,** *16,* 7-9.

46. Rothenberg, M.; et al. *J. Natl. Cancer Inst.* **1989,** *20,* 1523.

47. Guinosso, C. J.; et al. *Nucleosides Nucleotides* **1991,** *10,* 259.

48. Chakraborty, R.; Dasgupta, D.; Adhya, S.; Badu, M. K. *Biochem. J.* **1999,** *340,* 393-396.

49. Stetsenko, D. A.; Gait, M. J. A convenient solid-phase method for synthesis of 3′-conjugates of oligonucleotides. *Bioconjugate Chem.* **2001,** *12,* 576-586.

50. Habus, I.; Zhao, Q.; Agrawal, S. Synthesis, hybridization properties, nuclease stability, and cellular uptake of the oligonucleotide-amino-beta-cyclodextrins and adamantane conjugates. *Bioconjugate Chem.* **1995,** *6,* 327-331.

51. Raouane, M.; Desmaele, D.; Urbinati, G.; Massaad-Massade, L.; Couvreur, P. Lipid Conjugated Oligonucleotides: A Useful Strategy for Delivery. *Bioconjugate Chem.* **2012,** *23,* 1091-1104.

52. Farooqui, F.; et al. *Bioconjugate Chem.* **1991,** *2,* 422.

53. Gmeiner, W. H.; et al. *Bioorg. Med. Chem. Lett.* **1991,** *1,* 487.

54. Will, D. W.; Brown, T. *Tetrahedron Lett.* **1991,** *33,* 2729.

55. Yamada, T.; Peng, C. G.; Matsuda, S.; Addepalli, H.; Jayaprakash, K. N.; Alam, M. R.; Mills, K.; Maier, M. A.; Charisse, K.; Sekine, M.; Manoharan, M.; Rajeev, K. G. Versatile site-specific conjugation of small molecules to siRNA using click chemistry. *J. Org. Chem.* **2011,** *76* (5), 1198-1211.

56. Lehman, T. J.; Engels, J. W. *Bioorg. and Med. Chem.* **2001,** *9,* 1827-1835.

57. Starke, D.; et al. *Bioorg. and Med. Chem.* **2001,** *11,* 945-949.

58. Kim, S. J.; et al. *ChemBioChem* **2004,** *5,* 1517-1522.

59. Gonzales-Carmona, M. A.; Quasdorff, M.; Vogt, A.; Tamke, A. Inhibition of hepatitis C virus RNA translation by antisense bile acid conjugated phosphorothioate modified oligodeoxynucleotides (ODN). *Antiviral Research* **2013,** *97,* 49-59.

60. Merrifield, R. B. *J. Am. Chem. Soc.* **1963,** *85,* 2149.

61. Caruthers, M. H.; Matteucci, M. D. *J. Am. Chem. Soc.* **1981,** *103,* 3185-3191.

62. Adams, S. R.; Kauka, K. S.; Wikes, E. J.; Holder, S. B.; Galuppi, G. R. *J. Am. Chem. Soc.* **1983,** *105,* 661.

63. Goodman, M.; Felix, A.; Moroder, L.; Toniolo, C. *Synthesis of Peptides and Peptidomimetics;* Houben Weyl: New York, **2002**; Vol. 22.

64. Hovinen, J.; Guzaev, A.; Azhayev, A.; Lonberg, H. Novel solid supports for the preparation of 3′-derivatized oligonucleotides: Introduction of 3′-alkylphosphate tether groups bearing amino, carboxy, carboxamido and mercapto functionalities. *Tetrahedron* **1994,** *50,* 7203-7218.

65. Wuts, P. G. M.; Greene, T. W. *Greene's Protective Groups in Organic Synthesis,* 4th ed.; Wiley: Hoboken, New Jersey, **2007**.

66. Caruthers, M. H.; McBride, L. J. *Tetrahedron Lett.* **1983,** *24,* 245-248.

67. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1989,** *22,* 1859.

68. Krotz, A.; Carty, R.; Moore, M.; Scozzari, A.; Cole, D.; Ravikumar, V. Synthesis of

oligonucleotides using environmentally friendly and safe deprotection. *Green Chemistry* **1999,** 277-281.

69. Tensamani, J.; Kubert, M.; Agrawal, S. *Nucleic Acids Res.* **1995,** *23,* 1841-1844.

70. Fearon, K. L.; Stults, J. T.; Bergot, B. J.; Christensen, L. M.; Raible, A. M. *Nucleic Acids Res.* **1995,** *23,* 2754-2761.

71. Suzuki, T.; Ohsumi, S.; Makino, K. *Nucleic Acids Res.* **1994,** *22,* 4997-5003.

72. Hakala, H.; Oivanene, M.; Saloniemi, E. *J. Phys. Org. Chem.* **1992,** *5,* 824-828.

73. Beaucage, S. L.; Iyer, R. P. *Tetrahedron* **1992,** *48,* 2223-2231.

74. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981,** *22,* 1859-1862.

75. Welz, R.; Muller, S. *Tetrahedron Lett.* **2002,** *43,* 795-797.

76. Wu, X.; Pitsch, S. *Nucleic Acids Res.* **1998,** *26,* 4315-4323.

77. Pitsch, S.; et al. *Helv. Chim. Acta* **1999,** *82,* 1753-1761.

78. Reddy, K. S. *U.S. Patent 7* **2008,** *052 B2,* 339.

79. Threlfall, R. N.; et al. *Org. Biomol. Chem.* **2012,** *10,* 746-754.

80. Karmakar, S.; et al. *J. Org. Chem.* **2011,** *76,* 7119-7131.

81. Wolter, A.; Leuck, M. *U.S. Patent 0* **2006,** *431,* 247.

82. Berner, S.; Muehlegger, L.; Seliger, H. *Nucleic Acids Res.* **1983,** *17,* 853-864.

83. *Bioorg. Med. Chem. Lett.* **2006,** *16,* 2513-2517.

84. Krotz, A. H.; Gorman, D.; Mataruse, P.; Foster, C.; et al. *Org. Proc. Res. and Develop.* **2004,** *8,* 852-858.

85. Cheruvallath, Z. S.; Carty, R. L.; Moore, M. N.; Capaldi, D. C. *Org. Process Res. Dev.* **2000,** *4,* 199-204.

86. Robert, J.; Anouti, M. *Chem. Soc. Perkin Trans.* **1997,** *2,* 473-478.

87. Capaldi, D. C.; et al. *Org. Proc. Res. and Devel.* **2003,** *7,* 832-838.

88. Bergman, E. D.; Ginsberg, D.; Pappo, R. *Org. React.* **1959,** *10,* 179.

89. Holmberg, L. Large scale oligonucleotides synthesis. *IBC Meeting* **1997**.

90. Holmberg, L. *IBC Conference* **1998**.

91. Boal, J. H.; et al. *Nucleic Acids Res.* **1996,** *24,* 3115-3117.

92. Reddy, M. P.; et al. *Tetrahedron Lett.* **1995,** *36,* 8929-8932.

93. Warren, W. J.; Vella, G. *Protocols for Oligonucleotides Conjugates* **1994,** 233-264.

94. Kato, Y.; et al. *J. Chromatogr.* **1989,** *478,* 264.

95. Ausserer, W. A.; Biros, M. L. *BioTechniques* **1995,** *19,* 136.

96. Jilge, G.; et al. *J. Chromatogr.* **1989,** *476,* 37.

97. Warren, W. J.; Vella, G. *Mol. Biotechnol.* **1995,** *4,* 179.

98. Huber, G. C. *J. Chromatogr. A.* **1998,** *806,* 3.

99. Uesugi, T.; Sano, K.; Uesawa, Y.; et al. *J. Chromatogr. B.* **1997,** *703,* 63.

100. Krotz, A. H.; et al. *Org. Proc. Res. & Devel.* **2003,** *7,* 47-52.

7 Design, synthesis and biological evaluation of *2′-O-methyl-phosphorothioate* antisense oligonucleotides conjugated with lipophilic compounds to improve exon-skipping approaches

7.1 Introduction

The ability of antisense oligonucleotides (AONs) to regulate the expression of a gene through the hybridization mechanism between the AON and a target sense RNA sequence is well known [1]. Among the various mechanisms that use antisense technology (see paragraph 6.2), particularly important for therapeutic purposes is the strategy of exon skipping. In this case an AON binds to a nuclear pre-mRNA molecule (the target), to modify the splicing of the exons and usually generate a shorter mRNA sequence, but able to be translated into an alternative and functional isoform of the protein (see paragraph 6.2.4).

The exon skipping by AON is one of the most interesting methods for the treatment of certain genetic diseases, such as Duchenne muscular dystrophy (DMD) [2]. This disease constitutes 50% of all dystrophic forms and affects a male in 3500 in the first years of life. The females are generally not affected by the disease, but they are healthy carriers, as the disease is inherited as a recessive trait linked to the X chromosome. Only in rare cases the females show mild forms (muscle weakness), for random inactivation of the second chromosome X: in these situations we speak of manifesting carriers [3]. In some cases DMD is caused by the lack of synthesis of the dystrophin protein, determined by a mutation present on the 51 exon of the dystrophin gene located on the X chromosome. In the muscle, dystrophin is located on the cytoplasmic side of the sarcolemma and connects the cytoskeleton to the extracellular matrix in the muscle fibers (Figure 7.1). The absence of this protein leads the muscle cell to separate from the extracellular matrix resulting in muscle necrosis due to lack of nutrients [4].
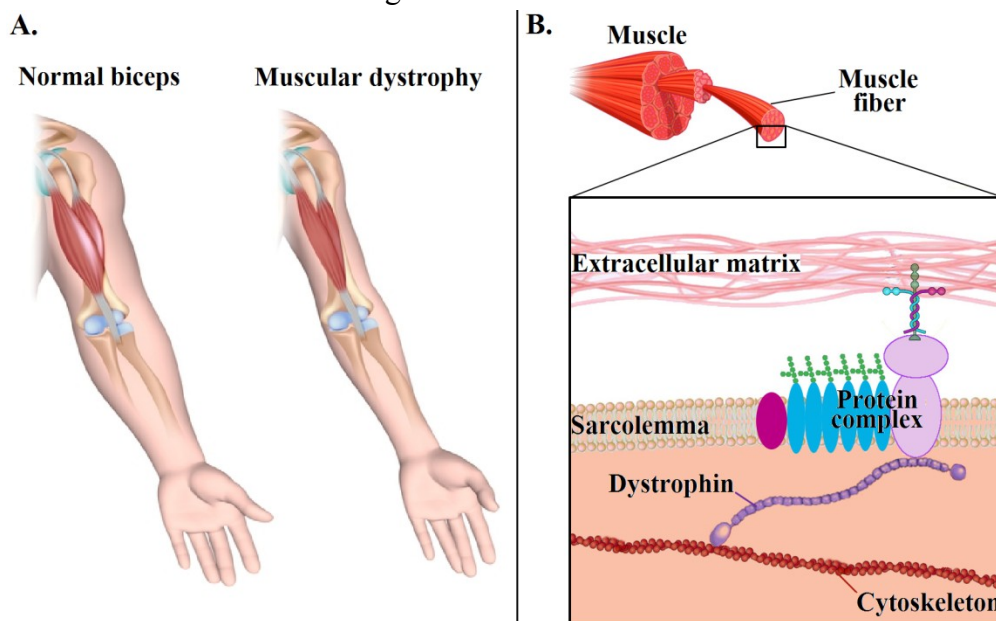

Figure 7.1 *Comparison between a normal biceps and one with muscular dystrophy [5] (A) and role of dystrophin that connects the cytoskeleton to the extracellular matrix in muscle fibers [6] (B).*

To date, two types of highly modified AONs have been used for studies in animal models and are currently in clinical phase III for the treatment of DMD: a modified *2′-O*-methyl-

oligoribonucleotide (2'-*O*Me-AON) with a fully phosphorothioate skeleton (PS) and a morpholino phosphorodiamidite oligomer (PMO, Figure 7.2). Despite the progress made by chemistry in the field of oligonucleotides, recent studies have highlighted the lack of considerable therapeutic benefits for both AONs, probably due to an insufficient *in vivo* exon skipping process [7]. A strategic point to take into account to increase antisense efficacy of oligonucleotides is their low ability to permeate cellular and nuclear membranes and reach the target pre-mRNA in the nucleus. Although nucleobases are hydrophobic, due to the presence of phosphorothioate bonds and sugars, the 2'-*O*MePS-oligonucleotide results to be a hydrophilic molecule. This feature, in addition to the anionic nature of phosphorothioate bridges, reduces cell permeability. In any case, the simple elimination of the anionic charges alone does not increase cellular uptake, as evidenced by oligomers with neutral structures such as methylphosphonates and peptide-nucleic acids [8]. To increase the systemic distribution of the AONs, new carrier systems have been developed, such as glycol-polyethylene chains, peptides capable of penetrating into cells (CPP) and polymeric nanoparticles [9,10].
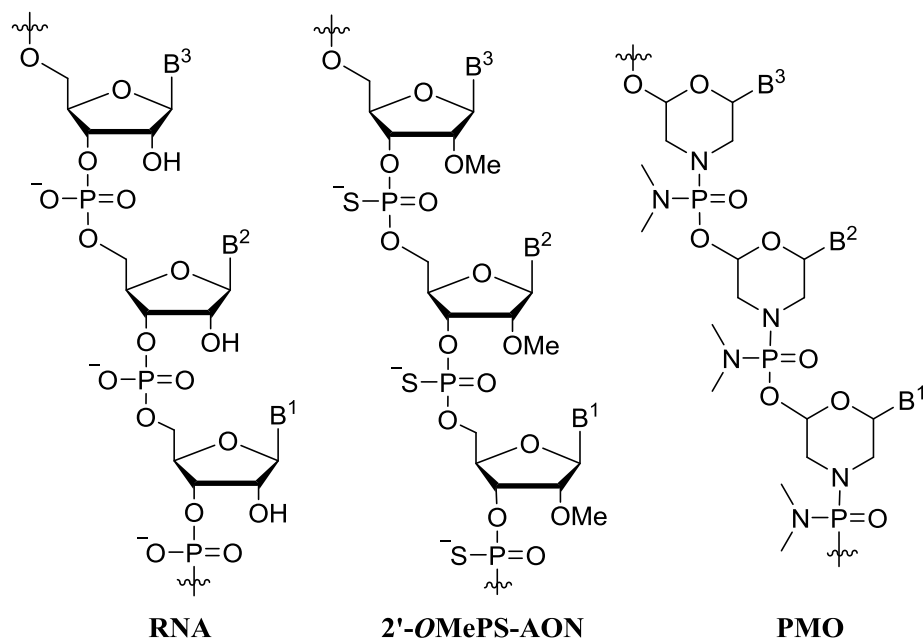

RNA **2'-*O*MePS-AON** **PMO**

Figure 7.2 Comparison of RNA, 2'-OMe-AON and PMO structures. On the left, the RNA structure with the characteristic ribose rings bound by phosphodiester bridges. In the center, the 2'-OMe-AON with phosphorothioate bonds that join the ribose rings. On the right, the PMO incorporates morpholine rings in place of ribose rings, conjugated by phosphorodiamidite bonds.

In this work, new synthetic approaches have been developed for the conjugation of a lipophilic molecule to the 20-mer 2'-*O*MePS-AON that is able to associate with exon 51 (2'-*O*MePS-Antisense51, that we will shorten as **AON51**) of the human dystrophin gene responsible for DMD, in order to improve the exon skipping approach. The sequence of **AON51** is 5'-UCAAGGAAGAUGGCAUUUCU-3'. The lipophilic molecule chosen for coupling with the AON is ursodeoxycholic acid (UDCA), for the following reasons:

➢ an increase in the lipophilicity of the conjugated AON is expected;
➢ cellular and nuclear membranes present receptors for some bile acids [11];
➢ UDCA is widely used in the medical field [12,13] (trade names: *Actigall*®, *Ursofalk*®, *Ursosan*®, *De-ursil*®, *Ursilon*®, *Ursacol*®, *Ursobil*®...);
➢ UDCA has already been used in recent years as a drug carrier [14,15,16];
➢ it has chemopreventive properties [17,18,19];

➢ it is a natural compound, with good availability and low cost.

In the literature it is reported that conjugation at the 3′-end of the AON confers greater stability of the compound towards nucleases [20] and in some cases leads to greater uptake compared to conjugations in other positions. On the other hand, the conjugation at the 5′-end is technically easier to make and allows to quickly build libraries of oligonucleotides where the conjugated moiety changes from time to time. To explore all possibilities, in this work UDCA has been conjugated to the **AON51** both at the 3′- and 5′-ends, as well as at both simultaneously.

We have also developed a library of conjugates between the same **AON51** and other lipophilic molecules, illustrated in Figure 7.3.
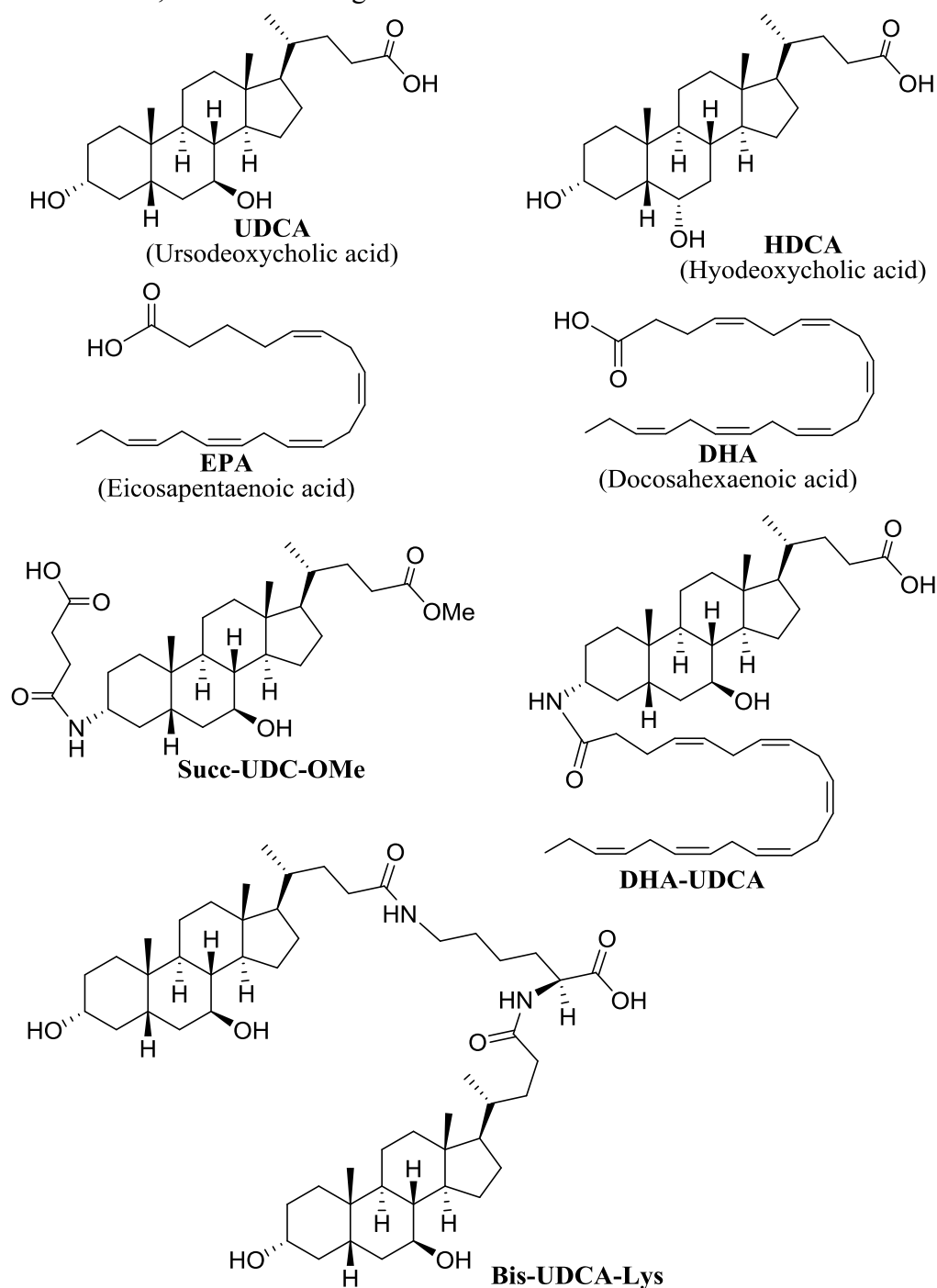

Figure 7.3 Liphophilic compounds conjugated with AON51.

The selected lipophilic molecules are non-toxic bile acids (UDCA and HDCA), ω-3 essential fatty acids (EPA and DHA) and derivatives of these same molecules. All new conjugated AONs were appropriately purified and characterized, then tested *in vitro* using immortalized myogenic cells of patients with DMD, in order to evaluate their efficacy in restoring dystrophin expression. The conjugate that showed the best results in *in vitro* tests was subsequently subjected to *in vivo* tests on animal model.

7.2 Results and discussion

Various methods are known for functionalizing the AONs in solution [21]. Some of them involve exchanges with disulfides, Diels-Alder cycloadditions and maleimide-thiol conjugations, but require laborious multistep procedures and usually have low yields.

For these reasons we have decided to develop solid phase coupling methods to realize our AON-conjugates. The coupling between the **AON51** and the lipophilic molecules was made by forming an amide bond, produced by the reaction between the carboxylic acid functional group present on the lipophilic molecule and the amino group inserted at the end of the AON by a suitable linker.

To insert the linker at the 5′-end of **AON51**, a commercially available amidite (*ssH-Linker*^{TM}, Figure 7.4) was used, which is linked to the last nucleotide of the sequence during the automated solid phase synthesis with a normal coupling cycle. The subsequent coupling with the lipophilic molecule was performed by maintaining the AON in the solid phase.
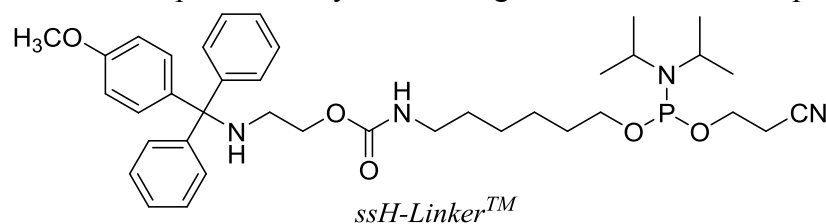

Figure 7.4 The ssH-Linker™ amidite. The amino functionality is protected by a monomethoxytrityl group.

More complex is the discourse regarding the coupling at the 3′-end, for which two synthetic procedures have been followed:

➢ oligonucleotide solid phase synthesis starting from a commercially available solid support already containing the amino linker (*Custom Primer Support*^{TM} *C6 Amino*) and subsequent coupling with the lipophilic molecule in solution;

➢ derivatization of an adequate solid support (*Primer Support*^{TM} *5G Amino*) with the lipophilic molecule and subsequent oligonucleotide solid phase synthesis on it.

The strategies followed to synthetize the AON51-UDCA conjugates will be described below.

7.2.1 Synthesis of AON51-5′-UDCA

To realize the **AON51-5′-UDCA**, the solid phase synthesis of the oligonucleotide in DMT-off mode (**1**) was performed using the automatic synthesizer "*ÄKTA oligopilot plus*" (see paragraph 6.4). The used solid support was the *Primer Support*^{TM} *5G*. The amino linker (*ssH-Linker*^{TM}) was inserted after the last nucleotide of the sequence, then the final detritylation was performed to obtain the AON DMT-off **2** (Scheme 7.1), with the free amino group for the subsequent coupling reaction with UDCA.
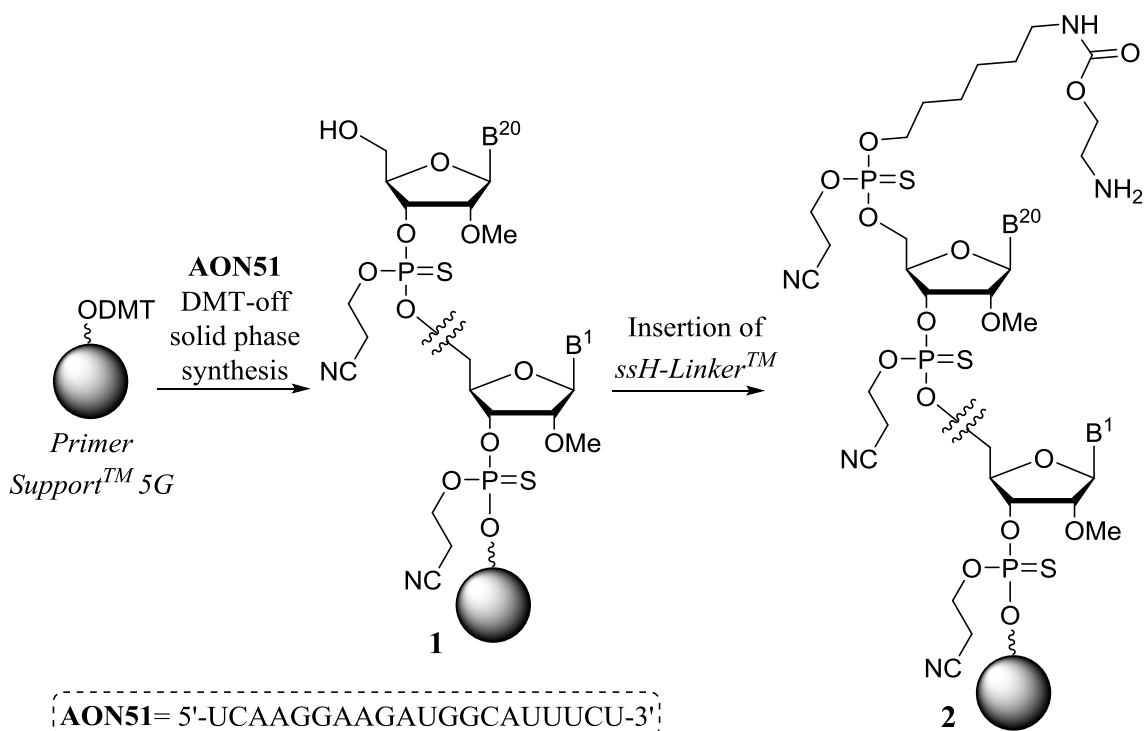
Scheme 7.1 *Solid phase synthesis of AON51 with amino linker at the 5′-end.*

AON51 = 5'-UCAAGGAAGAUGGCAUUUCU-3'

To allow the formation of the amide bond by reaction with the amine, the carboxylic acid group of UDCA was previously activated as ester. Three activation conditions have been tested, shown in the Table 7.1.

Entry	Reagents	Condensing agent or catalyst	Conditions	Yield
1 [22]	UDCA (1eq) + NHS (1.2eq)	DCC (1.2eq)	THF, 25°C, 18h	97%
2 [23]	UDCA (1eq) + NHS (1.2eq)	EDC (1.3eq)	THF, 25°C, 18h	-
3 [24]	UDCA (1eq) + PFP-TFA (2eq)	DIPEA (4eq)	THF, 25°C, 1h	20%

Table 7.1 *Conditions tested for the activation of bile acid. NHS= N-hydroxysuccinimide, DCC= dicyclohexylcarbodiimide, EDC= 1-ethyl-3-(3-dimethylaminopropyl)carbodiimides, PFP-TFA= pentafluorophenyl trifluoroacetate.*

The conditions that lead to a higher yield are the first [22], illustrated in Scheme 7.2. DCC has the function of condensing agent (see paragraph 1.1.1), which promotes the attack of NHS to carboxylic acid by incorporating a molecule of H_2O and thus becoming dicyclohexylurea (DCU), which is insoluble in THF and is easily eliminated by filtration on celite.
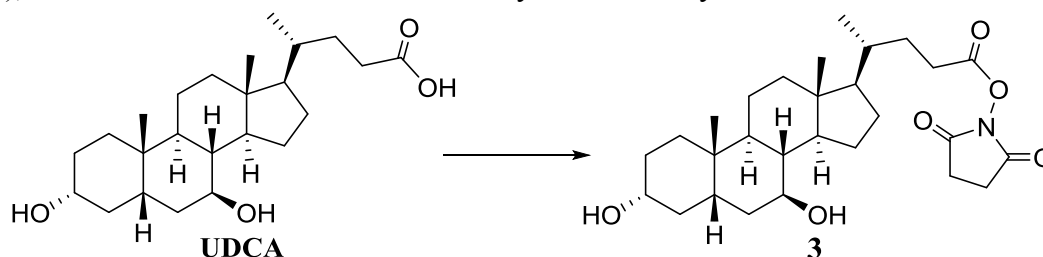
Scheme 7.2 *Activation of UDCA as N-succinimidyl ester **3**. Reactives and conditions: NHS (1.2eq), DCC (1.2eq), anhydrous THF, 25°C, 18h (97% yield).*

The coupling between AON51 functionalized with amino linker **2** and the active ester of UDCA **3** was then performed in solid phase (Scheme 7.3). The AON was held anchored to the solid support within the reactor, while compound **3** (10eq) was dissolved in DMSO with
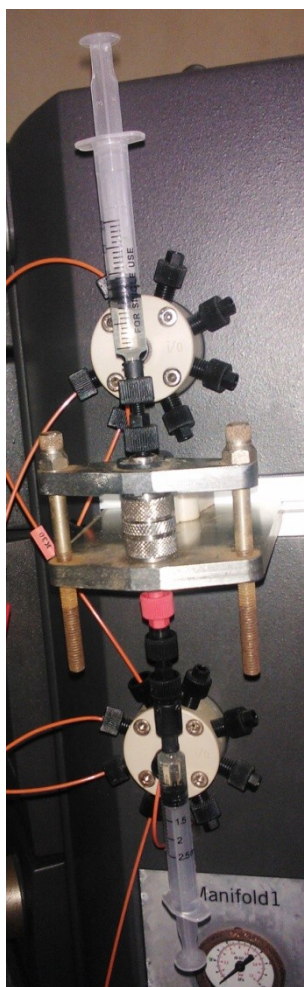
DIPEA (40eq) and loaded into a syringe. This syringe and another empty one were connected to the inlet and outlet of the reactor by means of suitable fittings, as shown in Figure 7.5. The solution was then made to flow through the reactor for about 4 hours, then the solid contained in the column was filtered under vacuum on sintered glass, washed with CH_3CN and transferred to a conical flask with screw cap.

Figure 7.5 System used for solid phase coupling at the 5'-end of the AON.
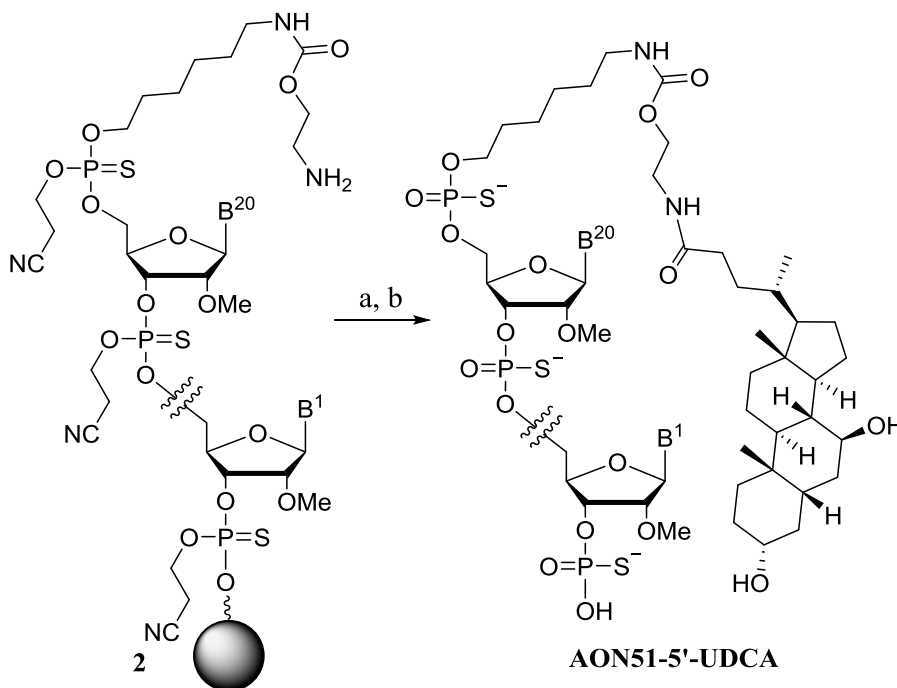
Scheme 7.3 Synthesis of **AON51-5'-UDCA**. Reagents and conditions: (a) **3** (10eq), DIPEA (40eq), DMSO/CH₃CN 1:1 (v/v), 25°C, 4h; (b) NH₄OH 33%, 50°C, 24h.

To release the oligonucleotide from the solid support and remove the protecting groups, NH₄OH 33% was then added into the flask, which was capped and left in an oven at 50°C for 24 hours (Scheme 7.3).
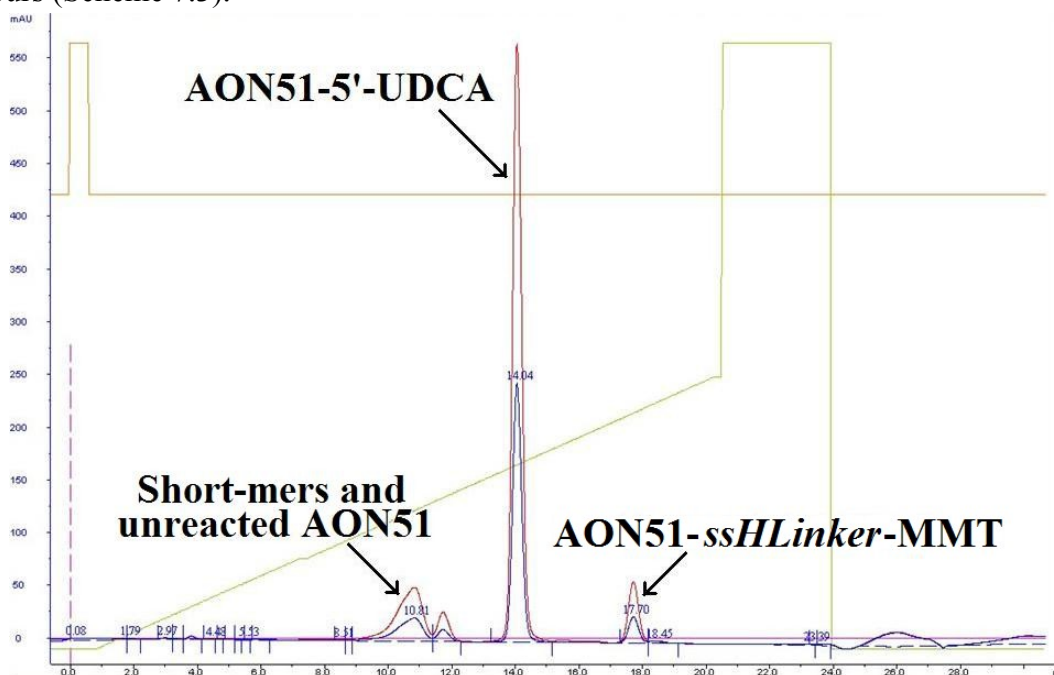

Figure 7.6 Chromatographic profile of purification of **AON51-5'-UDCA** by RP-HPLC.

The contents of the flask were then filtered under vacuum onto sintered glass and washed with EtOH, then diluted 4 times with the buffer (TEAA pH=8 + 5% CH$_3$CN) used as eluent in the subsequent purification step by RP-HPLC (Figure 7.6, see paragraph 6.5). The **AON51-5′-UDCA** conjugate was obtained with a total yield of 72% after chromatographic purification.
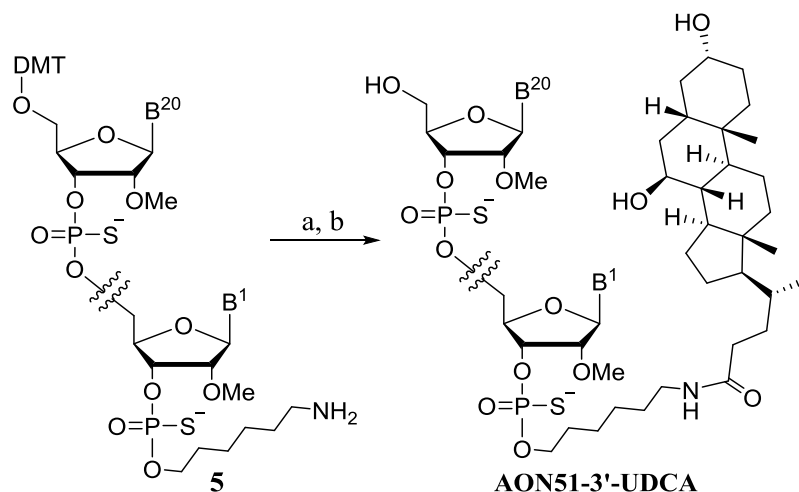
7.2.2 Synthesis of AON51-3′-UDCA

7.2.2.1 Coupling at the 3′-end of AON in solution

The first way for the synthesis of **AON51-3′-UDCA** provided for the solid phase synthesis of the oligonucleotide on *Custom Primer Support^{TM} C6 Amino*, a commercially available support containing the amino linker, in DMT-on mode (**4**, Scheme 7.4).
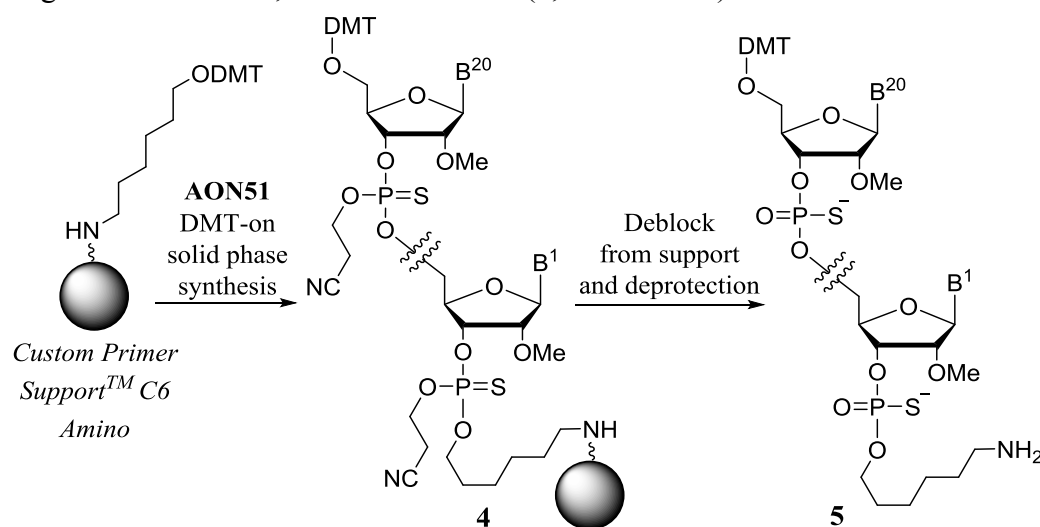
Scheme 7.4 *Solid phase synthesis of AON51 with amino linker at the 3′-end.*

Subsequently, the oligonucleotide was treated with NH$_4$OH 33% at 50°C for 24h to release it from the solid support and remove the protecting groups (**5**, Scheme 7.4), purified by RP-HPLC and subjected to the detritylation reaction in solution, by the sodium acetate buffer at pH = 3 (see paragraph 6.6). After a second purification by RP-HPLC, coupling in solution was performed with the active ester of UDCA **1**, using H$_2$O/1,3-dioxolane 1:1 as solvent in presence of DIPEA (Scheme 7.5). Conjugate **AON51-3′-UDCA** was then obtained with a total yield of 45% after further chromatographic purification.

Scheme 7.5 *Detritylation and synthesis of **AON51-3′-UDCA**. Reagents and conditions: (a) CH$_3$COONa buffer pH=3, 25°C, 90min (95% yield); (b) **3** (10eq), DIPEA (40eq), H$_2$O/1,3-dioxolane 1:1 (v/v), 25°C, 16h (81% yield).*
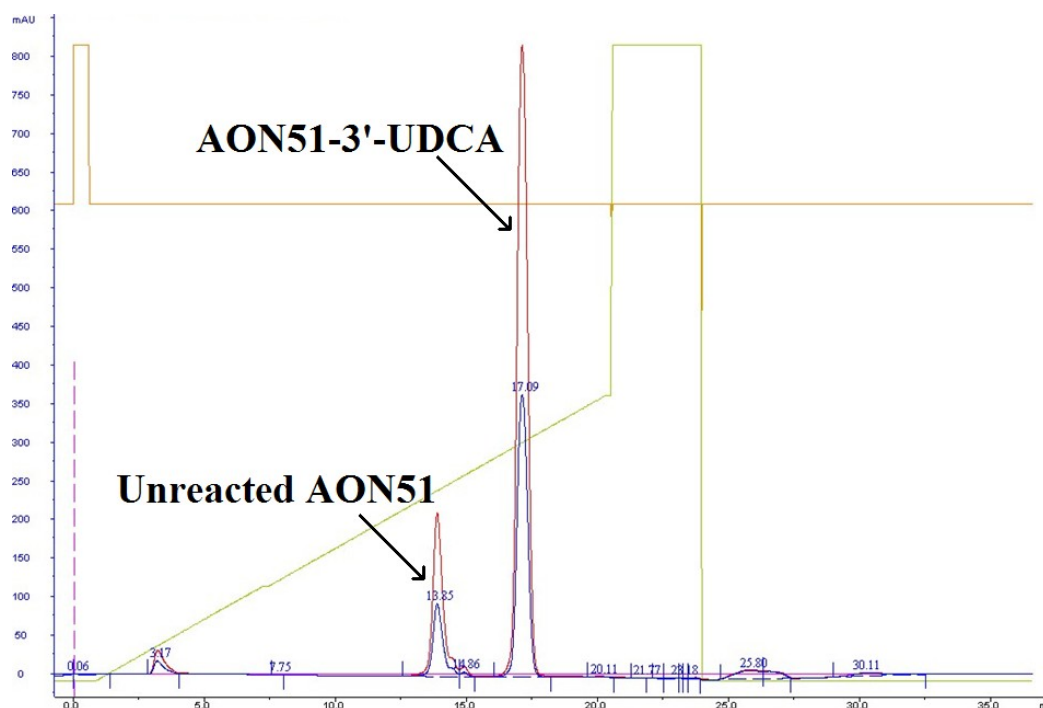
Figure 7.7 *Chromatographic profile of final purification of **AON51-3′-UDCA** by RP-HPLC.*

7.2.2.2 AON solid phase synthesis on support derivatized with UDCA

The second strategy developed for the synthesis of **AON51-3′-UDCA** can be divided into three steps:
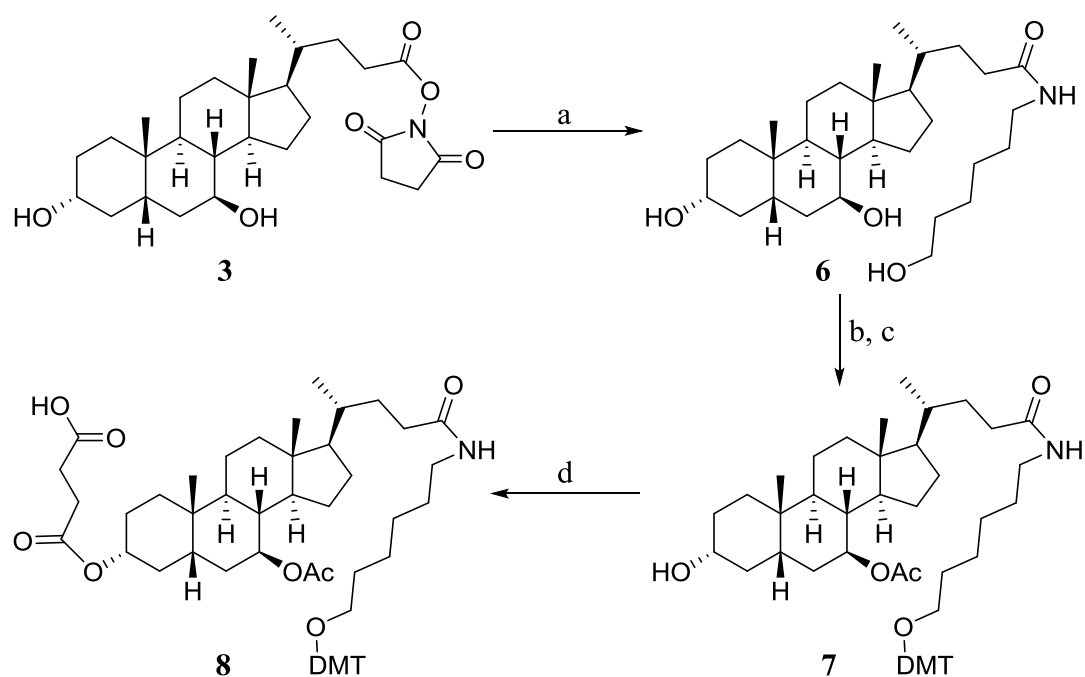
1. functionalization of UDCA;
2. derivatization of solid support with functionalized UDCA;
3. solid phase synthesis of AON51 on derivatized solid support.

Functionalization of UDCA

The synthesis of the ursodeoxycholic acid derivative occurred through a series of reactions indicated in Scheme 7.6.

In particular, the modifications made to the structure of ursodeoxycholic acid are two: esterification with succinic anhydride in position 3 and amidation of the carboxylic acid in position 24 with 6-amino-1-hexanol, subsequently protected with the dimethoxytrityl group.

The 6-amino-1-hexanol has the role of spacer for promoting the entry of dimethoxytrityl, which is characterized by a considerable steric encumbrance. In turn, it has the function of protecting group of the hydroxyl, which otherwise would also be subject to attack by the succinic group. We have chosen the dimethoxytrityl group because it is the same protector group present in the phosphoramidites used in the synthesis of the oligonucleotide. The hydroxyl group present in position 7 of the bile acid has been protected with an acetyl group because the dimethoxytrityl is too cumbersome to access this position. The only hydroxyl group which remains free is that in position 3 where, as previously said, succinic acid is bound. Succinic acid has the function of linker between the support and the bile acid molecule.

Scheme 7.6 *Functionalization of UDCA. Reagents and conditions: (a) 6-amino-1-hexanol, DMAP, DMF, 25°C, 18h; (b) 4,4′-dimethoxytrityl chloride, Py, 25°C, 18h; then acetic anhydride, DMAP, 25°C, 18h; (c) KOH, EtOH, 25°C, 3h (23% yield in four steps); (d) succinic anhydride, DMAP, Py, 70°C, 18h (80% yield).*

Starting from the *N*-succinimidyl ester **3** (see Scheme 7.2), the 6-amino-1-hexanol was first linked to the amide bond, using DMAP as catalyst, to form the compound **6** (Scheme 7.6).

Through a one-pot/two-step reaction, the primary hydroxyl was protected with dimethoxytrityl, using 4,4′-dimethoxytrityl chloride in pyridine, then the hydroxyls at positions 3 and 7 were simultaneously protected with acetyl groups adding acetic anhydride and DMAP to the reaction mixture. The acetyl in 3 was then selectively removed with basic hydrolysis performed by a KOH in EtOH solution, following the conditions reported in the literature [25,26], to give compound **7** (Scheme 7.6); in this step it is important to stop the reaction after 3 hours to avoid that deacetylation also occurs in 7.

The subsequent reaction between 3-OH and succinic anhydride leads to the formation of an ester bond and to the insertion of the carboxylic acid functional group (see paragraph 1.2.3), necessary to bind the steroid scaffold of compound **8** (Scheme 7.6) to the solid support.

Derivatization of solid support with functionalized UDCA

The compound **8**, obtained with the reactions described above, was used for the functionalization of the *Primer Support*[TM] *5G Amino* in liquid phase. In particular, the carboxylic function of the compound **8** reacts with the amine groups present on the *Primer Support*[TM] *5G Amino* leading to the formation of an amide. This forming bond is rather stable and remains so throughout the solid phase synthesis of the oligonucleotide. When the synthesis is terminated, the ester linkage at position 3 of the bile acid is cleaved to release the oligonucleotide now bound to the ursodeoxycholic acid derivative.

The *Primer Support*[TM] *5G Amino* is sold in inactivated form, *i.e.* the amine groups present on its surface are conjugated with a protection group, whose structure is not made known by the manufacturing company (*GE Healthcare*). To activate it we have followed a procedure provided by the same company that consists in making the following solvents flow through the support within the synthesizer:

➢ CH₃CN for 10 minutes with a flow of 2 ml / min;

> diisopropylethylamine (DIPEA) at 2% in CH$_3$CN for 30 minutes with a flow of 1 ml / min; following the reaction with the appropriate software it is possible to see that the conductivity curve has increased;

> CH$_3$CN again with a flow of 2 ml / min until the conductivity curve has stabilized at 0 μS.

The first attempts to perform the coupling reaction were performed following a procedure present in literature [27], by reacting 1eq of **8** with variable amounts of *Primer Support^{TM} 5G Amino*, 0.9 eq of HBTU (*O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate) and 3eq of DIPEA, as described in Table 7.2. The solvent used was CH$_3$CN/DMF in 1:1 ratio. The mixture was left inside an oscillating agitator at 25°C overnight.

Entry	Compound 8	*Primer Support^{TM} 5G Amino*	HBTU	DIPEA
1	1 eq	1 eq	0,9 eq	3 eq
2	1 eq	0,75 eq	0,9 eq	3 eq
3	1 eq	0,50 eq	0,9 eq	3 eq

Table 7.2 First functionalization tries of Primer Support^{TM} 5G Amino.

The work-up performed for these first tries consists of a series of washes of the obtained compound; in particular, two washes were carried out with CH$_2$Cl$_2$, two with a mixture CH$_2$Cl$_2$/MeOH 1:1, two with a mixture CH$_2$Cl$_2$ 10% in petrol ether, then the mixture was concentrated under reduced pressure. From the TLC control of the reaction it was found that a considerable amount of starting (compound **8**) had not reacted. To confirm this, from the concentration of the washing solvents almost all of the compound **8** was found. To calculate precisely the amount of compound **8** that was bound to the support would have been necessary measurement of the loading, but since the procedure is laborious and expensive and, based on the considerations just made, very low values were expected, we decided not to take the measurement and instead set a new reaction by making some changes to the conditions described above.

We have therefore modified the reagent equivalents, the solvent and the condensing agent: HBTU has been replaced with HCTU (*O*-(1*H*-6-chlorobenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate). A drawback caused by the use of an excess of this type of condensing agent is the possible side reaction with the amine groups present, which leads to the formation of a guanidine group in place of the desired amide group [28,29]. To avoid this it is necessary to use an excess of the compound containing the carboxylic group, in our case **8**, with respect to the condensing agent.

As described in Scheme 7.7, we have made reacted 1.1 eq of **8** and 1 eq of *Primer Support^{TM} 5G Amino* in anhydrous CH$_3$CN/DMF 1:1 with 2 eq of DIPEA and 1 eq of HCTU, overnight at 25°C in an oscillating stirrer. The compound was filtered, washed with CH$_3$CN and CH$_2$Cl$_2$ (solvent that promotes swelling) and finally dried at reduced pressure. From the TLC control of the reaction, it seemed that the starting had reacted almost completely. We then have obtained the loading value of the derivatized support **5** (through a procedure provided to us by the manufacturing company of *Primer Support^{TM} 5G Amino*, see "Loading calculation" in paragraph 7.4.5.2 in Experimental part), which was 289 μmol/g. Considering that the loading value reported for the *Primer Support^{TM} 5G Amino* is 350-400 μmol/g, the yield calculated for our derivatization is 60-69%.
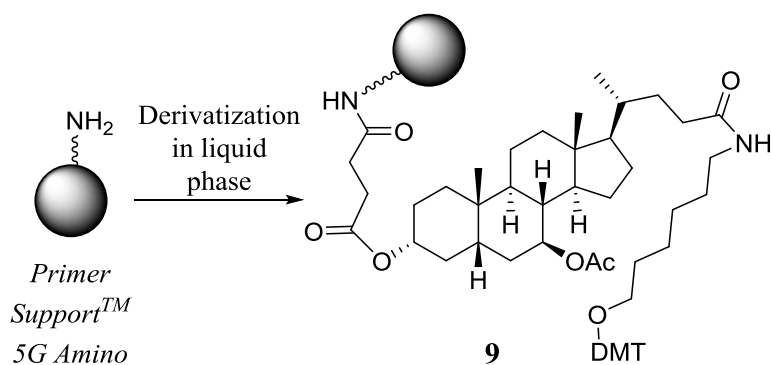
Scheme 7.7 *Support derivatization with bile acid moiety. Reagents and conditions: 8 (1.1eq), HCTU(1eq), DIPEA(2eq), CH₃CN/DMF 1:1 (v/v), 25°C, 18h (60-69% yield).*

Not all the amine groups of *Primer Support*^TM *5G Amino* were derivatized with **8** through the coupling reaction just described. This turns out to be a problem, because the amine functions remained free, in the course of the synthesis of the oligonucleotide, can react with the amidites, giving rise to secondary reactions and by-products. To overcome this problem, it is necessary to protect the amine groups with acetyl groups (Scheme 7.8). The amino groups that previously reacted with the compound **8** were not subject to the reaction.
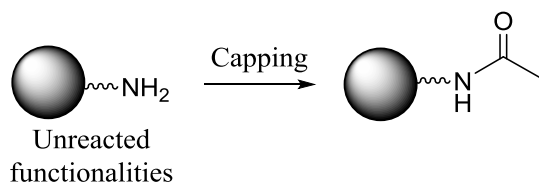
Scheme 7.8 *Capping of unreacted amino functionalities. Reagents and conditions: CapA (solution of N-ethylimidazole 20% in CH₃CN), CapB (solution 1:1 of CapB1 (acetic anhydride 40% in CH₃CN) and CapB2 (2,6-lutidine 60% in CH₃CN)), 25°C, 18h.*

Solid phase synthesis of AON51 on derivatized solid support

Before carrying out the synthesis of the oligonucleotide of our interest, a coupling test of the derivatized support **9** with the phosphoramidite U was made (Scheme 7.9), to see if on the functionalized support **9** it is possible to carry out a coupling reaction with a good yield.
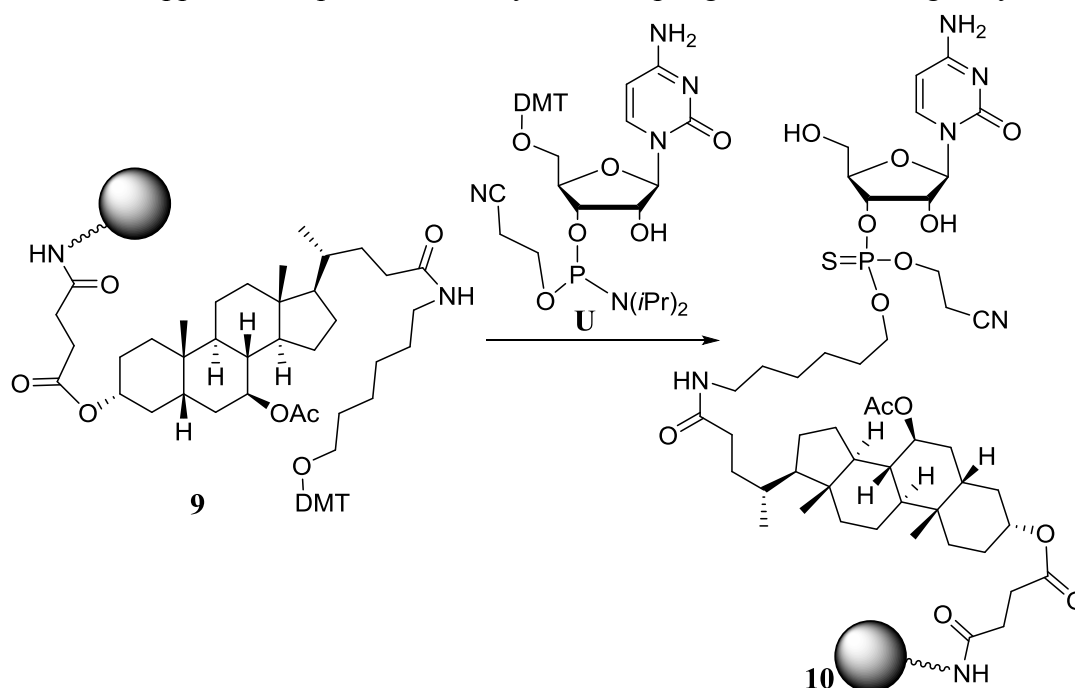
Scheme 7.9 *Test reaction of coupling between derivatized support 9 and phosphoramidite U.*

Through the appropriate software connected to the synthesizer it was possible to check the values of the detritylation areas, which corresponds to an estimate of the efficiency of the couplings. In relation to the synthesis scale, each type of amidite (U, A, C, G) presents a characteristic detritylation area, which is known to my research group thanks to the experience gained in the field. On the basis of this, we could state that the detritylation value of the amidite U was in line with the expected one and therefore the coupling reaction was performed with a good efficiency.

Since the coupling test with amidite U was successful, we proceeded with the solid phase synthesis of **AON51** in DMT-on mode on derivatized support **9**. After the automated synthesis, the oligonucleotide was treated with NH_4OH 33% at 50°C for 24h to release it from the solid support and remove the protecting groups, purified by RP-HPLC and subjected to the detritylation reaction in solution (Scheme 7.10), solubilizing in H_2O mQ and leaving it at r.t. for 18 h, then centrifuged to separate the supernatant from precipitate (4,4'-dimethoxitrityl alcohol; see paragraph 6.6). After decantation, the supernatant was lyophilized. Conjugate **AON51-3′-UDCA** was in this way obtained with a total yield of 30% after final chromatographic purification.
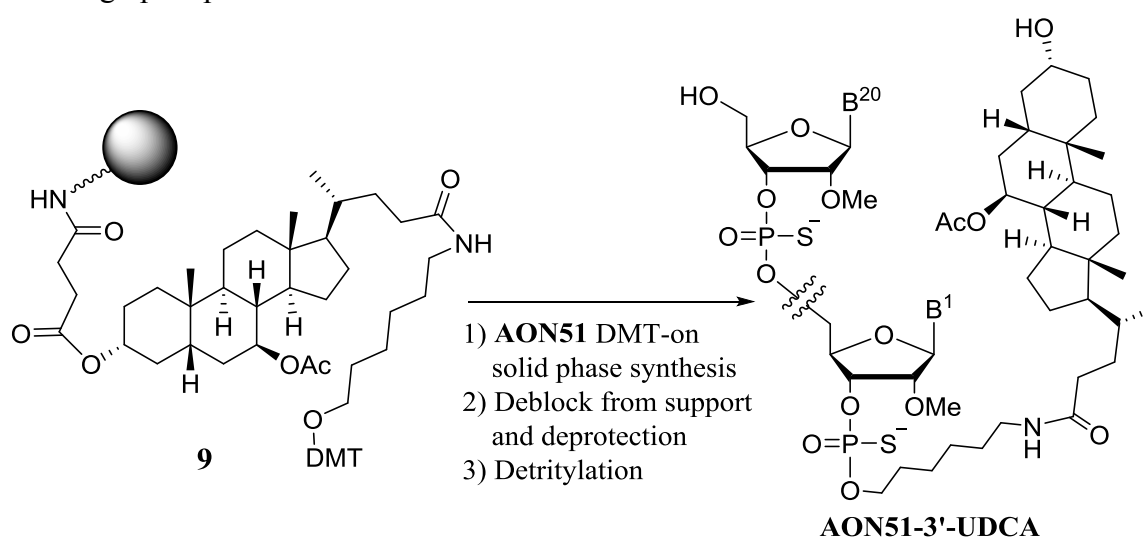

Scheme 7.10 *Synthesis of **AON51-3′-UDCA** through solid phase synthesis on derivatized support **9**.*

7.2.3 Synthesis of AON51-3′,5′-UDCA

The synthesis of the AON51 conjugate with two bile acid moieties at both ends 3' and 5' was achieved by combining the strategies seen for conjugations in 5' through the use of *ssH-Linker*[TM] and in 3' by synthetizing the oligonucleotide on *Custom Primer Support*[TM] *C6 Amino*. For the 3' conjugation the strategy of coupling in solution was preferred, since it is more technically simple and leads to a higher yield compared to solid phase synthesis on derivatized support.

AON51 was therefore synthesized in solid phase on *Custom Primer Support*[TM] *C6 Amino* in DMT-off mode, inserting the *ssH-Linker*[TM] after the last nucleotide of the sequence. The coupling between AON51 functionalized with amino linker at the 5′-end and the active ester of UDCA **3** was then performed in solid phase (**11**, Scheme 7.11), using the same two syringes method seen at paragraph 7.2.1. Subsequently, the oligonucleotide was treated with NH_4OH 33% at 50°C for 24h to release it from the solid support and remove the protecting groups and purified by RP-HPLC. At this point coupling at the 3′-end in solution was performed with the active ester of UDCA **3**, using the same conditions seen at paragraph

7.2.2.1. Conjugate **AON51-3′,5′-UDCA** (Scheme 7.11) was then obtained with a total yield of 58% after further chromatographic purification (Figure 7.8).
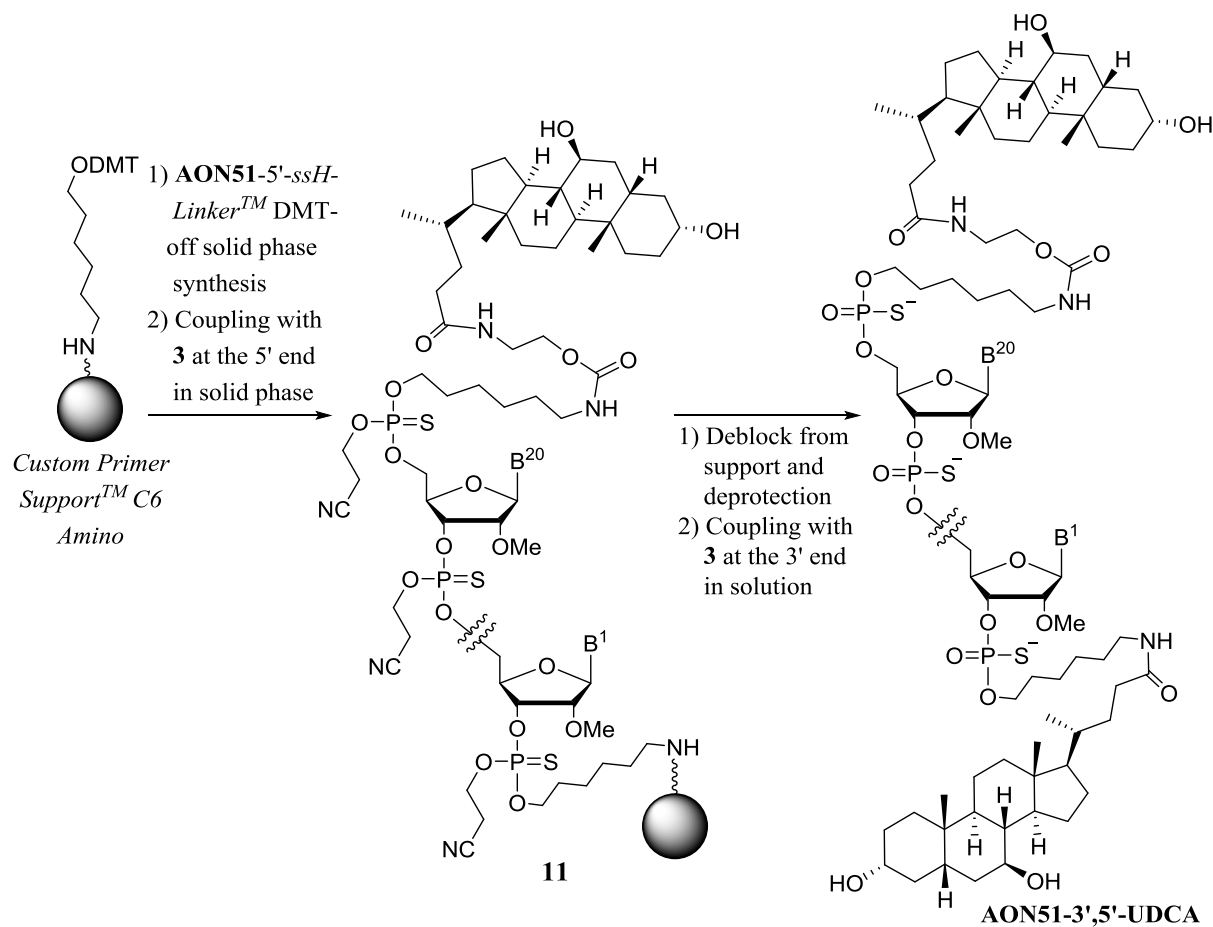
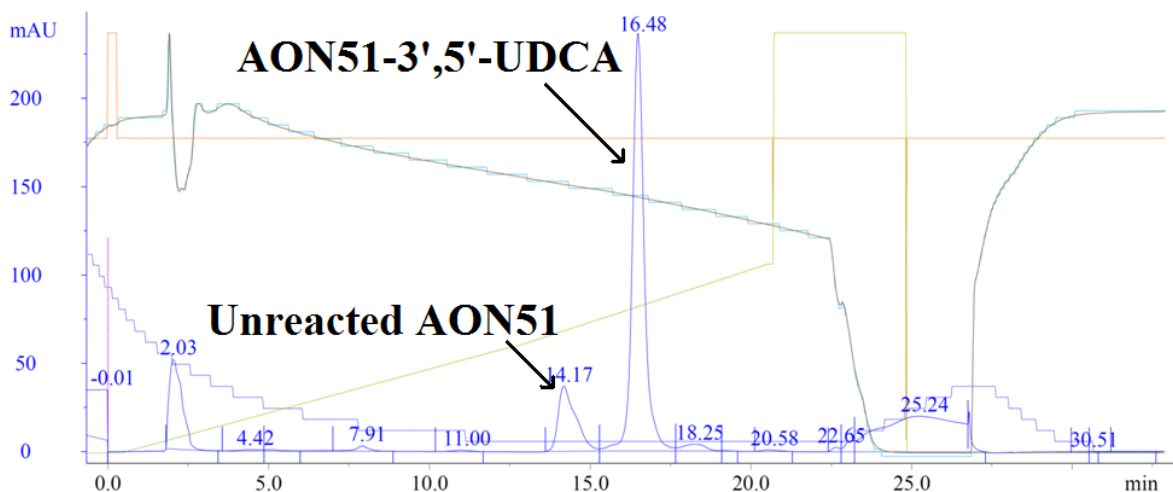**Scheme 7.11** *Synthesis of **AON51-3′,5′-UDCA** conjugate.*

Figure 7.8 *Chromatographic profile of final purification of **AON51-3′,5′-UDCA** by RP-HPLC.*

7.2.4   Ion exchange

All the synthesized oligonucleotides are subjected to ion exchange before being used for biological tests. The triethylammonium ions (TEA$^+$), present as counterions of the thiophosphate groups, have been exchanged with Na$^+$ ions using ion-exchange resin (see paragraph 6.7). *HiTrap Capto S* columns, installed on the HPLC apparatus and loaded with a NaCl 2M solution, have been employed.

7.2.5 *In vitro* biological assays of AON51-UDCA conjugates

All new conjugated AONs were tested *in vitro* using immortalized myogenic cells of patients with DMD, in order to evaluate their efficacy in restoring dystrophin expression. All *in vitro* tests have been performed by dr. Matteo Bovolenta at the Genethon Institute of Paris.

The first *in vitro* tests were performed on the **AON51-5′-UDCA**, **AON51-3′-UDCA** and **AON51-3′,5′-UDCA** conjugates. Exon skipping induced by AONs directed against exon 51 of the dystrophin gene in immortalized myogenic cells derived from a patient with DMD was quantified. These tests were performed treating cells with a solution of AON 100μM and using *JetPEI*® as transfecting reagent for 48h. Unconjugated **AON51** was used as a reference. From the analysis of the dystrophin transcript it was found that conjugated antisense oligonucleotides **AON51-5′-UDCA** and **AON51-3′-UDCA** are both more effective than unconjugated **AON51** in inducing exon skipping (Figure 7.9). The conjugate **AON51-3′,5′-UDCA** instead proved to be inefficient (Figure 7.10).

Figure 7.9 *Percentage of skipping induced by **AON51-5′-UDCA** and **AON51-3′-UDCA**, using unconjugated **AON51** as a reference. Conditions: 2μL AON 100μM, 4μL JetPEI®, 48h.*

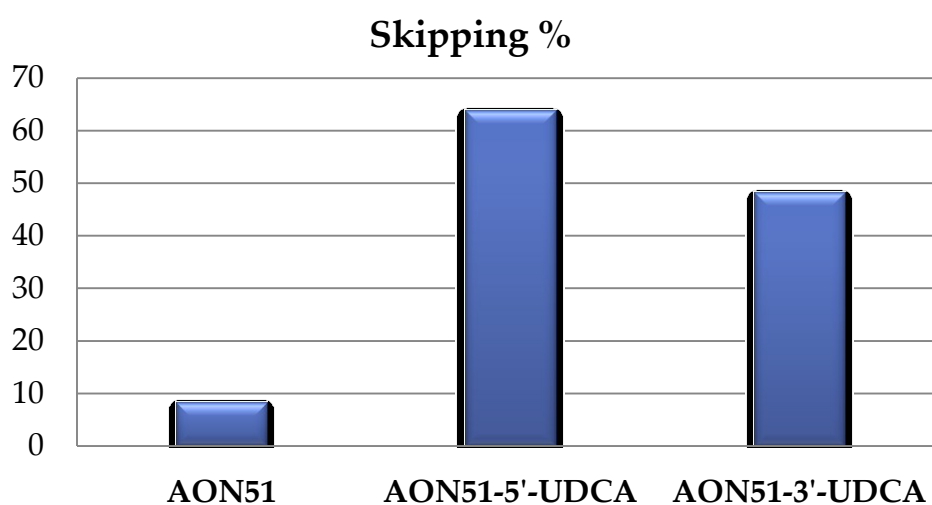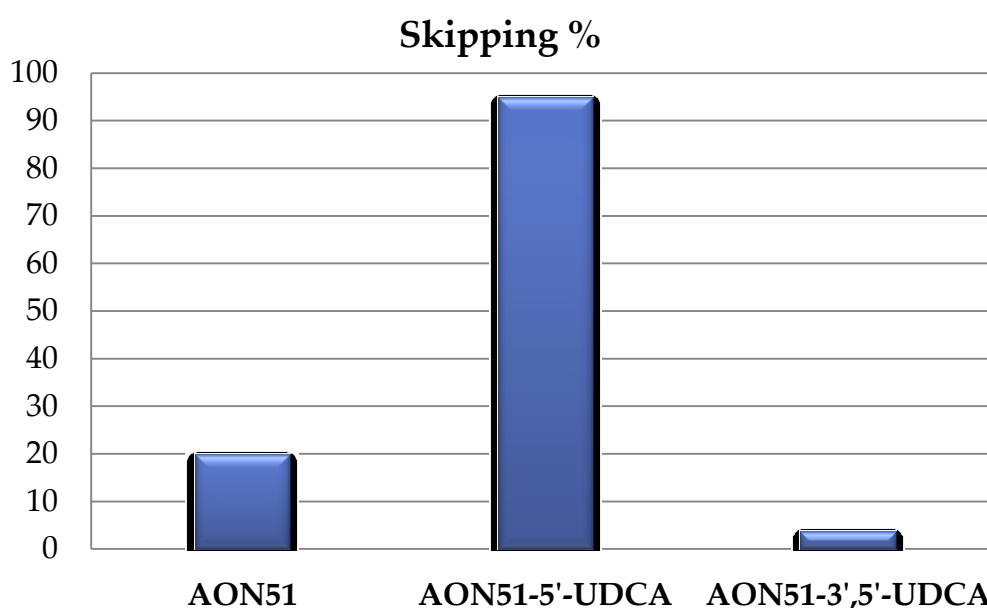

Figure 7.10 *Percentage of skipping induced by **AON51-5′-UDCA** and **AON51-3′,5′-UDCA**, using unconjugated **AON51** as a reference. Conditions: 4μL AON 100μM, 8μL JetPEI®, 48h.*

In particular, **AON51-5′-UDCA** and **AON51-3′-UDCA** led to an increase in skipping of 7.65 times and 5.88 times respectively compared to the unconjugated **AON51**. The best result is therefore associated with **AON51-5′-UDCA**.

Immunofluorescent analyzes of dystrophin (in red in Figure 7.11) were also performed in immortalized myogenic cells derived from a patient with DMD treated with **AON51**, **AON51-5′-UDCA** and **AON51-3′-UDCA**. From these analyzes the restored expression of dystrophin and correct localization to sarcolemma only in myotubes treated with conjugated antisense oligonucleotides were detected.
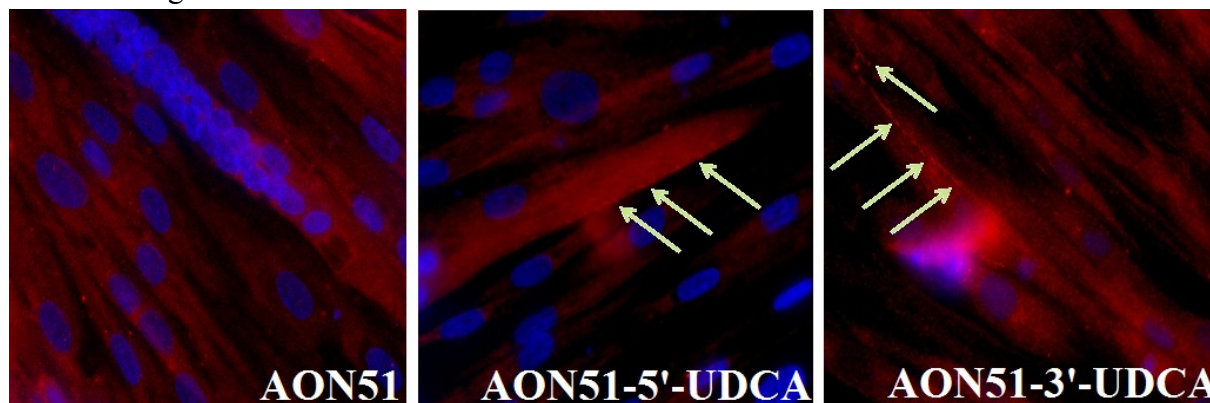

*Figure 7.11 Immunofluorescent analyzes of dystrophin (in red) in immortalized myogenic cells derived from a patient with DMD treated with **AON51**, **AON51-5′-UDCA** and **AON51-3′-UDCA**. Conditions: 2μL AON 100μM, 4μL JetPEI®, 5 days.*

7.2.6 Synthesis of conjugates between AON51 and other lipophilic molecules

Considering the results of the *in vitro* tests on the AON51-UDCA conjugates, which showed that the best results derived from the conjugate at the 5′-end (see paragraph 7.2.5), and since the conjugation in this position is also the most technically easy to achieve, we have developed a library of conjugates between the same **AON51** with amino linker at the 5′-end and other lipophilic molecules (Figure 7.3). This library of conjugates was constructed to evaluate whether the increased activity of **AON51-5′-UDCA** compared to **AON51** seen in *in vitro* tests was due solely to the lipophilic nature of UDCA or to its particular structure.

The synthesis of the conjugates with the natural compounds HDCA, EPA and DHA and with the synthesized compounds **Succ-UDC-OMe** and **DHA-UDCA** has foreseen the activation of their carboxylic acid functional groups as *N*-succinimidyl esters, exactly as done for UDCA (Scheme 7.2). The methodology used for the coupling was identical to that used to make **AON51-5′-UDCA** (paragraph 7.2.1).

The preparation of *N*-succinimidyl esters of **Succ-UDC-OMe** (**14**) and **DHA-UDCA** (**17**) is shown in Scheme 7.12; compound **13** represents a common intermediate between the two synthetic routes. The coupling between **13** and active ester of DHA **15** was performed to obtain the hybrid **16**.
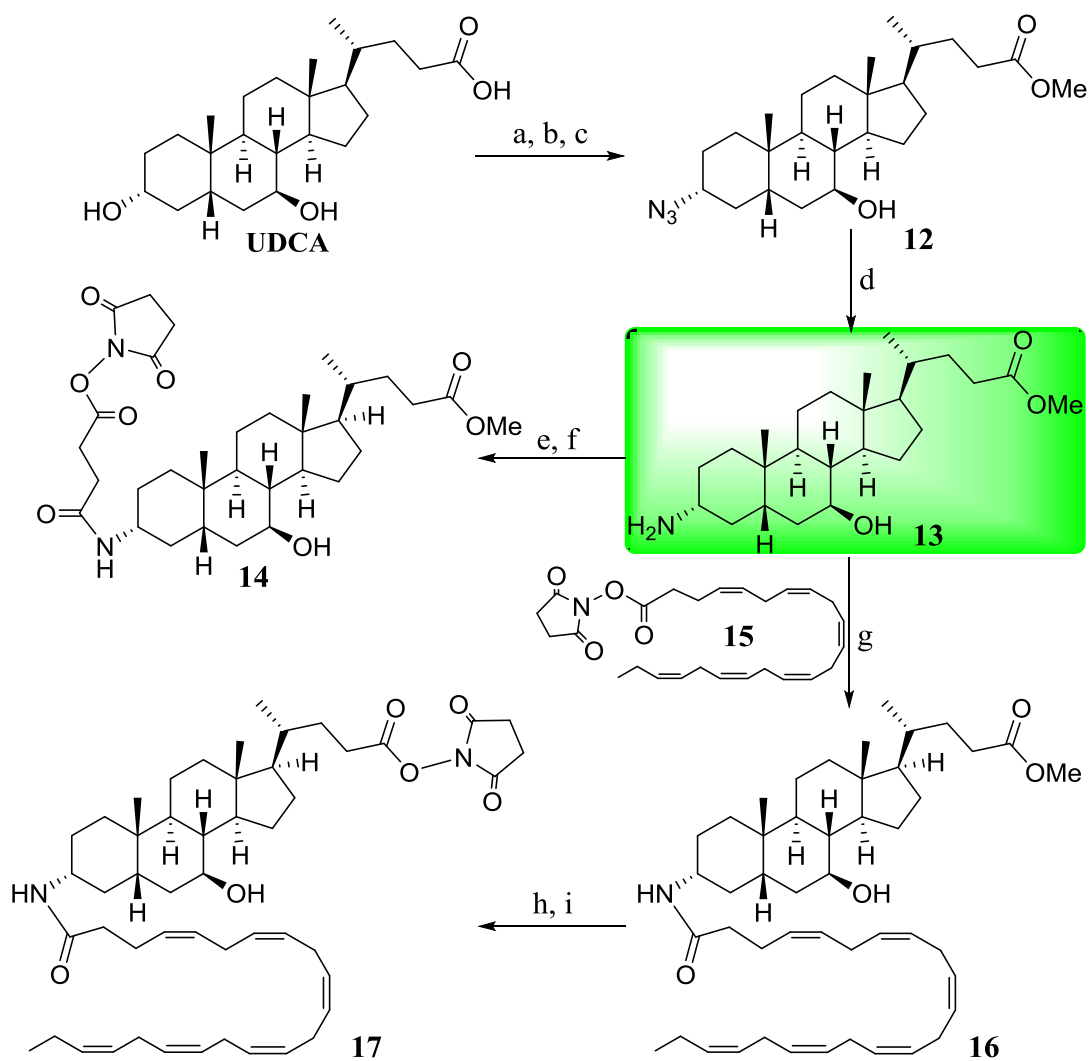
Scheme 7.12 Syntheses of N-succinimidyl esters of Succ-UDC-OMe (**14**) and DHA-UDCA (**17**). Reagents and conditions: (a) MeOH, H_2SO_4, 80°C, 2h; (b) N-bromosuccinimide, PPh_3, THF, 25°C, 90min; (c) NaN_3, DMF, 30°C, 18h (62% yield in three steps); (d) $NH_4^+COO^-$, Pd/C, AcOEt, MeOH, 70°C, 18h (78% yield); (e) succinic anhydride, DMAP, Py, 115°C, 18h; (f) NHS, DCC, THF, 25°C, 18h (39% yield in two steps); (g) DIPEA, DMF, 25°C, 18h (19% yield); (h) LiOH, H_2O, MeOH, 25°C, 18h; (i) NHS, DCC, THF, 25°C, 18h (25% yield in two steps).

For the synthesis of conjugate with **Bis-UDCA-Lys**, a different coupling method has been used instead, as shown in Scheme 7.13. HBTU was used as the activating agent for the coupling, since the usual NHS / DCC system proved ineffective in this case. This conjugate was synthesized to verify if the presence of two units of UDCA at the same end of the oligonucleotide further increase its activity. Lysine has been chosen as a linker because it has two amino functions and a carboxyl.
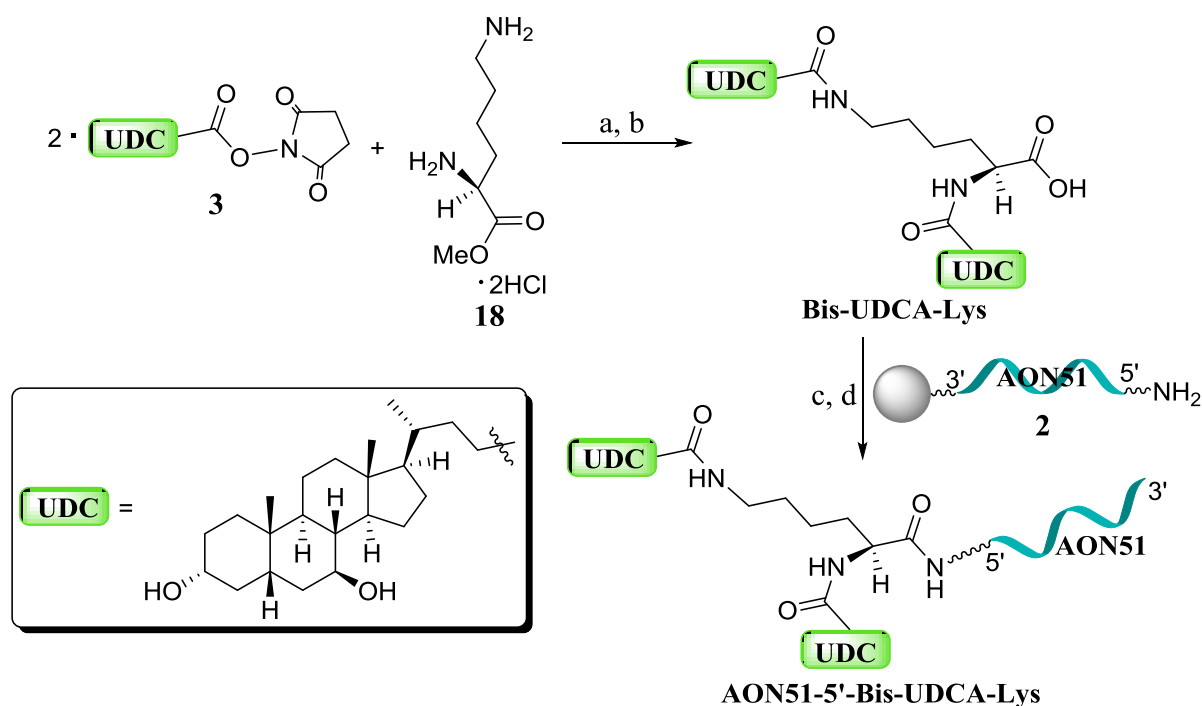
Scheme 7.13 *Synthesis of **AON51-5′-Bis-UDCA-Lys**. Reagents and conditions: (a) DIPEA, CH₂Cl₂, 25°C, 18h; (b) NH₄OH 33%, MeOH, 60°C, 18h (75% yield in two steps); (c) **2**, HBTU, DIPEA, DMSO, 25°C, 3h; (d) NH₄OH 33%, 50°C, 24h (20% yield in two steps).*

7.2.7 *In vitro* biological tests of conjugates between AON51 and other lipophilic molecules

In vitro tests were then performed on conjugates containing fatty acids: **AON51-5′-EPA, AON51-5′-DHA** and **AON51-5′-DHA-UDCA**. These tests were performed treating cells with a solution of AON 100μM and using *TurboFect*® as transfecting reagent for 72h. **AON51-5′-UDCA** has been retested to confirm the excellent results of the previous tests. Unconjugated **AON51** was used as a reference (Figure 7.12).
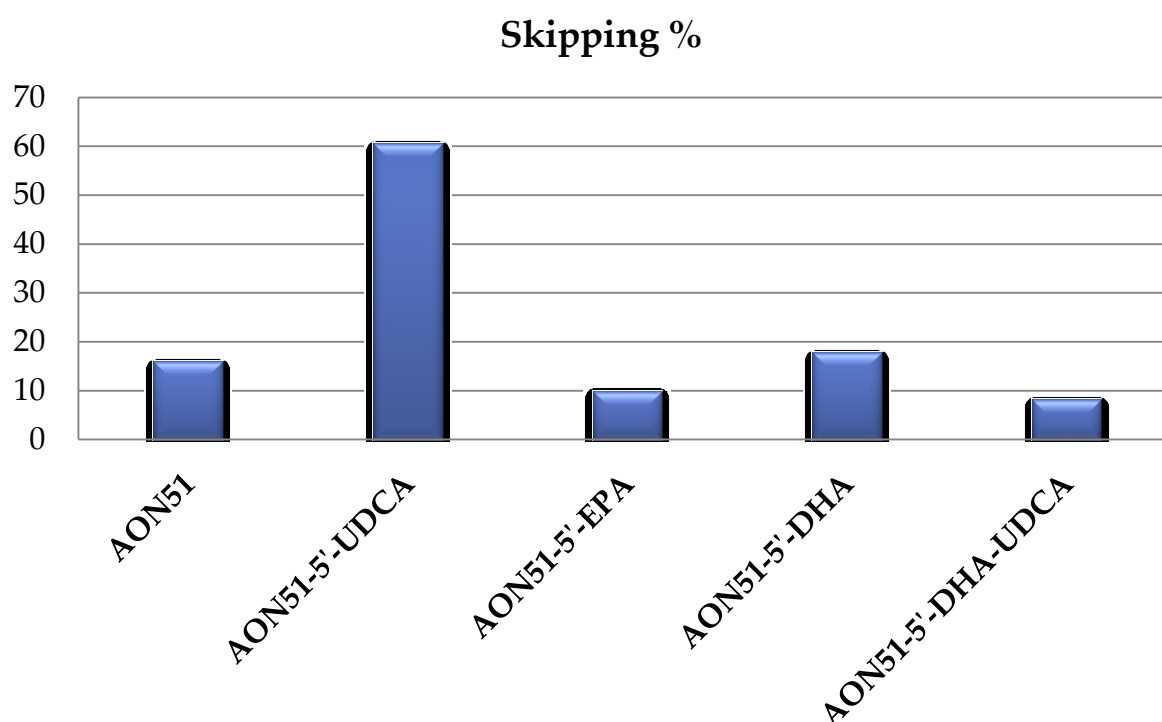

Figure 7.12 *Percentage of skipping induced by **AON51-5′-UDCA, AON51-5′-EPA, AON51-5′-DHA** and **AON51-5′-DHA-UDCA** using unconjugated **AON51** as a reference. Conditions: 2μL AON 100μM, 4μL TurboFect®, 72h.*

As shown in Figure 7.12, results obtained for the conjugates containing fatty acid moieties were disappointing, while **AON51-5′-UDCA** confirmed itself as the most effective compound in the induction of skipping.

Finally, gymnotic delivery tests (*i.e.* in the absence of a transfecting reagent) were performed of **AON51-5′-UDCA**, **AON51-5′-HDCA**, **AON51-5′-Succ-UDC-OMe** and **AON51-5′-Bis-UDCA-Lys** conjugates. These tests were performed treating cells with a concentrated solution of AON to make a final concentration of 50μM in each well, for 72h. Unconjugated **AON51** was used as a reference. An untreated sample with zero induced skipping is also reported (Figure 7.13).

As can be seen in Figure 7.13, conjugates containing UDCA moieties have shown similar induced skipping values. The HDCA conjugate also showed better induced skipping compared to unconjugated **AON51**, but lower compared to UDCA conjugates. Also in this case the best result was obtained for **AON51-5′-UDCA**, which confirmed itself as the most effective compound in the induction of skipping.
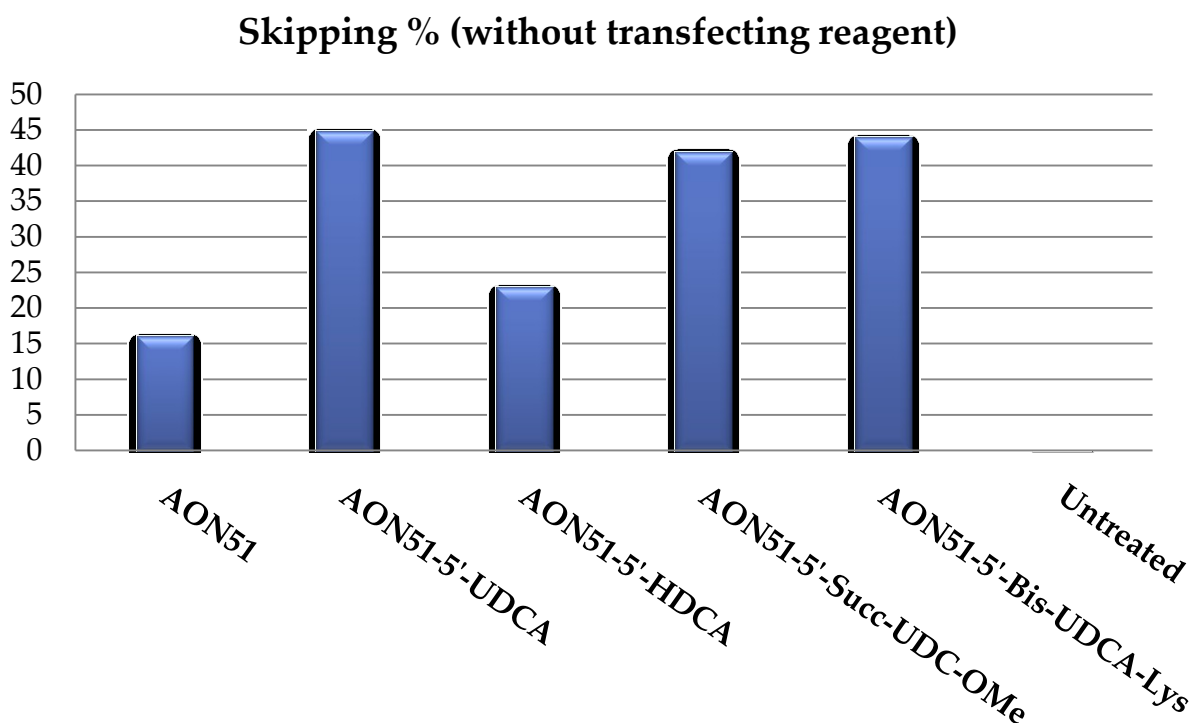
Skipping % (without transfecting reagent)

Figure 7.13 Percentage of skipping induced by **AON51-5′-UDCA**, **AON51-5′-HDCA**, **AON51-5′-Succ-UDC-OMe** and **AON51-5′-Bis-UDCA-Lys** using unconjugated **AON51** as a reference. An untreated sample is also reported. Conditions: AON 50μM (final concentration in each well), 72h.

7.2.8 *In vivo* biological tests

The effectiveness in inducing skipping of antisense oligonucleotides conjugated with UDCA was tested *in vivo* on mdx dystrophic mice. These mice present the genetic defect that causes muscular dystrophy on the exon 23 of their dystrophin gene. The nucleotide sequence of the AON that binds to this exon, which we will call **m23-AON**, is 5′-GGCCAAACCUCGGCUUACCU-3’. The conjugate at the 5′-end of this AON with UDCA (**m23-AON-5′-UDCA**) has been synthesized in an analog way to **AON51-5′-UDCA**, using the same *ssH-Linker^{TM}*. All *in vivo* tests have been performed by dr. Paola Braghetta at the Department of Molecular Medicine of Padua.

The oligonucleotides **m23-AON** and **m23-AON-5′-UDCA** are injected intraperitoneally at a dose of 200 mg/kg with a one-week administration regimen for 12 weeks in 2-month

C57BL/10ScSn-Dmd^{mdx}/J male mice. Mice injected with PBS (phosphate buffer saline) were used as controls. One week after the last treatment, the mice were sacrificed to collect samples of the heart, diaphragm, gastrocnemius and tibialis anterior. Muscles have been divided and fragmented to conduct transcript analysis and protein quantification.

With the exception of the heart, in all the analyzed muscles the treatment with **m23-AON-5′-UDCA** induced higher levels of skipping than the unconjugated **m23-AON**, with the highest levels of skipping identified in the diaphragm (Figure 7.14).

The quantification of the protein by western blot showed an increase in dystrophin produced in both treatments, with a greater quantity in mice treated with the antisense **m23-AON-5′-UDCA** (Figure 7.15).
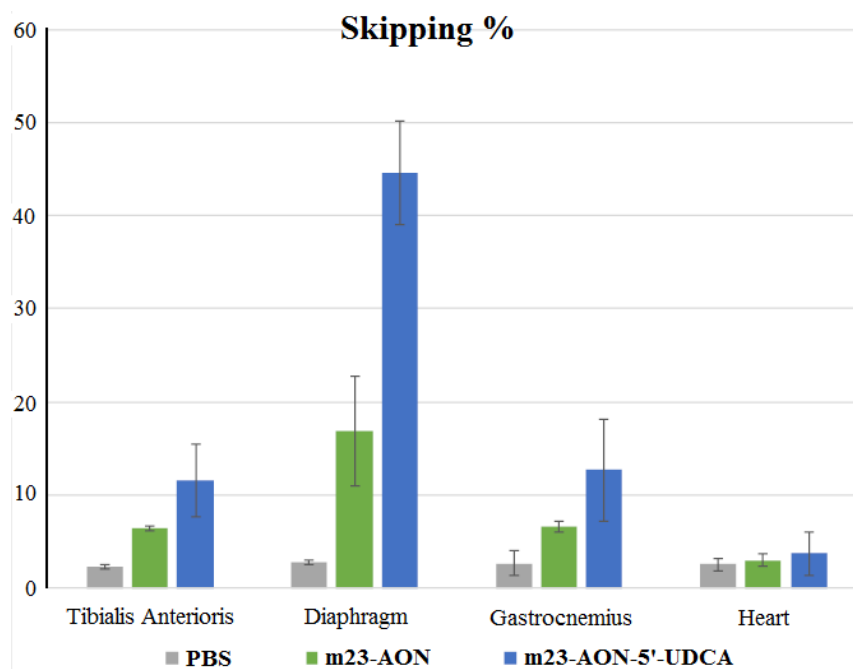

Figure 7.14 *Percentage of skipping exon 23 of the dystrophin gene in the muscles of C57BL/10ScSn-Dmd^{mdx}/J mice treated for 12 weeks.*
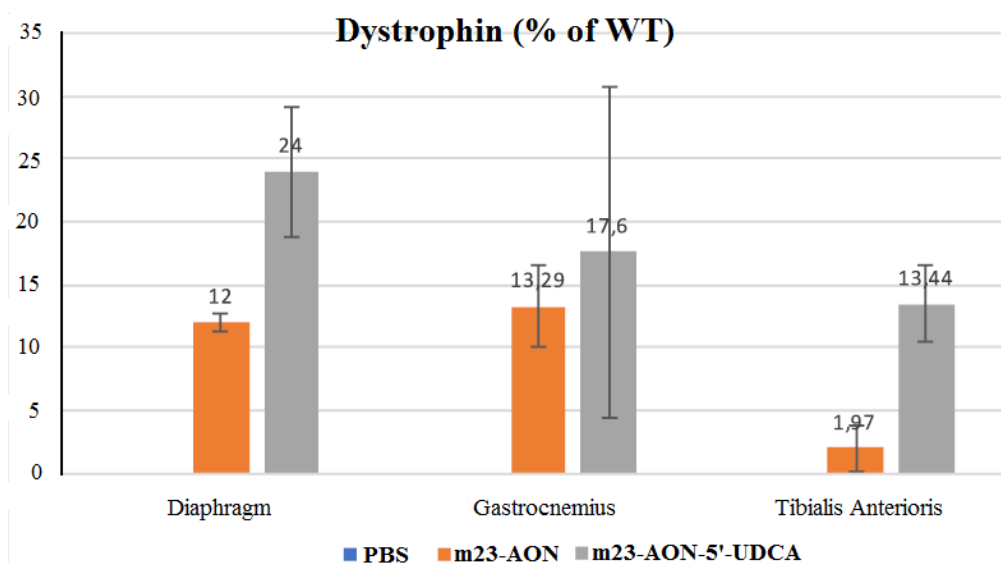

Figure 7.15 *Quantification of the dystrophin protein in the muscles of C57BL/10ScSn-Dmd^{mdx}/J mice treated for 12 weeks.*

7.3 Conclusion

In this work, new synthetic approaches have been developed for the conjugation of a lipophilic molecule to the 2′-*O*MePS-AON that is able to associate with exon 51 (**AON51**) of the human dystrophin gene responsible for DMD, in order to improve the exon skipping approach.

First the conjugates with UDCA are synthesized: **AON51-5′-UDCA**, **AON51-3′-UDCA** (with 2 different strategies) and **AON51-3′,5′-UDCA**. These compounds were tested *in vitro* using immortalized myogenic cells of patients with DMD, in order to evaluate their efficacy in restoring dystrophin expression. **AON51-5′-UDCA** and **AON51-3′-UDCA** are both resulted more effective than unconjugated **AON51** in inducing exon skipping, while **AON51-3′,5′-UDCA** proved to be inefficient, probably due to the excessive steric encumbrance given by the two bile acid units that interferes with the recognition between bases according to Watson and Crick. The best result is associated with **AON51-5′-UDCA**.

In consideration of these results, and since the conjugation at the 5′-end is also the most technically easy to achieve, we have synthesized a series of conjugates between the same **AON51** at the 5′-end and other lipophilic molecules, in order to evaluate whether the increased activity of **AON51-5′-UDCA** compared to unconjugated **AON51** was due solely to the lipophilic nature of UDCA or to its particular structure. The new conjugates **AON51-5′-EPA, AON51-5′-DHA**, **AON51-5′-DHA-UDCA**, **AON51-5′-HDCA**, **AON51-5′-Succ-UDC-OMe** and **AON51-5′-Bis-UDCA-Lys** were all tested *in vitro*: these tests showed the ineffectiveness of the compounds containing fatty acid moieties, confirming the importance of bile acid moiety to increase the exon skipping. Conjugates containing UDCA moieties have shown similar induced skipping values. The HDCA conjugate also showed better induced skipping compared to unconjugated **AON51**, but lower compared to UDCA conjugates. The most effective compound in the induction of skipping in *in vitro* tests turned out to be **AON51-5′-UDCA**.

To test *in vivo* the efficacy of antisense oligonucleotides conjugated at the 5′-end with UDCA, **m23-AON-5′-UDCA** has been synthesized in an analog way to **AON51-5′-UDCA** and administrated to mdx dystrophic mice. With the exception of the heart, in all the analyzed muscles of sacrificed mice the treatment with **m23-AON-5′-UDCA** induced higher levels of skipping than the unconjugated **m23-AON**, with the highest levels of skipping identified in the diaphragm. The quantification of the protein showed a greater quantity of dystrophin produced in mice treated with the antisense **m23-AON-5′-UDCA** compared with the ones treated with unconjugated **m23-AON**.

In light of the results obtained, we can state that the conjugation of ursodeoxycholic acid at the 5′-end of the antisense oligonucleotide is an effective method to improve the exon skipping approach.

The data reported regarding this project, in agreement with I.C.E. S.p.A. funding company of Reggio Emilia, are classified and currently awaiting patent approval.

7.4 Experimental

7.4.1 General information

7.4.1.1 Reactions in solution

Reactions were monitored by TLC on pre-coated silica gel plates (thickness 0.25mm, Merck, Darmstadt, Germany), and phosphomolybdic acid solution was used as the spray reagent to visualize the steroids. Flash column chromatography was performed on *Silica gel 60* (230–400mesh, Sigma-Aldrich). ESI-MS were acquired on a *Thermo Finnigan LCQ DUO ion trap* spectrometer (Thermo Fisher Scientific, Waltham, Massachussets, USA), using methanol as solvent. NMR spectra were recorded with a *Mercury Plus 400 MHz* instrument (Varian, Palo Alto, CA, USA); the used solvent is indicated from time to time. Bile acids are commercially available compounds that were used without further purification. Methyl 3α-azido-7β-hydroxy-5β-cholan-24-oate (**12**) was prepared according to the literature procedures [18,30].

7.4.1.2 Functionalization of Primer Support[TM] 5G Amino

For preliminary activation of *Primer Support[TM] 5G Amino* was used the oligosynthesizer *Äkta oligopilot 10 Plus* (GE Healthcare), following a standard procedure provided by GE Healthcare. The reactions in liquid phase were performed in a thermostatic oscillating stirrer *Asal DVRL 711/CT* and monitored by TLC on pre-coated silica gel plates (thickness 0.25mm, Merck). For the loading calculation of functionalized support spectrophotometric analyzes were performed, following a standard procedure provided by GE Healthcare. Spectrophotometric analyzes were performed by a *Varian CARY 100 Bio* spectrophotometer. DMF, CH_3CN and the solution of DIPEA 2% in CH_3CN were anhydrificated using molecular sieves *EZ DRY moisture trap* (emp Biotech), while *O*-(1*H*-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) was coevaporated three times with anhydrous CH_3CN for the synthesis of oligonucleotides under atmosphere of N_2. The used material for the reactions (needles, syringes, tips, etc.) was dried leaving in a heater at 60°C o.n., then cooled at r.t. and stocked in a drier. The reagents used for capping reaction corresponds to: CapA= solution of *N*-methylimidazole 20% (V:V) in CH_3CN; CapB= solution 1:1 of CapB1 (acetic anhydride 40% (V:V) in CH_3CN) and CapB2 (2,6-lutidine 60% (V:V) in CH_3CN).

7.4.1.3 Oligonucleotide solid phase synthesis

The synthesis in solid phase was performed with oligosynthesizer *Äkta oligopilot 10 Plus* (GE Healthcare). Analysis and purification of the crude product were made by the HPLC apparatus *Äkta purifier* on reverse phase column *Resource RPC 3 mL* (GE Healtcare), based on polystyrene/divinylbenzene. The ionic exchange was performed on columns *HiTrap Capto S* (GE Healthcare). Oligosynthesizer and HPLC apparatus were interfaced with computer by *Unicorn 5.10* software (GE Healthcare). Lyophilizations were made by lyophilizer *Labconco FreeZone 1* connected to the centrifuge *Christ RVC 2-18 CDplus*. ESI-MS were acquired on a *Thermo Finnigan LCQ DUO ion trap* spectrometer (Thermo Fisher Scientific), using water as solvent. ^{31}P-NMR spectra were recorded with a *Mercury Plus 400 MHz* instrument (Varian, Palo Alto, CA, USA), using D_2O. Spectrophotometric analyzes were performed by a *Varian CARY 100 Bio* spectrophotometer.

### 7.4.2 Derivatization of bile and fatty acids

7.4.2.1 General procedure for the synthesis of active esters

N-succimidyl esters of carboxylic acids were prepared adding *N*-hydroxy succinimide (1.528 mmol) to a solution of the appropriate carboxylic acid (**UDCA**, **HDCA**, **EPA**, **DHA**, **DHA-UDCA** or **Succ-UDC-OMe**; 1.274 mmol) in dry THF (10 mL). To the resulting homogeneous solution, dicyclohexyl-carbodiimide (1.528 mmol) in dry THF (5 mL) was added at 0°C. The mixture was stirred at 25°C for 18 h and the precipitated *N,N*-dicyclohexyl urea was removed by filtration. THF was removed under reduced pressure and the residue was extracted with ethyl acetate (20 mL) and washed successively with aqueous $NaHCO_3$ (10 mL), water (10 mL) and then with brine (10 mL). The extract was dried over Na_2SO_4 and AcOEt was removed under reduced pressure to get crude product.
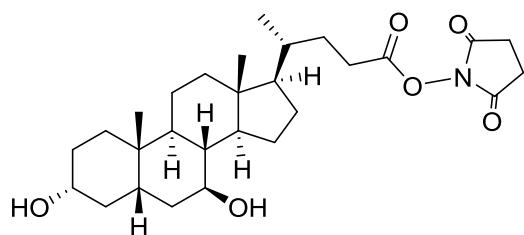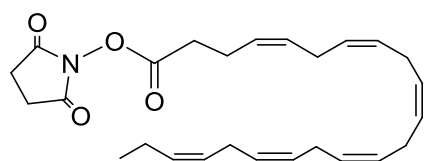
N-succimidyl ester of UDCA **3**

Purified by crystallization from AcOEt, white amorphous solid, yield ≥95%. ^{1}H-NMR (400 MHz, $CDCl_3$): δ = 3.64 – 3.51 (m, 2H, 3β- and 7α-H), 2.83 (s, 4H, $2CH_2$ of NHS), 2.72 – 2.45 (m, 2H, 23-CH_2), 2.06 – 0.90 (m, 30H), 0.68 (s, 3H, 18-CH_3). ^{13}C-NMR (101 MHz, $CDCl_3$) δ = 169.21, 169.07, 71.44, 71.35, 55.66, 54.75, 45.77, 43.74, 42.40, 40.08, 39.13, 37.24, 36.79, 35.04, 34.90, 34.06, 30.61, 30.29, 28.55, 28.00, 26.85, 25.58, 25.38, 23.37, 21.15, 18.28, 12.10, 8.60. MS (ESI, ES+): calculated for $C_{28}H_{43}NO_6$ 489.65; found 512.27 $[M+Na]^+$.
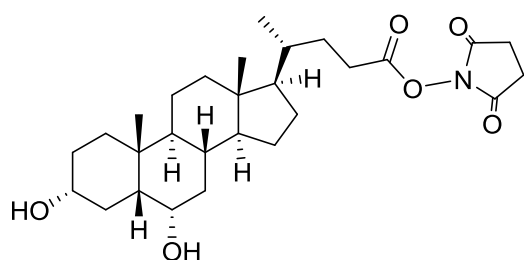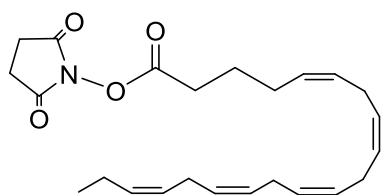
N-succimidyl ester of HDCA **19**

Purified by crystallization from AcOEt, white amorphous solid, yield 79%. ^{1}H-NMR (400 MHz, $CDCl_3$): δ= 4.08 – 3.99 (m, 1H, 6β-H), 3.66 – 3.55 (m, 1H, 3β-H), 2.83 (d, *J* = 4.4 Hz, 4H, $2CH_2$ of NHS), 2.69 – 2.60 (m, 1H, 23-CH_{2a}), 2.57 – 2.46 (m, 1H, 23-CH_{2b}), 1.99 – 0.99 (m, 24H), 0.94 (d, *J* = 6.3 Hz, 3H, 21-CH_3), 0.89 (s, 3H, 19-CH_3), 0.64 (s, 3H, 18-CH_3). ^{13}C-NMR (101 MHz, $CDCl_3$): δ= 169.21, 169.07, 71.44, 71.35, 55.66, 54.75, 45.77, 43.74, 42.40, 40.08, 39.13, 37.24, 36.79, 35.04, 34.90, 34.06, 30.61, 30.29, 28.55, 28.00, 26.85, 25.58, 25.38, 23.37, 21.15, 18.28, 12.10, 8.60. MS (ESI, ES+): calculated for $[C_{28}H_{43}NO_6 + Na]^+$ 512.64; found 512.40; calculated for $[2·C_{28}H_{43}NO_6 + H]^+$ 980.31; found 979.27. MS (ESI, ES-): calculated for $C_{28}H_{43}NO_6$ 489.65; found 488.47 $[M-H]^-$, 977.27 $[2M-H]^-$.

N-succimidyl ester of DHA **15**

Flash chromatography (Cyclohexane/AcOEt 4:1), orange syrup, yield 78%. ^{1}H-NMR (400 MHz, $CDCl_3$): δ= 5.59 – 5.21 (m, 12H, alkenic H), 2.93 – 2.74 (m, 14H, $2CH_2$ of NHS and =CH-CH_2-CH=), 2.67 (t, J = 7.1 Hz, 2H, 2-CH_2), 2.50 (q, J = 7.2 Hz, 2H, 3-CH_2), 2.07 (p, J = 7.5 Hz, 2H, 21-CH_2), 0.97 (t, J = 7.5 Hz, 3H, 22-CH_3). ^{13}C-NMR (101 MHz, $CDCl_3$): δ= 169.08, 168.06, 132.02, 130.34, 128.53, 128.42, 128.24, 128.06, 127.87, 127.78,

126.99, 126.46, 30.91, 25.57, 22.33, 20.55, 14.27. MS (ESI, ES+): calculated for $C_{26}H_{35}NO_4$ 425.57; found 425.47 [M+H]$^+$.

N-succimidyl ester of EPA 20

Flash chromatography (Cyclohexane/AcOEt 4:1), orange syrup, yield 78%. ^{1}H-NMR (400 MHz, CDCl$_3$): δ= 5.48 – 5.26 (m, 10H, alkenic H), 2.89 – 2.75 (m, 12H, 2CH$_2$ of NHS and CH-CH_2-CH), 2.61 (t, J = 7.5 Hz, 2H, 2-CH$_2$), 2.19 (q, J = 7.4 Hz, 2H, 4-CH$_2$), 2.07 (p, J = 6.8 Hz, 2H, 3-CH$_2$), 1.82 (p, J = 7.5 Hz, 2H, 19-CH$_2$), 0.96 (t, J = 7.5 Hz, 3H, 20-CH$_3$). ^{13}C-NMR (101 MHz, CDCl$_3$): δ= 169.24, 169.10, 71.56, 68.03, 56.09, 55.77, 48.35, 42.88, 39.89, 39.75, 35.92, 35.52, 35.12, 34.98, 34.80, 30.54, 30.20, 29.17, 28.00, 24.16, 23.47, 20.72, 18.15, 12.00. MS (ESI, ES+): calculated for $C_{24}H_{33}NO_4$ 399.53; found 400.80 [M+H]$^+$.
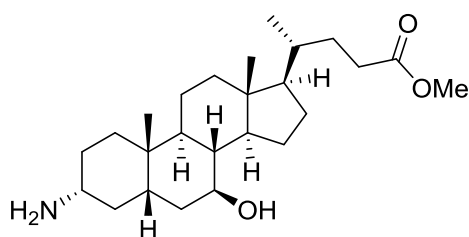
3α-NH₂-UDC-OMe 13

To a solution of 3α-N$_3$-UDC-OMe 12 (1.043 mmol) and NH$_4^+$HCOO$^-$ (10.430 mmol) in AcOEt/MeOH 1:1 (10 mL), Pd/C (2.086 mmol) in MeOH (5 mL) was added slowly. The mixture was stirred at 70°C for 18 h. Pd/C was removed by filtration. The solvent was removed under reduced pressure and the residue was extracted with CH$_2$Cl$_2$ (15 mL) and washed with brine (10 mL). The extract was dried over Na$_2$SO$_4$ and CH$_2$Cl$_2$ was removed under reduced pressure to get crude solid. Additional purification was not required. White amorphous solid, yield 78%. ^{1}H NMR (400 MHz, CDCl$_3$): δ= 3.62 (s, 3H, OMe), 3.59 – 3.50 (m, 1H, 7α-H), 2.64 (bs, 1H, 3β-H), 2.37 – 2.27 (m, 1H, 23-CH_{2a}), 2.23 – 2.13 (m, 1H, 23-CH_{2b}), 1.99 – 0.85 (m, 30H), 0.64 (s, 3H, 18-CH$_3$). ^{13}C NMR (101 MHz, CDCl$_3$): δ= 174.66, 70.93, 55.83, 54.93, 51.45, 51.24, 43.73, 43.65, 42.89, 40.15, 39.23, 38.36, 37.19, 35.66, 35.32, 34.11, 31.10, 31.01, 28.62, 26.93, 23.63, 21.16, 18.35, 12.11. MS (ESI, ES+): calculated for $C_{25}H_{43}NO_3$ 405.62; found 406.33 [M+H]$^+$, 811.27 [2M+H]$^+$, 1215.87 [3M+H]$^+$. MS (ESI, ES-): found 404.40 [M-H]$^-$, 805.07 [2M-H]$^-$.
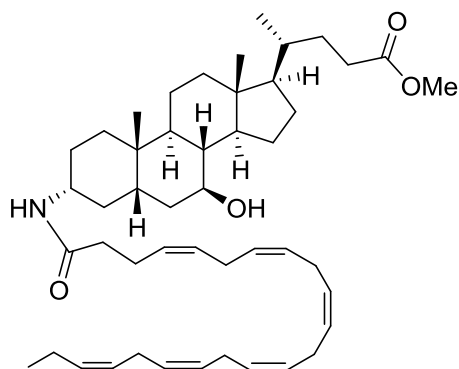
DHA-UDC-OMe 16

To a solution of 15 (1.752 mmol) in DMF (10 mL), 13 (1.752 mmol) and DIPEA (3.504 mmol; 491 μL) were added. The mixture was stirred at 25°C for 18 h, then 10 mL of aqueous HCl at pH=5 were added. The mixture was extracted with CH$_2$Cl$_2$ (30 mL) and washed with aqueous NaHCO$_3$ (3·10 mL). The extract was dried over Na$_2$SO$_4$ and CH$_2$Cl$_2$ was removed under reduced pressure to get crude solid. Flash chromatography (AcOEt/cyclohexane 1:1), yellow amorphous solid, yield 19%. ^{1}H NMR (400 MHz, CDCl$_3$): δ= 5.46 – 5.23 (m, 12H, alkenic H), 3.71 (bs, 1H, 3β-H), 3.66 (s, 3H, OMe), 3.57 – 3.48 (m, 1H, 7α-H), 2.89 – 2.73 (m, 10H, =CH-CH_2-CH=), 2.45 – 2.29 (m, 3H), 2.26 – 2.14 (m, 3H), 2.12 – 0.83 (m, 35H), 0.67 (s, 3H, 18-CH$_3$). ^{13}C NMR (101 MHz, CDCl$_3$): δ= 174.72, 171.43, 132.04, 129.23, 128.56, 128.26, 128.06, 127.84, 126.97, 71.34, 55.84, 54.99, 51.54, 49.14,

43.73, 42.76, 40.15, 39.28, 36.66, 35.43, 35.26, 34.63, 34.05, 31.06, 31.01, 28.60, 27.81, 26.87, 25.63, 25.54, 24.89, 23.55, 21.13, 20.56, 18.37, 14.29, 12.11. MS (ESI, ES+): calculated for $C_{47}H_{73}NO_4$ 716.10; found 479.50 $[2M+3H]^{3+}$, 503.37 $[2M+3Na]^{3+}$.
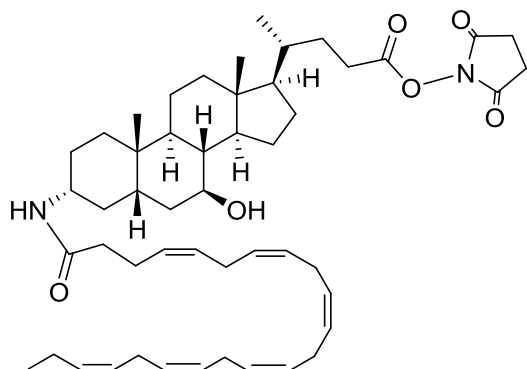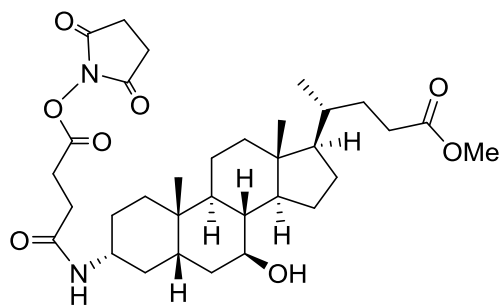
N-succinimidyl ester of DHA-UDCA **17**

To a solution of **16** (0.304 mmol) in MeOH (2.5 mL), an aqueous solution of LiOH 1.5M (3.652 mmol; 2.43 mL) was added. The mixture was stirred at 25°C for 18 h, then other LiOH 1.5M (3.652 mmol; 2.43 mL) was added. After others 18 h stirring, aqueous HCl 5% was added until pH=4. The mixture was extracted with AcOEt (3·8 mL). The extract was dried over Na_2SO_4 and the solvent was removed *in vacuo* to get crude solid. The crude product was used without further purification in the subsequent reaction of activation of the carboxylic acid with NHS (see "General procedure for the synthesis of active esters"). Flash chromatography (AcOEt/cyclohexane 2:1), yellow amorphous solid, yield 25% in 2 steps. ^{1}H NMR (400 MHz, CDCl$_3$): δ= 5.46 – 5.25 (m, 12H, alkenic H), 3.79 – 3.60 (m, 1H, 3β-H), 3.57 – 3.39 (m, 1H, 7α-H), 2.91 – 2.76 (m, 14H, 2CH$_2$ of NHS and =CH-C*H$_2$*-CH=), 2.70 – 2.60 (m, 3H), 2.57 – 2.47 (m, 3H), 2.39 (dd, J = 13.9, 7.4 Hz, 3H), 2.28 – 0.84 (m, 32H), 0.69 (s, 3H, 18-CH$_3$). ^{13}C NMR (101 MHz, CDCl$_3$): δ= 171.42, 169.22, 169.07, 132.04, 129.24, 128.57, 128.27, 128.08, 127.85, 126.99, 71.31, 55.82, 54.86, 49.14, 43.79, 43.71, 42.76, 40.15, 39.25, 36.71, 35.44, 35.07, 34.64, 34.05, 32.90, 32.21, 30.63, 28.58, 27.99, 27.81, 26.84, 25.59, 25.15, 24.74, 24.44, 24.11, 23.55, 23.49, 21.13, 20.56, 18.31, 14.31, 12.10. MS (ESI, ES+): calculated for $C_{50}H_{74}N_2O_6$ 799.15; found 799.93 $[M+H]^+$.

N-Succimidyl ester of Succ-UDC-OMe **14**

To a solution of **13** (0.678 mmol) in Py (4 mL), succinic anhydride (3.390 mmol) and catalytic DMAP were added. The mixture was stirred at 115°C for 18 h, then cooled at r.t., diluted with AcOEt (15 mL) and washed with aqueous HCl 5% (3·5 mL) and H$_2$O (5 mL). The extract was dried over Na_2SO_4 and the solvent was removed *in vacuo* to get crude solid. The crude product was used without further purification in the subsequent reaction of activation of the carboxylic acid with NHS (see "General procedure for the synthesis of active esters"). Flash chromatography (AcOEt/cyclohexane 5:1), white amorphous solid, yield 39% in 2 steps. ^{1}H-NMR (400 MHz, CDCl$_3$): δ= 5.66 (d, J = 7.9 Hz, 1H, NH), 3.77 – 3.67 (m, 1H, 7α-H), 3.65 (s, 3H, OMe), 3.57 – 3.41 (m, 1H, 3β-H), 2.96 (t, J = 7.1 Hz, 2H, succinic CH$_2$), 2.85 (s, 4H, 2CH$_2$ of NHS), 2.53 (t, J = 7.1 Hz, 2H, succinic CH$_2$), 2.40 – 2.29 (m, 1H, 23-C*H$_{2a}$*), 2.26 – 2.16 (m, 1H, 23-C*H$_{2b}$*), 2.07 – 0.99 (m, H), 0.94 (s, 3H, 19-CH$_3$), 0.91 (d, J = 6.4 Hz, 3H, 21-CH$_3$), 0.66 (s, 3H, 18-CH$_3$). ^{13}C NMR (101 MHz, CDCl$_3$): δ= 174.72, 169.11, 168.98, 168.16, 71.37, 55.84, 55.03, 51.51, 50.43, 49.47, 43.71, 42.74, 40.17, 39.26, 36.71, 35.38, 35.28, 34.30, 34.04, 32.88, 31.09, 31.04, 28.60, 27.52, 27.04, 26.88, 25.57, 25.18, 24.49, 23.52,

21.14, 18.37, 12.11. MS (ESI, ES+): calculated for $C_{33}H_{50}N_2O_8$ 602.77; found 625.27 [M+Na]$^+$, 1227.13 [2M+Na]$^+$. MS (ESI, ES-): found 601.47 [M-H]$^-$, 1238.93 [2M+Cl]$^-$.
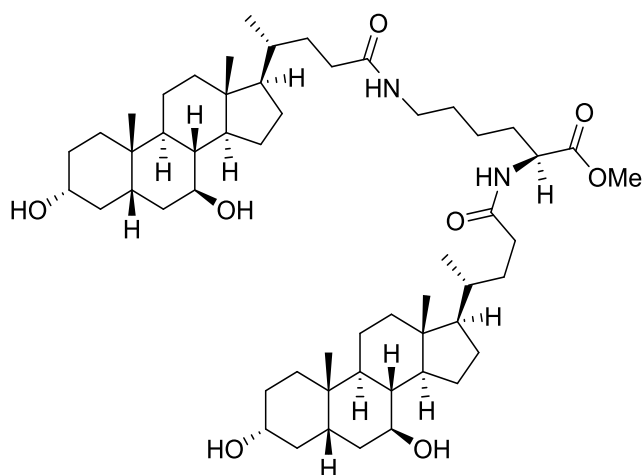
Bis-UDC-Lys-OMe 21

To a solution of **3** (2.247 mmol) in CH_2Cl_2 (20 mL), L-Lysine methyl ester dihydrochloride **18** (0.899 mmol) and DIPEA (4.494 mmol; 785 μL) were added at 0°C. The mixture was stirred at 25°C for 18 h, then 10 mL of aqueous HCl 5% were added. The so formed precipitate was recovered by filtration, redissolved in MeOH and concentrated *in vacuo* to get crude solid. Additional purification was not required. White amorphous solid, yield ≥95%. ^{1}H-NMR (400 MHz, DMSO-d_6): δ = 8.14 (d, J = 7.5 Hz, 1H, N*H*CH), 7.78 (t, J = 5.6 Hz, 1H, N*H*CH$_2$), 4.47 (t, J = 3.8 Hz, 2H, 2OH), 4.19 – 4.09 (m, 1H, NHC*H*), 3.88 (d, J = 6.8 Hz, 1H, OH), 3.60 (s, 3H, OMe), 3.33 – 3.22 (m, 4H, 3β-, 3'β-, 7α- and 7'α-H of UDC moieties), 3.16 (d, J = 4.7 Hz, 1H, OH), 3.05 – 2.94 (m, 2H, NHC*H*$_2$), 2.28 – 0.84 (m, 68H), 0.61 (d, J = 3.2 Hz, 6H, 18- and 18'-CH$_3$ of UDC moieties). ^{13}C-NMR (101 MHz, DMSO-d_6) δ = 172.77, 172.30, 69.58, 69.33, 55.76, 54.60, 51.69, 51.56, 48.47, 42.95, 42.88, 42.05, 38.61, 38.11, 37.77, 37.60, 37.15, 34.86, 34.71, 33.64, 32.36, 31.92, 31.58, 31.43, 30.28, 30.13, 28.56, 28.08, 26.61, 24.77, 23.21, 22.62, 20.74, 18.36, 11.91. MS (ESI, ES+): calculated for $C_{55}H_{92}N_2O_8$ 909.35; found 909.53 [M+H]$^+$, 1818.87 [2M+H]$^+$, 931.80 [M+Na]$^+$, 1840.93 [2M+Na]$^+$. MS (ESI, ES-): found 943.53 [M+Cl]$^-$.
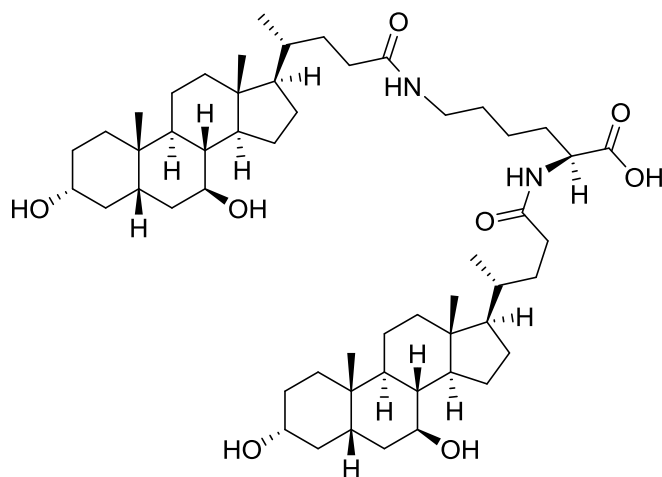
Bis-UDC-Lys

To a solution of **21** (0.649 mmol) in MeOH (7 mL), NH$_4$OH (7 mL) was added. The mixture was stirred at 60°C for 36 h, then NH$_3$ and MeOH was removed under reduced pressure. Aqueous HCl 5% (5 mL) was added and the mixture was filtered on Büchner. The solid was dried up in the heater at 80°C for 24 h. Additional purification was not required. White amorphous solid, yield 79%. ^{1}H-NMR (400 MHz, DMSO-d_6): δ = 7.94 (d, *J* = 7.4 Hz, 1H, N*H*CH), 7.75 (t, *J* = 8.2 Hz, 1H, N*H*CH$_2$), 7.27 (s, 1H), 6.93 (s, 1H), 4.50 (s, 2H, 2OH), 4.19 – 4.03 (m, 1H, NHC*H*), 3.90 (d, *J* = 6.3 Hz, 1H, OH), 3.41 – 3.21 (m, 4H, 3β-, 3'β-, 7α- and 7'α-H of UDC moieties), 3.16 (s, 1H, OH), 2.98 (bs, 2H, NHC*H*$_2$), 2.24 – 0.80 (m, 68H), 0.60 (d, *J* = 1.1 Hz, 6H, 18- and 18'-CH$_3$ of UDC moieties). ^{13}C-NMR (101 MHz, DMSO-d_6) δ = 173.96, 173.85, 172.56, 172.42, 69.66, 69.40, 55.82, 54.64, 52.07, 51.68, 43.02, 42.93, 42.10, 38.69, 38.05, 37.95, 37.64, 37.18, 35.06, 34.94, 34.75, 33.69, 32.42, 32.26, 32.13, 31.63, 31.54, 30.60, 30.16, 28.75, 28.66, 28.15, 26.66, 23.26, 22.76, 22.71, 20.80, 18.43, 11.97. MS (ESI, ES+):

calculated for $C_{54}H_{90}N_2O_8$ 895.32; found 895.47 $[M+H]^+$, 1789.80 $[2M+H]^+$, 917.67 $[M+Na]^+$, 1811.87 $[2M+Na]^+$. MS (ESI, ES-): found 893.67 $[M-H]^-$, 1788.73 $[2M-H]^-$.
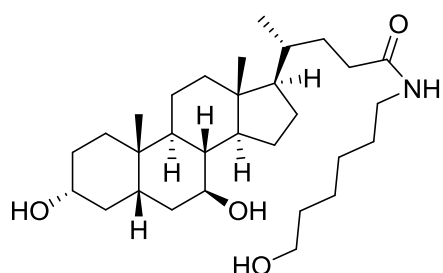
Compound 6

To a solution of **3** (1.421 mmol) in DMF (14 mL), 6-amino-1-hexanol (4.265 mmol) and DIPEA (2.844 mmol; 0.495 mL) were added. The mixture was stirred at 25°C for 24 h, then aqueous HCl at pH=5 (14 mL) was added. Th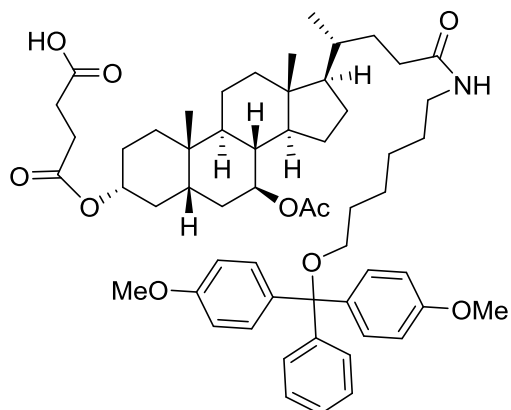e mixture was extracted by CH_2Cl_2 (3·10 mL). The extract was dried over Na_2SO_4 and the solvent was removed *in vacuo* to get crude solid. Additional purification was not required. White amorphous solid, yield 93%. ^{1}H-NMR (400 MHz, CDCl$_3$): δ = 5.52 (t, J = 5.6 Hz, 1H, NH), 3.66 – 3.60 (m, 2H, CH_2OH), 3.60 – 3.53 (m, 2H, 3β- and 7α-H), 3.27 – 3.18 (m, 2H, NHCH_2), 2.27 – 2.16 (m, 2H, 23-CH$_2$), 2.11 – 0.90 (m, 38H), 0.66 (s, 3H, 18-CH$_3$). ^{13}C-NMR (101 MHz, CDCl$_3$): δ = 173.57, 71.41, 71.34, 62.63, 55.72, 54.88, 43.73, 42.40, 40.13, 39.26, 39.16, 37.26, 36.82, 35.39, 34.90, 34.05, 33.65, 32.52, 31.85, 30.29, 29.66, 29.49, 28.69, 26.88, 26.48, 25.27, 23.37, 21.15, 18.49, 12.12. MS (ESI, ES+): calculated for $C_{30}H_{53}NO_4$ 491.76; found 492.60 $[M+H]^+$, 984.25 $[2M+H]^+$, 1475.71 $[3M+H]^+$.
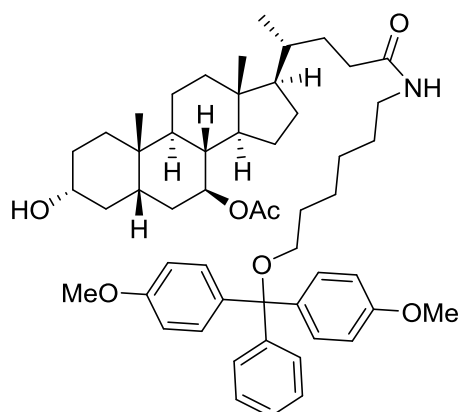
Compound 7

Compound **6** (0.665 mmol) was coevaporated with anhydrous CH$_3$CN (2·5 mL), then dissolved in Py (2 mL). 4,4'-dimethoxytrityl chloride (0.997 mmol) was added to the solution in three steps interspersed by 15 minutes. The mixture was stirred at 25°C for 18 h, then acetic anhydride (2.660 mmmol; 252 µL) and catalytic DMAP were added. The mixture was stirred at 25°C for 18 h, then MeOH (0.3 mL) was added. After 30 minutes stirring, the mixture was concentrated *in vacuo* and the residue was extracted with CH_2Cl_2 (10 mL) and washed with H_2O (2·5mL). The extract was dried over Na_2SO_4 and the solvent was removed *in vacuo*. The resulting solid was dissolved in a solution of KOH 0.089M in EtOH (0.783 mmol; 8.8 mL). The mixture was stirred at 25°C for 3 h, then a phosphate buffer ($Na_2HPO_4 \cdot H_2O$ 0,07M + KH$_2$PO$_4$ 0,07M in 4:1 ratio) was added until pH between 7 and 8. The mixture was extracted with AcOEt (4·10 mL). The extract was dried over Na_2SO_4 and the solvent was removed *in vacuo* to get crude solid. Flash chromatography (AcOEt/cyclohexane 3:1 + 3‰ Et$_3$N), white amorphous solid, yield 23% in 3 steps. ^{1}H-NMR (400 MHz, CD$_3$OD): δ= 7.40 (dt, J = 8.5, 1.8 Hz, 2H, 3''- and 5''-H of DMT), 7.31 – 7.23 (m, 6H, 2-, 2'-, 2''-, 6-, 6'- and 6''-H of DMT), 7.18 (ddd, J = 7.2, 3.8, 1.3 Hz, 1H, 4''-H of DMT), 6.86 – 6.81 (m, 4H, 3-, 3'-, 5- and 5'-H of DMT), 4.80 – 4.71 (m, 1H, 7α-H), 3.77 (s, 6H, 2 OMe of DMT), 3.54 – 3.43 (m, 1H, 3β-H), 3.21 – 3.00 (m, 4H, NHCH_2 and CH_2ODMT), 2.24 – 2.15 (m, 1H, 1H of 23-CH$_2$), 2.13 – 1.99 (m, 1H, 1H of 23-CH$_2$), 1.92 (s, 3H, CH$_3$ of acetyl), 1.88 – 0.98 (m, 32H), 0.96 (d, J = 6.5 Hz, 3H, 21-CH$_3$), 0.92 (s, 3H, 19-CH$_3$), 0.69 (s, 3H, 18-CH$_3$). ^{13}C-NMR (101 MHz, CD$_3$OD): δ= 176.55, 172.56, 159.94, 146.91, 137.85, 131.16, 129.28, 128.64, 127.61, 113.94, 87.01, 75.32, 71.92, 64.29, 56.70, 56.42, 55.67, 44.74, 43.64, 41.26, 40.73, 40.26, 37.77, 36.53, 35.88, 35.04, 34.10,

33.33, 31.08, 30.92, 30.39, 29.51, 27.86, 27.27, 26.93, 23.76, 22.30, 21.78, 18.92, 14.47, 12.58. MS (ESI, ES+): calculated for $C_{53}H_{73}NO_7$ 836.17; found 857.73 $[M+Na]^+$, 1696.47 $[2M+Na]^+$.

Compound 8

To a solution of **7** (0.154 mmol) in Py (2 mL), succinic anhydride (1.234 mmol) and DMAP (0.617 mmol) were added. The mixture was stirred at 70°C for 18 h, then concentrated *in vacuo*. The residue was extracted with AcOEt (5 mL) and washed successively with aqueous $NaHCO_3$ (5 mL) and a solution of citric acid 0.5% in H_2O (5 mL). The extract was dried over Na_2SO_4 and the solvent was removed *in vacuo*. The resulting solid did not require further purification. White amorphous solid, yield 80%. ^{1}H-NMR (400 MHz, CD₃OD): δ= 7.43 – 7.38 (m, 2H, 3''- and 5''-H of DMT), 7.32 – 7.24 (m, 6H, 2-, 2′-, 2''-, 6-, 6'- and 6''-H of DMT), 7.23 – 7.07 (m, 1H, 4''-H of DMT), 6.88 – 6.79 (m, 4H, 3-, 3′-, 5- and 5′-H of DMT), 4.79 – 4.59 (m, 2H, 3β- and 7α-H), 3.77 (s, 6H, 2 OMe of DMT), 3.22 – 3.00 (m, 4H, NHCH_2 and CH_2ODMT), 2.60 – 2.52 (m, 4H, succinic H), 2.25 – 2.15 (m, 1H, 1H of 23-CH₂), 2.14 – 1.99 (m, 1H, 1H of 23-CH₂), 1.92 (s, 3H, CH₃ of acetyl), 1.89 – 1.02 (m, 32H), 0.96 (d, *J* = 6.5 Hz, 3H, 21-CH₃), 0.94 (s, 3H, 19-CH₃), 0.69 (s, 3H, 18-CH₃). ^{13}C-NMR (101 MHz, CD₃OD): δ= 176.58, 173.76, 172.55, 159.96, 150.02, 146.92, 137.88, 131.17, 129.30, 129.21, 128.66, 127.62, 113.95, 87.03, 82.08, 75.34, 75.13, 64.31, 56.58, 56.34, 55.68, 44.74, 43.42, 41.23, 40.63, 40.27, 36.52, 35.46, 35.07, 34.10, 33.88, 33.31, 31.11, 30.40, 29.78, 29.50, 27.89, 27.38, 27.30, 26.92, 23.63, 22.31, 21.76, 18.93, 12.56. MS (ESI, ES+): calculated for $C_{57}H_{77}NO_{10}$ 936.24; found 934.67 $[M-H]^-$, 1870.67 $[2M-H]^-$.

7.4.3 Oligonucleotide solid phase synthesis

7.4.3.1 General procedure for oligonucleotide solid phase synthesis

To calculate the amount of functionalized support to be weighed, the used formula was:

$$synthesis\ scale\ (\mu mol) = amount\ of\ support\ (g) \cdot loading(\mu mol/g)$$

$$amount\ of\ support\ (g) = \frac{synthesis\ scale\ (\mu mol)}{loading\ (\mu mol/g)}$$

The weighted support was packaged in the column with CH₃CN, the column was placed in the oligosynthesizer and the synthesis was started. For detritylation step the used reagent was *DCA deblock* (solution of toluene/dichloroacetic acid 97:3 (V:V), Merck). For coupling steps, solutions 0.1M of phosphoramidites in anhydrous CH₃CN were prepared. Used phosphoramidites are: DMT-2′-*O*-methyl-rA(bz) phosphoramidite (bz = benzyl), DMT-2′-*O*-methyl-rG(ib) phosphoramidite (ib = isobutyryl), DMT-2′-*O*-methyl-rC(tac) phosphoramidite (tac = tert-butylphenoxyacetyl) and DMT-2′-*O*-methyl-rU phosphoramidite. For each coupling reaction, 5 eq of appropriate phosphoramidite were used, with recycling time of 10 minutes. The utilized activator was 5-(benzylthio)-1*H*-tetrazole (BTT), in solution 0.3M in CH₃CN. For thio-oxidation a column volume (CV) of solution of phenylacetyl disulfide (PADS) in CH₃CN/Py 1:1, with a flux of 0.4 mL/min. The reagents used for capping reaction corresponds to: CapA= solution of *N*-methylimidazole 20% (V:V) in CH₃CN; CapB= solution

1:1 of CapB1 (acetic anhydride 40% (V:V) in CH_3CN) and CapB2 (2,6-lutidine 60% (V:V) in CH_3CN). The total used volume of CapA + CapB for a capping step corresponded to 2CV, with a contact time of 0.3 min. The final treatment was carried out with a solution of diethylamine (DEA) 20% (V:V) in CH_3CN. During the whole synthesis process it was necessary to work in an inert nitrogen atmosphere, constantly maintaining the synthesis system at a pressure of 0.25-0.30 bar. Finished the synthesis, the support was vacuum-dried using the steel column as filter.

7.4.3.2 Release of the oligonucleotide from support

The compound was transferred in a flask and treated with CH_3CN (to promote swelling) and an aqueous solution of NH_4OH 33%. The flask was closed with screw cap and placed in a heater at 50°C for 24 h, then the mixture was cooled at r.t. and vacuum-filtered, washing the solid with a solution of H_2O/EtOH 1:1. The filtered liquid contained the crude product.

7.4.3.3 General method for analyzing and purifying oligonucleotides by RP-HPLC

Used eluents for analyses and purifications of crude oligonucleotides were CH_3CN and a buffer solution of triethylammonium acetate (TEAA) at pH=8 + 5% CH_3CN. The crude oligonucleotide was diluted in the TEAA buffer before being injected. The method had expected an increasing concentration gradient of CH_3CN in the TEAA buffer.

7.4.3.4 General method for ionic exchange of oligonucleotides by HPLC

The column resin was charged by flowing at 5 mL/min 5 CV of a 2M NaCl solution (the conductivity reached a plateau) and 5 CV of H_2O mQ; then the flux of H_2O was continued at 3 mL/min until conductivity was decreased below 0.017 μS/cm. Oligonucleotide was dissolved in H_2O mQ and injected in column, being careful not to exceed the load capacity of the column (it could be necessary to divide the sample and perform more injections). Once the oligonucleotide was collected, the column was washed with H_2O mQ (until chromatographic peak disappeared) and recharged with 2M NaCl solution to be ready for a new ion exchange cycle.

AON51

Sequence: 5′-UCAAGGAAGAUGGCAUUUCU-3'. Amorphous white solid, 75% yield. ^{31}P-NMR (122 MHz; D_2O): δ= 58.9-57.2 (m), 56.8-55.8 (m). MS (ESI, ES-): calculated 6977.1; found 1161.60 $[M-6H]^{6-}$, 1394.73 $[M-5H]^{5-}$, 1742.87 $[M-4H]^{4-}$.

m23-AON

Sequence: 5′-GGCCAAACCUCGGCUUACCU-3'. Amorphous white solid, 75% yield. ^{31}P-NMR (122 MHz; D_2O): δ= 56.0-53.8 (m). MS (ESI, ES-): calculated 6887.1; found 1376.83 $[M-5H]^{5-}$, 1720.97 $[M-4H]^{4-}$.

7.4.4 Coupling at the 5′-end of oligonucleotide

The coupling between the antisense oligonucleotide functionalized with amino linker at the 5′-end and the *N*-succimidyl esters **3**, **14**, **15**, **17**, **19** and **20** was performed in solid phase. The AON (7 theoretical μmol) was held anchored to the solid support within the reactor, while the appropriate *N*-succimidyl ester (10eq) was dissolved in 1 mL of DMSO with DIPEA (40eq) and loaded into a syringe. This syringe and another empty one were connected to the inlet and outlet of the reactor by means of suitable fittings. The solution was then made to flow through
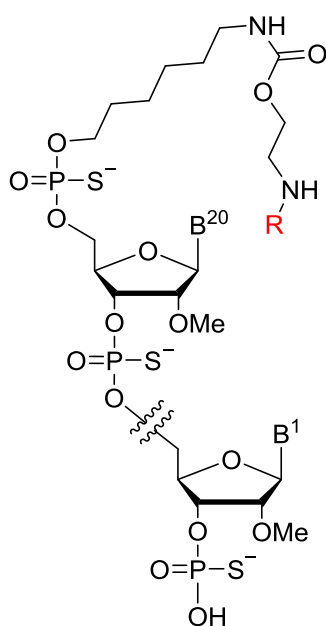
the reactor for about 4 hours, then the solid contained in the column was filtered under vacuum on sintered glass, washed with CH₃CN and transferred to a conical flask with screw cap to be released from the support following the indications in the paragraph 7.4.3.2. The two syringes-method was used also for the synthesis of conjugate with **Bis-UDCA-Lys**, adding HBTU (10eq) to the solution.

AON51-5′-UDCA

Amorphous white solid, 80% yield. ^{31}P-NMR (122 MHz; D$_2$O): δ= 58.2-56.1 (m). MS (ESI, ES-): calculated 7633.67; found 763.09 [M-10H]$^{10-}$, 847.47 [M-9H]$^{9-}$, 953.37 [M-8H]$^{8-}$.
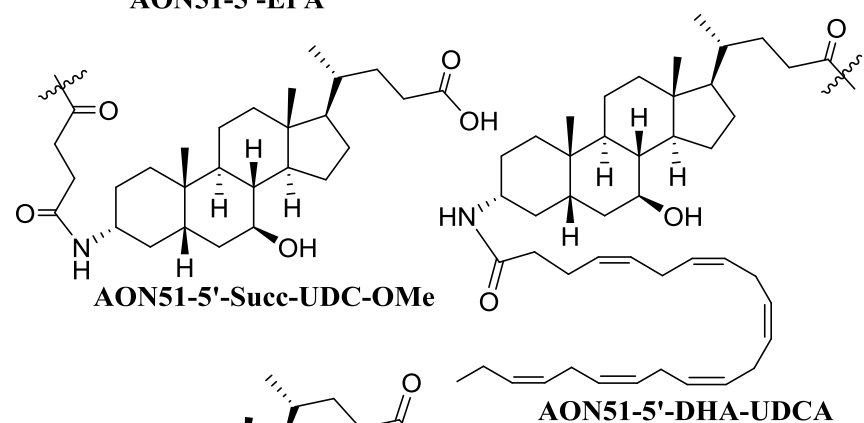
AON51-5′-HDCA

Amorphous white solid, 80% yield. MS (ESI, ES-): calculated 7633.67; found 1270.14 [M-6H]$^{6-}$, 1527.09 [M-5H]$^{5-}$, 1907.91 [M-4H]$^{4-}$.
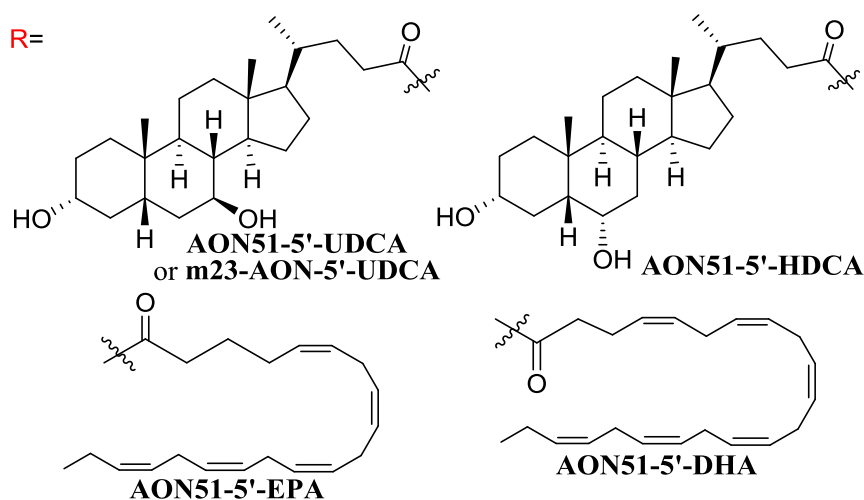
AON51-5′-EPA

Amorphous white solid, 70% yield. MS (ESI, ES-): calculated 7544.46; found 1257.00 [M-6H]$^{6-}$, 1507.88 [M-5H]$^{5-}$, 1885.98 [M-4H]$^{4-}$.

AON51-5′-DHA

Amorphous white solid, 70% yield. MS (ESI, ES-): calculated 7570.35; found 944.82 [M-8H]$^{8-}$, 1080.67 [M-7H]$^{7-}$, 1511.51 [M-5H]$^{5-}$, 1890.56 [M-4H]$^{4-}$.

AON51-5′-DHA-UDCA

Amorphous white solid, 40% yield. MS (ESI, ES-): calculated

R=

AON51-5'-UDCA
or **m23-AON-5'-UDCA**

AON51-5'-HDCA

AON51-5'-EPA

AON51-5'-DHA

AON51-5'-Succ-UDC-OMe

AON51-5'-DHA-UDCA

AON51-5'-Bis-UDCA-Lys
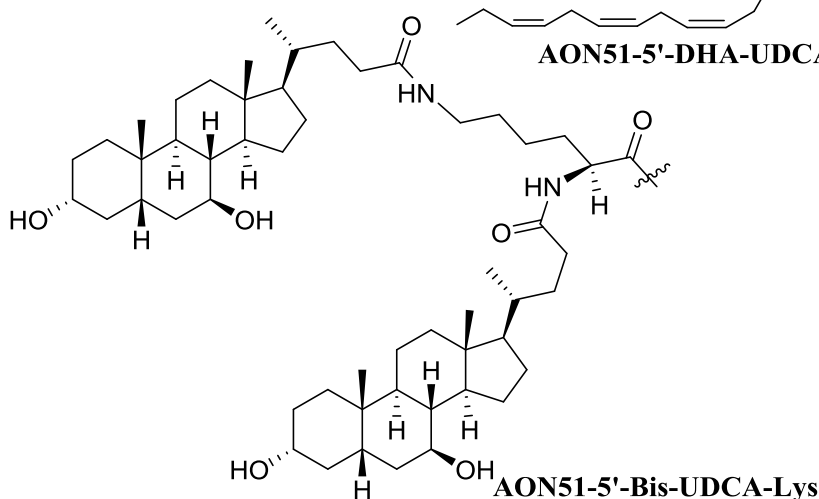
7943.93; found 990.35 [M-8H]$^{8-}$, 1136.66 [M-7H]$^{7-}$, 1322.40 [M-6H]$^{6-}$, 1588.28 [M-5H]$^{5-}$.

AON51-5′-Succ-UDC-OMe

It has been observed that the methyl ester is hydrolyzed in the treatment step with ammonium hydroxide. The final product therefore has in reality the carboxylic acid. Amorphous white solid, 40% yield. MS (ESI, ES-): calculated 7733.52; found 772.87 [M-10H]$^{10-}$, 858.57 [M-9H]$^{9-}$, 966.06 [M-8H]$^{8-}$, 1103.55 [M-7H]$^{7-}$, 1287.02 [M-6H]$^{6-}$, 1546.19 [M-5H]$^{5-}$, 1930.27 [M-4H]$^{4-}$.

AON51-5′-Bis-UDCA-Lys

Amorphous white solid, 20% yield. MS (ESI, ES-): calculated 8137.17; found 1015.89 [M-8H]$^{8-}$, 1162.03 [M-7H]$^{7-}$, 1355.43 [M-6H]$^{6-}$, 1626.64 [M-5H]$^{5-}$.

m23-AON-5′-UDCA

Amorphous white solid, 75% yield. MS (ESI, ES-): calculated 7543.66; found 685.47 [M-11H]$^{11-}$, 753.93 [M-10H]$^{10-}$, 836.36 [M-9H]$^{9-}$, 942.53 [M-8H]$^{8-}$, 1076.53 [M-7H]$^{7-}$, 1255.59 [M-6H]$^{6-}$, 1507.78 [M-5H]$^{5-}$, 1885.08 [M-4H]$^{4-}$.
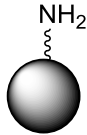
7.4.5 Coupling at the 3′-end of oligonucleotide

7.4.5.1 Coupling at the 3'-end of AON in solution
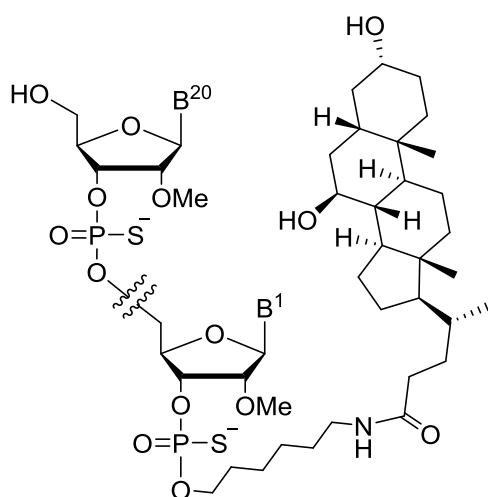
After solid phase synthesis of AON51 on *Custom Primer Support^{TM} C6 Amino*, detritylation and purification, coupling in solution was performed with the active ester of UDCA **3** (10eq), using H$_2$O/1,3-dioxolane 1:1 (*v/v*) as solvent in presence of DIPEA (40eq). Conjugate **AON51-3′-UDCA** was then obtained as amorphous white solid with a total yield of 45% after RP-HPLC purification. MS (ESI, ES-): calculated 7545.85; found 684.18 [M-11H]$^{11-}$, 752.81 [M-10H]$^{10-}$, 836.37 [M-9H]$^{9-}$, 940.95 [M-8H]$^{8-}$, 1075.46 [M-7H]$^{7-}$, 1255.03 [M-6H]$^{6-}$, 1506.21 [M-5H]$^{5-}$.

7.4.5.2 Oligonucleotide synthesis on derivatized support

Activation of Primer Support^{TM} 5G Amino

770 mg of *Primer Support^{TM} 5G Amino* (loading: 350 ÷ 400 μmol/g; 0.272 mmol) were placed in a 6.2 mL steel column. The steel column was inserted in the oligosynthesizer and CH$_3$CN (2 mL/min for 10 minutes) and a solution of DIPEA 2% in CH$_3$CN (1 mL/min for 30 minutes) were fluxed; it was thus possible to see that the conductivity curve increased. Then again CH$_3$CN (2 mL/min) was fluxed until conductivity decreased to 0 μS. The so activated support was dried in column under N$_2$.
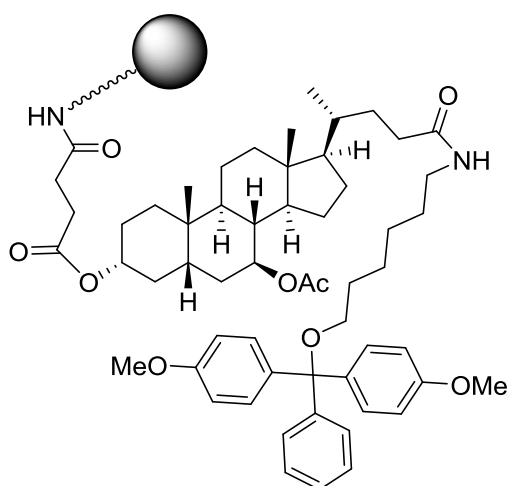
Functionalization of *Primer Support*[TM] *5G Amino* with compound **8**

Compound **8** (0.299 mmol) was coevaporated with anhydrous CH_3CN (3·17.5 mL) in a two-necked flask, connecting the rotavapor with flux of N_2. The previously activated *Primer Support*[TM] *5G Amino* (0.272 mmol) and CH_3CN (6 mL) were added. The mixture was stirred at r.t. for 5 minutes, then anhydrous DIPEA (0.598 mmol; 104 µL) was added. After 15 minutes stirring, a solution of HCTU (0.299 mmol) in anhydrous CH_3CN (2 mL) and anhydrous DMF (2 mL) were added *via cannula*. The mixture was stirred at 25°C for 18 h in a thermostatic oscillating stirrer. The mixture was next filtered, using the 6.2 mL steel column as filter and washing with CH_3CN (2·5 mL) and CH_2Cl_2 (2·5 mL) the solid. This last was dried in a vacuum gun at 40°C for 18 h. 725 mg of functionalized support was obtained.

Loading calculation

This procedure has been provided to us by the manufacturing company of *Primer Support*[TM] *5G Amino* (*GE Healthcare*) and must be followed scrupulously to obtain reliable results. To 101 mg of functionalized support, a solution of *p*-toluenesulphonic acid 0.1M was added until 500 mL of total volume. The mixture was placed in an ultrasonic bath for 10 minutes, then left to decant for 1 h. 1 mL of supernatant was analyzed by UV-Vis spectrophotometer at the wavelength of 411 nm, using as white the solution of *p*-toluenesulphonic acid 0.1M. Four absorbance measurements were performed (1.6476, 1.6537, 1.6312, 1.6445), whose calculated average was Ā=1.6442. This value was inserted in the following formula to calculate the loading:

$$c = \frac{\bar{A} \cdot V \cdot 35.5}{m} = \frac{1.6442 \cdot 500 \cdot 35.5}{101} = 289 \ \mu mol/g$$

This value of loading indicates how many micromoles of derivatized bile acid were tied per gram of support.

Capping of unreacted –NH₂

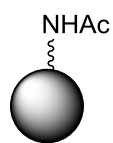725 mg of functionalized support were placed in a 250 mL flask with 5 mL of CapA solution and 5 mL of CapB solution. The mixture was stirred at 25°C for 18 h in a thermostatic oscillating stirrer. The mixture was next filtered, using the 6.2 mL steel column as filter and washing the solid with 5 mL of each of the following solvents in succession: CH_3CN, CH_2Cl_2, MeOH, H_2O and again CH_2Cl_2 (to promote swelling). The solid was dried in a vacuum gun at 40°C for 18 h. The procedure for loading calculation was then repeated with the following values: m=105 mg, V=500 mL, Ā=1.4272.

$$c = \frac{1.4272 \cdot 500 \cdot 35.5}{105} = 241 \ \mu mol/g$$

After two days, the absorbance measurements were performed again (Ā=1.4393) and the calculation repeated:

$$c = \frac{1.4393 \cdot 500 \cdot 35.5}{105} = 243 \ \mu mol/g$$

The loading value was substantially unchanged.
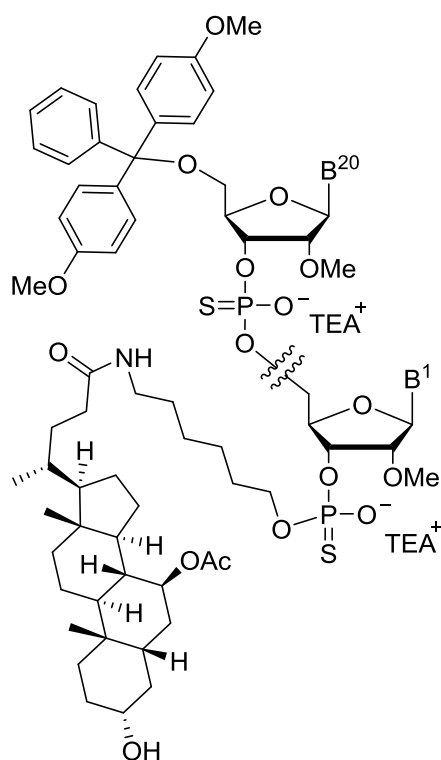
Oligonucleotide solid phase synthesis on functionalized support and purification

The oligonucleotide solid phase synthesis was performed on a 50 μmol scale. To calculate the amount of functionalized support to be weighed, the used formula was (taking $c = 242$ μmol/g):

$$amount\ of\ support = \frac{50}{242} = 0.207\ g$$

Therefore 207 mg of functionalized support were weighted in a 1.2 mL steel column and the synthesis was started (see "Oligonucleotide solid phase synthesis"). After solid phase synthesis, treatment with 1.2 mL of CH_3CN to promote swelling and 10 mL of NH_4OH 33% at 60°C for 20 h to release the oligonucleotide from support, the mixture was cooled at r.t. and vacuum-filtered, washing the solid with 70 mL of solution of H_2O/EtOH 1:1. 5 μL of solution of crude product were taken for UV-Vis analysis at λ=260 nm. Physiological solution was used as white. From optical density values (0.291, 0.329 and 0.268) it was calculated the average optical density $\overline{OD} = 0.296$, and consequently the OD per mL:

$$OD\ per\ mL = 0.296 \cdot \frac{1005}{5} = 59.5$$

From this value it was possible to calculate concentration and amount of crude product:

$$c = \frac{OD\ per\ mL}{\varepsilon \cdot l} = \frac{59.5}{204.5 \cdot 1} = 0.291\ mM$$
$$c \cdot V_{tot} = 0.291 \cdot 78 = 22.69\ \mu mol$$

The crude yield was 45%. The mixture was diluted with a buffer solution of triethylammonium acetate (TEAA) at pH=8 + 5% CH_3CN until 100 mL of total volume. The crude product was purified by RP-HPLC (see "General method for analyzing and purifying oligonucleotides by RP-HPLC"). From obtained solution of purified oligonucleotide (145 mL), 10 μL were taken to spectrophotometric analysis. From the average of the performed measurements ($\overline{OD} = 0.205$), the amount of product was derived from the following calculations:

$$OD\ per\ mL = 0.205 \cdot \frac{1010}{10} = 20.7$$

$$c = \frac{OD\ per\ mL}{\varepsilon \cdot l} = \frac{20.7}{204.5 \cdot 1} = 0.101\ mM$$
$$c \cdot V_{tot} = 0.101 \cdot 145 = 14.65\ \mu mol$$

The total yield was 30%. The mixture was then transferred in a flask and concentrated *in vacuo*, being careful not to exceed 40°C.

Detritylation and ionic exchange

The purified oligonucleotide was solubilized in 40 mL of H_2O mQ and leaved at r.t. for 18 h, then centrifuged to separate the supernatant from precipitate (4,4'-dimethoxitrityl alcohol, DMT-OH). After decantation, the supernatant was concentrated *in vacuo* to 10 mL and
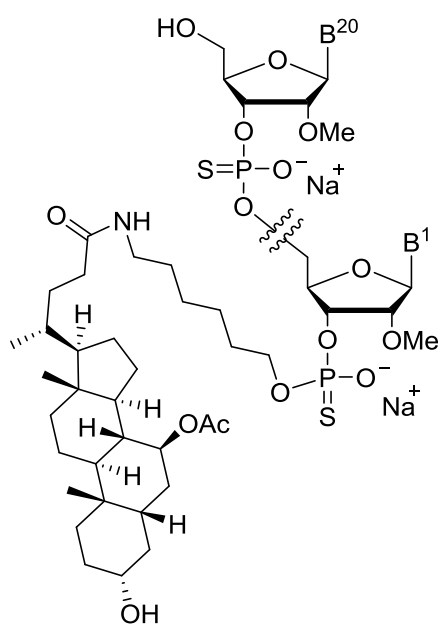
lyophilized. To remove the TEAA residue, the compound was redissolved in H_2O mQ and lyophilized again twice. It was then performed the ionic exchange, to substitute TEA^+ with Na^+ as counterion (see "General method for ionic exchange of oligonucleotides by HPLC"). Finally, the product was lyophilized to obtain 98 mg of white amorphous solid. MS (ESI, ES-): calculated 7587.85; found 1518.04 $[M-5H]^{5-}$, 1895.87 $[M-4H]^{4-}$.

7.4.6 Synthesis of AON51-3′,5′-UDCA

AON51 was synthesized in solid phase on *Custom Primer Support*^{TM} *C6 Amino* in DMT-off mode, inserting the *ssH-Linker*^{TM} after the last nucleotide of the sequence. The coupling between AON51 functionalized with amino linker at the 5′-end and the active ester of UDCA **3** was then performed in solid phase (paragraph 7.4.4). Subsequently, the oligonucleotide was released from the solid support (paragraph 7.4.3.2) and purified by RP-HPLC (paragraph 7.4.3.3). At this point coupling at the 3′-end in solution was performed with the active ester of UDCA **3**, using the same conditions seen at paragraph 7.4.5.1. Conjugate **AON51-3′,5′-UDCA** was then obtained as amorphous white solid with a total yield of 58% after further chromatographic purification.

MS (ESI, ES-): calculated 8203.30; found 910.49 $[M-9H]^{9-}$, 1024.37 $[M-8H]^{8-}$, 1171.22 $[M-7H]^{7-}$, 1368.65 $[M-6H]^{6-}$, 1639.61 $[M-5H]^{5-}$.
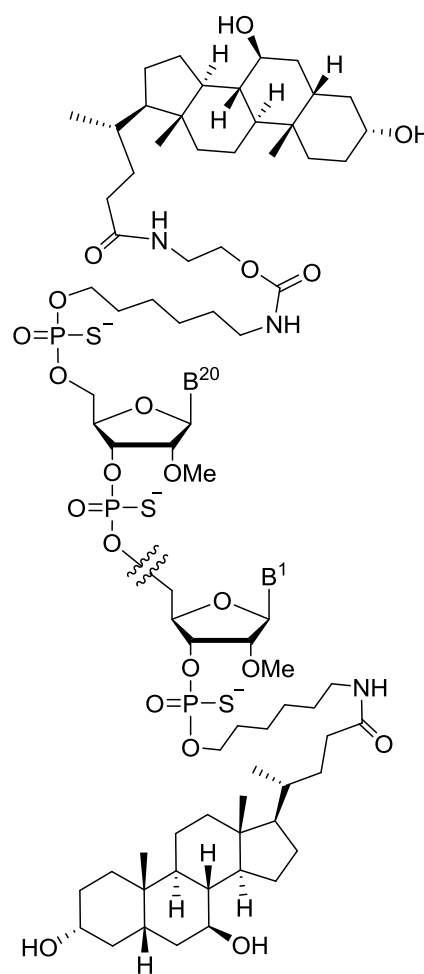
7.4.7 *In vitro* biological assays

To perform *in vitro* tests, myotubes obtained by differentiating a cell line of immortalized human myoblasts derived from a patient with an exon 52 deletion of the dystrophin gene (DMD), which breaks the reading frame, were treated with a solution of antisense oligonucleotide 5′-UCAAGGAAGAUGGCAUUUCU-3' complementary to exon 51 of the DMD gene, conjugated at the 5′- and/or at the 3′-end with a lipophilic molecule or non-conjugated.

The first experiment was performed with 2 μl of a 100 μM solution of **AON51**, **AON51-5′-UDCA** or **AON51-3′-UDCA**, using 4 μl of *JetPEI*® transfecting reagent in 24-well plates and cells collected 48 hours or five days later for RNA extraction or for immunofluorescence analysis respectively. The RNA was quantified and retrotranscribed for subsequent amplification with complementary primers exon 50 (Ex50F) and 54 (Ex54R) for 28 cycles of RT-PCR. One microliter of the RT-PCR was subsequently analyzed by *Bioanalyser 2100 Agilent* to quantify the products related to the transcript of the exon 52 debride (control) and

the transcript resulting from the skipping of the exon 51 induced by the antisense oligonucleotides. For immunofluorescence analysis, myoblasts were treated for five days with 4 μl of a 100 μM solution for each of the antisense oligonucleotides in appropriate plates, before being fixed and labeled with NCL-DYS2 and DAPI antibody.

The second experiment was performed with 4 μl of a 100 μM solution of **AON51**, **AON51-5′-UDCA** or **AON51-3′,5′-UDCA**, using 8 μl of *JetPEI*® transfecting reagent in 24-well plates and cells collected 48 hours later for RNA extraction.

The third experiment was performed with 2 μl of a 100 μM solution of **AON51**, **AON51-5′-UDCA**, **AON51-5′-EPA, AON51-5′-DHA** or **AON51-5′-DHA-UDCA**, using 4 μl of *TurboFect*® transfecting reagent and cells collected 72 hours later for RNA extraction.

For the last experiment (gymnotic delivery), a concentrated solution for each AON (**AON51**, **AON51-5′-UDCA**, **AON51-5′-HDCA**, **AON51-5′-Succ-UDC-OMe** and **AON51-5′-Bis-UDCA-Lys**) has been added in the differentiation medium without any transfection reagent to make a final concentration of 50 uM in each well. After three days, cells were collected for the RNA extraction.

7.4.8 *In vivo* tests

The oligonucleotides **m23-AON** and **m23-AON-5′-UDCA** were injected intraperitoneally at a dose of 200 mg/kg with a one-week administration regimen for 12 weeks in 2-month C57BL/10ScSn-Dmd^{mdx}/J male mice. Mice injected with PBS (phosphate buffer saline) were used as controls. One week after the last treatment, the mice were sacrificed to collect samples of the heart, diaphragm, gastrocnemius and tibialis anterior. Muscles had been divided and fragmented to conduct transcript analysis and protein quantification. Exon skipping was evaluated by RT-PCR performed with primers complementary to exons 20 and 26 of the murine dystrophin transcript able to amplify the 1098 base pair fragment corresponding to the complete transcript and 885 base pairs corresponding to the transcript without exon 23.

Muscles collected for semiquantitative analysis of dystrophin by western blot were homogenized in RIPA buffers and protease inhibitors to be subsequently quantified. Thirty micrograms of protein were mixed with *NuPage LDS buffer 4x* added with 50 mM DTT, heated for two minutes at 85°C before being loaded onto a *Novex 3% -8% Tris-Acetate* gel and migrated for 70 minutes at 150V. The proteins were then transferred to PVDF membranes using the 70V iBLOT system for 7 minutes and hybridized with antibodies to the carboxyterminal region of dystrophin (NCL-DYS2) and against alpha-actinin as a loading control.

7.5 Bibliography

1. Bennett, C. F.; Swayze, E. E. *Annu. Rev. Pharmacol. Toxicol.* **2010,** *50,* 259-293.

2. Spitali, P.; Rimessi, P.; Fabris, M.; Perrone, D.; Falzarano, S.; Bovolenta, M.; Trabanelli, C.; Mari, L.; Bassi, E.; Tuffery, S.; Gualandi, F.; Maraldi, N. M.; Sabatelli-Giraud, P.; Medici, A.; Merlini, L.; Ferlini, A. *Hum. Mutat.* **2009,** *30,* 1527-1534.

3. Segen, J. C. *Dizionario di medicina moderna;* McGraw-Hill, **2006**; p 225.

4. Genetics Home Reference - NIH U.S. National Library of Medicine. https://ghr.nlm.nih.gov/condition/duchenne-and-becker-muscular-dystrophy (accessed November 15, **2018**).

5. Genetics Home Reference - NIH U.S. National Library of Medicine.

https://ghr.nlm.nih.gov/art/large/dystrophic-arm-muscle.jpeg (accessed November 13, **2018**).

6. DISTROFIA MUSCOLARE DI DUCHENNE E BECKER: IL DOTT. PREVITALI ALLA XV CONFERENZA INTERNAZIONALE - UniSCIENZA&RICERCA. http://scienzaericerca.unisr.it/wp-content/uploads/2017/03/distrofina-e-muscolo.png (accessed November 13, **2018**).

7. Lu, Q.-L.; Cirak, S.; Partridge, T. *Mol. Ther. Nucleic Acids* **2014,** *3,* e152.

8. Gray, G. D.; Basu, S.; Wickstrom, E. *Biochem. Pharmacol.* **1997,** *15* (53), 1465-1476.

9. Rimessi, P.; Sabatelli, P.; Fabris, M.; Braghetta, P.; Bassi, E.; Spitali, P.; Vattemi, G.; Tomelleri, G.; Mari, L.; Perrone, D.; Medici, A.; Neri, M.; Bovolenta, M.; Martoni, E.; Maraldi, N. M.; Gualandi, F.; Merlini, L.; Ballestri, M.; Tondelli, L.; Sparnacci, K.; Bonaldo, P.; Caputo, A.; Laus, M.; Ferlini, A. *Mol. Ther.* **2009,** *17,* 820-827.

10. Shokrzadeh, N.; Winkler, A. M.; Mehrdad, D.; Winkler, J. *Bioorg. Med. Chem. Lett.* **2014,** *24,* 5758-5761.

11. Takaharu, M.; et al. *Biochem and Biophys. Res.* **2002,** *329,* 386-390.

12. Hofmann, A. F. *Ital. J. Gastroenterol* **1995,** *27,* 106-113.

13. *International Journal of Urology* **2002,** *9,* 42-46.

14. Dalpiaz, A.; Paganetto, G.; Pavan, B.; Fogagnolo, M.; Medici, A.; Beggiato, S.; Perrone, D. Zidovudine and Ursodeoxycholic Acid Conjugation: Design of a New Prodrug Potentially Able To Bypass the Active Efflux Transport Systems of the Central Nervous System. *Molecular Pharmaceutics* **2012,** *9* (4), 957-968.

15. Dalpiaz, A.; Ferraro, L.; Perrone, D.; Leo, E.; Iannuccelli, V.; Pavan, B.; Paganetto, G.; Beggiato, S.; Scalia, S. Brain Uptake of a Zidovudine Prodrug after Nasal Administration of Solid Lipid Microparticles. *Molecular Pharmaceutics* **2014,** *11* (5), 1550-1561.

16. Dalpiaz, A.; Contado, C.; Mari, L.; Perrone, D.; Pavan, B.; Paganetto, G.; Hanuskova, M.; Vighi, E.; Leo, E. Development and characterization of PLGA nanoparticles as delivery systems of a prodrug of zidovudine obtained by its conjugation with ursodeoxycholic acid. *Drug Delivery* **2014,** *21* (3), 221-232.

17. Ackerman, H. D.; Gerhard, G. S. Bile Acids in Neurodegenerative Disorders. *Frontiers in Aging Neuroscience* **2016,** *8,* 263.

18. Perrone, D.; Bortolini, O.; Fogagnolo, M.; Marchesi, E.; Mari, L.; Massarenti, C.; Navacchia, M. L.; Sforza, F.; Varani, K.; Capobianco, M. L. Synthesis and in vitro cytotoxicity of deoxyadenosine-bile acid conjugates linked with 1,2,3-triazole. *New J. Chem.* **2013,** No. 37, 3559-3567.

19. Singh, M.; Singh, A.; Kundu, S.; Bansal, S.; Bajaj, A. Deciphering the role of charge, hydration and hydrophobicity for cytotoxic activities and membrane interactions of bile acid based facial amphiphiles. *Biochimica et Biophysica Acta* **2013,** *1828,* 1926-1937.

20. Stetsenko, D. A.; Gait, M. J. A convenient solid-phase method for synthesis of 3′-conjugates of oligonucleotides. *Bioconjugate Chem.* **2001,** *12,* 576-586.

21. Singh, Y.; Spinelli, N.; Defrancq, E. *Curr. Org. Chem.* **2008,** *12,* 263-290.

22. Hazra, B. G.; Pore, V. S.; Dey, S. K.; Datta, S.; Darokar, M. P.; Saikia, D. *Bioorg. Med. Chem. Lett.* **2004,** *14,* 773-777.

23. Zhang, X.; Zeng, Y.; Yu, T.; Chen, J. *Langmuir* **2014,** *30,* 718-726.

24. Østergaard, M. E.; Yu, J.; Kinberger, G. A.; Wan, W. B.; Migawa, M. T.; Vasquez, G.; Schmidt, K.; Gaus, H. J.; Murray, H. M.; Low, A.; Swayze, E. E.; Prakash, T. P.; Seth, P. P. Efficient synthesis and biologicl evaluation of 5′-galNAc conjugate antisense oligonucleotides. *Bioconjugate Chemistry* **2015,** *26* (8), 1451-1455.

25. Stastna, E. Synthesis of C3, C5, and C7 pregnane derivatives and their effect on NMDA receptor responses in cultured rat hippocampal neurons. *Steroids* **2009,** *74* (2), 256-263.

26. Kohout, L. On steroids. CCCXX. Solvolytic reactions of epimeric 3-methanesulfonyloxy-7-benzoyloxy-5β,6β-cyclopropanocholestanes. *Collection of Czechoslovak Chemical Communications* **1986,** *51* (2), 436-446.

27. *J. Am. Chem. Soc.* **2014,** *136,* 16958-16961.

28. Marder, O.; Shvo, Y.; Albericio, F. *Chimica Oggi* **2002,** *20*.

29. Marder, O.; Albericio, F. *Chimica Oggi* **2003,** *21*.

30. Massarenti, C.; Bortolini, O.; Fantin, G.; Cristofaro, D.; Ragno, D.; Perrone, D.; Marchesi, E.; Toniolo, G.; Massi, A. Fluorous-tag assisted synthesis of bile acid–bisphosphonate conjugates via orthogonal click reactions: An access to potential anti-resorption bone drugs. *Org. Biomol. Chem.* **2017,** *15,* 4907–4920.

31. Singh, Y.; Spinelli, N.; Defrancq, E. *Curr. Org. Chem.* **2008,** *12,* 263-290.

Third part: Nucleoside bioconjugates as inhibitors of glycosyltransferases

Glycosyltransferases (GTs) are enzymes that carry out the protein-glycosylation [1], by transferring a glycosidic unit from a donor substrate (glycosyl nucleotides in most cases) to an acceptor substrate (Scheme III). The biosynthesis of glycoprotein is involved in a variety of biological functions, such as nutrient signaling, regulation of gene transcription, cellular response to stress and cell division. Given their importance, their dysregulation is implicated in disorders such as diabetes, in some neurodegenerative diseases (such as Alzheimer's) and in cancer: inhibiting these enzymes therefore provides a therapeutic target against these diseases [2].
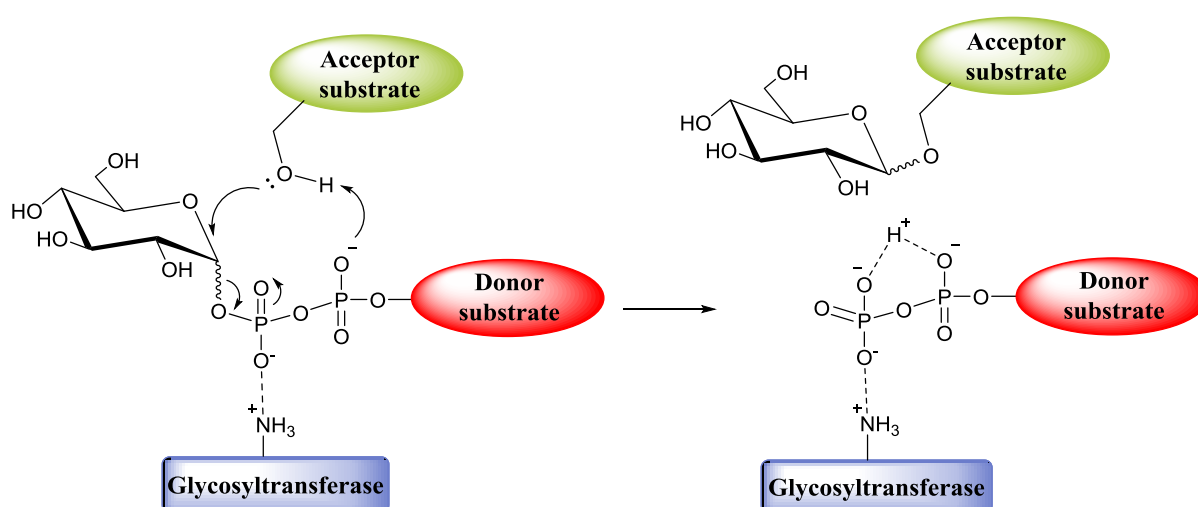
Scheme III *Example of transfer of a glycosidic unit from a donor substrate to an acceptor substrate, catalyzed by a generic GT.*

8 UDP-GlcNAc analogues as inhibitors of *O*-GlcNAc transferase (OGT): spectroscopic, computational, and biological studies

8.1 Introduction

Protein glycosylation, carried out by glycosyltransferases (GTs) [3,4,5], is a key post-translational modification in mammals [1]. The carbohydrate being transferred to the protein is activated by binding to a nucleotide diphosphate that interacts with the donor active site of the enzyme [6,7,8]. In the case of protein glycosylation, only nine nucleotide sugar donors (known as Leloir donors) containing uridine or guanine (with the exception of CMP-sialic acid, the substrate for sialyltransferases) are substrates for glycosyltransferases [9]. Numerous pathological states, including cancer [10,11] and neurological disorders [12,13], are closely related to the biosynthesis of glycoproteins. Inhibiting this essential biological process provides an obvious therapeutic target against these diseases [2]. Currently, the design and synthesis of inhibitors of glycosyltransferases is a hot topic of great interest [14,15,16,17,18]. As illustrated in Figure 8.1, several strategies are used to inhibit protein glycosylation [19]. The most widely used approach involves modifying the carbohydrate unit (**1**) [20,21,22,23]. Modifications at the ribose ring (**2**) and the nucleobase (**3**) of the nucleotide have also been reported, but to a lesser extent [24].
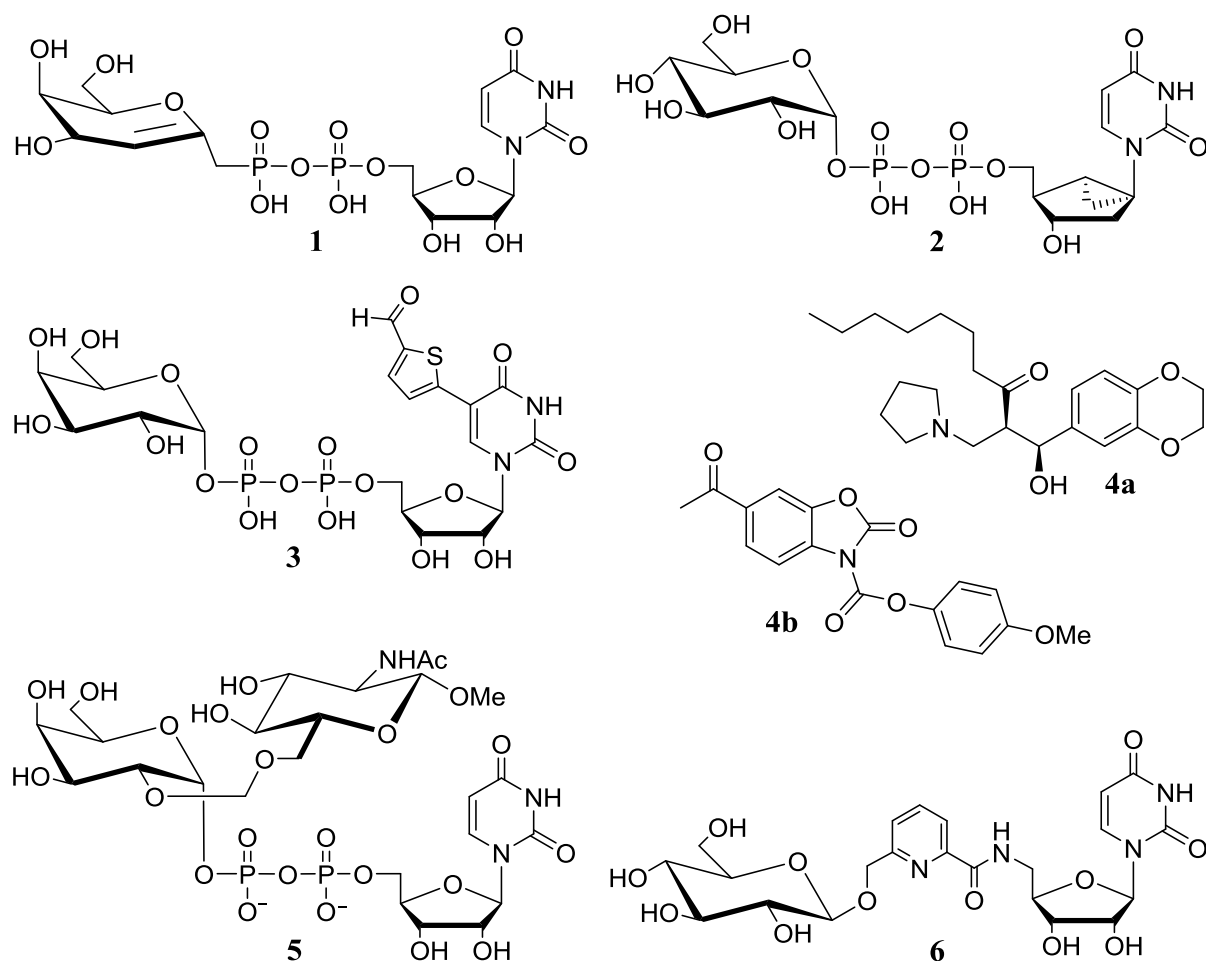

Figure 8.1 Strategies towards the design of GT inhibitors.

More recently, new chemotypes such as **4a-b**, far from GT substrate-like compounds [25,26], and bisubstrate inhibitors such as **5** [27,28], consisting of two conjugated fragments interacting with both donor and acceptor sites of the enzyme, have emerged as promising alternatives (Figure 8.1). Modification at the pyrophosphate linkage is also possible. Indeed, it is known that pyrophosphate derivatives have poor membrane permeability and hence low bioavailability. Consequently, conjugated sugar-nucleotides linked by neutral surrogates of the pyrophosphate unit such as **6** have also been designed as GT inhibitors [26,29,30] (Figure 8.1).

In this context, a recent study was conducted on less polar UDP-Gal and UDP-GalNAc analogues, in which the β-phosphate has been replaced by an alkyl chain [31] (Figure 8.2). As a model of GT, GalNAc-T2 was chosen; this GT is widely distributed in human tissues that plays an important role in health and disease [32,33,34]. The study showed the analogues to be weaker binders of the enzyme than the natural substrate UDP-GalNAc, and from the crystal structure it was inferred that the β-phosphate was required for binding to the metal ion.
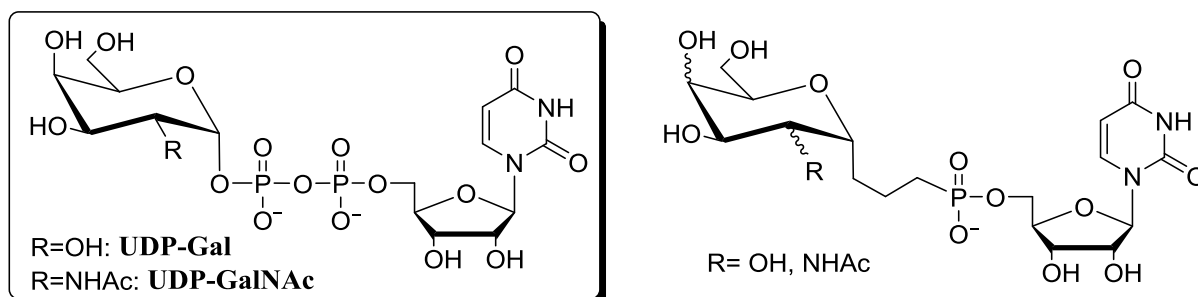

Figure 8.2 *Structure of natural substrates UDP-Gal and UDP-GalNAc and its analogues in which the β-phosphate has been replaced by an alkyl chain* [31].

Herein, we report the use of *O*-GlcNAc transferase (OGT) as a model system, an enzyme that plays a crucial role in a variety of biological functions within the human body [35,36,37,38] and does not require the presence of a metal [39]. The development of new inhibitors of OGT has been pursued for several years [40,41]. However, although promising results have recently been reported [42,43,44], finding specific and cell-permeable inhibitors for use in animal models still remains a challenge [45].

In the present study, we have examined a total of 14 candidates (Figure 8.3). UDP-Glc and UDP-GlcNAc analogues **7** and **8**, respectively, were chosen by analogy with the previous study on GalNAc-T2 [31]. In some previously reported inhibitors, the sugar moiety has been replaced by a different structure, since it is known that the sugar unit has some conformational freedom [46,47,24,48]. On the other hand, it has recently been reported that the precise conformation of the sugar unit could play a significant role in catalysis [49]. To study the effects of sugar substitution, we have considered compounds **9–16**, in which the sugar moiety has been replaced by a pyrrolidine unit bearing functional groups suitably predisposed for H-bonding (as in **9**, **11**, **13**, and **15**) or hydrophobic interactions (as in **10**, **12**, **14**, and **16**). Finally, analogues **17–20** containing a triazole ring in place of the β-phosphate unit have also been prepared (Figure 8.3).

Replacement of both phosphate groups by a triazole moiety has been reported previously, but the heterocyclic ring was placed close to the nucleoside and no biological studies were carried out [50]. More recently, Vidal and co-workers have reported the synthesis of neutral analogues

containing glycosyltriazoles, in which both phosphate groups were replaced, and suggested that the presence of the triazole ring results in additional interactions when a divalent metal ion is required for catalysis [29,42]. Our approach also employs glycosyltriazoles; however, only the β-phosphate has been replaced in order to determine the exact equivalence between the β-phosphate and the triazole ring in enzymes that do not require a metal ion for catalysis. Chemical syntheses, biological evaluation, and computational studies of the prepared compounds **7–20** are discussed herein.
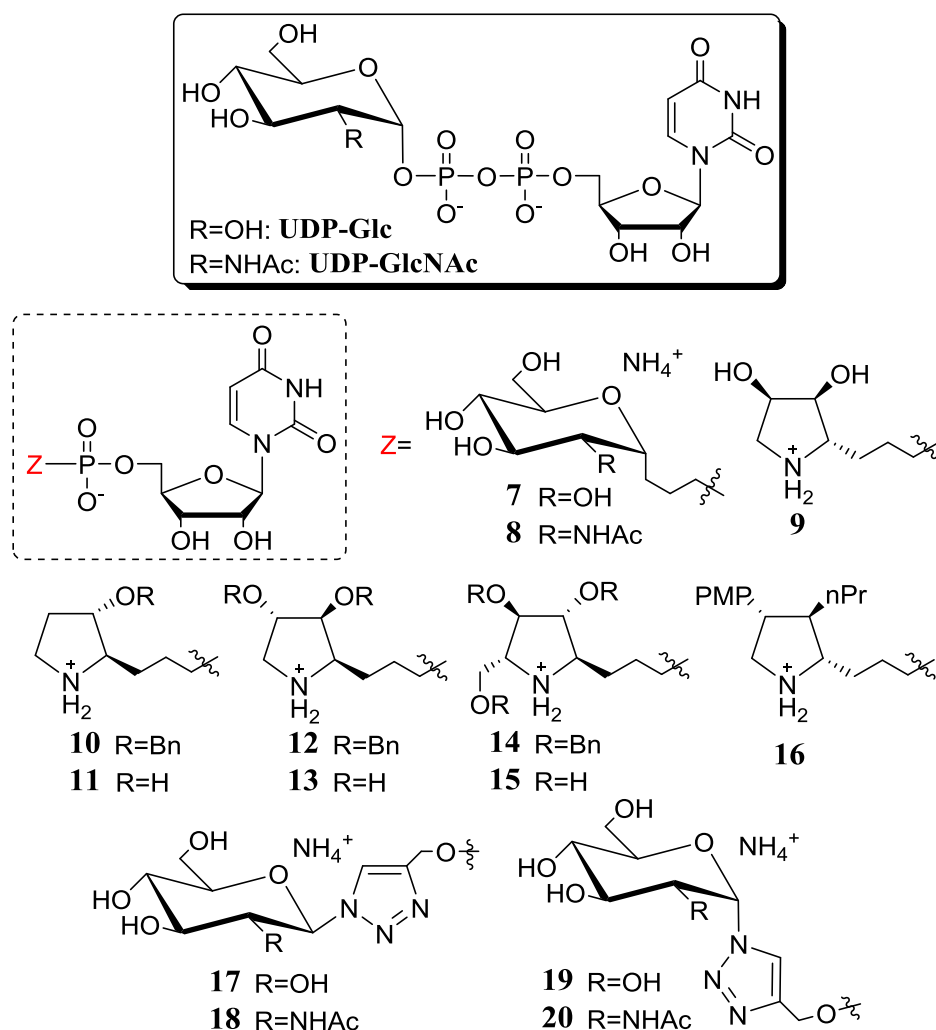

Figure 8.3 *Structure of natural substrates UDP-Glc and UDP-GlcNAc and its less polar analogues synthetized.*

8.2 Results and discussion

### 8.2.1 Synthesis of UDP-GlcNAc analogues

Compounds **7** and **8** were synthetized following the procedure reported in literature [31] (Scheme 8.1). The starting allyl sugars **21a** and **21b** were prepared respectively from D-glucose [51,52] and D-glucosamine [53,54]. Protected diacetyl uridine **23** was prepared from commercially available uridine *via* the trityl derivative [55,56] (see Scheme 8.5).
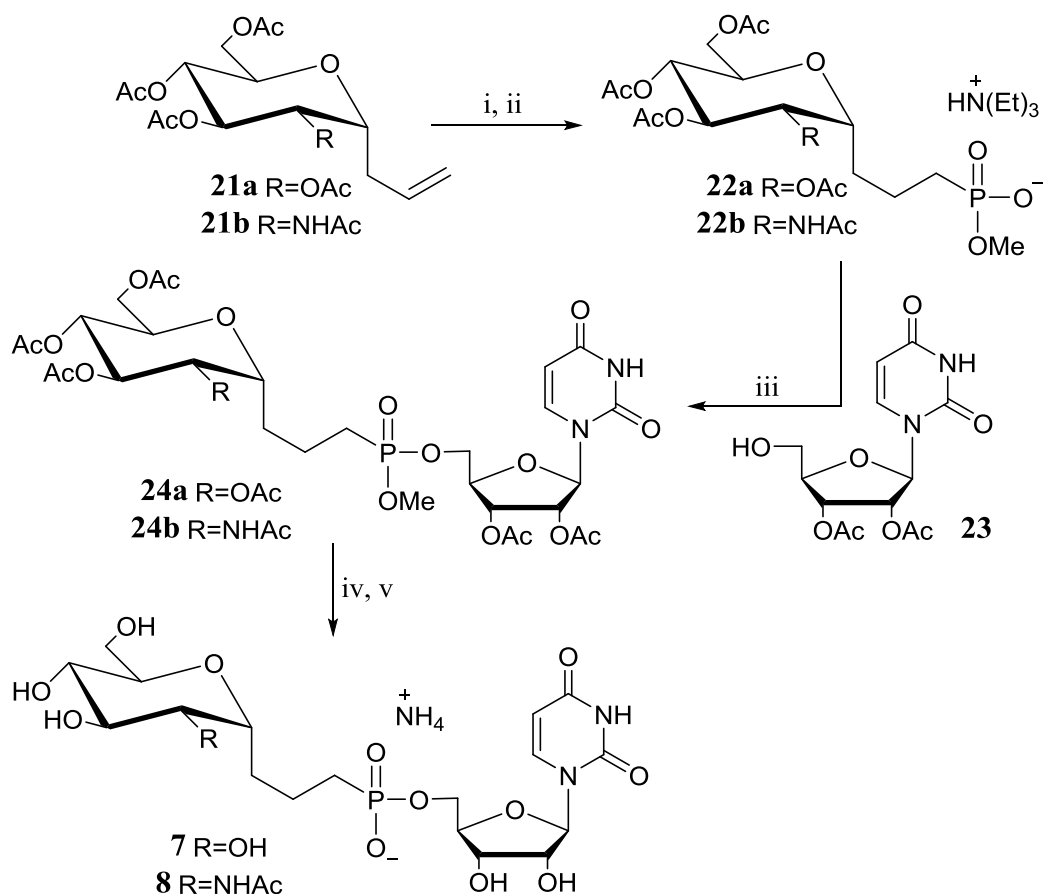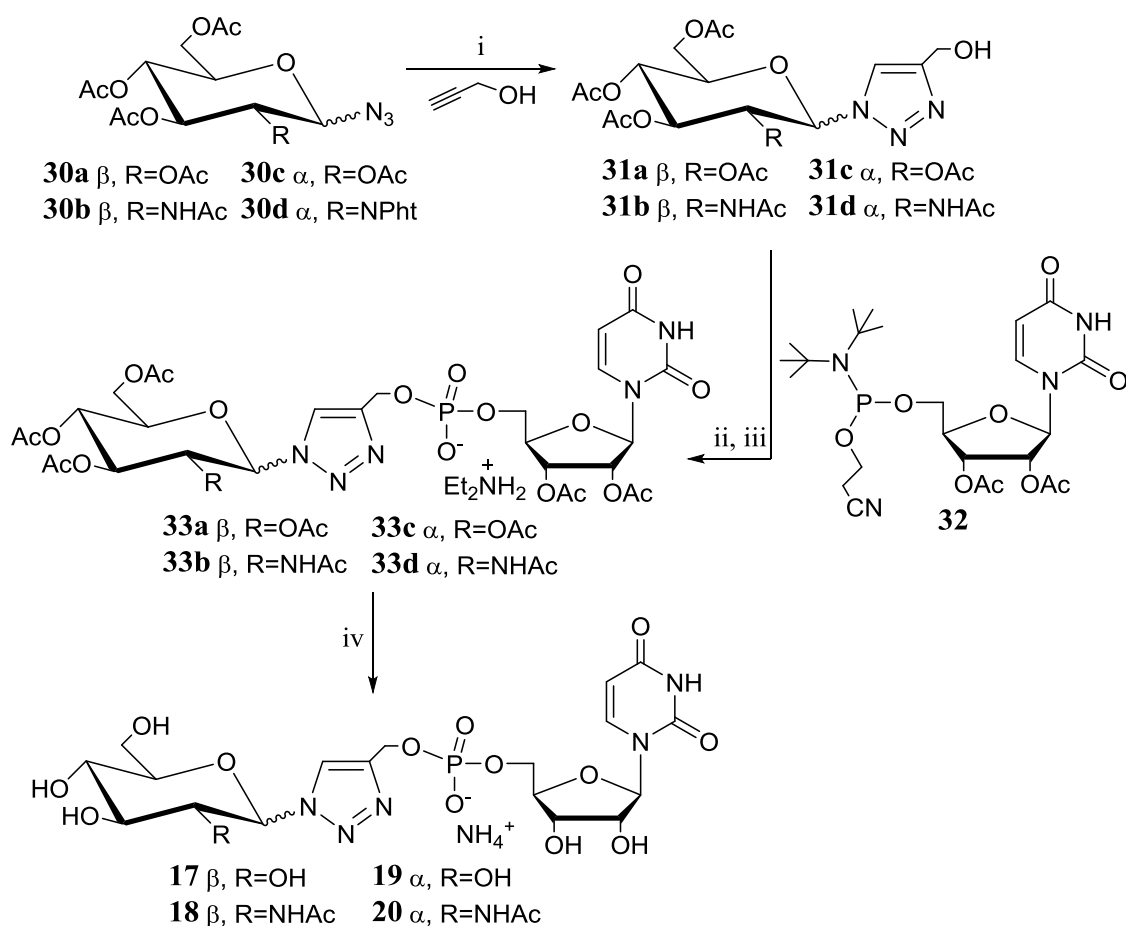
Scheme 8.1 *Syntheses of compounds **7** and **8**. Reagents and conditions: (i) dimethyl phosphite, diethyl phosphate acetophenone (DPAP), hv, r.t., 30min (73-94% yield); (ii) Me₃SiCl, NaBr, MeCN, 40°C, 1h; then r.t., 2h; then 5:1 H₂O/EtOAc, r.t., 2h; then NH₄OH, r.t., 2h (67-89% yield); (iii) PhSH, Et₃N, dioxane, r.t., 4h (44-80% yield); (iv) 10 (1.0 eq), (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (BOP), DMF, iPrEt₂N, r.t.,4h (64-93% yield); (v) 25% NH₄OH(aq), MeOH, r.t., 7h (63-65% yield).*

Novel compounds **9–16** were synthesized from nitrones **25a–e** through a highly diastereoselective allylation [57,58] and photoinduced free-radical hydrophosphonylation [59] as the key steps (Scheme 8.2). Selectively monodeprotected phosphonates **26a-e** were coupled with partially protected uridine to give advanced intermediates **27a-e**. Deprotection of the phosphate moiety and acid treatment afforded analogues **9**, **10**, **12**, **14** and **16** in good chemical yields. The fully deprotected analogues were synthesized starting from hydrophosphonylated intermediates **26b–d**. The benzyl groups were replaced by acetyl groups through hydrogenation in the presence of acetic anhydride; compounds **28b–d** were thereby obtained in excellent yields after partial deprotection of the phosphonate function. Coupling with partially protected uridine afforded intermediates **29b–d**. Complete deprotection of these intermediates yielded deprotected analogues **11**, **13** and **15** in good chemical yields.
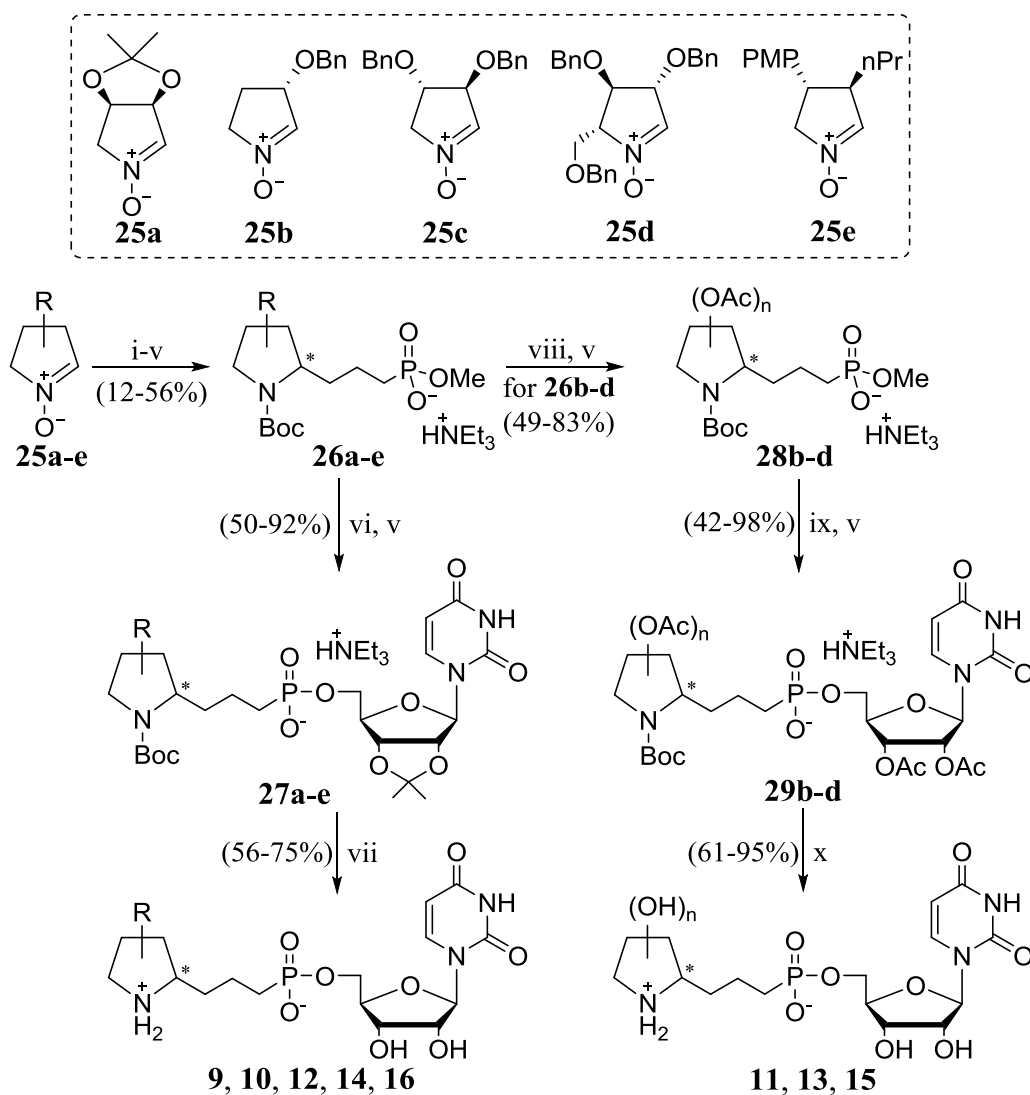
Scheme 8.2 Syntheses of compounds **9-16**. Reagents and conditions: (i) allylmagnesium bromide, THF, -80°C, 2h; (ii) Zn, AcOH, H_2O, r.t., 6h; (iii) Boc_2O, pyridine, CH_2Cl_2, r.t., 16h; (iv) dimethyl phosphite, diethyl phosphate acetophenone (DPAP), hv, r.t., 30min; (v) PhSH, Et_3N, dioxane, r.t., 4h; (vi) 2',3'-di-O-isopropylidene-uridine, Castro's reagent [Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate; BOP], DMF, iPr_2EtN, r.t., 4h; (vii) TFA, H_2O, r.t., 3h; (viii) H_2, $Pd(OH)_2$-C, Ac_2O, pyridine, r.t., 12h; (ix) 2',3'-di-O-acetyl-uridine, BOP, DMF, iPr_2EtN, r.t., 4h; (x) 25% NH_4OH, MeOH, r.t., 6h.

The syntheses of compounds **17–20** started from azides **30a-d** (Scheme 8.3), which were obtained from the corresponding carbohydrates (Scheme 8.4). Intermediates **31a-d** were obtained through a typical CuAAC "click" reaction [60,61,62,63,64], as has been widely employed in carbohydrate chemistry [65]. It was necessary to modify the reaction conditions slightly for each substrate. Whereas the classical copper(II) sulfate/sodium ascorbate system performed well for α-anomers, the reactions leading to β-anomers showed better results with the system $(EtO)_3P \cdot CuI$. In the case of compound **30d**, additional transformation of the phthalimido group into an acetamido group was necessary (Scheme 8.3). Coupling between glycosyltriazoles **31a-d** and the nucleotide unit **32** was realized using the phosphoramidite methodology [66,67,68]. The resulting intermediates **33a-d** were further transformed into the target analogues **17-20**.

30a β, R=OAc   **30c** α, R=OAc
30b β, R=NHAc   **30d** α, R=NPht

31a β, R=OAc   **31c** α, R=OAc
31b β, R=NHAc   **31d** α, R=NHAc

33a β, R=OAc   **33c** α, R=OAc
33b β, R=NHAc   **33d** α, R=NHAc

32

17 β, R=OH   **19** α, R=OH
18 β, R=NHAc   **20** α, R=NHAc

Scheme 8.3 *Syntheses of compounds **17-20**. Reagents and conditions: (i) for **30a** and **30b**: (EtO)$_3$P·CuI, 1:1 toluene/THF, reflux, 24h (75-85% yield); for **30c**: CuSO$_4$ (0.2 eq), sodium ascorbate (0.4 eq), THF, H$_2$O, reflux, 25min (90% yield); for **30d**: CuSO$_4$ (0.2 eq), sodium ascorbate (0.4 eq), tBuOH, H$_2$O, 50°C, 20h; then DMTrCl, pyridine, r.t., 18h, then 35% aq MeNH$_2$, MeOH, 70°C, 18h, then Ac$_2$O, pyridine, r.t., 20h, then 80% aq AcOH, MeOH, 50°C, 1h (75% yield); (ii) 4,5-dicyanoimidazole (DCI), MeCN, r.t., 45min; (iii) tBuOOH, decane, MeCN, r.t., 45min; then Et$_2$NH, MeCN, r.t., 45min(35-54% yield in 2 steps); (iv) NH$_4$OH, MeCN, r.t., 18h (95-97% yield).*

34   **30a** (52%)

35   **36** (69%)   **30b** (66%)

37 (64%)   **30d** (45%)

38   **30c** (43%)
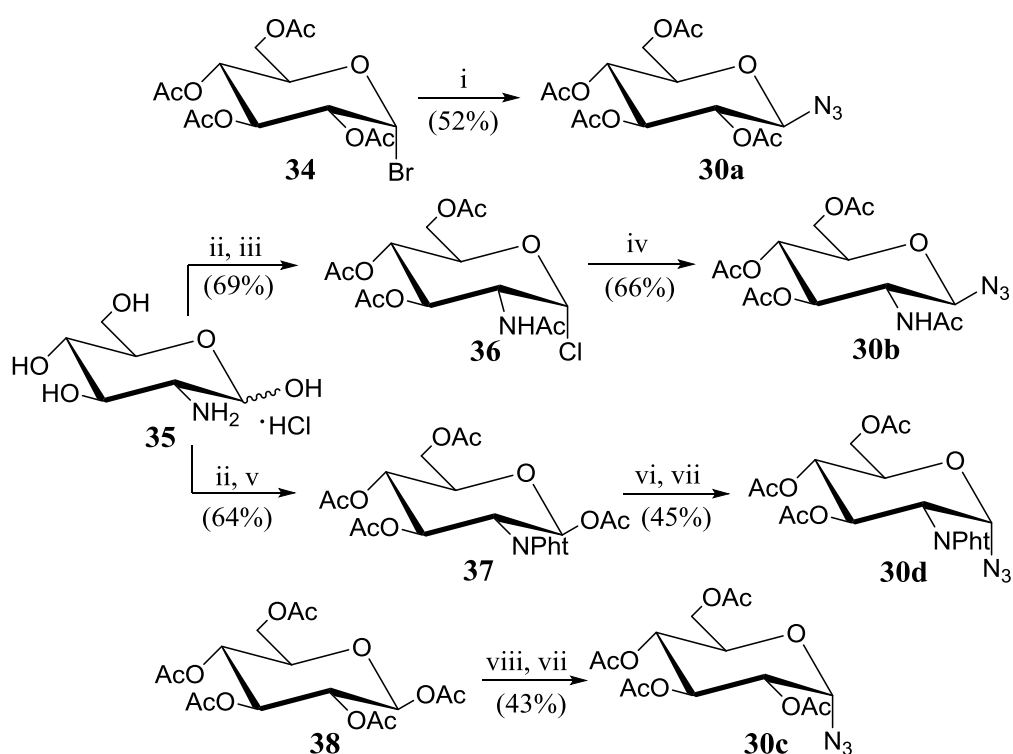
Scheme 8.4 *Syntheses of azides **30a-d**. Reagents and conditions: (i) NaN$_3$, DMSO, r.t, 18h; (ii) Na, MeOH, from 0°C to r.t., 1h; (iii) Ac$_2$O, r.t., 20h; then AcCl, r.t., 48h; (iv) NaN$_3$, tetrabutylammonium hydrogensulfate, saturated NaHCO$_3$,*

CH₂Cl₂, r.t., 1h; (v) phthalic anhydride, Et₃N, MeOH, 50°C, 2h; then Ac₂O, pyridine, 100°C, 2h; (vi) HBr 33% in AcOH, r.t., 90min; (vii) trimethylsilyl azide, tetrabutylammonium fluoride trihydrate, THF, 65°C, 24h; (viii) AlCl₃, CH₂Cl₂, r.t., 2h.

The nucleotide unit **32** was obtained from commercially available uridine (Scheme 8.5), protecting it as diacetyl uridine **23** *via* the trityl derivative [55,56] and causing it to react with 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite using the phosphoramidite methodology [66,67,68].
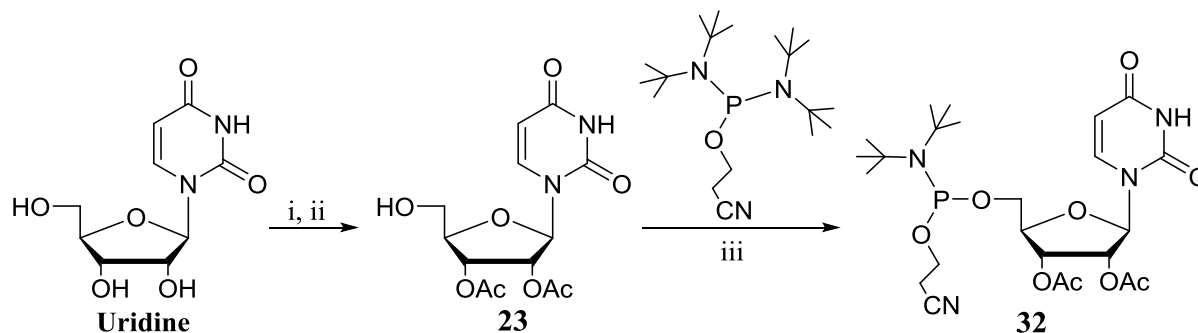
Scheme 8.5 *Synthesis of nucleotide unit **32**. Reagents and conditions: (i) 4,4'-dimethoxytrityl chloride, pyridine, r.t., 18h; then Ac₂O, r.t., 18h; (ii) AcOH aq 80%, MeOH, 50°C, 18h (74% yield in 2 steps); (iii) 4,5-dicyanoimidazole, MeCN, r.t., 1h (82% yield).*

8.2.2 Biological assays

Compounds **7–20** were subjected to enzyme-binding studies to human *O*-GlcNAc transferase (hOGT) based on fluorescence monitoring (Table 8.1). Binding affinities were determined for all of the compounds **7–20** by evaluating their efficacy in the displacement of a fluorescent probe [69]. The non-fluorescein modified probe was used as a positive control to verify the displacement. Most of the compounds showed no inhibition, and only compounds **14** and **18** showed moderate binding. Due to the low binding affinity of the compounds, 100% displacement was mimicked by introducing an artificial concentration of 10 M, which allowed the fitting of a four-parameter nonlinear regression curve and extrapolating to apparent K_i values. Among the sugar analogues studied (**7**, **8**, **17–20**), the best result was found for the UDP-GlcNAc analogue **18** having a β-configuration at the anomeric center, whereas the analogue **20** with the α-configuration (as in the natural substrate) proved to be a poor inhibitor. In fact, analogues **7** and **8**, both with the α-configuration, showed rather high K_i values. Another noteworthy finding was the high K_i value for compound **8**, which differs from the natural substrate only in the lack of the β-phosphate moiety. This finding clearly illustrates the key role of the phosphate unit in binding.

Pyrrolidinyl analogues showed high K_i values, with the exception of compound **14**, which showed a reasonable value of K_i=102.5 μM, the tightest binder among all of the UDP-GlcNAc analogues evaluated (the most potent inhibitor reported to date is the product of the reaction, UDP, showing K_d=0.5 μM [70]). Presumably, the presence of benzyl groups might promote hydrophobic interactions that stabilize the ligand. Indeed, successive elimination of benzyl groups to afford compounds **12** and **10** resulted in a notable increase in K_i

Inhibitor	K_i (μM)
7	920.7 ± 58
8	4769.0 ± 905
9	5452.0 ± 1875
10	718.0 ± 33
11	2218.0 ± 595
12	1140.0 ± 473
13	1804.0 ± 110
14	102.5 ± 19
15	1814.0 ± 132
16	2262.4 ± 10
17	1439.5 ± 249
18	231.9 ± 4
19	1048.0 ± 475
20	589.1 ± 63

Table 8.1 *Enzymatic inhibition of human O-GlcNAc transferase (hOGT). Apparent K_i values were calculated by introducing a 100% displacement at a putative concentration of 10 M.*

values. In the case of compound **12**, the opposite configurations of the stereogenic centers of the pyrrolidine ring with respect to **10** and **14** might also adversely affect the binding. The high values observed for analogues **11**, **13** and **15** suggest that H-bonding interactions are not essential in the relevant section of the binding pocket. Attempts to crystallize compound **14** complexed with hOGT revealed high electron density for the UMP part in the binding site, but low electron density for the rest of the ligand. Presumably, the high flexibility of the latter leads to a lack of electron density, thus precluding model building. Nevertheless, given the correct orientation of UMP, it is possible to infer the orientation of the benzyl groups toward a hydrophobic pocket, as supported by computational calculations (see "Computational studies" below).

8.2.3 Spectroscopic studies

Saturation transfer difference (STD) NMR spectroscopy [71,72] has been demonstrated to be of great utility in evaluating ligand–protein binding affinities [73]. In particular, STD-NMR allows the identification of binding epitopes [74], which reflect the distances between ligand protons and protons of binding site residues. STD-NMR experiments were carried out on selected compounds to cover a range of K_i values, from the best ligand **14** (K_i=102 μM) to **12** (K_i=1.14 mM). UMP was selected as a reference instead of UDP since the low K_d (0.5 μM) of the latter would prevent the observation of any displacement by the added ligand. Moreover, since the studied ligands lack the β-phosphate moiety, the use of UMP as a reference will provide a correct evaluation of the interactions of the glycomimetic moiety. To obtain the epitope mappings of the selected ligands, UMP, **10**, **14** and **18** were incubated with hOGT at pH 7.4 and the STD-NMR spectra were acquired at 25°C. Due to the low affinity, the experiments were carried out with a 1:50 enzyme/ligand ratio, which yielded STD spectra with the best quality. Analysis of the STD intensities allowed mapping of the epitopes of UMP and compounds **10**, **14**, and **18** (Figure 8.4). The uridine moiety consistently showed the same trend in STD intensities, with 100% relative STD enhancement for H-1 in all cases. The best results, with strong interactions for all of the uridine protons, were found for the best ligand, **14**. On the other hand, compound **18** only showed interactions for H-1, H-2, and H-1'. Remarkably, compound **14** showed better recognition of the uridine moiety than UMP. Indeed, when **14** was added to a solution containing the complex UMP-hOGT, only signals resembling an hOGT **14** complex were observed with higher intensity, showing that **14** displaces UMP and binds hOGT with a higher affinity. These results are consistent with the crystallographic structure of the complex UDPGlcNAc/hOGT [37], in which the uridine protons directed towards the inner part of the protein show the highest intensities. The most intense STD signals for the glycomimetic unit were also observed with **14**, corresponding to the three phenyl rings, thus suggesting the presence of important hydrophobic interactions.

In fact, elimination of two benzyl groups, as in **10**, resulted in a considerable loss of affinity, implying that the remaining benzyl group at C-3 of the pyrrolidine ring was not essential in the binding. In the case of UDP-GalNAc analogue **18**, only H-1 and H-2 of the galactosamine unit showed a strong interaction with the protein. The triazole ring did not show any remarkable interaction, indicating that it does not occupy a favorable position for binding. No strong interactions were observed for the alkyl chain that replaced the β-phosphate unit, confirming that the loss of binding was due to elimination of the latter.
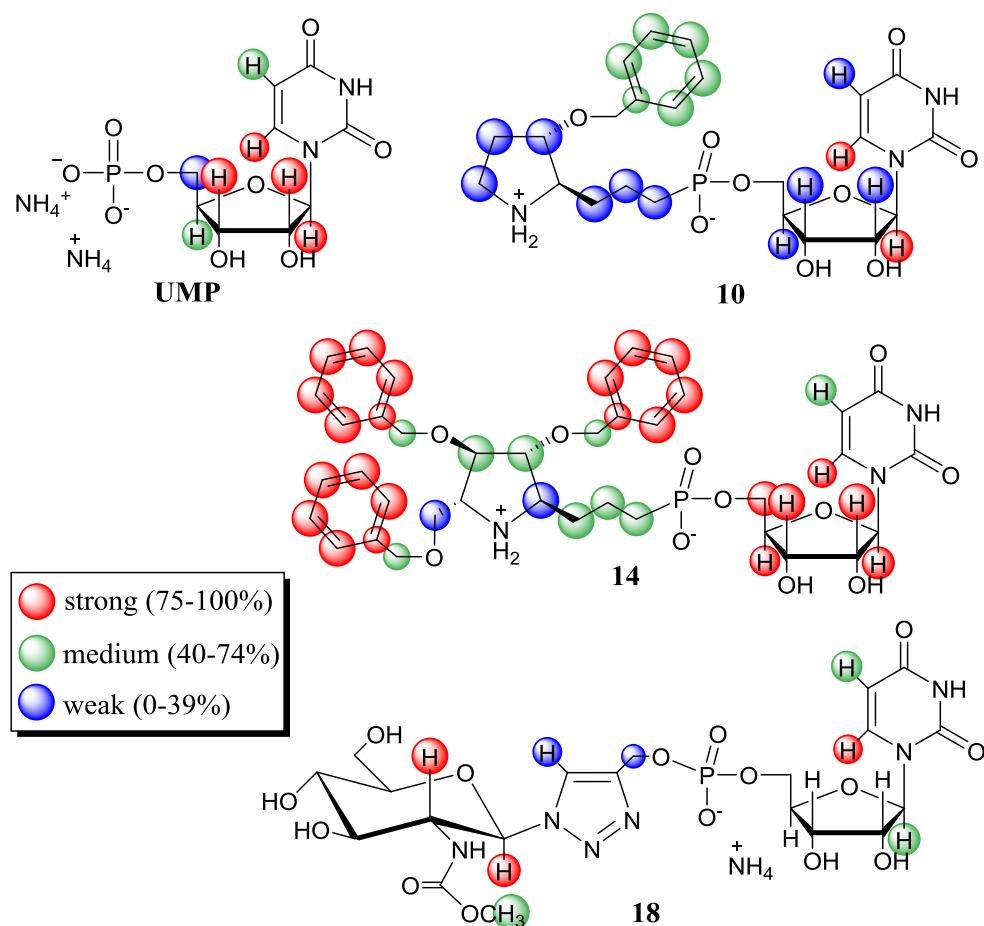
Figure 8.4 *Epitope mapping of UMP and compounds 10, 14 and 18.*

Nevertheless, whereas only a weak interaction was observed for **10**, some limited interaction was observed in the case of **14**. STD data confirmed that a pyrrolidine ring with the appropriate configuration, as seen in **14**, is an adequate scaffold for engaging the aromatic residues. However, other systems would also be suitable for the same task. These experiments suggest the presence of a hydrophobic pocket on hOGT and the importance of establishing hydrophobic interactions with the surrounding area of the sugar residue binding site.

8.2.4 Computational studies

To shed additional light on the protein–ligand binding modes, computational studies with the best ligands **14** and **18** were carried out using the known structure of OGT complexed with UDP-GlcNAc (PDB ID: 4GZ5) [37] as a template (comparative studies were also carried out with UDP-GlcNAc, **10** and **12**). Published crystal structures of UDP-GlcNAc and UDP complexed with hOGT [37] show interactions between the base moiety and Ala896 (H-bond) and His901 (π,π-interactions). Previously reported docking studies using the crystal structure of OGT complexed with UDP-GlcNAc as a template indicated the presence of four possible hydrogen bonds between the GlcNAc residue and OGT, that is C6-OH/Thr560, C4-OH/Leu653, C3-OH/Gly654 and C2-*N*-acyl/His920) [75]. The same studies suggested that the backbone carbonyl oxygen atom of Leu653 and the hydroxyl group of Thr560 contribute to binding of UDP-GlcNAc through key H-bonding interactions, and that the C2-acetamido group is directed towards a hydrophobic pocket constituted by Met501, Leu502, and Tyr841. Docking studies with β-1-triazolyl-GlcNAc analogues suggested interactions between the triazole ring and Thr921 and Thr924 resembling those involving the β-phosphate of UDP-

GlcNAc [42]. All of these computational studies were based on docking calculations with the *Schrödinger 2017* software package [44]. Docking with **14** and **18** was performed by including both ligands in the UDP-GlcNAc binding site. In both cases, it was observed that the orientation of the nucleoside-α-phosphate unit resembled that of the natural substrate, and the same interactions were observed for both UDP-GlcNAc and other ligands.

Molecular dynamics simulations allowed study of the evolution of the complex and offered more valuable and representative information than static or minimum-energy docking calculations. In this context, MD simulations have previously been used for elucidating the catalytic mechanism of OGT [76], further supported by structural data [37]. These studies pointed to the key role of α-phosphate and Asp554 as the catalytic bases, and this was further confirmed as the most energetically favorable pathway [77].

On the other hand, when Asp554 was mutated out, complete loss of activity was not observed, making it more plausible that α-phosphate could act as the catalytic base [70]. Recent structural studies carried out by one of us established that the peptide-binding site imposes size and conformational restrictions [35]. In our studies, we used the generated orientations in the docking studies to run MD simulations for 100 ns. A preliminary analysis showed that, after 50 ns, the situation had stabilized and essentially the same orientation was found for the ligands (Figure 8.5A).
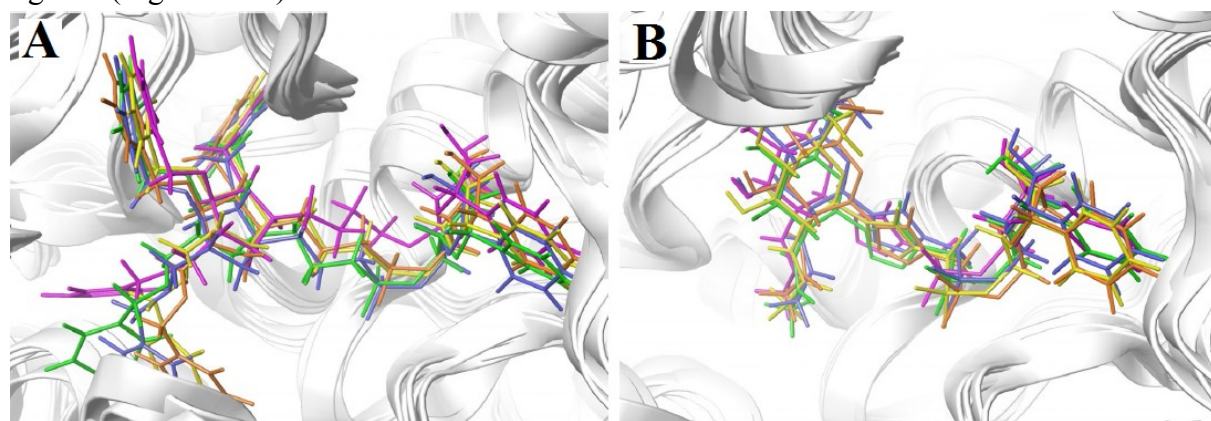

Figure 8.5 *Superimposition of ten structures of OGT/**14** (A) and OGT/**18** (B) complexes along the MD simulation from 50 to 100 ns; intermediate structures are shown graded from orange (starting geometry), through yellow, green and blue (intermediate geometries), to magenta (final target geometry)* [78].

For compound **14**, however, different orientations were observed for the CH$_2$OBn group at C-5 of the pyrrolidine ring, whereas the other benzyloxy groups at C-3 and C-4 remained in the same orientations. Compound **18** was consistently in the same orientation (Figure 8.5B).
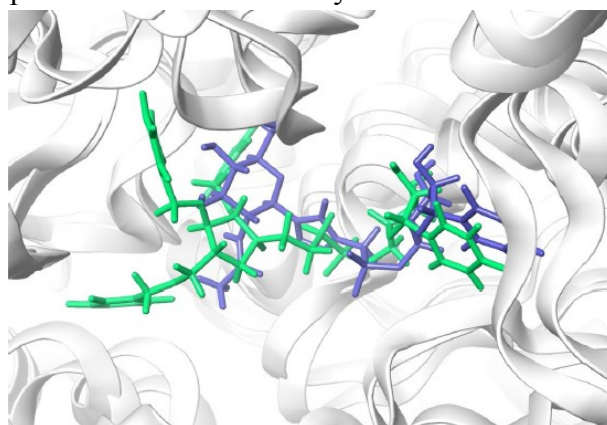

Figure 8.6 *Superimposition of the OGT/**14** (green) and OGT/**18** (blue) complexes found at 100 ns* [78].

A comparison between **14** and **18** (Figure 8.6) revealed that both have the same preferred conformation of the mononucleotide unit (exo orientation of the nucleobase with respect to the ribose unit with a dihedral angle C(C=O)-N(glycosidic)-C(anomeric)-O(ribose) of around 180°). On the other hand, such a unit is oriented in a different way inside the binding pocket, although the same interactions were observed for this part of the molecule (see below). Notably, the glycomimetic unit in **8** (glycosyltriazole moiety) was seen to point in the same direction as the less fixed CH_2OBn group at C-5 of the pyrrolidine ring of **14**. This result is consistent with the observed lower binding of **18** and supports the hypothesis that benzyl groups at C-3 and C-4 are required for binding, particularly that at C-4. Indeed, compound **12**, with only a single benzyl group at C-3, showed a markedly lower affinity.

A detailed inspection of the interactions between OGT and **14** revealed that they were the same as those observed in the crystal structure of the mononucleotide unit (Figure 8.7A).
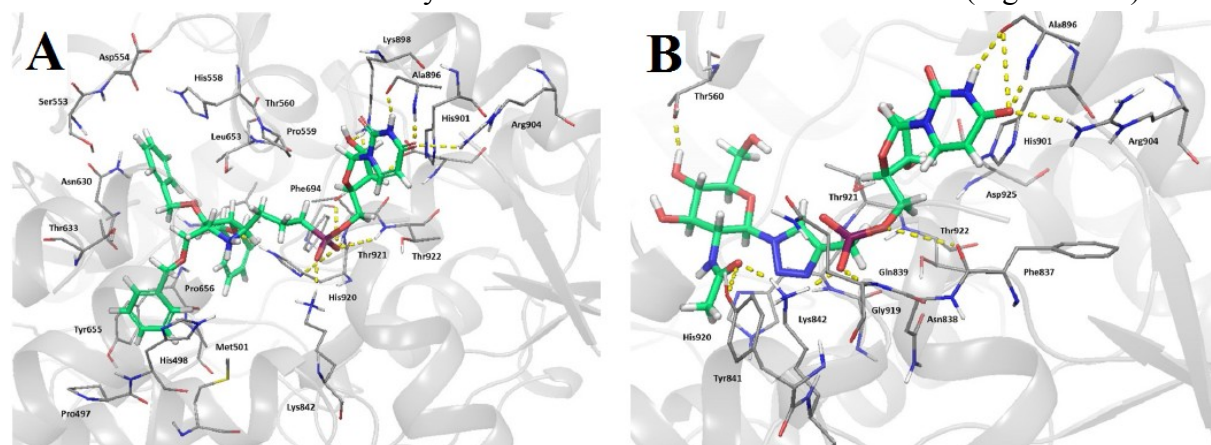

Figure 8.7 *MD simulations for compounds **14** (A) and **18** (B). Structures correspond to stable situations after 100 ns* [78].

The nucleobase shows H-bonding interactions with Ala896 and Arg904, and hydrophobic interactions with His901. The ribose unit shows H-bonding interactions with residues Lys898 (through O at C-2' and C-3') and Asp925 (through OH at C-3'). The phosphate group is engaged in H-bonding interactions with Lys842, Cys911, His920, Thr921 and Thr922. The three benzyl groups, responsible for the observed affinity, are installed in hydrophobic pockets. An aromatic group at C-3 of the pyrrolidine ring is surrounded by Leu653, Phe694 and Thr560. The benzyl group at C-4 is directed towards a cavity formed by Ser553, Asp554, Thr560 and His558. The benzyloxymethyl group at C-5 is oriented towards the hydrophobic region formed by Pro497, His498, Met501, Thr633, Tyr655 and Pro656. Compound **18** (Figure 8.7B) shows the same interactions for the nucleobase (Ala896, Arg904 and His901) and H-bonding to Asp925 for the ribose unit.

On the other hand, the interactions of the NHAc group of the glucosamine moiety with His498 and His920, present in the crystal structure and MD simulations, are lost in **18**, which forms H-bonding interactions with Tyr841 and Lys842. In contrast to previous docking studies [29], but in agreement with the STD-NMR experiments, no interactions were found for the triazole ring.

8.3   Conclusion

In the search for less polar compounds to increase the bioavailability of potential OGT inhibitors, we have used click chemistry procedures based on photoinduced free-radical hydrophosphonylation and copper-catalyzed alkyne azide cycloaddition (CuAAC) to prepare

a series of glycomimetics of UDP-GlcNAc, in which the β-phosphate group has been replaced by an alkyl chain. It has been reported that both the acetamido and β-phosphate groups play key roles in binding to hOGT, whereas α-phosphate is crucial for catalysis [70]. The high K_i value found for compound **7** (OH instead NHAc) supports the key role of the acetamido group in the binding [37]. Indeed, it has been reported that OGT can transfer UDP-GalNAc, but not UDP-Glc [37]. However, the higher K_i value for **8**, which bears an NHAc group, suggests that the absence of the β-phosphate is responsible for a wrong orientation of the carbohydrate unit. The β-phosphate moiety provides key hydrogen bonds (with Lys842 and three contiguous residues His920, Thr921, and Thr922), which are observed in both the studied OGT-UDP-GlcNAc complex (PDB: 4GZ5) and the ternary complex OGTUDP-GlcNAc peptide (PDB: 4GYW) [37]. Both experimental and computational results have confirmed the crucial role of β-phosphate in the binding of hOGT in a similar way to other glycosyltransferases requiring the presence of a metal ion, such as GalNAc-T2 [31]. Consequently, the elimination of the β-phosphate requires groups in the analogues capable of providing interactions to counterbalance those lost. In this respect, glycomimetics with free hydroxyl groups have proved not to be good candidates, suggesting that H-bonding interactions are not the key in this case. On the other hand, a triazole ring engages in some interactions in the case of compound **18** (K_i=231.9 μM), in agreement with the findings of Vidal and coworkers [29]. However, the best result was found for compound **14** (K_i=102.5 μM), with a value of the order of those reported in the literature [30]. This result was supported by STD-NMR experiments and computational (docking and MD) calculations, and points out the importance of hydrophobic interactions in the vicinity of the active site occupied by the carbohydrate for the design of an inhibitor.

Our results are consistent with the lower affinity of compounds in which the β-phosphate unit has been replaced by a different moiety, and reaffirm the importance of the interactions of such a unit, even in the absence of a metal, for achieving good binding. This is true in the present case of hOGT, for which it has been suggested that Lys842 adopts the role of the metal ion. At the same time, our study has shown the possibility of introducing hydrophobic groups at the carbohydrate surrogate to increase affinity. In other words, the lower affinity caused by the absence of β-phosphate is counterbalanced by the hydrophobic interactions of the poly-*O*-benzylated pyrrolidine acting as a surrogate of the carbohydrate unit. A combination of this approach with the replacement of the β-phosphate by a functional group that provides interactions capable of competing with those observed in the natural substrate might be the way to find inhibitors of hOGT at a nanomolar level [78].

8.4 Experimental

8.4.1 General information

The reaction flasks and other glass equipment were heated in an oven at 130°C overnight and assembled in a stream of Ar. All reactions were monitored by TLC on silica gel 60 F254; the position of the spots was detected with 254 nm UV light or by spraying with 5% ethanolic phosphomolybdic acid. Column chromatography was carried out in a *Buchi 800 MPLC* system using silica gel 60 microns or *Combiflash MPLC* system with pre-charged cartridges for flash chromatography. Melting points were uncorrected. ^{1}H- and ^{13}C-NMR spectra were recorded on *Bruker Avance 400* and *Varian Mercury Plus 400* instruments in the stated solvent. Chemical shifts are reported in ppm (δ) relative to TMS as external reference. Optical

rotations were taken on a *JASCO P-1020 DIP-370* polarimeter. Elemental analyses were performed on a *Perkin Elmer 240B* microanalyzer or with a *Perkin-Elmer 2400* instrument. ESI-HRMS were acquired on an *Agilent Dual ESI Q TOF 6520*, in positive-ion mode, using methanol.

This work was carried out in collaboration with the universities of Zaragoza (Spain) and Dundee (U.K.); below are reported the synthesis procedures and the characterizations of the compounds synthesized by my research group at the University of Ferrara.

8.4.2 Synthesis of uridine-phosphoramidite
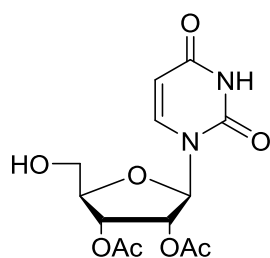
2',3'-di-*O*-acetyl uridine 23

Uridine (2.0 g, 8.2 mmol) was coevaporated twice with anhydrous CH_3CN, dissolved in 20 mL of anhydrous pyridine and 4,4'-dimethoxytrityl chloride (3.33 g, 9.8 mmol) was added under argon. Upon 18 h stirring at r. t., the reaction was complete and acetic anhydride (2.3 mL, 24.6 mmol) was added and the reaction stirred overnight, then treated with 0.5 mL of CH_3OH. The mixture was concentrated *in vacuo* and redissolved in 20 mL of CH_2Cl_2. The organics were washed with saturated $NaHCO_3$ (2 x 15 mL) and the water layers were extracted with 20 mL of CH_2Cl_2. The combined organic layers were dried with Na_2SO_4 and concentrated in *vacuo* giving 5.75 g of crude 2',3'-diacetyl-5'-trityl uridine as a pale yellow foam which was dissolved in CH_3OH and treated with 56 mL of aqueous acetic acid 80%. The solution was stirred at 50°C for 18 h, then concentrated in *vacuo* and purified by flash chromatography (CH_2Cl_2/CH_3OH 95:5) to give 2.0 g (74%) of pure compound **23** as a white wax solid. $[\alpha]_D^{25} \cong +2$ (*c* 0.7, DMSO). ^{1}H-NMR(400 MHz; $CDCl_3$): δ= 9.76 (bs, 1H, NH); 7.80 (d, 1H, *J*=8.1 Hz, H-6); 6.07 (d, 1H, *J*=5.6 Hz, H-1'); 5.79 (d, 1 H, *J*=8.1 Hz, H-5); 5.48-5.41 (m, 2H, H-2', H-3'); 4.21-4.17 (m, 1H, H-4'); 3.91 (dd, 1H, *J*=12.2, 2.2 Hz, H-5'$_A$); 3.83 (dd, 1H, *J*=12.2, 2.2 Hz, H-5'$_B$); 3.63 (bs, 1H, OH); 2.11 (s, 3H, CH_3); 2.06 (s, 3H, CH_3). ^{13}C-NMR (101 MHz; $CDCl_3$): δ= 170.2 (CO); 169.9 (CO); 163.8 (CO); 150.7 (CO); 141.0 (HC=); 103.1 (HC=); 87.3 (CH); 83.5 (CH); 73.1 (CH); 71.3 (CH); 61.7 (CH_2); 20.7 (CH_3); 20.5 (CH_3). HRMS Calcd for $[C_{13}H_{16}N_2O_8 + H]^+$ 329.0980, Found 329.1006.
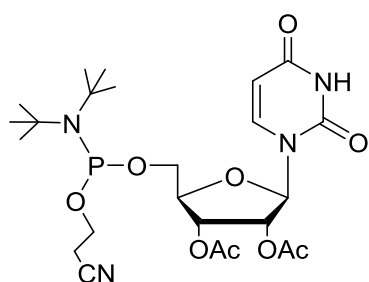
2',3'-di-*O*-acetyl-5'-phosphoramidite uridine 32

2',3'-diacetyl uridine **23** (0.40 g, 1.22 mmol) and 4,5-dicyanoimidazole (0.14 g, 1.16 mmol) were coevaporated twice with anhydrous CH_3CN, dissolved in 3 mL of anhydrous CH_3CN and then added *via cannula* to a stirred solution of 2-cyanoethyl *N,N,N',N'*-tetraisopropylphosphorodiamidite (0.37 g, 1.22 mmol) in 1 mL of CH_3CN. After one hour stirring, the solvent was removed in *vacuo* and the crude residue diluted with CH_2Cl_2 (10 mL) and washed with saturated $NaHCO_3$ (2 x 15 mL). The water layers were extracted with 10 mL of CH_2Cl_2 and the combined organic layers dried with Na_2SO_4. After removing the solvent, the crude compound was purified by flash chromatography (AcOEt/cyclohexane/Et_3N 60:35:5) to give 0.53 g (82%) of pure **32** as a white wax solid. The NMR analysis showed the presence of **32** as a mixture of two diastereomers at approximately 55:45 ratio. ^{1}H-NMR (400 MHz; DMSO-

d_6): δ= 11.20 (bs, 2H, NH); 7.74 (d, 1H, J= 8.1 Hz, H-6); 7.70 (d, 1H, J= 8.1 Hz, H-6); 5.99 (d, 1H, J=6.1 Hz, H-1'); 5.97 (d, 1H, J=6.1 Hz, H-1'); 5.37-5.27 (m, 4H, H-2', H-3'); 4.28-4.24 (m, 2H, H-4'); 3.85-3.68 (m, 8H, H-5'$_A$, H-5'$_B$, CH$_2$OP); 3.62-3.50 (m, 4H, NCH); 2.80-2.73 (m, 4H, CH$_2$CN); 2.08 (s, 3H, COCH$_3$); 2.07 (s, 3H, COCH$_3$); 2.01 (s, 6H, COCH$_3$); 1.19-1.08 (m, 24H, NCH(CH_3)$_2$). ^{13}C-NMR (101 MHz; DMSO-d_6): δ= 169.9 (CO); 169.8 (CO); 163.4 (CO); 163.3 (CO); 150.9 (CO); 140.7 (HC=); 119.4 (CN); 119.3 (CN); 103.1 (HC=); 102.9 (HC=); 86.5 (CH-1'); 86.3 (CH-1'); 81.9 (CH-4'); 81.8 (CH-4'); 72.5 (CHOAc); 72.4 (CHOAc); 71.1 (CHOAc); 70.9 (CHOAc); 63.3 (CH$_2$-5'); 62.9 (CH$_2$-5'); 59.1 (OCH$_2$CH$_2$); 58.9 (OCH$_2$CH$_2$); 43.1 (CH(CH$_3$)$_2$); 42.9 (CH(CH$_3$)$_2$); 42.8 (CH(CH$_3$)$_2$); 24.9 (CH$_3$); 24.8 (CH$_3$); 23.0 (CH$_3$); 20.8 (CH$_3$); 20.6 (CH$_3$); 20.3 (OCH$_2$CH$_2$). ^{31}P-NMR (122 MHz; DMSO-d_6): δ= 147.3; 147.1. HRMS Calcd for [C$_{22}$H$_{33}$N$_4$O$_9$P + H]$^+$ 529.2058, Found 529.2046.

8.4.3 Synthesis of glycosyl azides
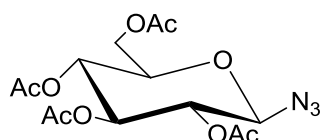
2,3,4,6-tetra-O-acetyl-Glc β-azide 30a

To a solution of peracetylated glucosyl α-bromide **34** (2.04 g, 4.96 mmol), in dry DMSO (7 mL) was added sodium azide (0.39 g, 5.95 mmol) and the reaction was allowed to steer at room temperature for 18 h. The reaction mixture was diluted with water (100 mL) and extracted with AcOEt (2 x 25 mL). The organic layer, dried with Na$_2$SO$_4$, was evaporated to dryness to give peracetylated glucosyl β-azide as brown solid. Crystallization from AcOEt and cyclohexane gives **30a** (960 mg, 52%) as a white solid. Mp=126-127°C. [α]D^{25} ≅ -33 (c 0.5, DMSO). Analytical and spectroscopic were found to be identical to those reported in the literature [79].
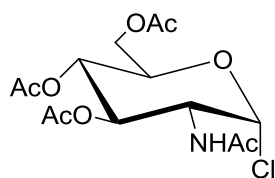
**3,4,6-tri-O-acetyl-GlcNAc α-chloride 36**

Sodium (114 mg, 4.9 mmol) was added in small pieces to 5 mL of methanol in a roundbottomed flask with cooling in an ice-bath. The solution was brought to room temperature and glucosamine **35** (1.04 g, 5 mmol) was added. The mixture was gently swirled for 1h, then treated with 0.6 mL (6.5 mmol) of acetic anhydride, and the flask was cooled to moderate the initial reaction. The solution was kept for 20 h at room temperature. The mixture was concentrated under reduced pressure and the crude evaporated twice with toluene (2 · 7 mL) and resuspended in the minimum amount of AcCl. The suspension was stirred for 48 h at room temperature monitoring with TLC (AcOEt) until formation of expected product. To the suspension 40 mL of dichloromethane were added and the solution was poured with vigorous stirring onto 40 ml of solution of NaHCO$_3$ and ice. The organic layer was dried over Na$_2$SO$_4$, concentrated to 2 mL on a rotary evaporator and added with 10 mL of dry ether. The solid **36** (1.28 g, 69%) was collected by filtration, washed with dry ether and dried well. [α]D^{25} ≅ +125 (c 0.7, DMSO). Analytical and spectroscopic were found to be identical to those reported in the literature [80].
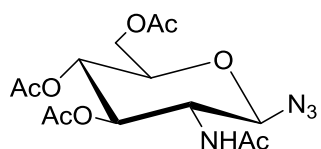
**3,4,6-tri-O-acetyl-GlcNAc β-azide 30b**

A solution of chloride **36** (500 mg, 1.36 mmol) containing sodium azide (267 mg, 4.1 mmol) and tetrabutylammonium hydrogensulfate (464 mg, 1.36 mmol) in CH$_2$Cl$_2$ (5 mL) was prepared at room temperature.

Saturated NaHCO$_3$ (5 mL) was added and the mixture was stirred vigorously for 1 h; the reaction was monitored by TLC (AcOEt). The mixture was diluted with CH$_2$Cl$_2$ (5 mL) and washed with water, saturated NaHCO$_3$ and brine. The organic layer was dried over Na$_2$SO$_4$, filtered and the solvent was evaporated under reduced pressure. Compound **30b** (334 mg, 66%) was obtained as a white solid after flash chromatography (AcOEt/Cy/MeOH 50:50:2). $[\alpha]_D^{25} \cong$ -45 (*c* 0.6, DMSO). Analytical and spectroscopic were found to be identical to those reported in the literature [79].
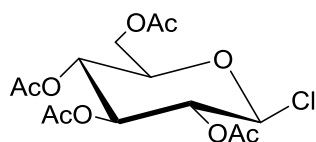
2,3,4,6-tetra-*O*-acetyl-Glc β-chloride 39

To a solution 0.5 M of **38** (1 g, 2.56 mmol) in dry CH$_2$Cl$_2$ (5.12 mL) aluminium trichloride (171 mg, 1.28 mmol) was added under argon at room temperature. After 2 h stirring, the reaction mixture was diluted with 100 mL of cyclohexane. The resulting white solid was filtered on celite and washed with CH$_2$Cl$_2$ (5 mL) and 2:3 AcOEt/Cy (100 mL). The solvent was evaporated under reduced pressure and the crude product **39** used in the next step without further purification.
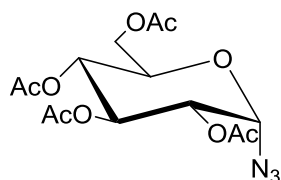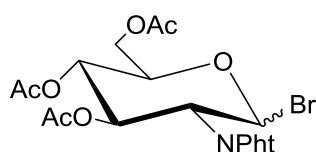
2,3,4,6-tetra-*O*-acetyl-Glc α-azide 30c

To a solution 0.1 M in THF (25 mL) of crude **39** were added 475 μl of trimethylsilyl azide (3.58 mmol) and 1.13 g (3.58 mmol) of tetrabutylammonium fluoride trihydrate. The solution was stirred at 65°C for 24 h, until disappearance of starting material (TLC AcOEt/Cy 2:3). The solvent was evaporated under reduced pressure and the crude was purified by flash chromatography using 1:3 AcOEt/Cy as eluent. Compound **30c** (410 mg, 43%, 2 steps) was obtained as a wax solid. Analytical and spectroscopic were found to be identical to those reported in the literature [81].
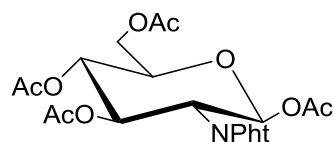
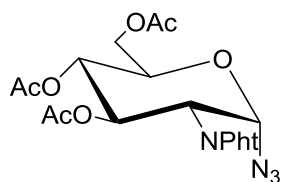### 1,3,4,6-tetra-O-acetyl-GlcNPht 37

Sodium (114 mg, 4.9 mmol) was added in small pieces to 5 mL of methanol in a roundbottomed flask with cooling in an ice-bath. The solution was brought to room temperature and glucosamine **35** (1.04 g, 5 mmol) was added. The mixture was gently swirled for 1h, then treated with 639 μL (5.5 mmol) of phtalic anhydride and 693 μL (5 mmol) of TEA. The solution was kept for 2 h at 50°C. The mixture was concentrated under reduced pressure and the crude evaporated twice with toluene (2 · 7 mL) and resuspended in 4.6 mL of Py. Ac$_2$O (3.78 mL; 40 mmol) was added and the solution was stirred for 2 h at 100°C, monitoring by TLC (AcOEt/Cy 3:7). The mixture was concentrated *in vacuo*, diluted with CH$_2$Cl$_2$ (50 mL) and washed with H$_2$O (50 mL). The organic layer was separated and dried over Na$_2$SO$_4$. The solution was filtered and the solvent was evaporated under reduced pressure. Compound **37** (1.527 g, 64% in 2 steps) was obtained as a white amorphous solid after flash chromatography (AcOEt/Cy 2:3). Analytical and spectroscopic were found to be identical to those reported in the literature [80].

3,4,6-tetra-*O*-acetyl-GlcNPht bromide 40

A solution of **37** (1.2 g, 2.52 mmol) in a 33% *w/v* solution of hydrogen bromide in glacial acetic acid (9.6 ml 39.06 mmol) was prepared. This solution was stirred for 1.5 h at room temperature and the reaction was

monitored by TLC (AcOEt/Cy 3:7). The solution was then diluted with CH_2Cl_2 (50 mL) and poured into saturated ice-cold $NaHCO_3$. The pH was adjusted with solid $NaHCO_3$ and the organic layer separated and dried over Na_2SO_4. The solution was filtered and the solvent was evaporated under reduced pressure. The crude product **40α/β** (1.23 g) was used without any further purification.

3,4,6-tri-*O*-acetyl-GlcNPht α-azide **30d**

To a solution of crude **40α/β** in THF (25 mL, 0.1 M), trimethylsilyl azide (467μL, 3.53 mmol) and 1.11 g tetrabutylammonium fluoride (3.53 mmol) were added. The solution was heated to 65°C and stirred for 6 h. The reaction was monitored by TLC (7:3 Cy/AcOEt). The solvent was evaporated under reduced pressure and the crude was purified by flash chromatography using 2:3 AcOEt/Cy as the eluent. Product **30d** (0.47 g) was obtained as a wax solid 45% (over 2 steps). Analytical and spectroscopic were found to be identical to those reported in the literature [80].

8.4.4 Synthesis of glycosyl triazoles
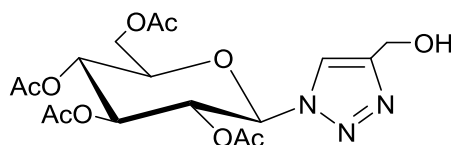
2,3,4,6-tetra-*O*-acetyl-Glc β-triazole **31a**

To a vigorous stirring suspension of **30a** (200 mg, 0.53 mmol) in *tert*-butyl alcohol (1.1 mL) was added propargyl alcohol (131 μl, 2.24 mmol). The reaction started with 1.7 ml of 0.1 M aqueous solution of $CuSO_4 \cdot 5\ H_2O$ (0.107 mmol) and 0.43 ml of 0.5 M solution of sodium ascorbate (0.214 mmol). The yellow suspension was stirred at 50°C until TLC (AcOEt/Cy 3:2) indicated reaction completion (3 h). The suspension was diluted with water (15 mL) and extracted twice with CH_2Cl_2 (2·10mL). The organic layer was dried, filtered and evaporated to give **31a** (196 mg, 85%) as a white solid. Mp=154-157°C. $[\alpha]_D^{25} \cong$ -31 (*c* 0.5, DMSO). Analytical and spectroscopic were found to be identical to those reported in the literature [82].
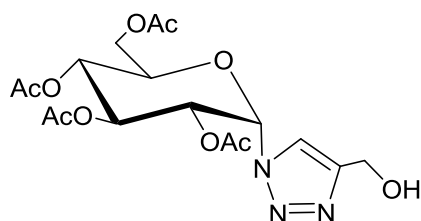
8.4.4.1 General procedure for "click" reactions with (EtO)₃P·CuI

The glycosyl azides (1 eq), were treated with propargyl alcohol (1.2 eq), using the copper catalyst (EtO)₃P·CuI (0.1 eq) in refluxing toluene/THF (1:1) for the appropriate time, until TLC showed complete disappearance of the limiting reagent. The mixture was concentrated *in vacuo* to give the corresponding glycosyl triazoles which were purified by flash chromatography.
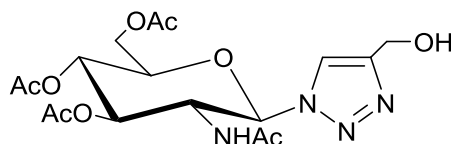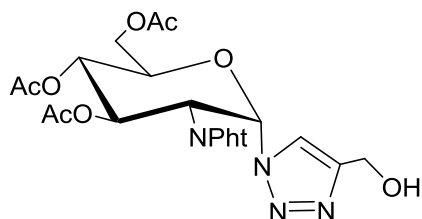
3,4,6-tri-*O*-acetyl-GlcNAc β-triazole **31b**

"General procedure for "click" reactions with (EtO)₃P·CuI" applied to **30b** yields **31b** (97 mg, 75%) as a white solid after flash chromatography (AcOEt/MeOH 8:2). Analytical and spectroscopic were found to be identical to those reported in the literature [80].

2,3,4,6-tetra-*O*-acetyl-Glc α-triazole **31c**

"General procedure for "click" reactions with (EtO)₃P·CuI" applied to **30c** yields **31c** (290 mg, 90%) as a white wax solid after flash chromatography (AcOEt/cyclohexane 3:1).

Analytical and spectroscopic were found to be identical to those reported in the literature [(82)].

3,4,6-tri-*O*-acetyl-GlcNPht α-triazole 41

"General procedure for "click" reactions with (EtO)$_3$P·CuI" applied to **30d** yields **41** (166 mg, 80% yield) as a white wax after flash chromatography (AcOEt/cyclohexane/MeOH 70:30:3). [α]D^{25}≅ +101 (*c* 0.5, DMSO). ^{1}H-NMR (400 MHz; DMSO-*d$_6$*): δ= 8.05 (s, 1H, C=CH-N); 7.84-7.78 (m, 4H, Pht); 6.81 (dd, 1H, *J*=11.9, 8.9 Hz, H-3); 6.49 (d, 1H, *J*=5.1 Hz, H-1); 5.25(dd, 1H, *J*=10.2, 9.0 Hz, H-4); 5.16 (bs, 1H, CH$_2$O*H*); 4.97 (dd, 1H, *J*=11.9, 5.1 Hz, H-2); 4.83 (ddd, 1H, *J*=10.2, 4.9, 2.2 Hz, H-5); 4.45 (d, 1H, *J*=13.4 Hz, C*H$_A$*OH); 4.40 (d, 1H, *J*=13.4 Hz, C*H$_B$*OH); 4.25 (dd, 1H, *J*=12.5, 4.9 Hz, H-6$_A$); 4.03 (dd, 1H, *J*=12.5, 2.2 Hz, H-6$_B$); 2.01 (2s, 6H, 2 Ac); 1.77 (s, 3H, Ac). ^{13}C-NMR (101 MHz; DMSO-*d$_6$*): δ= 170.1 (CO); 169.6 (CO); 168.8 (CO); 166.4 (2 NC=O); 147.1 (=CN); 135.1 (CHPht); 130.4 (CPht); 125.0 (=CHN); 123.6 (CHPht); 81.6 (CH-1); 72.4 (CH); 69.7 (CH); 66.5 (CH); 62.1 (CH$_2$); 54.8 (CH$_2$); 52.0 (CH); 20.6 (CH$_3$); 20.5 (CH$_3$); 20.4 (CH$_3$). Anal Calcd for C$_{23}$H$_{24}$N$_4$O$_{10}$: C, 53.49; H, 4.68; N, 10.85. Found: C, 53.25; H, 7.93; N, 11.02.
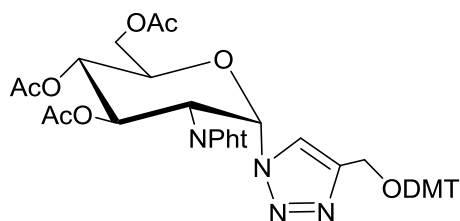
### 3,4,6-tri-*O*-acetyl-GlcNPht α-triazole-*O*-DMT 42

To a solution of **41** (158 mg, 0.3 mmol) in dry pyridine (1.6 mL) 4,4'-dimethoxytrityl chloride (124 mg, 0.36 mmol) was added and the reaction was stirred at r.t. until TLC (AcOEt/cyclohexane 7:3) showed complete disappearance of the starting reagent. After 18 h the mixture was concentrated under reduced pressure and the crude product was purified by flash chromatography on silica gel using cyclohexane/AcEtO/Et$_3$N (50:50:0.1). Compound **42** (220 mg, 88% yield) was obtain as a wax solid. [α]D^{25} ≅ +124 (*c* 0.5, DMSO). ^{1}H-NMR (400 MHz; CD$_3$OD): δ= 7.93 (s, 1H, C=CHN); 7.76-7.70 (m, 4H, Pht); 7.34-7.31 (m, 2H, DMT); 7.23-7.16 (m, 7H, DMT); 6.94 (dd, 1H, *J*=11.7 Hz, 9,0 Hz, H-4); 6.80-6.75 (m, 4H, DMT); 6.42 (d, 1H, *J*=5.3 Hz, H-1); 5.31 (dd, 1H, *J*=10.4, 9.0 Hz, H-3); 5.09-5.04 (m, 2H, H-2 and H-5); 4.39 (dd, 1H, *J*=12.6, 4.2 Hz, H-6$_A$); 4.16 (dd, 1H, *J*=12.6, 2.2 Hz, H-6$_B$); 4.15 (d, 1H, *J*=11.5 Hz, C*H$_A$*ODMT); 4.09 (d, 1H, *J*=11.5 Hz, C*H$_B$*ODMT); 3.76, 3.75 (2s, 6H, 2 OMe); 2.06 (s, 6H, 2 Ac); 1.82 (s, 3H, Ac). ^{13}C-NMR (101 MHz; CD$_3$OD): δ= 172.3 (Ac); 171.6 (Ac); 171.0 (Ac); 168.2 (2 C=O Pht); 160.2 (2 C-OMe); 146.3 (C-N); 146.1 (C DMT); 137.0 (C DMT); 136.9 (C DMT); 136.0 (2 CH Pht); 132.2 (2 C Pht); 131.2 (2 CH Pht, 4CH DMT); 129.2 (2 CH DMT); 128.8 (2 CH DMT); 127.9 (CH DMT); 126.2 (CHN); 124.6 (2 CH Pht); 114.1 (4 CH DMT); 88.2 (C DMT); 84.3 (CH-1); 74.6 (CH-5); 71.3 (CH-4); 68.4 (CH-3); 63.2 (CH$_2$-6); 58.7 (CH$_2$-O); 55.7 (CH$_3$-O DMT); 54.3 (CH$_3$-O DMT); 20.7 (CH$_3$); 20.6 (2 CH$_3$). HRMS Calcd for [C$_{44}$H$_{41}$N$_4$O$_{12}$ + H]$^+$ 841.2691, Found 841.2699.
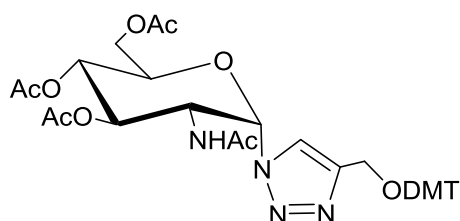
3,4,6-tri-*O*-acetyl-GlcNAc α-triazole-*O*-DMT 43

To a solution of **42** (345 mg, 0,42 mmol) in 8 ml of MeOH, a solution of MeNH$_2$ 35% (8 mL) was added and the reaction was stirred at 70°C for 18 h. The

solution was concentrated under reduced pressure and the solid obtained coevaporated several times with toluene. To the crude compound dissolved in pyridine (4 ml), Ac_2O (0.32 ml, 3.37 mmol) was added and the reaction was stirred at r.t for 20 h. The solvents were removed *in vacuo*, the crude material was diluted with saturated $NaHCO_3$ (20 mL) and then extracted with AcOEt (3 x 10 mL). The combined organic extracts were dried with anhydrous Na_2SO_4 and concentrated under reduced pressure. Compound **43** (340 mg, 89% yield) was obtained as a sticky oil pure enough to be used in the next step without further purification. ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 8.42 (s, 1H, C=CH-N); 7.96 (d, 1H, J=8.0 Hz, NH); 7.45-7.42 (m, 2H, DMT); 7.33-7.29 (m, 6H, DMT); 7.25-7.21 (m, 1H, DMT); 6.92-6.88 (m, 4H, DMT); 6.38 (d, 1H, J=5.9 Hz, H-1); 5.98 (dd, 1H, J=11.2, 9.0 Hz, H-3); 5.10 (dd, 1H, J=10.1, 9.0 Hz, H-4); 4.64-4.58 (m, 1H, H-2); 4.15 (dd, 1H, J=12, 4.4 Hz, H-6$_A$); 4.12-4.09 (m, 3H, H-5 and CH_2ODMT); 3.95 (dd, 1H, J=12 Hz, 1.7 Hz, H-6$_B$); 3.73 (s, 6H, 2 OMe); 1.98, 1.97, 1.93, 1.70 (4s, 12H, 4 Ac). ^{13}C-NMR (101 MHz; DMSO-d_6): δ= 170.4 (Ac); 170.0 (Ac); 169.7 (Ac); 169.4 (Ac); 158.2 (2 C-OMe); 144.7 (C-N); 144.1 (C DMT); 135.4 (2C DMT); 129.7 (4 CH DMT); 128.0 (2 CH DMT); 127.6 (2 CH DMT); 126.8 (CH DMT); 126.1 (CH-N); 113.3 (4 CH DMT); 86.1(C DMT); 82.0 (CH-1); 70.6 (CH-5); 70.1 (CH-4); 68.8 (CH-3); 61.6 (CH$_2$-6); 57.7 (CH$_2$-O); 55.1 (CH$_3$-O); 54.9 (CH$_3$-O); 50.1 (CH-2); 22.0 (CH$_3$) 20.5 (3 CH$_3$). Anal Calcd for $C_{38}H_{42}N_4O_{11}$: C, 62.46; H, 5.79; N, 7.67. Found: C, 62.72; H, 6.02; N, 7.98.
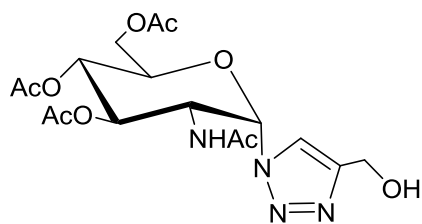
### 2,3,4,6-tri-*O*-acetyl-GlcNAc α-triazole **31d**

To a solution of **43**, obtained from the previous step, in MeOH (10 mL) AcOH (10 ml, 80% aq. solution) was added and the reaction was stirred at 50° for 1 h. The disappearence of the starting material was monitored by TLC (AcOEt/cycloesane/Et$_3$N 50:50:0,1). Then the solution was neutralized with NH$_4$OH (38% aq. soln) and evaporated under reduced pressure. The reaction mixture was poured into water (15 mL) and extracted three times with AcOEt (5 mL). The combined organic extracts were dried with anhydrous Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by flash chromatography starting with AcOEt 100 up to AcOEt/MeOH 90:10. Compound **31d** (136 mg, 75% yield) was obtained as a white solid, mp=171-174°C. $[\alpha]_D^{25} \cong +109$ (*c* 0.3, DMSO). ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 8.16 (s, 1H, C=CH-N); 7.92 (d, 1H, J=8.0 Hz, NH); 6.37 (d, 1H, J=5.9 Hz, H-1); 5.98 (dd, 1H, J=11.3, 9.1 Hz, H-3); 5.32-5.25 (m, 1H, CH$_2$OH); 5.08 (dd, 1H, J=10.3, 9.1 Hz, H-4); 4.62-4.54 (m, 3H, H-2 and CH_2OH); 4.14 (dd, 1H, J=12.5, 4,3 Hz, H-6$_A$); 4.04 (ddd, 1H, J=10.3, 4.3, 2.3 Hz, H-5); 3.91 (dd, 1H, J=12.5, 2.3 Hz, H-6$_B$); 1.98, 1.97, 1.93, 1.68 (4s, 12H, 4 Ac). ^{13}C-NMR (101 MHz; DMSO-d$_6$): δ= 170.4 (Ac); 170.0 (Ac); 169.7 (Ac); 169.4 (Ac); 147.76 (CN); 125.5 (CH-N); 81.8 (CH-1); 70.4 (CH); 70.2 (CH); 68.8 (CH-3); 61.6 (CH$_2$-6); 54.9 (CH$_2$-OH); 50.0 (CH-2); 22.0 (CH$_3$); 20.5 (3 CH$_3$). HRMS calculated for $[C_{17}H_{24}N_4O_9 + H]^+$ 429.1616, found 429.1624.

8.4.5 Synthesis of glycosyl uridine

8.4.5.1 General procedure for couplings

To a solution of phosphoramidite **32** (1 eq) in dry CH$_3$CN (3 mL) a solution of 4,5-dicyanoimidazole (DCI, 2 eq) and 1.3 eq of the approriate glycosyl triazole **31a-d** in dry CH$_3$CN (3 mL) was added *via cannula* under argon atmosphere. The mixture was stirred at

r.t. for 45 min and then a solution of *tert*-butylhydroperoxide in decane (5.5 M, 4 eq) was added. After 45 min stirring, the reaction was concentrated under reduced pressure and the residue dissolved in CH_2Cl_2 and washed with H_2O. The organic layer was dried with anhydrous Na_2SO_4 and concentrated *in vacuo*. To the crude product, dissolved in the minimum quantity of CH_3CN, a solution of diethylamine (20% in CH_3CN, 42 eq) was added and the reaction stirred at r. t. monitoring by TLC (AcOEt/MeOH/NH_4OH 70:30:0.3). The mixture was concentrated under reduced pressure and the residue filtered on a short pad of silica gel (AcOEt/MeOH/NH_4OH 70:30:0.3) affording the adducts **33a-d** which were pure enough to be employed in the step of deacetylation.
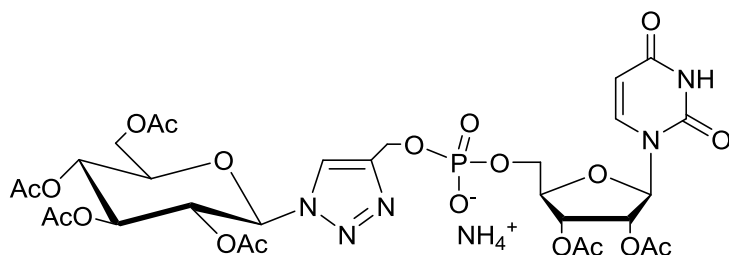
2,3,4,6-tetra-*O*-acetyl-Glc 2',3'-di-*O*-acetyl-uridine β 33a

"General procedure for couplings" applied to **31a** yields **33a** (110 mg, 37%, 3 steps) as a wax solid. ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 11.40 (s, 1H, NH); 8.31 (s, 1H, C=CHN); 8.05-7.92 (m, 1H, H-6); 6.32 (d, 1H, J=9.18 Hz, H-1g); 6.06-5.97 (m, 1H, H-1'); 5.69- 5.58 (m, 1H, H-2g); 5.53 (t, 1H, J=9.37 Hz, H-3g); 5.40-5.27 (m, 2H, H-2', H-5); 5.16 (t, 1H, J=9.96 Hz, H-4g); 4.81-4.70 (m, 2H, CH$_2$OP); 4.37-4.30 (m, 1H, H-5g); 4.27-4.21 (m, 1H, H-3'); 4.15-4.09 (m, 1H, H-4'); 4.08-4.01 (m, 2H, H-6g); 3,92-3,76 (m); 3.14 (d, 2H, J=4.69 Hz, H-5'$_A$, H-5'$_B$); 2.17-1.87 (m, 15H, 5 Ac); 1.78 (s, 3H, Ac). ^{13}C-NMR (101 MHz; DMSO-$d6$): δ= 170.1 (Ac); 169.6 (Ac); 169.4 (Ac); 168.6 (Ac); 163.0 (C=O); 150.7 (C=O); 146.0 (=C-N); 140.7 (CH-5u); 122.5 (=CH-N); 102.7 (CH-6u); 85.1 (CH-1'); 83.8 (CH-4'); 81.8 (CH-1g); 73.2 (CH); 72.2 (CH); 71.4 (CH); 70.1 (CH); 67.5 (CH); 64.0 (CH$_2$); 61.8 (CH$_2$); 57.9 (CH$_2$); 20.5 (2 CH$_3$); 20.3 (2 CH$_3$); 20.2 (CH$_3$); 19.9 (CH$_3$). ^{31}P-NMR (122 MHz; DMSO-$d6$): δ= -1,6. HRMS Calcd for [$C_{30}H_{38}N_5O_{20}P$ + H]$^{+}$ 820.1921, Found 820.1946.
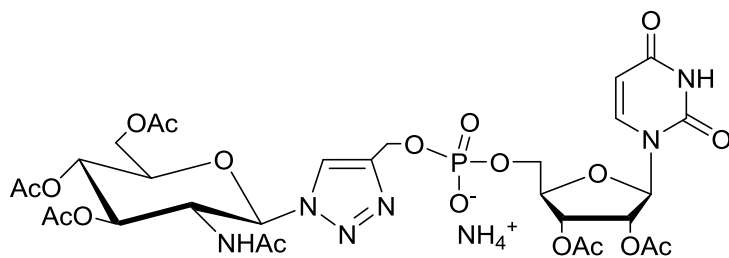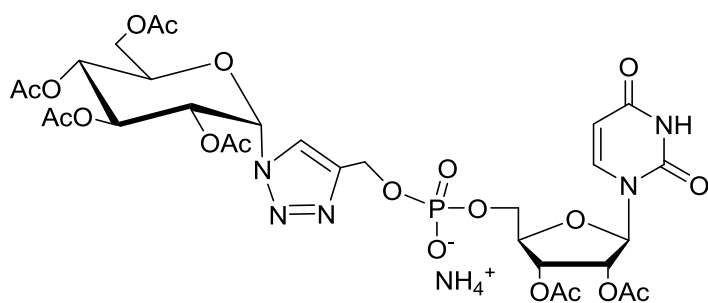

3,4,6-tri-*O*-acetyl-GlcNAc 2',3'-di-*O*-acetyl-uridine β 33b

"General procedure for couplings" applied to **31b** yields **33b** (122 mg, 36%, 3 steps) as a wax solid. ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 11.40 (bs, 1H, NH); 8.26 (s, 1H, C=CH-N); 8.02 (d, 1H, J=8.0 Hz, H-6); 6.58 (d, 1H, J=6.0 Hz, H-1g); 6.10 (dd, 1H, J=10.2, 9,4 Hz, H-3g); 6.02 (d, 1H, J=5.9 Hz, H-1'); 5.62 (d, 1H, J=8.0 Hz, H-5); 5.42 (dd, 1H, J=10.2, 6.0 Hz, H-2g); 5.38-5.33 (m, 2H, H-2' and H-3'); 5.20 (dd, 1H, J=10.2, 9.4 Hz, H-4g); 4.81 (bs, 2H, C*H$_2$*OP); 4.26-4.22 (m, 2H, H-5g and H-4'); 4.14 (dd, 1H, J=12.6, 4.4 Hz, H-6$_A$g); 3.93 (dd, 1H, J=12.6, 2.2 Hz, H-6$_B$g); 3.90-3.84 (m, 2H, H-5'$_A$ e H-5'$_B$); 2.08, 2.00, 1.99, 1.96, 1.95, 1.80 (6s, 18H, 6 Ac). ^{13}C-NMR (101 MHz; DMSO-d_6): δ= 170.0 (Ac); 169.6 (Ac); 169.5 (Ac); 169.3 (3Ac); 162.9 (C=O); 150.6 (C=O); 144.7 (=C-N); 140.7 (CH-5u); 126.4 (=CH-N); 102.6 (CH-6u); 85.3 (CH-1'); 81.7 (CH-4'); 80.5 (CH-1g); 72.1 (CH); 71.2 (CH); 70.8 (CH); 69.9 (CH); 68.1 (CH); 67.8 (CH); 64.0 (CH$_2$); 61.5 (CH$_2$); 57.8 (CH$_2$); 20.5 (2 CH$_3$); 20.4 (2 CH$_3$); 20.2 (CH$_3$); 20.1 (CH$_3$). ^{31}P-NMR

(122 MHz; DMSO-d_6): δ= -2,07. HRMS Calcd for $[C_{30}H_{39}N_6O_{19}P + H]^+$ 819.2081, Found 819.2046.

2,3,4,6-tetra-*O*-acetyl-Glc 2′,3′-di-*O*-acetyl-uridine α 33c

"General procedure for couplings" applied to **31c** yields **33c** (134 mg, 54%, 3 steps) as a wax solid. ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 11.40 (bs, 1H, NH); 8.20 (s, 1H, C=CH-N); 8.15 (d, 1H, *J*= 9.2 Hz, N*H*Ac); 8.02 (d, 1H, *J*=8.1 Hz, H-6); 7.48-7.05 (m, 4H, NH$_4^+$); 6.10 (d, 1H, *J*=10.0 Hz, H-1g); 6.05-6.01 (m, 1H, H-1'); 5.64 (d, 1H, *J*=8.1 Hz, H-5); 5.39-5.32 (m, 3H, H-2', H-3', H-3g); 5.07 (t, 1H, *J*=9.7 Hz, H-4g); 4.76 (d, 1H, *J*=12.8 Hz, C*H$_A$*OP); 4.71 (d, 1H, *J*=12.8 Hz, C*H$_B$*OP); 4.60-4.51 (m, 1H, H-2g); 4.26-4.23 (m, 1H, H-4'); 4.20 (ddd, 1H, *J*=10.0, 5.1, 2.1 Hz, H-5g); 4.13 (dd, 1H, *J*=12.4, 5.1 Hz, H-6$_A$g); 4.02 (dd, 1H, *J*=12.4, 2.1 Hz, H-6$_B$g); 3.92-3.81 (m, 2H, H-5'$_A$, H-5'$_B$); 2.09, 2.00, 1.99, 1.97, 1.92, 1.58 (6s, 18H, 6 Ac). ^{13}C-NMR (101 MHz; DMSO-d_6): δ=170.1 (Ac); 169.6 (3 Ac); 169.4 (3 Ac); 163.0 (C=O); 150.7 (C=O); 145.3 (=C-N); 140.7 (CH-5u); 122.4 (=CH-N); 102.7 (CH-6u); 85.2 (CH-1'); 84.6 (CH-4'); 81.8 (CH-1g); 73.3 (CH); 72.4 (CH); 72.2 (CH); 71.3 (CH); 68.0 (CH); 64.1 (CH$_2$); 61.8 (CH$_2$); 57.9 (CH$_2$); 52.1 (CH-2g); 22.4 (CH$_3$); 20.5 (CH$_3$); 20.4 (2 CH$_3$); 20.3 (CH$_3$); 20.2 (CH$_3$). ^{31}P-NMR (122 MHz; DMSO-d_6): δ= -1,67. HRMS Calcd for $[C_{30}H_{38}N_5O_{20}P + H]^+$ 820.1921, Found 820.1939.
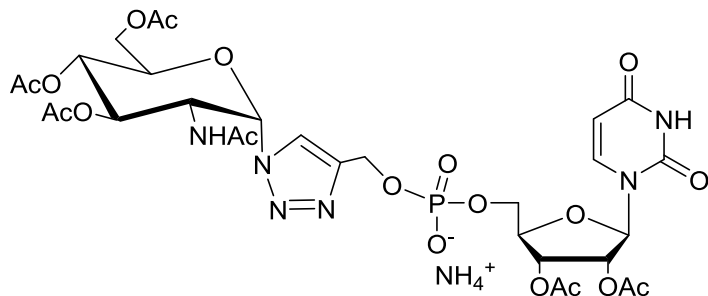

3,4,6-tri-*O*-acetyl-GlcNAc 2′,3′-di-*O*-acetyl-uridine α 33d

"General procedure for couplings" applied to **31d** yields **33d** (104 mg, 43%, 3 steps) as a wax solid. ^{1}H-NMR (400 MHz; DMSO-*d6*): δ= 8.22 (s, 1H, C=CH-N); 8.05 (d, 1H, *J*=8.01 Hz, H-6); 7.90 (d, 1H, *J*=8.1 Hz, N*H*Ac); 7.66-6.90 (m, 4H, NH$_4^+$); 6.37 (d, 1H, *J*=5.7 Hz, H-1g); 6,02 (d, 1H, *J*=6.0, H-1'); 5.98 (dd, 1H, *J*=11.3, 9.2 Hz, H-3g); 5.60 (d, 1H, *J*=8.1 Hz, H-5); 5.38-5.32 (m, 2H, H-2' and H-3'); 5.08 (dd, 1H, *J*=10.2, 9.2 Hz, H-4g); 4.78 (d, 2H, *J*=6.2 Hz, C*H$_2$*OP); 4.59 (ddd, 1H, *J*=11.3, 8.1, 5.7 Hz, H-2g); 4.24-4.21 (m, 1H, H-4'); 4.14 (dd, 1H, *J*=12.4, 4,2 Hz, H-6$_A$g); 4.07-4.03 (m, 1H, H-5g); 3.90 (dd, 1H, *J*=12.4, 2.1 Hz, H-6$_B$g); 3.87-3.80 (m, 2H, H-5'$_A$ and H-5'$_B$); 2.09, 1.99, 1.98, 1.97, 1.92, 1.68 (6s, 18H, 6 Ac). ^{13}C-NMR (101 MHz; DMSO-*d6*): δ= 170.4 (Ac); 170.0 (Ac); 169.7 (Ac); 169.5 (Ac); 169.3 (Ac); 162.9 (C=O); 150.6 (C=O); 145.0 (=C-N); 140.8 (CH-5u); 126.3 (=CH-N); 102.6 (CH-6u); 85.2 (CH-1'); 81.9 (CH-4', CH-1g); 72.2 (CH); 71.4 (CH); 70.4 (CH); 70.2 (CH); 68.7 (CH); 64.0 (CH$_2$); 61.5 (CH$_2$); 57.8 (CH$_2$); 50.0 (CH-2g); 24.1 (CH$_3$); 22.0 (CH$_3$); 20.4 (2 CH$_3$); 20.2 (CH$_3$). ^{31}P-NMR (122 MHz; DMSO-*d6*): δ= -1,51. HRMS Calcd for $[C_{30}H_{39}N_6O_{19}P + H]^+$ 819.2081, Found 819.2106.
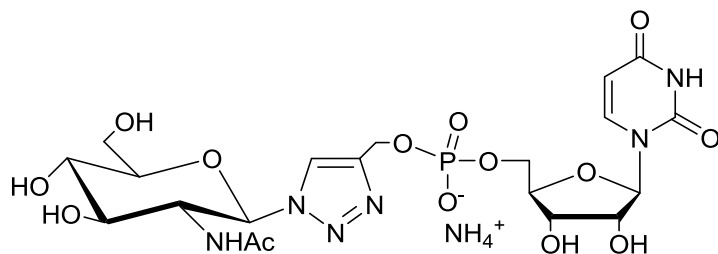
8.4.5.2 General procedure for deacetylations

Compounds **33a-d** were dissolved in CH_3CN and treated with NH_4OH solution (38% NH_3 in H_2O). After 18 h stirring at r.t., the mixture was concentrated under reduced pressure, the residue dissolved in H_2O and CH_2Cl_2. The aqueous phase was concentrated *in vacuo* to give the corresponding coupling products **17-20**.
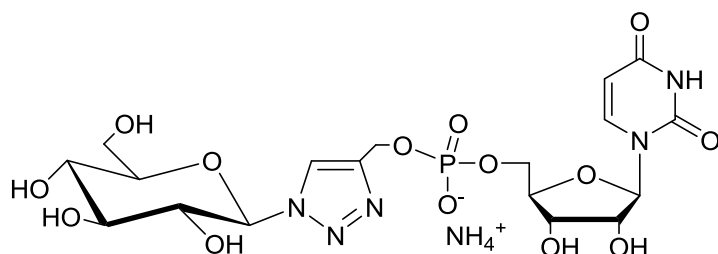
Glc uridine β 17

"General procedure for deacetylations" applied to **33a** yields **17** (80 mg, 95% yield) as a sticky oil. $[\alpha]_D^{25} \cong -2$ (*c* 1.2, H_2O). ^{1}H-NMR (400 MHz; DMSO-d_6):

δ= 8.23 (s, 1H, C=CH-N); 7.89 (d, 1H, *J*=8.01 Hz, H-6); 7.31 (s, 3H, 3OH); 6.69 (s, 3H, 3OH); 5.79 (d, 1H, *J*=5.66 Hz, H-1'); 5.62 (d, 1H, *J*=8.01 Hz, H-5); 5.50 (d, 1H, *J*=9.37 Hz, H-1g); 4.76 (d, 2H, *J*=6.05 Hz, CH₂OP); 4.07 (t, 1H, 5.27 Hz, H-2'); 4.01 (t, 1H, *J*=3.51 Hz, H-3'); 3.94 (t, 1H, *J*=2.15 Hz, H-4'); 3.90-3.80 (m, 2H, H-5'); 3.75 (t, 1H, *J*=9.18 Hz, H-2g); 3.65 (d, 2H, *J*=10.15 Hz, H-6g); 3.46-3.37 (m, 2H, H-3g, H-5g); 3.25-3.18 (m, 1H, H-4g); ^{13}C-NMR (101 MHz; DMSO-d_6): δ= 163.27 (C=O); 150.93 (C=O); 144.81 (CH-5u); 140.96 (=C-N); 123.05 (=CH-N); 102.03 (CH-6u); 87.50 (CH-1', CH-4'); 83.62 (CH-1g); 79.99 (CH); 77.04 (CH); 73.33 (CH); 72.10 (CH); 70.44 (CH); 69.60 (CH); 64.39 (CH₂); 60.80 (CH₂); 58.00 (CH₂). ^{31}P-NMR (122 MHz; DMSO-d_6): δ= -1.79. HRMS Calcd for $[C_{18}H_{26}N_5O_{14}P + H]^+$ 568.1286, Found 568.1288.

GlcNAc uridine β 18

"General procedure for deacetylations" applied to **33b** yields **18** (83 mg, 97% yield) as a sticky oil. $[\alpha]_D^{25} \cong +42$ (*c* 1.6, H_2O). ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 8.15 (s, 1H, C=CH-N); 7.89 (d, 1H, *J*=8.1 Hz, H-6); 7.30 (br s, 4H, NH_4^+); 6.12 (d, 1H, *J*=5.9 Hz, H-1g); 5.79 (d, 1H, *J*=5.7 Hz, H-1'); 5.61 (d, 1H, *J*=8.1 Hz, H-5); 5.57 (d, 1H, *J*=4.9 Hz, OH); 5.49 (d, 1H, *J*=5.5 Hz, OH); 5.48-5.43 (m, 1H, OH); 5.21 (d, 1H, *J*=4.8 Hz, OH); 5.18 (d, 1H, *J*=5.9 Hz, OH); 4.81-4.72 (m, 2H, C*H₂*OP); 4.56 (t, 1H, *J*=5.8 Hz, OH); 4.10-4.05 (m, 1H, H-2'); 4.05-4.00 (m, 2H, H-3', H-3g); 3.96-3.92 (m, 1H, H-4'); 3.89-3.78 (m, 2H, H-5$_A$' e H-5$_B$'); 3.76-3.69 (m, 1H, H-2g); 3.62 (ddd, 1H, *J*=9.9, 5.3, 2.0 Hz, H-5g); 3.58-3.52 (m, 1H, H-6$_A$g); 3.47-3.41 (m, 1H, H-6$_B$g); 3.28-3.21 (m, 1H, H-4g). ^{13}C-NMR (101 MHz; DMSO-d_6): δ= 163,3 (C=O); 151,0 (C=O); 143,9 (=C-N); 141,0 (CH-5u); 126,0 (=CH-N); 102.0 (CH-6u); 87.5 (CH-1'); 85.2 (CH-4'); 83.7 (CH-1g); 76.5 (CH); 73.4 (CH); 72.9 (CH); 70.6 (CH); 70.4 (CH); 69.9 (CH); 64.4 (CH₂); 60.8 (CH₂); 58.0 (CH₂). ^{31}P-NMR (122 MHz; DMSO-d_6): δ= -1.68. HRMS Calcd for $[C_{18}H_{26}N_5O_{14}P + H]^+$ 568.1286, Found 568.1289.
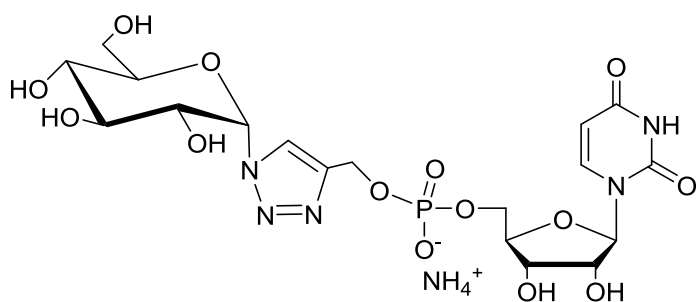
Glc uridine α 19

"General procedure for deacetylations" applied to **33c** yields **19** (97 mg, 97% yield) as a sticky oil. $[\alpha]_D^{25} \cong -6$ (*c* 2.7, H_2O). ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 8.08 (s, 1H, C=CH-N); 7.98 (d, 1H, *J*=9.1 Hz, N*H*Ac); 7.90 (d, 1H, *J*=8.1 Hz, H-6); 7.30 (br s, 4H, NH_4^+); 5.79 (d, 1H, *J*=5.8 Hz, H-1g); 5.70 (d, 1H, *J*=9.9 Hz, H-1'); 5.60 (d, 1H, *J*=8.1 Hz, H-5); 5.53-5.30 (m, 4H, 4 OH); 4.82-4.71 (m, 1H, OH); 4.76-4.66 (m, 2H, C*H₂*OP); 4.09-3.99 (m, 3H, H-2', H-3', H-2g); 3.95-3.91 (m, 1H, H-4'); 3.87-3.76 (m, 2H, H-5$_A$' e H-5$_B$'); 3.67 (d, 1H, J=10.8 Hz, H-6$_A$g); 3.55 (t, 1H, J=9,2 Hz, H-3g); 3.50-3.36 (m, 2H, H-5g e H- 6$_B$g); 3.27 (t, 1H, J=9.1 Hz, H-4g); 1.62 (s, 3H, Ac). ^{13}C-NMR (101 MHz; DMSO-d_6): δ= 169.5 (C); 163.2 (C=O); 150.9 (C=O); 145.0 (CH-5u); 141.0 (=C-N); 122.3 (=CH-N); 102.0 (CH-6u); 87.5 (CH); 86.0 (CH); 83.6 (CH); 80.1 (CH); 74.0 (CH); 73.4 (CH); 70.5 (CH); 70.0 (CH); 64.4 (CH_2); 60.8 (CH_2); 58.0 (CH_2); 54.6 (CH-2g); 22.8 (CH_3). ^{31}PNMR (122 MHz; DMSO-d_6): δ= -1.32. HRMS Calcd for $[C_{20}H_{29}N_6O_{14}P + H]^+$ 609.1552, Found 609.1556.
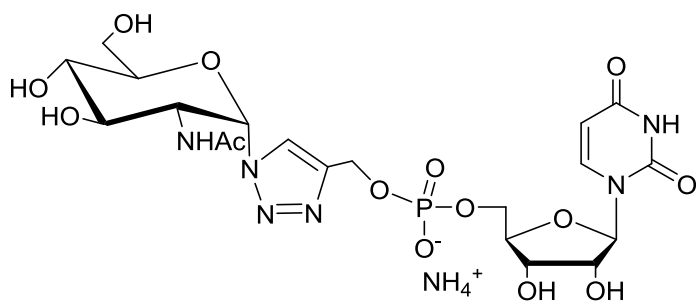
GlcNAc uridine α 20

"General procedure for deacetylations" applied to **33d** yields **20** (74 mg, 95% yield) as a sticky oil. $[\alpha]_D^{25} \cong +70$ (*c* 2.4, H_2O). ^{1}H-NMR (400 MHz; DMSO-d_6): δ= 8.16 (s, 1H, C=CHN); 7.90 (d, 1H, *J*=8.1 Hz, H-6); 7.82 (d, 1H, *J*=6.9 Hz, N*H*Ac); 7.30 (bs, 4H, NH_4^+); 6.22 (d, 1H, *J*=5,8 Hz, H-1g); 5.78 (d, 1H, *J*=5.7 Hz, H-1'); 5.58 (d, 1H, *J*=8.1 Hz, H-5); 4.80-4.71 (m, 2H, C*H₂*OP); 4.25 (dd, 1H, *J*=10.8, 8.5 Hz, H-3g); 4.10-4.03 (m, 2H, H-2', H-2g); 4.03-3.98 (m, 1H, H-3'); 3.95-3.90 (m, 1H, H-4'); 3.88-3.76 (m, 2H, H-5$_A$' e H-5$_B$'); 3.53 (dd, 1H, H=14, 4.7 Hz, H-6$_A$g); 3.47-3.40 (m, 2H, H-5g e H-6$_B$g); 3.32 (t, 1H, *J*=8.9 Hz, H-4g); 1.67 (s, 3H, Ac). ^{13}C-NMR (101 MHz; DMSO-d_6): δ= 170.4 (Ac); 163.2 (C=O); 150.9 (C=O); 144.4 (=C-N); 141.0 (CH-5u); 125.8 (=CH-N); 102.0 (CH-6u); 87.5(CH-1'); 83.6 (CH-4'); 83.0 (CH-1g); 76.2 (CH); 73.4 (CH); 73.0 (CH); 70.5 (CH); 70.2 (CH); 64.4 (CH_2); 60.6 (CH_2); 57.9 (CH_2); 53.4 (CH-2g); 22.3 (CH_3). ^{31}P-NMR (122 MHz; DMSO-d_6): δ= -1.35. HRMS Calcd for $[C_{20}H_{29}N_6O_{14}P + H]^+$ 609.1552, found 609.1556.

8.4.6 Binding assays

A truncated construct of the human *O*-GlcNAc transferase (312-1031) was recombinantly expressed and purified as described before [76,35]. Displacement of the fluorescein labeled bisubstrate conjugate (Figure 8.8) by compounds **7-20** was measured *via* fluorescence polarimetry [69].
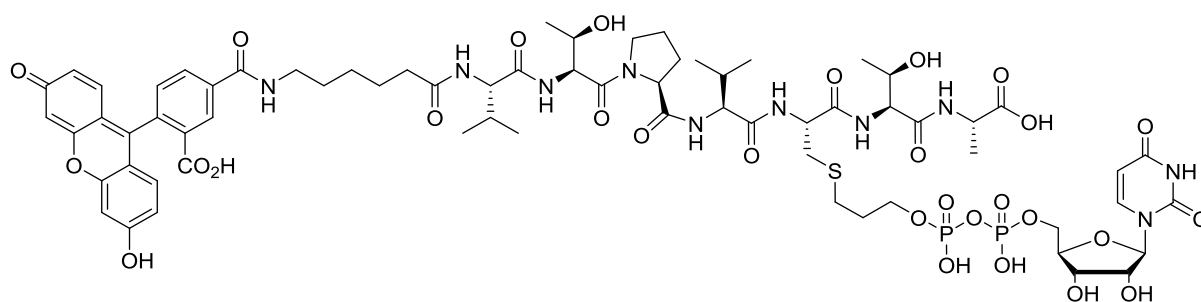
Figure 8.8 *5(6)-fluorescein carboxamide (Floc) labelled derivative used as a probe for fluorescence measurements* [69].

Assay conditions contained 0.8 μM hOGT, 0.75 μM fluorescent probe, 0.1 M Tris-HCl pH 7.4, 0.15 M NaCl, 0.5 mM TCEP and 5% DMSO and varying amounts of compounds **7-20**. Reaction mixtures were incubated in the dark for 30 minutes. Dose-dependent reduction of fluorescent polarimetry was measured on a *PHERAstar* multi-mode plate reader (*BMG Labtech*). Apparent IC_{50} values were calculated using a 4-point non-linear regression curve fit (*GraphPad Prism*). Due to the low binding affinity, a total displacement was simulated by introducing a 0-point at 10 M concentration, allowing calculation of apparent IC_{50} values. Apparent K_i values were calculated using the equation reported by Nikolovska-Coleska [83].

Compound	IC_{50} (mM)	K_i (μM)
7	1.43	920.7 ± 44
8	9.25	4769.0 ± 530
9	8.43	5452.0 ± 68
10	1.1	718.0 ± 15
11	3.37	2218.0 ± 268
12	1.7	1140.0 ± 98
13	2.7	1804.0 ± 168
14	0.18	102.5 ± 8
15	2.8	1814.0 ± 74
16	3.44	2262.4 ± 243
17	2.23	1439.5 ± 59
18	0.38	231.9 ± 16
19	2.28	1048.0 ± 38
20	0.88	589.1 ± 19

Table 8.2 *Apparent binding affinities of compounds **7-20** towards the human O-GlcNAc transferase.* K_i *values were calculated by introducing a fictive 100% displacement at 10 M.*

8.5 Bibliography

1. Kellokumpu, S.; Hassinen, A.; Glumoff, T. *Cell. Mol. Life Sci.* **2016,** *73,* 305-325.

2. Sun, X.-L. *Med. Chem.* **2013,** *3,* e106.

3. Voglmeir, J.; Flitsch, S. L. *Sci. Synth. Biocatal. Org. Synth.* **2015,** *1,* 507-542.

4. Liang, D.-M.; Liu, J.-H.; Wu, H.; Wang, B.-B.; Zhu, H.-L.; Qiao, J.-J. *Chem. Soc. Rev.* **2015,** *44,* 8350-8374.

5. Palcic, M. M. *Curr. Opin. Chem. Biol.* **2011,** *15,* 226-233.

6. Horton, D. *Carbohydrate Chemistry, Biology and Medical Applications;* Elsevier: Oxford, 2008; pp 1-28.

7. Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S. *Nat. Chem. Biol.* **2011,** *7,* 685-691.

8. Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I.-K.; Li, L.; Thorson, J. S. *Science* **2006,** *313,* 1291-1294.

9. Etzler, M. E.; Varki, A.; Cummings, R. L.; Esko, J. D.; Freeze, H. H.; Hart, G. W. *Essentials of Glycobiology,* 2nd ed.; Cold Spring Harbor Laboratory Press: New York, 2008.

10. Ma, Z.; Vosseller, K. *Amino Acids* **2013,** *45,* 719-733.

11. Lynch, T. P.; Reginato, M. J. *Cell Cycle* **2011,** *10,* 1712-1713.

12. Zhu, Y.; Shan, X.; Yuzwa, S. A.; Vocadlo, D. J. *J. Biol. Chem.* **2014,** *289,* 34472-34481.

13. Yuzwa, S. A.; Vocadlo, D. J. *Chem. Soc. Rev.* **2014,** *43,* 6839-6858.

14. Wang, S.; Vidal, S. *Carbohydr. Chem.* **2013,** *39,* 78-101.

15. Roychoudhury, R.; Pohl, N. L. B. *Curr. Opin. Chem. Biol.* **2010,** *14,* 168-173.

16. Schutzbach, J.; Brockhausen, I. *Methods in Molecular Biology* **2009,** *534,* 359-373.

17. Walker-Nasir, E.; Ahmad, I.; Saleem, M.; Hoessli, D. C. *Curr. Org. Chem.* **2007,** *11,* 591-607.

18. Jung, K.-H.; Schmidt, R. R. *Carbohydr.-Based Drug Discov.* **2003,** *2,* 609-659.

19. Kajimoto, T.; Node, M. *Synthesis* **2009,** 3179-3210.

20. Qian, X.; Palcic, M. M. *Carbohydr. Chem. Biol.* **2000,** *3,* 293-312.

21. Whalen, L. J.; Greenberg, W. A.; Mitchell, M. L.; Wong, C.-H. *Iminosugars* **2007,** 153-176.

22. Cipolla, L.; La Ferla, B.; Gregori, M. *Comb. Chem. High Throughput Screening* **2006,** *9,* 571-582.

23. Compain, P.; Martin, O. R. *Bioorg. Med. Chem.* **2001,** *9,* 3077-3092.

24. Merino, P.; Delso, I.; Tejero, T.; Ghirardello, M.; Juste-Navarro, V. *Asian J. Org. Chem.* **2016,** *5,* 1413-1427.

25. Tedaldi, T.; Wagner, G. K. *MedChemComm* **2014,** *5,* 1106-1125.

26. Jiang, J.; Lazarus, M. B.; Pasquina, L.; Sliz, P.; Walker, S. *Nat. Chem. Biol.* **2012,** *8,* 72-77.

27. Smithen, D. A.; Forget, S. M.; McCormick, E.; Syvitski, R. T.; Jakeman, D. L. *Org. Biomol. Chem.* **2015,** *13,* 3347-3350.

28. Izumi, M.; Yuasa, H.; Hashimoto, H. *Curr. Top. Med. Chem.* **2009,** *9,* 87-105.

29. Wang, S.; Cuesta-Seijo, J. A.; Lafont, D.; Palcic, M. M.; Vidal, S. *Chem. Eur. J.* **2013,** *19,* 15346-15357.

30. Wang, S.; Cuesta-Seijo, J. A.; Striebeck, A.; Lafont, D.; Palcic, M. M.; Vidal, S. *ChemPlusChem* **2015,** *80,* 1525-1532.

31. Ghirardello, M.; de las Rivas, M.; Lacetera, A.; Delso, I.; Lira-Navarrete, E.; Tejero, T.; Martin-Santamaria, S.; Hurtado-Guerrero, R.; Merino, P. *Chem. Eur. J.* **2016,** *22,* 7215-7224.

32. Lira-Navarrete, E.; de las Rivas, M.; Companon, I.; Pallares, M. C.; Kong, Y.; Iglesias-Fernandez, J.; Bernardes, G. A. J. L.; Peregrina, J. M.; Rovina, C.; Bernado, P.; Bruscolini, P.; Clausen, H.; Lostao, A.; Corzana, F.; Hurtado-Guerrero, R. *Nat. Commun.* **2015,** *6,* 6937.

33. Hurtado-Guerrero, R. *Biochem. Soc. Trans.* **2016,** *44,* 61-67.

34. Madariaga, D.; Martinez-Saez, N.; Somovilla, V. J.; Garcia-Garcia, L.; Alvaro Berbis, M.; Valero-Gonzales, J.; Martin-Santamaria, S.; Hurtado-Guerrero, R.; Asensio, J. L.; Jimenez-Barbero, J.; Avenoza, A.; Busto, J. H.; Corzana, F.; Peregrina, J. M. *Chem. Eur. J.* **2014,** *20,* 12616-12627.

35. Pathak, S.; Alonso, J.; Schimpl, M.; Rafie, K.; Blair, D. E.; Borodkin, V. S.; Schuttelkopf, A. W.; Albarbarawi, O.; van Aalten, D. M. F. *Nat. Struct. Mol. Biol.* **2015,** *22,* 744-750.

36. Zhu, Y.; Liu, T.-W.; Cecioni, S.; Eskandari, R.; Zandberg, W. F.; Vocadlo, D. J. *Nat. Chem. Biol.* **2015,** *11,* 319-325.

37. Lazarus, M. B.; Jiang, J.; Gloster, T. M.; Zandberg, W. F.; Whitworth, G. E.; Vocadlo, D. J.; Walker, S. *Nat. Chem. Biol.* **2012,** *8,* 966-968.

38. Ostrowski, A.; van Aalten, D. M. F. *Biochem. J.* **2013,** *456,* 1-12.

39. Gloster, T. M.; Vocadlo, D. J. *Curr. Signal Transduction Ther.* **2010,** *5,* 74-91.

40. Gross, B. J.; Kraybill, B. C.; Walker, S. *J. Am. Chem. Soc.* **2005,** *127,* 14588-14589.

41. Dorfmueller, H. C.; Borodkin, V. S.; Blair, D. E.; Pathak, S.; Navratilova, I.; van Aalten, D. M. F. *Amino Acids* **2011,** *40,* 781-792.

42. Wang, S.; Shen, D. L.; Lafont, D.; Vercoutter-Edouart, A.-S.; Mortuaire, M.; Shi, Y.; Maniti, O.; Girard-Egrot, A.; Lefebvre, T.; Pinto, B. M.; Vocadlo, D. J.; Vidal, S. *MedChemComm* **2014,** *5,* 1172-1178.

43. Nagel, A. K.; Ball, L. E. *Amino Acids* **2014,** *46,* 2305-2316.

44. Schrodinger, Ed. *Schrodinger Release 2017-1;* LLC: New York, NY, 2017.

45. Trapannone, R.; Rafie, K.; van Aalten, D. M. F. *Biochem. Soc. Trans.* **2016,** *44,* 88-93.

46. Mitchell, M.; Qaio, L.; Wong, C.-H. *Adv. Synth. Catal.* **2001,** *343,* 596-599.

47. Burkart, M. D.; Vincent, S. P.; Duffels, A.; Murray, B. W.; Ley, S. V.; Wong, C. H. *Bioorg. Med. Chem.* **2000,** *8,* 1937-1946.

48. Merino, P.; Tejero, T.; Delso, I.; Hurtado-Guerrero, R.; Gomez-SanJuan, A.; Sadaba, D. *Mini-Rev. Med. Chem.* **2012,** *12,* 1455-1464.

49. Ardèvol, A.; Rovina, C. *J. Am. Chem. Soc.* **2015,** *137,* 7528-7547.

50. Yeoh, K. K.; Butters, T. D.; Wilkinson, B. L.; Fairbanks, A. J. *Carbohydr. Res.* **2009,** *344,* 586-591.

51. Horton, D.; Miyake, T. *Carbohydr. Res.* **1988,** *184,* 221-229.

52. Kramer, J. R.; Deming, T. J. *J. Am. Chem. Soc.* **2010,** *132,* 15068-15071.

53. Babic, A.; Gobec, S.; Gravier-Pelletier, C.; Le Merrer, Y.; Pecar, S. *Tetrahedron* **2008,** *64,* 9093-9100.

54. Cui, J.; Horton, D. *Carbohydr. Res.* **1998,** *309,* 319-330.

55. Damha, M. J.; Usman, N.; Ogilvie, K. K. *Can. J. Chem.* **1989,** *67,* 831-839.

56. Barthelemy, P.; Prata, C. A. H.; Filocamo, S. F.; Immoos, C. E.; Maynor, B. W.; Hashmi, S. A. N.; Lee, S. J.; Grinstaff, M. W. *Chem. Commun.* **2005,** 1261-1263.

57. Delso, I.; Tejero, T.; Goti, A.; Merino, P. *Tetrahedron* **2010,** *66,* 1220-1227.

58. Merino, P.; Delso, I.; Mannucci, V.; Tejero, T. *Tetrahedron Lett.* **2006,** *47,* 3311-3314.

59. Dondoni, A.; Staderini, S.; Marra, A. *Eur. J. Org. Chem.* **2013,** 5370-5375.

60. Finn, M. G.; Fokin, V. V. *Chem. Soc. Rev.* **2010,** *39,* 1231-1232.

61. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001,** *40* (11), 2004-2021.

62. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem.* **2001,** *113,* 2056-2075.

63. Wang, C.; Ikhlef, D.; Kahlal, S.; Saillard, J.-Y.; Astruc, D. *Coord. Chem. Rev.* **2016,** *316,* 1-20.

64. Liang, L.; Astruc, D. *Coord. Chem. Rev.* **2011,** *255,* 2933-2945.

65. Tiwari, V. K.; Mishra, B. B.; Mishra, K. B.; Mishra, N.; Singh, A. S.; Chen, X. *Chem. Rev.* **2016,** *116,* 3086-3240.

66. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981,** *22,* 1859-1862.

67. Caruthers, M. H.; Beaucage, S. L.; Becker, C.; Efcavitch, J. W.; Fisher, E. F.; Galluppi, G.; Goldman, R.; DeHaseth, P.; Matteucci, M. *Gene Amplif. Anal.* **1983,** *3,* 1-26.

68. Beaucage, S. L. *Methods Mol. Biol.* **1993,** *20,* 33-61.

69. Rafie, K.; Gorelik, A.; Trapannone, R.; Borodking, V.; van Aalten, D. M. F. *Bioconj. Chem.* **2018,** *29,* 1834-1840.

70. Schimpl, M.; Zheng, X.; Borodkin, V. S.; Blair, D. E.; Ferenbach, A. T.; Schuttelkopf, A. W.; Navratilova, I.; Aristotelous, T.; Albarbarawi, O.; Robinson, D. A.; Macnaughtan, M. A.; van Aalten, D. M. F. *Nat. Chem. Biol.* **2012,** *8,* 969-974.

71. Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed.* **1999,** *38,* 1784-1788.

72. Mayer, M.; Meyer, B. *Angew. Chem.* **1999,** *111,* 1902-1906.

73. del Carmen Fernandez-Alonso, M.; Diaz, D.; Berbis, M. A.; Marcelo, F.; Canada, J.; Jimenez-Barbero, J. *Curr. Protein Peptide Sci.* **2012,** *13,* 816-830.

74. Haselhorst, T.; Lamerz, A. C.; von Itzstein, M. *Methods Mol. Biol.* **2009,** *534,* 375-386.

75. Ma, X.; Liu, P.; Yan, H.; Sun, H.; Liu, X.; Zhou, F.; Li, L.; Chen, Y.; Muthana, M. M.; Chen, X.; Wang, P. G.; Zhang, L. *PLoS One* **2013,** *8,* e63452.

76. Lazarus, M. B.; Nam, Y. S.; Jiang, J. Y.; Sliz, P.; Walker, S. *Nature* **2011,** *469,* 564-567.

77. Kumari, M.; Kozmon, S.; Kulhanek, P.; Stepan, J.; Tvaroska, I.; Koca, J. *J. Phys. Chem. B* **2015,** *119,* 4371-4381.

78. Ghirardello, M.; Perrone, D.; Chinaglia, N.; Sadaba, D.; Delso, I.; Tejero, T.; Marchesi, E.; Fogagnolo, M.; Rafie, K.; van Aalten, D. M. F.; Merino, P. *Chem. Eur. J.* **2018,** *24,* 7264-7272.

79. Tropper, F. D.; Andersson, F. O.; Braun, S.; Roy, R. *Synthesis* **1992,** 618-620.

80. Ortega-Munoz, M.; Perez-Balderas, F.; Morales-Sanfrutos, J.; Hernandez-Mateo, F.; Isac-Garcia, J.; Santoyo-Gonzalez, F. *Eur. J. Org. Chem.* **2009,** 2454-2473.

81. Bianchi, A.; Bernardi, A. *J. Org. Chem.* **2006,** *71,* 4565-4577.

82. Dedola, S.; Hughes, D. L.; Nepogodiev, S. A.; Rejzek, M.; Field, R. A. *Carbohydrate Res.* **2010,** *345,* 1123–1134.

83. Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P.; Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S. *Anal. Biochem.* **2004,** *332,* 261-273.

Abbreviations

A: adenine

AAC: azide-alkyne cycloaddition or Huisgen's reaction

Ac: acetyl

AcOEt or **EtOAc**: ethyl acetate

Ala: alanine

AON or **ASO**: antisense oligonucleotide

Arg: arginine

Asp: aspartic acid

ATP: adenosine triphosphate

BA: bile acid

BM-PEG$_3$: bismaleimidotriethyleneglycol

Bn: benzyl

BOP: (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate

BPS: bathophenanthrolinedisulfonate disodium

BTT: 5-(benzylthio)-1H-tetrazole

***t*Bu**: *tert*-butyl

C: cytosine

CA: cholic acid

CDCA: chenodeoxycholic acid

^{13}C-NMR: carbon-13 nuclear magnetic resonance

CPG: controlled pore glass

CuAAC: copper(I) catalyzed azide-alkyne cycloaddition

Cy: cyclohexane

DCA: deoxycholic acid

DCC: dicyclohexylcarbodiimide

DCM: dichloromethane

DCU: dicyclohexylurea

DEA: diethylamine

DFT: density functional theory

DHA: docosahexaenoic acid

DIPEA: *N,N*-diisopropylethylamine

DMAP: 4-dimethylaminopyridine

DMD: Duchenne Muscular Dystrophy

DMF: *N,N*-dimethylformamide

DMSO: dimethyl sulfoxide

DMT or **DMTr**: 4,4'-dimethoxytriphenylmethyl or 4,4'-dimethoxytrityl

DMT-Cl: 4,4'-dimethoxytriphenylmethyl chloride or 4,4'-dimethoxytrityl chloride

DMT-OH: 4,4'-dimethoxytriphenylmethanol or 4,4'-dimethoxytritanol

DNA: deoxyribonucleic acid

DPAP: diethyl phosphate acetophenone

DSG: disuccinimidyl glutarate

EDC: 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

EDTA: ethylenediaminetetraacetic acid

EDU: 1-ethyl-3-(3-dimethylaminopropyl)urea

EPA: eicosapentaenoic acid

ESI-MS: electrospray ionization mass spectrometry

Et: ethyl

EtOH: ethanol

Et$_3$N or TEA: triethylamine

FBS: fetal bovine serum

FDA: Food and Drug Administration

G: guanine

Gly: glycine

GT: glycosyltransferase

HBTU: *O*-(benzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate

HCA: hyocholic acid

HCTU: *O*-(1*H*-6-chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HCV: *Hepatitis C* virus

HDCA: hyodeoxycholic acid

His: histidine

HNA: hexitol nucleic acid

^{1}H-NMR: proton nuclear magnetic resonance

hOGT: human *O*-(*N*-acetylglucosamine) transferase

HRMS: high resolution mass spectrometry

IEX-HPLC: ion exchange high performance liquid chromatography

LCA: lithocholic acid

Leu: leucine

Lys: lysine

αMCA: α-muricholic acid

βMCA: β-muricholic acid

ωMCA: ω-muricholic acid

MD: molecular dynamic

MDCA: murodeoxycholic acid

Me: methyl

MeOH: methanol

Met: methionine

MMT or MMTr: 4-methoxytriphenylmetyl or 4-monomethoxytrityl

MOE: methoxyethyl

mRNA: messenger ribonucleic acid

MS: mass spectrometry

MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

NHS: *N*-hydroxysuccinimide

***N*Pht**: phtalimide

OD: optical density

OGT: *O*-(*N*-acetylglucosamine) transferase

2′-*O*MePS-AON: 2′-*O*-methyl phosphorothioate antisense oligonucleotide

ON: oligonucleotide

PADS: phenylacetyl disulfide

PBS: phosphate buffer saline
PEG: polyethylene glycol
PFP-TFA: pentafluorophenyl trifluoroacetate
Ph: phenyl
Phe: phenylalanine
PMO: morpholino phosphorodiamidite oligomer
PMP: *para*-methoxyphenyl
^{31}P-NMR: phosphorus-31 nuclear magnetic resonance
PO: phosphate
Pro: proline
PS: phosphorothioate
Pr or ***n*Pr**: propyl or *normal*-propyl
***i*Pr**: *iso*-propyl
Py: pyridine
RNA: ribonucleic acid
ROS: reactive oxygen species
RP-HPLC: reverse phase high performance liquid chromatography
SAR: structure-activity relationship
SDS: sodium dodecyl sulfate
Ser: serine
siRNA: small interfering ribonucleic acid
SMA: spinal muscular atrophy
SMCC: succinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate
STD-NMR: saturation transfer difference nuclear magnetic resonance
T: timine
TBTA: tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine
TCA: taurocholic acid
TCEP: tris carboxyethyl phosphine
TEA$^+$: triethylammonium ion
TEAA: triethylammonium acetate
TFA: trifluoroacetic acid
THF: tetrahydrofuran
Thr: threonine
TLC: thin layer chromatography
TMS: tetramethylsilane
TUDCA: tauroursodeoxycholic acid
Tyr: tyrosine
U: uracile
UDCA: ursodeoxycholic acid
UDP: uridine diphosphate
UDP-Gal: uridine diphosphate galactose
UDP-GalNAc: uridine diphosphate *N*-acetylgalactosamine
UDP-Glc: uridine diphosphate glucose
UDP-GlcNAc: uridine diphosphate *N*-acetylglucosamine
UMP: uridine monophosphate

Ringraziamenti