Università degli Studi di Ferrara

DOCTORAL COURSE IN
"Chemical sciences"

CYCLE XXXVI

COORDINATOR Prof. Massi Alessandro

Design, synthesis, and purification of biopharmaceuticals, included highly modified oligonucleotides.

Scientific/Disciplinary Sector (SDS) CHIM/06

Candidate
Dott. Preti Lorenzo

Supervisor
Prof. Perrone Daniela

Years 2020/2023

SUMMARY

Part I: Introduction to bioconjugation

Bioconjugation

Bioconjugation consist in linking a biomolecule; that is "a molecule that is produced by a living organism" with another moiety. This concept is extremely broad as can be applied to a great variety of molecular entities, ranging from proteins conjugated to other proteins; passing to DNA fragments conjugated to small molecules such as fluorophores; arriving also at small molecules conjugates with drug or other small molecules, such as lipids, peptides and carbohydrates. Of course, the applications of bioconjugation are the most diverse; broadly speaking is common to see labelling or tagging biomolecules or linking them to a surface or one another with the use of various general reactions. Although is not a clear-cut distinction, there are traditional bioconjugation reactions and a series of more recent ones. The distinction lies in the fact that the traditional bioconjugation methods rely on well-known reactivities while the "nontraditional" ones take advance of recently developed, and more chemoselective, methods such as "click chemistry".[1–6]

Traditional bioconjugation reactions

Being called traditional takes nothing out of the usefulness of the reaction used for bioconjugation. In general, a traditional bioconjugation reaction to be used effectively must have two characteristics:

- Reaction with a native group of a biomolecule under mild conditions.
- Having been used by many researchers for many years and with continued application today.

For a reaction to occur with native biomolecular groups, it means that predominantly the chemistry of proteins, nucleic acids, carbohydrates, and lipids must be used. The typical functional groups that are commonly found in these molecular entities are amines, alcohol, thiols, aldehydes, ketones, and carboxylic acids. All this functional groups are either naturally nucleophilic or electrophilic and this allows for good modularity of the reactions as illustrated by the fluorescein example in Figure 1.

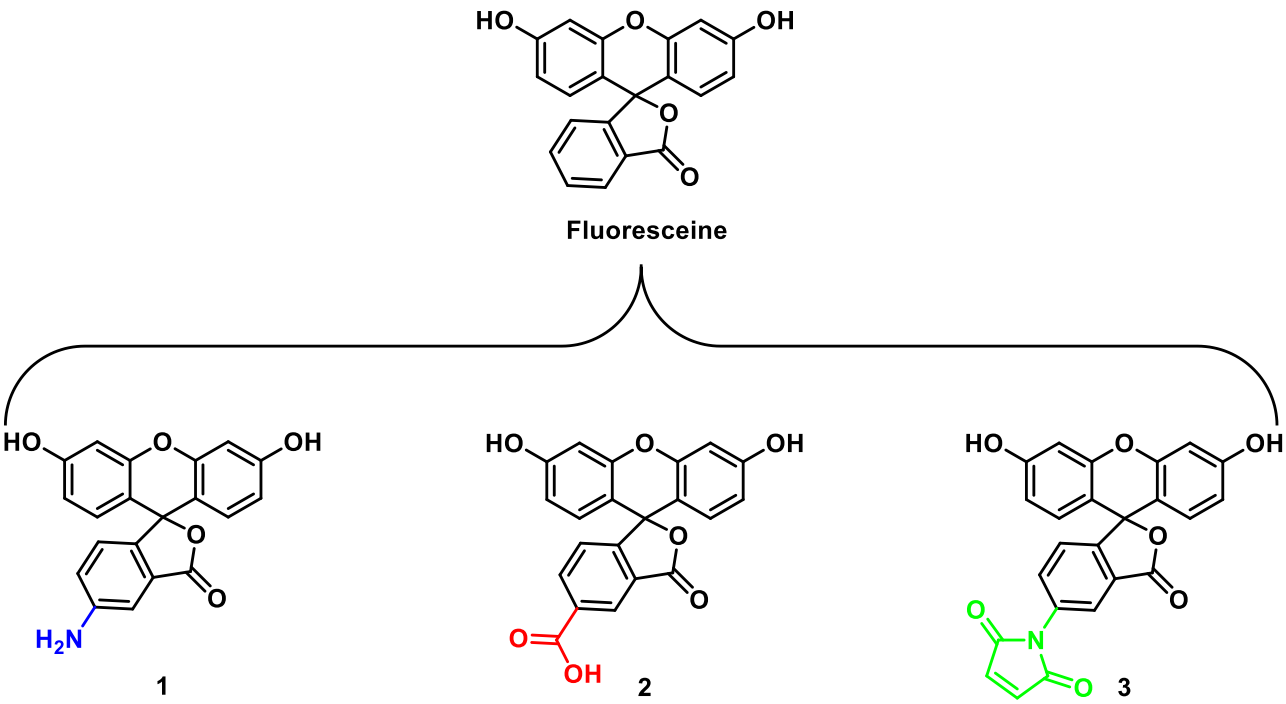


Figure 1. *Example of possible modification on fluoresceine to obtain conjugable fluorophores. Amine modified fluoresceine (**1**), carboxylic acid modified fluoresceine (**2**), and maleimide modified fluoresceine (**3**).*

It is possible to modify the fluoresceine structure without changing significantly its fluorescence properties; thus, allowing the introduction of bio-conjugable functional groups. In this case an amine can be used to form an amide bond with the biomolecule of interest. The carboxylic acid moiety can be employed for the same purpose on an amine containing molecule, while the maleimide is useful in Michael type bioconjugation of thiols. The modularity of traditional bioconjugation lies in the use of different functional groups used on different substrates. Of course, these are not the only possibilities as also thiocyanates, activated carboxylic acids and iodoacetamide can be employed for bioconjugation with amines and thiols respectively.[7]

In this chapter the most common traditional bioconjugate reactions will be discussed as some of them have been used for the aim of this thesis.

Reactions of amines

Among the nitrogen containing groups hydrazine and alkoxyamine can be exploited for immobilization of biomolecules as they selectively afford hydrazones and oximes, respectively, when reacted with a carbonyl group (Figure 2; A, B). However, both bonds are susceptible to hydrolysis particularly in acidic environment which is an important factor to be considered for many applications.[8] Nevertheless, the resulting bonds are much more stable to hydrolysis than the spontaneously forming imine (Schiff base);[9–11] even though the latter can be reduced with sodium cyanoborohydride to give a secondary amine after conjugation (Figure 2; C).[12,13]

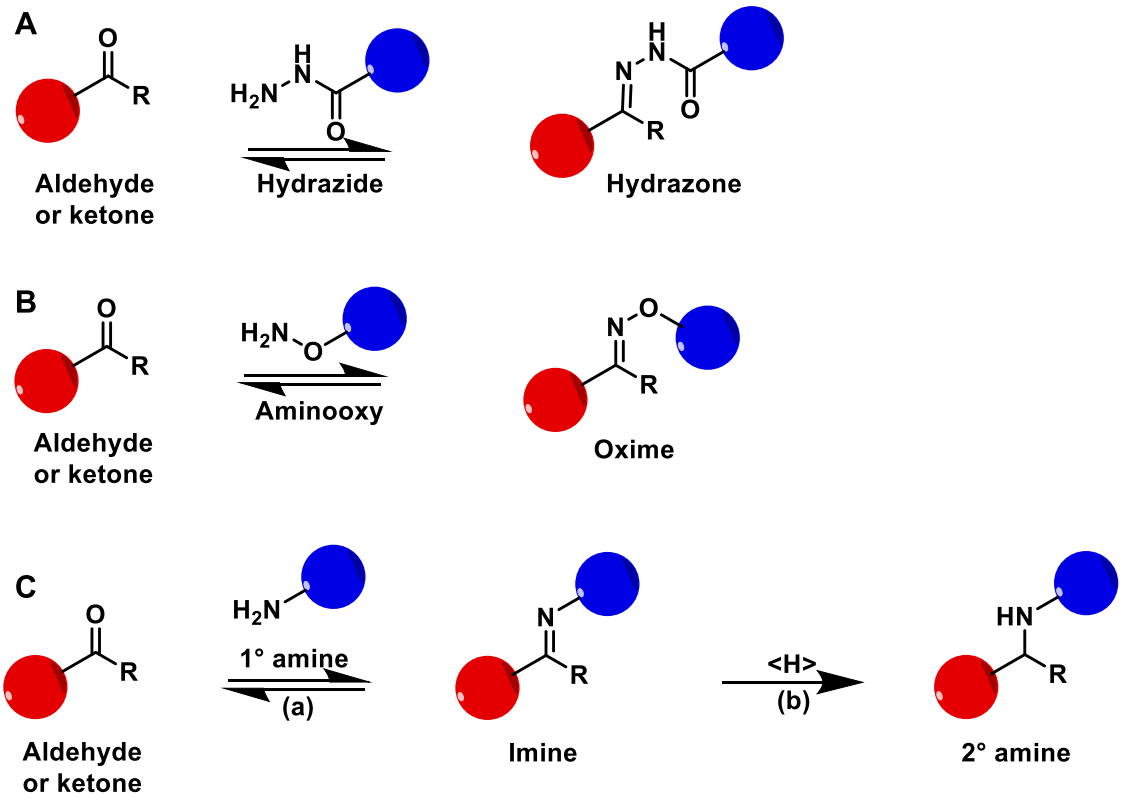


Figure 2. *A) carbonyl group transforming to hydrazone in the presence of hydrazide. B) carbonyl group converting to oxime in the presence of an aminooxy group. C) (a) reaction between a carbonyl group and an amine containing molecule to give the corresponding conjugated imine. (b) reduction of the imine with $NaCNBH_3$ to afford a stable secondary amine.*

Amines are very likely to occur in nature as they are present in all proteins and a variety of other natural products. They can react with a variety of electrophilic groups. In particular they can react with sulfonyl chlorides, carboxylic acids, activated esters, isothiocyanates, epoxides and imido esters. The necessary electrophile group can be native to the molecule of interest or introduced chemically.[2,14,15] Most of all, the amide linkage is very stable as it has a half-life of ca. 600 years in neutral solution at 25 °C.[16] For this reason amide formation is a very important reaction used in bioconjugation. the reaction can be accomplished with the use of carbodiimides as coupling agents and carboxylic acids as reactive groups. The most common reagents of that sort are Dicyclohexylcabodiimde (**4**)(DCC) and the water soluble EDCI (**5**) (N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide). Unfortunately, both reagents and the O-acylisoureas intermediates are unstable, and hydrolysis is a major competing reaction.[17,18] this drawback can in some cases be overcome by using excess reagent or when possible, by performing the reactions in organic solvents. Activation of the carboxylic acid is reported to be optimal at pH from 4.5 to 6.0.[19] However, the reaction remains feasible at pH between 7.0 and 7.5, which is often more suitable for target biomolecules. Another activation method often employed is the formation of succinimidyl esters. This reaction can be performed *in situ* by simple addition of one equivalent of NHS (**6**) (N-hydroxysuccinimide) to the O-acylisourea intermediate or as a two-step reaction in which the active ester is isolated, purified and then reacted with the desired amine. Although succinimidyl esters are more stable toward hydrolysis than O-acylisoureas, hydrolysis is still a competing reaction.[5] Rates of hydrolysis increase with increasing pH, as does amine reactivity,[20,21] the optimal Ph range for these reactions is therefore between 8.0 and 8.5 and the efficiency is severely limited above pH 9.[22] Succinimidyl esters have slow reaction rates with alcohols, phenols, and aromatic amines, such that there are usually minimal side reactions. Of course, carbodiimides are not the only amide forming existing reagents; indeed, a great

variety of such coupling agents exist, most notably uronium salts, and new ones are regularly reported in literature. [23–28]

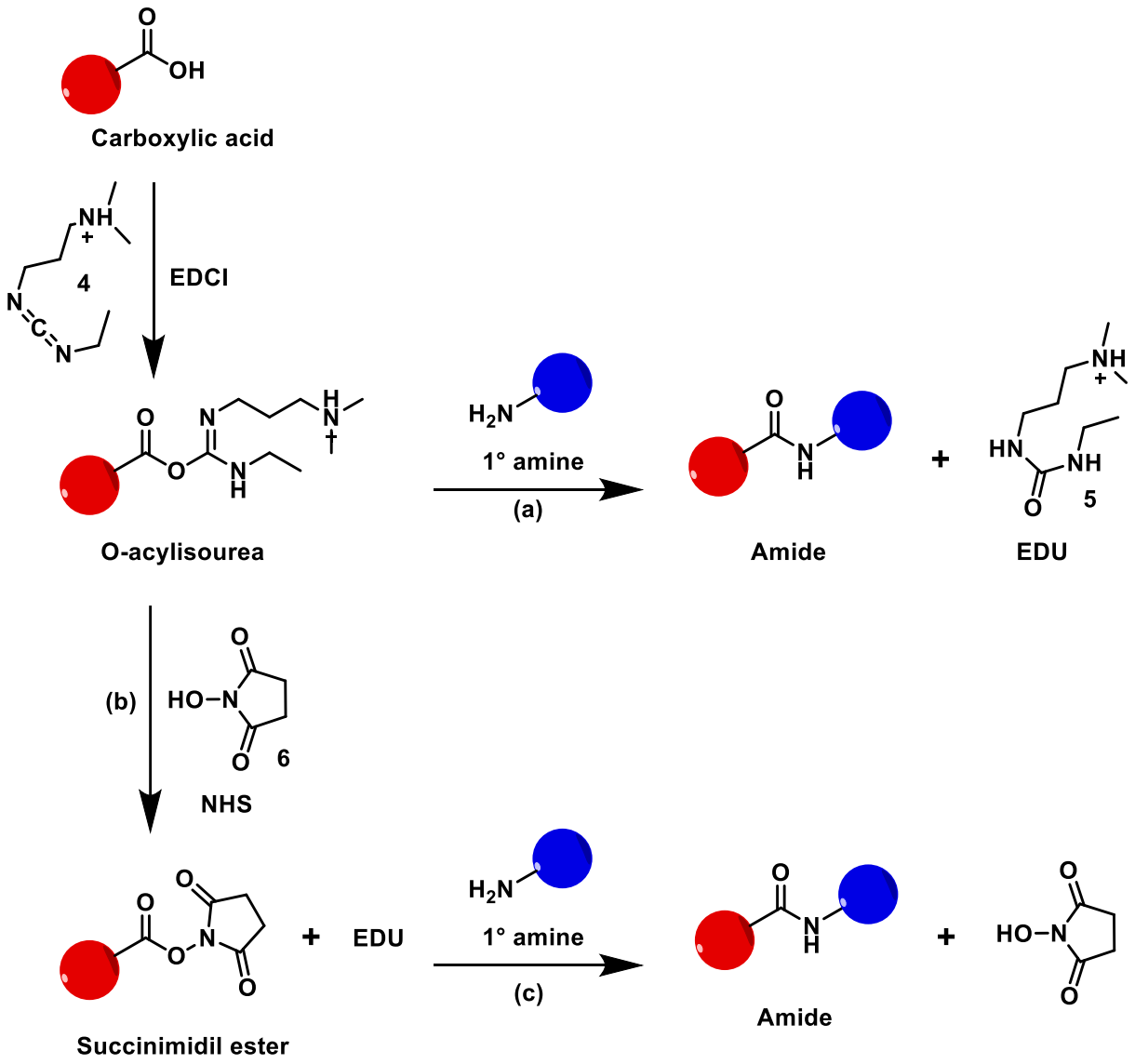


Figure 3*. N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide (EDCI) activate the carboxylic acid to form the O-acylisourea. This can then react directly with an amine (a) to give the corresponding amide and N-(3-Dimethylaminopropyl)-N'-ethylurea, or it can react with N-hydroxysuccinimide (b) to give the intermediate activate ester. This can be reacted immediately with the amine as a one pot strategy or can be isolated and reacted in a later occasion (c).*

Reaction of thiols

Thiols have been particularly attractive group for bioconjugation because they are present in proteins in but are quite rare, as opposed to amines. Cysteine is the second least abundant amino acid in proteins,[29] and this allows a certain specificity, even site specificity, when a bioconjugation reaction on protein is attempted.[8] If the right electrophilic counterpart is chosen, is possible to obtain a stable link between the two biomolecules and achieve a good chemoselectivity. As an example, thiols and thiolates, being the strongest protein nucleophiles,[30,31] can easily react with active esters to give unstable thioesters. Traditionally therefore, thiols have been conjugated to maleimides through thio-Michael reaction to form stable thioethers.[2,3,14] The reaction proceeds also in aqueous environment and is very selective between pH 6.5 to 7.5; above this range there is competition with the aza-Michael, this can be a problem, especially

in complicated substrates such as proteins.[32,33] Imide hydrolysis to give unreactive maleamic acids is also in competition with bioconjugation at elevated pH values.[34] In addition, the imide hydrolysis can also occur after conjugation giving rise to heterogeneity of the final product. Although in some cases this is not relevant since the thioether bond remain in place, mild methods for achieving complete hydrolysis and thus regaining product homogeneity have been developed.[35]

Alternatively, haloacetamides can be used as electrophiles. Chloro-acetamides, in particular, are selective towards thiols in cysteine residues. The reaction happens at neutral or slightly acidic pH at which amines are protonated and less nucleophilic. However, of iodo-derivatives react more than an order of magnitude faster than chloroacetamides, although they result more sensitive to light and require the bioconjugation to be carried out in the darkness to avoid iodine formation.[4]

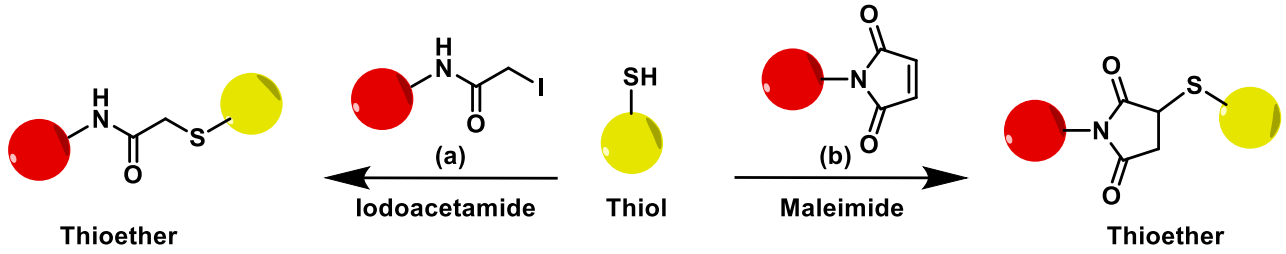


Figure 4. *Reactions of thiols to give thioethers. (a) reaction with iodoacetamide. (b) Thio-Michael reaction with maleimide.*

Another option is given by the possible formation of a disulfide bond between two thiols. The reaction requires the formation of an intermediate disulfide by action of reagents, such has 2,2'-Dipyridyl disulfide (**7**) (Figure 5; 1) or Ellman's reagent[36]. This can be later attacked by the second thiol to form the final bioconjugate disulfide. This bond is not particularly stable and results susceptible to cleavage in a reductive environment; this characteristic can be very advantageous if used for intracellular drug release by action of endogenous glutathione.[37]

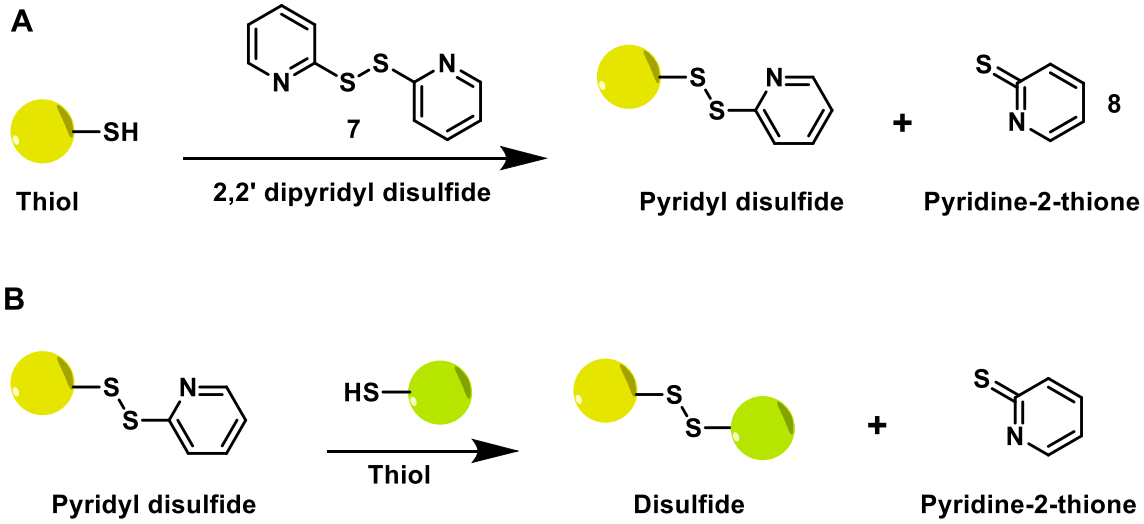


Figure 5. *(A). Reaction of thiol with 2,2' dipyridyl disulfide to afford pyridyl disulfide. (B) Reaction of pyridyl disulfide with another thiol containing molecule to give the bioconjugated disulfide*

Reactions of alcohols

The hydroxy group is found in a myriad of biological molecules, including approximately 65% of natural products; and more than a third of known drugs.[38] For this reason, the possible chemoselective bioconjugation of this group is highly desirable, but it remains a difficult task. Hydroxyls and phenol are generally not used in protein bioconjugation reactions since their nucleophilicity is lower than that of amine and thiols. The reaction is therefore not selective and not particularly useful in most cases.[4] However, when smaller molecules are involved is possible to exploit hydroxyl groups as a useful molecular handle for bioconjugation and modification reactions. Esterification reaction with active esters is the predominant bioconjugation strategy, although it can be severely hindered by the presence of water. Esterification in acidic conditions is more selective since the amines tend to be protonated and less nucleophilic. More recently, the esterification has also been attempted with a biocatalytic approach since some lipases are selective for ester over amine formation and can work in organic solvents.[39] Moreover, the same reaction can be obtained with zinc catalysis.[40]

One other strategy for hydroxyl bioconjugation is the Ullmann-type coupling with aryl moieties, which can be tuned for hydroxyl over amine selectivity; in this case is possible to obtain a more stable ether instead of an ester.[41,42] Ethers can be selectively formed with silyl containing molecule due to the oxophilic character of the silicon atom.[43] Other methods based on the functional group interconversion are reported in literature.[44] A notable example of bioconjugation is the possibility of linking diols and boronic acids. This reaction is highly selective towards cis-1,2 diols to form cyclic boronates; although It occurs also with 1,3 diols and 1,3 hydroxy acids. The reaction is reversible and the boronic esters hydrolyse at low pH values.[45]

Cross linking strategies

One the main aims of bioconjugation is to bond together two biomolecules; a cross linking strategy is usually employed depending on the availability of functional group on the two molecules. Below, some of the more common strategies will be discussed.

Traceless ligation

Traceless ligation (or zero-length cross-linking) refers to the direct linking of two biomolecules through covalent bonds *via* activating agent or reactive group, which is not incorporated into the final conjugate, leaving no residual atoms and adding no length to the bond. One common example of traceless ligation is amide coupling through carbodiimide activation of carboxyl groups or via a succinimidyl ester. Another one is given by the thiol-disulfide exchange aided by the presence of pyridyl disulfide reagents. Advantages of traceless ligation include minimal 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 or are too sterically hindered will not be able to react.[4]

Homobifunctional and Heterobifunctional Linkers

Many bioconjugation reaction rely on the use of tuneable linkers, both in terms of length and polarity, to overcome the steric hindrance problems often encountered when conjugating large biomolecules to surfaces or other biomolecules. These linkers are seldom used in small molecule conjugation but can be employed if the naturally present groups require them. The linkers can be divided in homobifunctional, meaning that there is the same reactive group at both ends of the linker or heterobifunctional, with different groups and reactivities. This latter type can also be exploited to improve conjugation selectivity.

Often the linker main chain is alkyl or composed of PEG, depending on the required polarity and length of the cain required for the specific purpose.

Homobifunctional linkers commonly available are for example bis-NHS esters, bismaleimides, and glutaraldehyde. The reactions of bis-NHS and bismaleimide reagents are analogous to their monofunctional counterparts, cross-linking amine and thiol groups to form amide and thioether linkages, respectively. Glutaraldehyde is frequently employed for protein bioconjugation exploiting the imine formation between the aldehydic function and primary amines present in proteins. Some examples of these linker can be seen in Figure 6. Many traditional heterobifunctional cross-linkers combine amine and thiol reactivity, often in the form of succinimidyl ester and maleimide groups separated by spacers of different lengths. New linkers combine a good variety of different functional groups to achieve significant selectivity, like azide or alkyne with carboxyl or amine moieties. Many factors contribute to the linker choice; apart from the reactivity of the biomolecules also the purification ease and reaction condition requirement are important when choosing an appropriate bioconjugation linker. It is particularly important to decide the required final stability of the bond, indeed, is possible to create labile linkages between the desired molecules using disulfides, diols, or esters that can be cleaved by reduction, oxidation, or a strong nucleophile attack. Alternatively, is possible to link them indefinitely with amide or triazole moieties, or even to modulate the bond strength by choosing hydrolysable groups at low pH, such as hydrazone, oxime and acetals. Other stimuli can be exploited to break the linkage when needed. The use of UV light in combination with photocleavable groups such as o-nitrobenzyl derivatives; or a structure that is recognized as a substrate by hydrolytic enzymes.[46–49]

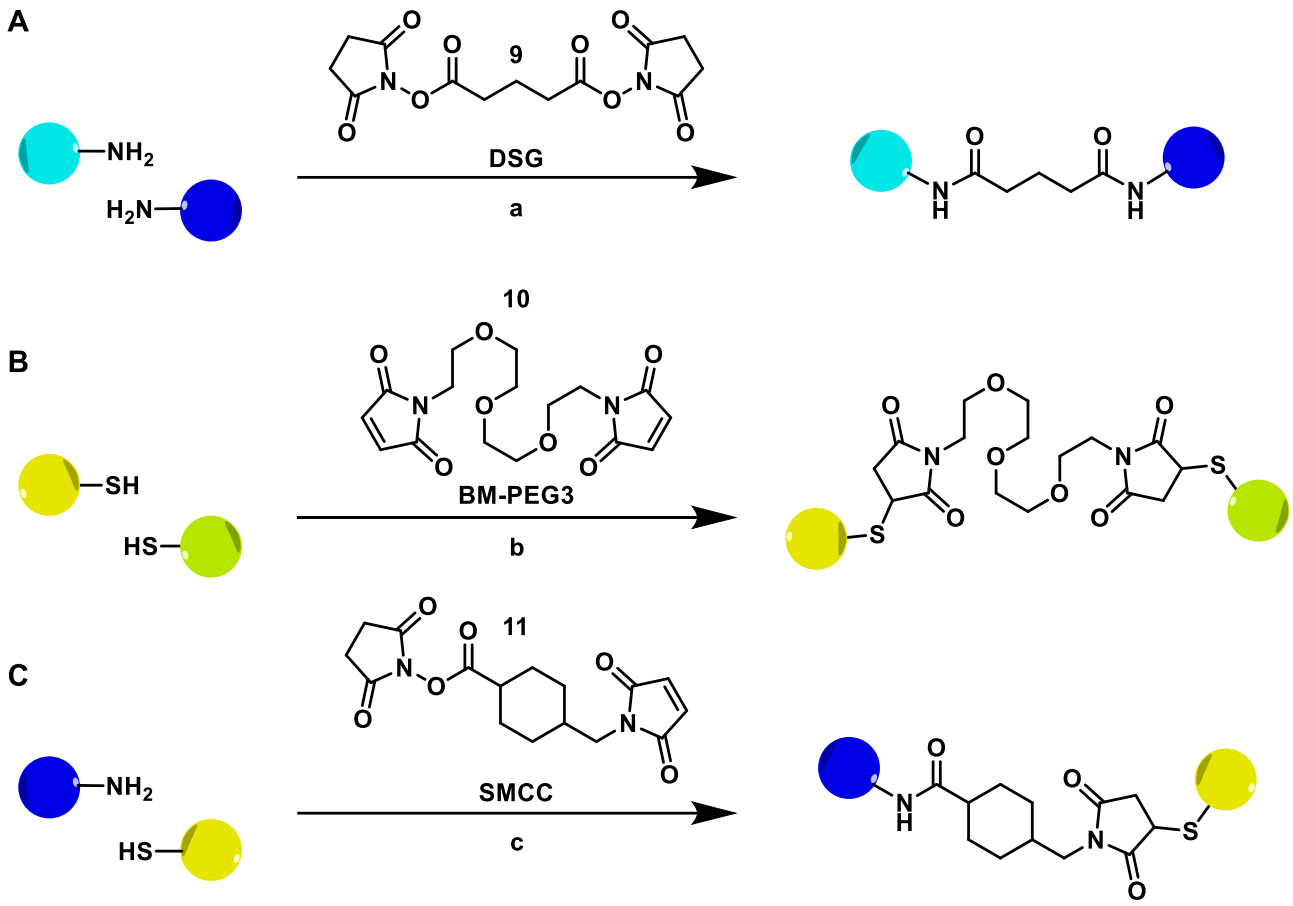


Figure 6. *Examples of Homobifunctional and heterobifunctional linkers used in bioconjugation. (a) Linking of two amine containing molecules through disuccinimidyl glutarate (DSG). (b) Conjugation of two thiols by bismaleimidotriethyleneglycol (BM-PEG₃). (c) conjugation of an amine and a thiol trough the heterobifunctional linker 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (SMCC).*

Functional group conversion

It may happen that the functional groups present in the biomolecule are not ideal for conjugation due to reaction conditions not compatible with the other molecule or lack of selectivity or stability. A good solution can be found in functional group conversion. This method often successfully applied in organic synthesis is precious also in bioconjugation. Ethylenediamine (**12**) and other bisamines (e.g., hexamethylenediamine) are, for example, a common route for converting electrophilic functional groups into amines.[5] carboxylic acids will form stable amides leaving an amine necessary for successive bioconjugation, while carbonyls will first form imines and then secondary amines after reduction. This allows to invert the electrophilicity of the naturally occurring groups adding a short or long appropriate linker. It is of course possible that undesired intramolecular or intermolecular crosslinking occurs when performing the reaction. In most cases, however, it is sufficient to use a large excess of the linker and purify the obtained product if needed. Short linkers further help minimize side reactions of this kind. Similar chemistry can be used to convert carbonyl and carboxyl groups into thiols using cystamine (**13**). 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 desired thiol.

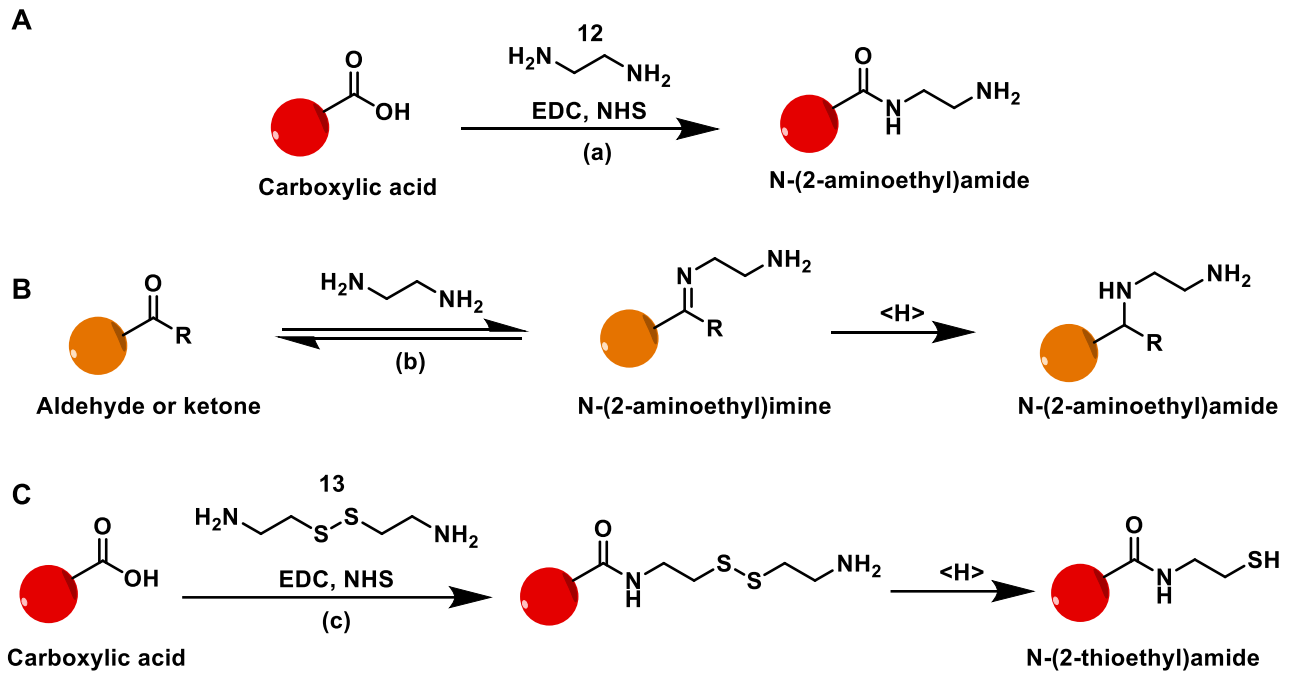


Figure 7. *A) Reaction of bifunctional amine (ethylenediamine) with carboxylic acids. B) Carbonyl compounds. C) Reaction of cystamine with a carboxylic acid to give and reduction to thioethyl amide.*

The opposite operation, the conversion of amines into carboxylic acids, can be accomplished using succinic anhydride. This reagent reacts also with thiols, imidazoles, phenols, and alcohols to afford less stable conjugates. In the case of competing amine and alcohol groups is possible to restore the selectivity by adding hydroxylamine. This breaks the ester reforming the original hydroxyl and giving NHS as a side product.[50] In every case the anhydride breaks and forms a terminal carboxylic acid, ready to be used in the next step. Homobifunctional succinimidyl ester cross-linkers can also be used to convert an amine to a carboxylic acid. In this case, one succinimidyl ester group reacts to form an amide bond and the other is allowed to hydrolyse to generate the carboxylic acid.

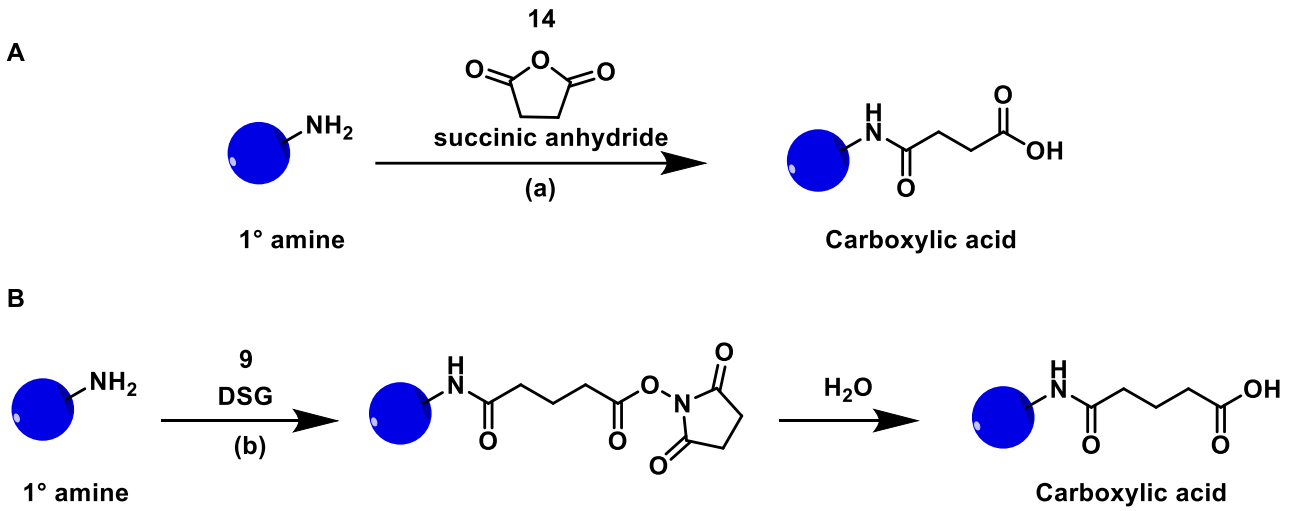


Figure 8. A) *Reaction of succinic anhydride with a primary amine to give a carboxylic acid. B) Reaction between primary amine and disuccinimidyl glutarate, followed by water hydrolysis to give a carboxylic acid.*

Traditional bioconjugation reaction challenges

Traditional bioconjugations have many challenges or drawbacks to be overcome. It is useful to establish what an ideal bioconjugation reaction looks like before describing the challenges faced by the traditional methods. Ideally, the reaction should be high yielding in mild conditions; chemo- and possibly regio-specific; occurring with widely available or easily prepared reagents. It would also be ideal if these reagents were safe, water soluble and easily removable in case of purification.

1. Conditions
    1.1. Fast reaction
    1.2. Stoichiometric efficiency
    1.3. Sub-micromolar concentration
    1.4. Physiological pH range
    1.5. Room or physiological temperature

2. Specificity
    2.1. Chemo-specific: only at target site
    2.2. No side reactions
    2.3. No matrix interference

3. Reagents
    3.1. Highly available reagents
    3.2. Stable
    3.3. Water soluble
    3.4. Safe
    3.5. Easily removable side-products

Unfortunately, there are few traditional reactions, if any, that satisfy these criteria. All the reactions seen above suffer from one or more drawbacks which renders them non ideal. Competing hydrolysis is often the first factor that limits reaction efficiency and usefulness, labels with succinimidyl esters are prime examples of this phenomenon. Since hydrolysis in water systems tends to be a reaction of pseudo-first order, while the bioconjugation is a reaction of second order, is frequent to have yield lower than stoichiometric.[34] The problem is, in some cases, circumventable by solvent change toward polar organic solvents like acetonitrile, dimethylformamide, dimethysuolfoxide or short chain alcohols. When the use of a different solvent or cosolvent is not sufficient to limit hydrolysis is possible to increase the ratio of labelling molecule or coupling agents to contrast the hydrolytic losses; frequently a large excess of reactants is required. Moreover, while the storage conditions of moisture-sensitive activating and coupling agents also play a role in their efficiency, the hydrolysis rate strongly depends on reaction pH and temperature. Since pH is determinant of both nucleophile reactivity (i.e amines) and hydrolysis rate (at higher pH the nucleophile reacts faster but so does water) it is of the outmost importance to find an optimal range for every substrate. Temperature control face a similar issue due to the faster conjugation kinetics at high temperature but also faster hydrolysis. Lastly a too lengthy reaction could allow side reactions to occur although seldomly. All these considerations force the chemist to empirically optimize every new bioconjugation, which can be a long process sometimes and negatively impacts the reproducibility of the research work. In addition, hydrolysis is not the only detrimental reaction in terms of selectivity. In complex substrates like proteins for example the strong nucleophilicity of thiols give a competing reaction when amines are the desired target and thus dictates the order of reaction: thiols first

and amines second. In some cases, also the aqueous buffer can be a problem; phosphate can interfere with carbodiimide coupling for example.[4] Yet another challenge, is the limited intrinsic diversity of biomolecules. The vast majority of them are indeed biopolymers with many repeated functional groups in different places on the molecule, for example amino acids in proteins, nucleotides in nucleic acids and saccharides in carbohydrates. This make a regioselective reaction virtually inaccessible. In the case of bifunctional linker, the reaction products can be further complicated by the combination of possible couplings. For clarification: if the molecule A is to be conjugated to the molecule B but both of them have the same reactivity in more than one site (i.e amines) a homodifunctionalized linker (i.e di-activated ester) will give rise the formation of the desired product A-B but also of the unsought A-A and B-B, without even counting the possible formation of oligomers (A_n-B_m).
Of all the traditional bioconjugation reactions, the thiol-maleimide and disulfide exchange are the two that come closest to the ideal. This is mostly due to the good nucleophilicity at neutral pH and the relatively scarcity of thiols in biomolecules. In addition, hydrolysis rates of maleimides are slower compared to other chemistries. There still are limitations like the need to use reducing agents for disulfide formation, that can interfere with biomolecules, and the two steps envisaged by the coupling procedure. Moreover, the hydrolysis of maleimides at low pH is still a problem even if slower than, for example, activated esters. Overall, these two reactions still represent a good standard desirable standard to be reached by new bioconjugation methodologies.[4]

Conclusion

The bioconjugation reactions discussed above are but some examples of the total available traditional strategies. Although everyone with their own specific limitations, these reactions remain of outmost importance, are still used daily and in some instances allow good selectivity and reactivity control. These reactions, moreover, remain convenient in many cases even when different approaches could be attempted. The limitations suffered by these reactions should be considered as a challenge to prompt research on discovery of new and improved conjugation methods. More ideal, biorthogonal, chemoselective and more sophisticated reactions have been emerging in the last decades and are still emerging now. These will be discussed below.[4]

New methods in bioconjugation

Click chemistry

The field of bioconjugation is in continuous expansion, driven by the desire for more selective and biorthogonal reactions. Scientist have developed many new procedures. Among these, prominent is the Staudinger ligation, both in traceless[51] and non-traceless version. This reaction consists of the formation of an amide bond starting from an azide and an acyl-triphenilphospshine.[8,52] another possibility is offered by the olefin metathesis reactions, applied also to proteins.[53,54] furthermore, there are squarate based methods which allows the reaction of two amino groups with **16** that results in their conjugation via two vinylogous amides.[55] A new methodology has also been developed for the bioconjugation through methionine residues in proteins.[56]

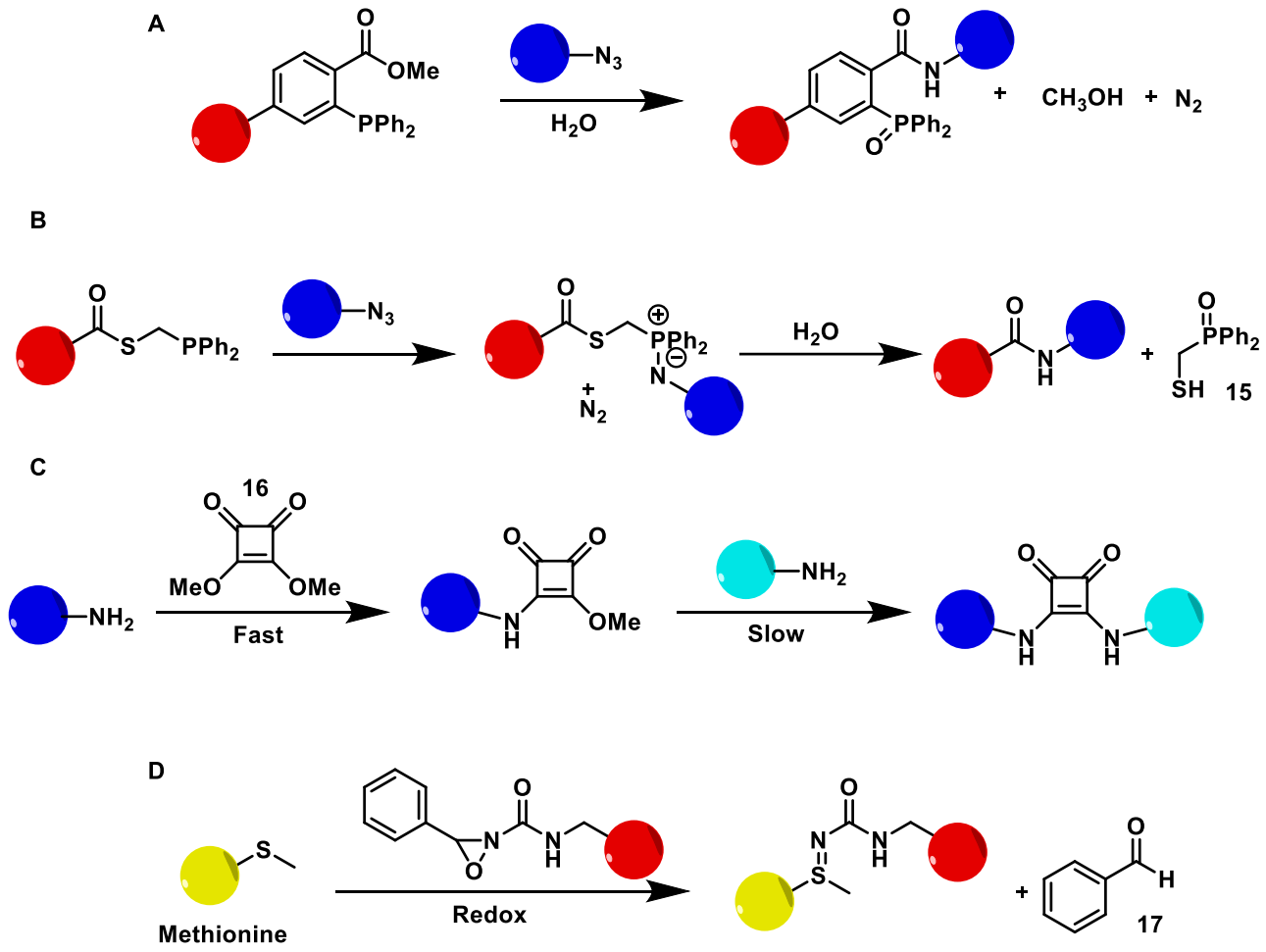


Figure 9. *Scheme representing the possible recently developed conjugation methodologies. A) Staudinger conjugation. B) traceless Staudinger ligation. C) conjugation through squaramide formation. D) conjugation of methionine residues in proteins.*

In addition to this, there are click reactions that can be used in bioconjugation. In 1999 K.B Sharpless defined click chemistry reactions as reactions in which the high energy stored in the reactants allows the formation of irreversible bonds between two complementary functional groups, much like a luggage strap connector, the two entities are clicked together by a stable bond. In 2022 the Nobel prize was assigned to him and co-workers for conceptualizing and developing such profoundly impacting reactions.[57] These reactions often give little or no by product and are biorthogonal. There are four kind of reactivity that allow click reactions: nucleophilic substitution; additions to C−C multiple bonds (i.e. Michael addition, epoxidation, dihydroxylation, aziridination); non-aldol like chemistry (i.e N- hydroxy succinimide active ester couplings); and cycloadditions (i.e Diels−Adler reaction, Huisgen's cycloaddition).[58] Many interesting reviews can be found on the myriad applications in which this kind of chemistry is used.[59–70]

In particular, (3+2) dipolar cycloaddition reactions (or 1,3 dipolar cycloadditions) are very useful since they allow great orthogonality and atom economy.[71,72] This kind of reaction, when not catalysed, involves the concerted rearrangement of two double bonds to give two new single bonds. This happens following the rules of [4π+2π] (like in the Diels-Alder reaction) to give pentatomic heterocycles. [73]

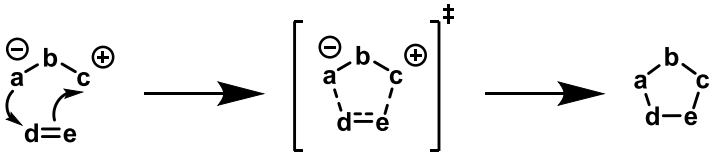


***Figure 10**. Mechanism of a general (3+2) cycloaddition. The dipolarophile and the dipolar molecule electrons are represented to move in a concerted fashion to give the intermediate. The molecular orbital of this intermediate is believed to have C_s symmetry.*

1,3-dipolar molecules are isoelectronic with the allyl anion and possess 4 π electrons; they have at least one resonant structure with separate formal charges. Some examples of 1,3-dipolar groups are reported below.

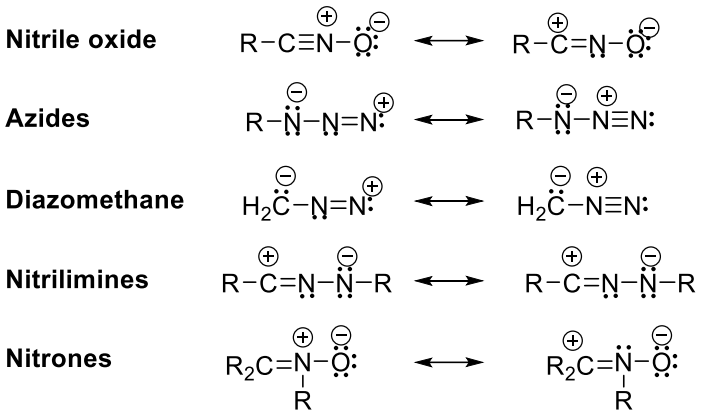


***Figure 11**. List of some common 1,3-dipolar groups with names.*

The 1,3 dipolar group reacts with a dipolarophile; usually an alkene or alkyne variously substituted. The reactivity depends on both the dipolarophile substituent and the nature of the dipole. Although many 1,3 dipolar groups have been found not all of them can react in this way, notable exceptions are ozone and nitrous oxide.[74] Many other groups, such as nitrones, diazo-compounds and azides are successfully exploited for creation of complex structures.

The CuAAC click reaction

Of particular interest in the bioconjugation field is the Huisgen azide-alkyne cycloaddition, introduced in 1967.[75] In this reaction both reactants are stable and easy to incorporate in biological molecules by means of nucleophilic substitution or other methods. The terminal alkyne and azide stability in water and oxygen rich environments, as well as in most synthetic conditions confers them orthogonality. In addition, they do not react in physiological conditions and are normally not enzymatically attacked; so they could also be suitable for biorthogonal reactions. However, their good stability also means that for them to react is necessary to furnish energy; commonly in form of heat.[76] The temperature must be raised around 100 °C thus making the reaction suitable only for very stable and simple compounds. Moreover, when they react in these conditions two regioisomers are formed; the 1,4 and the 1,5 triazole.[77–79] Fortunately, the discovery that copper could catalyse this reaction between alkyne and azide at room temperature resolved all the problems mentioned above and gave the chemist one the best reactions for bioconjugation currently available.

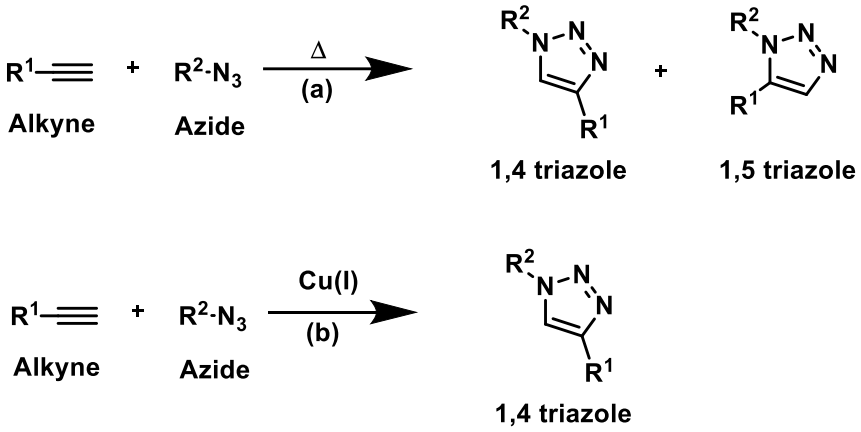


Figure 12. *Schematic representation of the uncatalyzed Huisgen cyclization (Above) and the copper catalysed azido alkyne cyclization. a) toluene, reflux. b) acetonitrile, CuI or water/ t-butanol, $CuSO_4$, sodium ascorbate.*

The copper catalysed version (CuAAC) of the azido alkyne cyclization allows biorthogonal conjugation. It is considered the best example of click reaction as it only affords the 1,4 regioisomer and is routinely done at room temperature. In fact, it incarnates all the requirement identifying click reactions, as expressed by Sharpless:[80]

- wide reaction scope.
- consistently high yielding reaction with diverse starting materials.
- Reaction must be easy to perform.
- Reaction must be insensitive to oxygen or water.
- Reaction should occur with readily available reagents.
- Simple reaction work-up and product isolation.

In addition, to these characteristics the CuAAC can take place in a variety of solvents and most notably in water. The use of ascorbate as a reducing agent to transform copper sulphate into active Cu(I) is a particularly advantageous method since it allows the use of water as a solvent. Many reactions between apolar substrates can still be carried out in this system in a biphasic solvent.[76]

Mechanism

The mechanism of this cycloaddition changes when catalysed by copper. Indeed, the reaction is no longer concerted but progresses through formation, in steps, of metal alkyne complexes and metal azide complexes. Many articles were published over the years trying to assess the mechanism of this reaction. In the first instance Sharpless proposed a mechanism based on the presence of one Cu (I) ion.[76] In this hypothesis the alkyne substrate (**19**) is deprotonated, and it coordinates to the copper (I) complex (**18**) by σ bond. The newly formed complex (**20**) then also exchanges one ligand for the azide substrate (**21**) which attacks from the formally negative nitrogen to give intermediate **22**. At this point the authors propose concomitant formation of the C-N bond and of a double bond between the copper atom and acetylene substrate. This forms an unusual six membered copper (III) metallacycle (**23**) which then undergoes ring contraction to give the copper triazolide (**24**). After protonation the 1,4 triazole ring (**25**) is released and the copper catalyst is reformed.

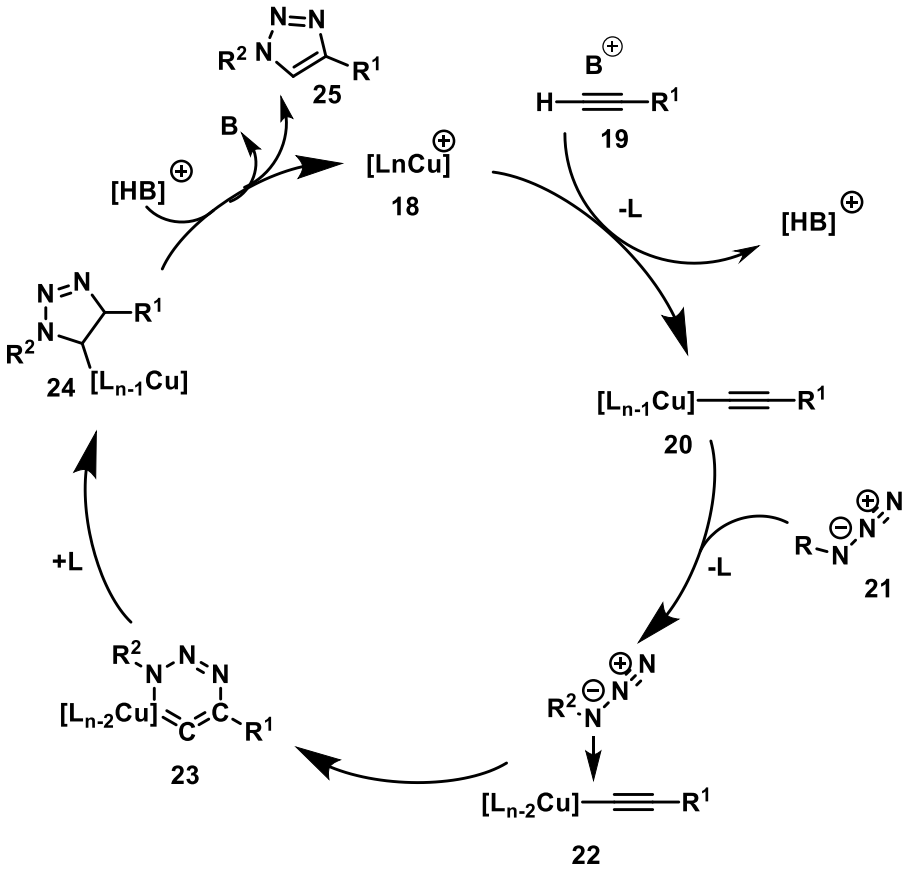


Figure 13. *First proposed catalytic cycle of the CuAAC. In this proposal, only one copper atom makes the active catalyst.*

The authors performed extensive DFT calculations which supported this stepwise mechanism opposed to a concerted Diels-alder like mechanism. The formation of the initial copper alkyne complex (**20**) is supported by the calculation and is a well-known reaction in many C-C bond formation by copper catalysis. Moreover, the study attributes to the initial formation of the π-complex between alkyne and copper the lowering of the alkyne C-H pKa by 9.8 units; which allows the deprotonation of the substrate in water.

Later studies by Fokin and Finn groups found that when the click reaction is performed with Cu(II) salts and sodium ascorbate as a reducing agent to form copper (I), the kinetics resulted of the second order for the catalyst. This suggested the possibility of two copper centres in the catalyst.[81] These findings were confirmed by Mykhalichko[82] and García-Granda[83] whom observed the presence of polynuclear copper(I) alkyne complexes, and by other studies on ligand free catalysis. Thus, a revised mechanism was presented by Fokin and Finn in 2005.[84] The mechanism with a di-nuclear copper catalyst was further confirmed by computational studies highlighting its increased rate of catalysis in respect to the mono-nuclear mechanism.[85,86]

In the first step, the alkyne (**19**) π-coordinates to the copper(I) centre and can be easily deprotonated to give the acetylide complex, for example by sodium ascorbate. A neighbouring copper(I) centre might also attract an acetylide ligand, which is shown in the equilibrium reaction on the right side, the product of this interaction (**27**) however, is inactive as a catalyst. At this point the di-nuclear copper complexes the azide. This forms the intermediate **28**, in which the azide is well positioned to attack the C2 of the alkyne substrate, probably giving a six membered ring intermediate similar to **23**. Although it is not reported in the work of Fokin and Finn[84] this intermediate is computationally proposed in the work of Straub[86]. In

any case, after the attack of the azide on the alkyne there is formation of the triazolide **29** and, in the same fashion as before, its protonation gives the product and restores the di-nuclear catalyst. A difference in this revised mechanism is given by the possible reversible addition of one more molecule of alkyne to the triazolide product **29** to give **30**.

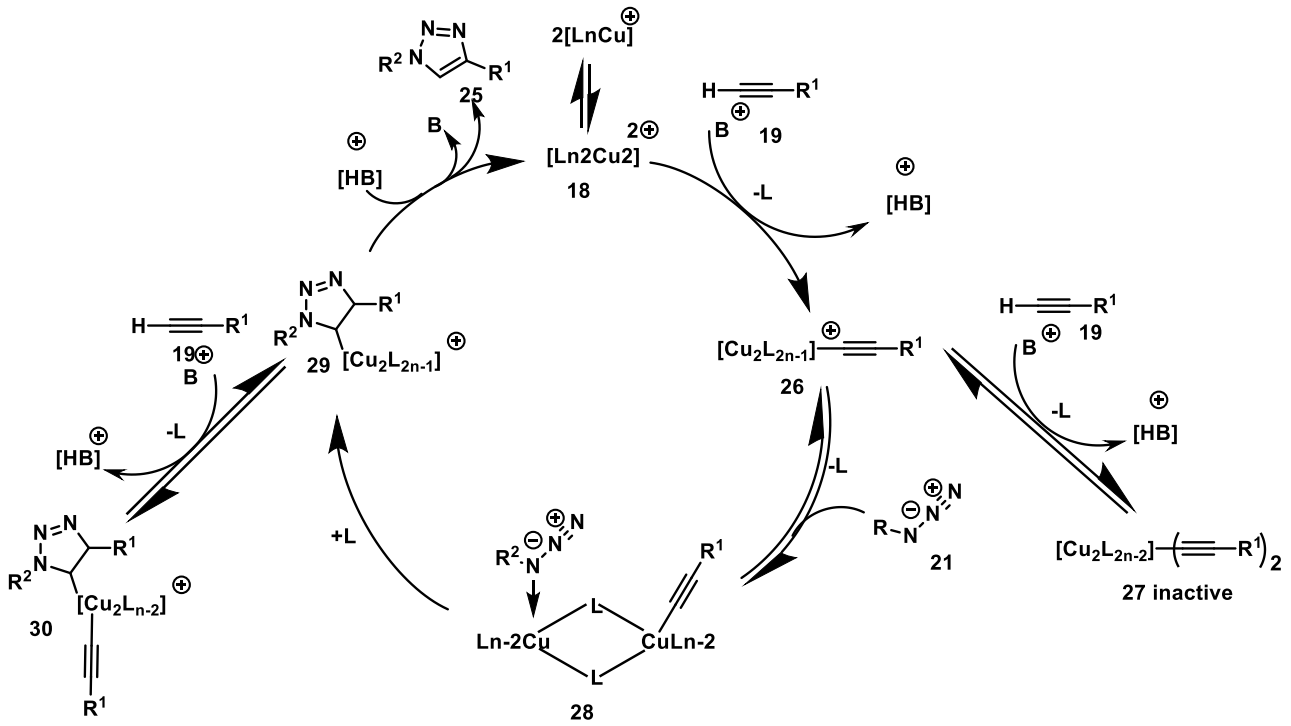


Figure 14. *Revised catalytic cycle for the CuAAC reaction. In this proposed catalytic cycle two copper atoms, linked by the relative ligands form the active catalyst and correctly position the intermediate for triazole formation.*

Both observation of yellow compounds formation in the reaction's early stages[87–91] and computational models of copper cluster formation, suggest that copper clusters are being formed, at least initially. So, it appears that the reality is more complicated than previously thought for this reaction.[92] Moreover, there was no experimental confirmation of the exact nature of the di-copper ligand complex in the studies regarding this second catalytic cycle. Recently, more evidence confirmed that the nature of CuAAC is dynamic; both mono and di-nuclear copper catalysis can form. The relative rates strongly depend on the reaction condition and ligands used.[91,93] In this case the catalysis always starts with the formation on the copper acetylide as in the previous mechanisms. At this point complex **20** can accommodate the azide to give the intermediate **22** and from there form the product as seen in the first mechanism illustrated. Alternatively, complex **20** can form a di-nuclear copper complex and lead to a new intermediate in which the two-copper atoms are bound to the triazolide in position 3 and 5. This then gets protonated by another alkyne substrate molecule and restart the cycle giving the copper catalyst already bound to the new substrate molecule; without needing a base. Overall, the authors report a much faster catalysis with the di-nuclear complex; almost one order of magnitude.[93]

Our understanding of the catalytic details of CuAAC has improved much since its discovery and the reaction is no longer a "black box". However, not all details are yet known regarding the structural nature of transition states and intermediates. More studies are regularly published shedding light on the possible structure.[94]

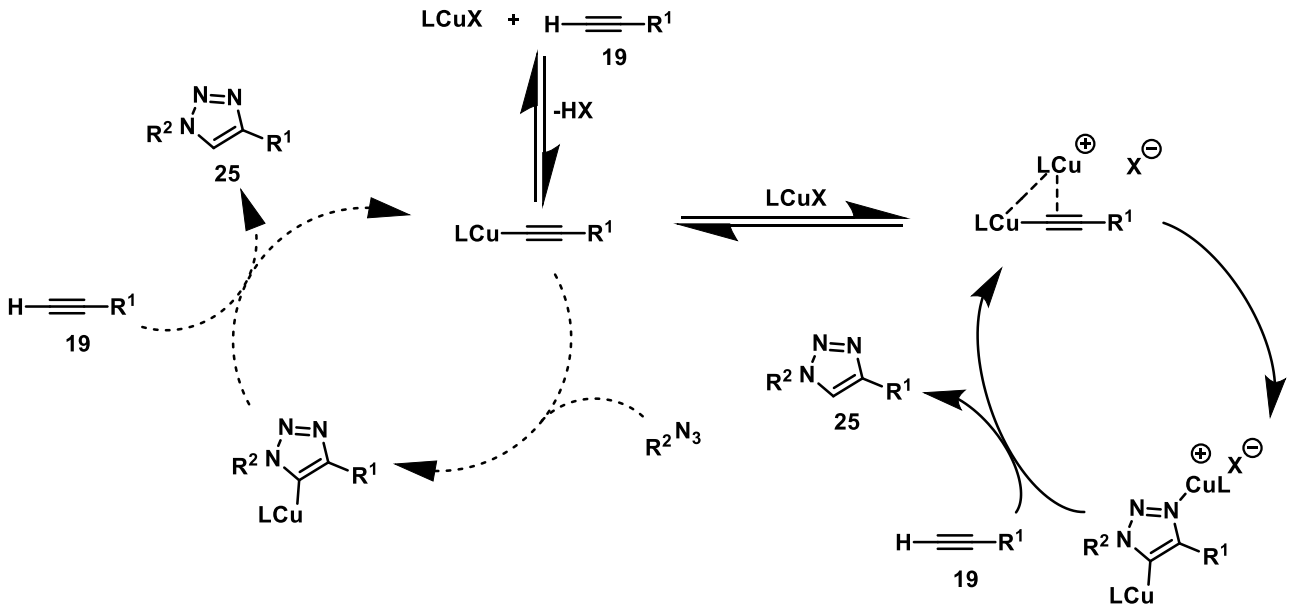


Figure 15. *Latest catalytic cycle of CuAAC proposed. In this instance the copper can act as catalyst both in a one atom fashion and as a bimetallic complex.*

Given all these favourable factors it is no surprise that this reaction has been used innumerable times in diverse fields with great success. The CuAAC still presents some limitations, for example, although the azide is a fairly stable group it can be reduced by soft nucleophiles, such as phosphines, to give a primary amine and release nitrogen gas. In addition, electron poor azides can undergo a photochemical transformation to form nitrenes.[95] Moreover, some small azides can be potentially explosive. Also, the alkyne group can sometimes be susceptible to unwanted reactions. Terminal alkynes can give coupling with free thiols[96] while the acyl-alkynyl moiety is a good Michael acceptor; thus, potentially susceptible to undesired nucleophilic attacks.[97]

Particularly demanding, is the field of *in vivo* bioconjugation where reactions must occur in the sensitive cellular environment. In this case the main limitation of copper catalysed azido alkyne cyclization is the necessary use of copper. While this is not a problem in many reactions, it is not ideal for conjugation in a cellular environment since the presence of this metal can lead to generation of free reactive oxygen species (ROS)[98,99]. This species can, in turn, damage biomolecules and are a known cause of cell toxicity.[100,101] The generation of ROS from Cu (I) follow the same patterns as the Fenton reaction with iron as illustrated in the Figure 16 below.[102]

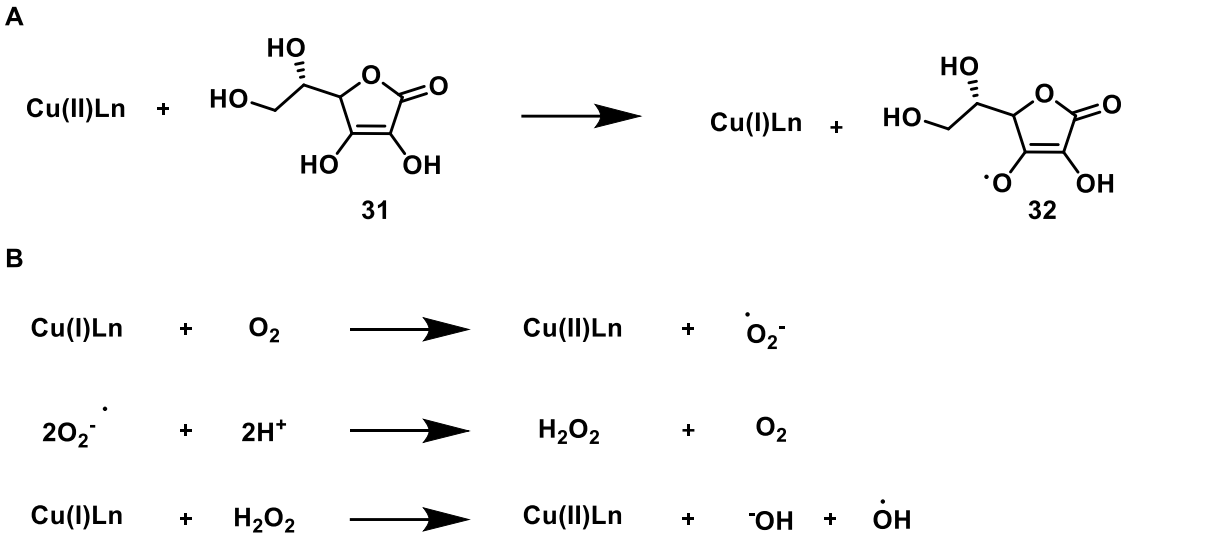


Figure 16*. Reduction of copper (II) complex by sodium ascorbate to give active catalyst Cu(I) (A). oxidation of Cu(I) to copper(II) complex with oxygen. This first reaction generates a superoxide radical. Two superoxide radicals are protonated to give hydrogen peroxide, which can further oxidize the copper (I) complex and release hydroxyl anion and radicals (B).*

Nonetheless, it is still possible, with appropriate care in conditions choice, to perform bioconjugation through CuAAC inside or on living cells.[102,103] A good 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. The first support ligand that showed to accelerate CuAAC for both synthetic applications and bioconjugations is tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine (TBTA).[104] After this, more water soluble ligands such as THPTA,[105] BTTAA [106,107] and TABTA[108] were developed. Ligands can protect copper(I) from oxidation, by complexation, and decrease its disproportionation rate in water, allowing the presence of more Cu(I) in solution overall. Moreover, ligands improve copper water solubility, and this accelerates the CuAAC reaction rate of almost 3 orders of magnitude in respect to the ligand free systems. This acceleration is also due to the ability of some ligands, particularly tridentate ones, to increase the rate of initial Cu(I)-acetylide formation; tetradentate ligands, not allowing this interaction, give slower reaction rates.[109] Nitrogen ligands can be classified based on the strength of their interaction with copper. Ligands with a stronger interaction can displace more easily solvent molecules and the azide substrate from the complex. This is useful when the solvent is a stronger donor than water such as DMSO or DMF. Ligands of Class I are characterized by weak donor abilities and are based on (triazolylmethyl)amine (Trz) skeletons. Class II ligands have two side arms exchanged either with benzimidazole (BimH) or pyridine (Py) units. this tends to give much stronger binding compared to the Trz ligands. By also replacing the third arm with BimH or Py, class III ligands are obtained. These binds so strongly that copper catalysis is inhibited. Accordingly, only ligands belonging to class I and II can be used to afford the efficient ligand accelerated CuAAC.

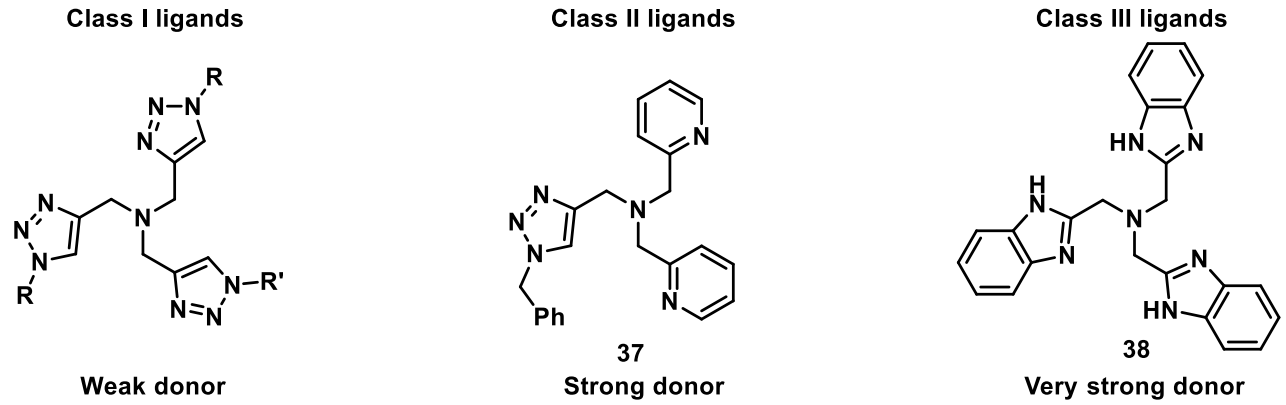


Figure 17. *Ligands commonly used during CuAAC reaction in aqueous solvents.*

Not all conjugation reactions are done in water and so, phosphine ligands have been developed to increase the solubility of copper catalyst in organic solvents like dichloromethane and toluene. They do so in the following order $(PPh_3)_3$ > PCy_3 > BINAP > dppe.[110] The Cu(I) is stabilized against oxidation by interaction with the phosphorous. However, the reaction rates are not as fast as the ones aided by nitrogen ligands and necessitate the addition of DBU (1,8-diazabicyclo[5.4.0]undec-7-ene), DIPEA (N,N-Diisopropylethylamine) and microwave irradiation.[109]

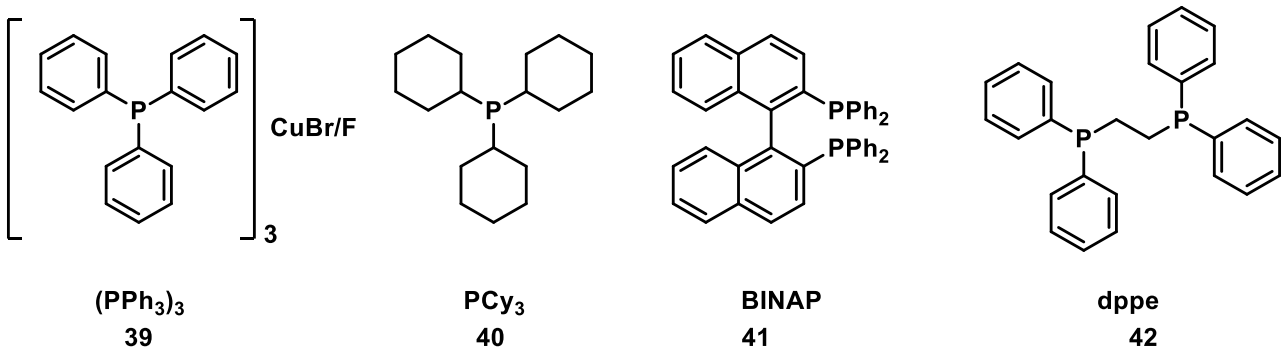


Figure 18. *Phosphorous ligands used for CuAAC. From left to right: triphenylphosphine complex with CuBr or CuF; tricyclohexylphospine; (2,2'-bis(diphenylphosphino)-1,1'-binaphthyl) (BINAP); 1,2-Bis(diphenylphosphino)ethane (dppe).*

RuAAC

Soon after having established the usefulness and scope of the CuAAC, Fokin, Jia and coworkers experimented with other transition metals to achieve the same reaction. Ruthenium is known to interact with alkynes, and they speculated that some ruthenium complexes could catalyse the Huisgen cycloaddition. This led to the discovery that Ruthenium does indeed provide a RuAAC reaction and more importantly, that some ligands direct its regioselectivity toward the 1,5 triazole product.[111] This gave an answer to the need of a reliable way to access the 1,5 regioisomer by a room temperature reaction; allowing large rings cyclization and more structure diversity for medicinal chemistry libraries.

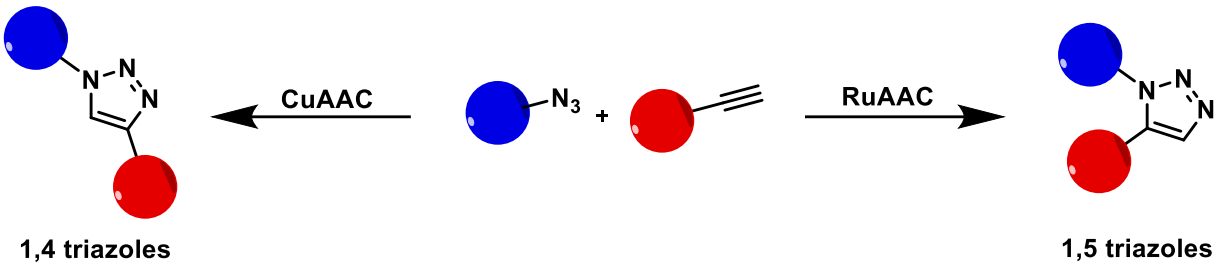


Figure 19. *Scheme representing the possible use of the same substrates in both CuAAC (left) and RuAAC (right), to give either 1,4 triazole or 1,5 triazole as linkers between the two conjugate moieties.*

In the study, not all ruthenium complexes performed well in term of yield and regioselectivity. The authors reported for example that phosphine-Ru complexes such as $RuCl_2(PPh_3)_3$ and $RuHCl(CO)(PPh_3)_3$ were quite ineffective, giving only a 20% yield of 1,4 triazole. On the other hand the addition of a cyclopentadiene to a phosphine complex, $CpRuCl(PPh_3)_2$, gave better yields and a mixture of 1,4 and 1,5 substituted triazole. Following this lead, complexes with the more sterically demanding pentamethylcyclopentadiene (Cp*), such as $Cp^*RuCl(PPh_3)_2$, gave complete conversion of the model reaction substrates to the 1,5 regioisomer. More catalysts gave similar results, among those $[Cp^*RuCl_2]_2$ and Cp*RuCl(COD) (COD = Cyclooctadiene). This last one gave the fastest reactions, although it is sensitive to oxygen and requires degassed conditions. All catalysts, however, were very tolerant toward reactive groups in the substrates and showed no reduction of reactivity in the presence of alkyl and aryl chlorides, boronic ester, alkenes, and polyalcohols, in both the azide and the alkyne reaction partners. Only in some cases, acidic functionalities such as carboxylic acids, boronic acids, and HCl/HBr salts of amines were found to be problematic.[112] About the reaction conditions required, microwave irradiation was often found to largely accelerate the reaction rate. This is not only due to an increased temperature; since the comparison with oil bath heating revealed that the irradiation technique was higher yielding and less prone to give byproducts. This also allowed an easier reaction optimization.[113,114] In general, many solvents can be used for this reaction: benzene, toluene, DCM, dioxane, N,N dimethylformamide (DMF) and tetrahydrofuran (THF). However, when pure protic solvents like water were used, the yield resulted substantially lower. On the other hand, the presence of water in aprotic solvents did not stop the reaction, which proceeded in good yields. A new and amazing feature of the RuAAC is that, unlike the CuAAC, is not limited to terminal alkynes and can give 1,4,5 trisubstituted triazoles.[111,115] This property is due the π activation of the vinyl group by the ruthenium, as opposed to the copper acetylide formation in the CuAAC and could be exploited to prepare bioconjugates with three different active substances.

Mechanism

Many researchers, studied the mechanism involved in this transformation and, in particular, the activation of alkynes, both terminal and internal, and the causes for regioselectivity. The studies rely on DFT (density functional theory) calculations and experimental results obtained by NMR, mass spectroscopy and X-ray crystal structure determination. The general mechanism for the RuAAC was first studied by Boren et al.[115] Through DFT calculations, the author reported four possible relative orientation of methyl azide and propyne when coordinated to the ruthenium atom of the complex [CpRuCl]. These orientations had very little difference in energy between them. Although two of these orientations would have brought to the formation of the 1,4 regioisomer, the lowest energy pathway brought to the 1,5 isomer and had its highest energy barrier at 13 kcal/mol; much lower than that of the uncatalyzed reaction. This demonstrated that ruthenium is a good catalyst for this reaction, reducing the energy barrier by several order of magnitude. This reaction was later revised in details and the calculations were

reiterated with different substrates and a Ru complex of pentamethylcyclopentadiene (Cp*) giving the same reaction paths as before and further explaining the selectivity toward the 1,5 triazole product.[116]

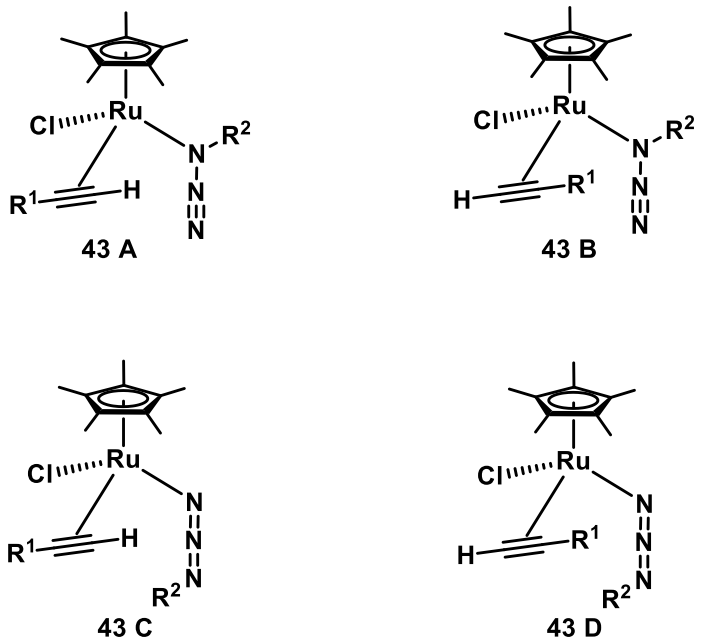


Figure 20. *Possible orientations of alkyne and azide complexed to ruthenium. Azide can be bound through the secondary nitrogen (A, B) or through the terminal nitrogen (C, D). The alkyne can have the terminal facing the azide moiety (A, C) or opposite to the azide (B, D).*

Concerning the mechanism, as a first step the Ruthenium complex frees two spectator ligands (in the case of this calculations phosphines) to accommodate both the azide and the triazole (**33**). Experimental proof established that both substrates have to be coordinated to start the catalysis. The presence of only azide actually inactivated the catalyst, for example (2-azidoethyl)benzene with Cp*RuCl(COD) has given inactive tetraazadiene complex.[117] With the two substrates present instead, there are four possible orientations of intermediate **33**. The ruthenium could coordinate the azide at the terminal nitrogen or the secondary one, but it is reported that the bonding proximal to carbon is more often observed.[118] The alkyne too could coordinate in two ways to the ruthenium; with the terminal carbon facing the azide or in the opposite direction; in this case the preference is determined by the alkyne substituents. Formation of complex **33** allows the terminal nitrogen atom on the azide group to attack the alkyne to form a ruthenacycle intermediate (**34**). In the next step, the alkyne group is separated from the metal and bonded to the azide, giving rise to the triazolide cycle **35**. In the last step the 1,5 triazole is released and the catalyst restored. The transition states involved in this mechanism have also been studied, along with possible orientations that could give rise to 1,4 substituted triazoles. These last orientations were found to have too high energy paths to be followed and thus the Cp* ruthenium complexes are confirmed to give 1,5 triazoles.

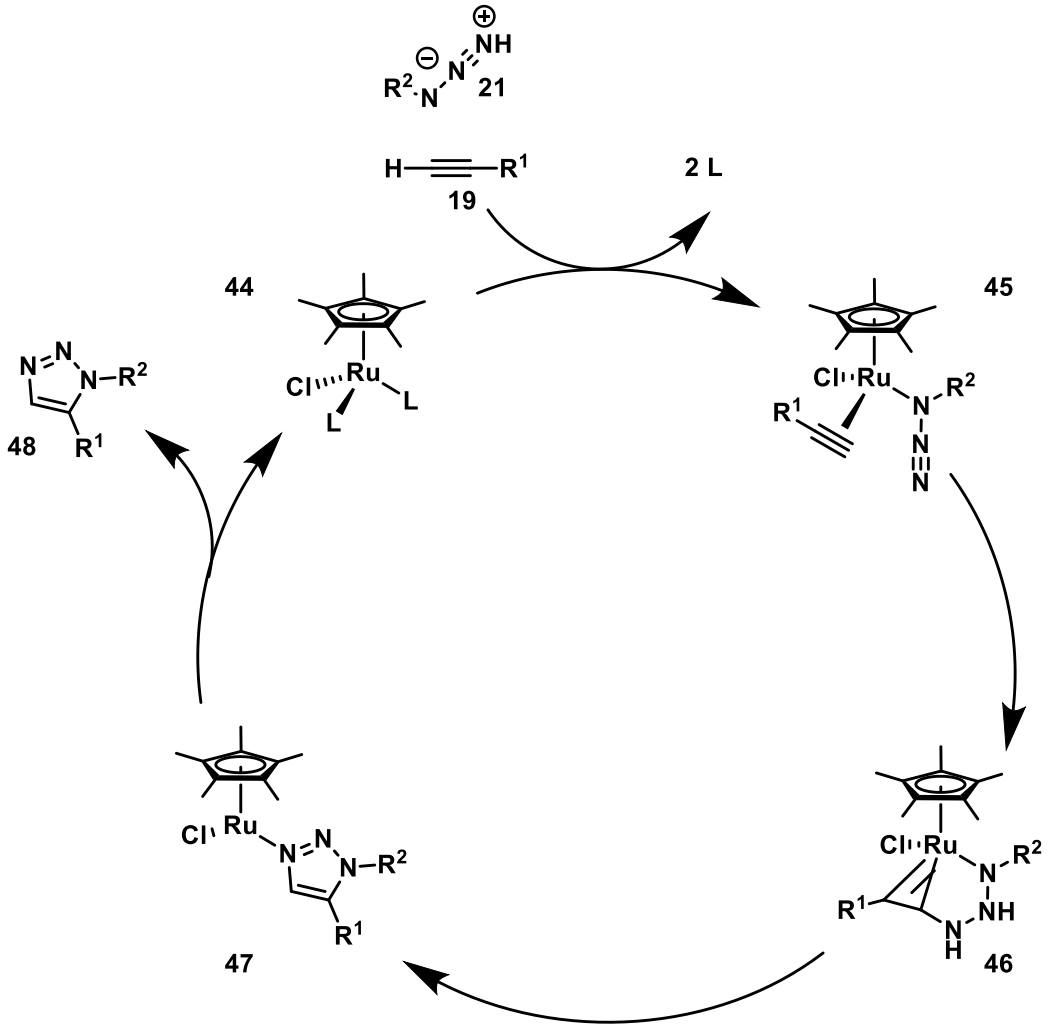


Figure 21. *Catalytic cycle of RuAAC.*[116]

As for the CuAAC the ruthenium catalysed formation of triazoles is of great synthetic value. It allows the orthogonal formation of a stable functionality with ease and consistent good yields in a good range of solvents. For chemical libraries application is a valuable complementary reaction to the CuAAC and offer the advantage of working also on internal alkynes. Unfortunately, the ruthenium catalysts do not work well in water; thus greatly restricting the bioconjugation potential of the reaction.

SPAAC

Another click reaction giving triazoles from alkynes and azides exist, and is particularly useful for bioconjugations in water and cellular environment.[119] It is the strain promoted alkyne azide cyclization. The development of this [3+2] cycloaddition was done by Bertozzi et al. in 2004.[120] No metal catalysis is needed for this reaction which is promoted by the strain energy of the alkyne in a small ring, such as cyclooctyne. As a consequence of the lack of catalysis, the reaction is much slower than the CuAAC (SPAAC k=0.0024 M^{-1} s^{1}; CuAAC k=10 M^{-1} s^{1}) but avoids all the problems mentioned above regarding the negative interference of copper in the cellular environment. Later development focused on reaction rate increase, this was achieved by using Dibenzocyclooctyne (DBCO) or bicyclo[6.1.0]nonyne (BCN).[121] Recently also modified aryl azides gave great increases in reaction rates.[122] Highly fluorinated azides, such as compound **51** have been recently developed to give a constant of reaction as high as 3.6 M^{-1} s^{-1} in protein conjugation.[123]

This reaction is thus biorthogonal and highly selective, although it usually gives two isomers unless symmetric substrates are used. Another disadvantage of this reaction is that the alkyne must come in quite a large moiety, and this can perturbate the chemical and physical characteristics of small molecule substrates. This can be less of a problem when conjugation of large entities is required. The SPAAC is now widely used in the biochemical and biomedical field, with great profit; especially for bioimaging of target of interest proteins.[124]

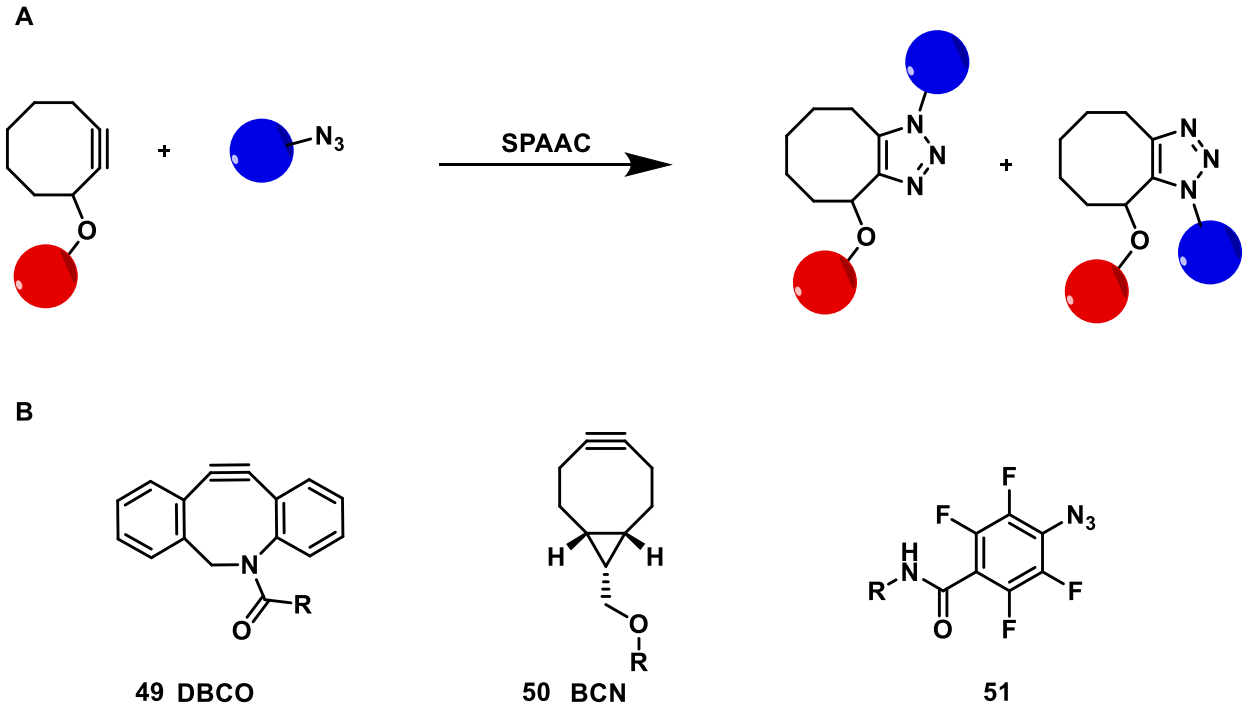


Figure 22. *Example of strain promoted azide alkyne cyclization (SPAAC) (A). examples of common cyclic alkyne structure used in SPAAC, Dibenzocyclooctyne (DBCO) and bicyclo[6.1.0]nonyne (BCN) and recently developed azide* ***18.***

Conclusion

New and powerful synthetic methodologies have been developed over the years to address many problems. The bioconjugation field has seen huge improvements over traditional techniques and chemistries with the advent of "click chemistry". The flagship reaction CuAAC, allows orthogonal conjugation of the most diverse and functionalised substrates at room temperature and in physiological conditions. Although a certain degree of optimization is required for every new bioconjugation, the development of new copper ligands promises faster and cleaner reactions on difficult and sensitive substrates and even *in vivo*. The 1,5 triazole regioisomer, inaccessible by CuAAC, can be obtained by ruthenium catalysis. This allows even more diversity to be obtained in chemical libraries and further expands the possibilities of bioconjugation. Despite the low yield given by RuAAC in water, most catalysts are robust and tolerate oxygen, as well as trace amounts of water in aprotic solvents. More importantly ruthenium gives access to 1,4,5 substituted triazoles previously not easily synthesized. Finally, the strain promoted azide alkyne cyclization, although slower than the CuAAC allows full biorthogonality, permitting functionalization in and on cells and even in living organism. The copper and ruthenium azide alkyne cyclization have been used extensively for this thesis work.

PART II: bile acids-small molecules conjugates

Introduction on bile acids

Bile acids

Bile acids origins

Bile acids are naturally present in many living animals, from nematode worms to mammalians. In humans, their biosynthesis starts in the liver with the catabolism of cholesterol to give, through many enzymatic steps, the two primary bile acids: cholic acid (CA) and chenodeoxycholic acid (CDCA). These can then undergo conjugation with either taurine or glycine amino acids 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. This functionality is often found as a sodium salt at physiological pH. These bile acids are stored in the gallbladder during fasting, where they make up a good part of the bile, which soluble fraction is composed of 67% them and other modified bile acids. The remaining part is 22% phospholipids (mainly lecithin), 4% cholesterol, 0.3% gall pigments and various proteins.[125] Bile is fundamental for the digestion process and during this phase is secreted into the small intestine where it absolves its emulsifying function and helps solubilize fats and liposoluble vitamins and other nutrients for intestinal absorption through micelle formation. After passive absorption in the intestine and active absorption in the ileum and colon, most of the bile acids are recovered and recirculated to the liver.[125,126] The active absorption of bile acids has been extensively studied and is ascertained to happen by means of a Na^+-dependent bile acid cotransport system. The proteins involved have been identified and found to have a high specificity for bile acids, in contrast to those present on liver cells membrane.[125,127–129] The enterohepatic circulation happens between 6 and 15 times a day and is very efficient, recovering almost 95% of the bile acids which and leaving only around 5% of the original portion to exit the body through the faecal excretion. This is attributed to Na^+ dependent Apical Sodium Dependent Bile acid Transporter (ASBT). The enterohepatic circulation allows the transport of around 10-40 grams of bile acids every day and is the main cholesterol homeostasis mechanism.

In the intestine and colon, primary bile acids can undergo transformation by the present flora. First, various hydrolases can remove the taurine or glycine moiety; then, secondary bile acids deoxycholate (DCA) and lithocholate (LCA) are formed by the action of bacterial 7α-dehydroxhylase. Oxidation of the C-7 position hydroxyl of chenodeoxycholic acids and successive reduction give the epimer ursodeoxycholic acid (UDCA), well-studied as a drug and of commercial interest. Further oxidases and reductases from the intestinal bacteria, as well as liver enzymes, can attack steroidal positions C-3, C-7 and C-12, leading to a complex pattern of tertiary bile acids; which composition depends on the prevailing bacterial population. Furthermore, in the liver, the hepatotoxic lithocholic acid is sulphated to increase its water-solubility and decrease its reabsorption: this results in the sulfolitocholic acid, another tertiary bile acid.[130] In Figure 23 are represented the main human bile acids along with their formation pathways. It is important to remember that the relative quantities of bile acids in the bile pool are around (~90%) of cholic acid, chenodeoxycholic acid and deoxycholic acid in a ratio of about 2:2:1; and that ursodeoxycholic acid and lithocholic acid are only found in the remaining 10% of the pool.[125]

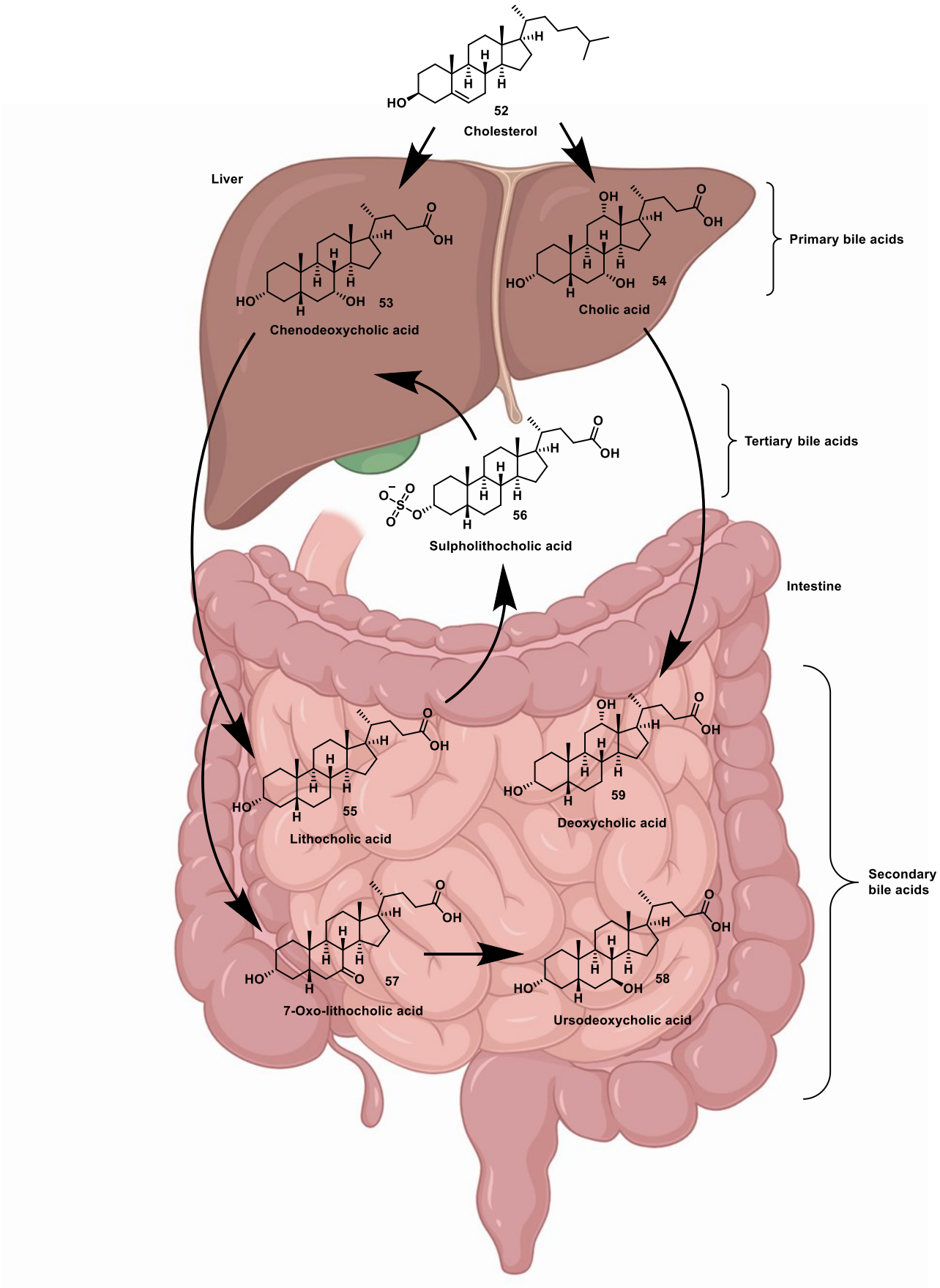


Figure 23. *Representation of the enterohepatic circulation. Primary bile acids (cholic and chenodeoxycholic acid) are biosynthesised in the liver from cholesterol and stored in the gallbladder, represented in green in the figure. During digestion these bile acids are poured into the into the intestine where they undergo transformation into the secondary bile acids (lithocholic, deoxycholic and ursodeoxycholic*

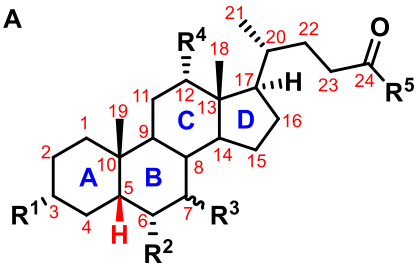
acid). secondary bile acids are then reabsorbed in the intestine and recirculated to the liver, where lithocholic acid undergo a further transformation to become the tertiary bile acid sulpholithocholic.

Curiously, bile acids names were commonly given form the animal species in which they have been isolate from the first time, or in which they are abundant. 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*).[131]

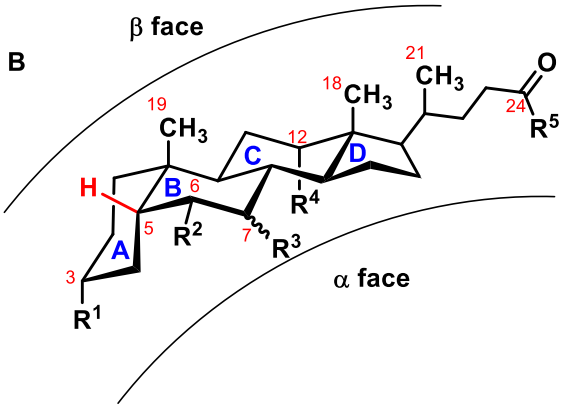
Bile acid structure and activity

Bile acids are a class of interesting molecules derived from cholesterol, their 24-atom carbon skeleton is a saturated tetracyclic hydrocarbon perhydrocyclopentanophenanthrene system, usually known as the steroid nucleus. Bile acids have methyl groups in positions C-10 and C-13; and a pentanoic acid side chain in C-17. As seen above there are many bile acids which are distinguished by the number and orientation of hydroxyl groups; these can be in positions C-3, C-6, C-7, C-12. The steroid scaffold is derived from cholesterol and features 8 stereocenters, not including the ones formed when hydroxyls are present. The scaffold is divided in 3 six membered rings, named A, B, C and one five membered ring named D. A distinctive trait of bile acids is the junction between rings A and B at C-5 which is *cis* and therefore gives the molecule a folded shape. This feature combined with the presence of hydroxyl groups and a carboxylic acid moiety gives the molecule a hydrophilic concave α-face and a lipophilic convex β-face. The lipophilic methyl substituents in positions C-10 and C-13 point toward the β-side, while the hydroxyls are oriented toward the hydrophilic α side. UDCA is the only naturally occurring bile acid which presents an exception in this regard since the C-7 hydroxyl is β-oriented. The detergent properties and biological relevance of bile acids are derived from this facial amphiphilicity which in turn is strongly influenced by number and orientation of hydroxyl groups.[132] The hydrophilicity of the common free and conjugated bile salts decreases in the order UDCA > cholic acid (CA) > chenodeoxycholic acid (CDCA) > deoxycholic acid (DCA) > lithocholic acid (LCA); and taurine-conjugated > glycine-conjugated > free species.[133] It is noteworthy that UDCA is more hydrophilic than cholic acid despite the latter having one more hydroxyl group. This is proof of how profound facial hydrophilicity influence on compound solubility is, since UDCA has a β facing hydroxyl in position C-7 that tend to disrupt the complete β face lipophilicity.

A



B



C

Name	R1	R2	R3	R4	R5
Cholanic acid	H	H	H	H	OH

Cholic acid	OH	H	OH	OH	OH
Chenodeoxycholic acid	OH	H	OH (α)	H	OH
Deoxycholic acid	OH	H	OH	OH	OH
Ursodeoxycholic acid	OH	H	OH (β)	H	OH
Lithocholic acid	OH	H	H	H	OH
Hyodeoxycholic acid	OH	OH	H	H	OH
Glicocholate	OH	H	OH	OH	$NHCH_2COO^-$
Taurocholate	OH	H	OH	OH	$NHCH_2CH_2SO_3^-$

Figure 24. *General structure of bile acids (A), position C-5 is evidenced in red, the junction between ring A and B is cis. General structure of bile acids in three-dimensional representation (B), the tree methyl group points all toward the β face, while the hydroxyl groups point toward the a face. In both representation the stereochemistry of position C-7 is omitted. The table (C) gives the names and substitution pattern of the most common bile acids. Glycine and taurine conjugate versions of every bile acid are possible, and cholate is only given as a prominent example.*

Bile acids have a length of approximately 20 Å and measure around 3.5 Å in diameter. As a consequence, they exhibit a great surface activity; and in aqueous solutions can form small aggregates or micelles of usually less than 10 monomers. Micelles only occur if their concentration as above a critical value called the critical micellar concentration (CMC).[134] Below the CMC, bile salts behave as 1 to 1 electrolytes, as demonstrated by freezing-point measurements.[135,136] The CMC value given for a certain bile is influenced primarily by its hydrophilicity; however, it is very depended on the method used to measure it, and in literature varying values have been proposed for the same bile acids. These values can be obtained by titration in water[133] or HPLC retention times[137]. As mentioned before these micelles are the main fat uptake mechanism in the intestine. 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 favour its elimination through the faeces. The emulsifying properties of bile acids play a major role in prevention and treatment of gallstones and other diseases. Gallstones can be formed by the excessive accumulation of cholesterol in the gallbladder and the presence of bile acids, particularly the most hydrophilic ursodeoxycholic acid and its taurine conjugate, helps dissolve the precipitate.[138]

UDCA is known in modern medicine since the seventies; however, bear bile containing mostly ursodeoxycholic acid, has been used for more than 3000 years in Asia and is reported as a remedy against hepatobiliary conditions since the first Chinese pharmacopeia (Tang Ban Cao).[139] UDCA, has well reported anti-inflammatory and cytoprotective properties.[140] Moreover, UDCA and its taurine conjugate TUDCA have been reported to have a neuroprotective role in the context of neurodegenerative disorders such as Alzheimer and Parkinson's disease.[141] In addition, ursodeoxycholic acids is used, due to its better toxicity profile in respect to other human bile acids, as treatment for primary biliary cholangitis[142] and to improve the liver biochemistry in primary biliary cirrhosis.[143] Furthermore, new uses of UDCA are being discovered due to its cytoprotective powers. For example, in a recent study it was found to help protect

against intestinal barrier breakdown in mice. This effect is given by its ability to help epithelial cells migrate and thus close the open wound. The author report that this effect is derived by the activation of various receptor implicated in cell migration among which cyclooxygenase 2 (COX-2) (also implicated in the inflammatory process) and epidermal growth factor receptor (EGFR).[144]

Apart from the lipid dissolving properties and classical use as drugs, bile acids have many roles in the body. They regulate the cholesterol homeostasis and have been found to serve specific, hormone-like functions in the control of glucose, lipids, and energy metabolisms.[126]

In 1999 bile acids have been found to interact with the transcription factor Farnesoid X receptor (FXR; also known as bile acid receptor).[145,146] This interaction is the basis of the mechanistic framework for bile acid sensing. Indeed, bile acids self-regulate since the FXR activation directly control the transcription of many genes, some of which, in turn, encode for proteins involved in bile acid synthesis, transport, and metabolism.[126,147] Expression of the FXR protein is high in organs that constitute the enterohepatic circulation (liver and intestine), but it is also present in white adipose tissue, kidney, adrenal gland, stomach, pancreas, endothelial cells, vascular smooth muscle cells and cells of the immune system. This distribution pattern indicates a broad spectrum of biological functions that seems to potentially involve cell types that lack high concentrations of bile acids. Not all bile acids act in the same way on this receptor; for example, highest activation is given by chenodeoxycholic acid, followed by deoxycholic acid, lithocholic acid and cholic acid.[145,146]

Shortly after the discovery of FXR a membrane G-coupled protein bile acid receptor was found, TGR5.[148,149] The receptor is highly expressed in liver cells other than hepatocytes, including Kupffer cells and cholangiocytes, and also in gallbladder epithelial cells and immune cells.[150] In addition, TGR5 is expressed in brown adipose tissue, the enteric nervous system, the central nervous system and in muscles.[151] It's activation by bile acids bring to its internalization with consequences variable depending on the tissue, in general it brings to anti-inflammatory effects, increased intestinal motility, and improved glucose metabolism and insulin sensitivity.

Given the presence of membrane receptor for bile acids in many cell types, but mostly in the liver and intestine, is possible to envisage bile acid conjugation as targeting agents to other drugs. This field of research has given promising results.

Bile acids conjugates

The treatment of chronic diseases in most cases involves long-term or even life-long use of drugs. Two aspects of long-term drug therapy are of major importance for patients and physicians: a site-specific drug action without adverse side effects; and non-invasive, preferably oral, administration of the drugs. The physiology of bile acid transport described above, with the predominant involvement of the liver and the small intestine, should therefore be ideally suited for the use of bile acids as putative shuttles of pharmaceuticals. Moreover, cytoprotective and anti-inflammatory UDCA is an appealing substrate for prodrug formation. The term prodrug refers to inactive derivatives of drug molecules which undergo enzymatic or biochemical transformation to release active pharmaceutical ingredients (API) *in vivo*. Over the past decades many bile acid-drug conjugates were synthesized with the aim of targeting intestinal and hepatic cells through Apical Sodium Dependent Bile acid Transporter (ASBT). Bile acids offer a range of possible conjugation sites on the molecule; esterification and amidation are possible at position C-24, while hydroxyls groups in position C-3, C-6, C-7 and C-12 can be used as nucleophiles to form esters or substituted with halogens to later give stable ethers; in addition, they can be oxidized or converted into amines. Modification can be done with direct drug conjugation or with the use of linkers. The

internalization mechanism of bile acids through ASBT has been studied, and several mechanisms have been proposed. Starting from a simple Na^+ regulated conformation switch, to a more complex three-dimensional rotation of the bile acids binding domain.[152,153] Although the actual process is not yet completely understood, it is known that ASBT follows an endocytosis process by multivalency interactions with bile acids and enters the subcellular organelles, and then returns to the membrane after having released the bile acid, in a still not precisely known way.[154,155] Many doubts arise also on the relationship between the bile acids structure and its uptake efficiency in liver and intestine. During the years, a good number of studies have shown discordant results in this regard. Lack and Weiner published the first results of a 1960's study where it is clear that it is possible to modify the bile acid through conjugation, but it is important to leave one negative charge in the side chain. In addition, it was found that tri-hydroxyl bile acids are better transported than di-hydroxyl ones; and that the transport a particular bile salt depressed the transport of another.[156] Later, it was found that the affinity and uptake rates of the bile acid analogues were primarily determined by the substituents on the steroid nucleus. Two hydroxyl groups at positions C3, C7, or C12 were optimal, whereas the presence of a third one decreased the activity, contradicting the previous study. Moreover, the affinity for C-6 hydroxylated bile acids was found to be high in hepatic cells and lower in the ileum tissue.[157] Based on this, and another, more theoretical pharmacophore model study,[158] Kramer proposed an adjourned mode of interaction of bile acids with the bile acid ileal cotransporter. In this model there are 5 important interaction sites: two hydrogen bond and three hydrophobic sites. One hydrogen bond acceptor is given by the negatively charged side chain; hydrogen bond donors are given by the α hydroxyl groups in position C-7 or C-12. Two out of three hydrophobic sites are present for cholyltaurine, and the other bile acids tried in the simulation. These are represented by the five membered ring D and the C-21or C-18 methyl groups. The authors emphasise that the *cis* junction between ring A and B and C-3 α hydroxy moiety are not necessary features for substrate recognition, thus advocating for drug conjugation at C-3 position. On the other hand, it was recently proved that negative charge at C-24 is not a prerequisite for transport, conjugating acyclovir to this position, *via* a valine linker, on both UDCA and CDCA.[159]

Kramer and Wess research group reported a series of drug-bile acids conjugates over the years. For example, the alkylating agent chlorambucil, used as a cytostatic in tumour treatment, was conjugated to the β C-3 hydroxyl of cholic acid through an ether linkage with an appropriate ω-aminoalkoxy linker. This combination gave strong inhibition of hepatic taurocholate transport, but the conjugate was adsorbed in rat ileum and secreted in the liver, retaining the alkylating properties of the drug. The authors also established that for active transport to occur the side chain of the bile acid must be free (or conjugated with taurine/glycine) as the methyl ester analogue did not give inhibition. The conjugate also gave uptake in hepatocellular carcinomas which expressed bile acid transport proteins (Sodium taurocholate co-transporting polypeptide (NTCP), both sodium dependent and independent.[160–162]

Other drugs were conjugated, in particular small peptides are very susceptible to degradation in the body. Especially if taken orally. In an attempt to increase the half-life of a collagen biosynthesis inhibitor, an oxapropylpeptide and its ter-butyl analogue were conjugated to the same bile acids as before. It was possible to demonstrate that intestinal adsorption occurred through the bile acid transport system and that the peptide conjugate was more resistant to degradation than the parent drug, as well as exerting the same function.[163]

Statins are normally used to lower the levels of LDL cholesterol (low density lipoproteins carriers of cholesterol) levels in blood. In this example the statins HR 780 and lovastatin were conjugated to bile acids in a prodrug approach. The drugs act by inhibition of the target HMG-CoA reductase enzyme (3-

hydroxy-3-methyl-glutaryl-coenzyme A reductase) the rate controlling of the mevalonate pathway producing cholesterol and isoprenoids. However, some adverse effects have been reported especially in muscle cells, due to the wide drugs distribution. The authors speculated that bile acids conjugation could target these drugs in the liver thus avoiding the undesired side effects. Selectivity which was confirmed by *in vivo* experiments.[164]

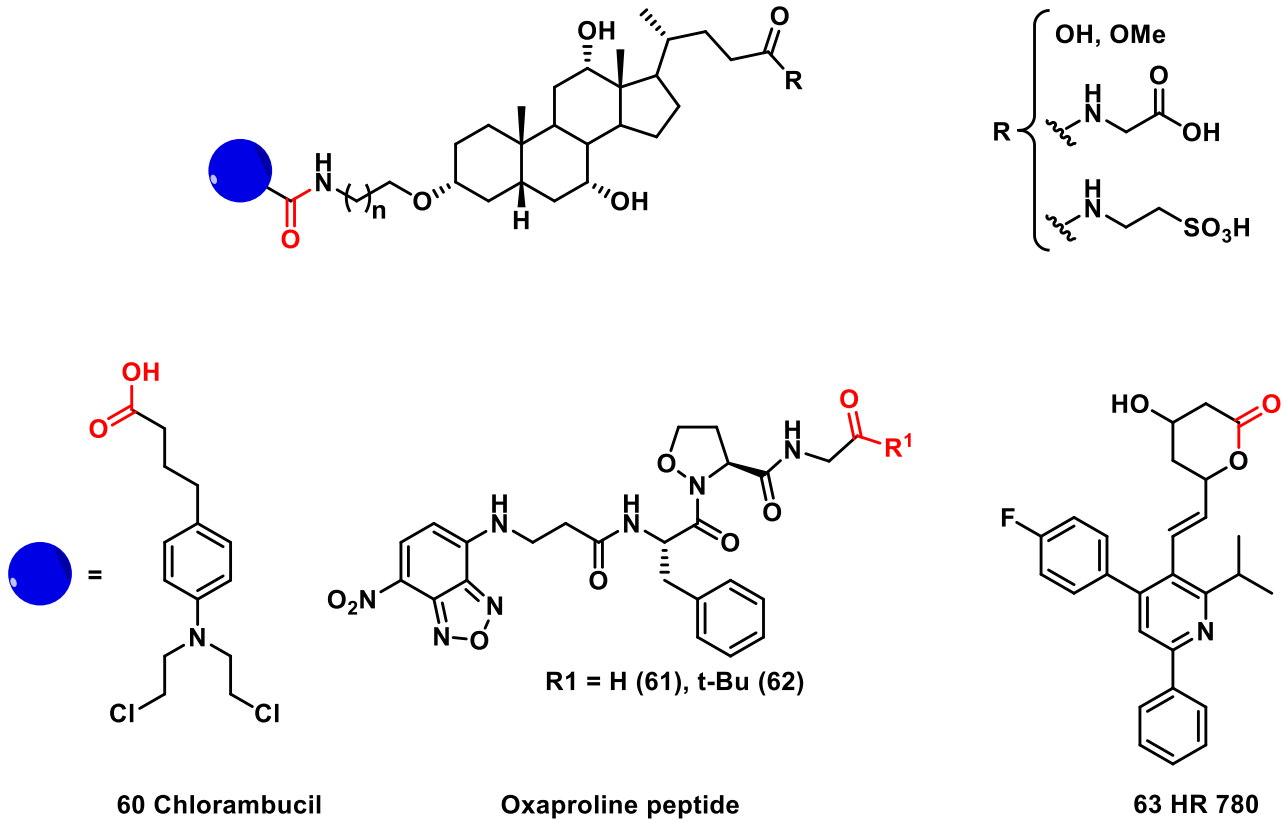
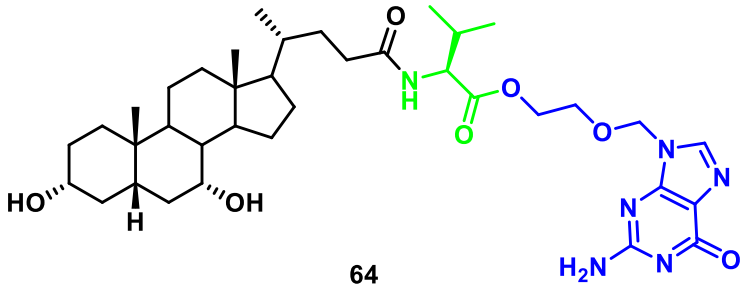


Figure 25*. Drugs conjugated to bile acids in order to exploit the prodrug targeting in the liver by active transport through ASBT.*

As mentioned above, conjugation can be fruitful also in position C-24, a recent example is the one provided by the conjugation of acyclovir to CDCA through a valine linker. Acyclovir inhibits herpes virus proliferation, and the conjugate valine-acyclovir is already a prodrug by itself that should improve the intestinal absorption of the parent drug. Four bile acids were used to prepare conjugates: CDCA, CA, DCA and UDCA. Of these, the chenodeoxycholate derivative exhibited the highest affinity for ASBT, comparable to that of cholic acid alone. The authors also proved that the parent drug was liberated by catalytical cleavage *via* esterase. The conjugate showed a higher uptake, both passive and active in respect to acyclovir and it was also possible to increase the bioavailability by twofold after oral administration.[159]



64

Chenodeoxycholic-valine-acyclovir

Figure 26. *Structure of the bioconjugate chenodeoxycholic acid-valine and acyclovir. Valine is represented in green while the acyclovir moiety is in blue.*

Synthesis and Biological Investigation of Bile Acid-Paclitaxel Hybrids

The work from this thesis was previously published[165] and some figures are extracted from the published material.

Introduction

Paclitaxel

Paclitaxel (PTX) or taxol is a cancer drug derived from the bark of the Pacific yew tree that has been shown to be effective in treating a variety of cancers. Its anticancer activity is linked to its ability to promote microtubule assembly, which leads to mitotic arrest.[166–170] With annual sales of more than $1 billion, PTX, which was approved by the Food and Drug Administration in 1992, is one of the most affordable and best-selling chemotherapy drugs. Many studies have been conducted over the years to obtain smarter targeting PTX, in order to overcome some limitations towards a broader framework of clinical application. The low water solubility of PTX drew a lot of attention in the early stages of its development studies. As a result, various prodrug strategies have been developed to improve the efficacy of PTX.[171] The critical role of taxoids in anticancer therapy, as well as their significant side effects and drug resistance, have prompted numerous efforts to investigate the structure-activity relationship (SAR) of such complex molecules. SAR studies have revealed that the C13 side chain, ester groups at C2 and C4, and the rigid core represented by the oxetane ring (Figure 27) are all inextricably linked to taxoids' biological activity.[172–176]

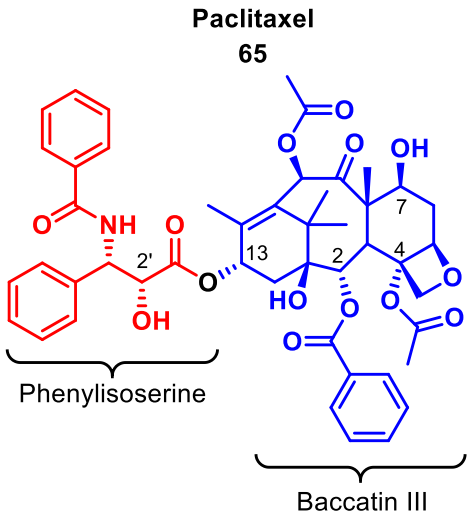
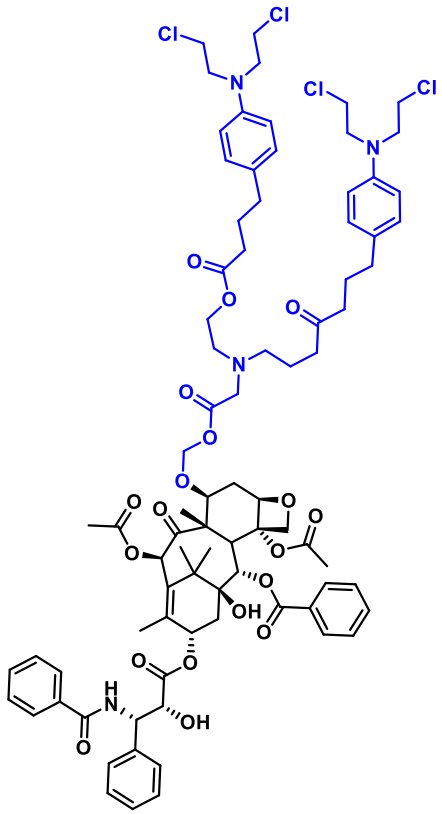


Figure 27. *Structure of taxol (paclitaxel) in red is highlighted the phenylisoserine side chain, while the baccatin core is blue coloured.*

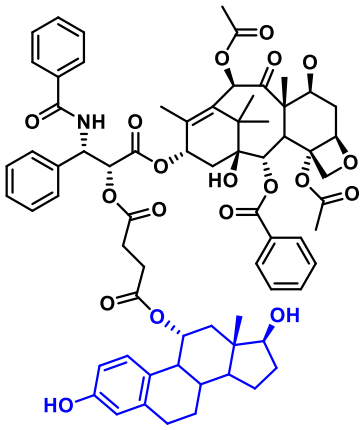
Paclitaxel modification

This discovery has prompted researchers to develop new PTX derivatives. Many efforts have been made to synthesize modified PTX on both the baccatin core and the phenylisoserine side chain. Kingston et al.[177] published a report on the synthesis and biological evaluation of a series of PTX-steroids conjugated at the 2'-OH against a variety of solid tumors. All new conjugates were found to be active against the ovarian cancer cells A2780, albeit at a lower concentration than PTX alone. Wittman et al.[178] reported the conjugation of PTX-chlorambucil hybrids at the baccatin core. The conjugate was successfully tested in vitro and in vivo in M109 and PTX resistant M109/taxlR cancer cells. A much less successful approach, on the other hand, is to replace the baccatin core with synthetic or natural products.[179–181] The new taxoid

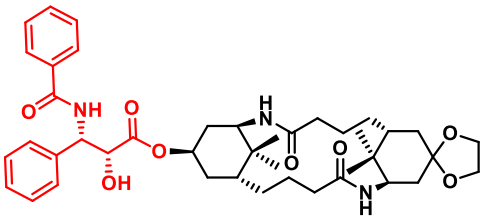
mimics (**68, 69**) obtained this way failed to outperform the parent PTX in both microtubule disassembly assays and cytotoxicity against a variety of cancer cell types.



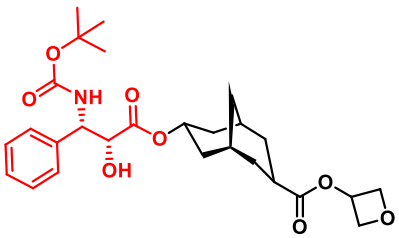
66 Paclitaxel-chlorambucil conjugate



67 Paclitaxel-estradiol conjugate



68 Botta and Corelli simplified structure



69 Zefirova's simplified structure

Figure 28*. Example structure of paclitaxel conjugates, the conjugate molecule is highlighted in blue. Examples of structures in which the baccatin is substituted by another structure, phenyl and terbutyl isoserine are red coloured.*

Many attempts have been made over the years to improve PTX anticancer activity and selectivity through conjugation, and research in this area has never stopped. In this light, our research team planned to conjugate PTX to two different bile acids in order to investigate the biological activity of BA-PTX in comparison to the parent PTX. As PTX conjugation partners, chenodeoxycholic acid (CDCA) and ursodeoxycholic acid (UDCA) were chosen. As previously stated, CDCA and UDCA differ in their absolute configuration at C7-OH and exhibit different physical-chemical properties, with CDCA being more lipophilic than UDCA.

Bile acids and their conjugation to paclitaxel

Bile acids, among other things, have the ability to inhibit cell proliferation on several cancer cell lines via a variety of mechanisms, including apoptosis, membrane alterations, nuclear receptor modulation, and oxidative stress.[182] Some reported evidence suggests that the hydrophobicity of bile acids is responsible for their cytotoxicity.[183] Furthermore, depending on the circumstances, the more hydrophilic bile acids, such as UDCA, can function as either anti- or pro-apoptotic molecules.[183,184]

However, the relatively low intrinsic cytotoxicity of bile acids (IC50 > 100 M) precludes clinical use and suggests conjugation. On the other hand, because of their low toxicity and biocompatibility, they have sparked considerable interest in drug delivery research. PTX delivery and selectivity have been improved by using conventional drug carrier systems, for instance, encapsulation into nanosystems such as nanomicelles[185] and solid lipid nanoparticles,[186] or polymers conjugation.[187] Bile acids, because of their amphiphilic properties, can improve drug bioavailability through the process of micellar solubilization and act as permeation-modifying agents.[188] As a result, bile acid conjugation with drugs with low aqueous solubility or membrane permeability can be effective in increasing bioavailability. In general, the hydrophobicity of bile acid is directly related to its increased bioavailability and oral adsorption. Furthermore, because BA can withstand enzymatic and gastric degradations and promote active absorption via the ileum apical sodium-dependent bile acid transporter (ASBT), chemical conjugation of drugs to BA may improve the oral bioavailability of parent compounds.[188]

Given the immense literature on paclitaxel modification and conjugation, suggesting the importance of the baccatin core and the phenylisoserine moiety for efficacy, our group created new PTX hybrids by preserving the baccatin core and using the 2'-OH of phenylisoserine moiety as a conjugation point of paclitaxel. The conjugation in that position should result in an inhibition of the efficacy of paclitaxel as long as it is conjugated. Consequently, the conjugates synthesized in this work can be considered prodrugs. On the other side of the hybrid, two distinct hydrophobic bile acids, CDCA and UDCA, were connected at the C24 position by a condensation reaction, generating the corresponding BA-PTX ester derivatives. The cytotoxicity of BA-PTX conjugated was evaluated towards two leukemic (HL60 and NB4) and two colon carcinoma (RKO and HCT116) cell lines.

The same conjugates bearing fluorescent probes were synthesized to better investigate the biological activity of BA-PTX hybrids. To create more drug-like probes, conjugated Pacific Blue (PB), a coumarin-based fluorophore, was conjugated to derivatives of both BA-PTX hybrids using the hydroxyl group at position C7 of the baccatin core, as described in the literature.[189]

Discussion

Synthesis of BA-PTX Hybrids

A condensation reaction between UDCA and CDCA and PTX, mediated by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCI) as a water-soluble coupling reagent, was used to prepare hybrids UDC-PTX (**70**) and CDC-PTX (**71**) (***Figure 29***).

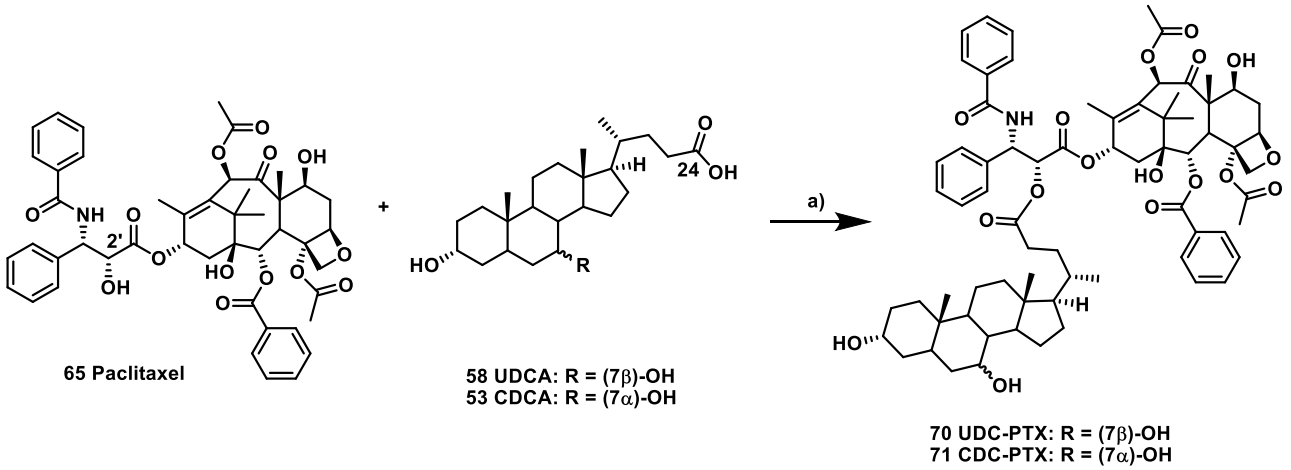


Figure 29. *Synthesis of UDC-PTX and CDC-PTX.BA hybrids by condensation reaction. (a) EDC hydrochloride, DMAP, DMF, 25 °C, 16 h.*

After chromatographic purification, the desired BA-PTX compounds were obtained in high yield.

Synthesis of BA-PTX-PB and PTX-PB

We chose Pacific Blue as a fluorophore to synthesize a fluorescent analogue of our hybrid compounds because of several properties of this coumarin dye. Essentially, as previously reported by Lee et al,[189] we determined that this small, non-polar fluorophore had no effect on hybrids' ability to enter cells. Furthermore, Pacific Blue is a bright blue, fluorescent dye that is best excited by the 405 nm line of violet laser in flow cytometry and has good photobleaching resistance.

The initial attempt at the synthesis saw the regiospecific protection of the bile acids UDCA and CDCA with tertbutyl-dimethyl-silyl chloride to give an easily removable group on the most reactive bile acid hydroxyl. To obtain the protected starting material a procedure from literature was followed.[190] However, the product was only obtained in low yields. After successful purification of the product, it was possible to conjugate the bile acid with taxol. This was done by esterification at the 2' position hydroxyl of PTX employing DCC (N,N'-Dicyclohexylcarbodiimmide) and DMAP as coupling agent and catalyst. It is important to take note that taxol is sensitive to basic conditions since the β keto hydroxy group in position 9 and 7 can undergo retro-aldol reaction and reform the baccatin core with inverted stereochemistry at the 7 positions. This epimerization must be avoided as it can negatively impact the efficacy of the drug. However, DMAP is not a strong base and has been used extensively in reactions with taxol. The retention of configuration was confirmed by ^{1}H-NMR analysis of the reaction product. At this point it was attempted to conjugate the 7-hydroxyl taxol position with Fmoc protected glycine with the same procedure as reported in literature.[189] Unfortunately, it was not possible to obtain the product as the reaction was not proceeding on both the UDC-PTX and CDC-PTX conjugate. It was hypothesized that maybe the fluorenylmethyloxycarbonyl steric hindrance, combined with the unknown spatial conformation of the new hybrids did not allow the reaction to occur. The low yield (37%) reported in literature for the same reaction on PTX alone corroborated this hypothesis. Therefore, coupling with Boc protected glycine was attempted. This gave the desired conjugates, in low yield. However, taxol is also sensitive to acid conditions and when deprotection of the boc-glycine conjugate with a mixture of DCM and trifluoro acetic acid was attempted complete degradation of the starting material was observed. At this point the synthetic strategy was revised to start form the PTX conjugated with glycine and then conjugate it with the bile acids. For this way a protection-deprotection approach of the 2' hydroxyl position of taxol was needed.

Following PTX silylation, the fluorenylmethoxycarbonyl (Fmoc) glycine derivative was obtained with improved yields when DCC was used in place of the previously reported EDCI hydrochloride coupling reagent.[189] The resulting dicyclohexylurea side-product, on the other hand, proved to be more difficult to remove from the reaction mixture than the water-soluble analogue EDC derived urea. The 2'-O-silyl group of compound **72** was removed using a hydrogen fluoride-pyridine (HF-Py) solution, this allowed to perform the conjugation with the carboxylic groups of unprotected UDCA and CDCA. The condensation reaction took place at room temperature within three hours in the presence of DCC, affording esters **74** and **75** in 86 and 74% yields, respectively, after chromatographic purification. After removing the Fmoc protecting group from glycine under mild conditions, amidation with Pacific Blue-N-hydroxysuccimide (PB-NHS) ester produced satisfactory yields of the desired fluorescent UDC- and CDC-PTX-PB hybrids **76** and **77**. Pacific Blue hybrids were purified using reverse-phase high performance liquid chromatography (HPLC) to achieve the purity level required for biological tests.

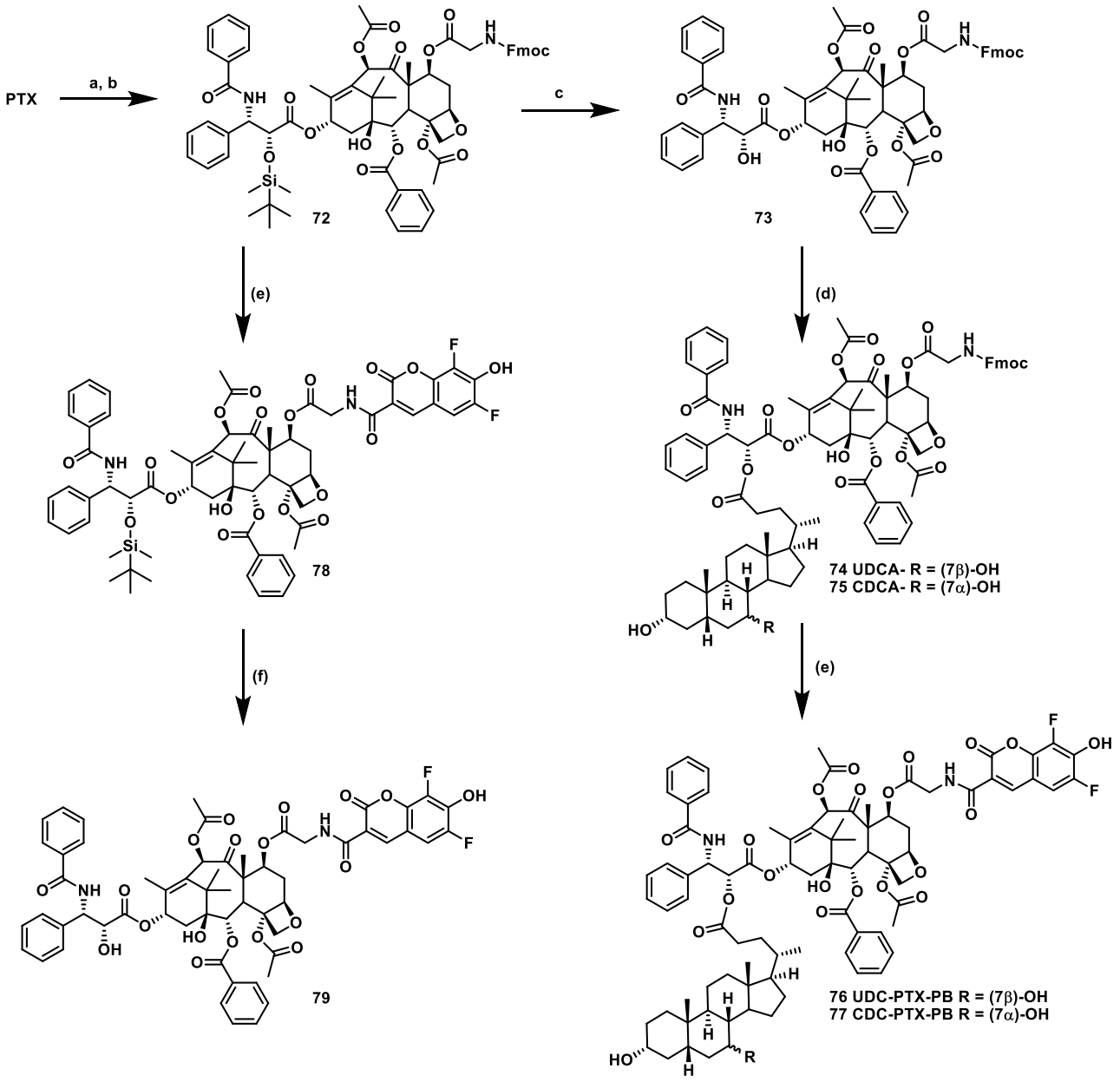


Figure 30. *Synthetic strategy for the compounds PTX-PB, UDC-PTX-PB and CDC-PTX-PB. (a) TBDMSiCl, Imidazole, Pyridine, DMF, rt, 3 h. 80% yield. (b) Fmoc-Gly-OH, DCC, DMAP, DMF, 35 °C, 16 h. 90% yield. (c) HF-Py, THF, 3 h.*

60% yield. (d) UDCA or CDCA, DCC, DMAP, DMF, rt, 3 h. 86% or 74% yield. (e) (1) DMF/Piperidine 20%, rt, 10 min. (2) PB-NHS, DIPEA, DMF, rt, 16 h. 63% yield. (f) TBAF, THF, 22 °C, 90 min.

The synthesis of PTX-PB depicted in Figure 30 was performed following the Lee et al. procedure.[189] Starting with compound **72**, amidation with Pacific Blue-NHS ester was accomplished by removing the Fmoc group, and the final desilylation step was performed using a tetrabutylammonium fluoride (TBAF) solution.

Chemical Stability of BA-PTX hybrids

HPLC analyses were used to determine the chemical stability of PTX, CDC-PTX, and UDC-PTX in MeOH/water solutions at pH = 3 and 8. All compounds tested were found to be stable in the pH range tested for up to 96 hours. Likewise, the chemical stability of PTX, CDC-PTX, and UDC-PTX in DMEM cell culture medium was investigated. A time course at 0, 4, and 24 hours revealed that all compounds were stable in cell culture medium as well.

Biological Evaluation

Comparative Effects of PTX and PTX Derivatives on Tumor Cell Lines

PTX derivatives were tested for cytotoxicity in leukemic (HL60 and NB4) and colon carcinoma (RKO and HCT116) cell lines. CDC-PTX, at a concentration of 2 M, caused a significant reduction in cell viability on HL60 and NB4 cell lines after 24 and 48 hours of treatment, as shown in Figure 31 (A, B) and this effect was comparable to that of PTX. In contrast, UDC-PTX derivative had no effect on NB4 cell viability at any of the time points studied but did slightly reduce HL60 viability at both 24 and 48 hours after treatment. MTT assay data analysis confirmed these findings. After 48 hours, PTX and CDC-PTX induced a significant and comparable apoptotic effect on HL60 and NB4 leukemic cells.

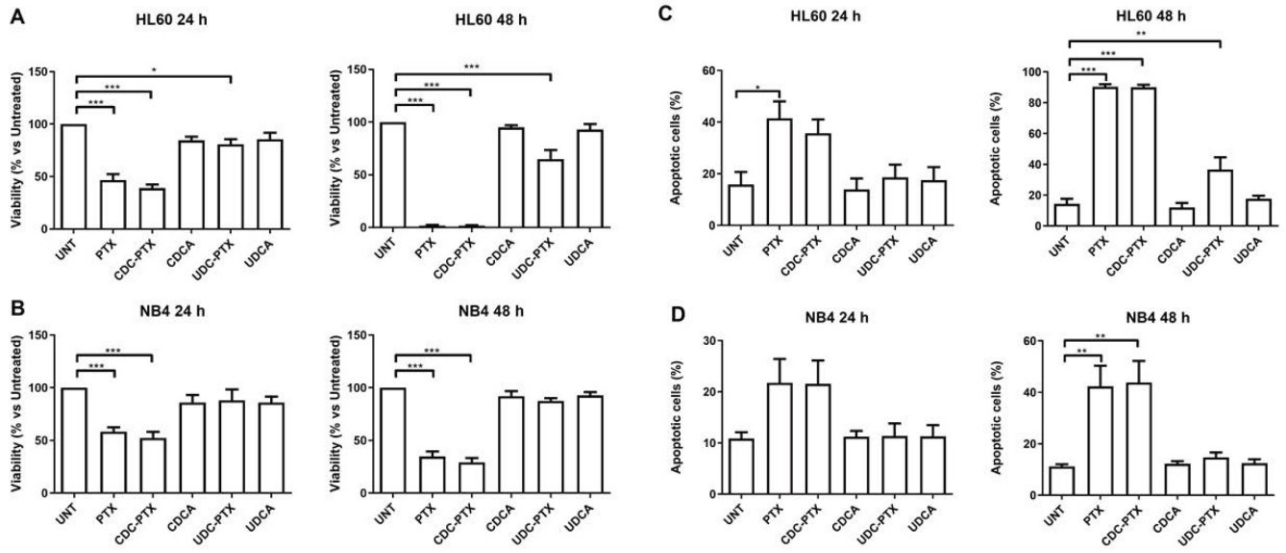


Figure 31. *Cell viability and apoptosis evaluation in HL60 (A,C) or NB4 (B,D) cells treated for 24 or 48 h with 2 μM of each of the following compounds: PTX; CDC-PTX, CDCA, UDC-PTX or UDCA. Data are reported as mean ± SEM of at least 3-independent experiments. Statistical analysis was performed by ANOVA followed by Bonferroni post hoc test for pairwise comparisons. * $p < 0.05$; ** $p < 0.01$ or *** $p < 0.001$*

The UDC-PTX derivative, on the other hand, had no significant apoptotic effect on RKO, but it appeared to be significantly effective in inducing apoptosis in HCT116 and inducing a reduction in cell viability on both cell lines, albeit much less than PTX, as indicated by the significance of the viability and apoptosis data of PTX vs. UDC-PTX.

It should be noted that neither CDCA nor UDCA were effective in inducing cytotoxicity in leukemic or colon carcinoma cell lines.

The data described for HCT116 cell lines were also supported by xCELLigence cell viability/proliferation analyses performed with 5 μM PTX and its derivatives. As shown in Figure 32, treatment with PTX and CDC-PTX resulted in a low significant difference in reducing the Cell Index compared to untreated cells (set to 1) only at 24 h, with no significant differences at 48 or 72 h. Furthermore, the effect induced by UDC-PTX, while significant in comparison to untreated cells, was lower in comparison to PTX and CDC-PTX, confirming the data obtained by viability and apoptosis evaluation.

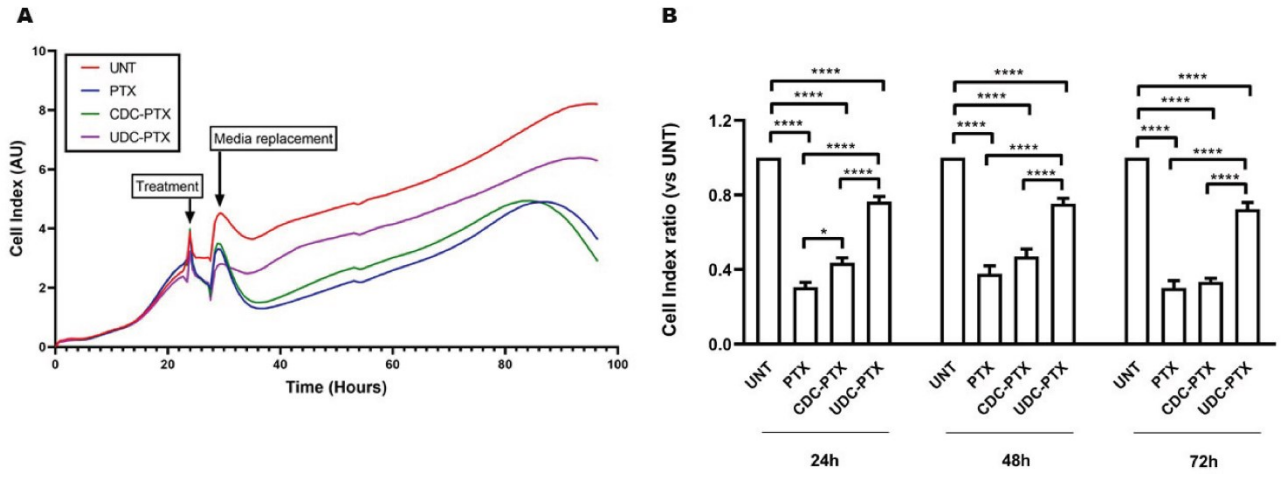


Figure 32*. Time course effects of PTX and of its derivatives on HCT116 evaluated by the xCELLigence RTCA DP Instrument. (A) Representative graph from RTCA software reporting real time cell index. (B) Analysis of HCT116 cell viability after treatment with 5 μM PTX, CDC-PTX-UDC-PTX for 4 h. Results are indicated as mean ± SEM of two replicates performed at least in quadruplicate. Statistical analyses were performed 24, 48 and 72 h after treatment with ANOVA and corrected with Bonferroni post-test. * $p < 0.05$; **** $p < 0.0001$.*

The CDC-PTX derivative Showed Lower Toxicity with Respect to PTX on NIH-3T3 healthy Cell Line

The cytotoxic effect of PTX and its derivatives on the NIH-3T3 fibroblast cell line, which served as the study's healthy control, was investigated in the following series of experiments. After 4 hours of treatment and 72 hours incubation in fresh medium, 5 μM PTX significantly reduced cell viability. At the same concentration and experimental conditions, CDC-PTX and UDC-PTX caused a low cell viability reduction that was not comparable to that caused by PTX. Furthermore, consistent with the cell viability results, PTX induced a potent apoptotic signal, resulting in a significant increase in the apoptotic rate, whereas BA-PTX derivates did not induce apoptosis. Even in this case, CDCA and UDCA had no effect on cell viability.

Condensation of Bile Acids to PTX Ameliorated Its Incoming into the Tumour Cells

In the final set of experiments, HCT116 cells were treated with Pacific Blue-conjugated compounds alongside non-conjugated compounds to test for potential differences in the ability of the different hybrids to pass through the plasmatic membrane. After treatment with CDC-PTX-PB and UDC-PTX-PB, the percentage of PB positive cells reached 99.92% ± 0.02 and 99.88% ± 0.04 respectively, while PTX-PB treatment caused the staining of only 34.69% ± 6.50 of the cells, as clearly shown in figure 33. These findings indicated that bile acid condensation to PTX significantly improved its entry into HCT116 cells, allowing the derivatives to enter the totality of the cells. Furthermore, the significantly different fluorescence intensity of cells treated with CDC-PTX-PB versus UDC-PTX-PB, as indicated by the MFI

(median fluorescence intensity) of these two cell populations, suggested that CDC-PTX-PB may have a greater ability to cross the plasmatic membrane than UDC-PTX-PB. Cells treated with non-fluorophore bearing compounds showed no fluorescence signal, as expected.

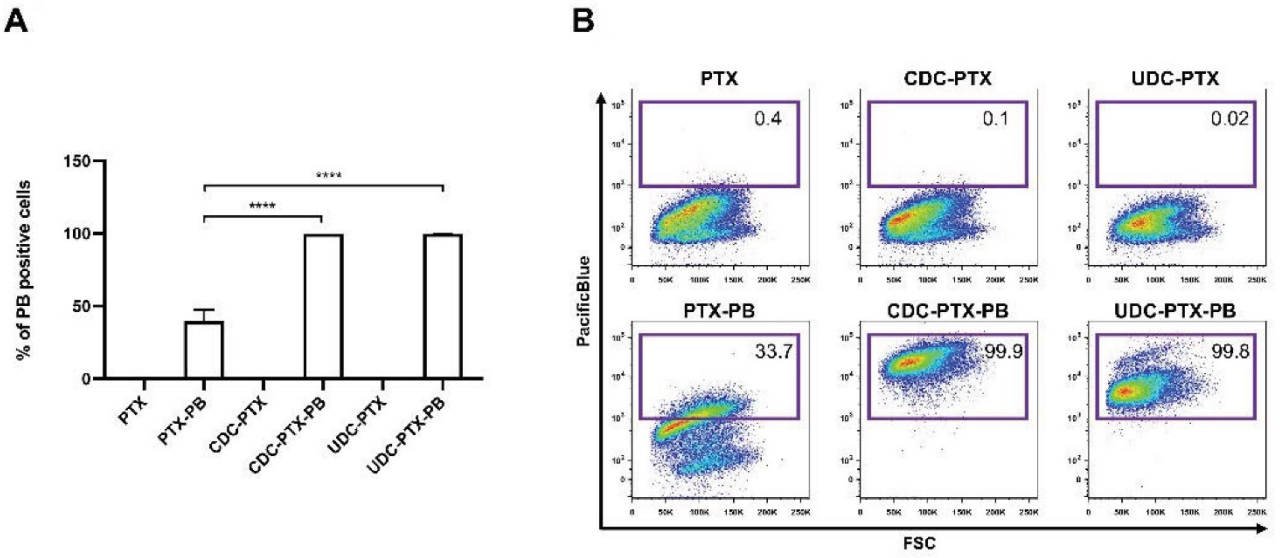


Figure 33. *Evaluation of the Pacific Blue positive HCT116 cells after 4 + 72 h of the indicated treatment (PTX, CDC-PTX and UDC-PTX: 5 μM; PTX-PB, CDC-PTX-PB and UDC-PTX-PB: 18 μM) analysed by flow cytometry. In (A), data are reported as mean ± SEM of at least 3 independent experiments. Statistical analysis was performed by ANOVA followed by Bonferroni post hoc test for pairwise comparisons. **** $p < 0.0001$. In (B), a representative panel of cytofluorimetric analysis of HCT116 cells is shown. The FSC axes scale is reported as linear, while the Pacific Blue axes scale is reported as logarithmic.*

Conclusion

In conclusion, two novel BA-PTX hybrids based on CDCA and UDCA were synthesized and tested for anticancer activity against a variety of leukaemia and colon cancer cells.

The synthetic procedure for the drug conjugates requires only one high yielding synthetic step followed by purification, making the conjugates affordable. The synthesis of the fluorophore bearing conjugates instead require more synthetic steps and needed optimization before achieving the final products required but was only needed for the *in vitro* studies.

CDC-PTX and PTX had similar anticancer activity against all cancer cell lines tested while UDC-PTX was found to be much less cytotoxic than CDC-PTX and PTX, with some anticancer activity only against HCT116 cells. Furthermore, both BA-PTX hybrids were significantly less toxic to NIH-3T3 non-tumoral cells than unconjugated PTX. The ability of BA-PTX hybrids to pass through the plasmatic membrane has also been tested using newly synthesized fluorescent analogue BA-PTX-PB hybrid compounds as well as PTX-PB for direct comparison. Interestingly, in the case of HCT116 cancer cells, both BA-PTX hybrids showed a higher ability to enter the cancer cells with PB positive cells up to 99.9% with respect to PTX-PB with PB positive cells up to only 35%.

The strikingly different cytotoxicity of CDC-PTX and UDC-PTX hybrids shows that the nature of the bile acid conjugated can influence PTX anticancer activity. The ability to enter cancer cells differs, demonstrating that the presence of a BA unit can influence some physical-chemical properties. The greater ability of the CDC hybrid to pass the plasmatic membrane compared to the UDC one may account, at least in part, for the higher anticancer activity of CDC-PTX compared to UDC-PTX.

Furthermore, hydrophobic BA, including CDCA, are known to be more cytotoxic than hydrophilic BA, such as UDCA. Given the high entry rate of the conjugated compounds in cells but their relative low efficacy compared to paclitaxel, it is possible to speculate that the conjugates remain fairly stable inside the cells, and that the prodrug strategy is not very efficacious in this case.

Overall, the data herein presented, even though very preliminary, pave the way to further studies directed to tuning PTX anticancer activity through bile acid conjugation. Diverse modes of conjugation can be considered both for a prodrug strategy and for a new drug strategy. chemical conjugation with bile acids could also be considered to investigate oral forms of toxoids administration thanks to some beneficial features of bile acids, such as their ability to target bile acid transporters and to localize in specific organs.

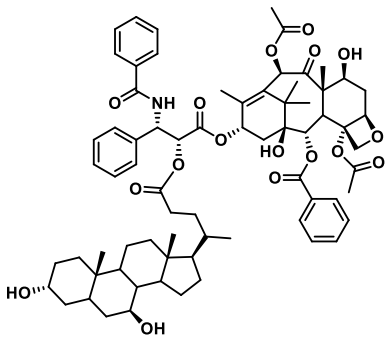
Materials and Methods

Commercial Paclitaxel (PTX) was purchased by Carbosynth (Compton, Berkshire, UK). Pacific blue activated as *N*-Hydroxysuccinimide ester (PB-NHS) was purchased from AAT Bioquest (Sunnyvale, CA, USA). UDCA and CDCA, were kindly furnished by ICE SpA (Reggio Emilia, Italy). All the chemicals were used without further purification. The reactions were monitored by thin layer chromatography (TLC) on pre-coated silica gel F_{254} plates (thickness 0.25 mm, Merck, Darmstadt, Germany). UV light spotted the presence of PTX and Pacific Blue, while phosphomolybdic acid solution was used as a spray to develop steroids spots. Flash column chromatography was performed on silica gel (60 a, 230–400 mesh). NMR spectra were recorded with a Varian Mercury 400 MHz instrument (Varian, Palo Alto, CA, USA), or Varian Mercury 300 MHz instrument (Varian, Palo Alto, CA, USA). HRMS spectra were acquired with a Waters Micromass ZQ instrument (Waters Corp, Milford, MA, USA). Preparative and analytical HPLC was executed on Xterra C18 column (Waters Corp., Milford, MA, USA) with a Jasco LC-2000 plus instrument (Jasco, Easton, MD, USA). Tetrahydrofuran (THF) was used as freshly distilled on sodium/benzophenone with standard procedures. *N*,*N*-dimethylformamide (DMF) was dried over freshly activated molecular sieves for at least one day before use.

General Procedure for the Synthesis of BA-PTX Esters

In a closed cap vial, a solution of 0.058 mmol (50 mg) of PTX and 0.064 mmol of the required bile acid in 3.5 mL of anhydrous DMF was prepared. The solution was stirred in an ice bath, and 0.064 mmol of 4-dimethylaminopyridine (DMAP) and EDC were added, then allowed to warm to room temperature and stirred overnight. The reaction mixture was then diluted with 15 mL of diethyl ether and poured into a separatory funnel. The organic phase was extracted three times with an equal volume of water in order to remove DMF and DMAP. The diethyl ether solution was dried by adding anhydrous sodium sulphate ad evaporated under vacuum. The resulting white solid was purified by flash chromatography to obtain the desired ester.

UDC-PTX

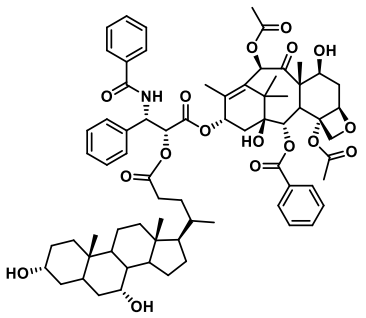


EtOAc/Cyclohexane 4:1; white amorphous solid, 86% yield.

^{1}H-NMR (400 MHz, $CDCl_3$)selected data: δ 8.15–8.11 (m, 2H), 7.77–7.70 (m, 2H), 7.61–7.30 (m, 11H), 6.88 (d, *J* = 9.2 Hz, 1H), 6.29 (s, 1H), 6.27–6.21 (m, 1H), 5.94 (dd, *J* = 9.2, 3.2 Hz, 1H), 5.68 (d, *J* = 7.2 Hz, 1H), 5.51 (d, *J* = 3.4 Hz, 1H), 4.97 (d, *J* = 7.6 Hz, 1H), 4.44 (dd, *J* = 10.9, 6.7 Hz, 1H), 4.31 (d, *J* = 8.3 Hz, 1H), 4.20 (d, *J* = 8.0 Hz, 1H), 3.81 (d, *J* = 7.0 Hz, 1H), 3.61–3.53 (m, *J* = 10.3, 5.4 Hz, 2H), 2.52 (dd, *J* = 21.7, 12.1 Hz, 2H), 2.44 (s, 3H), 2.40–2.27 (m, 3H), 2.23 (s, 3H), 1.94 (d, *J* = 1.2 Hz, 3H), 1.68 (s, 3H), 1.23 (s, 3H), 1.13 (s, 3H), 0.94 (s, 3H), 0.87 (d, *J* = 6.3 Hz, 3H), 0.64 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) selected data: δ 203.78, 171.23, 169.76, 168.07, 167.02, 166.99, 142.77, 137.01, 133.65, 132.20, 132.10, 132.01, 130.20, 129.03, 128.73, 128.43, 127.06, 126.52, 84.43, 81.03, 79.14, 76.43, 75.57, 75.07, 73.71, 72.09, 71.71, 71.40, 71.32, 58.49, 55.68, 54.80, 52.83, 45.56, 43.73, 43.16, 42.39, 40.11, 39.14, 37.22, 36.87, 35.55, 35.07, 34.89, 34.04, 30.73, 30.29, 28.59, 26.83, 23.35, 22.67, 22.10, 21.14, 20.83, 18.25, 14.83, 12.12, 9.59.

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{71}H_{90}NO_{17}]^+$ 1129.2887; found 1129.6195.

CDC-PTX

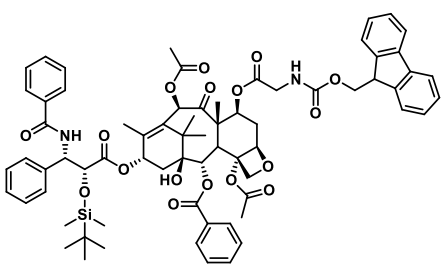


EtOAc/Cyclohexane 2:1 to 4:1; white amorphous solid, 94% yield.

^{1}H-NMR (400 MHz, $CDCl_3$) selected data: δ 8.17–8.09 (m, 2H), 7.74 (d, *J* = 7.2 Hz, 2H), 7.68–7.29 (m, 11H), 6.89 (d, *J* = 9.1 Hz, 1H), 6.30 (s, 1H), 6.28–6.17 (m, *J* = 8.3 Hz, 1H), 5.94 (dd, *J* = 9.2, 3.1 Hz, 1H), 5.68 (d, *J* = 7.1 Hz, 1H), 5.51 (d, *J* = 3.3 Hz, 1H), 4.97 (d, *J* = 7.8 Hz, 1H), 4.44 (dd, *J* = 10.8, 6.5 Hz, 1H), 4.31 (d, *J* = 8.3 Hz, 1H), 4.20 (d, *J* = 8.4 Hz, 1H), 3.86–3.77 (m, 2H), 3.51–3.39 (m, 1H), 2.56 (s, 1H), 2.45 (s, 3H), 2.41–2.25 (m, 3H), 2.24 (s, 3H), 1.94 (s, 3H), 1.68 (s, 3H), 1.23 (s, 3H), 1.13 (s, 3H) 0.90 (s, 3H), 0.87 (d, *J* = 6.3 Hz, 3H), 0.62 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) selected data: δ 203.78, 173.17, 171.27, 169.78, 168.08, 167.03, 166.98, 137.03, 133.66, 132.73, 132.21, 132.00, 130.20, 129.15, 129.02, 128.72, 128.42, 127.06, 126.54, 84.43, 81.03, 79.17, 76.42, 75.58, 75.07, 73.70, 72.07, 71.96, 71.70, 68.47, 58.50, 55.56, 52.82, 50.45, 45.56, 43.14, 42.64, 41.40, 39.83, 39.60, 39.36, 35.54, 35.28, 35.10, 35.01, 34.64, 32.82, 30.61, 30.52, 28.11, 26.80, 23.64, 22.74, 22.67, 20.82, 20.54, 18.12, 14.83, 11.75, 9.59.

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{71}H_{90}NO_{17}]^+$ 1129.2887; found 1129.6198.

TBDMSiO-PTX-Gly-Fmoc

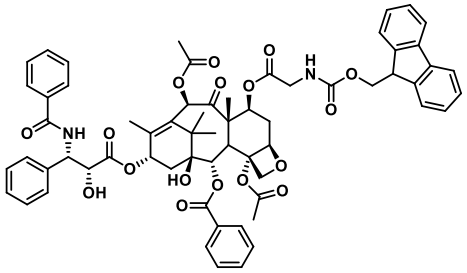


A solution of 0.682 mmol (66 mg) of TBDMSiO-PTX 0.273 mmol (81 mg) of Gly-Fmoc and 0.034 mmol (4 mg) of DMAP in anhydrous DMF (4 mL) was prepared in a closed cap vial. To this solution, 0.273 (56 mg) mmol of DCC were added, and the mixture was stirred at 35 °C. After two hours, the mixture was diluted with ethyl acetate (50 mL) and washed with saturated NH_4Cl (25 mL × 2), $NaHCO_3$ (25 mL × 3) and brine (25 mL). The organic phase was then dried by adding anhydrous sodium sulphate, filtered, and evaporated under vacuum. The resulting crude compound was purified by flash chromatography (EtOAc/cyclohexane 1:1,8) to give 78 mg of pure compound **1** in 90% yield as an amorphous solid.

^{1}H-NMR (400 MHz, $CDCl_3$) selected data. δ 8.13 (d, *J* = 7.3 Hz, 2H), 7.82–7.71 (m, 4H), 7.71–7.58 (m, 2H), 7.56–7.27 (m, 16H), 7.09 (d, *J* = 8.9 Hz, 1H), 6.27 (t, *J* = 8.8 Hz, 1H), 6.21 (s, 1H), 5.77–5.68 (m, 3H), 5.52 (dd, *J* = 7.1, 4.8 Hz, 1H), 4.99 (d, *J* = 8.6 Hz, 1H), 4.69 (d, *J* = 1.7 Hz, 1H), 4.48–4.35 (m, 3H), 4.30–4.24 (m, 1H), 4.21 (d, *J* = 8.5 Hz, 1H), 4.15–4.05 (m, 1H), 3.97 (d, *J* = 6.9 Hz, 1H), 3.89–3.79 (m, 1H), 2.66–2.51 (m, 4H), 2.49–2.35 (dd, *J* = 15.2, 9.4 Hz, 1H), 2.27–2.08 (m, 4H), 1.99 (s, 3H), 1.96–1.86 (m, 1H), 1.83 (s, 3H), 1.24 (s, 3H), 1.17 (s, 3H), 0.81 (s, 9H), −0.02 (s, 3H), −0.29 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) selected data. δ 201.82, 171.47, 169.92, 169.87, 169.65, 167.00, 166.88, 156.71, 144.06, 143.89, 141.26 (2 C), 138.20, 134.08, 133.76, 132.32, 131.80, 130.20, 129.01, 128.78, 128.75, 127.98, 127.71, 127.63, 127.03, 126.98, 126.38, 125.31, 125.22, 125.14, 119.97, 119.91, 83.89, 80.92, 78.60, 76.33, 75.69, 75.08, 74.39, 72.01, 71.29, 67.24, 56.10, 55.65, 47.14, 46.90, 43.34, 43.10, 35.58, 33.27, 32.59, 30.71, 26.37, 26.18, 25.52, 25.21, 24.68, 22.99, 21.38, 20.91, 18.13, 14.67, 10.87, −5.19, −5.80.

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{70}H_{79}N_2O_{17}Si]^+$ found 1248.1549 ($M + H^+$) calculated 1248.5182

PTX-Gly-Fmoc



In a polypropylene vial, a solution of 0.41 mmol of Si-PTX-Gly-Fmoc 1 in anhydrous THF (7.8 mL) was prepared and stirred under nitrogen atmosphere at 0 °C. To this solution, 2.6 mL of HF-Pyridine were added slowly with the aid of a (polypropylene) syringe. (FAVS, Bologna, Italy) The mixture was allowed to reach room temperature, stirred for five hours, then diluted with EtOAc and quenched with $NaHCO_3$ aqueous saturated solution. The organic phase was transferred to a separatory funnel and extracted with water and brine. The organic phase was dried with sodium sulphate and the solvent evaporated under vacuum. The resulting crude compound was purified by column chromatography (EtOAc/cyclohexane 1:1) to yield the product in 51% yield as a white powder.

^{1}H-NMR (400 MHz, $CDCl_3$) selected data. δ 8.17–8.03 (m, 2H), 7.86–7.71 (m, 4H), 7.69–7.28 (m, 17H), 7.10 (d, *J* = 9.0 Hz, 1H), 6.23–6.03 (m, 2H), 5.80 (dd, *J* = 8.9, 2.4 Hz, 1H), 5.66 (m, 2H), 5.54–5.42 (m, 1H), 4.94 (d, *J* = 8.1 Hz, 1H), 4.80 (s, 1H), 4.38 (d, *J* = 6.8 Hz, 2H), 4.31 (d, *J* = 8.5 Hz, 1H), 4.27 (t, *J* = 7.2 Hz, 1H), 4.18 (d, *J* = 8.4 Hz, 1H), 4.08 (dd, *J* = 17.9, 7.2 Hz, 1H), 3.92 (d, *J* = 6.9 Hz, 1H), 3.85 (dd, *J* = 17.8, 4.8 Hz, 1H), 3.71 (s, 1H), 2.66–2.50 (m, 1H), 2.38 (s, 1H), 2.36–2.29 (m, 3H), 2.20 (s, 1H), 1.91 (d, *J* = 12.4 Hz, 1H), 1.86 (s, 1H), 1.83 (s, 3H), 1.82 (s, 1H), 1.73 (s, 1H), 1.22 (s, 3H), 1.16 (s, 3H).

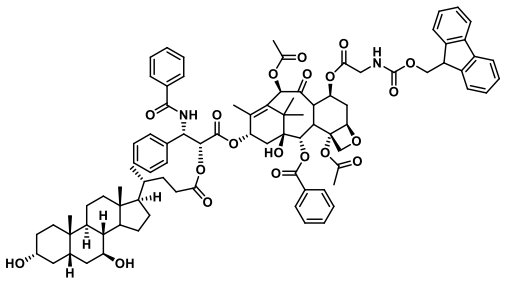
^{13}C-NMR (101 MHz, $CDCl_3$) δ 201.82, 172.60, 170.47, 169.97, 169.67, 167.11, 166.95, 156.80, 144.14, 143.98, 141.36, 140.79, 138.09, 133.91, 133.75, 132.72, 132.03, 130.26, 129.08, 128.84, 128.79, 128.42, 127.76, 127.15, 125.39, 125.32, 120.03, 83.90, 81.06, 78.57, 76.47, 75.82, 74.31, 73.29, 72.17, 67.34, 56.30, 55.03, 47.24, 47.09, 43.32, 43.18, 35.66, 33.38, 26.62, 22.63, 21.04, 20.93, 14.76, 10.91.

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{64}H_{65}N_2O_{17}]^+$ 1133.4239; found 1133.8087.

General Procedure for the Preparation of UDC-PTX-Gly-Fmoc 3 and CDC-PTX-Gly.Fmoc

To a solution of 0.39 mmol of the selected bile acid, 0.097 mmol of PTX-Gly-Fmoc 2 and 0.049 mmol of DMAP in 1 mL of anhydrous DMF in a closed cap vial, 0.39 mmol of DCC were added, and the mixture was stirred at room temperature. After 20 min a precipitate formed, and the completeness of the reaction was assessed by TLC after 1.5 h. The mixture was filtered through Celite and diluted with EtOAc (50 mL). The organic phase was extracted with NH_4Cl (10 mL × 3), $NaHCO_3$ (10 mL × 2), water (10 mL), and dried with anhydrous sodium sulphate. The solvent was removed under vacuum and the solid purified by column chromatography.

UDC-PTX-Gly-Fmoc

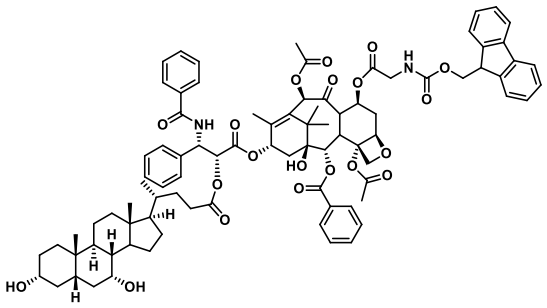


EtOAc/cyclohexane 3:1; white powder, 82% yield.

^{1}H-NMR (300 MHz, $CDCl_3$) selected data. δ 8.21–8.03 (m, 2H), 7.82–7.71 (m, 3H), 7.70–7.29 (m, 18H), 6.98–6.77 (m, 1H), 6.27–6.13 (m, 2H), 6.00–5.86 (m, 1H), 5.71–5.64 (m, 1H), 5.57–5.45 (m, 2H), 5.04–4.88 (m, 1H), 4.48–4.33 (m, 1H), 4.33–4.21 (m, 1H), 4.21–4.12 (m, 1H), 4.09–3.98 (m, 1H), 3.99–3.91 (m, 1H), 3.89–3.74 (m, 1H), 3.71–3.62 (m, 1H), 3.55–3.31 (m, 1H), 2.52–2.40 (m, 1H), 2.37–2.24 (m, 1H), 2.19 (s, 1H), 2.04 (s, 1H), 1.99–1.95 (m, 1H). ^{13}C-NMR (101 MHz, $CDCl_3$) δ 202.26, 173.45, 170.18, 169.97, 169.81, 168.60, 167.29, 157.07, 144.43, 144.28, 141.92, 141.60, 137.39, 134.12, 132.48, 132.35, 130.54, 129.38, 129.09, 128.78, 127.99, 127.41, 126.96, 125.71, 125.62, 120.25, 84.23, 81.14, 79.05, 76.06, 74.76, 73.90, 72.44, 71.92, 68.80, 67.61, 56.43, 55.88, 50.75, 50.45, 47.49, 47.29, 43.59, 43.46, 43.01, 41.81, 40.22, 39.93, 39.74, 35.67, 35.38, 33.55, 33.15, 30.97, 30.83, 28.42, 26.78, 25.65, 24.94, 24.01, 23.10, 22.91, 21.55, 21.26, 20.90, 18.50, 14.97, 12.12, 11.21.

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{88}H_{103}N_2O_{20}]^+$ 1508.7093, found 1508.5138.

CDC-PTX-Gly-Fmoc



EtOAc/cyclohexane 3:1; 74% yield white powder.

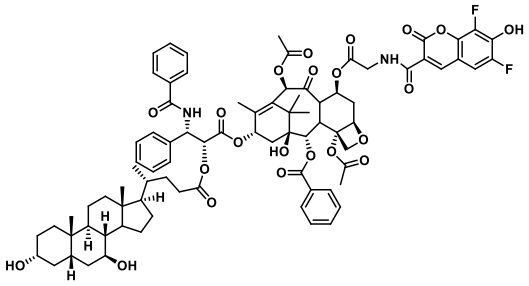
^{1}H-NMR (300 MHz, $CDCl_3$) selected data. δ 8.12 (d, *J* = 7.3 Hz, 2H), 7.81–7.71 (m, 3H), 7.71–7.30 (m, 18H), 7.05–6.91 (m, 1H), 6.34–6.12 (m, 2H), 6.01–5.90 (m, 1H), 5.70 (s, 1H), 5.56 (s, 1H), 4.94 (d, 1H), 4.37 (s, 2H), 4.33–4.23 (m, 1H), 4.22–4.15 (m, 1H), 4.13–4.02 (m, 1H), 3.98–3.86 (m, 1H), 3.85–3.77 (m, 1H), 3.55–3.36 (m, 1H), 2.68–2.46 (m, 2H), 2.42 (s, 1H), 2.37–2.24 (m, 1H), 2.19 (s, 1H), 0.93–0.80 (m, 1H), 0.62 (s, 3H). ^{13}C-NMR (75 MHz, $CDCl_3$) selected data δ 202.26, 173.45, 170.18, 169.97, 169.81, 168.60, 167.29, 157.07, 144.43, 144.28, 141.92, 141.60, 137.39, 134.12, 132.48, 132.35, 130.54, 129.38, 129.09, 128.78, 127.99, 127.41, 126.96, 125.71, 125.62, 120.25, 84.23, 81.14, 79.05, 76.06, 74.76, 73.90, 72.44, 71.92, 68.80, 67.61, 56.43, 55.88, 50.75, 50.45, 47.49, 47.29, 43.59, 43.46, 43.01, 41.81, 40.22, 39.93, 39.74, 35.67, 35.38, 33.55, 33.15, 30.97, 30.83, 28.42, 26.78, 25.65, 24.94, 24.01, 23.10, 22.91, 21.55, 21.26, 20.90, 18.50, 14.97, 12.12, 11.21.

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{88}H_{102}N_2O_{20}]^+$ 1509.7127, found 1509.7878.

General Procedure for the Preparation of UDC-PTX-Gly-PB and CDC-PTX-Gly-PB

8.7 μmol of the Fmoc protected glycine conjugate previously obtained were dissolved in a solution of 20% piperidine in DMF (2 mL) and stirred at room temperature for 10 min. The piperidine was completely removed under vacuum at 40 °C and to the remaining solution in DMF (approx. 0.1 mL) 3.97 μL (0.023 mmol) of DIPEA, and 2.5 mg (7.4 μmol) of PB-NHS dissolved in 0.5 mL of anhydrous DMF were added. The mixture was stirred overnight, and the solvent completely removed under vacuum. The resulting bright yellow solid was first purified by flash chromatography, then by reverse phase HPLC: A: deionized water, B: ACN + 0.1% TFA. t = 0 min A:B 60/40 t = 20 min A:B 0/100 T = 25 min A:B 0/100 t = 26 min A:B 60/40. 1.4 mL/min flow, r.t. λ = 260 nm.

UDC-PTX-PB



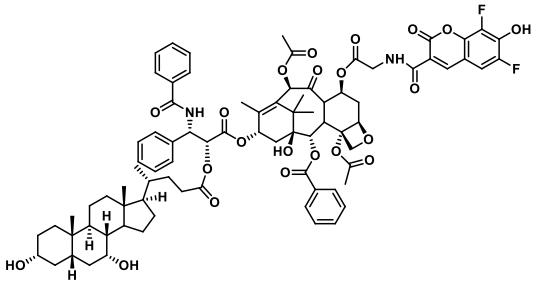
Flash chromatography: DCM/MeOH 8:2; yield 63%.

^{1}H-NMR (400 MHz, $CDCl_3$) selected data. δ 9.15 (dd, *J* = 7.2, 4.7 Hz, 1H), 8.76 (d, *J* = 1.3 Hz, 1H), 8.15–8.00 (m, 2H), 7.80–7.70 (m, 2H), 7.66–7.29 (m, 10H), 7.19 (dd, *J* = 9.4, 2.0 Hz, 1H), 6.97 (d, *J* = 8.8 Hz, 1H), 6.24–6.12 (m, 2H), 5.87 (dd, *J* = 8.8, 3.5 Hz, 1H), 5.74 (dd, *J* = 10.7, 7.2 Hz, 1H), 5.66 (d, *J* = 7.0 Hz, 1H), 5.47 (d, *J* = 3.6 Hz, 1H), 4.96 (d, *J* = 8.5 Hz, 1H), 4.42 (dd, *J* = 17.9, 8.0 Hz, 1H), 4.32 (d, *J* =

8.3 Hz, 1H), 4.16 (d, *J* = 8.4 Hz, 1H), 3.97 (dd, *J* = 17.4, 4.5 Hz, 1H), 3.91 (d, *J* = 6.7 Hz, 1H), 3.73–3.49 (m, 2H), 2.62–2.42 (m, 4H), 2.39 (s, 3H), 2.35–2.23 (m, 4H), 2.22 (s, 4H), 2.20–2.09 (m, 4H), 2.04 (s, 3H), 1.95 (s, 3H), 1.78 (s, 3H), 1.21 (s, 3H), 1.14 (s, 3H), 0.94 (s, 3H), 0.86 (d, *J* = 6.2 Hz, 3H), 0.62 (s, 3H). ^{13}C-NMR (101 MHz, $CDCl_3$) selected data δ 201.87, 173.08, 169.67, 169.39, 169.01, 168.46, 167.23, 166.87, 162.16, 159.66, 147.81, 141.74, 137.04, 133.69, 132.11, 132.01, 130.12, 128.98, 128.73, 128.43, 127.08, 126.66, 116.56, 109.98, 83.94, 80.80, 78.63, 76.45, 75.67, 74.34, 73.71, 72.28, 71.86, 71.66, 71.56, 55.93, 55.60, 55.01, 53.12, 47.18, 43.70, 43.64, 43.20, 42.39, 41.82, 40.05, 39.27, 36.92, 36.63, 35.40, 35.18, 34.82, 34.01, 30.62, 30.42, 30.04, 28.67, 26.73, 26.47, 23.37, 22.58, 21.16, 18.10, 14.73, 12.07, 10.78.

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{83}H_{95}F_2N_2O_{22}]^+$ 1509.6300, found 1509.7878.

CDC-PTX-PB



Flash chromatography: DCM/MeOH 9:1; yield 17%.

^{1}H-NMR (300 MHz, $CDCl_3$) selected data. δ 9.28–9.15 (m, 1H), 8.78 (d, *J* = 1.2 Hz, 1H), 8.12–8.04 (m, 2H), 7.81–7.72 (m, 2H), 7.66–7.29 (m, 10H), 7.24–7.16 (m, 1H), 7.00 (d, *J* = 9.0 Hz, 1H), 6.24–6.12 (m, 2H), 5.86 (dd, *J* = 8.9, 3.9 Hz, 1H), 5.73 (dd, *J* = 10.6, 7.0 Hz, 1H), 5.65 (d, *J* = 6.8 Hz, 1H), 5.50 (d, *J* = 3.8 Hz, 1H), 4.97 (d, *J* = 8.8 Hz, 1H), 4.48–4.36 (m, 1H), 4.32 (d, *J* = 8.5 Hz, 1H), 4.16 (m, 1H), 4.08–3.95 (m, 1H), 3.90 (d, *J* = 6.7 Hz, 1H), 3.87–3.77 (m, 1H), 3.58 (m, 1H), 2.74 (s, 1H), 2.37 (s, 1H), 1.25 (s, 3H), 1.20 (s, 3H), 1.14 (s, 3H), 0.62 (s, 3H).

HRMS (ESI, *m/z*) $[M + H]^+$ calculated for $[C_{83}H_{95}F_2N_2O_{22}]^+$ 1509.6333, found 1509.7889.

Synthesis of Dihydroartemisinin Bile Acids Conjugates for Biological Studies Against SARS-CoV-2

Part of the work present in this chapter will be published in a peer reviewed journal concomitantly with this thesis.

Introduction

Artemisinin

Artemisinin is the progenitor of a family of drugs widely used against malaria. The disease, although now almost eradicated in temperate climates and developed nations, is still a major death cause in tropical and developing countries. For the drug discovery the 2015 Nobel prize in physiology and medicine was awarded to professor Youyou Tu.[191] The sesquiterpene lactone molecule was first extracted from the leaves of *artemisia annua* in 1970s, and later characterized, leading to the discovery of a surprisingly stable peroxide bond and a unique tridimensional structure.[192] Despite its chemical synthesis being published in 1983[193], the extraction method, and the more efficient semi synthesis from engineered microorganisms remains more economical and efficient.[194] Many derivatives of artemisinin have been prepared over the years. Among these, dihydroartemisinin (DHA), has a stronger activity against malaria than the parent compound. Its synthesis involves a reduction of the lactone to give a hemiacetal bond at the carbon 10.[195] Apart from increasing the water solubility of the molecule, this transformation is very useful since it gives a molecular handle which allows easier modification of the compound without altering the original structure significantly. Artesunate (ART), is the product of the previous compound with succinic acid and has the advantage of being water soluble.[196] Another compound derived from dihydroartemisinin is artemether; this methyl ether is soluble in oil and more suitable for oral administration.[192] Many more modifications have been made on the artemisinin scaffold and studies of structure activity relationship (SAR) have been undertaken. The chemical structure of artemisinin and some of its derivatives are reported in the figure below.

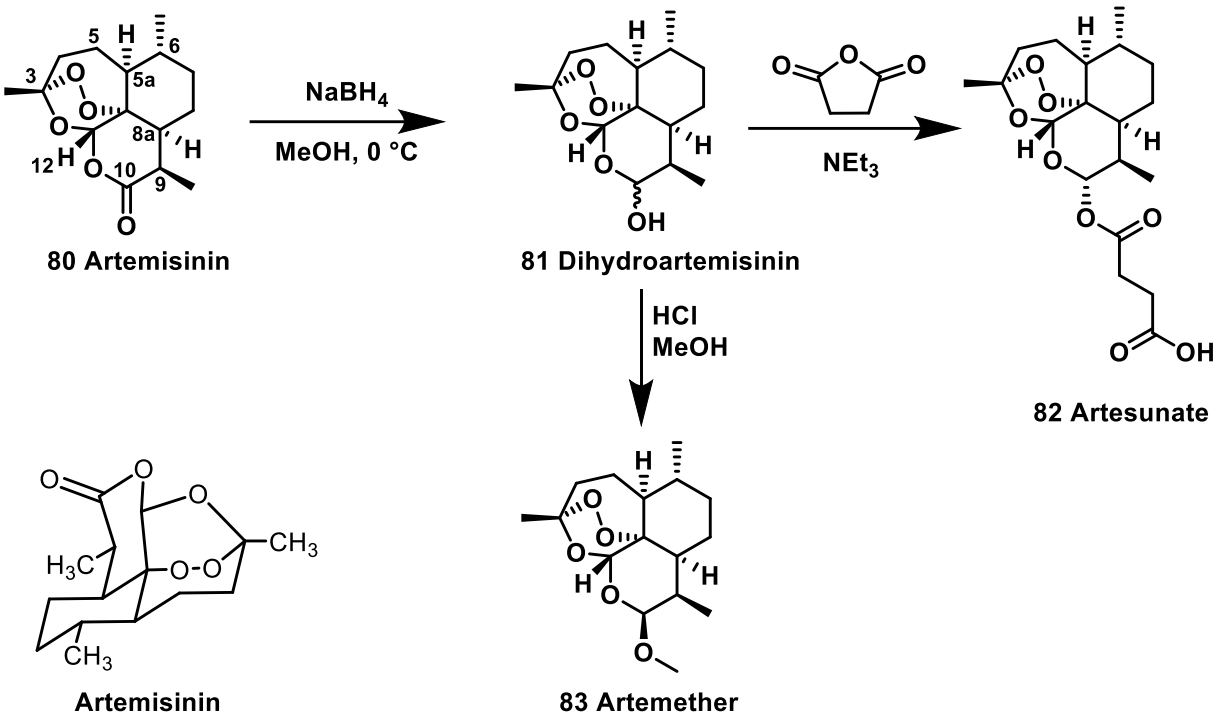


Figure 34. *Scheme representing artemisinin and its derivatives. Dihydroartemisinin is obtained by reduction of the lactone in position 10 to give a hemiacetal. Artesunate can be obtained by conjugation of DHA and succinic anhydride. Artemether is obtained by protection of the DHA hemiacetal with methanol, the corresponding product with ethanol (artether) is also commercial and used as antimalarial. Artemisinin is also represented with a three-dimensional structure.*

Artemisinin mechanism of action against Plasmodium Falciparum

The mechanism of action against the *Plasmodium Falciparum*, responsible for malaria, has been partially elucidated thanks to the collective effort of many scientists over the years. In a very crude summary, the molecule is able to act during the blood infectious phase of the parasite thanks to the endoperoxide activation by ferrous heme present in the attacked erythrocyte.[197,198] The metal ions are found in large quantities due to heme digestion from the *plasmodium* and they cleave the peroxide bond leaving an alkylating radical free to blindly attack proteins of the parasite itself and of the infected cell.[199–201] However, it is thought that other activation mechanisms are present and that also reactive oxygen species can play a crucial role for artemisinin activity.[202,203] This specific mechanism of activation is particularly suited to combat malaria disease since the molecule only acts on infected cells. Moreover, its unspecificity toward proteins and indiscriminate alkylation greatly decreases the chances of parasite adaptation to the drug mechanism of action; thus, circumventing the rising of a resistant strain.[197] To fully appreciate this fact is important to remember that most microorganism strains can evolve drug resistance by changing the conformation of the drug targeted proteins, or some other involved tract, if given enough time. This factor, along with the relative wide availability of artemisinin made it a great resource against drug resistant malaria strains, especially when used in combination with other drugs such as the previously ubiquitous chloroquine, mefloquine and piperaquine. The combination therapy approach (ACT) with other drugs is indeed the one recommended by world health organization due to the short half-life or artemisinin and its derivatives *in vivo*.

Drug repurposing

The generality of their action as alkylating agents confers artemisinin derivatives (ARTs) with multiple pharmacological activities, including anticancer,[204–209] anti-inflammatory and antiviral properties against

several DNA and RNA viruses by affecting viral protein synthesis.[210–215] Based on data reporting the antiviral potential of antimalarial agents on SARS-CoV-2 infection, ARTs have been considered as potential therapies against SARS-CoV-2[216] themselves and in combination therapies.[217,218] In particular, Zhong *et al.* recently demonstrated that (DHA), can inhibit SARS-CoV-2 replication *in vitro* by decreasing viral protein production; thus representing a possible candidate for SARS-CoV-2 treatment.[216]

The COVID-19 pandemic has dramatically affected the population worldwide, the WHO reports over 760 million confirmed cases and 6.9 million deaths to date.[219] Moreover, the long-COVID sequelae, particularly at enteric and cardiovascular level,[220–223] compromises the quality of life of millions of people all over the world and burdens on health care systems. Despite the unprecedented speed of vaccine development and global mass-vaccination efforts, a relevant proportion of global population remains unvaccinated. Moreover, the appearance of new SARS-CoV-2 variants threatens the effectiveness of existing vaccines that could be not able to prevent the disease occurrence efficaciously. Therefore, infected individuals still need treatment options. Among the different therapeutical strategies proposed to counteract SARS-CoV-2 infection, the use of repurposed drugs as antiviral agents, for instance Hydroxychloroquine and Remdesivir, reported unsuccessful or discordant results.[224–228] The development of small-molecule antivirals against SARS-CoV-2 can provide an important therapeutic treatment option and remain a pressing matter.[229]

Nevertheless, the poor stability, of ARTs prevents a large pharmacological application. To amend the stability issues of ARTs several approaches have been employed, such as the combination therapy (ACT) and the developments of ARTs hybrid molecules.[230] Molecular hybridization is based on the conjugation through covalent bonds of bioactive molecules aimed at improving the stability and the biological activity or expand the therapeutic area of the components.[231] As mentioned before, molecular hybrids can be achieved through cleavable linkers designed to release the two agents under physiological or enzymatic conditions (prodrug strategy), as well as stable linkers leading to hybrid drugs. In the framework of an extensive study on the design and synthesis of natural hybrid molecules with potential anticancer activity,[232–235] our research has been also focused on the design of DHA-bile acid (DHA-BA) hybrids characterized by both ester and 1,2,3-triazole linker.[236] These products were tested previously as anticancer agents[236] and one of them has been tried against viral replication in SARS-CoV-2 infection model in this work.

Discussion

Hybrid rationale

Virus-host cell interaction interference has already been reported to efficiently limit viral spread, for instance by affecting virus envelope conformation and receptor.[237–239] It is well-established that SARS-CoV-2 enters the host cells by hijacking the human angiotensin converting enzyme receptor (ACE2). For this reason, among the different strategies adopted to counteract infection by enveloped viruses, such as SARS-CoV-2, the interference with viral entry phase seems to be a promising one.[240] In literature, it is known that several naturally occurring and clinically available triterpenoids, such as glycyrrhetinic and oleanolic acids, as well as primary and secondary bile acids and their amidated derivatives, such as glyco-ursodeoxycholic acid, and semi-synthetic derivatives such as obeticholic acid, reduce significantly the RBD/ACE2 binding.[241] This highlights that bile acids, apart from their conventional role as surfactant, could exhibit important properties towards different cell types. As said before, bile acids are also considered signaling molecules interacting with different receptors, such as the G-protein-coupled bile acid receptor-1 (GPBAR-1, also known as TGR5), and the farnesoid X nuclear receptors (FXR).[242–244]

Interestingly, Brevini T. *et al.* demonstrated that ursodeoxycholic acid (UDCA) is able to regulate FXR, a direct regulator of ACE2 transcription in multiple COVID19-affected tissues, downregulating ACE2 levels, thus reducing susceptibility to SARS-CoV-2 in *ex vivo* experiments.[245]

Due to the potential antiviral activity of DHA and ursodeoxycholic bile acid (UDCA) our research group focused on repurposing a selection of DHA-UDCA based hybrids with anticancer activity against leukemia and hepatocellular carcinoma[236,246,247] with the aim of testing their antiviral activity against SARS-CoV-2. The idea of expanding the biological area of study of DHA-UDCA based hybrids was also strengthened by the observation that cancer patients are more susceptible to SARS-CoV-2 infection and to adverse outcomes with higher risk of death than the general population.[248–250] Although the biological interplay between cancer and COVID-19 is not currently fully understood, the development of multitarget bioactive molecules can be an interesting approach.

Lead compound identification

For the biological study two DHA-UDCA hybrids depicted in Figure 35 were selected. Hybrid DHA-UDC was synthesized as previously reported through a condensation reaction using EDCI (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide) leading to the labile ester linker[236] whereas DHA-t-UDCOMe was synthesized following an improved procedure herein reported via click chemistry (CuAAC). The click reaction was used to obtain a non-cleavable linkage through the formation of the 1,2,3-triazole ring. The introduction of the triazole moiety is also relevant for its recognized ability to improve pharmacological, pharmacokinetic, and physiochemical profiles of bioactive compounds.[251] Another hybrid (DHA-s-UDCOMe) was obtained by the conjugation of artesunate to UDCOMe; in this case the site of conjugation on the bile acid is the same as the one with the triazolyl linker. However, the succinimidyl linker between the DHA and bile acid this time conjugates the two molecules through esters, susceptible to esterase cleavage in the cellular environment. The last hybrid, as reported in the figure below, is given by the double addition of artesunate to UDCOMe.

BA-DHA hybrids displayed significant improvement in stability and anticancer activity toward selected leukemia and hepatocellular carcinoma cells respect to parent DHA, demonstrating the effectiveness of the conjugation.[236,246,247]

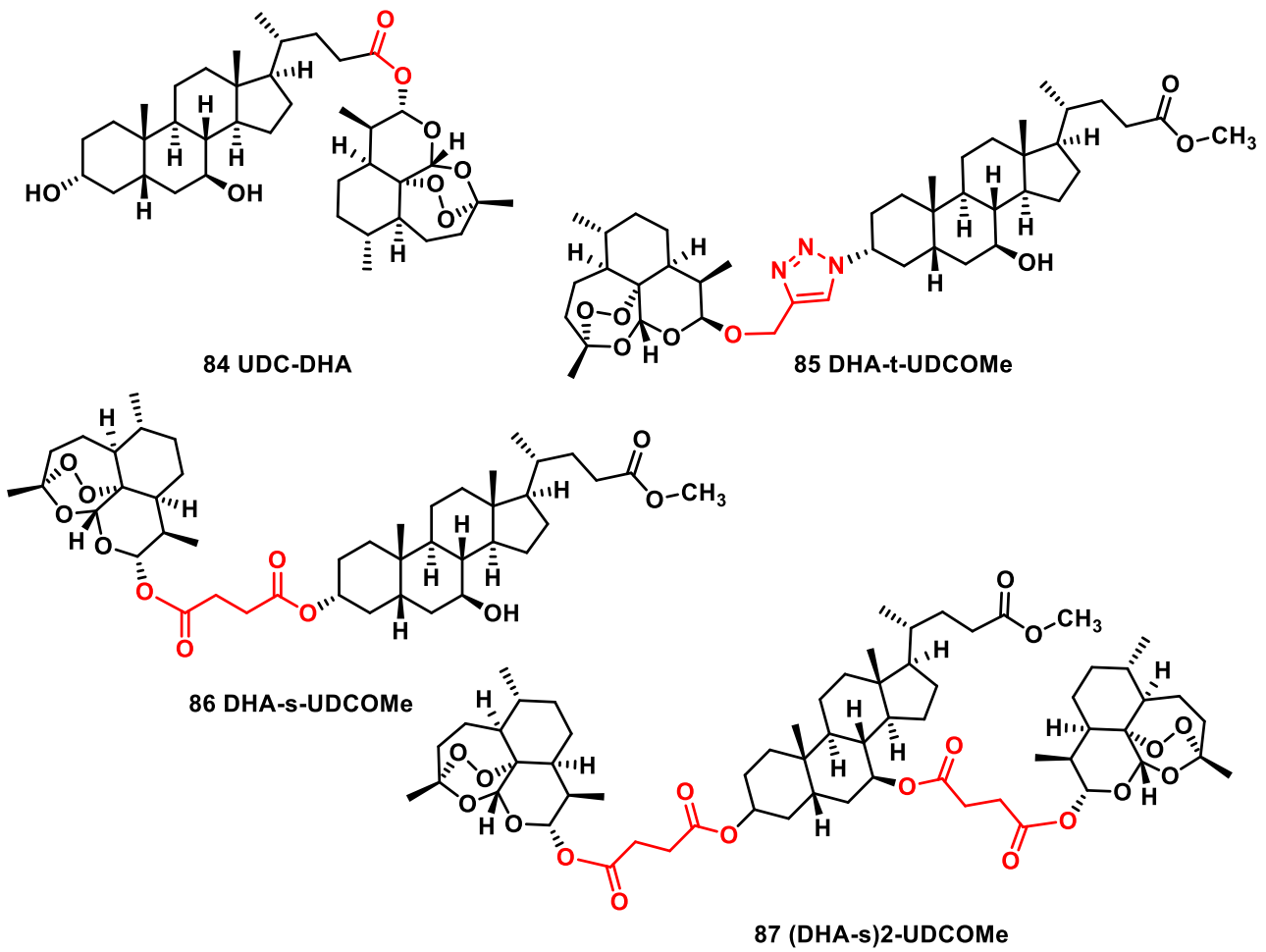


Figure 35. *Conjugates used for the biological studies. UDC-DHA is the hybrid between ursodeoxycholic acid and dihydroartemisinin obtained by esterification. DHA-t-UDCOMe is the hybrid obtained by click reaction (CuAAC) of a modified artemisinin in which an alkyne has been introduced and a modified bile acid bearing an azide. ART-UDCOMe is the hybrid between artesunate (succinic acid conjugate of DHA) and UDCA obtained by esterification; in this case the bile acid side chain is protected with a methyl ester. All the linker moieties are highlighted in red.*

The chemical stability of DHA and the hybrids in cell culture medium was previously evaluated by means of HPLC-MS/MS analyses.[246,247] Hybrids were found stable up to 24 hours incubation in cell culture medium whereas unconjugated DHA was reported to be far more unstable than the hybrid compounds being decomposed up to 70% after 6 hours in cell culture medium. Indeed, the hybridization allows to overcome the chemical instability of DHA.

Biological studies

The biological effects of the first hybrids were evaluated in collaboration with Prof. Rizzo Paola's research group and are reported here to help clarify the rest of the work. All compounds including unconjugated DHA were tested first for their cytotoxicity in the concentration range of 1-80 µM. DHA-t-UDCOMe showed the best safety profile with no significant cytotoxicity up to 80µM whereas hybrid DHA-UDC showed the highest cytotoxicity comparable to that of unconjugated DHA with a cell viability < 80% above 10 µM concentration. The biological work was carried out both on Vero E6 cells from African green monkey kidney (model used for the SARS-CoV-2 infection) and on CALU-3 cells form human lung adenocarcinoma. The cytotoxicity results were, in general, consistent among the two cells lines. The most promising conjugate, DHA-t-UDCOMe, on VERO E6 cells was also tested on CALU-3 and its

mechanism of action was assessed in further experiments. Here are only reported the activity results of the best hybrid obtained on the human cell line for brevity.

Evaluation of antiviral activity on Calu-3 cells

To confirm the antiviral effect of conjugate DHA-t-UDCOMe observed in Vero E6 cells, SARS-CoV-2-infected Calu-3 human lung cell lines were treated with DHA compounds at the concentration between 1 and 20 μM in a time-of-drug-addition assay and results compared to the antiviral effect observed in presence of the reference molecule DHA. In addition to viral titer measure by RT-qPCR (real time quantitative polymerase chain reaction) on extracted supernatants, the presence of infective viral particles was defined by plaque assay.

As depicted in Figure 36, DHA showed only a low viral replication interference of 1-1.2 logarithmic reduction (LR), independently of the time of administration, compared to untreated infected cells, at the highest non-toxic concentration (Figure 36, A). RT-q-PCR data were confirmed by plaque assay.

Conjugate DHA-t-UDCOMe presented antiviral activity dependent on the time of addiction. In fact, this conjugate caused a marked decrease of viral load when administrated in the post-infection stage, reaching 4.04 LR at the highest concentration tested, compared to untreated infected cells (Figure 36, B), as confirmed by plaque assay.

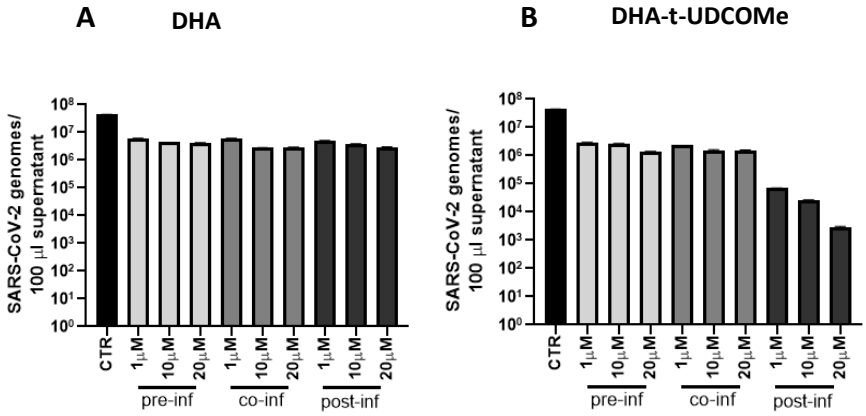


Figure 36*. Evaluation of antiviral activity of compounds DHA and DHA-t-UDCOMe on Calu-3 SARS-CoV-2 infected cells by RT-qPCR (A and B). Experiments were run in triplicates and reported as mean±SD.*

Based on these results, post-infection treatment with compound DHA-t-UDCOMe was selected for further investigations. Since SARS-CoV-2 infection requires the interaction between viral spike protein and cellular ACE2 receptor to allow viral entry and spread, the post-infection treatment with compound DHA-t-UDCOMe was selected to assess the effect on ACE2 expression, in order to clarify if the consistent reduction of viral replication reported involved this mechanism. As expected, the infection induced upregulation of ACE2 expression at both mRNA (Figure 37, A) and protein (Figure 37 B,C) level. On the contrary, post-infection treatment with 10-20μM hybrid DHA-t-UDCOMe led to a significant decrease of ACE2 expression in infected cells, in a dose dependent manner at both translational (Figure 37, A) and transductional level (Figure 37, B,C). This result suggests that the main reduction of SARS-CoV-2 infection observed after post-infection treatment with DHA-t-UDCOMe, was possibly due to a lower ACE2 availability on cell surface via FXR inhibition, which reduces Spike-ACE2 interaction rate, interfering with viral entry into the cells. Indeed, a recently published study demonstrated

that UDCA could act as a FXR inhibitor, a known ACE2 regulation factor, decreasing ACE2 expression.[245]

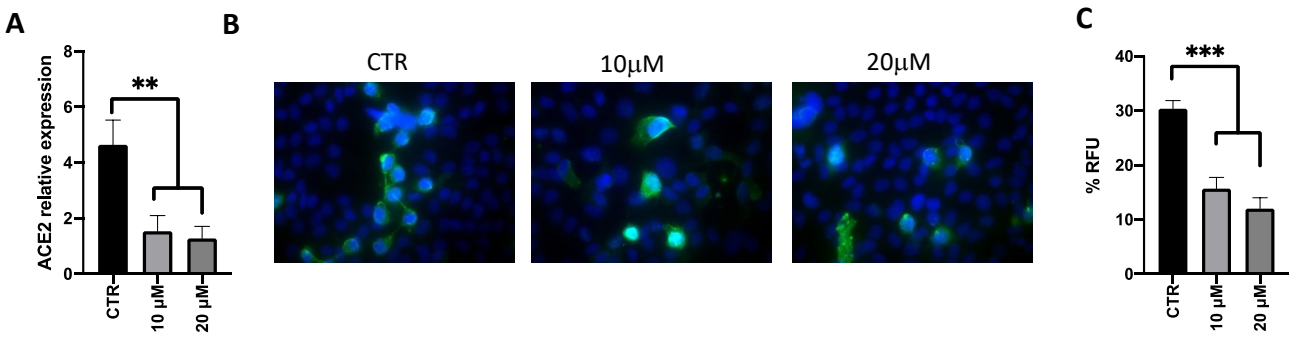


Figure 37. *ACE2 expression evaluation in SARS-CoV-2 infected Calu-3 cells after post-infection treatment with hybrid DHA-t-UDCOMe at 10 and 20µM at mRNA (A) and protein level by IF (B and C). RFU: Relative Fluorescent Units.*

Synthesis

Synthetic procedures for some of the compounds used for biological testing have already been published from our research group in the past.[236] However, the encouraging results obtained against SARS-CoV-2 prompted us to optimize the synthesis; in particular of the clicked conjugates, which was only obtained in low yield earlier. Having only low amounts of the conjugate available, the work started with the optimization of the precursors synthesis to culminate on the careful adjustment of the CuAAC conditions. In addition, viewing DHA-t-UDCOMe as a potential leading compound, new similar products were synthesized to explore, in the future, the relationship between the fine structure and activity of such conjugates. These conjugates were obtained in a similar fashion to the leading compound, requiring only slight variations on the synthetic conditions. The synthetic scheme for the new conjugates is reported in Figure 38 below.

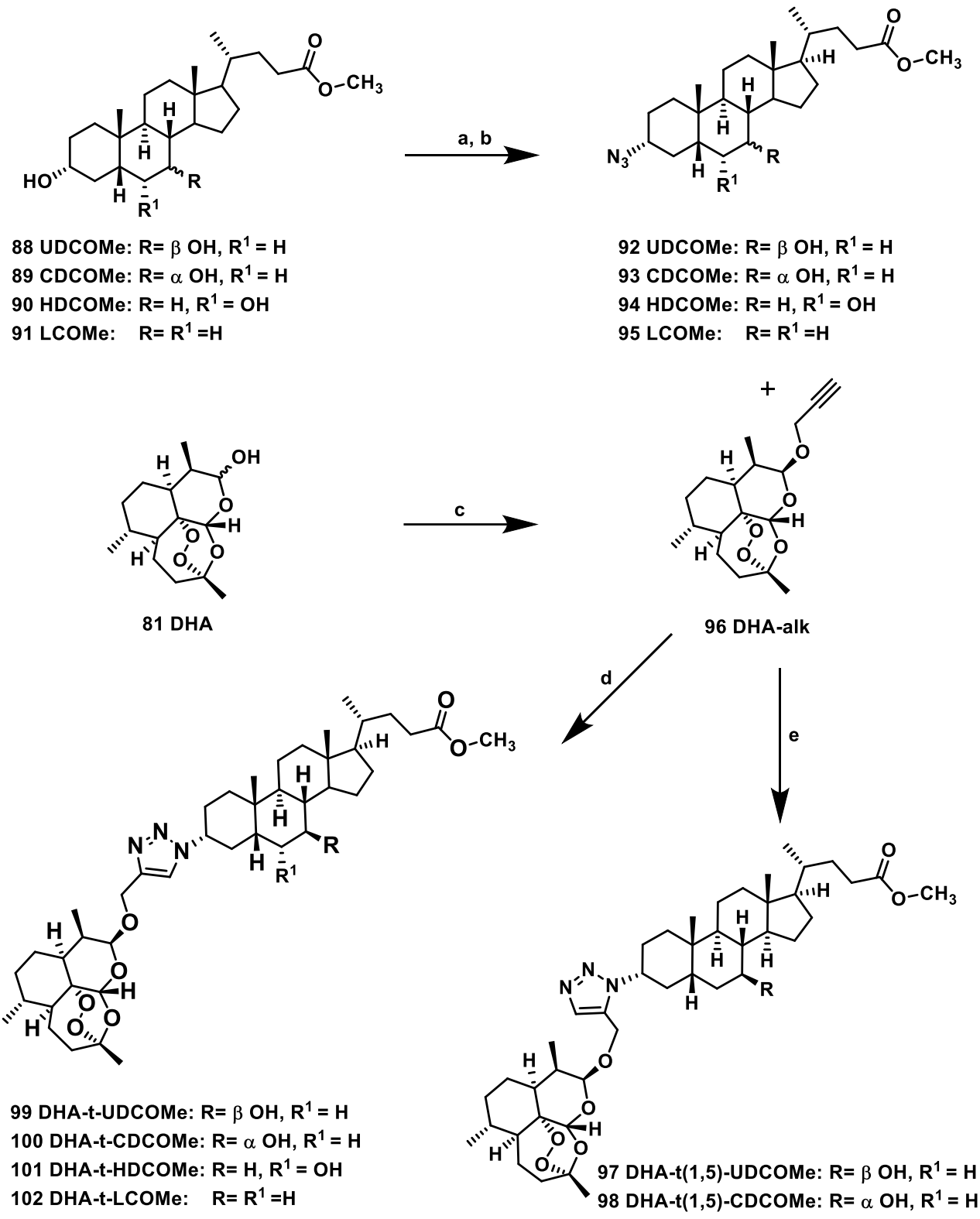


Figure 38. *Synthetic scheme of dihydroartemisinin-bile acids hybrids. Conditions: a) NBS, Ph_3P, THF, r.t., 60 min. b) NaN_3, DMF, 40 °C 18 h. c) propargyl alcohol, $BF_3{\cdot}Et_2O$, DCM, 0°C to rt, 1 h. d) $CuSO_4$, sodium ascorbate, $CH_2Cl_2/H_2O/CH_3CN$ (1:1:0.1) r.t., 4 h. e) Cp*RuCl$(PPh_3)_2$, THF, reflux, 18 h.*

As mentioned above the first issue was to optimize the steps conducing to DHA-t-UDCOMe conjugate. The first step is the protection of ursodeoxycholic acid, which is a high yielding Fischer esterification and needs no optimization. The second step is critical as it involves an Appel-type substitution of the 3-position hydroxy group, which must be regioselective to avoid waste of starting material and time in

purification steps. Thanks to the slightly different chemical environment provided by the three-dimensional shape of bile acids and their hydroxyls in position 3, 7, and 6 in the case of hyodeoxycholic acid; these molecules have slightly different rates of reaction in the conditions employed for this reaction. The reaction was at first executed on UDCOMe with the already reported conditions employing triphenylphosphine and imidazole as hydroxy activators and iodine for substitution.[252,253] Although these conditions allowed to obtain good yields, the scale up of the process was hindered by difficulties in controlling the reaction temperature with subsequent formation of doubly substituted side-product. It is to note that in this system the regioselectivity is given by the fact that the hydroxyl activation by triphenylphosphine is faster if the group is in a more accessible position. The relative steric hindrance of the various secondary hydroxyls on a bile acid scaffold is dependent on their position on the cyclopentanoperhydrophenanthrene ring as well as on the direction the group is pointing toward (stereochemistry). Wanting to avoid uneconomical and time-consuming chromatographic purification, N-bromo succinimide (NBS) was employed. Its slower reactivity, in respect to iodine, allowed us to improve the control over regioselectivity in gram scale reactions and demonstrated itself as advantageous in conversion of four different bile acids: namely UDCA, CDCA, HDCA and LCA. After work-up, precipitation of the residual triphenylphosphine oxide was performed by addition of $MgCl_2$[254] furnishing the 3-β -bromo derivatives of all bile acids in good yield and purity; except for HDCA, which still presented selectivity issues. It is not completely excluded that selectivity could be improved also on hyodeoxycholic methyl ester, but since enough product was obtained from the first reaction there was no attempt to specifically improve the procedure for this bile acid. Substitution of the secondary bromide with sodium azide gave the corresponding azide product in good and consistent yield for all bile acids modified in this work. This double step procedure was intended to give the azide with the same stereochemistry as the starting compound. Functionalization of DHA with the alkyne moiety was already reported in literature and proceeded through a Lewis acid catalysed glycosylation reaction in good yield. Having obtained the two conjugate units, the next critical step was the optimization of the click reaction conditions.

Click optimization

At first, the CuI was employed as a catalyst in anhydrous acetonitrile, but this method only afforded a maximum yield of 39% after removal of oxygen from the solvent and the reaction atmosphere (entry 2, Table 1). The addition of 10 % DMSO, which could act as a radical scavenger and prevent DHA-Alk degradation, did not improve the yield (entry 3, Table 1). Changing solvent to dichloromethane and adding a base was not beneficial at all, and no product formation was observed (entry 4, Table 1). The use of a biphasic catalysis in the presence of $CuSO_4$ and sodium ascorbate in commonly used conditions, afforded the desired compound in low to moderate yields depending on the reaction time (entries 5-6, Table 1). On the contrary, the addition of 10% acetonitrile to the biphasic system greatly improved the yield and diminished the reaction time (entry 7, Table 1). It is noteworthy that in the optimized conditions very little degradation of the starting material DHA-alkyne derivative took place. Since it is reported that artemisinin can degrade in the presence of ferrous ion, Fe(II)-heme, or biological reductants,[199,255] we speculated that the presence of copper (I), which is essential for the CuAAC catalysis, could also bring to the cleavage of the endoperoxide in DHA-alkyne derivative in reductive conditions as previously reported by others.[256] With this in mind, we also speculated that the optimized conditions allowed a better yield due to the faster reaction rate of the click in respect to the degradation reaction. In the conditions reported in entry 7, (Table 1) it was possible to obtain one gram of hybrid **99** (DHA-t-UDCOMe) per batch.

Table 1. Reaction conditions study for the CuAAC click reaction.

Entry[a]	DHA-Alk (eq.)	Solvent	Catalytic system (eq)	Reaction time (hours)	Yield (%)[c]
1	1	CH_3CN	CuI (0.1)	18.0	28 [d]
2	1	CH_3CN [b]	CuI (0.1)	2.0	39
3	1	CH_3CN:DMSO[b] (1:0.1)	CuI (0.1)	2.0	27
4	0.83	CH_2Cl_2/DIPEA [b]	CuI (0.5)	3.0	0
5	1.1	CH_2Cl_2:H_2O (1:1)	NaAsc (0.5) $CuSO_4$ (0.5)	3.5	25
6	1.1	CH_2Cl_2:H_2O (1:1)[b]	NaAsc (0.15) $CuSO_4$ (0.05)	8.0	18
7	1.1	CH_2Cl_2:H_2O:CH_3CN (1:1:0.1)	NaAsc (0.15) $CuSO_4$ (0.05)	1.0	72

[a]All reactions were carried out at room temperature by employing 1 equivalent of 3-a-N_3UDCMe. [b]CH_3CN and CH_2Cl_2 were freshly dried, then degassed by freeze pump thaw method and employed under argon atmosphere. [c]Yields obtained after chromatographic purification. [d]Conditions reported in ref [257]

Once optimized the click reaction it was possible to use the same conditions for the conjugation all the different bile acids. The reaction proceeded smoothly except for the lithocholic acid derivative, probably due to its higher lipophilicity, in respect to the other bile acids. Accounting for this difference, a larger amount of acetonitrile was used. However, the reaction still proceeded slowly and in low yields.

As mentioned before, the ruthenium catalysed azide alkyne cyclization brings to 1,5 triazole formation. Wanting to expand the number of hybrids with a triazole linker it seemed logical to employ this metal catalysis to obtain the conjugates of DHA with UDCA and CDCA. This different triazole linker should retain the hydrogen bonding abilities of the more common 1,4 triazole while giving the conjugate a different three-dimensional shape. To obtain the conjugates, ruthenium catalysts Cp*$RuCl(PPh_3)_2$ was allowed to coordinate the alkyne moiety at room temperature in tetrahydrofuran. After five minutes the azide substrate was added, and the reaction mixture heated to reflux. Ruthenium catalysis proceeds in slower fashion in respect to the copper one, and in this case was not optimized.

Until now, all conjugates were obtained with the acidic portion of the bile acid protected as a methyl ester. However, it could be possible that interaction with receptors for bile acids, particularly in cell lines which are commonly involved in in the gastrointestinal circulation could increase the efficacy of the conjugate by helping internalization. Initially, it was attempted to deprotect the bile acid after azide modification. Although the deprotection was successful, it was not possible to obtain the conjugate in good yields. Even though CuAAC is a robust methodology DHA degradation was observed when its reaction with the free bile acids was attempted. Thus, to obtain the free acid conjugates a deprotection in alkaline conditions was employed on the conjugate of interest, as reported in Figure 39.

Another point of variability is given by the stereochemistry of the C-3 azide on the bile acid. Indeed, inverting the configuration of this azide can drastically change the three-dimensional structure of the conjugate, putting the triazole on the β face of the bile acid. Only a slight deviation from the procedures showed above is necessary to obtain the azide in this configuration. Moreover, activation of the 3-position hydroxyl with tosyl chloride proceeds in good yield and good selectivity, at least for chenodeoxycholic acid. it was thus possible to obtain the conjugate.

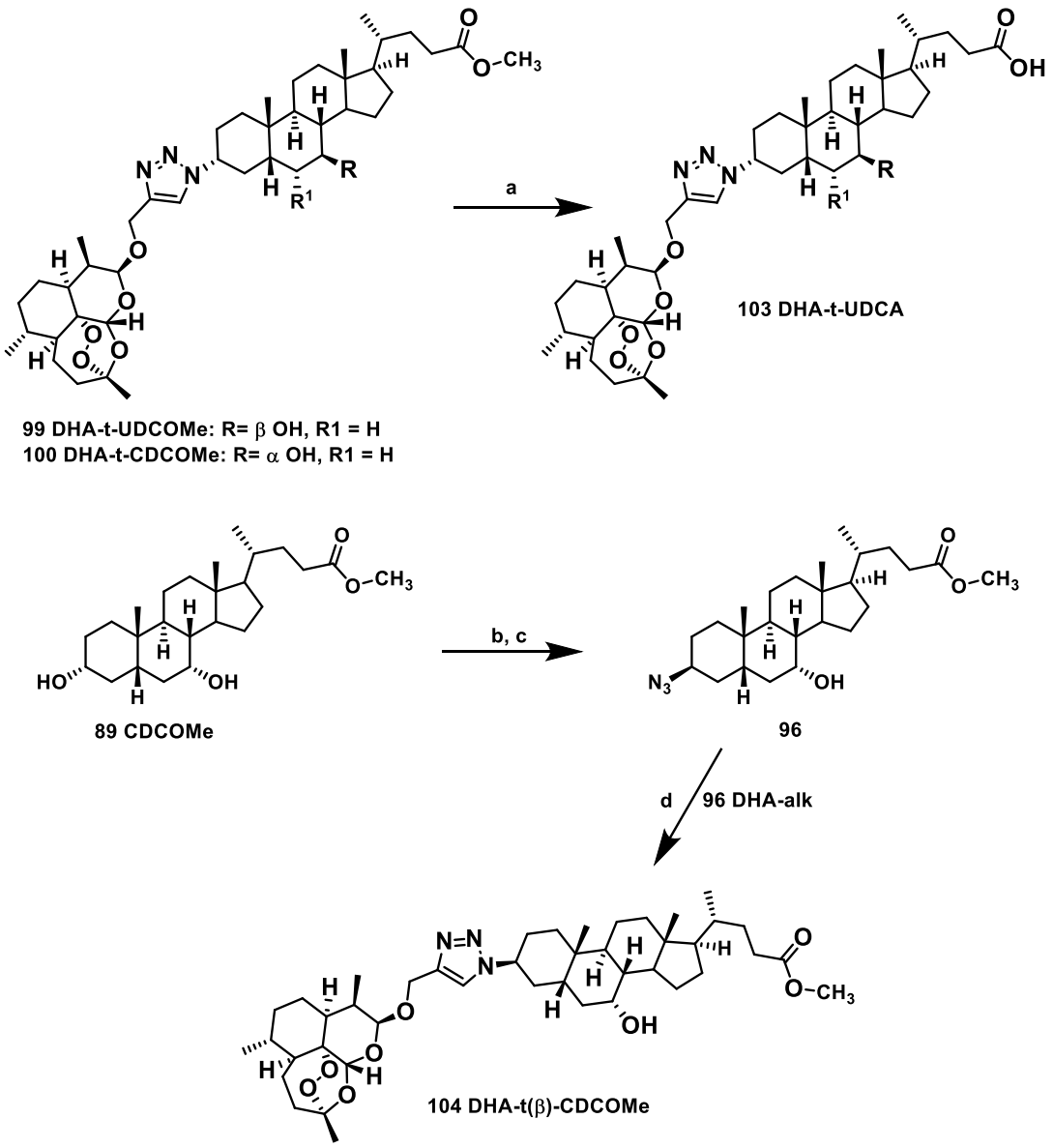


Figure 39. *Synthetic scheme of conjugates deprotection. And β face azide synthesis. a) LiOH, MeOH, rt, 24h. b) Ts-Cl, Py, rt, 4h. c) NaN_3, DMF, rt, 18 h. d) $CuSO_4$, sodium ascorbate, $CH_2Cl_2/H_2O/CH_3CN$ (1:1:0.1) r.t., 4 h.*

Conclusion

In the project outlined above the aim was to see if it was possible to repurpose conjugates between dihydroartemisinin and bile acids as antiviral agents against SAR-CoV-2. After the first preliminary biological assays determining the efficacy of one conjugate (DHA-t-UDCOMe), the need for a larger quantity of the same compound stimulated us to improve the synthetic procedure and optimize the reaction conditions of the Appel type substitution step as well as the CuAAC step. This allowed to obtain the desired product in good yield and gram quantities. After this optimization and the confirmation of

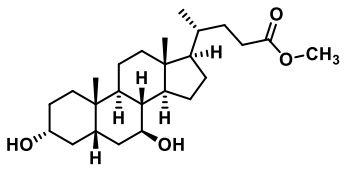
the biological results on human lung cell lines, our attention shifted on the synthesis of similar compounds to assess possible differences in activity at the slight structure variation of the compounds. For this purpose, more bile acids were modified in a similar fashion as before and more conjugates were obtained. For two bile acids, in particular CDCA and UDCA the hybrid was also formed using ruthenium catalysis thus giving a 1,5 triazole. Moreover, the corresponding free acid side chain was obtained on the conjugates. this should increase the water solubility of the compounds and could add a potential internalization mechanism to the conjugate through receptor ligation. Lastly, in one example (DHA-t(β)-CDCOMe) the triazole was made in order to point to the β face of the bile acid instead of the normal α one. Further biological testing will be necessary to evaluate the antiviral potential of the new conjugates. In addition, since DHA-t-UDCOMe was originally tested against cancer cells it may be possible to also test these new conjugates in the same assays.

Materials and methods

Commercial dihydroartemisinin (DHA) and EDCI were purchased from Carbosynth (Compton, Berkshire, UK). UDCA was kindly furnished by ICE SpA (Reggio Emilia, Italy). Sodium azide (NaN_3), sodium ascorbate, triphenylphosphine (PPh_3), propargylic alcohol and N-bromosuccinimide (NBS) and $Cp^*RuCl(PPh_3)_2$ were purchased from Sigma-Aldrich (St. Louis, MO, USA); copper sulphate ($CuSO_4.5H_2O$) was purchased from Fluka (Buchs, Switzerland); imidazole was obtained from Acros (Geel, Belgium). All the chemicals were used without further purification. The reactions were monitored by TLC on pre-coated silica gel F254 plates (thickness 0.25 mm, Merck), developed with phosphomolybdic acid solution. Flash column chromatography was performed on silica gel (60 Å, 230–400 mesh). NMR spectra were recorded with a Varian Mercury 300, 400 or 500 MHz Varian Mercury instrument in the stated solvent. Preparative and analytical HPLC was executed on Xterra C18 column with a mobile phase composed of water and CH_3CN in gradient at room temperature. Anhydrous solvents were freshly distilled on sodium/benzophenone with standard procedures or purchased anhydrous. Known compounds were synthetised according to literature procedures, and in all cases characterization data were in agreement with literature. [252,257–260]

Synthesis of conjugates

3α, 7β-dihydroxy 24-methyl cholate. (UDCOMe)

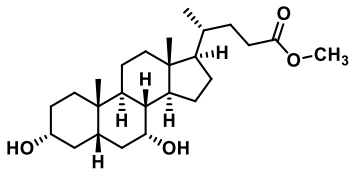


20.0g (51.0 mmol) of ursodeoxycholic acid (UDCA), were dissolved in the minimum amount of methanol in round bottom flask. 0.5 ml of concentrated sulphuric acid (H_2SO_4) were added dropwise while stirring. (H_2SO_4). The flak was equipped with a water condenser and the reaction mixture refluxed overnight. the reaction was monitored by TLC and at completion 100 ml of water were added. Methanol was evaporated under vacuum to give a white solid suspended in water. The solid was filtered on paper with a Buchner, washed water, and left to dry in air. After one day 18.59 g (45.7 mmol) of white powder were obtained in 89.7% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ:3.85 (q, 1H, H7, *J* = 3.0 Hz), 3.66 (s, 3H, CH_3, OMe), 3.46 (dt, 1H, H3, *J* = 11.0, 5.5 Hz), 2.40-0.93 (m, 24H), 0.92 (d, 3H, $CH_3$21, J=6.4 Hz), 0.90 (s, 3H, $CH_3$19), 0.65 (s, 3H,

$CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.7, 71.1, 60.8, 55.6, 54.8, 51.5, 43.7, 43.6, 42.7, 40.0, 39.1, 36.6, 35.2, 35.1, 34.1, 33.4, 31.0, 31.0, 28.6, 26.8, 26.6, 23.4, 21.1, 18.4, 12.1.

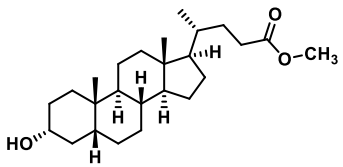
3α, 7α-dihydroxy 24-methyl cholate (CDCOMe)



10,0g of chenodeoxycholic acid (CDCA) (25.5 mmol), were dissolved in the minimum amount of methanol in round bottom flask. 0.5 ml of concentrated sulphuric acid (H_2SO_4) were added dropwise while stirring. (H_2SO_4). The flak was equipped with a water condenser and the reaction mixture refluxed overnight. the reaction was monitored by TLC and at completion 100 ml of water were added. Methanol was evaporated under vacuum to give a white solid suspended in water. The solid was filtered on paper with a Buchner, washed water, and left to dry in air. 8.55 g (21.04 mmol) of off-white powder were recovered in 82,6 % yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ:3.85 (q, 1H, H7, *J* = 3.0 Hz), 3.66 (s, 3H, OMe), 3.46 (dt, 1H, H3, *J* = 11.0, 5.5 Hz), 2.40-0.93 (m, 24H), 0.92 (d, 3H, $CH_3$21, J=6.4 Hz), 0.90 (s, 3H, $CH_3$19), 0.65 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.7, 68.2, 61.3, 55.7, 51.5, 50.3, 42.6, 41.8, 39.6, 39.5, 39.3, 35.5, 35.4, 35.3, 35.1, 34.4, 32.7,30.9, 28.1, 26.8, 23.7, 22.8, 20.5, 18.2, 11.7.

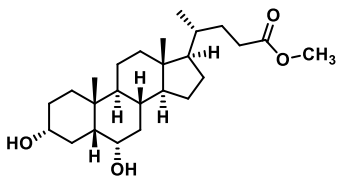
3α-hydroxy 24-methyl cholate (LCOMe)



5,00 g of lithocholic acid (LCA) (13.28 mmol) were dissolved in the minimum amount of methanol in round bottom flask. 0.5 ml of concentrated sulphuric acid (H_2SO_4) were added dropwise while stirring. (H_2SO_4). The flak was equipped with a water condenser and the reaction mixture refluxed overnight. the reaction was monitored by TLC and at completion 100 ml of water were added. Methanol was evaporated under vacuum to give a white solid suspended in water. The obtained off-white powder was dried in air to afford 5,06 g (12.96 mmol) of product with 97,6% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 3.66 (s, 3H, OMe), 3.61(m, 1H, H3), 2.39-0.96 (m, 25H), 0.91-0,90 (6H, $CH_3$19,21), 0.64 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.8, 61.6, 56.4, 55.9, 51.5, 42.7, 42.3, 40.4, 40.0, 35.8, 35.5, 35.4, 34.6, 32.4, 31.0, 30.9, 28.2, 27.1, 26.7, 26.3, 24.2, 23.4, 20.8, 18.3, 12.0.

3α, 6α-dihydroxy 24-methyl cholate (HDCOMe)

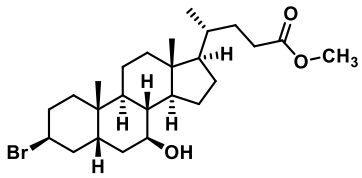


10,0g of hyodeoxycholic acid (HDCA) (25.5 mmol), were dissolved in the minimum amount of methanol in round bottom flask. 0.5 ml of concentrated sulphuric acid (H_2SO_4) were added dropwise while stirring.

(H_2SO_4). The flak was equipped with a water condenser and the reaction mixture refluxed overnight. the reaction was monitored by TLC and at completion 100 ml of water were added. Methanol was evaporated under vacuum to give a white solid suspended in water. The solid was filtered on paper with a Buchner, washed water, and left to dry in air. 8.93 g (22.0 mmol) of off-white powder were recovered in 86 % yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.01 (dt, J = 11.9, 4.7 Hz, 1H, H6), 3.64 (s, 3H, OMe), 3.62 – 3.52 (m, 1H, H3), 2.43 – 0.95 (m, 27H, aliphatic protons), 0.93 – 0.83 (m, 6H, CH_3 19,21), 0.62 (s, 3H, CH_3 18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.85, 71.58, 68.09, 56.32, 56.06, 51.61, 48.56, 42.95, 40.08, 39.99, 36.05, 35.73, 35.47, 34.93 (2C), 31.18, 31.05, 30.20, 29.36, 28.23, 24.32, 23.64, 20.87, 18.35, 12.12.

3β- Bromo, 7β-hydroxy 24-methyl cholate (3-β-BrUDCMe)

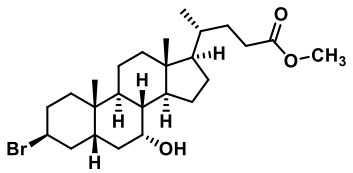


4.00 g (9.8 mmol) of UDC methyl ester (UDCOMe) and 3.85 g (14.7 mmol) of triphenylphosphine (PPh_3) were added to a two-necked 100 ml round bottom flask equipped with a thermometer. 20 ml of anhydrous tetrahydrofuran (THF) were added, the flask was stoppered, and the solution stirred until dissolution of the solids. The solution temperature was lowered with an ice bath until it reached 0 °C and 2.61 g (14.7 mmol) of N-bromosuccinimide (NBS) were added in three times. After 5 minutes the initial temperature raise was completely mitigated, and the flask was allowed to reach room temperature after removal of the ice bath. The solution was stirred for 35 minutes and then the reaction was quenched by addition of 30 ml 10% hydrogen peroxide solution (H_2O_2). After 10 minutes of stirring the solvent was removed under vacuum without heating. The resulting gummy reddish solid in water was dissolved in 50 ml of CH_2Cl_2 and the organic phase was washed five times with 10 ml of 10% $Na_2S_2O_5$ solution to remove excess bromine. The clear organic phase was dried by addition of anhydrous sodium sulphate, filtered and a part of solvent removed under vacuum to leave only 10 ml of solution. This was transferred to a 50 ml round bottom flask and 2.83 g (29.8 mmol) of $MgCl_2$ were added to it. The solid was stirred overnight and then filtered on silica to afford, after solvent removal, 3.40 g of the product as a yellowish oil in 74% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 4.73 (s, 1H), 3.68 (s, 3H), 3.56 – 3.29 (m, 1H), 2.44 – 1.05 (m, 28H), 1.03 (s, 3H), 0.92 (d, J = 6.3 Hz, 3H), 0.68 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 175.15, 72.11, 56.29, 56.09, 55.42, 51.97, 44.23, 44.15, 40.60, 40.10, 38.42, 36.60, 35.73, 35.31, 31.54, 31.50, 31.05, 30.00, 29.06, 27.33, 24.15, 21.74, 18.86, 12.61.

HRMS ESI (+) $C_{25}H_{41}BrO_3$: 469.2158 $[M+H]^+$, 491.2189 $[M+Na]^+$

3β-bromo, 7α-hydroxy 24-methylcholate (Br-CDCOMe)



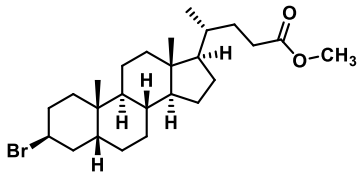
2.000 g of CDC-OMe (4.92 mmol) and 1.936 g triphenylphosphine (PPh_3) (7.378 mmol, 1.5 eq) were dissolved in 20 ml anhydrous THF under nitrogen atmosphere in a 100 ml round bottomed flask. The

temperature was brought to 0 °C with the aid of an ice bath and 1.313 g of N-Bromosuccinimide (NBS) (7,378mmol, 1,5 eq) were added in three portions to avoid a temperature increase. After complete addition the reaction mixture was allowed to reach room temperature and monitored by NMR and TLC.

After one hour the reaction was stopped by dilution with 20 ml 20% hydrogen peroxide (H_2O_2). THF was evaporated under vacuum and a dense solid started to form. 50 ml dichloromethane were added to, and the flask shaken until the solid dissolved. The mixture was poured into a 500 ml separatory funnel and the aqueous phase was discarded. The organic phase was washed with 20 ml water, 4X10 ml 5% solution of sodium metabisulphite, 15 ml brine. The organic phase was then dried with anhydrous sodium sulphate, filtered on cotton, and dried under vacuum to give a white powder. This was dissolved in the minimum amount of ethyl acetate (EtOAc) and 1.405 g of magnesium chloride were added in solution. the suspension was stirred overnight at room temperature, and then filtered on silica. The solvent was removed under vacuum to afford 1.400 g (2.98 mmol) of off-white powder in 60.5 % yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 4.76 (q, 1H, H3, J=3.1 Hz), 3.85 (d, 1H, H7, J=3.1Hz), 3.66 (s, 3H, OMe), 2.40 - 1.03 (m, 24H), 0.99 (s, 3H, $CH_3$19), 0.92 (d, 3H, $CH_3$21, J=6.3 Hz), 0.66 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.75, 68.41, 57.69, 55.85, 51.55, 50.51, 42.73, 39.65, 39.35, 38.34, 36.92, 35.83, 35.40, 34.05, 33.11, 31.05, 31.02, 30.91, 29.63, 28.19, 23.73, 23.07, 20.77, 18.32, 11.82.

3β- Bromo- 24-colanato di metile (Br-LCOMe)

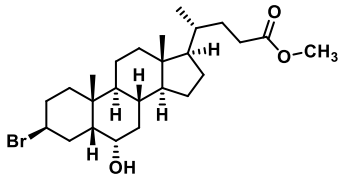


4,00 g LCOMe (10.24 mmol) and 4,03 g (15.36 mmol, 1,5 eq) triphenylphosphine (PPh_3) were dissolved in 40 ml anhydrous THF under nitrogen atmosphere in a 100 ml round bottomed flask. The temperature was brought to 0 °C with the aid of an ice bath and 2,73g di NBS (15,36 mmol, 1,5 eq) were added in three portions to avoid a temperature increase. After complete addition the reaction mixture was allowed to reach room temperature and monitored by NMR and TLC.

After one hour the reaction was stopped by dilution with 20 ml 20% hydrogen peroxide (H_2O_2). THF was evaporated under vacuum and a dense solid started to form. 50 ml dichloromethane were added to, and the flask shaken until the solid dissolved. The mixture was poured into a 500 ml separatory funnel and the aqueous phase was discarded. The organic phase was washed with 20 ml water, 4X10 ml 5% solution of sodium metabisulphite, 15 ml brine. The organic phase was then dried with anhydrous sodium sulphate, filtered on cotton, and dried under vacuum to give a white powder. This was dissolved in the minimum amount of ethyl acetate (EtOAc) and 2,925 g (30.72 mmol, 3 eq) of magnesium chloride were added in solution. the suspension was stirred overnight at room temperature, and then filtered on silica. The solvent was removed under vacuum to afford 4.29 g (14.17 mmol) of off-white powder in 92.3 % yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ:4.79 (s, 1H, H3), 3.66 (s, 3H, H del CH_3), 2,40-1,03 (m, 25H), 1.00 (s, 3H, $CH_3$19), 0.90 (d, 3H, J=6,34 Hz, $CH_3$21), 0.64 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.77, 57.10 (C3), 56.58, 55.97, 51.49, 42.72, 40.83, 40.16, 37.36, 35.67, 35.36, 35.32, 31.04, 30.98, 30.87, 29.68, 28.16, 26.67, 26.33, 24.14, 23.66, 20.90, 18.25, 12.04.

3β-Bromo, 6α-hydroxy 24-methylcolate (BrHCOMe)

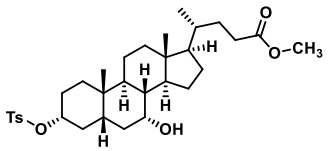


2.000 g of HDC-OMe (4.92 mmol) and 1.936 g triphenylphosphine (PPh_3) (7.378 mmol, 1.5 eq) were dissolved in 20 ml anhydrous THF under nitrogen atmosphere in a 100 ml round bottomed flask. The temperature was brought to 0 °C with the aid of an ice bath and 1.313 g of N-Bromosuccinimide (NBS) (7,378mmol, 1,5 eq) were added in three portions to avoid a temperature increase. After complete addition the reaction mixture was allowed to reach room temperature and monitored every 10 minutes by TLC (cyclohexane/EtOAc 5:1). Bromide substitution of both position 3 and 6 was observed and the reaction was allowed to procced for one hour. Two spots were observed on TLC the desired mono-brominated product in position 3 (Rf=0.3) and the dibrominated byproduct (Rf=75).

After one hour the reaction was stopped by dilution with 20 ml 20% hydrogen peroxide (H_2O_2). THF was evaporated under vacuum and a dense solid started to form. 50 ml dichloromethane were added to, and the flask shaken until the solid dissolved. The mixture was poured into a 500 ml separatory funnel and the aqueous phase was discarded. The organic phase was washed with 20 ml water, 4X10 ml 5% solution of sodium metabisulphite, 15 ml brine. The organic phase was then dried with anhydrous sodium sulphate, filtered on cotton, and dried under vacuum to give a white powder. This was dissolved in the minimum amount of ethyl acetate (EtOAc) and 1.405 g of magnesium chloride were added in solution. the suspension was stirred overnight at room temperature, and then filtered on silica. The solid was purified by flash chromatography (5:1 cy/EtOAc) to afford 655 mg of product in 28.4 % yield. in addition, 387 mg of byproduct were recovered in 14.8% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ:4.83 (s, 1H, H3), 4.16 (dt, 1H, H6), 3.66 (s, 3H, H OMe), 2,40-1,01(m, 24H), 0.99 (s, 3H, $CH_3$19), 0.90 (d, J=6,30 Hz, 3H, $CH_3$21), 0.64 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.70, 67.50, 56.23, 55.93, 55.63 (C3), 51.49, 44.06, 42.84, 40.19, 39.94, 36.62, 35.32, 34.95, 34.62, 31.03, 29.49, 28.37, 28.08, 26.89, 24.12, 23.80, 20.85, 18.25, 12.03.

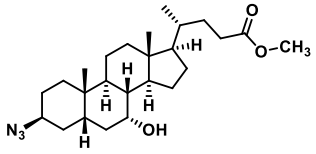
3α-tosyl, 7α-hydroxy 24-methylcholate (Ts-CDCOMe)



4,854 g of CDCOMe (11.94 mmol, 1 eq) and 4,552 g tosyl chloride (23.88 mmol, 2 eq) were dissolved in 60 ml anhydrous pyridine. The reaction mixture was stirred at room temperature for 4 hours and monitored by TLC. At reaction complete the mixture was transferred to a separatory funnel and diluted with 120 ml water and extracted with 3X 60 ml Et_20. The combined organic phases were washed with 3X 30 ml 5% HCl solution, dried with anhydrous sodium sulphate and filtered on cotton. The solvent was removed under vacuum to give 4.922 g of product in 96% yield. spectroscopic data were in agreement with the literature. [258]

^{1}H NMR (500 MHz $CDCl_3$): selected data δ 7.80 (2H, d, J = 8.5 Hz), 7.34 (2H, d, J= 8.5 Hz), 4.38 (1H, m), 3.83 (1H, s), 3.62 (3H, s), 2.40 (3H, s) 2.35 (1H, m), 2.23 (1H, m), 0.93 (3H, d, J = 6.4 Hz), 0.89 (3H, s), 0.65 (3H, s).

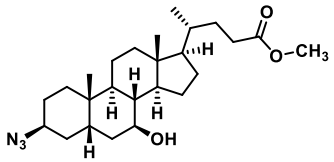
3β-azido, 7α-hydroxy 24-methylcholate (N_3-CDCOMe)



4,922 g di TsO-CDCOMe (8.7769 mmol, 1eq) were dissolved in 50 ml DMF and 1,711 g of NaN3 (26,33 mmol, 3eq) were added. The reaction mixture was warmed to 50 °C and stirred for 18 hours. Work-up was executed by diluting the reaction mixture with 120 ml diethyl ether and washing with 4X 30 ml 5% LiCl solution. The organic phase was dried with anhydrous sodium sulphate and filtered on cotton. The solvent was removed under vacuum to give 3.435 g of product in 91% yield.

1H NMR (400 MHz, CDCl3) selected data: δ 3.93 – 3.82 (m, 1H), 3.66 (s, 3H), 3.61 (s, 1H), 1.24 – 1.06 (m, 3H), 0.66 (s, 3H).

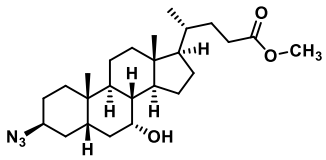
3α- azido, 7β-hydroxy 24-methylcholate (N_3-UDCOMe)



1,634 g of Br-UDCOMe (3,48mmol) were dissolved in 44,26 ml N,N-dimetilformammide (DMF) in a 100 ml round bottom flask, 680 mg (10,45 mmol) of sodium azide were added. The solution heated at 40 °C in an oil bath and stirred overnight. The reaction mixture was poured into a separatory funnel and 70 ml of ethyl ether (Et_2O) were added and the mixture was washed with 15 ml 5% lithium chloride (LiCl) solution twice. The organic phase was further washed with 2X 10 ml brine and dried with sodium sulphate anhydrous. The solvent was removed under vacuum to give 1.066 g (2,47 mmol) of product in 71% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ:3.53 (s, 3H, OMe), 3.42 (dq, 1H, H7, J = 9.0, 5.2 Hz), 3.17 (ddt, 1H, H in 3, J = 16.2, 10.6, 4.8 Hz), 2.28- 0.91 (m, 24H), 0.85 (s, 3H, $CH_3$19), 0.80 (d, 3H, $CH_3$21, J= 6.3Hz), 0.56 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.83, 71.30, 61.01, 55.75, 54.99, 51.65, 43.84 (2C), 42.86, 40.14, 39.29, 36.75, 35.37, 35.27, 34.28, 33.54, 31.17 (2C), 28.71, 26.98, 26.73, 23.60, 21.31, 18.52, 12.26.

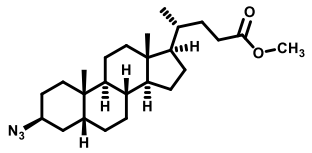
3α-azido, 7α-hydroxy 24-methylcholate (N_3-CDCOMe)



1.519 g of Br-CDCOMe (3.239 mmol) were dissolved in 41 ml DMF in a 100 ml round bottom flask. To this solution 632 mg NaN_3 (9.716mmol, 3 eq) were added and the mixture was heated at 40 °C in an oil bath. The reaction mixture was stirred overnight and poured into a separatory funnel. Into this were added 70 ml Et_2O and the organic phase was washed with 15 ml 5% lithium chloride (LiCl) solution twice and with 2X 10 ml brine. The organic solvent was dried with sodium anhydrous sulphate, filtered on cotton, and removed under vacuum to afford 1.136 g (2.63 mmol) of white powder as product in 81,3% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ:3.85 (d, 1H, H7, J=3.1Hz), 3.66 (s, 3H, OMe), 3.15 (tt, 1H, H3, J=11.9 4.2 Hz), 2.41-0.96 (m, 24H), 0.93- 0.91(d, 6H, $CH_3$19,21), 0.66 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, CDCl3) δ 174.88, 68.46, 61.49, 55.87, 51.63, 50.49, 42.81, 41.94, 39.64, 39.52, 35.69, 35.59, 35.49, 35.26, 34.57, 32.90, 31.13, 31.10, 28.27, 26.95, 23.83, 22.99, 20.70, 18.39, 11.90.

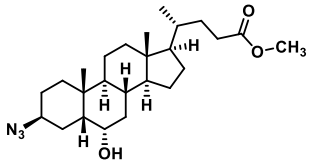
3α-azido -24methylcholate (N_3-LCOMe)



4,259 g of Br-LCOMe were dissolved in 65 ml DMF in a 100 ml round bottom flask. To this solution were added 1.832 g NaN_3 (28.179 mmol, 3 eq) and the reaction mixture was brought to °C with an oil bath. The mixture was stirred overnight and then poured into a separatory funnel, to this were added 100 ml ethyl ether and the organic phase was washed with 2x 20 ml 5% lithium chloride solution and 2X 30 ml brine. The organic solvent was dried with sodium sulphate anhydrous, filtered on cotton and evaporated under vacuum to give 3.353 (8.07 mmol) of white powder as a clean product in 85.9%.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 3.66 (s, 3H, OMe), 3.31(m, 1H, H3), 2.41-0.97 (m, 25H), 0.93 (s, 3H, $CH_3$19), 0.91 (d, 3H, J=6,23Hz, $CH_3$21), 0.64 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.75, 61.25 (C3), 56.37, 55.90, 51.46, 42.71, 42.36, 40.43, 40.04, 35.77, 35.54, 35.34, 34.62, 32.44, 31.04, 30.98, 28.15, 27.07, 26.72, 26.31, 24.16, 23.43, 20.80, 18.25, 12.03.

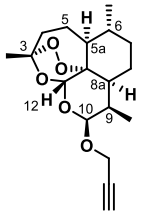
3α-azido, 6α-hydroxy 24-methylcholate (N_3-HDCOMe)



579 mg of Br-HCOMe (1.234 mmol) were dissolved in 15.67 ml DMF in a 50 ml round bottom flask. To this solution 241 mg (3.703 mmol, 3 eq) NaN_3 were added, and the reaction mixture was heated to 40°C. After overnight stirring, the solution was poured into a separatory funnel and 50 ml Et2O were added. The organic phase was washed with 2x 20 ml 5% lithium chloride solution and 2X 30 ml brine, dried with anhydrous sodium sulphate and filtered on cotton. The solvent was removed under vacuum to afford 440 mg (1.02 mmol) of product as a clear oil in 82,6% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ:4.07 (m, 1H, H6), 3.66 (s, 3H, OMe), 3.29 (m, 1H, H3), 2.39-1,02 (m, 24H), 0.92 (s, 3H, $CH_3$19), 0.90 (d, 3H, $CH_3$21), 0.64 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 179.88, 68.13, 61.12, 56.11, 55.97, 48.68, 42.96, 39.91, 36.07, 35.80, 35.41, 35.01, 34.87, 31.07, 30.83, 28.21, 26.81, 25.59, 24.28, 23.68, 20.85, 18.34, 12.16,

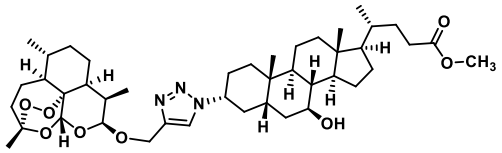
Alkyne derivative of dihydroartemisinin (DHA-alk)



2.00 g of dihydroartemisinin (DHA) (7,04 mmol) were dissolved in 20 ml anhydrous dichloromethane (DCM) in a 50 ml round bottom flask under nitrogen atmosphere. 1.60 ml (28 mmol, 4 eq) propargyl alcohol were added and the reaction mixture cooled to °C in an ice bath. 60 µl of BF_3 Et_2O (0.486mmol, 0.07 eq) were slowly added and the ice bath was removed. The reaction progress was monitored by ^{1}H-NMR, and after 45 minutes, the Lewis acid was quenched by addition of 25 ml saturated sodium bicarbonate solution ($NaHCO_3$). The mixture was poured into a separatory funnel and the aqueous phase discarded. The organic phase was dried with anhydrous sodium sulphate and filtered on cotton. The solvent was removed under vacuum to afford 2.00 g of product as a white powder in 88.5 % yield.

^{1}H-NMR (400 MHz, $CDCl_3$) δ : 5.39 (s, 1H, H12), 4.95 (d, 1H, J 3.3, H10), 4.29 (d, 2H, J 2.4), 2.65 (m, 1H), 2.40–2.30 (m, 2H), 2.05–1.17 (m, 10H), 1.42 (s, 3H), 0.93 (d, 3H, J 6.0), 0.91 (d, 3H, J 7.2) ^{13}C NMR (75 MHz, $CDCl_3$) δ : 104.1, 100.6, 88.0, 81.0, 79.7, 73.9, 54.9, 52.5, 44.3, 37.4, 36.4, 34.6, 30.6, 26.1, 24.7, 24.4, 20.3, 12.7.

3α-triazolyl-DHA,7β-hydroxy 24-methylcholate (DHA-t-UDCOMe)

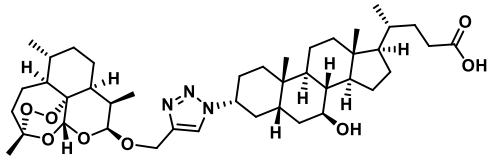


In a 10 ml round bottom flask were weighted 50 mg (0.116 mmol) of N_3-UDCOMe and 40 mg (0.127 mmol) of DHA-Alk. The compounds were dissolved in 0.5 ml of CH_2Cl_2 and to this solution were added 0.5 ml of water and 0.2 ml of CH_3CN. After nitrogen atmosphere was created inside the flask, 46 µl of a 0.5 M (0.05 eq.) solution of sodium ascorbate and 66 µl of a 0.1 M (0.15 eq) $CuSO_4$ solution were added to the mixture. After four hours of emulsification at room temperature the reaction mixture was poured into 20 ml of water and extracted with 20 ml CH_2Cl_2. The organic solvent was dried with anhydrous sodium sulphate and the solvent removed under vacuum. The crude product was then purified by flash chromatography (EtOAc/Cy 1:1) to give 64 mg of the product as a white powder in 73% yield.

^{1}H NMR (300 MHz, $CDCl_3$) selected data δ: 7.51 (s, 1H, H triazole), 5.36 (s, 1H, H12 DHA), 4.91 – 4.59 (m, 3H, H10 DHA, CH_2 propargyl), 4.37 (tt, J = 12.1, 4.2 Hz, 1H, H3 UDC-OMe), 3.64 (s, 3H, OMe), 3.61 – 3.51 (m, 1H, H7 UDC-OMe), 2.66 – 2.52 (m, 1H, H9 DHA), 1.03 (s, 3H), 0.93 (d, s, J = 5.8 Hz, 6H), 0.87 (d, J = 7.4 Hz, 4H), 0.67 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.82, 144.48, 120.85, 104.26, 101.61, 88.13, 81.29, 71.22, 61.66, 60.79, 55.70, 55.03, 52.66, 51.66, 44.53, 43.90, 43.26, 40.08, 39.45, 37.54, 36.65, 36.56, 35.60, 35.37, 34.89, 34.72, 34.45, 31.19, 31.19, 30.98, 28.70, 28.20, 26.97, 26.34, 24.85, 24.57, 23.65, 21.39, 20.48, 18.53, 13.12, 12.28.

HRMS ESI ($^+$) $C_{43}H_{67}N_3O_8$: 754.4992 $[M+H]^+$

3α-triazolyl-DHA,7β-hydroxy cholic acid (DHA-t-UDCA)

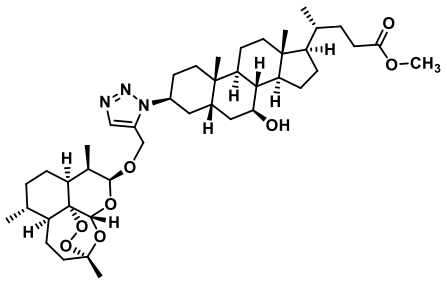


100 mg DHA-t-UDCOMe (0.1327 mmol, 1eq) were dissolved in 3.22 ml MeOH and 1.06 ml 1,5 M LiOH (1.60 mmol, 12eq) solution were added. The solution was stirred at room temperature for 24 hours. 10 ml of water were added to the mixture and the solution pH was brought to 7 by addition of 5% HCl

solution. methanol was removed under vacuum and 5% HCl was added until pH 3 was reached. The solution was extracted with ethyl acetate (50 ml) and the organic solvent was dried with anhydrous sodium sulphate and filtered on cotton. After evaporation of the solvent under vacuum 75 mg of product as a white powder were obtained in 75% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.55 (s, 1H, H triazole), 5.38 (s, 1H, H12 DHA), 4.95 – 4.87 (m, 2H, CH_2 propargyl), 4.72 (d, J = 12.8 Hz, 1H, H10 DHA), 4.47 – 4.34 (m, 1H, H3 UDC), 3.64 – 3.53 (m, 1H, H7 UDC), 2.64 (m, 1H, H9 DHA), 2.46 – 1.06 (m, 30H), 1.04 (s, 3H, CH_3), 0.98 – 0.91 (m, 9H), 0.88 (d, J = 7.3 Hz, 3H, $CH_3$21), 0.70 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 178.48, 144.13, 120.96, 104.14, 101.60, 87.99, 81.12, 71.07, 61.36, 60.88, 55.53, 54.88, 52.50, 44.36, 43.76 (2C), 43.10, 39.92, 39.30, 39.14, 37.37, 36.41 (2C), 35.43, 35.19, 34.69, 34.56, 34.28, 30.80 (2C), 28.56, 28.01, 26.80, 26.17, 24.69, 24.42, 23.48, 21.25, 20.33, 18.36, 12.97, 12.13.

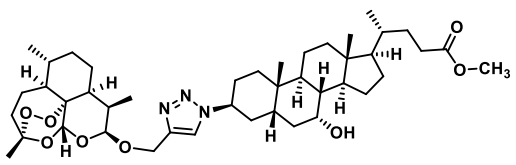
3α-(1,5)triazolyl-DHA,7β-hydroxy 24-methylcholate (DHA-t(1,5)-UDCOMe)



In a two necked round bottom flask were dissolved 825 mg of DHA-alk (2.436 mmol, 1.05 eq) in 17 ml anhydrous tetrahydrofuran (THF), under nitrogen atmosphere. 97 mg of Cp*RuCl(PPh_3)$_2$ (0.121 mmol, 0.05eq) in 17 ml anhydrous THF were added and the mixture was stirred for 5 minutes at room temperature. 1.051 g (2.436 mmol, 1eq) of N_3-UDCOMe in 17 ml dry THF were added, and the reaction mixture was refluxed for 18 hours. at reaction complete solvent is removed under vacuum to give 1.941 g of crude. This is purified by silica flash chromatography (2:3 Cy/EtOAc) to give 450 mg of product as a white foam in 25% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.61 (s, 1H, H triazole), 5.35 (s, 1H, H12 DHA), 4.91 (d, J = 12.8 Hz, 1H, CH_2 propargyl), 4.88 (d, J = 3.5 Hz, 1H, H10 DHA), 4.52 (d, J = 12.8 Hz, 1H, CH_2 propargyl), 4.29 – 4.16 (m, 1H, H3 UDCOMe), 3.66 (s, 3H, OMe), 3.64 – 3.55 (m, 1H, H7 UDCOMe), 2.75 – 2.63 (m, 1H, H9 DHA), 1.45 (s, 3H, CH_3 DHA), 2.43-1.50 (m, 34H, aliphatic protons), 1.03 (s, 3H, CH_3 UDCOMe), 0.93 (dd, J = 6.1, 2.0 Hz, 6H, CH_3 DHA, UDCOMe), 0.88 (d, J = 7.3 Hz, 3H, CH_3 DHA), 0.69 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.86, 133.30, 104.51, 101.84, 88.23, 80.99, 71.19, 59.23, 57.99, 55.43, 54.95, 52.53, 51.65, 44.20, 43.92, 43.70, 39.95, 39.19, 37.63, 36.62, 36.44, 36.06, 35.39, 35.18, 34.56, 34.47, 31.23, 31.16, 30.72, 28.71, 27.72, 26.98, 26.25, 24.79, 24.60, 23.71, 21.42, 20.37, 18.53, 12.97, 12.27.

3β-triazolyl-DHA,7α-hydroxy 24-methylcholate (DHA-tβ-CDCOMe)

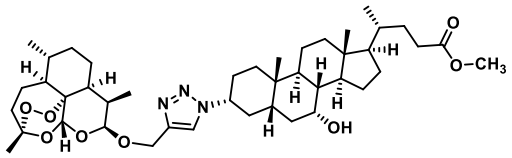


1 g of N_3-CDCOMe (2.317 mmol, 1eq) and 747 mg DHA-ALK (2.317 mmol, 1eq) were dissolved in 22 mixture of water DCM and acetonitrile (1:1:0.1). to this mixture were added 2,317 ml 0,1M $CuSO_4$ solution (0.2317 mmol, 0.1eq) and 2,231 ml 0.5M (1.158 mmol, 0.5eq) sodium ascorbate solution. The

mixture was stirred vigorously for 1 h and monitored by TLC (2:3 Cy/EtOAc). The mixture was poured into a separatory funnel and the aqueous phase was discarded. the organic solvent was washed with water 3X 10 ml, dried over anhydrous sodium sulphate, and filtered on cotton. The solvent was removed under vacuum to afford 1.650 g of crude. This was purified by flash column chromatography (1:1 to 3:2 Cy/EtOAc) to give 848 mg of pure product as a white powder in 51.5% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56 (s, 1H, H triazole), 5.38 (s, 1H, H12 DHA), 4.92 – 4.84 (m, 2H, CH_2 propargyl, H10 DHA), 4.69 (d, J = 12.6 Hz, 1H, CH_2 propargyl), 4.59 (s, 1H, H3 CDCOMe), 3.87 (d, J = 3.0 Hz, 1H, H7 CDCOMe), 3.74 (s, 3H, OMe) 2.99 – 2.86 (m, 1H), 2.67 – 2.55 (m, 1H, H9 DHA), 240-1.02 (m, 34H, aliphatic protons), 1.41 (s, 3H, CH_3), 0.90 (dd, J = 6.3, 2.1 Hz, 6H, CH_3 DHA, CDCOMe), 0.86 (s, 3H, CH_3), 0.84 (d, J = 7.3 Hz, 3H, CH_3), 0.64 (s, 3H, CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.85, 144.62, 104.24, 101.68, 88.13, 81.28, 68.66, 61.91, 56.75, 55.97, 52.67, 51.64, 50.59, 44.54, 42.87, 39.74, 39.47, 37.54, 37.00, 36.57, 35.51, 35.41, 34.73, 34.11, 32.97, 32.86, 31.16, 31.11, 30.98, 30.80, 28.28, 26.33, 24.99, 24.82, 24.58, 23.86, 23.28, 20.93, 20.47, 18.41, 13.14, 11.91.

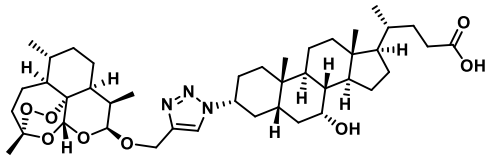
3α-triazolyl-DHA,7α-hydroxy 24-methylcholate (DHA-t-CDCOMe)



In a 10 ml round bottom flask were weighted 510 mg (1.182 mmol, 1.1 eq) of N_3-CDCOMe and 481 mg (1.333 mmol) of DHA-Alk. The compounds were dissolved in 5 ml of CH_2Cl_2 and to this solution were added 5 ml of water and 2 ml of CH_3CN. After nitrogen atmosphere was created inside the flask, 1.33 ml of a 0.5 M (0.5 eq.) solution of sodium ascorbate and 1.33 ml of a 0.1 M (0.1 eq) $CuSO_4$ solution were added to the mixture. After four hours of emulsification at room temperature the reaction mixture was poured into 50 ml of water and extracted with 50 ml CH_2Cl_2. The organic solvent was dried with anhydrous sodium sulphate and the solvent removed under vacuum to give 752 mg of crude product. The crude was then purified by flash chromatography (EtOAc/Cy 2:3 to 1:1) to give 225 mg of the product as a white powder in 25% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.57 – 7.52 (s, 1H, H triazole), 5.44 (s, 1H, H12 DHA), 4.90 (dd, J = 8.0, 4.4 Hz, 2H, CH_2 propargyl, H10 DHA), 4.64 (d, J = 12.4 Hz, 1H, CH_2 propargyl), 4.41 – 4.31 (m, 1H, H3 CDCOMe), 3.88 (s, 1H, H7 CDCOMe), 3.66 (s, 3H, OMe), 2.74 (q, J = 13.0 Hz, 1H), 2.68 – 2.57 (m, 1H, H9 DHA), 2.46 – 1.05 (m, 30H, aliphatic protons), 1.45 (s, 1H, CH_3) 0.99 (s, 3H, CH_3), 0.93 (dd, J = 6.4, 1.5 Hz, 6H, CH_3 21 CDCOMe, DHA), 0.87 (d, J = 7.3 Hz, 3H, CH_3), 0.68 (s, 3H, CH_3 18 CDCOMe). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.86, 104.25, 101.45, 88.13, 81.33, 68.49, 61.59, 61.18, 55.95, 52.71, 51.65, 50.52, 44.57, 42.86, 42.28, 39.61, 39.55, 37.48, 37.02, 36.59, 35.96, 35.50, 35.40, 34.73, 34.44, 33.07, 31.15, 31.11, 30.98, 28.44, 28.27, 26.34, 24.84, 24.59, 23.83, 23.02, 20.78, 20.49, 18.42, 13.14, 11.94.

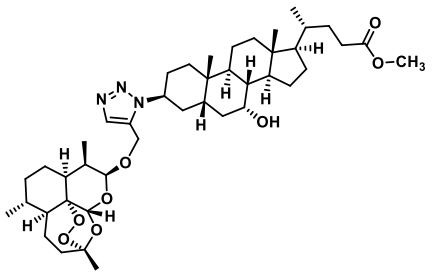
3α-triazolyl-DHA,7α-hydroxy 24-cholic acid (DHA-t-CDCA)



217 mg of DHA-t-CDCOMe (0.2883 mmol, 1eq) were dissolved 7 ml MeOH and 2.30 ml 1,5M LiOH (3.460 mmol, 12eq) solution were added. The reaction mixture was stirred at room temperature for 24 hours. 15 ml water were added, and the pH was brought to 7 by addition of 5% HCl solution and methanol was removed under vacuum. The aqueous solution pH was adjusted to 3 with 5% HCl solution and extracted with 50 ml EtOAc. The organic solvent was dried with anhydrous sodium sulphate, filtered on cotton, and removed under vacuum to afford 195 mg of product as a white powder in 89% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.55 (s, 1H, triazole), 5.43 (s, 1H, H12 DHA), 4.94 – 4.84 (m, 2H, CH_2 propargyl H10 DHA), 4.63 (d, J = 12.4 Hz, 1H, CH_2 propargyl), 4.44 – 4.27 (m, 1H, H3 CDCA), 3.88 (m, 1H, H7 CDCA), 2.74 (q, J = 13.0 Hz, 1H), 2.66 – 2.56 (m, 1H, H9 DHA), 1.46 (m, 0H), 1.44 (s, 3H, CH_3 DHA), 0.99 (s, 3H, CH_3), 0.94 (dd, J = 6.4, 3.7 Hz, 6H, CH_3 21 CDCA, CH_3 DHA), 0.86 (d, J = 7.3 Hz, 3H), 0.68 (s, 3H, CH_3 18 CDCA). ^{13}C NMR (101 MHz, $CDCl_3$) δ 179.11, 144.20, 120.58, 104.25, 101.46, 88.12, 81.31, 68.48, 61.53, 61.22, 55.96, 52.68, 50.47, 44.55, 42.85, 42.25, 39.59, 39.52, 37.45, 36.98, 36.57, 35.95, 35.49, 35.38, 34.72, 34.42, 33.03, 31.05, 30.96, 30.93, 28.42, 28.26, 26.31, 24.82, 24.57, 23.79, 23.00, 20.77, 20.48, 18.40, 13.12, 11.93.

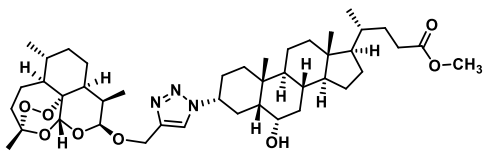
3α-(1,5)triazolyl-DHA,7α-hydroxy 24-methylcholate (DHA-t(1,5)-CDCOMe)



In a two necked round bottom flask were dissolved 392 mg of DHA-alk (1.216 mmol, 1.05 eq) in 8 ml anhydrous tetrahydrofuran (THF), under nitrogen atmosphere. 46 mg of $Cp^*RuCl(PPh_3)_2$ (0.058 mmol, 0.05eq) in 8 ml anhydrous THF were added and the mixture was strirred for 5 minutes at room temperature. 500 mg (1.158 mmol, 1eq) of N_3-CDCOMe in 8 ml dry THF were added, and the reaction mixture was refluxed for 18 hours. at reaction complete solvent is removed under vacuum and the crude is purified by silica flash chromatography (1:1 to 2:3 Cy/EtOAc) to give 175 mg of product as a white foam in 25% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.62 (s, 1H, triazole), 5.36 (s, 1H, H12 DHA), 4.97 – 4.85 (m, 2H), 4.53 (d, J = 12.7 Hz, 1H), 4.22 – 4.10 (m, 1H), 3.84 (s, 1H), 3.67 (d, J = 1.0 Hz, 3H), 3.01 – 2.87 (m, 1H), 2.69 (dt, J = 7.9, 4.2 Hz, 1H), 2.45 – 2.17 (m, 4H), 2.14 – 1.02 (m, 21H), 0.99 (s, 3H), 0.94 (d, J = 5.9 Hz, 8H), 0.88 (d, J = 7.4 Hz, 4H), 0.67 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.88, 133.65, 132.48, 104.46, 101.82, 88.23, 81.06, 68.38, 59.76, 58.30, 55.88, 52.55, 51.64, 50.34, 44.26, 42.89, 42.75, 39.72, 39.52, 37.60, 37.38, 36.47 (2C), 35.51, 35.45, 34.58, 34.22, 32.86, 31.16, 31.09, 30.77, 28.28, 27.58, 26.26, 24.80, 24.59, 23.84, 23.10, 20.77, 20.37, 18.39, 12.98, 11.96.

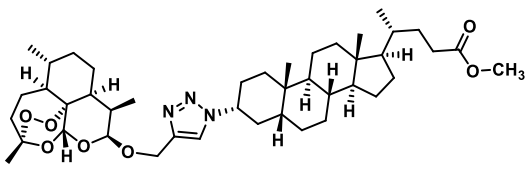
3α-triazolyl-DHA,6α-hydroxy 24-methylcholate (DHA-t-HCOMe)



In a 10 ml round bottom flask were weighted 200 mg (0.463 mmol, 1 eq) of N_3-HDCOMe and 150 mg (0.463 mmol) of DHA-Alk. The compounds were dissolved in 5 ml of CH_2Cl_2 and to this solution were added 5 ml of water and 2 ml of CH_3CN. After nitrogen atmosphere was created inside the flask, 463 µl of a 0.5 M (0.5 eq.) solution of sodium ascorbate and 463 µl of a 0.1 M (0.1 eq) $CuSO_4$ solution were added to the mixture. After four hours of emulsification at room temperature the reaction mixture was poured into 50 ml of water and extracted with 50 ml CH_2Cl_2. The organic solvent was dried with anhydrous sodium sulphate and the solvent removed under vacuum to give crude product. This was then purified by flash chromatography (EtOAc/Cy 3:2) to give 133 mg of the product as a white powder in 39% yield.

^{1}H NMR (400 MHz, $CDCl_3$) selected data δ 7.55 (s, 1H, H triazole), 5.39 (s, 1H, H12 DHA), 4.94 – 4.86 (m, 3H, H10 DHA, CH_2 propargyl), 4.68 (d, J = 12.1 Hz, 1H), 4.40 (s, 1H, H3), 4.14 (t, J = 6.1 Hz, 1H, H6), 3.66 (s, 3H, H3 HDC-OMe), 2.63 (s, 1H, H9 DHA), 0.89 – 0.84 (m, 6H), 0.66 (s, 3H, $CH_3$18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.85, 104.27, 101.66, 88.14, 81.31, 67.93, 61.73, 60.91, 56.10, 56.05, 52.68, 51.65, 49.21, 44.56, 42.99, 40.11, 39.89, 37.52, 36.58, 36.26, 36.20, 35.47, 35.17, 34.94, 34.72, 31.18, 31.09, 31.00, 28.23, 27.21, 26.35, 24.85, 24.59, 24.30, 23.77, 20.96, 20.48, 18.40, 13.14, 12.18.

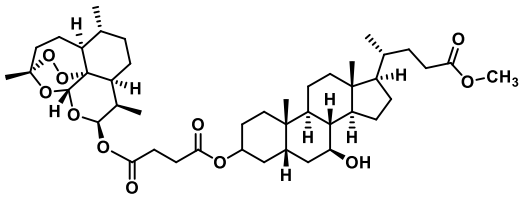
3α-triazolyl-DHA,6α-hydroxy 24-methylcholate (DHA-t-LCOMe)



200 mg of N_3-LCOMe (0.481 mmol, 1 eq) and 155 mg of DHA-alk (0.481 mmol, 1 eq) were dissolved in 6 ml water/DCM/ACN (1:1:1). To this mixture were added 962 µl of 0.1 M $CuSO_4$ (0.0962 mmol, 0.2 eq) solution and 962 µl 0.5 M sodium ascorbate solution (0.481 mmol, 1 eq). The reaction mixture was stirred vigorously for 18 hours and monitored by TLC. The mixture was diluted with 200 ml DCM and 20 ml water and poured in a separatory funnel; the aqueous phase was discarded, and the organic solvent removed under vacuum. The reaction crude (381 mg) was purified by flash column chromatography (cy/EtOAc 3:1) to yield 97 mg of white powder as a product in 27.3% yield.

^{1}H NMR (300 MHz, $CDCl_3$) selected data δ: 7.55 (s, 1H, H triazole), 5.37 (s, 1H, H12 DHA), 4.93 – 4.63 (m, 3H, H10 DHA, CH_2 propargyl), 4.45 (ddt, J = 11.8, 8.6, 4.3 Hz, 1H, H3 LCOMe), 3.66 (s, 3H, OMe), 2.71 – 2.56 (m, 1H, H9 DHA), 0.66 (s, 3H, CH_3 18). ^{13}C NMR (101 MHz, $CDCl_3$) δ 174.89, 144.44, 120.76, 104.25, 101.70, 88.12, 81.29, 61.73, 61.13, 56.50, 56.12, 52.66, 51.63, 44.55, 42.89, 40.79, 40.14, 37.54, 36.58, 35.97 (2C), 35.50, 34.94, 34.73, 34.13, 31.20, 31.14, 31.00, 28.31 (2C), 27.15, 26.46, 26.34, 24.86, 24.57, 24.31, 23.63, 21.04, 20.47, 18.42, 13.13, 12.20.

DHA-s-UDCOMe and (DHA-s)$_2$-UDCMe



Artesunate was obtained from DHA and purified as described by Presser et al.[260] In a round bottom reaction flask were added 0.10 g (0.246 mmol) of UDCOMe, 0.19 g of artesunate (0.492 mmol), 0.02 g (0.172 mmol) of DMAP and 0.09 g (0.492 mmol) of EDCI. Under nitrogen atmosphere were added 10

ml of dry CH2Cl2 and the reaction mixture was stirred at room temperature. After 18 hours the reaction mixture was diluted with 60 ml of CH2Cl2 and washed three times with 20 ml of saturated solution of NH4Cl each, and 20 ml of brine. The organic phase was dried with anhydrous sodium sulphate and the solvent evaporated under vacuum to obtain a crude mixture. The white solid was purified by silica gel column chromatography (Cy/EtOAc 7:4) to obtain 0.09 g of the product in 48% yield and 0.07 g of (DHA-s)$_2$-UDCMe as a side product in 25% yield. The product was further purified by RP-HPLC on a C18 column.

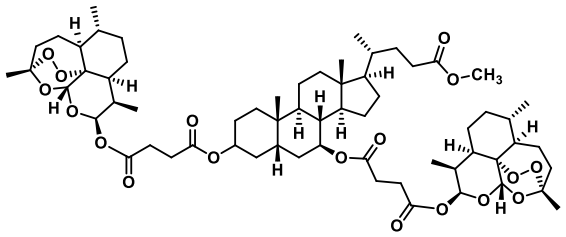
^{1}H NMR (300 MHz, $CDCl_3$) selected data: δ 5.80 (d, J = 9.9 Hz, 1H), 5.44 (s, 1H), 4.79 – 4.56 (m, 1H), 3.66 (s, 3H), 3.64 – 3.49 (m, 1H), 2.81 – 2.66 (m, 2H), 2.66 – 2.47 (m, 3H), 2.44 – 2.30 (m, 2H), 2.29 – 2.14 (m, 1H), 1.43 (s, 3H), 0.94 (d,d,s, J = 10.7, 6.9 Hz, 3 CH3), 0.85 (d, J = 7.1 Hz, CH3), 0.67 (s, CH3). ^{13}C NMR (75 MHz, CDCl3) δ 175.06, 171.89, 171.49, 104.80, 92.47, 91.85, 80.46, 74.70, 74.53, 71.58, 56.01, 55.26, 51.89, 45.57, 44.06, 42.60, 40.41, 39.50, 37.61, 36.92, 36.56, 35.59, 34.94, 34.43, 33.39, 32.17, 31.39, 29.62, 29.57, 28.93, 27.19, 26.76, 26.30, 24.92, 23.67, 22.35, 21.54, 20.55, 18.74, 12.47.

HPLC conditions: Xterra C18 column, CH_3CN/H_2O 10:90 to 90:10 in 15 minutes, tr= 15.9 min

MS (ESI, ES+) $C_{44}H_{68}O_{11}$: 795 $[M+Na]^+$

HRMS ESI (+) $C_{44}H_{68}O_{11}$: 773.4829 $[M+H]^+$, 790.5095 $[M+NH_4]^+$, 795.4646 $[M+Na]^+$

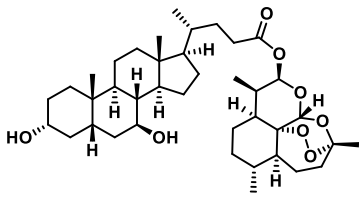
(DHA-s)$_2$-UDCOMe



^{1}H NMR (300 MHz, $CDCl_3$) δ 5.80 (2d, J = 4.8 Hz, 2H), 5.43 (2s, 2H), 4.93 – 4.48 (2m, 2H), 3.66 (s, 2H), 2.87 – 1.47 (m, 65H), 1.42 (s, 6H), 1.38 – 0.98 (m, 8H), 0.98 – 0.94 (m, 6H), 0.91 (d, J = 6.1 Hz, 9H), 0.86 (d, J = 7.1 Hz, 3H), 0.67 (s, 3H). ^{13}C NMR (101 MHz, DMSO-d6): δ= 174.15, 171.72, 171.49, 171.31, 171.17, 103.99, 92.19, 91.01, 80.28, 73.87, 73.76, 54.76, 51.66, 51.55, 44.96, 43.51, 41.62, 39.05, 36.35, 35.04, 34.27, 34.15, 34.00, 32.96, 32.83, 32.08, 31.03, 30.83, 30.69, 29.36, 29.13, 28.91, 28.46, 25.95, 24.78, 24.64, 23.27, 21.43, 21.18, 20.51, 18.66, 12.43, 12.24, 12.21.

HRMS ESI (+) $C_{63}H_{94}O_{18}$: 1156.6773 $[M+NH_4]^+$, 1161.6323 $[M+Na]^+$

DHA-UDCOMe



A solution of 118 mg DHA (0.18 mmol, 1 eq) and 116 mg N_3-UDCOMe (0.27 mmol, 1.5 eq) in 1 ml anhydrous DMF was cooled to 0°C with an ice bath, under an argon atmosphere. Then 32.94 mg dimethyl amino pyridine (DMAP) (0.27 mmol 1.5 eq) and 48 mg EDCI (0.31 mmol, 1.72 eq) were added. After stirring for 10 minutes at 0 °C, the reaction was warmed to room temperature and stirred for 18

hours, then diluted with 15 ml water and extracted with 3X 5 ml Et_2O. The organic layers were combined, dried over anhydrous Na_2SO_4 and concentrated under vacuum. The residue was purified by flash chromatography EtOAc/cyclohexane 3:2, containing 1% Et_3N to afford the product as a colourless syrup in 60% yield.

^{1}H NMR (400 MHz, DMSO-d6): δ 5.64 (d, J=9.8 Hz, 1H), 5.54 (s, 1H), 4.43 (d, J=4.6 Hz, 1H; D_2O exch), 3.87 (d, J=6.9 Hz, 1H; D_2O exch), 3.31–3.20 (m, 2H), 2.47–2.37 (m, 2H), 2.33–2.10 (m, 3H), 2.02–0.80 (m, 45H), 0.75 (d, J=7.1 Hz, 3H), 0.60 ppm (s, 3H); ^{13}C NMR δ (101 MHz, DMSO-d6): 172.11, 103.47, 91.39, 90.46, 79.79, 69.63, 69.35, 59.68, 55.74, 54.49,51.04, 44.50, 43.01, 42.91, 42.08, 37.64, 37.18, 35.86, 35.80, 34.74, 34.65, 33.67, 31.48, 30.46, 30.17, 28.07, 26.65, 25.43, 24.11, 23.23,20.91, 20.76, 19.99, 18.18, 14.01, 11.95, 11.76 ppm.

MS (ESI+): m/z calculated for $C_{39}H_{62}O_8^+$ Na^+: 681.43 $[M+Na]^+$; found: 681.33, 1339.40 $[2M+Na]^+$.

elemental analysis calculated (%) for $C_{39}H_{62}O_8$: C 71.09, H 9.48; found: C 70.84, H 9.83.

PART III: Introduction to oligonucleotides

Structure of oligonucleotides

Oligonucleotides (ONs) represent brief sequences of nucleotides, which are the monomers constituting the nucleic acids DNA and RNA. Typically, they encompass several tens of nucleotides, although their length can extend to several hundreds.

Nucleotides comprise three essential components: a nitrogenous base, a pentose sugar, and a phosphate group.[261,262] These nitrogenous bases can be categorized into two families: purines and pyrimidines. Pyrimidines are characterized by a hexatomic ring composed of carbon and nitrogen atoms; within this category, is possible to find cytosine (C), thymine (T), and uracil (U). In contrast, the purines, adenine (A) and guanine (G), consist of a fused pentatomic ring coupled with a pyrimidine ring. Diverse purines and pyrimidines are distinguished by their associated functional groups.

It's important to clarify that thymine is exclusive to DNA, while uracil is specific to RNA. Moreover, in RNA, the sugar component is ribose, while in DNA, it is deoxyribose. The only distinction between these two sugars lies in deoxyribose lacking the hydroxyl group at position 2'.

The connection between the sugar and a nitrogenous base, enabled by a β-N-glycosidic bond, gives rise to a nucleoside. If a phosphate group additionally binds to the carbon atom at the 5' position of the sugar, the resulting molecule is a nucleotide or nucleoside monophosphate. These nucleotides are interconnected by covalent bonds, termed phosphodiester bonds, which exist between the 5'-phosphate of one nucleoside and the 3'-hydroxyl group of the neighbouring nucleotide. This arrangement forms a sugar-phosphate backbone, with the nitrogenous bases extending from it.

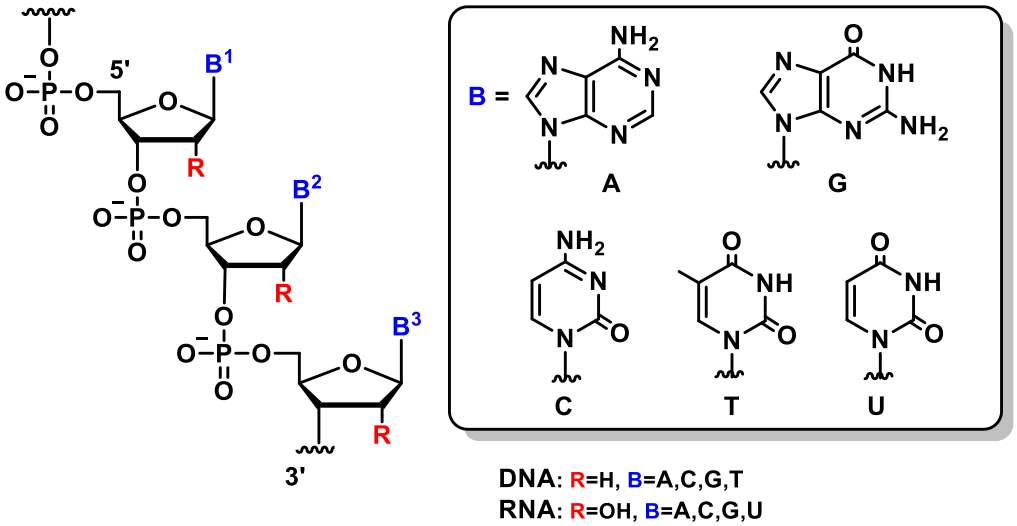


Figure 40. *general structure of oligonucleotides.*

Gene modulation by oligonucleotides

Oligonucleotides, follow the same rules as longer DNA or RNA; crucially, they can specifically bind to complementary sequences thanks to Watson-Crick base pairing rules.[263] The rules state that, naturally, Adenine (A) bind with Thymine (T) while Guanine (G) pairs with Cytidine (C) due to the hydrogen bond interaction between these specific couples.[264]

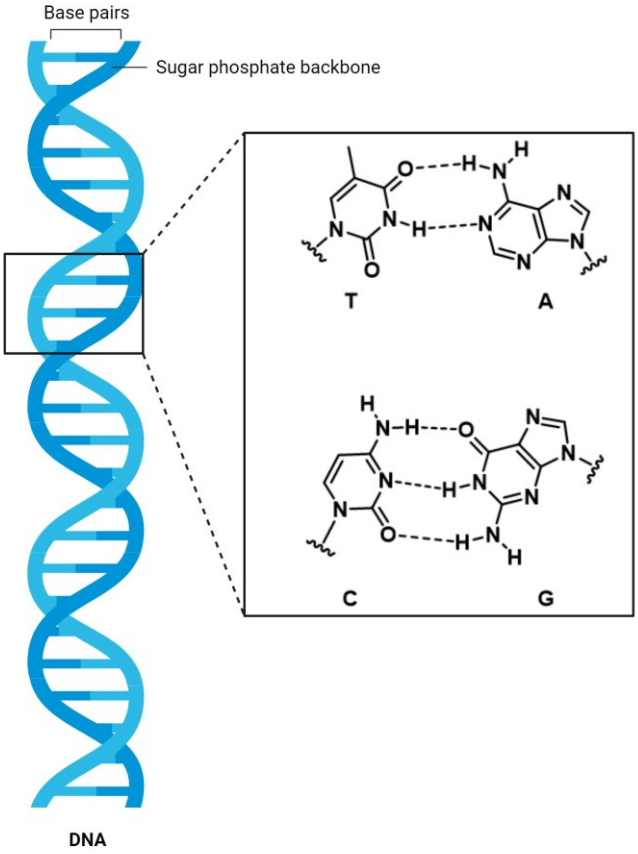


Figure 41. *Base pairing according to the Watson-Crick model.*

The study of nucleic acids and their interactions with proteins forms the foundation of molecular biology, a field dedicated to understanding the molecular mechanisms of life. Central to this discipline is the concept that, starting from DNA, a process of transcription occurs, leading to the synthesis of messenger RNA (mRNA). In eukaryotes, the mRNA undergoes several essential modifications within the nucleus, including the addition of a 5'-end cap, the attachment of a poly-A tail at the 3'-end, and splicing. In the cytoplasm, mature mRNA is used as the template for protein synthesis through the process of translation[261,262] (as illustrated in Figure 42). The relationship between genes and proteins, which forms the basis of this dogma, was initially proposed by Garrod in the early 1900s and later developed and confirmed by Beadle and Tatum in the 1940s.[265,266]

Many of the drugs in use today primarily target the final stages of the process mentioned above or act on specific enzymes that play crucial roles in our bodies. In recent years, numerous studies have aimed to create molecules that can target different components other than proteins. Notably, among these molecules are oligonucleotides, which possess the unique ability to selectively bind to DNA and RNA. They exert their effects through various mechanisms of action collectively referred to as "code blockers".[267–270]

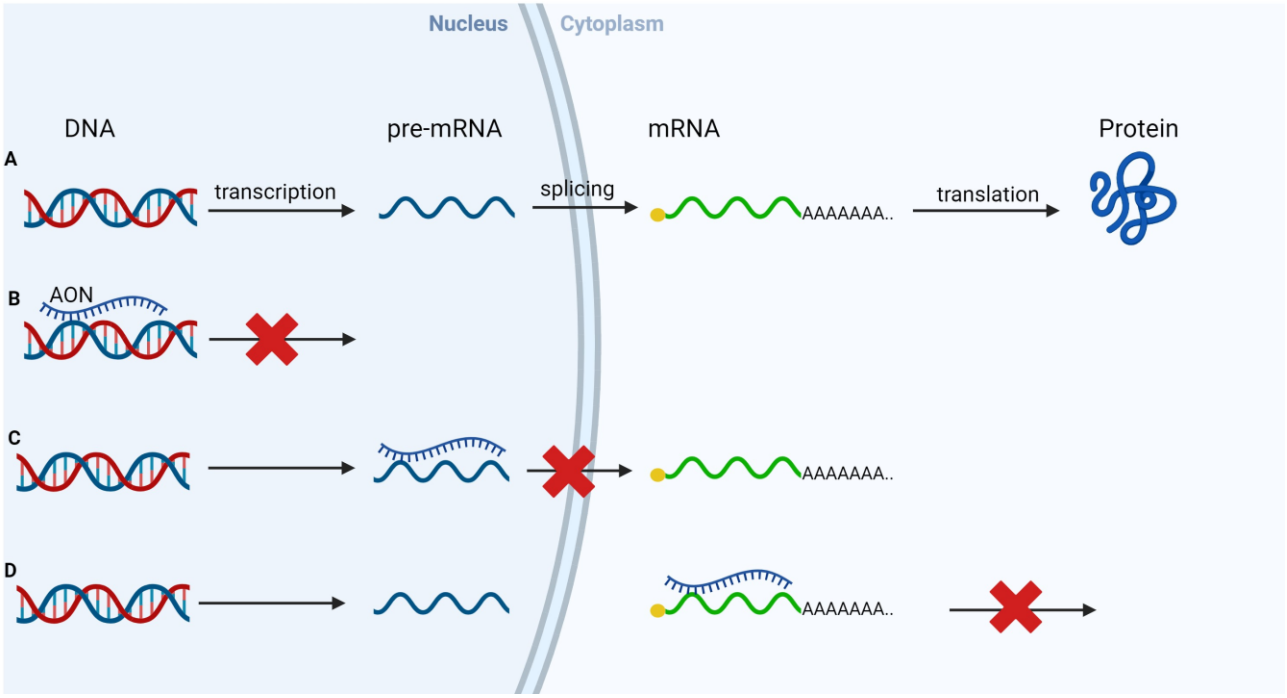


Figure 42. *Normal process of protein synthesis from the DNA through transcription, splicing and translation (A) and mechanisms of action of AON with steric block (B, C and D).*

The first mechanism, developed in 1978, is known as antisense technology.[271–273] This approach involves the synthesis of single-stranded oligonucleotides capable of binding to specific sequences of DNA or RNA, which are complementary to them, and serve as target sequences through Watson-Crick base pairing. These antisense oligonucleotides (abbreviated as AONs or ASOs) possess the remarkable ability to inhibit the expression of the target gene. They achieve this by hybridizing with the "sense" sequences, which would typically undergo transcription or translation in the absence of AONs. The target DNA or RNA sequences may belong to endogenous cells containing mutations that could be detrimental to the organism or may belong to tumour cells.

AONs exert their effects through four primary mechanisms of action: steric hindrance of transcription, splicing or translation processes,[270] activation of RNase H,[274] induction of gene silencing via small interfering RNA (siRNA),[275,276] and regulation of alternative splicing.[277]

Steric block of transcription, splicing or translation

Steric hindrance, resulting from AON binding, can impede several critical processes. Within the nucleus, AONs can hybridize with DNA, blocking transcription, or they can hybridize with pre-mRNA, thereby preventing splicing and the synthesis of mature mRNA (Figure 42 B, C). In the cytoplasm, AONs can bind to the mRNA site where ribosome binding typically occurs. This prevents the ribosome from attaching to the mRNA and carrying out translation, ultimately leading to a deficiency in protein synthesis (Figure 42 D).[278]

Activation of RNase H

Another way to block the formation of a protein is to degrade the mRNA exploiting the natural action of enzymes present in the cell, such as Rnase H.[279] This enzyme is ubiquitous in cells and is present both inside the nucleus and in the cytosol. Its function is to cleave DNA/RNA hybrids which could otherwise jeopardize genome stability by accumulation of R-loops (RNA/DNA-DNA hybrids).[280] To do so, it requires a minimum of four RNA nucleotides to recognise the double strand; and it does so with

micromolar affinity. The affinity for DNA/DNA and RNA/RNA double strands is respectively 25-fold and 100-fold higher, thus giving good selectivity toward mixed hybrids.[281] Normally the cleavage site is at the centre of the RNA sequence. The mechanism underlying the activation of RNase H is therefore distinctly different from the steric block. Indeed, this enzyme can recognize the AON-RNA hybrid, whether it is in the form of mature mRNA or pre-mRNA, and subsequently degrade the RNA molecule. Hydrolases such as RNase H can hydrolyse the phosphodiester bonds within RNA, releasing oligonucleotides and mononucleotides with a 5'-phosphate-3'-OH configuration. Once the hydrolytic process is complete, the AON is liberated and can bind to a new RNA strand, thus acting through a catalytic mechanism (Figure 43). Therefore, the efficacy of this mechanisms is higher than that of steric block and can give a downregulation of the target protein as high as 95%.[282] Over the years enormous efforts have been devoted into developing a gapmer technology, which couples the stability of modified nucleotides, strategically placed toward the ends of the oligonucleotide, with the RNase H activation mechanism given by the DNA nucleotides at the centre of the sequence.[283,284]

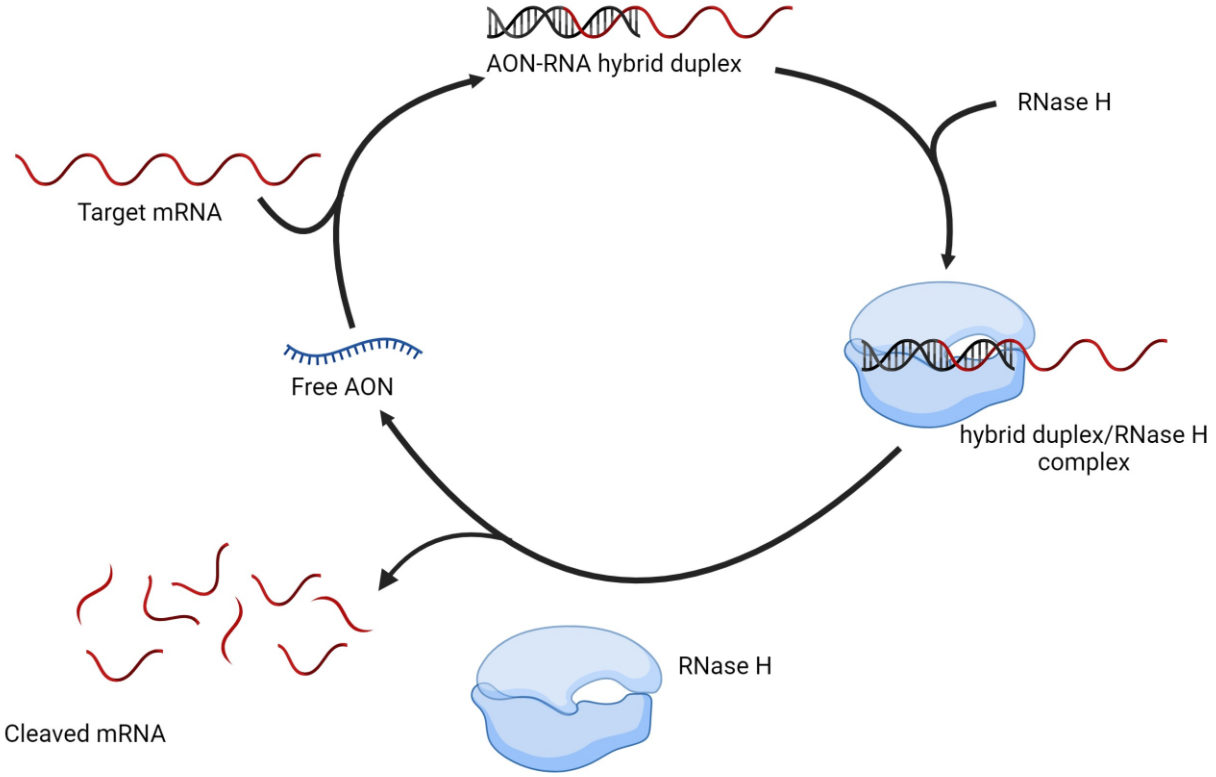


Figure 43. *schematic mechanism of RNase H activation and protein translation downregulation.*

Gene silencing induced by siRNA

Another significant mechanism involves gene silencing induced by siRNA (small interfering RNA), which consists of short double-stranded RNA fragments typically comprising 21-23 nucleotides.[276] These siRNAs have the ability to suppress the expression of specific mRNA targets involved in the pathogenesis of significant diseases through degradation. As depicted in Figure 44, an ATP-dependent helicase identifies these short double helices and separates the siRNA molecule into its two individual RNA strands: one referred to as the "sense" strand and the other as the "antisense" strand. The antisense strand becomes part of the protein complex RISC (RNA-induced silencing complex), serving as a guide for recognizing the complementary mRNA target sequence and subsequently cleaving it. After the mRNA

molecule is dismantled, the RISC complex maintains its enzymatic activity, identifying and degrading other molecules with the same mRNA target in multiple cycles of catalytic action. siRNA is still a developing filed of research and there are some unknown in the mechanism of action. The technology has great therapeutic potential, even against cancer, although is faced with the same delivery problems as all oligonucleotides.[285,286]

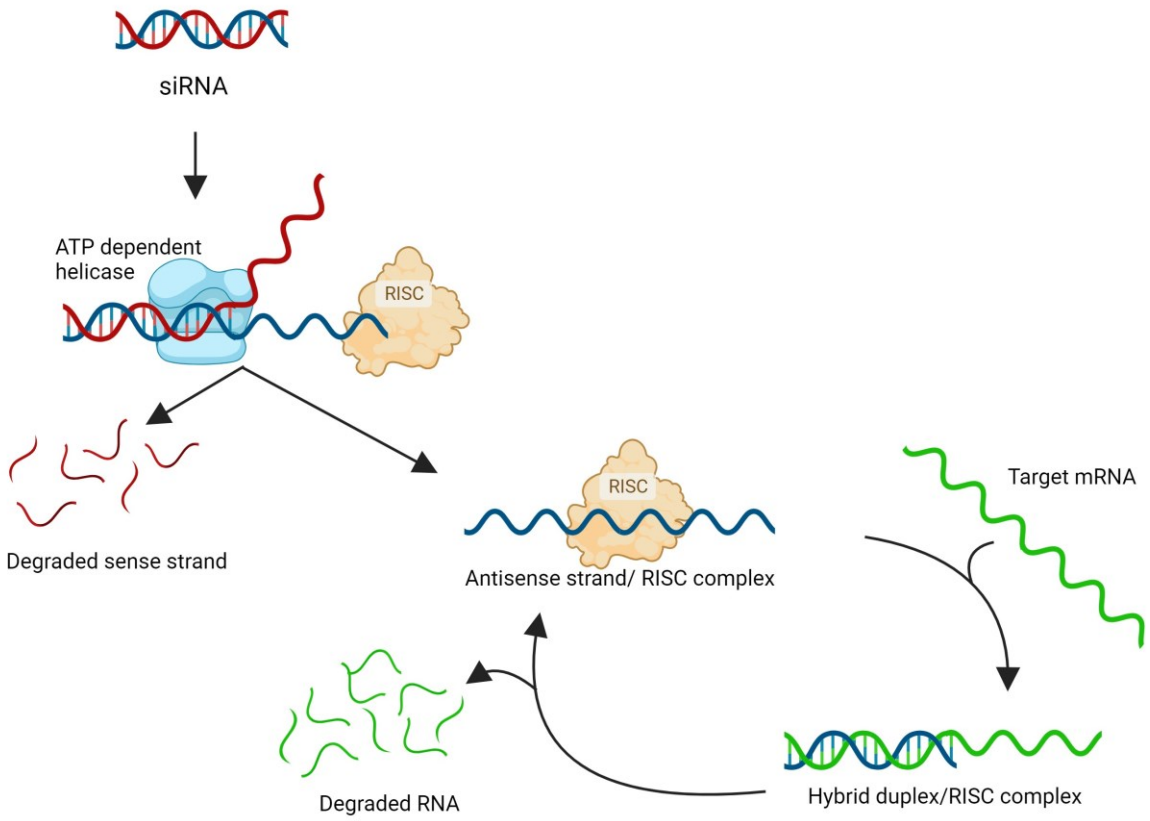


Figure 44. *Schematic representation of siRNA action mechanism.*

Splicing modulation

Splicing is a natural process that can be exploited to modulate gene expression, to understand how this works is important to know the basic concepts of splicing first. The central dogma in biology states that DNA is first transcribed into pre-mRNA, then this matures into mRNA before exiting the nucleus, and goes in the cytosol to provide the code for protein synthesis. Splicing occurs during the maturation of pre-mRNA into mRNA. In fact, the pre-mRNA comprises exons, which are coding regions containing genetic information, interspersed with introns; non-coding sequences that play a regulatory role. Before exiting the nucleus, the pre-mRNA undergoes splicing by enzymes that excise the introns from the molecule and connect the exons, resulting in the formation of a mature mRNA molecule with a continuous coding sequence. Translation, initiated from this mature mRNA, leads to the synthesis of the corresponding protein, as illustrated in Figure 45. Through the splicing process, a single gene can give rise to various types of proteins by selecting which exons ends up in the final mRNA; this process is for example responsible for the synthesis of various protein isoforms in different tissues.

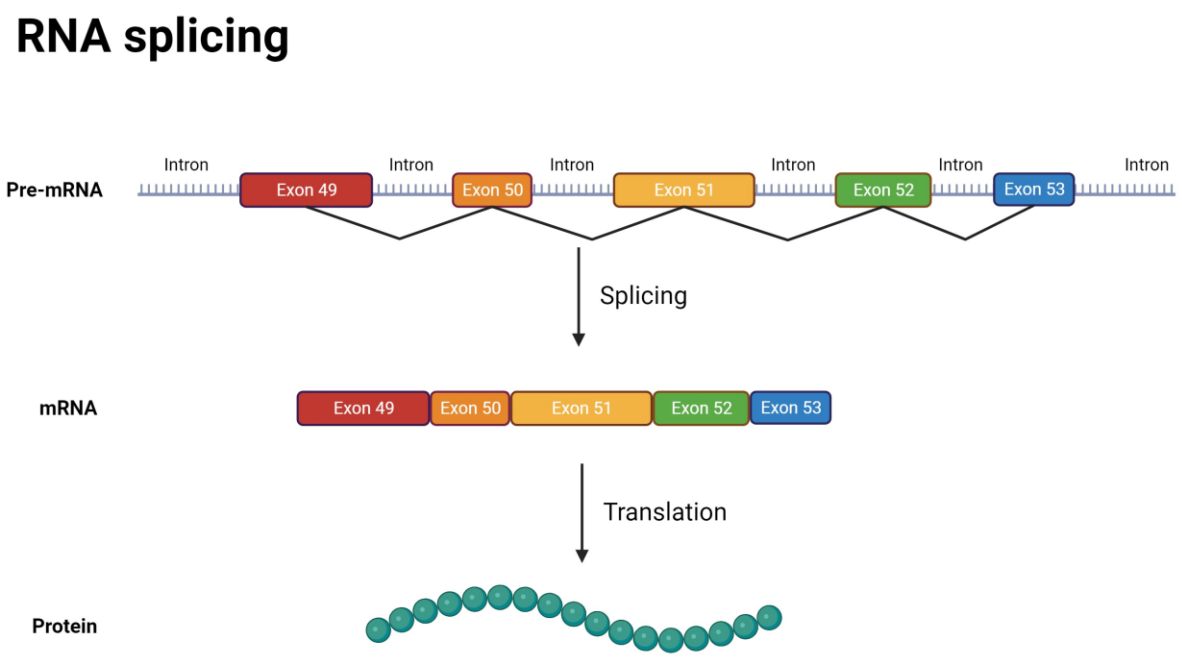


Figure 45. *Schematic representation of the splicing process used to synthetise different isoform of a protein.*

Certain genetic diseases arise from DNA alterations, that result in the production of malfunctioning proteins or a lack of protein production. For instance, in the case of Duchenne muscular dystrophy (DMD), a single point mutation in one of the exons leads to a shift of reading frame, which then brings to the generation of mature mRNA lacking functional significance. This mRNA is not translated, resulting in the absence of protein synthesis, as depicted in Figure 46.[287]

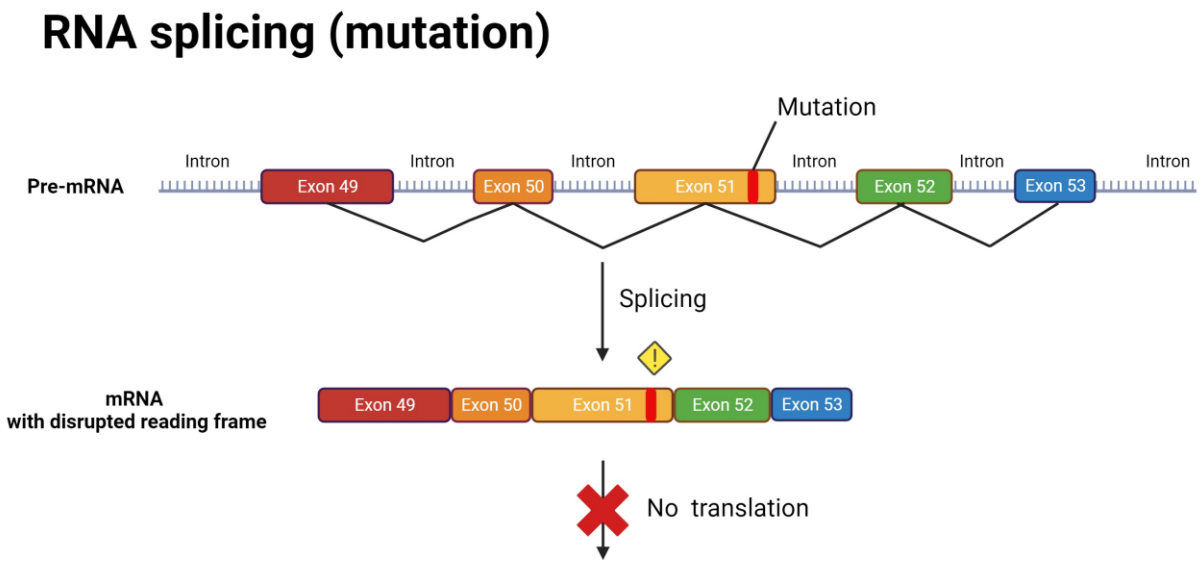


Figure 46. *Schematic representation of the downregulation of protein due to genetic defect influenced by splicing.*

For such pathologies, a potential therapeutic approach involving AONs harnesses a form of alternative splicing known as exon skipping. Exon skipping is a gene therapy strategy designed to mitigate molecular defects by targeting the pre-mRNA responsible for encoding the specific protein. [288–291] In this process, the AON binds to the exon containing the mutation, resulting in its exclusion from the mature mRNA. As a result, the mutation is eliminated from the mRNA, which can then be translated into the respective protein. While the resulting protein may lack a portion of an exon, and differ to some degree from the original protein, it remains functional, as illustrated in Scheme 6.6.

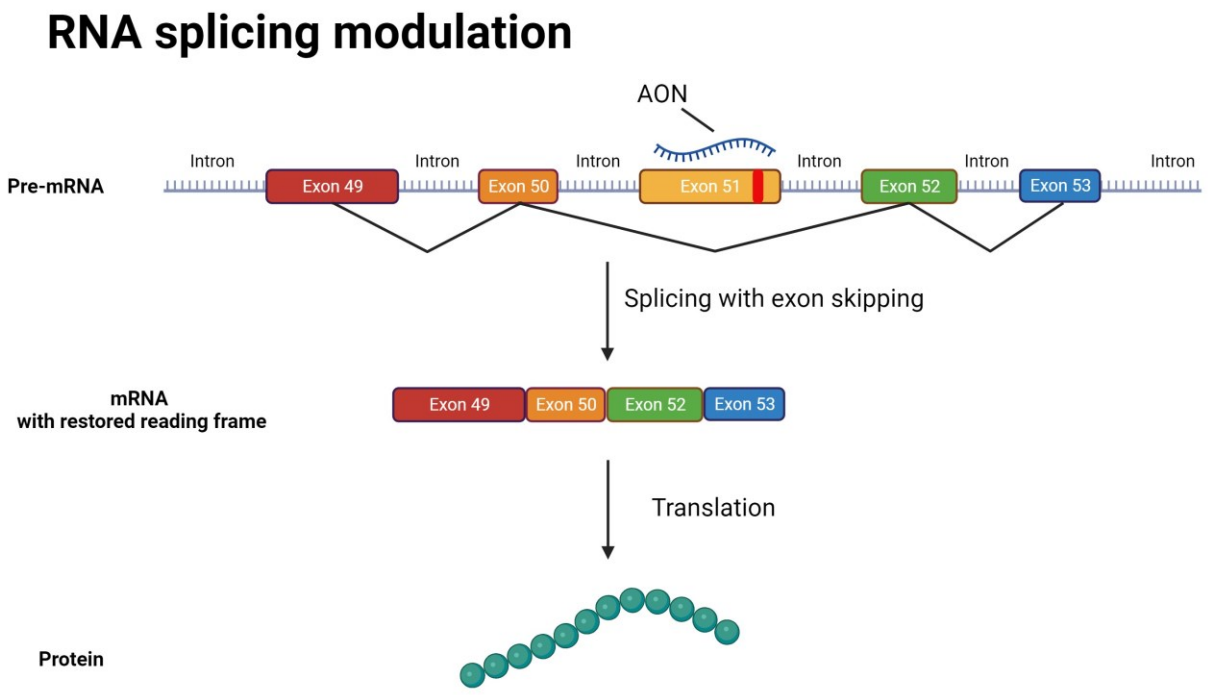


Figure 47. *Schematic representation of AON modulation of exon skipping and restoration of a partly functional protein of the dystrophin gene.*

AONs in gene therapy

For antisense oligonucleotides (AONs) to be applicable in gene therapy, they must undergo rigorous *in vitro* and *in vivo* testing to ensure several critical properties: effective delivery, non-toxicity, stability and long-term efficacy. Delivery encompasses the AONs' capability to reach the precise site where the disease-causing genes are located; and is of today one the major issues holding back oligonucleotide therapeutic agents.[292] Non-toxicity signifies the absence of adverse effects on endogenous cells, especially long term ones. Furthermore, these molecules must demonstrate stability in biological fluids, at least until they have reached their intended target and explicated their function. While DNA-based AONs tend to exhibit greater stability compared to RNA-based AONs, both categories are susceptible to degradation. Consequently, stability remains a significant limitation for unmodified antisense oligonucleotides. Over the years, researchers have developed more stable analogues through structural modifications (See section: Modified oligonucleotides).

In recent times, there has been a surge of interest in these innovative molecules capable of modulating gene expression, holding promise for applications in therapy and diagnostics. Numerous preclinical and clinical studies have explored the potential of these agents for treating conditions ranging from cancer,[293,294] cardiopathies and muscular dystrophies, to infectious diseases, inflammatory disorders (both viral and genetic), and metabolic illnesses. Some AONs have advanced to the most mature clinical phases, as outlined in Table 2.[295] Other AONs have been already approved, such as SPINRAZA®, that has been introduced for the treatment of spinal muscular atrophy (SMA) in paediatric and adult patients and has MOE (*2' methoxyethyl*) modified sugars. For Duchenne muscular dystrophy, four different AONs have been approved by the FDA (Food and Drug Administration): Eleplinsen, Golodirsen, Vitolarsen and

more recntly Casimersen. All these AON have PMO (phosphorodiamidate morpholino) based backbones.[296]

Table 2. AON currently under clinical testing.

Therapeutic area	Drug name	chemistry	target	inidcation	Clinical phase
Neurological	Eplontersen	MOE, DNA, GalNac	TTR	Hereditary ATTR polyneuropathy	registration
	Ulefnersen	MOE, DNA	FUS	Amyotrophic Lateral Sclerosis	3
	Tofersen	MOE, DNA	SOD1	Amyotrophic Lateral Sclerosis	3
	Zilganersen	MOE, DNA	GFAP	Alexander's disease	3
	IONIS-MAPT$_{Rx*}$	MOE, DNA	TAU	Alzheimer's disease	2
	ION859	MOE, DNA	LRRK2	Parkinson's disease	2
Cardiovascular	Eplontersen	MOE, DNA, GalNac	TTR	Transthyretin Amyloid Cardiomyopathy	3
	Olezarsen	MOE, DNA, GalNac	ApoC-III	Severely elevated triglycerides	3
	Fesomersen	MOE, DNA, GalNAc	Factor XI	Thrombotic disorders	2
	Pelacarsen	MOE, DNA, GalNAc	APO-a	Cardiovascular diseases	3
Rare diseases	Donidalorsen	MOE, DNA	PKK	Hereditary angioedema	3
	Sapablursen	unknown	TMPRSS6	Polycythemia Vera	2
Antiviral	Bepirovirsen	MOE, DNA, GalNAc	Hepatitis B Virus	Hepatitis B Virus Infection	3
Kidney disease	ION532	APOL1	APOL1	Chronic Kidney Disease	2
Cancer[293]	Oblimersen	DNA	BCL-2	Lymphoma, colorectal cancer	3
	Apatorsen	DNA	HSP-27	Lung, prostate cancer	2
	Custirsen	DNA	clusterin	Prostate, bladder, breast, kidney, ovarian cancer	3

Table 2. *Examples of AONs currently undergoing clinical trials.*[295] MOE mean 2' methoxyethyl modification of the ribose. GalNAc refers to the conjugation of N-Acetylgalactosamine.

Modified oligonucleotides

The negatively charged phosphate groups present in oligonucleotides confer them a polyanionic character, which in turn gives difficulties in crossing both the cytoplasmic and nuclear membranes to reach their intended site of action.[292] AONs also face instability in the bloodstream, as they are recognized and removed by the endothelial reticulum system, an essential component of the body's defence mechanisms. Subsequently, they undergo degradation by nucleases. In addition, they can combine with plasma proteins, ultimately leading to renal excretion, resulting in a relatively short plasma half-life, measured in just minutes. In some instances, AONs may also exhibit inadequate affinity for their target RNA.

To enhance the stability and cellular uptake of AONs, the literature outlines two primary approaches: the incorporation of structural chemical modifications and/or conjugation with lipids, peptides, polymeric molecules, nanoparticles, and liposomes.[297–299] These strategies offer several potential advantages:

- greater chemical stability
- increased stability towards endo- and exo- nucleases
- improved delivery into appropriate cellular compartments, such as cytoplasm or nucleus
- improved ADME (absorption, distribution, metabolism, and excretion); faster passage from the blood to the tissues where they are widely distributed.

However, chemical modifications of nucleotide structures should preserve their capacity to recognize complementary bases, thereby enabling them to hybridize with the target molecules. Key sites for introducing these modifications include the backbone phosphate group, the ribose sugar, and the nitrogenous bases. Figure 48 provides a schematic representation of the specific sites amenable to modification.

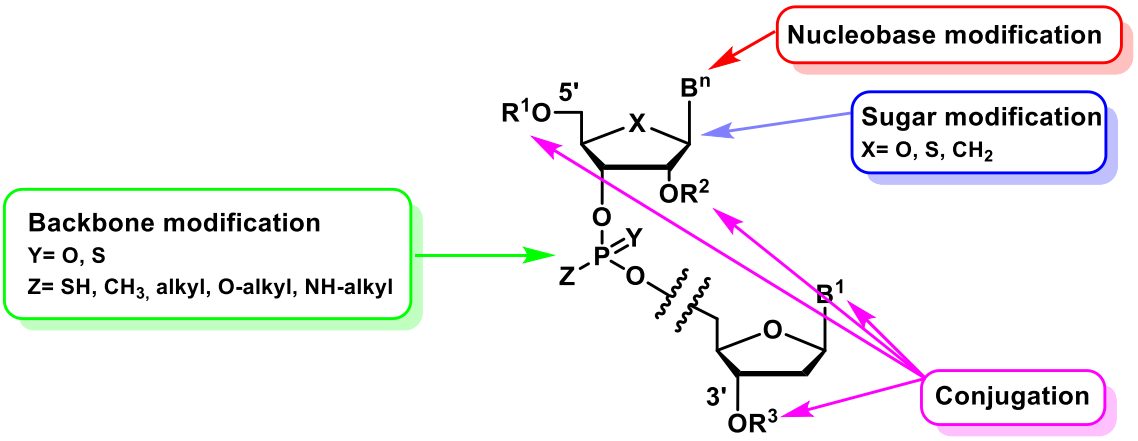


Figure 48. *Possible modification sites present in oligonucleotides.*

Backbone modifications

The first chemical modification of oligonucleotides was the substitution of the phosphate backbone with a methylphosphonate equivalent. This modification erases the negative charge of the phosphate and makes the oligonucleotide more lipophilic and stable. However, this also reduces the solubility and delivery. Moreover, the modification destabilizes the duplex with the target sequence.

Nowadays, the most prevalent modifications of the backbone involves replacing a non-bridging oxygen atom of the phosphate group with a sulphur atom, resulting in the formation of phosphorothioate

oligonucleotides (Figure 49).[300–302] This modification offers several advantages, including increased stability against nucleases,[303] enhanced cellular uptake,[304–307] good hybridization capacity, and the ability to activate RNase H as a mechanism of action.[308] When both the oxygens not bound to nucleoside units are replaced with sulphur, phosphorodithioates are formed (Figure 49). While they are resistant to nucleases, they have limited affinity for the target and, consequently, are not widely used.[309] Boranophosphate oligonucleotides substitute oxygen with a borane group (BH_3^-, Figure 49). They exhibit greater stability against nucleases than DNA and are slightly more stable than phosphorothioates. Additionally, they have been shown to activate Escherichia coli's RNase H. Another type of modification to the backbone involves the complete replacement of the phosphate group with other functional groups. For example, in literature is reported the substitution with an amide group along with the insertion of a methoxy group at position 2'. While this enhances affinity towards complementary RNA to some extent, the amide group exhibits high resistance to nucleases as it differs from their typical target, the phosphodiester group.[308] A more recent substitution of the phosphate group with both 1,4 and 1,5 triazoles has been studied. This modification allows to obtain a very stable linkage between two oligonucleotides. It is therefore useful for creating longer oligonucleotides while conserving the oligonucleotide biocompatibility with enzymes such as DNA polymerase.[310]

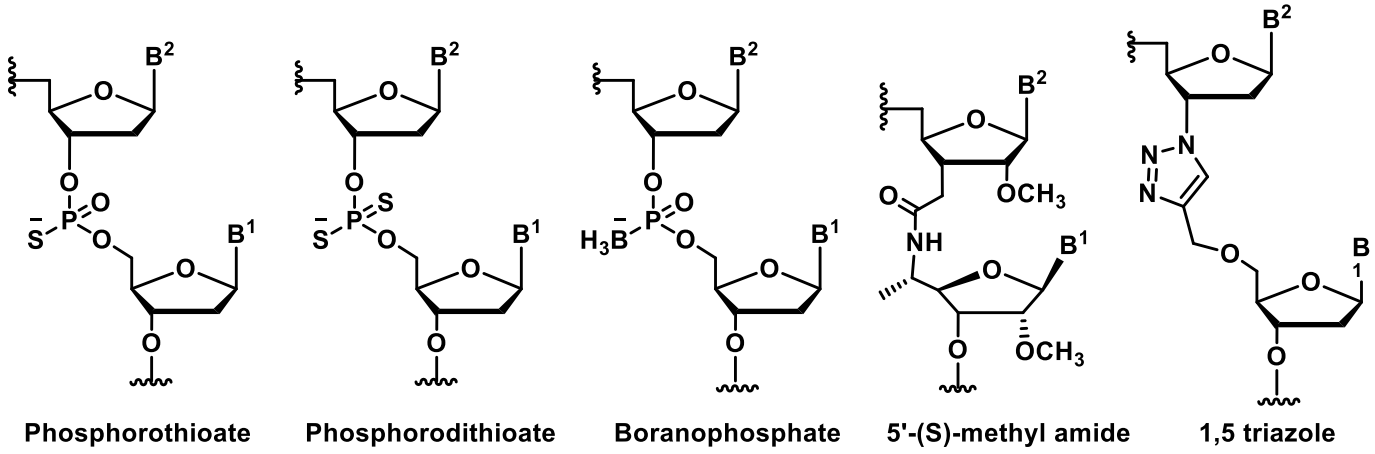


Figure 49. *Examples of backbone modifications.*

Sugar modifications

The 2' position in the pentose sugar can undergo numerous modifications to enhance the pharmacological properties of oligonucleotides. Modifications at this position can offer two main advantages: increased affinity and reduced susceptibility to nucleases, owing to the proximity of the substituent at 2' to the phosphate group at position 3'. However, such modifications often significantly diminish or completely inhibit the ability to activate RNase H, thereby disrupting this antisense mechanism.[311]

The most common modifications at the 2' position include replacing the hydroxyl group with a fluorine atom or introducing 2'-O-alkyl groups such as O-methyl and O-methoxyethyl (MOE, Figure 50). The high electronegativity of the fluorine atom results in increased affinity for RNA, which is instead relatively lower in the presence of 2'-O-alkyl groups. The 2'-O-methoxyethyl modification is presently the most advanced among these types of alterations and has undergone numerous clinical trials as seen above. AONs with this modification exhibit further increased resistance to nucleases and a potential reduction in toxicity.[312]

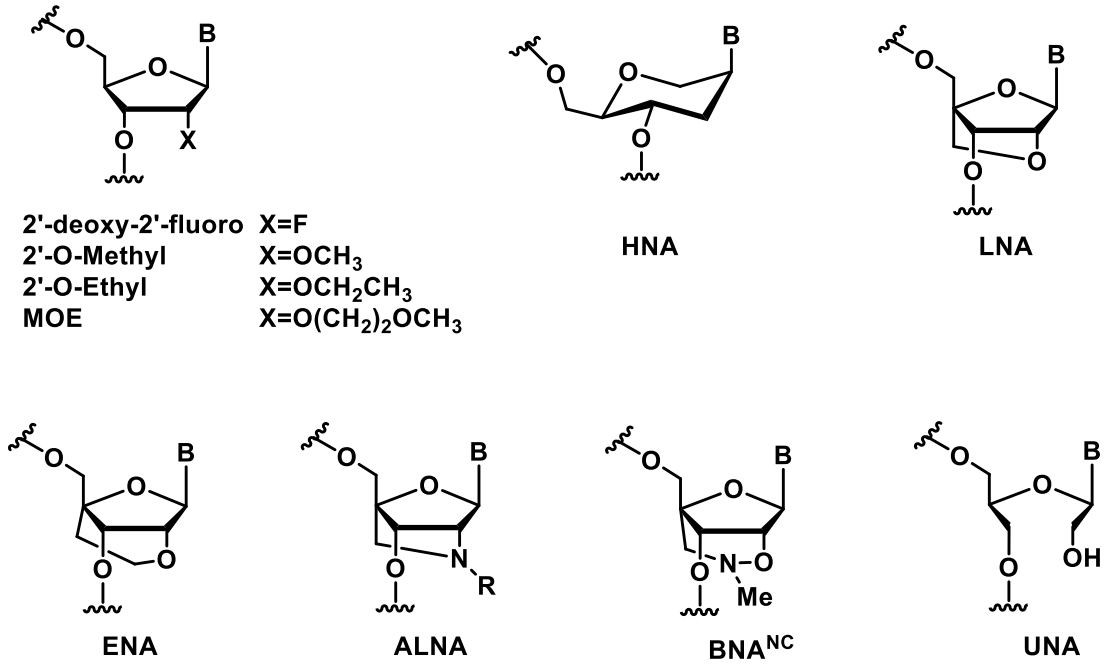
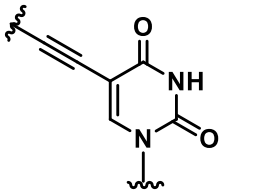


Figure 50. *Examples of sugar modifications.*

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, is of interest the hexitol nucleic acid (HNA), which is able to hybridize in a stable way with the RNA and which has a high resistance to nucleases.[313] Even more interesting is the locked nucleic acid modification of the ribose sugar, in this case a bicyclo ribose is formed by the conjugation of the 2' and 4' position of the sugar, which can happen through a oximethylen (LNA or α-LNA), oxyethylen (ENA),[314] aminomethylen group (ALNA) [313,315] or 2'-O,4'-aminoethylene bridged nucleic acid (BNA^{NC})[316]. Moreover, is also possible to modify the ribose and remove the bond between position 2' and 3' to get unlocked nucleic acids (UNA).[317] In the case of oxymethylen locked nucleic acids increase in duplex stability and nuclease resistance has been reported.[318,319] Moreover, their use as modified nucleotides for gapmer technology is gaining increased attention. More on this modification will be explained in this thesis. Other modifications are possible are there is a great plethora of literature on this reagard.[320,321]

Nucleobase modifications

There is a great deal of literature on nucleobase modification, both on conjugation on natural nucleobases, on synthetic nucleobases and of direct substitution of a nucleobase with another molecular entity. One of the prominent modifications made to the nitrogenous bases involves insertion of a linker moiety, specifically a propynyl group, at position 5 in pyrimidine-based bases, such as uracil (Figure 51). This group, rich in π electrons, enhances stacking interactions between the bases, thereby increasing the stability of the duplex. Stacking interactions are aromatic interactions between adjacent and parallel nitrogenous bases, which serve to promote the stability of the double helix. Importantly, this modification preserves the Watson and Crick-type base recognition.[322] Even more importantly, it allows further modification through click chemistry. Similar linkers exist with protected amines useful for conjugation through amide bond.



5-Propynyl uracil
105

Figure 51. *example of nucleobase modification.*

Conjugation with neutral lipids

To enhance cellular uptake, a common approach involves the use of cationic lipids or polymers that interact electrostatically with oligonucleotides, forming nanoplexes with sizes ranging from 10 to a few hundred nanometers. These formulations offer advantages such as increased stability in biological fluids, improved cellular uptake, and enhanced biodistribution. However, a major drawback of this strategy is the toxicity associated with the cationic nature of lipids or polymers, making them less suitable for biomedical applications.[323] Given these considerations, it is intriguing to explore the use of non-ionic lipids, which, being of natural origin, are non-toxic and biocompatible. Unfortunately, the association of AONs with nanoparticles composed of neutral lipids through physical interactions, such as adsorption, appears to be less efficient compared to cationic lipids. To address this issue, one potential approach is to covalently conjugate AONs with neutral lipids.[324] This covalent bonding can be accomplished at either the 3' or 5' positions of the oligonucleotide, each with its own set of advantages and disadvantages:

- Conjugation at the 3' position increases stability against nucleases,[325,326] but it involves greater technical complexity in the synthesis.[327]
- Conjugation at the 5' position offers ease of technical execution and the ability to rapidly construct libraries of oligonucleotides by changing only the conjugate at the end of the synthesis. However, it has a less pronounced impact on nuclease stability and shows variability in uptake.

It's also possible to incorporate a neutral molecule within the oligonucleotide chain. To do this, units that compose the AON should already be conjugated with the neutral molecule or possess a suitable linker to which the lipid can be attached post-synthetically.

Another consideration is whether to carry out the coupling in solid phase or in solution. Conducting coupling in solution (Figure 52) involves removing the synthesized oligonucleotide from the support, purifying it if necessary, combining it with the neutral lipid, and purifying it again. This technique can be more labour-intensive than solid-phase conjugation due to the complexity of the purification process. In solid-phase coupling, purification is simpler, as it is achieved through the washing of reagents left in the reaction environment. Furthermore, in solution-phase coupling, large excesses of the conjugating group must be used, and the chemical reactions employed must be compatible with the nature of the AON.

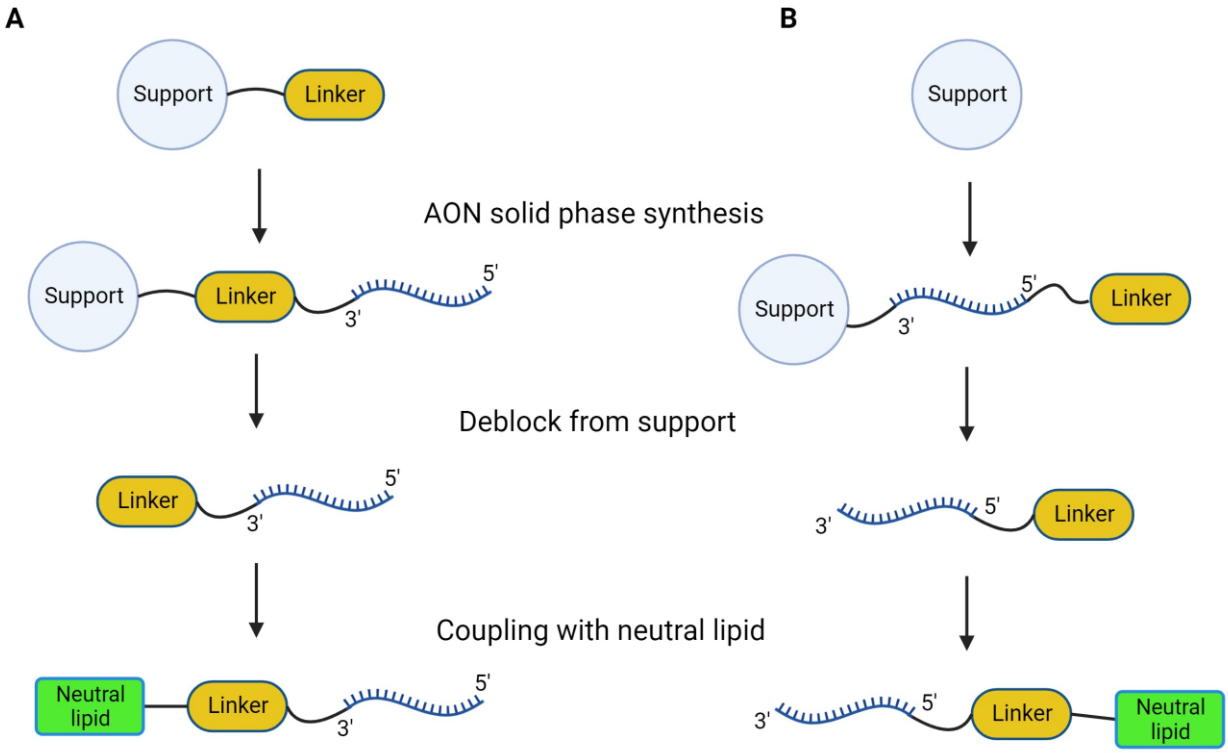


Figure 52. *A) Syntheses of AON-3'-neutral lipid conjugate B) AON-5'-neutral lipid conjugates with coupling in solution.*

Using a solid support can help overcome challenges related to the poor solubility of one of the reagents. The covalent bonds most commonly employed to link a lipid molecule to the oligonucleotide include thioether, disulfide, phosphate, and phosphoramidate bonds. To avoid potential steric hindrance issues, a simple alkyl chain is often used as a linker between the oligonucleotide and the lipid molecule.[328]

For solid-phase coupling at the 3' position, a support that is already functionalized with the lipophilic molecule is required. The entire oligonucleotide is synthesized starting from this support, and at the end of the synthesis, the lipophilic-oligonucleotide conjugate is detached from the support through an ammonolysis reaction (Figure 53, A). Conversely, in the case of coupling at the 5' position, the suitably functionalized lipophilic molecule is conjugated to the oligonucleotide at the end of the solid-phase synthetic process (Figure 53 ,B).

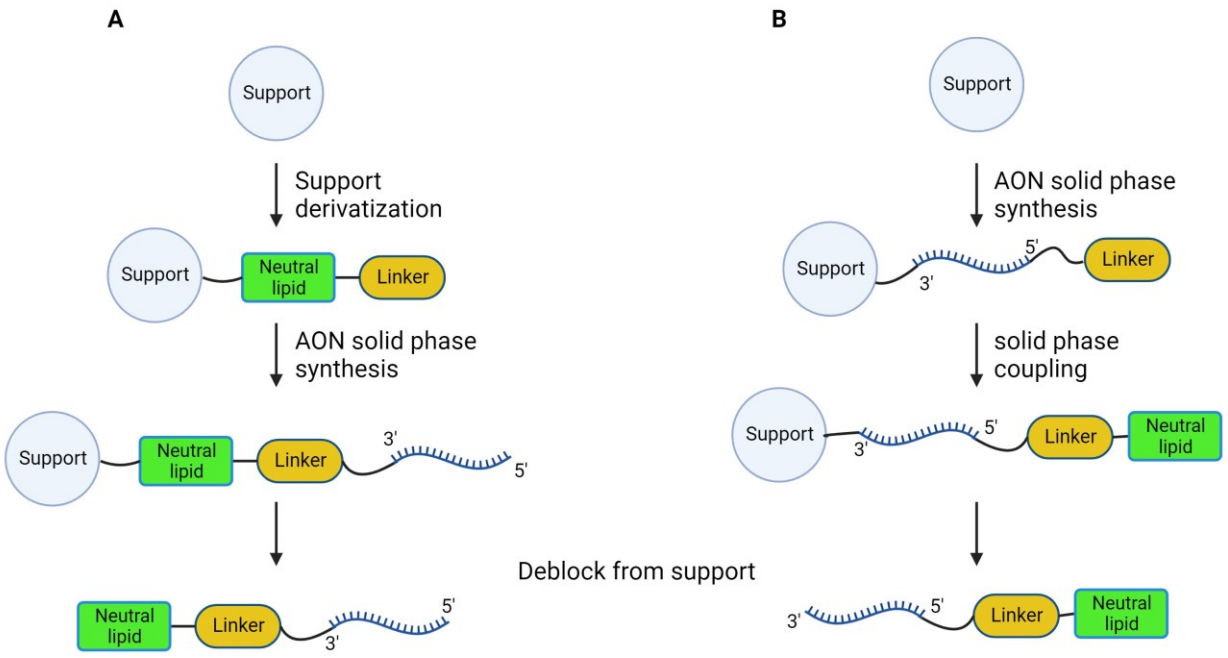


Figure 53. *A) Syntheses of AON-3'-neutral lipid in solid phase. B) AON-5'-neutral lipid) conjugates with solid phase coupling*

Solid phase synthesis of oligonucleotides

Solid-phase synthesis is a technique utilized for constructing biomacromolecules from monomers. The initial monomer in the sequence is attached to a solid support, and a solution containing the necessary reagents is passed over it. Through various steps, the intended compound is synthesized.

This method was pioneered by Robert Bruce Merrifield for peptide synthesis,[329] earning him the Nobel Prize in Chemistry in 1984. Nowadays, it finds applications in various fields, including the synthesis of oligonucleotides and combinatorial chemistry. While oligonucleotides can also be synthesized in solution, solid-phase synthesis offers several advantages. It simplifies the purification process after each step, and it can be automated using synthesizers. Furthermore, it requires smaller quantities of reagents compared to liquid-phase synthesis yet allows the use of excess reagents when needed to drive reactions to completion, which can be easily removed.[330]

Solid support

The solid support must possess a series of requirements:

- sufficient rigidity and pressure resistance
- possibility of derivatization with functional groups
- insolubility and inertia towards reagents and solvents used
- high porosity and sufficient pore size to facilitate access to reagents and allow the chain growth

Various materials have the characteristics mentioned above, in first instance Merrifield resins. They usually consist of crosslinked polystyrene microspheres; the crosslinking degree is controlled in order to ensure solvent swelling and open the polymer pores. The polymer contains around 5% of pendant chloromethyl phenyl groups to be functionalized.[329] The functionalization of these resins with appropriate carboxylic acid bearing linkers give rise to commercially available support functionalized with

the nucleotides required to start the synthesis. As an example, 2'OMe U will be used in both the 200 μmol/g and 40 μmol/g later in this work.

Some solid supports are functionalized with spacer or linker molecules, on which the oligonucleotide will be constructed (Figure 54). The purpose of these spacers or linkers is to provide a functional group at the 3' position of the oligonucleotide. This functional group, known as the linker, facilitates the subsequent attachment of other molecules using a solution phase functionalization strategy. At the end of the solid-phase synthesis, the bond between the support and the linker is cleaved, while the bond between the linker and the oligonucleotide remains intact.[331] such supports are for example *Primer Support 5G Amino™* and *Custom Primer Support C6 Amino ™* which can used to functionalize the oligonucleotide with an amine at 3' end.[327,332]

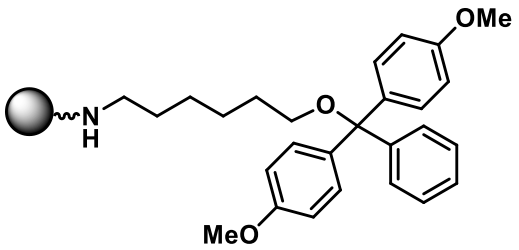


Figure 54. *Example of support functionalized with an amino linker (Custom Primer Support C6 Amino ™)*

UniSupport™ and UniLinker[333] are universal supports which allow starting the synthesis with a desired nucleotide and are available in various loadings. These supports employ a special linker which give de-phosphorilation of the 3' nucleotide upon cleavage from the support.

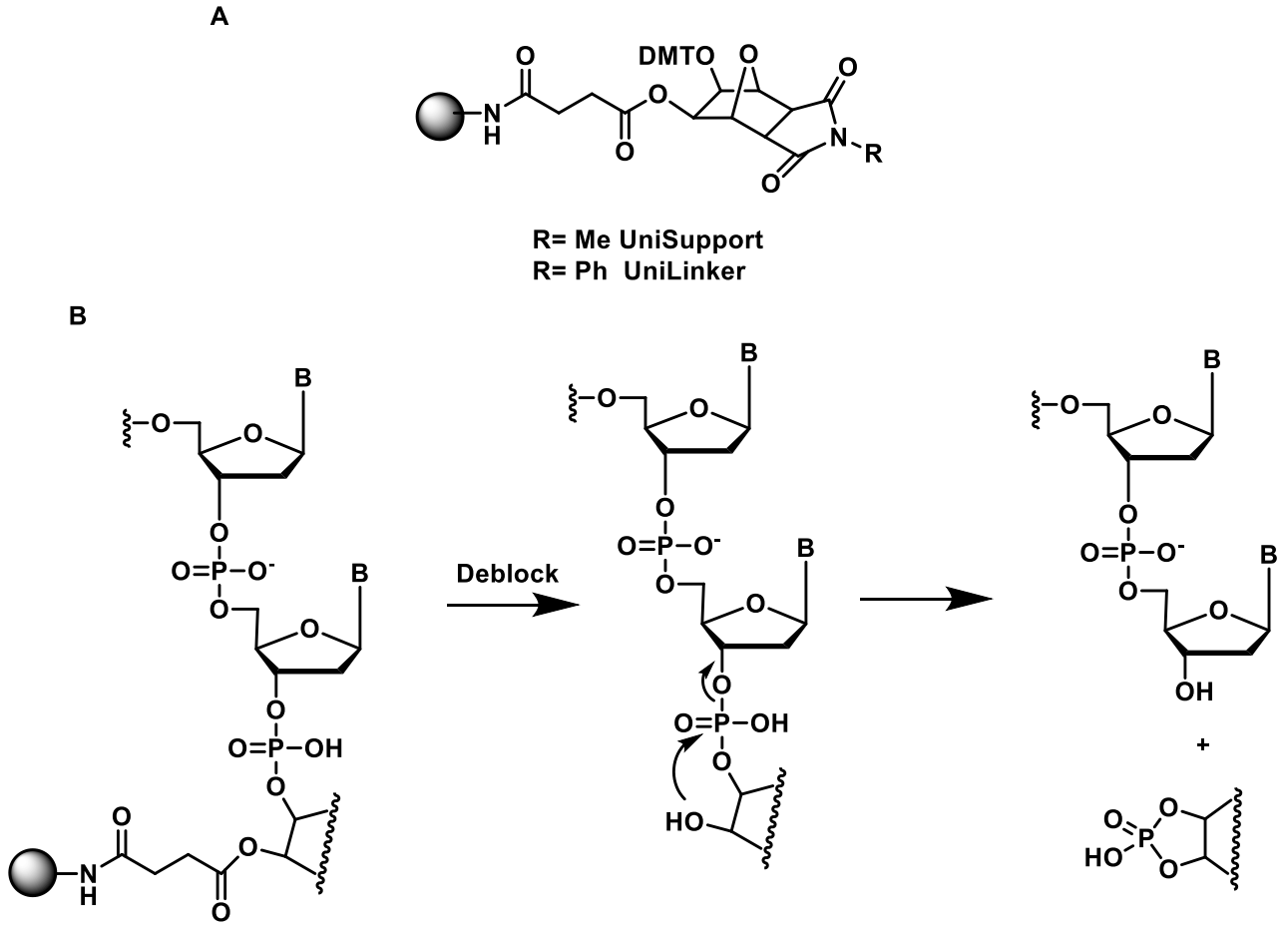


Figure 55. *Schematic representation of UniLinker and UniSupport (top). Cleavage and de-phosphorilation mechanism (bottom).*

Another material used as support is controlled pore glass (CPG), in this case is easy to see how the glass is very chemically resistant to the various conditions that can be applied during the synthesis. Typically , these resins have pores of 50 nm, in which oligonucleotides as long as 40 nucleobases can be synthesised. Furthermore, a controlled porosity with pore size of 100 nm allows the construction of long oligonucleotides (100 nucleobases) even if these supports tend to be more fragile. These supports are commonly available with a loading ranging from 0.4 to 50 μmol/g.

The primary distinction between polystyrene and CPG supports lies in how they interact with synthetic solvents. Polystyrene supports swell upon contact with solvents, which can lead to issues with overpressure in the synthesizer. However, polystyrene supports offer significantly higher loading values than CPG, allowing larger synthesis scales. They also exhibit greater moisture exclusion properties in respect to CPG, thus increasing the oligonucleotide yield. Additionally, some solid-phase syntheses are incompatible or less effective when performed on CPG. As a result, polystyrene resins are better suited for large-scale synthesis of short oligonucleotides (up to 25 nucleotides).

An essential characteristic of a solid support is its loading, typically expressed in μmoles per gram. This value is determined by the number of functional groups present on the microspheres. Given the loading value, one can calculate the amount of support required to achieve a specific synthesis scale, which is expressed in μmoles. The equation linking these two values is as follows:

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

Conditions for efficient solid phase synthesis

As mentioned earlier, the building blocks of oligonucleotides are nucleotides that are linked together through a phosphodiester bond formed between the 3'-OH and 5'-OH functional groups. In solid-phase synthesis, oligonucleotide growth proceeds in the 3'→5' direction, contrary to the natural DNA polymerase process that operates in the 5'→3' direction. This reversal in synthesis direction is employed to take advantage of the higher reactivity of the primary hydroxyl group at the 5' position in comparison to the secondary hydroxyl group at the 3' position during the coupling reactions. To ensure that only the phosphoramidite group at the 3' position would be capable of nucleophilic attack, it is essential to protect the 5'-OH group of the same molecule. Commonly used protective groups are the triphenylmethyl groups, including trityl, monomethoxytrityl (MMT), and dimethoxytrityl (DMT). The number of methyl groups on the phenyl ring affects removal rate of the protecting group, with dimethoxytrityl being more easily removed under mild conditions, typically using trichloroacetic or dichloroacetic acid in non-aqueous solvents.[334]

Following the addition of each nucleotide, the end of the oligonucleotide chain needs to be deprotected to allow the attachment of the next nucleoside phosphoramidite. Since deprotection is performed after each coupling cycle, the described protecting groups are often referred to as "temporary." In contrast, "semipermanent" protective groups are removed only at the end of the entire oligonucleotide chain synthesis. These are used to protect the amino groups present on adenine, cytosine, and guanine bases. Otherwise, these amines could compete in coupling reactions between the phosphoramidite and the 5' alcohol. It's important to emphasize the significance of this protection/deprotection system for solid-phase synthesis, as it minimizes the likelihood of unintended couplings.

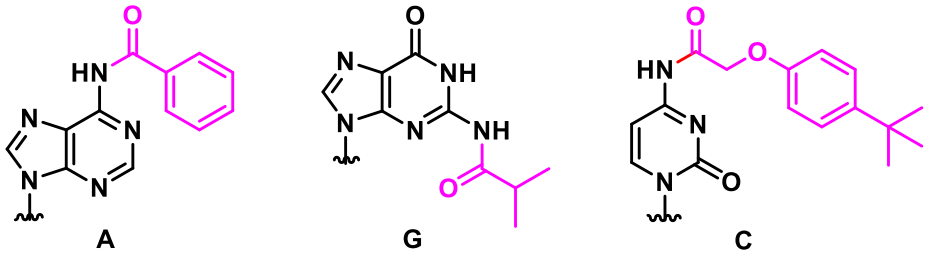


Figure 56. *Semipermanent protecting groups used for the protection of adenine, guanine, and cytosine: respectively benzoyl, isobutyryl and tert-butylphenoxyacetyl. Uracil and thymine do not need protection.*

Another potential nucleophile that can compete with amines is water, even in small quantities, present in reagents, solvents, or in the air; water could couple with the amidite instead of the intended 5' hydroxyl. To overcome this challenge, very dry reagents and solvents for oligonucleotide synthesis are utilized, containing a water concentration of less than 30 ppm. Additionally, to maintain anhydrous conditions, it is essential to keep the synthesizer continuously pressurized with anhydrous inert gas such as nitrogen (N2) or argon (Ar).

Automatic Synthesiser

The chemical assembly of the oligonucleotide chain occurs within an automatic synthesizer, which interfaces with software, such as Unicorn in our case. This software enables the programming of the synthesis and real-time monitoring of critical parameters related to the synthesis process. These parameters include system pressure, reagent conductivity, UV-Visible absorption of reagents, and the efficiency of the coupling reactions. The most advanced synthesizers are capable of synthesizing both classical phosphodiester DNA/RNA and modified oligonucleotides.

The primary components of the oligosynthesizers in use at our laboratory; the "ÄKTA oligopilot plus" by GE Healthcare (Figure 6.11), are:

- Pump P-900, a high-pressure pump system; similar the one used in HPLC.
- UV-900 monitor, with multiple wavelengths, which allows simultaneous monitoring of up to three wavelengths between 190 and 700 nm
- PH/C-900 monitor, for conductivity and pH monitoring

Figure 57. *Image of the oligosyntesiser ÄKTA oligopilot plus 10.*

The two pumps draw and distribute the amidites, and all other various reagents necessary for the synthesis through a series of valves and tubes ultimately connected to the solid support containing column. The phosphoramidite need to be freshly dissolved (dry acetonitrile) in appropriate bottles which are then attached to special connections on the outside of the instrument, for convenience. The other reagents are instead prepared in variously sized Winchester bottles placed near the instrument and connected through appropriate tubing and hermetic caps. The instrument operates in an inert atmosphere (either anhydrous argon or nitrogen) to maintain low humidity levels during reactions. An external gas source supplies constant gas pressure through a gauge and distributes gas to the various reagents through two manifolds. The presence of inert gas lines aids the pumps in creating a consistent flow (by creating pressure on the liquids to be drawn), thereby enhancing the efficiency of the synthesis process. Outside the synthesizer, there is the reactor housing where the support is loaded, and synthesis occurs.

There are two types of reactors that vary in terms of material and size, depending on the desired synthesis scale. The first type consists of columns made of polymeric material (small cartridges) with a fixed volume of 0.6 ml, and these are used for synthesizing in the range of 1-7 μmolar scale. The second type is composed of steel columns, available in various volumes, and are used for larger-scale syntheses. In any case the chosen support can be weighted inside the column.

Phosphoramidite method

Since the 1950s, various researchers have dedicated themselves to finding an effective method for the synthesis of oligonucleotides. This pursuit eventually led to the development of the phosphoramidite method; today well-established method commonly used in research and commercial applications. This method, conceived by Marvin Caruthers in the 1980s, involves the addition of a single nucleotide phosphoramidite during each cycle of synthesis.[330,335,336] each cycle is structured in four phases: deprotection of the 5' hydroxyl on solid phase; coupling of the new amidite; oxidation of the phosphite triester group to the corresponding phosphate (or phosphorothioate); capping of the eventually unreacted 5' hydroxyls.

A phosphoramidite nucleoside is composed of a nitrogenous base linked to a pentose sugar, with the 5' hydroxyl group protected by a dimethoxytrityl group. Unlike natural nucleotides, the 3' hydroxyl group is not directly linked to the phosphate group, as this connection is not useful in the chemical synthesis of oligonucleotides. Instead, it is attached to a phosphite triester group (P III), which contains a β-cyanoethyl protecting group that is removed only at the end of the synthesis. Additionally, it features a di-isopropyl amino group (Figure 58). This structurally modified compound exhibits high reactivity in the presence of activator agents but is still stable enough to be stored in the solid state for several months at -20°C. when dissolved in anhydrous solvents the phosphonamidites are usually stable for the time needed for synthesis completion, which is usually not longer than 2 days.

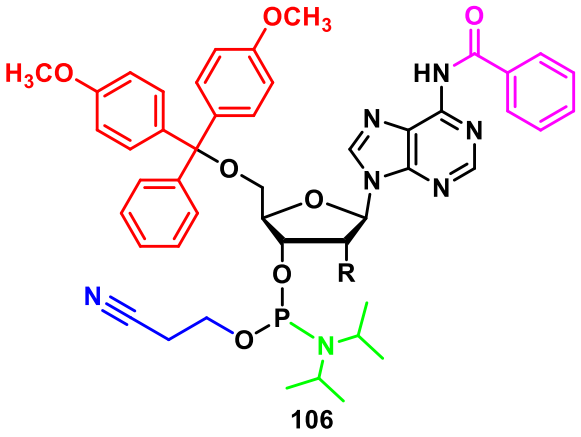


Figure 58. *Example of a nucleoside phosphoramidite structure with explicitated protecting groups. In red the acid cleavable DMT (dimethoxytrityl) group. In pink the adenine semipermanent benzoyl protecting group. In blue the 2-cyanoethyl phosphite semipermanent protecting group, and in green the di-isopropyl amine leaving group.*

As mentioned above the synthetic cycle of the phosphoramidite method starts with the cleavage of the DMT group from the solid supported oligonucleotide 5' under acidic conditions and proceeds with the coupling of the next nucleoside under strictly anhydrous conditions. An oxidation step follows the coupling and gives a stable phosphorous V link between the two nucleotides. The capping step has the purpose of avoiding chain growth on sequences lacking on nucleotide, thus reducing the impurities in the final product mixture, especially the ones similar to the product itself. All these steps will be now described in more details.

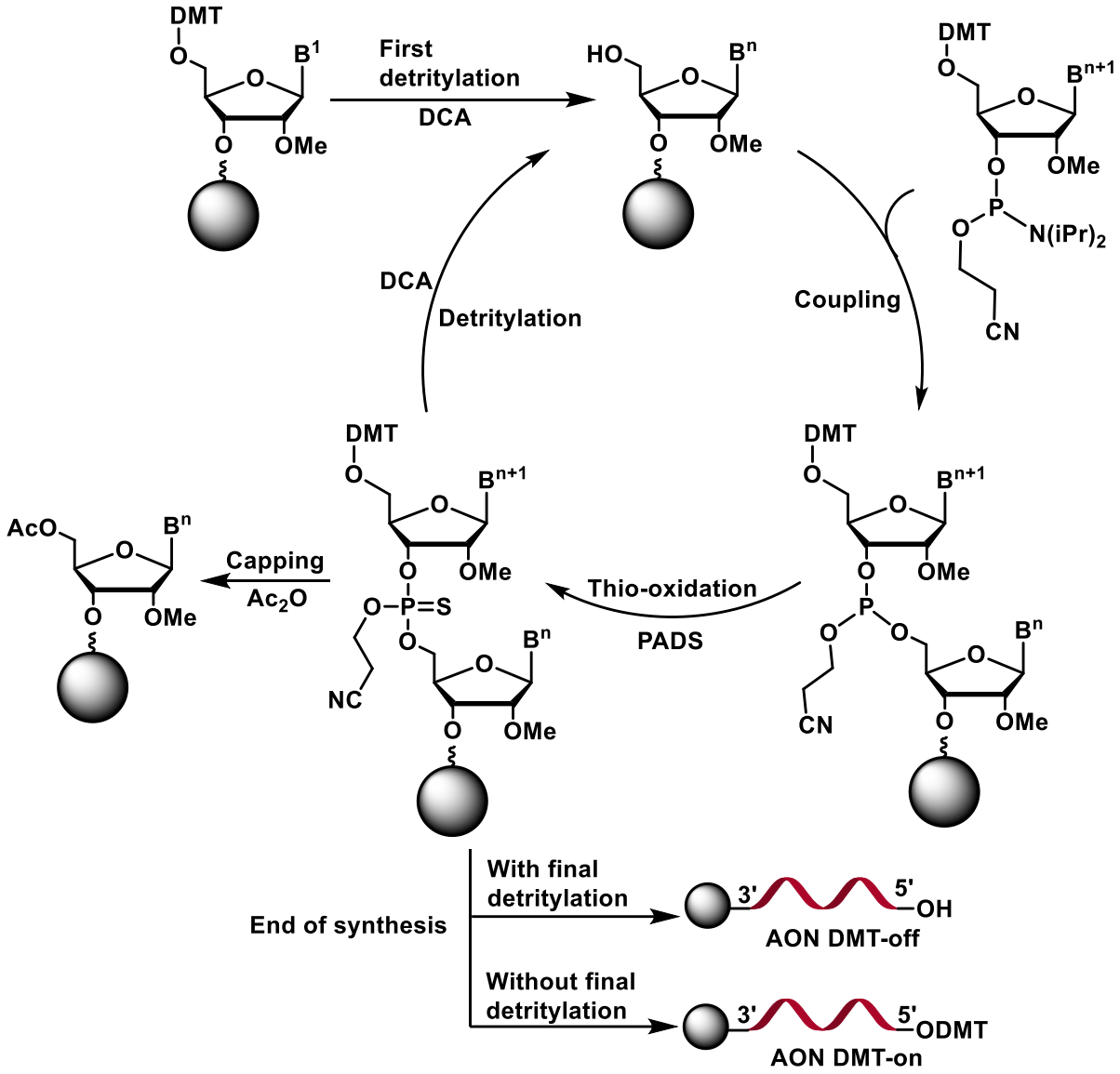


Figure 59. *phosphoramidite method reaction cycle.*

Detritylation

Removal of the dimethoxytrityl group occurs in the presence of a 3% dichloroacetic acid solution in toluene.[337] As an alternative to toluene, methylene chloride or other chlorinated solvents could be used. However, since they are less environmentally friendly and have higher disposal costs, it is preferred to avoid their use. The deprotection releases the hydroxyl group in 5', which will be involved in the subsequent coupling step. Detritylation is a fundamental step of the synthesis, because if the reaction does not go to completion, some sites will remain protected and block the chain growth thus ultimately leading to a low yield of the overall synthesis. Unfortunately, it is known that the detritylation reaction is very often incomplete.[338,339] Moreover, it not possible to simply opt for more drastic conditions since depurination can happen as a side reaction in acidic environment. Depurination is the splitting of the β-N-glycosidic bond that causes the detachment of the purine bases adenine and guanine.[340,341]

To monitor the efficiency of each individual detritylation reaction in real time, UV-VIS measurements are employed with the assistance of the software. The dimethoxytrityl cation, released in an acidic environment, serves as a chromophore group that absorbs at λ_{max}=498 nm. It is characterized by a

red/orange colour given by to the total delocalization of the positive charge on the aromatic rings composing it.

Detritylation is performed after each individual phosphoramidite coupling, with the exception of the last nucleotide in the chain. This last nucleotide remains tritylated to ease the following purification process. This gives rise to an antisense oligonucleotide (AON) referred to as "DMT-on." Alternatively, is possible to detritylate also the last nucleotide, for example when conjugation of other molecules in solid phase is envisaged or if the product is complete in that way and is easy to purify for other reasons, such as already having a lipophilic group at the 3' extremity. In these cases, the oligonucleotide is termed "DMT-off".

Coupling

The coupling step starts with a nucleophilic attack by the 5'-OH of the supported nucleotide on the phosphite triester of the incoming phosphoramidite (Figure 59). The outgoing group of this nucleophilic substitution is the di-isopropyl-amine. However, the amine is good leaving group only when already protonated to give the quaternary salt. Therefore, is necessary to employ an activating reagent that protonates the di-isopropyl-amine and makes it a better leaving group than when it is in the amine form.

The first activating agent used in solid-phase synthesis was 1H-tetrazole (Figure 60).[342] This agent has good activating properties but has some limitations, predominantly low solubility in acetonitrile, especially at temperatures below 20°C (the typical synthesis temperature), which can result in its crystallization. Crystallization of reagents is deleterious for the automatic synthesizer as it can block tubes, valves and pumps. Additionally, 1H-tetrazole is too acidic (pKa=4.89) and can lead to the premature removal of the DMT group. To overcome these issues, other activating agents can be used, such as 5-benzylthio-1H-tetrazole (BTT)[343,344] and 5-(bis-3,5-trifluoromethylphenyl)-1H-tetrazole, commonly referred to as "Activator 42" [343,345,346] (Figure 60). These activating agents offer improved solubility and stability while facilitating the coupling reaction.

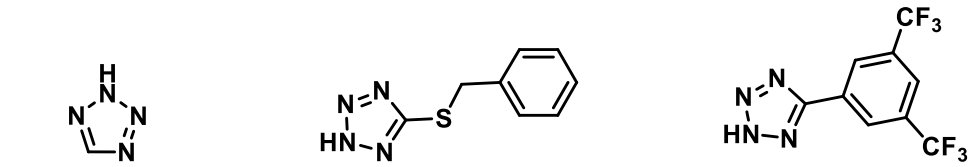


Figure 60. *Structure of common phosphoramidite coupling activators.*

The activating agent acts as a donor of protons first to activate the amine leaving group and as nucleophile after to substitute the leaving group, as shown in Scheme 6.10.

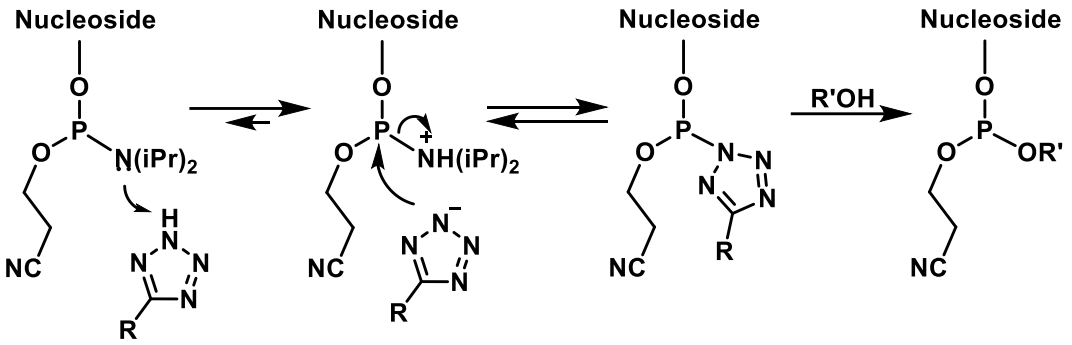


Figure 61. *Coupling mechanism with tetrazole activation.*

The time for each coupling reaction to occur with a specific amidite is chosen in the synthesis method and is known as the "recycle time." This parameter is crucial in the synthesis process and must be carefully selected to achieve optimal reaction efficiency. If the recycle time is too short, the phosphoramidite will not have sufficient time to couple with all available sites, resulting in the formation of some nucleotide sequences shorter than the intended ones. On the other hand, if the recycle time is too long, there's a risk of the phosphoramidite re-coupling at the site where it has just reacted (multiple couplings), leading to the formation of sequences longer than planned. This phenomenon is more likely to occur when the sequences to be synthesized are rich in nucleoside-phosphoramidites of the "g" type, which are more susceptible to tetrazolic acidity, making them prone to multiple couplings within the same coupling cycle.

Coupling efficiency is indeed an important parameter in the synthesis of oligonucleotides. It reflects the efficiency of the synthesizer in adding new bases to the growing oligonucleotide chain. A coupling efficiency of 100% would mean that each available base on the chain has successfully reacted with the new input base. However, in practice, achieving 100% coupling efficiency is impossible, and small deviations occur. If, for example, the maximum coupling efficiency achieved is 99%, it means that at each coupling step, approximately 1% of the available bases fail to react with the new base. Maximizing coupling efficiency is essential for obtaining the desired oligonucleotide sequence with high quality and yield. In fact, the total yield is given by the product of each cycle yield; if for example a 20-mer sequence is needed having a yield of 99% in each cycle will only give a theoretical ≈81% total yield.

Oxidation/Thio-oxidation

The thio-oxidation reaction is usually the third step in the solid-phase synthesis of oligonucleotides. It transforms the phosphite triester (P III), which is obtained in the coupling reaction, into the more stable thiophosphate (P V). The thiophosphate bond contains a sulphur atom instead of an oxygen atom, making it different from the classic phosphate bond found in DNA and RNA. This modification on the backbone contributes to the stability of synthetic oligonucleotides. However, it is easy to achieve oxidation to obtain the naturally occurring phosphate bond, and the two reactions can also be intercalated in the same synthesis to obtain a mixed backbone oligonucleotide.

The most commonly used thio-oxidizing reagent is phenylacetyl disulfide (PADS).[347,348] To enhance the performance of this reagent, organic bases such as pyridine are often added to the acetonitrile solution. These bases serve as co-solvents. Additionally, it's important to note that the solutions of these reagents need a period of aging, typically at least one day, before they are used. However, these solutions have a limited shelf life and should be used within a relatively short time, usually about fifteen days, because they tend to degrade quickly. Efficient thio-oxidation is essential for the synthesis of high-quality oligonucleotides.

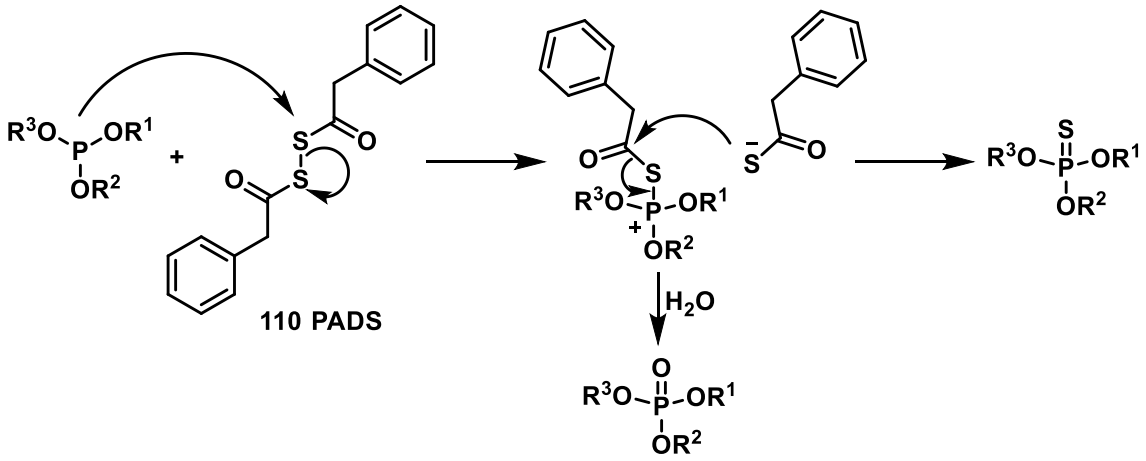


Figure 62. *Thio-oxidation mechanism by PADS. The presence of water in the reaction environment can the synthesis of phosphate by-products.*

The probable mechanism of PADS oxidation starts with 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 ensues. Water presence during the reaction can lead to phosphate side product, as reported in the figure above.

Capping

The capping phase is the last step in the solid-phase synthesis of oligonucleotides. This phase serves to block the reactive 5'-OH functions left by incomplete coupling and thus preventing failed sequences from continuing to grow. Blocking is done by protecting the hydroxyl groups at the 5' ends with an acetyl group. It's important to note that only the failed sequences have free 5'-OH groups that can be acetylated, while the correct sequences have their 5' ends protected as -ODMT and are immune to capping.

The acetylation reaction is carried out using two solutions, Capping A and Capping B. Capping A is usually 20% N-methylimidazole solution in CH_3CN, while Capping B is a mixture of two solutions, B1 and B2. B1 is a 40% acetic anhydride solution in CH_3CN, and B2 is a 60% solution of 2,6-lutidine, or pyridine in CH_3CN. The capping stage helps ensure that only the correct sequences continue to grow, making purification easier and more effective. After capping, the cycle resumes with the detritylation of the newly attached nucleotide.

DEA treatment

After completion of the sequence synthesis is possible to deblock the oligomer from the support and deprotect the nucleobases and phosphate groups. This is done in a basic solution, however, deprotection of the 2 cyanoethyl group lead to formation of acrylonitrile in solution. if the nucleobases are deprotected at same time they can attack the acrylate and give byproducts. To avoid this issue is common to trat the support with a 20% diethylamine (DEA) solution in anhydrous acetonitrile before deblock.[349]

During this phase, the cyanoethyl protecting groups located on the thiophosphate bridges are removed, releasing the acrylonitrile (AN) group, which is then washed away. The removal of these protecting groups and potential by-products is an essential step in achieving a pure and high-quality oligonucleotide.

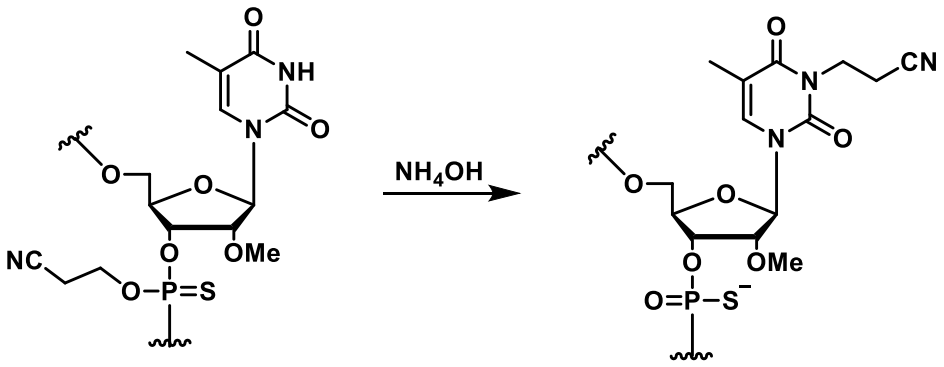


Figure 63. *Formation of by-products by alkylation in position 3 of thymidine residues during treatment with NH_4OH.*

Deblock and deprotection of the oligonucleotide

After the DEA treatment and washing, the support is extracted from the reactor and subjected to a deprotection process. This involves treating the oligonucleotide with a concentrated solution of ammonium hydroxide at 33% and maintaining it at 60°C for at least 10 hours,[350,351] although depending on the protocol followed these conditions may vary. During this process, the oligonucleotide is detached from the solid support, and the exocyclic amino functions present on the nitrogenous bases undergo deprotection. This step is essential for the final release of the fully synthesized and deprotected oligonucleotide. As an alternative to ammonium hydroxide, gaseous ammonia[352] or methylamine[353] can also be used in this deprotection process. When some specific linkers or other molecules are conjugated to the oligonucleotide, these conditions can be adjusted to ensure simultaneous deprotection of the desired added molecule or to avoid its degradation.

Purification of oligonucleotides

Purification by IP-HPLC

After synthesis and deprotection the oligonucleotide crude normally contains all of the residual cleaved protecting groups, as well as the product and a, possibly small, percentage of shorter oligonucleotide sequences. These shorter sequences are the one capped during the synthesis, in each cycle is possible to accumulate a certain amount of a shorter sequence. These oligonucleotides "shortmers" range from one nucleotide to n-1 where the n is the final number of nucleotides present in the product. In addition, there can also be presence of longer than wanted oligonucleotides (n+1) "longmers", due to the possible over coupling seen above. Other impurities can be due to depurination, which can occur during the synthesis or during the deblock; and substitution of a phosphorothioate with a phosphate group, which could happen during the oxidation phase. Lastly, especially if a DEA treatment was not executed before deprotection, is possible to have some bases modified by acrylonitrile. Of course, is necessary to purify the product from all these impurities.[354]

The purification of DNA- or RNA-based oligonucleotides typically involves three main techniques: reverse phase high-performance liquid chromatography (RP-HPLC)[355] typically in the presence of pairing agents (IP-HPLC),[356] ion-exchange high-performance liquid chromatography (IEX-HPLC), or a combination of both methods.[354,357–359]

IP-HPLC relies on the separation of products based on their different polarities, similarly to RP-HPLC. When using this technique, the detritylation of the last amidite of the sequence during solid-phase

synthesis is not performed, resulting in a DMT-on oligonucleotide that is much less polar than all the shorter chains that lack the dimethoxytrityl group. This difference in polarity facilitates the separation of the product from impurities. Especially from the n-1 oligonucleotide which is the one impurity that elute nearest the product peak. The presence of a DMT group is usually, with an optimized chromatographic method enough also to separate the product from n+1 oligonucleotide. However, it requires a subsequent detritylation reaction in solution. Generally, the stationary phase used for this purpose is a polystyrene, a C8 or C18.[360] Since the oligonucleotides are normally soluble in water ion pairing agents are added with the double intent of controlling the pH and improve the separation. The most diffused methods involve the use of triethyl ammonium acetate (TEAA) or bicarbonate (TEAB) buffers with acetonitrile gradients. The triethylammonium (TEA^+) electrostatically interacts with the phosphate groups of the oligonucleotide, effectively surrounding them and making the whole molecule more lipophilic. in a similar way the triethylammonium side chains interact with the stationary phase adding positive charges to it. This kind of reverse phase chromatography is known as ion pairing. The chemist also has the option of choosing more lipophilic buffer salts such as tributyl ammonium (TBu^+) when the oligonucleotide to purify is particularly hydrophilic and elutes too early. Alternatively, is possible to employ ammonium buffer for especially lipophilic oligonucleotides.[361]

IEX-HPLC,[362,363] on the other hand, exploits the interaction between ions present on the oligonucleotide and those of opposite charge on the resin. This technique eliminates the need for DMT-on oligonucleotides and the subsequent detritylation reaction in solution. It is advantageous because it reduces the number of purification steps and allows work in the absence of organic solvents. However, IEX-HPLC requires large quantities of aqueous salt solutions to elute the oligonucleotide from the chromatographic column. This can lead to incomplete recovery of the oligonucleotide and necessitates a time-consuming desalting phase that requires high amounts of aqueous solvents. Typically, this technique is preferred for long oligonucleotides, around 40-80 nucleobases because it gives better n-1/product resolution in respect to IP-HPLC.

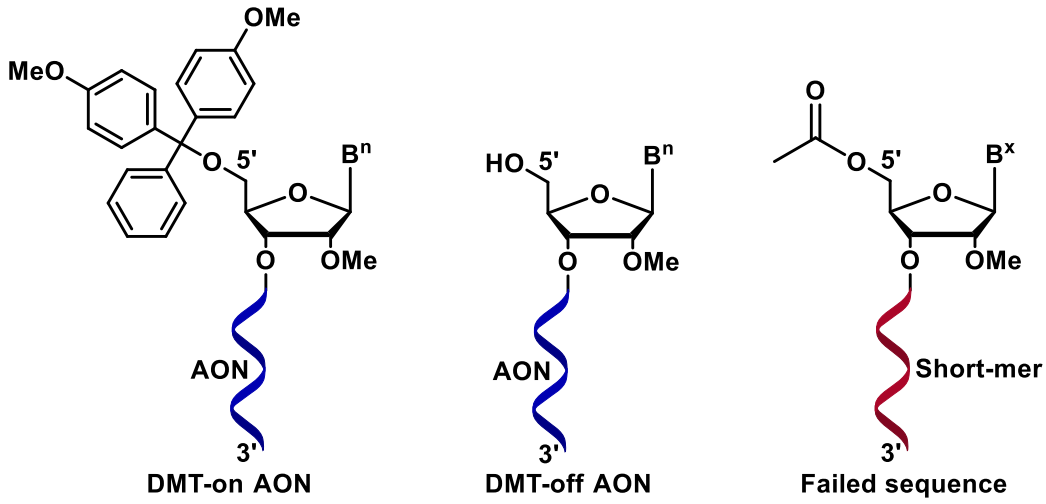


Figure 64. *Example of the different compounds to be separated in IP-HPLC. At the left is a DMT protected lipophilic oligonucleotide. In the centre a DMT-off sequence, and at the right a failed sequence capped.*

Recently, more advanced techniques have been developed for chromatographic purification of oligonucleotides and peptides. Of particular interest is an approach called Multicolumn Countercurrent Solvent Gradient Purification (MCSGP). This technique uses two identical columns to simulate a moving bed of stationary phase; and conceptually divides the reaction crude into three distinct parts. The first is

the product which elutes at a certain solvent strength, the second is given by the "strong impurities" that is to say all of those impurities that elute after the product and the third one is given by the "weak impurities" which elute before the product. In a normal situation the product peak will overlap on both sides with the impurity's peaks; to obtain a high purity product only the heart of the product peak is isolated while the rest goes to waste with the impurities. What the MCSGP system does in practice, is to start the elution on the first column, let the head of the weak impurities peak go to waste when it starts exiting the column. With the aid of valves, the head of the product peak (still containing weak impurities) goes to the second column, while the heart of the product is collected. The product peak tail (containing strong impurities) is sent to the second column, at the same time with new feed solution. The tail of the strong impurities peak is of course, discarded. This way the product normally wasted by a single purification is put back (with fresh one) in a second column and another purification cycle can start. After some cycles, a work regimen is reached, and is possible to purify large volumes of compound with much smaller columns in respect to a single elution purification. Moreover, the total yield of the product is increased, for the same purity level, in respect to a single elution purification due to a recovery of the product normally lost under the impurities' peaks. This methodology however is predominantly used for peptides and is suitable for large scale purifications.[357,364–367]

Detritylation in solution

When purification is done following the DMT-on strategy, the final compound obtained after chromatography contains only the full-length oligonucleotide chain with the 5' hydroxyl group still protected by the DMT group (DMT-on). After lyophilization, the detritylation of the compound is performed in solution using a saline buffer (sodium acetate at pH = 3) or alternatively 20% acetic acid. This reaction must occur within a short time (about an hour) to avoid depurination caused by the acidic environment.[340] The detritylation reaction has optimal kinetics when working at certain concentrations, so the AON concentration should be determined by spectrophotometric analysis. The optical density (OD) value of the sample can be used to calculate the amount of water needed to dissolve the AON, ensuring that a concentration of 1000 OD per ml of H_2O is maintained. Excessive dilution can slow down the detritylation reaction and increase the likelihood of depurination.[340,341,368]

An alternative method for detritylation in solution involves dissolving the AON in water and allowing it to stand at room temperature for 18 hours. The mild acidity of the water is sometimes sufficient to remove the dimethoxytrityl group, which precipitates as 4,4'-dimethoxytritanol (DMT-OH). This precipitate can be easily removed by centrifugation and decanting the mixture. This alternative method avoids the need for precise concentration adjustments and minimizes the risk of depurination.

Purification of the DMT-off oligonucleotide

After deprotection and eventual further coupling in solution with other molecules is necessary to purify the oligonucleotide again. As before is possible to use the same IP-HPLC technique, and because the oligonucleotide has been already purified once there should be no need to separate it from shortmers, but only from depurinated sequences if they occur, and residual DMT-on oligonucleotides. After each purification step the oligonucleotide must be either precipitated with ethanol and dried or lyophilised.

Ionic exchange

If the synthesised oligonucleotide is to be used for *in vitro* tests, it is necessary to exchange the buffer ions (TEA^+), present as counterions of the thiophosphate or phosphate groups, with sodium ions, because the first ones show cellular toxicity. This can be done by filtration on ion-exchange resins with. In a

typical procedure the oligonucleotide solution is eluted through an ion exchange resin (such as Cytiva Capto S) charged with sodium ions beforehand.

Oligonucleotide analysis

Purity of the obtained oligonucleotide can be assessed by analytical HPLC profile. The presence of only one peak in the chromatogram indicates product purity, in some cases smaller peaks are seen near the main one, but if the purification phase yielded a good product, they are usually very small. An ideal chromatogram of a pure oligonucleotide can be seen in the figure below.

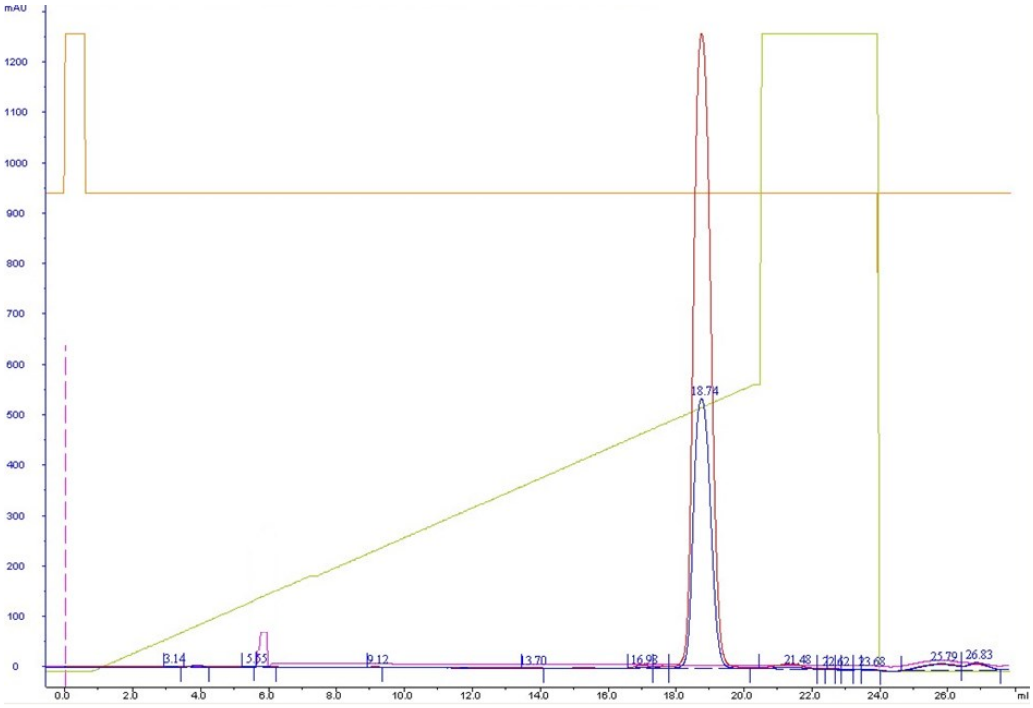


Figure 65. *Chromatogram of a purified 21mer oligonucleotide on a resource RPC 15 ml (15 μm polystyrene) column. TEAA and acetonitrile gradient at room temperature; λ=260 nm.*

Mass spectroscopy is then required to confirm the identity of the product. Various detectors can be used for this purpose, such as TOF (time of flying), of quadrupoles. With appropriate software is possible to identify the molecular peak of the oligonucleotide. Alternatively, is common to see an ionization pattern giving the molecule with various negative charges, one for every phosphate group present. It is always possible to calculate the mass of the product from the peaks of this pattern, as can be seen below (Figure 66).

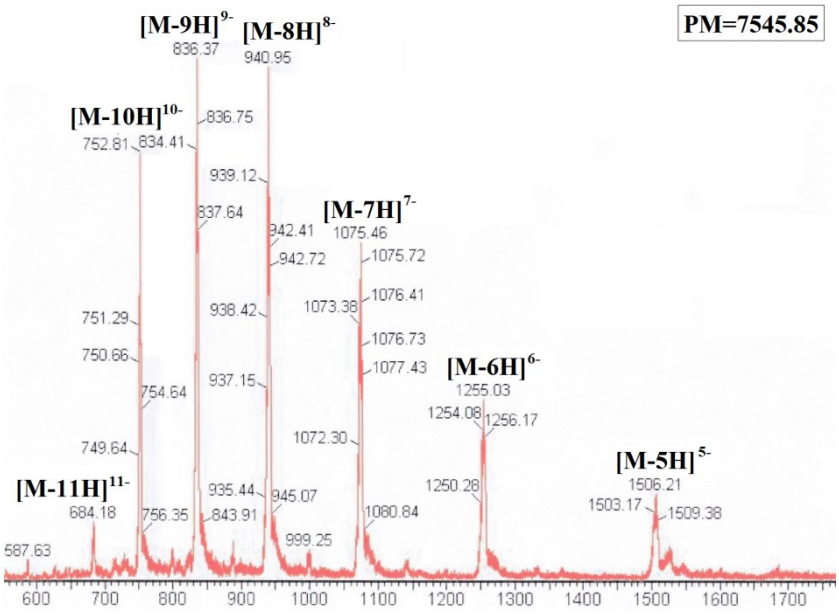


Figure 66*. example of Mass spectrum of a 21-mer oligonucleotide.*

If needed, is also possible to obtain an ^{31}P-NMR spectrum of the oligonucleotide. This way, it is possible to estimate the ratio between phosphate and phosphorothioate groups. In ^{31}P-NMR, the chemical shift of phosphorothioate groups typically falls in the range of 60-50 ppm, while that of phosphodiesters falls between 0 and -5 ppm.

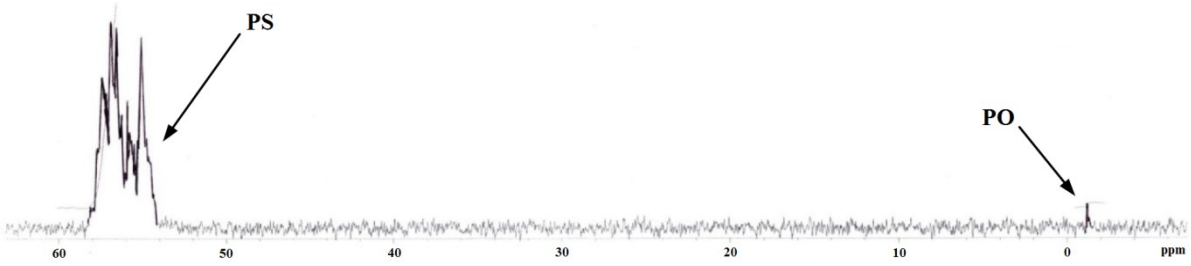


Figure 67. *^{31}P-NMR spectrum of a phosphorothioate oligonucleotide with phosphate impurities.*

PART IV: Oligonucleotide-small molecules conjugates

Synthesis of modified oligonucleotide-bile acids conjugates for study of the linker influence on exon skipping activity

Introduction

Duchenne muscular dystrophy

Duchenne Muscular Dystrophy is a rare, degenerative, and neuromuscular genetic disease related to the X chromosome. From an epidemiological perspective, it primarily affects males (1 in every 3500-6000 live male births) and manifests around the age of three with motor difficulties and weakness in skeletal muscles. Over the years, there is a progressive loss of the ability to walk, making the individual dependent on a wheelchair. Additionally, the progressive degeneration of muscle tissue also affects the heart muscle, which will eventually require pharmacological support therapies to manage cardiomyopathies. Another involuntary muscle subject to necrosis is the diaphragm, so the care provided to these patients includes mechanical ventilation. For these reasons, individuals affected by Duchenne Muscular Dystrophy unfortunately have a grim prognosis, and current life expectancy reaches around 50 years with optimal medical care.

In contrast to females, males have only one X chromosome, so a mutation in this chromosome leads to the manifestation of the disease. Females are rarely affected; more frequently, they are referred to as carriers, as they carry the mutated allele on one of their two X chromosomes.[369]

Becker Muscular Dystrophy represents a milder form compared to Duchenne Muscular Dystrophy. Therefore, the symptoms of BMD are less severe than those of DMD, and life expectancy is higher.[370]

The gene affected by genetic mutations that lead to Duchenne Muscular Dystrophy is located in the Xp21 locus. It has a size of approximately 2.2 megabases, containing 79 exons and numerous introns. This gene encodes a protein called dystrophin, with a size of 427 kilodaltons. This protein is present in various isoforms expressed in different tissues. Its primary function is to connect F-actin of the cytoskeleton to the extracellular matrix through its N-terminal and C-terminal domains. Due to the enormous size of this gene, it has the highest mutation frequency in the entire human genome.[371]

Statistically, two regions have been identified as more frequently associated with mutations, known as minor and major hot spot regions. The minor hot spot is located around intron 7, while the major hot spot is situated between exons 40-55.[372] It has been observed that these regions are prone to deletions, or the start and/or stop codons within them can be split between two adjacent exons, altering the reading frame of the produced mRNA (out-of-frame mutations). This results in the expression of a non-functional and unstable protein, leading to function loss in connecting subcellular units and providing mechanical support for the sarcolemma.[373]

In contrast to DMD, BMD is characterized by deletions of entire exons that do not alter the reading frame (in-frame exons). This allows the translation of a dystrophin protein that retains the correct amino acid sequence in the C-terminal and N-terminal domains, making it partially functional. In both cases, however, the mutated dystrophin is cause of inflammation, ischemia, and cellular necrosis resulting from the release of excessive calcium from mitochondria. Additionally, there is a high level of oxidative stress

caused by reactive oxygen species (ROS). The diaphragm is typically the first muscle to be severely affected in both forms of the disease.[374]

There is currently no definitive cure for DMD and BMD, so the suggested treatments provide support for vital functions. In the treatment of respiratory insufficiency, tracheostomy and mechanical ventilation are often necessary. In addition to these treatments, current standard pharmacological care is aimed at:

- Supporting cardiac function in the presence of cardiomyopathies. Diuretics, ACE inhibitors, and digitalis glycosides are administered.[369]
- Blocking or at least limiting inflammatory, fibrotic, and muscular degeneration processes. The administration of prednisone or deflazacort is recommended, as they can increase muscle strength and delay the loss of ambulation.[375]

In addition to symptom-focused therapies, innovative treatments have been studied over the years and have received significant attention from the scientific community. These include cellular therapy, gene therapy, and exon skipping.[376]

Therapies against DMD

Cellular therapy involves the heterologous (from a donor) or autologous (from the patient) transplantation of myoblasts, muscle stem cells, or satellite cells into damaged muscle tissue or organs. The goal is to restore muscle morphology and function by externally supplying the dystrophin that the muscle needs. Despite its immense potential, this technique, has not yet been applied at the clinical level, has numerous limitations, and requires further optimization.

Gene therapy is an approach that involves transfecting the dystrophin gene into the patient's cells. This is done through specially modified viral vectors designed to deliver the dystrophin gene into the human genome containing the mutated or partially deleted gene. This technology is still in the research phase due to the challenges it presents.

Among the other possible options, exon skipping is a highly promising therapeutic approach for DMD. This strategy is based on the observation that individuals with BMD, unlike those with DMD, express a truncated yet functional protein. In Figure 68, an example of a mutation present in DMD is shown: the deletion of exon 50 leads to the production of a non-functional mRNA, with premature termination of protein translation. Therefore, the idea is to intervene at the mRNA splicing stage using small RNA sequences called antisense oligonucleotides (AONs). According to the Watson-Crick model, these sequences pair with the target strand, such as exon 51, responsible for altering the reading frame. Due to the presence of the duplex, the splicing molecular machinery will be unable to read this piece of *pre*-mRNA and it will skip to the next exon. This will restore the reading frame and produce mature mRNA that codes for a protein that has an internal deletion compared to the wild-type dystrophin in a healthy individual but is more functional than the one observed in DMD. The result is a protein similar to that found in individuals with BMD, leading to improvements in the clinical condition of the individual.[288]

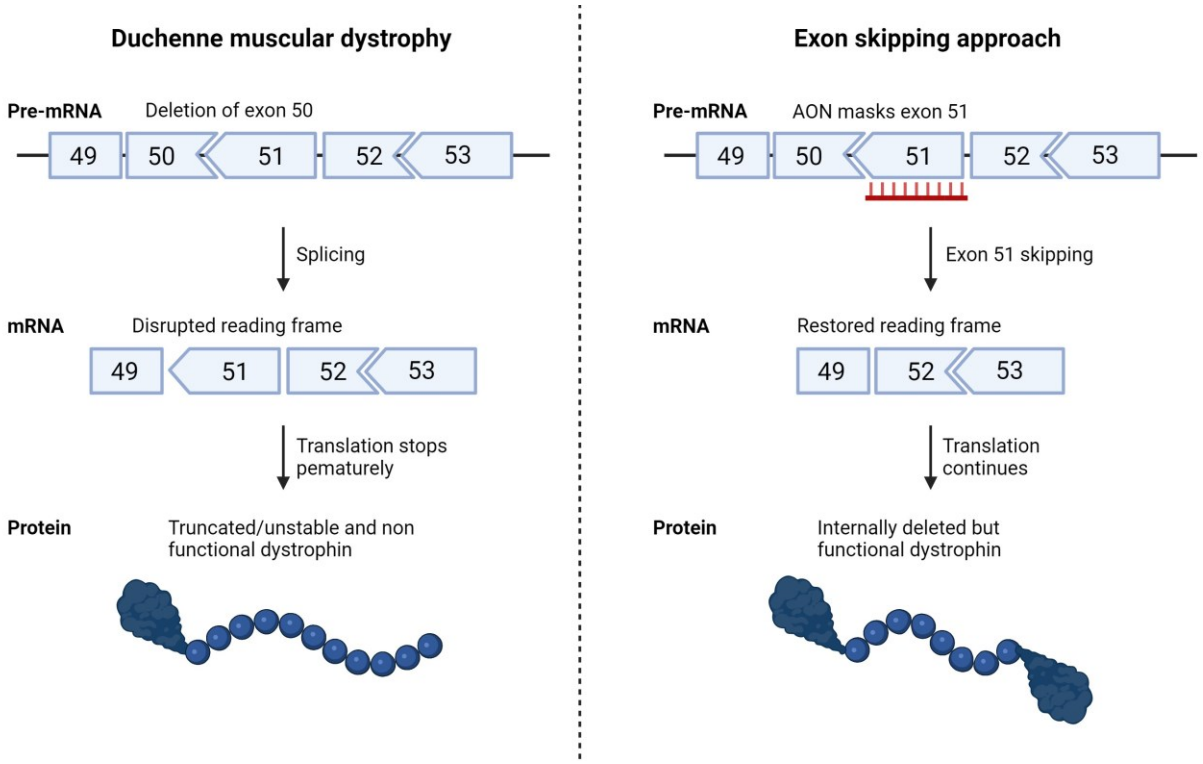


***Figure 68**. Schematic representation of the Duchenne muscular dystrophy disruption of reading frame (left) and exon skipping approach for reading frame restoration (right).*

Antisense oligonucleotides against DMD

The first AON to enter clinical trials was Drisapersen, which targets exon 51 of the human dystrophin gene, this oligonucleotide had 2'Ome modification on the ribose and had a fully phosphorothioate backbone. Over the years, more therapeutic oligonucleotides were clinically tried and approved, some are reported in Table 3. However, some were discontinued due to toxicity related issues. Until now the FDA has conditionally approved four AONs eteplirsen, viltolarsen, golodirsen, and casimersen for the treatment of DMD.[377–380] all four oligonucleotides are modified with an all phosphorodiamidate morpholino backbone (PMO), a six-sided morpholine ring resistant to enzymatic degradation. Although the clinical trial is completed for these oligonucleotides, evaluation of the long-term benefit for patients is still underway and will terminate toward the end of 2024. These studies will be very important to demonstrate the feasibility of AON in exon skipping and treatment of diseases. However, it important to notice that all the most advance antisense oligonucleotides currently have the same chemistry, this could be both an advantage or a disadvantage depending on the long-term efficacy and toxicity result. In any case, there still exist plenty of space for different chemistries and modification to be tested. In this regard, there is great variation, among which, appear PPMO (peptide conjugated morpholino oligomers) and bridged nucleic acids.

Table 3. Recent exon skipping oligonucleotides against DMD

Therapeutic Target	Name	AON Chemistry	Sponsor	Status
Exon 53 Skipping	viltolarsen	PMO	NS Pharma	Conditionally Approved

Therapeutic Target	Name	AON Chemistry	Sponsor	Status
	golodirsen	PMO	Sarepta Therapeutics	Conditionally Approved
	WVE-N531	phosphoryl guanidine (PN) backbone	Wave Life Sciences	Phase I/II
Exon 51 Skipping	eteplirsen	PMO	Sarepta Therapeutics	Conditionally Approved
	SRP-5051	PPMO	Sarepta Therapeutics	Phase II
	PGN-EDO51	PPMO	PepGen	Phase I
	DYNE-251	Antibody-PMO	Dyne Therapeutics	Phase I
Exon 45 Skipping	casimersen	PMO	Sarepta Therapeutics	Conditionally Approved
	DS-5141B	2'-O,4'-C-ethylene-bridged nucleic acid (ENA)	Daiichi Sankyo	Phase II
Exon 44 Skipping	NS-089/NCNP-02	Unknown	NS Pharma	Phase II
	AOC 1044	Antibody-PMO	Avidity Biosciences	Phase I
	ENTR-601-44	PPMO	Entrada Therapeutics	Preclinical
Exon 2 Skipping	SCAAV9.U7.ACCA	AAV U7snRNA	Astellas Pharma	Phase I/II

Table 3. *List of currently approved antisense oligonucleotides against Duchenne muscular dystrophy. Oligonucleotides under clinical evaluation are reported with the relative phase progress. Chemistry of oligonucleotides is reported.*[381]

Conjugation of AON and lipophilic molecules

Apart from backbone and chemistry modification of the sugar, is possible to change the character of oligonucleotides by chemical conjugation of different entities. Starting with the known sequence of AON against exon 51 (Drisapersen)[382] and following the 2'OMe modification with all phosphorothioate backbone a 21 mer oligonucleotide is obtained. The AON possess 19 negative charges and is naturally soluble in water. However, to cross the cell membrane this is a disadvantage, since normally the compounds able to cross the membrane are more lipophilic it seems natural to cogitate lipophilic molecules to AON to increase their membrane permeation. Our research group focused on conjugation of different lipophilic molecules, among which bile acids to oligonucleotides tested against Duchenne muscular dystrophy, the results are summarized in a recent publication.[332]

The conjugation of bile acids to antisense oligonucleotides is an innovative and relatively unexplored approach. In this study the following bile acids were considered: ursodeoxycholic acid (UDCA), hyodeoxycholic acid (HDCA), and tauroursodeoxycholic acid (TUDCA), along with the omega-3 fatty acids docosahexaenoic acid (DHA) and eicosapentaenoic acid (EPA), all conjugated at the 5'. Additionally, the AON51 sequence was synthesized with conjugation to ursodeoxycholic acid at both the 5' and 3' ends.

All the conjugates efficacy was compared using myotubes, and the results of the first set of experiments with transfection agents indicate that the most efficient compounds for exon skipping are 5'-UDC-AON51, 5'-N-UDC-AON51 (a hybrid consisting of the oligonucleotide sequence conjugated at the 5' end to one of the two ends of succinic acid; the other end forms an amide bond with ursodeoxycholic acid where the hydroxyl group at position 3 has been replaced with a primary amine), and 5'3'-bis-UDC-AON51. Compounds with skipping efficiency exceeding 40% were also tested without transfection agents. In this case, 5'3'-bis-UDC-AON51, 5'HDC, and 5'UDC yielded the best results, highlighting the necessity of good cellular uptake in exon skipping efficacy.[292]

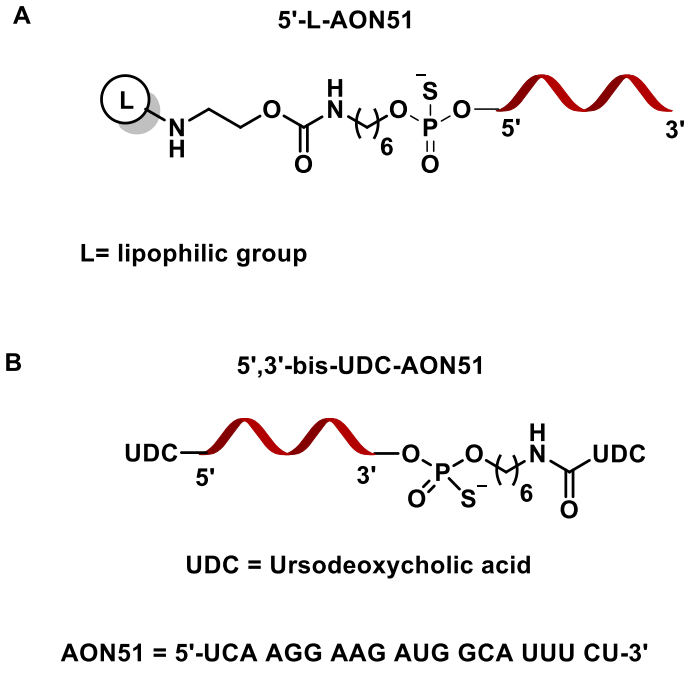


Figure 69. *A) schematic representation of AON51 conjugated with lipophilic molecules through ssH linker. B) schematic representation of the AON51 double conjugated with ursodeoxycholic acids. In 5' the lipophilic molecule is conjugated via ssH linker as before, while in 3' is conjugated through C6 amine linker.*

To better understand the mechanism by which this occurs, given the current knowledge that conjugates between lipophilic entities (such as bile acids) and hydrophilic sequences (the oligonucleotide) form supramolecular aggregates, experiments were set up using Photon Correlation Spectroscopy (PCS) and Transmission Electron Microscopy (TEM). These experiments revealed the presence of both individual molecules in solution and the formation of nanoparticle aggregates (150 nm for 5'UDC-AON51 and 150-200 nm for 5'3'-bis-UDC-AON51). Further data will be needed to understand whether the mechanism of cellular internalization occurs through the entry of individual molecules via receptors or membrane transporters or through the endocytosis of nanoparticles.

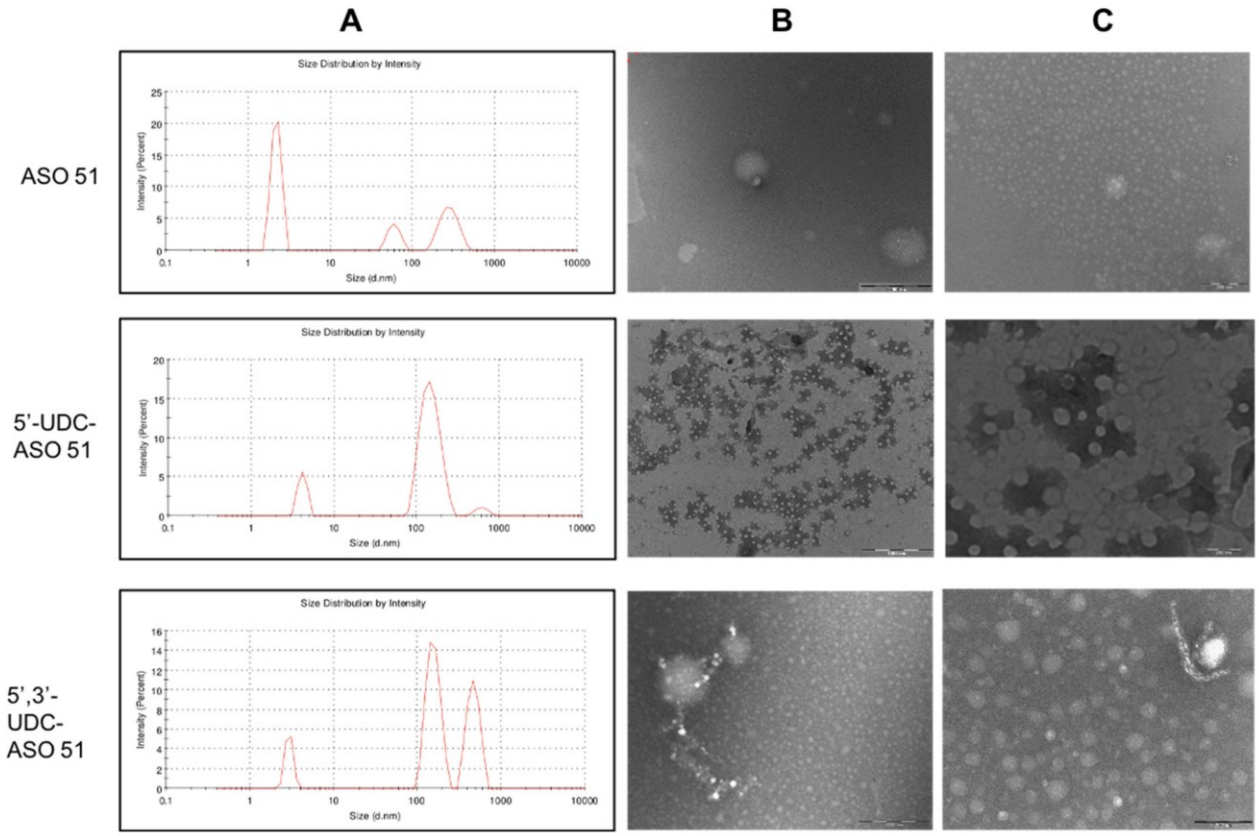


Figure 70*. Physicochemical properties of AON 51, 50-UDC-AON 51 and 50,30-bis-UDC-AON 51. (A): Size distribution by intensity detected by PCS; (B): Negatively stained transmission electron micrographs (scale bar: 1000 nm); (C): Magnifications of negatively stained transmission electron micrographs (scale bar: 200 nm).*

In the study all lipophilic compounds were conjugated to 5' via the same linker moiety (commercial ssH linker) and to 3' end by C6 amine linker. Since there appears to be a small difference in the exon skipping properties of the two conjugates our research group wondered what the influence of linker moieties could be on the final conjugate. Moreover, is not yet clear in which way the ursodeoxycholic acid performs its function so linker length could also be an important factor in determining it efficacy.

Discussion

Wanting to expand the knowledge of the oligonucleotide-ursodeoxycholic acid system our research group decided to synthesize new conjugates with different linkers. This work is aimed at determining the optimal distance between the two units of the hybrid concerning the AON51-5'-UDC conjugate. The purpose is to ascertain whether bringing the ursodeoxycholic acid (UDCA) closer to or further from the oligonucleotide sequence affects the aggregation properties of the previously investigated AON51-5'-ssH-UDC conjugate. This investigation also seeks to understand whether such proximity or separation of UDCA from the oligonucleotide sequence alters exon skipping through interactions with the enzymatic complexes involved in pre-mRNA splicing. The following linkers will be evaluated:

- C3 linkers, short aliphatic chain amine
- C6 medium length aliphatic chain amine
- C12 long aliphatic chain amine
- TEG long tri-ethylene-glycol linker more hydrophilic

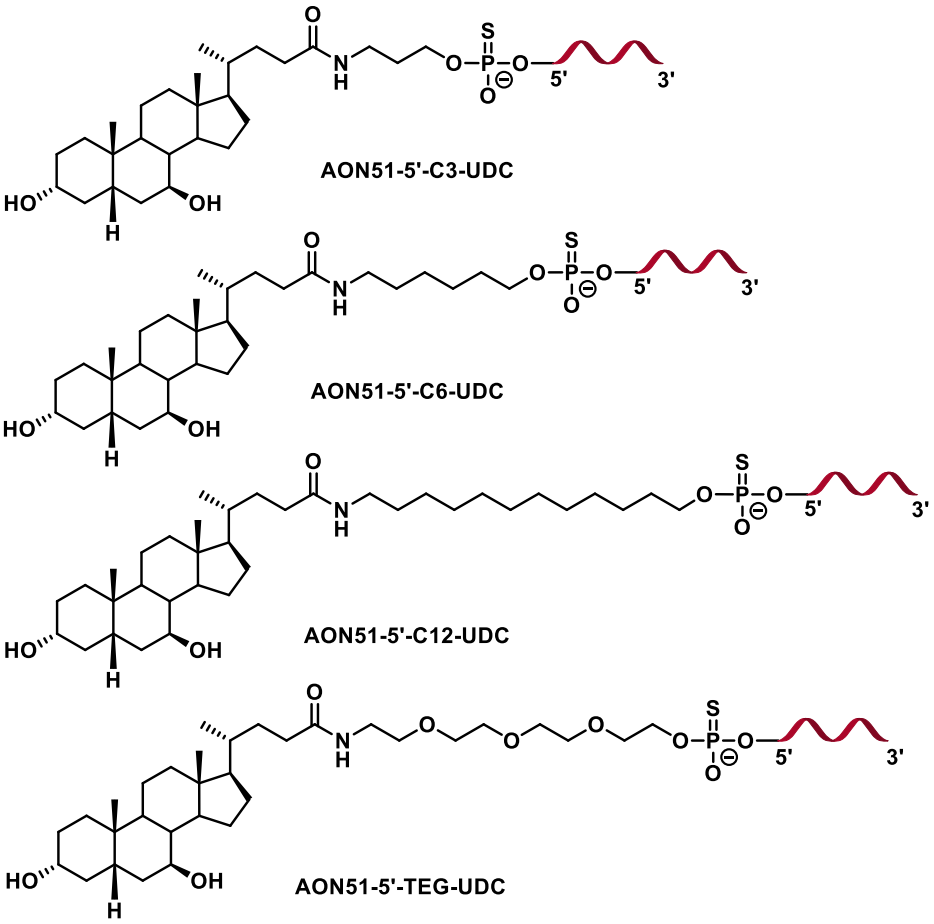


Figure 71*. scheme representing the different linkers to be used to conjugate ursodeoxycholic acid to AON51 sequence 5' position.*

Synthesis

The synthetic approach for coupling bile acids to the oligonucleotide sequence through the linker is of the post-synthetic type (as shown in Figure 72). Briefly, after the synthesis of the nucleotide sequence and linker coupling, the neutral lipid activated in the form of an active ester is reacted with the primary amino group exposed on the linker attached to the antisense oligonucleotide to form an amide bond. To do this, the synthetic pathway involves the solid-phase synthesis of the oligonucleotide sequence, which remains constant in all the conjugates in this work and corresponds to the antisense oligonucleotide targeting exon 51 of the human dystrophin gene. The sequence of this oligonucleotide is 5'-UCAAGGAAGAUGGCAUUUCU-3'. This step is followed by the deprotection of the 5'-OH terminal, followed by the coupling of the chosen linker, protected on the primary amino group with a specific protecting group. Subsequently, for the conjugation of the neutral lipid in solution, it is necessary to carry out the deprotection of the amino group, detachment of the oligonucleotide from the support, and purification of the batch to remove reaction by-products such as shortmers, corresponding to shorter sequences than the desired one, and longmers, which are longer sequences than the desired one. after this is possible to proceed with the conjugation of the chosen active ester and subsequent purification. As an example, the conjugation with ursodeoxycholic acid has been reported.

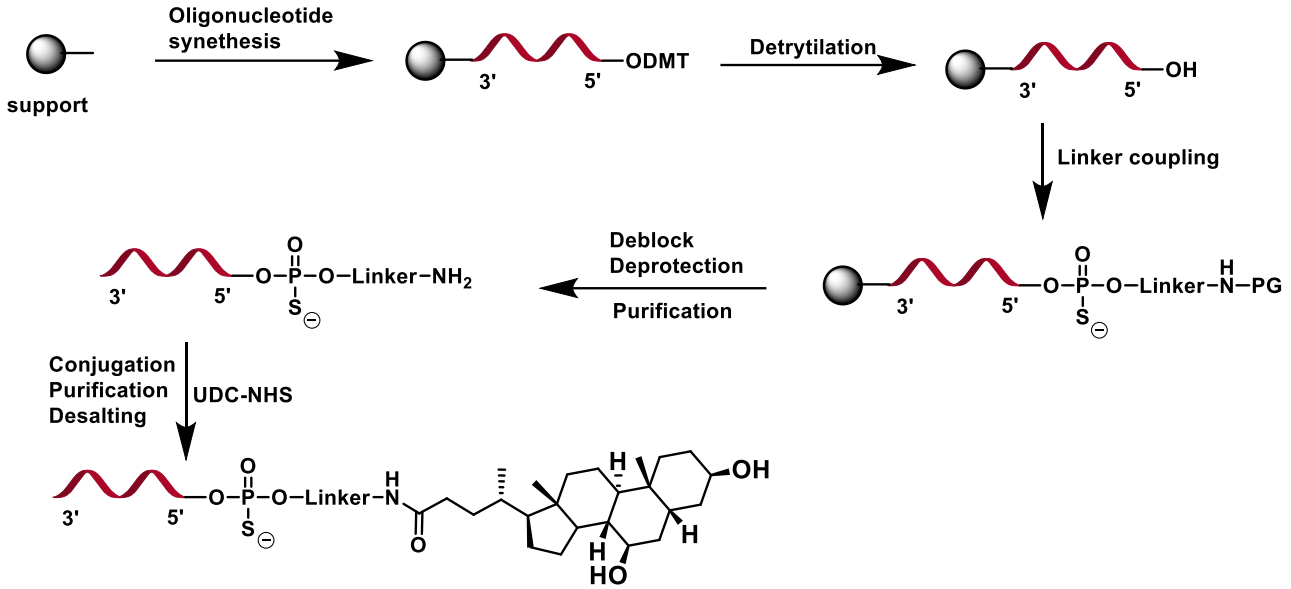


Figure 72. *Synthetic scheme for in solution conjugation of UDCA to oligonucleotide linker.*

Activation of UDCA has been carried out as reported in materials and methods section. There has been a little optimization of the procedure, and the reported synthesis is the optimized one. Initially, ester activation was carried out with DCC (N,N'-Dicyclohexylcarbodiimide) and NHS (N-hydroxysucinimide) in THF but this reaction was not giving good results. Since the reaction with DCC was not proceeding to completion, EDCI (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide) was used as a condensing agent. However, EDCI is not very soluble in THF, and this factor affected the yield. To improve the solubility of EDCI, CH_2Cl_2 was used as the reaction solvent, which provided better results in terms of yield. An attempt was also made with DMF as the reaction solvent, which could dissolve all the reagents and condensing agents, the reaction kinetics were particularly long, resulting in lower yields compared to reactions conducted in DCM. Therefore, this solvent was discarded.

Hypothesizing that the reaction would proceed to completion in a basic environment, DMAP was added to the reaction mixture in the initial attempts. However, literature suggests that this is not necessary to have a basic reaction pH. Therefore, DMAP was removed from the reaction conditions, and it was shown that almost complete conversion could be achieved without it.

The ideal conditions for obtaining a quantitative yield of the product are those reported in metrials and method section. After addition of the solvent, the reaction environment appeared turbid due to the limited solubility of EDCI. However, as the reaction proceeds, a dynamic solubility equilibrium is established, resulting in the complete dissolution of all species in solution. This leads to a clear reaction mixture, which is one of the indicators that the reaction has reached completion. the reaction however was monitored by TLC or ^{1}H-NMR. The work-up was also optimized. Initially the organic solvent was washed with 5% HCl to remove EDCI, however, this led to the hydrolysis of the active ester, thus the optimized procedure entails brine washings a and a recrystallization in toluene to obtain the pure product.

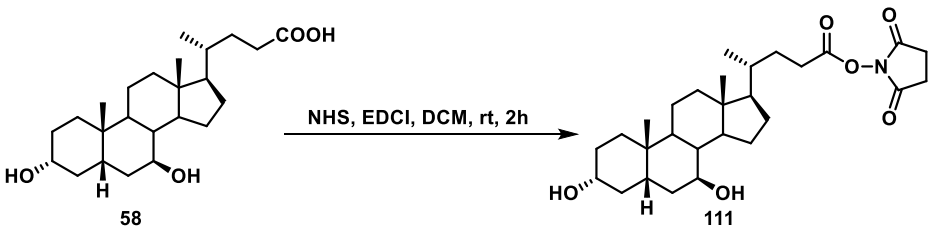


Figure 73. *Ursodeoxycholic acid activation reaction. Conditions: UDCA (1 eq), NHS (1,5 eq), EDCI (2 eq) in CH_2Cl_2; rt, 2 h.*

General synthesis of the oligonucleotide sequence AON51

The synthetic pathway to AON51 was performed on solid-phase to obtain the 20-mer RNA 2'-OMe oligonucleotide with a phosphorothioate bridge. The synthesis was automated using the "ÄKTA Oligopilot 10" oligosynthesizer from GE Healthcare equipped with the UNICORN software, following procedures previously optimized by the manufacturer, on a 1.2 mL steel column. The support used was commercially available and was Primer Support™ 5G U80 for 20 μM scales and U42 for 10 μM scales. Acetonitrile was purchased anhydrous and molecular sieves were added. Phosphoramidite solutions were freshly prepared at 0.05 M concentration and EZ Dry moisture traps with molecular sieves (EMP Biotech) were added before starting the synthesis. The phosphoramidites used in this synthesis were DMT-20-OMe-rU-phosphoramidite, DMT-20-OMe-rA-(bz)-phosphoramidite, DMT-20-OMe-rG-(ibu)-phosphoramidite, and DMT-20-OMe-rC-(ac)-phosphoramidite.

Reagents for activation, detrytilation and thiooxidation were prepared and used as recommended by the manufacturer, and were: For detritylation, DCA deblock (DCA solution in toluene 3% V:V) from Proligo Reagents; as activator freshly prepared 5-benzylthio-1H-tetrazole (BTT) solution in anhydrous ACN (0.3 M), with addition of molecular sieves; thiooxidant was a 1:1 solution of acetonitrile and a 0.2 M phenylacetyldisulfide (PADS) in pyridine.

Capping reagent consisted of a 1:1 mixture of Cap A and Cap B solution prepared directly during the synthesis by mixing in the Oligosynthesiser. Where Cap A was 20% V/V N-methylimidazole in anhydrous acetonitrile and Cap B was a freshly prepared 1:1 mixture of 40% V/V Ac_2O in anhydrous acetonitrile and 60% V/V pyridine in anhydrous acetonitrile.

In all cases after oligonucleotide synthesis and before linker coupling a DEA treatment was executed to remove the cyanoethyl phosphate protecting group. This was done using a 20 % V/V dimethylamine solution in ACN. The purpose of the treatment was to remove the majority of cyanoethyl protecting group to avoid successive formation of cyanoacrylate during deblocking and nucleobase deprotection, which could otherwise have reacted with nucleophilic amines.

Each synthesis cycle corresponded to the addition of a single nucleoside to the sequence, with the software connected to the machine providing real-time data for each of the four reactions that make up the synthesis cycle. At the end of the synthesis, the following chromatogram (Figure 74) and Table 4, which shows the areas of detritylation for each nucleoside and the overall synthesis efficiency, were obtained. The synthesis efficiency should never be lower than the 80% threshold. In this case, all synthesis were done between 7 and 20 μM, and the synthesis efficiency was always above 99% (software calculation). It is important to notice that in the table efficiency is calculated based on the previous nucleobase detritylation area.

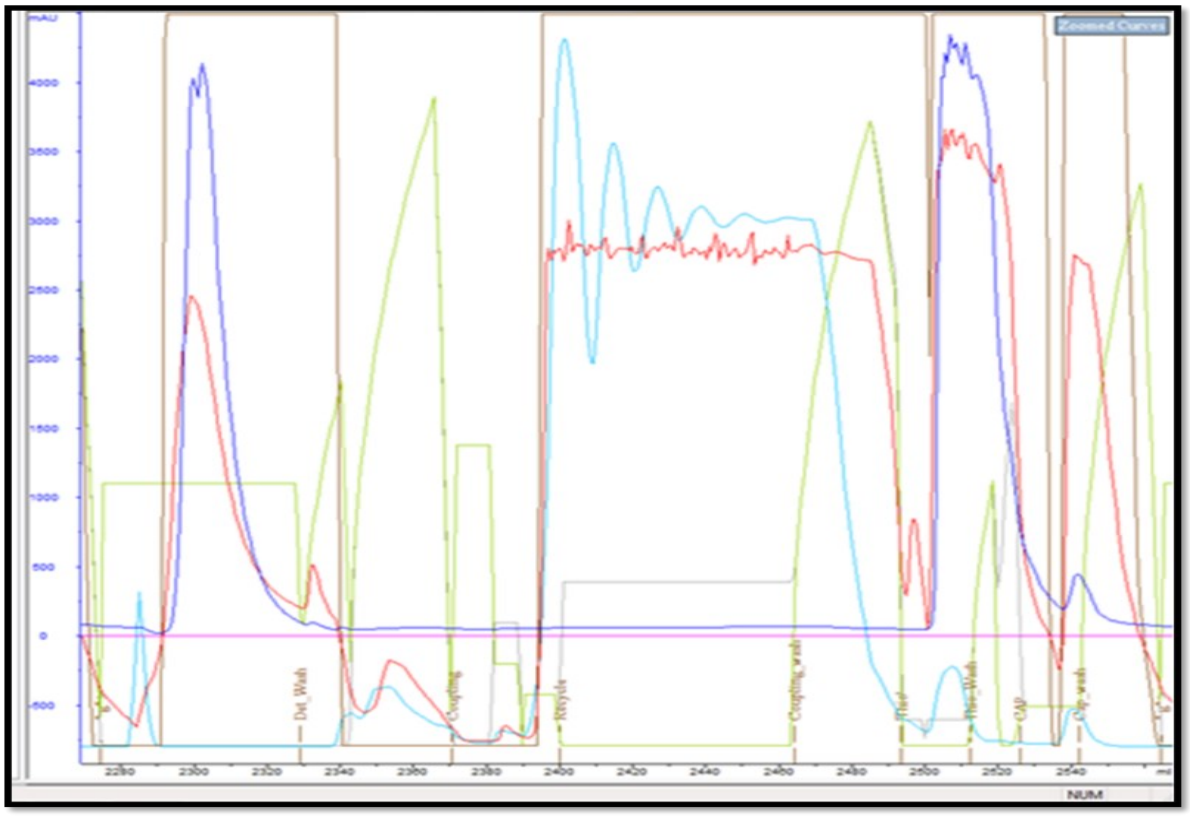


Figure 74. *chromatogram of one synthetic cycle of AON51 synthesis, each group of lines indicate a cycle. Red line indicates the UV absorbance at 280 nm; blue line is the trityl monitor (436 nm). Green line is the solution conductivity (μS/cm). light green lines indicate the solution flow (ml/min).*

Table 4. *coupling efficiencies calculated with unicorn software for the synthesis of AON 51. Underlined letters indicate RNA in this case modified with 2'Ome. The efficiency of a cycle (nucleotide coupling) is calculated by monitoring the detritylation area. Efficiency is calculated in comparison to the previous nucleobase.* Data for the AON51 synthesis

Efficiency

Average efficiency: 100.6 %
Total yield: 100.0 %

Table

Pos.	Base.	Retention ml	Duration ml	Peak Ht	Detrit Area	Last eff. %	Ave eff. %
1	u	4.41	4.50	4341.0	11113.5	100.0	100.0
2	c	4.59	6.15	4256.0	15614.7	100.0	100.0
3	u	4.87	5.80	4361.0	14858.5	100.0	100.0
4	u	6.17	6.46	4421.0	16063.5	108.1	102.0
5	u	6.58	6.94	4394.0	16780.1	104.5	102.5
6	a	6.33	6.46	4174.0	15588.3	100.0	102.1
7	c	5.85	7.51	4274.0	18212.8	103.1	102.2
8	g	5.97	6.68	4322.0	14987.6	100.0	102.0
9	g	5.33	7.24	4423.0	15533.3	103.6	102.1
10	u	7.39	8.64	4414.0	19618.2	103.2	102.3
11	a	5.97	8.42	4363.0	18473.0	103.5	102.4
12	g	5.67	8.34	4094.0	16192.7	101.4	102.3
13	a	5.49	8.68	4212.0	18381.8	99.8	102.1
14	a	6.21	8.47	4200.0	17496.0	95.2	101.6
15	g	5.38	8.07	4122.0	15072.4	97.6	101.3
16	g	5.61	8.03	4080.0	14716.1	97.6	101.1
17	a	5.55	8.34	4313.0	16637.3	98.3	100.9
18	a	5.97	7.86	4256.0	15986.4	96.1	100.7
19	c	6.38	8.34	4281.0	17279.5	99.6	100.6

Table 4*. coupling efficiencies calculated with unicorn software for the synthesis of AON 51. Underlined letters indicate RNA in this case modified with 2'Ome. The efficiency of a cycle (nucleotide coupling) is calculated by monitoring the detritylation area. Efficiency is calculated in comparison to the previous nucleobase.*

Cleavage from the support was performed with a 33% aqueous NH_4OH solution (10 mL/g of resin), and the product was filtered under reduced pressure and washed with a 1:1 solution of water and ethanol. The analysis of the produced batches was conducted using the "ÄKTA Purifier" apparatus with a "Resource RPC 3mL" reverse-phase column with cross-linked polystyrene with divinylbenzene stationary phase. Purification was performed using the "ÄKTA Purifier" apparatus with a "HiScale™ 26/20 column" packed with "SOURCE 15RPC" matrix (Cytiva). Ion exchange was carried out using "HiTrap Capto S" columns (Cytiva) suitable for HPLC. Samples were typically diluted with TEAA buffer 0.1 M pH 8 + 5% ACN and microfiltered before HPLC injection. Quantification was performed through UV spectrophotometric analysis using the "Varian CARY 100 Bio" instrument in water for injectable preparations (Fresenius Kabi). ESI-MS spectra were recorded using the "Thermo Finnigan LCQ Duo Ion Trap" mass spectrometer.

Synthesis of AON51-C3-UDC

The synthesis of AON51-C3-UDC was carried out on a μ20 M scale, oligonucleotide AON51 was synthesised in good yield according to the software calculations (100.6% total yield). After DEA treatment and 5' and washing the linker 5'-Amino modifier C3-TFA was dissolved in acetonitrile at a concentration of 0.1 M and connected to a specially designated slot. This phosphoramidite linker followed the same synthesis cycle as nucleoside phosphoramidite.

At the end of the process, the resin was dried under vacuum to eliminate any traces of solvent. Subsequently, the product still attached to the support was transferred to a vial with a screw cap, 7 mL of 33% ammonium hydroxide solution were added, the reaction was carried out at 50 °C for 18 hours.

after which sample of the solution was taken, diluted in a TEAA 0.1 M pH 8 + 5% ACN buffer solution, and an analytical check was performed using HPLC on a reverse-phase Resource RPC 3 mL with a buffer composed of two solutions: solution A containing 0.1M triethylammonium acetate (TEAA) at pH 8 + 5% ACN and solution B containing acetonitrile. The programmed run was executed with two acetonitrile gradients the first from 0 to 25% in 6 column volumes (CV), followed by a second gradient from 25% to 45% in 4 column volumes (method 1). The resulting chromatogram was as follows (Figure 75).

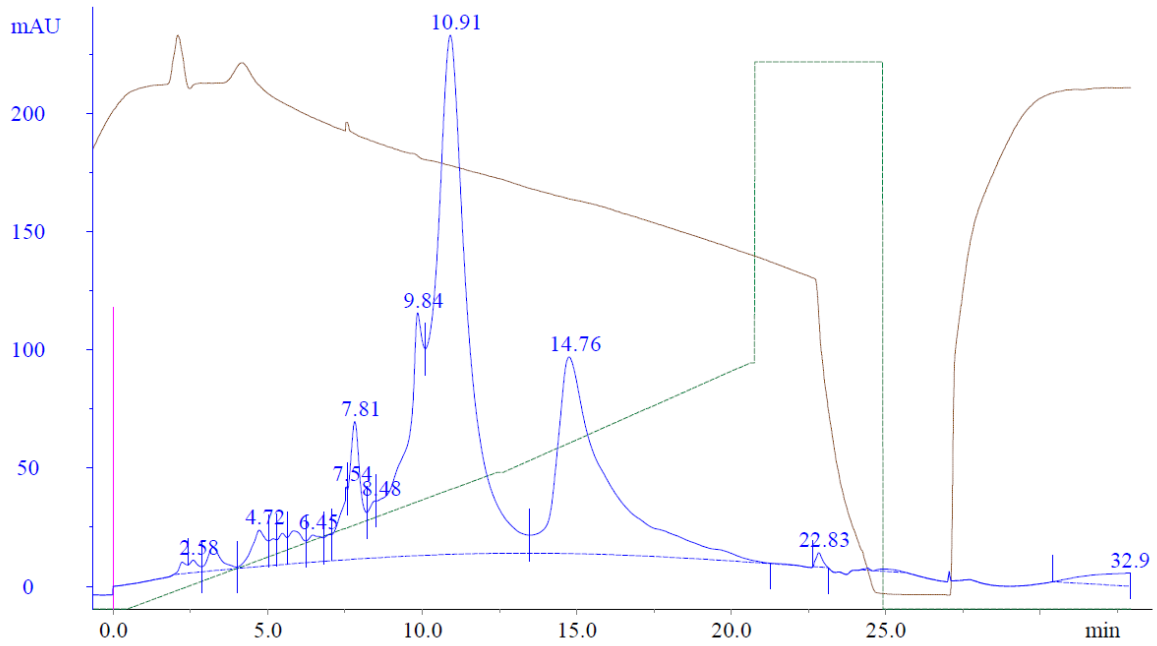


Figure 75. *chromatogram of crude after synthesis of AON51-C3. The blue line indicated the UV trace, green line refers to the gradient and red line is the conductivity of the solution.*

The peak appearing at 10.91 minutes could correspond to the product AON51-5'-C3-NH2. The small difference in the number of carbon atoms between the "naked" oligonucleotide AON51 and the one linked to the C3 linker makes these two products almost indistinguishable. The peak appearing at 14.76 minutes could correspond to AON51-5'-C3-NH-TFA (the product is still protected).

This reaction crude was diluted in a TEAA 0.1 M pH 8 + 5% ACN buffer solution, loaded onto the HPLC apparatus HiScale™ 26/20 column packed with SOURCE 15RPC matrix, and purified; fractions corresponding to the following peaks were collected:

- Fraction (1A) containing impurities.

- Fraction (1B) containing the presumed product.

- Fraction (2) containing the oligonucleotide linked to the linker while still protected.

All fractions were lyophilised and reaction with N-hydroxy succinimide activated ursodeoxycholic acid (UDC-NHS) was carried out of the presumed product fraction by dissolving the oligonucleotide in DMSO to a final concentration of 5 mM and adding 10 equivalents of DIPEA and 2 equivalent of activated ursodeoxycholic ester.

The reaction was monitored at regular intervals through HPLC sampling on a reverse-phase Resource RPC 3 mL column, as before, with a different gradient consisting of 0 to 15% B in 2 column volumes (CV), followed by a second gradient from 15% to 45% in 4 column volumes (method 2). The analyte

was diluted in a TEAA 0.1 M pH 8 + 5% ACN buffer solution and loaded onto the column. The resulting chromatogram was as follows:

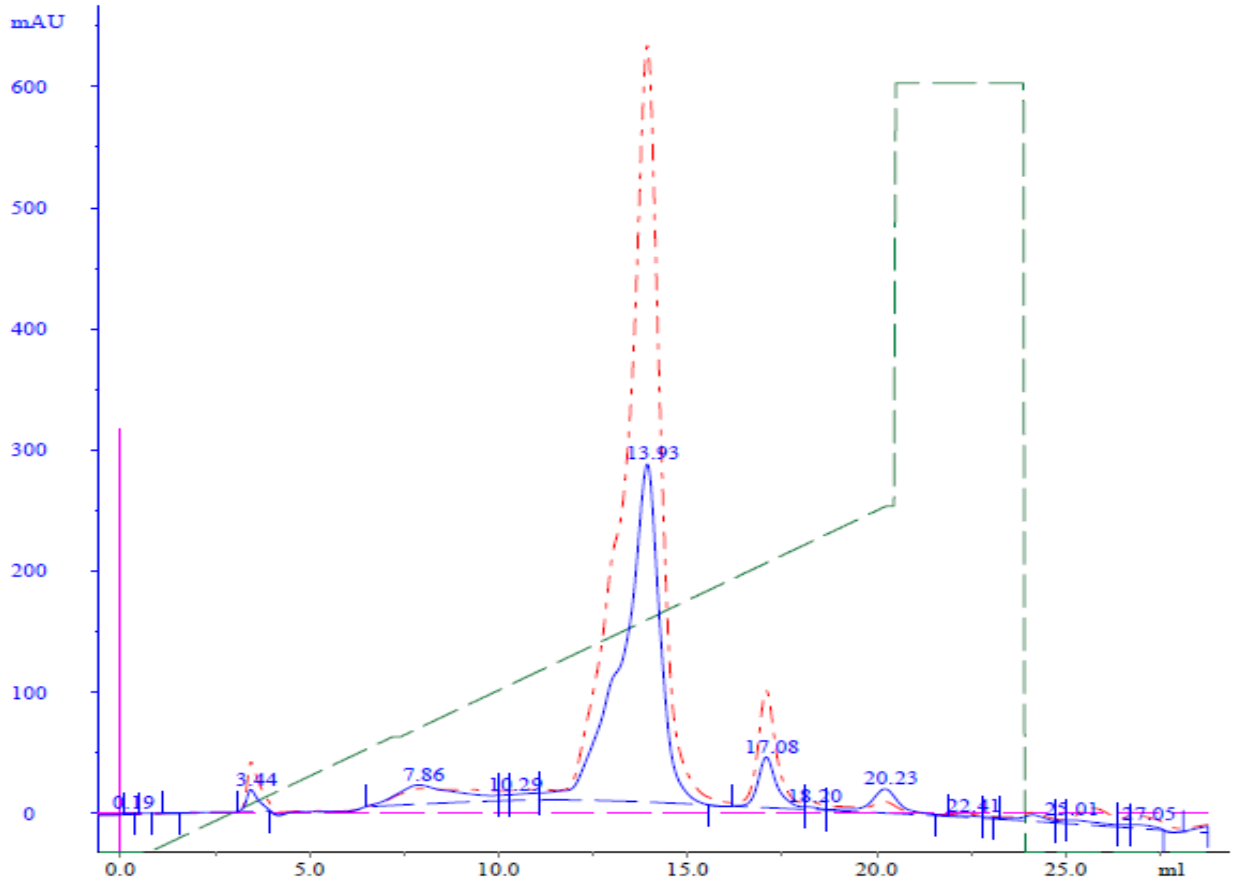


Figure 76*. Chromatogram of reaction crude. peak with Ret(mL) = 13.93 mL, which could correspond to the product AON51-5'-C3-NH2.2. A peak with Rt = 17.08 mL, which could correspond to the product I (AON51-5'-C3-UDC).*

The chromatogram profile does not show quantitative conversion of the reagent into the product. One possible reason for the incomplete conversion could be the poor efficiency of linker coupling. If the linker is not effectively coupled to the oligonucleotide, it may result in incomplete reactions. Another possibility is that the C3 linker may have too few carbon atoms, leading to insufficient spacing between the neutral molecule and the oligonucleotide sequence. This lack of spacing could create steric hindrance, making the reaction thermodynamically unfavourable.

Since the chromatographic profile is not ideal and does not allow for the purification of the isolated product, further optimization of the synthetic route is needed to achieve a high degree of purity in the product. This optimization may involve adjusting reaction conditions, linker design, or purification methods to improve the overall efficiency of the reaction. However, when the reaction was reiterated with newly purchased linker similar results were obtained, this reinforced the possibility of a too large steric hindrance for the amide coupling to happen efficiently.

For the synthesis of AON51-C6 the on support conjugation strategy was employed. In this strategy is possible to avoid one purification step since the conjugation is done directly on support and there is no need to purify the product after deblocking and deprotection. On the other hand, the reaction on solid support is more difficult to monitor and in general tends to be slower than in solution. The scheme below shows the general steps followed for this kind of reaction.

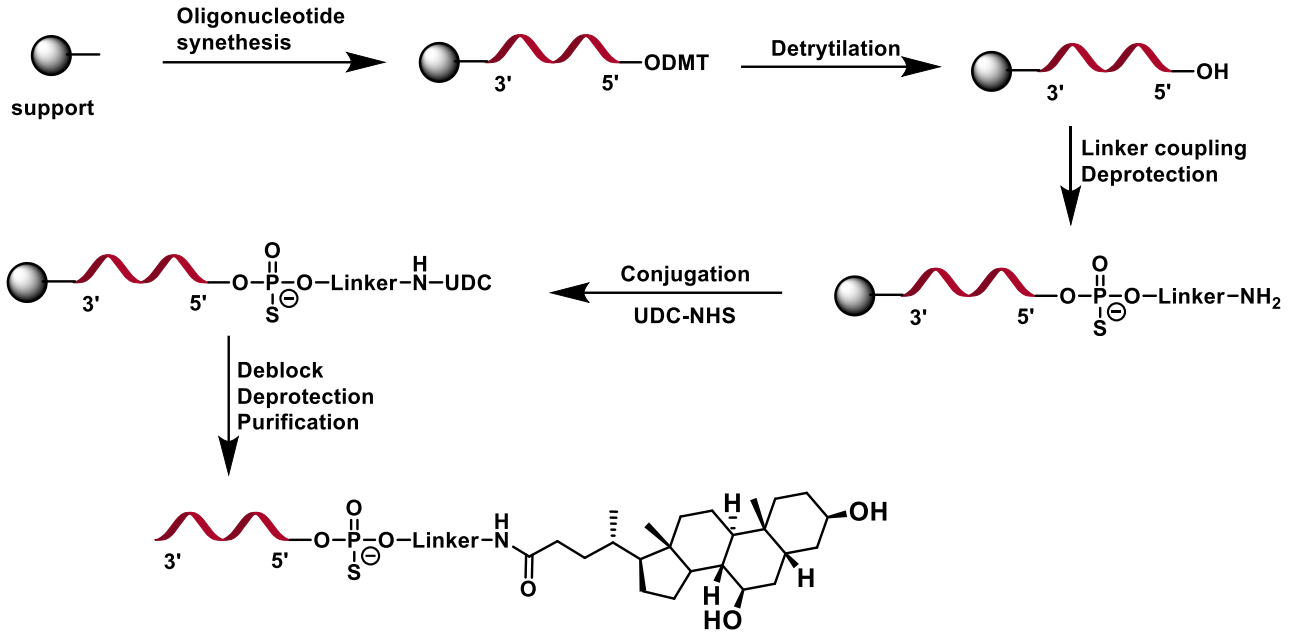


Figure 77. *synthetic scheme for on-support synthesis of conjugate oligonucleotides.*

Synthesis of AON51-C6-UDC

The oligonucleotide sequence was obtained as described before in good efficiency and yield on a 10 μM scale. The linker conjugated to the oligonucleotide sequence was the 5'-Amino-Modifier C6-MMt. It was reacted with the free 5'-OH terminus using a protocol similar to the one described earlier. For this synthesis an on-support strategy has been followed, so after the linker coupling cycle, deprotection was manually carried out setting the trityl monitor wavelength at 472 nm. Indeed, MMT (monomethoxytrityl) protecting group is commonly used in oligonucleotide synthesis and can be cleaved with dichloroacetic acid as for DMT. The support was first washed for 2 minutes with acetonitrile (5ml/min), then detritylation solution was flowed for two minutes, then again, the resin was washed with acetonitrile. This process was repeated two times until no more yellow colour could be seen eluting from the column.

Following deprotection and cleavage with an ammonium hydroxide solution (60°C, 18 hours), an analytical check was performed using HPLC on a reverse-phase column "Resource RPC 3 mL" with the method 1. The analyte was diluted in a buffer solution of TEAA 0.1 M at pH 8 with 5% ACN and loaded onto the column. The resulting chromatogram shows a small percentage of the oligonucleotide still protected.

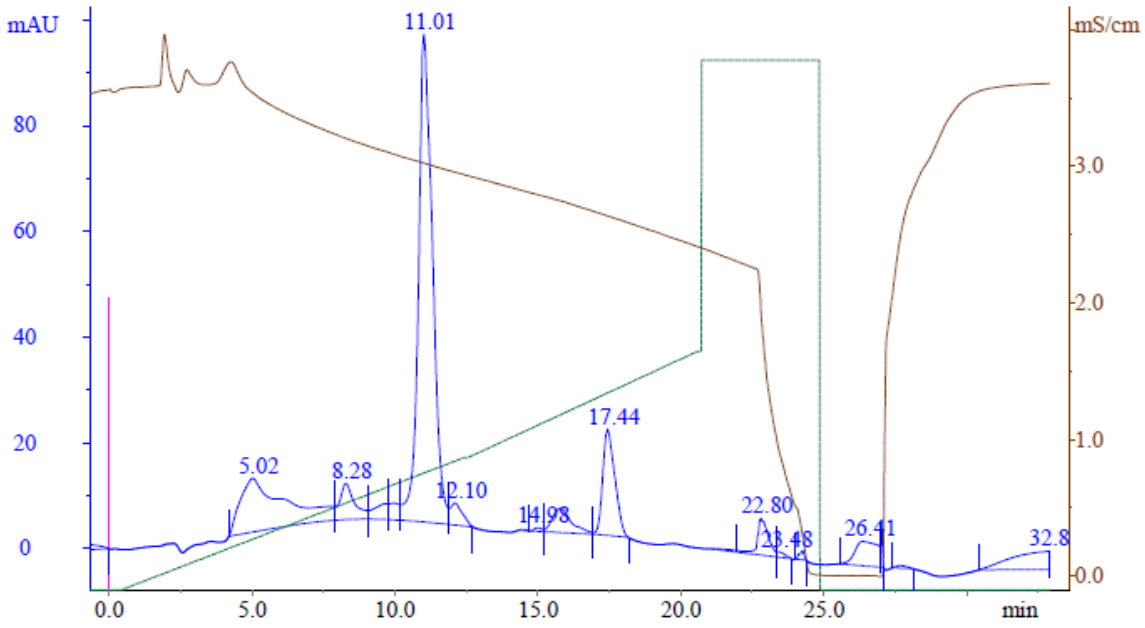


***Figure 78**. The chromatogram shows the raw synthesis of the oligonucleotide. The peak with a retention time (Tr) of 11 minutes and an Area/Peak area % of 47.8 corresponds to AON51-C6-NH_2. The peak with a Tr of 17.4 minutes and an Area/Peak area % of 8.87 corresponds to AON51-C6-MMT (the product is still protected).*

The reaction with the UDC-NHS active ester was performed on the solid support. A cartridge was used, inside of which the solid support linked to the hybrid consisting of the newly synthesized oligonucleotide sequence and the chosen linker was weighed (30 mg). This column was equipped with filters and connected at both ends to two syringes that allow the creation of a flow of the reaction mixture used.

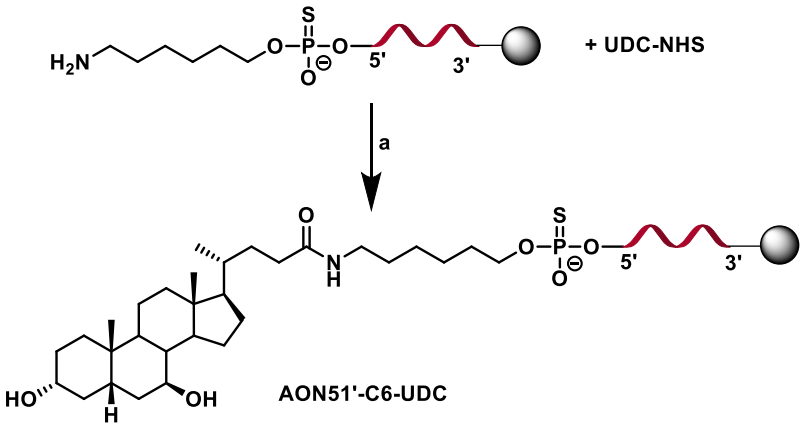


***Figure 79.** condensation reaction of AON51-5'-C6-NH_2. a) AON51-5'-C6-NH_2 (1 eq); UDC-NHS (10 eq); DIPEA (40 eq); ACN/DMSO 3:1; rt, 3 hours.*

At the end of the process, the syringes were disconnected from the reactor, and the solid support was washed successively with 5 ml of DMSO and 5 ml of acetonitrile. Subsequently, the product still attached to the support was transferred to a screw-capped vial, and a volume of 7 mL of a 33% aqueous solution of NH_4OH was added. The mixture was then kept at 60°C for 18 hours.

Afterward, the solid support was filtered, washed with a 1:1 solution of water/ethanol, and the organic solvent was evaporated.

The sample was subjected to an analytical check using HPLC on a reverse-phase column with the method 1. The analyte was diluted in a buffer solution of 0.1 M TEAA at pH 8 + 5% ACN and loaded onto the column. The resulting chromatogram was as follows:

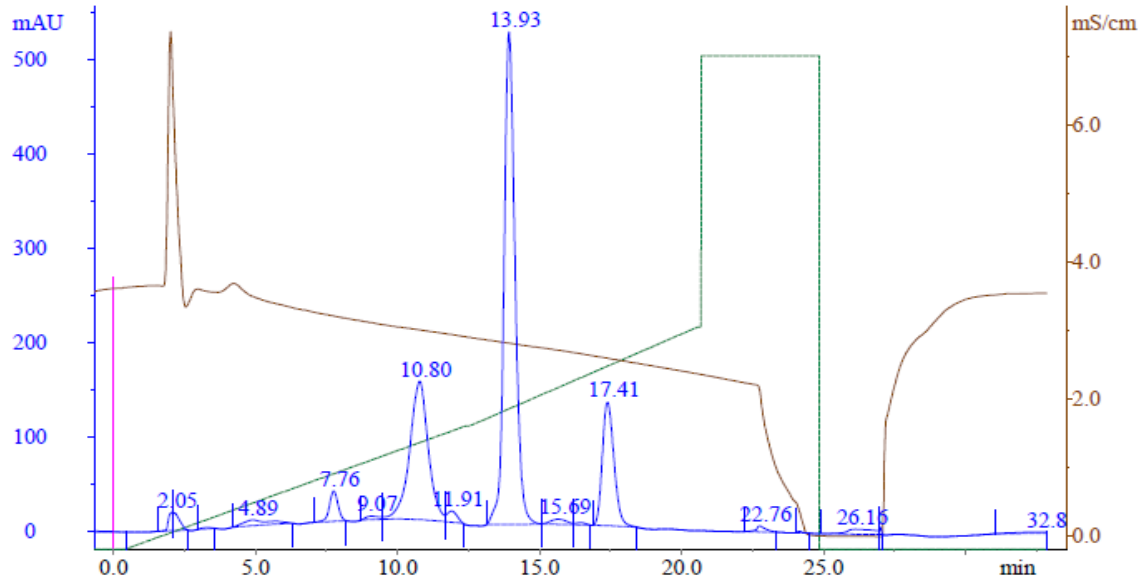


***Figure 80**. Chromatogram after on support reaction crude with UDC-NHS. The peak at Tr = 10.8 minutes with an Area/Peak area % of 24 corresponds to AON51-5'-C6-NH2. The peak at Tr = 13.9 minutes with an Area/Peak area % of 50.9 corresponds to AON51-5'-C6-UDC. The peak at Tr = 17.4 minutes with an Area/Peak area % of 13 corresponds to AON51-C6-NH-MMT (product still blocked).*

Although the reaction was not fully complete with this system it was possible to roughly estimate starting conversion to a 65% using the relative area of the peaks and keeping the area of the MMT protected product as an internal standard.

The support was then subjected to the standard procedure for deblocking and deprotection, followed by purification and lyophilization. Mass analysis confirmed the identity of the product calculated 7547,88; found 942,58 $[M-8H]^{8-}$, 1077,26 $[M-7H]^{7-}$, 1259,89 $[M-6H]^{6-}$, 1507,76 $[M-H5]^{5-}$. Product fraction was desalted using HiTrap Capto S cartridge. At the end of the whole process 2.92 µmol (22.3 mg) of sodium salt AON51-C6-UDC were obtained in approximately 26 % total yield.

Synthesis of AON51-C12-UDC

The scale of AON51 synthesis was again 10 µM for the product AON51-C12-UDC. The linker conjugated to the oligonucleotide sequence is the 5'-Amino-modifier C12-MMt and was reacted with the free 5'-OH terminus using the same protocol described previously for C3 and C6 linkers.

After oligonucleotide synthesis the support was dried under vacuum and transferred to a screw-capped vial, and 7 mL of an aqueous solution of NH_4OH at 33% is added, followed by incubation at 60°C overnight.

Afterward, the solid support was filtered, washed with a 1:1 solution of water and ethanol, and the solvent was evaporated, yielding 7.5 mL in total. A sample of analyte was diluted in a buffer solution of 0.1 M TEAA at pH 8 with 5% ACN and loaded onto the column. The sample was subjected to analytical control using an HPLC column with a reverse-phase with method 1. From the chromatogram in Figure 81 was possible to see that the coupling reaction with C12 linker was not very good and only afforded a

good amount of the desired product. In any case, it was still possible to continue with the synthesis once the desired product was purified.

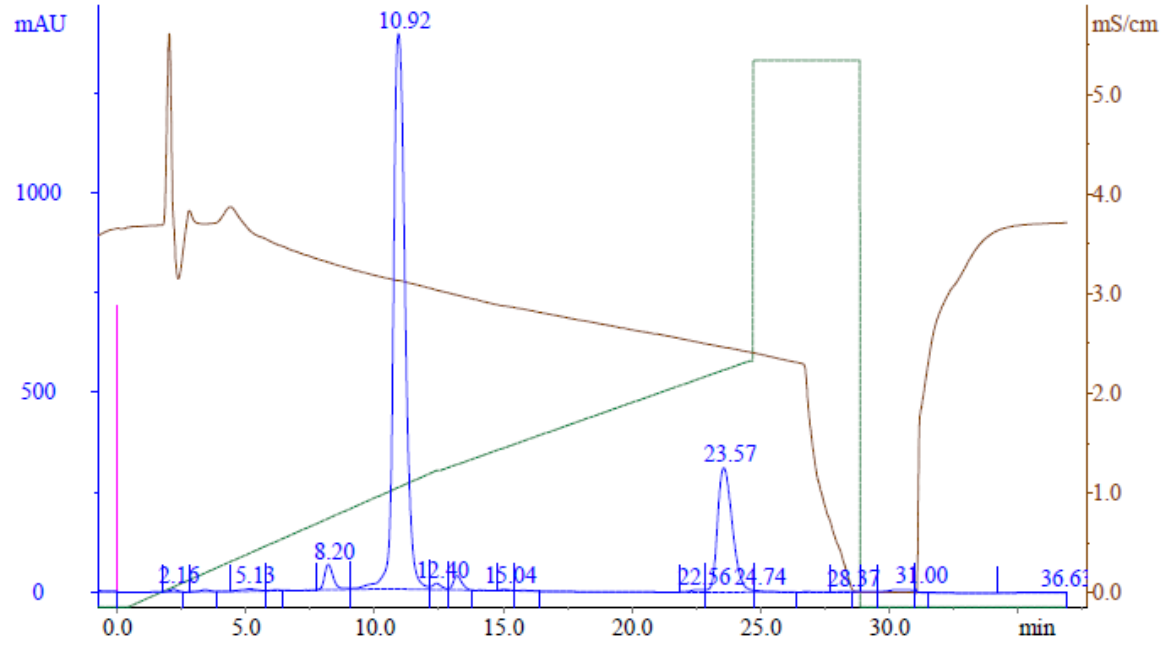


Figure 81*. chromatogram of reaction crude for AON51-C12-NH. Peak at 10.92 corresponds to AON51 while the peak at 23.57 is AON51-C12-MMT.*

The purified product fraction was lyophilised and retaken in a 20% acetic acid solution in water. The solution was kept at room temperature for 20 minutes and 0.1 M TEAA (pH=8) buffer was added. The aqueous solution was extracted three times with the same amount of ethyl acetate, each time the organic solvent was removed by pipetting. The aqueous phase was then lyophilised again. This treatment removed the MMT protecting group allowing the subsequent reaction with UDC-NHS. As before the reaction was carried out using 2 equivalents of UDC-NHS for six hours at room temperature in DMSO.

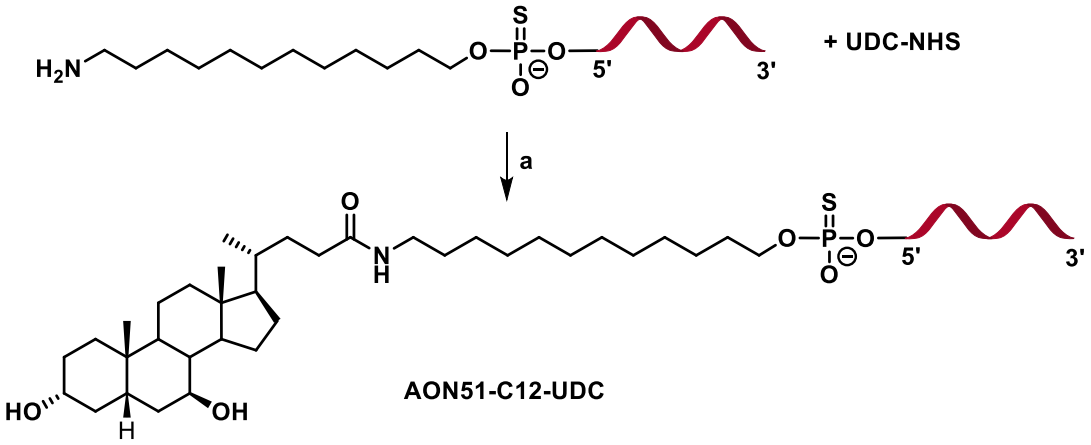


Figure 82. *Reaction of UDC-NHS on AON51-5'-C12-NH₂. a): AON51-5'-C12-NH₂ (1 eq); UDC-NHS (2 eq); DIPEA (2 eq); DMSO; rt, 6 hours.*

The reaction was followed by sampling and running analytical HPLC. After 6 hours addition of UDC-NHS and DIPEA was required. The reaction proceeded overnight. At the end of the process, the

chromatogram (Figure 83) highlights the presence of the conjugation product of ursodeoxycholic acid on AON51-5'-C12-NH_2, with a retention time (Tr) of 16.65 minutes.

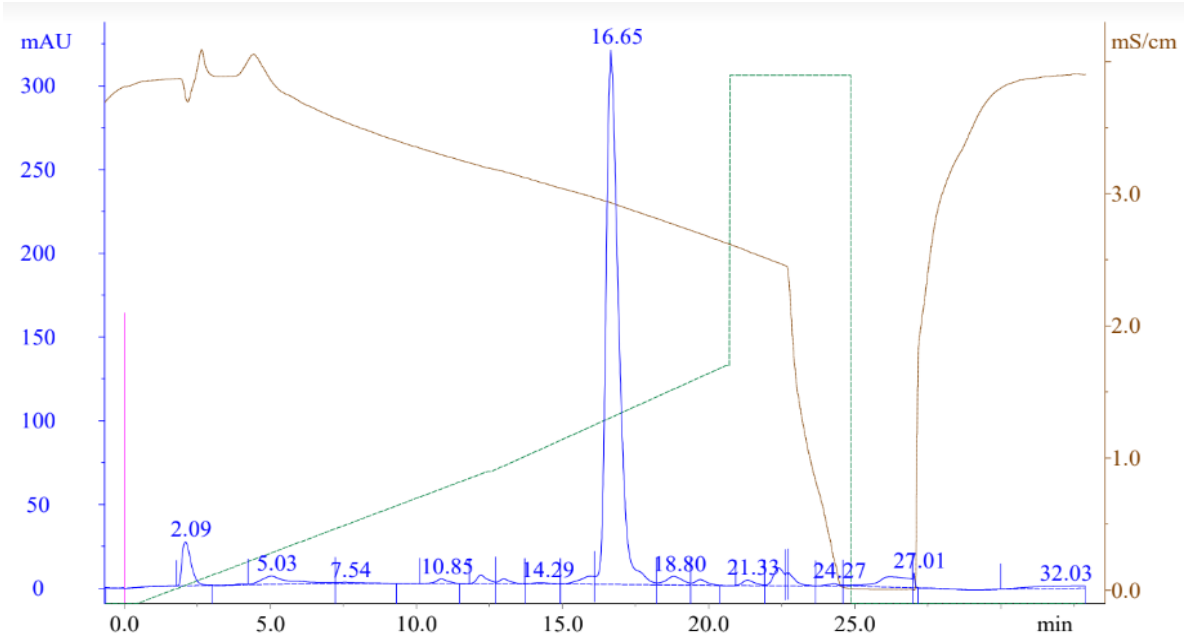


***Figure 83**. chromatogram of reaction crude (overnight) of AON51-C12-UDC. AON51-C12-UDC is the peak at 16.65.*

The product was purified on a large scale using a HiScale™ 26/20 column packed with SOURCE 15RPCresin, following the same method that was used for the analytical control of this reaction. The fraction corresponding to that peak was collected, and the solvent was evaporated using a lyophilizer.

Mass spectrometry confirmed the identity of the desired product MS (ESI, ES^-): calculated 7632,04; found 761,68 $[M-10H]^{10-}$, 846,65 $[M-9H]^{9-}$, 953,25 $[M-8H]^{8-}$, 1089,62 $[M-7H]^{7-}$, 1270,86 $[M-6H]^{6-}$, 1525,19 $[M-5H]^{5-}$. The product was then desalted as before to afford 0.45 μmol (3.66 mg) of product in a 73% yield form the linker isolated fractions and approximately in 4% total yield.

Synthesis of AON51-TEG-UDC

the oligonucleotide sequence was synthetised following the same procedures applied previously on a 10 μM scale, final yield was good from the software. In this case the linker was an amino modified triethyleneglycol protected with PDA (phtalic acid diamide), this novel protecting group gives the linker very favourable characteristics. Most of all, the linker is in solid form, as opposed to the viscous oils seen before and is easier to manipulate and aliquot. The linker was conjugated to the 5' end of oligonucleotide still on solid support with the phosphoramidite method, at the termination of the synthesis the support was dried and suspended in 7 ml AMA solution, heated to 60 °C for 10 minutes and filtered. The solution was lyophilized, and an analytical profile of the product was acquired.

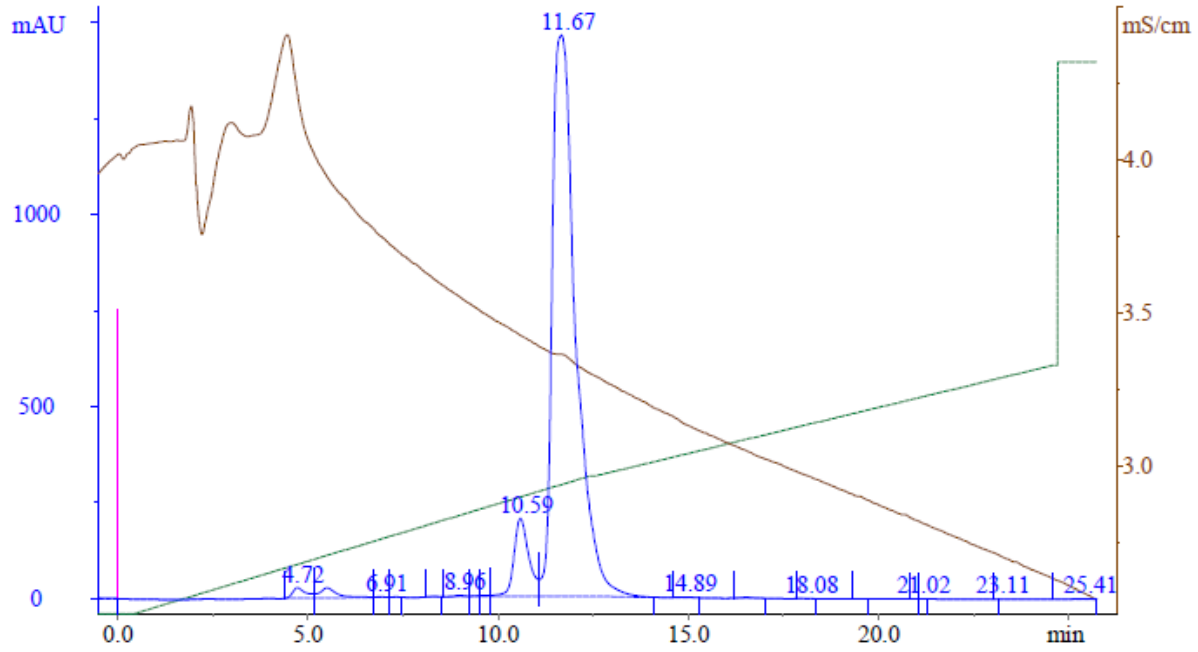


***Figure 84**. chromatogram of crude after synthesis product AON51-TEG-NH2.*

Since the analytical profile resulted quite clean, we tried to avoid one purification step and employed the crude AON51-TEG-NH2 directly in UDCA conjugation. As before the coupling reaction was set up in DMSO and with 2 equivalents of UDC-NHS and monitored at regular intervals.

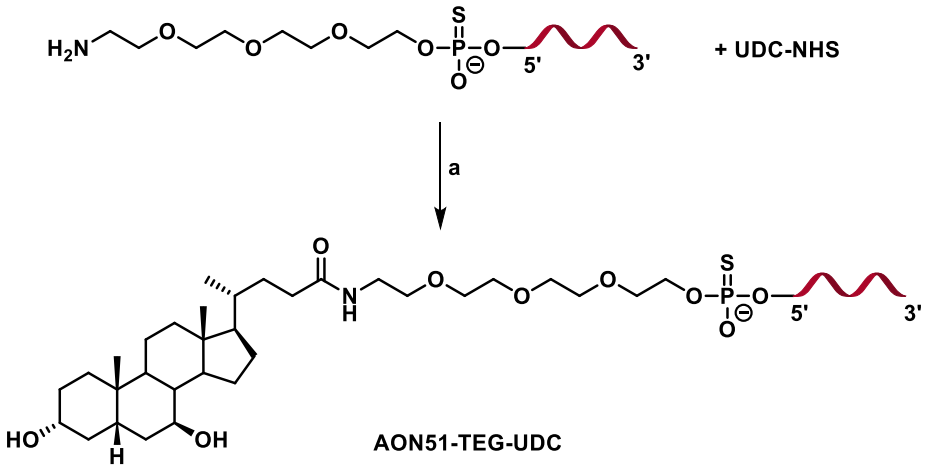


***Figure 85**. UDCA conjugation om AON51-TEG-NH2. a) UDC-NHS (30 eq), DIPEA (30 eq), DSMO, rt, 5 hours.*

However, there appeared to be no reaction occurring. After addition of a large quantity of UDC-NHS (30 eq) the conjugation was successful. We speculate that some methylamine residue could have reacted with the active ester thus blocking the reaction with the oligonucleotide.

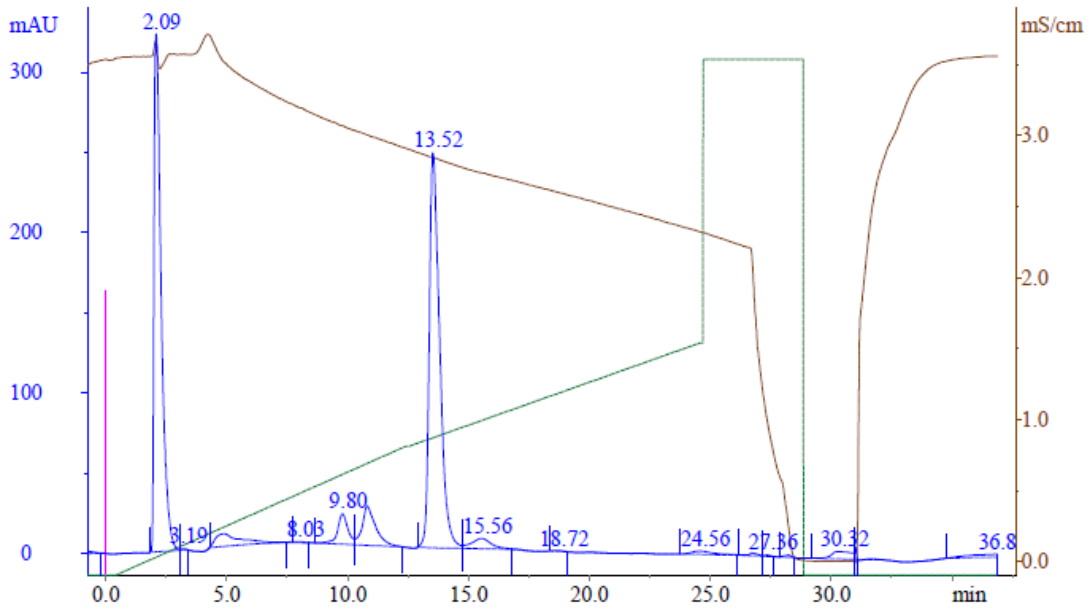


***Figure 86**. sampling chromatogram of the reaction between AON51-TEG-NH$_2$ and UDC-NHS. The peak at 13.52 min is the desired product AON51-TEG-UDC.*

After reaction the product was purified on Rsource 15RPC column as described before, the product fraction was lyophilised. Mass analysis confirmed the product identity, MS (ESI, ES$^-$): calculated 7623,93; found 761,62 [M-10H]$^{10-}$, 846,40 [M-9H]$^{9-}$, 952,31 [M-8H]$^{8-}$, 1088,05 [M-7H]$^{7-}$, 1269,55 [M-6H]$^{6-}$, 1523,56 [M-5H]$^{5-}$.the product was desalted with the help of a Hitrap Capto S cartridge to give 22.3 mg of sodium ion oligonucleotide in 29% yield (2.92 μmol).

Synthesis of AON51-ssH-CDC

The linker ssH, with a carbamate moiety used in the previous work by our group was used to obtain the conjugate AON51-ssH-UDC as a reference compound for the biological experiments on exon skipping. Moreover, the same linker was used to obtain a new conjugate with chenodeoxycholic acid. In this case, the linker was protected as monomethoxytrityl (MMt). The linker was made to react with the free 5'-OH terminus of the oligonucleotide sequence using the same protocol as previously described. Upon completion of the linker coupling cycle, the MMt group was deprotected by applying specific manual parameters as described before. This deprotection step rendered the primary amine group of the ssH linker available for the subsequent on-support reactions with chenodesoxycholic acid.

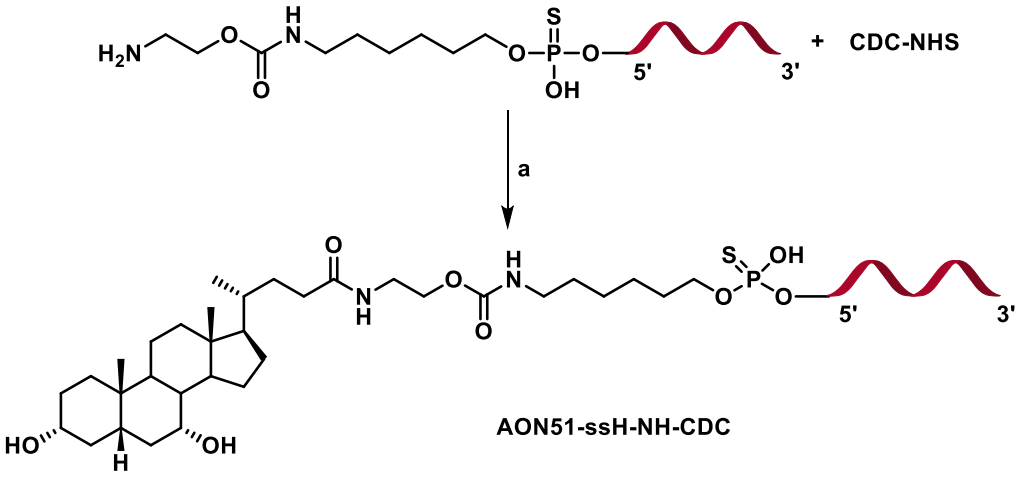


Figure 87. *reaction scheme of AON51-ssH-CDC. a) AON51-5'-ssH-NH₂ (1 eq); CDC-NHS (10 eq); DIPEA (40 eq); DMSO:ACN (1:1); rt, 3 hours.*

The reaction with chenodeoxycholic acid, CDCA, was carried out on support (Scheme 16): a cartridge was employed, within which the solid support (0.04 g, loading 80 µmol/g) linked to the hybrid composed of the newly synthesized oligonucleotide sequence and the chosen linker was weighed. This column was equipped with filters and connected at both ends to two syringes, allowing for the creation of a flow for the reaction mixture used. The reaction mixture was slowly flowed back and forth over three hours manually.

At the end of the process, the syringes were disconnected from the reactor, and the support was sequentially washed with 5 ml of DMSO and 5 ml of acetonitrile, connecting a new syringe to only one end of the reactor and removing the solvent from the other end. Afterwards, the support was transferred to a vial with a screw cap, and a volume of 7 mL of an aqueous solution of 33% ammonium hydroxide was added. The entire mixture was left to incubate in an oven at 60°C for 18 hours. The sample was diluted in a buffer solution of 0.1 M TEAA at pH 8 with 5% ACN and loaded onto a Resource RPC 3 mL reverse-phase column to get the analytical chromatogram reported below.

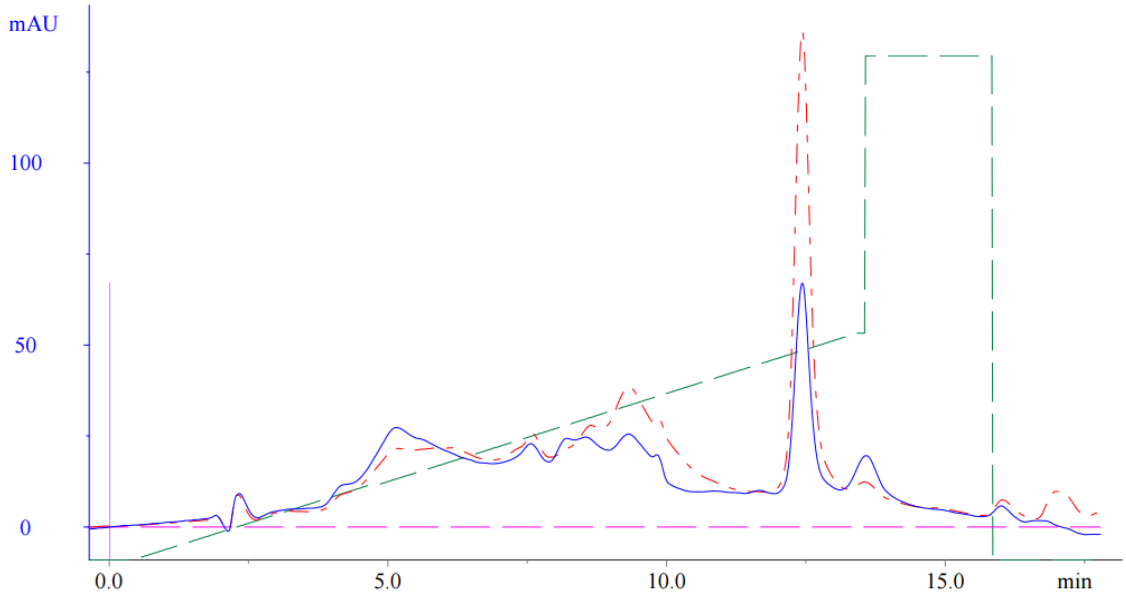


Figure 88*. chromatogram of crude AON51-ssH-CDC. peak at Rt= 18.80 ml is the desired product.*

The product was purified as usual collecting the fraction with the retention time (Tr) corresponding to the isolated product in analytical HPLC. Mass spectrometry analysis confirmed the identity of the product: calculated 76333,67; found 847,28 $[M-9H]^{9-}$; 953,18 $[M-8H]^{8-}$; 1089,37 $[M-7H]^{7-}$; 1271,18 $[M-6H]^{6-}$; 1525,82 $[M-5H]^{5-}$. The counter-ion (TEA^{+}) was exchanged *via* ion-exchange chromatography with sodium cations using a HiTrap Capto S column (Cytiva). A total of 10.6 mg (1.3 μmoles) of the product was obtained in 60% yield.

Biological tests

Exon skipping efficiency assays were executed in Matteo Bovolenta's laboratory to test the influence of the various linkers on the oligonucleotide-bile acid conjugate. The experiment was performed on immortalized myoblasts obtained from a patient with exon 52 deletion and differentiated into myotubes. The myotubes were treated with 2 μg of oligonucleotide using 4 μl of JetPEI as a transfecting agent. When a transfecting agent is employed in the experiment virtually all oligonucleotides can cross the cellular membrane, due to formation of pores given by the cationic polymer. In such an experiment therefore, there is no effect due to the cellular uptake and the differences in exon skipping are only due to the AON interaction with the RNA and relative splicing proteins, and in part from the crossing of the nuclear membrane.

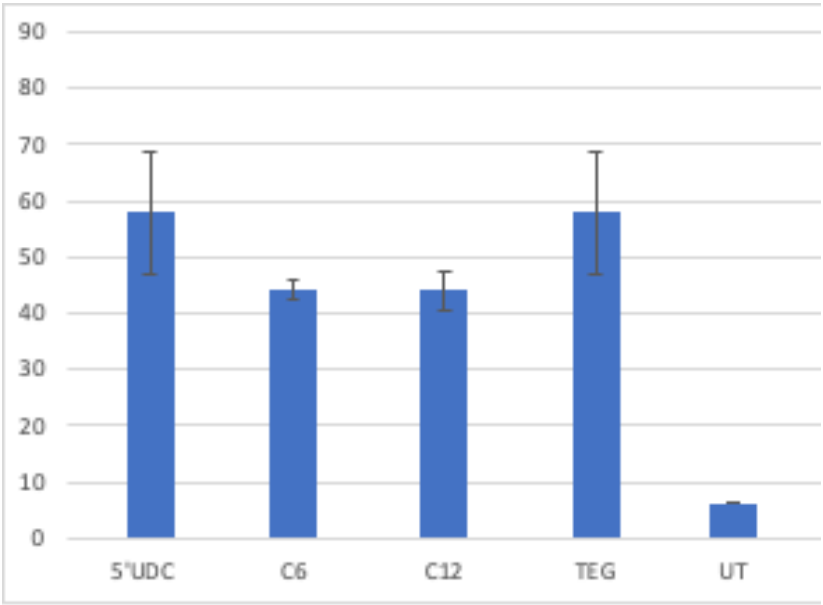


***Figure 89**. Exon skipping efficacy results of the various oligonucleotides tested. The bars indicate the standard deviation of the obtained data. The efficacy is expressed as a percentage of skipping, where 100% would give only skipped RNA.*

The graph illustrates the percentages of skipping in different cell samples treated with AON51-UDC obtained using various linkers (C6, C12, TEG). AON51-ssH-UDC treated sample and the untreated sample are used as references. Although these are preliminary results is possible to see that the original ssH linker already had good characterises in terms of length and lipophilicity which allowed a good level of exon skipping. Linker C6 is shorter than ssH and more lipophilic and gives a slightly smaller exon skipping percentage (45% against almost 60%); this could be due to a more difficult interaction with the slicing proteins. An almost identical result is given by incorporation of C12 linker unit. The C12 linker and ssH have similar length but differ in hydrophilicity, because the ssH linker possess a carbamate group which makes the chain more water-soluble. Again, it could be possible that the alkyl chain linked to an

already lipophilic bile acids give a conjugate too lipophilic to have fruitful interaction with the proteins of the spliceosome. When the TEG linker is considered however the situation changes and the exon skipping efficacy raises to practically the same level as the ssH linker. Since the ssH and the TEG have almost the same length, is possible that polarity of the chain is a strong influencing factor of exon skipping efficacy. Indeed, both the triethylen glycol chain and the ssH linker are more water soluble that the C12 linker while having approximately the same length.

These results, however, are still preliminary and other experiments to confirm this behaviour will be done in the future. Moreover, other conjugates such as the already synthesises AON51-CDC will be tested for their efficacy. It will also be important to assess the efficacy in gymnotic experiments, in absence of transfecting agents, to ascertain if the cellular uptake id influenced by the different linkers.

Conclusion

The work presented involved the design and synthesis of new candidates as potential modulators of human Dystrophin gene splicing. These molecules are conjugates of a modified antisense oligonucleotide, called AON51, linked at the 5' end to a linker of varying length and polarity, which is further linked via an amide bond to bile acids UDCA and CDCA.

Using an automatic oligonucleotide synthesizer, the synthesis of the oligonucleotide sequence was performed, and the coupling of the chosen linker was executed. UDCA was derivatized as N-succinimide esters for the subsequent conjugation reaction. Conjugations to obtain derivatives AON51-C12-UDC, AON51-TEG-UDC and AON51-C3-UDC occurred in solution, although for the latter it was not possible to isolate a significant amount of the product. While for derivatives AON51-C6-UDC, AON51-ssH-UDC and AON51-ssH-CDC the on-support approach was chosen.

Additionally, derivative AON51-C3-UDC does not exhibit a sufficient degree of purity to conduct biological tests. All obtained derivatives were purified using RP-HPLC to remove unwanted sequences and reaction by-products. Subsequently, ion exchange was performed to ensure the presence of sodium ions as counterions for the phosphorothioate group.

All conjugates were characterized via MS-ESI to confirm the presence of the expected product. Following the ion exchange, the products obtained were sent to laboratories where in vitro tests were conducted on myotubes of myoblasts from a Duchenne muscular dystrophy patient with an exon 50 deletion from the dystrophin gene.

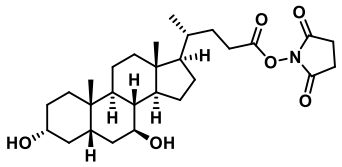
The preliminary results obtained indicate an influence of the linker on the overall exon skipping efficiency, the differences attributed of the various linkers could be explained by their hydrophily; a hydrophilic linker seem to facilitate exon skipping. However, more tests will be needed to confirm this behaviour and to assess difference in cellular uptake.

Materials and methods

Commercial EDCI and HBTU were purchased from Carbosynth (Compton, Berkshire, UK). UDCA and CDCA were kindly furnished by ICE SpA (Reggio Emilia, Italy). Protected nucleobase amidites were purchased from ChemGenes (Billerica, MA, USA). All the chemicals were used without further purification. The reactions were monitored by TLC on pre-coated silica gel F254 plates (thickness 0.25 mm, Merck), developed with phosphomolybdic acid solution. NMR spectra were recorded with a Varian Mercury 300 MHz instrument in the stated solvent. Preparative and analytical IP-HPLC was executed on Resource column with a mobile phase composed of triethylammonium acetate and CH_3CN in the stated

gradient at room temperature. mass spectra were executed on a *Thermo Finnigan LCQ Duo Ion Trap* instrument. Anhydrous solvents were freshly distilled on sodium/benzophenone with standard procedures or purchased anhydrous.

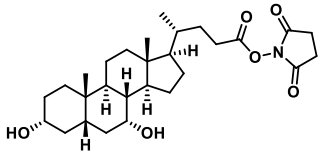
Synthesis of UDC-NHS



5 grams of ursodeoxycholic acid (12.74 mmol; 1 eq), 2.2 grams of N-hydroxysuccinimide (19.1 mmol; 1.5 eq), and 4.883 grams of EDCI (25.72 mmol; 2 eq) were weighed and placed in a 100 mL round-bottom flask. An anhydrous atmosphere was created using a vacuum pump, and the flask was purged with nitrogen. The substances were dissolved in 40 mL of anhydrous CH_2Cl_2, resulting in a suspension. The mixture was stirred for two hours at room temperature, and the formation of a clear solution indicated the completion of the reaction. Control checks were performed using TLC until the starting material spot disappeared, using 100% EtOAc as the mobile phase. The solution was transferred into a separatory funnel and washed with three successive portions of 15 mL each of saturated aqueous NaCl solution. The organic phase was collected, dried with anhydrous Na_2SO_4 and filtered through cotton. The solvent was removed under reduced pressure using a rotary evaporator. This yielded 6.223 grams of crude reaction product as a white, solid. 500 mg of the obtained solid were dissolved in the smallest possible amount of boiling toluene and allowed to slowly cool at 4°C. The mixture was filtered through Gooch filter (porosity 3) and left to dry at room temperature under a fume hood. This process yielded 283 mg of round crystals with dimensions of 1-2 mm. The crystallization yield was 57%.

^{1}H NMR (300 MHz, $CDCl_3$), selected data: δ 3.70 – 3.50 (m, 2H, H3, H7), 2.83 (s, 4H, 2 CH_2 NHS), 2.73 – 2.44 (m, 2H), 2.08 – 1.74 (m, 6H), 1.72 – 1.55 (m, 5H), 1.54 – 1.39 (m, 10H), 1.31 (dd, J = 23.0, 12.6 Hz, 5H), 1.08 (dd, J = 23.7, 14.6 Hz, 2H), 0.99 – 0.92 (m, 6H), 0.69 (s, 3H, H18).

Synthesis of CDC-NHS



5 grams of chenodeoxycholic acid (6.37 mmol; 1 eq), 2.2 grams of N-hydroxysuccinimide (9.55 mmol; 1.5 eq), and 4.883 grams of EDCI (12.85 mmol; 2 eq) were weighed and placed in a 100 mL round-bottom flask. An anhydrous atmosphere was created using a vacuum pump, and the flask was purged with nitrogen. The substances were dissolved in 40 mL of anhydrous CH_2Cl_2, resulting in a suspension. The mixture was stirred for two hours at room temperature, and the formation of a clear solution indicated the completion of the reaction. Control checks were performed using TLC until the starting material spot disappeared, using 100% EtOAc as the mobile phase. The solution was transferred into a separatory funnel and washed with three successive portions of 15 mL each of saturated aqueous NaCl solution. The organic phase was collected, dried with anhydrous Na_2SO_4 and filtered through cotton. The solvent was removed under reduced pressure using a rotary evaporator. This yielded 3.1 grams of crude reaction product as a white, solid. 300 mg of the obtained solid were dissolved in the smallest possible amount of boiling toluene and allowed to slowly cool in at 4°C. The mixture was filtered through Gooch

filter (porosity 3) and left to dry at room temperature under a fume hood. This process yielded 140 mg of small crystals in 60% yield.

^{1}H NMR (300 MHz, $CDCl_3$), selected data: δ 3.85 (m, 1H, H7), 3.48 (m, 1H, H3) 2.83 (s, 4H, 2 CH_2 NHS), 2.73 – 2.44 (m, 2H), 2.08 – 1.74 (m, 6H), 1.72 – 1.55 (m, 5H), 1.54 – 1.39 (m, 10H), 1.31 (dd, J = 23.0, 12.6 Hz, 5H), 1.08 (dd, J = 23.7, 14.6 Hz, 2H), 0.99 – 0.92 (m, 6H), 0.69 (s, 3H, H18).

Oligonucleotides conjugated to thiazole-orange for mismatch recognition and improved cellular delivery

Introduction

Oligonucleotides conjugated with fluorophores

Oligonucleotides (ON) have been conjugated with fluorophores for various reasons, such as cellular localization experiments, binding assays in cells, detection of G quadruplex or other DNA structures both *in vitro* and *in cellulo*; detection of small molecules or ions through aptamers or specific sequences; in PCR product detection and also as a safer alternative to radioactive labelling for various purposes, some of which listed above. In recent years, following the completion of the Human Genome Project, there has been an ever-increasing demand for fast and accurate detection and quantification of DNA and RNA sequences. For this aim the hybridisation of a small synthetic oligonucleotide with a nucleic acid target represents a powerful approach. The use of fluorescent labels allows to combine the specificity of the oligonucleotide for its target with the sensitivity of the fluorescence reading. Over the years probes, such as Molecular Beacon (MB)[383] Scorpion Primers,[384] and Hybeacon[385,386] have been developed for this purpose. Another advantage of fluorescence labelling is the possibility of utilizing the reading in both solution and solid phase, and even inside the cellular environment. Labelling can be done in different ways and with different families of fluorophores, so there exist different classes of fluorescent oligonucleotides based on hybridisation methods.[387] There is a great diversity in oligonucleotide fluorophore conjugates due to different parameters, such as the nature of the label and the different binding interaction with the target. From the subtle changes in fluorescence and the knowledge of the different mechanisms involved is possible to gain information on structure, mobility, distances, macromolecular size, conformational rearrangements, ligation, digestion, or cleavage of labelled ONs as well as on their binding to different ligands.

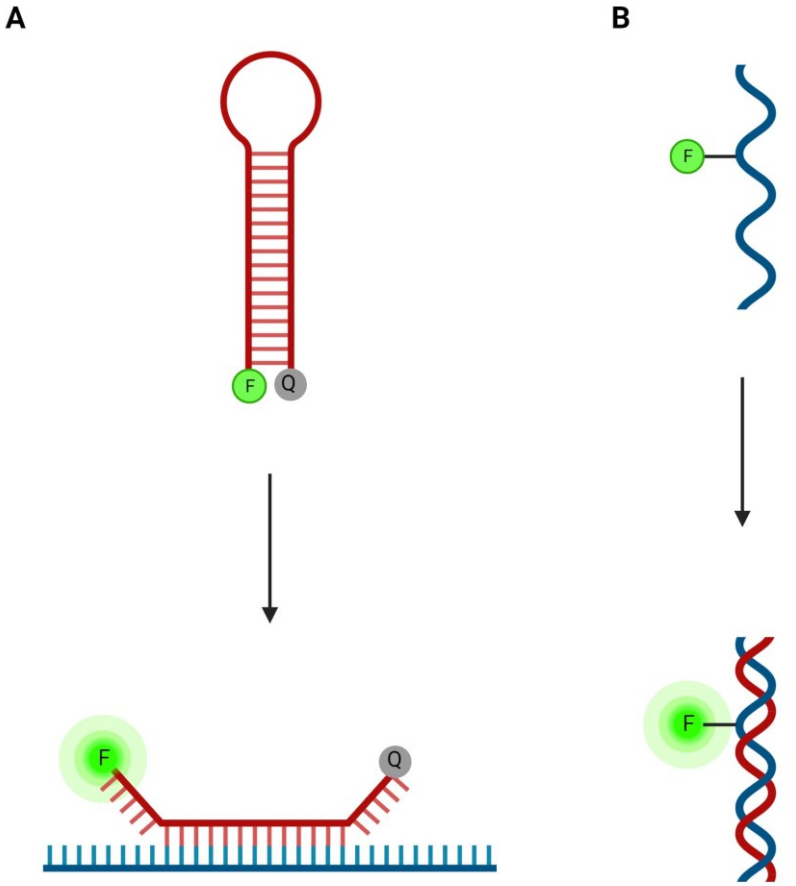
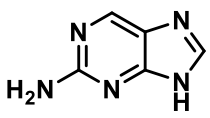
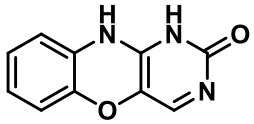


***Figure 90**. A) Molecular beacon: in analyte absence the fluorophore is quenched. When the target sequence is added the hairpin conformation is lost, fluorophore and quencher are separated, and a signal is detected. B) Hybeacon probe: the fluorophore is inactive in the single strand probe. When target strand is added the fluorophore emits a signal.*

One way to label oligonucleotides is to substitute a nucleobase with a fluorescent analogue. This approach allows for the creation of sequences sensitive to the duplex and triplex environments.[388–390] Theoretically, since the label is attached in a specific site, they should be able to discriminate between nucleobases. This capacity, however, is dependent on the label's ability to form a specific base pair. Over the years, many nucleobase analogues have been produced. The most commonly used is 2-aminopurine as it forms a stable base pair with cytosine and thymidine.[391] 1,3-diaza-2-oxo-phenothiazine, a tricyclic cytosine analogue is instead able to selectively pair with guanine and adenine. Other labels do not have this ability although they possess a good fluorescence in terms of quantum yield. Generally, the fluorescence intensity of the label is high when the oligonucleotide is alone in solution and drops when there is binding with the target, such as the case of pteridines, for example.[392,393] In another case, the fluorescence remains high when the double strand is in place and decreases upon the insertion of another complementary sequence, providing sensitivity when a triplex is formed.[394]



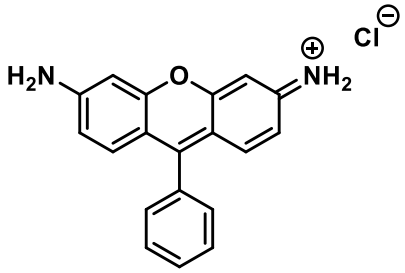
112 2-aminopurine



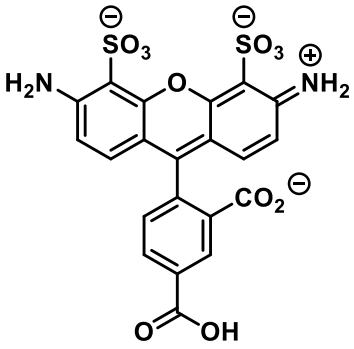
113 1,3-diaza-2-oxo-phenothiazine

Figure 91. *Structure of two un-natural nucleobases.*

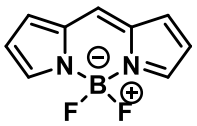
Naturally, it is possible to conjugate fluorophores to oligonucleotides ends. There are various commercially available fluorescent molecules including fluorescein, rhodamine, Alexa Fluor 488, BODIPY, and various cyanine dyes, all conveniently modified to ensure rapid and high-yield conjugation with either activated ester, azide, or amine groups. The fluorophore can be chosen based on various factors, for example fluorescein is pH sensitive and although prone to photobleaching has been used in many applications, such as real time PCR.[395] Rhodamine is less subject to bleaching and is employed similarly. BODIPY, Alexa fluor and cyanine dyes series are available and cover much of the visible spectrum, in addition they are not sensitive to solution pH. These fluorophores are commonly used to label oligonucleotides post synthetically.



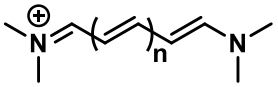
114 rhodamine



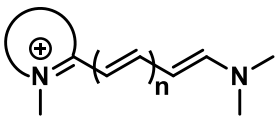
115 Alexa Fluor 488



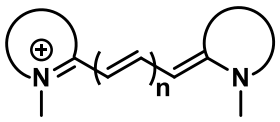
116 bodipy



Generic streptocyanine



Generic hemicyanine



Generic cyanine

Figure 92*. Structure of fluorophores commonly used on oligonucleotides.*

Another important class of fluorophores is represented by molecules able to intercalate in double stranded nucleic acids. Among these molecules are listed ethidium bromide, pyrene, perylene, acridine and. These dyes are commonly used in staining of double stranded nucleic acids as they are retained inside the helix. When such dyes are conjugated to oligonucleotides through flexible linkers, fluorescence changes occur upon hybridization with a complementary strand. Some can even reveal triplexes when conjugate to the 5' end of oligonucleotides.[387] For most of these dyes the fluorescence signal is quenched on duplex formation, but for some the effect is opposite. For example, ethidium bromide fluorescence increases when it binds to DNA, probably due to the protection of the amine hydrogens from solvent exchange when intercalation occurs.[396] The fluorescence of acridine instead only increases when intercalating between adenine and thymine and is quenched when near guanidine. [397] Interestingly, while most of cyanine dyes do not intercalate in nucleic acids the monomethine bridged ones can do so efficiently. They are highly fluorescent when they interact with the double-stranded DNA target. It has

been proposed that the strong fluorescence enhancement of these monomethine cyanine dyes upon binding to DNA is caused by a reduced degree of rotation around the internuclear bridge formed by the benzoxazole and quinoline rings.[398–400] The most studied and understood fluorophores of this class are thiazole orange (TO) and oxazole yellow (YO).

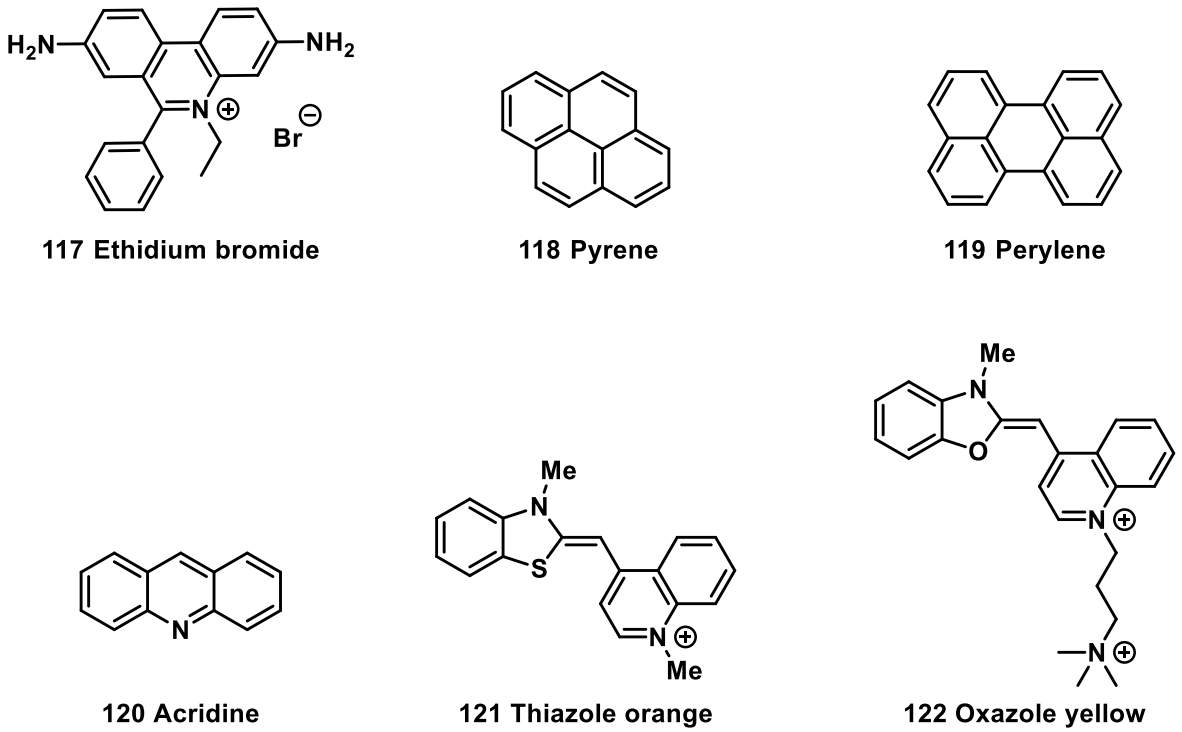


***Figure 93**. Structures of intercalating dies. Thiazole orange and oxazole yellow, are both hemicyanine dyes.*

Thiazole orange

As mentioned above, Thiazole orange is an asymmetric cyanine dye with a monomethine bridge connecting two different heterocyclic moieties, a benzothiazole and a quinoline; with an absorbance maximum at 510 nm. The double bond can thus assume a cis or trans configuration and its rotation allows a rapid non-radiative decay when the molecule is dissolved in water. When the rotation is limited in some way, by viscosity, aggregation, or as in this case, by intercalation in a double stranded nucleic acid, the luminesce is greatly enhanced due to forced planarity of the molecule. This turn-on mechanism gives a thousand-fold increase in fluorescence and a high signal to noise ratio which allows its use as a staining dye in gel electrophoresis of nucleic acids. Although other staining dyes, such as ethidium bromide and DAPI (4',6-Diamidino-2-phenylindole) are commercially available, TO has the advantage of absorbing visible light, thus avoiding all the problems associated with UV irradiation in cells, like DNA damage.

However, thiazole orange shares with the other dyes the drawback of being unselective, and for this reason is has been conjugated to oligonucleotide strands which can target it specifically to a certain sequence, similarly to molecular beacons. Molecular beacons consist of a hairpin forming oligonucleotide with a conjugate fluorophore at one end and a quencher at the opposite end. When the sequence is alone in solution the hairpin shape allows fluoresce resonance energy transfer (FRET) to quench the fluorescence signal by non-radiative decay. Conversely, a signal is obtained when the oligonucleotide anneals with a complementary strand separating the quencher from the reporter dye. Typically, the system

employs fluoresceine or other commercial dyes depending on the wavelength needed for the analysis.[401] The use of thiazole orange in a similar fashion gave rise to Light-Up probes. These probes have the advantage of providing a lower background signal since there is no need for a quencher molecule, which can sometimes have its own signal or may not completely suppress the reporter signal. Initially these probes were constructed with a peptide nucleic acid (PNA) backbone and a flexible linker by Kubista and colleagues.[402,403] The PNA being a neutral backbone with improved binding affinity for DNA, gave good fluorescence increase by permitting strong binding, thus allowing good intercalation of TO, and preventing intramolecular interaction of phosphate groups with the positively charged thiazole orange. The fifty-fold increase in fluorescence obtained with this system allowed its use in PCR product detection.[404] However, the fluorescence intensity of thiazole orange is inversely proportional to temperature, thus reducing the possible use in DNA melting applications.[405]

Another way of conjugating thiazole orange to an oligonucleotide is to substitute a nucleobase with it. With this approach the hope is to impart more specificity for the desired target by avoiding binding to mismatched sequences. Seitz and colleagues developed a probe in which thiazole orange is linked to the PNA skeleton through a short linker by the quinoline moiety, resembling a nucleobase.[406] This kind of probes were termed forced intercalation probes (FIT-probes).[407] As previously, the fluoresce intensity of the probe increases with upon binding to the target, this time by twenty or thirty-fold, but with sensitivity to single mismatches. Indeed, FIT-probes bind better to fully complementary targets than to DNA containing mismatched bases adjacent to the thiazole orange incorporation. It is believed that mismatched bases allow more space for torsional quenching of TO.[408] However, the nature of the nucleobases adjacent to thiazole orange influences the quantum yield. As a trend, guanine quenches the excited state by acting as an energy transfer receiver, giving no useful radiation in return. But time resolved measures can be used to circumvent this issue. In addition, although PNA based probes can be good choices for diagnostic applications, their chemistry is expensive and not widely available, they are also more hydrophobic than DNA based oligonucleotides; this can lead to aggregation and unspecific binding in biological mixtures and *in vitro* applications. Compared to PNA the DNA oligonucleotide more closely resemble natural nucleic acids, this facilitates the probe use in cellular experiments. As mentioned above, the presence of negatively charged phosphate group can however increase the background signal of thiazole orange.

Thiazole orange can also be conjugated to an internal nucleobase as exemplified in ECHO probes. In this approach two thiazole orange units are conjugated to the same uridine nucleoside. The quenching effect of the free TO in solution enhanced by dimer formation due to an exciton coupling effect that occurs upon parallel aggregation of the two units (H-aggregation). This further reduced emission in solution remains also when a single oligonucleotide strand is constructed. However, upon hybridization with the target the two thiazole orange units intercalate and emit their fluorescence signal. In this way a 160-fold increase in fluoresce is possible, greatly enhancing the signal delta obtained with single thiazole orange incorporation at the same nucleobase.[409] In this case DNA and 2'OMe RNA backbone is used.

Pyrene

Pyrene is a very interesting molecule, widely available in modified forms, such as di- and monocarboxylic acids, NHS ester, hydrazide, alcohol, maleimide but also as an amine, azide and alkyne bearing and even phosphoramidite, often with both short or long polyethylene glycol linkers.[410,411] Its fluorescence properties have been thoroughly studied as is very sensitive to the microenvironment and finds application in a variety of fields such as sensors, polymer chemistry, nanotechnology, cell membrane and protein studies and, of course, nucleic acid research.[412–416] Pyrene has a surface area similar to that of the

natural Watson-Crick base pairs (pyrene ≈ 184 Å^2, A:T base pair ≈ 223 Å^2).[417] Therefore, the molecule has a tendency to π-stack with other aromatic units, base pairs or another pyrene, and can intercalate in nucleic acid duplexes. In particular, pyrene shows better intercalation in DNA duplexes than in RNA ones due to a more compressed structure and less efficient nucleobase overlap in the latter.[418–420] The attention gained by pyrene in so many fields is partly due to its stability and availability, but mostly to sensitivity and ability to form excimer bands upon stacking with other pyrenes. Although the fluorophore has five emission maxima, normally, only vibronic emission bands I and III can be seen due to signal broadening. The ratio between these bands provides information on the polarity of the environment, higher I_{III}/I_I ratios are observed in apolar solvents. When intercalated, its emission can be quenched by neighbouring nucleobases, as with thiazole orange, guanine is the strongest in this regard, followed by cytosine and thymine to similar extent and adenines as the worst quencher.[421,422] Given this knowledge, a higher fluorescence signal is to be expected when pyrene is in polar grooves than when intercalated in GC-rich duplexes. Broad, red-shifted, solvent-dependent, excimers bands form when two pyrenes are stacked at the optimal distance of ≈ 3.4 Å; thus, construction of systems exploiting these bands is beneficial for *in vivo* application. Furthermore, the fluorescence properties of pyrene can be modulated by extension of its π-conjugation, for example by coupling a pyren-yl moiety directly to the 5 position of uracil nucleobase, through Sonogashira coupling reaction.

Similarly, to thiazole orange, pyrene can be incorporated into an oligonucleotide by substitution of a nucleobase, at the oligonucleotide ends or by modification of the sugar or nucleobase. The attachment point and length of the linker can influence the characteristics of the pyrene label. When pyrene is used as a nucleobase substitute the duplex is either destabilized or stabilized in respect to the normal control sequence, depending on the way it is attached. Pyrene modification is stabilizing if it is conjugated to uridine through a long flexible linker in 5 position of uracil (up to +7.1 °C),[423] but not if conjugation is achieved by a short linker (-4.6 °C).[424]

Locked nucleic acids

locked nucleic acids (LNA) were independently invented in the 1990' by Wengel[425] and Imanishi[426] groups. this modification consists of a ribose with a C2'-C4' oxymethylene bridge that effectively locks the furanose ring in a C3' *endo* conformation, thus resembling the conformation typical of the sugar moiety in RNA duplexes.[427] Locked nucleic acids are therefore considered an RNA analogue, with the advantage of having a protected 2' hydroxyl group and making the oligonucleotide much more stable toward nuclease degradation. In addition, the bicyclic LNA skeleton restricts the glycosidic torsion angle because of steric clash between the H6 position of pyrimidines, or H8 position of purines, and the H3', which is locked in pseudo-axial position, due to the endo puckering of LNA.[428] These special characteristics of the locked nucleic acids confers preorganisation also to adjacent nucleotides, making them acquire an increased N-type conformation in both the single strand and duplex oligonucleotides with DNA and RNA.[427,429,430] This in turn, confers the duplex more A-type character and increases the thermal affinity of LNA oligonucleotides because the restrictions imposed by the locked nucleosides affects also neighbouring nucleosides and mechanically couple them, dampening the overall duplex dynamic equilibrium between N-type and S-type nucleosides.[431] Increasing the number of LNAs in an oligonucleotide and studying their influence on RNA structure reveals that the structural changes affect the neighbouring bases, especially toward the 3' end of the strand. A saturation of the effect is reached before complete modification of the strand, with the maximum thermal stability observed for less than 50% modification of the nonamer studied. Indeed, the maximum thermal stability, giving A-type LNA:RNA duplex, was found for less than 50% modification of the nonamer studied.[432] LNAs are

known to increase the melting temperature (T_m) of double helix nucleic acids by 3-9 °C for RNA targets and from 1-8 °C for DNA, compared to non-modified sugar backbones. They also exhibit higher specificity, and better mismatch discrimination.[432,433] For these reasons LNAs have been extensively studied in research, especially in gene modulation, since a better affinity for the target sequence means a possibly shorter therapeutic oligonucleotide to be synthesized; thus, ameliorating the delivery issues associated with oligonucleotides.[434] In this regard the antisense approach could draw benefit from the use of LNA mixmers,[435,436] but LNA have also been used to protect the ends of gapmer polynucleotides for the RNase-H gene modulation approach.[319] LNA have also found good applications on diagnostics, for example this modification is applied in PCR probes and gives better efficiency by lowering the cycle threshold required for analyte detection. LNA probes and primers can also increase amplification specificity, giving a better reading quality. In addition, thanks to the increased melting temperature of LNA containing duplexes is possible to reduce the amount of the primers used by tenfold.[436] For the same reasons LNA has been used extensively for single nucleotide polymorphism (SNP) identification.[437] indeed, LNA exhibits superior single-nucleotide mismatch discrimination compared to DNA; thus, the melting temperature difference between a perfect match and a single-nucleotide mismatch is greater for LNA oligomers than it is for DNA oligomers.[432]

LNA modification has also been used in the tentative to ameliorate the duplex destabilizing effect of pyrene incorporation at the nucleobases. Since LNA modification increases the duplex melting temperature even when non-aromatic moieties are linked to C5 position of uridine, the authors speculated that this effect could hold true for pyrene functionalized LNA.[438] Unfortunately, the melting temperature obtained for the modified oligonucleotides was comparable to the DNA analogue.[437]

Discussion

The versatility of thiazole orange has also been exploited for the construction of combination probes.[439] These probes aim to combine the most favourable feature of light.up probes seen above, and HyBeacon probes.[440,441] Single-stranded hairpin oligonucleotides with two or more fluorophores covalently bonded to internal nucleotides are known as HyBeacon probes. Elevated fluorescence is detected upon hybridization of these probes with complementary DNA targets, as opposed to low fluorescence when they are single-stranded, because of collisional quenching. These probes did not employ thiazole orange, but fluorophores such as FAM (6-carboxyfluorescein) and other modified fluoresceins. The combination of thiazole orange and a quencher dye, both internally linked into an oligonucleotide sequence, allows to more efficiently quench the fluorescence of the single stranded probe. On the other hand, in the duplex state the fluoresce of thiazole orange is enhanced by Förster resonance energy transfer (FRET) and allows monitoring at different wavelengths.

Inspired by the development of combination probes and by studies on the conjugation mode of thiazole orange to the uridine nucleoside this works proposes the development of a locked nucleic acid uridine monomer modified with an alkyne linker for the construction of oligonucleotide sequences modified with thiazole orange. The aim of this study is to investigate the impact of incorporating locked nucleic acid (LNA) and thiazole orange (TO) on the melting temperature of the resulting duplex. To accomplish this, the synthesis of the locked nucleic acid iodo-uridine was performed to allow the conjugation of a protected octadinyl unit through a Sonogashira reaction. This monomer was subsequently transformed into a phosphoramidite and used in solid phase oligonucleotide synthesis with standard protocols. After deblocking and purification by sephadex (NAP) cartridge, the modified sequences were reacted in a CuAAC protocol in solution phase with thiazole orange or pyrene azide. Having obtained the conjugated oligonucleotides. Melting temperatures experiments on DNA and RNA targets were performed, thiazole

orange incorporation on locked nucleic acid nucleoside increased the melting temperature of all tested duplexes in respect to the unmodified control ones, by 10-19 °C depending on the sequence. Additionally, melting temperature was assessed for mismatched sequences, with mismatches positioned either corresponding to the modification position or the neighbouring ones for both TO and pyrene modified oligonucleotides. Fluorescence measurements were taken for the single-stranded modified sequences and during annealing. For thiazole orange modified oligonucleotides an increase in emission was registered in each case with variations due to sequence specificity.

Monomer synthesis

As mentioned above the modified oligonucleotide's structure involves a nucleobase modification and a sugar modification on the same nucleoside. The synthesis started by following literature protocols to obtain the 5-iodo locked uridine.[442] In particular, form the commercially available compound **123** was possible to obtain **124** in good yield by replacing the acetal protecting group of 1' and 2' hydroxyl groups with aceto-ester protection. subsequently, Vorbrüggen reaction with uracil afforded the protected uridine **126**.[443] Deprotection of 2' hydroxyl and locking or the ribose was executed in one step under basic conditions to give **127** in good yield. At this stage, it became necessary remove the unreacted 5' mesyl group; this could be accomplished under strongly basic conditions or by successive substitution and cleavage of a benzoate ester. Some starting material degradation was observed when attempting the removal with sodium hydroxide, while the benzoate method gave better yield although being more time consuming.

Removal of the benzyl group was attempted using ammonium formate as a hydrogen forming reagent in refluxing methanol. However, the reaction proved much cleaner when hydrogen gas was employed at 40 °C. Once the unprotected, locked uridine **130** was obtained, iodination in 5 position was executed with cerium ammonium nitrate (CAN) and iodine in acetic acid as reported in literature. after 40 minutes the reaction reached completion, as assessed by TLC. Prolonging the reaction at 80 °C led to formation of polar impurities, lowering the yield, and causing a more difficult product purification.

After dimethoxy trityl (DMT) protection of 5' hydroxyl position it was possible to proceed with Sonogashira coupling of octadiyne. Initially the reaction was attempted with ten equivalents of octadiyne but was unsuccessful. However, after protection of one alkyne unit[444] the reaction, retried at higher temperatures, gave the product in good yield. The TMS protecting group introduced on the alkyne has been demonstrated to be conveniently cleaved during oligonucleotide deblocking as later demonstrated by mass analysis of the oligonucleotides. Compound **133** was then reacted in degassed and anhydrous conditions to give the phosphoramidite reagent for oligonucleotide synthesis.

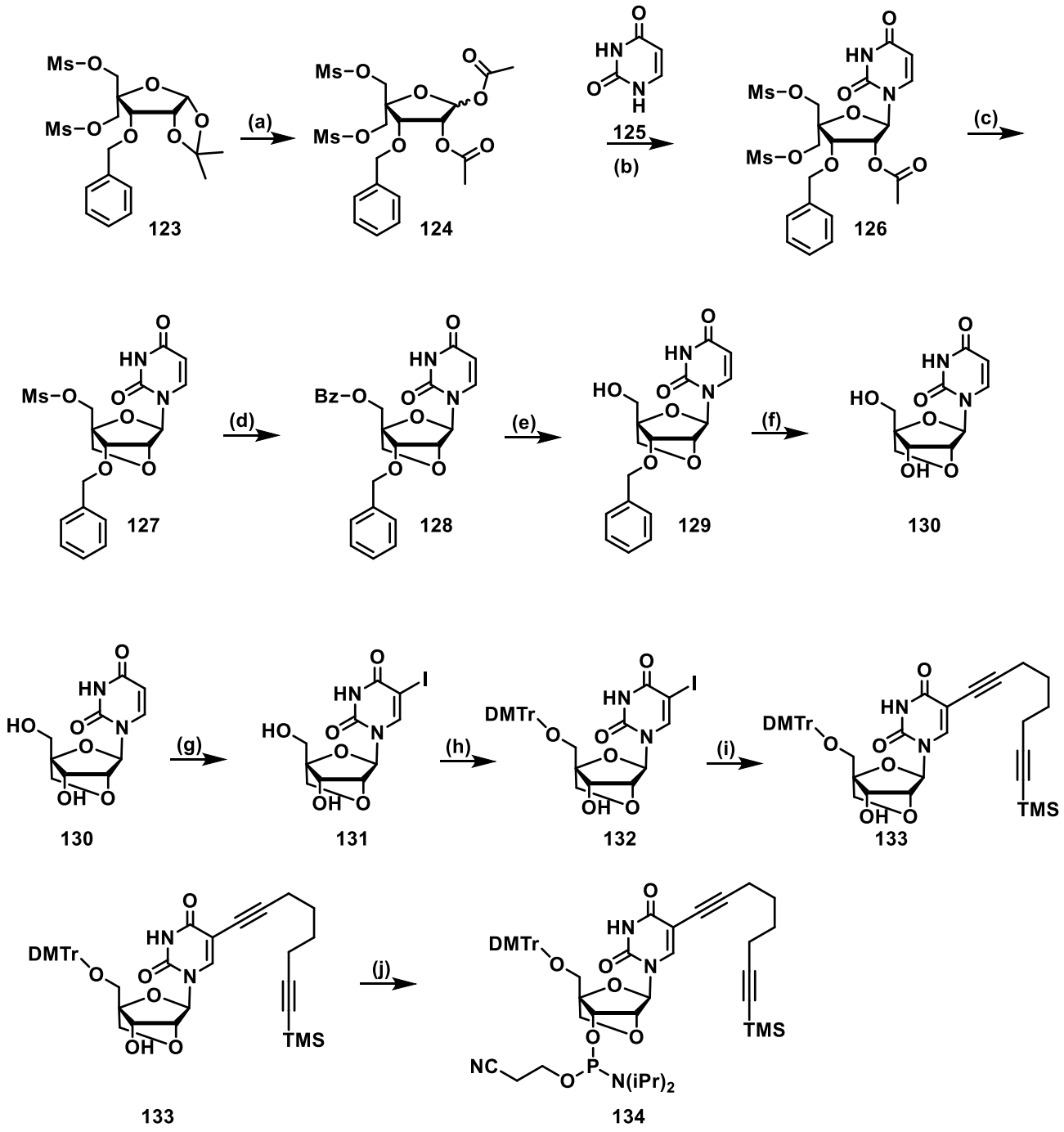


Figure 94*. (a): 1) TFA 80%, rt, 1h. 2) Ac_2O (4 eq), Py, rt, overnight. (b): uracil, BSA, TMSOTf, ACN overnight. (c): water/dioxane (1:1), NaOH 2M, rt, 1h. (d): NaOBz, DMF, 100 °C, 5h. (e): water/THF (1:1), 2M NaOH, rt, 1h. F: MeOH, Pd/C, H_2, 30 °C, overnight. (g): CH_3COOH, I_2, CAN, 80 °C, 40 min. (h): Pyridine, DMT-Cl, rt, 1h. (i): trimethyl(octa-1,7-diyn-1-yl)silane, $Pd(PPH_3)_4$, CuI, TEA, DMF, 40 °C, 1 h. (j): PCl reagent, TEA, DCM rt, 1h.*

Oligonucleotide synthesis

Oligonucleotides were synthesised with standard procedures on an Applied Biosystems 394 automated DNA/RNA synthesizer. DNA, 2'OMe phosphoramidites and solid supports were purchased from Link Technologies and Applied Biosystems Ltd. 1.0 μmol phosphoramidite cycle of acid catalysed detritylation, coupling, capping and iodine oxidation or thioxidation. Stepwise coupling efficiencies and overall yields were determined by an automated trityl cation conductivity monitoring facility and in all

cases were >98.0%. β-Cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C and T monomers was 35 s while the coupling time for locked nucleic acid was increased to three minutes. Cleavage of the oligonucleotides from the solid support was achieved by exposure to concentrated aqueous ammonia solution for 1 hour at room temperature. Deprotection was carried out with the same solution for 5 hours at 55 °C. After ammonia evaporation under vacuo, the oligonucleotides were purified by filtration on GE Healthcare Life Sciences illustra NAP 10 gel filtration columns or reverse-phase IP-HPLC on a Gilson system using a Luna 10 μm C8 100 Å pore Phenomenex 10x250 mm column. The buffer was composed of: buffer A: 0.1 M triethylammonium bicarbonate, pH 7.0 and buffer B: 0.1 M triethylammonium bicarbonate, pH 7.0, with 50% acetonitrile. Gradients were varied based on the polarity of the oligonucleotide to be purified. Typically, for non-modified oligonucleotides was used a 0% to 100% buffer B gradient over 20 min. For thiazole orange or pyrene modified oligonucleotides the gradient was 25% to 100% buffer B. Alternatively, TEAA buffer was used with the same methods; The buffer was composed of buffer A: 0.1 M triethylammonium acetate, pH 7.0 and buffer B: 0.1 M triethylammonium acetate, pH 7.0, with 50% acetonitrile. Elution was monitored by UV absorption at 260 or up to 299 nm if needed due to sample concentration. After purification, oligonucleotides were freeze-dried and in case of TEAA buffer were desalted with GE Healthcare Life Sciences illustra NAP 10 gel filtration columns. After IP-HPLC purification, all oligonucleotides were characterised by electrospray mass spectrometry using a UPLC-MS Waters XEVO G2-QTOF instrument in ESI- mode. Data were processed using MaxEnt or MassLynx v4.1. Obtained masses for the oligonucleotides are reported in Table 5 below.

General procedure for post-synthetic oligonucleotide functionalization

To a 20 nmol solution of alkyne modified oligonucleotide in 200 μL distilled sterile water were added 200 μL 2 M TEAA buffer (pH = 7.0), 288 μL DMSO and 12 μL stock solution of azide dyes (15 nmol/μL in DMSO, 7.5 eq). The obtained solution was degassed for 5 minutes by argon bubbling. 200 μL of 10 mM sodium ascorbate solution and 100 μL of 10 mM CuII:THPTA solution were added. The mixture was vortexed at 30 °C for 3 hours. Copper catalyst and ligand complex solution CuII:THPTA was prepared by mixing tris(3-hydroxypropyltriazolylmethyl)amine and $CuSO_4$ in DMSO/H_2O (55:45).

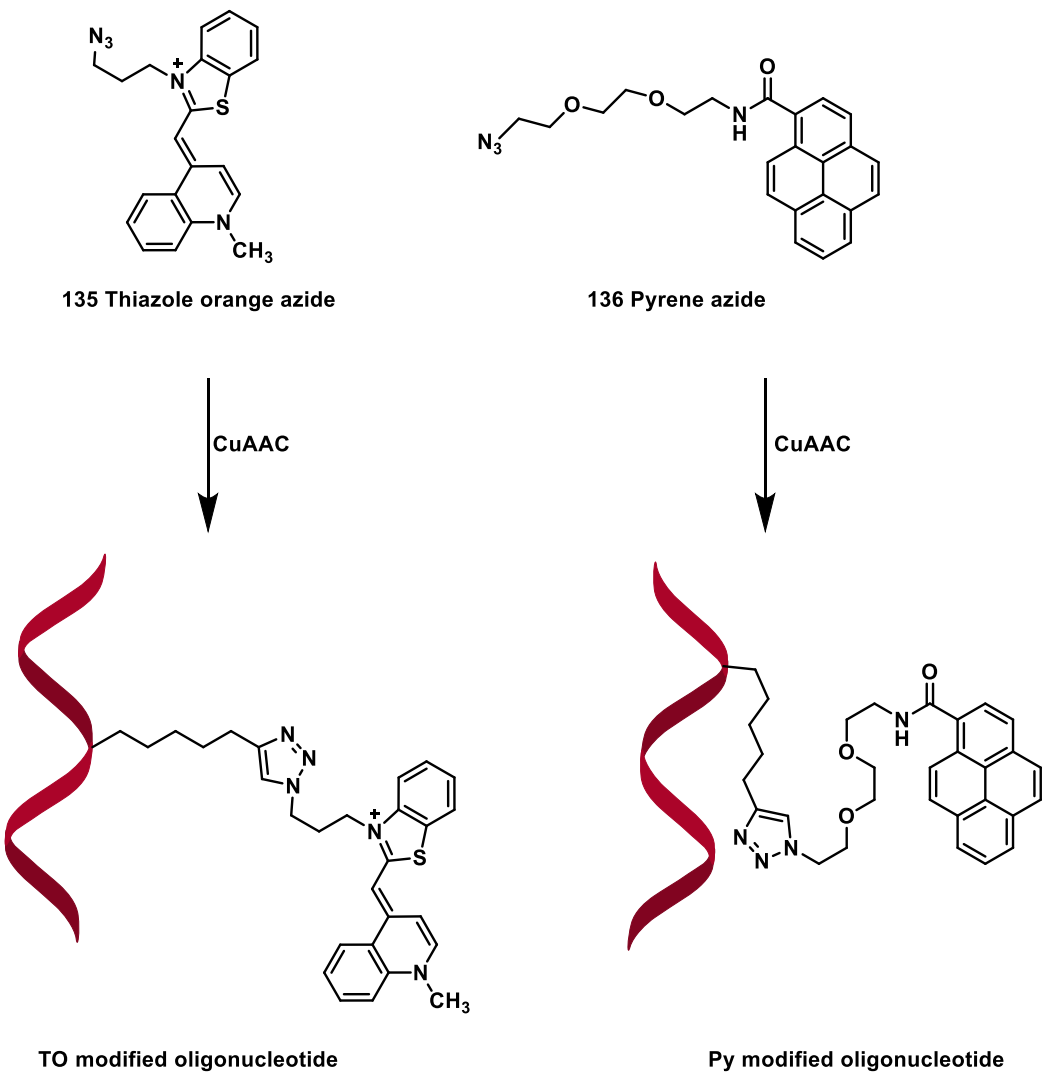


Figure 95. *Reaction scheme of oligonucleotide functionalization through click reaction.*

The mixture was desalted by NAP 10 gel filtration following the producer protocol, lyophilised, and purified by IP-HPLC to obtain the desired labelled oligonucleotide.

Table 5. Oligonucleotide sequences

Reference	Oligonucleotide sequence	Mass calculated	Mass found
AXA	GCAGAXAGACT	3483.3	3483.0
TXT	GCAGTXTGACT	3465.2	3465.5
GXG	GCAGGXGGACT	3515.1	5315.5
CXC	GCAGCXCGACT	3435.2	3435.5
TXG	GCAGCTXGACT	3450.2	3450.5
GXT	GCAGXTCGACT	3450.2	3450.5

Comp TXG	AGTCAAGTCGC	3341.2	3341.0
Comp GXT	AGTCGAACTGC	3341.2	3341.5
Control TXG	GCAGCTTGACT	3332.2	3332.5
Control GXT	GCAGTTGCACT	3332.2	3332.5
Comp Mis +1	AGTCGGACTGC	3357.0	3357.0
Comp Mis -1	AGTCGAAATGC	3365.0	3365.5
Comp Mis 0	AGTCGAGTCGC	3357.0	3357.0
TU^{TO}T	GCAGTU^{TO}TGACT	3881.2	3880.5
GU^{TO}G	GCAGGU^{TO}GGACT	3931.1	3930.5
CU^{TO}C	GCAGCU^{TO}CGACT	3851.2	3850.5
TU^{TO}G	GCAGCTU^{TO}GACT	3866.2	3865.5
GU^{TO}T	GCAGU^{TO}TCGACT	3866.2	3865.5
AU^{TO}A	GCAGAU^{TO}AGACT	3899.3	3898.5
AU^{py}A	GCAGAU^{py}AGACT	3885.3	3886.0
TU^{py}T	GCAGTU^{py}TGACT	3867.7	3868.0
GU^{py}G	GCAGGU^{py}GGACT	3917.6	3917.5
CU^{py}C	GCAGCU^{py}CGACT	3837.7	3838.0
TU^{py}G	GCAGCTU^{py}GACT	3852.7	3853.0
GU^{py}T	GCAGU^{py}TCGACT	3852.7	3853.0
RNA Mis pos +1	agucggacugc	3505.0	3505.5
RNA Mis pos -1	agucgaaaugc	3513.0	3513.5
RNA Mis pos 0	agucgagcugc	3505.0	3490
RNA comp TXG	agucaagcugc	3489.0	3489.0
RNA comp GXT	agucgaacugc	3489.0	3489.5
Luc 0	ccucuuaccucaguuaca	6096.0	6100.0
Luc 1	ccucuuaccXcaguuaca	6198.0	6202.0

Luc 2	ccXcuuaccucaguXaca	6300.0	6304.0
Luc 3	ccXcuuaccXcaguXaca	6402.0	6406.0
Luc 1 TO	ccucuuaccU^{TO}caguuaca	6614; 6738 (2Cu); 6800 (3Cu)	6741.5; 6803.5
Luc 2 TO	ccU^{TO}cuuaccucaguU^{TO}aca	7132; 7256 (2Cu); 7320 (3Cu)	7261.0; 7322.0
Luc 3 TO	ccU^{TO}cuuaccU^{TO}caguU^{TO}aca	7650; 7776 (2Cu)	7778.0

Table 5*. Sequences synthesised for this work with relative name. all masses were found to be a good match of the calculated ones and confirmed the correct identity of the oligonucleotide. Uppercase letters represent DNA amidite, RNA is denoted in lowercase letters locked nucleic acid is noted as X or U^{to} or U^{by} . Phosphate backbone was used for all oligonucleotides except for the luciferase assay sequences, which have an all phosphorothioate backbone.*

UV melting studies

UV melting measurements were made on a Varian Cary 4000 UV-Vis spectrophotometer with a Cary temperature controller. The samples were analysed in 1 mL quartz cuvettes with 1 cm path length and the absorbance was measured at 260 nm. Oligonucleotide probe concentration was made up to 3.0 μM and 3.3 μM oligonucleotide target concentration in a buffer composed of 10 mM phosphate and 200 mM NaCl at pH 7.0. The samples were initially denatured by heating to 85 °C at 10 °C min^{-1} then cooled to 15 °C at 1 °C per minute and then maintained at 15 °C for 2 min before heating to 85 °C at 1 °C per minute. UV absorption was recorded every 0.1 °C and the cycle was repeated 6 times. The melting temperature (Tm) values were obtained from the derivatives of the melting curves and calculated at 260 nm using Cary Win UV Thermal application software. Six successive melting curves were measured and averaged. All the experiments were done in duplicate.

Modification of deoxy uridine with a linker of similar length in position C5 is reported in literature, with relative study of melting temperatures and fluorescence, and in some cases this data will be referred to in order to speculate on the effect of locked nucleic acid modification.[445]

The melting temperatures obtained with thiazole orange modified oligonucleotides resulted, as expected, higher in respect to the reference duplexes. I particular the increase in melting temperature is strongly dependent on the sequence, and as can be seen in the table below there is a variation between +19.1 °C for the modification with two adjacent adenine nucleobases and +9.6 °C when the modification is enclosed by two thymine nucleobases. A similar sequence dependence could be due to π-stacking of the thiazole orange in the local environment, as the linker probably only permits directional intercalation of the thiazole orange. This however is just speculative, double strand structural analysis thorough X ray diffraction could help investigate this fact in the future. To verify the contribute of thiazole orange to the duplex stabilization the melting temperature was measured also on the LNA oligonucleotides modified only with the alkyne moiety. As mentioned above LNA modification alone has been ascertained to increase duplex stability by 1-9 °C, while the combination of C5 uridine modification and LNA ribose is reported to give a final result based on the modification nature, and in general a stabilization effect when an aliphatic moiety is conjugated. In this case the octadiynyl linker appears to give a generally stabilizing sequence specific effect, with the exception of TU^{TO}T which has a -1.3 °C melting temperature decrease in respect to the reference. From the delta in Tm differences between the thiazole orange modified oligonucleotides and the alkyne bearing ones, it appears clear however that the influence of thiazole orange on duplex stabilization is greater than that given by the locked ribose. By comparison with previously reported data on similar uridine thiazole orange modification[445] it appears that the

addition of a locked ribose is further beneficial for increasing the melting temperature of the duplexes on average.

Table 6. DNA target melting temperatures

Entry	Oligo LNA-TO	Tm reference	Tm match	Δ Tm match[a]	Oligo LNA-alkyne	Δ Tm match alkyne[b]	Δ Δ Tm Alkyne-TO
1	**AU^{TO}A**	41.2	60.3	19.1	**AXA**	8.0	11.2
2	**GU^{TO}G**	55.2	66.1	10.9	**GXG**	2.5	8.4
3	**CU^{TO}C**	59.2	73.2	14.0	**CXC**	1.8	12.1
4	**TU^{TO}T**	51.7	61.3	9.6	**TXT**	-1.3	10.9
5	**CTU^{TO}**	50.7	66.5	15.8	**CTX**	5.6	10.3
6	**U^{TO}TC**	49.1	65.6	16.6	**XTC**	2.4	14.22

Table 6. *All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) calculated by difference between the melting temperature of modified thiazole orange (U^{TO}) sequence with its unmodified DNA complementary sequence and the reference sequence obtained with unmodified DNA probe. b) calculated by difference between the melting temperature of modified alkyne (X) sequence and unmodified DNA complementary. All sequences differed only for three central nucleotides and follow the general formula: GCAG"YXY"GACT where Y represents one of the four nucleobases as indicated in the oligo entry.*

Next, the melting temperatures of mismatched sequences were evaluated. For this purpose, two modified sequences were synthesized with a mismatch in respect to the control sequence, in this instance the thiazole orange bearing nucleotide is shifted one nucleotide in 5' (CT**U^{TO}**) or one in 3' (**U^{TO}**TC) direction in respect to the normal sequence (complementary of CXC), entry 1 and 2 of table 3 respectively. There is still a considerable duplex stabilization effect due to thiazole orange incorporation, but the melting temperature obtained for TO modified duplexes is lower than that obtained before for the matched sequences. Moreover, the position influence of mismatch was evaluated. The mismatched nucleobase was placed either corresponding to the thiazole orange modification (Table 7, entry 4), one nucleobase in the 5' direction (Table 7, entry 3) or one nucleobase in the 3' direction (Table 7, entry 5). The stability is the lowest when the mismatch is positioned immediately after TO modification in the 5' direction. On the other hand, positioning the mismatch at the 3' position next to TO is less detrimental for stability. The trend for the mismatch positional influence on stability, seems to be a progressive reduction in stabilization from 3' to 5', as the melting temperature obtained when the mismatch is corresponding to thiazole orange position, falls between the ones obtained for the two cases mentioned above. However, it should be considered that the mismatch in position -1 is given by two purines and not by a purine and pyrimidine as for the other mismatches, this would naturally lead to more destabilization. By comparison of the difference in match and mismatch temperatures of the modified oligonucleotides and reference sequences, is possible to see an improvement, almost double in percentage, of mismatch discrimination upon oligonucleotide modification. This in turn, can be advantageous in avoiding potential off-target side effects during *in vitro* exon skipping test.

Table 7. DNA target mismatch melting temperatures

Entry	Oligo-LNA-TO	Tm mismatch reference	Tm mismatch DNA	Δ Tm mismatch DNA	Δ Tm match - mismatch	Δ Tm match – mismatch reference
1	**CTU^{TO}** [a]	45.9	57.8	11.9	8.7	3.4
2	**mismatch pos +1**[b]	45.7	57.5	11.8	8.1	3.4
3	**mismatch pos 0**[b]	44.9	57.3	12.4	8.3	4.2
4	**mismatch pos -1**[b]	41.4	55.3	13.9	10.3	7.7

Table 7. *All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) Melting temperature of modified thiazole orange (U^{TO}) sequence with an unmodified DNA double mismatched sequence and the reference sequence obtained with unmodified DNA probe. b) melting temperature of thiazole orange (U^{TO}) sequence and single mismatch in different positions against DNA sequence. Position 6 has the mismatch one nucleotide in 5' direction respect to TO modification, position 5 has the mismatched nucleotide on the modification, position 4 has it one nucleotide 3' of modification. ⊿ Tm match – mismatch is calculated by subtracting the melting temperature of mismatched sequences to the reference match sequence with thiazole orange.*

The same evaluations as before were done with RNA targets. This time the variation observed in melting temperatures values was less pronounced and ranged between 13.3 and 10.7 °C. As for DNA, there are sequence specific variations, for example duplex stability when TO is between two guanidine is lower than the average. However, this time the same effect is not present when the nucleobases are thymines. Again, good stabilization occurs when TO modification is encompassed by adenine of cytosine, with higher-than-average Tm values both for entry 1 and 3 in Table 8. Overall, the stabilization given by thiazole orange and locked ribose seem a bit lower in DNA:RNA than in DNA:DNA duplexes but is consistent and less dependent on adjacent nucleobases. In addition, the increase in melting temperature given just by the C5 octadiynyl locked uridine is higher than that obtained for DNA targets for almost all sequences, (except entry 6).

Table 8. RNA target melting temperatures

Entry	Oligo-LNA-TO	Tm reference RNA	Tm match RNA (°C)	Δ Tm match[a]	Oligo LNA-alkyne	Δ Tm match alkyne[b]	Δ Δ Tm Alkyne-TO
1	**AU^{TO}A**	40.8	53.8	13.0	**AXA**	7.6	5.4
2	**GU^{TO}G**	53.1	64.7	11.5	**GXG**	5.3	6.2
3	**CU^{TO}C**	56.5	69.8	13.3	**CXC**	5.2	8.0
4	**TU^{TO}T**	46.3	59.7	13.3	**TXT**	3.1	10.2
5	**CTU^{TO}**	50.7	63.9	13.3	**CTX**	6.0	7.3
6	**U^{TO}TC**	49.1	59.8	10.8	**XTC**	-1.4	12.1

Table 8*. All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) calculated by difference between the melting temperature of modified thiazole orange (U^{TO}) sequence with its unmodified RNA complementary sequence and the reference sequence obtained with unmodified RNA probe. b) calculated by difference between the melting temperature of modified alkyne (X) sequence and unmodified DNA complementary. All sequences differed only for three central nucleotides and follow the general formula: GCAG"YXY"GACT where Y represents one of the four nucleobases as indicated in the oligo entry.*

Mismatch stability was evaluated with the same experiments as before, giving more oscillating results. Similar result as the one obtained for DNA targets were obtained for mismatched nucleobase in the 5' position next to TO modification (Table 9, entry 3). However, when the mismatch is positioned directly opposite the modification, there seem to be a complete loss of discriminating ability since the melting temperature is practically the same as the matched reference (Table 9, entry 3). This is consistent with previous reports.[445] As before, the duplex with mismatch in position -1 is a lot less stable than the other mismatched duplexes, this time the difference in melting temperature is more than double. This is again probably due to the steric clash between the adenine and guanine for this sequence. In general, it seems that an RNA target is more sensitive to the modifications and the various effects are enhanced. Like before the best discrimination ability is obtained when the mismatch is not positioned directly on the modification (entry 3, Table 9).

Table 9. RNA target mismatch melting temperatures

Entry	Oligo-LNA-TO	Tm mismatch RNA	Tm mismatch reference	Δ Tm mismatch RNA	Δ Tm match - mismatch	Δ Tm match – mismatch reference
1	**CTU^{TO}** [a]	54.6	52.8	1.8	9.4	-2.1
2	**mismatch pos +1** [b]	52.8	45.7	7.1	7.0	3.4
3	**mismatch pos 0** [b]	59.5	49.2	10.3	0.3	-0.1
4	**mismatch pos -1** [b]	44.2	31.0	13.2	15.6	18.0

Table 9. *All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) Melting temperature of modified thiazole orange (U^{TO}) sequence with an unmodified DNA double mismatched sequence and the reference sequence obtained with unmodified DNA probe. b) melting temperature of thiazole orange (U^{TO}) sequence and single mismatch in different positions against DNA sequence. Position 6 has the mismatch one nucleotide in 5' direction respect to TO modification, position 5 has the mismatched nucleotide on the modification, position 4 has it one nucleotide 3' of modification. ⊿ Tm match – mismatch is calculated by subtracting the melting temperature of mismatched sequences to the reference match sequence with thiazole orange.*

When the same tests were repeated with pyrene modified oligonucleotides, a general duplex stabilization was obtained. This is encouraging given that duplex destabilization (up to -5 °C) has been reported in literature upon pyrene modification of LNA oligonucleotides using triazole functionalized C5 uridine.[420] In this instance however, smaller differences in modified oligonucleotides and reference melting temperatures were observed, signifying a smaller stabilization of the duplex in respect to thiazole orange. In general duplex stabilization appears to be approximately half that obtained with the thiazole orange modification. some sequence specific trends, such as a relatively high stabilization of AXA sequence and a relatively low stabilization of TXT sequence are confirmed. In addition, the duplex stabilization contribute of pyrene, calculated by subtraction of the ΔTm obtained with alkyne alone to ΔTm obtained with pyrene, is lower than that observed with thiazole orange. Little variation in the ΔTm match alkyne confirms that stabilization given by locked nucleic acids remains approximately the same for the two sequences.

Table 10. DNA target melting temperatures

Entry	Oligo LNA-Py	Tm reference	Tm match	Δ Tm match[a]	Oligo LNA-alkyne	Δ Tm match alkyne[b]	Δ Δ Tm Alkyne-Py
1	**AU^{Py}A**	41.2	54.6	13.4	**AXA**	7.93	5.5
2	**GU^{Py}G**	55.2	61.4	6.2	**GXG**	2.5	3.7
3	**CU^{Py}C**	59.2	64.6	5.4	**CXC**	1.82	3.6
4	**TU^{Py}T**	51.7	54.8	3.1	**TXT**	-1.31	4.4
5	**CTU^{Py}**	52.8	60.7	7.9	**CTX**	3.43	4.5
6	**U^{Py}TC**	53.5	59.5	6.0	**XTC**	-2.05	8.1

Table 10*. All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) calculated by difference between the melting temperature of modified pyrene (U^{Py}) sequence with its unmodified DNA complementary sequence and the reference sequence obtained with unmodified DNA probe. b) calculated by difference between the melting temperature of modified alkyne (X) sequence and unmodified DNA complementary. All sequences differed only for three central nucleotides and follow the general formula: GCAG"YXY"GACT where Y represents one of the four nucleobases as indicated in the oligo entry.*

As before, the mismatch discrimination ability was evaluated by difference between the match melting temperature of the modified pyrene oligonucleotides and the mismatch one. The smaller increase in melting temperature obtained before with pyrene in respect with thiazole orange is replicated on mismatched sequences giving a smaller increase in Tm. This in turn, allows to obtain a similar ΔTm match-mismatch to the one obtained with thiazole orange. However, this time, the discriminating ability result similar to the one obtained by unmodified oligonucleotides.

Table 11. DNA mismatch target melting temperatures

Entry	Oligo-LNA-Py	Tm mismatch reference	Tm mismatch DNA	Δ Tm mismatch DNA	Δ Tm match - mismatch	Δ Tm match – mismatch reference
1	**CTU$^{Py\ a}$**	46.0	53.1	7.1	7.6	6.8
2	**mismatch pos +1[b]**	45.8	50.9	5.2	8.6	7.8
3	**mismatch pos 0[b]**	44.9	50.8	5.9	8.7	8.6
4	**mismatch pos -1[b]**	41.4	48.0	6.6	11.5	12.1

Table 11. *All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) Melting temperature of modified pyrene (U^{Py}) sequence with an unmodified DNA double mismatched sequence and the reference sequence obtained with unmodified DNA probe. b) melting temperature of thiazole orange (U^{Py}) sequence and single mismatch in different positions against DNA sequence. Position 6 has the mismatch one nucleotide in 5' direction respect to Py modification, position 5 has the mismatched nucleotide on the modification, position 4 has it one nucleotide 3' of modification. ⊿ Tm match – mismatch is calculated by subtracting the melting temperature of mismatched sequences to the reference match sequence with pyrene.*

The results concerning the RNA target melting temperatures are again similar to the ones obtained before with TO for RNA sequences. Once again there is a smaller stabilization contribution from pyrene than from TO and the differences in melting temperatures are pretty consistent for all sequences. Against RNA the sequence specific effects of improved AXA stabilization and lower TXT duplex stabilization are mitigated completely. Only for entry 5 and 6 in Table 12 there is an increase in melting temperature difference, which however stays below the range of that reached with DNA modification.

Table 12. RNA target melting temperatures

Entry	Oligo LNA-Py	Tm reference	Tm match	Δ Tm match[a]	Oligo LNA-alkyne	Δ Tm match alkyne[b]	Δ Δ Tm Alkyne-Py
1	**AU^{Py}A**	40.8	47.1	6.3	**AXA**	7.6	-1.3
2	**GU^{Py}G**	53.2	58.4	5.3	**GXG**	5.3	0.0
3	**CU^{Py}C**	56.5	62.1	5.6	**CXC**	5.2	0.4
4	**TU^{Py}T**	46.3	50.8	4.4	**TXT**	3.1	1.3
5	**CTU^{Py}**	50.7	63.9	13.3	**CTX**	6.0	7.3
6	**U^{Py}TC**	49.1	59.8	10.8	**XTC**	-1.4	12.1

Table 12*. All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) calculated by difference between the melting temperature of modified thiazole orange (U^{Py}) sequence with its unmodified RNA complementary sequence and the reference sequence obtained with unmodified RNA probe. b) calculated by difference between the melting temperature of modified alkyne (X) sequence and unmodified DNA complementary. All sequences differed only for three central nucleotides and follow the general formula: GCAG"YXY"GACT where Y represents one of the four nucleobases as indicated in the oligo entry.*

RNA mismatch studies paint a similar picture to the one observed for thiazole orange modification. This time there is a stronger discrimination ability of the pyrene modified oligonucleotide when compared to the reference. However, this is only true when the mismatch is positioned 5' to the modification in target strand. When is directly against the modification or 3' shifted the discriminating power is lost. This could indicate a preferential positioning of pyrene upon intercalation.

Table 13. RNA mismatch target melting temperatures

Entry	Oligo-LNA-Py	Tm mismatch reference	Tm mismatch RNA	Δ Tm mismatch RNA	Δ Tm match - mismatch	Δ Tm match – mismatch reference
1	**CTU^{Py}** [a]	46.1	51.7	5.6	12.3	3.0
2	**mismatch pos +1**[b]	45.7	52.8	7.1	7.0	3.4
3	**mismatch pos 0**[b]	49.2	59.5	10.3	0.3	-0.1

4	**mismatch pos -1[b]**	31.0	44.2	13.2	15.6	18.0

Table 13*. All melting temperatures are given in Celsius and are measured as reported above in duplicate. a) Melting temperature of modified thiazole orange (U^{Py}) sequence with an unmodified DNA double mismatched sequence and the reference sequence obtained with unmodified DNA probe. b) melting temperature of pyrene (U^{Py}) sequence and single mismatch in different positions against DNA sequence. Position 6 has the mismatch one nucleotide in 5' direction respect to Py modification, position 5 has the mismatched nucleotide on the modification, position 4 has it one nucleotide 3' of modification. ⊿ Tm match – mismatch is calculated by subtracting the melting temperature of mismatched sequences to the reference match sequence with pyrene.*

Fluorescence studies

The fluorescence behaviour of thiazole orange and pyrene were evaluated on a Perkin Elmer LS50B luminescence spectrometer fitted with Perkin Elmer PTP-1 Peltier temperature controller. The experiments were set with FLWinlabTempScan software and envisaged the irradiation at 484 nm and recording of emission spectra from 510 nm for thiazole orange and irradiation at 336 nm and recording of emission starting at 360 nm for pyrene, with an excitation amplitude slit of 7 nm, and an emission slit of 7 nm at 20 °C. For each experiment a blank sample of 10 mM phosphate and 200 mM NaCl at pH 7.0 was recorded to assure no interference was present. The fluorescence was measured at a concentration of 0.25 μM probe and the target strand added at 0.28 μM concentration. The obtained data were then plotted with excel graphic elaboration software and integrated to give the total area. The ratio between the area of the double stranded and singe stranded oligonucleotide fluorescence was calculated and is reported in the tables below. The ratio between the intensities is given at the maximum emission intensity registered for each spectrum, which falls between 520-560 nm for thiazole orange and 380-400 nm for pyrene.

In general, an increase in fluorescence is to be expected upon hybridization with the presence in thiazole orange. This behaviour is verified for all oligonucleotide sequences in varying degrees. For instance, when the thiazole orange is placed between adenine the ratio of increase is DNA is 4.3; while the effect with guanine (entry 2, Table 14) is only 1.5. However, this last behaviour can be explained by the partial increase in emission in the single strand due to the π-stacking of thiazole orange on guanine. The results are in general agreement with similar modification on DNA reported in literature, although with somewhat smaller ratios.[445] This, in turn, means that the presence of the locked nucleic acid could in some way reduce the intercalating ability of thiazole orange while still maintaining a good duplex stability, as seen by the melting temperature experiments. Similar observations were made for RNA, in this case however, the ratios are higher and more favourable. In particular, the same sequence specific effects could be seen for $GU^{TO}G$ and $CU^{TO}C$ sequences, but the ratio between the maximum intensity of the double strand and single strand give in general a good discrimination, comparable to the one obtained with PNA oligonucleotides when the thiazole orange is linked by the thiazole moiety.[446]

Table 14. Fluorescence ratios measurements matched sequences

Entry	Oligo LNA-TO	Fds/Fss DNA	Ids/Iss DNA	Fds/Fss RNA	Ids/Iss RNA
1	**$AU^{TO}A$**	4.3	4.8	6.2	7.1

2	**GU^{TO}G**	1.5	1.5	2.0	2.0
3	**CU^{TO}C**	1.7	1.7	2.8	3.0
4	**TU^{TO}T**	2.8	3.4	4.7	5.5
5	**CTU^{TO}**	4.0	4.8	4.4	5.0
6	**U^{TO}TC**	3.9	4.0	6.2	6.3

Table 14. *Fds/Fss is the ratio between the total area of the emission graph of double/single stranded oligonucleotides. Ids/Iss is the ration between the maximum emission intensity of double/single stranded oligonucleotides.*

Fluorescence was also measured for mismatched sequences, as before the ratio between double and single stranded oligonucleotides is reported in Table 11. Moreover, the ratio for both DNA and RNA targets is greater, on average, than the one reported in table 10; thus, confirming the specific influence of neighbouring nucleobases on thiazole orange modification. The fluorescent ratio between match (entry 6, Table 15) and mismatch sequence was evaluated on maximum emission wavelength. In general, mismatched sequences gave lower fluorescence than matched ones, with comparable ratios, when possible, to the ones obtained in literature, albeit a little lower. There seems to be little variation caused by mismatch positioning. At the same time is confirmed the lower thiazole orange mismatch discrimination ability when the mismatched nucleotide in the target sequence is positioned directly in front of probe sequence modified nucleotide (Entry 3, Table 15).

Table 15. Fluorescence ratios measurements mismatched sequences

Entry	Oligo LNA-TO	Fds/Fss DNA	Ids/Iss DNA	Im/Imm DNA	Fds/Fss RNA	Ids/Iss RNA	Im/Imm RNA
1	**CTU^{TO}**	3.5	3.9	1.3	4.9	5.3	1.9
2	**mismatch pos +1**	5.8	6.8	1.0	4.2	4.5	2.8
3	**mismatch pos 0**	5.0	5.4	0.8	5.7	5.8	2.0
4	**mismatch pos -1**	3.6	3.7	1.1	4.1	4.1	2.8

Table 15. *Fds/Fss is the ratio between the total area of the emission graph of double/single stranded oligonucleotides. Ids/Iss is the ration between the maximum emission intensity of double/single stranded oligonucleotides.Im/Imm is the ratio between the match sequence fluorescence taken at its maximum and the mismatched sequence one.*

Fluorescent behaviour was also studied on pyrene modified oligonucleotides. Pyrene is an intercalator agent but in general its fluorescence is quenched by π-stacking, so in the table the fluorescence ratio is reported between single strand and the double stranded oligonucleotides. While for DNA target the fluorescent quenching is more pronounced, giving ratios between 2.7 and 1.1 (entry 6 and 3, Table 16), for the RNA targets this effect is less visible giving ratios no higher than 2.0 (entry 1, Table 16). Moreover, for RNA targets there seems to be an enhancement of the fluorescence signal upon duplex formation (entry 3,4,5, Table 16). During the emission measurements there was no recording of excimers bands due to pyrene-pyrene interaction, as they were not intended to happen in this instance.

Table 16. Fluorescence ratios measurements matched sequences

Entry	Oligo LNA-Py	Fss/Fds DNA	Iss/Ids DNA	Fss/Fds RNA	Iss/Ids RNA
1	**AU^{Py}A**	2.6	2.8	2.0	2.1
2	**GU^{Py}G**	1.5	1.4	0.7	0.7
3	**CU^{Py}C**	1.1	1.2	0.7	0.7
4	**TU^{Py}T**	1.2	1.1	0.9	1.0
5	**CTU^{Py}**	1.4	1.4	1.0	1.0
6	**U^{Py}TC**	2.7	2.2	1.8	1.7

Table 16. *Fss/Fds is the ratio between the total area of the emission graph of single/double stranded oligonucleotides. Iss/Ids is the ration between the maximum emission intensity of single/double stranded oligonucleotides.*

As before, mismatch discrimination was evaluated, and in all cases, there was little variation between the fluorescence signal of the matched sequence and the mismatched one, thus making a single pyrene modification not very useful in this regard. Moreover, there seem to be no clear trend toward an increased or decreased fluorescence signal between matched and mismatched sequence.

Table 17. Fluorescence ratios measurements mismatched sequences

Entry	Oligo LNA-TO	Fss/Fds DNA	Iss/Ids DNA	Im/Imm DNA	Fss/Fds RNA	Iss/Ids RNA	Im/Imm RNA
1	**CTU^{Py}**	0.6	0.6	1.4	1.0	1.0	1.0
2	**mismatch pos +1**	2.8	2.3	1.0	1.2	1.2	0.7
3	**mismatch pos 0**	3.4	2.7	1.2	1.6	1.6	0.8
4	**mismatch pos -1**	1.4	1.3	0.5	1.2	1.1	0.6

Table 17. *Fss/Fds is the ratio between the total area of the emission graph of single/double stranded oligonucleotides. Iss/Ids is the ration between the maximum emission intensity of single/double stranded oligonucleotides. Im/Imm is the ratio between the match sequence fluorescence taken at its maximum and the mismatched sequence one.*

Conclusion

Incorporation of thiazole orange into oligonucleotides can be achieved on the uridine nucleobase even when the relative sugar is modified, as in this case by locking to the 4' ribose position. The synthesis of the novel locked ribose alkyne modified uridine phosphoramidite is reported. Thiazole orange makes the oligonucleotide more lipophilic and adds a positive charge. At the same time, being an intercalator it increases the melting temperature of the duplexes with target sequences. This effect is further enhanced by the locked ribose giving differences in melting temperature as high as + 19 °C in respect to the control. The ability to discriminate between fully matched sequence and a single mismatched one was evaluated and is very good. This characteristic can hopefully be exploited to avoid off-target effect of therapeutic

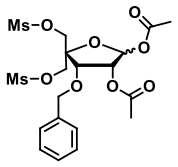
sequences. The results seem to be true for both DNA and RNA targets. Moreover, similar results were obtained for pyrene modification, although with less mismatch discrimination ability and smaller duplex stabilization. The fluorescent properties of thiazole orange are in general conserved and in some cases provide good signal enhancement upon duplex formation. When a comparison with literature is possible however, there seems to be a slight intercalation difficulty that could be given by the locked ribose modification. Fluorescence studied on pyrene modification gave little variations between single and double stranded oligonucleotides. Oligonucleotide sequences studied for exon skipping on a luciferase assay were synthetised with one, two or three modified uridines, with a phosphorothioate backbone and 2'Ome modification of the sugars. These sequences were purified by ion-pairing HPLC and clicked with thiazole orange following the same protocol as the other DNA sequences synthesised. Once further purified to remove residual copper ions, the sequences will be used to test the delivery on cells, through the luciferase assay, thus assessing the increased lipophilicity given by the thiazole orange and the effect of the hypothesised increase in duplex stability obtained by the fluorophore incorporations.

Materials and methods

All reagents were purchased from Sigma-Aldrich, Acros Organics, Lonza, Invitrogen or Fisher Scientific and used without further purification, or dried as described below. 3 Å molecular sieves (beads, 4 – 8 mesh, Sigma-Aldrich) were used to dry MeOH, DMF and ACN. Dry solvents were collected from a Grubbs-type SPS. Thin layer chromatography (TLC) was performed using Merck TLC silica gel 60 F254 plates (0.22 mm thickness, aluminium backed) and the compounds were visualized by irradiation at 254/365 nm and stained with *p*-anisaldehyde, potassium permanganate or sulfuric acid solutions. ^{1}H NMR spectra were measured at 400 MHz on a Bruker DPX400 (AVIIIHD 400) spectrometer. ^{13}C NMR spectra were measured at 101 MHz on a Bruker DPX400 spectrometer. ^{1}H were internally referenced to the appropriate residual un-deuterated solvent signal; ^{13}C NMR spectra were referenced to the deuterated solvent. Assignment of the signals was aided by COSY (^{1}H - ^{1}H), HSQC-DEPT, HSQC (^{1}H - ^{13}C) and HMBC (^{1}H - ^{13}C) experiments. Low-resolution mass spectra (LRMS) were recorded using electrospray ionisation (ESI^{+} or ESI^{-}) on a Waters ZMD quadrupole mass spectrometer in HPLC grade methanol. High-resolution mass spectra (HRMS) were recorded in HPLC grade methanol using electrospray ionisation (EI) on a Bruker APEX III FT-ICR mass spectrometer.

Monomer synthesis

(3R,4S)-4-(benzyloxy)-5,5-bis(((methylsulfonyl)oxy)methyl)tetrahydrofuran-2,3-diyl diacetate (124)

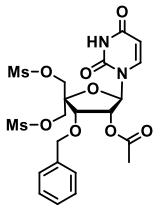


Literature procedures from Koshkin et al[447] were followed to obtain this compound.

Briefly, 10.00 grams (21.46 mmol) of compound 1 were dissolved in 63 ml 80% TFA and stirred at room temperature for 1 h. The solvents were removed under vacuum and the resulting solid was dissolved in 125 ml DCM and washed twice with 125 ml saturated $NaHCO_3$. The organic phase was dried with anhydrous sodium sulphate and the solvent was removed under vacuum. The obtained reaction crude was dissolved in dry pyridine and the solvent was evaporated. This procedure was repeated twice. The

obtained oily crude was taken in 50 ml dry pyridine. To this reaction mixture 7.54 ml (74 mmol) of acetic anhydride and the reaction was stirred overnight. the reaction was quenched by addition of 160 ml saturated $NaHCO_3$ and the compound extracted with 2X 125 ml EtOAc. The organic phase was washed with 100 ml brine and dried with anhydrous sodium sulphate. The solvent was removed under vacuum and the oily residue co-evaporated with toluene twice. The product was dried overnight in high vacuum to afford 9,85 g of white solid in 90% yield as a mixture of diastereomers. The product was used for the next reaction without further purification.

(2R,3R,4S)-4-(benzyloxy)-2-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-5,5-bis(((methylsulfonyl)oxy)methyl)tetrahydrofuran-3-yl acetate (126)

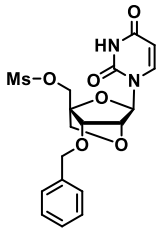


The reaction was performed following the procedures from Kumar et al.[448]

Briefly, 8.00 grams (15.68) of compound **125** and 2.18 grams (19.46) of uracil were placed in a 250 ml two necked round bottom flask, equipped with a condenser, and suspended in 80 ml dry acetonitrile under argon atmosphere. To this were added 10.70 ml (40.34 mmol) of N,O-Bis(trimethylsilyl)trifluoroacetamide and the reaction mixture was stirred at reflux for 1 hour, after which the reaction mixture turned clear. The mixture was allowed to cool to room temperature and 3.76 ml (20.66 mmol) of Trimethylsilyl trifluoromethanesulfonate were added slowly. The reaction mixture was refluxed for 20 hours. After assessment of reaction completion by TLC (1:1 EtOAc/PET) the reaction was quenched by addition of 100 ml saturated $NaHCO_3$. To this mixture were added 100 ml DCM and the resulting emulsion was filtered on a celite pad. The phases were separated and the aqueous one was extracted with 2 x 100 ml DCM. The combined organic phases were washed with 2 x 100 ml $NaHCO_3$, the solvent was dried by addition of anhydrous sodium sulphate, filtered on cotton ad dried to obtain 8.60 grams of crude off-white solid. 6.60 grams (11.63 mmol) of product were obtained after silica gel column purification (0-2% MeOH in DCM) in 75% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 9.34 (s, 1H, NH), 7.41 – 7.28 (m, 5H, Ph), 7.24 (d, J = 8.1 Hz, 1H, H6), 5.71 (d, J = 8.0 Hz, 1H, H5), 5.67 (d, J = 3.19 Hz, 1H, H1') 5.58 (dd, J = 6.5, 3.2 Hz, 1H, H2'), 4.70 (d, J = 6.6 Hz, 1H, 5'), 4.63 – 4.50 (m, 3H, 5"/CH_2Ph), 4.40 – 4.27 (m, 3H, 3', 5', 5"), 3.00 (s, 3H, CH_3SO_2), 2.99 (s, 3H, CH_3SO_2), 2.11 (s, 3H, CH_3CO). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.16 ($COCH_3$), 163.02 (C4), 150.02 (C2), 142.58 (C6), 136.70 (Ph), 128.82 (Ph), 128.69 (Ph), 128.57 (Ph), 103.25 (C5), 93.23 (C1'), 84.31, 78.01 (C3'), 74.92 (C4'), 73.80 (C2'), 68.30 (CH_2Ph), 67.60 (C5'), 37.84(CH_3SO_2), 37.75 (CH_3SO_2), 20.85 (CH_3CO).

((1R,3R,4R,7S)-7-(benzyloxy)-3-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl methanesulfonate (127)

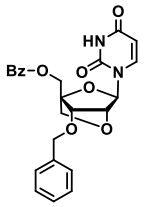


The reaction was performed by slightly varying the procedures from Kumar et al.[448]

7.50 grams (13.34 mmol) of compound **126** were dissolved in 43 ml dioxane. To this solution 43 ml water and 34 ml 2 M NaOH were added in succession. The solution turned yellow upon base addition. The reaction mixture was stirred for 1 h at rt, after which the rection completion was confirmed by TLC (EtOAc/PET 8:2). The pH of the solution was carefully brought to 7 by addition of concentrated HCl and a precipitate formed. The solution was placed a 4 °C overnight and then filtered on sintered glass. The white powder obtained was dissolved in methanol and the solvent evaporated. To afford 5,2 grams (12.26 mmol) of pure product in 92% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 (s, 1H, NH), 7.54 (d, J = 8.3 Hz, 1H, H6), 7.41 – 7.27 (m, 5H, Ph), 5.74 (d, J = 8.2 Hz, 1H, H5), 5.66 (s, 1H, H1'), 4.66 (d, J = 11.4 Hz, 1H, H5'), 4.62 – 4.51 (m, 4H, H2', H5', CH_2Ph), 4.11 (d, J = 7.9 Hz, 1H, H5"), 3.90 (d, J = 7.9 Hz, 1H, H5"), 3.86 (s, 1H, H3'), 3.07 (s, 3H Ms). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.17, 163.16, 150.12, 142.61, 136.72, 128.81, 128.67, 128.55, 103.24, 93.25, 84.32, 78.05, 77.48, 77.16, 76.84, 74.91, 73.80, 68.34, 67.62, 53.57, 37.81, 37.72, 20.84.

((1R,3R,4R,7S)-7-(benzyloxy)-3-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2,5-dioxabicyclo[2.2.1]heptan-1-yl)methyl benzoate (128).



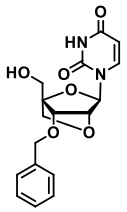
The reaction was performed following the procedures from Kumar et al.[448]

5.00 grams (11.79 mmol) of compound **127** were dissolved in 110 ml of dry DMF under argon atmosphere in a 250 ml round bottom flask equipped with an air condenser. 5.09 grams (35.07 mmol) of sodium benzoate were added to the solution and the reaction mixture was stirred at 90 °C overnight. After reaction completion (TLC EtOAc/PET 8:2) the mixture was allowed to reach room temperature and 30 ml of distilled water were added. The mixture was stirred until complete dissolution of the white precipitate. The solution was poured in 250 ml of cold water (approx. 5 °C) while stirring and the formed precipitate was filtered on glass (porosity 4). The powder was washed with 120 ml of cold water and dried under vacuum to afford 4.88 grams (10.80 mmol) of product in 92% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.92 (dd, J = 7.1, 1.2 Hz, 2H, *o*Bz), 7.64 (tt, J = 7.0, 1.2 Hz, 1H, *p*Bz), 7.48 (m, 3H, *m*Bz, H6), 7.28 (s, 5H, Ph), 5.64 (s, 1H, H1'), 5.43 (dd, J = 8.2, 2.1 Hz, 1H, H5), 4.80 (d, J = 12.7 Hz, 1H, H5'), 4.72 (d, J = 11.8 Hz, 1H, CH_2Ph), 4.62 (s, 1H, H3'), 4.59 (d, J = 12.7 Hz, 1H, H5'), 4.53 (d, J = 11.8 Hz, 1H, CH_2Ph), 4.19 (d, J = 7.9 Hz, 1H, H5"), 3.97 (d, J = 7.8 Hz, 1H, H5"), 3.87 (s, 1H,

H2'). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.68, 162.84, 138.44, 136.45, 133.92, 129.66, 128.86, 128.80, 128.56, 128.10, 101.84, 87.74, 86.45, 77.48, 77.16, 76.84, 76.55, 75.78, 72.48, 72.20, 59.14.

1-((1S,3R,4R,7S)-7-(benzyloxy)-1-(hydroxymethyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)pyrimidine-2,4(1H,3H)-dione (129)

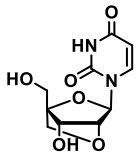


The reaction was performed by slightly modifying the procedures from Kumar et al.[448]

Briefly, 4.88 grams (10.80 mmol) of compound **128** were dissolved in 85 ml THF and to this solution 85 ml H_2O and 30 ml 2M NaOH were added. The reaction mixture was stirred for 2 hours at room temperature. The pH brought to 8 by careful addition of concentrated hydrochloric acid. The mixture was left at 4 °C overnight and then the precipitate was filtered on glass (porosity 4). The filtrate was washed with cold water and dried to afford 3,44 grams (9.93 mmol) of product in 92% yield.

^{1}H NMR (400 MHz, DMSO) δ 7.73 (d, J = 8.1 Hz, 1H, H6), 7.38 – 7.23 (m, 5H, Ph), 5.62 (d, J = 8.1 Hz, 1H, H5), 5.48 (d, J = 0.7 Hz, 1H, H1'), 5.26 (t, 1H, ex, J = 5.5 Hz, 5'-OH) 4.60 (s, 2H, CH_2Ph), 4.44 (s, 1H, H2'), 3.92 (s, 1H, H3'), 3.86 (d, J = 7.9 Hz, 1H, H5"), 3.78 (d, J = 4.2 Hz, 2H, H5'), 3.69 (d, J = 7.9 Hz, 1H, H5"). ^{13}C NMR (101 MHz, DMSO) δ 163.38 (C4), 150.02 (C2), 139.04 (C6), 137.92 (Ph), 128.26 (2C Ph), 127.60 (Ph), 127.42 (2C Ph), 100.90 (C5), 88.33 (C4'), 86.50 (C1'), 76.49 (C2'), 75.80 (C3'), 71.59 (C5"), 71.13 (CH_2Ph), 55.99 (C5').

1-((1S,3R,4R,7S)-7-hydroxy-1-(hydroxymethyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)pyrimidine-2,4(1H,3H)-dione (130)

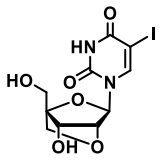


The reaction was performed by slightly modifying the procedures from Kumar et al.[448]

Briefly, 3 grams (8.66 mmol) of **129** were suspended in 83 ml methanol in a 250 ml round bottom flask. An argon atmosphere was created and 1 gram of 10% Pd on carbon (0.86 mmol) was added. The atmosphere was replaced with a hydrogen one and the reaction mixture was stirred overnight at 40 °C. TLC 9:1 DCM/MeOH confirmed the reaction completion. The suspension was filtered on wet celite to remove the carbon and most of the solvent was evaporated under vacuum. To the solution was added silica and the remaining solvent was evaporated. The crude was purified by flash column chromatography (DCM/MeOH 0-20 % v/v) to give 1.80 grams (6.79 mmol) of white solid in 78% yield.

^{1}H NMR (400 MHz, DMSO) δ 11.36 (s, 1H, ex, NH), 7.76 (d, J = 8.1 Hz, 1H, H6), 5.67 (d, J = 4.3 Hz, 1H, 3'-OH), 5.63 (d, J = 8.1 Hz, 1H, H5), 5.42 (d, J = 0.7 Hz, 1H, H1'), 5.15 (t, J = 5.5 Hz, 1H, 5'-OH), 4.14 (s, 1H, H2'), 3.88 (d, J = 4.3 Hz, 1H, H3'), 3.83 (d, J = 7.8 Hz, 1H, H5"), 3.75 (d, J = 5.6 Hz, 2H, H5'), 3.64 (d, J = 7.8 Hz, 1H, H5"). ^{13}C NMR (101 MHz, DMSO) δ 163.32 (C4), 150.03 (C2), 139.19 (C6), 100.85 (C5), 88.90 (C4'), 86.38 (C1'), 78.91 (C2'), 71.03 (C5"), 68.65 (C3'), 55.97 (C5').

1-((1S,3R,4R,7S)-7-hydroxy-1-(hydroxymethyl)-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-iodopyrimidine-2,4(1H,3H)-dione (131)

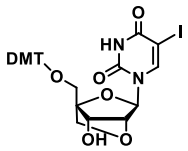


The reaction was performed following the procedures from Kumar et al.[448]

1.68 grams (6.56 mmol) of compound **130** were dissolved in 68 ml acetic acid. To this solution were added 1.79 grams (3.26 mmol) of Cerium ammonium nitrate (CAN) and 998 mg (3.93 mmol) of iodine. The reaction mixture was stirred at 80 °C and monitored by TLC (DCM/MeOH 9:1). After 40 minutes the reaction reached completion and the solvent was evaporated under vacuum. The solid residue was taken in 80 ml MeOH, and the solvent completely evaporated. This procedure was repeated other two times. The crude solid was purified by flash column chromatography (DCM/MeOH 0-20% v/v) to give 1.75 grams (4.58 mmol) of product in 70% yield.

^{1}H NMR (400 MHz, DMSO) δ 11.69 (s, 1H, ex, NH) 8.13 (s, 1H, H6), 5.68 (d, J = 4.4 Hz, 1H, 3'- OH), 5.40 (s, 1H, H1'), 5.30 (t, J = 5.3 Hz, 1H, 5'-OH), 4.14 (s, 1H, H2'), 3.91 (d, J = 4.5 Hz, 1H, H3'), 3.80 (d, J = 7.8 Hz, 1H, H5"), 3.74 – 3.70 (m, 2H, H5'), 3.60 (d, J = 7.7 Hz, 1H, H5"). ^{13}C NMR (101 MHz, DMSO) δ 160.77 (C4), 149.82 (C2), 143.68 (C6), 88.99 (C4'), 86.58 (C1'), 78.71 (C2'), 70.98 (C5"), 68.68 (C3'), 68.36 (C5), 55.43 (C5').

1-((1R,3R,4R,7S)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-iodopyrimidine-2,4(1H,3H)-dione (132)



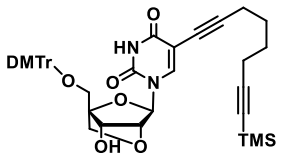
The reaction was performed by slightly modifying the procedures from Kumar et al.[448]

1.70 grams (4.45 mmol) of **131** were dissolved in 35 ml dry pyridine under argon atmosphere. To this solution 1.99 grams (5.87 mmol) of DMTr-Cl (4,4'-Dimethoxytrityl chloride) were added the reaction mixture was stirred for at room temperature. After 1 hour reaction was complete by TLC (EtOAc/PET 7:3) and the solvent was evaporated under vacuum. The residue was dissolved in 150 ml EtOAc and washed with 2 x 150 ml $NaHCO_3$. The aqueous phase was back extracted with 2 x 100 ml EtOAc. The combined organic phases were dried with anhydrous sodium sulphate, filtered and the solvent was removed under vacuum. The crude solid was purified by flash column chromatography (DCM/MeOH 0-5% v/v + 1% Py) to give 2.45 grams (3.58 mmol) of product in 80% yield.

^{1}H NMR (400 MHz, DMSO) δ 11.74 (s, 1H, NH), 7.96 (s, 1H, H6), 7.47 – 7.40 (m, 2H, DMTr), 7.39 – 7.11 (m, 9H, DMTr), 6.95 – 6.87 (m, 4H, DMTr), 5.74 (d, J = 4.6 Hz, 1H, 3'-OH), 5.43 (s, 1H, H1'), 4.23

(s, 1H, H2'), 4.06 (d, J = 4.7 Hz, 1H, H3'), 3.75 (s, 2H, H5''), 3.74 (s, 6H, 2xOCH$_3$), 3.40 (d, J = 11.1 Hz, 1H, H5'), 3.29 (d, J = 11.1 Hz, 1H, H5'). ^{13}C NMR (176 MHz, DMSO) δ 158.61 (C4), 150.32 (C2), 143.28 (C6), 128.50 (DMTr), 128.09 (DMTr), 127.22(DMTr), 113.82 (DMTr), 88.06, 87.45, 86.14 (C1'), 79.35 (C2'), 71.87 (C5''), 69.92 (C3'), 69.58 (DMTr), 60.23, 59.48 (C5'), 55.54 (2xOCH$_3$).

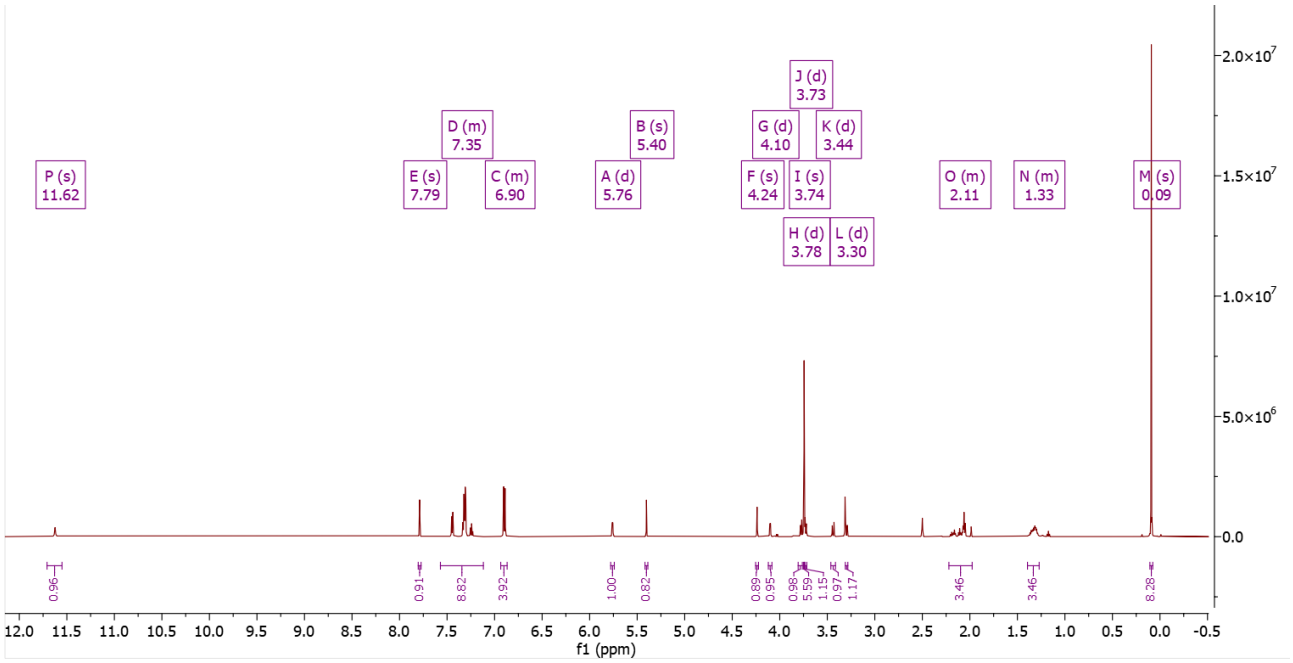
1-((1R,3R,4R,7S)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-7-hydroxy-2,5-dioxabicyclo[2.2.1]heptan-3-yl)-5-(8-(trimethylsilyl)octa-1,7-diyn-1-yl)pyrimidine-2,4(1H,3H)-dione (133)



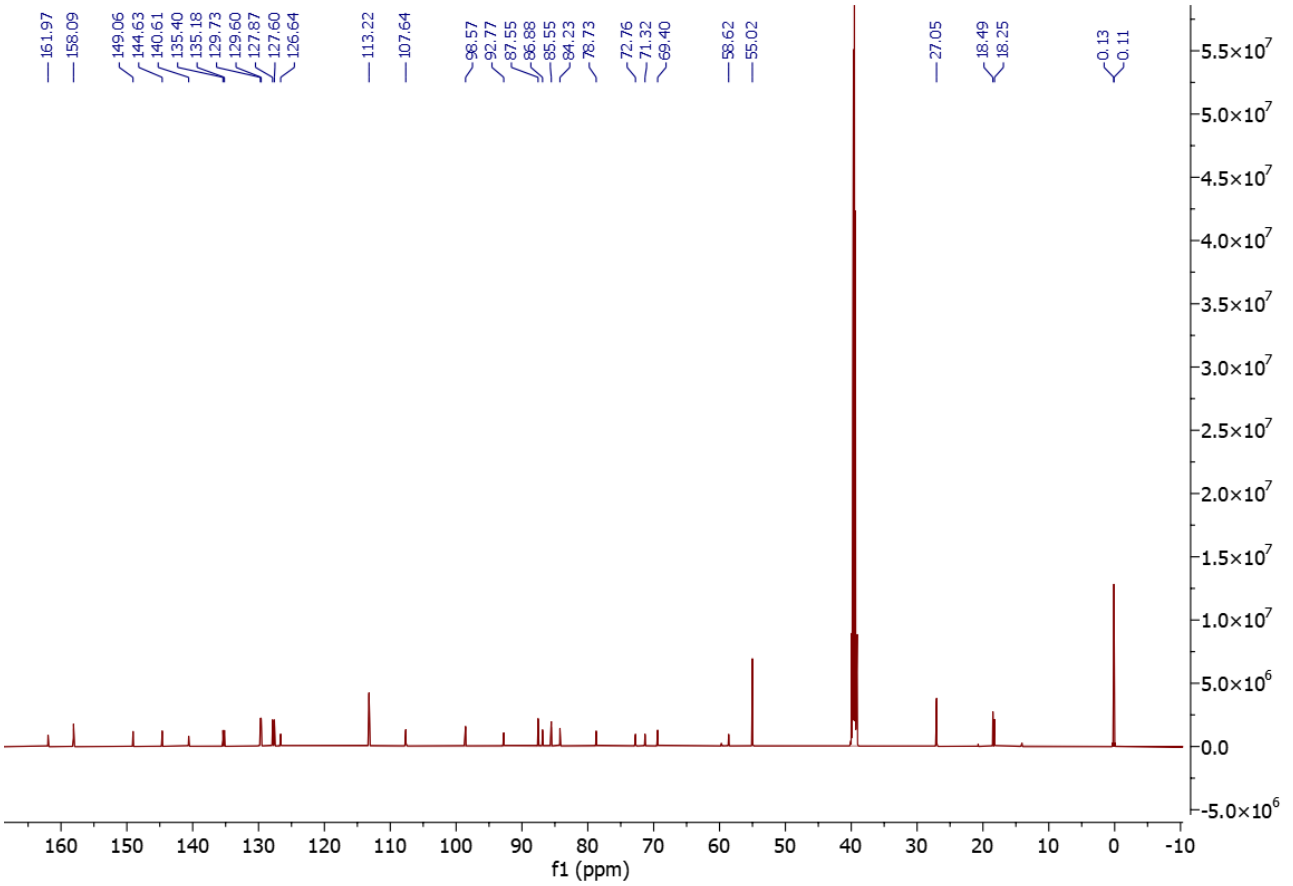
990 mg (1.44 mmol) of compound **132** were dissolved in 20 ml DMF in a 25 ml round bottom flask. To this solution 603 µl (4.33 mmol) of triethylamine (TEA) and 708 µl (2.89 mmol) of trimethyl(octa-1,7-diyn-1-yl)silane[449] were added and the solution was degassed by the freeze-pump-thaw method three times. At the end of the degassing procedure an argon atmosphere was created in the reaction flask. 136 mg (0.72 mmol) of CuI and 167 mg (0.14 mmol) of Pd(PPh$_3$)$_4$ were added together in the reaction flask. Argon atmosphere was recreated, and the reaction was monitored by TLC (DCM/MeOH 9:1 + 1% TEA). After 1 hour the solvent was removed under vacuum. The crude oil was dissolved in 50 ml EtOAc and washed with 3x 50 ml saturated NaHCO$_3$ and 3 x saturated EDTA (pH=7.5). the organic solvent was removed under vacuum and the crude product purified by flash column chromatography (DCM/MeOH 0-20% v/v + 2% TEA) to afford 852 mg (1.16 mmol) of pure product in 80% yield.

MS m/z [$C_{42}H_{47}N_2O_8Si$]$^+$ 735.3118; calculated mass 735.3096.

^{1}H NMR (600 MHz, DMSO) δ 11.62 (s, 1H, NH), 7.79 (s, 1H, H6), 7.57 – 7.12 (m, 9H, DMTr), 6.94 – 6.87 (m, 4H, DMTr), 5.76 (d, J = 4.6 Hz, 1H, 3-OH), 5.40 (s, 1H, H1'), 4.24 (s, 1H, H2'), 4.10 (d, J = 4.5 Hz, 1H, H3'), 3.78 (d, J = 7.9 Hz, 1H, H5''), 3.74 (s, 6H, 2OCH$_3$), 3.73 (d, J = 7.9 Hz, 1H, H5''), 3.44 (d, J = 11.1 Hz, 1H, H5'), 3.30 (d, J = 11.3 Hz, 1H, H5'), 2.22 – 1.98 (m, 4H, alk H3,6), 1.39 – 1.27 (m, 3H, alk H4,5), 0.09 (s, 9H, TMS). ^{13}C NMR (151 MHz, DMSO) δ 161.97 (C4), 158.09 (C2), 149.06 (C5), 144.63 (DMTr), 140.61 (C6), 135.40 (DMTr), 135.18 (DMTr), 129.73 (DMTr), 129.60 (DMTr), 127.87 (DMTr), 127.60 (DMTr), 126.64 (DMTr), 113.22 (DMTr), 107.64 (alk), 98.57 (alk), 92.77 (alk), 87.55 (C4'), 86.88 (C1'), 85.55 (DMTr), 84.23 (alk), 78.73 (C2'), 72.76 (alk), 71.32 (C5''), 69.40 (C3'), 58.62 (C5'), 55.02 (2OCH$_3$), 27.05 (alk C4,5), 18.49 (alk, C6), 18.25 (alk C3), 0.13 (TMS).

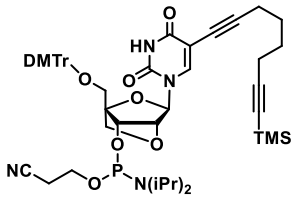


^{1}H-NMR of compound **134**



^{13}C-NMR of compound **134**

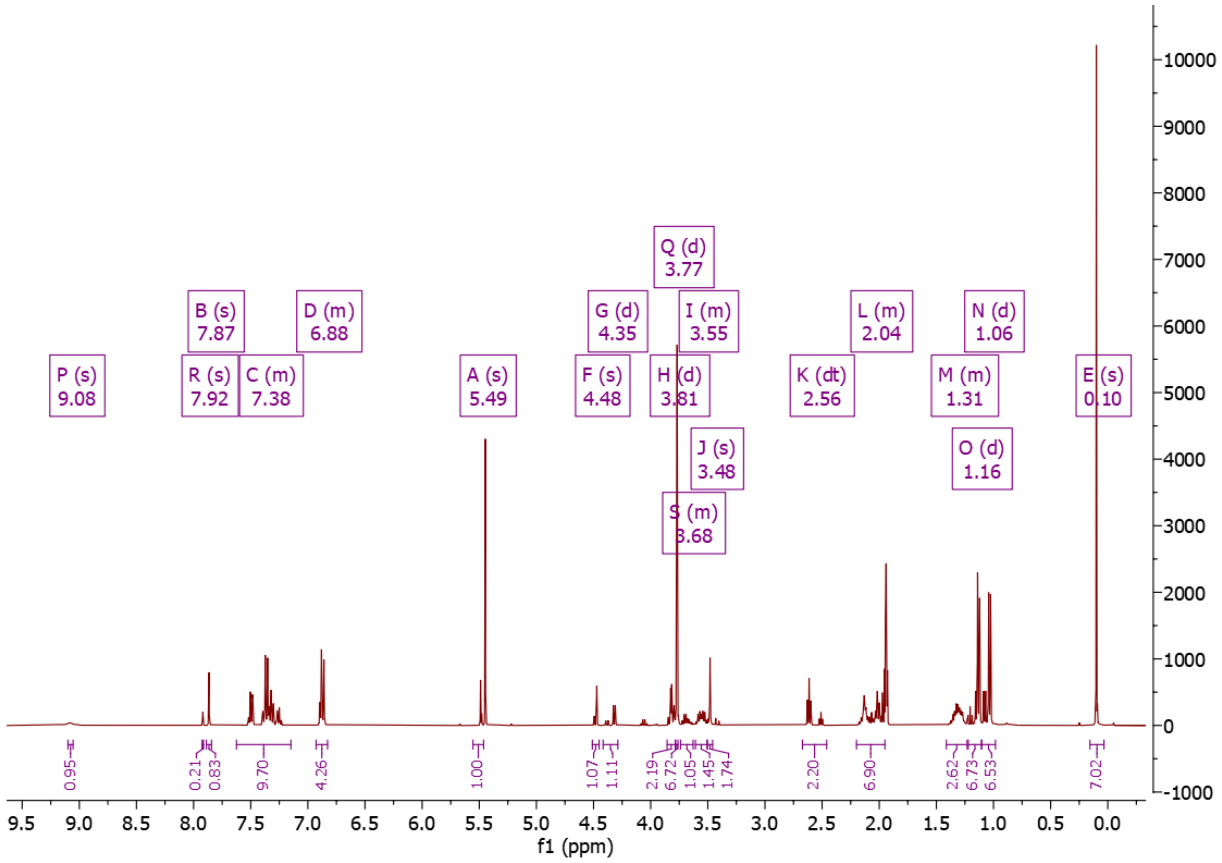
(1R,3R,4R,7S)-1-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-3-(2,4-dioxo-5-(8-(trimethylsilyl)octa-1,7-diyn-1-yl)-3,4-dihydropyrimidin-1(2H)-yl)-2,5-dioxabicyclo[2.2.1]heptan-7-yl (2-cyanoethyl) diisopropylphosphoramidite (134)



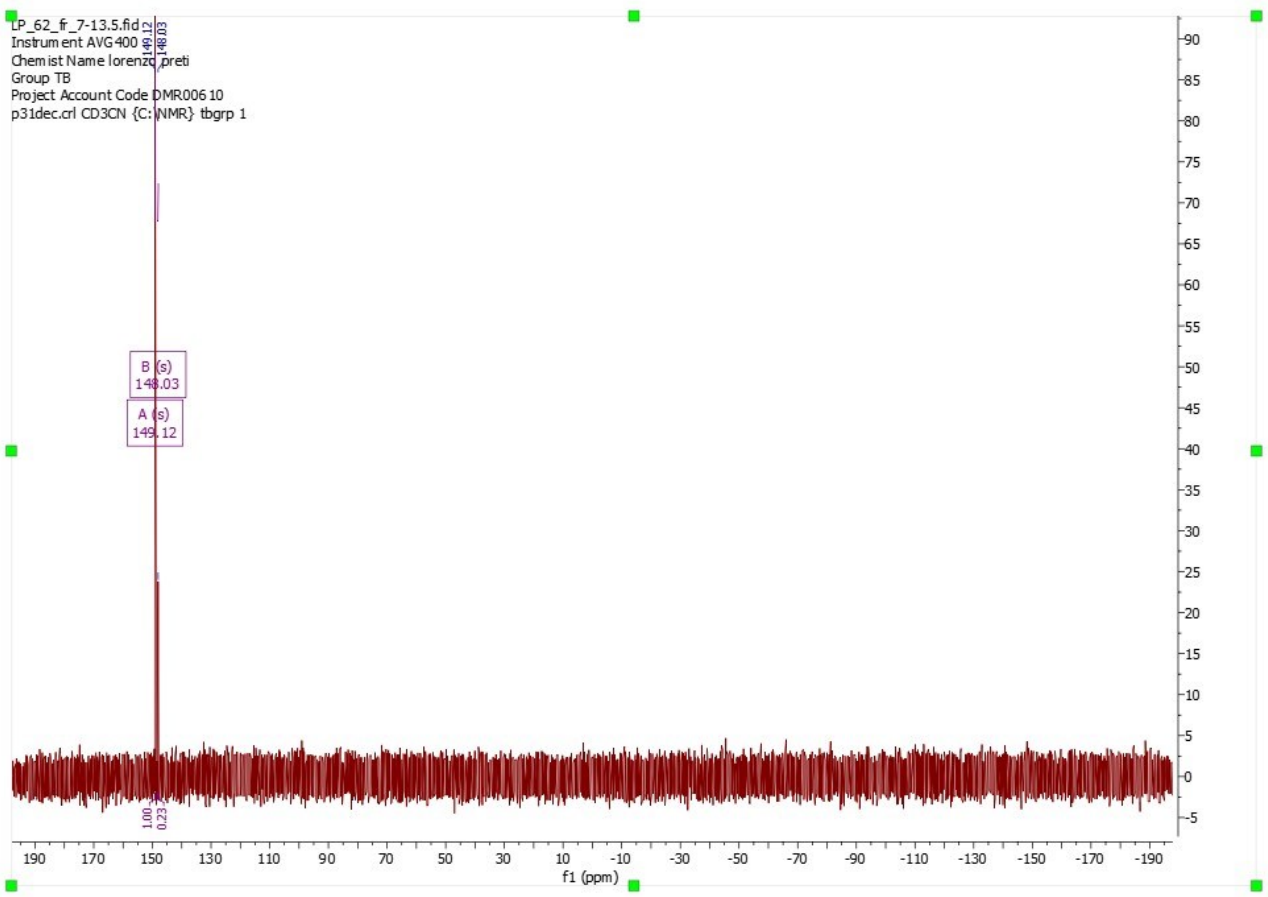
350 mg (0.476 mmol) of compound **133** were dissolved in dry DCM and dried overnight under high vacuum. The compound was redissolved in 10 ml dry DCM under argon atmosphere. To this solution were added 331 µl (2.38 mmol) of dry TEA and 158 µl (0.714 mmol) 3-((chloro(diisopropylamino)phosphaneyl)oxy)propanenitrile. The reaction mixture was stirred at room and monitored by TLC (DCM/MeOH 98/2 v/v + 1 % TEA) and after 2 hours was complete. the reaction mixture was diluted with 50 ml DCM and washed with 50 ml saturated KCl solution under argon atmosphere. the aqueous phase was exctracted with 2 x 25 ml DCM and the organic phases combined and dried with sodium sulphate anhydrous. The solvent was removed under vacuum and flask filled with argon gas. The crude product was purified by a short silica column (DCM 0 to 2 v/v MeOH + 1 % TEA) under argon to afford 360 mg (0.385 mmol) of product as two diastereomers (8:2) as a white foam in 80% yield.

^{1}H NMR (400 MHz, CD_3CN) δ 9.08 (s, 1H, NH), 7.87 (s, 1H, H6), 7.62 – 7.15 (m, 10H, DMTr), 6.93 – 6.83 (m, 4H, DMTr), 5.49 (s, 1H, H1'), 4.48 (s, 1H, H2'), 4.35 (d, J = 7.6 Hz, 1H, H3'), 3.81 (d, J = 3.4 Hz, 2H, H5''), 3.77 (s, 6H, 2xOCH_3), 3.74 – 3.63 (m, 1H, iPr (H)), 3.61 – 3.51 (m, 1H, iPr (H)), 3.48 (s, 2H, H5'), 2.56 (t, 5.9 Hz, 2H, CNEt), 2.20 – 1.95 (m, 6H, alk, CNEt), 1.41 – 1.23 (m, 4H, alk), 1.16 (d, J = 6.9 Hz, 6H, 2x iPr (CH_3)), 1.06 (d, J = 6.8 Hz, 6H, 2x iPr (CH_3)), 0.10 (s, 9H, TMS).

^{31}P NMR (CD_3CN) δ 149.12, 148,12.



^{1}H-NMR of compound **134**



^{31}P-NMR of compound **134**

Bibliography

1. Fantoni, N. Z., El-Sagheer, A. H. & Brown, T. A Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. *Chem Rev* **121**, 7122–7154 (2021).
2. Narain, R. *Chemistry of Bioconjugates*. (Wiley, 2014). doi:10.1002/9781118775882.
3. Wong, S. S. & Jameson, D. M. *Chemistry of Protein and Nucleic Acid Cross-Linking and Conjugation*. (CRC Press, 2011). doi:10.1201/b11175.
4. Algar, W. R. A Brief Introduction to Traditional Bioconjugate Chemistry. in *Chemoselective and Bioorthogonal Ligation Reactions* 1–36 (Wiley, 2017). doi:10.1002/9783527683451.ch1.
5. Hermanson, G. T. Bioconjugate Techniques: Third Edition. *Bioconjugate Techniques: Third Edition* 1–1146 (2013) doi:10.1016/C2009-0-64240-9.
6. Algar, W. R., Dawson, P. E. & Medintz, I. L. *Chemoselective and bioorthogonal ligation reactions. volume 1 and 2 : concepts and applications*.
7. Medintz, I. L. *Chemoselective and Bioorthogonal Ligation Reactions*. (2017).
8. Kalia, J. & Raines, R. Advances in Bioconjugation. *Curr Org Chem* **14**, 138–147 (2010).
9. Kalia, J. & Raines, R. T. Hydrolytic Stability of Hydrazones and Oximes. *Angewandte Chemie International Edition* **47**, 7523–7526 (2008).

10. Wilbur, D. S. Radiohalogenation of Proteins: An Overview of Radionuclides, Labeling Methods and Reagents for Conjugate Labeling. *Bioconjug Chem* **3**, 433–470 (1992).

11. Brinkley, M. A Brief Survey of Methods for Preparing Protein Conjugates with Dyes, Haptens and Crosslinking Reagents. *Bioconjug Chem* **3**, 2–13 (1992).

12. Baslé, E., Joubert, N. & Pucheault, M. Protein chemical modification on endogenous amino acids. *Chem Biol* **17**, 213–227 (2010).

13. Jentoft, N. & Dearborn, D. G. Protein labeling by reductive methylation with sodium cyanoborohydride: Effect of cyanide and metal ions on the reaction. *Anal Biochem* **106**, 186–190 (1980).

14. Hermanson, G. T. Bioconjugate techniques Chapter 3 - The Reactions of Bioconjugation. *Bioconjugate Techniques (Third edition)* 229–258 (2013).

15. Wong, S. S. & Jameson, D. M. *Chemistry of Protein and Nucleic Acid Cross-Linking and Conjugation*. (CRC Press, 2011). doi:10.1201/b11175.

16. Dawson, P. E. & Kent, S. B. H. Synthesis of native proteins by chemical ligation. *Annu Rev Biochem* **69**, 923–960 (2000).

17. Gilles, M. A., Hudson, A. Q. & Borders, C. L. Stability of water-soluble carbodiimides in aqueous solution. *Anal Biochem* **184**, 244–248 (1990).

18. Nakajima, N. & Ikada, Y. Mechanism of Amide Formation by Carbodiimide for Bioconjugation in Aqueous Media. *Bioconjug Chem* **6**, 123–130 (1995).

19. Sehgal, D. & Vijay, I. K. A method for the high efficiency of water-soluble carbodiimide-mediated amidation. *Anal Biochem* **218**, 87–91 (1994).

20. ANJANEYULU, P. S. R. & STAROS, J. V. Reactions of N-hydroxysulfosuccinimide active esters. *Int J Pept Protein Res* **30**, 117–124 (1987).

21. Nojima, Y., Iguchi, K., Suzuki, Y. & Sato, A. The pH-dependent formation of PEGylated bovine lactoferrin by branched polyethylene glycol (PEG)-N-hydroxysuccinimide (NHS) active esters. *Biol Pharm Bull* **32**, 523–526 (2009).

22. Wiederschain, G. Ya. The Molecular Probes handbook. A guide to fluorescent probes and labeling technologies. *Biochemistry (Moscow)* **76**, 1276–1276 (2011).

23. Valeur, E. & Bradley, M. Amide bond formation: Beyond the myth of coupling reagents. *Chem Soc Rev* **38**, 606–631 (2009).

24. El-Faham, A. & Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chem Rev* **111**, 6557–6602 (2011).

25. BODANSZKY, M. In search of new methods in peptide synthesis. A review of the last three decades. *Int J Pept Protein Res* **25**, 449–474 (1985).

26. Manne, S. R., De La Torre, B. G., El-Faham, A. & Albericio, F. OxymaPure Coupling Reagents: Beyond Solid-Phase Peptide Synthesis. *Synthesis (Stuttg)* **52**, 3189–3210 (2020).

27. Davies, J. J., Christopher Braddock, D. & Lickiss, P. D. Silicon compounds as stoichiometric coupling reagents for direct amidation. *Org Biomol Chem* **19**, 6746–6760 (2021).

28. Magano, J. Large-Scale Amidations in Process Chemistry: Practical Considerations for Reagent Selection and Reaction Execution. *Org Process Res Dev* **26**, 1562–1689 (2022).

29. McCaldon, P. & Argos, P. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* **4**, 99–122 (1988).

30. Kalia, J. & Raines, R. Advances in Bioconjugation. *Curr Org Chem* **14**, 138–147 (2010).

31. De Graaf, A. J., Kooijman, M., Hennink, W. E. & Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconjug Chem* **20**, 1281–1295 (2009).

32. Wilbur, D. S. Radiohalogenation of proteins: an overview of radionuclides, labeling methods, and reagents for conjugate labeling. *Bioconjug Chem* **3**, 433–470 (1992).

33. Means, G. E. & Feeney, R. E. Chemical Modifications of Proteins: History and Applications. *Bioconjug Chem* **1**, 2–12 (1990).

34. Brinkley, M. A Brief Survey of Methods for Preparing Protein Conjugates with Dyes, Haptens and Crosslinking Reagents. *Bioconjug Chem* **3**, 2–13 (1992).

35. Kalia, J. & Raines, R. T. Catalysis of imido group hydrolysis in a maleimide conjugate. *Bioorg Med Chem Lett* **17**, 6286–6289 (2007).

36. Ellman, G. L. Tissue sulfhydryl groups. *Arch Biochem Biophys* **82**, 70–77 (1959).

37. Lei, J. *et al.* Drug Release from Disulfide-Linked Prodrugs: Role of Thiol Agents. *Mol Pharm* **18**, 2777–2785 (2021).

38. Henkel, T., Brunne, R. M., Müller, H. & Reichel, F. Statistical Investigation into the Structural Complementarity of Natural Products and Synthetic Compounds. *Angew. Chem. Int. Ed* **38**, (1999).

39. Gardossi, L., Bianchi, D. & Klibanov, A. M. Selective Acylation of Peptides Catalyzed by Lipases in Organic Solvents. *J Am Chem Soc* **113**, 6328–6329 (1991).

40. Ohshima, T., Iwasaki, T., Maegawa, Y., Yoshiyama, A. & Mashima, K. Enzyme-like chemoselective acylation of alcohols in the presence of amines catalyzed by a tetranuclear zinc cluster. *J Am Chem Soc* **130**, 2944–2945 (2008).

41. Shafir, A., Lichtor, P. A. & Buchwald, S. L. N- versus O-arylation of aminoalcohols: Orthogonal selectivity in copper-based catalysts. *J Am Chem Soc* **129**, 3490–3491 (2007).

42. Monnier, F. & Taillefer, M. Catalytic C-C, C-N, and C-O Ullmann-type coupling reactions. *Angew Chem Int Ed Engl* **48**, 6954–6971 (2009).

43. Odendaal, A. Y., Trader, D. J. & Carlson, E. E. Chemoselective enrichment for natural products discovery. *Chem Sci* **2**, 760 (2011).

44. Trader, D. J. & Carlson, E. E. Chemoselective hydroxyl group transformation: an elusive target. *Mol Biosyst* **8**, 2484 (2012).

45. Hall, D. G. *Boronic acids : preparation and applications in organic synthesis and medicine*. (Wiley-VCH Verlag, 2005).

46. Christie, R. J., Anderson, D. J. & Grainger, D. W. Comparison of hydrazone heterobifunctional cross-linking agents for reversible conjugation of thiol-containing chemistry. *Bioconjug Chem* **21**, 1779–1787 (2010).

47. Siegel, D. Applications of reversible covalent chemistry in analytical sample preparation. *Analyst* **137**, 5457–5482 (2012).

48. Leriche, G., Chisholm, L. & Wagner, A. Cleavable linkers in chemical biology. *Bioorg Med Chem* **20**, 571–582 (2012).

49. Holmes, C. P. Model Studies for New o-Nitrobenzyl Photolabile Linkers: Substituent Effects on the Rates of Photochemical Cleavage. *J Org Chem* **62**, 2370–2380 (1997).

50. Boyd, D. B. *Computational Medicinal Chemistry for Drug Discovery Edited by Patrick Bultinck, Hans De Winter, Wilfried Langenaeker, and Jan P. Tollenaere. Marcel Dekker, New York. 2004. xiv + 794 pp. 18 × 26 cm. ISBN 0-8247-4774-7. $195.00. Journal of Medicinal Chemistry* vol. 47 (American Chemical Society (ACS), 2004).

51. Soellner, M. B., Nilsson, B. L. & Raines, R. T. Reaction mechanism and kinetics of the traceless Staudinger ligation. *J Am Chem Soc* **128**, 8820–8828 (2006).

52. Bednarek, C., Wehl, I., Jung, N., Schepers, U. & Bräse, S. The Staudinger Ligation. *Chem Rev* **120**, 4301–4354 (2020).

53. Lin, Y. A., Chalker, J. M. & Davis, B. G. Olefin metathesis for site-selective protein modification. *Chembiochem* **10**, 959–969 (2009).

54. Binder, J. B. & Raines, R. T. Olefin Metathesis for Chemical Biology. *Curr Opin Chem Biol* **12**, 767 (2008).

55. Tietze, L. F. *et al.* Anticancer Agents, 15. Squaric Acid Diethyl Ester: A New Coupling Reagent for the Formation of Drug Biopolymer Conjugates. Synthesis of Squaric Acid Ester Amides and Diamides. *Chem Ber* **124**, 1215–1221 (1991).

56. Lin, S. *et al.* Redox-based reagents for chemoselective methionine bioconjugation. *Science (1979)* **355**, 597–602 (2017).

57. Press release: The Nobel Prize in Chemistry 2022 - NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/2022/press-release/.

58. *Click Chemistry: Diverse Chemical Function from a Few Good Reactions - Kolb - 2001 - Angewandte Chemie International Edition - Wiley Online Library*.

59. Fantoni, N. Z., El-Sagheer, A. H. & Brown, T. A Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. *Chem Rev* **121**, 7122–7154 (2021).

60. A Not-So-Ancient Grease History: Click Chemistry and Protein Lipid Modifications. *Chem Rev* **121**, 7178–7248 (2021).

61. Light-Triggered Click Chemistry. *Chem Rev* **121**, 6991–7031 (2021).

62. Bioorthogonal Reactions Utilizing Nitrones as Versatile Dipoles in Cycloaddition Reactions. *Chem Rev* **121**, 6699–6717 (2021).

63. Shieh, P., Hill, M. R., Zhang, W., Kristufek, S. L. & Johnson, J. A. Clip Chemistry: Diverse (Bio)(macro)molecular and Material Function through Breaking Covalent Bonds. *Chem Rev* **121**, 7059–7121 (2021).

64. Worch, J. C., Stubbs, C. J., Price, M. J. & Dove, A. P. Click Nucleophilic Conjugate Additions to Activated Alkynes: Exploring Thiol-yne, Amino-yne, and Hydroxyl-yne Reactions from (Bio)Organic to Polymer Chemistry. *Chem Rev* **121**, 6744–6776 (2021).

65. Porte, K., Riomet, M., Figliola, C., Audisio, D. & Taran, F. Click and Bio-Orthogonal Reactions with Mesoionic Compounds. *Chem Rev* **121**, 6718–6743 (2021).

66. Bioorthogonal Reactions of Triarylphosphines and Related Analogues. *Chem Rev* **121**, 6802–6849 (2021).

67. Mechanisms and Substituent Effects of Metal-Free Bioorthogonal Reactions. *Chem Rev* **121**, 6850–6914 (2021).

68. Fairbanks, B. D. *et al.* Photoclick Chemistry: A Bright Idea. *Chem Rev* **121**, 6915–6990 (2021).

69. Albada, B., Keijzer, J. F., Zuilhof, H. & van Delft, F. Oxidation-Induced "One-Pot" Click Chemistry. *Chem Rev* **121**, 7032–7058 (2021).

70. Click Chemistry with Cyclopentadiene. *Chem Rev* **121**, 6777–6801 (2021).

71. Scinto, S. L. *et al.* Bioorthogonal chemistry. *Nature Reviews Methods Primers* **1**, 30 (2021).

72. Shih, H. W., Kamber, D. N. & Prescher, J. A. Building better bioorthogonal reactions. *Curr Opin Chem Biol* **21**, 103–111 (2014).

73. Hashimoto, T. & Maruoka, K. Recent Advances of Catalytic Asymmetric 1,3-Dipolar Cycloadditions. *Chem Rev* **115**, 5366–5412 (2015).

74. Gothelf, K. V. & Jørgensen, K. A. Asymmetric 1,3-dipolar cycloaddition reactions. *Chem Rev* **98**, 863–909 (1998).

75. Huisgen, R., Szeimies, G. & Möbius, L. 1.3-Dipolare Cycloadditionen, XXXII. Kinetik der Additionen organischer Azide an CC-Mehrfachbindungen. *Chem Ber* **100**, 2494–2507 (1967).

76. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise huisgen cycloaddition process: Copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angewandte Chemie - International Edition* **41**, 2596–2599 (2002).

77. Padwa, A. 1,3-dipolar cycloaddition chemistry. (1984).

78. Huisgen, R. 1,3-Dipolar Cycloadditions. Past and Future. *Angewandte Chemie International Edition in English* **2**, 565–598 (1963).

79. Sustmann, R. A simple model for substituent effects in cycloaddition reactions. I. 1,3-dipolar cycloadditions. *Tetrahedron Lett* **12**, 2717–2720 (1971).

80. Kolb, H. C. & Sharpless, K. B. The growing impact of click chemistry on drug discovery. *Drug Discov Today* **8**, 1128–1137 (2003).

81. Lewis, W. G., Magallon, F. G., Fokin, V. V. & Finn, M. G. Discovery and characterization of catalysts for azide-alkyne cycloaddition by fluorescence quenching. *J Am Chem Soc* **126**, 9152–9153 (2004).

82. Mykhalichko, B. M., Temkin, O. N. & Mys'kiv, M. G. Polynuclear complexes of copper(I) halides: Coordination chemistry and catalytic transformations of alkynes. *Usp Khim* **69**, 1068–1070 (2000).

83. Diez, J. *et al.* Synthesis and characterization of triangulo copper(I) complexes containing mono- and bicapping systems of asymmetric .mu.3-.eta.1-acetylide ligands: molecular structures of [Cu3(.mu.3-

.eta.1-C.tplbond.CPh)(.mu.-dppm)3][BF4]2, [Cu3(.mu.3-.eta.1-C.tplbond.CPh)2(.mu.-dppm)3][BF4], and [Cu3(.mu.3-.eta.1-C.tplbond.CPh)(.mu.3-Cl)(.mu.-dppm)3][BF4] [dppm = bis(diphenylphosphino)methane]. *Organometallics* **12**, 2213–2220 (1993).

84. Rodionov, V. O., Fokin, V. V. & Finn, M. G. Mechanism of the Ligand-Free CuI-Catalyzed Azide–Alkyne Cycloaddition Reaction. *Angewandte Chemie International Edition* **44**, 2210–2215 (2005).

85. Ahlquist, M. & Fokin, V. V. Enhanced reactivity of dinuclear copper(I) acetylides in dipolar cycloadditions. *Organometallics* **26**, 4389–4391 (2007).

86. Straub, B. F. μ-Acetylide and μ-alkenylidene ligands in “click” triazole syntheses. *Chemical Communications* 3868 (2007) doi:10.1039/b706926j.

87. Berg, R. *et al.* Highly Active Dinuclear Copper Catalysts for Homogeneous Azide–Alkyne Cycloadditions. *Adv Synth Catal* **354**, 3445–3450 (2012).

88. Buckley, B. R., Dann, S. E., Harris, D. P., Heaney, H. & Stubbs, E. C. Alkynylcopper(I) polymers and their use in a mechanistic study of alkyne–azide click reactions. *Chemical Communications* **46**, 2274–2276 (2010).

89. Shao, C. *et al.* Copper(I) Acetate: A Structurally Simple but Highly Efficient Dinuclear Catalyst for Copper-Catalyzed Azide-Alkyne Cycloaddition. *Adv Synth Catal* **352**, 1587–1592 (2010).

90. Buckley, B. R., Dann, S. E. & Heaney, H. Experimental Evidence for the Involvement of Dinuclear Alkynylcopper(I) Complexes in Alkyne–Azide Chemistry. *Chemistry – A European Journal* **16**, 6278–6284 (2010).

91. Berg, R. & Straub, B. F. Advancements in the mechanistic understanding of the copper-catalyzed azide–alkyne cycloaddition. *Beilstein Journal of Organic Chemistry* **9**, 2715–2750 (2013).

92. Makarem, A., Berg, R., Rominger, F. & Straub, B. F. A Fluxional Copper Acetylide Cluster in CuAAC Catalysis. *Angewandte Chemie International Edition* **54**, 7431–7435 (2015).

93. Jin, L., Tolentino, D. R., Melaimi, M. & Bertrand, G. Isolation of bis(copper) key intermediates in Cu-catalyzed azide-alkyne "click reaction. *Sci Adv* **1**, (2015).

94. Héron, J. & Balcells, D. Concerted Cycloaddition Mechanism in the CuAAC Reaction Catalyzed by 1,8-Naphthyridine Dicopper Complexes. *ACS Catal* **12**, 4744–4753 (2022).

95. Keana, J. F. W. & Cai, S. X. New Reagents for Photoaffinity Labeling: Synthesis and Photolysis of Functionalized Perfluorophenyl Azides. *Journal of Organic Chemistry* **55**, 3640–3647 (1990).

96. Hoyle, C. E. & Bowman, C. N. Thiol-ene click chemistry. *Angew Chem Int Ed Engl* **49**, 1540–1573 (2010).

97. Declerck, V., Martinez, J. & Lamaty, F. Aza-Baylis - Hillman reaction. *Chem Rev* **109**, 1–48 (2009).

98. Brewer, G. J. Risks of copper and iron toxicity during aging in humans. *Chem Res Toxicol* **23**, 319–326 (2010).

99. Halliwell, B. & Gutteridge, J. M. C. [1] Role of free radicals and catalytic metal ions in human disease: An overview. *Methods Enzymol* **186**, 1–85 (1990).

100. Li, S. *et al.* Extent of the Oxidative Side Reactions to Peptides and Proteins During the CuAAC Reaction. *Bioconjug Chem* **27**, 2315–2322 (2016).

101. Lallana, E., Riguera, R. & Fernandez-Megia, E. Reliable and efficient procedures for the conjugation of biomolecules through Huisgen azide-alkyne cycloadditions. *Angew Chem Int Ed Engl* **50**, 8794–8804 (2011).

102. Kennedy, D. C. *et al.* Cellular Consequences of Copper Complexes Used To Catalyze Bioorthogonal Click Reactions. *J Am Chem Soc* **133**, 17993–18001 (2011).

103. Li, S. *et al.* Copper-catalyzed click reaction on/in live cells. *Chem Sci* **8**, 2107 (2017).

104. Chan, T. R., Hilgraf, R., Sharpless, K. B. & Fokin, V. V. Polytriazoles as copper(I)-stabilizing ligands in catalysis. *Org Lett* **6**, 2853–2855 (2004).

105. Hong, V., Steinmetz, N. F., Manchester, M. & Finn, M. G. Labeling live cells by copper-catalyzed alkyne-azide click chemistry. *Bioconjug Chem* **21**, 1912–1916 (2010).

106. Soriano Del Amo, D. *et al.* Biocompatible copper(I) catalysts for in vivo imaging of glycans. *J Am Chem Soc* **132**, 16893–16899 (2010).

107. Besanceney-Webler, C. *et al.* Increasing the Efficacy of Bioorthogonal Click Reactions for Bioconjugation: A Comparative Study. *Angewandte Chemie International Edition* **50**, 8051–8056 (2011).

108. Rudolf, G. C. & Sieber, S. A. Copper-Assisted Click Reactions for Activity-Based Proteomics: Fine-Tuned Ligands and Refined Conditions Extend the Scope of Application. *ChemBioChem* **14**, 2447–2455 (2013).

109. Neumann, S. *et al.* The CuAAC: Principles, Homogeneous and Heterogeneous Catalysts, and Novel Developments and Applications. *Macromol Rapid Commun* **41**, 1900359 (2020).

110. Gonda, Z. & Novák, Z. Highly active copper-catalysts for azide-alkyne cycloaddition. *Dalton Trans* **39**, 726–729 (2010).

111. Zhang, L. *et al.* Ruthenium-catalyzed cycloaddition of alkynes and organic azides. *J Am Chem Soc* **127**, 15998–15999 (2005).

112. Johansson, J. R., Lincoln, P., Nordén, B. & Kann, N. Sequential one-pot ruthenium-catalyzed azide-alkyne cycloaddition from primary alkyl halides and sodium azide. *Journal of Organic Chemistry* **76**, 2355–2359 (2011).

113. Rasmussen, L. K., Boren, B. C. & Fokin, V. V. Ruthenium-catalyzed cycloaddition of aryl azides and alkynes. *Org Lett* **9**, 5337–5339 (2007).

114. Pradere, U., Roy, V., McBrayer, T. R., Schinazi, R. F. & Agrofoglio, L. A. Preparation of ribavirin analogues by copper- and ruthenium-catalyzed azide-alkyne 1,3-dipolar cycloaddition. *Tetrahedron* **64**, 9044–9051 (2008).

115. Boren, B. C. *et al.* Ruthenium-catalyzed azide-alkyne cycloaddition: scope and mechanism. *J Am Chem Soc* **130**, 8923–8930 (2008).

116. Boz, E. & Tüzün, N. Ş. Reaction mechanism of ruthenium-catalyzed azide–alkyne cycloaddition reaction: A DFT study. *J Organomet Chem* **724**, 167–176 (2013).

117. Risse, J., Scopelliti, R. & Severin, K. Beyond click-chemistry: Transformation of azides with cyclopentadienyl ruthenium complexes. *Organometallics* **30**, 3412–3418 (2011).

118. Lamberti, M. *et al.* Coordinatively unsaturated ruthenium complexes as efficient alkyneazide cycloaddition catalysts. *Organometallics* **31**, 756–767 (2012).

119. Kim, E. & Koo, H. Biomedical applications of copper-free click chemistry: in vitro, in vivo, and ex vivo. *Chem Sci* **10**, 7835–7851 (2019).

120. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J Am Chem Soc* **126**, 15046–15047 (2004).

121. Sletten, E. M. & Bertozzi, C. R. From mechanism to mouse: A tale of two bioorthogonal reactions. *Acc Chem Res* **44**, 666–676 (2011).

122. Dommerholt, J. *et al.* Highly accelerated inverse electron-demand cycloaddition of electron-deficient azides with aliphatic cyclooctynes. *Nature Communications 2014 5:1* **5**, 1–7 (2014).

123. Kang, X., Cai, X., Yi, L. & Xi, Z. Multifluorinated Aryl Azides for the Development of Improved H2S Probes, and Fast Strain-promoted Azide-Alkyne Cycloaddition and Staudinger Reactions. *Chem Asian J* **15**, 1420–1429 (2020).

124. Reiner, T., Earley, S., Turetsky, A. & Weissleder, R. Bioorthogonal small-molecule ligands for PARP1 imaging in living cells. *Chembiochem* **11**, 2374–2377 (2010).

125. Enhsen, A., Kramer, W. & Wess, G. Bile acids in drug discovery. *Drug Discov Today* **3**, 409–418 (1998).

126. Kuipers, F., Bloks, V. W. & Groen, A. K. Beyond intestinal soap—bile acids in metabolic control. *Nature Reviews Endocrinology 2014 10:8* **10**, 488–498 (2014).

127. investigation, D. S.-T. J. of clinical & 1997, undefined. Point mutations in the ileal bile salt transporter cause leaks in the enterohepatic circulation leading to severe chronic diarrhea and malabsorption. *Am Soc Clin Investig* **21**, 109–114 (1995).

128. Petzinger, E. Transport of organic anions in the liver. An update on bile acid, fatty acid, monocarboxylate, anionic amino acid, cholephilic organic anion, and anionic drug transport. *REV.PHYSIOL.BIOCHEM.PHARMACOL.* **123**, 47–211 (1994).

129. Stedronsky, E. R. Interaction of bile acids and cholesterol with non-systemic agents having hypocholesterolemic properties. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1210**, 255–287 (1994).

130. Radominska, A. *et al.* Human liver steroid sulphotransferase sulphates bile acids. *Biochemical Journal* **272**, 597–604 (1990).

131. Carey, M. C. Cheno and Urso: What the Goose and the Bear Have in Common. *New England Journal of Medicine* **293**, 1255–1257 (1975).

132. Monte, M. J. *et al.* Bile acids: Chemistry, physiology, and pathophysiology. *World J Gastroenterol* **15**, (2009).

133. Hofmann, A. F. & Roda, A. Physicochemical properties of bile acids and their relationship to biological properties: an overview of the problem. *J Lipid Res* **25**, 1477–1489 (1984).

134. Hofmann, A. F. Bile acids: The good, the bad, and the ugly. *News in Physiological Sciences* **14**, 24–29 (1999).

135. Coello, A., Meijide, F., Rodríguez Núñez, E. & Tato, J. V. Aggregation Behavior of Sodium Cholate in Aqueous Solution. *J. Phys. Chem* **97**, 10186–10191 (1993).

136. Coello, A., Meijide, F., Rodríguez Núñez, E. & Vázquez Tato, J. V. Aggregation behavior of bile salts in aqueous solution. *J Pharm Sci* **85**, 9–15 (1996).

137. Armstrong, M. J. & Carey, M. C. The hydrophobic-hydrophilic balance of bile salts. Inverse correlation between reverse-phase high performance liquid chromatographic mobilities and micellar cholesterol-solubilizing capacities. *J Lipid Res* **23**, 70–80 (1982).

138. Tint, G. S. *et al.* Ursodeoxycholic acid: a safe and effective agent for dissolving cholesterol gallstones. *Ann Intern Med* **97**, 351–356 (1982).

139. Keely, S. J., Steer, C. J. & Lajczak-McGinley, N. K. Ursodeoxycholic acid: A promising therapeutic target for inflammatory bowel diseases? *Arq Neuropsiquiatr* **317**, G872–G881 (2019).

140. Keely, S. J., Steer, C. J. & Lajczak-McGinley, N. K. Ursodeoxycholic acid: A promising therapeutic target for inflammatory bowel diseases? *Arq Neuropsiquiatr* **317**, G872–G881 (2019).

141. Ackerman, H. D. & Gerhard, G. S. Bile Acids in Neurodegenerative Disorders. *Bile Acids in Neurodegenerative Disorders. Front. Aging Neurosci* **8**, 263 (2016).

142. Melchor-Mendoza, Y. K. *et al.* Ursodeoxycholic Acid Therapy in Patients with Primary Biliary Cholangitis with Limited Liver Transplantation Availability. *Ann Hepatol* **16**, 430–435 (2017).

143. Rudic, J. S., Poropat, G., Krstic, M. N., Bjelakovic, G. & Gluud, C. Ursodeoxycholic acid for primary biliary cirrhosis. *Cochrane Database Syst Rev* **2012**, (2012).

144. Golden, J. M. *et al.* Ursodeoxycholic acid protects against intestinal barrier breakdown by promoting enterocyte migration via EGFR-and COX-2-dependent mechanisms. *Am J Physiol Gastroin-test Liver Physiol* **315**, 259–271 (2018).

145. Wang, H., Chen, J., Hollister, K., Sowers, L. C. & Forman, B. M. Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. *Mol Cell* **3**, 543–553 (1999).

146. Parks, D. J. *et al.* Bile acids: natural ligands for an orphan nuclear receptor. *Science* **284**, 1365–1368 (1999).

147. Halilbasic, E., Claudel, T. & Trauner, M. Bile acid transporters and regulatory nuclear receptors in the liver and beyond. *J Hepatol* **58**, 155–168 (2013).

148. Kawamata, Y. *et al.* A G protein-coupled receptor responsive to bile acids. *J Biol Chem* **278**, 9435–9440 (2003).

149. Maruyama, T. *et al.* Identification of membrane-type receptor for bile acids (M-BAR). *Biochem Biophys Res Commun* **298**, 714–719 (2002).

150. Keitel, V. & Häussinger, D. Perspective: TGR5 (Gpbar-1) in liver physiology and disease. *Clin Res Hepatol Gastroenterol* **36**, 412–419 (2012).

151. Thomas, C. *et al.* Article TGR5-Mediated Bile Acid Sensing Controls Glucose Homeostasis. *Cell Metab* **10**, 167–177.

152. Zhou, X. *et al.* Structural basis of the alternating-access mechanism in a bile acid transporter. *Nature* **505**, 569–573 (2014).

153. Hu, N. J., Iwata, S., Cameron, A. D. & Drew, D. Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature* **478**, 408–411 (2011).

154. Kim, K. S., Suzuki, K., Cho, H., Youn, Y. S. & Bae, Y. H. Oral Nanoparticles Exhibit Specific High-Efficiency Intestinal Uptake and Lymphatic Transport. *ACS Nano* **12**, 8893–8900 (2018).

155. Al-Hilal, T. A. *et al.* Functional transformations of bile acid transporters induced by high-affinity macromolecules. *Sci Rep* **4**, (2014).

156. Lack, L. & Weiner, I. M. Intestinal bile salt transport: structure-activity relationships and other properties. *Am J Physiol* **210**, 1142–1152 (1966).

157. Kramer, W. *et al.* Substrate specificity of the ileal and the hepatic Na(+)/bile acid cotransporters of the rabbit. I. Transport studies with membrane vesicles and cell lines expressing the cloned transporters. *J Lipid Res* **40**, 1604–17 (1999).

158. Baringhaus, K. H., Matter, H., Stengelin, S. & Kramer, W. Substrate specificity of the ileal and the hepatic Na(+)/bile acid cotransporters of the rabbit. II. A reliable 3D QSAR pharmacophore model for the ileal Na(+)/bile acid cotransporter. *J Lipid Res* **40**, 2158–68 (1999).

159. Tolle-Sander, S., Lentz, K. A., Maeda, D. Y., Coop, A. & Polli, J. E. Increased acyclovir oral bioavailability via a bile acid conjugate. *Mol Pharm* **1**, 40–48 (2004).

160. Kramer, W. & Wess, G. Bile acid transport systems as pharmaceutical targets. *Eur J Clin Invest* **26**, 715–732 (1996).

161. Kramer, W. *et al.* Liver-specific drug targeting by coupling to bile acids. *Journal of Biological Chemistry* **267**, 18598–18604 (1992).

162. Kullak-Ublick, G. A. *et al.* Chlorambucil-taurocholate is transported by bile acid carriers expressed in human hepatocellular carcinomas. *Gastroenterology* **113**, 1295–1305 (1997).

163. Kramer, W. *et al.* Intestinal absorption of peptides by coupling to bile acids. *Journal of Biological Chemistry* **269**, 10621–10627 (1994).

164. KRAMER, W. *et al.* Bile acid derived HMG-CoA reductase inhibitors. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1227**, 137–154 (1994).

165. Melloni, E. *et al.* Synthesis and Biological Investigation of Bile Acid-Paclitaxel Hybrids. *Molecules* **27**, 471 (2022).

166. Rohena, C. C. & Mooberry, S. L. Recent progress with microtubule stabilizers: New compounds, binding modes and cellular activities. *Nat Prod Rep* **31**, 335–355 (2014).

167. Dumontet, C. & Jordan, M. A. Microtubule-binding agents: A dynamic field of cancer therapeutics. *Nat Rev Drug Discov* **9**, 790–803 (2010).

168. Derry, W. B., Wilson, L. & Jordan, M. A. Substoichiometric Binding of Taxol Suppresses Microtubule Dynamics. *Biochemistry* **34**, 2203–2211 (1995).

169. Jordan, M. A., Toso, R. J., Thrower, D. & Wilson, L. Mechanism of mitotic block and inhibition of cell proliferation by taxol at low concentrations. *Proc Natl Acad Sci U S A* **90**, 9552–9556 (1993).

170. Xiao, H. *et al.* Insights into the mechanism of microtubule stabilization by Taxol. *Proc Natl Acad Sci U S A* **103**, 10166–10173 (2006).

171. Meng, Z. *et al.* Prodrug strategies for paclitaxel. *Int J Mol Sci* **17**, (2016).

172. Zhang, S. *et al.* Poly(ethylene oxide)-block-polyphosphester-based paclitaxel conjugates as a platform for ultra-high paclitaxel-loaded multifunctional nanoparticles. *Chem Sci* **4**, 2122–2126 (2013).

173. Correard, F. *et al.* Delaying Anticancer Drug Delivery by Self-Assembly and Branching Effects of Minimalist Dendron-Drug Conjugates. *Chemistry* **25**, 9586–9591 (2019).

174. Wang, F., Porter, M., Konstantopoulos, A., Zhang, P. & Cui, H. Preclinical development of drug delivery systems for paclitaxel-based cancer chemotherapy. *Journal of Controlled Release* **267**, 100–118 (2017).

175. Jordan, M. A. & Wilson, L. Microtubules as a target for anticancer drugs. *Nat Rev Cancer* **4**, 253–265 (2004).

176. Gelderblom, H., Verweij, J., Nooter, K. & Sparreboom, A. Cremophor EL: The drawbacks and advantages of vehicle selection for drug formulation. *Eur J Cancer* **37**, 1590–1598 (2001).

177. Liu, C. *et al.* Design, Synthesis, and Bioactivities of Steroid-Linked Taxol Analogues as Potential Targeted Drugs for Prostate and Breast Cancer. *J Nat Prod* **67**, 152–159 (2004).

178. Wittman, M. D. *et al.* Synthesis and antitumor activity of novel paclitaxel-chlorambucil hybrids. *Bioorg Med Chem Lett* **11**, 811–814 (2001).

179. Kingston, D. G. I. & Snyder, J. P. The quest for a simple bioactive analog of paclitaxel as a potential anticancer agent. *Acc Chem Res* **47**, 2682–2691 (2014).

180. Daniel, J. *et al.* Hydrophilic Fluorescent Nanoprodrug of Paclitaxel for Glioblastoma Chemotherapy. *ACS Omega* **4**, 18342–18354 (2019).

181. Roussi, F., Ngo, Q. A., Thoret, S., Guéritte, F. & Guénard, D. The Design and Synthesis of New Steroidal Compounds as Potential Mimics of Taxoids. *European J Org Chem* **2005**, 3952–3961 (2005).

182. Salvador, J. A. R. *et al.* Anticancer steroids: Linking natural and semi-synthetic compounds. *Nat Prod Rep* **30**, 324–374 (2013).

183. Barrasa, J. I., Olmo, N., Lizarbe, M. A. & Turnay, J. Bile acids in the colon, from healthy to cytotoxic molecules. *Toxicology in Vitro* **27**, 964–977 (2013).

184. Goossens, J. F. & Bailly, C. Ursodeoxycholic acid and cancer: From chemoprevention to chemotherapy. *Pharmacol Ther* **203**, (2019).

185. Negahban, Z., Shojaosadati, S. A. & Hamedi, S. A novel self-assembled micelles based on stearic acid modified schizophyllan for efficient delivery of paclitaxel. *Colloids Surf B Biointerfaces* **199**, (2021).

186. Arduino, I. *et al.* Preparation of cetyl palmitate-based PEGylated solid lipid nanoparticles by microfluidic technique. *Acta Biomater* **121**, 566–578 (2021).

187. Grigoletto, A. *et al.* Folic acid-targeted paclitaxel-polymer conjugates exert selective cytotoxicity and modulate invasiveness of colon cancer cells. *Pharmaceutics* **13**, (2021).

188. Pavlović, N. *et al.* Bile acids and their derivatives as potential modifiers of drug release and pharmacokinetic profiles. *Front Pharmacol* **9**, (2018).

189. Lee, M. M., Gao, Z. & Peterson, B. R. Synthesis of a Fluorescent Analogue of Paclitaxel That Selectively Binds Microtubules and Sensitively Detects Efflux by P-Glycoprotein. *Angewandte Chemie - International Edition* **56**, 6927–6931 (2017).

190. Zhang, J. liang *et al.* Structural analysis and antitussive evaluation of five novel esters of verticinone and bile acids. *Steroids* **74**, 424–434 (2009).

191. The Nobel Prize in Physiology or Medicine 2015. https://www.nobelprize.org/prizes/medicine/2015/summary/.

192. Chang, Z. The discovery of Qinghaosu (artemisinin) as an effective anti-malaria drug: A unique China story. *Sci China Life Sci* **59**, 81–88 (2016).

193. Schmid, G. & Hofheinz, W. Total Synthesis of Qinghaosu. *J Am Chem Soc* **105**, 624–625 (1983).

194. Arsenault, P., Wobbe, K. & Weathers, P. Recent Advances in Artemisinin Production Through Heterologous Expression. *Curr Med Chem* **15**, 2886–2896 (2008).

195. Kumar, A. & Bishnoi, A. K. One-pot green synthesis of β-artemether/arteether. *RSC Adv* **4**, 31973–31976 (2014).

196. Presser, A., Feichtinger, A. & Buzzi, S. A simplified and scalable synthesis of artesunate. *Monatsh Chem* **148**, 63 (2017).

197. Wang, J. *et al.* Artemisinin, the Magic Drug Discovered from Traditional Chinese Medicine. *Engineering* **5**, 32–39 (2019).

198. Zhang, S. & Gerhard, G. S. Heme activates artemisinin more efficiently than hemin, inorganic iron, or hemoglobin. *Bioorg Med Chem* **16**, 7853–7861 (2008).

199. Jefford, C. W. *et al.* The Deoxygenation and Isomerization of Artemisinin and Artemether and Their Relevance to Antimalarial Action. *Helv Chim Acta* **79**, 1475–1487 (1996).

200. Creek, D. J., Chiu, F. C. K., Prankerd, R. J., Charman, S. A. & Charman, W. N. Kinetics of iron-mediated artemisinin degradation: Effect of solvent composition and iron salt. *J Pharm Sci* **94**, 1820–1829 (2005).

201. O'Neill, P. M., Barton, V. E. & Ward, S. A. The Molecular Mechanism of Action of Artemisinin—The Debate Continues. *Molecules 2010, Vol. 15, Pages 1705-1721* **15**, 1705–1721 (2010).

202. Wang, J. *et al.* Artemisinin Directly Targets Malarial Mitochondria through Its Specific Mitochondrial Activation. *PLoS One* **5**, e9582 (2010).

203. Sun, C., Li, J., Cao, Y., Long, G. & Zhou, B. Two distinct and competitive pathways confer the cellcidal actions of artemisinins. *Microbial Cell* **2**, 14 (2015).

204. Chaturvedi, D., Goswami, A., Saikia, P. P., Barua, N. C. & Rao, P. G. Artemisinin and its derivatives: a novel class of anti-malarial and anti-cancer agents. *Chem Soc Rev* **39**, 435–454 (2010).

205. Singh, N. P. & Lai, H. C. Artemisinin Induces Apoptosis in Human Cancer Cells. *Anticancer Res* **24**, 2277–2280 (2004).

206. Efferth, T. Mechanistic perspectives for 1,2,4-trioxanes in anti-cancer therapy. *Drug Resistance Updates* **8**, 85–97 (2005).

207. Cabello, C. M. *et al.* The redox antimalarial dihydroartemisinin targets human metastatic melanoma cells but not primary melanocytes with induction of NOXA-dependent apoptosis. *Invest New Drugs* **30**, 1289 (2012).

208. Efferth, T. Willmar Schwabe Award 2006: antiplasmodial and antitumor activity of artemisinin--from bench to bedside. *Planta Med* **73**, 299–309 (2007).

209. Jung, M., Lee, K., Kim, H. & Park, M. Recent Advances in Artemisinin and Its Derivatives as Antimalarial and Antitumor Agents. *Curr Med Chem* **11**, 1265–1284 (2012).

210. Efferth, T. *et al.* The antiviral activities of artemisinin and artesunate. *Clin Infect Dis* **47**, 804–11 (2008).

211. Raffetin, A. *et al.* Use of artesunate in non-malarial indications. *Med Mal Infect* **48**, 238–249 (2018).

212. Oiknine-Djian, E. *et al.* The Artemisinin Derivative Artemisone Is a Potent Inhibitor of Human Cytomegalovirus Replication. *Antimicrob Agents Chemother* **62**, (2018).

213. Han, Y., Pham, H. T., Xu, H., Quan, Y. & Mesplède, T. Antimalarial drugs and their metabolites are potent Zika virus inhibitors. *J Med Virol* **91**, 1182–1190 (2019).

214. Oiknine-Djian, E. *et al.* Artemisone demonstrates synergistic antiviral activity in combination with approved and experimental drugs active against human cytomegalovirus. *Antiviral Res* **172**, 104639 (2019).

215. Farmanpour-Kalalagh, K., Beyraghdar Kashkooli, A., Babaei, A., Rezaei, A. & van der Krol, A. R. Artemisinins in Combating Viral Infections Like SARS-CoV-2, Inflammation and Cancers and Options to Meet Increased Global Demand. *Front Plant Sci* **13**, 2 (2022).

216. Cao, R. *et al.* Anti-SARS-CoV-2 potential of artemisinins in vitro. *ACS Infect Dis* **6**, 2524–2531 (2020).

217. Soni, R., Shankar, G., Mukhopadhyay, P. & Gupta, V. A concise review on Artemisia annua L.: A major source of diverse medicinal compounds. *Ind Crops Prod* **184**, 115072 (2022).

218. Gendrot, M. *et al.* Antimalarial artemisinin-based combination therapies (ACT) and COVID-19 in Africa: In vitro inhibition of SARS-CoV-2 replication by mefloquine-artesunate. *International Journal of Infectious Diseases* **99**, 437–440 (2020).

219. Weekly epidemiological update on COVID-19 - 1 September 2023. https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---1-september-2023.

220. Wang, C. *et al.* Long COVID: The Nature of Thrombotic Sequelae Determines the Necessity of Early Anticoagulation. *Front Cell Infect Microbiol* **12**, 408 (2022).

221. Zamboni, P. *et al.* Bowel ischemia as onset of COVID-19 in otherwise asymptomatic patients with persistently negative swab. *J Intern Med* **291**, 224–231 (2022).

222. Mehandru, S. & Merad, M. Pathological sequelae of long-haul COVID. *Nature Immunology 2022 23:2* **23**, 194–202 (2022).

223. Bortolotti, D. *et al.* Relevance of VEGF and CD147 in different SARS-CoV-2 positive digestive tracts characterized by thrombotic damage. *The FASEB Journal* **35**, e21969 (2021).

224. Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results. *New England Journal of Medicine* **384**, 497–511 (2021).

225. COVID-19 Treatment Guidelines. https://www.covid19treatmentguidelines.nih.gov/.

226. Cheong, D. H. J., Tan, D. W. S., Wong, F. W. S. & Tran, T. Anti-malarial drug, artemisinin and its derivatives for the treatment of respiratory diseases. *Pharmacol Res* **158**, (2020).

227. Uzun, T. & Toptas, O. Artesunate: could be an alternative drug to chloroquine in COVID-19 treatment? *Chin Med* **15**, 54 (2020).

228. Ho, W. E., Peh, H. Y., Chan, T. K. & Wong, W. S. F. Artemisinins: pharmacological actions beyond anti-malarial. *Pharmacol Ther* **142**, 126–139 (2014).

229. Gambino, D. Editorial: Development/repurposing of drugs to tackle the multiple variants of SARS-CoV-2. *Frontiers in Drug Discovery* **3**, 5 (2023).

230. Xu, C., Zhang, H., Mu, L. & Yang, X. Artemisinins as Anticancer Drugs: Novel Therapeutic Approaches, Molecular Mechanisms, and Clinical Trials. *Front Pharmacol* **11**, 1608 (2020).

231. Meunier, B. Hybrid molecules with a dual mode of action: Dream or reality? *Acc Chem Res* **41**, 69–77 (2008).

232. Perrone, D. *et al.* Synthesis and in vitro cytotoxicity of deoxyadenosine–bile acid conjugates linked with 1,2,3-triazole. *New Journal of Chemistry* **37**, 3559–3567 (2013).

233. Navacchia, M. L. *et al.* NO Photoreleaser-Deoxyadenosine and -Bile Acid Derivative Bioconjugates as Novel Potential Photochemotherapeutics. *ACS Med Chem Lett* **7**, 939–943 (2016).

234. Navacchia, M. L. *et al.* Rational Design of Nucleoside–Bile Acid Conjugates Incorporating a Triazole Moiety for Anticancer Evaluation and SAR Exploration. *Molecules : A Journal of Synthetic Chemistry and Natural Product Chemistry* **22**, (2017).

235. Melloni, E. *et al.* Synthesis and Biological Investigation of Bile Acid-Paclitaxel Hybrids. *Molecules* **27**, (2022).

236. Marchesi, E. *et al.* Dihydroartemisinin-Bile Acid Hybridization as an Effective Approach to Enhance Dihydroartemisinin Anticancer Activity. *ChemMedChem* **14**, 779–787 (2019).

237. Teissier, E., Penin, F. & Pécheur, E. I. Targeting cell entry of enveloped viruses as an antiviral strategy. *Molecules* **16**, 221–250 (2010).

238. Wisskirchen, K., Lucifora, J., Michler, T. & Protzer, U. New pharmacological strategies to fight enveloped viruses. *Trends Pharmacol Sci* **35**, 470 (2014).

239. Rizzo, R. *et al.* HHV-6A/6B Infection of NK Cells Modulates the Expression of miRNAs and Transcription Factors Potentially Associated to Impaired NK Activity. *Front Microbiol* **8**, (2017).

240. Yang, J. *et al.* Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor. *Nature Communications 2020 11:1* **11**, 1–10 (2020).

241. Carino, A. *et al.* Hijacking SARS-CoV-2/ACE2 receptor interaction by natural and semi-synthetic steroidal agents acting on functional pockets on the receptor binding domain. *Front Chem* **8**, 1–15 (2020).

242. Fiorucci, S. & Distrutti, E. The Pharmacology of Bile Acids and Their Receptors. *Handb Exp Pharmacol* **256**, 3–18 (2019).

243. Li, T. & Chiang, J. Y. L. Bile acid signaling in metabolic disease and drug therapy. *Pharmacol Rev* **66**, 948–983 (2014).

244. Copple, B. L. & Li, T. Pharmacology of bile acid receptors: Evolution of bile acids from simple detergents to complex signaling molecules. *Pharmacol Res* **104**, 9–21 (2016).

245. Brevini, T. *et al.* FXR inhibition may protect from SARS-CoV-2 infection by reducing ACE2. *Nature 2022 615:7950* **615**, 134–142 (2022).

246. Huang, T. E. *et al.* Evaluation of the Anticancer Activity of a Bile Acid-Dihydroartemisinin Hybrid Ursodeoxycholic-Dihydroartemisinin in Hepatocellular Carcinoma Cells. *Front Pharmacol* **11**, (2020).

247. Hsu, Y.-F. ; *et al.* Anticancer Activity and Molecular Mechanisms of an Ursodeoxycholic Acid Methyl Ester-Dihydroartemisinin Hybrid via a Triazole Linkage in Hepatocellular Carcinoma Cells. *Molecules 2023, Vol. 28, Page 2358* **28**, 2358 (2023).

248. Bakouny, Z. *et al.* COVID-19 and Cancer: Current Challenges and Perspectives. *Cancer Cell* **38**, 629–646 (2020).

249. Bora, V. R. & Patel, B. M. The Deadly Duo of COVID-19 and Cancer! *Front Mol Biosci* **8**, 196 (2021).

250. Saini, G. & Aneja, R. Cancer as a prospective sequela of long COVID-19. *BioEssays* **43**, 2000331 (2021).

251. Xu, Z., Zhao, S. J. & Liu, Y. 1,2,3-Triazole-containing hybrids as potential anticancer agents: Current developments, action mechanisms and structure-activity relationships. *Eur J Med Chem* **183**, (2019).

252. Massarenti, C. *et al.* Fluorous-tag assisted synthesis of bile acid–bisphosphonate conjugates via orthogonal click reactions: an access to potential anti-resorption bone drugs. *Org Biomol Chem* **15**, 4907–4920 (2017).

253. Perrone, D. *et al.* Synthesis and in vitro cytotoxicity of deoxyadenosine–bile acid conjugates linked with 1,2,3-triazole. *New Journal of Chemistry* **37**, 3559–3567 (2013).

254. Batesky, D. C., Goldfogel, M. J. & Weix, D. J. Removal of triphenylphosphine oxide by precipitation with zinc chloride in polar solvents. *Journal of Organic Chemistry* **82**, 9931–9936 (2017).

255. Parapini, S., Olliaro, P., Navaratnam, V., Taramelli, D. & Basilico, N. Stability of the antimalarial drug dihydroartemisinin under physiologically relevant conditions: Implications for clinical treatment and pharmacokinetic and in vitro assays. *Antimicrob Agents Chemother* **59**, 4046–4052 (2015).

256. Kapkoti, D. S., Singh, S., Luqman, S. & Bhakuni, R. S. Synthesis of novel 1,2,3-triazole based artemisinin derivatives and their antiproliferative activity. *New Journal of Chemistry* **42**, 5978–5995 (2018).

257. Marchesi, E. *et al.* Dihydroartemisinin-Bile Acid Hybridization as an Effective Approach to Enhance Dihydroartemisinin Anticancer Activity. *ChemMedChem* **14**, 779–787 (2019).

258. Festa, C. *et al.* Exploitation of cholane scaffold for the discovery of potent and selective farnesoid X receptor (FXR) and G-protein coupled bile acid receptor 1 (GP-BAR1) ligands. *J Med Chem* **57**, 8477–8495 (2014).

259. Navacchia, M. L. *et al.* Rational Design of Nucleoside–Bile Acid Conjugates Incorporating a Triazole Moiety for Anticancer Evaluation and SAR Exploration. *Molecules : A Journal of Synthetic Chemistry and Natural Product Chemistry* **22**, (2017).

260. Presser, A., Feichtinger, A. & Buzzi, S. A simplified and scalable synthesis of artesunate. *Monatsh Chem* **148**, 63–68 (2017).

261. Lewin, B. *Genes seven*. (Oxford University Press, 2000).

262. Alberts, B. *et al. DNA Repair*. (Garland Science, 2002).

263. Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature 1953 171:4356* **171**, 737–738 (1953).

264. Mahmuh, J., Rownd, R. & Schildkbatjt, C. L. Denaturation and Renaturation of Deoxyribonucleic Acid12. *Prog Nucleic Acid Res Mol Biol* **1**, 231–300 (1963).

265. Beadle, G. W. & Tatum, E. L. Genetic Control of Biochemical Reactions in Neurospora*. *Proceedings of the National Academy of Sciences* **27**, 499–506 (1941).

266. Strauss, B. S. Biochemical Genetics and Molecular Biology: The Contributions of George Beadle and Edward Tatum. *Genetics* **203**, 13–20 (2016).

267. Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33 (2000).

268. Crooke, S. T. Molecular mechanisms of action of antisense drugs. *Biochim Biophys Acta* **1489**, 31–44 (1999).

269. Flanagan, W. M. *et al.* A cytosine analog that confers enhanced potency to antisense oligonucleotides. *Proceedings of the National Academy of Sciences* **96**, 3513–3518 (1999).

270. Liebhaber, S. A., Cash, F. E. & Shakin, S. H. Translationally associated helix-destabilizing activity in rabbit reticulocyte lysate. *Journal of Biological Chemistry* **259**, 15597–15602 (1984).

271. Szymkowski, D. E. Developing antisense oligonucleotides from the laboratory to clinical trials. *Drug Discov Today* **1**, 415–428 (1996).

272. Wagner, R. W. Gene inhibition using antisense oligodeoxynucleotides. *Nature* **372**, 333–335 (1994).

273. Zamecnik, P. C. & Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc Natl Acad Sci U S A* **75**, 280 (1978).

274. Pang, J., Guo, Q. & Lu, Z. The catalytic mechanism, metal dependence, substrate specificity, and biodiversity of ribonuclease H. *Front Microbiol* **13**, 1034811 (2022).

275. Soutschek, J. *et al.* Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **432**, 173–178 (2004).

276. Dana, H. *et al.* Molecular Mechanisms and Biological Functions of siRNA. *Int J Biomed Sci* **13**, 48–57 (2017).

277. Bennett, C. F. & Swayze, E. E. RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu Rev Pharmacol Toxicol* **50**, 259–293 (2010).

278. Hélène, C. & Toulmé, J. J. Specific regulation of gene expression by antisense, sense and antigene nucleic acids. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1049**, 99–125 (1990).

279. Cerritelli, S. M. & Crouch, R. J. Ribonuclease H: The enzymes in eukaryotes. *FEBS Journal* **276**, 1494–1505 (2009).

280. Broccoli, S. *et al.* Effects of RNA polymerase modifications on transcription-induced negative supercoiling and associated R-loop formation. *Mol Microbiol* **52**, 1769–1779 (2004).

281. Hyjek, M., Figiel, M. & Nowotny, M. RNases H: Structure and mechanism. *DNA Repair (Amst)* **84**, 102672 (2019).

282. Dias, N. & Stein, C. A. Antisense oligonucleotides: Basic concepts and mechanisms. *Mol Cancer Ther* **1**, 347–355 (2002).

283. Lima, W. F. *et al.* The Positional Influence of the Helical Geometry of the Heteroduplex Substrate on Human RNase H1 Catalysis. *Mol Pharmacol* **71**, 73–82 (2007).

284. Monia, B. P. *et al.* Evaluation of 2'-modified oligonucleotides containing 2'-deoxy gaps as antisense inhibitors of gene expression. *Journal of Biological Chemistry* **268**, 14514–14522 (1993).

285. Kumar, P. *et al.* Transvascular delivery of small interfering RNA to the central nervous system. *Nature 2007 448:7149* **448**, 39–43 (2007).

286. Kumar, P. *et al.* T cell-specific siRNA delivery suppresses HIV-1 infection in humanized mice. *Cell* **134**, 577–586 (2008).

287. Spitali, P. *et al.* Exon skipping-mediated dystrophin reading frame restoration for small mutations. *Hum Mutat* **30**, 1527–1534 (2009).

288. Echevarría, L., Aupy, P. & Goyenvalle, A. Exon-skipping advances for Duchenne muscular dystrophy. *Hum Mol Genet* **27**, R163–R172 (2018).

289. Goyenvalle, A. *et al.* Functional correction in mouse models of muscular dystrophy using exon-skipping tricyclo-DNA oligomers. *Nat Med* **21**, 270–275 (2015).

290. Shimo, T., Maruyama, R. & Yokota, T. Designing Effective Antisense Oligonucleotides for Exon Skipping. in *Methods in Molecular Biology* vol. 1687 143–155 (Humana Press Inc., 2018).

291. Li, D., Mastaglia, F. L., Fletcher, S. & Wilton, S. D. Precision Medicine through Antisense Oligonucleotide-Mediated Exon Skipping. *Trends Pharmacol Sci* **39**, 982–994 (2018).

292. Juliano, R. L. The delivery of therapeutic oligonucleotides. *Nucleic Acids Res* **44**, 6518–6548 (2016).

293. Bartolucci, D., Pession, A., Hrelia, P. & Tonelli, R. Precision Anti-Cancer Medicines by Oligonucleotide Therapeutics in Clinical Research Targeting Undruggable Proteins and Non-Coding RNAs. *Pharmaceutics* **14**, (2022).

294. Zhang, Z. *et al.* Dumbbell-Shaped Antisense Oligonucleotide Prodrugs Showed Improved Antinuclease Stability and Anticancer Efficacy. *Mol Pharm* **19**, 3915–3921 (2022).

295. Antisense Pipeline | Ionis. https://www.ionispharma.com/ionis-technology/antisense-pipeline/.

296. Egli, M. & Manoharan, M. Chemistry, structure and function of approved oligonucleotide therapeutics. *Nucleic Acids Res* **51**, 2529–2573 (2023).

297. Shokrzadeh, N., Winkler, A. M., Dirin, M. & Winkler, J. Oligonucleotides conjugated with short chemically defined polyethylene glycol chains are efficient antisense agents. *Bioorg Med Chem Lett* **24**, 5758 (2014).

298. Rimessi, P. *et al.* Cationic PMMA nanoparticles bind and deliver antisense oligoribonucleotides allowing restoration of dystrophin expression in the mdx mouse. *Molecular Therapy* **17**, 820–827 (2009).

299. Dean, D. A. Peptide nucleic acids: versatile tools for gene therapy strategies. *Adv Drug Deliv Rev* **44**, 81 (2000).

300. Stein, C. A. & Cheng, Y. C. Antisense oligonucleotides as therapeutic agents--is the bullet really magical? *Science* **261**, 1004–1012 (1993).

301. Matteucci, M. Structural modifications toward improved antisense oligonucleotides. *Perspectives in Drug Discovery and Design* **4**, 1–16 (1996).

302. Uhlmann, E. & Peyman, A. Antisense Oligonucleotides: A New Therapeutic Principle. *Chem Rev* **90**, 543–584 (1990).

303. Campbell, J. M., Bacon, T. A. & Wickstrom, E. Oligodeoxynucleoside phosphorothioate stability in subcellular extracts, culture media, sera and cerebrospinal fluid. *J Biochem Biophys Methods* **20**, 259–267 (1990).

304. Neckers, L. M. Cellular Internalization of Oligodeoxynucleotides. in *Gene Therapeutics* 180–192 (Birkhäuser Boston, 1994). doi:10.1007/978-1-4684-6822-9_10.

305. Jaroszewski, J. W. & Cohen, J. S. Cellular uptake of antisense oligodeoxynucleotides. *Adv Drug Deliv Rev* **6**, 235–250 (1991).

306. Crooke, R. M. In vitro toxicology and pharmacokinetics of antisense oligonucleotides. *Anticancer Drug Des* **6**, 609–46 (1991).

307. Smith, R. A. *et al.* Antisense oligonucleotide therapy for neurodegenerative disease. *J Clin Invest* **116**, 2290–2296 (2006).

308. Crooke, S. T. & Lebleu, B. *Antisense Research and Applications Edited by Library of Congress Cataloging-in-Publication Data*.

309. Ghosh, M. K., Ghosh, K., Dahl, O. & Cohen, J. S. Evaluation of some properties of a phosphorodithioate oligodeoxyribonucleotide for antisense application. *Nucleic Acids Res* **21**, 5761–5766 (1993).

310. Epple, S. *et al.* A New 1,5-Disubstituted Triazole DNA Backbone Mimic with Enhanced Polymerase Compatibility. *J Am Chem Soc* **143**, 16293–16301 (2021).

311. Inoue, H., Hayase, Y., Iwai, S. & Ohtsuka, E. Sequence-dependent hydrolysis of RNA using modified oligonucleotide splints and RNase H. *FEBS Lett* **215**, 327–330 (1987).

312. Hill, A. C. & Hall, J. The MOE Modification of RNA: Origins and Widescale Impact on the Oligonucleotide Therapeutics Field. *Helv Chim Acta* **106**, e202200169 (2023).

313. De Bouvere, B. *et al.* Hexitol Nucleic Acids (HNA): Synthesis and Properties. *Nucleosides Nucleotides Nucleic Acids* **16**, 973–976 (1997).

314. Morita, K. & Koizumi, M. Synthesis of ENA Nucleotides and ENA Oligonucleotides. *Curr Protoc Nucleic Acid Chem* **72**, 4.79.1-4.79.21 (2018).

315. Singh, S. K., Kumar, R. & Wengel, J. Synthesis of Novel Bicyclo[2.2.1] Ribonucleosides: 2‘-Amino- and 2‘-Thio-LNA Monomeric Nucleosides. *J Org Chem* **63**, 6078–6079 (1998).

316. Soler-Bistué, A., Zorreguieta, A. & Tolmasky, M. E. Bridged Nucleic Acids Reloaded. *Molecules* **24**, (2019).

317. Kotkowiak, W., Wengel, J., Scotton, C. J. & Pasternak, A. Improved RE31 Analogues Containing Modified Nucleic Acid Monomers: Thermodynamic, Structural, and Biological Effects. *J Med Chem* **62**, 2499–2507 (2019).

318. Watts, J. K. Locked nucleic acid: Tighter is different. *Chemical Communications* **49**, 5618–5620 (2013).

319. Hagedorn, P. H. *et al.* Locked nucleic acid: modality, diversity, and drug discovery. *Drug Discov Today* **23**, 101–114 (2018).

320. Liczner, C., Duke, K., Juneau, G., Egli, M. & Wilds, C. J. Beyond ribose and phosphate: Selected nucleic acid modifications for structure–function investigations and therapeutic applications. *Beilstein Journal of Organic Chemistry 17:76* **17**, 908–931 (2021).

321. Maiti, M. *et al.* Xylonucleic acid: synthesis, structure, and orthogonal pairing properties. *Nucleic Acids Res* **43**, 7189 (2015).

322. Crooke, S. T. *Antisense drug technology : principles, strategies, and applications*. (CRC Press, 2008).

323. Mönkkönen, J. & Urtti, A. Lipid fusion in oligonucleotide and gene delivery with cationic lipids. *Adv Drug Deliv Rev* **34**, 37–49 (1998).

324. Zhao, B., Tian, Q., Bagheri, Y. & You, M. Lipid–oligonucleotide conjugates for simple and efficient cell membrane engineering and bioanalysis. *Curr Opin Biomed Eng* **13**, 76–83 (2020).

325. Habus, I., Zhao, Q. & Agrawal, S. Synthesis, Hybridization Properties, Nuclease Stability, and Cellular Uptake of the Oligonucleotide-Amino-β-Cyclodextrins and Adamantane Conjugates. *Bioconjug Chem* **6**, 327–331 (1995).

326. Stetsenko, D. A. & Gait, M. J. A convenient solid-phase method for synthesis of 3'-conjugates of oligonucleotides. *Bioconjug Chem* **12**, 576–586 (2001).

327. Marchesi, E. *et al.* Synthesis and exon-skipping properties of a 3'-ursodeoxycholic acid-conjugated oligonucleotide targeting dmd pre-mrna: Pre-synthetic versus post- synthetic approach. *Molecules* **26**, 7662 (2021).

328. Raouane, M., Desmaële, D., Urbinati, G., Massaad-Massade, L. & Couvreur, P. Lipid conjugated oligonucleotides: A useful strategy for delivery. *Bioconjug Chem* **23**, 1091–1104 (2012).

329. Merrifield, R. B. Solid-phase peptide synthesis. *Adv Enzymol Relat Areas Mol Biol* **32**, 221–296 (1969).

330. Matteucci, M. D. & Caruthers, M. H. Synthesis of deoxyoligonucleotides on a polymer support. *J Am Chem Soc* **103**, 3185–3191 (1981).

331. Hovinen, J., Guzaev, A., Azhayev, A. & Lönnberg, 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* **50**, 7203–7218 (1994).

332. Marchesi, E. *et al.* Antisense Oligonucleotides Conjugated with Lipophilic Compounds: Synthesis and In Vitro Evaluation of Exon Skipping in Duchenne Muscular Dystrophy. *Int J Mol Sci* **23**, (2022).

333. Ravikumar, V. T. *et al.* UnyLinker: An Efficient and Scaleable Synthesis of Oligonucleotides Utilizing a Universal Linker Molecule: A Novel Approach To Enhance the Purity of Drugs. *Org Process Res Dev* **12**, 399–410 (2008).

334. Wuts, P. G. M. & Greene, T. W. Greene's Protective Groups in Organic Synthesis. *Greene's Protective Groups in Organic Synthesis* (2006) doi:10.1002/0470053488.

335. Beaucage, S. L. & Caruthers, M. H. DEOXYNUCLEOSIDE PHOSPHORAMIDITES-A NEW CLASS OF KEY INTERMEDIATES FOR DEOXYPOLYNUCLEOTIDE SYNTHESIS. *Tetrahedron Lett* **22**, 1859–1862 (1981).

336. McBride, L. J. & Caruthers, M. H. An investigation of several deoxynucleoside phosphoramidites useful for synthesizing deoxyoligonucleotides. *Tetrahedron Lett* **24**, 245–248 (1983).

337. Krotz, A. H. *et al.* Synthesis of antisense oligonucleotides. using environmentally friendly and safe deprotection procedures. *Green Chemistry* **1**, 277–282 (1999).

338. Fearon, K. L., Stults, J. T., Bergot, B. J., Christensen, L. M. & Raible, A. M. Investigation of the 'n-1' impurity in phosphorothioate oligodeoxynucleotides synthesized by the solid-phase beta-cyanoethyl phosphoramidite method using stepwise sulfurization. *Nucleic Acids Res* **23**, 2754 (1995).

339. Temsamani, J., Kubert, M. & Agrawal, S. Sequence identity of the n-1 product of a synthetic oligonucleotide. *Nucleic Acids Res* **23**, 1841 (1995).

340. Hakala, H., Oivanen, M., Saloniemi, E., Gouzaev, A. & Lönnberg, H. Acid-catalysed depurination of DI-, TRI- and polydeoxyribonucleotides: Effect of molecular environment on the cleavage of adenine residue. *J Phys Org Chem* **5**, 824–828 (1992).

341. Suzuki, T., Ohsumi, S. & Makino, K. Mechanistic studies on depurination and apurinic site chain breakage in oligodeoxyribonucleotides. *Nucleic Acids Res* **22**, 4997–5003 (1994).

342. Beaucage, S. L. & Caruthers, M. H. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett* **22**, 1859–1862 (1981).

343. Wu, X. & Pitsch, S. Synthesis and pairing properties of oligoribonucleotide analogues containing a metal-binding site attached to beta-D-allofuranosyl cytosine. *Nucleic Acids Res* **26**, 4315–4323 (1998).

344. Welz, R. & Müller, S. 5-(Benzylmercapto)-1H-tetrazole as activator for 2'-O-TBDMS phosphoramidite building blocks in RNA synthesis. *Tetrahedron Lett* **43**, 795–797 (2002).

345. Wei, X. Coupling activators for the oligonucleotide synthesis via phosphoramidite approach. *Tetrahedron* **69**, 3615–3637 (2013).

346. Threlfall, R. N., Torres, A. G., Krivenko, A., Gait, M. J. & Caruthers, M. H. Synthesis and biological activity of phosphonoacetate- and thiophosphonoacetate-modified 2'-O-methyl oligoribonucleotides. *Org Biomol Chem* **10**, 746–754 (2012).

347. Krotz, A. H. *et al.* Phosphorothioate oligonucleotides with low phosphate diester content: Greater than 99.9% sulfurization efficiency with 'aged' solutions of phenylacetyl disulfide (PADS). *Org Process Res Dev* **8**, 852–858 (2004).

348. Ravikumar, V. T., Andrade, M., Carty, R. L., Dan, A. & Barone, S. Development of siRNA for therapeutics: Efficient synthesis of phosphorothioate RNA utilizing phenylacetyl disulfide (PADS). *Bioorg Med Chem Lett* **16**, 2513–2517 (2006).

349. Capaldi, D. *et al.* Quality Aspects of Oligonucleotide Drug Development: Specifications for Active Pharmaceutical Ingredients. *Drug Inf J* **46**, 611–626 (2012).

350. Sanghvi, Y. S. *et al.* Chemical Synthesis and Purification of Phosphorothioate Antisense Oligonucleotides. in 3–23 (Springer, Boston, MA, 1999). doi:10.1007/978-1-4615-5067-9_1.

351. Anderson, N. G., Anderson, N. L., Taylor, J. & Goodman, J. Large-scale oligonucleotide synthesizers. I. Basic principles and system design. *Appl Biochem Biotechnol* **54**, 19–42 (1995).

352. Boal, J. H. *et al.* Cleavage of oligodeoxyribonucleotides from controlled-pore glass supports and their rapid deprotection by gaseous amines. *Nucleic Acids Res* **24**, 3115–3117 (1996).

353. Reddy, M. P., Farooqui, F. & Hanna, N. B. Methylamine deprotection provides increased yield of oligoribonucleotides. *Tetrahedron Lett* **36**, 8929–8932 (1995).

354. Pourshahian, S. THERAPEUTIC OLIGONUCLEOTIDES, IMPURITIES, DEGRADANTS, AND THEIR CHARACTERIZATION BY MASS SPECTROMETRY. *Mass Spectrom Rev* **40**, 75–109 (2021).

355. Warren, W. J. & Vella, G. Analysis and purification of synthetic oligonucleotides by high-performance liquid chromatography. *Methods Mol Biol* **26**, 233–264 (1994).

356. Azarani, A. & Hecker, K. H. RNA analysis by ion-pair reversed-phase high performance liquid chromatography. *Nucleic Acids Res* **29**, e7 (2001).

357. Biba, M., Foley, J. P. & Welch, C. J. Liquid chromatographic separation of oligonucleotides. in *Liquid Chromatography: Applications: Second Edition* vol. 2 159–182 (Elsevier Inc., 2017).

358. Goyon, A., Yehl, P. & Zhang, K. Characterization of therapeutic oligonucleotides by liquid chromatography. *J Pharm Biomed Anal* **182**, 113105 (2020).

359. Catani, M. *et al.* Oligonucleotides: Current Trends and Innovative Applications in the Synthesis, Characterization, and Purification. *Biotechnol J* **15**, (2020).

360. Huber, C. G., Oefner, P. J. & Bonn, G. K. High-resolution liquid chromatography of oligonucleotides on nonporous alkylated styrene-divinylbenzene copolymers. *Anal Biochem* **212**, 351–358 (1993).

361. Fornstedt, T. & Enmark, M. Separation of therapeutic oligonucleotides using ion-pair reversed-phase chromatography based on fundamental separation science. *Journal of Chromatography Open* **3**, 100079 (2023).

362. Sun, J. *et al.* Anion exchange chromatography of oligonucleotides under denaturing conditions. *Nucleosides Nucleotides Nucleic Acids* **39**, 818–828 (2020).

363. Ausserer, W. A. & Biros, M. L. High-resolution analysis and purification of synthetic oligonucleotides with strong anion-exchange HPLC. *Biotechniques* **19**, 136–139 (1995).

364. Aumann, L. & Morbidelli, M. A continuous multicolumn countercurrent solvent gradient purification (MCSGP) process. *Biotechnol Bioeng* **98**, 1043–1055 (2007).

365. Vogg, S., Ulmer, N., Souquet, J., Broly, H. & Morbidelli, M. Experimental Evaluation of the Impact of Intrinsic Process Parameters on the Performance of a Continuous Chromatographic Polishing Unit (MCSGP). *Biotechnol J* **14**, (2019).

366. Luca, C. De *et al.* From batch to continuous chromatographic purification of a therapeutic peptide through multicolumn countercurrent solvent gradient purification. *J Chromatogr A* **1625**, (2020).

367. Catani, M. *et al.* Oligonucleotides: Current Trends and Innovative Applications in the Synthesis, Characterization, and Purification. *Biotechnol J* **15**, 1900226 (2020).

368. Krotz, A. H., McElroy, B., Scozzari, A. N., Cole, D. L. & Ravikumar, V. T. Controlled detritylation of antisense oligonucleotides. *Org Process Res Dev* **7**, 47–52 (2003).

369. Emery, A. E. H. The muscular dystrophies. *Lancet* **359**, 687–695 (2002).

370. Mackenzie, S. J., Nicolau, S., Connolly, A. M. & Mendell, J. R. Therapeutic Approaches for Duchenne Muscular Dystrophy: Old and New. *Semin Pediatr Neurol* **37**, (2021).

371. Chamberlain, J. R. & Chamberlain, J. S. Progress toward Gene Therapy for Duchenne Muscular Dystrophy. *Mol Ther* **25**, 1125–1131 (2017).

372. Sironi, M. *et al.* A region in the dystrophin gene major hot spot harbors a cluster of deletion breakpoints and generates double-strand breaks in yeast. *The FASEB Journal* **20**, 1910–1912 (2006).

373. Kole, R., Krainer, A. R. & Altman, S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nature Reviews Drug Discovery 2012 11:2* **11**, 125–140 (2012).

374. Duan, D., Goemans, N., Takeda, S., Mercuri, E. & Aartsma-Rus, A. Duchenne muscular dystrophy. *Nature Reviews Disease Primers 2021 7:1* **7**, 1–19 (2021).

375. Waldrop, M. A. & Flanigan, K. M. Update in Duchenne and Becker muscular dystrophy. *Curr Opin Neurol* **32**, 722–727 (2019).

376. Sun, C., Serra, C., Lee, G. & Wagner, K. R. Stem cell-based therapies for Duchenne muscular dystrophy. *Exp Neurol* **323**, 113086 (2020).

377. Lim, K. R. Q., Maruyama, R. & Yokota, T. Eteplirsen in the treatment of Duchenne muscular dystrophy. *Drug Des Devel Ther* **11**, 533–545 (2017).

378. Roshmi, R. R. & Yokota, T. Viltolarsen for the treatment of Duchenne muscular dystrophy. *Drugs Today (Barc)* **55**, 627–639 (2019).

379. Anwar, S. & Yokota, T. Golodirsen for Duchenne muscular dystrophy. *Drugs Today (Barc)* **56**, 491–504 (2020).

380. Shirley, M. Casimersen: First Approval. *Drugs* **81**, 875–879 (2021).

381. Aboul-Fadl, T. Antisense oligonucleotides: the state of the art. *Curr Med Chem* **12**, 2193–2214 (2005).

382. Van Deutekom, J. *et al.* Next Generation Exon 51 Skipping Antisense Oligonucleotides for Duchenne Muscular Dystrophy. *Nucleic Acid Ther* **33**, 193–208 (2023).

383. Tyagi, S. & Kramer, F. R. Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* **14**, 303–308 (1996).

384. Whitcombe, D., Theaker, J., Guy, S. P., Brown, T. & Little, S. Detection of PCR products using self-probing amplicons and fluorescence. *Nature Biotechnology 1999 17:8* **17**, 804–807 (1999).

385. French, D. J., McDowell, D. G., Debenham, P., Gale, N. & Brown, T. HyBeacon probes for rapid DNA sequence detection and allele discrimination. *Methods Mol Biol* **429**, 171–185 (2008).

386. French, D. J., Archard, C. L., Brown, T. & McDowell, D. G. HyBeacon™ probes: A new tool for DNA sequence detection and allele discrimination. *Mol Cell Probes* **15**, 363–374 (2001).

387. Asseline, U. Development and Applications of Fluorescent Oligonucleotides. *Curr Org Chem* **10**, 491–518 (2006).

388. Rist, M. & Marino, J. Fluorescent Nucleotide Base Analogs as Probes of Nucleic Acid Structure, Dynamics and Interactions. *Curr Org Chem* **6**, 775–793 (2005).

389. Hawkins, M. E. Fluorescent pteridine nucleoside analogs: a window on DNA interactions. *Cell Biochem Biophys* **34**, 257–281 (2001).

390. Wojczewski, C., Stolze, K. & Engels, J. W. Fluorescent Oligonucleotides-Versatile Tools as Probes and Primers for DNA and RNA Analysis. 1667–1678 (1999).

391. Ward, D. C., Reich, E. & Stryer, L. Fluorescence Studies of Nucleotides and Polynucleotides. *Journal of Biological Chemistry* **244**, 1228–1237 (1969).

392. Driscoll, S. L., Hawkins, M. E., Balis, F. M., Pfleiderer, W. & Laws, W. R. Fluorescence properties of a new guanosine analog incorporated into small oligonucleotides. *Biophys J* **73**, 3277–3286 (1997).

393. Hawkins, M. E., Pfleiderer, W., Balis, F. M., Porter, D. & Knutson, J. R. Fluorescence properties of pteridine nucleoside analogs as monomers and incorporated into oligonucleotides. *Anal Biochem* **244**, 86–95 (1997).

394. Godde, F., Toulme, J. J. & Moreau, S. 4-amino-1H-benzo[g]quinazoline-2-one: a fluorescent analog of cytosine to probe protonation sites in triplex forming oligonucleotides. *Nucleic Acids Res* **28**, 2977–2985 (2000).

395. Crockett, A. O. & Wittwer, C. T. Fluorescein-labeled oligonucleotides for real-time pcr: using the inherent quenching of deoxyguanosine nucleotides. *Anal Biochem* **290**, 89–97 (2001).

396. Olmsted, J. & Kearns, D. R. Mechanism of Ethidium Bromide Fluorescence Enhancement on Binding to Nucleic Acids. *Biochemistry* **16**, 3647–3654 (1977).

397. Pachmann, U. & Rigler, R. Quantum yield of acridines interacting with DNA of defined sequence. A basis for the explanation of acridine bands in chromosomes. *Exp Cell Res* **72**, 602–608 (1972).

398. Fürstenberg, A. *et al.* Ultrafast Excited-State Dynamics of DNA Fluorescent Intercalators: New Insight into the Fluorescence Enhancement Mechanism. *J Am Chem Soc* **128**, 7661–7669 (2006).

399. Lee, L. G., Chen, C. -H & Chiu, L. A. Thiazole orange: a new dye for reticulocyte analysis. *Cytometry* **7**, 508–517 (1986).

400. Daniel, D. C., Thompson, M. & Woodbury, N. W. Fluorescence intensity fluctuations of individual labeled DNA fragments and a DNA binding protein in solution at the single molecule level: A comparison of photobleaching, diffusion, and binding dynamics. *Journal of Physical Chemistry B* **104**, 1382–1390 (2000).

401. Bidar, N. *et al.* Molecular beacon strategies for sensing purpose. *TrAC Trends in Analytical Chemistry* **134**, 116143 (2021).

402. Svanvik, N., Westman, G., Wang, D. & Kubista, M. Light-up probes: thiazole orange-conjugated peptide nucleic acid for detection of target nucleic acid in homogeneous solution. *Anal Biochem* **281**, 26–35 (2000).

403. O, S., F, B. & D, H. A Convergent Strategy for the Modification of Peptide Nucleic Acids: Novel Mismatch-Specific PNA-Hybridization Probes. *Angew Chem Int Ed Engl* **38**, 2203–2206 (1999).

404. Svanvik, N., Ståhlberg, A., Sehlstedt, U., Sjöback, R. & Kubista, M. Detection of PCR Products in Real Time Using Light-up Probes. *Anal Biochem* **287**, 179–182 (2000).

405. Nygren, J., Svanvik, N. & Kubista, M. The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* **46**, 39–51 (1998).

406. Köhler, O. & Seitz, O. Thiazole orange as fluorescent universal base in peptide nucleic acids. *Chemical Communications* **3**, 2938–2939 (2003).

407. Bethge, L., Jarikote, D. V. & Seitz, O. New cyanine dyes as base surrogates in PNA: forced intercalation probes (FIT-probes) for homogeneous SNP detection. *Bioorg Med Chem* **16**, 114–125 (2008).

408. Jarikote, D. V., Krebs, N., Tannert, S., Röder, B. & Seitz, O. Exploring Base-Pair-Specific Optical Properties of the DNA Stain Thiazole Orange. *Chemistry – A European Journal* **13**, 300–310 (2007).

409. Wang, D. O. & Okamoto, A. ECHO probes: Fluorescence emission control for nucleic acid imaging. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **13**, 112–123 (2012).

410. Casas-Solvas, J. M., Howgego, J. D. & Davis, A. P. Synthesis of substituted pyrenes by indirect methods. *Org Biomol Chem* **12**, 212–232 (2013).

411. Feng, X., Hu, J., Redshaw, C. & Yamato, T. Functionalization of Pyrene To Prepare Luminescent Materials—Typical Examples of Synthetic Methodology. *Chemistry – A European Journal* **22**, 11898–11916 (2016).

412. Karuppannan, S. & Chambron, J. C. Supramolecular Chemical Sensors Based on Pyrene Monomer–Excimer Dual Luminescence. *Chem Asian J* **6**, 964–984 (2011).

413. Somerharju, P. Pyrene-labeled lipids as tools in membrane biophysics and cell biology. *Chem Phys Lipids* **116**, 57–74 (2002).

414. Winnik, F. Photophysics of Preassociated Pyrenes in Aqueous Polymer Solutions and in Other Organized Media. *Chem Rev* **93**, 587–614 (1993).

415. Bains, G., Patel, A. B. & Narayanaswami, V. Pyrene: A Probe to Study Protein Conformation and Conformational Changes. *Molecules 2011, Vol. 16, Pages 7909-7935* **16**, 7909–7935 (2011).

416. Figueira-Duarte, T. M. & Müllen, K. Pyrene-Based Materials for Organic Electronics. *Chem Rev* **111**, 7260–7314 (2011).

417. Guckian, K. M. *et al.* Factors contributing to aromatic stacking in water: Evaluation in the context of DNA. *J Am Chem Soc* **122**, 2213–2222 (2000).

418. Dougherty, G. & Koller, T. Determination of the number of superhelical turns by the hyperchromicity of partially denatured covalently-closed DNA molecules. *Nucleic Acids Res* **10**, 525–538 (1982).

419. Lu, X.-J. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* **31**, 5108–5121 (2003).

420. Østergaard, M. E. & Hrdlicka, P. J. Pyrene-functionalized oligonucleotides and locked nucleic acids (LNAs): Tools for fundamental research, diagnostics, and nanotechnology. *Chem Soc Rev* **40**, 5771–5788 (2011).

421. Manoharan, M., Tivel, K. L., Zhao, M., Nafisi, K. & Netzel, T. L. Base-sequence dependence of emission lifetimes for DNA oligomers and duplexes covalently labeled with pyrene: Relative electron-transfer quenching efficiencies of A, G, C, and T nucleosides toward pyrene. *Journal of physical chemistry* **99**, 17461–17472 (1995).

422. Young, J. S., Jin, H. R. & Byeang, H. K. Quencher-free, end-stacking oligonucleotides for probing single-base mismatches in DNA. *Org Lett* **7**, 4931–4933 (2005).

423. Kottysch, T., Ahlborn, C., Brotzel, F. & Richert, C. Stabilizing or Destabilizing Oligodeoxynucleotide Duplexes Containing Single 2'-Deoxyuridine Residues with 5-Alkynyl Substituents. *Chemistry – A European Journal* **10**, 4017–4028 (2004).

424. Skorobogatyi, M. V. *et al.* Fluorescent 5-Alkynyl-2'-Deoxyuridines: High Emission Efficiency of a Conjugated Perylene Nucleoside in a DNA Duplex. *ChemBioChem* **7**, 810–816 (2006).

425. Koshkin, A. A. *et al.* LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **54**, 3607–3630 (1998).

426. Obika, S., Andoh, J. ichi, Sugimoto, T., Miyashita, K. & Imanishi, T. Synthesis of a conformationally locked AZT analogue, 3′-azido-3′-deoxy-2′-O,4′-C-methylene-5-methyluridine. *Tetrahedron Lett* **40**, 6465–6468 (1999).

427. Petersen, M. *et al.* The conformations of locked nucleic acids (LNA). *Journal of Molecular Recognition* **13**, 44–53 (2000).

428. Nielsen, J. T., Arar, K. & Petersen, M. NMR solution structures of LNA (locked nucleic acid) modified quadruplexes. *Nucleic Acids Res* **34**, 2006–2014 (2006).

429. Konorov, S. O. *et al.* Ultraviolet resonance Raman spectroscopy of locked single-stranded oligo(dA) reveals conformational implications of the locked ribose in LNA. *Journal of Raman Spectroscopy* **40**, 1162–1171 (2009).

430. Petersen, M., Bondensgaard, K., Wengel, J. & Peter Jacobsen, J. Locked nucleic acid (LNA) recognition of RNA: NMR solution structures of LNA:RNA hybrids. *J Am Chem Soc* **124**, 5974–5982 (2002).

431. Nielsen, K. E. & Spielmann, H. P. The structure of a mixed LNA/DNA:RNA duplex is driven by conformational coupling between LNA and deoxyribose residues as determined from 13C relaxation measurements. *J Am Chem Soc* **127**, 15273–15282 (2005).

432. Petersen, M. & Wengel, J. LNA: a versatile tool for therapeutics and genomics. *Trends Biotechnol* **21**, 74–81 (2003).

433. Singh, S. K. & Wengel, J. Universality of LNA-mediated high-affinity nucleic acid recognition. *Chemical Communications* 1247–1248 (1998) doi:10.1039/a801571f.

434. Godfrey, C. *et al.* Delivery is key: lessons learnt from developing splice-switching antisense therapies. *EMBO Mol Med* **9**, 545–557 (2017).

435. Braasch, D. A., Liu, Y. & Corey, D. R. Antisense inhibition of gene expression in cells by oligonucleotides incorporating locked nucleic acids: Effect of mRNA target sequence and chimera design. *Nucleic Acids Res* **30**, 5160–5167 (2002).

436. Kamali, M. J. *et al.* Locked nucleic acid (LNA): A modern approach to cancer diagnosis and treatment. *Exp Cell Res* **423**, 113442 (2023).

437. Mouritzen, P. *et al.* Single nucleotide polymorphism genotyping using locked nucleic acid (LNA™). *Expert Rev Mol Diagn* **3**, 27–38 (2003).

438. Kumar, P. *et al.* Synthesis and biophysical properties of C5-functionalized LNA (locked nucleic acid). *Journal of Organic Chemistry* **79**, 5047–5061 (2014).

439. Qiu, J., Wilson, A., El-Sagheer, A. H. & Brown, T. Combination probes with intercalating anchors and proximal fluorophores for DNA and RNA detection. *Nucleic Acids Res* **44**, e138–e138 (2016).

440. French, D. J., Archard, C. L., Andersen, M. T. & McDowell, D. G. Ultra-rapid DNA analysis using HyBeacon™ probes and direct PCR amplification from saliva. *Mol Cell Probes* **16**, 319–326 (2002).

441. French, D. J., Archard, C. L., Brown, T. & McDowell, D. G. HyBeacon™ probes: A new tool for DNA sequence detection and allele discrimination. *Mol Cell Probes* **15**, 363–374 (2001).

442. Kumar, P., Østergaard, M. E. & Hrdlicka, P. J. Preparation of C5-Functionalized Locked Nucleic Acids (LNAs). *Curr Protoc Nucleic Acid Chem* **44**, (2011).

443. Wang, Z. Vorbrüggen Glycosylation. in *Comprehensive Organic Name Reactions and Reagents* 2915–2919 (Wiley, 2010). doi:10.1002/9780470638859.conrr652.

444. Gierlich, J., Burley, G. A., Gramlich, P. M. E., Hammond, D. M. & Carell, T. Click chemistry as a reliable method for the high-density postsynthetic functionalization of alkyne-modified DNA. *Org Lett* **8**, 3639–3642 (2006).

445. Klimkowski, P., De Ornellas, S., Singleton, D., El-Sagheer, A. H. & Brown, T. Design of thiazole orange oligonucleotide probes for detection of DNA and RNA by fluorescence and duplex melting. *Org Biomol Chem* **17**, 5943–5950 (2019).

446. Köhler, O., Jarikote, D. V. & Seitz, O. Forced Intercalation Probes (FIT Probes): Thiazole Orange as a Fluorescent Base in Peptide Nucleic Acids for Homogeneous Single-Nucleotide-Polymorphism Detection. *ChemBioChem* **6**, 69–77 (2005).

447. Koshkin, A. A., Fensholdt, J., Pfundheller, H. M. & Lomholt, C. A simplified and efficient route to 2′-O, 4′-C-methylene-linked bicyclic ribonucleosides (locked nucleic acid). *Journal of Organic Chemistry* **66**, 8504–8512 (2001).

448. Kumar, P., Østergaard, M. E. & Hrdlicka, P. J. Preparation of C5-Functionalized Locked Nucleic Acids (LNAs). *Curr Protoc Nucleic Acid Chem* **44**, (2011).

449. Gramlich, P. M. E., Warncke, S., Gierlich, J. & Carell, T. Click-click-click: Single to triple modification of DNA. *Angewandte Chemie - International Edition* **47**, 3442–3444 (2008).