University “Federico II” of Naples

Department of Organic Chemistry and Biochemistry

PhD School in Chemical Sciences

XXIV course

PhD THESIS in Organic Chemistry

SYNTHESIS, CHARACTERIZATION AND APPLICATIONS OF NOVEL AMPHIPHILIC DERIVATIVES OF NUCLEOSIDES

Student: Dr. Luca Simeone

Tutor: Prof. Daniela Montasarchio

Supervisor: Dr. Alfonso Iadonisi

Index

Chapter 3

Synthesis and Characterization of Novel Amphiphilic Nucleolipid Complexes of Ruthenium (III) as Potential Anticancer Agents

Chapter 4

Design and Synthesis of a Novel Cationic Aminoacylnucleolipid as a Model Compound for Highly Functionalized Nucleolipids

Chapter 5

Design, Synthesis and Characterization of Novel Guanosine-based Amphiphiles

List of Publications

This thesis is based on the following papers:

1. **Design, Synthesis and Characterization of novel Guanosine-based Amphiphiles**

   L. Simeone, D. Milano, L. De Napoli, C. Irace, A. Di Pascale, M. Boccalon, P. Tecilla, D. Montesarchio

   Chemistry – A European Journal, **2011**, *17*, 13854-13865.

2. **Boc N-3 Protection of Thymidine for an easy Synthetic Access to Sugar-alkylated Nucleoside Analogs**

   L. Simeone, L. De Napoli, D. Montesarchio

   Chemistry & Biodiversity, *accepted* doi: 10.1002/cbdv.201100103.

3. **Nucleolipid Nanovectors as Molecular Carriers for Potential Applications in Drug Delivery**

   L. Simeone, G. Mangiapia, C. Irace, A. Di Pascale, A. Colonna, O. Ortona, L. De Napoli, D. Montesarchio, L. Paduano

   Molecular BioSystems **2011**, *7*, 3075–3086.

4. **A Cholesterol-Based Nucleolipid-Ruthenium Complex Stabilized by Lipid Aggregates for Antineoplastic Therapy**

   L. Simeone, G. Mangiapia, G. Vitiello, C. Irace, A. Di Pascale, A. Colonna, O. Ortona, L. De Napoli, D. Montesarchio, L. Paduano

   Bioconjugate Chemistry, submitted.

5. **Ruthenium-Based Amphyphylic Nanocarriers As A Novel Approach For Anticancer Therapy**

   L. Simeone *et al.*, manuscript in preparation.

Other papers not included in this thesis:

1. **On the Compatibility of Azides in Phosphoramidite-Based Couplings: Synthesis Of a Novel, Convertible Azido-Functionalized CyPLOS Analogue**

   C. Coppola, L. Simeone, L. De Napoli, D. Montesarchio

   European Journal of Organic Chemistry, **2011***, 6,* 1155-1165.

2. **Synthesis and NMR characterization of a novel crown ether ring-fused uridine analogue**

   C. Coppola, L. Simeone, R. Trotta, L. De Napoli, A. Randazzo and D. Montesarchio

   Tetrahedron, **2010**, *66*, 2769-6774.

Chapter 1.

Introduction

1. Amphiphilic molecules: structure, classification and biological role

Amphiphilic (from the Greek αμφις, amphis: both and φιλία, philia: love, friendship) is a term describing a chemical compound possessing both **hydrophilic** (water-loving, polar) and **hydrophobic** (fat-loving) properties (Figure 1). Such a compound is called amphiphilic or amphipathic. This forms the basis for a number of areas of research in chemistry and biochemistry, notably that of lipid polymorphism. The lipophilic groups of the amphiphiles are typically long hydrocarbon residues such as a long chain of the form $CH_3(CH_2)_n$, with n > 4. [1-3]

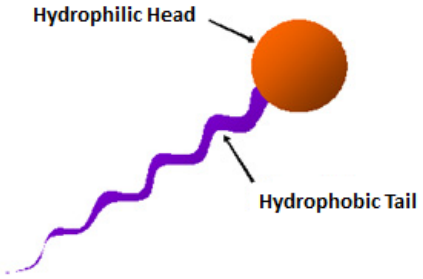


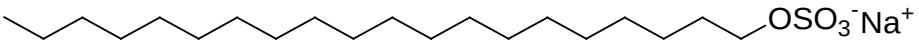
Figure 1. Schematic representation of an amphiphile molecule.

The hydrophilic groups can be classified in various categories:

1. Charged groups
    - Anionic (R = lipophilic part):
        - Carboxylates: $RCOO^-$
        - Sulfate groups: $ROSO_3^-$
        - Phosphate groups: $ROPO_3^{2-}$, present in phospholipids
    - Cationic:
        - Amines: RNH_3^+
2. Neutral (uncharged groups): examples are alcohols with large R groups, such as diacyl glycerol (DAG), and oligoethyleneglycols with long alkyl chains.

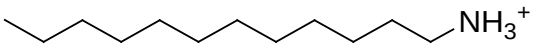
In Figure 2 same examples of amphiphilic molecules are reported:

Anionic



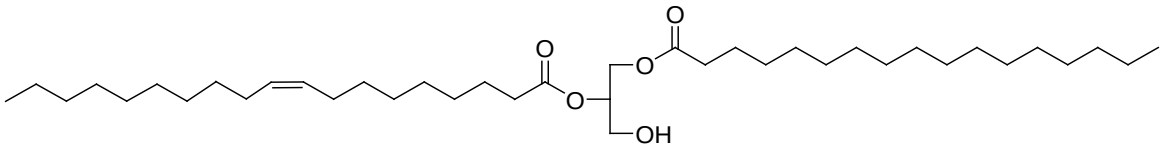
Sodium dodecyl sulfate (SDS)

Cationic



Ammonium dodecane

Neutral



1-Palmitoyl-2-oleoyl-glycerol

Figure 2. Some examples of amphiphilic molecules.

Amphiphiles orientate their domain with the highest affinity for the solvent toward the solvent molecules, while the other part of the molecule tends to fold so to avoid contact with the solvent. In this way amphiphilic molecules align according to well-defined patterns and give rise to a variety of different self-aggregates. The shape (=morphology) of the aggregates is determined by the ratio between the size of the hydrophilic and the hydrophobic part as well as the nature of the solvent. Scanning electron microscopy and other techniques have revealed various morphologies: micellar rods (**A**), micelles (**B**), bilayers (**C**), vesicles (**D**) and inverted aggregates (**E**), as shown in Figure 3.

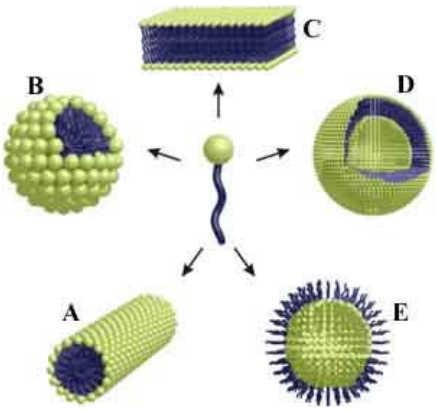


Figure 3. Possible self-aggregation modes of amphiphilic molecules.

For example, in water solution the polar heads are exposed to the solvent and the hydrophobic tails are not in contact with the solvent; in this way the amphiphiles are able to solubilize in water a large number of hydrophobic guest molecules (oils), otherwise insoluble in polar solvents, by encapsulation of the guest within the micelle core (Figure 4). This property is used in detergents. [3]

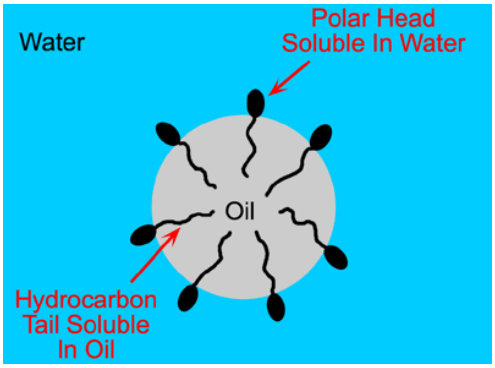
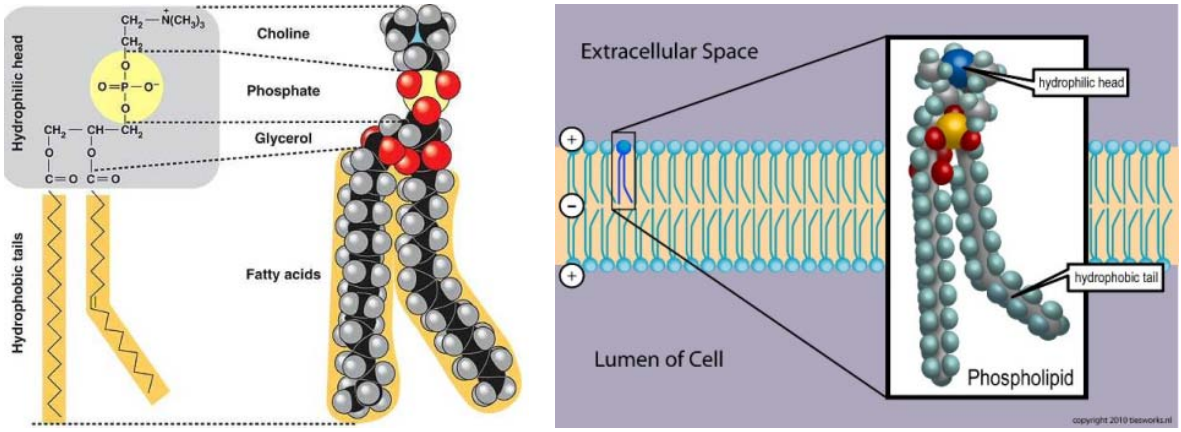


Figure 4. Solubilization of hydrophobic molecules in water.

Amphiphilic molecules are very common in nature. The phospholipids are an example of naturally occurring amphiphiles. They are a major component of all cell membranes as they can form lipid bilayers which enclose the cytoplasm and other contents of a cell. Most phospholipids contain a glycerol backbone, a phosphate group, and a structurally simple organic molecule such as choline. The hydrophilic head contains the negatively charged phosphate group, and may contain other polar groups. The hydrophobic tail usually consists of long fatty acid hydrocarbon chains. When placed in water, phospholipids form a variety of structures depending on the specific properties of

the phospholipids. These specific properties allow phospholipids to play an important role in the phospholipid bilayers. In cell membranes, the phospholipids often occur in combination with other molecules (*e.g.*, proteins, glycolipids, cholesterol). Lipid bilayers are obtained when hydrophobic tails line up one against the other, forming a membrane with hydrophilic heads on both sides facing the water (Figure 5). [3-4]



A **B**

Figure 5. (A) Phospholipids structure; **(B)** Phospholipid bilayer (cell membrane).

2. Nucleolipids: a class of amphiphilic molecules

In recent years, nucleic acids features have been used in the development of numerous artificial structures, including oligonucleotide [5-7] and nucleolipid analogs. [8-10] In these biomimetic approaches, two fundamental objectives have been pursued in combining nucleic acid and lipid families:

i) the development of new therapeutic strategies;

ii) the construction of new supramolecular assemblies. [11]

2.1 Natural nucleolipids

Nucleolipids are hybrid molecules composed of a lipid covalently linked to a nucleobase, a nucleoside or an oligonucleotide. Hybrid lipid–nucleoside structures occur in eukaryotic and prokaryotic cells. Tunicamicyns for instance possess antimicrobial, antifungal, antiviral and antitumor activities. [12-13] Cytidinediphosphate diacylglycerol coenzyme (CDP-DAG, Figure 6), a diphosphorylated nucleolipid (or liponucleotide), is present in mammalian cells and plays a central role in the metabolism of phosphoinosides and cardiolipins. [14] CDP-DAG and its analogue 2'-deoxy-CDP-DAG are also present in prokaryotic cells as key intermediates in the synthesis of all glycerophospholipids. In addition to their biochemical function, nucleolipids are unique in terms of molecular structure (containing different structural motifs, *i.e.* lipid, phosphate, sugar, heterocyclic moieties) and biophysical properties. Consequently, nucleolipid analogs have been investigated as anticancer drugs and pro-drugs. [11]

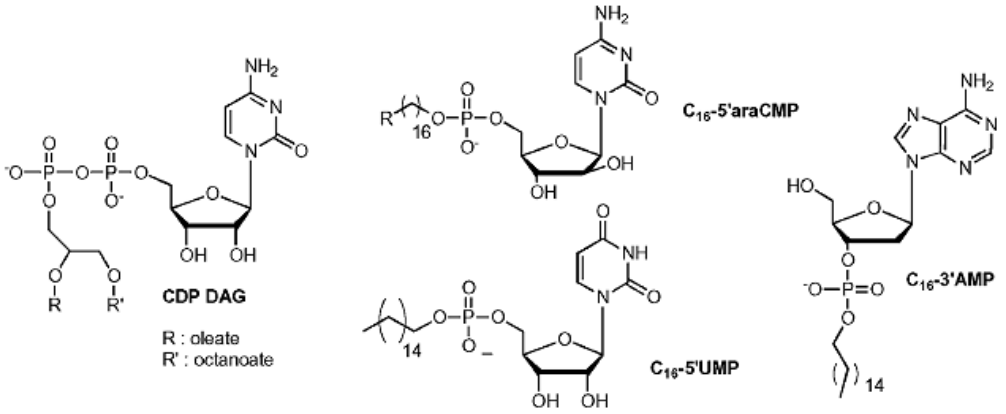


Figure 6. Chemical structures of CDP-DAG and synthetic amphipatic nucleotides.

2.2 Artificial nucleolipids

Nucleotides are natural substrates of transcriptase or reverse transcriptase. Accordingly, many analogs have been designed to block the enzymatic elongation of RNA chains and are used as anti-tumor or antiviral drugs. [11] Given their inherent toxicity, many prodrugs of these chemotherapeutic agents have been evaluated over the years. [15-20] Since the 80's, several research groups have devoted their efforts to combine the pharmacological activity of nucleosides with the aggregation properties of vesicle-forming lipids. It is well known that liposome-forming species can produce effective drug carrier systems, capable of incorporating large amounts of lipophilic drugs in aqueous solutions and efficiently transport them within cells by endocytosis (Figure 7). If the carrier system is covalently linked to the drug, a very efficient prodrug may be obtained, which, by self-structuring into liposomes, is expected to undergo only minimal leakages of the lipidic prodrug from the delivery system, thus ensuring an efficient transport and, in parallel, protecting the drug from enzymatic degradation. This simple concept is the basis for the increasing interest towards nucleolipids.

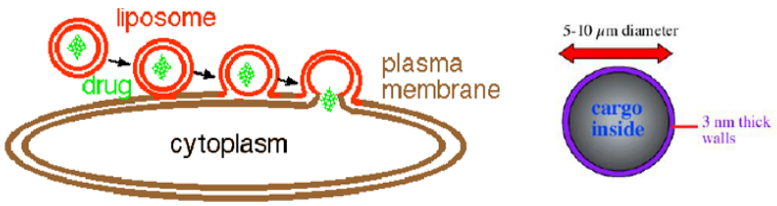


Figure 7. Transport of a drug into a cell by liposome formulation.

2.3 Supramolecular assemblies of nucleolipids

Nucleolipids, in contrast to classical detergents, possess a highly informative polar head (A, T, C, G, U or analogs) with additional H-bonding and π-stacking capabilities that can specifically base-pair with other nucleobases. The interplay between these specific base pairing interactions, [5-7] and the unique aggregation properties of lipids is very attractive for the design of supramolecular assemblies with new and interesting properties. These concepts of molecular recognition with nucleolipids were initially applied to a supramolecular film at the air–water interface. [8,21,22] More recently, base-pairing between complementary nucleolipids at mesoscopic interfaces such as those found in vesicular and micellar systems have been investigated. [8,9,23–28] Nucleolipid–nucleolipid interactions have been studied in different types of supramolecular systems including micelles, [29] vesicles, [30] and monolayers. [31-32] For these purposes a variety of synthetic strategies have been used to prepare nucleolipids. [15,33,34]

Since 2002, Barthélémy and coworkers have been designing novel amphiphilic structures derived from nucleobases. They have investigated how modifications within the molecular structure of the nucleolipid (hydrophobization of the 2' and/or 3' position in contrast to the classical 5'-modification) affect their physico-chemical and self-assembling properties. A large number of novel amphiphilic nucleolipids have been therefore synthesized, which can be classified according to their charge:

1) nonionic (compounds $\mathbf{l}_{1–3}$), [35,36]

2) zwitterionic (phosphocholine derivatives, $\mathbf{l}_{4–6}$), [37,38]

3) anionic [39] (single chain nucleotides $\mathbf{l}_7$)

4) cationic compounds ($\mathbf{l}_8$ and $\mathbf{l}_9$). [40-41]

Examples of nucleolipid structures described in the literature are shown in Figure 8:

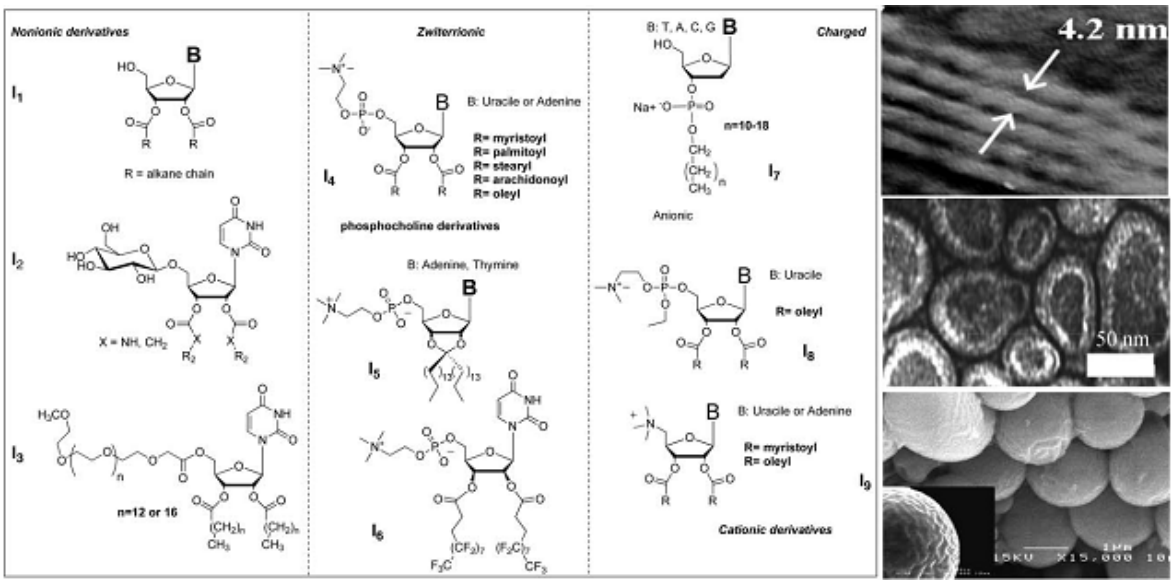


Figure 8. Left, examples of synthetic amphiphilic nucleolipids (I). Right, TEM and SEM images showing examples of different nucleolipid self-assemblies: nano-fibers, vesicles, and microspheres. [11]

The wide possibility to finely modulate the physico-chemical properties and three-dimensional structures of nucleolipids as a function of the nature of the inserted decorations renders these artificial compounds new, powerful tools for innovative developments both in biomedical applications and in material science.

3. References

1. W. Norde *Colloids and Interfaces in Life Sciences and Bionanotechnology*, CRC Press Inc, **2011.**
2. R. Nagarajan *Amphiphiles: molecular assembly and applications* (ACS symposium series, *vol. 1070*), **2011**.
3. C. Tanford *The Hydrophobic Effect: Formation of Micelles & Biological Membranes*, Wiley Publishing: New York, **1980**.
4. N.A. Campbell, B. Williamson, R.J. Heyden *Biology: Exploring Life*, Boston, Massachusetts: Pearson Prentice Hall, **2006**.
5. C. G. Overberger, Y. Inaki *J. Polym. Sci.: Polym. Chem. Ed.*, **1979**, *17*, 1739–1758.
6. C. G. Overberger, Y. Morishima *J. Polym. Sci. A1*, **1980**, *18*, 1247–1265.
7. J. Pitha, M. Akashi, M. Draminski in *Biomedical Polymers*, ed. E. P. Goldberg and A. Nakajima, Academic Press, New York, **1980**, p. 271.
8. M. Ahlers, H. Ringsdorf, H. Rosemeyer, F. Seela *Colloid Polym. Sci.,* **1990**, *268*, 132–142.
9. D. Berti, L. Franchi, P. Baglioni, P. L. Luisi *Langmuir*, **1997**, *13*, 3438–3444.
10. J. Huang, C. Li, Y. Liang *Langmuir*, **2000**, *16*, 3937–3940.
11. A. Gissot, M. Camplo, M. W. Grinstaff, P. Barthélémy *Org. Biomol. Chem.*, **2008**, *6*, 1324-1333.
12. D. C. Hiss, G. A. Gabriels, P. I. Folb *Cancer Cell Int.*, **2007**, *7*,5.
13. F. H. van Tiel, W. A. Boere, T. Harmsen, C. A. Kraaijeveld, H.Snippe *Antimicrob. Agents Chemother.*, **1985**, *27*, 802–805.
14. A. M. Heacock, B. W. Agranoff *Biochim. Biophys. Acta*, **1997**, *1348*, 166–172.
15. C. I. Hong, A. J. Kirisits, A. Nechaev, D. J. Buchheit, C. R. West *J. Med. Chem.*, **1985**, *28*, 171–177.
16. M. MacCoss, E. K. Ryu and T. Matsushita *Biochem. Biophys. Res. Commun.*, **1978**, *85*, 714–723.
17. J. G. Turcotte, S. P. Srivastava, W. A. Meresak, B. A. Rizkalla, F. Louzon, T. P. Wunz *Biochim. Biophys. Acta*, **1980**, *619*, 604–618.
18. J. G. Turcotte, S. P. Srivastava, J. M. Steim, P. Calabresi, L. M. Tibbetts, M. Y. Chu *Biochim. Biophys. Acta*, **1980**, *619*, 619–631.

19. T. Matsushita, E. K. Ryu, C. I. Hong, M. MacCoss *Cancer Res.*,**1981**, *41*, 2707–2713.
20. E. K. Ryu, R. J. Ross, T. Matsushita, M. MacCoss, C. I. Hong, C. R. West *J. Med. Chem.*, **1982**, *25*, 1322–1329.
21. K. Ariga, T. Kunitake *Acc. Chem. Res.*, **1998**, *31*, 371–378.
22. C. M. Paleos, D. Tsiourvas *Adv. Mater.*, **1997**, *9*, 695–710.
23. D. Berti, P. Baglioni, S. Bonaccio, G. Barsacchi-Bo, P. L. Luisi *J. Phys. Chem. B*, **1998**, 102, 303–308.
24. S. Bonaccio, P. Walde, P. L. Luisi, *J. Phys. Chem.*, **1994**, *98*, 6661–6663.
25. Y. Itojima, Y. Ogawa, K. Tsuno, N. Handa, H. Yanagawa *Biochemistry*, **1992**, *31*, 4757–4765.
26. T. Kawahara, K. Kurihara, T. Kunitake *Chem. Lett.*, **1992**, *21*, 1839–1842.
27. T. Shimizu, R. Iwaura, M. Masuda, T. Hanada, K. Yase *J. Am. Chem. Soc.*, **2001**, *123*, 5947–5955.
28. H. Yanagawa, Y. Ogawa, H. Furuta, K. Tsuno *J. Am. Chem. Soc.*, **1989**, *111*, 4567–4570.
29. J. S. Nowick, T. Cao, G. Noronha *J. Am. Chem. Soc.*, **1994**, *116*, 3285–3289.
30. M. Onda, K. Yoshihara, H. Koyano, K. Ariga, T. Kunitake *J. Am. Chem. Soc.*, **1996**, *118*, 8524–8530.
31. C. Li, J. Huang, Y. Liang *Langmuir*, **2000**, *16*, 7701–7707.
32. P. Berndt, K. Kurihara, T. Kunitake *Langmuir*, **1995**, *11*, 3083–3091.
33. T. Tanaka, K. Nakayama, K. Machida, M. Taniguchi *Microbiology*, **2000**, *146(Pt 2)*, 377–384.
34. G. M. T. Van Wijk, K. Y. Hostetler, H. Van den Bosch *Biochim. Biophys. Acta*, **1991**, 1084, 307–310.
35. J. Arigon, C. A. Prata, M. W. Grinstaff, P. Barthélémy *Bioconjugate Chem.*, **2005**, *16*, 864–872.
36. P. Barthélémy, C. A. Prata, S. F. Filocamo, C. E. Immoos, B. W. Maynor, S. A. Hashmi, S. J. Leeand, M. W. Grinstaff *Chem. Commun.*,**2005**, 1261–1263.
37. L. Moreau, M. W. Grinstaff, P. Barthélémy *Tetrahedron Lett.*, **2005**, *46*, 1593–1596.
38. L. Moreau, P. Barthélémy, M. El Maataoui, M. W. Grinstaff *J. Am. Chem. Soc.*, **2004**, *126*, 7533–7539.
39. N. Campins, P. Dieudonne, M. W. Grinstaff, P. Barthélémy *New J. Chem.*, **2007**, *31*, 1928–1934.

40. P. Barthélémy, S. J. Lee, M. Grinstaff *Pure Appl. Chem.*, **2005**, *77*, 2133–2148.

41. P. Barthélémy, M. Camplo *MRS Bull.*, **2005**, *30*, 647–653.

Chapter 2.

Synthesis and Characterization of Novel Nucleolipid Nanovectors as Molecular Carriers for Potential Applications in Drug Delivery

1. Introduction: the prodrugs

The development of prodrugs — chemically modified versions of the pharmacologically active agent that must undergo an *in vivo* transformation to release the active drug — is now a well established strategy to improve the physico-chemical, biomedical or pharmacokinetic properties of pharmacologically relevant compounds, and thereby increase the developability and usefulness of a potential drug (Figure 1). [1–9] For example, prodrugs provide possibilities to overcome various barriers to drug formulation and delivery, such as poor aqueous solubility, chemical instability, insufficient oral absorption, rapid pre-systemic metabolism, inadequate brain penetration, toxicity and local irritation. Prodrugs can also improve drug targeting, and the development of a prodrug of an existing drug with improved properties may represent a life-cycle management opportunity. In most cases, prodrugs are simple chemical derivatives that require only one or two chemical or enzymatic transformation steps to yield the active parent drug. In some cases, a prodrug may consist of two pharmacologically active drugs that are coupled together in a single molecule so that each drug acts as a promoiety for the other; such derivatives are called co-drugs. [10,11] Prodrugs have also been referred to as reversible or bioreversible derivatives, latentiated drugs or biolabile drug-carrier conjugates, [12–14] but the term prodrug is now standard. A bioprecursor prodrug is a prodrug that does not contain a carrier or promoiety, but results from a molecular modification of the active agent itself. This modification (for example, an oxidation or reduction) generates a new compound that can be transformed metabolically or

chemically, with the resulting compound being the active agent (it can also be referred to as an active metabolite). Finally, soft drugs, which are often confused with prodrugs, also find applications in tissue targeting. However, in contrast to prodrugs, soft drugs are active drugs that are designed to undergo a predictable and controllable deactivation or *in vivo* metabolism once their therapeutic effect is achieved. [15–17] Currently, 5–7% of the drugs approved worldwide can be classified as prodrugs, and approximately 15% of all new drugs approved in 2001 and 2002 were prodrugs. [2,6]

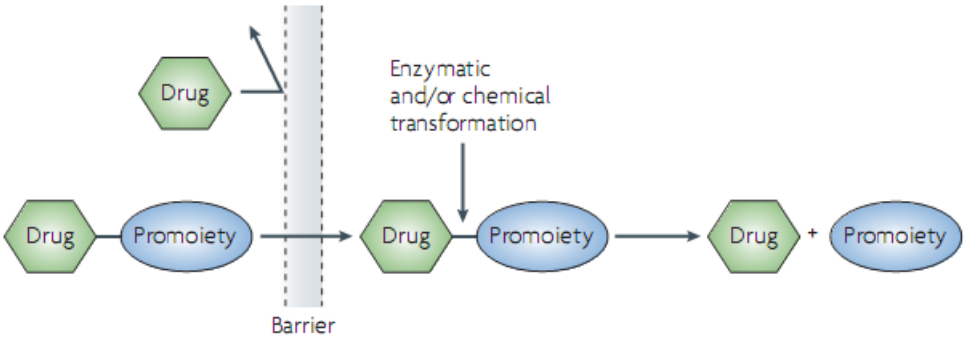


Figure 1. A simplified representative illustration of the prodrug concept. The drug–promoiety is the prodrug that is typically pharmacologically inactive. In broad terms, the barrier can be thought of as any liability or limitation of a parent drug that prevents optimal (bio)pharmaceutical or pharmacokinetic performance, and which has to be overcome for the development of a marketable drug. The drug and promoiety are covalently linked *via* bioreversible groups that are chemically or enzymatically labile, such as those shown here. The 'ideal' prodrug yields the parent drug with high recovery ratios, with the promoiety being non-toxic. [1]

2. Aims of the work

In recent years a number of amphiphilic nucleoside-based nanoaggregates have been proposed as transfection agents. In this frame, Berti and coworkers reported a promising strategy for the base-specific recognition between micelles composed of dioctanoylphosphatidyluridine, a negatively charged amphiphile, and short oligoadenylic acids. [18-21] Extensive and detailed highlights on this field have been recently presented by Gissot *et al*. [22] In this review several examples of nucleoside- or nucleotide- and oligonucleotide-based amphiphiles along with their peculiar properties are presented, which in principle may result into efficient pro-drugs. Insertion of functional reporter groups into amphiphilic nucleolipids in principle further expands the potential of this therapeutic strategy, allowing specific recognition/interaction with drugs. The complex between the functional nucleolipid and the drug can be *ad hoc* designed, so that rapid self-assembly is realized to give stable super-aggregates such as micelles, vesicles, liposomes, etc.

Nanostructuring in pseudo-physiological conditions leads to efficient cell uptake through the phospholipid bilayers and also confers protection to the drug from extracellular enzymatic degradation. On the other hand, full reversibility - within the cells- of the chemical bonds connecting the nucleoside decorations may be a powerful strategy to prevent toxicity, thus opening the way to a completely new approach for drug delivery. In this chapter we present the design, synthesis and physico-chemical characterization of novel nucleolipids along with the investigation of their cytotoxic profile by *in vitro* experiments.

The new described nucleolipids are characterized by:

- thymidine or uridine selected as the central scaffold;
- one pyridine-methyl group (linker) on the nucleobase as the ligand to successively attach the effective drug [a Ru(III) complex, which will be discussed in detail in Chapter 3];
- one hydrophilic oligo(ethylene glycol) unit in position 5’;
- one or two (in the case of uridine) lipophilic oleyl or cholesteroxyacetyl groups covalently anchored in position 3’, or 2’ and 3’.

A schematic representation of the novel derivatives here proposed is shown in Figure 2.

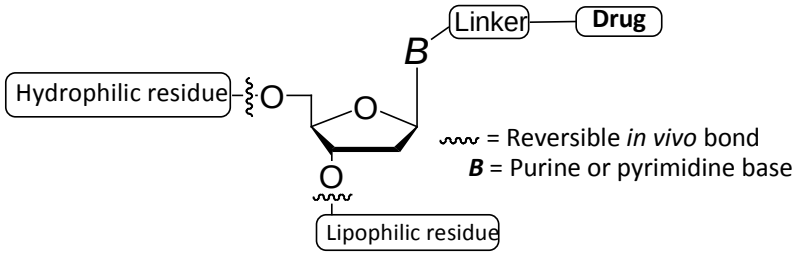


Figure 2. Representation of the new amphiphilic nucleolipids here described.

Specifically, as optimized substrates inspired to a prototype previously described by Paduano and Ruffo, [23] the nucleolipids here synthesized - depicted in Figure 3 - are: 3-(4-pyridylmethyl)-3'-*O*-oleyl-5'-*O*-(monomethoxy)triethylene glycol-acetyl-thymidine (ToThy, **1a**), 3-(4-pyridylmethyl)-3'-*O*-oleyl-5'-*O*-(benzyloxy)hexaethylene glycol-acetyl-thymidine (HoThy, **1b**), 3-(4-pyridylmethyl)-2',3'-di-*O*-oleyl-5'-*O*-(benzyloxy)hexaethylene glycol-acetyl-uridine (DoHu, **1c**) and 3-(4-pyridylmethyl), 3'-*O*-cholesteroxyacetyl, 5'-*O*-(monomethoxy)triethylene glycol-acetyl-thymidine (ToThyChol, **1d**). [24,25]

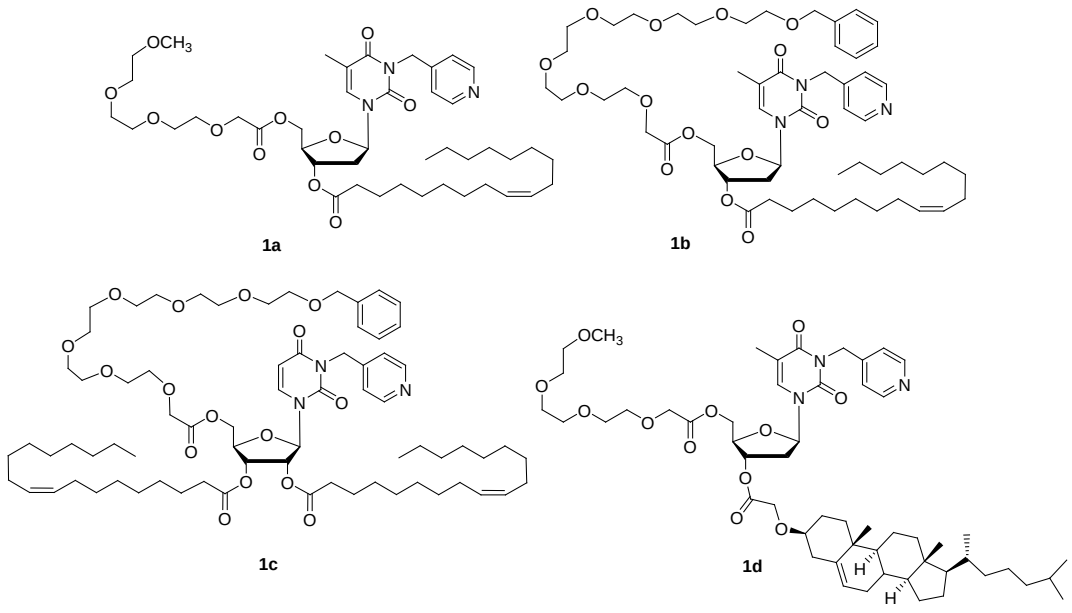


Figure 3. Functional nucleolipids: ToThy (**1a**), HoThy (**1b**), DoHu (**1c**) and ToThyChol (**1d**).

The physico-chemical characterization of these systems has been addressed to the study of the aggregation behaviour and of the morphology of the aggregates they form both in pure water and in pseudo-physiological conditions by means of dynamic light scattering (DLS) and small-angle neutron scattering (SANS). Cell culture experiments were

performed in order to get an insight into the *in vivo* potential of the developed nanocarriers on the basis of rightly interpreted *in vitro* results.

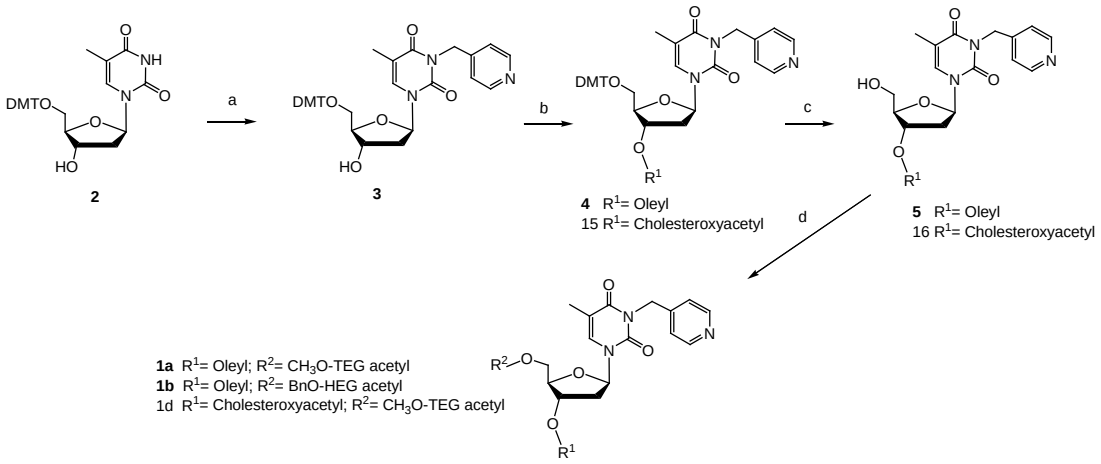
Nowadays a major concern of nanomedicine consists in understanding the contingent toxicity, and, more generally, the environmental impact of nanomaterials. Extensive research efforts are devoted to developing drug-transporting devices to target cells without affecting healthy ones and without showing undesired side effects. Hence, detailed biochemical and toxicological investigations are crucial to determine the feasibility of effective applications in biotechnological platforms for the new developed nanoaggregates. For this reason – in parallel with a detailed microstructural characterization of these systems, designed as new nanovectors - we focused on *in vitro* bioscreening by useful and standardised tumor and non tumor *in vitro* models, aiming at verifying if these nucleolipid aggregates affect cell growth and viability.

3. Results and discussion

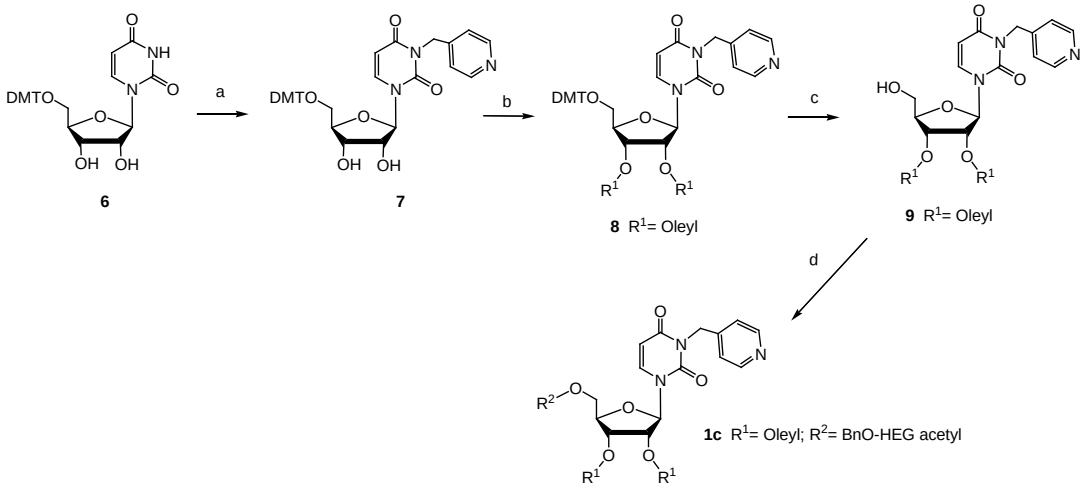
3.1 Synthesis of functional nucleosides ToThy, HoThy, DoHu and ToThyChol

The design of functional amphiphilic nucleosides **1a-d** (Figure 3) – following a previous work by Paduano and Ruffo [23] - was based on the usage of thymidine or uridine as the central scaffold, anchoring three diverse structural motifs: [24,25]

1) one pyridine-methyl arm, inserted at the *N*-3 position, as the functional reporter group, useful for successive manipulations and/or metal complexation;
2) one (in the case of thymidine derivatives **1a**, **1b** and **1d**) or two (in the case of uridine analog **1c**) lipophilic residues, attached at the ribose secondary hydroxyl group(s) to induce self-assembly in aqueous solutions;
3) one hydrophilic oligo(ethylene glycol) chain, linked at the 5'-OH ribose moiety; the role of the oligo(ethylene glycol) was here intended not only to optimize the "hydrophilic/lipophilic balance" within the hybrid molecules, but also – at a functional level - to prevent extracellular enzymatic degradation by the reticuloendothelial system.



Scheme 1. Synthetic scheme for the preparation of thymidine analogues **1a** and **1b**. Reaction conditions: **a)** 4-(bromomethyl)pyridine, K_2CO_3, DMF, 12 h, 60 °C; **b)** oleic acid, or cholesteroxy-acetic acid (**18**), DCC, DMAP, CH_2Cl_2, 12 h, r.t.; **c)** 1% TCA in CH_2Cl_2, 1 h, r.t.; **d)** CH_3O-TEG acetic acid (**11**), or BnO-HEG acetic acid (**14**), DCC, DMAP, CH_2Cl_2, 4 h, r.t.



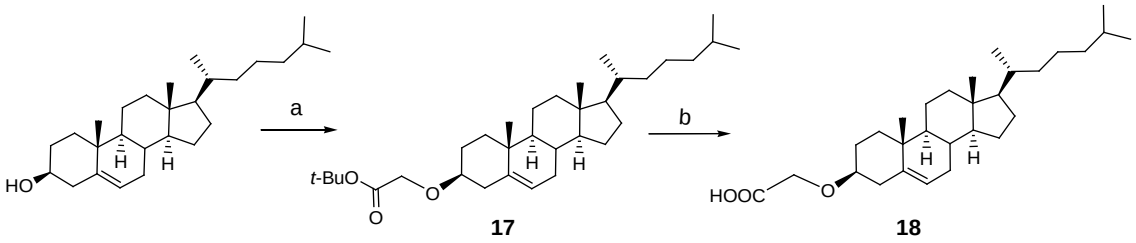
Scheme 2. Synthetic scheme for the preparation of uridine analog **1c**. Reaction conditions: **a)** 4-(bromomethyl)pyridine, K_2CO_3, DMF, 12 h, 60 °C; **b)** oleic acid, DCC, DMAP, CH_2Cl_2, 12 h, r.t.; **c)** 1% TCA in CH_2Cl_2, 2 h, r.t.; **d)** BnO-HEG acetic acid (**14**), DCC, DMAP, CH_2Cl_2, 3 h, r.t.

The pyridine residue, attached through a methylene linker, was here selected as the ligand allowing, in a successive step, for transition metals trapping into stable complexes (the synthesis and properties of which are discussed in detail in Chapter 3).

The synthesis of target molecules ToThy, HoThy, DoHu and ToThyChol was realized by a straightforward and high yielding strategy. Insertion of the pyridine moiety on the pyrimidine nucleobases was achieved in almost quantitative yields by treatment of the 5'-DMT protected nucleosides **2** or **6**, respectively, with 4-(bromomethyl)pyridine in DMF using K_2CO_3 as the base (Schemes 1 and 2). After a simple work-up, not requiring column chromatography, the alkylated nucleosides **3** and **7** were then subjected to the following reactions:

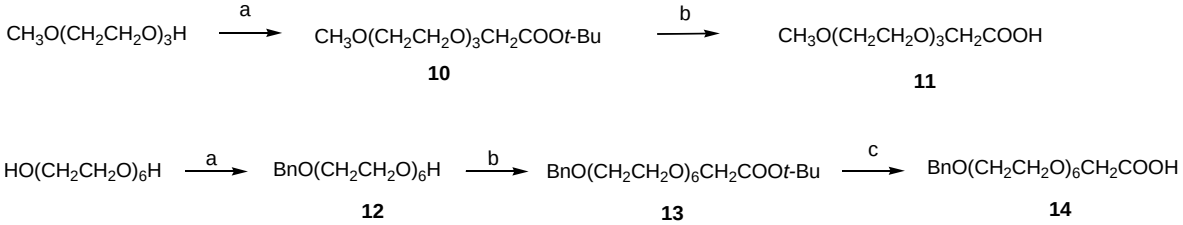
i) insertion of the fatty acid chain at the secondary hydroxyls to give **4**, **8** and **15**, respectively; **ii)** detritylation, yielding **5**, **9** and **16**, respectively, and, finally, **iii)** conjugation with the oligoethylene glycol chains at the 5'-end to give the target compounds **1a**, **1b**, **1c** and **1d**. In all cases, the hydrophilic and lipophilic components were attached to the sugar hydroxyl moieties through ester linkages, which are cleanly formed by classical DCC activation of the selected carboxylic acids. These chemical bonds are typically stable in neutral and extracellular media, but can be rapidly degraded in cells by esterases. For the lipophilic component, oleic acid and cholesteroxy-acetic acid were chosen as suitable building blocks due to both their structural similarity with lipid membranes constituents and easy availability. The synthesis of cholesteroxy-acetic acid

was performed starting from cholesterol, that was first converted into ester **17** by reaction with *tert*-butyl bromoacetate and NaH in THF, followed by treatment with formic acid, assuring *tert*-butyl removal and recovery of target acid **18** in almost quantitative yields (Scheme 3).



Scheme 3. Synthetic scheme for the preparation of cholesteroxy-acetic acid (**18**).
Reaction conditions: **a)** NaH, *tert*-butyl bromoacetate, THF, 12 h, r.t.; **b)** HCOOH, 2 h, r.t.

After DMT removal, the nucleobase-alkylated nucleolipids **5**, **9** and **16** were decorated with different hydrophilic groups at their primary OH groups. For a fine tuning of their hydrophilic/lipophilic balance, oligo(ethylene glycol) chains of different length were adopted. The starting materials were (monomethoxy)triethylene glycol (CH_3O-TEG) and hexaethylene glycol (HEG). These compounds were elaborated into acetic acid derivatives **11** and **14** (Scheme 4), following a straightforward and high yielding procedure, essentially similar to the synthetic scheme exploited to obtain cholesteroxy-acetic acid **18** from cholesterol. In the case of hexaethylene glycol, an additional step of benzylation to cap one OH group was required, leading to alcohol **12**, so to allow high yields in the following coupling with *tert*-butyl bromoacetate in the presence of NaH, furnishing ester **13**. After an acid treatment with formic acid, esters **10** and **13** were cleanly converted into target acids **11** and **14**, respectively, which were then coupled at the 5'-OH end of **5**, **9** or **16** , finally giving **1a**, **1b**, **1c** and **1d** in good overall yields (41-48 % for 4 steps).



Scheme 4. Synthetic scheme for the preparation of carboxylic acids **11** and **14**. Reaction conditions for **11**: **a)** NaH, *tert*-butyl bromoacetate, THF, 12 h, r.t.; **b)** HCOOH, 2 h, r.t. Reaction conditions for **14**: a) NaH, BnBr, THF, 12 h, r.t.; **b)** NaH, *tert*-butyl bromoacetate, THF, 12 h, rt.; **c)** HCOOH, 2 h, r.t.

3.2 Characterization of the aggregates formed by ToThy, HoThy, DoHu and ToThyChol

The characterization of the aggregates formed by ToThy, HoThy, DoHu and ToThyChol was accomplished in collaboration with the research group of Prof. Luigi Paduano of the Department of Chemistry "Paolo Corradini" in Naples.

The aggregation capabilities for all the systems were investigated by a combined use of DLS (*Dynamic Light Scattering*) and SANS (*Small Angle Neutron Scattering*) to obtain information about the dimension and morphology of the aggregates. DLS experiments were conducted in H_2O and in pseudo-physiological conditions to measure the diffusion coefficients and hydrodynamic radii of the aggregates. As reported in Figure 4, we can observe that while ToThy and ToThyChol present a mono-modal distribution both in water and in pseudo-physiological conditions, HoThy and DoHu show a bimodal or monomodal distribution of the aggregates depending on the environmental conditions. In pure water, DLS measurements showed the presence of two migrating species with different mean diffusion coefficients and in turn two aggregates size distributions were observed.

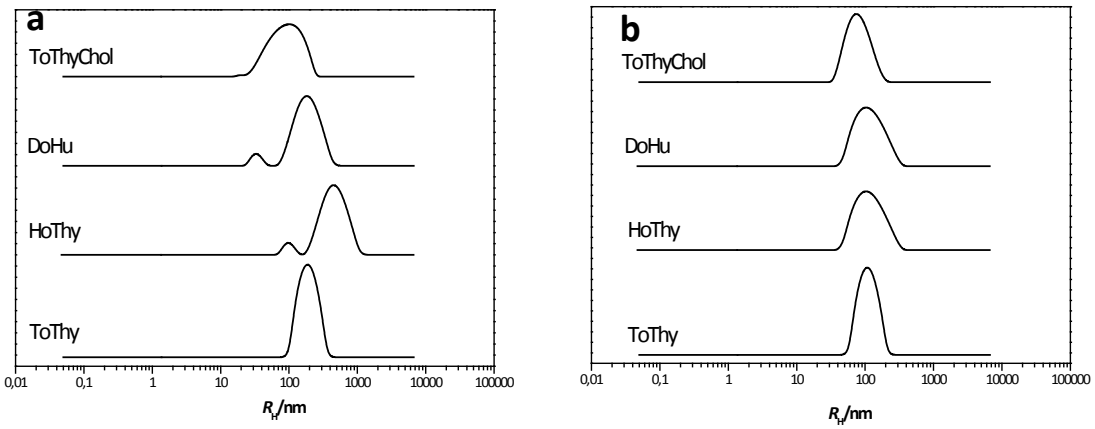


Figure 4. Examples of the hydrodynamic radii distribution of the aggregates formed by nucleolipids in pure water **(a)** and in pseudo-physiological conditions **(b)**. For all cases the concentration was 0.2 mmol kg^{-1}.

In pseudo-physiological conditions all the systems show only a single distribution and, notably, for ToThy a sensible reduction of the particles size has been observed. In this system, as well as in pure water, the size of the aggregates of the investigated

samples was found to be quite constant in the concentration range explored. The overview of these experimental results suggests that in pure water the molecules here presented, with the exception of ToThy and ToThyChol, show some tendency to form two aggregates with different morphology. This is very clear for DoHU where two dimensionally different aggregates are detected by DLS whereas for HoThy the presence of small aggregates is less evident. In both cases these supramolecular species show a hydrodynamic radius larger than the length of the molecule in a fully extended conformation. In fact, in the case of spherical micelles, the radius should be of the same order of the fully extended molecule, *i.e.* ~ 3.5 nm for HoThy and DoHu, well below the measured values (80-90 nm). On this basis, the small aggregates formed by HoThy and DoHu systems can be described as rod-like micelles. In contrast, in pseudo-physiological conditions all the systems exclusively show the presence of bigger aggregates, *i.e.* vesicles, while the distribution ascribable to the micelles disappears. This is confirmed by SANS results that do not reveal the presence of any correlation peak due to small aggregates and suggest the contemporary presence of a population of single and multilamellar vesicles, while for ToThyChol the aggregates are ascribed to rod-like micelles (see Table 1). As a general consideration, in pseudo-physiological conditions the radius of the vesicles is reduced with respect to that measured in pure water, though their size distribution broadens. This was also observed in other systems based on amphiphilic molecules carrying oligo(ethylene glycol) moieties in the head group. In fact, the presence of salts forces the molecules to aggregate in shape of vesicles, due to unfavorable interactions between the alkyl chains and the solvent, and the ethylene glycol chains and the solvent as well. The presence of salts increases the polar nature of the solvent and promotes the formation of a better packed structure of the alkyl chains, so that it is less accessible to ionic species. This is reinforced by the unfavorable interactions between the oligo(ethylene glycol) moieties and the salts. [26]

System in pseudo-physiological conditions, conc. 0.2 mmol kg^{-1}	Morphology	τ (nm)	*N*	*d* (nm)
ToThy	Vesicle	4.4±0.2	2±1	12±3
HoThy	Vesicle	5.0±0.1	4±1	24±3
DoHu	Vesicle	5.2±0.3	3±1	20±5
ToThyChol	Micelle	-	-	6±2

Table 1. SANS results for the investigated systems.

τ = layer thickness; ***N*** = number of consecutive bilayers; ***d*** = distance from the center of two consecutive bilayers.

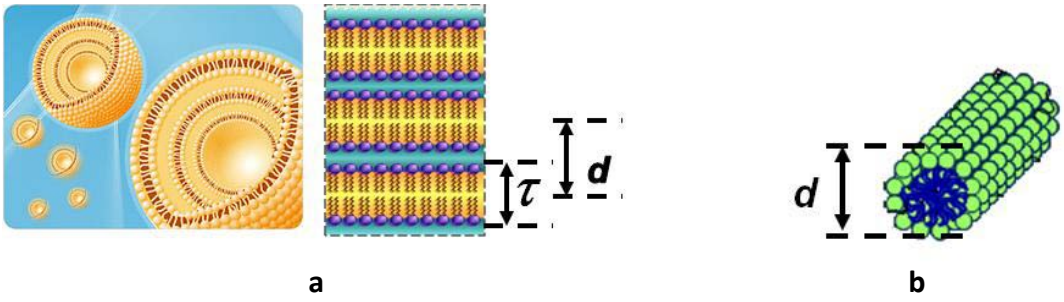


Figure 5. Schematic representation of the aggregates morphology.

a) Multilamellar vesicles; **b)** rod-like micelles.

3.3 Bioactivity and in vitro cytotoxicity

Biochemical and toxicological investigations were performed to screen and validate the new developed technological platform. In particular, we focused on the evaluation of cell growth and viability following the administration of the nucleolipid aggregates on different cultured tumor and non-tumor cell lines. For comparison, in the same experimental conditions, we evaluated the bioactivity of conventional POPC-based liposomes, which are currently used in a broad range of pharmaceutical applications for systemic drug delivery and for which no adverse effects have been reported to date.

As depicted in Table 2, DoHu, HoThy and ThoThyChol aggregates do not show toxicological relevance within the used cell line panel, even when administered at high concentrations (250 μM) for 72 h. These results are similar to those obtained from POPC bioscreening. The related IC_{50} values (see Table 2) were higher than 103 μM for all the examined cell populations. Typically, IC_{50} values of this magnitude indicate the absence of bioactivity and would suggest the theoretical safety of these nanoaggregates towards biostructures, thus allowing to use them in a wide range of concentrations. Conversely, ToThy showed a weak but significant biological activity. In fact, even though at high concentrations (250 μM), a noteworthy cytotoxicity was observed toward certain tumor cells as neuroblastoma (SH-SY5Y cell line), glioma (C6 cell line) and cervical cancer (HeLa cell line). This activity was associated with IC_{50} values ranging from 65 to 180 μM, indicating some ability to interfere with cell viability and/or proliferation (Figure 5). As a consequence, the ToThy nucleolipidic aggregate appears unsuitable for drug delivery applications because of its intrinsic cytotoxicity. However, this is not a surprising result since hybrid molecules as nucleolipids, besides occurring in eukaryotic and prokaryotic cells in which they play a central role in some metabolic pathways, frequently show biological features like antimicrobial, antifungal, antiviral and antitumor activities. [27-30] Indeed, several research groups have tried in the past to use the pharmacologically active nucleolipids as therapeutic agents and/or as pro-drugs. [31-33] In addition, the occurrence of cytotoxic and antiproliferative activities can be probably related to aggregate stability, uptake and internalization pathways. Future studies, aimed at understanding the mechanisms behind the bioactivity showed by ToThy, will provide precious information to obtain a picture of the structure-activity relationships of these

nucleolipids, so to develop a second generation of nucleolipids with improved biochemical and biotechnological properties. In this frame, considering that neuroblastoma and glioma cells were the most sensitive to ToThy treatments, the design of nucleolipid derivatives endowed with a certain degree of cell and tissue specificity may also be envisaged, which is a perspective of paramount importance in the research of supramolecular devices with specific *in vivo* target recognition.

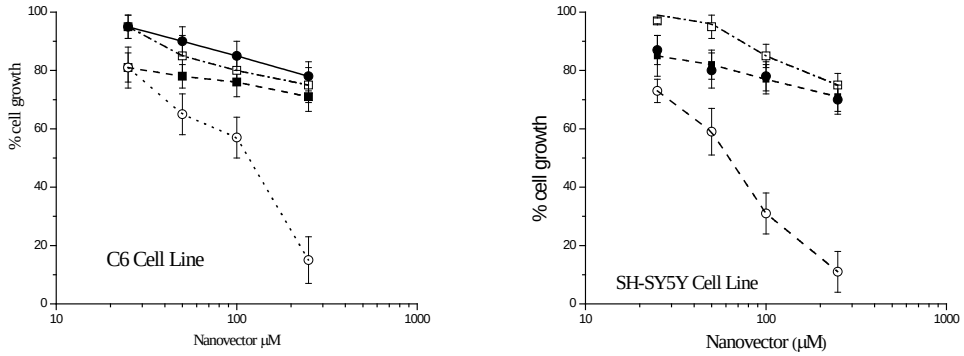


Figure 5. Examples of cell growth and viability evaluations performed on the indicated cell lines treated for 72 h with POPC (■), ToThy (○), HoThy (□) and DOHU (●) nanovectors

	IC_{50} (μM)				
Nanovectors	**Cell lines**				
	3T3-L1	**C6**	**CaCo-2**	**HeLa**	**SH-SY5Y**
POPC	$>10^3$	$> 10^3$	$>10^3$	$>10^3$	$>10^3$
ToThy	475 ± 35	130 ± 20	380 ± 25	180 ± 20	65 ± 15
HoThy	$>10^3$	$>10^3$	$>10^3$	$>10^3$	$>10^3$
DoHu	$>10^3$	$>10^3$	$>10^3$	$>10^3$	$>10^3$
ToThyChol	$>10^3$	$>10^3$	$>10^3$	$>10^3$	$>10^3$

Table 2. IC_{50} values for the investigated nanovectors.

Here tested cell lines: 3T3-L1, murine embryonic fibroblasts (non cancer cells); HeLa, human cervical cancer cells; C6, rat glioma cells; CaCo-2, human epithelial colorectal adenocarcinoma cells; SH-SY5Y, human neuroblastoma cells.

To the best of our knowledge, this is one of the first cytotoxicity investigations performed on amphiphilic nucleolipids. [34,35] Since the cytotoxicity of DoHu, HoThy and ToThyChol was shown to be virtually null, the supramolecular systems obtained can be regarded as promising tools, leading to new types of functional, biocompatible materials, as well as vehicles for biomolecules transport and drug delivery. Therefore, these results show nanocarriers based on HoThy, DoHu and ToThyChol nucleolipids as realistic, effective alternatives to conventional lipid-based micelles.

4. Conclusions

According to the large number of recent publications, amphiphilic hybrid molecules based on nucleolipids are attracting increasing interest in a variety of biotechnological fields, with relevant applications both in the design of artificial molecular devices and in novel therapeutic strategies. In this context, we here described the design, synthesis and supramolecular organization of three novel synthetic nucleolipids (here named ToThy, HoThy, DoHu and ToThyChol), as well as a preliminary investigation of their cytotoxic activity on a panel of tumor and non-tumor cell lines. Using thymidine and uridine as the central, polyfunctional scaffold, insertion of one oligo(ethylene glycol) residue and one or two lipophilic chains on the ribosidic moiety confer to these nucleolipids the desired amphiphilic behaviour in aqueous solutions. A pyridine-methyl arm was also introduced on the nucleobase as a functional ligand, available for successive complexation with transition metals, useful for diagnostic and therapeutic applications. Studies on the synthesis, biophysical characterization and biological evaluation of Ru(III) complexes of these molecules are reported in the Chapter 3.

ToThy, HoThy, DoHu, ToThyChol nucleolipids, prepared in few, straightforward and high yielding synthetic steps, showed the desired distinctive features of nanocarriers, being mainly organized as multilayer vesicles or large micelles in pseudo-physiological conditions. For *in vivo* applications, a vesicular nanocarrier is more efficient than a micellar one due to the bigger size, with a more efficient delivery effect obtained if the aggregates size is within the 50-1000 nm range. [36] As determined by DLS and SANS measurements, the hydrodynamic radii of ToThy, HoThy and DoHu vesicular carriers are between 50 and 400 nm. In pure water, HoThy and DoHu also form cylindrical micelles. For ToThyChol we observed in pseudo-physiological conditions the formation of rod-like micelles.

Bioscreening investigations revealed for HoThy, DoHu and ToThyChol the complete absence of significant cytotoxicity and toxicological effects on both normal and tumor cells. Given the utmost importance of delivering drugs and targeting specific cells, these nucleolipid-based amphiphiles with no associated cytotoxicity are very attractive candidate carriers, rationally designed to incorporate metal-based drugs in stable complexes. Conversely, due to the emergence of some cytotoxicity, further investigations

will be addressed to the elucidation of the molecular basis of ToThy bioactivity. Taken together, these results clearly indicate a novel and promising approach for the development of specific nanocarrier devices, consisting of finely tunable multifunctional materials with a versatile, *ad hoc* tailored chemical structure.

4. Experimental section

General Methods

All the reagents were of the highest commercially available quality and were used as received. TLC analyses were carried out on silica gel plates from Merck (60, F254). Reaction products on TLC plates were visualized by UV light and then by treatment with a 10 % $Ce(SO_4)_2/H_2SO_4$ aqueous solution. For column chromatography, silica gel from Merck (Kieselgel 40, 0.063-0.200 mm) was used. NMR spectra were recorded on Bruker WM-400, Varian Gemini 200 and Varian Inova 500 spectrometers, as specified. All the chemical shifts are expressed in ppm with respect to the residual solvent signal. Peak assignments have been carried out on the basis of standard ^1H-^1H COSY and HSQC experiments. The following abbreviations were used to explain the multiplicities: s = singlet; d = doublet; t = triplet; q = quartet; m = multiplet; b = broad; dd = double doublet. For the ESI MS analyses, a Waters Micromass ZQ instrument – equipped with an Electrospray source – was used in the positive and/or negative mode. MALDI TOF mass spectrometric analyses were performed on a PerSeptive Biosystems Voyager-De Pro MALDI mass spectrometer in the Linear mode using 3,4-dihydroxybenzoic acid as the matrix.

Synthesis of 3-(4-pyridylmethyl)-3'-*O*-oleyl-5'-*O*-(4,4'-dimethoxytriphenylmethyl) -thymidine (4).

5'-*O*-(4,4'-dimethoxytriphenylmethyl)thymidine (**2**) (350 mg, 0.64 mmol) was dissolved in 5 ml of dry DMF. K_2CO_3 (267 mg, 1.93 mmol) and 4-(bromomethyl)pyridine hydrobromide (245 mg, 0.97 mmol) were then added to the reaction mixture, left at 60 °C under stirring. After 12 h, TLC analysis indicated the presence of a single product in the reaction mixture. Alkylated compound **3** was not isolated, but directly converted into its 3'-conjugated derivative **4**. To this purpose, the solvent was removed under reduced pressure and the residual solid was taken up in DCM, washed with water, dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. Successively, the residue was dissolved in 4 ml of dry DCM and then DMAP (117 mg, 0.96 mmol), oleic acid (304 µL, 0.96 mmol) and DCC (198 mg, 0.96 mmol) were sequentially added. After stirring for 12 h at room

temperature the reaction mixture was concentrated under reduced pressure; the crude product was purified by chromatography on a silica gel column using *n*-hexane/AcOEt (7:3, v/v, containing 1% of TEA) as eluent, yielding desired compound **4** in 64% for the two steps (368 mg, 0.41 mmol).

4: oil**:** R_f = 0.8 ($CHCl_3/CH_3OH$, 95:5, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.53 [d, *J* = 6.0 Hz, 2H, (2x H_α Py)]; 7.64 (s, 1H, H-6); 7.37-6.82 (overlapped signals, 15H, 2x H_β Py and aromatic protons of *DMT*); 6.45 (dd, *J* = 8.0 and 6.5 Hz, 1H, H-1'); 5.45 (m, 1H, H-3'); 5.34 (m, 2H, H-9 and H-10 oleic acid); 5.10 (s, 2H, -C$\underline{H}_2$Py); 4.11 (m, 1H, H-4'); 3.78 [s, 6H, 2x (-OC$\underline{H}_3$) of *DMT*]; 3.46 (m, 2H, H_2-5'); 2.44 (m, 2H, H-2'); 2.30 (apparent t, *J* = 8.0 and 7.5 Hz, 2H, C$\underline{H}_2$-2 oleic acid); 2.00-1.26 [overlapped signals, 26H, 13x (-C$\underline{H}_2$- oleic acid)]; 1.41 (s, 3H, C$\underline{H}_3$-Thy); 0.87 (t, *J* = 6.5 and 6.5 Hz, 3H, C$\underline{H}_3$-18 oleic acid).

^{13}C NMR ($CDCl_3$, 100 MHz): δ 173.0 (C=O ester); 163.0 (C-4); 158.4, 144.0, 133.8, 127.9, 127.0 and 113.2 (aromatic carbons of *DMT*); 150.8 (C-2); 149.8 (2x C_α Py); 145.3 (C_γ Py); 135.0 (C-6); 129.9 and 129.5 (C-9 and C-10 oleic acid); 123.4 (2x C_β Py); 110.7 (C-5); 87.0 (quaternary carbon of *DMT*); 85.0 (C-1'); 84.0 (C-4'); 74.8 (C3'); 63.5 (C-5'); 55.1 [2x (O$\underline{C}H_3$) of *DMT*]; 43.4 (-$\underline{C}H_2$Py); 37.9 (C-2'); 34.0, 31.7, 29.5, 29.1, 28.9, 27.0, 24.6 and 22.5 (aliphatic carbons of oleic acid); 13.9 ($\underline{C}H_3$-18 oleic acid); 12.2 ($\underline{C}H_3$-Thy).

MALDI-MS (positive ions): calculated for $C_{55}H_{69}N_3O_8$ 899.51; found m/z: 901.93 ($M+H^+$).

Synthesis of 3-(4-pyridylmethyl)-3'-*O*-oleyl-thymidine (5).

4 (237 mg, 0.26 mmol) was dissolved in 3 ml of a 1% TCA solution in DCM and stirred for 1 h at room temperature. The solution was then concentrated *in vacuo* and purified by chromatography on a silica gel column eluted with $CHCl_3/CH_3OH$ (99:1, v/v), giving the desired compound **5** in almost quantitative yields (155 mg, 0.26 mmol).

5: oil**:** R_f = 0.4 ($CHCl_3/CH_3OH$, 95:5, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.67 [d, *J* = 4.0 Hz, 2H, (2x H_α Py)]; 7.63 (s, 1H, H-6); 7.57 [d, *J* = 5.0 Hz, 2H, (2x H_β Py)]; 6.29 (dd, *J* = 7.5 and 7.0 Hz, 1H, H-1'); 5.34 (overlapped signals, 3H, H-3', H-9 and H-10 oleic acid); 5.19 (s, 2H, -C$\underline{H}_2$Py); 4.08 (m, 1H, H-4'); 3.93 (m, 2H, H_2-5'); 2.40 (m, 2H, C$\underline{H}_2$-2 oleic acid); 2.33 (m, 2H, H-2'); 2.01 (m, 2H, C$\underline{H}_2$-8 oleic acid); 1.96 (s, 3H, C$\underline{H}_3$-Thy); 1.62 (m, 2H, C$\underline{H}_2$-11 oleic acid); 1.28 [m, 22H, 11x (-C$\underline{H}_2$- oleic acid)]; 0.88

(apparent t, *J* = 6.0 and 7.5 Hz, 3H, C<u>H</u>$_3$-18 oleic acid).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 173.4 (C=O ester); 162.9 (C-4); 150.7 (C-2); 146.1 (2x C_α Py); 145.0 (C_γ Py); 134.8 (C-6); 129.9 and 129.6 (C-9 and C-10 oleic acid); 124.8 (2x C_β Py); 110.5 (C-5); 86.4 (C-1'); 85.1 (C-4'); 74.2 (C3'); 62.4 (C-5'); 43.4 (-<u>C</u>H$_2$Py); 37.3 (C-2'); 34.0, 31.8, 29.5, 29.4, 29.2, 28.9, 27.0, 24.7 and 22.6 (aliphatic carbons of oleic acid); 14.0 (<u>C</u>H$_3$-18 oleic acid); 13.2 (<u>C</u>H$_3$-Thy).

MALDI-MS (positive ions): calculated for $C_{34}H_{51}N_3O_6$ 597.38; found m/z: 596.19 ($M+H^+$); 618.57 ($M+Na^+$); 635.86 ($M+K^+$)

Synthesis of 3-(4-pyridylmethyl)-3'-*O*-oleyl-5'-*O*-(monomethoxy)triethylene glycol acetyl-thymidine (1a).

5 (88 mg, 0.15 mmol) was dissolved in 1.5 ml of dry DCM and then DMAP (27 mg, 0.22 mmol), carboxylic acid **11** (49 mg, 0.22 mmol) and DCC (46 mg, 0.22 mmol) were added. The reaction mixture, left under stirring for 4 h at room temperature, was then concentrated under reduced pressure; the crude product was purified by chromatography on a silica gel column eluted with *n*-hexane/AcOEt (1:9, v/v), giving the desired compound **1a** in 67% yield (80 mg, 0.10 mmol).

1a: oil**:** R_f = 0.5 ($CHCl_3$/CH_3OH, 95:5, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz, 0.06 M): δ 8.53 [d, *J* = 4.0 Hz, 2H, (2x H_α Py)]; 7.36-7.34 [overlapped signals, 3H, H-6 and (2x H_β Py)]; 6.35 (dd, *J* = 6.0 and 6.0 Hz, 1H, H-1'); 5.33 (m, 2H, H-9 and H-10 oleic acid); 5.20 (m, 1H, H-3'); 5.11 (s, 2H, -C<u>H</u>$_2$Py); 4.42 (m, 2H, H$_2$-5'); 4.26-4.13 (overlapped signals, 3H, H-4' and -OC<u>H</u>$_2$COO-C5'); 3.73-3.50 [overlapped signals, 12H, 3x (-O-C<u>H</u>$_2$-C<u>H</u>$_2$-O-)]; 3.35 (s, 3H, C<u>H</u>$_3$O-TEG); 2.42-2.24 (m, 2H, H-2'); 2.31 (m, 2H, C<u>H</u>$_2$-2 oleic acid); 2.00 (overlapped signals, 4H, C<u>H</u>$_2$-8 and C<u>H</u>$_2$-11 oleic acid); 1.95 (s, 3H, C<u>H</u>$_3$-Thy); 1.60 (m, 2H, C<u>H</u>$_2$-3 oleic acid); 1.29-1.23 [overlapped signals, 20H, 10x (-C<u>H</u>$_2$- oleic acid)]; 0.86 (apparent t, *J* = 6.0 and 7.0 Hz, 3H, C<u>H</u>$_3$-18 oleic acid).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 173.1 (C=O oleic acid ester); 169.7 (C=O TEG ester); 162.8 (C-4); 150.7 (C-2); 149.0 (2x C_α Py); 146.2 (C_γ Py); 133.4 (C-6); 129.9 (C-9 and C-10 oleic acid); 129.5 (2x C_β Py); 110.7 (C-5); 85.4 (C-1'); 82.0 (C-4'); 74.1 (C-3'); 71.8, 71.0, 70.5 and 70.4 (O-<u>C</u>H$_2$-<u>C</u>H$_2$-O-TEG); 68.6 (--O<u>C</u>H$_2$COO-C5'); 64.1 (C-5'); 58.9 (<u>C</u>H$_3$O-TEG); 43.5 (-<u>C</u>H$_2$Py); 36.9 (C-2'); 33.9 (<u>C</u>H$_2$-2 oleic acid); 31.7 (<u>C</u>H$_2$-8 and <u>C</u>H$_2$-11 oleic acid); 29.5, 29.4, 29.2,

28.9, 27.0, 24.6 and 22.5 (aliphatic carbons of oleic acid); 14.0 (CH_3-18 oleic acid); 13.1 (CH_3-Thy).

MALDI-MS (positive ions): calculated for $C_{43}H_{67}N_3O_{11}$ 801.48; found m/z: 802.89 ($M+H^+$); 824.93 ($M+Na^+$); 840.90 ($M+K^+$).

Synthesis of 3-(4-pyridylmethyl)-3'-*O*-oleyl-5'-*O*-(benzyloxy)hexaethylene glycol-acetyl-thymidine (1b).

5 (64 mg, 0.11 mmol) was dissolved in 1.0 ml of dry DCM and then DMAP (20 mg, 0.16 mmol), carboxylic acid **14** (50 mg, 0.13 mmol) and DCC (34 mg, 0.16 mmol) were sequentially added. The reaction mixture was stirred for 4 h and then the solvent was removed *in vacuo*; the crude product was purified by chromatography on a silica gel column with AcOEt as eluent, giving the desired compound **1b** in 64% yields (69 mg, 0.070 mmol).

1b: oil**:** R_f = 0.6 (AcOEt/CH_3OH, 98:2, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz, 0.06 M): δ 8.53 (overlapped signals, 2H, 2x H_α Py); 7.35-7.26 (overlapped signals, 8H, H-6, 2x H_β Py and 5x aromatic protons of Bn); 6.36 (dd, *J* = 6.0 and 6.0 Hz, 1H, H-1'); 5.36 (m, 2H, H-9 and H-10 oleic acid); 5.21 (m, 1H, H-3'); 5.10 (s, 2H, -CH_2Py); 4.56 (s, 2H, -CH_2-Ph); 4.44 (m, 2H, H_2-5'); 4.26-4.15 (overlapped signals, 3H, H-4' and -OCH_2COO-C5'); 3.74-3.57 [overlapped signals, 24H, 6x (-O-CH_2-CH_2-O-)]; 2.42-2.24 (m, 2H, H-2'); 2.32 (m, 2H, CH_2-2 oleic acid); 2.00 (overlapped signals, 4H, CH_2-8 and CH_2-11 oleic acid); 1.96 (s, 3H, CH_3-Thy); 1.61 (m, 2H, CH_2-3 oleic acid); 1.30-1.25 [overlapped signals, 20H, 10x (-CH_2- oleic acid)]; 0.87 (apparent t, *J* = 6.5 and 7.0 Hz, 3H, CH_3-18 oleic acid).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 173.1 (C=O oleic acid ester); 169.7 (C=O HEG ester); 162.8 (C-4); 150.7 (C-2); 149.8 (2x C_α Py); 145.2 (C_γ Py); 138.2 (-CH_2-Ph); 133.2 (C-6); 129.9 (C-9 and C-10 oleic acid); 129.5 (2x C_β Py); 128.2 and 127.6 (aromatic carbons of Bn); 110.7 (C-5); 85.4 (C-1'); 82.0 (C-4'); 74.0 (C-3'); 73.1 (-CH_2Ph); 71.0, 70.5 and 69.3 (O-CH_2-CH_2-O-HEG); 68.6 (-OCH_2COO-C5'); 64.1 (C-5'); 43.5 (-CH_2Py); 37.0 (C-2'); 34.0 (CH_2-2 oleic acid); 31.7 (CH_2-8 and CH_2-11 oleic acid); 29.6, 29.4, 29.2, 29.0, 27.1, 24.6 and 22.5 (aliphatic carbons of oleic acid); 14.0 (CH_3-18 oleic acid); 13.1 (CH_3-Thy).

ESI-MS (positive ions): calculated for $C_{55}H_{83}Cl_4N_3O_{14}$, 1009.6; found *m/z*: 1010.8 ($M+H^+$);

1048.7 ($M+K^+$).

Synthesis of 3-(4-pyridylmethyl)-2',3'-di-*O*-oleyl-5'-*O*-(4,4'-dimethoxytriphenylmethyl)-uridine (8).

5'-*O*-(4,4'-dimethoxytriphenylmethyl)uridine **6** (326 mg, 0.59 mmol) was dissolved in 5 ml of dry DMF. K_2CO_3 (370 mg, 2.68 mmol) and 4-(bromomethyl)pyridine hydrobromide (450 mg, 1.77 mmol) were added and the reaction mixture was left under stirring at 60 °C. After 12 h, TLC analysis indicated the presence of a single product in the reaction mixture. Alkylated compound **7** was not isolated, but directly converted into its 2',3'-bis-conjugated derivative **8**. To this purpose, the solvent was removed *in vacuo* and the residual solid was taken up in DCM, washed with water, dried over anhydrous Na_2SO_4 and then concentrated under reduced pressure. Successively the residue was dissolved in 2.5 ml of dry DCM and then DMAP (216 mg, 1.77 mmol), oleic acid (473 µL, 1.49 mmol) and DCC (365 mg, 1.77 mmol) were sequentially added to the reaction mixture. After 12 h under stirring at room temperature, the solvent was removed *in vacuo* and the crude product was purified by chromatography on a silica gel column with *n*-hexane/AcOEt (6:4, v/v, containing 1% of TEA) as eluent, affording pure compound **8** in 60% yields for the two steps (416 mg, 0.36 mmol).

8: oil**:** R_f = 0.4 (*n*-hexane/AcOEt, 1:1, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.52 [d, J = 5.5 Hz, 2H, (2x H_α Py)]; 7.76 (d, J = 8.5 Hz, 1H, H-6); 7.38-6.83 (overlapped signals, 15H, aromatic protons of *DMT* and 2x H_β Py); 6.18 (d, J = 5.5 Hz, 1H, H-1'); 5.57 (overlapped signals, 2H, H-2' and H-3'); 5.43 (d, J = 8.0 Hz, 1H, H-5); 5.36-5.31 [overlapped signals, 4H, 2x (H-9 and H-10 oleic acid)]; 5.05 (m, 2H,-C$\underline{H}_2$Py); 4.21 (m, 1H, H-4'); 3.78 [s, 6H, 2x (-OC$\underline{H}_3$) of *DMT*]; 3.48 (m, 2H, H_2-5'); 2.29 [m, 4H, 2x (C$\underline{H}_2$-2 oleic acid)]; 1.99 [overlapped signals, 8H, [2x (C$\underline{H}_2$-8) and 2x (C$\underline{H}_2$-11) oleic acid]; 1.59 [overlapped signals, 4H, 2x (C$\underline{H}_2$-3 oleic acid)]; 1.94-1.91, 1.71-1.65, 1.39-1.05 [overlapped signals, 40H, 2x (10x -C$\underline{H}_2$- oleic acid)]; 0.92-0.86 [overlapped signals, 6H, 2x (C$\underline{H}_3$-18 oleic acid)].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 172.4 and 172.2 (C=O esters); 162.0 (C-4); 158.7, 130.1, 130.0, 129.6, 128.1, 127.2, 123.1 and 113.3 (aromatic carbons of *DMT* and 2x C_α Py);

149.8 (2x C_β Py); 145.2 (C_γ Py); 143.8 (C-2); 138.1 (C-6); 134.8 and 134.6 (C-9 and C-10 oleic acid); 102.3 (C-5); 87.5 (C-1'); 86.7 (quaternary carbon of *DMT*); 82.0 (C-4'); 73.1 and 70.6 (C-2' and C3'); 62.3 (C-5'); 55.2 [2x (-O<u>C</u>H_3) of *DMT*]; 43.2 (-<u>C</u>H_2Py); 33.9 and 33.7 [2x (-<u>C</u>H_2-2 oleic acid)]; 31.8, 29.7, 29.5, 29.3, 29.1, 29.0, 27.2, 24.7, 24.6 and 22.6 (aliphatic carbons of oleic acid); 14.1 [2x (<u>C</u>H_3-18 oleic acid)].

ESI-MS (positive ions): calculated for $C_{72}H_{99}N_3O_{10}$, 1165.7; found *m/z*: 1167.0 ($M+H^+$); 1205.0 ($M+K^+$).

Synthesis of 3-(4-pyridylmethyl)-2',3'-di-*O*-oleyl-uridine (9).

Nucleoside **8** (400 mg, 0.34 mmol) was dissolved in 6 ml of a 1% TCA solution in DCM and stirred for 2 h at room temperature. The solution was then concentrated *in vacuo* and purified by chromatography on a silica gel column with *n*-hexane/AcOEt (3:7, v/v) as eluent, yielding the desired compound **9** in almost quantitative yields (290 mg, 0.34 mmol).

9: oil**:** R_f = 0.4 (AcOEt/*n*-hexane, 3:2, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.69 [broad signal, 2H, (2x H_α Py)]; 7.79 (d, *J* = 8.0 Hz, 1H, H-6); 7.59 [broad signal, 2H, (2x H_β Py)]; 5.97 (d, *J* = 5.5 Hz, 1H, H-1'); 5.89 (d, *J* = 8.0 Hz, 1H, H-5); 5.49-5.44 (overlapped signals, 2H, H-2' and H-3'); 5.36-5.29 [overlapped signals, 4H, 2x (H-9 and H-10 oleic acid)]; 5.18 (m, 2H, -<u>C</u>H_2Py); 4.21 (m, 1H, H-4'); 3.99-3.84 (m, 2H, H_2-5'); 2.35 and 2.29 [t's, *J* = 7.5 and 7.5 Hz, 2H each, 2x (<u>C</u>H_2-2 oleic acid)]; 1.99 [overlapped signals, 8H, 2x (<u>C</u>H_2-8) and 2x (<u>C</u>H_2-11) oleic acid]; 1.59 [overlapped signals, 4H, 2x (<u>C</u>H_2-3 oleic acid)]; 1.31-1.26 [overlapped signals, 40H, 2x (10x -<u>C</u>H_2- oleic acid)]; 0.89-0.86 [overlapped signals, 6H, 2x (<u>C</u>H_3-18 oleic acid)].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 172.7 and 171.2 [2x (C=O esters)]; 162.1 (C-4); 151.0 (2x C_α Py); 145.4 (C_γ Py); 143.8 (C-2); 139.6 (C-6); 130.0 and 129.6 (C-9 and C-10 oleic acid); 125.3 (2x C_β Py); 102.3 (C-5); 88.8 (C-1'); 83.6 (C-4') ; 73.0 and 70.9 (C-2' and C3'); 60.4 (C-5'); 43.2 (-<u>C</u>H_2Py); 33.9 [overlapped signals, 2x (-<u>C</u>H_2-2 oleic acid)]; 31.8, 29.7, 29.5, 29.3, 29.1, 29.0, 27.2, 24.8 and 22.6 (aliphatic carbons of oleic acid); 14.1 and 14.0 [2x (<u>C</u>H_3-18 oleic acid)].

ESI-MS (positive ions): calculated for $C_{51}H_{81}N_3O_8$, 863.6; found *m/z*: 864.8 ($M + H^+$).

Synthesis of 3-(4-pyridylmethyl)-2',3'-di-*O*-oleyl-5'-*O*-(benzyloxy)hexaethylene glycol-acetyl-uridine (1c).

Nucleoside **9** (67 mg, 0.078 mmol) was dissolved in 1 ml of dry DCM and then DMAP (15 mg, 0.12 mmol), carboxylic acid **14** (40 mg, 0.09 mmol) and DCC (24 mg, 0.16 mmol) were sequentially added. The reaction mixture was stirred for 3 h and then the solvent was removed *in vacuo*; the crude product was purified by chromatography on a silica gel column with AcOEt as eluent, yielding the desired compound **1c** in 80% yields (80 mg, 0.063 mmol).

1c: oil**:** R_f = 0.6 (AcOEt/CH_3OH, 98:2, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.53 [d, *J* = 5.5 Hz, 2H, (2x H_α Py)]; 7.54 (d, *J* = 8.5 Hz, 1H, H-6); 7.34-7.26 [overlapped signals, 7H, (2x H_β Py) and (5x aromatic proton of Bn)]; 6.07 (d, *J* = 6.0 Hz, 1H, H-1'); 5.89 (d, *J* = 8.5 Hz, 1H, H-5); 5.39-5.37 (overlapped signals, 2H, H-2' and H-3'); 5.34 [overlapped signals, 4H, 2x (H-9 and H-10 oleic acid)]; 5.07 (m, 2H,-C$\underline{H}_2$Py); 4.56 (s, 2H, -C$\underline{H}_2$-Ph); 4.42-4.37 (m, 2H, H_2-5'); 4.34 (m, 1H, H-4'); 4.25 (s, 2H, -OC$\underline{H}_2$CO-HEG); 3.74-3.63 [overlapped signals, 24H, 6x (-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)]; 2.35 and 2.26 [t's, *J* = 7.5 Hz, 2H each, 2x (C$\underline{H}_2$-2 oleic acid)]; 2.01 and 1.99 [overlapped signals, 8H, [2x (C$\underline{H}_2$-8) and 2x (C$\underline{H}_2$-11) oleic acid]; 1.58 [overlapped signals, 4H, 2x (C$\underline{H}_2$-3 oleic acid)]; 1.30-1.26 [overlapped signals, 40H, 2x (10 x -C$\underline{H}_2$- oleic acid)]; 0.89-0.86 [overlapped signals, 6H, 2x (C$\underline{H}_3$-18 oleic acid)].

^{13}C NMR (CD_3OD, 50 MHz): δ 172.6, 170.4 and 163.0 (C=O esters); 158.7 (C-4); 151.2 (2x C_α Py); 149.1 (C_γ Py); 147.3 (C-2); 140.1 (C-6); 138.5 (-$\underline{C}H_2$-Ph); 129.8 (C-9 and C-10 oleic acid); 129.6 (2x C_β Py); 128.2, 127.7 and 127.5 (aromatic carbons of Bn); 101.2 (C-5); 89.4 (C-1'); 80.5 (C-4'); 73.0 and 70.8 (C-2' and C-3'); 70.4 (-$\underline{C}H_2$Ph); 69.5 (O-$\underline{C}H_2$-$\underline{C}H_2$-O- HEG); 68.4 (-O$\underline{C}H_2$CO- HEG); 63.3 (C-5'); 43.1 (-$\underline{C}H_2$Py); 33.6 [overlapped signals, 2x(-$\underline{C}H_2$-2 oleic acid)]; 31.9, 29.7, 29.5, 29.2, 29.1, 27.0, 25.6, 24.9 and 22.6 (aliphatic carbons of oleic acid); 13.3 [2x ($\underline{C}H_3$-18 oleic acid)].

ESI-MS (positive ions): calculated for $C_{72}H_{113}N_3O_{16}$, 1275.8; found *m/z*: 1277.2 ($M+H^+$); 1315.1 ($M+K^+$).

Synthesis of *tert*-butyl (monomethoxy)triethylene glycol acetate (10).

Monomethoxytriethylene glycol (1.5 g, 9.13 mmol) was dissolved in 6 ml of dry THF and NaH 60% p.p. (730 mg, 18.3 mmol) and then *tert*-butyl bromoacetate (3.37 ml, 22.8 mmol) were sequentially added. The solution was stirred at room temperature for 12 h, then CH_3OH (1 ml) was added and the solvent removed *in vacuo*. The crude product, dissolved in $CHCl_3$ and filtered on celite, was then purified by chromatography on a silica gel column with $CHCl_3$ as eluent, yielding the desired ester **10** in 98% yields (2.5 g, 8.96 mmol).

10: oil**:** R_f = 0.6 ($CHCl_3/CH_3OH$, 95:5, v/v).

^{1}H NMR ($CDCl_3$, 400 MHz): δ 3.91 (s, 2H, -C$\underline{H}_2$COO-*tert*-butyl); 3.60-3.44 [overlapped signals, 12H, 3x (-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)]; 3.27 (s, 3H, C$\underline{H}_3$O-); 1.37 [s, 9H, 3x (C$\underline{H}_3$ *tert*-butyl)].

^{13}C NMR ($CDCl_3$, 100 MHz): δ 169.3 (C=O ester); 81.0 (quaternary carbon of *tert*-butyl); 68.7 (-$\underline{C}$H$_2$COO-*tert*-butyl); 71.6 and 70.2 [3x (-O-$\underline{C}$H$_2$-$\underline{C}$H$_2$-O-)]; 58.6 ($\underline{C}$H$_3$O-); 27.7 [3x ($\underline{C}$H$_3$ *tert*-butyl)].

MALDI-MS (positive ions): calculated for $C_{13}H_{26}O_6$ 278.17; found m/z: 279.37 ($M+H^+$); 301.40 ($M+Na^+$); 317.38 ($M+K^+$).

Synthesis of (monomethoxy)triethylene glycol acetic acid (11).

Ester **10** (630 mg, 2.26 mmol) was dissolved in 1.5 ml of formic acid and stirred at room temperature for 2 h. The solvent was then removed *in vacuo* and the residue was coevaporated three times with $CHCl_3$ (3x 3 ml), yielding the desired carboxylic acid **11** in 99% yields (500 mg, 2.25 mmol).

11: oil**:** R_f = 0.4 ($CHCl_3/CH_3OH$, 95:5, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz): δ 4.13 (s, 2H, -C$\underline{H}_2$COOH); 3.70-3.50 [overlapped signals, 12H, 3x (-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)]; 3.34 (s, 3H, C$\underline{H}_3$O-).

^{13}C NMR ($CDCl_3$, 125 MHz): δ 170.9 (C=O carboxylic acid); 71.6, 70.6 and 70.3 [3x (-O-$\underline{C}$H$_2$-$\underline{C}$H$_2$-O-)]; 68.3 (-$\underline{C}$H$_2$COOH); 58.7 ($\underline{C}$H$_3$O-).

MALDI-MS (positive ions): calculated for $C_9H_{18}O_6$ 222.11; found m/z: 223.5 ($M+H^+$).

Synthesis of (monobenzyloxy)hexaethylene glycol (12).

Hexaethylene glycol (2.0 g, 7.1 mmol) was dissolved in 8 ml of dry THF and NaH 60% p.p. (170 mg, 4.25 mmol) and then benzylbromide (3.37 ml, 22.8 mmol) were sequentially added. The reaction mixture was stirred at room temperature for 12 h, then CH_3OH (1 ml) was added and the solvent removed *in vacuo*. The crude product was dissolved in $CHCl_3$, filtered on celite and then purified by chromatography on a silica gel column with AcOEt/CH_3OH (9:1, v/v) as eluent, yielding the desired compound **12** in 65% yields (1.7 g, 4.6 mmol).

12: oil**:** R_f = 0.5 (AcOEt/CH_3OH, 9:1, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.34-7.33 (overlapped signals, 5H, aromatic protons); 4.56 (s, 2H, -C$\underline{H}_2$-Ph); 3.66-3.60 [overlapped signals, 24H, 6x (-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 138.2, 128.3, 127.7 and 127.6 (aromatic carbons); 73.2 (-$\underline{C}H_2$-Ph); 72.5, 70.6, 70.2 and 69.4 [5x (-O-$\underline{C}H_2$-$\underline{C}H_2$-O-) and (-O-$\underline{C}H_2$-CH_2-OH)]; 61.7 (O-CH_2-$\underline{C}H_2$-OH).

ESI-MS (positive ions): calculated for $C_{19}H_{32}O_7$, 372.2; found *m/z*: 373.2 ($M+H^+$); 395.1 ($M+Na^+$).

Synthesis of *tert*-butyl (monobenzyloxy)hexaethylene glycol acetate (13).

Alcohol **12** (155 mg, 0.42 mmol) was dissolved in 1 ml of dry THF and NaH 60% p.p. (33 mg, 0.83 mmol) and then *tert*-butyl bromoacetate (154 μL, 1.0 mmol) were sequentially added. The reaction mixture was stirred at room temperature for 12 h, then few drops of CH_3OH were added and the solvent removed *in vacuo*. The crude product was dissolved in $CHCl_3$, filtered on celite and then purified by chromatography on a silica gel column eluted with AcOEt, yielding the desired compound **13** in 86% yields (175 mg, 0.36 mmol).

13: oil**:** R_f = 0.6 (AcOEt/CH_3OH, 95:5, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.34-7.32 (overlapped signals, 5H, aromatic protons); 4.56 [s, 2H, -C$\underline{H}_2$-Ph]; 4.01 (s, 2H, -C$\underline{H}_2$COO-*tert*-butyl); 3.69-3.64 [overlapped signals, 24H, 6x (-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)]; 1.46 [s, 9H, 3x (C$\underline{H}_3$ *tert*-butyl)].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 169.6 (C=O); 138.3, 128.3, 127.7 and 127.5 (aromatic carbons of Bn); 81.4 (quaternary carbon of *tert*-butyl); 73.2 (-$\underline{C}H_2$-Ph); 70.6 and 69.4 [6x (-O-$\underline{C}H_2$-$\underline{C}H_2$-O-)]; 69.0 (-$\underline{C}H_2$COO-*t*-Bu); 28.1 [3x (C$\underline{H}_3$ *tert*-butyl)].

ESI-MS (positive ions): calculated for $C_{25}H_{42}O_9$, 486.3; found *m/z*: 509.2 ($M+Na^+$); 525.2 ($M+K^+$).

Synthesis of (monobenzyloxy)hexaethylene glycol acetic acid (14).

Ester **13** (95 mg, 0.19 mmol) was dissolved in 1.5 ml of HCOOH and stirred at room temperature for 2 h. The solvent was then removed *in vacuo* and the residue was coevaporated three times with $CHCl_3$ (3x 3 ml), yielding the desired compound **14** in almost quantitative yields (82 mg, 0.19 mmol).

14: oil**:** R_f = 0.2 (AcOEt/CH_3OH, 95:5, v/v).

^{1}H NMR (CD_3OD, 200 MHz): δ 7.36-7.33 (overlapped signals, 5H, aromatic protons); 4.55 [s, 2H, -C$\underline{H}_2$-Ph]; 4.12 (s, 2H, -C$\underline{H}_2$COOH); 3.66-3.62 [overlapped signals, 24H, 6x (-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)].

^{13}C NMR (CD_3OD, 50 MHz): δ 174.0 (C=O); 139.7, 129.4, 129.0 and 128.7 (aromatic carbons); 74.1 (-$\underline{C}H_2$-Ph); 71.7, 71.6 and 70.6 [6x (-O-$\underline{C}H_2$-$\underline{C}H_2$-O-)]; 69.1 (-$\underline{C}H_2$COOH).

ESI-MS (positive ions): calculated for $C_{21}H_{34}O_9$, 430.2; found *m/z*: 431.1 ($M+H^+$); 453.1 ($M+Na^+$); 459.1 ($M+K^+$).

Synthesis of 3-(4-pyridylmethyl), 3'-*O*-cholesteroxyacetyl, 5'-*O*-(4,4'-dimethoxy-triphenylmethyl)thymidine (**15**)

5'-*O*-(4,4'-dimethoxytriphenylmethyl)thymidine (**2**) (396 mg, 0.73 mmol) was dissolved in 5 mL of dry DMF. K_2CO_3 (302 mg, 2.18 mmol) and 4-(bromomethyl)pyridine hydrobromide (276 mg, 1.09 mmol) were sequentially added to the reaction mixture, left under stirring at 60 °C. After 12 h, TLC analysis indicated the presence of one main product in the reaction mixture. Alkylated compound **3** was not isolated, but directly converted into its 3'-conjugated derivative **15**. To this purpose, the solvent was removed and the residual solid was taken up in DCM, washed with water, dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. Successively the residue was dissolved in 3 mL of dry DCM and then DMAP (32 mg, 0.266 mmol), cholesteroxy-acetic acid **18** (118 mg, 0.266 mmol) and DCC (55 mg, 0.266 mmol) were sequentially added. The reaction mixture was stirred for 12 h at room temperature and then the solvent was removed *in vacuo*; the

crude product was purified by chromatography on a silica gel column with *n*-hexane/AcOEt (1:1, v/v, containing 1% of TEA) as eluent, yielding the desired compound **15** in 19% yields for the two steps (144 mg, 0.14 mmol).

15: oil: R_f= 0.3 (AcOEt/*n*-hexane 1:1, v/v).

^{1}H-NMR ($CDCl_3$, 500 MHz): significant signals at δ 8.57 (d, *J* = 4.5 Hz, 2H, 2x H_αPy); 7.66 (s, 1H, H-6); 7.49 (d, *J* = 4.5 Hz, 2H, 2x H_βPy); 7.36-6.82 (overlapped signals, 13H, aromatic protons of *DMT*); 6.44 (dd, *J* = 6.0 and 5.5 Hz, 1H, H-1'); 5.53 (d, *J* = 5.0 Hz, 1H, H-3'); 5.36 (m, 1H, H-6 Chol); 5.15 (s, 2H, -CH_2-Py); 4.13 (overlapped signals, 3H, -CH_2C=O and H-4'); 3.79 [s, 6H, 2x (-OCH_3) of *DMT*]; 3.47 (m, 2H, H-5'); 3.23 (m, 1H, H-3 Chol); 2.48 (m, 2H, H-2'); 1.42 (s, 3H, CH_3-Thy); 0.99 (s, 3H, CH_3-19 Chol); 0.91 (d, *J* = 6.5 Hz, 3H, CH_3-21 Chol); 0.86 (coincident d's, *J* = 4.5 Hz, 6H, CH_3-26 and CH_3-27 Chol); 0.67 (s, 3H, CH_3-18 Chol).

^{13}C-NMR ($CDCl_3$, 50 MHz): significant signal δ 170.3 (C=O); 163.0 (C-4); 158.7, 144.1, 133.8, 130.0, 128.0, 127.2 and 113.3 (aromatic carbons of *DMT*); 150.8 (C-2); 149.9 (2x C_αPy); 145.4 (C_γPy); 140.2 (C-5 Chol); 135.1 (C-6); 123.5 (2x C_βPy); 122.1 (C-6 Chol) 110.8 (C-5); 87.0 (quaternary carbon of *DMT*); 85.0 (C-1'); 83.9 (C-4'); 80.1 (C-3 Chol); 75.6 (C-3'); 65.4 (-CH_2C=O); 63.5 (C-5'); 55.2 [2x (-OCH_3) of *DMT]*; 43.5 (-CH_2Py); 38.6 (C-2'); 22.7 and 22.5 (C-27 and C-26 Chol); 19.2 (C-19 Chol); 18.6 (C-21 Chol); 12.3 (CH_3-Thy); 11.8 (C-18 Chol).

ESI-MS (positive ions): calculated for $C_{66}H_{83}N_3O_9$: 1061.61 found *m/z*: 1062.93 ($M+H^+$); 1101.02 ($M+K^+$).

Synthesis of 3-(4-pyridylmethyl), 3'-*O*-cholesteroxy-acetyl thymidine (**16**)

15 (140 mg, 0.13 mmol) was dissolved in 3 mL of a 1% TCA solution in DCM and stirred for 1 h at room temperature. The reaction mixture was then concentrated *in vacuo* and purified by chromatography on a silica gel column with $CHCl_3$/CH_3OH (99:1, v/v) as eluent, giving the desired compound **16** in almost quantitative yields (100 mg, 0.13 mmol).

16: oil**:** R_f= 0.4 (AcOEt/*n*-hexane 7:3, v/v).

^{1}H-NMR ($CDCl_3$, 500 MHz): significant signals at δ 8.55 (d, *J* = 4.5 Hz, 2H, 2x H_αPy); 7.55 (s, 1H, H-6); 7.36 (d, *J* = 4.5 Hz, 2H, 2x H_βPy); 6.27 (dd, *J* = 6.5 and 7.5 Hz, 1H, H-1'); 5.45 (m, 1H, H-3'); 5.36 (m, 1H, H-6 Chol); 5.12 (s, 2H, -CH_2-Py); 4.15 (s, 2H, -CH_2C=O); 4.11 (bs, 1H,

H-4'); 3.94 (bs, 2H, H-5'); 3.24 (m, 1H, H-3 Chol); 2.42 (m, 2H, H-2'); 1.95 (s, 3H, C<u>H</u>$_3$-Thy); 1.00 (s, 3H, H-19); 0.91 (d, *J* = 6.5 Hz, 3H, H-21 Chol); 0.86 (concident d's, *J* = 4.5 Hz, 6H, H-26 and H-27 Chol); 0.67 (s, 3H, H-18 Chol).

^{13}C-NMR ($CDCl_3$, 50 MHz): significant signals δ 170.6 (C=O); 163.0 (C-4); 150.8 (C-2); 149.2 (2x C$_\alpha$Py); 146.2 (C$_\gamma$Py); 140.2 (C-5 Chol); 134.6 (C-6); 123.6 (2x C$_\beta$Py); 122.1 (C-6 Chol) 110.4 (C-5); 86.1 (C-1'); 85.1 (C-4'); 80.2 (C-3 Chol); 75.3 (C-3'); 65.3 (-C<u>H</u>$_2$C=O); 62.1 (C-5'); 43.4 (-C<u>H</u>$_2$Py); 38.6 (C-2'); 22.7 and 22.5 (C-27 and C-26 Chol); 19.2 (C-19 Chol); 18.6 (C-21 Chol); 13.1 (C<u>H</u>$_3$-Thy); 11.8 (C-18 Chol).

ESI-MS (positive ions): calculated for $C_{45}H_{65}N_3O_7$: 759.48; found *m/z*: 760.88 ($M+H^+$); 799.08 ($M+K^+$).

Synthesis of 3-(4-pyridylmethyl), 3'-*O*-cholesteroxy-acetyl, 5'-*O*-monomethoxytri-ethylene glycol acetyl thymidine (1d, ToThyChol).

16 (96 mg, 0.13 mmol) was dissolved in 2 mL of dry DCM and then DMAP (46 mg, 0.38 mmol), carboxylic acid **14** (41 mg, 0.18 mmol) and DCC (38 mg, 0.18 mmol) were sequentially added. The reaction mixture was stirred for 12 h at room temperature and the solvent was then removed *in vacuo*; the crude product was purified by chromatography on a silica gel column eluted with *n*-hexane/acetone (1:1, v/v), giving the desired compound **1d** in 85% yields (106 mg, 0.11 mmol).

6: oil**:** R_f= 0.3 (AcOEt).

^{1}H-NMR ($CDCl_3$,500 MHz): significant signals at δ 8.55 (bs, 2H, 2x H$_\alpha$Py); 7.36 (bs, 3H, H-6 and 2x H$_\beta$Py); 6.36 (dd, *J* = 5.5 and 9.0 Hz, 1H, H-1'); 5.37 (bs, 1H, H-6 Chol), 5.30 (apparent d, *J* = 5.0 Hz, 1H, H-3'); 5.12 (s, 2H, -C<u>H</u>$_2$-Py); 4.43 (AB part of an ABX system, *J* = 12.0 and 3.5 Hz, 2H, H-5'$_a$ and H-5'$_b$); 4.24 (s, 2H, -C<u>H</u>$_2$C=O TEG); 4.16 (overlapped signals, 3H, -C<u>H</u>$_2$C=O and H-4'); 3.76-3.50 (overlapped signals, 12H, 3x –OC<u>H</u>$_2$-C<u>H</u>$_2$O-); 3.36 (s, 3H, -OC<u>H</u>$_3$); 3.24 (m, 1H, H-3 Chol); 2.42 (m, 2H, H-2'); 1.96 (s, 3H, C<u>H</u>$_3$-Thy); 1.00 (s, 3H, C<u>H</u>$_3$-19); 0.91 (d, *J* = 6.5 Hz, 3H, C<u>H</u>$_3$-21 Chol); 0.86 (coincident d's, *J* = 4.5 Hz, 6H, C<u>H</u>$_3$-26 and C<u>H</u>$_3$-27 Chol); 0.67 (s, 3H, C<u>H</u>$_3$-18 Chol).

^{13}C-NMR ($CDCl_3$, 125 MHz): significant signals δ 170.6 (C=O Chol ester); 169.7 (C=O TEG ester); 162.8 (C-4); 150.7 (C-2); 149.0 (2x C$_\alpha$Py); 146.0 (C$_\gamma$Py); 140.1 (C-5 Chol); 133.5 (C-6); 123.7 (2x C$_\beta$ Py); 122.1 (C-6 Chol); 110.8 (C-5); 85.4 (C-1'); 81.9 (C-4'); 80.2 (C-3 Chol);

74.8 (C-3'); 71.8, 71.0, 70.7 and 70.4 (carbons of TEG); 68.7 (-$\underline{C}H_2C$=O TEG); 65.3 (-$\underline{C}H_2C$=O Chol); 64.0 (C-5'); 58.9 (-O$\underline{C}H_3$); 43.5 (-$\underline{C}H_2$Py); 38.6 (C-2'); 22.7 and 22.5 (C-27 and C-26 Chol); 19.2 (C-19 Chol); 18.6 (C-21 Chol); 13.1 ($\underline{C}H_3$-Thy); 11.8 (C-18 Chol).

ESI-MS (positive ions): calculated for $C_{54}H_{81}N_3O_{12}$ 963.58; found m/z: 963.97 ($M+H^+$); 985.97 ($M+Na^+$); 1002.1 ($M+K^+$).

Synthesis of *tert*-butyl cholesteroxy-acetate (**17**)

Cholesterol (963 mg, 2.5 mmol) was dissolved in 4 mL of dry THF and NaH 60% p.p. (200 mg, 4.98 mmol) and then *tert*-butyl bromoacetate (920 μL, 6.22 mmol) were sequentially added to the reaction mixture taken at 0 °C. After 24 h at room temperature, the reaction was quenched by addition of CH_3OH (2 mL) and the solvent removed *in vacuo*. The residual solid was taken up in DCM, washed with water, dried over anhydrous Na_2SO_4 and concentrated *in vacuo*. The residue was then purified by chromatography on a silica gel column using *n*-hexane/AcOEt (9:1, v/v) as eluent, yielding the desired compound **17** in 31 % yields (392 mg, 0.77 mmol).

17: amorphous solid**:** R_f= 0.8 (AcOEt/*n*-hexane 1:3, v/v).

^{1}H-NMR ($CDCl_3$, 200 MHz): significant signals at δ 5.36 (d, *J* = 5.0 Hz, 1H, H-6); 4.20 (s, 2H, $C\underline{H}_2C$=O); 3.23 (m, 1H, H-3); 2.31 (m, 2H, H_2-4); 1.48 (s, 9H, $C\underline{H}_3$*t*-Bu); 1.02 (s, 3H, H-19); 0.91 (d, *J* = 6.5 Hz, 3H, H-21); 0.86 (coincident d's, *J* = 4.5 Hz, 6H, H-26 and H-27); 0.67 (s, 3H, H-18).

^{13}C-NMR ($CDCl_3$, 50 MHz): significant signals at δ 170.1 (C=O); 140.6 (C-5); 121.8 (C-6); 81.3 (C quaternary of *t*-Bu); 79.8 (C-3); 66.1 ($\underline{C}H_2C$=O); 28.1 (3x $\underline{C}H_3$*t*-Bu); 22.8 and 22.5 (C-27 and C-26); 19.3 (C-19); 18.7 (C-21); 11.8 (C-18).

ESI-MS (positive ions): calculated for $C_{33}H_{56}O_3$ 500.42; found *m/z*: 521.25 ($M+Na^+$); 537.45 ($M+K^+$).

Synthesis of cholesteroxy-acetic acid (**18**)

Ester **17** (300 mg, 0.59 mmol) was dissolved in 3.0 mL of formic acid and stirred at room temperature for 2 h. The solvent was then removed *in vacuo* and the residue coevaporated three times with $CHCl_3$ (3x 3 mL), yielding the desired compound **18** in almost quantitative yields (266 mg, 0.59 mmol).

8: oil**:** R_f= 0.2 (AcOEt/*n*-hexane 1:3, v/v).

^{1}H-NMR ($CDCl_3$,200 MHz): significant signals at δ 5.34 (d, *J* = 4.4 Hz, 1H, H-6); 4.17 (s, 2H, C$\underline{H_2}$C=O); 3.28 (m, 1H, H-3); 0.91 (d, *J* = 6.5 Hz, 3H, H-21); 0.86 (coincident d's, *J* = 4.5 Hz, 6H, H-26 and H-27); 0.68 (s, 3H, H-18).

^{13}C-NMR ($CDCl_3$, 50 MHz): significant signals at δ 175.3 (C=O); 140.1 (C-5); 122.2 (C-6); 80.3 (C-3); 65.1 ($\underline{C}H_2$C=O); 22.8 and 22.5 (C-26 and C-27); 19.3 (C-19); 18.7 (C-21); 11.8 (C-18).

ESI-MS (positive ions): calculated for $C_{29}H_{48}O_3$ 444.36; found *m/z*: 464.72 ($M+Na^+$); 480.88 ($M+K^+$).

5. References

1. J. Rautio, H. Kumpulainen, T.Heimbach, R. Oliyai *Nature*, **2008**, *7*, 255-270.
2. V. Stella in Optimizing the "Drug‐Like" Properties of Leads in Drug Discovery, 221–242 (Springer, New York, **2006**).
3. K. Beaumont, R. Webster, I. Gardner, K. Dack *Curr. Drug Metab.* **2003**, *4*, 461–485.
4. A comprehensive review on ester prodrugs for improved oral absorption is: P. Ettmayer, G. L. Amidon, B. Clement, B. Testa *J. Med. Chem.*, **2004**, *47*, 2393–2404.
5. T. Jarvinen, J. Rautio, M. Masson, T. Loftsson in Drug Discovery Handbook (ed. Gad, S.) 733–79 (John Wiley & Sons, Hoboken, **2005**).
6. V. Stella, Prodrugs as therapeutics. *Expert Opin. Ther. Pat.*, **2004**, *14*, 277–280.
7. B. Testa *Biochem. Pharmacol.*, **2004**, *68*, 2097–2106.
8. V. Stella, Prodrugs: Challenges and Rewards (AAPS Press/Springer, New York, **2007**).
9. V. Stella, K. W. Nti-Addae *Adv. Drug Deliv. Rev.* **2007**, *59*, 677–694.
10. P. K. Kiptoo, M. O. Hamad, P.A. Crooks, A. L. Stinchcomb *J. Control. Release* **2006**, *113*, 137–145.
11. J. Leppanen, J. Huuskonen, T. Nevalainen, J. Gynther, H. Taipale, T. Järvinen *J. Med. Chem.* **2002**, *45*, 1379–1382.
12. R. E. Notari *Pharmacol. Ther.* **1981**, *14*, 25–53.
13. A. A. Sinkula, S. H. Yalkowsky *J. Pharm. Sci.*, **1975**, *64*, 181–210.
14. V. Stella, W. N. Charman, V. H. Naringrekar *Drugs*, **1985**, *29*, 455–473.
15. N. Bodor, P. Buchwald, *Pharmacol. Ther*. **1997** ,*76*, 1–27.
16. N. Bodor, P. Buchwald *Med. Res. Rev.* **2000**, *20*, 58–101.
17. T. Thorsteinsson, T. Loftsson, M. Masson *Curr. Med. Chem.* **2003** ,*10*, 1129–1136.
18. P. Baglioni, D. Berti *Current Opinion in Colloid & Interface Science*, **2003**, *8*, 55-61.
19. M. Banchelli, D. Berti, P. Baglioni, *Angew. Chem. Int. Ed.*, **2007**, *46*, 3070-3073.
20. M. Fortini, D. Berti, P. Baglioni, B. W. Ninham *Current Opinion in Colloid & Interface Science* **2004**, *9*, 168-172.
21. S. Milani, D. Berti, S. Dante, T. Hauss, P. Baglioni *Langmuir*, **2009**, *25*, 4084-4092.
22. A. Gissot, M. Camplo, M. W. Grinstaff, P. Barthélémy *Org. Biomol. Chem.*, **2008**, *6*, 1324-1333.

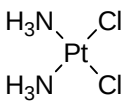
23. M. Vaccaro, R. Del Litto, G. Mangiapia, A. M. Carnerup, G. D'Errico, F. Ruffo, L. Paduano *Chem. Commun.*, **2009**, 1404-1406.
24. L. Simeone, G. Mangiapia, C. Irace, A. Di Pascale, A. Colonna, O. Ortona, L. De Napoli, D. Montesarchio, L. Paduano *Mol. BioSyst.*, **2011**, *7*, 3075–3086.
25. L. Simeone, G. Mangiapia, G. Vitiello, C. Irace, A. Di Pascale, A. Colonna, O. Ortona, L. De Napoli, D. Montesarchio, L. Paduano *Bioconjugate Chemistry, submitted.*
26. A. Accardo, D. Tesauro, G. Morelli, E. Gianolio, S. Aime, M. Vaccaro, G. Mangiapia, L. Paduano K. Schillen *JBIC, Journal of Biological Inorganic Chemistry*, **2007**, *12*, 267-276.
27. H. Rosemeyer *Chemistry & Biodiversity*, **2005**, *2*, 977-1062.
28. A. M. Heacock, B. W. Agranoff *Biochimica et Biophysica Acta, Lipids and Lipid Metabolism*, **1997**, *1348*, 166-172.
29. C. Hiss Donavon, A. G. Gabriels, I. P. Folb, *Cancer Cell International*, **2007**, *7*, 5.
30. F. H. Van Tiel, W. A. M. Boere, T. Harmsen, C. A. Kraaijeveld, H. Snippe *Antimicrobial Agents and Chemotherapy*, **1985**, *27*, 802-805.
31. M. B. Hansen, S. E. Nielsen, K. Berg, *Journal of Immunological Methods*, **1989**, *119*, 203-210.
32. R. A. Schwendener, A. Supersaxo, W. Rubas, H. G. Weder, H. R. Hartmann, H. Schott, A. Ziegler, H. Hengartner, *Biochem. Biophys. Res. Comm.*, **1985**, *126*, 660-666.
33. C. J. Welch, A. Larsson, A. C. Ericson, B. Oeberg, R. Datema, J. Chattopadhyaya, *Acta Chemica Scandinavica, Series B: Organic Chemistry and Biochemistry*, **1985**, *B39*, 47-54.
34. P. Chabaud, M. Camplo, D. Payet, G. Serin, L. Moreau, P. Barthélémy, M. W. Grinstaff, *Bioconjugate Chem.*, **2006**, *17*, 466-472.
35. L. Moreau, P. Barthélémy, Y. Li, D. Luo, C. A. H. Prata, M. W. Grinstaff, *Mol. BioSyst.*, **2005**, *1*, 260-264.
36. V. V. Ranade, M. A. Hollinger *Drug Delivery Systems.*, CRC, Boca Raton, **1996**.

Chapter 3.

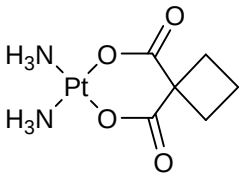
Synthesis and Characterization of Novel Amphiphilic Nucleolipid Complexes of Ruthenium (III) as Potential Anticancer Agents

1. Introduction: Ruthenium(III) complexes as anticancer drugs

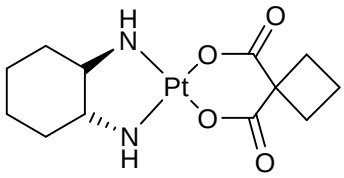
In the last decades, research has proposed a growing number of antineoplastic agents. Among these, transition metal-based complexes represent, nowadays, a very important class of chemotherapeutics, intensively used for clinical treatments. The anticancer activity of the most important of these, Cisplatin, was discovered in 1969 [1] and in thirty years this complex became one of the most used drugs in the treatment of some tumoral diseases, such as testicular, breast, uterine and ovarian cancers. Ever since, a large variety of Cisplatin analogues have been tested and some of them have been approved as drugs, such as Carboplatin and Oxaliplatin. [2]



Cisplatin



Carboplatin



Oxaliplatin

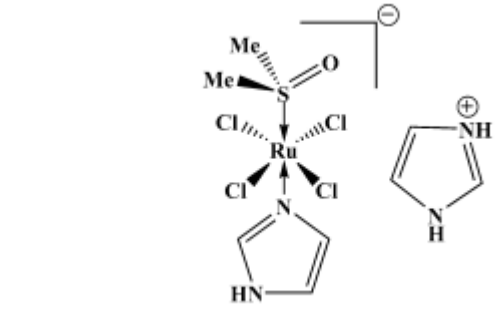
Although these heavy metal agents are active against a variety of cancers, their use is associated with severe side effects including gastrointestinal symptoms (nausea, vomiting, diarrhea, abdominal pain), renal tubular injury, neuromuscular complications, and ototoxicity. In addition, the use of platinum is limited in many tumor types by primary and acquired resistance to this agent. [3] This has led to an ongoing quest for the

discovery of non-platinum metals that may extend the spectrum of activity of metal-based drugs [4].

A possible improvement of the antitumoral drug properties is represented by the efficacy toward the formation and growth of metastases. In fact, many tumors can develop metastases often very extended at the diagnosis time, making scarcely effective the surgical treatment. [5] If the primary tumor can be surgically removed, the pharmacological therapy seems to be the best choice for the metastases treatment, because of their wide delocalization. [6] On the other hand, unlike the primary tumor cells, metastases are not responsive to several chemotherapeutics, probably because of their different proliferation kinetics. [7,8] This is the reason why many widely employed drugs show only limited effects on metastatic disease attack. These limitations have prompted the research to develop more effective and less toxic metal-based antitumor agents.

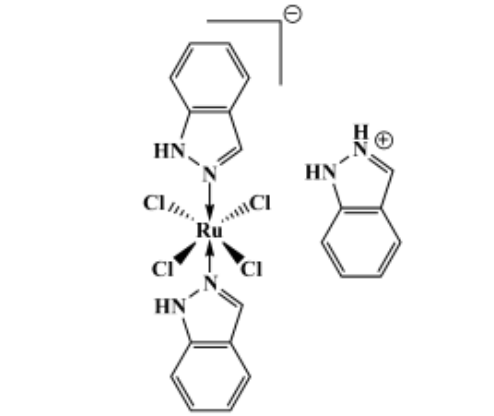
In this perspective, ruthenium complexes have attracted much interest as a promising alternative to platinum, showing a remarkable anti-tumoral and anti-metastatic activity, together with lower toxicity. Ruthenium, a transition metal of the platinum group (group 8B of the Periodic Table), was first hypothesized to exert its anticancer effects by direct interaction with DNA as observed with platinum. However, it is now known that ruthenium shows a number of differences with respect to platinum. [9] First, ruthenium appears to accumulate preferentially in neoplastic masses rather than normal tissues possibly by using transferrin to enter into tumors. [10] It has been proposed that transferrin–ruthenium complexes are actively transported into neoplastic tissues containing high transferrin receptor densities. Once bound to the transferrin receptor, it is assumed that the complex liberates ruthenium that is then internalized by the tumor. [11] Secondly, ruthenium remains in its relatively inactive Ru(III) oxidation state until it reaches the tumor site. In this environment, where there is a lower oxygen content and higher acidity compared to normal tissues, reduction to the more reactive Ru(II) occurs. [12] This reaction, termed *"activation-by-reduction"*, not only results in selective tumor targeting but may also direct cytotoxic activity toward hypoxic tumors that are more likely to be resistant to chemotherapy and radiations. [13]

The most promising and remarkable Ru-based complexes have been described by Sava and Alessio. [14-20] Since the beginning of 90's, Sava and Alessio have been pioneers in studying the perspectives of transition-metal complexes as potential anticancer agents, in particular discovering a very active Ru complex, NAMI-A. Such compound, together with KP1019, developed by Keppler *et al.*, recently passed Phase I clinical trials with good outcome. [21]



Imidazolium-*trans*-dimethylsulfoxide-imidazole-tetrachlororuthenate

NAMI-A



Indazolium-*trans*-bis(1*H*-indazole)-tetrachlororuthenate

KP1019

2. Design of novel Ruthenium(III) complexes

A large number of studies on Ru complexes have been published in recent years; in all cases, scarce stability in aqueous media of these compounds, undergoing replacement of the chloride ligands with hydroxide ions and/or water molecules and further consequent degradation processes, has been observed. To overcome these drawbacks, an innovative approach to vehiculate in cells transition metal-containing drugs for anticancer therapy has been very recently proposed in the literature by Paduano and Ruffo, based on the use of amphiphilic nanovectors carrying Ruthenium complexes. [22,23] As a continuation of this work, the synthesis and characterization of a mini-library of novel Ru-containing nucleolipids is here described.

As widely known, amphiphilic self-assembly allows an efficient bottom-up strategy in order to obtain nanosized aggregates whose size and shape are quite easily tunable. Further benefits related to the use of nanostructures concern:

i) the possibility to transport a larger quantity of the metal inside the blood stream;
ii) the possibility to obtain “stealth” aggregates, which are not detected by the human immune system, thus specifically increasing their residence time in the blood;
iii) the possibility to obtain aggregates selectively captured by specific tumoral cells lines by inserting within ruthenium aggregates some "marker" molecules, able to be recognized by protein receptors over-expressed by the tumoral cells;
iv) the possibility to tune the shape and size of the aggregates by acting either on their molecular structure or on physico-chemical parameters, such as pH and ionic strength. These effects may be very useful in a stimuli-responsive scenario.

With these aims in mind, we have used the functionalized nucleolipids ToThy, HoThy, DoHu and ToThyChol (the preparation of which has been discussed in Chapter 2) to synthesize and characterize a series of amphiphilic Ru complexes, designed as spontaneously self-aggregating systems in structures such as micelles, liposomes or even more complex structures.

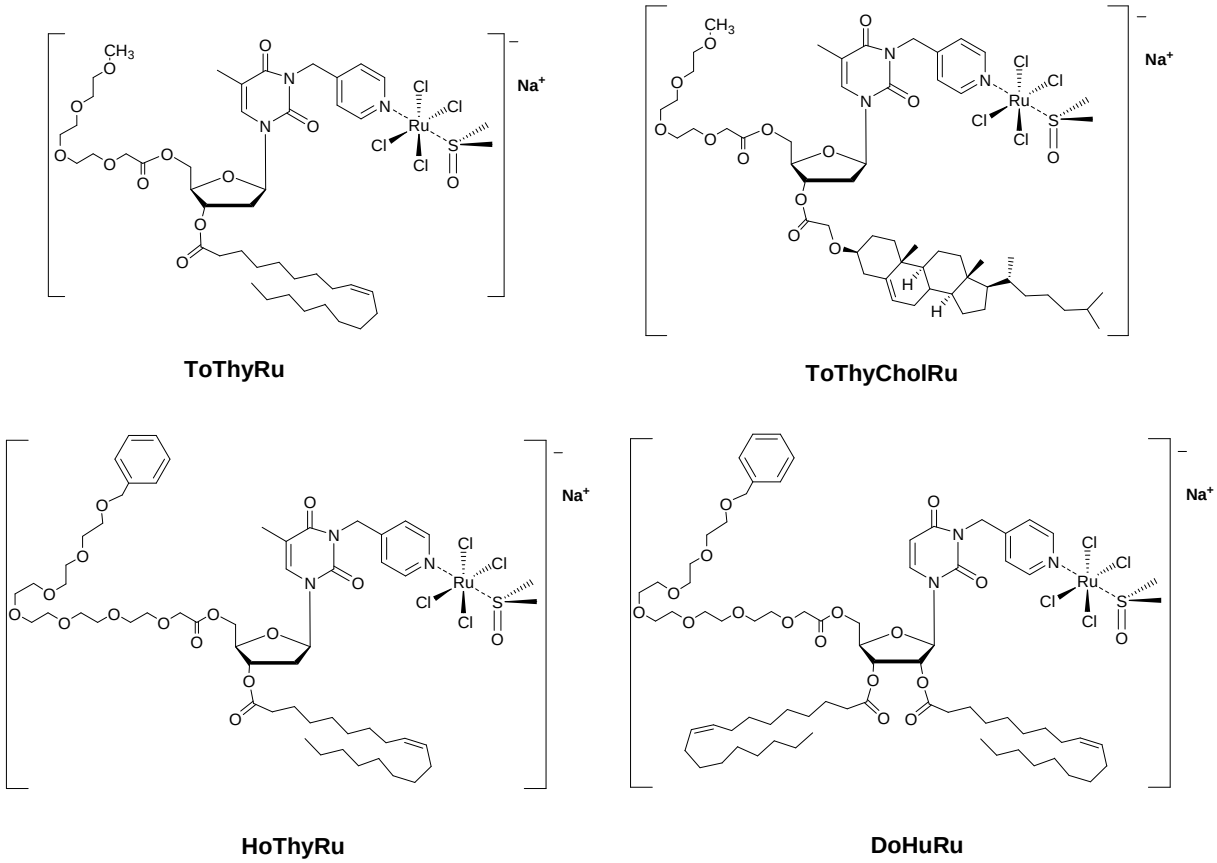


Figure 1. Functional nucleolipids: ToThyRu, HoThyRu, DoHuRu and ToThyCholRu.

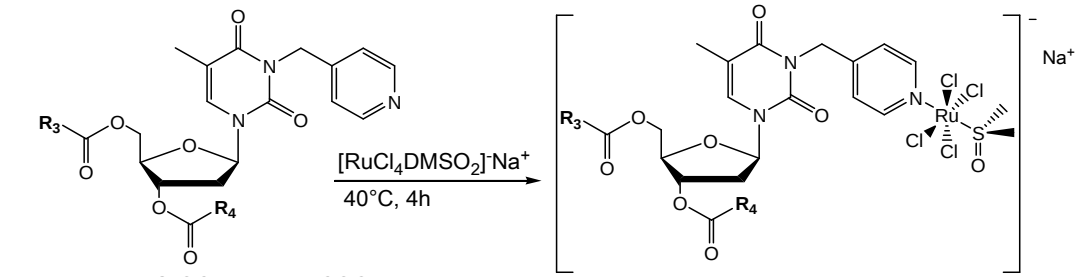
The functional nucleolipids reported in Figure 1 have been built starting from a pyrimidinic 2'-deoxyribonucleoside (Thymidine, in the case of ToThyRu, HoThyRu and ToThyCholRu) or ribonucleoside (Uridine, for DoHuRu). The nucleoside has been chosen as the poly-functional main scaffold, allowing high diversification. In fact one or two oleic acid chains have been inserted onto the nucleoside, as well as a poly(ethylene oxide) chain able to make the aggregates more resistant toward the enzymatic degradation. Finally, a residue of pyridine - as a chelating moiety able to complex Ru(III) ions - has been attached together with the metallic species.

The use of hybrid molecules carrying a double functionality, *i.e.* nucleosides and lipid chains, has gained a considerable interest because of the capability of these molecules to mimic the molecular organizations of the biological systems, as well as for the possibility to form a wide variety of supramolecular systems such as liposomes/vesicles, cubic phases, ribbons, etc. [24-26]

Furthermore, the use of formulations of Ru complexes with phospholipids allows to modulate the metal amount to be administered, as well as to preserve nanoaggregates from degradation. In fact, double chained phospholipids in aqueous solutions adopt an ordered arrangement by forming double layers in which Ru complexes are easily lodged. Among phospholipids, palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine, POPC, is particularly interesting being one of the main components of cell membranes.

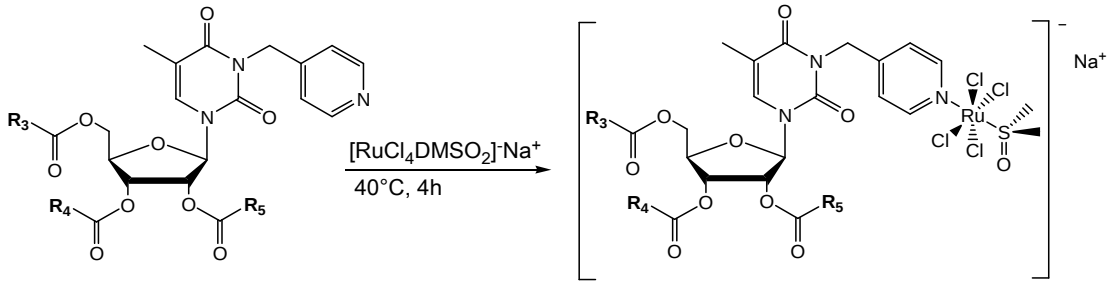
2.1 Synthesis of Ruthenium complexes ToThyRu, HoThyRu, DoHuRu and ToThyCholRu

The synthesis of ruthenium complexes ToThyRu, HoThyRu, DoHuRu and ToThyCholRu was accomplished in one step starting from the corresponding nucleosidic nanovectors ToThy, HoThy, DoHu and ToThyChol described in Chapter 2. The reaction was carried out by mixing the nucleolipid and the ruthenium salt $[RuCl_4(DMSO)_2]^-Na^+$, prepared following literature procedures, [27] in the same molar ratio in the appropriate solvent (anhydrous CH_2Cl_2 or CH_3CN) and heated at 40 °C for 4 h, as reported in Scheme 1. The reactions, monitored by TLC on alumina, showed in 4 h the total disappearance of the starting nucleosides. In all cases, the desired products could be isolated in quantitative yields simply by removing the solvent under reduced pressure, thus not requiring further purification. [28,29]



ToThy R_3 = TEG-OCH_3 R_4COO = oleyl residue
HoThy R_3 = HeptaEG-OBn R_4COO = oleyl residue
ToThyChol R_3 = TEG-OCH_3 R_4 = Cholesteryl-3-OCH_2

ToThyRu R_3 = TEG-OCH_3 R_4COO = oleyl residue
HoThyRu R_3 = HeptaEG-OBn R_4COO = oleyl residue
ToThyCholRu R_3 = TEG-OCH_3 R_4 = Cholesteryl-3-OCH_2



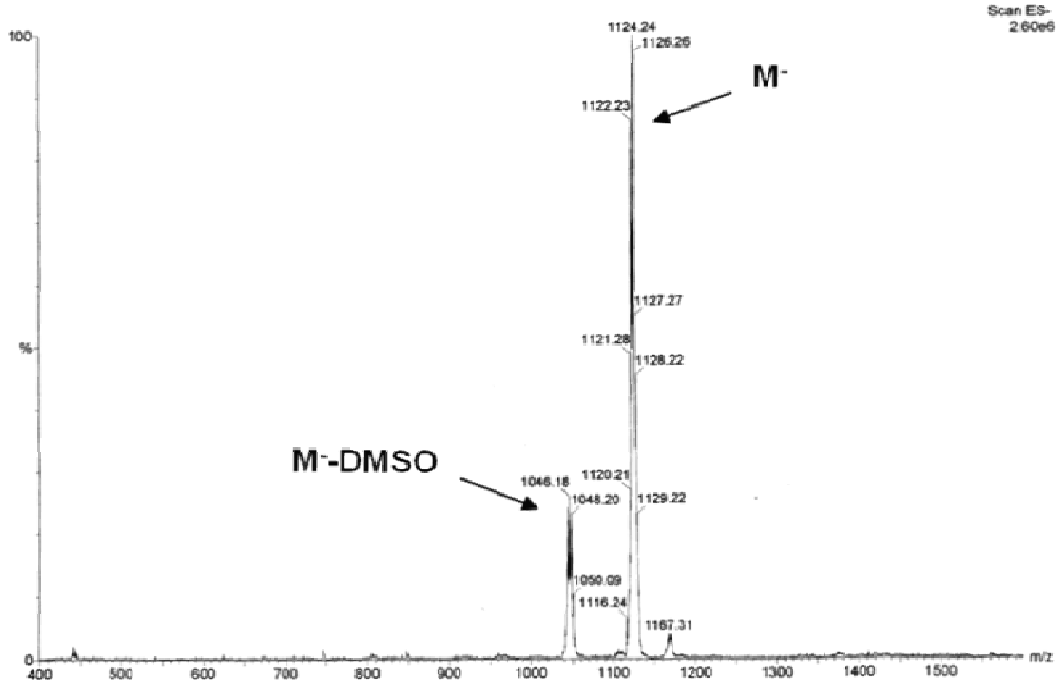
DoHu R_3= HeptaEG-OBn R_4 and R_5COO = oleyl residue

DoHuRu R_3= HeptaEG-OBn R_4 and R_5COO = oleyl residue

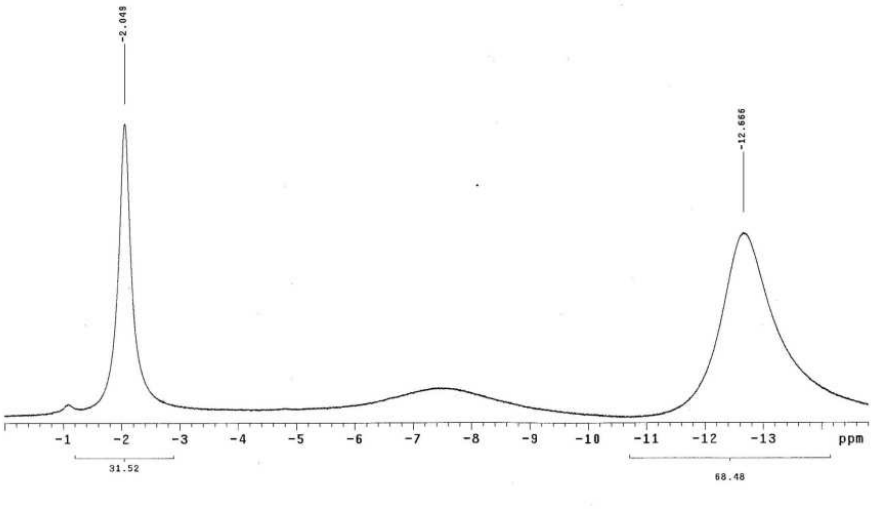
Solvent reaction: CH_3CN for ToThyRu and ToThyCholRu
CH_2Cl_2 for DoHuRu and HoThyRu

Scheme 1. Synthesis of the ruthenium complexes ToThyRu, HoThyRu, DoHuRu and ToThyCholRu.

The obtained ruthenium complexes were characterized by NMR and ESI-MS. The ^{1}H and ^{13}C-NMR spectra showed broadened signals, due to the presence of the paramagnetic ruthenium(III) nucleus. In the ^{1}H-NMR spectrum - registered in $CDCl_3$ - of the salts, this also induced a dramatic upfield shift of the pyridine protons and of the methyl protons of the DMSO ligand, respectively here found at δ = - 2.0 ppm, and δ = - 12.00 ppm *ca.*, as very broadened signals, which can be considered diagnostic of the effective complex formation. [30,31] The ESI-MS spectra were recorded in negative ions mode, showing m/z signals only for the expected molecular ions and thus confirming the identity of the target structures. The ^{1}H-NMR and ESI-MS spectra for ToThyRu are reproduced in Figure 2.



a



b

Figure 2. a) ESI-MS (negative ions) spectrum of ToThyRu; **b**) Upfield expansion for the ^{1}H NMR spectrum of ToThyRu (500 MHz, $CDCl_3$)

2.2 Characterization of ToThyRu, HoThyRu, DoHuRu, and ToThyCholRu

In analogy to what previously mentioned for NAMI-A, also the here synthesized Ru complexes, ToThyRu, HoThyRu, DoHuRu and ToThyCholRu, have showed limited stability when left in prolonged contact with aq. solutions. In fact, depending on the physico-chemical conditions, a clear change in the properties of the Ru-complex is observed, with formation of insoluble green precipitates in a period going from few hours, in phosphate buffers, to few days in water (see Figure 3). [28, 29]

Figure 3. a. ToThyCholRu in H_2O (stable for 72 h); **b.** ToThyCholRu in a pseudo-physiological solution (stable for 4 h).

The degradation process is imputable to the replacement of the ligands of the ruthenium complex by water molecules and/or hydroxide ions, well known for NAMI-A and related analogs. [32] Also in the present case Ru complexes undergo this process, as confirmed by the changes of the UV/Vis spectra with time. In Figure 4 the time evolution of the UV/Vis spectra for ToThyCholRu in H_2O and pseudo-physiological conditions is reported.

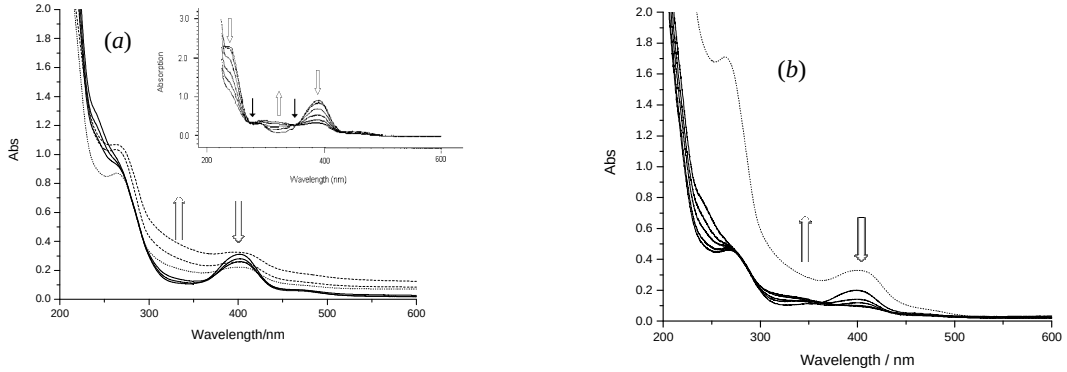


Figure 4. Absorbance evolution with time of 0.1 mmolkg^{-1} ToThyCholRu solution in: (*a*) water (72 h), (*b*) pseudo-physiological conditions (4 h). Two isosbestic points are present at 274 and 358 nm, respectively. The dashed lines refer to the formation of dark particles in solution and the arrows indicate the direction of evolution with time. The inset represents the absorbance evolution with time of NAMI-A solution. [32]

As expected, the degradation phenomena are more marked at alkaline pH. To sensibly reduce the aquation processes of the Ru complexes, as well as to increase the biocompatibility of the formulations and modulate the amount of Ruthenium carried, the synthesized amphiphilic molecules were let to co-aggregate with amounts of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC). This commercially available phospholipid - as previously mentioned, widely represented in eukaryotic cell membranes - was here chosen as the stabilizing system for these Ru-complexes in view of its well ascertained biocompatibility. The resulting co-aggregate systems were then studied in pure water and in a buffer containing NaCl and K_2HPO_4 at physiological ionic strength (here cited as "pseudo-physiological conditions"). As a matter of fact, the presence of POPC increases the aggregates stability, since no precipitation due to poly-oxo species formation is observed. Nevertheless, in pseudo-physiological conditions the stability of the Ru-complex aggregates in mixture with POPC is guaranteed for several days only if the Ru amount does not exceed 15% in molar fraction inside the aggregates (see Figure 5).

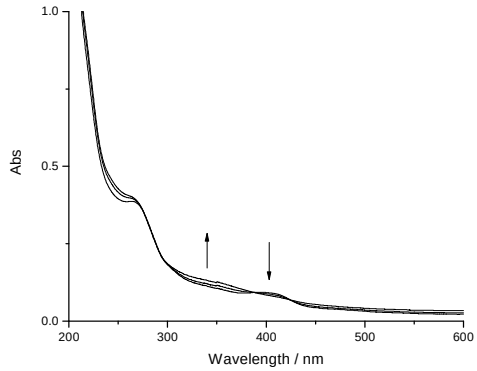


Figure 5. UV-VIS time evolution of the ToThyCholRu-POPC 15:85 sample at 0.3 mmolkg^{-1} total concentration in pseudo-physiological conditions over a period of 150 h. The arrows indicate the sense of the spectra evolution with time.

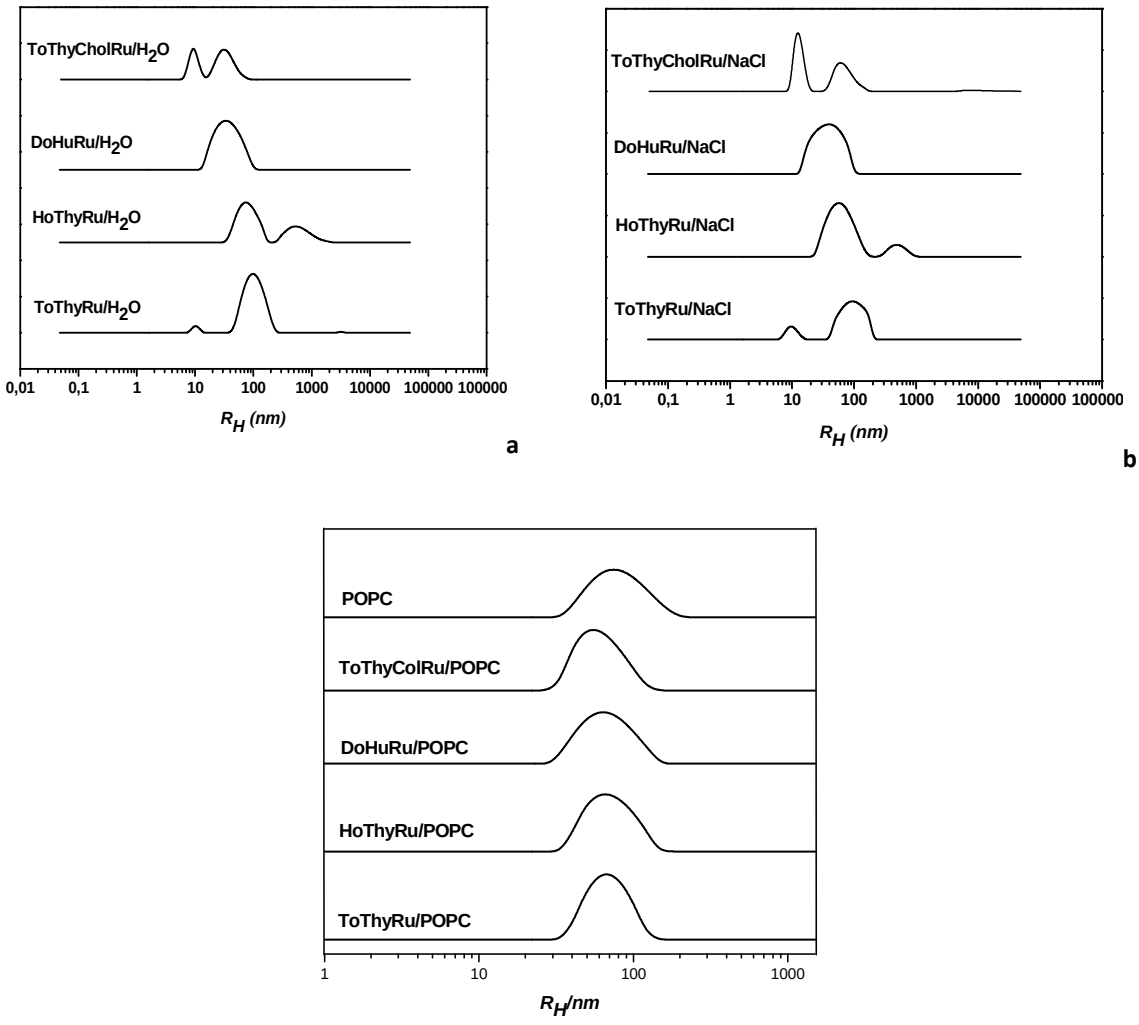


Figure 6 Example of the hydrodynamic radii distribution of the aggregates formed by Ru-complexes in:
a) pure water; **b)** physiological ionic strength NaCl solution; **c)** pseudo-physiological conditions in POPC formulation (Ru-complex/POPC 15:85 molar ratio). For all cases, the concentration of the investigated samples was 0.1 mmol kg^{-1}.

As pure compounds, Ru-containing amphiphilic complexes, namely ToThyRu, HoThyRu, DoHuRu and ToThyCholRu, were investigated by DLS and SANS in bidistilled water and at physiological ionic strength, *i.e.* NaCl 0.9% wt solution. The same systems in pseudo-physiological conditions, as previously hinted, could not be investigated in detail because of the fast hydrolysis and polymerization phenomena.

The hydrodynamic radii distributions of ToThyRu, HoThyRu, DoHuRu and ToThyCholRu pure aggregates in water and NaCl solution are reported in Figures 6a and 6b. From inspection of the Figures for the single-tailed molecules ToThyRu, HoThyRu, and ToThyCholRu, a bi-modal well-separated distribution of the hydrodynamic radii is observed. However, while for the ToThyRu system the smallest objects represent only a negligible percentage of the bi-modal distribution, the contrary is true for HoThyRu and ToThyCholRu. Furthermore, a perusal of Figure 6a shows that the dimensions of the biggest aggregates in ToThyRu solution qualitatively correspond to those of the smallest ones for the HoThyRu system. For ToThyCholRu the presence of two different aggregates is observed, with mean hydrodynamic radii of ~ 10 and ~ 30 nm, respectively. In the case of the double tailed DoHuRu aggregates, the mean value of its monomodal radius distribution is similar to the value shown by the biggest aggregates formed by ToThyRu and the smallest ones formed by HoThyRu. Similar results are observed in NaCl solution, except for ToThyCholRu, that shows a major population for the smallest aggregates (see Figure 6b).

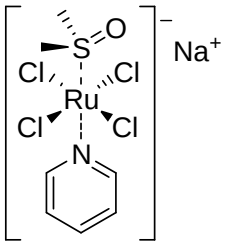
In pseudo-physiological conditions in POPC formulation (Ru complex/POPC 15:85 molar ratio) all the complexes show a monomodal distribution of the aggregates that present a hydrodynamic radii distribution similar to pure POPC (Figure 6c). Analyzing also the SANS results we can deduce that the aggregation is essentially driven by POPC, as shown in Table 1.

Sample in pseudo-physiological conditions (0.1 mmol kg-1)	**Morphology**	**τ/Å**	***N***	**d_L/Å**
ToThyRu/POPC	vesicles	4.5±0.5	12±1	6.5±0.3
HoThyRu/POPC	vesicles	4.4±0.3	25±1	5.4±0.1
DoHuRu/POPC	vesicles	4.2±0.7	10±1	6.7±0.4
ToThyCholRu/POPC	vesicles	4.5±0.7	1-2	78±7
POPC	vesicles	4.4±0.1	5±1	8.7±0.5

Table 1. SANS results for the investigated systems.

2.3 In Vitro Bioactivity Study

To gain an insight into their antiproliferative activity, the cytotoxicity profile of these novel nucleolipid-ruthenium complexes toward human cancer cells was analyzed in parallel with a ruthenium complex, that we named AziRu. This novel salt, designed as an analog of NAMI-A with one pyridine ligand replacing the imidazole residue – *i.e.*, the same ruthenium-surrounding subsystem present in ToThyRu, HoThyRu, DoHuRu and ToThyCholRu -, was here prepared to evaluate, by direct comparison, the benefits due to the presence of the nucleolipids within the structure of the complexes.



AziRu (NAMI-A-like ruthenium complex)

Cell lines were exposed to AziRu for 72 h in a growth inhibition assay and the cytotoxicity was determined in terms of IC_{50} values. The detected cytotoxicity values were rather moderate and the IC_{50} values for AziRu were 900, 690, 515, 405 and 1500 µM in HeLa, CaCo-2, WiDr, MCF-7 and SH-SY5Y cells, respectively. First, under the same incubation conditions these results indicate a significant difference in ruthenium cytotoxicity exclusively depending on the cell line, suggesting a selective cell-dependent mechanism of action. [33] Moreover, in accordance to many published studies on ruthenium complexes, AziRu cytotoxicity was much lower than that of a well known reference drug (data not shown) based on transition metals as the antitumor cisplatin. [34] However, concerning NAMI-A, it should be considered that the concentration needed to reach *in vitro* cytotoxicity is much higher than the NAMI-A concentration that is capable to inhibit the formation of metastases, suggesting that the recognized antimetastatic activity of NAMI-A is probably not due exclusively to a reduction in cell viability. [35]

A complete and comprehensive study on the *in vitro* bioactivity was carried out on the ToThyRu/POPC, HoThyRu/POPC, DoHuRu/POPC and ToThyCholRu/POPC formulations. Focusing on the cells (MCF-7 and WiDr lines) more sensitive to the ruthenium treatment in our *in vitro* experimental model, we assessed the potential cytotoxic effect of ruthenium-containing liposomes, and interestingly we found that the studied complexes in POPC formulation were definitively more active than AziRu in both cell lines. A graphical representation of these results is reported in Figure 7 for the ToThyCholRu/POPC liposome.

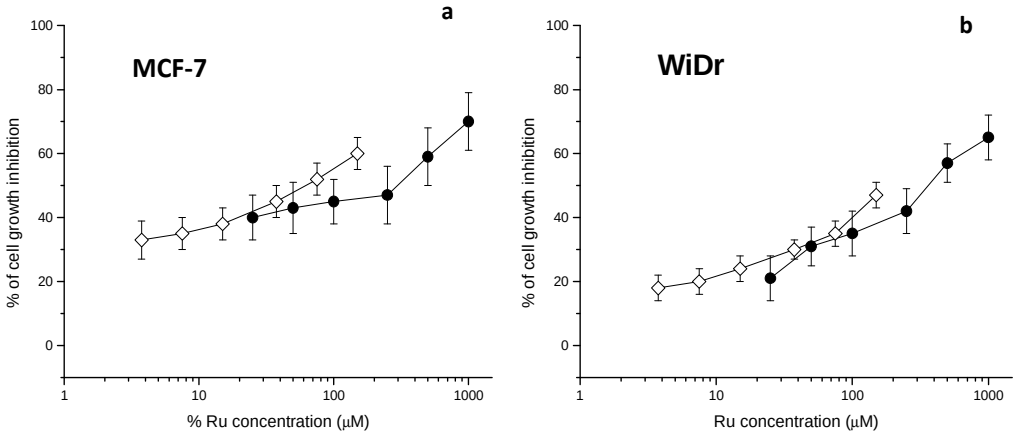


Figure 7. Growth inhibition of MCF-7 (panel A) and WiDr cells (panel B) treated for 72 h with AziRu (•) and ToThy-CholRu/POPC 15:85 liposome (◊), evaluated by MTT assay and cell counting. Data are expressed as percentage of untreated cells and are reported as mean of three independent experiments ± SEM.

The latter is a noteworthy finding in the design and biological evaluation of innovative supramolecular systems for the anticancer drugs vectorization. The calculated IC_{50} values reported in Table 2 for the real amount of ruthenium delivered to cells after 72 h of incubation in cell culture media also suggested that complexes in POPC formulation were more effective than AziRu in inducing growth inhibition in MCF-7 and WiDr cell lines. In addition, the evaluation of the potentiating factor (PF) values demonstrated that ruthenium vectorization by this liposome potentiated the growth inhibition induced by AziRu, reaching values higher than 3 in cancer cells of different histogenesis (see Table 2). Particularly interesting are the results obtained on the MCF-7 cell lines for the formulations ToThyRu/POPC and HoThyRu/POPC with IC_{50} values of 9 and 15 μM, respectively, and high potentiating factors (45 and 27, respectively). Similar *in vitro* results

emerged from bioassays on other human cancer cell lines for all the complexes, albeit generally less sensitive to the action of the antiproliferative ruthenium. In Figure 8, as a representative example, the results obtained on other cell lines for the ToThyCholRu/POPC system are reported.

Cell Lines	IC_{50}[a] (P.F.[b])				
	AziRu	**ToThyRu/POPC**	**HoThyRu/POPC**	**DoHuRu/POPC**	**ToThyCholRu/POPC**
MCF-7	405 ± 10	9±4 (45)	15±2 (27)	71 ± 6 (5.7)	70 ± 12 (5.7)
WiDr	515 ± 15	75±4 (6.9)	40±5 (12.9)	99 ± 5 (5.2)	165 ± 10 (3.1)

Table 2. IC_{50} values (μM) of Azi-Ru (a NAMI-A-like Ru complex) and of ToThyRu/POPC, HoThyRu/POPC, DoHuRu/POPC and ToThyCholRu/POPC liposomes in the indicated cell lines following 72 h of incubation, and the corresponding potentiating factors (P.F.). (a) IC_{50} values are reported as the mean value ± SEM; (b) potentiating factor calculated as the ratio of IC_{50} values of the complexes in POPC formulation with respect to the IC_{50} of AziRu complex. MCF-7, human breast adenocarcinoma cells; WiDr, human epithelial colorectal adenocarcinoma cells.

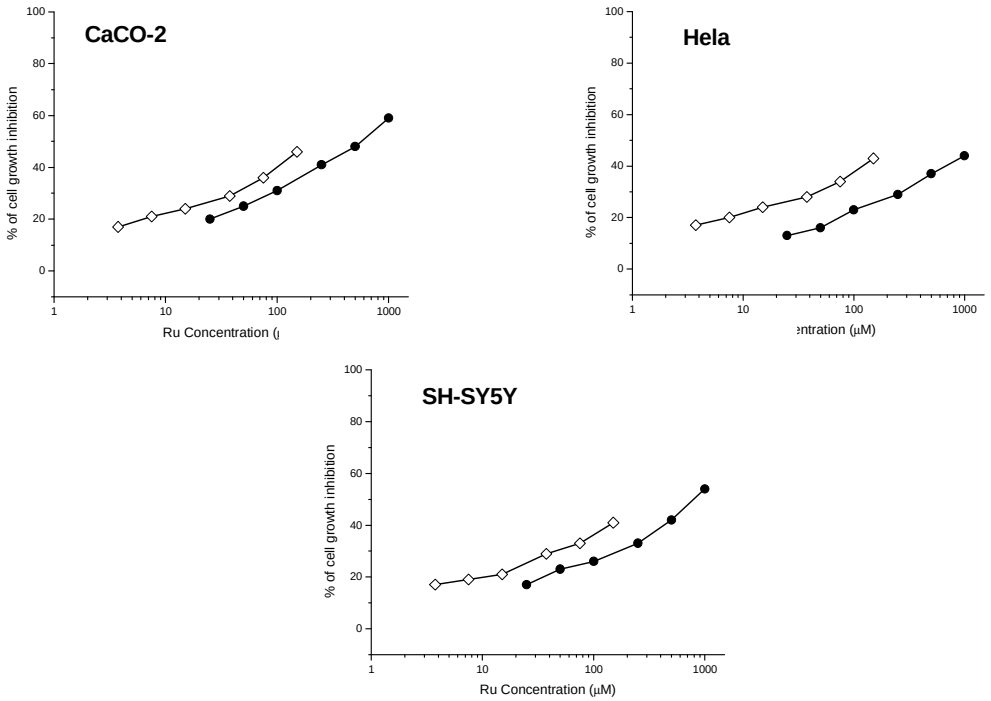


Figure 8. Growth inhibition of CaCo-2, Hela and SH-SY5Y cells treated for 72 h with AziRu (•) and ToThy-CholRu/POPC 15:85 liposome (◊), evaluated by MTT assay and cell counting. Data are expressed as percentage of untreated cells and are reported as the mean of three independent experiments ± SEM.
CaCo-2, human epithelial colorectal adenocarcinoma cells; HeLa, human cervical cancer cells; SH-SY5Y, human neuroblastoma cells.

Hence, consistently with the data previously described in Chapter 2 for the nucleolipid nanovectors, [36] different aggregate compositions differentially impacted cell viability. As demonstrated by cytotoxic activities higher than in the case of AziRu, the ruthenium-containing nucleolipidic liposomes stabilized with POPC seem to deliver the transition metal into the cells in a very effective way. Within this context, it is believed that free ruthenium complexes such as NAMI-A enter cells *via* a passive or facilitated passive transport molecular mechanism, albeit apparently with more difficulty than other antitumor drugs based on transition metals, such as cisplatin. [37] Together with other considerations, this outcome may account for the lower cytotoxicity of emerging ruthenium-based molecules assayed for antineoplastic activity, thus emphasizing the critical role of a proper vehiculation. The importance of the liposome stabilization imparted by the presence of the naturally occurring POPC is also supported by the fact that pure Ru complexes, even if *per se* organized in liposomial super-structures, resulted in a significantly lower cytotoxic effect. It can be therefore concluded that the absence of POPC within these novel liposomal formulations always reduces the effectiveness of the vehiculated anti-cancer drug.

3. Synthesis of a fluorescently-labelled nucleolipid Ruthenium(III) complex

For a detailed investigation of the *in vivo* mechanism of the synthesized complexes, including the internalization process into cells and their metabolic fate, we have designed a novel nucleolipid, built using HoThy as the main scaffold and further functionalized with a fluorescent label. To this purpose we have selected the dansyl group. This is a fluorescent molecule widely used in biochemistry and chemistry to label organic compounds. It offers several advantages: dansyl is a residue with limited steric hindrance compared to other common fluorescent dyes; it is easy to install and chemically stable in a large variety of reaction conditions; moreover, it is very sensitive to the solvent polarity conditions, thus offering relevant information on the local environment in which it is found. [38] Fluorescently-labelled nucleolipid **9** thus consists of:

- one pyridine-methyl arm at the N-3 position, as the “chemical handle” for the successive metal complexation;
- a dansyl group attached *via* a sulphate linkage in position 2’;
- one oleic acid residue in position 3’;
- one hepta(ethylene glycol) chain in position 5’.

Following the design we have adopted for ToThyRu, HoThyRu, DoHuRu and ToThyCholRu nanovectors, the latter residues have been attached through ester linkages, chemically stable but in principle easily reversible *in vivo* by esterases degradation (see Figure 9).

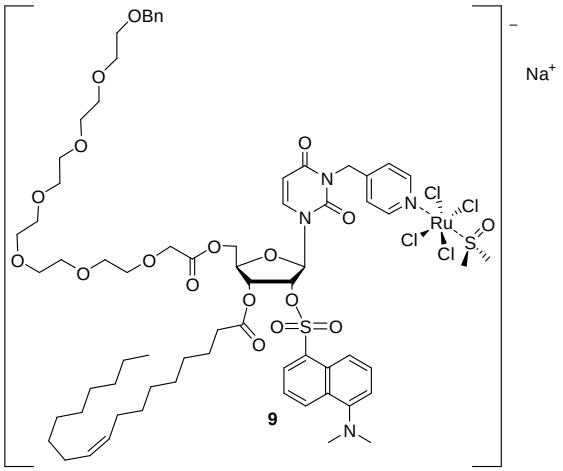
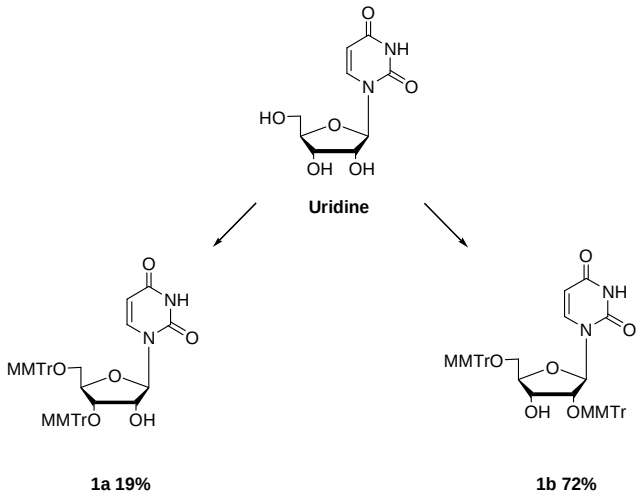


Figure 9. Fluorescently labelled Ruthenium(III) complex **9**, here designed as an analog of HoThyRu.

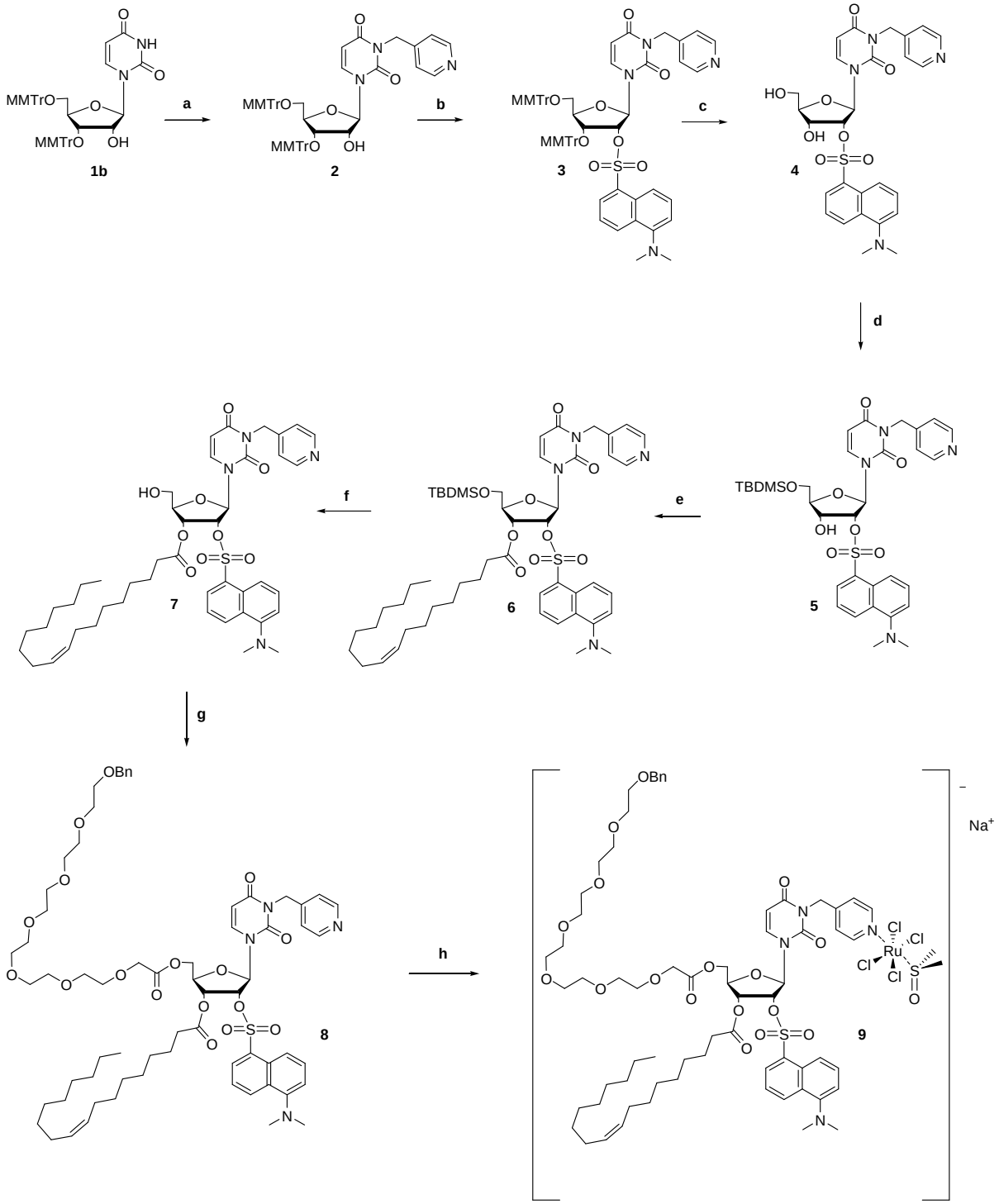
The synthesis of complex **9** was accomplished in 9 steps starting from uridine. The first step ensures the regioselective protection of the hydroxyl groups in positions 3' and 5', so to successively allow the efficient insertion of the fluorescent label in the 2'-OH position. To this purpose we have selected the acid-labile 4-monomethoxytrityl (MMTr) protective group. Following the procedure described by Ogilvie *et al.* to obtain dimethoxytritylether derivatives of uridine, [39] we have treated this nucleoside with 2.8 equivalents of MMTrCl and $AgNO_3$ in the presence of pyridine, obtaining two regioisomers, **1a** and **1b**, with an overall yield of 91%, which were easily purified by silica gel column chromatography. In this procedure, regioisomer **1a** was obtained in 19% yield while protected nucleoside **1b** in 72% isolated yield (Scheme 2). The latter derivative has been exploited as the starting material in the design of novel nucleolipid - Ru(III)-complexes, discussed in the following section; in parallel, **1a** was used as the precursor of target derivative **9**, the synthesis of which is here presented.



Scheme 2. Synthesis of MMTr-protected uridine derivatives **1a** and **1b.**

Reaction conditions: Uridine, MMTrCl (2.8 equiv), $AgNO_3$ (2.8 equiv), Py, THF, 12 h, r.t.

Regioisomer **1a** was then functionalized with a pyridine-methyl group in position N-3 by reaction with 4-(bromomethyl)pyridine hydrobromide in the presence of K_2CO_3, in analogy with the synthetic scheme used for the preparation of ToThy, HoThy, DoHu and ToThyChol. Successively, the N-3 alkylated nucleoside **2** was reacted with dansyl chloride, to give dansyl-labelled derivative **3**, with the dye inserted in 2' position (Scheme 3). Ditritylated nucleoside **3** was deprotected by acidic treatment with TFA in the presence of TIS to obtain deprotected nucleoside **4**. The successive reaction with a stoichiometric amount of TBDMSCl in the presence of imidazole allowed the selective protection of the position 5'. Subsequent esterification of **5** with oleic acid in the presence of DCC as the condensing agent led to nucleolipid **6**. This was then deprotected at the 5' position through a mild desilylation procedure, involving the use of $Et_3N{\cdot}3HF$ in THF, to obtain alcohol **7.** This nucleolipid was then esterified by reaction of the 5'-OH group with a residue of $BnO(CH_2CH_2O)_6CH_2COOH$ in the presence of DCC, used as the condensing agent, giving derivative **8**. As a final step, desired ruthenium complex **9** was obtained by reaction of **8** with an equimolar amount of $[RuCl_4(DMSO)_2]^-Na^+$ for 1 h at 40 °C in CH_2Cl_2. As in the case of ToThyRu, HoThyRu, DoHuRu and ToThyCholRu, complex **9** was characterized by ESI-MS analysis in the negative mode, showing only two diagnostic peaks, corresponding to the expected m/z ions (respectively at 1487.78 [M^--DMSO] and 1566.76 [M^-]), thus fully confirming the identity and purity of the desired Ru-complex-containing nucleolipid.



Scheme 3. Synthesis of ruthenium complex **9**. **a)** 4-(bromomethyl)pyridine hydrobromide, K_2CO_3, DMF, 48 h, 60 °C; **b)** Dansyl chloride, DMAP, TEA, CH_2Cl_2, 15 min, r.t.; **c)** TFA, TIS, 1 h, r.t.; **d)** TBDMSCl, imidazole, DMF, 2 h, r.t.; **e)** Oleic acid, DCC, DMAP, CH_2Cl_2, 1 h, r.t.; **f)** $Et_3N{\cdot}3HF$, THF, 12 h, r.t.; **g)** $BnO(CH_2CH_2O)_6CH_2COOH$, DCC, DMAP, CH_2Cl_2, 12 h, r.t.; **h)** $[RuCl_4(DMSO)_2]^-Na^+$, CH_2Cl_2, 1 h, 40 °C.

4. Synthesis of a second generation of nucleolipid Ruthenium(III) complexes

To expand the repertoire of available nucleolipid ruthenium(III) complexes and thus define the best structural requirements to obtain efficient anticancer agents, we have also explored a novel design, based on a different decoration of the nucleosidic skeleton. In the first series of compounds we have prepared (ToThyRu, HoThyRu, DoHuRu and ToThyCholRu, discussed in section 2.1), remarkably the pyridine-methyl group, chosen as the privileged ligand for ruthenium, is anchored to the N-3 position of the nucleobase. In principle, in order to exploit the recognition abilities of the nucleobases, these should not be blocked by hindered groups. Therefore, on the basis of the promising results obtained for the first generation of ruthenium(III) complexes, we here propose a novel nucleolipid carrying the pyridine-methyl group on the sugar moiety, as a model for a second generation of metal-complexed nucleolipids. The synthesized molecule **21** consists of:

- one pyridine-methyl arm at the 3' position, linked through an amide bond, as the ligand for the metal complexation;
- one oleic acid residue in position 2';
- one hepta(ethylene glycol) chain in position 5'.

A wide range of analogs of **21** can be easily accessible by varying the nature of the hydrophilic and lipophilic chains attached to the furanosic moiety (Figure 10).

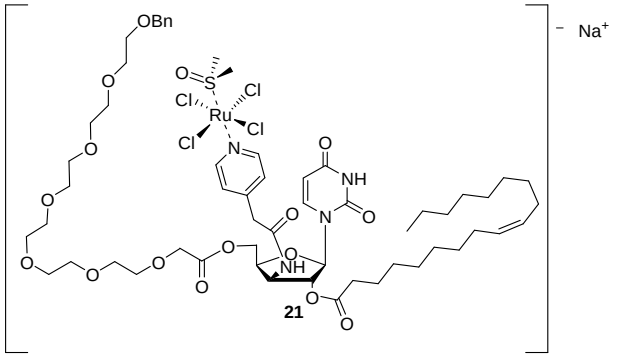
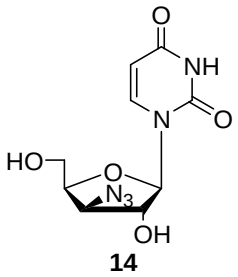


Figure 10. Novel nucleolipid ruthenium(III) complex **21**.

Key intermediate for the synthesis of this compound is the modified nucleoside 3′-azido-3′-deoxy-1-β-D-xylofuranosyluracil **14** showed in Figure 11. This is a very simple building block, having three functional groups which can be easily differentiated: two hydroxyls, one primary and one secondary function, and one azido group, which can be considered as a masked amino group, being easily reduced to a primary amine. By exploiting a rational scheme of protecting groups to temporarily deactivate these functions, it is therefore possible to insert on each of them a specific decoration in a very selective and efficient manner. Nucleoside **14** was already known in the literature, but the previously described synthesis involved a large number of synthetic steps and low yields. [40] In this work we have adopted a new, optimized protocol, requiring only straightforward and high yielding steps.



14

Figure 11. 3′-azido-3′-deoxy-1-β-D-xylofuranosyluracil **14**.

As a starting material we have selected nucleoside **1b**, carrying one free hydroxyl group in the 3’ position. The key step for the conversion of the 3’-OH group into an azide involves the use of the Mitsunobu reaction which leads to inversion of configuration at C-3’. [41] To this purpose, our first goal has been the development of a synthetic strategy for the efficient and easy protection of the nucleobase in N-3 position to prevent, under the Mitsunobu reaction conditions, the formation of the corresponding 2,3’-anhydro-derivative (Figure 12). [42]

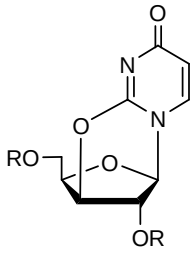
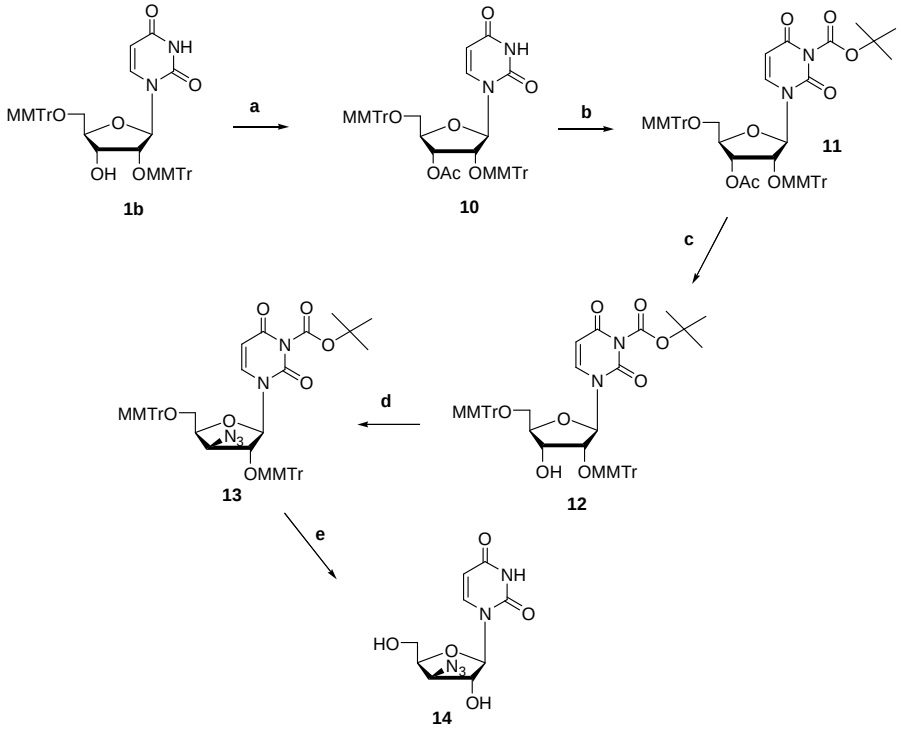


Figure 12. 2,3’-anhydro-derivative.

This target was accomplished by the use of the Boc, that was inserted as the N-3 protective group by a simple synthetic manipulation developed in our research group (see Chapter 6). [43]

Nucleoside **1b** was first protected in 3' position by insertion of an acetyl group; successively, nucleoside **10** was protected in N-3 position with a Boc group by reaction with Boc_2O, TEA and DMAP in DCM obtaining nucleoside **11** in 80% isolated yield. The 3'-acetyl group was then removed by treatment with catalytic K_2CO_3 in CH_3OH/TEA obtaining nucleoside **12** in quantitative yield. The acetylation/deacetylation procedures required prolonged reaction times (compared to the classically reported procedures), probably as a consequence of the high steric hindrance of the 3' position (Scheme 4).

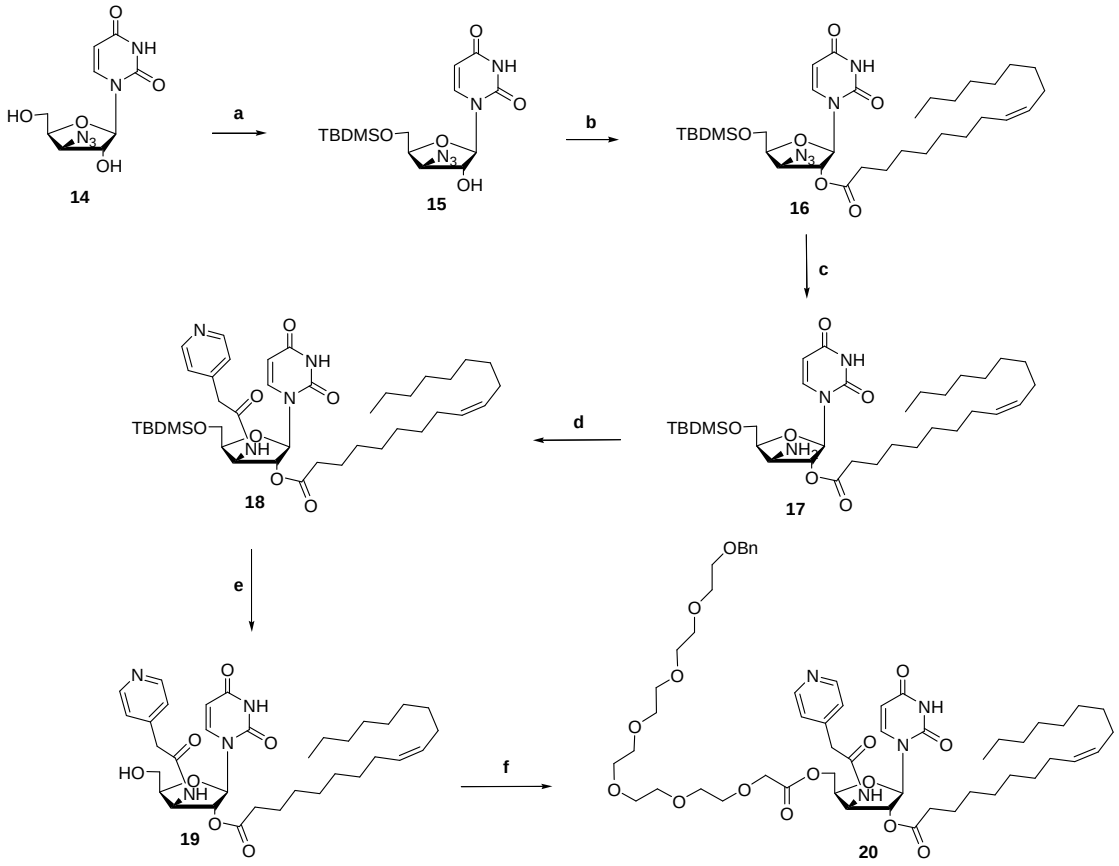


Scheme 4. Synthesis of derivative **14**. **a)** Ac_2O, Py, 12 h, r.t.; **b)** Boc_2O, TEA, DMAP, DCM, 1 h, r.t.; **c)** K_2CO_3, MeOH / Et_3N (3:1), 24 h, r.t.; **d)** PPh_3, DIAD, DPPA , THF, 24 h, r.t.; **e)** TFA, TIS/DCM (1:9), 1 h, r.t.

The Mitsunobu reaction was carried out treating nucleoside **12** with PPh_3, DIAD and DPPA, leading to the 3'-azido-3'-deoxyxylofuranosyl nucleoside **13** in 92% isolated yield; on this substrate the reaction with 10% TFA and TIS in DCM was carried out to achieve the fully deprotected nucleoside **14**. Following this synthetic scheme, the desired

nucleoside 3′-azido-3′-deoxy-1-β-D-xylofuranosyluracil was obtained in 6 steps with 50% overall yields starting from uridine.

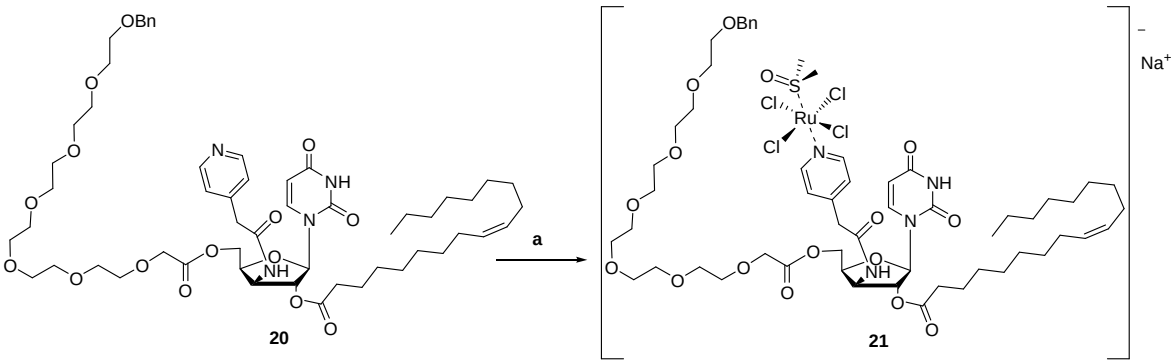
The obtained nucleoside **14** was easily manipulated to obtain the desired complex **21**. As the first step, we have protected the 5′ position with the TBDMS group and then esterified the 2′ position with oleic acid, using DCC as the condensing agent, thus obtaining derivative **16** (Scheme 5).



Scheme 5. Synthesis of nucleolipid **19. a)** TBDMSCl, imidazole, DMF, 1 h, r.t.; **b)** oleic acid; DCC, DMAP, CH_2Cl_2, 2 h, r.t.; **c)** H_2 1 atm, Pd/C 10% p.p., AcOEt, 24 h, r.t.; **d)** 2-(pyridin-4-yl)acetic acid hydrochloride; DCC, DMAP, CH_2Cl_2, 12 h, r.t.; **e)** $Et_3N{\cdot}HF$, THF, 3 h, r.t.; **f)** $BnO(CH_2CH_2O)_6CH_2COOH$, DCC, DMAP, CH_2Cl_2, 2 h, r.t.

The reduction of the azide group to amine was carried out by reaction with H_2 (P= 1 atm) in the presence of Pd/C for 24 h. The amino-nucleoside **17** was obtained in almost quantitative yields in a pure form, simply filtering off and washing the catalyst. Then, the amino group was reacted with a residue of 2-(pyridin-4-yl)acetic acid hydrochloride to obtain derivative **18.** The subsequent reaction with $Et_3N{\cdot}3HF$ liberates the 5′-OH group in

nucleoside **19** that was then functionalized with a residue of benzyl-capped hexa(ethylene glycol) acetic acid, providing nucleolipid **20**. The reaction of **20** with an equimolar amount of $[RuCl_4(DMSO)_2]^-Na^+$ for 1 h at 40 °C in CH_2Cl_2 led to complex **21** in quantitative yields, obtained in a pure form by TLC control and ESI-MS analysis, thus not requiring any further purification (Scheme 6).



Scheme 6. Synthesis of Ru(III) complex **21**. **a)** $[RuCl_4(DMSO)_2]^-Na^+$, CH_2Cl_2, 1 h, 40 °C.

5. Conclusions

In this Chapter the synthesis and characterization of a novel family of Ruthenium(III) complexes based on amphiphilic nucleolipids, as potential anticancer drugs, has been described. The first four derivatives, ToThyRu, HoThyRu, DoHuRu and ToThyCholRu, were synthesized starting from the corresponding nucleolipid nanovectors already described in Chapter 2, by complexation with the easily accessible ruthenium salt $[RuCl_4(DMSO)_2]^-Na^+$. These novel amphiphilic complexes were characterized by ESI-MS data and then studied in their self-assembling properties in aqueous solutions and pseudo-physiological conditions. A detailed study of their aggregative behaviour has shown that the synthesized compounds form a wide variety of supramolecular aggregates. Aggregates formed by the pure Ru-containing complexes show a fast degradation kinetics that, depending on pH, is almost complete from few hours to some days. This behaviour is absolutely comparable to the one reported in the literature for NAMI-A and other Ru-complexes currently in clinical trials; as a matter of fact, the scarce stability in physiological conditions is mentioned as the major, crucial limit to possible future exploitations of all the known Ru-based candidates as marketable drugs. We therefore addressed this main issue by testing these complexes in formulation with natural, biocompatible phospholipids. First, very promising results have been obtained by mixing the here synthesized Ru-complexes with POPC. Particularly, we observed that, when the ruthenium complexes are lodged in POPC liposomes up to a 15% in moles, the degradation in physiological media is virtually null for several days and the formulation is stable for at least several weeks.

It is worth mentioning that *in vitro* bioscreenings based on the evaluation of concentrations-effect curves reveal that complexes in POPC formulation are much more effective in inhibiting the growth of human cancer cells of different histogenesis with respect to the reference ruthenium-complex NAMI-A-like Azi-Ru. In particular, promising results were obtained with ToThyRu/POPC and HoThyRu/POPC formulations on MCF-7 cancer cells. Therefore, this study opens new perspectives in the synthesis and characterization of highly biocompatible transition metal-based aggregates having potential applications as antineoplastic agents.

Successively, for a detailed investigation of the *in vivo* mechanism of action of the synthesized complexes, including studies on their internalization process into cells and their metabolic fate, we have designed nucleolipid-complex **9**, built using HoThy as the main scaffold and further functionalized with the fluorescent dansyl group. On this novel Ru(III) nucleolipid-complex and its precursor **8** studies are actually in progress on their cytotoxic activity on a panel of cancer cells lines, and on the *in vivo* mechanism of action, carried out by taking advantage also of fluorescence microscopy techniques.

As a further improvement, to expand the repertoire of available nucleolipid ruthenium(III) complexes and thus define the best structural requirements to obtain efficient anticancer agents, I have also explored a novel design, based on a different decoration of the nucleosidic skeleton. In particular, novel nucleolipid **21** carrying the pyridine-methyl group on the sugar moiety was synthesized, as a model compound for a second generation of metal-complexed nucleolipids. To this purpose, I selected 3′-azido-3′-deoxy-1-β-D-xylofuranosyluracil as the key nucleosidic scaffold to obtain the target, highly functionalized nucleolipid. A novel synthetic procedure was developed and optimized to prepare this modified nucleoside, obtained in only six straightforward synthetic steps in 50% yield starting from Uridine. Successive synthetic manipulation of this modified nucleosidic building block, requiring seven, simple and efficient steps, afforded the model nucleolipid **21** in 20% overall yields.

6. Experimental Section

General synthetic methods.

All the reagents were of the highest commercially available quality and were used as received. TLC analyses were carried out on silica gel plates from Merck (60, F254). Reaction products on TLC plates were visualized by UV light and then by treatment with a 10 % $Ce(SO_4)_2/H_2SO_4$ aqueous solution. For column chromatography, silica gel from Merck (Kieselgel 40, 0.063-0.200 mm) was used. NMR spectra were recorded on Bruker WM-400, Varian Gemini 200 and Varian Inova 500 spectrometers, as specified. All the chemical shifts are expressed in ppm with respect to the residual solvent signal. Peak assignments have been carried out on the basis of standard H-H COSY and HSQC experiments. The following abbreviations were used to explain the multiplicities: s = singlet; d = doublet; t = triplet; q = quartet; m = multiplet; b = broad; dd = double doublet. For the ESI MS analyses, a Waters Micromass ZQ instrument – equipped with an Electrospray source – was used in the positive and/or negative mode. MALDI TOF mass spectrometric analyses were performed on a PerSeptive Biosystems Voyager-De Pro MALDI mass spectrometer in the Linear mode using 3,4-dihydroxybenzoic acid as the matrix.

General procedure for the synthesis of Nucleoside-Ru complexes ToThyRu, HoThyRu, DoHuRu and ToThyCholRu

The selected nucleolipid (0.033 mmol) was dissolved in 1 mL of the appropriate dry solvent (CH_3CN for ToThy; CH_2Cl_2 for HoThy, DoHu and ToThyChol) and then [*trans*-$RuCl_4(DMSO)_2]^-$ Na^+ (0.033 mmol) was added. The reaction mixture was stirred at 40 °C and the solvent was then removed under reduced pressure. The reaction, monitored by TLC on alumina, showed in 4 h the total disappearance of the starting material to give the desired salt in almost quantitative yields.

- **ToThyRu**

Red powder: R_f = 0.5 [$CHCl_3$/CH_3OH, 95:5 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): significant signals at δ 7.39 (s, 1H, H-6); 6.37 (m, 1H, H-1'); 5.40-5.37 (overlapped signals, 2H, H-9 and H-10 oleic acid); 5.30 (s, 2H, -C$\underline{H}_2$Py); 5.23 (m, 1H, H-3'); 4.99 (s, 1H, -OC$\underline{H}_2$CO- TEG); 4.30 (broad signals, 2H, H_2-5)'; 3.74 (m, 1H, H-4'); 3.67-3.37 [overlapped signals, 15H, 3x(-O-C$\underline{H}_2$-C$\underline{H}_2$-O-) and C$\underline{H}_3$O- TEG]; 2.34 (m, 2H, H-2'); 2.07-1.93 (overlapped signals, 4H, C$\underline{H}_2$-8 and C$\underline{H}_2$-11 overlapped signals); 1.61-1.31 [broad and overlapped signals, 27H, 12x(-C$\underline{H}_2$- oleic acid), C$\underline{H}_3$-Thy]; 0.89 (apparent triplet, 3H, C$\underline{H}_3$-18 oleic acid); -2.05 (very broad signal, Py-Ru); -12.7 (very broad signal, C$\underline{H}_3$ DMSO).

^{13}C NMR ($CDCl_3$, 50 MHz): significant signals at δ 173.0 (C=O oleic acid); 169.5 (C=O TEG); 160.9 (C-4); 154.4 (C-2); 133.5 (C-6); 129.8 and 129.5 (C-9 and C-10 oleic acid); 110.6 (C-5); 84.6 (C-1'); 81.6 (C-4'); 73.7 (C-3'); 71.6, 70.9 and 70.3 (O-$\underline{C}H_2$-$\underline{C}H_2$-O- TEG); 68.6 (-OCH_2CO- TEG); 63.8 (C-5'); 58.8 ($\underline{C}H_3$O- TEG); 44.8 (-$\underline{C}H_2$Py); 33.8 (C-2'); 31.6, 29.6, 29.4, 29.0, 28.8, 26.9, 24.4 and 22.4 (aliphatic carbons of oleic acid); 14.2 ($\underline{C}H_3$-18 oleic acid); 12.4 (-$\underline{C}H_3$-Thy).

ESI-MS (negative ion): calculated for $C_{45}H_{73}Cl_4N_3O_{12}RuS$, 1123.27; *m/z*, found: 1124.24 (M^-).

- **HoThyRu**

Red powder: R_f = 0.4 [AcOEt/CH_3OH, 1:1 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz, 0.06 M): significant signals at δ 7.50-7.26 [broad and overlapped signals, 6H, H-6 and (5 x aromatic protons of Bn)]; 6.34 (broad signal, 1H, H-1'); 5.38 and 5.34 (m, 2H, H-9 and H-10 oleic acid); 4.56 (s, 2H, -C$\underline{H}_2$-Ph); 4.47 (m, 2H, H_2-5'); 4.36-4.15 (broad and overlapped signals, 3H, H-4' and -OC$\underline{H}_2$CO-HEP); 3.80-3.51 [broad and overlapped signals, 24H, 6x(-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)]; 2.35-2.28 (overlapped signals, 4H, H_2-2' and C$\underline{H}_2$-2 oleic acid); 2.05-1.94 (broad signals, 4H, C$\underline{H}_2$-8 and C$\underline{H}_2$-11 oleic acid); 1.58-1.31 [broad and overlapped signals, 27H, 12x(-C$\underline{H}_2$- oleic acid), C$\underline{H}_3$-Thy]; 0.89 (broad signal, 3H, C$\underline{H}_3$-18 oleic acid); -1.98 (very broad signal, Py-Ru); -13.0 (very broad signal, 6H, C$\underline{H}_3$ DMSO).

^{13}C NMR ($CDCl_3$, 50 MHz): significant signals at δ 173.0 (C=O oleic acid); 169.8 (C=O HEP); 163.1 (C-4); 153.8 (C-2); 138.0 (C quaternary aromatic carbon of Bn); 133.6 (C-6); 129.7 (C-9 and C-10 oleic acid); 128.1, 127.5 and 127.3 (aromatic carbons Bn); 110.7 (C-5); 83.6 (C-1'); 82.0 (C-4'); 74.0 (C-3'); 73.1 (-$\underline{C}H_2$Ph); 71.0, 70.6 and 69.4 (O-$\underline{C}H_2$-$\underline{C}H_2$-O- HEP); 68.1 (-O$\underline{C}H_2$CO- HEP); 64.3 (C-5'); 44.7 (-$\underline{C}H_2$Py); 37.1 (C-2'); 31.6, 29.4, 29.2, 29.0, 28.8, 26.9, 24.4 and 22.3 (aliphatic carbons of oleic acid); 14.1 ($\underline{C}H_3$-18 oleic acid); 12.6 (-$\underline{C}H_3$-Thy).

ESI-MS (negative ions): calculated for $C_{57}H_{89}Cl_4N_3O_{15}RuS$, 1331.38; *m/z*, found: 1332.86 (M^-).

- **DoHuRu**

Red powder: R_f = 0.3 [AcOEt/CH_3OH, 1:1 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): significant signals at δ 7.53 (d, *J* = 8.5 Hz, 1H, H-6); 7.36-7.26 [overlapped signals, 5H, aromatic proton of Bn)]; 6.03 (d, *J* = 5.0 Hz, 1H, H-1'); 5.89 (d, *J* = 8.0 Hz, 1H, H-5); 5.37 (overlapped signals, 2H, H-2' and H-3'); 5.34 [overlapped signals, 4H, 2 x (H-9 and H-10 oleic acid)]; 5.00 (broad singlet, 2H,-C$\underline{H}_2$Py); 4.56 (s, 2H, -C$\underline{H}_2$-Ph); 4.23-4.01 (broad and overlapped signals, 3H, H_2-5', H-4', -OC$\underline{H}_2$CO- HEP); 3.74-3.63 [overlapped signals, 24H, 6x(-O-C$\underline{H}_2$-C$\underline{H}_2$-O-)]; 2.34 and 2.25 [overlapped signals, 4H, 2x(C$\underline{H}_2$)-2 oleic acid)]; 2.21 and 2.01 [overlapped signals, 8H, [2x(C$\underline{H}_2$-8) and 2x(C$\underline{H}_2$-11) oleic acid]; 1.56-1.27 [overlapped signals, 48H, 2x(12 x -C$\underline{H}_2$-) oleic acid]; 0.89-0.87 [overlapped signals, 6H, 2x(C$\underline{H}_3$)-18 oleic acid)]; -1.97 (very broad signal, Py-Ru); -12.9 (very broad signal, C$\underline{H}_3$ DMSO).

^{13}C NMR ($CDCl_3$, 50 MHz): significant signals at δ 171.5, 170.9 and 163.8 (C=O esters); 157.7 (C-4); 147.8 (C-2); 140.7 (C-6); 138.9 (quaternary carbon of Bn); 130.0 (C-9 and C-10 oleic acids); 128.0, 127.2 and 127.0 (aromatic carbons of Bn); 101.0 (C-5); 90.0 (C-1'); 80.9 (C-4'); 73.2 and 70.5 (C-2' and C3'); 70.5 (-$\underline{C}H_2$Ph); 70.0 (O-$\underline{C}H_2$-$\underline{C}H_2$-O- HEP); 68.7 (-O$\underline{C}H_2$CO- HEP); 63.8 (C-5'); 43.0 (-$\underline{C}H_2$Py); 33.8 [overlapped signals, 2x(-$\underline{C}H_2$)-2 oleic acid)]; 31.6, 29.9, 29.6, 29.3, 29.0, 27.5, 25.3, 24.1 and 22.3 (aliphatic carbons of oleic acid); 13.0 [2x($\underline{C}H_3$-18 oleic acid)].

ESI-MS (negative ions): for $C_{74}H_{119}Cl_4N_3O_{17}RuS$, calculated 1595.61; *m/z*, found: 1596.9 (M^-).

- **ToThyCholRu**

red powder: *Rf*: 0.4 ($CHCl_3$/CH_3OH, 95:5, v/v).

^{1}H NMR ($CDCl_3$, 500 MHz): *δ* significant signals at δ 9.00 (d, *J* = 6 Hz, 1H, H-6); 6.69 (bs, 1H, H-1'); 5.54 (bs, 1H, H-6 Chol), 5.41 (bs, 1H, H-3'); 5.08 (s, 2H, -C$\underline{H_2}$-Py); 4.25-4.00 (overlapped signals, 7H, H-5'$_a$ and H-5'$_b$, -C$\underline{H_2}$C=O TEG, C$\underline{H_2}$C=O and H-4'); 3.65-3.3.51 (overlapped signals, 12H, 3x –OC$\underline{H_2}$-C$\underline{H_2}$O-); 3.40 (s, 3H, -OC$\underline{H_3}$); 3.23 (bs, 1H, H-3 Chol); 2.38 (m, 2H, H-2'); 1.00 (s, 3H, C$\underline{H_3}$-19); 0.95 (d, *J* = 6.0 Hz, 3H, C$\underline{H_3}$-21 Chol); 0.86 (coincident d's, *J* = 4.5 Hz, 6H, C$\underline{H_3}$-26 and C$\underline{H_3}$-27 Chol); 0.72 (s, 3H, C$\underline{H_3}$-18 Chol), -1.87 (bs, Py protons); -12.66 (bs, C$\underline{H_3}$ of DMSO).

^{13}C NMR ($CDCl_3$, 125 MHz): *δ* significant signals δ 170.0 (C=O Chol ester); 169.3 (C=O TEG ester); 160.8 (C-4); 139.9 (C-5 Chol); 132.1 (C-6); 121.9 (C-6 Chol); 109.1 (C-5); 84.4 (C-1'); 82.2 (C-4'); 81.4 (C-3 Chol); 74.3 (C-3'); 71.8, 70.8, 70.3 (carbons of TEG); 68.4 (-$\underline{C}H_2$C=O TEG); 65.1 (-$\underline{C}H_2$C=O Chol); 63.7 (C-5'); 59.1 (-O$\underline{C}H_3$); 49.8 (-$\underline{C}H_2$Py); 42.0 (C-2'); 22.5 and 22.1 (C-27 and C-26 Chol); 19.0 (C-19 Chol); 18.4 (C-21 Chol); 14.2 ($\underline{C}H_3$-Thy); 11.6 (C-18 Chol).

ESI-MS (negative ions): for $C_{56}H_{87}Cl_4N_3O_{13}RuS^-$ calculated 1283.37; found *m/z*: 1285.18 (M^-); 1209.12 (M^--DMSO).

- **Synthesis of AziRu complex**

In analogy to the synthesis of **1**, pyridine (3.0 µl, 0.037 mmol) was dissolved in 0.3 ml of dry acetonitrile and then [trans-$RuCl_4(DMSO)_2]^-$ Na^+ (15.6 mg, 0.037 mmol) was added. The reaction mixture was stirred at 50 °C for 2 h and the solvent was then removed *in vacuo*, giving the desired product AziRu, identified on the basis of its NMR and ESI-MS data (data not shown), in quantitative yields.

Synthesis of fluorescent Ruthenium complex 9.

Synthesis of compounds 1a and 1b.

Uridine (1.00 g, 4.09 mmol) was suspended in 120 mL of anhydrous THF. Pyridine (3.30 mL, 40.9 mmol), $AgNO_3$ (1.95 g, 11.5 mmol) and MMTrCl (3.54 g, 11.5 mmol) were sequentially added to the solution. The resulting mixture was stirred at room temperature for 12 h, then the suspension was filtered and dried under reduced pressure. The crude was dissolved in CH_2Cl_2 and then washed with a water solution of $NaHCO_3$ 5% p.p.; the collected organic phases were dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column, eluted with *n*-hexane/AcOEt (1:1, v/v + 1% TEA). Nucleoside **1a** was obtained in 19% isolated yield (614 mg, 0.778 mmol), while its regioisomer **1b** in 72% yield (2.32 g, 2.95 mmol).

1a: white solid; R_f = 0.6 [*n*-hexane/AcOEt, 4:6 v/v].

1b: white solid; *Rf* = 0.8 [*n*-hexane/AcOEt 4:6 v/v].

Regioisomer **1a**:

^{1}H-NMR (C_6D_6, 200 MHz): δ 7.78-6.7 (29H, overlapped signals, H-6 and aromatic protons of *MMTr*); 6.45 (1H, d, *J* = 4.6 Hz, H-1'); 5.27 (1H, d, *J* = 8.6 Hz, H-5); 5.02 (1H, d, *J* = 7.2 Hz, H-3'); 4.60 (1H, dd, *J* = 4.4 and 4.6 Hz, H-2'); 4.14 (1H, broad singlet, H-4'); 3.52-3.11 (2H, m, H-5'$_a$ and H-5'$_b$); 3.35 (6H, s, 2x OC*H*$_3$).

^{13}C-NMR (C_6D_6, 50 MHz): δ 163.3 (C-4); 159.5, 145.2, 144.5, 135.8, 135.4, 131.4, 130.9, 133.6 (aromatic carbons of *MMTr*); 151.4 (C-2); 140.2 (C-6); 102.7 (C-5); 90.2 (C-1'); 88.2 and 87.7 (quaternary carbons of *MMTr*); 83.3 (C-4'); 75.2 (C-3'); 73.8 (C-2'); 63.2 (C-5'); 54.8 (O*C*H$_3$).

ESI-MS (positive ions): for $C_{49}H_{44}N_2O_8$, calculated 788.8823; found m/z: 811.94 [M+Na^+]; 827.79 [M+K^+].

Regioisomer **1b:**

^{1}H NMR (C_6D_6, 200 MHz): *δ* 7.70 (1H, d, *J* = 8.2 Hz, H-6); 7.64-6.64 (28H, overlapped signals, aromatic protons of *MMTr*); 6.61 (1H, d, *J* = 7.2 Hz, H-1'); 5.22 (1H, d, *J* = 8.4 Hz, H-5); 4.70 (1H, dd, *J* = 4.4 and 4.6 Hz, H-2'); 4.03 (1H, bs, H-3'); 3.33 (3H, s, OC*H*$_3$); 3.29 (3H, s, OC*H*$_3$); 3.13-2.92 (3H, overlapped signals, H-5'$_a$, H-5'$_b$ and H-4').

^{13}C NMR (C_6D_6, 50 MHz): δ 163.3 (C-4); 159.7, 159.4, 149.5, 144.9, 144.3, 143.9, 135.9, 135.2, 131.0, 130.6, 128.7, 128.5, 128.2, 128.0, 127.5, 123.6, 114.1, 113.6 (aromatic carbons of *MMTr*); 151.8 (C-2); 140.8 (C-6); 103.2 (C-5); 88.0 and 87.8 (quaternary carbons of *MMTr*); 87.1 (C-1'); 84.7 (C-4'); 78.2 (C-3'); 71.3 (C-2'); 64.7 (C-5'); 54.9 (O*C*H_3).
ESI-MS (positive ions): for $C_{49}H_{44}N_2O_8$, calculated 788.8823; found m/z: 811.97 [M+Na^+]; 827.84 [M+K^+].

Synthesis of compound 2.

1a (600 mg, 0.760 mmol) was dissolved in 4 mL of anhydrous DMF. K_2CO_3 (420 mg, 3.04 mmol) and 4-(bromomethyl)pyridine hydrobromide (384 mg, 0.152 mmol) were then added to the solution, which was stirred at 60 °C for 48 h. The solution was then concentrated under reduced pressure, diluted with CH_2Cl_2 and washed three times with H_2O. The collected organic phases were dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column, eluted with *n*-hexane/AcOEt (35:65, v/v + 1% TEA). Nucleoside **2** was obtained in 35% isolated yield (237 mg, 0.269 mmol).

2: white solid; R_f = 0.4 [*n*-hexane/AcOEt, 4:6 v/v].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 8.38 (2H, d, *J* = 5.8 Hz, 2x H_αPy); 7.63-6.51 (31H, overlapped signals, H-6, 2x H_βPy and aromatic protons of *MMTr*); 6.41 (1H, d, *J* = 4.8 Hz, H-1'); 5.28 (1H, d, *J* = 8.0 Hz, H-5); 4.72 (2H, m, C*H*$_2$Py); 4.54 (1H, dd, *J* = 4.6 and 4.4 Hz, H-3'); 4.04 (1H, broad singlet, H-4'); 3.90 (1H, m, H-2'); 3.49-3.07 (2H, m, H-5'$_a$ and H-5'$_b$); 3.26 and 3.21 (6H, s, 2x OC*H*$_3$).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 161.9 (C-4); 159.8, 146.1, 144.5, 135.8, 135.3, 131.2, 130.9, 129.0, 128.5, 128.1 128.0 127.5, 113.6 (aromatic carbons of *MMTr*); 151.3 (C-2); 150.0 (C_αPy); 145.1 (C_γPy); 139.0 (C-6); 123.7 (C_βPy); 101.6 (C-5); 91.4 (C-1'); 88.0 and 87.7 (quaternary carbons of *MMTr*); 83.3 (C-4'); 74.8 (C-3'); 74.0 (C-2'); 63.8 (C-5'); 54.8 (O*C*H_3); 43.0 (*C*H_2Py).

ESI-MS (positive ions): for $C_{55}H_{49}N_3O_8$, calculated 879.9929; found m/z: 881.73 [M+H^+].

Synthesis of compound 3.

2 (104 mg, 0.118 mmol) was dissolved in anhydrous CH_2Cl_2 (1.0 mL). TEA (66 µL, 0.472 mmol), DMAP (1.0 mg, 0.0060 mmol) and dansyl chloride (41 mg, 0.153 mmol) were then added to the solution. The reaction mixture was stirred at room temperature for 15 min, then the solvent was removed under reduced pressure and the residue purified on a silica gel column, eluted with *n*-hexane/AcOEt (7:3, v/v + 1% TEA). The product **3** was obtained in 87% isolated yield (114 mg, 0.103 mmol).

3: yellow solid; R_f= 0.4 [*n*-hexane/AcOEt 3:7 v/v]

^{1}H-NMR ($CDCl_3$, 500 MHz): δ 8.47 (2H, d, *J* = 5.5 Hz, 2x H_αPy); 7.86 (2H, d, *J* = 6.0 Hz, 2x H_βPy); 7.54-6.69 (35H, overlapped signals of *Dansyl* group, *MMTr* and H-6); 6.54 (1H, d, *J*= 9.0 Hz, H-1'); 5.38 (1H, d, *J* = 8.0 Hz, H-5); 5.22-4.97 (3H, overlapped signals, C*H*$_2$Py and H-2'); 4.41 (1H, m, H-3'); 4.23 (1H, d, *J*= 5.0 Hz, H-4'); 3.81-3.71 (8H, overlapped signals, 2xOC*H*$_3$, H-5'$_a$ and H-5'$_b$); 3.00 (6H, s, N(C*H*$_3$)$_2$).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 161.0 (C-4); 158.6, 150.2, 143.8, 143.0, 135.9, 134.4, 131.7, 131.4, 130.4, 130.2, 129.8, 129.3, 128.4, 127.9, 127.1, 119.4, 115.6, 113.2 (aromatic carbons of *MMTr* and *Dansyl* group); 151.6 (C-2); 150.0 (C_αPy); 145.7 (C_γPy); 136.8 (C-6); 123.1 (C_βPy); 101.7 (C-5); 88.2 and 87.3 (quaternary carbons of *MMTr*); 83.8 (C-1'); 83.1 (C-4'); 77.9 (C-2'); 73.0 (C-3'); 62.8 (C-5'); 55.0 (O*C*H$_3$); 45.2 [N(*C*H$_3$)$_2$]; 42.9 (*C*H$_2$Py).

ESI-MS (positive ions): for $C_{67}H_{60}N_4O_{10}S$, calculated 1113.2791; found m/z: 1113.06 [M+H$^+$]; 1135.05 [M+Na$^+$]; 1151.99 [M+K$^+$].

Synthesis of compound 4.

3 (108 mg, 0.097 mmol), dissolved in anhydrous CH_2Cl_2 (0.90 mL), was treated with 50 µL of TFA and 50 µL of TIS. The reaction mixture was stirred at room temperature for 1 h, then the solvent was removed under reduced pressure and the residue purified on a silica gel column, eluted with AcOEt/CH_3OH (8:2, v/v). Product **4** was obtained in quantitative yields (55 mg, 0.097 mmol).

4: yellow solid; R_f= 0.3 [AcOEt/CH_3OH, 95:5 v/v].

^{1}H-NMR (CD_3OD, 200 MHz): δ 8.60 (2H, d, *J* = 8.6 Hz, 2x H_α Py); 8.27-7.23 (9H, overlapped signals, protons of *Dansyl* group, H-6 and, 2x H_β Py); 6.14 (1H, d, *J* = 7.8 Hz, H-1'); 5.21 (1H,

d, *J* = 8.2 Hz, H-5); 5.0 (2H, m, C*H*$_2$Py); 4.76 (1H, dd, *J* = 4.6 Hz and 4.8 Hz, H-2'); 4.38 (1H, d, *J* = 4.6 Hz, H-3'); 4.03 (1H, bs, H-4'); 3.63 (2H, m, H-5'$_a$ and H-5'$_b$); 2.87 (6H, s, N(C*H*$_3$)$_2$).

^{13}C-NMR (CD_3OD, 50 MHz): δ 161.0 (C-4); 156.5, 133.4, 132.6, 131.2, 130.7, 129.8, 126.3, 120.1, 116.8 (aromatic carbons of *Dansyl* group); 153.1 (C-2); 151.9 (C$_\alpha$ Py); 144.5 (C$_\gamma$ Py); 139.6 (C-6); 124.3 (C$_\beta$ Py); 102.6 (C-5); 88.1 (C-1'); 86.2 (C-4'); 81.1 (C-2'); 71.6 (C-3'); 62.6 (C-5'); 45.7 [N(*C*H$_3$)$_2$]; 44.6 (*C*H$_2$Py).

ESI-MS (positive ions): for $C_{27}H_{28}N_4O_8S$, calculated 568.5982; found m/z: 568.86 [M+H$^+$]; 590.78 [M+Na$^+$]; 606.68 [M+K$^+$].

Synthesis of compound 5.

4 (51 mg, 0.093 mmol), dissolved in anhydrous DMF (0.5 mL), was reacted with TMBDMSCl (17 mg, 0.112 mmol) and imidazole (13 mg, 0.186 mmol). The solution was stirred for 2 h at room temperature; then the solvent was removed under reduced pressure and the residue purified on a silica gel column, eluted with AcOEt/CH_3OH (95:5, v/v). Compound **5** was obtained in 69% isolated yield (55 mg, 0.064 mmol).

5: yellow solid; R_f = 0.6 [AcOEt/CH_3OH, 95:5 v/v].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 8.56 (1H, d, *J* = 9.0 Hz, H-6); 8.45 (2H, d, *J* = 6.0 Hz, 2x H$_\alpha$Py); 8.16-7.13 (8H, overlapped signals, protons of *Dansyl* group, 2x H$_\beta$Py); 6.21 (1H, d, *J* = 6.0 Hz, H-1'); 5.19 (1H, d, *J* = 8.0 Hz, H-5); 4.82-4.65 (2H, m, C*H*$_2$Py); 4.55 (1H, m, H-2'); 4.37 (1H, m, H-3'); 4.13 (1H, bs, H-4'); 3.73 (2H, m, H-5'$_a$ and H-5'$_b$); 2.84 (6H, s, N(C*H*$_3$)$_2$); 0.84 (SiC(C*H*$_3$)$_3$); 0.02 and -0.02 (Si(C*H*$_3$)$_2$).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ 161.0 (C-4), 150.2, 132.4, 131.0, 129.9, 129.7, 129.4, 128.9 (aromatic carbons of *Dansyl* group); 151.9 (C-2); 149.2 (C$_\alpha$ Py); 145.6 (C$_\gamma$ Py); 136.9 (C-6); 123.5 (C$_\beta$ Py); 101.8 (C-5); 85.3 (C-1'); 84.6 (C-4'); 85.3 (C-2'); 70.5 (C-3'); 63.4 (C-5'); 45.2 [N(*C*H$_3$)$_2$]; 43.1 (*C*H$_2$Py); 25.9 (SiC(*C*H$_3$)$_3$); 18.3 (Si*C*(CH$_3$)$_3$) -5.6 (Si(*C*H$_3$)$_2$).

ESI-MS (positive ions): for $C_{33}H_{42}N_4O_8SSi$, calculated 682.8591; found m/z: 682.84 [M+H$^+$]; 704.79 [M+Na$^+$].

Synthesis of compound 6.

5 (40 mg, 0.059 mmol), dissolved in anhydrous CH_2Cl_2 (1.0 mL), was treated with oleic acid (24 μL, 0.077 mmol), DMAP (3.6 mg, 0.030 mmol) and DCC (24 mg, 0.12 mmol). The solution was stirred for 1 h at room temperature; then the solvent was removed under reduced pressure and the residue purified on a silica gel column, eluted with *n*-hexane/AcOEt (2:8, v/v). Nucleolipid **6** was obtained in almost quantitative yields (56 mg, 0.059 mmol).

6: yellow solid; R_f = 0.8 [AcOEt].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 8.60 (1H, d, *J* = 8.4 Hz, H-6); 8.47 (2H, d, *J* = 5.8 Hz, 2x H_αPy); 8.18-7.11 (8H, overlapped signals, protons of *Dansyl* group, 2x H_βPy); 6.22 (1H, d, *J* = 7.2 Hz, H-1'); 5.35-5.24 (4H, overlapped signals, H-5, H-3', H-9 and H-10 oleic acid); 4.87 (2H, m, C*H*$_2$Py); 4.78 (1H, dd, *J* = 5.6 and 5.4 Hz, H-2'); 4.09 (1H, bs, H-4'); 3.78 (2H, m, H-5'$_a$ and H-5'$_b$); 2.87 (6H, s, N(C*H*$_3$)$_2$); 2.42 (2H, t, *J* = 7.4 and 7.4 Hz, C*H*$_2$C=O); 2.00-1.25 (20H, overlapped signals, aliphatic protons of oleic acid); 0.92-0.83 (12H, overlapped signals, SiC(C*H*$_3$)$_3$ and C*H*$_3$ oleic acid); 0.09 and 0.04 (Si(C*H*$_3$)$_2$).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ 169.3 (C=O); 161.0 (C-4), 149.6, 139.5, 136.3, 132.1, 130.6, 130.1, 129.7, 129.6, 129.5, 128.5, 122.6, 118.8, 115.3, 109.4 (aromatic carbons of *Dansyl* group); 151.7 (C-2); 150.1 (C_αPy); 144.7 (C_γPy); 136.3 (C-6); 123.1 (C_βPy); 102.0 (C-5); 84.7 (C-1'); 83.8 (C-4'); 76.3 (C-2'); 70.8 (C-3'); 62.8 (C-5'); 45.1 [N(*C*H$_3$)$_2$]; 43.0 (*C*H$_2$Py); 34.7, 33.7, 32.5, 31.7, 30.7, 29.4, 29.1, 28.9, 28.8, 27.0, 26.2, 25.4, 24.7, 24.0, 22.4 (aliphatic carbons of oleic acid); 25.7 (SiC(*C*H$_3$)$_3$); 18.2 (Si*C*(CH$_3$)$_3$); 14.5 (*C*H$_3$ oleic acid); -5.8 and -5.8 (Si(*C*H$_3$)$_2$).

ESI-MS (positive ions): for $C_{51}H_{74}N_4O_9SSi$, calculated 947.3052; found m/z: 947.12 [M+H$^+$]; 969.08 [M+Na$^+$]; 985.14 [M+K$^+$].

Synthesis of compound 7.

6 (50 mg, 0.053 mmol), dissolved in anhydrous THF (1.0 mL), was reacted with $Et_3N{\cdot}3HF$ (34 μL, 0.212 mmol). The solution was stirred for 12 h at room temperature; then the solvent was removed under reduced pressure and the residue purified on a silica gel column, eluted with *n*-hexane/AcOEt (2:8, v/v). Nucleolipid **7** was obtained in 68% isolated yield (30 mg, 0.036 mmol).

7: yellow solid; R_f = 0.6 [*n*-hexane/acetone, 1:1 v/v].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 8.60 (1H, d, *J* = 8.6 Hz, H-6); 8.52 (2H, bs, 2x H_αPy); 8.18-7.13 (8H, overlapped signals, protons of *Dansyl* group, 2xH_βPy); 5.93 (1H, d, *J* = 6.8 Hz, H-1'); 5.43-5.33 (4H, overlapped signals, H-5, H-3', H-9 and H-10 oleic acid); 5.18 (1H, dd, *J* = 7.0 and 5.6 Hz, H-2'); 4.81 (2H, m, C*H*$_2$Py); 4.10 (1H, bs, H-4'); 3.77 (2H, bs, H-5'$_a$ and H-5'$_b$); 2.89 (6H, s, N(C*H*$_3$)$_2$); 2.27 (2H, t, *J* = 7.4 and 7.4 Hz, C*H*$_2$C=O); 2.0-1.27 (20H, overlapped signals, aliphatic protons of oleic acid); 0.89 (3H, m, C*H*$_3$ oleic acid).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 172.6 (C=O); 161.3 (C-4), 150.4, 133.8, 132.3, 131.0, 130.0, 129.7, 128.8, 123.0, 118.8, 115.5 (aromatic carbons of *Dansyl* group); 152.0 (C-2); 149.0 (C_αPy); 146.0 (C_γPy); 138.8 (C-6); 123.7 (C_βPy); 102.4 (C-5); 88.6 (C-1'); 84.3 (C-4'); 76.5 (C-2'); 71.0 (C-3'); 61.7 (C-5'); 45.3 [N(*C*H$_3$)$_2$]; 43.1 (*C*H$_2$Py); 33.9, 31.9, 29.7, 29.5, 29.3, 29.1, 27.2, 25.6, 24.9, 24.7, 22.7 (aliphatic carbons of oleic acid); 14.1 (*C*H$_3$ oleic acid).

ESI-MS (positive ions): for $C_{45}H_{60}N_4O_9S$, calculated 833.0443; found m/z: 833.08 [M+H$^+$]; 855.07 [M+Na$^+$]; 870.98 [M+K$^+$].

Synthesis of compound 8.

7 (28 mg, 0.034 mmol), dissolved in anhydrous CH_2Cl_2 (0.8 mL), was reacted with BnO(CH_2CH_2O)$_6$$CH_2$COOH (19 mg, 0.044 mmol), DMAP (2.0 mg, 0.017 mmol) and DCC (14 mg, 0.068 mol). The solution was stirred for 12 h at room temperature; then the solvent was removed under reduced pressure and the residue purified on a silica gel column, eluted with *n*-hexane/acetone (1:1, v/v). The product **8** was obtained in almost quantitative yields (42 mg, 0.034 mmol).

8: yellow solid; R_f = 0.4 [*n*-hexane/acetone, 1:1 v/v].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 8.54 (1H, d, *J* = 8.4 Hz, H-6); 8.48 (2H, d, *J* = 4.8 Hz, 2x H_αPy); 8.11-6.92 (13H, overlapped signals, protons of *Dansyl* group, 2x H_βPy and protons of Bn); 5.80 (1H, d, *J* = 5.4 Hz, H-1'); 5.50 (1H, d, *J* = 8.2 Hz, H-5); 5.30 (2H, m, H-10 and H-9 oleic acid); 5.17 (1H, dd, *J* = 6.2 and 4.0 Hz, H-3'); 5.08 (1H, dd, *J* = 5.6 and 5.8 Hz, H-2'); 4.82 (2H, m, C*H*$_2$Py); 4.26-4.12 (5H, overlapped signals, H-4', C*H*$_2$C=O, H-5'$_a$ and H-5'$_b$); 4.50 (2H, s, C*H*$_2$Bn); 3.69-3.58 (24H, overlapped signals, OC*H*$_2$C*H*$_2$O), 2.83 (6H, s, N(C*H*$_3$)$_2$); 2.14-1.27 (22H, overlapped signals, aliphatic protons of oleic acid); 0.82 (3H, m, C*H*$_3$ oleic acid).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 172.6 and 169.4 (2xC=O); 161.3 (C-4), 150.1, 138.5, 132.4, 130.7, 130.5, 130.0, 129.8, 129.6, 129.6, 128.9, 128.3, 127.7, 127.5, 122.9, 118.7, 115.5 (aromatic carbons of *Dansyl* group and Bn); 152.1 (C-2); 149.1 (C_αPy); 145.8 (C_γPy); 138.2 (C-6); 123.5 (C_βPy); 102.4 (C-5); 89.3 (C-1'); 80.1 (C-4'); 76.7 (C-2'); 73.1 (*C*H_2Bn); 71.0 (C-3'); 70.5, 69.6, 69.4 (O*C*H_2*C*H_2O); 63.1 (C-5'); 45.3 [N(*C*H_3)$_2$]; 43.2 (*C*H_2Py); 33.9, 33.5, 32.6, 31.8, 30.6, 29.7, 29.6, 29.4, 29.2, 29.0, 28.9, 27.1, 26.2, 25.6, 24.9, 22.6 (aliphatic carbons of oleic acid); 14.0 (*C*H_3 oleic acid).

ESI-MS (positive ions): for $C_{66}H_{92}N_4O_{17}S$, calculated 1245.5183; found m/z: 1245.31 [$M+H^+$]; 1267.18 [$M+Na^+$].

Synthesis of Ru complex 9.

Nucleolipid **8** (31 mg, 0.025 mmol), dissolved in 0.8 mL of anhydrous DCM, was treated with $[RuCl_4(DMSO)_2]^-Na^+$ (10.6 mg, 0.025 mmol) and the resulting solution was taken under stirring at 40 °C for 1 h. TLC monitoring on alumina showed in 1 h the total disappearance of the starting material; the reaction mixture was therefore taken to dryness, giving the desired product **9** in quantitative yields, obtained in a pure form, not requiring further purification (40 mg, 0.025 mmol).

^{1}H-NMR ($CDCl_3$, 500 MHz): significant signals at *δ* 7.36-7.26 (aromatic proton of *Bn* and *Dansyl* group); 5.38-3.75 (very broad signals: protons of ribose, H-5 and polyether); 2.04-0.90 (overlapped signals, protons of oleic acid); -1.85 (very broad signal, protons of Py); -12.49 (very broad signal, 2x C*H*$_3$ DMSO).

^{13}C-NMR ($CDCl_3$, 50 MHz): significant signals at *δ* 129.2; 128.9; 127.7; 127.0; 71.2; 48.3; 33.3; 28.9; 28.6; 28.4; 26.4; 24.8; 24.2; 13.3.

ESI-MS (negative ions): for $C_{68}H_{98}N_4O_{18}RuS_2$, calculated -1566.5337; found m/z: -1487.78 [M^--DMSO]; 1566.76 [M^-].

Synthesis of Ruthenium complex 21

Synthesis of compound 10.

1b (870 mg, 1.10 mmol) was dissolved in Py (3.0 mL) and Ac_2O (1.5 mL)**.** The solution was stirred at room temperature for 12 h, then the reaction was quenched with 2 mL of MeOH. The crude was washed with DCM/H_2O; the collected organic phase was dried over anhydrous Na_2SO_4, concentrated under reduced pressure to give compound **10** in 98% isolated yield (898 mg, 1.08 mmol).

10: white solid, R_f = 0.4 [*n*-hexane/AcOEt, 1:1 v/v].

^{1}H-NMR ($CDCl_3$, 500 MHz): δ 7.60 (1H, d, *J* = 8.5 Hz, H-6); 7.48-6.71 (28H, overlapped signals, aromatic protons of *MMTr*); 6.55 (1H, d, *J* = 7.5 Hz, H-1'); 5.12 (1H, d, *J* = 8.5 Hz, H-5); 4.48 (1H, dd, *J* = 8.0 Hz, H-2'); 3.83 (2H, bs, H-3' and H-4'); 3.79 (3H, s, OCH_3); 3.78 (3H, s, OCH_3); 3.20 (2H, ABX system, *J* = 11.0 and 1.5 Hz, H-5'$_a$ and H-5'$_b$); 1.99 (3H, s, CH_3C=O).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 169. 7 (*C*=O); 163.6 (C-4); 159.7, 159.4, 149.5, 144.9, 144.4, 144.3, 143.9, 135.9, 135.1, 131.1, 130.6, 129.3, 128.8, 128.5, 128.2, 128.0, 127.5, 125.7, 123.8, 113.9, 113.7 (aromatic carbons of *MMTr*); 151.6 (C-2); 140.3 (C-6); 103.2 (C-5); 87.9 and 87.7 (quaternary carbons of *MMTr*); 86.8 (C-1'); 84.4 (C-4'); 75.9 (C-3'); 72.8 (C-2'); 64.2 (C-5'); 54.9 (O*C*H_3); 20.5 (*C*H_3C=O).

ESI-MS (positive ions): for $C_{51}H_{46}N_2O_9$, calculated 830.32; found m/z: 854.10 [M+Na^+]; 870.09 [M+K^+].

Synthesis of compound 11.

10 (889 mg, 1.07 mmol), dissolved in 6 mL of anhydrous CH_2Cl_2, was treated with TEA (448 µL, 3.21 mmol), Boc_2O (467 mg, 2.14 mmol) and DMAP (65 mg, 0.535 mmol). The solution was stirred at room temperature for 1 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (6:4, v/v + a few drops of TEA). The desired product **11** was obtained in 80% isolated yield (793 mg, 0.856 mmol).

11: white solid, R_f = 0.7 [*n*-hexane/AcOEt 1:1 v/v]

^{1}H-NMR ($CDCl_3$, 500 MHz): δ 7.57 (1H, d, *J* = 8.5 Hz, H-6); 7.50-6.72 (28H, overlapped signals, aromatic proton of *MMTr*); 6.52 (1H, d, *J* = 8.0 Hz, H-1'); 5.06 (1H, d, *J* = 8.0 Hz, H-

5); 4.42 (1H, dd, *J* = 6.0 and 5.5 Hz, H-2'); 3.80 (2H, bs, H-3' and H-4'); 3.79 (3H, s, OC*H*$_3$); 3.77 (3H, s, OC*H*$_3$); 3.19 (2H, ABX system, *J* = 9.5 and 2 Hz, H-5'$_a$ and H-5'$_b$); 2.00 (3H, s, C*H*$_3$C=O); 1.60 (9H, s, 3x C*H*$_3$ Boc).

^{13}C-NMR (CDCl$_3$, 50 MHz): δ 169. 9 (*C*=O); 160.2 (C-4); 159.1, 158.7, 143.9, 143.6, 143.5, 143.0, 134.5, 134.4, 130.4, 130.1, 128.2, 128.1, 128.0, 127.9, 127.8, 127.3, 113.4, 113.2 (aromatic carbons of *MMTr*); 148.9 (C-2); 147.6 (*C*=O Boc); 139.5 (C-6); 102.1 (C-5); 87.6 (C-1'); 87.0 (quaternary carbon of *MMTr*); 86.6 (*C*(CH$_3$)$_3$ Boc); 84.3 (C-4'); 75.6 (C-2'); 72.5 (C-3'); 63.7 (C-5'); 55.2 (O*C*H$_3$); 27.3 (C(*C*H$_3$)$_3$ Boc); 20.5 (*C*H$_3$C=O).

ESI-MS (positive ions): for $C_{56}H_{54}N_2O_{11}$, calculated 930.37; found m/z: 954.11 [M+Na$^+$]; 970.30 [M+K$^+$].

Synthesis of compound 12.

11 (965 mg, 1.036 mmol) was dissolved in 3 mL of MeOH and 1 mL of TEA. K_2CO_3 (29 mg, 0.207 mmol) was added and the resulting solution was stirred at room temperature for 12 h. Then the solvent was removed under reduced pressure and the crude was washed with DCM/H_2O; the collected organic phase was dried over anhydrous Na_2SO_4, concentrated under reduced pressure to give the compound **12** in almost quantitative yields (921 mg, 1.035 mmol).

12: white solid, R_f = 0.7 [*n*-hexane/AcOEt, 1:1 v/v].

^{1}H-NMR (CDCl$_3$, 200 MHz): δ 7.79 (1H, d, *J* = 8.2 Hz, H-6); 7.60-6.62 (28H, overlapped signals, aromatic protons of *MMTr*); 6.00 (1H, d, *J* = 5.4 Hz, H-1'); 5.11 (1H, d, *J* = 8.2 Hz, H-5); 4.50 (1H, m, H-2'); 4.10 (1H, bs, H-3'); 3.81 (6H, overlapped singlets, 2x OC*H*$_3$); 3.03 (3H, m, H-4', H-5'$_a$ and H-5'$_b$); 1.65 (9H, s, 3x C*H*$_3$ Boc).

^{13}C-NMR (CDCl$_3$, 50 MHz): δ 160.3 (C-4); 159.2, 158.7, 143.9, 143.4, 143.3, 143.0, 134.3, 134.1, 130.3, 130.1, 128.2, 128.0, 127.8, 127.4, 127.3, 113.7, 113.1 (aromatic carbons of *MMTr*); 149.0 (C-2); 147.6 (*C*=O Boc); 140.0 (C-6); 102.0 (C-5); 87.6 (C-1'); 86.6 (quaternary carbons of *MMTr* and *C*(CH$_3$)$_3$ Boc); 84.6 (C-4'); 78.0 (C-2'); 70.7 (C-3'); 64.2 (C-5'); 55.1 (O*C*H$_3$); 27.3 (C(*C*H$_3$)$_3$ Boc).

ESI-MS (positive ions): for $C_{54}H_{52}N_2O_{10}$, calculated 888.36; found m/z: 911.2 [M+Na$^+$]; 927.3 [M+K$^+$].

Synthesis of compound 13.

12 (464 mg, 0.520 mmol), dissolved in 5 mL of anhydrous THF, was reacted with PPh_3 (410 mmg, 1.56 mmol) and DIAD (307 µL, 1.56 mmol). The resulting solution was stirred at room temperature for 5 minutes, then DPPA (336 µL, 1.560 mmol) was added and the reaction mixture stirred overnight at room temperature. The solvent was then removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (75:25, v/v + TEA). The desired product **13** was obtained in 92% isolated yield (440 mg, 0.481 mmol).

13: white solid, R_f = 0.7 [*n*-hexane/AcOEt, 3:2 v/v].

^{1}H-NMR ($CDCl_3$, 500 MHz): δ 7.48 (1H, d, *J* = 8.0 Hz, H-6); 7.34-6.81 (28H, overlapped signals, aromatic protons of *MMTr*); 6.4 (1H, bs, H-1'); 5.61 (1H, d, *J* = 8.0 Hz, H-5); 4.10 (1H, m, H-2'); 3.94 (1H, bs, H-4'); 3.79 (6H, overlapped singlets, 2x O*C*H_3); 3.52 (1H, m, H-5'$_a$); 3.07 (1H, m, H-5'$_b$), 2.63 (1H, d, *J* = 3.0Hz, H-3');1.64 (9H, s, 3x CH_3 Boc).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 159.8 (C-4); 159.2, 158.6, 143.9, 143.6, 143.2, 134.7, 134.5, 130.4, 130.1, 127.9, 127.7, 127.4, 126.9, 113.5, 112.9 (aromatic carbons of *MMTr*); 148.3 (C-2); 147.4 (*C*=O Boc); 140.0 (C-6); 102.6 (C-5); 90.5 (C-1'); 88.7 (C-4'); 86.8 and 86.6 (quaternary carbons of *MMTr*); 83.4 (*C*$(CH_3)_3$ Boc); 78.9 (C-2'); 66.6 (C-3'); 60.9 (C-5'); 54.9 (O*C*H_3); 27.3 (C(*C*$H_3)_3$ Boc).

ESI-MS (positive ions): for $C_{54}H_{51}N_5O_9$, calculated 913.37; found m/z: 814.90 [M-Boc+H^+].

FT-IR: 2112.9 (N_3 stretching).

Synthesis of compound 14.

13 (440 mg, 0.481 mmol), dissolved in 2.7 mL of anhydrous CH_2Cl_2, was reacted with 0.15 mL of TFA and 0.15 mL of TIS. The resulting solution was stirred at room temperature for 1 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (1:4, v/v). The desired product **14** was obtained in 97% isolated yield (126 mg, 0.468 mmol).

14: white solid, R_f= 0.1 [*n*-hexane/AcOEt 3:7 v/v]

^{1}H-NMR (CD_3OD, 500 MHz): δ 7.86 (1H, d, *J* = 8.0 Hz, H-6); 5.79 (1H, d, *J* = 3.0 Hz, H-1'); 5.71 (1H, d, *J* = 8.0 Hz, H-5); 4.37 (1H, m, H-4'); 4.35 (1H, dd, *J* = 3.5 and 3.0 Hz, H-2'); 4.18 (1H, dd, *J* = 4.0 and 3.0, H-3'); 3.85 (1H, d, *J* = 4.5 Hz, H-5'$_a$ and H-5'$_b$).

^{13}C-NMR (CD_3OD, 125 MHz): 166.8 (C-4); 152.9 (C-2); 142,8 (C-6); 102.9 (C-5); 92.2 (C-1'); 82.9 (C-4'); 80.3 (C-2'); 68.3 (C-3'); 61.7 (C-5').

ESI-MS (positive ions): for $C_9H_{11}N_5O_5$, calculated 269.08; found m/z: 270.0 [M+H$^+$]; 292.0 [M+Na$^+$].

Synthesis of compound 15.

14 (197 mg, 0.732 mmol), dissolved in 1 mL of anhydrous DMF, was reacted with imidazole (119 mg, 1.76 mmol) and TBDMSCl (132 mg, 0.878 mmol). The solution was stirred at room temperature for 1 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (1:1, v/v). The desired product **15** was obtained in 80% isolated yield (131 mg, 0.586 mmol).

15: oil, *Rf* = 0.5 [*n*-hexane/AcOEt, 1:1 v/v].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 7.73 (1H, d, *J* = 8.4 Hz, H-6); 5.79 (1H, d, *J* = 1.4 Hz, H-1'); 5.71 (1H, d, *J* = 8.0 Hz, H-5); 4.45 (overlapped signals, H-2' and H-4'); 4.20 (1H, m, H-3'); 3.93 (2H, d, *J* = 5.2 Hz, H-5'$_a$ and H-5'$_b$), 0.921 (9H, s, SiC(CH_3)$_3$), 0.12 (6H, s, Si(CH_3)$_2$).

^{13}C-NMR ($CDCl_3$, 50 MHz): 164.0 (C-4); 151.2 (C-2); 140.3 (C-6); 101.8 (C-5); 91.8 (C-1'); 81.7 (C-4'); 79.4 (C-2'); 65.9 (C-3'); 61.0 (C-5'); 25.8 (SiC(*C*H$_3$)$_3$); 18.2 (Si*C*(CH$_3$)$_3$); -5.5 and -5.6 (Si(*C*H$_3$)$_2$).

ESI-MS (positive ions): for $C_{15}H_{25}N_5O_5Si$, calculated 383.4750; found m/z: 384.14 [M+H$^+$]; 406.06 [M+Na$^+$]; 422.00 [M+K$^+$].

Synthesis of compound 16.

15 (100 mg, 0.261 mmol), dissolved in 2 mL of anhydrous DCM, was reacted with DMAP (16 mg, 0.131 mmol), oleic acid (88 mg, 0.316 mmol) and DCC (81 mg, 0.392 mmol). The solution was stirred at room temperature for 2 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (7:3, v/v). The desired product **16** was obtained in 74% isolated yield (125 mg, 0.193 mmol).

16: oil, *Rf* = 0.7 [*n*-hexane/AcOEt, 1:1 v/v].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 7.70 (1H, d, *J* = 8.2 Hz, H-6); 6.05 (1H, d, *J* = 3.8 Hz, H-1'); 5.76 (1H, dd, *J* = 1.4 and 8.2 Hz, H-5); 5.33 (2H, m, H-9 and H-10 oleic acid); 4.4 (1H, t, *J* = 3.2 and 3.2 Hz, H-2'); 4.22 (2H, m, H-3' and H-4'); 3.91 (2H, d, *J* = 4.8 Hz, H-5'$_a$ and H-5'$_b$), 2.5 (2H, dd, *J* = 6.8 and 7.8 Hz, 2x H-2 oleic acid); 2.02 (4H, m, 2x H-8 and 2x H-11 oleic acid); 1.63 (2H, m, H-3 oleic acid), 1.28 (20H, overlapped signals of oleic acid); 0.92 (9H, s, SiC(CH_3)$_3$), 0.87 (3H, t, *J* = 6.6 Hz, CH_3 oleic acid); 0.12 (6H, s, Si(CH_3)$_2$).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 172.4 (C=O), 163.1 (C-4); 150.3 (C-2); 139.8 (C-6); 130.0 and 129.6 (C-9 and C-10 oleic acid); 102.9 (C-5); 87.2 (C-1'); 80.2 (C-4'); 79.5 (C-2'); 64.6 (C-3'); 61.0 (C-5'); 33.8, 31.8, 29.7, 29.6, 29.4, 29.2, 29.0, 28.9, 27.1, 24.6, 22.6 (aliphatic carbons of oleic acid); 25.8 (SiC(*C*H_3)$_3$); 18.2 (Si*C*(CH$_3$)$_3$); 14.0 (-*C*H_3 oleic acid); -5.5 and -5.7 (Si(*C*H_3)$_2$).

ESI-MS (positive ions): for $C_{33}H_{57}N_5O_6Si$, calculated 647.9211; found m/z: 648.12 [M+H^+]; 670.01 [M+Na^+]; 686.07 [M+K^+].

Synthesis of compound 17.

16 (31 mg, 0.048 mmol) was dissolved in anhydrous AcOEt (0.7 mL) and then Pd/C 10% p.p. (10 mg, 0.0.01 mmol) was added. The solution was stirred at room temperature for 12 h under pressure of H_2 (1 atm), then the solvent was removed under reduced pressure and the crude was filtered on a short column of silica gel, eluted with AcOEt. The desired product **17** was obtained in almost quantitative yields (30 mg, 0.048 mmol).

17: oil, *Rf* = 0.2 [*n*-hexane/AcOEt, 3:2 v/v].

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 8.04 (1H, d, *J* = 8.4 Hz, H-6); 6.03 (1H, d, *J*= 4.8 Hz, H-1'); 5.72 (1H, d, *J* = 8.4 Hz, H-5); 5.35 (2H, m, H-9 and H-10 oleic acid); 5.0 (1H, t, *J* = 5.2 and 5.2 Hz, H-2'); 4.20 (2H, m, H-4'); 3.93 (2H, m, H-5'$_a$ and H-5'$_b$), 3.68 (1H, dd, *J* = 6.4 and 5.8 Hz, H-3'); 2.36 (2H, t, *J* = 7.4 Hz, 2x H-2 oleic acid); 2.02 (4H, m, 2x H-8 and 2x H-11 oleic acid); 1.58 (2H, m, H-3 oleic acid), 1.27 (20H, overlapped signals oleic acid); 0.93 (12H, overlapped signals, CH_3 oleic acid and SiC(CH_3)$_3$), 0.14 (6H, s, Si(CH_3)$_2$).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 173.2 (C=O), 162.7 (C-4); 150.2 (C-2); 140.8 (C-6); 130.0 and 129.6 (C-9 and C-10 oleic acid); 102.4 (C-5); 86.2 (C-1'); 81.5 (C-4'); 80.2 (C-2'); 61.9 (C-5'); 56.6 (C-3'); 33.9, 31.8, 29.6, 29.4, 29.2, 29.0, 27.1, 24.6, 22.5 (aliphatic carbons of oleic acid); 25.8 (SiC(*C*H_3)$_3$); 18.0 (Si*C*(CH$_3$)$_3$); 14.0 (-*C*H_3 oleic acid); -5.7 (Si(*C*H_3)$_2$).

ESI-MS (positive ions): for $C_{33}H_{59}N_3O_6Si$, calculated 621.9236; found m/z: 622.15 [M+H^+]; 644.07 [M+Na^+]; 660.13 [M+K^+].

Synthesis of compound 18.

Nucleolipid **17** (48 mg, 0.077 mmol), dissolved in 1 mL of anhydrous DCM, was treated with DMAP (28 mg, 0.231 mmol), 2-(pyridin-4-yl)acetic acid hydrochloride (20 mg, 0.116 mmol) and DCC (40 mg, 0.193 mmol). The solution was stirred at room temperature for 12 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (2:8, v/v). The desired product **18** was obtained in 99% isolated yield (56 mg, 0.076 mmol).

^{1}H-NMR (CD_3OD, 200 MHz): δ 8.47 (2H, d, *J* = 4.4 Hz, 2x H_αPy); 7.82 (1H, d, *J* = 8.0 Hz, H-6); 7.39 (2H, d, *J* = 5.8 Hz, 2x H_βPy); 5.81 (1H, d, *J* = 5.4 Hz, H-1'); 5.70 (1H, d, *J* = 8.2 Hz, H-5); 5.35 (2H, m, H-9 and H-10 oleic acid); 4.72 (1H, dd, *J* = 6.2 and 7.2 Hz, H-2'); 4.35 (1H, m, H-3'); 3.93 (5H, overlapped signals, H-5'$_a$ and H-5'$_b$, H-4', CH_2Py), 2.36 (2H, dd, *J* = 7.4 and 7.0 Hz, 2x H-2 oleic acid); 2.02 (4H, m, 2x H-8 and 2x H-11 oleic acid); 1.60 (2H, m, H-3 oleic acid), 1.29 (20H, overlapped signals oleic acid); 0.90 (12H, overlapped signals, CH_3 oleic acid and SiC(CH_3)$_3$), 0.07 (6H, s, Si(CH_3)$_2$).

^{13}C-NMR (CD_3OD, 50 MHz): δ 174.7 (C=O oleyl), 171.8 (C=O amide); 165.8 (C-4); 152.2 (C-2); 150.2 (C_αPy); 146.8 (C_γPy); 143.2 (C-6); 130.9 (C_βPy); 126.2 (C-9 and C-10 oleic acid); 103.2 (C-5); 89.9 (C-1'); 80.6 (C-4'); 79.4 (C-2'); 63.2 (C-5'); 55.7 (C-3'); 42.8 (*C*H_2Py); 34.7, 34.6, 33.0, 30.8, 30.6, 30.4, 30.3, 30.2, 30.1, 28.1, 26.7, 26.0, 25.7, 23.7, 19.2 (aliphatic carbons of oleic acid); 26.5 (SiC(*C*H_3)$_3$); 18.0 (Si*C*(CH$_3$)$_3$); 14.5 (-*C*H_3 oleic acid); -5.2 and -5.3 (Si(*C*H_3)$_2$).

ESI-MS (positive ions): for $C_{40}H_{64}N_4O_7Si$, calculated 741.0443; found m/z: 741.19 [M+H^+]; 763.17 [M+Na^+]; 768.02 [M+K^+].

Synthesis of compound 19.

Nucleolipid **18** (56 mg, 0.076 mmol), dissolved in 1 mL of anhydrous THF, was treated with 50 µL of Et_3N·3xHF (0.30 mmol). The solution was stirred at room temperature for 3 h, then the solvent was removed under reduced pressure and the crude was purified on a

silica gel column, eluted with AcOEt/CH_3OH (95:5, v/v). The desired product **19** was obtained in 86% isolated yield (41 mg, 0.065 mmol).

^{1}H-NMR ($CDCl_3$, 200 MHz): δ 8.55 (2H, d, *J* = 4.4 Hz, 2x H_αPy); 7.68 (1H, d, *J* = 8.0 Hz, H-6); 7.29 (2H, d, *J* = 6.2 Hz, 2x H_βPy); 5.71 (1H, d, *J* = 7.8 Hz, H-5); 5.45-5.24 (5H, overlapped signals, H-1', CH_2Py, H-9 and H-10 oleic acid); 4.75 (1H, bs, H-2'); 4.23 (1H, bs, H-3'); 3.64 (3H, bs, H-5'$_a$ and H-5'$_b$, H-4'); 2.34 (2H, t, *J* = 7.4 and 7.4 Hz, CH_2C=O); 1.99-1.27 (26H, overlapped signals oleic acid); 0.89 (3H, m, CH_3 oleic acid).

^{13}C-NMR ($CDCl_3$, 50 MHz): δ 173.4 (C=O oleyl), 170.3 (C=O amide); 163.1 (C-4); 150.3 (C-2); 149.7 (C_αPy); 144.0 (C_γPy); 143.3 (C-6); 130.1 (C_βPy); 129.6 (C-9 and C-10 oleic acid); 103.2 (C-5); 93.5 (C-1'); 79.5 (C-4'); 78.6 (C-2'); 60.2 (C-5'); 54.9 (C-3'); 42.8 (*C*H_2Py); 33.8, 31.9, 29.7, 29.3, 29.1, 27.2, 24.6, 22.7 (aliphatic carbons of oleic acid); 14.1 (-*C*H_3 oleic acid).

ESI-MS (positive ions): for $C_{34}H_{50}N_4O_7$, calculated 626.3679; found m/z: 627.38 [M+H^+].

Synthesis of compound 20.

19 (17 mg, 0.027 mmol), dissolved in 0.7 mL of anhydrous DCM, was reacted with DMAP (2.0 mg, 0.013), BnO(CH_2CH_2O)$_6$$CH_2$COOH (15 mg, 0.035 mmol) and DCC (8.0 mg, 0.040 mmol). The solution was stirred at room temperature for 2 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/acetone (2:3, v/v). The desired product **20** was obtained in 70% isolated yield (20 mg, 0.019 mmol).

^{1}H-NMR ($CDCl_3$, 500 MHz): δ 8.65 (2H, bs, 2x H_αPy); 7.37-7.25 (7H, overlapped signals, aromatic proton of Bn, H-6, 2x H_βPy); 5.75 (1H, d, *J* = 8.0 Hz, H-5); 5.30-5.27 (5H, overlapped signals, H-1', CH_2Py, H-9 and H-10 oleic acid); 4.76 (1H, bs, H-2'); 4.55 (2H, s, CH_2Bn); 4.39 (1H, bs, H-3'); 4.18-4.30 (5H, bs, H-5'$_a$ and H-5'$_b$, H-4', CH_2C=O); 3.77-3.63 (24H, OCH_2CH_2O); 2.33 (2H, dd, *J* = 8.0 and 7.0 Hz, CH_2C=O); 1.93-1.08 (26H, overlapped signals oleic acid); 0.87 (3H, dd, *J* = 6.0 and 7.0 Hz, CH_3 oleic acid).

^{13}C-NMR ($CDCl_3$, 125 MHz): δ 173.2, 172.3 and 169.6 (C=O ester, C=O amide), 162.5 (C-4); 156.7 (C_αPy); 150.0 (C-2); 142.8 (C_γPy); 138.1 (quaternary carbon of Bn); 128.3, 127.6, 127.5 (C_βPy, C-9 and C-10 oleic acid); 103.2 (C-5); 93.3 (C-1'); 80.9 (C-4'); 79.6 (C-2'); 73.1 (*C*H_2Bn); 70.4, 70.3 (O*C*H_2*C*H_2O); 69.3 (*C*H_2C=O); 61.9 (C-5'); 54.7 (C-3'); 49.0 (*C*H_2Py); 33.8,

29.6, 29.3, 29.2, 29.1, 29.0, 25.5, 24.8, 24.5, 22.6 (aliphatic carbons of oleic acid); 14.0 (-*C*H_3 oleic acid).

ESI-MS (positive ions): for $C_{55}H_{82}N_4O_{15}$, calculated 1039.2574; found m/z: 1041.27 [M+H^+]; 1063.20 [M+Na^+]; 1079.24 [M+K^+].

Synthesis of compound 21.

Nucleolipid **20** (18 mg, 0.017 mmol), dissolved in 0.4 mL of anhydrous DCM, was treated with $[RuCl_4(DMSO)_2]^-Na^+$ (7.0 mg, 0.017 mmol) and the resulting solution was taken under stirring at 40 °C for 2 h. TLC monitoring on alumina showed in 4 h the total disappearance of the starting material; the reaction mixture was therefore taken to dryness, giving the desired product **21** in quantitative yields, obtained in a pure form, not requiring further purification (23 mg, 0.017 mmol).

^{1}H NMR ($CDCl_3$, 500 MHz): significant signals at *δ* 7.55-7.30 (protons of Bn and H-6); 4.96-3.86 (very broad signals: protons of ribose, H-5 and polyether); 2.7-0.9 (overlapped signals, protons of oleic acid); -1.98 (protons of Py); -9.59 (2x C*H*$_3$ DMSO).

^{13}C NMR ($CDCl_3$, 125 MHz): significant signals at *δ* 128.2, 127.9, 127.4 (carbons of Bn, C-9 and C-10 oleic acid); 49.3 (*C*H_2Py); 33.9, 31.3, 29.1, 28.8, 25.0, 24.7, 22.1 (aliphatic carbons of oleic acid); 13.5 (-*C*H_3 oleic acid).

ESI-MS (negative ions): for $C_{57}H_{88}N_4O_{16}RuS$, calculated -1360.2728; found m/z: -1284.37 [M^--DMSO]; -1359.89 [M^-].

7. References

1. M. Galanski, M.A. Jakupec, B. Keppler *Curr. Med. Chem.,* **2005**, *12*, 2075-2094.
2. M. Galanski, *Recent Pat. Anticancer Drug Discov.,* **2006,** *1*, 285-295.
3. V. Brabec, J. Kasparkova *Drug Resist. Updat,* **2005**, *8*, 131–146.
4. I. Ott, R. Gust *Arch Pharm (Weinheim)* **2007**, *340*, 117–126.
5. S. Ramaswamy, K. N. Ross, E. S. Lander, T. R. Golub *Nature Genetics*, **2003**, *33*, 49-54.
6. K. Pantel, R. H. Brakenhoff *Nature Reviews Cancer*, **2004**, *4*, 448-456.
7. A. A. Adjei, E. K. Rowinsky *Cancer Biology & Therapy*, **2003**, *2*, S5-S15.
8. E. Germanov, J. N. Berman, D. L. Guernsey *International Journal of Molecular Medicine*, **2006**, *18*, 1025-1036.
9. P. HeVeter, U. Jungwirth, M. Jakupec, C. Hartinger, M. Galanski *Drug Resist Updat,* **2008** *11*, 1–16.
10. G. Sava, S. Zorzet, T. Giraldi, G. Mestroni, G. Zassinovich *Eur. J. Cancer Clin. Oncol.*, **1984**, *20*, 841–847.
11. G. Sava, A. Bergamo *Int. J. Oncol.,* **2000**, *17,* 353–365.
12. P. Schluga, C. G. Hartinger, A. Egger, E. Reisner, M. Galanski *Dalton Trans*, **2006**, *14*, 1796–1802.
13. S. Rockwell, I. T. Dobrucki, E. Y. Kim, S. T. Marrison, V. T. Vu *Curr. Mol. Med.*, **2009**, *9*, 442–458.
14. A. Bergamo, M. Cocchietto, I. Capozzi, G. Mestroni, E. Alessio, G. Sava *Anti-Cancer Drugs*, **1996**, *7*, 697-702.
15. G. Sava *Metal Compounds in Cancer Therapy*, **1994**, 65-91.
16. G. Sava, S. Pacor, F. Bregant, V. Ceschia *Anticancer Res.*, **1991**, *11*, 1103-1107.
17. G. Sava, S. Pacor, F. Bregant, V. Ceschia, G. Mestroni *Anti-Cancer Drugs*, **1990**, *1*, 99-108.
18. G. Sava, S. Pacor, G. Mestroni, E. Alessio *Clinical & Experimental Metastasis*, **1992**, *10*, 273-280.
19. G. Sava, S. Pacor, S. Zorzet, E. Alessio, G. Mestroni *Pharmacol. Res.*, **1989**, *21*, 617-628.
20. E. S. Antonarakis, A. Emadi *Cancer Chemother. Pharmacol.,* **2010**, *66*, 1–9.

21. F. Lentz, A. Drescher, A. Lindauer, M. Henke, R.A. Hilger, C. G. Hartinger, M. E. Scheulen, C. Dittrich, B. K. Keppler, U. Jaehde *Anticancer Drugs*, **2009**, *20(2)*, 97-103
22. W.-H. Ang, A. Casini, G. Sava, P. J. Dyson *Journal of Organometallic Chemistry*, **2011**, *696*, 989-998.
23. M. Vaccaro, R. Del Litto, G. Mangiapia, A. M. Carnerup, G. D'Errico, F. Ruffo, L. Paduano *Chem. Commun.*, **2009**, 1404-1406.
24. P. Barthélémy *Comptes Rendus Chimie*, **2009**, *12*, 171-179.
25. D. Berti, F. B. Bombelli, M. Fortini, P. Baglioni *Journal of Physical Chemistry B*, **2007**, *111*, 11734-11744.
26. M. Fortini, D. Berti, P. Baglioni, B. W. Ninham *Current Opinion in Colloid & Interface Science*, **2004**, *9*, 168-172.
27. Alessio, E., Balducci, G., Calligaris, M., Costa, G., Attia, W. M., Mestroni, G. *Inorganic Chemistry* **1991,** *30*, 609-18.
28. L. Simeone, G. Mangiapia, G. Vitiello, C. Irace, A. Colonna, O. Ortona, D. Montesarchio, L. Paduano *Bioconjugate Chem.*, *submitted*.
29. L. Simeone *at al.* manuscript in preparation.
30. A. H. Velders, A. Bergamo, E. Alessio, E. Zangrando, J. G. Haasnoot, C. Casarsa, M. Cocchietto, S. Zorzet, G. Sava *Journal of Medicinal Chemistry*, **2004,** *47*, 1110-1121.
31. M. W. Grinstaff, P. Barthélémy, C. Prata, L. Moreau, Patent WO, **2005**, 2005007810.
32. M. Bouma, B. Nuijen, M. T. Jansen, G. Sava, A. Flaibani, A. Bult, J. H. Beijnen *International Journal of Pharmaceutics*, 2002, **248**, 239-246.
33. G. Sava, S. Pacor, F. Bregant, V. Ceschia, G. Mestroni *Anti-Cancer Drugs*, **1990**, *1*, 99-108.
34. D. Pluim, R. Van Waardenburg, H. Beijnen Jos, H. M. Schellens Jan *Cancer Chemotherapy and Pharmacology*, **2004**, *54*, 71-8.
35. G. Sava, K. Clerici, I. Capozzi, M. Cocchietto, R. Gagliardi, E. Alessio, G. Mestroni, A. Perbellini *Anti-Cancer Drugs*, **1999**, *10*, 129-138.
36. L. Simeone, G. Mangiapia, C. Irace, A. Di Pascale, A. Colonna, O. Ortona, L. De Napoli, D. Montesarchio, L. Paduano *Mol. BioSyst.,* **2011**, *7*, 3075-86.
37. A. Bergamo, S. Zorzet, M. Cocchietto, M. E. Carotenuto, M. Magnarin, G. Sava *Anticancer Res.*, **2001**, *21*, 1893-1898.

38. For applications of dansyl groups in recognition/sensing systems, see for example: (a) C. Coppola, A. Paciello, G. Mangiapia, S. Licen, M. Boccalon, L. De Napoli, L. Paduano, P. Tecilla, D. Montesarchio *Chem. Eur. J*., **2010**, *16(46*), 13757-13772.b) B. Branchi, P. Ceroni, V. Balzani, G. Bergamini, F. G. Klaerner, F. Voegtle *Chem. Eur. J*., **2009**, *15(32*), 7876-7882; (c) M. Montalti, L. Prodi, N. Zaccheroni, G. Battistini, S. Marcuz, F. Mancin, E. Rampazzo, U. Tonellato *Langmuir*, **2006**, *22(13*), 5877-5881; (d) S. Pagliari, R. Corradini, G. Galaverna, S. Sforza, A. Dossena, M. Montalti, L. Prodi, N. Zaccheroni, R. Marchelli *Chem. Eur. J*., **2004**, *10(11*), 2749-2758; (e) A. Ueno, A. Ikeda, I. Ikeda. T. Ikeda, F. Toda *J. Org. Chem.,* **1999**, *64*, 382-387.
39. G. H. Hakimelahi, Z. A. Proba, K. K. Ogilvie *Can. J. Chem.*, **1982**, *60:(9)*, 1106-1113.
40. J. Grant Buchanan, D. R. Clark *Carbohydr. Res.,* **1979**, *68*, 331-341.
41. H. O. Kim, Y. H. Park, H. R. Moon, L. S. Jeong *Bioorg. Med. Chem. Lett*. **2002**, *12*, 2403–2406.
42. L. B. Townsend, *Chemistry of Nucleosides and Nucleotides*, 1994 Plenum Press New York.
43. L. Simeone, De Napoli, D. Montesarchio *Chem. Biodiv.*, *accepted* doi: 10.1002/cbdv.201100103.

Chapter 4.

Design and Synthesis of a Novel Cationic Aminoacylnucleolipid as a Model Compound for Highly Functionalized Nucleolipids

1. Catanionic vesicles: structure, properties and applications

Catanionic amphiphiles are generally bicatenar systems that result from mixing oppositely charged surfactants in water (Figure 1). [1]

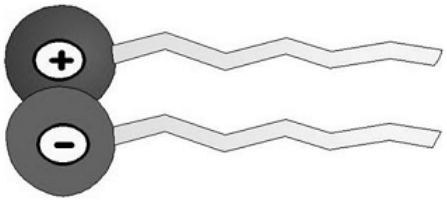


Figure 1. A bicatenary catanionic amphiphile.

When discussing the preparation methods of catanionic amphiphiles, it is necessary to distinguish between catanionic mixtures and catanionic surfactants (pure ion pairs). [2] In the latter category, the inorganic counterions associated with the amphiphiles are eliminated, whereas in catanionic mixtures, counterions remain in solution. For use in vectorization, the presence of residual salts is to be avoided, as some salts may be toxic.

Catanionic surfactants can spontaneously form vesicles in water. [3,4] However, it should be noted that the catanionic surfactants usually precipitate when the two oppositely charged amphiphiles are mixed in equimolar quantities. [3] In fact, the electrostatic interaction between charges leads to a shrinking of the polar head and therefore a decrease in the hydrophilicity of the system. The weakened solvation sphere

makes more difficult the solubilization. Thus catanionic vesicles are generally formed with an excess of either positive or negative charge. In 1997, Menger and coworkers [5] proposed the first example of a water-soluble catanionic surfactant based on a glycosidic amphiphile (Figure 2).

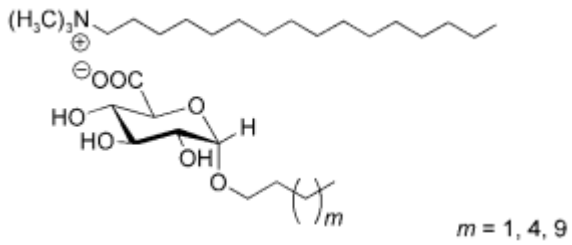


Figure 2. Structure of the first sugar-derived catanionic surfactant described in the literature.

If not for sporadic examples, catanionic vesicles have not been explored in vectorization. [4] A new concept has been developed that involves the direct association of a potentially ionizable drug with a surfactant that gives rise to a catanionic active entity. [6,7] Catanionic vesicles, such as membranes, creating a compartmentalization that allows the trapping of molecules or macromolecules, are therefore used as nanoreactors or as agents of transport and release (*"delivery vehicles"*) of drugs, DNA and receptors in various biomedical non-invasive experiments. As multicharged systems, the strong electrostatic interactions that are established between the cell membrane and the outer surface of the vesicle can induce fusion processes, creating an efficient transport system. [1] One of the great benefits that makes catanionic vesicles ideal candidates for drug delivery is that, unlike liposomes, the preparation of which requires the use of an organic solvent and specific preliminary manipulations, catanionic vesicles are spontaneously formed in water, thus avoiding the use of organic solvents or other, potentially toxic, products. As reported by M. Blanzat, I. Rico-Lattes and coworkers, only one study, carried out in their laboratories, is undergoing industrial development. This research involves the cutaneous delivery of an anti-inflammatory drug by direct association with a sugar-derived amphiphile forming a catanionic surfactant (Figure 3).

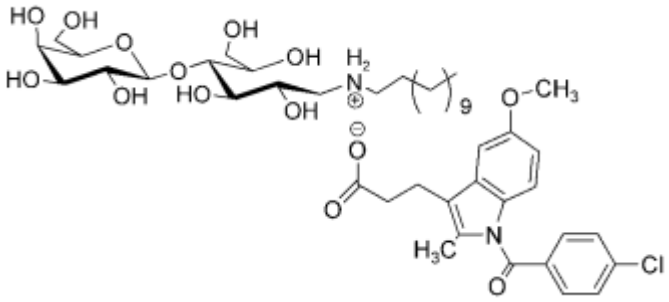


Figure 3. Structure of a catanionic surfactant from the association of an anti-inflammatory drug and a sugar-derived surfactant.

The catanionic assembly, which spontaneously forms vesicles, ensures an increased anti-inflammatory activity of the active principle together with a controlled and prolonged release through the skin. It also protects the drug from harmful irradiation effects. [8,9] In consideration of the intrinsic anionic nature of the Ru(III) complexes we have studied as potential anticancer agents (here described in Chapter 3), it has been considered the possibility to use these systems in mixture with cationic nucleolipids, so to obtain peculiar catanionic vesicles, for an optimized cell delivery and prolonged *in vivo* half-life. In this frame, the design and synthesis of a novel model cationic nucleolipid has therefore been undertaken and is here described.

2. Cationic nucleolipids: design of a novel cationic aminoacylnucleolipid

Nucleolipid-based liposomes may typically be exploited as drug-carriers, being able to transport both hydrophilic compounds, encapsulated in the aqueous compartment, and lipophilic species, inserted within the lipid layers. Liposomes are generally able to incorporate drugs in high concentration, protecting them against extracellular enzymatic degradation and allowing efficient delivery within cells by endocytosis.

A recent work by Yang *et al*. [10] showed very interesting examples of cationic nucleolipids obtained from uridine linking one arginine residue at the 5' position and two fatty acid chains at the 2' and 3' OH attached through carbamate linkages, chemically stable but easily degradable *in vivo* (Figure 4).

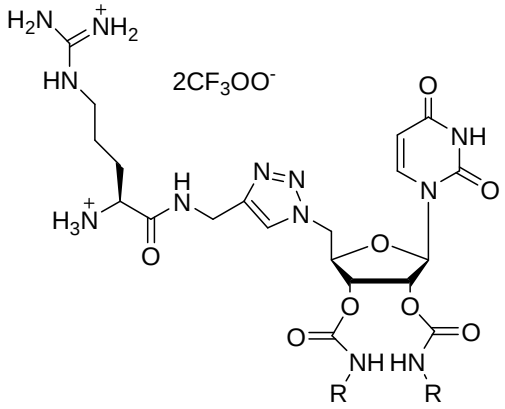


Figure 4. Cationic nucleolipids synthesized by Yang *et al.* [10]

These derivatives, exhibiting poor cytotoxicity, were able to self-aggregate into stable liposomes and thus efficiently vehiculate siRNA *in vitro***Errore. Il segnalibro non è definito.**.

Inspired by these systems, we reasoned that a synthetic platform for a more general and versatile design allowing a large variety of diverse cationic nucleolipids would be desirable. Therefore we searched for structural modifications of the original nucleoside backbone allowing a higher degree of functionalization than it is generally allowed starting from a natural ribo- or deoxyribonucleoside. More specifically, our idea

was to select as the starting material a nucleoside equipped with at least three easily differentiable functional groups, so to attach at each of them – using very simple, efficient and selective procedures and by inserting chemical bonds of different lability in physiological conditions - a different residue, endowed with a specific property. Considering that the nucleobases have to be kept unchanged, since in principle it is desirable to maintain unaltered their recognition abilities, all the required modifications must be introduced at the level of the sugar moiety. An ideal starting scaffold - satisfying these requirements - is the modified nucleoside 3-amino-3-deoxy-1-β-D-xylofuranosyluracil, easily obtained in 6 steps and 50% overall yields starting from uridine following the protocol previously described in Chapter 3, that we have exploited in parallel also for the synthesis of a novel nucleolipid-Ru(III) complex. In three, straightforward steps, this amino-nucleoside can be easily transformed into the derivative 5-*O*-*tert*-butyldimethylsilyl-3-amino-3-deoxy-2-oleyl-1-β-D-xylofuranosyluracil, the synthesis of which has been previously described in Chapter 3.

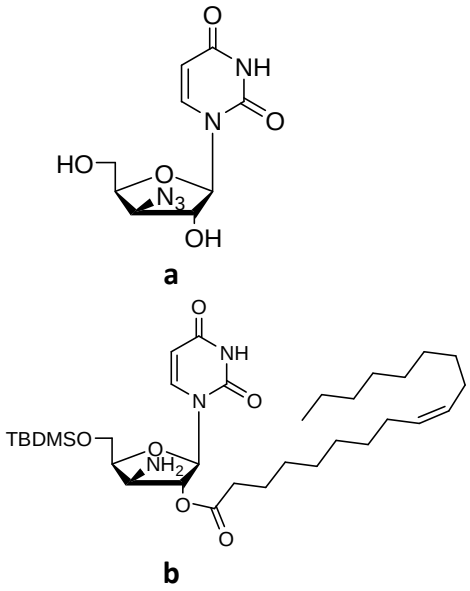
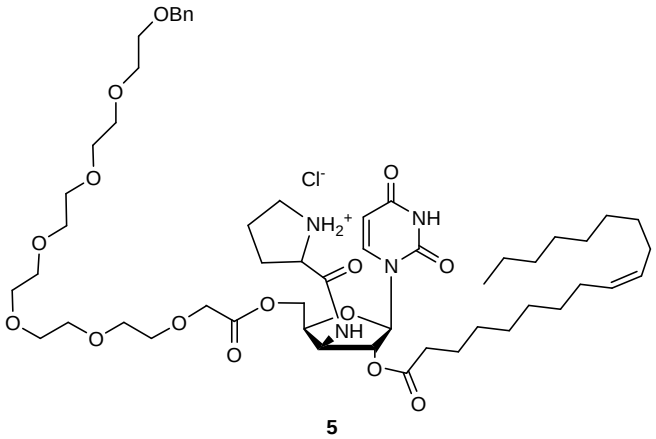


Figure 5. **a.** 3-amino-3-deoxy-1-β-D-xylofuranosyluracil;
b. 5-*O*-TBDMS-3-amino-3-deoxy-2-oleyl-1-β-D-xylofuranosyluracil.

In order to obtain a suitable model compound for a novel class of cationic nucleolipids, we have here exploited this nucleoside analog to build aminoacylnucleolipid **5**, decorated with the following groups:

I. a hydrophilic, polyether chain attached at the 5' position;
II. a lipophilic fatty acid chain at the 2' position, ensuring the desired amphiphilicity to the final compound – and intrinsically the ability to generate large aggregates in aqueous conditions;
III. a proline residue at the 3' position, ionized in physiological conditions and therefore conferring the desired positive charge.

The synthesized molecule **5** is reported in Figure 6. Compared with the previously described cationic nucleolipids reported by Yang *et al.* and depicted in Figure 4, this design offers in principle one additional functional group for diversification. In the here synthesized compound, this diversification site was exploited to insert a hydrophilic, polyether chain, so to obtain an amphiphile with a better balanced lipophilicity/hydrophilicity ratio in its backbone, and expectedly endowed with better nanostructuring performances in aqueous media with respect with known literature examples.

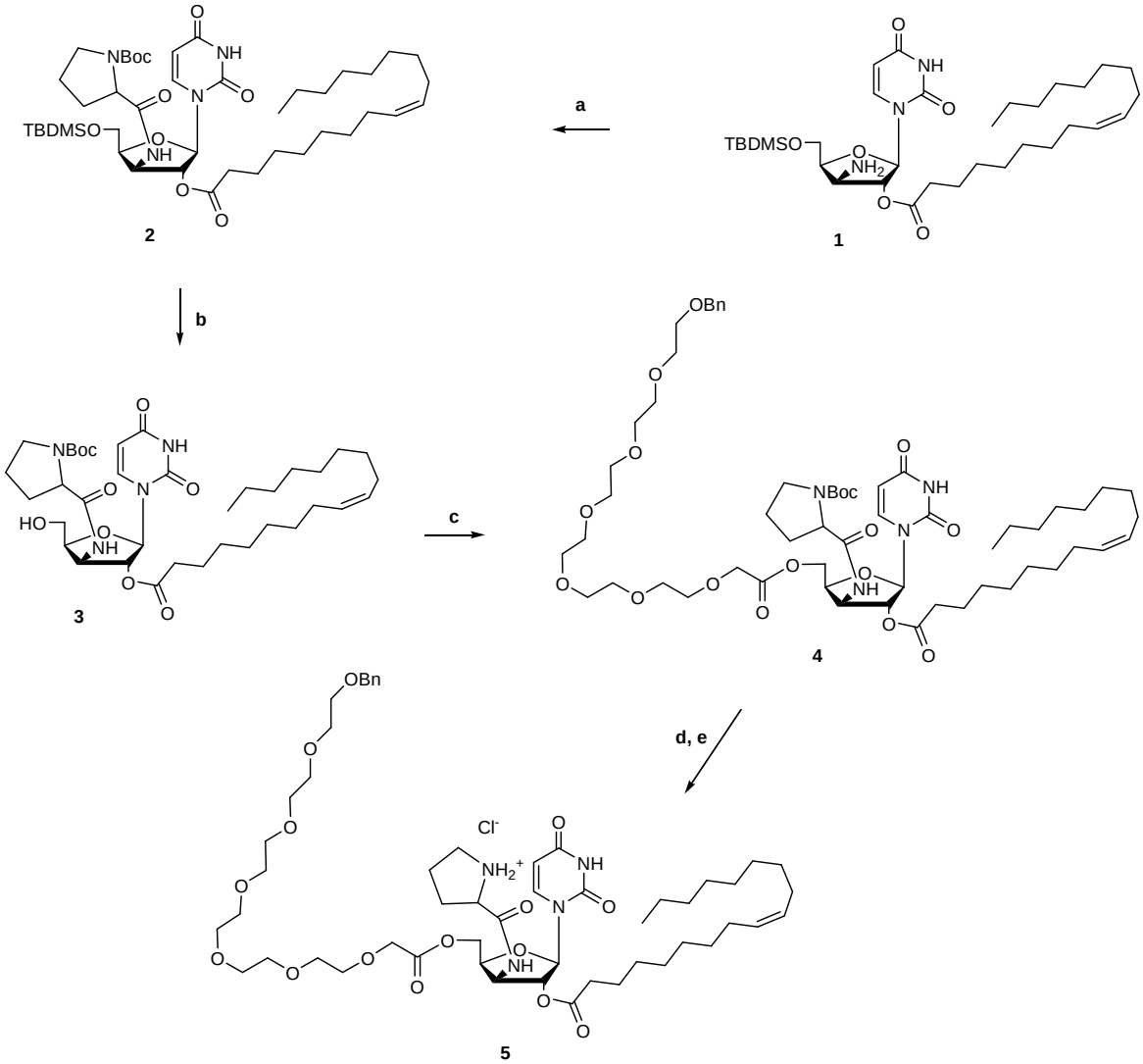


5

Figure 6. Cationic aminoacylnucleolipid **5**.

3. Synthesis of cationic aminoacylnucleolipid

The synthetic scheme adopted for the synthesis of aminoacylnucleolipid **5** is based on the use, as the starting material, of the nucleoside 5'-*O*-TBDMS-3'-amino-3'-deoxy-2'-oleyl-1-β-D-xylofuranosyluracil, obtained following the procedures described in Chapter 3.



Scheme 1. **a.** Boc-Pro, DCC, DMAP, DCM, 12 h, r.t.; **b.** $Et_3N{\cdot}3HF$, THF, 12 h, r.t.; **c.** $BnO(CH_2CH_2O)_6CH_2COOH$, DCC, DMAP, DCM, 2 h, r.t.; **d.** TFA, 10% TIS, DCM, 1 h, r.t.; **e.** $NH_4^+Cl^-/H_2O$, lyophilizations.

As shown in Scheme 1, the first step involves the condensation of a residue of *N*-Boc Proline with the amino group at the 3' position of nucleoside **1**, realized by classical DCC activation, giving aminoacylnucleolipid **2**. The desired amide was then deprotected at

the 5' position by treatment with $Et_3N{\cdot}3HF$ in THF to give nucleoside **3** in almost quantitative yields. At this stage a hydrophilic residue of $BnO(CH_2CH_2O)_6CH_2COOH$ was attached at the 5' position through an ester linkage, thus obtaining nucleoside **4.** The final step requires a simple manipulation by a mild acidic treatment (10% TFA in DCM) ensuring the clean Boc removal, accomplished in almost quantitative yield to give aminonucleolipid **5** as a trifluoroacetate salt. To obtain the aminoacylnucleolipid in a better biocompatible form we have then replaced the trifluoroacetate counterion with a chloride anion: this was realized by treatment of **5** with an equimolar amount of $NH_4^+Cl^-$ and subsequent repeated lyophilizations from water. The absence of CF_3COO^- was confirmed by ^{13}C-NMR analysis, which in concentrated samples after extremely long acquisition times (>24 h) did not show the diagnostic signals of the carbons of trifluoroacetate, even in traces.

Detailed studies on the biocompatibility and ability to generate stable self-aggregates in aqueous solution are currently in progress in collaboration with specialized laboratories.

Moreover, the aminoacylnucleolipid **5** will be studied in formulation with anionic derivatives. Particularly, our current interests are focused on the investigation of catanionic vesicles formed by amphiphilic nucleolipids Ru(III) complexes (anionic derivatives) described in Chapter 3, in mixture with the here described cationic nucleolipid **5** with the final aim to generate more stable formulations for the ruthenium-containing nucleolipids for *in vivo* experiments.

4. Conclusions

A novel cationic aminoacylnucleolipid is here described. This prolyl-nucleolipid based on Uridine was obtained by exploiting a straightforward, high yielding and versatile synthetic protocol, starting from the easily accessible 5'-*O*-TBDMS-3'-amino-3'-deoxy-2'-oleyl-1-β-D-xylofuranosyluracil key intermediate. Target nucleoside **5** has been designed as a model compound, synthesized as the first member of a more general class of highly functionalized nucleolipids carrying positive charges, in principle allowing a large variety of different, related analogs with a high degree of molecular diversity.

Studies on a detailed physico-chemical characterization of this nucleolipid are currently in progress in the laboratories of proff. Luigi Paduano and Gerardino D'Errico of the Dept. of Chemistry "Paolo Corradini" of "Federico II" University of Naples, carried out by using a combined approach, including DLS, surface tension and conductibility measurements in aqueous solution. Moreover, tests on the cytotoxicity activity on various human cells lines and bacteria are also underway in collaboration with specialized laboratories (more in detail: with the research group of prof. Alfredo Colonna, of the Dept. of Pharmacology of "Federico II" University of Naples, as far as the general cytotoxicity analyses and the biocompatibility studies are concerned, and with Dr. Monica Benincasa of the Dept. of Life Sciences of University of Trieste, as far as the antimicrobial and antifungal tests are concerned).

Successively, the ability of the here synthesized cationic nucleolipid **5** to generate catanionic vesicles when mixed in formulation with the ruthenium(III)-containing nucleolipids, showed in Chapter 3, will be investigated in detail with the final goal to increase the biostability and also enhance the bioactivity of the Ru(III) complexes.

5. Experimental section

General Methods.

All the reagents were of the highest commercially available quality and were used as received. TLC analyses were carried out on silica gel plates from Merck (60, F254). Reaction products on TLC plates were visualized by UV light and then by treatment with a 10 % $Ce(SO_4)_2/H_2SO_4$ aq. solution. For column chromatography, silica gel from Merck (Kieselgel 40, 0.063-0.200 mm) was used. NMR spectra were recorded on Varian XR 200 and Varian Inova 500 spectrometers, as specified. All the chemical shifts are expressed in ppm with respect to the residual solvent signal. Peak assignments have been carried out on the basis of standard ^1H-^1H COSY and HSQC experiments. The following abbreviations were used to explain the multiplicities: s = singlet; d = doublet; t = triplet; q = quartet; m = multiplet; b = broad; dd = double doublet. For the ESI MS analyses, a Waters Micromass ZQ instrument – equipped with an Electrospray source.

Synthesis of compound 2.

Nucleolipid **1** (28 mg, 0.045 mmol), dissolved in anhydrous CH_2Cl_2 (0.50 mL), was reacted with DMAP (3.0 mg, 0.025 mmol), Boc-Proline (13 mg, 0.060 mmol) and DCC (14 mg, 0.068 mmol). The resulting solution was stirred at room temperature for 12 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (7:3, v/v). The desired product **2** was obtained in 98% isolated yield (36 mg, 0.044 mmol).

2: oil, *Rf* = 0.5 [*n*-hexane/AcOEt (1:1, v/v)].

^{1}H NMR (DMSO-d_6, 500 MHz): δ 11.41 (1H, bs, N*H*-3); 7.96 (1H, d, *J* = 7.0 Hz, N*H*-3'); 7.75 (1H, m, H-6); 5.83 (1H, d, *J* = 6.0 Hz, H-1'); 5.74 (1H, bs, H-5); 5.56 (1H, d, *J* = 8.0 Hz, H_α Proline); 5.31 (2H, m, H-9 and H-10 oleic acid); 5.24 (1H, bs, H-2'); 4.62 (1H, m, H-3'); 4.27 (1H, bs, H-4'); 3.70 (2H, m, H-5'$_a$ and H-5'$_b$); 3.30 (2H, overlapped signals, 2x H_β Proline); 2.27 (2H, bs, CH_2C=O oleic acid); 1.97 (4H, d, *J* = 6.0 Hz, 2x H-8 and 2x H-11 oleic acid); 1.76 (2H, m, H_γ Proline); 1.48 – 1.24 (33H, overlapped signals, aliphatic protons of oleic acid, Boc and 2x H_δ Proline); 0.85 (12H, overlapped signals, CH_3 oleic acid and SiC(CH_3)$_3$), 0.07 and 0.05 (6H, s's, Si(CH_3)$_2$).

^{13}C NMR (DMSO-d_6, 125 MHz): δ 172.3 (C=O), 162.9 (C-4); 156.5 (C=O Proline); 153.1 (C=O Boc); 150.4 (C-2); 140.6 (C-6); 129.5 (C-9 and C-10 oleic acid); 102.1 (C-5); 85.8 (C-1'); 78.6 (C-2' and C-4'); 77.3 (C_α Proline); 62.3 (C-5'); 60.0 (C-3'), 53.1 (C_δ Proline); 47.5 (C_γ Proline); 33.3 (C_β Proline); 31.2, 29.1, 28.8, 28.6, 28.2, 28.0, 26.6, 25.3, 24.4, 24.2, 22.1 (aliphatic carbons of oleic acid and Boc -C(CH_3)$_3$); 25.8 (SiC(CH_3)$_3$); 17.8 (SiC(CH_3)$_3$); 14.0 (-CH_3 oleic acid); -5.5 (Si(CH_3)$_2$).

ESI-MS (positive ions): for $C_{43}H_{74}N_4O_9Si$, calcd. 818.5225; found m/z: 820.41 [M+H^+]; 842.13 [M+Na^+]; 858.25 [M+K^+].

Synthesis of compound 3.

Nucleolipid **2** (60 mg, 0.073 mmol), dissolved in anhydrous THF (1.0 mL), was treated with $Et_3N{\cdot}3HF$ (35 µL, 0.22 mmol). The solution was stirred at room temperature for 12 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/AcOEt (1:4, v/v). The desired product **3** was obtained in 93% isolated yield (48 mg, 0.068 mmol).

3: oil, *Rf*= 0.2 [*n*-hexane/AcOEt (3:7, v/v)].

^{1}H NMR (DMSO-d_6, 500 MHz): δ 11.34 (1H, bs, N*H*-3); 8.06 (1H, bs, N*H*-3'); 7.99 (1H, d, *J* = 8.0 Hz, H-6); 5.89 (1H, bs, H-1'); 5.73 (1H, d, *J* = 7.5 Hz, H-5); 5.56 (1H, d, *J* = 8.0 Hz, H_α Proline); 5.32 (3H, m, H-2', H-9 and H-10 oleic acid); 5.17 (1H, m, O*H*-5'); 4.74 (1H, m, H-3'); 4.24 (1H, bs, H-4'); 3.57 (2H, m, H-5'$_a$ and H-5'$_b$); 3.29 (2H, overlapped signals, 2x H_β Proline); 2.27 (2H, bs, CH_2C=O oleic acid); 1.97 (4H, d, *J* = 5.5 Hz, 2x H-8 and 2x H-11 oleic acid); 1.73 (2H, m, H_γ Proline); 1.52 – 1.24 (33H, overlapped signals, aliphatic protons of oleic acid, Boc and 2x H_δ Proline); 0.85 (3H, t, *J* = 6.5 and 7.0 Hz, CH_3 oleic acid).

^{13}C NMR (DMSO-d_6, 125 MHz): δ 172.7 (C=O); 163.1 (C-4); 156.8 (C=O Proline); 153.4 (C=O Boc); 150.6 (C-2); 140.9 (C-6); 129.8 (C-9 and C-10 oleic acid); 102.4 (C-5); 85.2 (C-1'); 78.9 (C-4'); 78.5 (C-2'); 77.2 (C_α Proline); 60.5 (C-5'); 60.0 (C-3'); 52.9 (C_δ Proline); 47.7 (C_γ Proline); 33.5 (C_β Proline); 31.5, 29.2, 29.0, 28.9, 28.7, 28.6, 28.4, 28.2, 26.8, 25.5, 24.6, 24.4, 22.3 (aliphatic carbons of oleic acid and Boc -C(CH_3)$_3$); 14.2 (-CH_3 oleic acid).

ESI-MS (positive ions): for $C_{37}H_{60}N_4O_9$, calcd. 704.4360; found m/z: 705.98 [M+H^+]; 728.03 [M+Na^+]; 743.97 [M+K^+].

Synthesis of compound 4.

Nucleolipid **3** (45 mg, 0.064 mmol), dissolved in anhydrous CH_2Cl_2 (1.0 mL), was treated with DMAP (4.0 mg, 0.032 mmol), $BnOCH_2CH_2(OCH_2CH_2)_5OCH_2COOH$ (36 mg, 0.083 mmol) and DCC (20 mg, 0.096 mmol), added in the order. The solution was stirred at room temperature for 2 h, then the solvent was removed under reduced pressure and the crude was purified on a silica gel column, eluted with *n*-hexane/acetone (1:1, v/v). The desired product **4** was obtained in 84% isolated yield (60 mg, 0.054 mmol).

4: oil, *Rf* = 0.4 [*n*-hexane/acetone (1:1, v/v)].

^{1}H NMR (DMSO-d_6, 500 MHz): δ 11.40 (1H, bs, N*H*-3); 8.23 (1H, bs, N*H*-3'); 7.70 (1H, d, *J* = 7.5 Hz, H-6); 7.34 – 7.27 (5H, m, aromatic protons of Bn); 5.87 (1H, d, *J* = 5.5, H-1'); 5.78 (1H, bs, H-5); 5.32 (3H, d, *J* = 8.0 Hz, H_α Proline, H-9 and H-10 oleic acid); 5.42 (1H, dd, *J* = 6.5 and 6.0 Hz, H-2'); 5.33 (2H, m, C*H*$_2$Bn); 4.70 (1H, m, H-3'); 4.48 (3H, bs, H-4' and C*H*$_2$C=O); 4.10 (2H, m, H-5'$_a$ and H-5'$_b$); 3.50 (24H, overlapped signals, -OC*H*$_2$C*H*$_2$O-); 3.30 (2H, overlapped signals, 2x H_β Proline); 2.27 (2H, m, C*H*$_2$C=O oleic acid); 1.97 (4H, d, *J* = 5.5 Hz, 2x H-8 and 2x H-11 oleic acid); 1.72 (2H, m, H_γ Proline); 1.52 – 1.00 (33H, overlapped signals, aliphatic C*H*$_2$ protons of oleic acid, Boc and 2x H_δ Proline); 0.85 (3H, t, *J* = 6.5 and 7.0 Hz, C*H*$_3$ oleic acid).

^{13}C NMR (CD_3OD, 50 MHz): δ 172.1 and 170.7 (2x C=O), 166.1 (C-4); 156.8 (C=O Proline); 152.7 (C=O Boc); 146.6 (C-2); 140.2 (C-6); 139.8 (quaternary carbon of Bn group); 131.4, 131.3, 129.9, 129.4, 129.2 (C-9 and C-10 oleic acid and aromatic *C*H carbons of Bn); 103.9 (C-5); 82.5 (C-1'); 81.8 (C-4'), 79.0 (C-2'); 74.6 (O*C*H$_2$C=O); 72.4 (C_α Proline); 72.0, 71.6, 71.1 (overlapped signals, O*C*H$_2$*C*H$_2$O); 69.8 (-O*C*H$_2$Bn); 63.8 (C-5'); 63.2 (C-3'), 56.3 (C_δ Proline); 33.5 (C_β Proline); 31.5, 29.2, 29.0, 28.9, 28.7, 28.6, 28.4, 28.2, 26.8, 25.5, 24.2 (aliphatic *C*H$_2$ carbons of oleic acid and Boc-C(*C*H$_3$)$_3$); 15.0 (-*C*H$_3$ oleic acid). C_γ Proline is buried under the solvent signal.

ESI-MS (positive ions): for $C_{58}H_{92}N_4O_{17}$, calcd. 1116.6457; found m/z: 1118.88 [M+H$^+$]; 1140.66 [M+Na$^+$]; 1156.65 [M+K$^+$].

Synthesis of compound 5.

Nucleolipid **4** (54 mg, 0.048 mmol), dissolved in anhydrous DCM (0.85 mL), was reacted with TFA (0.15 mL). The reaction was stirred at room temperature for 1 h, then the

solvent was removed under reduced pressure and the crude was coevaporated with 2-propanol (1.0 mL for 3 times). The desired final compound was thus obtained in almost quantitative yields in the form of trifluoroacetate salt (55 mg, 0.048 mmol).

Successive treatment with one equivalent (2.6 mg, 0.048 mmol) of ammonium chloride, followed by repeated lyophilizations from water, gave target compound **5** in the form of chloride salt.

5: white powder, *Rf* = 0.2 [*n*-hexane/acetone (1:4, v/v)].

^{1}H NMR (DMSO-d$_6$, 500 MHz): δ 11.40 (1H, bs, N*H*-3); 8.81 (1H, d, *J* = 8.0 Hz, N*H*-3'); 7.70 (1H, d, *J* = 8.0 Hz, H-6); 7.34 – 7.26 (5H, complex signals, aromatic protons of Bn); 5.89 (1H, d, *J* = 6.0, H-1'); 5.78 (1H, d, *J* = 7.5 Hz, H-5); 5.30 (3H, m, H-2', H-9 and H-10 oleic acid); 4.72 (1H, m, H-3'); 4.84 (3H, bs, H-4' and OC*H*$_2$C=O); 4.15 (3H, overlapped signals, H$_\alpha$ proline, H-5'$_a$ and H-5'$_b$); 3.51 (24H, overlapped signals, OC*H*$_2$C*H*$_2$O); 3.34 (2H, m, C*H*$_2$C=O oleic acid); 2.30 (2H, m, 2 x H$_\beta$ proline); 1.96 (4H, m, 2x H-8 and 2x H-11 oleic acid); 1.50 (2H, overlapped signals, H$_\gamma$ proline and C*H*$_2$CH$_2$C=O); 1.26 – 1.00 (24H, overlapped signals, aliphatic C*H*$_2$ protons of oleic acid and 2x H$_\delta$ proline); 0.83 (3H, m, C*H*$_3$ oleic acid).

^{13}C NMR (CD$_3$OD, 50 MHz): δ 172.5 and 170.5 (2x C=O); 166.2 (C-4); 160.3 (C=O Proline); 152.6 (C-2); 144.3 (C-6); 139.9 (quaternary carbon of Bn); 131.4 131.3, 130.0, 129.5, 129.3 (C-9 and C-10 oleic acid and aromatic *C*H carbons of Bn); 104.1 (C-5); 91.2 (C-1'); 80.3 (C-4'); 78.0 (C-2'); 74.5 (C$_\alpha$ Proline); 72.2, 71.7, 70.9 (-O*C*H$_2$*C*H$_2$O-); 69.8 (-O*C*H$_2$Bn); 69.5 (O*C*H$_2$CO); 64.3 (C-3'); 61.7 (C-5'), 56.8 (C$_\delta$ Proline); 48.0 (C$_\gamma$ Proline); 35.2 (C$_\beta$ Proline); 33.5, 32.3, 31.3, 31.1, 30.8, 28.6, 27.5, 27.2, 27.0, 26.5, 26.2, 25.5, 24.2 (aliphatic *C*H$_2$ carbons of oleic acid); 15.0 (-*C*H$_3$ oleic acid).

ESI-MS (positive ions): for $C_{53}H_{84}N_4O_{15}$, calcd. 1016.5933; found m/z: 1018.57 [M+H$^+$]; 1040.56 [M+Na$^+$]; 1056.44 [M+K$^+$].

6. References

1. E. Soussan, S. Cassel, M. Blanzat, I. Rico-Lattes *Angew. Chem. Int. Ed.*, **2009**, *48*, 274 and literature cited therein.
2. C. Tondre, C. Caillet *Adv. Colloid Interface Sci.*, **2001**, *93*, 115.
3. E. W. Kaler, A. Kamalakara Murthy, B. E. Rodriguez, J. A. N. Zasadzinski *Science* **1989**, *245*, 1371.
4. T. Bramer, N. Dew, K. Edsman *J. Pharm. Pharmacol.*, **2007**, *59*, 1319.
5. F. M. Menger, W. H. Binder, J. S. Keiper *Langmuir*, **1997**, *13*, 3247.
6. T. Bramer, M. Paulsson, K. Edwards, K. Edsman *Pharm. Res.*, **2003**, *20*, 1661.
7. T. Bramer, N. Dew, K. Edsman *J. Pharm. Sci.*, **2006**, *95*, 769.
8. S. Consola, M. Blanzat, E. Perez, J. C. Garrigues, P. Bordat, I. Rico-Lattes *Chem. Eur. J.*, **2007**, *13*, 3039.
9. S. Consola, M. Blanzat, I. Rico-Lattes, E. Perez, P. Bordat *EP2006/064502*, **2006**; WO2007039561, 2007.
10. H. W. Yang., J. W. Yi, E. K. Bang, E. M. Jeon, B. H. Kim *Org. Biomol. Chem*., **2011**, *9*, 291.

Chapter 5.

Design, Synthesis and Characterization of Novel Guanosine-based Amphiphiles

1. Introduction: Guanosine and G-assembly

Guanosine is a special nucleoside: it can expand the Watson-Crick interactions capability having a higher number of acceptor/donor H-bonding sites. In particular, guanine moieties may give rise to the Hoogsteen interactions involving positions 6 and 7 as acceptors (**A**), and 1 and 2 as donors (**D**) (Figure 1).

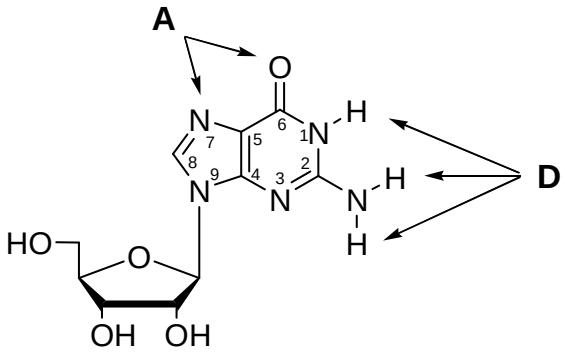


Figure 1. Hoogsteen's interaction sites of Guanosine.

Guanosine can self-assemble in various supramolecular architectures. The most relevant is the G-quartet, a cyclic planar array of four guanosines, each of which gives and receives two hydrogen bonds. These aggregates are stabilized by the coordination of a metal ion (generally, sodium or potassium) with the carbonyl group in position 6 (Figure 2). At high concentrations, the G-quartet can form G-quadruplex structures by overlaying several plans of G-quartets in a cylindrical or helical symmetry (Figure 3). [1]

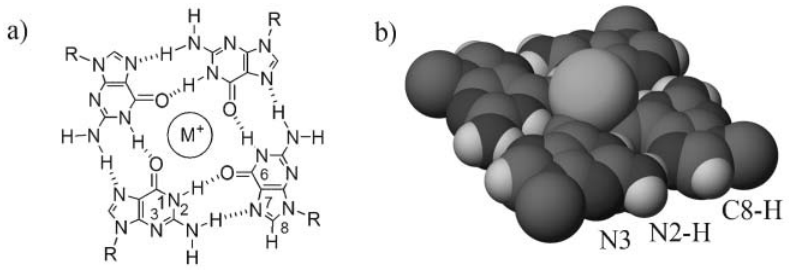


Figure 2. G-quartet arrangement.

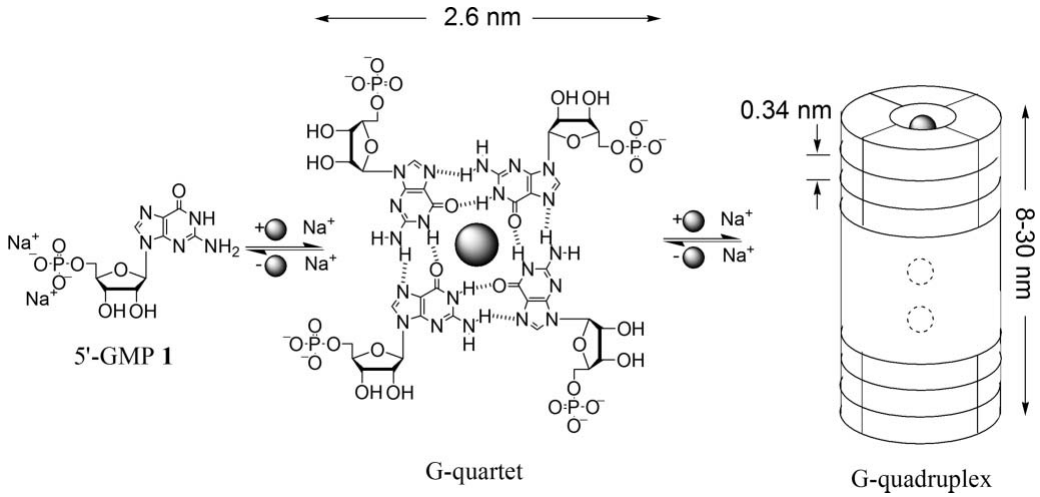


Figure 3. G-quadruplex formation.

In the absence of cations guanosines, when derivatized with lipophilic groups, can form another supramolecular architecture: the G-ribbon. The G-ribbons are organized in sheet-like bidimensional assemblies involving another donor site on the nucleobase: the position 3 (Figure 4). [2,3]

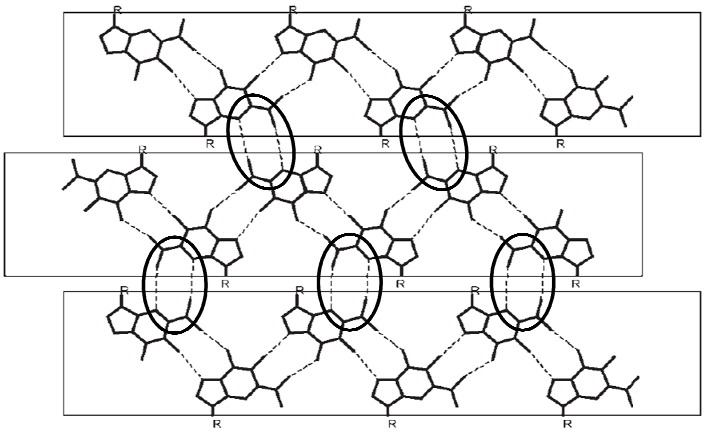


Figure 4. G-ribbon structure.

In Figure 5 some examples of lipophilic guanosines described in the literature are reported. These molecules are able to self-organize in G-quartet or G-ribbon-based architectures in relation to the presence or absence of metal cations. Milestones in this research field are the works of G. P. Spada [1b,3-5], J. T. Davis [1,6-8] and K. Araki [9-11], who have studied several compounds obtained from guanosine by inserting various lipophilic chains on the ribose moiety.

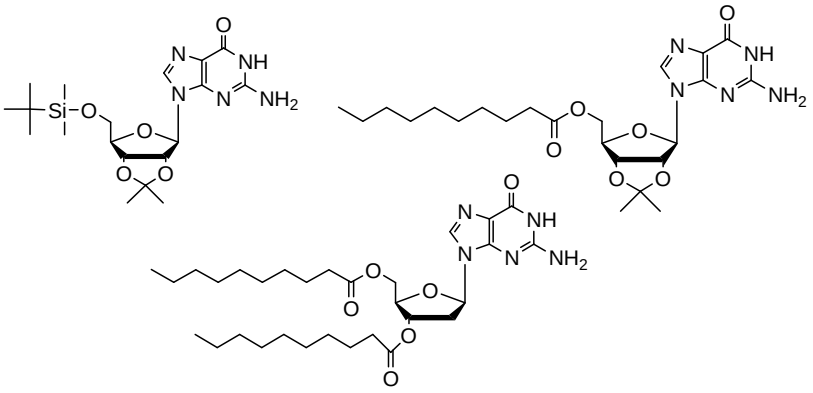


Figure 5. Examples of lipophilic guanosine derivatives reported in the literature [1b].

It is to be mentioned that all the synthetic schemes so far proposed for their preparation are not general in their scope, not allowing an easy access to highly functionalized derivatives; particularly it is not possible to obtain compounds carrying a large diversity of functionalizations on the sugar motif.

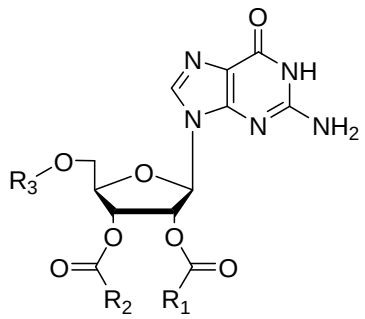
2. Design of a library of amphiphilic guanosine derivatives

This part of my research has been focused on the design, synthesis and characterization of a mini-library of amphiphilic, sugar-modified guanosine derivatives.

If several amphiphilic nucleolipids based on adenosine, cytosine, uridine or thymidine are described in the literature, very few examples are known concerning guanosine [12-16] and almost completely unexplored are the potential applications of guanosine-based amphiphiles in biomedical applications and/or as self-assembling materials. Still, guanosine-based amphiphiles in principle are extremely intriguing compounds, filling the gap between the hydrophilic guanosine or guanosine phosphate derivatives and related G-quadruplex systems, on one side, and the known lipophilic guanosine analogs, on the other. In this scenario, it was urgent to preliminary provide a general synthetic scheme allowing for a high diversification of the ribose functionalizations in the guanosine backbone.

R_3= Polar Group

- Polyethers
- Aminoacids
- Charged groups
- Carbohydrates



R_1, R_2= lipophilic groups: fatty acids

A variety of synthetic strategies have been proposed to prepare nucleolipids. In most cases, the design is based on the chemical modification of the sugar functions, thus minimally perturbing the nucleobase moieties, and keeping unaltered their critical recognition sites. Ester functions are typically considered as ideal chemical connections for nucleosidic ribose derivatizations, offering several advantages: these linkages can be obtained through straightforward and high yielding condensations, not requiring prior modifications of the nucleoside moieties, and, remarkably, are rapidly cleaved by cell esterases, liberating the “free” nucleosides within the cells once the transport function through lipid bilayers is achieved. [17]

The general design here presented for the preparation of amphiphilic guanosine-derivatives is based on the insertion of a saturated fatty acid residue – *i.e.,* myristic acid – as the lipophilic chain. In the synthesized compounds, here named **G1**-**G8** (Figure 6), this group was attached at the ribose secondary hydroxyls, to give bi-tailed compounds **G1**-**G5**, or at the 5’-OH end, to give mono-tailed derivative **G7**. [18] Higher molecular diversity

was introduced at the level of the hydrophilic groups, chosen among polyethers (introduced in **G1**, **G6**, **G7** and **G8**), charged functional groups (in **G3**), or natural α-aminoacids (in **G2** and **G4**) and carbohydrates (in **G5**). Aiming at a simple, versatile and finely tunable synthetic scheme, the library of sugar-modified amphiphilic analogs was prepared starting from a unique key intermediate (**3**, Figure 6), exhaustively protected on the nucleobase, [19] and exhibiting three ribosidic OH groups available for condensations with various appendages.

In the series **G1**-**G5**, following the Grinstaff and Berthélémy's approach, [12] the ribose secondary OH functions have been exploited to insert fatty acid residues, while the 5'-OH is derivatized with different hydrophilic groups, introduced to balance the hydrophobic contribution of the lipid chains. **G7** was designed as the **G1** analog having the same lipophilic and hydrophilic tails but with inverted ratio, while **G6** is decorated uniquely with polyether groups. Finally, the **G8** derivative was designed as a **G7** analog; it was obtained by insertion - in position 5' - of 5-doxyl-stearic acid, bearing the stable doxyl spin label group, for applications in studies of interactions with cell membranes, monitored by ESR experiments.

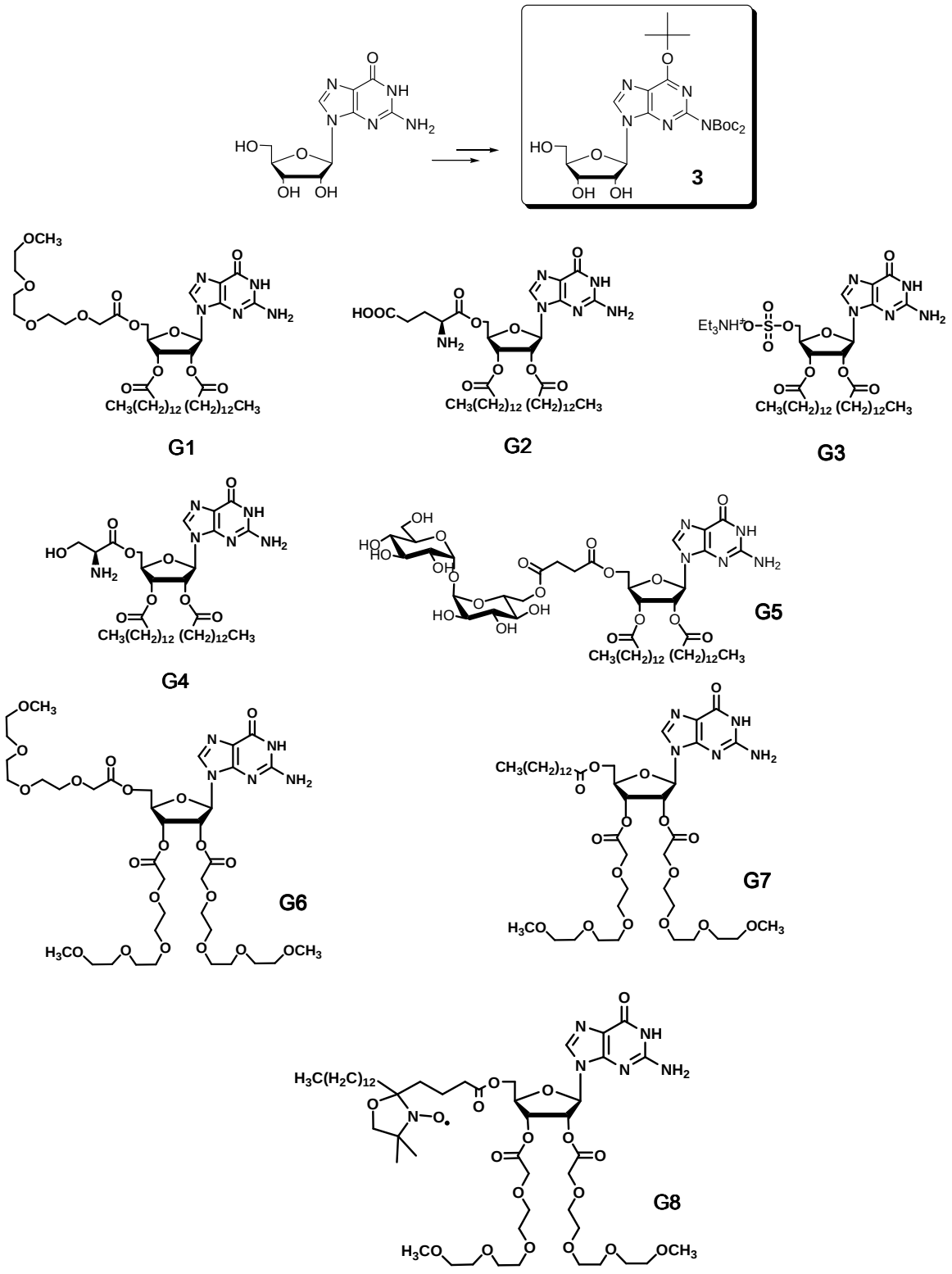
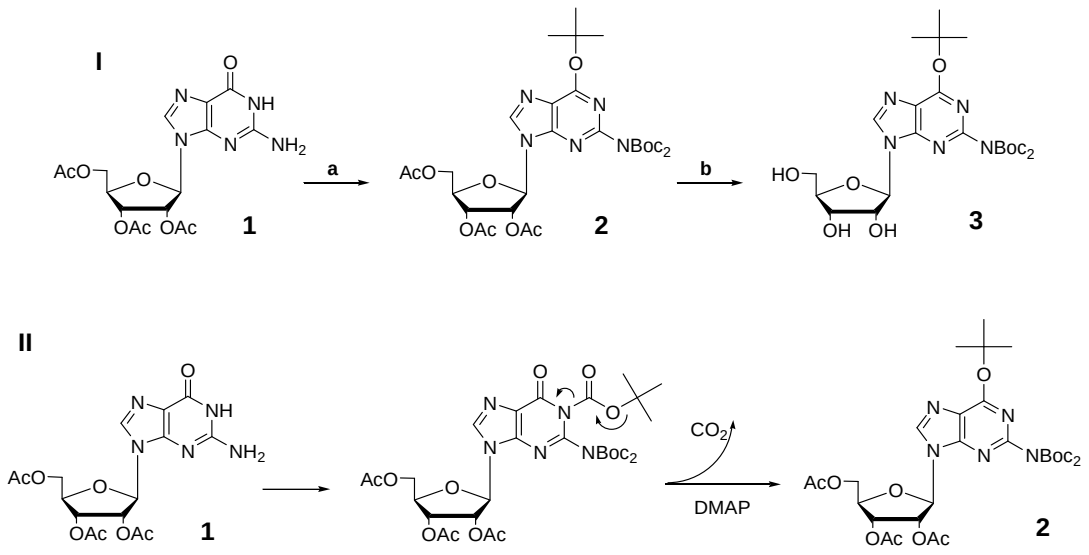


Figure 6. Library of the here studied compounds **G1-G8.** [18]

In this manner, the series of compounds **G1**, **G7**, **G6** was generated, with respectively 2, 1 or 0 myristic acid tails, and in turn 1, 2 or 3 triethylene glycol groups.

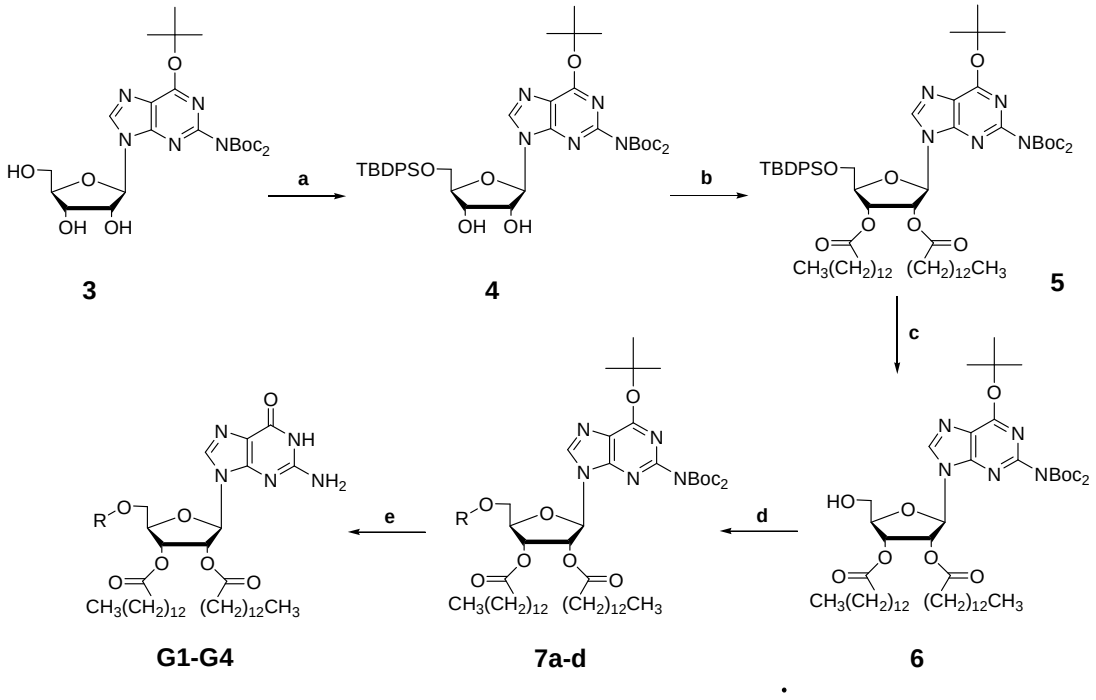
3. Synthesis of compounds G1-G8

The synthetic access to the target sugar-modified guanosine analogs was thus realized starting from compound **3**, obtained in three straightforward steps from guanosine (Scheme 1). [18] The Boc group, though seldom used for nucleobase protection in nucleoside chemistry, [20] was here selected to protect the guanine moiety in consideration of the easy installation procedure, its lipophilic character, facilitating guanosine manipulations in organic solvents, and, above all, the convenient removal conditions. As a matter of fact, Boc deprotection can be achieved by mild acidic treatments (dilute TFA solutions in anhydrous organic solvents), fully compatible with the ester linkages, and eventually originates only volatile side products, thus not requiring column chromatography to give the pure, deprotected compounds. In order to obtain **3**, guanosine was first regioselectively protected [21] by treatment with acetic anhydride in CH_3CN and TEA; tri-*O*-acetylated derivative **1** was then reacted with an excess of di-*tert*-butyl dicarbonate (Boc_2O) in CH_3CN in the presence of TEA and DMAP, giving fully protected nucleoside **2**. Interestingly, this compound, obtained in acceptable yields (57%) compared to previously reported procedures, [20] showed a *tert*-butyl group linked to the O-6 position. This, in analogy with a previous report on guanine protection, [19] may be ascribed to spontaneous loss of CO_2, promoted by DMAP, from the original *N*-1 Boc carbamate, undergoing an internal transposition to the O^6-*tert*-butyl ether (Scheme 1). Differently from the above cited case, we could not isolate the tris-Boc protected intermediate and only compound **2** was recovered after column chromatography. This derivative was then exhaustively deacetylated by treatment with methanolic ammonia, giving target compound **3**, obtained in 56% overall yields for three steps.



Scheme 1. I) Synthesis of intermediate **3**. Reaction conditions: **a)** Boc_2O, TEA, DMAP, CH_3CN, r.t., 72 h, 57%; **b)** NH_3, CH_3OH, 12 h, r.t., quant. **II)** Rationalization of the formation of the O^6-*tert*-*b*utyl derivative.

The synthesis of compounds **G1**-**G5** was accomplished in two simple manipulations starting from derivative **6**, in turn obtained in three steps from key intermediate **3** (Scheme 2). Reaction of **3** with TBDPSCl and imidazole in DMF provided 5'-protected nucleoside **4**, which was then condensed with myristic acid in the presence of DCC, cleanly giving 2',3'-di-*O*-myristoyl derivative **5**. TBDPS group removal was achieved by treatment of **5** with the complex $Et_3N{\cdot}3HF$ in THF, leading to target compound **6** in 70% overall yield from **3**. **G1** was then obtained in almost quantitative yields in two steps, involving first the coupling of 5'-OH deprotected nucleoside **6** with monomethoxy(triethylene glycol) acetic acid, prepared as described in Chapter 2, [17] using DCC as the condensing agent, followed by an acidic treatment with 10% TFA in CH_2Cl_2 for 2 h at r.t., to achieve full nucleobase deprotection.

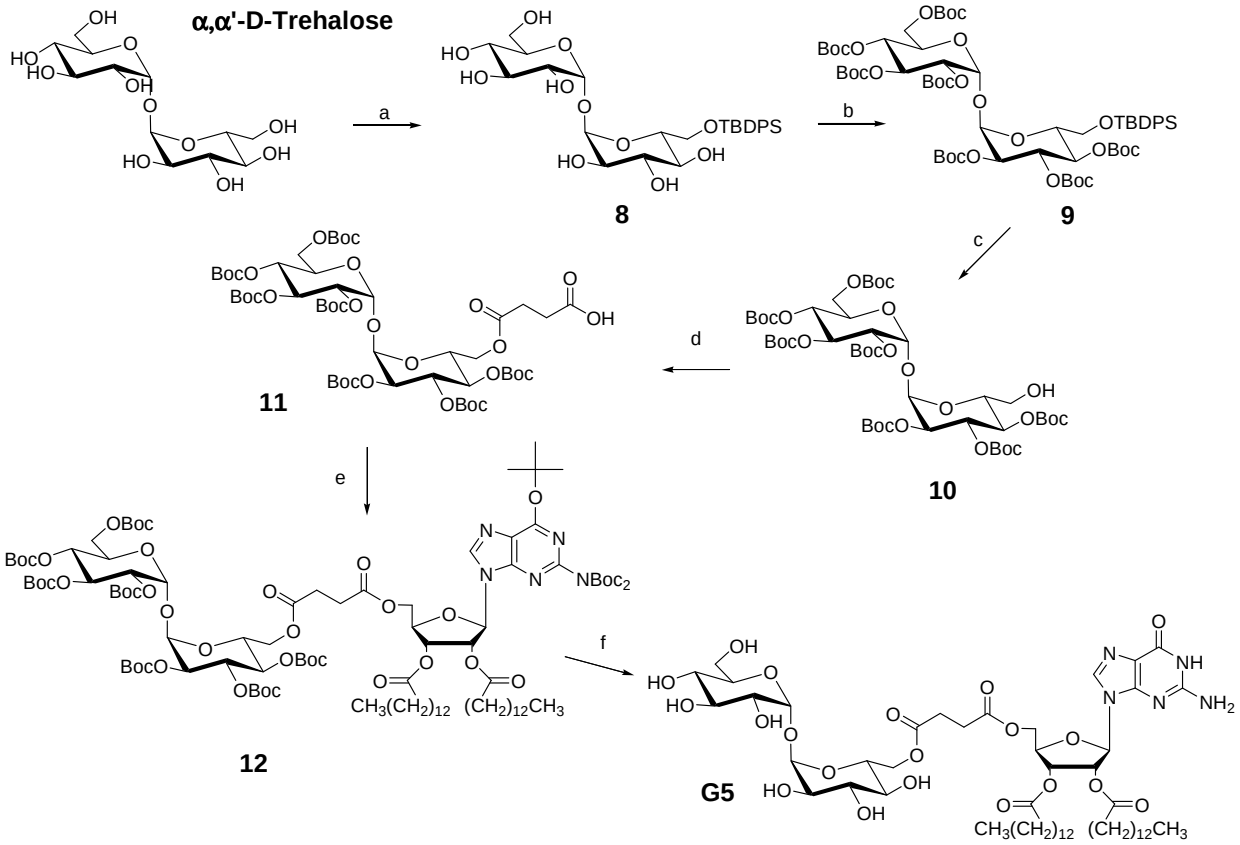


Scheme 2. Synthesis of **G1**-**G4**. Reaction conditions: **a)** TBDPSCl, imidazole, DMF, r.t., 2 h, 92%; **b)** $CH_3(CH_2)_{12}COOH$, DCC, DMAP, 12 h, r.t., 95%; **c)** $Et_3N{\cdot}3HF$, THF, 48 h, r.t., 80%; **d)** coupling with: for **7a** $CH_3O(CH_2CH_2O)_3CH_2COOH$; for **7b** Fmoc-Glu(O*t*Bu)OH; for **7d** Fmoc-Ser(OTrt)OH in CH_2Cl_2, DCC, 1.5 h, r.t.; for **7c** SO_3 x Et_3N, DMF, 45 min, r.t., 70%; **e)** i. Only for **7b** and **7d**: Fmoc removal with 10% piperidine in DMF, 20-40 min, r.t., ii. 10% TFA in CH_2Cl_2, 2 h, r.t., quant.

Hybrid nucleosides **G2** and **G4** were obtained using a similar procedure, based on the coupling of derivative **6** with commercially available Fmoc-protected α-amino acids - respectively, Fmoc-Glu(O*t*Bu)-OH and Fmoc-Ser(OTrt)-OH -, also in these cases mediated by DCC activation. The choice of Fmoc in lieu of Boc-protection could allow, in principle, the possibility of further derivatization of the final compounds, as, for instance, peptide elongation or attachment of special reported groups. For both compounds, Fmoc removal, cleanly achieved by reaction with 10% piperidine in DMF, was followed by treatment with 10% TFA in CH_2Cl_2 for 2 h at r.t., ensuring an effective one-pot deprotection of the guanine and of the side-chains of glutamic acid and serine residues, respectively masked as *tert*-butyl ester and trityl ether.

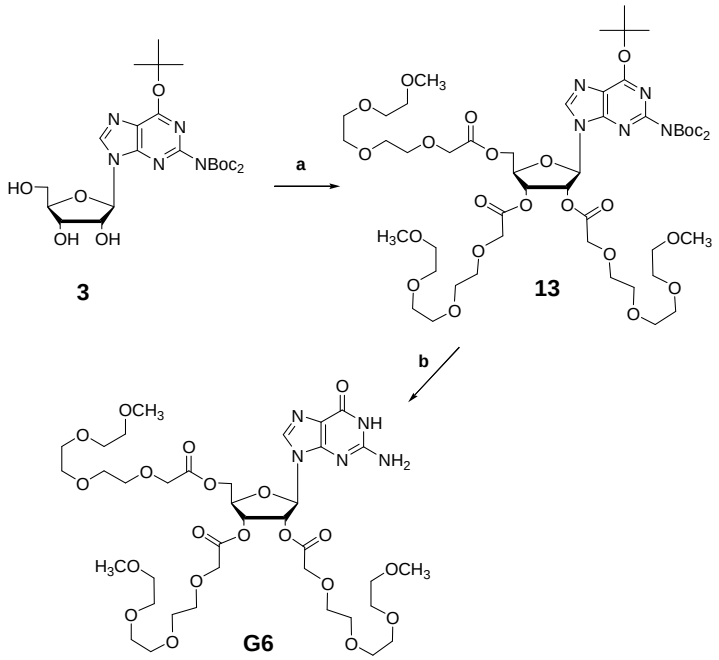
Compound **G3** was prepared reacting **6** with the commercially available complex $SO_3{\cdot}Et_3N$ in DMF, followed by final TFA deprotection, carried out as above described, thus leading to the desired 5'-O-sulfate derivative in 70% yields for the two steps.

Preparation of glyco-nucleolipid hybrid **G5** – incorporating a disaccharide residue, here chosen to confer a higher hydrophilic character to the final derivative compared to the series **G1**-**G4** - required prior synthetic elaboration of the disaccharide building block. To this purpose, α,α'-D-trehalose was selected on the basis of its favourable properties: it is an easily available, symmetrical, non-reducing disaccharide, not undergoing mutarotation. The latter issue is of the utmost relevance from a synthetic point of view, considering that the overall strategy to obtain the here described guanosine-containing nucleolipids is based on a final acidic treatment. The succinic group was here chosen as the linker providing the covalent attachment of the trehalose residue to the 5'-OH end of the guanosine scaffold. This was introduced on one primary hydroxyl of the disaccharide, previously subjected to the following derivatizations, depicted in Scheme 3: i) mono-silylation of the starting disaccharide, carried out with TBDPSCl and imidazole in DMF at 0 °C, giving **8** in 86 % yield; ii) exhaustive Boc protection of the remaining seven hydroxyls, by treatment with Boc_2O in CH_3CN, in the presence of TEA and DMAP, leading to **9** in 58 % yield; iii) removal of the *tert*-butyldiphenylsilyl protecting group by reaction with $Et_3N{\cdot}3HF$ in THF, providing **10** with 80 % yield; iv) reaction with succinic anhydride and cat. DMAP in THF, allowing the target succinylated trehalose derivative **11** in 78 % yield. This building block was finally used for the DCC-promoted coupling with nucleoside **6**, yielding target **12** in 86 % yield. As in the case of **G1**-**G4**, target compound **G5** was then obtained upon TFA treatment of **12**, quantitatively allowing in a single step the complete sugar and nucleobase deprotection.



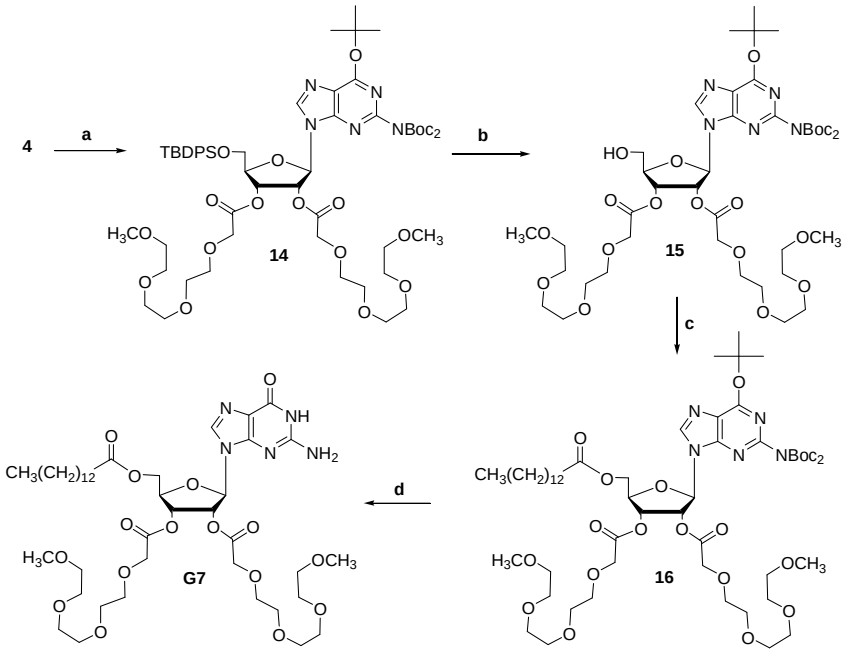
Scheme 3. Synthesis of **G5.** Reaction conditions: **a)** TBDPSCl, imidazole, DMF, 50 min, 0 °C, 86%; **b)** Boc_2O, TEA, DMAP, CH_3CN, 48 h, r.t., 58%; **c)** Et_3N x 3HF, THF, 72 h, r.t., 80%; **d)** succinic anhydride, DMAP, THF, 24 h, 78%; **e)** coupling with **6**, DCC, CH_2Cl_2, 3 h, r.t., 86%; **f)** 10% TFA in CH_2Cl_2, 3 h, r.t., quant.

Derivative **G6** was obtained in two steps starting from **3**, involving first exhaustive derivatization with monomethoxy(triethylene glycol) acetic acid, [17] in the presence of DCC, followed by TFA treatment of **13** for the nucleobase Boc removal, as depicted in Scheme 4.



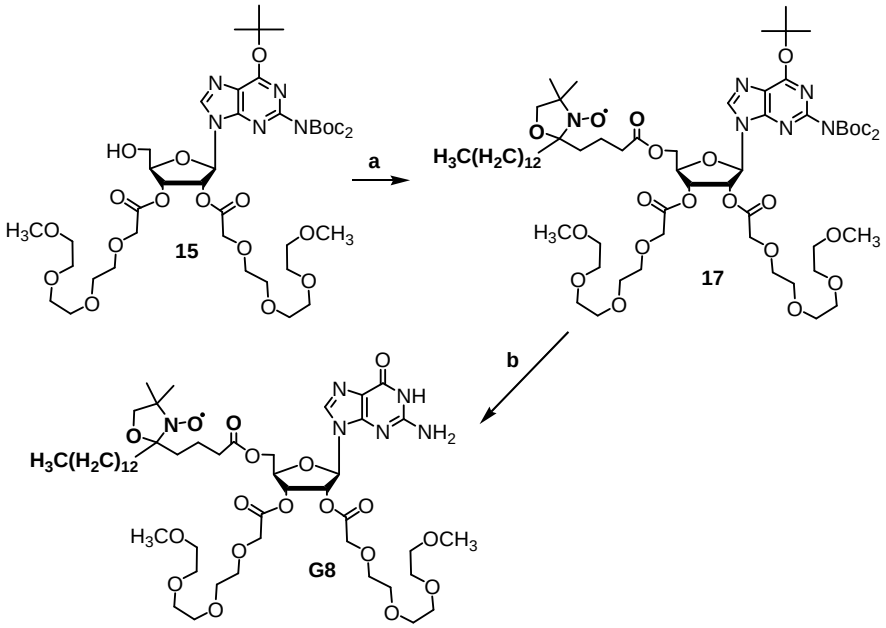
Scheme 4. Synthesis of **G6**. Reaction conditions: **a)** $CH_3O(CH_2CH_2O)_3CH_2COOH$, DCC, DMAP, CH_2Cl_2, 48 h, r.t., 42%; **b)** 10% TFA in CH_2Cl_2, 3 h, r.t., quant.

Finally, derivative **G7** was prepared from nucleoside **4**, upon esterification of both secondary hydroxyls with monomethoxy(triethylene glycol) acetic acid, [17] leading to compound **14** which was successively desilylated, coupled with myristic acid at the 5'-OH group and successively deprotected on the guanine moiety (Scheme 5).



Scheme 5. Synthesis of **G7**. Reaction conditions: **a)** $CH_3O(CH_2CH_2)_3CH_2COOH$, DCC, DMAP, CH_2Cl_2 1 h, r.t., 80%; **b)** $Et_3N{\cdot}3HF$, THF, 18 h, r.t., 56%; **c)** $CH_3(CH_2)_{12}COOH$, DCC, DMAP, CH_2Cl_2, 2 h, r.t., quant.; **d)** TFA 10% in CH_2Cl_2, 2.5 h, r.t., quant.

Guanosine derivative **G8**, bearing a spin label group on the lipophilic residue in position 5', was obtained starting from intermediate **15**. This was coupled with 5-doxyl-stearic acid to give intermediate **17,** that was then deprotected on the nucleobase as previously described (Scheme 6).



Scheme 6. Synthesis of **G8**. Reaction conditions: **a)** 5-Doxyl-stearic acid, DCC, DMAP, CH_2Cl_2, 3 h, r.t., quant.; **b)** 10% TFA in CH_2Cl_2, quant.

All the intermediate products here synthesized were purified by silica gel column chromatography and then fully characterized by ^{1}H and ^{13}C-NMR spectroscopy, as well as by ESI mass spectrometry. Final compounds **G1**-**G8**, obtained from the corresponding nucleobase-protected precursors by a simple treatment with 10% TFA in CH_2Cl_2 followed by repeated coevaporations from isopropanol till complete disappearance of excess TFA, were identified on the basis of their NMR and ESI-MS spectra.

4. Characterization of G1-G7

4.1 Study of the gelling abilities

Preliminary structural investigations have been attempted on **G1**-**G7** by means of NMR spectroscopy. [18] In all cases, NMR analysis showed for all the here prepared nucleosides dissolved in $CDCl_3$ dramatically broadened and very badly resolved signals in the ^{1}H and ^{13}C NMR spectra. In addition, a typical concentration-dependent response was obtained, which precluded the possibility to gain useful structural information. With time, the signals broadening was more and more marked, particularly in the imino protons region of ^{1}H NMR spectra, suggesting the presence of growing superstructures subjected to dynamic equilibria. Upon changing the solvent from $CDCl_3$ to the less structuring CD_3OD and/or varying the temperature in the range 288-328 °K, only minor benefits - in terms of resolution and line sharpness - could be observed. These results suggested for these compounds a strong tendency to form highly aggregated systems.

As a first step to determine the properties of the newly synthesized amphiphilic derivatives, we have then investigated their solubility and gelling abilities, and the related data are summarized in Table 1.

Strong tendency to aggregation – attributable to the well-known ability of guanine-based systems to realize a tight network of intermolecular hydrogen bonds - has been confirmed for all the synthesized samples, and their capacity as organogelators evaluated by the inversion method. [22] In all cases, analysis of these systems at concentrations lower than 20 mM did not indicate any tendency to form stable gels, and higher concentrations were therefore tested.

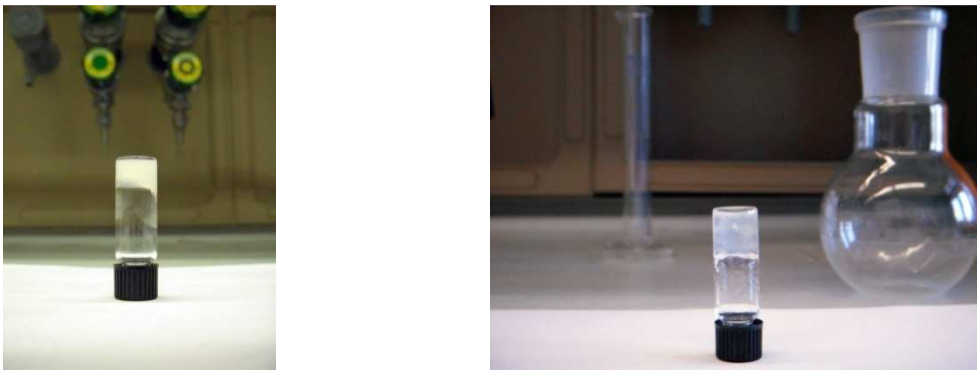
G1 and **G2** – examined in the range 3-7 % w/w (corresponding to *ca.* 25-70 mM) gave stable gels in polar solvents, as methanol, ethanol and CH_3CN (pictures of the gels obtained from **G1** and **G2** are shown in Figure 2). In comparable concentration conditions (5% w/w), lipophilic guanosine derivatives based on 2',3'-*O*-isopropylidenguanosine, carrying different *n*-alkylsilyl ethers at the 5'-position, synthesized by Araki and coworkers, [9] were found to gelate from *n*-decane, cyclohexane and *n*-hexane. Thus the here presented molecules offer new examples of low molecular weight organogelators,

which may profitably complement both the known lipophilic guanosines, and the hydrophilic, naturally occurring guanosine mono- and polyphosphates, covering a completely different polarity range of solvents, intermediate between highly apolar solvents and water. Analyzing the obtained results from a qualitative point of view, comparison of the systems **G1**-**G7**-**G6** indicates that two long aliphatic chains are essential for gelation; on the other hand, the presence of charged (as in **G3**) or markedly hydrophilic groups (as in **G5**) seems to negatively affect the formation of stable gels. **G4** is apparently a unique case, being essentially insoluble in all the most commonly used organic solvents, with the only exception of DMSO.

Table 1. Solubility properties of derivatives **G1-G7** in selected solvents

Solvent	**G1** 26 mM	**G1** 57 mM	**G2** 57 mM	**G2** 67 mM	**G3** 57mM	**G5** 57 mM	**G6** 57 mM	**G7** 57 mM
CH_3OH	**G[a] 3%**	**G 6%**	S	**G[a] 7%**	S	S	S	S
Ethanol	S	S	**G[a] 6%**	S	S	I	S	S
CH_3CN	S	**G**	S	S	I	I	S	I
n-hexane	S	S	P	P	P	S	P	S
Acetone	S	S	S	S	P	S	S	S
Cyclohexane	P	P	P	P	P	S	P	P
$CHCl_3$	S	S	S	S	P	S	S	S
THF	S	S	S	S	P	P	S	S
Dioxane	S	S	S	S	P	S	S	S
H_2O	P	P	P	P	P	P	P	P
Propanol	S	S	S	S	P	P	S	S
iso-propanol	S	S	S	S	P	P	S	S
CH_2Cl_2	S	S	S	S	P	S	S	S
AcOEt	S	S	S	S	P	S	S	P

[a] rapid gelification was achieved leaving the compound at +4 °C.; G = gel; S = homogeneous solution; I = insoluble.



a **b**

Figure 7. a) Gel of **G1** 6 wt% in CH_3CN; **b)** Gel of **G2** 6 wt% in EtOH.

4.2 CD studies

In order to get useful information on the conformational behaviour in organic solvents of the synthesized amphiphilic compounds, CD spectra were registered in the non structuring solvent $CHCl_3$. [18] For **G4**, only soluble in DMSO and therefore not directly comparable with the other systems under investigation, CD data were not acquired.

To preliminarily assess the ability of the here investigated systems to form G-tetrads, qualitative potassium picrate tests were carried out, following the procedure described for lipophilic guanosine derivatives. [23] In all cases, these colorimetric tests confirmed for the here investigated compounds, dissolved in $CHCl_3$, the ability to extract K^+ ions from aqueous solutions containing the yellow salt potassium picrate (Pic^-K^+), and transfer them in the organic phase. The yellow colour of the organic phases after this treatment was therefore indicative of G-tetrads formation by the guanosine derivatives: these attract from aqueous solutions K^+ ions, in turn transporting with them, to ensure electroneutrality, also picrate anions, otherwise highly insoluble in organic solvents. As reported in Figure 8, in sample **1** there is a control aqueous solution of Pic^-K^+ - markedly coloured in yellow - in contact with $CHCl_3$. In sample **2** there is a chloroform solution of a guanosine derivative (**G1**) which is left in contact with the aqueous Pic^-K^+ solution: here the transfer of the yellow colour from the aqueous to the chloroform phase is observed as a consequence of the transfer of Pic^-K^+. This qualitative test - indicative of G-tetrads formation by the guanosine derivatives – was positive for all the compounds here investigated, in the series **G1** – **G7**.

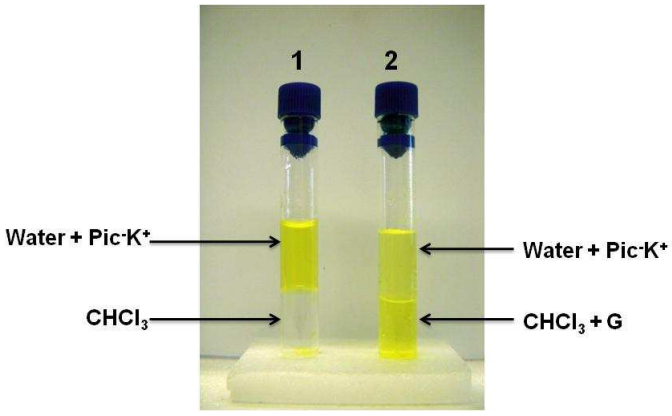


Figure 8. Test of potassium picrate on guanosine derivative **G1** here synthesized, showing its capability to extract potassium ions from the aqueous solution and transfer them into the $CHCl_3$ solution.

Then, compounds **G1**-**G3** and **G5**-**G7** were analyzed by CD spectroscopy in dilute $CHCl_3$ solutions, before and after the addition of potassium picrate. 2',3',5'-tri-*O*-acetylguanosine was also studied in parallel as a reference compound, having sugar modifications conferring high solubility in $CHCl_3$, but no relevant contribution to the self-assembling capability of the nucleoside; it can be therefore assumed – at least on a first basis - that its overall structuring process is only guided by the guanosine moiety and not influenced by the sugar decorations. This nucleoside, found to be positive to the potassium picrate test, showed a CD spectrum having a single negative band at 259 nm, very close to the maximum exhibited in the absorbance spectrum, registered in the same solvent (262 nm). This correspondence between the CD and UV spectra (see Figures 9 and 10) would point to a random distribution of the G-tetrads generated by this nucleoside in solution, not producing ordered three-dimensional structures, with characteristic orientations of the chiral centres.

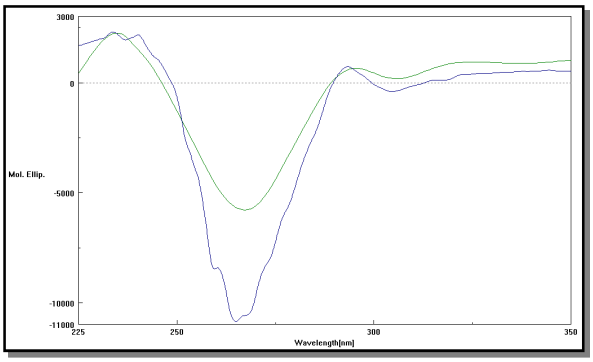


Figure 9. CD spectra of 2',3',5'-tri-*O*-acetyl guanosine (2.7 mM solution in $CHCl_3$). Green curve: 2',3',5'-tri-*O*-acetyl guanosine before treatment with potassium picrate; blue curve: 2',3',5'-tri-*O*-acetyl guanosine after treatment with potassium picrate. Negative band at λ = 262 nm.

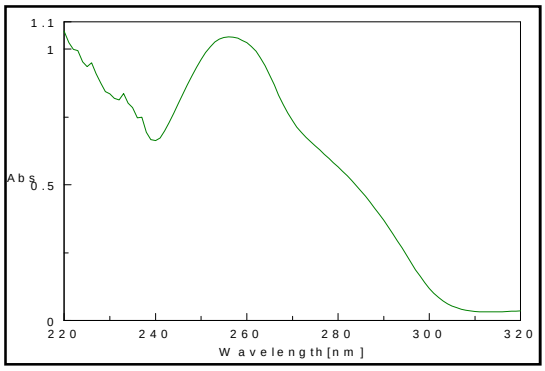


Figure 10. UV spectrum of 2',3',5'-tri-*O*-acetyl guanosine (0.12 mM solution in $CHCl_3$). Maximum absorption at λ = 259 nm.

Of the investigated amphiphilic derivatives, only **G3** and **G5** showed CD spectra reflecting – apart from the sign – the UV-vis ones, in analogy with 2',3',5'-tri-*O*-acetylguanosine (see Figures 11 and 12).

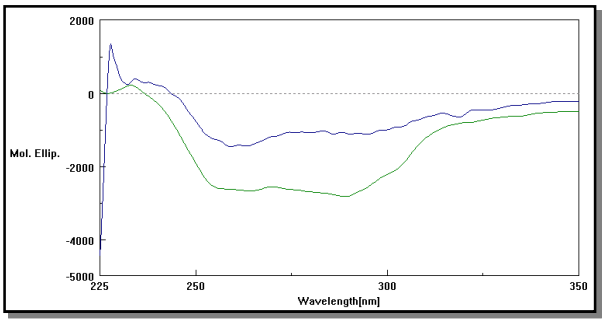


Figure 11. CD spectra of compound **G3** (2.7 mM solution in $CHCl_3$). Green curve: **G3** solution before treatment with potassium picrate; blue curve: **G3** solution after treatment with potassium picrate.

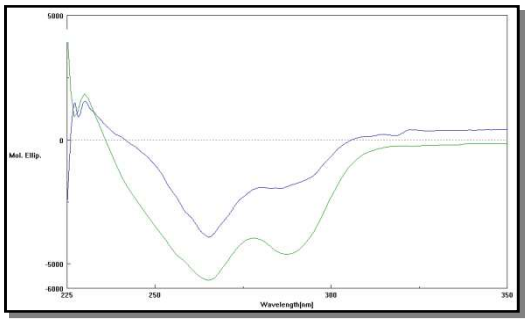


Figure 12. CD spectra of compound **G5** (2.7 mM solution in $CHCl_3$). Green curve: **G5** solution before treatment with potassium picrate; blue curve: **G5** solution after treatment with potassium picrate. Negative bands at λ = 265, 289 nm.

On the contrary, **G1**, **G2**, **G6** and **G7** – all having UV-vis profiles very similar to the reference compound (data not shown) - exhibited distinctive CD spectra with two intense bands of opposite sign, diagnostic of regular G-tetrads stacking organized in well defined octameric complexes, with C_4 or D_4 symmetry (Figures 13-16). Following the rationalization introduced by A. Randazzo and G. P. Spada, [24] the position of the opposite sign bands in the CD spectra of guanosine-based systems can provide information on the G-tetrads stacking not only in G-rich oligonucleotides – i.e. in systems where the stacked guanines within each strand are covalently linked by the sugar-phosphate oligonucleotide backbone – but also in non-covalently linked guanosine monomers. So the presence of a negative band centered at *ca.* 240 nm and a positive band at 260 nm is diagnostic of G-tetrads organized in homopolar stacking (also referred to as "H-to-T", with H and T representing the "head" and the "tail" surfaces of a G-tetrad, [25] respectively, here facing each other). A blue-shift of 20-30 nm of both bands is diagnostic of heteropolar, "H-to-H" or "T-to-T", stackings.

Inspection of the spectra allows us to conclude that **G6** exhibits homopolar G-tetrad stacking, while **G1**, **G2** and **G7** tend to generate octamers containing only "H-to-H"

or "T-to-T" G-tetrads stacking. Interestingly, **G1**, **G6** and **G7** showed spectra diagnostic of structuring of the G-tetrads even before addition of K^+ cations. Since their G-tetrad-forming ability has been previously proved by positive potassium picrate tests, these results can be explained only assuming that these systems display an extraordinary avidity for K^+, which could be extracted from the traces present in solvents, glassware, etc. and thus incorporated into stable complexes even before direct contact with potassium solutions. To confirm this hypothesis, the guanosine analogs **G1**, **G6** and **G7**, after the potassium picrate addition, were treated with an excess of the cryptand 4,7,13,16,21,24-hexaoxa-1,10-diazabicyclo[8.8.8]hexacosane, sequestering K^+ ions, and analyzed at CD, showing only residual CD signals.

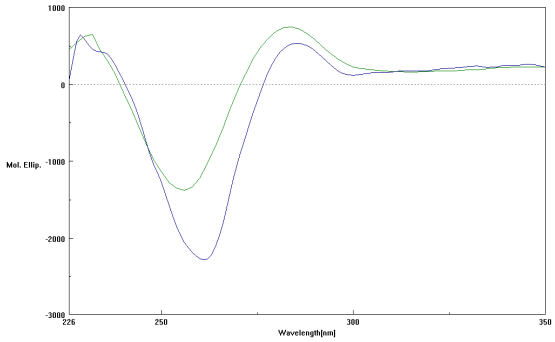


Figure 13. CD spectra of compound **G1** at 2.7 mM $CHCl_3$; green curve, before addition of potassium picrate, λ_{max} = 256, 283 nm; blue curve, after addition of potassium picrate, λ_{max} = 262, 285 nm.

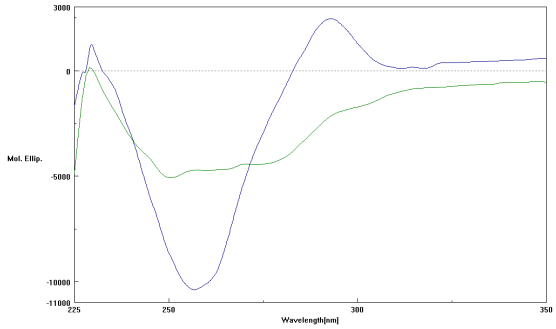


Figure 14. CD spectra of compound **G2** at 2.7 mM $CHCl_3$; green curve, before addition of potassium picrate; blue curve, after addition of potassium picrate, λ_{max} = 257, 293 nm

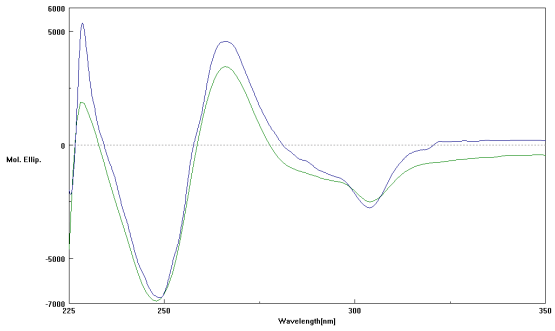


Figure 15. CD spectra of compound **G6** at 2.7 mM $CHCl_3$; green curve, before addition of potassium picrate; blue curve, after addition of potassium picrate; λ_{max} = 249, 265 nm.

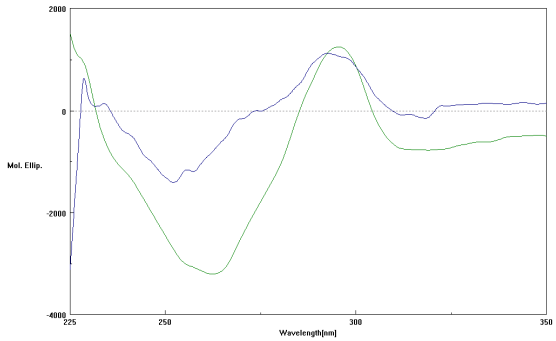


Figure 16. CD spectra of compound **G7** at 2.7 mM $CHCl_3$; green curve, before addition of potassium picrate; blue curve, after addition of potassium picrate; λ_{max} = 264, 295 nm.

CD melting experiments were carried out on **G1**, **G2**, **G6** and **G7** to get information on the thermal stability of the stacked G-tetrad systems. For all these compounds, under the studied conditions, a significant, non-linear decrease of the CD signal intensity on

increasing the temperature was observed, even if not in the form of well defined S-shaped curves, referring to a single transition with a clearly identified melting temperature. This would be expected in the case of a single, columnar G-quadruplex structure, involving all the G-tetrads stacked in a high hierarchical order. In all cases, curves ascribable to the superimposition of multiple transitions were registered, indicating the coexistence in solution of non homogeneous dissociating systems, probably containing a variable number of stacked G-tetrads, however essentially stable at room temperature (Figures 17 and 18).

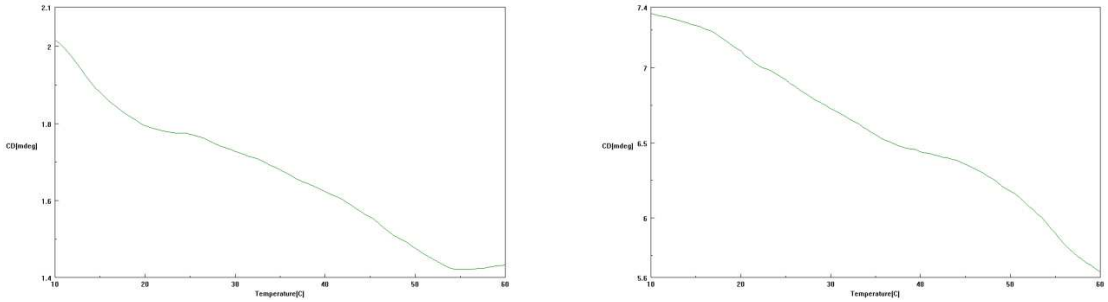


Figure 17. CD-monitored melting denaturation curves of compound **G1** (on the left, λ= 285 nm) and **G2** (on the right, λ= 293 nm), scan rate = 0.5 °C/min; temperature range: 10 – 60 °C.

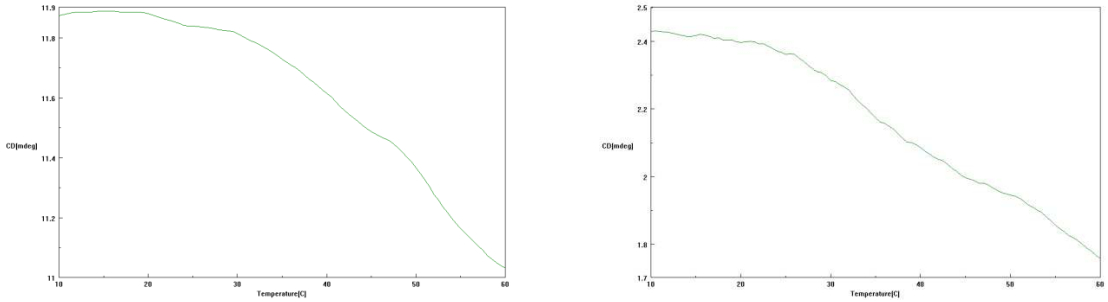


Figure 18. CD-monitored melting denaturation curves of compound **G6** (on the left, λ= 265 nm) and **G7** (on the right, λ= 295 nm), scan rate = 0.5 °C/min; temperature range: 10 – 60 °C

Taken together, these results show that in guanosine-based self-assembly of compounds **G1**-**G3** and **G5**-**G7** in $CHCl_3$, a non-negligible role is played by the sugar functionalization. When the ribose decorations are not able to elicit intermolecular interactions, as is the case of 2',3',5'-tri-*O*-acetylguanosine, no specific three-dimensional structuring is observed. For **G1**, **G2**, **G6** and **G7** CD spectra showed the presence of regular G-tetrads stacking. However, CD-monitored thermal denaturation curves do not support

the existence of single, well defined G-quadruplex complexes. Interestingly, of these four nucleosides, only **G6** - exclusively carrying polyether chains on the ribose moiety - is able to give homopolar stacking. On the contrary, the asymmetrically substituted guanosine derivatives, carrying both lipophilic and hydrophilic groups, tend to form heteropolar G-tetrads stacking. This different behaviour can be explained in terms of relevant contributions to the overall self-assembly given by alkyl and polyether chains, which, contributing to self-assembly in solution, may favour a peculiar three-dimensional arrangement of the stacked G-tetrads. Following this interpretation, **G3** and **G5** do not show a regular G-tetrads stacking in $CHCl_3$ solution, probably because of steric hindrance of the substituents, in **G5**, and presence of net negative charges, in **G3**, which prevent the formation of ordered aggregates.

4.3 Ionophoric activity

The group of J. T. Davis has recently demonstrated that potassium promoted self-assembly of lipophilic guanosine derivatives may contribute to stabilize ion channels active in phospholipid membranes. [6-8]

In consideration of the amphiphilic backbone of the here synthesized derivatives, we therefore investigated the ionophoric activity of compounds **G1**-**G7** using HPTS experiments, standard base-pulse assays which directly report H^+/OH^- transport and indirectly cation/anion transport. These studies were carried out in collaboration with the research group of prof. Paolo Tecilla (Department of Chemical Sciences, University of Trieste). [18]

For these experiments, liposomes (100 nm diameter) containing the pH-sensitive dye HPTS (8-hydroxypyrene-1,3,6-trisulfonic acid trisodium salt) were prepared in a pH 7.0 HEPES buffer containing 100 mM NaCl (or MCl and NaX in the experiments for cation and anion selectivity, respectively) and, once the ionophore was added, the external pH was suddenly brought to 7.6 by addition of NaOH (or MOH in experiments for cation selectivity) and the fluorescence emission of the dye recorded. The emission maximum for HPTS is about 510 nm in both the acidic and conjugate base forms. However, the excitation wavelength of the acidic form (403 nm) is significantly different from that of the conjugate base form (460 nm) so that the acid/base ratio is directly reflected in the emission intensity modulation produced by alternating excitation at the two wavelengths. [24] In essence, the modulated emission signal, obtained cycling the excitation wavelength and recording the emission at 510 nm, reports the effective pH within the liposome. The increase of HPTS fluorescence emission indicates basification of the liposome inner water pools which may be correlated to the H^+/OH^- transport and to the associated cation/anion symport or antiport. [26] The results obtained with compounds **G1**-**G7** are reported in Figure 19. Amphotericin B (AmB), a naturally occurring ionophore, was a positive control in these experiments.

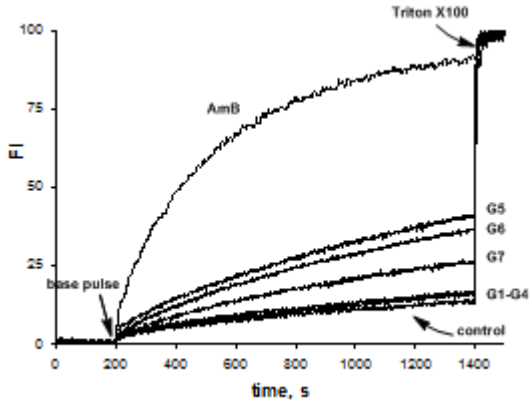


Figure 19. Normalized fluorescence change in HPTS emission (FI, λ_{ex} 403 and 460 nm, λ_{em} 510 nm) as a function of time after addition of the base (50 μL of 0.5 M NaOH) to 95:5 EYPC/EYPG LUVs (100 nm diameter) loaded with HPTS (0.1 mM HPTS, 0.17 mM total lipid concentration, 25 mM HEPES, 100 mM NaCl, pH 7.0, total volume 3 mL, 25 °C), in the presence of 2% **G1**-**G7** derivatives and Amphotericin B (AmB, 1%). The concentration of the ionophores is given in percent with respect to the total concentration of lipids. The control trace has been recorded in the absence of ionophore. EYPC = egg yolk phosphatidylcholine; EYPG egg yolk phosphatidylglycerol.

Amphiphilic guanosine derivatives **G1**-**G4** are mostly inactive, while an increasing ability to discharge the pH gradient is observed with compounds **G7**, **G6** and **G5**. The most active derivative is **G5** followed by **G6**, both characterized by a large hydrophilic portion, the disaccharide residue in **G5** and three monomethoxy(triethylene glycol) chains in **G6**. The relationship activity/hydrophilicty is further confirmed by the trend observed with **G1**, **G6** and **G7** which differ for the number of monomethoxy(triethylene glycol) residues: on increasing the number of polar chains the activity increases. A similar behaviour was previously observed in somehow structurally related CyPLOS derivatives studied in my research group. [27] Also in that case the activity was strongly dependent on the number of polar ethylene glycol chains appended to a rigid macrocyclic scaffold and the activity was ascribed to a destabilization of the phospholipid bilayer caused by the insertion of the polar chains. The ionophoric activity of compounds **G1**-**G7** is not influenced by the presence of K^+ ion. Experiments performed in the conditions of Figure 19 but using KCl instead of NaCl and KOH instead of NaOH, and therefore in the presence of a high concentration of potassium ion as the only externally added cation, show kinetic profiles for the pH discharge which are practically superimposable to those obtained in the presence of Na^+ ion (Figure 20).

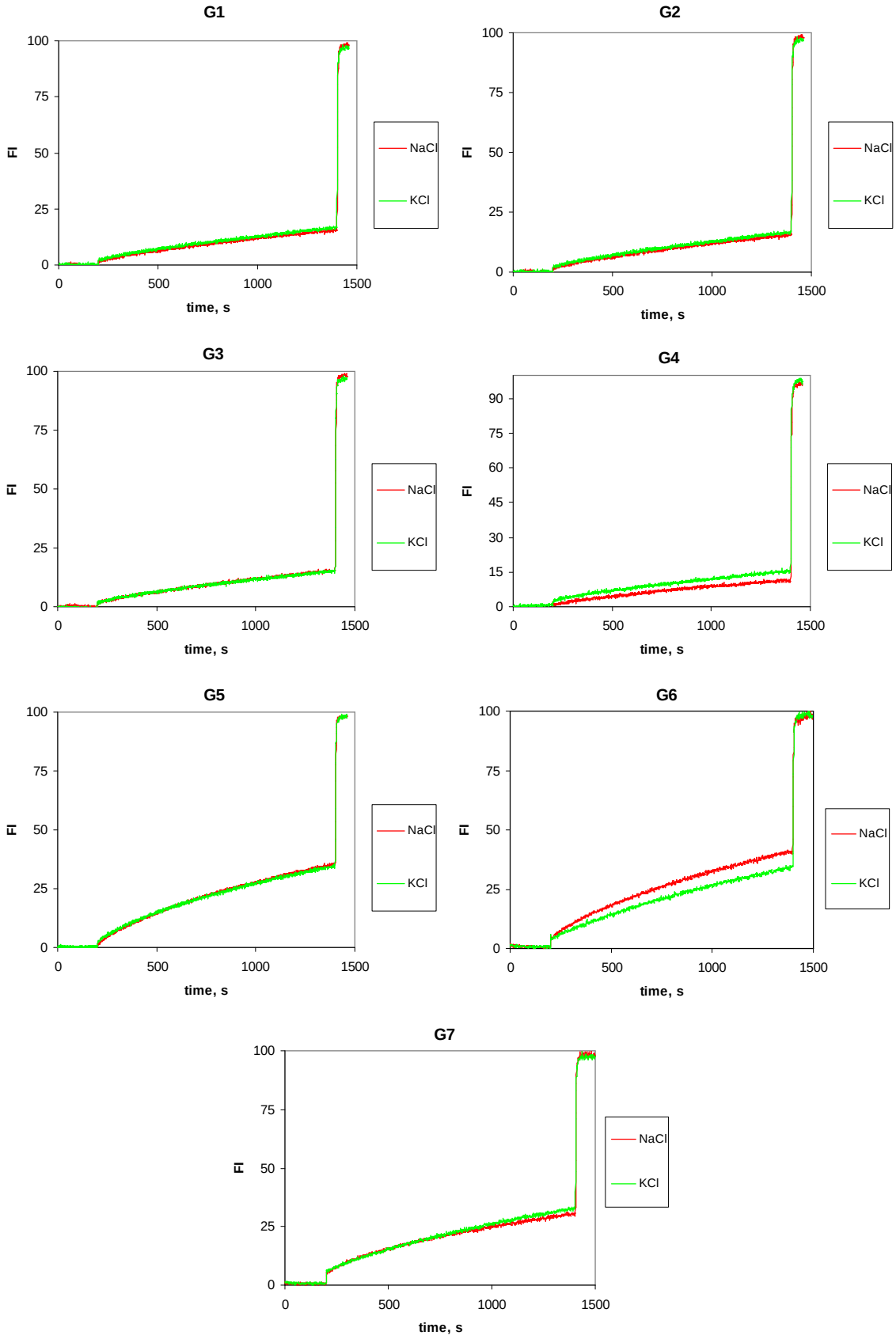


Figure 20. Ionophoric activity of **G1**-**G7** in the presence of Na^+ and K^+. Comparison between the activity of compounds **G1**-**G7** (2%) measured in buffer containing Na^+ (red curve) or K^+ (green curve). The experimental conditions are the same ones reported in Figure 19.

This suggests that in the phospholipid membranes the potassium ion is not able to promote the assembly of the guanosine derivatives as, on the contrary, is observed in more concentrated chloroform solutions. The low tendency of simple guanosine derivatives to form G-quartets in phospholipid membranes was not really unexpected and parallels the findings reported by the group of J. T. Davis. Stable G-quartets were indeed reported only in systems in which the guanosine moieties are covalently linked to a proper scaffold [6] or inserted in derivatives in which self-assembly is mainly driven by the formation of hydrogen bonds between bis-urea [7] or bis-carbamate subunits. [8] Apparently the sole interaction between potassium ion and guanosine is not strong enough to promote the self-assembling process in membranes, possibly due to the localization of the guanosine moieties close to the surface of the bilayer where water may compete efficiently with the formation of the G-tetrads.

In order to better characterize the ionophoric activity of the amphiphilic guanosine derivatives we investigated the influence on the transport process of the nature of the cation or anion present in solution in the case of **G5**, which is the most active derivative. The HPTS assay, here used, may indeed give indirect information on the cation and/or anion selectivity of the transport process. The transmembrane discharge of the externally built on pH gradient, which is signalled by HPTS fluorescence emission increase, may derive by H^+ efflux from the liposome inner water pool or by OH^- influx from the bulk water to the inner water pool of the liposome. In any case, this ion traffic has to be counterbalanced and this occurs through four overall processes: H^+/Na^+ antiport, OH^-/Cl^- antiport, H^+/Cl^- symport and Na^+/OH^- symport. Therefore, from the comparison of kinetic experiments performed in the presence of different cations and anions it is possible to gain indirect evidence on their effect on the transport process. The results obtained for the cation and anion selectivity experiments using as representative examples the group I alkali metals and the halogen anions are reported in Figure 21a and 21b, respectively.

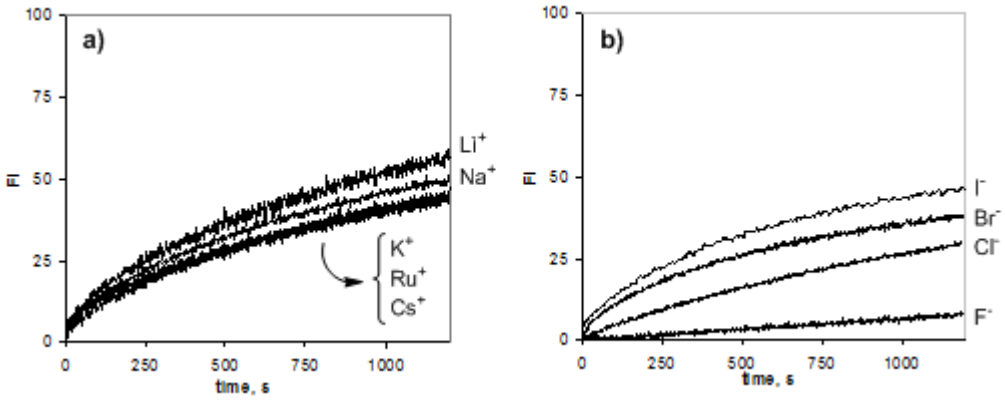


Figure 21. a) Cation selectivities for ionophore **G5** (3 % concentration), using the HPTS assay (100 mM MCl, pH 7.0, base pulse by addition of 50 µL of 0.5 M MOH; b) anion selectivities for ionophore **G5** (2 % concentration), using the HPTS assay (100 mM NaX, pH 7.0, base pulse by addition of 50 µL of 0.5 M NaOH). The kinetics are corrected for the spontaneous permeation of the different anions.

Inspection of Figure 21 shows that the rate of transport little depends on the cation present while it is strongly influenced by the anions and increases on going from fluoride to iodine, following the lyotropic sequence. This suggests that the transport process is governed by the translocation of the anion and in particular limited by its dehydration cost.

Taken together, the obtained results indicate that the amphiphilic guanosines **G1**-**G7** partition in the membrane probably positioning the guanine moiety and the sugar ring close to the surface of the liposome and the appended chains dipped in the membrane. When the ribose decorations have a balanced amphiphilic character, as in the case of **G5** and **G6**, they destabilize the membrane thus allowing the transit of small ions in a non selective process mainly governed by the dehydration cost of the anions. The ionophore acts in a monomeric form even in the presence of potassium ion which, under the studied conditions, is apparently unable to promote the aggregation of the guanosine derivatives. The ionophore is too short to span the membrane and the ionophoric activity is probably related to the formation of a disordered zone in the membrane characterized by a higher permeability. Ions are therefore able to cross this disordered zone in a process which is intrinsically little selective and strongly correlated to the lipophilicity of the anions.

4.4 In vitro screening of their antiproliferative activity

Antiviral [28] and anti-cancer [29] activity has been discovered in several G-rich oligonucleotides, and associated to the peculiar ability of these molecules to self-assemble in G-quartet-based superstructures, *in vivo* recognized by specific proteins. On the other hand, biological activities are searched in modified nucleosides since almost half a century, and, currently, large biomedical interest is also associated to nanoaggregated systems. [30] Since novel guanosine derivatives **G1**-**G7** seem to combine both the ability to form G-tetrads and to generate large nanoaggregates, i*n vitro* experiments have been carried out on a panel of cancer and non-cancer cell lines in order to investigate their bioactivity in a preliminary screening.

To this aim, cell lines were treated for 48 h with **G1**-**G7** at various concentrations in a growth inhibition assay and the cytotoxicity was determined in terms of IC_{50} (Table 2). [18]

The antiproliferative activity of these novel molecules was investigated in collaboration with the research group of Dr. Carlo Irace (Department of Experimental Pharmacology, Naples).

The results showed for **G1**-**G5** a moderate to weak selective cytotoxicity against both human MCF-7 breast adenocarcinoma and WiDr epithelial colorectal adenocarcinoma cells. In particular, **G3** and **G5** proved to be the most active compounds, exhibiting a significant antiproliferative profile against MCF-7 cells, with IC_{50} values of 22 and 17 μM, respectively. IC_{50} values within the micromolar range are generally consistent with the ability to interfere with cell viability and/or proliferation. Interestingly, in contrast to several anti-proliferative drugs, these compounds did not show *in vitro* unspecific cytotoxicity toward tumor and non-tumor cell lines. This might suggest potential specific interactions with the biological targets. Detailed structure-activity relationship studies will be carried out in the frame of an extensive investigation aimed at the development of nucleolipids with selective antineoplastic activity.

Table 2. Cytotoxicity profile of compounds **G1**-**G6** against cancer and non-cancer cell lines, IC_{50} (μM) [a].

Cell line [b]	**G1**	**G2**	**G3**	**G4**	**G5**	**G6**	**G7**
HeLa	$> 10^3$	198 ± 4	445 ± 6	130 ± 5	436 ± 9	$> 10^3$	$> 10^3$
WiDr	82 ± 6	90 ± 6	135 ± 10	130 ± 11	136 ± 10	$> 10^3$	1550 ± 15
MCF-7	185 ± 12	96 ± 5	22 ± 4	46 ± 7	17 ± 5	$> 10^3$	$> 10^3$
C6	$> 10^3$	245 ± 4	616 ± 5	152 ± 8	302 ± 5	880 ± 14	860 ± 18
3T3-L1	$> 10^3$	$> 10^3$	750 ± 4	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$

[a] IC_{50} values are expressed as mean ± SEM (n = 24) of three independent experiments. Bold values show IC_{50} < 100 μM. [b] HeLa, human cervical cancer cells; MCF-7, human breast adenocarcinoma cells; WiDr, human epithelial colorectal adenocarcinoma cells; C6, rat glioma cells; 3T3-L1, murine embryonic fibroblasts (no cancer cells).

5. Conclusions

In this study, a general and versatile synthetic strategy has been developed to obtain a library of amphiphilic, sugar-modified analogs of guanosine, generated by very simple and high yielding manipulations starting from a unique precursor, fully protected on the nucleobase. Large molecular diversity is ensured by insertion of different hydrophilic groups, as aminoacids, carbohydrates, oligoethers, and lipophilic residues, as fatty acid chains. In all cases, these appendages were attached to the ribose moiety through ester linkages, obtained under mild and very effective coupling conditions and, in principle, easily cleaved inside cells by esterases.

Qualitative potassium picrate tests confirmed for all the synthesized compounds the ability to extract potassium ions from aq. solutions into $CHCl_3$, and, thus, to form G-tetrads. CD analysis on dilute $CHCl_3$ solutions of these derivatives showed for **G1**, **G2**, **G6** and **G7** CD spectra distinctive of different G-tetrad self-assemblies: for **G6**, the CD spectrum was diagnostic of homopolar stacking of the G-quartets, while for the other three compounds the CD bands supported for G-quartets heteropolar stacking. Two derivatives, **G1** and **G2**, showed peculiar gelling abilities in polar solvents, as methanol, ethanol and acetonitrile, thus complementing lipophilic guanosine derivatives, known to form stable organogels in highly apolar solvents. When analyzed in their ion transport abilities, the activity was strongly correlated to the presence of a large hydrophilic portion in the molecule, with **G5** shown to be the most active compound. An interesting antiproliferative activity was found for **G3** and **G5** when tested on MCF-7 cancer cells, with IC_{50} values in the 20 μM range, while no cytotoxicity emerged on normal, control cells. Taken together, these data show that guanosine-based amphiphiles display a variety of peculiar properties, largely and finely tunable as a function of the nature and number of the ribose substituents, which render this class of compounds of great interest for both their biological/biomedical potential and innovative applications related to the development of novel self-assembling materials.

6. Experimental section

General Methods

All the reagents were of the highest commercially available quality and were used as received. TLC analyses were carried out on silica gel plates from Merck (60, F254). Reaction products on TLC plates were visualized by UV light and then by treatment with a 10 % $Ce(SO_4)_2/H_2SO_4$ aqueous solution. For column chromatography, silica gel from Merck (Kieselgel 40, 0.063-0.200 mm) was used. NMR spectra were recorded on Bruker WM-400, Varian Gemini 200 and Varian Inova 500 spectrometers, as specified. All the chemical shifts are expressed in ppm with respect to the residual solvent signal. Peak assignments have been carried out on the basis of standard ^{1}H-^{1}H COSY and HSQC experiments. The following abbreviations were used to explain the multiplicities: s = singlet; d = doublet; t = triplet; q = quartet; m = multiplet; b = broad; dd = double doublet. For the ESI MS analyses, a Waters Micromass ZQ instrument – equipped with an Electrospray source – was used in the positive and/or negative mode. MALDI TOF mass spectrometric analyses were performed on a PerSeptive Biosystems Voyager-De Pro MALDI mass spectrometer in the Linear mode using 3,4-dihydroxybenzoic acid as the matrix.

Synthesis of compound 2.

Nucleoside **1** (986 mg, 2.41 mmol) was dissolved in anhydrous CH_3CN (20 mL), then TEA (3.00 mL, 21.7 mmol), DMAP (147 mg, 1.20 mmol) and Boc_2O (4.21 g, 19.3 mmol) were sequentially added. The mixture was stirred at room temperature for 72 h, then the solvent was removed under reduced pressure and the residue purified on a silica gel column (eluent: *n*-hexane/acetone 65:35, v/v). The desired product **2** was obtained in 57% isolated yield (909 mg, 1.37 mmol).

2: oil, R_f = 0.5 [acetone/*n*-hexane, 2:3 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.09 [s, 1H, H-8]; 6.19 [d, *J* = 5.2 Hz, 1H, H-1']; 5.79 [dd, *J* = 5.4 and 4.4 Hz, 1H, H-2']; 5.56 [dd, *J* = 4.4 and 4.8 Hz, 1H, H-3']; 4.46 - 4.36 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 2.13 [s, 6H, 2x C$\underline{H}_3$CO]; 2.06 [s, 3H, C$\underline{H}_3$CO]; 1.71 [s, 9H, C(C$\underline{H}_3$)$_3$]; 1.41 [s, 18H, 2x OCOC(C$\underline{H}_3$)$_3$].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 170.2, 169.4 and 169.0 [3x $CH_3\underline{C}O$]; 161.1 [C-6]; 151.9 and 151.4 [C-2, C-4]; 150.6 [2x O$\underline{C}$OC(CH$_3$)$_3$]; 140.3 [C-8]; 121.4 [C-5]; 86.2 [C-1']; 84.3

[O<u>C</u>(CH$_3$)$_3$]; 82.9 [C-4']; 80.3 [2x OCO<u>C</u>(CH$_3$)$_3$]; 73.4 [C-2']; 70.7 [C-3']; 63.2 [C-5']; 28.3 [C(<u>C</u>H$_3$)$_3$]; 27.8 [2x OCOC(<u>C</u>H$_3$)$_3$]; 20.7, 20.5 and 20.3 [3x <u>C</u>H$_3$CO].

ESI-MS (positive ions): calcd. for $C_{30}H_{43}N_5O_{12}$ = 665.2908; found *m/z* 666.1 [M + H$^+$]; 688.1 [M + Na$^+$]; 704.1 [M + K$^+$]. **HRMS** (MALDI-TOF): calcd. for $C_{30}H_{43}N_5O_{12}Na$: 688.2806; found *m/z* 688.2819 (M+Na$^+$).

Synthesis of compound 3.

Nucleoside **2** (909 mg, 1.37 mmol) was dissolved in CH_3OH (8.0 mL) and 2.0 mL of a 29.1 % NH_4OH aq. solution were added. The solution was stirred at room temperature for 12 h, then the solvent was removed at reduced pressure and the residue washed three times with CH_2Cl_2/H_2O. The collected organic phases were dried over anhydrous Na_2SO_4. The desired product **3** was obtained in almost quantitative yields.

3: oil, R_f = 0.3 [$CHCl_3/CH_3OH$, 95:5 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.10 [s, 1H, H-8]; 5.87 [d, *J* = 6.0 Hz, 1H, H-1']; 4.78 [dd, *J* = 5.6 and 5.6 Hz, 1H, H-2']; 4.41 [broad signal, 1H, H-3']; 4.22 [broad signal, 1H, H-4']; 3.77 [m, 2H, H-5'$_a$ and H-5'$_b$]; 1.66 [s, 9H, C(C<u>H</u>$_3$)$_3$]; 1.39 [s, 18H, 2x OCOC(C<u>H</u>$_3$)$_3$].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 161.1 [C-6]; 151.1 [C-2]; 150.8 [C-4]; 150.6 [2x O<u>C</u>OC(CH$_3$)$_3$]; 142.8 [C-8]; 121.8 [C-5]; 91.0 [C-1']; 87.2 [C-4']; 84.5 [<u>C</u>(CH$_3$)$_3$]; 83.5 [2x OCO<u>C</u>(CH$_3$)$_3$]; 74.3 [C-2']; 71.6 [C-3']; 62.6 [C-5']; 28.2 [C(<u>C</u>H$_3$)$_3$]; 27.8 [2x OCOC(<u>C</u>H$_3$)$_3$].

ESI-MS (positive ions): calcd. for $C_{24}H_{37}N_5O_9$ = 539.2591; found *m/z* 561.76 [M+Na$^+$], 577.76 [M+K$^+$]. **HRMS** (MALDI-TOF): calcd. for $C_{24}H_{37}N_5O_9Na$ = 562.2489; found *m/z* 562.2452 [M+Na$^+$].

Synthesis of compound 4.

Nucleoside **3** (400 mg, 0.742 mmol) was dissolved in anhydrous DMF (3.0 mL), then TBDPSCl (227 μL, 0.890 mmol) and imidazole (116 mg, 1.70 mmol) were added. The solution was stirred for 2 h and then taken to dryness under reduced pressure. The crude was washed three times with CH_2Cl_2/H_2O, the collected organic phases were dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: *n*-hexane/AcOEt 55:45, v/v). The desired product **4** was obtained in 92% isolated yield (530 mg, 0.682 mmol).

4: oil, R_f = 0.3 [AcOEt/*n*-hexane, 1:1 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.11 [s, 1H, H-8]; 7.59 – 7.31 [m, 10H, aromatic proton]; 5.93 [d, *J* = 6.0 Hz, 1H, H-1']; 4.56 [dd, *J* = 5.5 and 6.0 Hz, 1H, H-2']; 4.46 [m, 1H, H-3']; 4.36 [m, 1H, H-4']; 3.82 [m, 2H, H-5'$_a$ and H-5'$_b$]; 1.71 [s, 9H, C(C$\underline{H}_3$)$_3$]; 1.47 [s, 18H, 2x OCOC(C$\underline{H}_3$)$_3$]; 0.93 [s, 9H, SiC(C$\underline{H}_3$)$_3$].

^{13}C NMR ($CDCl_3$, 100 MHz): δ 160.9 [C-6]; 150.7 [C-2, C-4]; 150.6 [2x O$\underline{C}$OC(CH$_3$)$_3$]; 139.9 [C-8]; 135.4, 132.4, 130.7, 129.8, 128.7, 127.7 [aromatic carbons]; 120.8 [C-5]; 90.6 [C-1']; 87.0 [C-4']; 84.0 [$\underline{C}$(CH$_3$)$_3$]; 83.4 [2x OCO$\underline{C}$(CH$_3$)$_3$]; 72.1 [C-2']; 68.0 [C-3']; 63.9 [C-5']; 28.2 [C($\underline{C}$H$_3$)$_3$]; 27.8 [2x OCOC($\underline{C}$H$_3$)$_3$]; 26.6 [SiC($\underline{C}$H$_3$)$_3$]; 19.0 [Si$\underline{C}$(CH$_3$)$_3$].

ESI-MS (positive ions): calcd. for $C_{40}H_{55}N_5O_9Si$ = 777.3769; found *m/z*: 778.04 [M + H^+]; 799.90 [M + Na^+]; 815.87 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{40}H_{55}N_5O_9SiNa$ = 800.3667; found *m/z*: 800.3691 [M + Na^+].

Synthesis of compound 5.

4 (467 mg, 0.600 mmol) was dissolved in anhydrous CH_2Cl_2 (5.0 mL), then myristic acid (411 mg, 1.80 mmol), DCC (372 mg, 1.80 mmol) and DMAP (109 mg, 0.900 mmol) were sequentially added. The solution was stirred for 12 h at room temperature, then the solvent was removed at reduced pressure and the residue purified on a silica gel column (eluent: *n*-hexane/AcOEt 4:1, v/v). The desired compound was obtained in 95% isolated yield (680 mg, 0.570 mmol).

5: oil, R_f = 0.3 [AcOEt/*n*-hexane,1:4 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.18 [s, 1H, H-8]; 7.70 - 7.36 [overlapped signals, 10H, aromatic protons]; 6.35 [d, *J* = 6.5 Hz, 1H, H-1']; 5.72 - 5.67 [overlapped signals, 2H, H-2' and H-3']; 4.22 [broad signal, 1H, H-4']; 3.91 [m, 2H, H-5'$_a$ and H-5'$_b$]; 2.35 [dd, *J* = 8.0 and 7.5 Hz, 2H, -C$\underline{H}_2$CO]; 2.24 [m, 2H, -C$\underline{H}_2$CO]; 1.72 [s, 9H, C(C$\underline{H}_3$)$_3$]; 1.63 [m, 4H, -C$\underline{H}_2$CH$_2$CO]; 1.40 [s, 18H, 2x OCOC(C$\underline{H}_3$)$_3$]; 1.30 - 1.24 [overlapped signals, 40H, -CH$_2$]; 0.88 [overlapped triplets, *J* = 7.0 and 6.5 Hz, 6H, 2x CH$_3$].

^{13}C NMR ($CDCl_3$, 100 MHz): relevant signals at δ 172.1, 171.4 [2x $\underline{C}$OOCH$_2$-]; 150.6 [2x O$\underline{C}$OC(CH$_3$)$_3$]; 142.0 [C-8]; 135.2, 132.1, 130.4, 129.0, 128.6, 127.4 [aromatic carbons]; 90.3 [C-1']; 88.2 [C-4']; 85.9 [$\underline{C}$(CH$_3$)$_3$]; 80.5 [2x OCO$\underline{C}$(CH$_3$)$_3$]; 72.8 [C-2']; 69.1 [C-3']; 64.0 [C-5']; 34.2, 33.9 [2x $\underline{C}$H$_2$COO myristic acid]; 31.8, 29.9, 29.7, 29.6, 29.5, 29.2, 28.8, 28.0, 27.5 [22x -CH$_2$, C($\underline{C}$H$_3$)$_3$, 2x OCOC($\underline{C}$H$_3$)$_3$]; 26.5 [SiC($\underline{C}$H$_3$)$_3$]; 22.5 [CH$_3$]; 19.0 [Si$\underline{C}$(CH$_3$)$_3$].

ESI-MS (positive ions): calcd. for $C_{68}H_{107}N_5O_{11}Si$ = 1197.7736; found *m/z* = 1198.6 [M + H^+]; 1222.1 [M + Na^+]. **HRMS** (MALDI-TOF): calcd. for $C_{68}H_{107}N_5O_{11}SiNa$ = 1220.7634; found *m/z*: 1220.7700 [M + Na^+].

Synthesis of compound 6.

To a solution of **5** (673 mg, 0.561 mmol), dissolved in 3.0 mL of anhydrous THF, $Et_3N{\cdot}3HF$ (1.83 mL, 11.2 mmol) was added under stirring at room temperature. After 48 h, the mixture was dried under reduced pressure, redissolved in CH_2Cl_2, transferred into a separatory funnel and then washed three times with H_2O/CH_2Cl_2. The collected organic phases were dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: *n*-hexane/AcOEt 7:3, v/v). The desired product **6** was obtained in 80% isolated yield (432 mg, 0.450 mmol).

6: oil, R_f = 0.7 [AcOEt/*n*-hexane, 3:7 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.93 [s, 1H, H-8]; 5.99 [overlapped signals, 2H, H-1' and H-2']; 5.65 [m, 1H, H-3']; 4.29 [broad signal, 1H, H-4']; 3.85 [m, 2H, H-5'$_a$ and H-5'$_b$]; 2.37 [t, *J* = 7.5 and 7.5 Hz, 2H, -C$\underline{H}_2$CO]; 2.20 [t, *J* = 7.5 and 7.5 Hz, 2H, -C$\underline{H}_2$CO]; 1.71 [s, 9H, C(C$\underline{H}_3$)$_3$]; 1.65 [m, 4H, -C$\underline{H}_2$CH$_2$CO]; 1.42 [s, 18H, 2x OCOC(C$\underline{H}_3$)$_3$]; 1.25 - 1.22 [overlapped signals, 40H, -CH_2]; 0.87 [overlapped triplets, *J* = 5.5 and 6.5 Hz, 6H, 2x CH_3].

^{13}C NMR ($CDCl_3$, 125 MHz): δ 172.3, 171.3 [2x $\underline{C}$OOCH$_2$-]; 161.4 [C-6]; 151.0 [C-2, C-4]; 150.3 [2x O$\underline{C}$OC(CH$_3$)$_3$]; 141.9 [C-8]; 122.5 [C-5]; 88.2 [C-1']; 86.2 [C-4']; 84.5 [$\underline{C}$(CH$_3$)$_3$]; 83.1 [2x OCO$\underline{C}$(CH$_3$)$_3$]; 72.4 [C-2']; 72.3 [C-3']: 62.4 [C-5']; 34.0, 33.5 [2x $\underline{C}$H$_2$COO myristic acid]; 31.8, 29.5, 29.4, 29.3, 29.2, 29.0, 28.9, 28.2, 27.7 [22x -CH_2, C($\underline{C}$H$_3$)$_3$, 2x OCOC($\underline{C}$H$_3$)$_3$]; 22.5 [CH_3].

ESI-MS (positive ions): calcd. for $C_{52}H_{89}N_5O_{11}$ = 959.6559; found *m/*z = 959.92 [M + H^+]. **HRMS** (MALDI-TOF): calcd. for $C_{52}H_{89}N_5O_{11}Na$ = 982.6456; found *m/z*: 982.6483 [M + Na^+].

Synthesis of compound 7a.

6 (216 mg, 0.225 mmol) was dissolved in 2.5 mL of anhydrous CH_2Cl_2. To the solution DCC (93 mg, 0.45 mmol), DMAP (5.0 mg, 0.04 mmol) and $CH_3O(CH_2CH_2O)_3CH_2COOH$ [19] (75 mg, 0.34 mmol) were added and the resulting solution was stirred for 1.5 h at room temperature. The crude was concentrated under reduced pressure and purified on a silica

gel column (eluent: *n*-hexane/acetone 65:35, v/v). The desired product **7a** was obtained in 97% isolated yield (254 mg, 0.218 mmol).

7a: oil, R_f = 0.5 [acetone/*n*-hexane, 35:65 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.17 [s, 1H, H-8]; 6.19 [d, *J* =5.6 Hz, 1H, H-1']; 5.82 [t, *J* = 5.6 and 5.6 Hz, 1H, H-2']; 5.56 [m, 1H, H-3']; 4.42 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 4.23 [s, 2H, -OCH_2CO-]; 3.72 - 3.40 [overlapped signals, 12H, 3x -OCH_2CH_2O-]; 3.35 [s, 3H, CH_3O-]; 2.37-2.20 [overlapped signals, 4H, 2x $COCH_2$-]; 1.69 - 1.22 [overlapped signals, 71H, $C(C\underline{H}_3)_3$, 2x $OCOC(C\underline{H}_3)_3$, -CH_2-]; 0.85 [overlapped triplets, *J* = 5.4 and 6.8 Hz, 6H, 2x CH_3].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 172.2, 171.8 [2x -$CH_2\underline{C}OO$ myristoyl residue]; 169.7 [-$OCH_2\underline{C}OO$-]; 161.0 [C-6]; 152.1 [C-2]; 151.3 [C-4]; 150.6 [2x $O\underline{C}OC(CH_3)_3$]; 140.7 [C-8]; 121.2 [C-5]; 85.8 [C-1']; 84.1 [C-4']; 82.9 [2x $OCO\underline{C}(CH_3)_3$]; 80.5 [$\underline{C}(CH_3)_3$]; 72.9 [C-2']; 71.8 [C-3']; 70.9, 70.7, 70.6, 70.5 [3x -OCH_2CH_2O-]; 68.6 [-$O\underline{C}H_2COO$-]; 63.6 [C-5']; 58.9 [CH_3O-]; 33.9, 33.8 [2x -$\underline{C}H_2COO$ myristoyl residue]; 31.8, 29.6, 29.4, 29.3, 29.2, 29.2, 29.1, 29.0, 28.2, 27.8 [22x -CH_2, $C(\underline{C}H_3)_3$, 2x $OCO\underline{C}(CH_3)_3$]; 22.6 [2x CH_3].

ESI-MS (positive ions): calcd. for $C_{61}H_{105}N_5O_{16}$ = 1163.7556; found *m/z* = 1164.50 [M + H^+]; 1186.47 [M + Na^+]; 1202.49 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{61}H_{105}N_5O_{16}Na$ = 1186.7454; found *m/z*: 1186.7479 [M + Na^+].

Synthesis of compound G1.

7a (254 mg, 0.218 mmol), dissolved in 2.0 mL of anhydrous CH_2Cl_2, was reacted with 0.25 mL TFA. The solution was stirred at room temperature for 2 h, then the solvent was removed under reduced pressure and the residue was coevaporated three times with *i*-PrOH. The desired product **G1** was thus obtained in quantitative yield as a pure compound (198 mg, 0.218 mmol).

G1: oil, R_f = 0.3 [acetone/*n*-hexane, 8:2 (v/v)].

^{1}H NMR (CD_3OD, 10.8 mM, 500 MHz): δ 7.92 [s, 1H, H-8]; 6.06 [d, *J* = 5.5 Hz, 1H, H-1']; 5.87 [t, *J* = 5.5 and 6.0 Hz, 1H, H-2']; 5.69 [t, *J* = 5.5 and 5.5 Hz, 1H, H-3']; 4.55 - 4.40 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 4.23 [m, 2H, -$OC\underline{H}_2CO$-]; 3.74 – 3.68, 3.65 – 3.63, 3.54 – 3.51 [m's, 4H each, 3x -OCH_2CH_2O-]; 3.59 [s, 3H, CH_3O-]; 2.44 - 2.32 [overlapped signals, 4H, 2x $OOCC\underline{H}_2$-]; 1.69 - 1.54 [overlapped signals, 4H, 2x $OOCCH_2C\underline{H}_2$-]; 1.43 - 1.22 [overlapped signals, 40H, 20x -CH_2-]; 0.86 [overlapped triplets, *J* = 5.5 and

6.5 Hz, 6H, 2x CH_3]. **^{13}C NMR** (CD_3OD, 10.8 mM, 100 MHz): δ 174.4, 174.2 [2x -CH_2<u>C</u>OO myristoyl residue]; 171.9 [-OCH_2<u>C</u>OO-]; 160.9 [C-6]; 157.1 [C-2]; 152.4 [C-4]; 138.3 [C-8]; 118.7 [C-5]; 89.0 [C-1']; 82.8 [C-4']; 75.1 [C-2']; 73.4 [C-3']; 72.3, 71.9, 71.5 [3x -OCH_2CH_2O-]; 70.0 [-O<u>C</u>H_2COO-]; 64.8 [C-5']; 58.2 [CH_3O-]; 35.2, 33.6 [2x -<u>C</u>H_2COO myristoyl residue]; 31.9, 31.3, 31.0, 30.7, 27.3, 27.0, 26.9, 26.5, 24.2 [22x -CH_2]; 14.9 [2x CH_3]. **ESI-MS** (positive ions): calcd. for $C_{47}H_{81}N_5O_{12}$ = 907.5882; found *m/z* = 908.57 [M + H^+]. **HRMS** (MALDI-TOF): calcd. for $C_{47}H_{81}N_5O_{12}Na$ = 931.1621; found *m/z*: 931.1666 [M + Na^+].

Synthesis of compound 7b.

6 (184 mg, 0.192 mmol) was dissolved in 2.0 mL of anhydrous CH_2Cl_2. To the solution, DCC (59 mg, 0.29 mmol), DMAP (1.2 mg, 0.0096 mmol) and Fmoc-Glu(O*t*Bu)OH (106 mg, 0.250 mmol) were added. The mixture was stirred for 3 h at room temperature, then the solvent was removed under reduced pressure and the residue was purified on a silica gel column (eluent: *n*-hexane/AcOEt 4:1, v/v). The desired product **7b** was obtained in 93% isolated yield (244 mg, 0.178 mmol).

7b: oil, R_f = 0.4 [acetone/*n*-hexane, 2:3 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.12 [s, 1H, H-8]; 7.73 - 7.21 [overlapped signals, 8H, Fmoc aromatic protons]; 6.10 [d, *J* = 4.8 Hz, 1H, H-1']; 5.77 [dd, *J* = 5.2 and 5.8 Hz, 1H, H-2']; 5.61 [dd, *J* = 4.2 and 5.0 Hz, 1H, H-3']; 4.54 - 4.27 [overlapped signals, 7H, H-4', H-5'$_a$ and H-5'$_b$, H-α Glu, CH_2 Fmoc, CH Fmoc]; 2.35 - 2.20 [overlapped signals, 8H, 2x CH_2CO myristoyl residue, CH_2-β and CH_2-γ Glu]; 1.67 [s, 9H, $C(CH_3)_3$]; 1.39 [s, 27H, 3x $OCOC(C\underline{H}_3)_3$]; 1.22 [s, 44H, 22x CH_2 myristoyl residues]; 0.87 [overlapped signals, *J* = 5.4 and 6.4 Hz, 6H, 2x CH_3].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 171.9, 171.8, 171.7, 171.5 [<u>C</u>OO myristoyl and Glu residues]; 160.9 [C-6]; 155.9 [<u>C</u>O Fmoc]; 151.8, 151.1 [C-2 and C-4]; 150.5 [2x O<u>C</u>$OC(CH_3)_3$]; 143.5, 127.5, 126.9, 125.0, 119.7 [Fmoc aromatic carbons]; 141.1 [C-8]; 121.1 [C-5]; 83.9 [C-1']; 82.7 [C-4']; 80.7, 80.2 [2x $OCO\underline{C}(CH_3)_3$ and $\underline{C}(CH_3)_3$]; 73.1 [C-2']; 70.3 [C-3']; 67.0 [CH_2 Fmoc]; 64.2 [C-5']; 53.6 [C-α Glu]; 46.9 [CH Fmoc]; 33.6, 33.5, 31.7, 29.5, 29.3, 29.2, 29.0, 28.1, 27.9, 27.7, 24.6, 24.5, 22.5 [2x <u>C</u>H_2CO myristoyl residues, CH_2-β and CH_2-γ Glu, 22x CH_2 myristoyl residues, $C(\underline{C}H_3)_3$, 2x $OCO\underline{C}(CH_3)_3$]; 13.9 [2x CH_3].

ESI-MS (positive ions): calcd. for $C_{76}H_{114}N_6O_{16}$ = 1366.8291; found *m/z* = 1311.02 [M - *t*-Bu + H^+]; 1333.13 [M - *t*-Bu + Na^+]; 1368.24 [M+H^+]. **HRMS** (MALDI-TOF): calcd. for $C_{76}H_{114}N_6O_{16}Na$ = 1389.8189; found *m/z*: 1389.8203 [M + Na^+].

Synthesis of compound G2.

7b (64 mg, 0.060 mmol), dissolved in anhydrous DMF (0.45 mL), was reacted with piperidine (50 µL). The solution was stirred at room temperature for 20 min and then the solvent was removed under reduced pressure. The residue was purified on a silica gel column (eluent: *n*-hexane/AcOEt 1:1, v/v). Then the product was dissolved in 0.90 mL of anhydrous CH_2Cl_2 and 0.10 mL of TFA. The solution was stirred at room temperature for 2 h, then the solvent was removed under reduced pressure and the residue coevaporated three times with *i*-PrOH without further purification. The desired product **G2** was obtained in 80% overall yield for the two steps (40 mg, 0.048 mmol).

G2: amorphous solid, R_f = 0.1 [$CHCl_3$/CH_3OH, 9:1 (v/v)].

^{1}H NMR (CD_3OD, 500 MHz): δ 7.90 [s, 1H, H-8]; 6.08 [d, *J* = 5.5 Hz, 1H, H-1']; 6.00 [apparent triplet, *J* = 5.5 Hz, 1H, H-2']; 5.68 [apparent triplet, *J* = 5.0 Hz, 1H, H-3']; 4.60 [m, 2H, H-5'$_a$ and H-5'$_b$]; 4.47 [broad signal, 1H, H-4']; 4.19 [t, *J* = 6.5 Hz, 1H, H-α Glu]; 2.52 [m, 2H, CH_2-γ Glu]; 2.44 - 2.34 [m, 4H, 2x CH_2CO myristoyl residues]; 2.27 - 2.14 [m, 2H, CH_2-β Glu]; 1.65 – 1.54 [m, 4H, 2x C$\underline{H}_2CH_2CO$ myristoyl residues]: 1-40 - 1.28 [overlapped signals, 40H, 20x CH_2]; 0.90 [overlapped triplets, *J* = 6.0 and 7.0 Hz, 6H, 2x CH_3].

^{13}C NMR (CD_3OD, 100 MHz): δ 174.5, 174.2, 170.8 [$\underline{C}OO$ myristoyl and Glu residues]; 163.2 [C-6]; 159.7 [C-2]; 156.1 [C-4]; 138.7 [C-8]; 118.3 [C-5]; 86.3 [C-1']; 81.7 [C-4']; 74.5 [C-2']; 72.1 [C-3']; 67.0 [C-5']; 53.8 [C-α Glu]; 35.2, 35.1 [2x $\underline{C}H_2CO$ myristoyl residues]; 33.6 [CH_2-γ Glu]; 31.3, 31.0, 30.8, 30.7, 30.7, 30.4, 28.8, 26.4, 24.2 [CH_2-β Glu, 22x CH_2 myristoyl residues]; 15.0 [2x CH_3].

ESI-MS (positive ions): calcd. for $C_{43}H_{72}N_6O_{10}$ = 832.5310; found *m/z* = 871.50 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{43}H_{72}N_6O_{10}Na$ = 855.5208; found *m/z*: 855.5253 [M + Na^+].

Synthesis of compound 7c.

6 (121 mg, 0.126 mmol) was dissolved in 1.0 mL of anhydrous DMF and then treated with the complex SO_3·Et_3N (25 mg, 0.139 mmol). The solution was stirred at room temperature for 45 min and then the solvent was removed under reduced pressure. The residue was

purified on a silica gel column (eluent: AcOEt/CH_3OH 95:5, v/v). The desired product **7c** was obtained in the form of triethylammonium salt in 70% isolated yield (91 mg, 0.088 mmol).

7c: oil, R_f = 0.6 [AcOEt/CH_3OH, 9:1 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): δ 10.85 [broad signal, 1H, $(CH_3CH_2)_3N\underline{H}^+$]; 9.70 [s, 1H, H-8]; 6.16 [d, *J* = 5.5 Hz, 1H, H-1']; 5.74 [dd, *J* = 5.5 and 5.0 Hz, 1H, H-2']; 5.65 [broad signal, 1H, H-3']; 4.44 - 4.31 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 3.15 [m, 6H, $(CH_3C\underline{H}_2)_3NH^+$]; 2.32 - 2.25 [overlapped signals, 4H, 2x CH_2CO myristoyl residues]; 1.61 - 1.54 [overlapped signals, 27H, $C(CH_3)_3$, 2x $OCOC(CH_3)_3$]; 1.40 [t, *J* = 7.5 and 7.5 Hz, 9H, $(C\underline{H}_3CH_2)_3NH^+$]; 1.28 - 1.24 [overlapped signals, 44H, 22x CH_2]; 0.88 [overlapped triplet, *J* = 6.5 and 6.0 Hz, 6H, 2x CH_3].

^{13}C NMR ($CDCl_3$, 100 MHz): significant signals a δ 172.6, 172.2 [2x $\underline{C}OCH_2$]; 87.2 [C-1']; 83.4 [C-4']; 75.2 [C-2']; 73.7 [C-3']: 66.2 [C-5']; 46.0 [$N(\underline{C}H_2CH_3)_3$]; 33.8 [2x $\underline{C}H_2CO$ overlapped signals]; 31.8, 29.6, 29.2, 27.9, 27.5, 24.7, 22.6 [22x -CH_2, $C(\underline{C}H_3)_3$, 2x $OCOC(\underline{C}H_3)_3$]; 14.0 [2x CH_3]; 8.6 [$N(CH_2\underline{C}H_3)_3$].

ESI-MS (negative ions): calcd. for $C_{52}H_{88}N_5O_{14}S^-$ = 1038.6054; found *m/z* = 1037.42 [M^-].

HRMS (MALDI-TOF): calcd. for $C_{52}H_{88}N_5O_{14}S^-$ = 1038.6054; found *m/z*: 1038.6081 [M^-].

Synthesis of compound G3.

7c (91 mg, 0.088 mmol), dissolved in 0.90 mL of anhydrous CH_2Cl_2, was reacted with 0.10 mL TFA. The solution was stirred at room temperature for 1.5 h, then the solvent was removed under reduced pressure and the residue coevaporated three times with *i*-PrOH. The desired product **G3** was obtained in quantitative yield as a pure compound (69 mg, 0.088 mmol).

G3: amorphous solid, R_f = 0.2 [AcOEt/CH_3OH, 9:1 (v/v)].

^{1}H NMR (CD_3OD, 500 MHz): δ 9.10 [broad signal, 1H, $(CH_3CH_2)_3N\underline{H}^+$]; 8.73 [s, 1H, H-8]; 6.18 d, *J* = 5.5 Hz, 1H, H-1']; 5.87 [m, 1H, H-2']; 5.66 [broad signal, 1H, H-3']; 4.33 - 4.06 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 3.23 [m, 6H, $(CH_3C\underline{H}_2)_3NH^+$]; 2.42 - 2.27 [overlapped signals, 4H, 2x CH_2CO]; 1.66 – 1.46 [broad signal, 4H, 2x $C\underline{H}_2CH_2CO$]; 1.32 - 1.18 [overlapped signals, 20H, 10x CH_2]; 1.14 [m, 9H, $(C\underline{H}_3CH_2)_3NH^+$]; 0.89 [overlapped triplets, *J* = 6.0 and 7.0 Hz, 6H, 2x CH_3].

^{13}C NMR (CD_3OD, 100 MHz): δ 174.5 [2x CO myristoyl residues]; 161.8 [C-6]; 155.5 [C-2]; 152.1 [C-4]; 138.7 [C-8]; 119.4 [C-5]; 80.5 [C-1']; 76.7 [C-4']; 73.9 [C-2']; 71.6 [C-3']; 70.1 [C-5']; 48.4 [$(CH_3\underline{C}H_2)_3NH^+$]; 35.3 [2x $\underline{C}H_2CO$]; 33.6 [2x $\underline{C}H_2CH_2CO$]; 31.3, 31.1, 30.7, 29.5, 28.8, 26.5, 24.2, 22.6, 22.1 [20x CH_2]; 15.0 [2x CH_3]; 9.7 [$(\underline{C}H_3CH_2)_3NH^+$].

ESI-MS (negative ions); calcd. for $C_{38}H_{64}N_5O_{10}S^-$ = 782.4379; found *m/z* = 782.11 [M^-].

HRMS (MALDI-TOF): calcd. for $C_{38}H_{64}N_5O_{10}S^-$ = 782.4379; found *m/z*: 782.4354 [M^-].

Synthesis of compound 7d.

6 (71 mg, 0.074 mmol) was dissolved in 2.5 mL of anhydrous CH_2Cl_2. To the solution DCC (23 mg, 0.11 mmol), DMAP (1.0 mg, 0.0037 mmol) and Fmoc-Ser(OTrt)-OH (51 mg, 0.089 mmol) were sequentially added. The solution was stirred at room temperature for 1.5 h, then the solvent was removed under reduced pressure and the residue was purified on a silica gel column (eluent: AcOEt/*n*-hexane 3:7, v/v). The desired product **7d** was obtained in quantitative yield (112 mg, 0.074 mmol).

7d: oil, R_f = 0.5 [AcOEt/*n*-hexane, 4:6 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.98 [s, 1H, H-8]; 7.75 - 7.15 [overlapped signals, 23H, aromatic protons of Trt and Fmoc]; 6.00 [d, *J* = 4.2 Hz, 1H, H-1']; 5.75 - 5.57 [overlapped signals, 2H, H-2' and H-3']; 4.66 - 4.21 [overlapped signals, 7H, H-4', CH_2 and CH Fmoc, CH-α and CH_2-β Ser]; 3.55 [m, 2H, H-5'$_a$ and H-5'$_b$]; 2.40 - 2.22 [overlapped signals, 4H, 2x CH_2CO myristoyl residues]; 1.71 [s, 9H, $C(C\underline{H}_3)_3$]; 1.41 [s, 18H, 2x $OCOC(C\underline{H}_3)_3$]; 1.36 - 1.16 [overlapped signals, 44H, 22x CH_2]; 0.88 [overlapped triplets, *J* = 6.0 and 6.8 Hz, 6H, 2x CH_3].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 171.9, 170.1 [3x $\underline{C}OO$ myristoyl and Ser residues]; 160.9 [C-6]; 155.9 [$\underline{C}O$ Fmoc]; 151.8, 151.5 [C-2 and C-4]; 150.6 [2x $O\underline{C}OC(CH_3)_3$]; 143.7, 143.2, 140.7, 128.4, 127.6, 127.2, 127.0, 125.1, 119.9 [Trt and Fmoc aromatic carbons]; 141.2 [C-8]; 121.4 [C-5]; 86.8 [C-1']; 84.0 [$\underline{C}(CH_3)_3$]; 82.9 [C-4']; 80.1 [2x $OCO\underline{C}(CH_3)_3$]; 73.3 [C-2']; 71.4 [C-3']; 67.3 [CH_2 Fmoc]; 64.3 [C-5']; 63.8 [C-β Ser]; 54.5 [C-α Ser]; 47.0 [CH Fmoc]; 33.7 [2x $\underline{C}H_2CO$ myristoyl residues]; 31.9, 29.6, 29.4, 29.3, 28.3, 27.8, 24.7, 22.6 [22x -CH_2, $C(\underline{C}H_3)_3$, 2x $OCOC(\underline{C}H_3)_3$]; 14.0 [2x CH_3].

ESI-MS (positive ions): calcd. for $C_{89}H_{118}N_6O_{15}$ = 1510.8655; found *m/z* = 1550.21 [$M + K^+$].

HRMS (MALDI-TOF): calcd. for $C_{89}H_{118}N_6O_{15}Na$ = 1533.8553; found *m/z*: 1533.8582 [$M + Na^+$].

Synthesis of compound G4.

7d (106 mg, 0.070 mmol), dissolved in 900 μL of anhydrous DMF, was treated with 100 μL of piperidine. The solution was stirred at room temperature for 40 min, then the solvent was removed under reduced pressure and the residue purified on a silica gel column (eluent: AcOEt/*n*-hexane 1:4, v/v). The obtained product was successively dissolved in 0.90 mL of anhydrous CH_2Cl_2 and 0.10 mL of TFA and stirred at room temperature for 2 h. Successively the solvent was removed under reduced pressure, the residue coevaporated three times with *i*-PrOH and then washed with acetone. The desired product was obtained in 46% overall yield for the two steps (25 mg, 0.032 mmol).

G4: amorphous solid, R_f = 0.1 [AcOEt/CH_3OH, 9:1 (v/v)].

^{1}H NMR (DMSO-d6, 500 MHz): δ 10.8 [s, 1H, NH-1]; 7.96 [s, 1H, H-8]; 6.56 [broad signal, 2H, NH_2-2]; 5.97 [d, *J* = 6.5 Hz, 1H, H-1']; 5.77 [dd, *J* = 6.0 and 6.5 Hz, 1H, H-2']; 5.59 [broad signal, 1H, OH]; 5.55 [apparent triplet, *J* = 4.0 Hz, 1H, H-3']; 4.55 [m, 1H, H-4']; 4.38 [m, 2H, H-5'$_a$ and H-5'$_b$]; 4.25 [broad signal, 1H, CH-α Ser]; 3.84 [m, 2H, CH_2-β Ser]; 2.42 - 2.22 [overlapped signals, 4H, 2x CH_2CO myristoyl residues]; 1.55 - 1.44 [overlapped signals, 4H, CH_2-β myristoyl residue]; 1.26 – 1.18 [overlapped signals, 44H, 22x CH_2]; 0.85 [overlapped triplets, *J* = 6.5 and 7.0 Hz, 6H, 2x CH_3].

^{13}C NMR (DMSO-*d6*, 125 MHz): δ 171.9, 171.6 [2x COO myristoyl residues]; 167.9 [COO Ser]; 156.6 [C-6]; 154.0 [C-2]; 151.2 [C-4]; 144.3 [C-8]; 116.7 [C-5]; 83.9 [C-1']; 79.4 [C-4']; 71.8 [C-2']; 69.6 [C-3']; 64.9 [C-5']; 59.5 [CH_2-β Ser]; 54.2 [CH-α Ser]; 33.2, 33.0 [2x CH_2CO myristoyl residues]; 32.9, 31.3, 29.0, 28.9, 28.8, 28.7, 28.5, 28.3, 24.4, 24.2, 22.1 [22x - CH_2]; 13.9 [2x CH_3].

ESI-MS (positive ions): calcd. for $C_{41}H_{70}N_6O_9$ = 790.5204; found *m/z* = 791.12 [M + H^+].

HRMS (MALDI-TOF): calcd. for $C_{41}H_{70}N_6O_9Na$ = 814.0191; found *m/z*: 814.0210 [M + Na^+].

Synthesis of compound 8.

Commercially available α,α'-D-trehalose 2xH_2O (1.14 g, 3.33 mmol) was previously dehydrated by heating under vacuum at 100 °C for 30 min. Then the solid was cooled at room temperature in anhydrous N_2 atmosphere and successively suspended in 10 mL of anhydrous DMF. To the stirred suspension imidazole (136 mg, 2.00 mmol) and TBDPSCl (432 μL, 1.66 mmol) were sequentially added. The reaction mixture was stirred at 0 °C for 50 min, then the solvent was removed under vacuum and the residue purified on a silica

gel column (eluent: $CHCl_3/CH_3OH$ 85:15, v/v). The desired product **8** was obtained in 86% isolated yield (828 mg, 1.43 mmol).

8: amorphous solid, R_f = 0.6 [$CHCl_3/CH_3OH$, 7:3 (v/v)].

^{1}H NMR (CD_3OD, 500 MHz): δ 7.72-7.37 [overlapped signals, 10H, aromatic protons of TBDPS]; 5.18 [d, *J* = 3.7 Hz, 1H, H-1]; 5.16 [d, *J* = 3.7 Hz, 1H, H-1']; 3.95 - 3.81 [overlapped signals, 7H, H-3, H-5, H-6_a, H-6_b, H-3', H-5', H-$6'_a$]; 3.71 [dd, *J* = 5.5 and 5.5 Hz, 1H, H-$6'_b$]; 3.56-3.52 [overlapped signals, 2H, H-2, H-4]; 3.48 [dd, *J* = 3.5 and 3.5 Hz, 1H, H-2']; 3.35 [buried under the residual solvent signal, H-4']; 1.03 [s, 9H, 3x $SiC(C\underline{H}_3)_3$].

^{13}C NMR (CD_3OD, 125 MHz): δ 137.3, 137.2, 136.3, 135.6, 135.2, 131.2, 129.1, 121.9 [aromatic carbons of TBDPS]; 95.4 [C-1]; 95.2 [C-1']; 75.2 [C-3]; 75.1 [C-3']; 74.5 [C-5]; 74.3 [C-5']; 73.7 [C-2]; 73.6 [C-2']; 72.4 [C-4]; 72.2 [C-4']; 64.8 [C-6]; 63.1 [C-6']; 27.8 [$SiC(\underline{C}H_3)_3$]; 20.6 [$Si\underline{C}(CH_3)_3$].

ESI-MS (positive ions): calcd. for $C_{28}H_{40}O_{11}Si$ = 580.2340; found *m/z* = 604.12 [M + Na^+]; 650.00 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{28}H_{40}O_{11}SiNa$ = 603.6859; found *m/z*: 603.6880 [M + Na^+].

Synthesis of compound 9.

8 (363 mg, 0.62 mmol) was dissolved in 4 mL of anhydrous CH_3CN and then DMAP (8.0 mg, 0.062 mmol), TEA (1.4 mL, 9.92 mmol), and Boc_2O (2.0 mL, 8.68 mmol) were sequentially added. The reaction mixture was stirred for 48 h at room temperature, then the solvent was removed under reduced pressure and the residue purified on a silica gel column (eluent: AcOE/*n*-hexane 1:4, v/v). The desired compound **9** was obtained in 58% isolated yield (460 mg, 0.36 mmol).

9: oil, R_f = 0.4 [AcOEt/*n*-hexane, 1:4 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.67 - 7.30 [overlapped signals, 10H, aromatic protons of TBDPS]; 5.33 [apparent triplet, *J* = 10.5 and 9.0, 1H, H-3']; 5.27 [d, *J* = 4.0 Hz, 1H, H-1']; 5.22 [t, *J* = 10.5 and 10.0 Hz, 2H, H-5 and H-5']; 5.18 [d, *J* = 4.0 Hz, 1H, H-1]; 4.94 [apparent triplet, *J* = 10.5 and 9.5 Hz, 1H, H-3]; 4.90 and 4.79 [dd's, *J* = 4.0 and 10.0 Hz, 1H each, H-2 and H-2']; 4.36 [dd, *J* = 2.0 and 11.0 Hz, 1H, H-$6'_a$]; 4.26 [d, *J* = 10.5 Hz, 1H, H-4]: 4.10 [d, *J* = 10.5 Hz, 1H, H-4']; 4.01 [dd, *J* = 2.0 and 11.0 Hz, 1H, H-$6'_b$]; 3.80 [dd, *J* = 2.0 and 11.0 Hz, 1H, H-6_a]; 3.63 [d, *J* = 11.0 Hz, 1H, H-6_b]; 1.49, 1.47, 1.45 [singlets, 63H, 7x $OCOC(CH_3)_3$]; 1.01 [s, 9H, $SiC(CH_3)_3$].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 153.1, 152.5, 152.1, 152.0, 151.9, 151.8 [7x O<u>C</u>OC$(CH_3)_3$]; 135.6, 135.4, 133.4, 132.8, 129.4, 127.5 [aromatic carbons of TBDPS]; 94.3 [C-1]; 93.9 [C-1']; 82.7, 82.6, 82.5, 82.3, 82.2, 82.0 [7x OCO<u>C</u>$(CH_3)_3$]; 73.7, 72.9, 72.5, 72.0 [C-5, C-5', C-3, C-3']; 70.5, 70.3 [C-2 and C-2']; 70.1 [C-4']; 68.0 [C-4]; 63.5 [C-6]; 61.4 [C-6']; 27.7, 27.6, 27.5, 27.3 [7x OCOC(<u>C</u>$H_3)_3$]; 26.5 [SiC(<u>C</u>$H_3)_3$]; 19.2 [Si<u>C</u>$(CH_3)_3$].

ESI-MS (positive ions): calcd. for $C_{63}H_{96}O_{25}Si$ = 1280.6010; found *m/z* = 1303.00 [M + Na^+]; 1320.00 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{63}H_{96}O_{25}SiNa$ = 1303.5908; found *m/z*: 1303.5933 [M + Na^+].

Synthesis of compound 10.

9 (460 mg, 0.36 mmol), dissolved in 2.0 mL of anhydrous THF, was reacted with $Et_3N{\cdot}3HF$ (1.16 mL, 7.12 mmol). The solution was stirred at room temperature for 72 h, then the crude was washed three times with H_2O/CH_2Cl_2; the collected organic phases were dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: *n*-hexane/AcOEt 4:1, v/v). The desired product **10** was obtained in 80% isolated yield (300 mg, 0.29 mmol).

10: oil, R_f = 0.3 [AcOEt/*n*-hexane, 3:7 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): δ 5.28 - 5.24 [overlapped signals, 3H, H-1, H-1', H-3']; 5.20 [apparent triplet, *J* = 10.5 and 9.5 Hz, 1H, H-3]; 4.96 [apparent triplet, *J* = 9.5 and 10.5 Hz, 1H, H-4]; 4.84 - 4.79 [overlapped signals, 3H, H-2, H-2', H-4']; 4.36 [dd, *J* = 2.5 and 12.0 Hz, 1H, H-6_a]; 4.26 [d, *J* = 10.5 Hz, 1H, H-5]; 4.06 [d, *J* = 10.0 Hz, 1H, H-5']; 4.00 [dd, *J* = 2.0 and 12.0 Hz, 1H, H-6_b]; 3.71-3.59 [m, 2H, H-6'$_a$, H-6'$_b$]; 1.50-1.44 [overlapped signals, 63H, 7x OCOC$(CH_3)_3$].

^{13}C NMR ($CDCl_3$, 125 MHz): δ 153.4, 153.1, 152.2, 152.1, 151.9, 151.8 [7x O<u>C</u>OC$(CH_3)_3$]; 94.7 and 94.6 [C-1 and C-1']; 83.4, 83.0, 82.7, 82.6, 82.4, 82.1 [7x OCO<u>C</u>$(CH_3)_3$]; 72.9, 72.4, 72.2, 72.1, 71.1, 70.4, 70.3, 68.1 [C-3, C-3', C-5, C-5', C-2, C-2', C-4, C-4']; 63.3 [C-6]; 60.4 [C-6']; 27.6 [7x OCOC(<u>C</u>$H_3)_3$].

ESI-MS (positive ions): calcd. for $C_{47}H_{78}O_{25}$ = 1042.4832; found *m/z* = 1065.69 [M + Na^+]; 1082.42 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{47}H_{78}O_{25}Na$ = 1065.4730; found *m/z*: 1065.4754 [M + Na^+].

Synthesis of compound 11.

10 (125 mg, 0.12 mmol) was dissolved in 1.0 mL of anhydrous THF. To the solution, DMAP (15 mg, 0.12 mmol) and succinic anhydride (24 mg, 0.24 mmol) were sequentially added and the resulting mixture was taken under stirring at room temperature. After 24 h the reaction mixture was taken to dryness and the crude, transferred into a separatory funnel, was washed three times with H_2O/CH_2Cl_2; the collected organic phases were dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: *n*-hexane/AcOEt 3:7, v/v). The desired product **11** was obtained in 78% isolated yield (107 mg, 0.094 mmol).

11: oil, R_f = 0.4 [AcOEt/*n*-hexane, 1:1 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): δ 5.24 - 5.22 [overlapped signals, 2H, H-1, H-1']; 5.20, 5.17 [triplets, *J* = 10.0 and 10.0 Hz, 1H each, H-3, H-3']; 4.94 [overlapped triplets, *J* = 10.0 and 9.5 Hz, 2H, H-4, H-4']; 4.89 - 4.84 [overlapped signals, 2H, H-2, H-2']; 4.41 - 4.35 [overlapped signals, 2H, H-6_a, H-$6'_a$]; 4.28 - 4.22 [overlapped signals, 2H, H-5, H-5']; 4.08 [m, 1H, H-$6'_b$]; 3.99 [m, 1H, H-6_b]; 2.68 [m, 4H, C$\underline{H}_2$C$\underline{H}_2$COOH]; 1.48 - 1.42 [overlapped signals, 63H, 7x OCOC(C$\underline{H}_3$)$_3$].

^{13}C NMR ($CDCl_3$, 100 MHz): δ 175.5, 171.5 [2x OO$\underline{C}$CH_2CH_2]; 153.0, 152.2, 151.8 [7x O$\underline{C}$OC$(CH_3)_3$]; 94.2 [C-1, C-1']; 82.9, 82.6, 82.3, 82.0 [7x OCO$\underline{C}(CH_3)_3$]; 72.9 [C-3, C-3']; 72.0 [C-2, C-2'], 70.3 [C-4, C-4']; 68.0 [C-5, C-5']; 63.3 [C-6]; 61.2 [C-6']; 29.5, 28.7 [2x CO$\underline{C}H_2CH_2$]; 28.5, 27.5 [7x OCOC($\underline{C}H_3)_3$].

ESI-MS (positive ions): calcd. for $C_{51}H_{82}O_{28}$ = 1142.4993; found *m/z* = 1144.10 [M + H^+]; 1166.06 [M + Na^+]; 1182.03 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{51}H_{82}O_{28}Na$ = 1165.4890; found *m/z*: 1165.4910 [M + Na^+].

Synthesis of compound 12.

11 (88 mg, 0.077 mmol) was dissolved in 0.80 mL of anhydrous CH_2Cl_2 and to the resulting solution DCC (29 mg, 0.14 mmol), DMAP (1.0 mg, 0.077 mmol) and **6** (70 mg, 0.070 mmol) were sequentially added. The reaction mixture was stirred at room temperature for 3 h, then the solvent was removed under vacuum and the residue purified on a silica gel column (eluent: *n*-hexane/AcOEt 4:1, v/v). the desired product **12** was obtained in 86% isolated yield (124 mg, 0.066 mmol).

12: oil, R_f = 0.4 [acetone/*n*-hexane, 35:65 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.03 [s, 1H, H-8]; 6.11 [d, *J* = 5.0 Hz, 1H, H-1']; 5.74 [dd, *J* = 5.6 and 5.0 Hz, 1H, H-2']; 5.53 [m, 1H, H-3']; 5.22 - 5.12 [overlapped signals, 4H, H-a, H-a', H-c, H-c']; 4.98 - 4.79 [overlapped signals, 4H, H-b, H-b', H-d, H-d']; 4.42 - 3.94 [overlapped signals, 9H, H-4, H_2-5', H-f_a, H-f'$_a$, H-e, H-e', H-f_b, H-f'$_b$]; 2.65 [s, 4H, OCC$\underline{H}_2$C$\underline{H}_2$CO succinic acid]; 2.28 [m, 4H, -C$\underline{H}_2$COO myristoyl residues]; 1.68 [s, 9H, C(CH_3)$_3$]; 1.59 - 1.12 [overlapped signals, 125H, 9x OCOC(CH_3)$_3$, 22x CH_2]; 0.84 [overlapped signals, *J* = 5.8 and 5.8 Hz, 6H, 2x CH_3].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 171.8, 171.3 [2x $CH_2\underline{C}OO$ myristoyl and 2x $CH_2\underline{C}OO$ succinoyl residue]; 161.0 [C-6]; 153.1, 152.1, 151.8, 151.2, 150.5 [C-2, C-4, 9x O$\underline{C}$OC(CH_3)$_3$]; 140.3 [C-8]; 121.3 [C-5]; 94.1 [C-a, C-a']; 86.4 [C-1']; 83.8 [C-4']; 82.6, 82.2, 81.9, 80.3 [$\underline{C}$(CH_3)$_3$, 9x OCO$\underline{C}$(CH_3)$_3$]; 73.0 [C-c, C-c', C-2']; 72.2 [C-b, C-b', C-3']; 70.4 [C-d, C-d']; 68.1 [C-e, C-e']; 63.5 [C-f, C-5']; 61.4 [C-f']; 33.7, 33.5, 31.7 [2x $\underline{C}H_2$COO myristoyl residues, 2x $\underline{C}H_2$COO succinoyl residue]; 29.4, 29.1, 28.5, 28.2, 27.7, 27.5, 25.4, 24.6, 24.5, 22.4 [22x CH_2, -C($\underline{C}H_3$)$_3$, 9x OCOC($\underline{C}H_3$)$_3$]; 13.8 [2x CH_3].

MALDI-TOF (positive ions): calcd. for $C_{103}H_{169}N_5O_{38}$ = 2084.1446; found *m/z* = 2085.18 [M + H^+]. **HRMS** (MALDI-TOF): calcd. for $C_{103}H_{169}N_5O_{38}Na$ = 2107.1343; found *m/z*: 2107.1369 [M + Na^+].

Synthesis of compound G5.

12 (111 mg, 0.053 mmol), dissolved in 0.90 mL of anhydrous CH_2Cl_2, was treated with 0.10 mL of TFA. The solution was stirred at room temperature for 3 h; then the solvent was removed under vacuum and the residue coevaporated three times with *i*-PrOH. The desired product **G5** was obtained in quantitative yield without further purification (60 mg, 0.053 mmol).

G5: amorphous solid, R_f = 0.1 [$CHCl_3$/CH_3OH, 7:3 (v/v)].

^{1}H NMR (CD_3OD, 500 MHz): δ 8.72 [s, 1H, NH-1]; 8.21 [s, 1H, H-8]; 6.09 [d, *J* = 4.5 Hz, 1H, H-1']; 5.95 [t, *J* = 5.5 and 5.5 Hz, 1H, H-2']; 5.66 [t, *J* = 5.0 and 5.0 Hz, 1H, H-3']; 5.08 - 5.06 [overlapped signals, 2H, H-a and H-a']; 4.50 – 4.38 [overlapped signals, 4H, H-f'$_a$, H-4' and H_2-5']; 4.21 – 4.18 [m, 1H, H-f'$_b$]; 4.01 [m, 1H, H-e']; 4.00 – 3.78 [overlapped signals, 4H, H-f_a, H-e, H-c and H-c']; 3.68 – 3.47 [m, 1H, H-f_b]; 3.49 – 3.43 [overlapped signals, 2H, H-b and H-b']; 2.67 [s, 4H, COC$\underline{H}_2$C$\underline{H}_2$CO]; 2.44 - 2.32 [overlapped signals, 4H, 2x C$\underline{H}_2$COO myristoyl residues]; 1.86 – 1.54 [overlapped signals, 4H, 2x C$\underline{H}_2$$CH_2$COO myristoyl

residues]; 1.38 - 1.10 [overlapped signals, 20H, 10x CH_2 myristoyl residues]; 0.90 [overlapped triplets, *J* = 6.5 and 7.0 Hz, 6H, 2x CH_3]. H-d and H-d' signals are buried under the residual solvent signal.

^{13}C NMR (CD_3OD, 100 MHz): significant signals at δ 174.5, 174.4, 174.3 [4x $CH_2\underline{C}OO$]; 138.5 [C-8]; 116.6 [C-5]; 95.8, 95.7 [C-a and C-a']; 88.7 [C-1']; 82.2 [C-4']; 75.1 [C-2']; 74.3, 73.7 [C-e and C-e']; 72.5 [C-b and C-b']; 71.9 [C-3']; 65.5 [C-d and C-d']; 64.7 [C-5']; 63.2 [C-f and C-f']; 35.3, 33.6 [2x $\underline{C}H_2COO$ myristoyl residues]; 31.3, 31.0 [2x $\underline{C}H_2COO$ succinoyl residue]; 27.2, 26.5, 24.2 [22x CH_2]; 14.9 [CH_3].

ESI-MS (positive ions): calcd. for $C_{54}H_{89}N_5O_{20}$ = 1127.6101; found *m/z* = 1128.30 [M + H^+].

HRMS (MALDI-TOF): calcd. for $C_{54}H_{89}N_5O_{20}Na$ = 1150.5999; found *m/z*: 1150.5617 [M + Na^+].

Synthesis of compound 13.

3 (129 mg, 0.239 mmol) was dissolved in 1.0 mL of anhydrous CH_2Cl_2. To the solution DCC (222 mg, 1.98 mmol), DMAP (3 mg, 0.024 mmol) and $CH_3O(CH_2CH_2O)_3CH_2COOH$ [19] (212 mg, 0.956 mmol) were sequentially added under stirring at room temperature. After 48 h the solvent was removed under vacuum and the residue purified on a silica gel column (eluent: acetone/*n*-hexane 55:45, v/v). The desired product **13** was obtained in 42% isolated yield (114 mg, 0.099 mmol).

13: oil, R_f = 0.3 [acetone/*n*-hexane, 2:3 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.16 [s, 1H, H-8]; 6.17 [d, *J* = 5.4 Hz, 1H, H-1']; 5.97 [dd, *J* = 5.2 and 5.6 Hz, 1H, H-2']; 5.68 [m, 1H, H-3']; 4.43 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 4.21 - 4.11 [overlapped signals, 6H, 3x $OC\underline{H}_2CO$]; 3.70 - 3.49 [overlapped signals, 36H, 9x OCH_2CH_2O]; 3.34 and 3.33 [singlets, 9H, 3x OCH_3]; 1.67 [s, 9H, $C(C\underline{H}_3)_3$]; 1.39 [s, 18H, 2x $OCOC(C\underline{H}_3)_3$].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 169.6, 169.1, 168.9 [3x $OCH_2\underline{C}OO$]; 160.6 [C-6]; 151.9 [C-2]; 151.2 [C-4]; 150.5 [2x $O\underline{C}OC(CH_3)_3$]; 140.8 [C-8]; 121.2 [C-5]; 85.9 [C-1']; 84.1 [C-4']; 82.9 [$\underline{C}(CH_3)_3$]; 80.1 [2x $OCO\underline{C}(CH_3)_3$]; 73.0 [C-2']; 71.7, 71.1, 70.9, 70.4 [C-3', 9x $O\underline{C}H_2\underline{C}H_2O$]; 68.4, 68.0, 67.8 [3x $O\underline{C}H_2CO$]; 63.4 [C-5']; 58.8 [3x OCH_3]; 28.1 [$C(\underline{C}H_3)_3$] 27.7 [2x $OCOC(\underline{C}H_3)_3$].

ESI-MS (positive ions): calcd. for $C_{51}H_{85}N_5O_{24}$ = 1151.5584; found *m/z* = 1152.77 [M + H^+]; 1174.71 [M + Na^+]; 1190.77 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{51}H_{85}N_5O_{24}Na$ = 1174.5482; found *m/z*: 1174.5501 [M + Na^+].

Synthesis of compound G6.

13 (108 mg, 0.090 mmol), dissolved in 0.90 mL of anhydrous CH_2Cl_2, was treated with 0.10 mL of TFA. The solution was stirred at room temperature for 3 h, then taken to dryness and coevaporated three times with *i*-PrOH. The desired product **G6** was obtained in quantitative yield without further purification (81 mg, 0.090 mmol).

G6: oil, R_f = 0.2 [AcOEt/CH_3OH, 7:3 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.77 [s, 1H, H-8]; 6.79 [broad signal, 2H, NH_2-2]; 6.12 [d, *J* = 5.6 Hz, 1H, H-1']; 5.97 [m, 1H, H-2']; 5.67 [m, 1H, H-3']; 5.47 [m, 1H, H-4']; 4.24 - 4.10 [overlapped signals, 8H, H-5'$_a$ and H-5'$_b$, 3x OCH_2COO]; 3.70 - 3.49 [overlapped signals, 36H, 9x OCH_2CH_2O]; 3.38 - 3.35 [overlapped signals, 9H, 3x OCH_3].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 172.0, 169.3, 169.2 [3x $OCH_2\underline{C}OO$]; 159.5 [C-6]; 155.4 [C-2]; 153.6 [C-4]; 136.2 [C-8]; 117.8 [C-5]; 87.6 [C-1']; 81.2 [C-4']; 73.1 [C-2']; 71.7 [C-3']; 70.7, 70.2 [9x $O\underline{C}H_2\underline{C}H_2O$]; 68.5, 68.0 [3x $O\underline{C}H_2COO$]; 63.0 [C-5']; 58.8 [3x OCH_3].

ESI-MS (positive ions): calcd. for $C_{37}H_{61}N_5O_{20}$ = 895.3910; found *m/z* = 896.2 [M + H^+]; 918.2 [M + Na^+]; 934.2 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{37}H_{61}N_5O_{20}Na$ = 918.3808; found *m/z*: 918.3831 [M + Na^+].

Synthesis of compound 14.

4 (87 mg, 0.11 mmol) was dissolved in 0.50 mL of anhydrous CH_2Cl_2. To the solution DCC (68 mg, 0.33 mmol), DMAP (3 mg, 0.022 mmol) and $CH_3O(CH_2CH_2O)_3CH_2COOH$ [19] (73 mg, 0.33 mmol) were sequentially added and stirred at room temperature for 1 h; then the solvent was removed under vacuum and the residue purified on a silica gel column (eluent: acetone/*n*-hexane 1:1, v/v). The desired product **14** was obtained in 80% isolated yield (105 mg, 0.088 mmol).

14: oil, R_f = 0.8 [acetone/*n*-hexane, 2:3 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.16 [s, 1H, H-8]; 7.67 - 7.31 [overlapped signals, 10H, TBDPS aromatic protons]; 6.30 [d, *J* = 6.8 Hz, 1H, H-1']; 5.81 [apparent triplet, *J* = 5.2 and 6.6 Hz, 1H, H-2']; 5.70 [m, 1H, H-3']; 4.25 - 4.22 [m, 1H, H-4']; 4.14, 4.08 [singlets, 2H each, 2x

OCH_2CO]; 3.89 [m, 2H, H-5'$_a$ and H-5'$_b$]; 3.71 - 3.48 [overlapped signals, 24H, 6x OCH_2CH_2O]; 3.35, 3.33 [singlets, 3H each, 2x OCH_3]; 1.69 [s, 9H, $C(C\underline{H}_3)_3$]; 1.38 [s, 18H, 2x $OCOC(C\underline{H}_3)_3$]; 1.09 [s, 9H, $SiC(C\underline{H}_3)_3$].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 169.3, 168.8 [2x $OCH_2\underline{C}OO$]; 160.9 [C-6]; 152.3 [C-2]; 151.3 [C-4]; 150.5 [2x $O\underline{C}OC(CH_3)_3$]; 139.7 [C-8]; 135.5, 135.3, 132.2, 131.8, 130.0, 129.9 [TBDPS aromatic carbons]; 120.9 [C-5]; 84.1 [C-1']; 84.0 [C-4']; 83.3 [$\underline{C}(CH_3)_3$]; 82.8 [2x $OCO\underline{C}(CH_3)_3$]; 73.8 [C-2']; 71.8, 71.2, 70.9, 70.8, 70.4 [C-3', 6x OCH_2CH_2O]; 68.5, 68.0, 67.7 [2x $O\underline{C}H_2CO$]; 63.2 [C-5']; 58.8 [2x OCH_3]; 33.8 [$SiC(\underline{C}H_3)_3$]; 28.2 [$C(\underline{C}H_3)_3$] 27.7 [2x $OCOC(\underline{C}H_3)_3$]; 11.4 [$Si\underline{C}(CH_3)_3$].

ESI-MS (positive ions): calcd. for $C_{58}H_{87}N_5O_{19}Si$ = 1185.5765; found *m/z* = 1186.28 [M + H^+]; 1208.13 [M + Na^+]; 1224.15 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{58}H_{87}N_5O_{19}SiNa$ = 1208.5662; found *m/z*: 1208.5688 [M + Na^+].

Synthesis of compound 15.

14 (100 mg, 0.080 mmol), dissolved in 0.30 mL of anhydrous THF, was reacted with $Et_3N\cdot 3HF$ (260 µL, 1.6 mmol). The solution was stirred at room temperature for 18 h, then the solvent was removed under vacuum and the residue purified on a silica gel column (eluent: acetone/*n*-hexane 1:1, v/v). The desired product **15** was obtained in 56% isolated yield (47 mg, 0.045 mmol).

15: oil, R_f = 0.2 [acetone/*n*-hexane, 3:2 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.04 [s, 1H, H-8]; 6.07 [overlapped signals, 2H, H-1' and H-2']; 5.73 [broad signal, 1H, H-3']; 4.34 [broad signal, 1H, H-4']; 4.23 [broad signal, 4H, 2x OCH_2CO]; 4.06 [m, 2H, H-5'$_a$ and H-5'$_b$]; 3.72 - 3.50 [overlapped signals, 24H, 6x OCH_2CH_2O]; 3.35, 3.33 [singlets, 3H each, 2x OCH_3]; 1.69 [s, 9H, $C(C\underline{H}_3)_3$]; 1.40 [s, 18H, 2x $OCOC(C\underline{H}_3)_3$].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 169.8, 168.8 [2x $OCH_2\underline{C}OO$]; 161.7 [C-6]; 152.3 [C-2]; 151.3 [C-4]; 150.7 [2x $O\underline{C}OC(CH_3)_3$]; 142.6 [C-8]; 122.7 [C-5]; 88.1 [C-1']; 86.0 [C-4']; 84.9 [$\underline{C}(CH_3)_3$]; 83.5 [2x $OCO\underline{C}(CH_3)_3$]; 73.8 [C-2']; 73.0 [C-3']; 72.1, 71.2, 70.7, 70.6 [6x OCH_2CH_2O]; 68.5, 68.1 [2x $O\underline{C}H_2COO$]; 62.5 [C-5']; 59.2 [2x OCH_3]; 28.5 [$C(\underline{C}H_3)_3$] 28.1 [2x $OCOC(\underline{C}H_3)_3$].

ESI-MS (positive ions): calcd. for $C_{42}H_{69}N_5O_{19}$ = 947.4587; found *m/z* = 948.07 [M + H^+]; 969.85 [M + Na^+]; 985.94 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{42}H_{69}N_5O_{19}Na$ = 970.4484; found *m/z*: 970.4509 [M + Na^+].

Synthesis of compound 16.

15 (29 mg, 0.030 mmol) was dissolved in 0.80 mL of anhydrous CH_2Cl_2, then was added DCC (18 mg, 0.090 mmol), DMAP (0.4 mg, 0.0030 mmol) and myristic acid (10 mg, 0.045 mmol). The solution was stirred for 2 h at room temperature, then the solvent was removed under vacuum and the residue purified on a silica gel column (eluent: acetone/*n*-hexane 1:1 v/v). The desired product **16** was obtained in quantitative yield (36 mg, 0.030 mmol).

16: oil, R_f = 0.8 [acetone/*n*-hexane, 6:4 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.04 [s, 1H, H-8]; 6.16 [d, *J* = 4.8 Hz, 1H, H-1']; 5.90 [apparent triplet, *J* = 5.2 and 4.8 Hz, 1H, H-2']; 5.67 [apparent triplet, *J* = 5.2 and 4.8 Hz, 1H, H-3']; 4.45 - 4.38 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 4.20, 4.14 [singlets, 2H each, 2x OCH_2COO]; 3.70 - 3.51 [overlapped signals, 24H, 6x OCH_2CH_2O]; 3.36, 3.35 [singlets, 3H each, 2x OCH_3]; 2.35 [t, *J* = 7.4 and 7.4 Hz, 2H, $C\underline{H}_2COO$-C5']; 1.69 [broad signal, 9H, $C(C\underline{H}_3)_3$]; 1.40 [s, 18H, 2x $OCOC(C\underline{H}_3)_3$; 1.24 [broad signal, 22H, 11x CH_2]; 0.86 [t, *J* = 6.6 and 6.6 Hz, 3H, CH_3].

^{13}C NMR ($CDCl_3$, 50 MHz): δ 172.9 [$CH_2\underline{C}OO$-C5']; 169.8, 168.8 [2x $OCH_2\underline{C}OO$]; 161.0 [C-6]; 152.3 [C-2]; 151.7 [C-4]; 150.5 [2x $O\underline{C}OC(CH_3)_3$]; 140.4 [C-8]; 121.3 [C-5]; 86.2 [C-1']; 84.1 [C-4']; 82.9 [$\underline{C}(CH_3)_3$]; 80.1 [2x $OCO\underline{C}(CH_3)_3$]; 73.4 [C-2']; 71.8 [C-3']; 70.9, 70.4 [6x OCH_2CH_2O]; 68.0, 67.8 [2x $O\underline{C}H_2COO$]; 62.8 [C-5']; 58.8 [2x OCH_3]; 48.9 [$\underline{C}H_2COO$-C5']; 33.8, 31.8, 29.5, 29.2, 28.9, 28.1, 27.7, 25.5, 24.8, 24.6, 22.5 [$C(\underline{C}H_3)_3$; 2x $OCOC(\underline{C}H_3)_3$; 11x CH_2]; 13.9 [CH_3].

ESI-MS (positive ions); calcd. for $C_{56}H_{95}N_5O_{20}$ = 1157.6570; found *m/z* = 1158.38 [M + H^+]; 1180.30 [M + Na^+]; 1196.28 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{56}H_{95}N_5O_{20}Na$ = 1180.6468; found *m/z*: 1180.6490 [M + Na^+].

Synthesis of compound G7.

16 (31 mg, 0.027 mmol) was dissolved in 0.90 mL of anhydrous CH_2Cl_2 and 0.10 mL of TFA. The solution was stirred at room temperature for 2.5 h, the solvent removed under vacuum and coevaporated three times with *i*-PrOH. The desired product **G7** was obtained in quantitative yield without further purification (24 mg, 0.027 mmol).

G7: oil R_f = 0.2 [AcOEt/CH_3OH, 8:2 (v/v)].

^{1}H NMR ($CDCl_3$, 200 MHz): δ 8.04 [s, 1H, H-8]; 6.03 [d, *J* = 4.8 Hz, 1H, H-1']; 5.90 [apparent triplet, *J* = 5.2 and 4.8 Hz, 1H, H-2']; 5.67 [apparent triplet, *J* = 5.2 and 4.8 Hz, 1H, H-3']; 4.45 - 4.38 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 4.20 - 4.14 [overlapped signals, 4H, 2x OCH_2CO]; 3.70 - 3.51 [overlapped signals, 24H, 6x OCH_2CH_2O]; 3.36, 3.35 [singlets, 3H each, 2x OCH_3]; 2.35 [t, *J* = 7.4 and 7.4 Hz, 2H, $C\underline{H}_2COO$-C5']; 1.24 [broad signals, 22H, 11x CH_2]; 0.86 [t, *J* = 6.6 and 6.6 Hz, 3H, CH_3].

^{13}C NMR ($CDCl_3$, 100 MHz): δ 173.4, 173.3 [2x $OCH_2\underline{C}OO$]; 169.8 [$CH_2\underline{C}OO$-C5']; 161.5 [C-6]; 156.7 [C-2]; 153.6 [C-4]; 140.8 [C-8]; 118.3 [C-5]; 88.8 [C-1']; 84.1 [C-4']; 73.4 [C-2']; 72.7 [C-3']; 71.7, 70.7, 70.3 [6x OCH_2CH_2O]; 68.4, 68.1 [2x $O\underline{C}H_2COO$]; 62.8 [C-5']; 58.8 [2x OCH_3]; 49.1 [$\underline{C}H_2COO$-C5']; 33.8, 31.8, 30.2, 29.5, 29.2, 25.5, 24.8, 22.5 [11x CH_2]; 13.9 [CH_3].

ESI-MS (positive ions): calcd. for $C_{42}H_{71}N_5O_{16}$ = 901.4896; found *m/z* = 901.92 [M + H^+]; 923.98 [M + Na^+]; 939.94 [M + K^+]. **HRMS** (MALDI-TOF): calcd. for $C_{42}H_{71}N_5O_{16}Na$ = 924.4794; found *m/z*: 924.4808 [M + Na^+].

Synthesis of compound 17.

15 (7.0 mg, 0.0075 mmol) was dissolved in 200 μL of anhydrous CH_2Cl_2. DCC (2.9 mg, 0.014 mmol), DMAP (0.10 mg, 0.00070 mmol) and 5-doxyl-stearic acid (3.8 mg, 0.010 mmol) were sequentially added to the solution, which was stirred at room temperature. After 3 h the solvent was removed under reduced pressure and the residue purified on a silica gel column, eluted with acetone/*n*-hexane (1:1, v/v). The desired product **17** was obtained in quantitative yields (10 mg, 0.0070 mmol).

17: oil, R_f = 0.6 [acetone/*n*-hexane 6:4 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): relevant signals at *δ* 8.06 [s, 1H, H-8]; 6.16 [d, *J* = 5.0 Hz, 1H, H-1']; 5.94 [dd, *J* = 5.0 and *J* = 5.5 Hz, 1H, H-2']; 5.69 [bs, 1H, H-3']; 4.44 - 4.39 [overlapped signals, 3H, H-4', H-5'$_a$ and H-5'$_b$]; 4.22 - 4.14 [overlapped signals, 4H, 2x -OCH_2CO-]; 3.74 - 3.64 [overlapped signals, 24H, 6x -OCH_2CH_2O-]; 3.54 [bs, 2H, -CH_2 doxyl]; 3.38, 3.36 [overlapped singlet, 6H, 2x -OCH_3]; 2.40 – 2.35 [broad signals, 2H, CH_2-α stearic acid]; 1.95 - 1.84 [broad signals, 2H, CH_2-β stearic acid]; 1.71 [s, 18H, 2x -$OCOC(CH_3)_3$]; 1.58 [s, 9H, -$C(CH_3)_3$]; 1.48 - 1.23 [overlapped signals, 26H, 13x -CH_2-]; 1.17 - 1.10 [overlapped signals, 6H, 2x CH_3 doxyl]; 0.87 [t, *J* = 6.5 and 6.5 Hz, 3H, -CH_3].

^{13}C NMR ($CDCl_3$, 100 MHz): relevant signals δ 168.9, 168.7 [3x -CO-]; 160.8 [C-6]; 151.5 [C-2]; 151.0 [C-4]; 150.3 [2x -O<u>C</u>OC$(CH_3)_3$]; 121.3 [C-5]; 86.4 [C-1']; 83.9 [C-4']; 82.7 [<u>C</u>$(CH_3)_3$]; 80.0 [2x -OCO<u>C</u>$(CH_3)_3$]; 73.6 [C-2']; 71.7 [C-3']; 71.1, 70.3 [6x -OCH_2CH_2O-]; 68.4, 67.8 [2x -O<u>C</u>H_2CO-, -<u>C</u>H_2CO-]; 63.4 [C-5']; 58.8 [2x -OCH_3]; 48.8 [<u>C</u>H_2COO-C5']; 33.6, 31.6, 30.3, 29.4, 29.0, 28.0, 27.8, 25.3, 24.6, 22.4 [14x -CH_2-, -C(<u>C</u>$H_3)_3$, 2x -OCOC(<u>C</u>$H_3)_3$, 2x CH_3 doxile]; 13.8 [CH_3].

ESI-MS (positive ions): calcd. for $C_{64}H_{109}N_6O_{22}$= 1313.7595; found *m/z*= 1314.40 [M + H^+]; 1336.44 [M + Na^+]; 1352.34 [M + K^+].

Synthesis of G8.

17 (9.3 mg, 0.0070 mmol) was dissolved in 225 µL of anhydrous CH_2Cl_2 and 25 µL of TFA. The solution was stirred at room temperature for 2.5 h, the solvent removed under vacuum and coevaporated three times with *i*-PrOH. The desired product **G8** was obtained as a pure compound, not requiring further purification, in quantitative yields (7.4 mg, 0.0070 mmol).

G8 : solid, R_f = 0.3 [AcOEt/CH_3OH, 8:2 (v/v)].

^{1}H NMR ($CDCl_3$, 500 MHz): relevant signals at δ 8.12 - 8.06 [broad signals, 1H, H-8]; 6.19 - 6.07 [broad signals, 1H, H-1']; 6.07 – 6.02 [dd, *J* = 4.5 and 5.0 Hz, 1H, H-2']; 5.80 – 5.73 [broad signals, 1H, H-3']; 4.49 - 4.21 [broad signals, 5H, H-4', H_2-5', 2x -OCH_2CO-]; 3.75 - 3.47 [overlapped signals, 26H, 6x -OCH_2CH_2O-, -CH_2 doxyl]; 3.37, 3.34 [singlets, 6H, 2x -OCH_3]; 2.50 - 2.35 [broad signals, 2H, CH_2-α stearic acid]; 1.86 - 1.67 [broad signals, 2H, CH_2-β stearic acid]; 1.61 - 1.02 [overlapped signals, 34H, 14x CH_2, 2x CH_3 doxyl]; 0.96 - 0.88 [broad signals, 3H, -CH_3].

^{13}C NMR ($CDCl_3$, 100 MHz): relevant signals at δ 174.9 [-CH_2<u>C</u>O-]; 171.8, 171.6 [2x -OCH_2<u>C</u>O-]; 160.3 [C-6]; 140.3 [C-8]; 88.5 [C-1']; 81.9 [C-4']; 80.0, 79.9 [2x -OCO<u>C</u>$(CH_3)_3$]; 75.0 [C-2']; 73.4 [C-3']; 72.5, 72.0 [6x -OCH_2CH_2O-]; 69.6 [2x -O<u>C</u>H_2CO-]; 69.5 [-<u>C</u>H_2CO-]; 64.5 [C-5']; 59.6 [2x -OCH_3]; 44.0 [-O<u>C</u>H_2- doxyl]; 35.2, 34.4, 33.6, 32.1, 31.2, 28.9, 25.4, 24.2, 20.4 [14x -CH_2-, -C(<u>C</u>$H_3)_3$, 2x -OCOC(<u>C</u>$H_3)_3$, 2x CH_3 doxyl]; 14.9 [-CH_3].

ESI-MS (positive ion): calcd. for $C_{50}H_{85}N_6O_{18}$= 1057.5920; found *m/z*= 1059.91 [M + H^+]; 1094.86 [M + K^+].

7. References

1. For recent reviews, see: a) J. T. Davis *Angew. Chem. Int. Ed.*, **2004**, *43*, 668 – 698 and literature cited therein; b) J. T. Davis, G. P. Spada *Chem. Soc. Rev.*, **2007**, *36*, 296–313 and literature cited therein.
2. S. Lena, S. Masiero, S. Pieraccini, G.P. Spada *Chem. Eur. J*., **2009**, *15*, 7792-7806.
3. G. Gottarelli, G.P. Spada *The Chemical Record*, **2004**, *4*, 39-49.
4. S. Lena, S. Masiero, S. Pieraccini, G. P. Spada *Chem. Eur. J.*, **2009**, *15*, 7792-7806 and literature cited therein.
5. S. Lena, S. Masiero, S. Pieraccini, G. P. Spada *Mini-Reviews in Organic Chemistry*, **2008**, *5*, 262-273.
6. M. S. Kaucher, W. A. Harrell, J. T. Davis *J. Am. Chem. Soc.*, **2006**, *128*, 38-39.
7. L. Ma, M. Melegari, M. Colombini, J. T. Davis *J. Am. Chem. Soc.* **2008**, *130*, 2938-2939.
8. L. Ma, W. A. Harrell, J. T. Davis *Org. Lett.*, **2009**, *11*, 1599-1602.
9. I. Yoshikawa, S. Yanagi, Y. Yamaji, K. Araki *Tetrahedron,* **2007**, *63*, 7474–7481.
10. K. Araki, I. Yoshikawa *Topics in Current Chemistry*, **2005**, *256*, 133-165.
11. I. Yoshikawa, J. Sawayama, K. Araki *Angew. Chem. Int. Ed.*, **2008**, *47*, 1038-1041.
12. A. Gissot, M. Camplo, M. W. Grinstaff, P. Berthélémy, *Org. Biomol. Chem.*, **2008**, *6*, 1324-1333 and literature cited therein.
13. H. Rosemeyer, *Chemistry & Biodiversity*, **2005**, *2*, 977-1062 and literature cited therein.
14. A. H. Farji, P. Wipf, *Bioorg. Med. Chem.*, **2009**, *17*, 2950-2962.
15. For recent examples, see: a) H. W. Yang, W. Jeong, E.-K. Bang, E. M. Jeon, B. H. Kim, *Org. Biomol. Chem.,* **2011**, *9(1*), 291-296; b) C. Ceballos, S. Khiati, C. A. H. Prata, X.-X. Zhang, S. Giorgio, P. Marsal, M. W. Grinstaff, P. Berthélémy, M. Camplo, *Bioconjugate Chem.*, **2010**, *21(6*), 1062-1069; c) S. Khiati, N. Pierre, S. Andriamanarivo, M. W. Grinstaff, N. Arazam, F. Nallet, L. Navailles, P. Berthélémy, *Bioconjugate Chem.*, **2009**, *20(9*), 1765-1772.
16. One example of guanosine-based amphiphiles is described in: H. R. Jung, T. K. Vu, S. K. Choi, S. M. Park, R. H. Kim, *Bull. Korean. Chem. Soc.*, **2010**, *31*, 549-550. Here the authors produced 5 nucleobase-protected amphiphilic derivatives of guanosine in

17% overall yields on average. The presented synthetic scheme does not allow for final guanine deprotection of the guanosine derivatives.

17. L. Simeone, G. Mangiapia, C. Irace, A. Di Pascale, A. Colonna, O. Ortona, L. De Napoli, D. Montesarchio, L. Paduano *Mol. BioSyst.*, **2011**, *7*, 3075–3086.
18. L. Simeone, D. Milano, L. De Napoli, C. Irace, A. Di Pascale, M. Boccalon, P. Tecilla, D. Montesarchio *Chem. Eur. J.*, **2011**, *17*, 13854-13865.
19. R. W. Wang, B. Gold *Org. Lett.*, **2009**, *11*, 2465-2468.
20. S. A. Sikchi, P. G. Hultin *J. Org. Chem.*, **2006**, *71*, 5888-5891.
21. A. Matsuda, M. Shinozaki, M. Suzuki, K. Watanabe, T. Miyasaka *Synthesis*, **1986**; 385-386.
22. for example, see: a) R. Iwaura, K. Yoshida, M. Masuda, K. Yase, T. Shimizu *Chem. Mater.*, **2002**, *14*, 3047-3053; b) R. Iwaura, K. Yoshida, M. Masuda, M. Ohnishi-Kameyama, M. Yoshida, T. Shimizu *Angew. Chem. Int. Ed.*, **2003**, *42*, 1009-1012.
23. G. Gottarelli, S. Masiero, G. P. Spada *J. Chem. Soc., Chem. Commun.* **1995**, *24*, 2555-2557.
24. S. Masiero, R. Trotta, S. Pieraccini, S. De Tito, R. Perone, A. Randazzo, G. P. Spada *Org. Biomol. Chem.* **2010**, *8(12)*, 2683-2692.
25. Following the classification used in ref. 22, a G-tetrad shows the “head” surface (or H) if it is oriented with the H-bonds network going clockwise from the donor to the neighbour acceptor; “tail” (or T) is defined the opposite one, with the donor-to-acceptor H-bonding system running counter-clockwise.
26. T. M. Fyles, H. Luong *Org. Biomol. Chem.*, **2009**, *7*, 733–738.
27. a) S. Licen, C. Coppola, J. D’Onofrio, D. Montesarchio, P. Tecilla *Org. Biomol. Chem.*, **2009**, *7*, 1060-1063; b) C. Coppola, A. Paciello, G. Mangiapia, S. Licen, L. De Napoli, L. Paduano, P. Tecilla, D. Montesarchio *Chem Eur. J.*, **2010**, *16*, 13757–13772.
28. For a recent review, see: B. Gatto, M. Palumbo, C. Sissi *Curr. Med. Chem.*, **2009**, *16*, 1248-1265; for recent examples, see: a) G. Di Fabio, J. D' Onofrio, M. Chiapparelli, B. Hoorelbeke, D. Montesarchio, J. Balzarini, L. De Napoli *Chem. Commun.* **2011**, *47*, 2363-2365. b) E. Pedersen, J. T. Nielsen, C. Nielsen, V. V. Filichev *Nucleic Acids Res.*, **2011**, *39*, 2470-2481.
29. a) S. Balasubramanian, L. H. Hurley, S. Neidle, *Nature Rev. Drug Disc.*, **2011**, *10*, 261-275. b) P. J. Bates, E. W. Choi, L. V. Nayak in *G-rich oligonucleotides for cancer*

treatment, Methods in Molecular Biology, Totowa, NJ, United States, **2009**, 542 (Gene Therapy of Cancer (2nd Ed.), pp. 379-392. c) A. C. Girvan, Y. Teng, L. K. Casson, S. D. Thomas, S. Jueliger, M. W. Ball, J. B. Klein, W. M. Pierce, Jr., S. S. Barve, P. J. Bates *Mol. Cancer Ther.*, **2006**, *5*, 1790-1799.

30. S. R. Meyers, F. S. Juhn, A. P. Griset, N. R. Luman, M. W. Grinstaff *J. Am. Chem. Soc.*, **2008**, *130*, 14444–14445.

Chapter 6.

On the Use of Boc as a Thymine Protecting Group for the Synthesis of Deoxyribose-Alkylated Thymidine Analogs

1. Introduction

Boc is a widely used protecting group, [1] first developed for the peptide synthesis and successively exploited in a wide range of synthetic organic chemistry elaborations. Its application in nucleoside or oligonucleotide chemistry and related mimics has been reported only in sporadic cases, as for the protection of exocyclic amino groups in cytidine, adenosine and guanosine derivatives [2] and the synthesis of modified PNA. [3]

To the best of our knowledge, its use for masking the thymine base has been reported only by Moon *et al.*, who exploited *N*-3 Boc-protected thymidine building blocks to obtain a 3'-deoxy, 3'-[^{18}F]fluorinated thymidine analog in high radiochemical yield. [4] These researchers prepared the Boc-protected nucleoside by reaction with Boc anhydride in the presence of stoichiometric amounts of DMAP in THF, which required long reaction times (5 h) and led to the target compound in only 53% yields. A similar protocol was used by Yu *et al.* to protect uracil in the synthesis of 5-fluoro (FaraU) and (E)-5-(2-fluorovinyl) arabinosyl uridine (FVAU) *via* 5-trimethylstannyl and (E)-5-(2-tributylstannylvinyl) arabinosyl uridine analogs. [5] Also in this case, insertion of the Boc group was obtained in low overall yields (50 %), leading to a mixture of *N*- and *O*-derivatized nucleosides.

More efficient procedures to protect thymine and uracil, useful to give synthetic access to a variety of sugar-modified thymidine and uridine analogs, have been explored by other research groups. Among others, benzyl (removable by catalytic hydrogenation) [6], *p*-methoxybenzyl (removable by ceric ammonium nitrate treatment) [7] and *p*-

nitrophenylethyl (removable by treatment with DBU in pyridine) [8] groups have been adopted to this purpose.

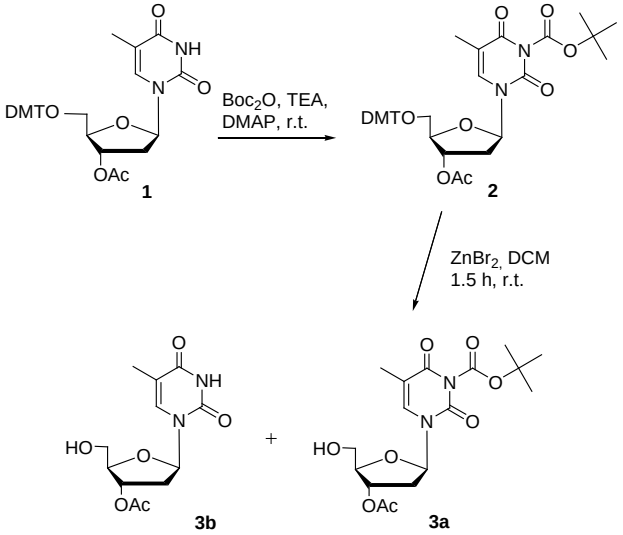
In the context of studies on nucleoside analogs, a novel protection for the thymine base in nucleoside and oligonucleotide synthetic elaborations was recently developed in this laboratory, based on the 2-(phenylthio)ethyl group, inserted at the 3-*N* position by a Mitsunobu reaction with 2-(phenylthio)ethanol. [9] In an efficient "two-stage" system, after oxidation of the thioether to sulfone, this protecting group could be totally removed by a β-elimination mechanism using 0.1 M NaOH aq. solution. The use of 2-(phenylthio)ethyl, stable to strongly basic conditions, allowed to selectively achieve *O*-alkylation of the ribose moieties in satisfactory yields, [10] and thus to efficiently obtain the anti-HIV active Hotoda's $^{5'}$TGGGAG$^{3'}$ sequence, [11] alkylated at the 5'-OH of the thymidine unit with the 3,4-dibenzyloxybenzyl group.

Aiming at expanding the repertoire of available sugar-alkylated thymidine analogs, of interest as potential thymidine kinase inhibitors [12] and functional building blocks for the synthesis of sugar-modified oligonucleotides, we were intrigued by fast and simple synthetic methods to prepare easy-to-handle nucleoside scaffolds. Following the renovated interest in Boc protection, stimulated by the discovery of novel, very mild procedures for its removal, [13] we decided to re-investigate its applicability as a thymine protecting group in nucleoside analogs chemistry.

2. Use of Boc as a thymine protecting group in nucleoside chemistry

As model compounds for this study we chose 5'-*O*-DMT, 3'-*O*-acetyl-thymidine (**1**, Scheme 1) and 5'-*O*-TBDMS, 3'-*O*-acetyl-thymidine (**4**, Scheme 2), which allowed us to explore the compatibility of Boc with the most commonly used protecting groups in nucleoside manipulations, *i.e.* the acid-labile 4,4'-dimethoxytriphenylmethyl (DMT), the fluoride-labile *tert*-butyl dimethylsilyl (TBDMS) and the base-labile acetyl (Ac) group.

Insertion of the Boc group was first optimized on **1**, testing a very simple procedure, involving the use of 2 equiv of Boc anhydride in the presence of TEA (3 equiv) and catalytic amounts of DMAP in different solvents. [14] At this stage the solvent was found to dramatically affect the course of the reaction. When the above mentioned reagents were dissolved in dioxane or acetonitrile, no transformation occurred, even forcing the system conditions; in fact, no benefit could be obtained by prolonging the reaction times, or using high temperatures, or adding to the mixtures bases stronger than TEA, as DIPEA. On the contrary, in a large variety of other solvents, the Boc protection was always complete in 1 h at r.t., with very good-to-excellent yields (see Table 1). [15] In the latter cases, target compound **2** was isolated as a stable compound after a simple work-up, in yields always higher than 90%.



Scheme 1. Synthesis of Boc-protected 5'-O-DMT-thymidine derivative 2 and conversion into its detritylated analog **3a.**

Table 1. Yields of insertion of Boc group on **1** as a function of the used solvent

Solvent	Yields
Acetonitrile	0%
Benzene	92%
1,2-Dichloroethane	92%
Dichloromethane	92%
Diethyl ether	90%
N,*N*-Dimethylformamide	90%
1,4-Dioxane	0%
Pyridine	99%

The best conditions for the Boc installation on **1** were then extended to 5'-*O*-TBDMS, 3'-*O*-acetyl-thymidine **4**. This nucleoside, dissolved in pyridine, was therefore reacted with Boc anhydride (2 equiv) and TEA (3 equiv) in the presence of 0.5 equiv of DMAP, leading to the target compound **5** in almost quantitative yields.

^{1}H and ^{13}C NMR spectra of **2** and **5** clearly indicated that the Boc group was selectively inserted on the 3-*N* position. In fact, no signal attributable to the N-H group could be found in the ^{1}H NMR spectra of these nucleosides dissolved in DMSO-d_6; in parallel, the ^{13}C NMR spectra of both these compounds showed a unique signal at ca. 148 ppm, clearly attributable to the sp^2 carbon of the Boc carbonyl linked to a nitrogen atom, with the concomitant absence of signals at ca. 155 ppm, expected in case of its attachment to an oxygen atom. In contrast with previous findings, [5] we never observed migration of the Boc group from the *N*-3 to the *O*-4 or *O*-2 position, both in this and all the successive manipulations. In addition, no transposition of carbamates to the corresponding *O-tert*-butyl derivatives [2b] was observed, as confirmed by MS analysis of the reaction crudes and of the isolated reaction products.

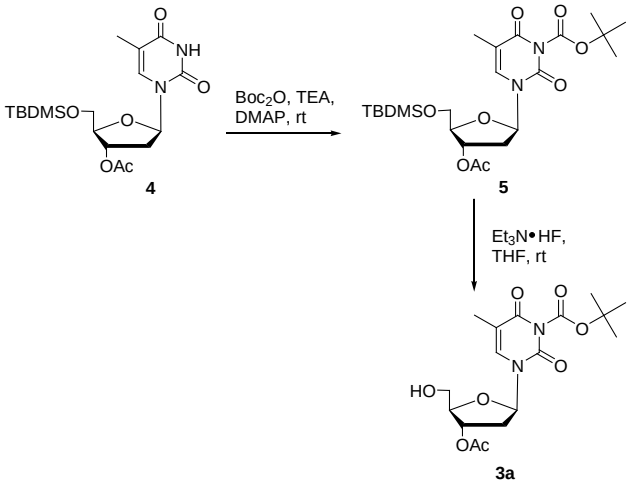
With the two Boc-protected thymidine derivatives in hand, we searched, as a first goal, for efficient conditions allowing the selective removal of the DMT group from **2**. Having *a priori* discarded the use of even diluted, weak protic acids - not able to guarantee the requested selectivity between the two acid-labile protecting groups Boc and DMT - anhydrous Lewis acids, and particularly $ZnBr_2$ [16] and $CuSO_4$ [17] were investigated in a variety of reaction systems. The use of the latter reagent always caused partial cleavage of the Boc group. On the contrary, satisfactory results were obtained with anhydrous $ZnBr_2$, which led to **3a** in 83% yields. The isolation of only traces of **3b** from the reaction mixture essentially confirmed that this treatment did not markedly affect the Boc group. The best conditions for the preparation of **3a** involved the use of 2 equiv of $ZnBr_2$ in anhydrous CH_2Cl_2, taking the reaction mixture under stirring at r.t. for 1.5 h. Prolonging the reaction times, even using a lower excess of the Lewis acid, was detrimental for the success of the reaction, causing also partial Boc removal; similar results were obtained when anhydrous CH_3OH was used in lieu of CH_2Cl_2 as the solvent (see Table 2).

Table 2. Conditions tested for the DMT removal from nucleoside **2** and relative yields

equiv. $ZnBr_2$	Solvent	Reaction Time (h)	Isolated yield (%)	
			3a	**3b**
2.5	DCM	1.5	80	10
3.0	DCM	2.5	60	32
1.5	DCM	2.5	60	32
2.5[a]	CH_3OH	1.25	55	0

[a] $ZnBr_2$ was added as a saturated methanolic solution

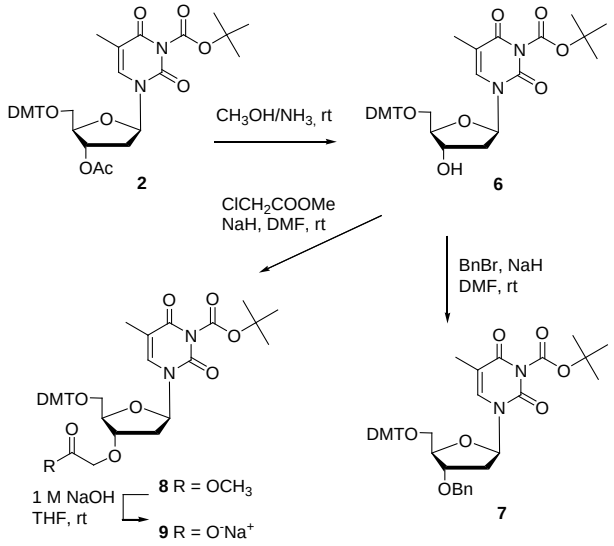
Removal of TBDMS from **5** was carried out first by treatment with standard 1 M TBAF solution in THF, which, quite unexpectedly, was found to partially cleave the Boc group. On the contrary, full selectivity was obtained with the milder desilylating reagent $Et_3N{\cdot}3HF$, [18] which gave the desired Boc-protected nucleoside **3a** in quantitative yields.



Scheme 2. Synthesis of Boc-protected 5'-*O*-TBDMS-thymidine derivative **5** and its conversion into **3a**.

In order to prepare 3'-*O*-alkylated thymidine analogs, the acetyl group at the 3'-OH group had to be removed first. For this step, a standard treatment with conc. aq. ammonia was tested on **2**, which after 2.5 h at r.t. fully converted the starting material into desired **6**. A model reaction for the 3'-OH alkylation was carried out first by reacting **6** with benzyl bromide and NaH (1.2 equiv each) in DMF, giving the target adduct **7** in 85% isolated yields. Then, the alkylation reaction on **6** was carried out with methyl chloroacetate. Due to the lower reactivity of alkyl chlorides with respect to bromides, a higher excess of the alkylating agent and the base (2.5 equiv) was here required for the

reaction to go to completion, which after chromatography gave methyl ester **8** in 82% yields. Hydrolysis with 1 M NaOH (30 min, r.t.) led to target 3'-*O*-carboxymethyl derivative **9**, sodium salt, obtained in 82% yields as a pure compound by TLC and NMR after a simple extraction (66% overall yields for 4 steps starting from **1**). This thymidine analog, as a convenient precursor of 3'-*O*-carboxymethyl-thymidine, [19] is of interest as a versatile building block for backbone-modified oligonucleotides; for example it can be attached at the NH_2-end of a PNA chain in the synthesis of DNA-PNA chimera. [20]



Scheme 3. Synthesis of 3'-*O*-alkylated thymidine analogs **7**, **8** and **9**.

3. Conclusions

In this study we have described the use of Boc to conveniently protect thymine for a simple and general synthetic access to a large variety of sugar-alkylated thymidine analogs. Conditions for its quantitative installation on sugar-protected thymidine derivatives were found, as well as for the selective removal of commonly used protecting groups as DMTr, TBDMS and acetyl group on Boc-protected nucleosides. These high-yielding procedures in principle allow to easily obtain useful and versatile nucleoside building blocks of use in a variety of base-promoted sugar modifications. To demonstrate the feasibility of this approach, two sugar-alkylated thymidine analogs, **5** and **7**, have been synthesized and characterized. The use of **9** as a useful monomer linker to be attached at the NH_2-end of a PNA chain in the synthesis of DNA-PNA chimera is currently underway in our laboratory.

4. Experimental section

Materials and Methods.

TLC analyses were carried out on silica gel plates from Merck (60, F254). Reaction products on TLC plates were visualized by UV light and then by treatment with a 10 % $Ce(SO_4)_2/H_2SO_4$ aqueous solution. For column chromatography, silica gel from Merck (Kieselgel 40, 0.063-0.200 mm) was used. For the ESI MS analyses a Waters Micromass ZQ instrument – equipped with an Electrospray source – was used in the positive and/or negative mode. NMR spectra were recorded on Bruker WM-400 or Varian XR-200 spectrometers. All the chemical shifts are expressed in ppm with respect to the residual solvent signal; *J* values are given in Hz. The following abbreviations were used to explain the multiplicities: s = singlet; d = doublet; t = triplet; q = quartet; m = multiplet; b = broad; dd = double doublet. Peak assignments have been carried out on the basis of standard ^{1}H-^{1}H COSY and HSQC experiments.

General procedure for the insertion of Boc on 1: synthesis of 2.

5'-*O*-DMT-3'-*O*-acetyl-thymidine **1** (53 mg, 0.090 mmol), dissolved in the appropriate solvent, was reacted with Boc anhydride (40 mg, 0.18 mmol), TEA (38 μL, 0.27 mmol) and DMAP (5.5 mg, 0.045 mmol). After 1 h under stirring at r.t., the solvent was removed under reduced pressure and the residue was purified on a silica gel column (eluent: *n*-hexane/AcOEt 7:3, v/v). The yields obtained for each reaction are shown in Table 1 in the main text.

2. foam. Rf = 0.6 (*n*-hexane/AcOEt 1:1, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.55 (d, *J* = 1.2 Hz, 1H, H-6); 7.44-6.84 (overlapped signals, 13H, aromatic protons of *DMT*), 6.40 (dd, *J* = 7.4 and 6.0 Hz, 1H, H-1'); 4.61 (m, 1H, H-3'); 4.08 (m, 1H, H-4'); 3.82 (s, 9H, 2 x –OC$\underline{H}_3$), 3.10 (ABX system, *J* = 2.8 and 10.4 Hz, 2H, H-5'$_a$ and H-5'$_b$); 2.42 (m, 2H, H-2'$_a$ and H-2'$_b$); 2.10 (s, 3H, C$\underline{H}_3$ acetyl); 1.63 [s, 9H, -C(C$\underline{H}_3$)$_3$]; 1.45 (d, *J* = 1.0 Hz, 3H, -CH_3 Thy).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 161.1 (C-4); 158.7, 144.2, 135.0, 130.0, 128.1, 128.0, 127.1, 113.2 (aromatic carbons of *DMT*), 148.5 (C-2); 148.0 (CO carbamate); 135.3 (C-6); 110.9 (C-5); 86.9 (quaternary carbon of *DMT*); 86.6 [-$\underline{C}$(CH_3)$_3$]; 86.2 (C-1'); 85.0 (C-4'); 72.1 (C-3'); 63.4 (C-5'); 55.2 (-O$\underline{C}H_3$); 41.1 (C-2'); 27.4 [-C($\underline{C}H_3$)$_3$]; 21.0 (CO$\underline{C}H_3$); 11.7 (-$\underline{C}H_3$ Thy).

ESI-MS (positive ions): for $C_{38}H_{42}N_2O_{10}$, calculated 686.28; found m/z: 686.09 [M+H$^+$].

Procedure for the removal of the DMT group from 5'-*O*-DMT-3'-*O*-acetyl-3-*N*-Boc-thymidine: synthesis of 3a.

5'-*O*-DMT-3'-*O*-acetyl-3-*N*-Boc-thymidine **2** (30 mg, 0.044 mmol) was dissolved in 1 mL of anhydrous DCM, then dry $ZnBr_2$ (25 mg, 0.110 mmol) was added. The mixture was stirred at rt for 1.5 h, then the reaction was quenched by addition of few drops of MeOH. The crude, diluted with DCM, was washed twice with H_2O; the organic phase was dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: *n*-hexane/AcOEt 2:3, v/v). Products **3a** and **3b** were isolated in 80 and 10 % isolated yield, respectively.

Other reaction conditions were tested (see Table 2), but lower yields were always observed.

3a oil. Rf = 0.4 (*n*-hexane/AcOEt 3:7, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.55 (d, *J* = 1.2 Hz, 1H, H-6); 6.25 (dd, *J* = 7.4 and 6.8 Hz, 1H, H-1'); 5.33 (m, 1H, H-3'); 4.10 (m, 1H, H-4'); 3.92 (m, 2H, H-5'$_a$ and H-5'$_b$); 2.39 (m, 2H, H-2'$_a$ and H-2'$_b$); 2.09 (s, 3H, C$\underline{H}_3$ acetyl); 1.93 (d, *J* = 1.4 Hz, 3H, -CH_3 Thy); 1.60 [s, 9-H, -C(C$\underline{H}_3$)$_3$].

^{13}C NMR ($CDCl_3$, 100 MHz): δ 170.6 (C=O); 161.1 (C-4); 148.5 (C-2); 147.7 (C carbamate); 135.3 (C-6); 110.9 (C-5); 86.7 [-$\underline{C}$(CH_3)$_3$]; 86.0 (C-1'); 85.0 (C-4'); 74.5 (C-3'); 62.4 (C-5'); 37.3 (C-2'); 27.3 [-C($\underline{C}H_3$)$_3$]; 20.8 (CO$\underline{C}H_3$); 12.5 (-$\underline{C}H_3$ Thy).

ESI-MS (positive ions): for $C_{17}H_{24}N_2O_8$, calculated 384.15; found m/z: 407.1 [M+Na$^+$]; 423.0 [M+K$^+$]; 791.2 [2xM+Na$^+$].

3b was identified by TLC and NMR comparison with an authentic sample.

Procedure for the removal of the acetyl group from 5'-*O*-DMT-3'-*O*-acetyl-3-*N*-Boc-thymidine: synthesis of 4.

5'-*O*-DMT-3'-*O*-acetyl-3-*N*-Boc-thymidine **2** (61 mg, 0.089 mmol) was dissolved in 800 μL of CH_3OH and 200 μL of a NH_4OH solution (29.1 % p.p.) were added to the solution. After stirring for 2.5 h at r.t., the reaction mixture was concentrated under reduced pressure. The crude was then washed with DCM/H_2O; the organic phase was dried over anhydrous

Na_2SO_4 and concentrated under reduced pressure. The desired product **4** was obtained in almost quantitative yields without any observable side reaction.

4. oil. Rf = 0.4 (*n*-hexane/AcOEt 1:1, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.55 (d, *J* = 1.2 Hz, 1H, H-6); 7.44-6.84 (overlapped signals, 13H, aromatic protons of *DMT*), 6.40 (dd, *J* = 7.4 and 6.0 Hz, 1H, H-1'); 4.61 (m, 1H, H-3'); 4.08 (m, 1H, H-4'); 3.82 (s, 9H, 2 x –OC$\underline{H}_3$), 3.10 (ABX system, *J* = 2.8 and 10.4 Hz, 2H, H-5'$_a$ and H-5'$_b$); 2.42 (m, 2H, H-2'$_a$ and H-2'$_b$); 1.63 [s, 9-H, -C(C$\underline{H}_3$)$_3$]; 1.45 (d, *J* = 1.0 Hz, 3H, -CH_3 Thy).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 161.1 (C-4); 158.7, 144.2, 135.0, 130.0, 128.1, 128.0, 127.1, 113.2 (aromatic carbons of *DMT*), 148.5 (C-2); 148.0 (C carbamate); 135.3 (C-6); 110.9 (C-5); 86.9 (quaternary carbon of *DMT*); 86.65 [-$\underline{C}$(CH_3)$_3$]; 86.2 (C-1'); 85.0 (C-4'); 72.1 (C-3'); 63.4 (C-5'); 55.2 (-O$\underline{C}H_3$); 41.1 (C-2'); 27.4 [-C($\underline{C}H_3$)$_3$]; 11.7 (-$\underline{C}H_3$ Thy).

ESI-MS (positive ions): for $C_{36}H_{40}N_2O_9$, calculated 644.27; found m/z: 644.97 [M+H^+]; 666.94 [M+Na^+]; 672.95 [M+K^+].

Procedure for synthesis of 5'-*O*-DMT-3'-*O*-benzyl-3-*N*-Boc-thymidine (5).

5'-O-DMT-3-N-Boc-thymidine **4** (50 mg, 0.078 mmol) was dissolved in 400 μL of anhydrous DMF and cooled at 0 °C. After 5 min BnBr (11 μL, 0.094 mmol) and NaH (60% p.p., 4.0 mg, 0.094 mmol) were added in the order. The mixture was stirred at r.t. for 4 h, then a few drops of CH_3OH were added at 0 °C to quench the reaction. The suspension was concentrated under reduced pressure and the crude was then washed with DCM/H_2O; the organic phase was dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: n-hexane/AcOEt 2:3, v/v). The desired product **5** was obtained in 85% isolated yield (49 mg, 0.066 mmol).

5. oil. Rf = 0.8 (*n*-hexane/AcOEt 1:1, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.63 (d, *J* = 1.2 Hz, 1H, H-6); 7.36-6.79 (overlapped signals, 18H, aromatic protons of *DMT* and Bn), 6.34 (dd, *J* = 7.6 and 5.8 Hz, 1H, H-1'); 4.58-4.40 (AB system, *J*= 11.8 Hz, 2H, -OC*H*$_2$Bn); 4.32 (m, 1H, H-3'); 4.18 (m, 1H, H-4'); 3.79 (s, 9H, 2 x –OC$\underline{H}_3$), 3.30 (ABX system, *J* = 2.9 and 10.6 Hz, 2H, H-5'$_a$ and H-5'$_b$); 2.61-2.17 (m, 2H, H-2'$_a$ and H-2'$_b$); 1.61 [s, 9-H, -C(C$\underline{H}_3$)$_3$]; 1.42 (d, *J* = 1.0 Hz, 3H, -CH_3 Thy).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 161.3 (C-4); 158.7, 144.2, 137.3, 130.0, 128.1, 128.5, 128.1, 127.9, 127.6, 127.1, 113.2 (aromatic carbons of *DMT* and *Bn*), 148.5 (C-2); 148.0 (C

carbamate); 135.3 (C-6); 110.9 (C-5); 86.9 (quaternary carbon of *DMT*); 86.2 (C-1'); 85.2 [(-$\underline{C}(CH_3)_3$)]; 84.2 (C-4'); 78.3 ($\underline{C}H_2Ph$), 71.3 (C-3'); 63.4 (C-5'); 55.2 (-O$\underline{C}H_3$); 46.2 (C-2'); 27.4 (-C($\underline{C}H_3)_3$); 11.4 (-$\underline{C}H_3$).

ESI-MS (positive ions): for $C_{43}H_{46}N_2O_9$, calculated 734.32; found m/z: 735.91 [M+H^+]; 756.77 [M+Na^+]; 772.67 [M+K^+].

General procedure for the insertion of Boc on 8: synthesis of 9.

5'-*O*-TBDMS-3'-*O*-acetyl-thymidine **8** (53 mg, 0.090 mmol) was dissolved in 1 mL of dry pyridine, then Boc-anhydride (40 mg, 0.180 mmol), TEA (38 μl, 0.270 mmol) and DMAP (5.5 mg, 0.045 mmol) were added in the order. The mixture was stirred at r.t. for 1 h, the solvent was removed under reduced pressure and the residue was purified on a silica gel column (*n*-hexane/AcOEt 7:3, v/v). The desired compound **9** was obtained in 98% isolated yield (60 mg, 0.088 mmol).

9. oil. Rf = 0.8 (*n*-hexane/AcOEt 4:6, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.54 (d, *J* = 1.2 Hz, 1H, H-6); 6.33 (dd, *J* = 5.2, 1H, H-1'); 5.23 (m, 1H, H-3'); 4.10 (m, 1H, H-4'); 3.90 (m, 2H, H-5'$_a$ and H-5'$_b$); 2.42 (m, 2H, H-2'$_a$ and H-2'$_b$); 2.08 (s, 3H, C$\underline{H}_3$ acetyl); 1.93 (d, *J* = 1.2 Hz, 3H, -CH_3); 1.61 [s, 9-H, -C(C$\underline{H}_3)_3$]; 0.92 [s, 9H, -SiC(C$\underline{H}_3)_3$]; 0.13 [s, 6H, -Si(C$\underline{H}_3)_2$].

^{13}C NMR ($CDCl_3$, 100 MHz): δ 170.6 (C=O); 161.2 (C-4); 148.5 (C-2); 147.9 (C carbamate); 134.3 (C-6); 110.9 (C-5); 86.7 [-$\underline{C}(CH_3)_3$]; 85.4 (C-1'); 85.0 (C-4'); 75.3 (C-3'); 63.5 (C-5'); 37.3 (C-2'); 27.3 [-C($\underline{C}H_3)_3$]; 25.9 [-SiC($\underline{C}H_3)_3$]; 20.9 (CH_3 acetyl); 18.3 [-Si$\underline{C}(CH_3)_3$]; 12.5 (-$\underline{C}H_3$ Thy); -5.6 [-Si(C$\underline{H}_3)_2$].

ESI-MS (positive ions): for $C_{23}H_{38}N_2O_8Si$, calculated 498.23; found m/z: 499.14 [M+H^+]; 520.92 [M+Na^+]; 536.82 [M+K^+].

Procedure for the removal of TBDMS group from 5'-*O*-TBDMS-3'-*O*-acetyl-3-*N*-Boc-thymidine (9).

5'-*O*-TBDMS-3'-*O*-acetyl-3-N-Boc-thymidine **9** (20 mg, 0.040 mmol), dissolved in 200 μL of anhydrous THF, was reacted with triethylamine trihydrofluoride (20 μL, 0.120 mmol). The reaction mixture was stirred at r.t. for 1 h, then the solution, diluted with DCM, was washed twice with H_2O; the organic phase was dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: n-

hexane/AcOEt 2:3, v/v). The desired product was obtained in quantitative yield, and proved to be identical to nucleoside **3a**, isolated from the detritylation of **2**.

Procedure for synthesis of 5'-*O*-DMT-3'-*O*-methoxycarbonylmethyl-3-*N*-Boc-thymidine (6).

5'-O-DMT-3-N-Boc-thymidine **4** (40 mg, 0.062 mmol) was dissolved in 0.5 mL of dry DMF, then methyl chloroacetate (14 μL, 0.15 mmol) and NaH (60% p.p., 8.0 mg, 0.19 mmol) were added. The reaction mixture was stirred at r.t. for 24 h, then a few drops of CH_3OH were added at 0 °C to quench the reaction. The suspension was taken to dryness, then the crude was redissolved in DCM and washed twice with H_2O; the organic phase was dried over anhydrous Na_2SO_4, concentrated under reduced pressure and purified on a silica gel column (eluent: n-hexane/AcOEt 3:7, v/v). The desired ester **6** was obtained in 80% isolated yield (36 mg, 0.050 mmol).

6. oil. Rf = 0.5 (*n*-hexane/AcOEt 1:1, v/v).

^{1}H NMR ($CDCl_3$, 200 MHz): δ 7.63 (s, H-6); 7.44-6.84 (overlapped signals, 13H, aromatic protons of *DMT*), 6.33 (dd, *J* = 5.6 and 6.0 Hz, 1H, H-1'); 4.33-4.10 (overlapped signals, 4H, H-3', H-4', and $C\underline{H}_2COOCH_3$); 3.81 (s, 9H, 2 x $-OC\underline{H}_3$), 3.75 (s, 3H, $COOC\underline{H}_3$); 3.41 (ABX system, *J* = 2.8 and 10.4 Hz, 2H, H-5'$_a$ and H-5'$_b$); 2.58-2.26 (m, 2H, H-2'$_a$ and H-2'$_b$); 1.61 [s, 9-H, $-C(C\underline{H}_3)_3$]; 1.44 (s, 3H, $-CH_3$ Thy).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 170.7 (C=O); 161.1 (C-4); 158.7, 144.2, 135.0, 130.0, 128.1, 128.0, 127.1, 113.2 (aromatic carbons of *DMT*), 148.5 (C-2); 147.3 (C carbamate); 135.3 (C-6); 110.6 (C-5); 86.8 (quaternary carbon of *DMT*); 84.9 [$-\underline{C}(CH_3)_3$]; 84.2 (C-1'); 80.2 (C-4'); 79.2 (C-3'); 63.4 (C-5'); 55.2 ($-O\underline{C}H_3$); 37.1 (C-2'); 27.4 [$-C(\underline{C}H_3)_3$]; 12.6 ($-COO\underline{C}H_3$); 11.8 ($-\underline{C}H_3$ Thy).

ESI-MS (positive ions): for $C_{39}H_{44}N_2O_{11}$, calculated 716,29; found m/z: 716.94 [$M+H^+$]; 738.73 [$M+Na^+$]; 754.75 [$M+K^+$].

Procedure for the synthesis of 5'-*O*-DMT-3'-*O*-carboxymethyl-3-*N*-Boc-thymidine, sodium salt (7).

5'-*O*-DMT-3'-*O*-methoxycarbonylmethyl-3-*N*-Boc-thymidine **6** (12 mg, 0.017 mmol) was dissolved in 100 μL of dry THF, then NaOH (0.7 mg, 0.019 mmol) in 20 μL of water was added. The reaction was kept under stirring at r.t. for 30 min, then the mixture was taken

to dryness, redissolved in DCM and washed twice with H_2O; the organic phase was dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. The desired salt **7** was obtained in 80% isolated yield (10 mg, 0.014 mmol).

7. oil. Rf = 0.1 (AcOEt).

^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.61 (s, 1H, H-6); 7.31- 6.77 (overlapped signals, 13H, aromatic protons of *DMT*); 6.34 (broad signal, 1H, H-1'); 4.17 (overlapped signals, 2H, H-3' and H-4'); 3.79 (s, 2H, $C\underline{H_2}CO$); 3.73 (s, 6H, 2 x $OC\underline{H_3}$); 3.30 (m, 2H, H-5_a' and H-5_b'); 2.55 – 2.14 (m, 2H, H-2_a' and H-2_b'); 1.49 [(s, 9H, -$C(C\underline{H_3})_3$)]; 1.25 (s, 3H, -CH_3 Thy).

^{13}C NMR ($CDCl_3$, 50 MHz): δ 175.6 (C=O); 161.1 (C-4); 158.7, 144.0, 129.9, 127.9, 127.0, 113.2 (aromatic carbons of *DMT*), 148.6 (C-2); 147.8 (C carbamate); 135.3 (C-6); 111.1 (C-5); 86.9 (quaternary carbon of *DMT*); 86.5 [-$\underline{C}(CH_3)_3$]; 85.2 (C-1'); 83.9 (C-4'); 80.9 (C-3'); 63.8 (C-5'); 55.1 (-$O\underline{C}H_3$); 37.3 (C-2'); 27.2 [-$C(\underline{C}H_3)_3$]; 11.5 (-$\underline{C}H_3$ Thy).

ESI-MS (negative ions): for $C_{38}H_{41}N_2O_{11}^-$, calculated 701.27; found m/z: 700.88 [M^-].

5. References

1) Green, T. W.; Wuts, P. G. M. In *Protective Groups in Organic Synthesis. 4th Ed. Wiley-Interscience: New York*, **2005**; b) Kocienski, P. J. *In Protecting Groups, 3rd Ed. Thieme, Stuttgart,* **2005**; c) Agami, C.; Couty, F. *Tetrahedron,* **2002**, *58*, 2701-2724.
2) a) Jung, H. R.; Vu, T. K.; Choi, S. K.; Park, S. M.; Kim, B. H. *Bull. Kor. Chem. Soc.*, **2010**, *31*, 549-550; b) Wang, R.-W.; Gold, B. *Org. Lett.*, **2009**, *11*, 2465-2468; c) Michel, B. Y.; Strazewski, P. *Tetrahedron*, **2007**, *63*, 9836-9841; d) Sikchi, S. A.; Hultin, P. G. *J. Org. Chem.*, **2006**, *71*, 5888-5891; e) Huang, Y.; Dey, S.; Garner, P. *Tetrahedron Lett.*, **2003**, *44*, 1441-1444.
3) de la Torre, B. G.; Eritja, R. *Bioorg. Med. Chem. Lett.*, **2003**, *13*, 391-393.
4) Yun, M.; Oh, S. J.; Ha, H.-J.; Ryu, J. S.; Moon, D. H. *Nucl. Med. Bio.*, **2003**, *30*, 151-157.
5) Yu, C.-S.; Chiang, L.-W.; Wu, C.-H.; Hsu, Z.-K.; Lee, M.-H.; Pan, S.-D.; Pei, K. *Synthesis*, **2006**, *22*, 3835-3840.
6) Luzzio, F. A.; Menes, M. E. *J. Org. Chem.*, **1994**, 59, 7267–7272.
7) Van Aershot, A.; Jie, L.; Herdewijn, P. *Tetrahedron Lett.*, **1991**, *32*, 1905-1908.
8) Schultz, B. S.; Pfleiderer, W. *Tetrahedron Lett.*, **1983**, *24*, 3587-3590.
9) D'Onofrio, J.; De Napoli, L.; Di Fabio, G.; Montesarchio, D. *Synlett*, **2006**, *6*, 845-848.
10) D'Onofrio, J.; Petraccone, L.; Erra, E.; Martino, L.; Di Fabio, G.; De Napoli, L.; Giancola, C.; Montesarchio, D. *Bioconjugate Chem.*, **2007**, *18*, 1194-1204.
11) Hotoda, H., Koizumi, M., Koga, R., Kaneko, M., Ohmine, T., Furukawa, H., Hagatsuma, T., Nishigaki, T., Sone, J., Tsutsumi, S., Kosaka, T., Abe, K., Rimura, S., Shimada, K. *J. Med. Chem.*, **1998**, *41*, 3655-3663.
12) For a recent example, see: a) Van Poecke, S.; Negri, A.; Janssens, J.; Solaroli, N.; Karlsson, A.; Gago, F.; Balzarini, J.; Van Calenbergh, S. *Org. Biomol. Chem.*, **2011**, *9*, 892-901.
13) For recent examples, see: a) Nadim S. S.; Anil S. G.; Vishnu H. D.; Ashutosh V. B. *Tetrahedron Lett.*, **2000**, *41*, 385-387. b) Marcantoni, E.; Massaccesi, M.; Torregiani, E.; Bartoli, G.; Bosco, M.; Sambri, L. *J. Org. Chem.*, **2001**, *66*, 4430-4432. c) Raghavendra S. N.; Kumar B. P.; Longqin H. *Tetrahedron Lett.*, **2006**, *47*, 389-393. d) Thaqi, A.; McCluskey, A.; Scott, J. L. *Tetrahedron Lett.*, **2008**, *49*, 6962-6964; e) Wang,

J.; Liang, Y.-L.; Qu, J. *Chem. Commun.*, **2009**, *34*, 5144-5146. g) Wang, G.; Li, C.; Li, J.; Jia, X. *Tetrahedron Lett.*, **2009**, *50*, 1438-1440.

14) Simeone, L.; De Napoli, L.; Montesarchio, D. *Chemistry & Biodiversity*, accepted doi: 10.1002/cbdv.201100103.

15) Solvent-controlled alkylation of 5'-O-protected thymidine derivatives has been previously described in: a) Teste, K.; Colombeau, L.; Hadj-Bouazza, A.; Lucas, R.; Zerrouki, R.; Krausz, P.; Champavier, Y. *Carbohydr. Res.*, **2008**, *341*, 1490-1495; b) Wu, J. C.; Xi, Z.; Gioeli, C.; Chattopadhyaya, J. *Tetrahedron*, **1991**, *47*, 2237-2254. Following Zerrouki et al., when allyl, propargyl and benzyl bromide were reacted in stoichiometric amounts with 5'-*O*-TBDMS-thymidine in the presence of NaH in THF under ultrasound activation, no reaction was observed. In the same solvent, using 2.5 equiv of the alkylating agent and of the base, exclusively the 3'-O-alkylated derivative was found. On the contrary, only the corresponding 3-N adduct could be isolated from the reaction carried out in DMF. This behavior was explained in terms of different dielectric constants of the solvents, with low dielectric constant solvents favoring the O-alkylation and high dielectric constant solvents promoting N-alkylation. In our case, the dielectric constant of the solvent cannot account for the different course of the Boc insertion on **1**, since the only observed failures occurred in dioxane and acetonitrile, the first one having a very low (2.21) and the latter a very high (37.5) dielectric constant, respectively.

16) Kohli, V.; Blöcker, H.; Köster, H. *Tetrahedron Lett.*, **1980**, *23*, 2683-2686.

17) Randazzo, G.; Capasso, R.; Cicala, M. R.; Evidente, A. *Carbohydr. Res.*, **1980**, *85*, 298-301.

18) Westman, E.; Stromberg, R. *Nucleic Acids Res.*, **1994**, *22*, 2430-2431.

19) a) Edge, M. D.; Jones, A. S. *J. Chem. Soc. (C)*, **1971**, 1933-1939; b) Halford, M. H.; Jones, A. S., *Chem. Soc. C*, **1968**, 2667-2670.

20) For a recent example of linker monomers in the synthesis of DNA-PNA, see: Musumeci, D.; Roviello, G. N.; Valente, M.; Sapio, R.; Pedone, C.; Bucci, E. M., *Biopolymers*, **2004**, *76*, 535-542.

Summary

My PhD thesis has been focused on the synthesis and characterization of a number of novel amphiphilic nucleolipids.

In a first part of my work, preliminary to all the successive investigations, novel thymidine- or uridine-based nucleolipids, containing one hydrophilic oligo(ethylene glycol) chain and one or two oleic acid residues (called ToThy, HoThy, DoHu and ToThyChol), have been synthesized with the aim to develop biocompatible nanocarriers for drug delivery and/or produce pro-drugs. Studies of microstructural characterization of the corresponding aggregates have been carried out in pure water and in pseudo-physiological conditions through DLS and SANS experiments. For ToThy, HoThy and DoHu stable vesicles, with mean hydrodynamic radii ranging between 120 nm and 250 nm have been revealed. In the case of ToThyChol, stable micelles were observed under the same experimental conditions, with mean hydrodynamic radii of 100 nm. Biological validation of the nucleolipidic nanocarriers was ensured by evaluation of their toxicological profiles, performed by administration of the nanoaggregates to a panel of different cell lines. ToThy exhibited a weak cytotoxicity and, at high concentration, some ability to interfere with cell viability and/or proliferation. In contrast, DoHu, HoThy and ToThyChol exhibited no toxicological relevance, behaving similarly to POPC-based liposomes, widely used for systemic drug delivery. Taken together, these results show that the here synthesized nucleolipid-based nanocarriers are finely tunable, self-assembling materials, potentially suitable for the *in vivo* transport of biomolecules or drugs.

In a successive study, these nucleosidic nanovectors have been exploited as multifunctional ligands to obtain the corresponding ruthenium(III) salts, of interest as potential anticancer drugs. The obtained amphiphilic nucleosidic complexes were then studied in their self-aggregation properties by DLS and SANS techniques in aqueous solutions and in pseudo-physiological conditions. In analogy with the behaviour reported in the literature for the known Ru(III) complexes, the studied complexes ToThyRu, HoThyRu, DoHuRu and ToThyCholRu showed a limited stability in aqueous solutions, producing in few hours green precipitates. Therefore we studied how to achieve the complete stabilization in physiological media of these compounds. This goal was

accomplished by use of the nucleosidic Ru(III) complexes in formulation with POPC in molar ratio 15:85. Under these conditions, the complexes were completely stable in pseudo-physiological solutions for several weeks. A complete and comprehensive study on the *in vitro* bioactivity was performed on the synthesized complexes in POPC formulation, particularly examining their growth inhibition ability on MCF-7 and WiDr cell lines. Most remarkably, very promising results were observed on ToThyRu/POPC and HoThyRu/POPC formulations, showing IC_{50} values of 9 and 15 μM, respectively, on MCF-7 cell lines. *In vitro* bioscreening studies on other cancer cell lines are currently in progress, with the peculiar aim to evaluate specificity effects.

To investigate the *in vivo* mechanism of action of the synthesized compounds, a novel amphiphilic nucleosidic Ru(III)-complex, bearing the fluorescent dansyl group, was designed and synthesized, essentially built around the same basic skeleton present in HoThyRu, for applications in fluorescence microscopy.

In a successive work, I have then investigated a novel design for the amphiphilic nucleosidic complexes starting from a highly functionalized uridine-based nucleolipid. In this optimized scaffold, the pyridine ligand for the metal complexation was attached on the sugar skeleton, in lieu of the N-3 position on the nucleobase, as in the case of ToThyRu, HoThyRu, DoHuRu and ToThyCholRu complexes.

A modified nucleoside, 3-azido-3-deoxy-1-β-D-xylofuranosyluracil - here prepared following a new, simple and very convenient synthetic procedure - was also here exploited as a suitable key intermediate to obtain a cationic aminoacylnucleolipid, of interest *per se*, as a model compound of a valuable class of novel biocompatible, highly functionalized nucleolipids, and to be used in mixture with the naturally negatively charged ruthenium(III) complexes, thus producing catanionic vesicles.

For all the latter compounds, including the nucleolipid-ruthenium(III) complexes and the cationic aminoacynucleolipid, a detailed microstructural characterization as well as biological activity assays are currently underway, in collaboration with specialized laboratories, and the related results will be presented in due course.

In the frame of the design, synthesis and characterization of novel nucleolipids, a relevant part of my efforts have been then addressed to the study of guanosine-containing amphiphiles, almost unexplored compounds of interest in the development of

smart, novel self-assembling materials as well as for their potential biological activity. A small library of sugar-modified guanosine derivatives has been prepared, starting from a common intermediate, fully protected on the nucleobase. Insertion of myristoyl chains and of diverse hydrophilic groups, such as an oligoethylene glycol, an amino acid or a disaccharide chain, connected through *in vivo* reversible ester linkages, or of a charged functional group provided different examples of amphiphilic guanosine analogs, named **G1**-**G7** herein. All of the sugar-modified derivatives were positive in the potassium picrate test, showing a marked ability to form G-tetrads. CD spectra demonstrated that, as dilute solutions in $CHCl_3$, distinctive G-quadruplex systems may be formed, with spatial organisations dependent upon the structural modifications. Two compounds, **G1** and **G2**, proved to be good low-molecular-weight organogelators in polar organic solvents, such as methanol, ethanol and acetonitrile. Ion transportation experiments through phospholipid bilayers were carried out to evaluate their ability to mediate H^+ transportation, with **G5** showing the highest activity within the investigated series. Moreover, **G3** and **G5** exhibited a significant cytotoxic profile against human MCF-7 cancer cells in *in vitro* bioassays with IC_{50} values in the 20 μM range, while no cytotoxic activity was observed on normal, control cells.

Finally, to expand the knowledge about the available protective groups for nucleosides, the use of Boc as a thymine protecting group in the synthesis of sugar-alkylated (or, more generally, sugar-modified) thymidine analogs was here described. Boc was easily inserted at the 3-N position in high yields and found to be stable to standard treatments for the removal of acetyl and TBDMS groups, as well as to $ZnBr_2$-mediated DMTr deprotection. Boc protection proved to be completely resistant to the strong basic conditions required to selectively achieve 3'-O-alkylation. This acid-labile group, unprecedentedly used for the masking of pyrimidine nucleobases, was then here exploited to allow the synthesis of 3-azido-3-deoxy-1-β-D-xylofuranosyluracil in six, high yielding steps from uridine, previously mentioned as a valuable starting material for the synthesis of novel, highly functionalized nucleolipids.

From the full comprehension of the structure-activity relationships of the here described compounds, a relevant contribution to the knowledge of nucleolipid ruthenium(III)-complexes, as promising anticancer agents, and more generally of novel

nucleoside-based amphiphiles as attractive, bioinspired building blocks of use in the production of innovative, smart materials for biomedical applications is expected.

List of Abbreviations

The following abbreviations were used throughout the text:

Ac = acetyl

AcOEt = ethyl acetate

Bn = benzyl

Boc = *tert*-butoxycarbonyl

Boc_2O = Di-*tert*-butyl dicarbonate

DCC = dicyclohexylcarbodiimide

DIAD = diisopropylazodicarboxylate

DMAP = 4-(*N*,*N*-dimethylamino)pyridine

DMF = *N*,*N*-dimethylformamide

DMSO = dimethylsulfoxide

DMT = 4,4-dimethoxy-triphenylmethyl

DPPA = diphenylphosphorylazide

EtOH = ethanol

Fmoc = 9-fluorenylmethoxycarbonyl

HEG = hexaethylene glycol

HEPES = 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HPTS = (8-hydroxypyrene-1,3,6-trisulfonic acid trisodium salt

Im = imidazole

i-PrOH = isopropanol

MMTr = 4-monomethoxytriphenylmethyl

Py = pyridine

Pic = picrate

t-Bu = *tert*-butyl

TBDMS = *tert*-butyldimethylsilyl

TBDPS = *tert*-butyldiphenylsilyl

TCA = trichloroacetic acid

TEA = triethylamine

TEG = triethylene glycol

TFA = trifluoroacetic acid

THF = tetrahydrofuran

Thy = thymidine residue

TIS = triisopropylsilane

Trt = trityl