UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN

"SCIENZA E TECNOLOGIA DEI MATERIALI"

CICLO XXX

CALIX[4]ARENE DERIVATIVES AS ARTIFICIAL PHOSPHODIESTERASES AND LIGANDS FOR PROTEINS

Coordinatore:

Chiar.mo Prof. Enrico Dalcanale

Tutore:

Chiar.mo Prof. Alessandro Casnati

Dottorando: Stefano Volpi

Anno 2014-2017

Contents

List of abbreviations

The following list reports the significance of various abbreviations and acronyms present in this thesis. Other non-standard acronyms and abbreviated names for compounds or experimental techniques which are not in the list have been directly reported in the following chapters.

AcOEt: Ethyl Acetate

Ar: Aromatic

Boc: tert-Butyloxycarbonyl

DCM: Dichloromethane

DMF: Dimethylformamide

DMSO: Dymethyl sulfoxide

DNA: Deoxyribonucleic Acid

ESI-MS: Electron Spray Ionization - Mass Spectroscopy

EtOH: Ethanol

HPLC: High Performance Liquid Chromatography

MeOH: Methanol

NEt_3: Triethylamine

NMR: Nuclear Magnetic Resonance

RNA: Ribonucleic Acid

TES: Triethylsilane

TFA Trifluoroacetic acid

TLC: Thin Layer Chromatography

UV-VIS: Ultraviolet-visible spectroscopy

Abstract

The work of this PhD thesis was focused on the synthesis of novel calix[4]arene derivatives and on their use as supramolecular catalysts and as ligands for proteins.

The *cone*-calix[4]arene scaffold has been often described as an appropriate platform for the development of effective artificial phosphodiesterases, by decorating its upper rim with guanidinium groups or metal-ion-containing ligating moieties. However, the possibility to exploit the cooperation between these two different types of catalytically active units in the cleavage of phosphodiester bonds, when present on the same macrocyclic scaffold, was not evaluated before this thesis work. Thus, we synthesized two new *cone*-calix[4]arene derivatives, bearing at the upper rim a guanidinium group and a triazacyclononane moiety in 1,3 distal- or 1,2 vicinal- position, as ligating unit for Zn(II) or Cu(II) ions. After the synthesis, the purification and the characterization of these compounds, detailed kinetic studies were performed to investigate the catalytic activity of the corresponding metal complexes. The Cu(II)-complexes of both the distal and vicinal catalysts gave rise to a high acceleration in the cleavage of 2-hydroxypropyl-*p*-nitrophenylphosphate (HPNP) and bis-*p*-nitrophenylphopshate (BNPP), which can be considered as activated RNA and DNA model compounds, respectively. The rate enhancements of the studied processes were of 5-6 orders of magnitude respect to the background reaction. The same Cu(II)-based catalysts were also tested in the cleavage of different diribonucleoside monophosphates, showing a noteworthy phosphodiesterase activity, with acceleration of the reaction, in the best case, as high as 10^7-folds compared to the background. The combination of the data collected from the kinetic and potentiometric investigations suggested that the studied artificial catalysts exploit a general-base action or a nucleophilic attack of a metal-bound hydroxide ion to allow hydrolysis to occur. Quite relevant appears to be the electrostatic stabilization of the transition state provided by electrostatic interactions involving the guanidinium unit and the phosphate groups of the reacting partners.

In order to mimic the catalytic triad present in the active site of the human Topoisomerase I enzyme, composed by a tyrosine and two arginine units, we also synthesized a third calix[4]arene-based phosphodiesterase model, by functionalizing its upper rim with a phenolic hydroxyl group and two proximal guanidinium units. The bis-protonated form of this compound was tested in the cleavage of the DNA model compound BNPP (bis-*p*-nitrophenyl phosphate) and was found to promote the *p*-nitrophenol liberation from the substrate with a rate enhancement of $6.5 \cdot 10^4$-folds compared to the background hydrolysis. The experimental data indicate that the three active units of this compound cooperate in a reaction sequence involving a phosphoryl transfer process from BNPP to the nucleophilic phenolate moiety of the calix[4]arene-derivative, followed by a further liberation of *p*-nitrophenol from

the phosphorylated intermediate, which takes place thanks to the electrophilic activation by the neighboring guanidine/guanidinium dyad. At the end of the process, the resulting phosphorylated calix[4]arene is rather inert toward dephosphorylation and, consequently, has no turnover. Thus, the starting bisguanidinium-phenolic derivative should be more properly described as a promoter rather than as a catalyst.

The calix[4]arene scaffold can also be effectively used for the creation of ligands able to recognize the surface of different proteins, influencing their physico-chemical properties or biological activity. Accordingly, some calix[4]arene derivatives were tested as ligands for different model proteins, during a five months-placement period spent in Galway at the National University of Ireland, under the supervision of Dr. Peter Crowley. The first investigated protein-ligand couple was composed by cytochrome *c* (cyt *c*) and a calix[4]arene derivative bearing at its upper rim four C-terminal alanine residues, which were considered as appropriate binding units for the protein surfaces, mainly positively charged at neutral pH. ^{15}N-HSCQC NMR titrations pointed out that the ligand is able to recognize different lysine side chains on the protein surfaces (i.e. Lys4 and one of the two neighboring Lys87 and Lys 89), in agreement with what reported for other negatively charged calix[4]arene derivatives. The fitting of the chemical shift perturbations (CSP) experienced by the amide resonances upon increasing the ligand concentration, allowed to demonstrate the presence of at least two binding sites on the protein surfaces, with values for the corresponding dissociation constants (K_d) of 0.4 mM and 1.2 mM. Attempts to get more detailed thermodynamic and structural information for this binding event by ITC or co-crystallization experiments were unsuccessful, since, in the tested experimental conditions, the calorimetric investigations yielded negligible heat of reaction, while in the co-crystallization studies only jellification of the protein was observed in the presence of the ligand.

A second studied protein-ligand system was formed by the B1 immunoglobulin-binding domain of streptococcal protein G (GB 1) and a small library of different tetraguanidinium calix[4]arenes. The corresponding binding events were investigated by ITC experiments, with the aim to get additional data to support those obtained by NMR titrations in a previous work. In the tested conditions, feasible data were acquired only when a solution of GB 1 was injected in a solution of a calix[4]arene functionalized with four 3-guanidiniumpropoxy- moieties at the lower rim. The K_d determined by ITC (8.9 μM) as fitting parameter of the resulting binding isotherm, is slightly lower than that previously estimated by NMR (80 μM), while the stoichiometry of binding (n= 0.85) suggests the coexistence of 1:1 and 1:2 protein to ligand complexes in solution. The analysis of the different contributions to the overall binding free energy pointed out that both the enthalpy and the entropy terms are favorable for the recognition event and similar in magnitude, thus indicating that both electrostatic and hydrophobic interactions are important to allow the binding process to take place. Taking into account also the results of the NMR titrations, it was proposed that the electrostatic interactions could involve the guanidinium pendant

groups at the lower rim of the macrocycle and negatively charged or polar amino acid residues (i. e. Asp36, Glu15, Glu42, Gln32 and Asn35) on the protein surfaces, while hydrophobic, C-H/π or cation/π contacts could take place between the calix[4]arene backbone and aliphatic/aromatic lateral groups (i. e. Thr18, Phe30) or positively charged amino acid side (i. e Lys13).

Chapter 1

Enzymes and artificial catalysts for the cleavage of phosphodiester bonds

1.1 Introduction

Enzymes are supramolecular bio - catalysts that accelerate a wide variety of chemical reactions which take place in cells. They show high substrate regio- and stereospecificity and have an impressive ability in catalyze, under physiological conditions, reactions that normally would not be fast enough to sustain life[1]. Excluding a few catalytic RNA molecules called rybozimes, enzymes are proteins with a large and complicated three – dimensional structure. However, only specific regions of their surface, the *active sites*, are directly involved in the catalytic process. In these sites, a small number of highly preorganized active units (i. e. amino acid residues, metal cations and other co-factors) cooperate to selectively bind the substrate and to convert it into the corresponding product[2].

The fascinating properties of enzymes have always attracted supramolecular chemists that, in last decades, have focused their attention in the creation of artificial catalysts which try to reproduce the activity of their natural analogues. These systems, called enzyme models[3], are generally organic molecules designed to mimic the catalytic units present in the active site of enzymes and to obtain synthetic catalysts with similar activity[4]. Phosphodiesterase enzymes catalyze the hydrolytic cleavage of phosphodiester bonds contained in DNA, RNA and other important biomolecules, resulting, in terms of acceleration over the uncatalyzed reaction[5], one of the most efficient enzyme family in nature. A particular subclass of them, called nuclease, takes part in biological processes where DNA and RNA are manipulated[6], and the creation of biomimetic analogues has also attracted much attention because they offer the possibility to deepen our knowledge about the mechanism of action of these crucial enzymes. Despite artificial nucleases are at the moment far to reach enzyme–like activity, the effort spent in their development is still intense because, in future, researchers hope to use them in biotechnology- or health- related applications, like gene manipulation, DNA sequencing or antisense therapy[7,8].

In the first part of this chapter, a general introduction on enzyme catalysis, supported by the description of the mechanism of action of some natural phosphodiesterases, will be reported. Then, in the second part, the main results obtained in the field of enzyme mimic will be described, with a short review of the most important examples of artificial phosphodiesterases and of their efficiency in the cleavage of DNA and/or RNA model compounds.

1.2 Enzyme catalysis

1.2.1 Enzyme kinetics

The Michaelis - Menten model[1,2] predicts the rate of an enzyme–catalyzed reaction as a function of the substrate concentration. The model describes the enzymatic transformation as a two- step process (**Scheme 1**) where the enzyme E reversibly binds a substrate S to form an intermediate complex ES, which is then irreversibly converted in the product while the enzyme is regenerated in its original form.

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E+P \qquad \textbf{Scheme 1}$$

The rate of product formation for this system is related to the enzyme-substrate complex concentration [*ES*], which is not directly measurable (**Eqn. 1**).

$$v = \frac{d[P]}{dt} = k_2[ES] \quad \textbf{Eqn. 1}$$

To convert this equation in a more manageable form it is necessary to apply the steady state approximation where it is assumed that, excluding the initial part of the reaction where [*ES*] increases very rapidly, [*ES*] remains approximately constant during the overall process, while [*S*] decreases and [*P*] increases (**Eqn. 2**).

$$\frac{d[ES]}{dt} = 0, \;\; k_1[E][S] = (k_{-1} + k_2)[ES] \quad \textbf{Eqn. 2}$$

By considering that the free enzyme concentration is [*E*] = $[E]_{tot}$– [*ES*] we can rearrange these two equations to determine [*ES*] and use it to obtain the Michaelis – Menten equation in the best known form reported in **Eqn. 3**, where $v_{max}= k_2[ES]=k_2[E]_{tot}$ and $K_m = \frac{k_{-1}+k_2}{k_1}$.

$$v = \frac{d[P]}{dt} = k_2[ES] = \frac{v_{max}\,[S]}{K_m + [S]} \quad \textbf{Eqn. 3}$$

In this equation the rate of product formation depends only on experimentally measurable parameters. In fact, in the most common case in which the substrate is in large excess respect to the enzyme, the free- substrate concentration [*S*] is approximately equal to the initial substrate concentration $[S]_{tot}$ (free ligand approximation), v_{max} is the maximum reaction velocity under substrate saturation conditions and the Michaelis-Menten constant K_m corresponds to the [*S*] value that generates a half maximum process rate ($v=v_{max}/2$).

Alternatively, the Michaelis – Menten equation can also be obtained by applying the rapid equilibrium approximation, in which it is assumed that the conversion of the ES complex to yield the product and regenerate the free enzyme is much slower than its dissociation to E + S. Under these conditions, valid during the very early part of the enzymatic process, the concentration of the free substrate does not change significantly, if it is in large excess respect to enzyme, and the amount of the forming product is negligible in the investigated time range. In other words, it is assumed that $k_2 << k_{-1}$ and [*P*] ~ 0 to calculate the relative ratios of E, S and ES directly from the expression of the equilibrium constant for the association (K_a) or the dissociation (K_d) of the enzyme-substrate complex (**Eqn. 4**).

$$K_a = \frac{1}{K_d} = \frac{k_1}{k_{-1}} = \frac{[ES]}{[E]\,[S]} \quad \textbf{Eqn. 4}$$

Thus, by considering that $[E] = [E]_{tot} - [ES]$ and by solving for $[ES]$, the velocity of the studied enzymatic reaction is given by **Eqn. 5**, whose last term merges to **Eqn. 3** since $K_d \sim K_m$ in the limiting case of $k_2 << k_{-1}$.

$$v = \frac{d[P]}{dt} = k_2[ES] = \frac{k_2 K_a [S][E]_{tot}}{1 + K_a[S]} = \frac{k_2[S][E]_{tot}}{K_d + [S]} \quad \textbf{Eqn. 5}$$

The plot of the Michaelis – Menten equation (**Figure 1**) gives a hyperbolic curve where the reaction rate, respect to the substrate concentration, first increases linearly and then reaches a plateau around its maximum value v_{max}. In fact, at low $[S]$ values, the last term of **Eqn. 3** can be approximated as $\frac{v_{max}[S]}{K_m + [S]} \sim \frac{v_{max}[S]}{K_m}$ and the reaction follows a first order kinetics respect to the substrate concentration, while at high [S] values the enzyme becomes saturated and the reaction rate approach asymptotically v_{max}, following a zero order kinetic law.

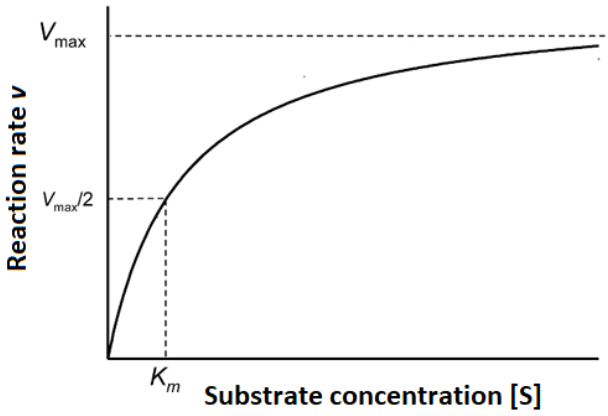


Figure 1: Example of experimental plot of the Michaelis – Menten equation.

The parameters which appear in the Michaelis – Menten equation can be used to obtain very important information about the kinetics of the enzymatic process. K_m, for example, is an index of the affinity of the enzyme for the substrate. As just mentioned, in fact, under rapid equilibrium approximation it corresponds to the dissociation constant $K_d = \frac{K_{-1}}{K_1}$ of the ES complex and, thus, lower K_m values correspond to tighter binding of the substrate by the enzyme. k_2, instead, corresponds to the catalytic rate constant generally indicated as k_{cat}, which is also the turnover frequency $k_{cat} = \frac{v_{max}}{[E]_{tot}}$ of the enzyme, or the number of substrate molecules that can be converted in the product by one molecule of catalyst per unit of time. Another important parameter is the specificity constant k_2/K_m. When the substrate concentration is very low compared to K_m, all the ES complex is readily converted in the product and its formation becomes the rate determining step. In these conditions, as mentioned before, $\frac{[S]}{K_m + [S]} \sim \frac{[S]}{K_m}$, and $[E]_{tot} \sim [E]$, thus the Michaelis – Menten equation can be approximated as follows in **Eqn. 6**, where the reaction rate depends both on the enzyme and on the substrate initial concentrations and the specificity constant is the pseudo-second-order rate constant of the process. The specificity constant is a measure of the catalytic efficiency. In fact, the highest it is, the most effective is the

substrate binding by the enzyme, with maximum efficiency for $k_2 >> k_{-1}$. In this case $k_2/K_m = k_1$, thus the reaction is diffusion limited and the enzyme reaches the maximum theoretical efficiency possible.

$$v \sim \frac{k_2}{K_m}[E][S] \quad \textbf{Eqn. 6}$$

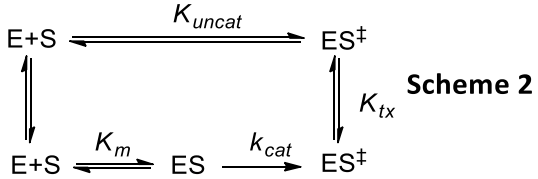
Although many enzymatic processes can be more complicated, the Michaelis – Menten model is a good approximation to describe many of these reactions, especially for single-substrates biological transformations. Moreover, it results to be generally correct in predicting the kinetics of enzyme model-catalyzed reactions. These systems, however, are often studied by using an excess of catalyst, in order to simplify the measurements of the corresponding kinetic constants, whose absolute values are generally low. In these cases, thus, it is more convenient to predict the velocity of the catalytic process with **Eqn. 7**, in which the rate of product formation is expressed as a function of the free catalyst concentration [*Cat*], which can be approximated to the total catalyst concentration $[Cat]_{tot}$ if the concentration of the substrate is significantly lower. This equation is analogous to **Eqn. 5**, since it can be obtained under rapid equilibrium approximation as described before, but tacking into account that the reactant in excess is the catalyst instead of the substrate.

$$v = \frac{k_2 K_a [S]_{tot}[Cat]}{1 + K_a[Cat]} \sim \frac{k_2 K_a [S]_{tot}[Cat]_{tot}}{1 + K_a[Cat]_{tot}} \quad \textbf{Eqn. 7}$$

1.2.2 Origin of enzyme catalysis

The first considerations on the origin of enzyme catalysis were formulated in the well-known *key and lock model* by Emil Fischer,[9] who suggests that the substrate (the key) is activated by the active site of the enzyme (the lock) as a consequence of their specific binding. This model was then modified[10] by proposing that the key does not perfectly fit in the lock, originating a certain strain that destabilizes its ground state and increase its rate of conversion into the product. This proposal is nowadays known as the *reactant* (ground) *state destabilization* theory (RSD). However, the hypothesis which is currently considered the most correct one to describe the origin of the enzyme catalysis is the *transition state stabilization theory*[11] (TSS), which was introduced by Linus Pauling. This theoretical model proposes that the enzyme binds the substrate in the transition state better than in the ground state, thus lowering the activation energy of the process.

The pseudo-thermodynamic cycle shown in **Scheme 2** can be used to compare the energetics of a reaction in water solution to the corresponding enzyme-catalyzed one.



Scheme 2

The activation free energy in the enzymatic reaction can be divided in two contributions: the free energy for the formation of the Michaelis complex ΔG_{bind}, which is related to K_m, and the free energy for the formation of the transition state $\Delta g^{\ddagger}_{cat}$, which is related to k_{cat}. As shown in **Figure 2**, $\Delta g^{\ddagger}_{cat}$ turns to be much smaller than the corresponding free energy for the formation of the transition state in water $\Delta g^{\ddagger}_{uncat}$ both in the RSD and in the TSS theory, but in the first case this result is due to the formation of an unstable ES complex ($\Delta G_{bind}>0$), while in latter to the stabilization of the transition state ES‡ ($\Delta G_{bind}<0$). The affinity of the enzyme for the transition state can be estimated by calculating $K_{tx}= k_{non} \cdot K_m / k_{cat}$,[12] which gives an indication on the acceleration of the enzymatic reaction respect to the uncatalyzed one and thus on the catalytic power of the system.

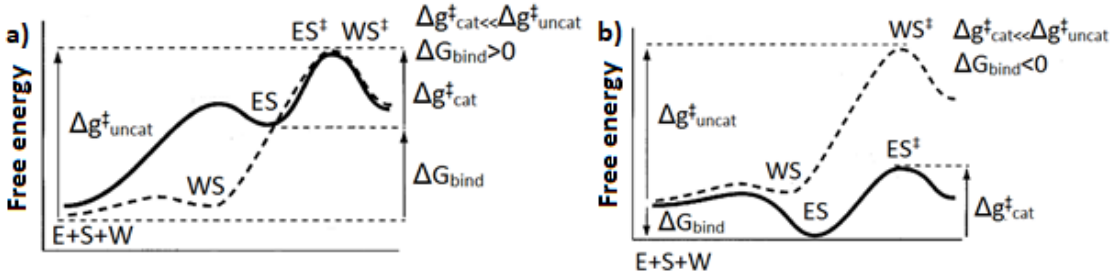


*Figure 2: comparison of the free energy surface for an enzymatic reaction (solid lines) and the corresponding uncatalyzed reaction in solution (dashed lines) for the **a)** RSD theory and the **b)** TSS theory. E, S, W are enzyme, substrate, water molecules respectively.*

This representation of the catalytic event allows us to describe the catalytic power in terms of reduction of free activation energy, but it does not tell us how it is originated. Several researchers, by combining available structural, kinetic, thermodynamic data and accurate computer simulations, have proposed several explanation of enzyme catalysis, which will be now rapidly described.

Some studies have pointed out a connection between enzyme catalysis and intramolecularity[13–15], by observing that many intermolecular reactions turn to be much faster when converted to the corresponding intramolecular ones and that a reaction of an enzyme and its bound substrate reminds an intramolecular reaction. Jenks and Page suggested[16], in a pioneering work, that a large part of the catalytic power gained in this way would derive from entropic factors, in particular from the loss of translational and rotational degrees of freedom during the substrate-enzyme binding. However, Villà et al.[17] demonstrated that the calculation presented in this paper were based on an incomplete pseudo-thermodynamic cycle and that the formation of the transition state in enzymatic reactions does not lead to a substantial loss of degree of freedom when compared to the corresponding uncatalyzed reactions.

In fact, it is now considered more correct that enzyme catalysis mainly originates from enthalpic contributions [17,18]. Menger et al., always by exploiting the concept of intramolecularity, introduced the *spatiotemporal principle*[19,20], according to which fast enzymatic reactions take place if the reactive groups of the substrate are held by the enzyme at “contact distance” and with a proper reciprocal

orientation, where for "contact distance" was meant a distance small enough to avoid the presence of the solvent in between (<3 Å, that is the diameter of a water molecule). A similar theory was proposed by Bruice and co-workers[21–23] by introducing the definition of *Near Attack Conformation* (NAC). NACs are the conformations that allow the reactant to reach the transition state and, for intramolecular reactions, the activation energy depends on how high is the fraction of NACs among the overall reactant conformations. In this model, the ability of an enzyme to bring the substrate in its NAC, is considered proportional to its activity.

Unless both the spatiotemporal principle and the NAC effect have generally been presented as examples of RSD effect[21,24], they are now though to be consistent with the TSS theory. The mechanism of action of Chorismate Mutase (CM)[20,23] has been reported as one of the main example to support these theories because this enzyme binds the substrate in a conformation very close to that of the transition state and it was considered to activate the substrate by inducing a strong steric strain on it. However, Warshel et al.[25], by performing detailed computational calculations, pointed out that CM binds the transition state of the substrate much more effectively than its ground state and thus that the enzyme works by TSS. In that paper, the apparent NAC effect reported by Bruice was described as a consequence of the preorganization of the active site of the enzyme that drives the substrate in a conformation able to maximize the TSS effect by electrostatic interactions.

According to the latter example, the most recent investigations in the field of enzyme catalysis agree with the idea that its main source is the electrostatic stabilization of the transition state[26–28]. Warshel et al. have reviewed the most important evidences that support this thesis in **Ref. 28**, by showing the results of accurate quantum-mechanic calculations. According to these results, the catalytic power of enzymes was attributed to the preorganization of their active site, that provides a network of electrostatic interactions already partially oriented to stabilize the transition state. In the corresponding reaction in water, in fact, the solvent must pay a high energy cost to reorganize the polar environment in order to stabilize the transition state charges, while in enzymes this cost has already been paid during the folding of the active site. In the same study alternative proposal for the origin of enzyme catalysis have been also evaluated, but they were found to be marginal when compared to TSS effect (covalent catalysis[29], RSD by steric strain effect[19-24] or desolvation[30], other entropic contributions[17]) or to be associated with it (NAC effect[25]).

Finally, some researchers have suggested that quantum-mechanic tunneling[31], deviation from equilibrium distribution in phase space[32] and dynamic effects can be related to enzyme catalysis[33], because are known to play an important role in their mechanism of action. These phenomena do not deal with the TSS theory but, by proper quantum-mechanic investigations, their magnitude in enzymatic processes and in the corresponding reaction in solution was reported to be generally similar[27]. Thus, their role in enzyme catalysis is thought to be limited.

1.3 Cleavage of phosphodiester bonds

Phosphate esters are part of the backbone of many important biomolecules which are involved in several physiological transformation[34]. For example, phosphodiester bonds are the building blocks chosen by nature to link the nucleotides in DNA and RNA strands, which are natural oligomers responsible of the conservation and the transmission of genetic code, respectively. Because of the crucial role played by these compounds in living organisms, it is not surprising to notice that they can be considered inert towards spontaneous hydrolysis under mild conditions (water solution, room temperature, neutral pH).

For instance, the cleavage of DNA strands by the intermolecular attack of a generic hydroxide ion on the phosphodiester backbone (**Figure 3a**) is so slow that the corresponding rate constant has never been directly measured. In fact, the most accurate estimations[35], extrapolated from high temperature experiments on simple model compounds, gave a half-life of 31 million of years for this process. The case of RNA is rather different because, excluding when it is cleaved by some rybozymes, it undergoes to hydrolysis by intramolecular attack of the ribose 2'OH[36] (**Figure 3b**), which is kinetically much favored. As a consequence, RNA turns to be much more reactive than DNA, but appropriate studies[37] on model compounds showed that the half-life for the phosphodiester hydrolysis is still quite high, being around 110 years at neutral pH.

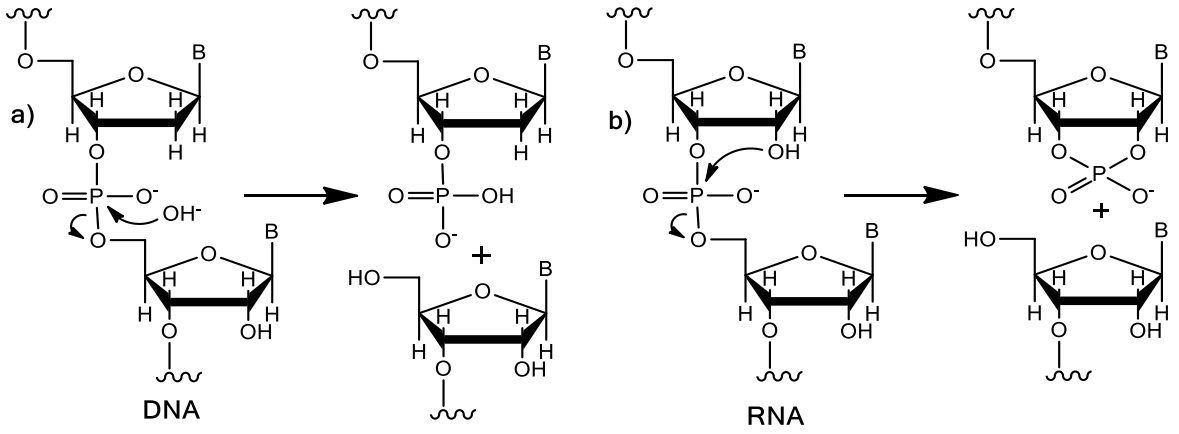


*Figure 3: Generic representation for the uncatalyzed hydrolysis of phosphodiester backbone of **a)** DNA and **b)** RNA.*

Despite their kinetic inertness, phosphate diester bonds have to be hydrolyzed in several biological processes[6] (i. e. DNA and RNA expression, DNA duplication and repair, elimination of foreign DNA and RNA) and living organisms use phospodiesterase enzymes (also called nuclease when the substrate is a nucleic acid) as natural and extremely efficient catalysts to achieve this purpose.

1.3.1 Enzymatic cleavage

Phosphate ester hydrolysis reactions are nucleophilic substitution at the phosphorous which can follow three limit mechanisms[38–40] (**Figure 4**):

- A dissociative mechanism ($D_N + A_N$) where the leaving group departure comes before the attack of the nucleophile with the formation of a metaphosphate intermediate
- An associative mechanism ($A_N + D_N$) where, in two subsequent steps, the nucleophile attacks the phosphate group to form a phosphorane intermediate before the departure of the leaving group
- A concerted mechanism (D_NA_N) in which the attack of the nucleophile and the departure of the leaving group are synchronous and no intermediates can be observed.

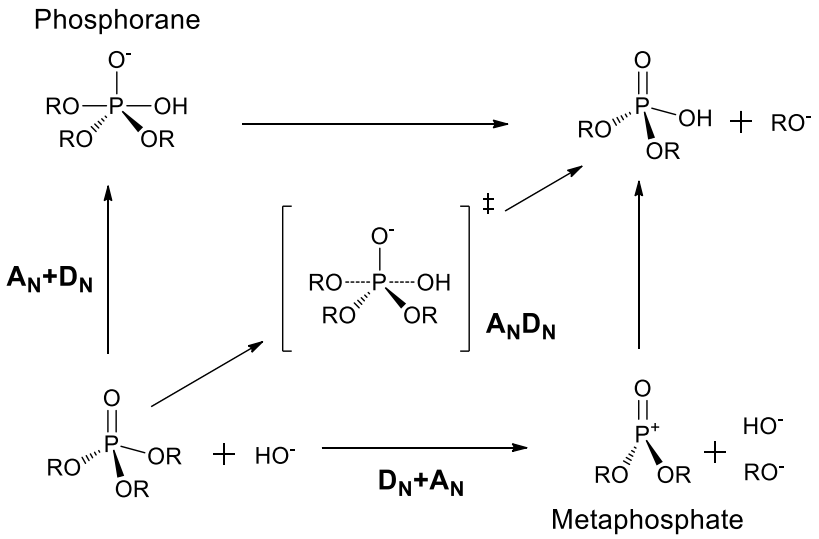


Figure 4: Diagram showing the three limit mechanism for the hydrolysis of phosphate esters.

The real reaction, of course, can also follow a mixed pathway which can be described as a A_ND_N mechanism involving a pentavalent transition state where the entering and the leaving groups are in axial positions, with $D_N + A_N$ and $A_N + D_N$ contributions that depend on the structural features of the substrate. In fact, moving from dianionic phosphate monoesters to neutral phosphate triesters, the attack of the nucleophile is expected to be gradually faster than the departure of the leaving group and the transition state is expected to become more associative in character. On the other hand, also the nature of the leaving group can influence the transition state form[41,42]. In fact, good leaving groups are expected to increase its dissociative contribution, while poor leaving groups should increase its associative character. Accordingly, DNA and RNA, being phosphate diesters with poor leaving groups, are generally believed to undergo to hydrolysis via a concerted mechanism with an important associative contribution to the transition state[39], but a definite and fully convincing reaction pathway is not yet available and the question is still under debate.

Metal ions are known to have a prominent role as active units of nuclease enzymes[43–46], and can assist the cleavage of the phosphate groups in the following ways (**Figure 5**):

I. activation of the phosphate group and stabilization of the phosphorane-like transition state by electrophilic/electrostatic interactions
II. intramolecular activation of the nucleophile (generally a metal-bound water molecule, whose pK_a is lowered by the coordination to the metal cation)
III. stabilization of the leaving group by coordination to the alcoholic oxygen (by decreasing the pKa of the corresponding leaving alcoholate anion)
IV. indirect activation of the nucleophile by a metal-bound molecule (general base catalysis)
V. indirect activation of the leaving group by a metal-bound molecule (general acid catalysis)

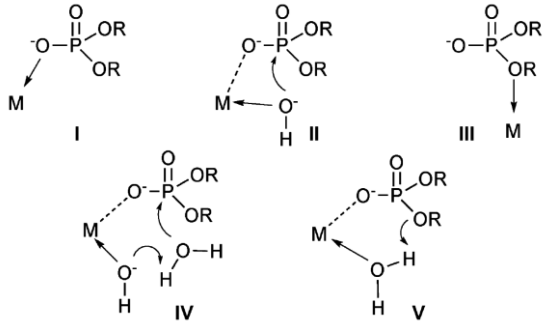


*Figure 5: Possible mechanisms of metal-ion catalysis for the cleavage of a generic phosphate diester (figure taken from **Ref. 47**).*

According to the available X-ray structure of phosphodiesterase enzymes, it is known that their active site can contain from one to three Mg(II), Ca(II), Zn(II) and, sometimes, Mn(II), Fe(II) and Fe(III) ions. These observations led many researchers to deduce that enzymes can exploit the cooperation of two or more metal ions in the cleavage of their substrates. However, it is necessary to be extremely careful in the use of information from crystallographic data to suggest the mechanism of action of an enzyme[44,47,48]. In fact, to induce their crystallization it is often required a high salt concentration and thus it is not sure that all the metal ions found in the solid state are also present *in vivo*. Moreover, metal ions, even if close to the active site, can have simply a structural role and can be scarcely or not involved in the catalytic process. In fact, despite some examples of tri- or dimetallic catalysis have been demonstrated, phosphodiesterases use most frequently one or even nor metal ions for catalysis.

A mechanism exploiting a dimetallic catalysis has been demonstrated for Purple Acid Phosphatases (PAPs) from pig and kidney bean[49] (**Figure 6a**). These enzymes are generally involved in the cleavage of phosphate monoesters, but can also be active with phosphate diesters. At their active site is present a two-centers binding site for a Fe(III) and a M(II) ions (M(II)= Fe(II), Mn(II), Zn(II)) coordinated to two bridging oxygens, one belonging to a μ-hydroxide ion and one to an aspartate residue. In the case of phosphate diesters the enzyme follows a processive mechanism in which the phosphate group is coordinated to the divalent cation and undergoes a nucleophilic attack by a Fe(III)-bound hydroxide ion. Then, the phosphate monoester intermediate is further activated by coordination to the Fe(III) cation

and is hydrolyzed by the bridging μ-hydroxide. *Eco*RI and *Eco*RV are instead example of single-metal catalysis[46,50]. These enzymes are restriction endonuclease that cleave double stranded DNA with high sequence specificity. Both of them need a Mg(II) ion as cofactor to activate the substrate by coordination of one of the non-bridging oxygens of the phosphate during the nucleophilic attack by a water molecule, which in the case of *Eco*RI is directly bound to metal center (**Figure 6b**) while in *Eco*RV is coordinated to the phosphate group present at the 3'O side of the scissile one (**Figure 6c**). In both enzymes the binding site for the metal ion is constituted by two aspartate residues, but, in the case *Eco*RV, the presence of an additional glutamate in close proximity has led some researchers to suggest the participation of a second Mg(II) to the catalytic process[51]. This fact has been for long time an object of controversy, but, thanks to accurate kinetic and computational investigations, the enzyme is now thought to use a single metal ion, which might move towards the two binding sites during the catalytic process[52,53].

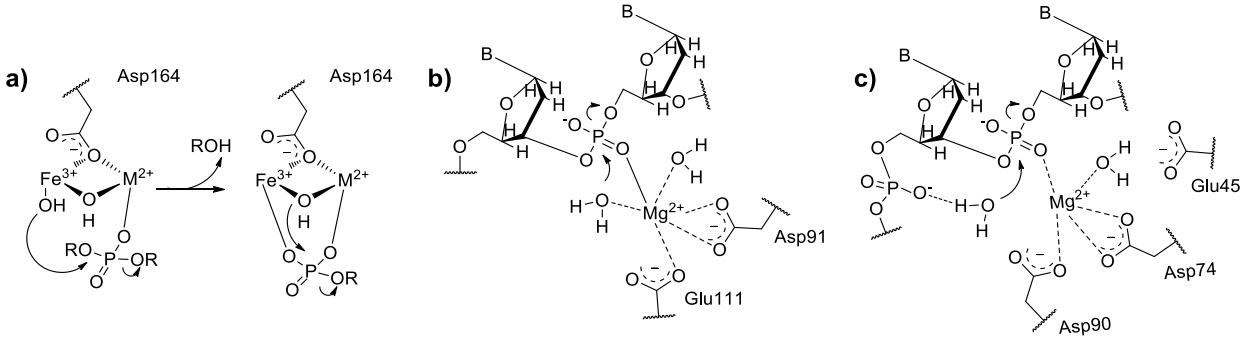


*Figure 6: Schematic representation of the active sites of **a)** Purple Acid Phosphatase, **b)** EcoRI, **c)** EcoRV and corresponding mechanism of action.*

Metal ions are not the only active units present in enzymes. In fact, they often cooperate with amino acid residues that can also have an equivalent catalytic action[43–45] and eventually can substitute them[54–56]. The amino acid residues generally found in the active sites of enzymes are His, Lys, His, Glu, Asp, Tyr, Thr, Ser, Cys, Asn, Gln and can contribute to the catalytic process in very different ways, by exploiting the properties of their side chains. For example, protonated amino acid residues (i. e. Lys, Arg, His, Asp, Glu) are often found to act as Lewis or general acids to activate the substrate and stabilize the transition state or to protonate the leaving group, respectively. On the other hand, in their deprotonated form, amino acids (i. e. Asp, Glu, His) can also work as Lewis or general bases to coordinate or deprotonate a water molecule, both metal-bound or not, thus providing a suitable nucleophile to cleave the substrate. Additional hydrogen bonds can be also provided by the amidic C=O and NH moieties of the enzyme backbone, in order to achieve a more complete pattern of electrostatic interaction able to increase the overall catalytic activity. Moreover, amino acids with nucleophilic side-chains (i. e. Tyr, Ser, Cys) can act directly as nucleophile by attacking the phosphate group of the substrate and form a covalent adduct with it.

Two remarkable examples of cooperation between metal ions and amino acid residues are given by the Alkaline Phosphatase and Staphylococcal Nuclease. Alkaline Phosphatase from *E. coli*[57] is a dimeric monophosphatase enzyme with a low substrate selectivity, and is thus able to cleave also phosphate diesters. Its mechanism of action is currently understood to involve two Zn(II) cations and an arginine side-chain to bind the substrate and stabilize the pentavalent intermediate which is formed after the nucleophilic attack by a serine residue on the phosphate group. Then, a metal-bound water molecule hydrolyze the phosphoseryl intermediate to yield the product and regenerate the free enzyme (**Figure 7a**). An additional Mg(II) ion is thought to be indirectly involved in the catalytic process by inducing the structural rearrangement required by the two subunits of the enzyme to maximize the affinity for the substrate and the subsequent release of the product. Staphylococcal Nuclease from *Staphylococcus aereus*[58] cleaves single strand DNA and RNA chains thanks to the synergic action of a Ca(II) ion and two arginine residues that stabilize the transient phosphorane specie formed by the attack of a water molecule, deprotonated by a neighboring aspartate residue (**Figure 7b**). Moreover, although it is not definitely demonstrated, one of the two guanidinium ions may act as general acid by protonating the alcoholic oxygen of the leaving group and thus facilitating its departure.

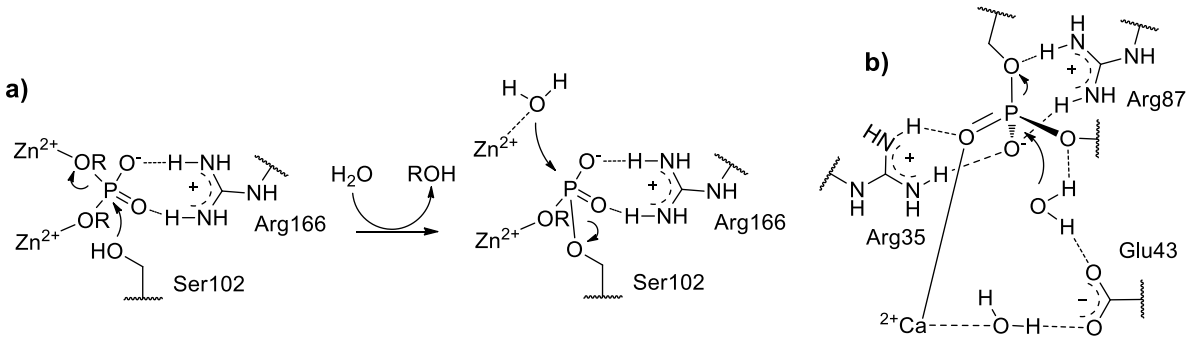


*Figure 7: Schematic representation of the active sites of **a)** Alkaline Phoshatase, **b)** Staphylococcal Nuclease and corresponding mechanism of action.*

According to the crucial role played by amino acid residues in phosphodiesterases, it is not surprising to report that some of these enzymes use them as single type of catalytic units, without the help of other cofactors. One of the most important cases is human Topoisomerase I[55,59], a monomeric enzyme that solves topological problems generated by key nuclear processes such as DNA replication, transcription, recombination, repair, chromatin assembly, and chromosome segregation. More in details, the enzyme allows the disentanglement of supercoiled DNA by transiently breaking one strand of the target duplex, thus generating a gate through which another region of DNA can be passed. The required strand cleavage is achieved by exploiting the nucleophilic attack of a tyrosine residue to form a covalently bound intermediate, whose pentavalent phosphorane group is stabilized by hydrogen bonding of the non-bridging oxygens to two arginine and a histidine residues, respectively (**Figure 8a**). The histidine residue is also close to the scissile 5'O of the substrate and thus could also act as general acid by

protonating the leaving group and facilitating its departure. Another noteworthy example is bovine RNAse A[56]. This enzyme cleaves single strand RNAs by exploiting the action of a couple of catalytically active histidines (**Figure 8b**). In a first step His 12 acts as general base to deprotonate the ribose 2'OH and promotes its attack on the phosphate group, while His 119 acts as general acid by protonating the 5'O of the leaving ribose. In a second step the acid/base role of the two histidines is reversed. In fact, His 119 acts as general base by deprotonating a water molecule that attacks on the just formed pentavalent intermediate and His 12 reprotonates the ribose 2'OH to form the product and the free enzyme.

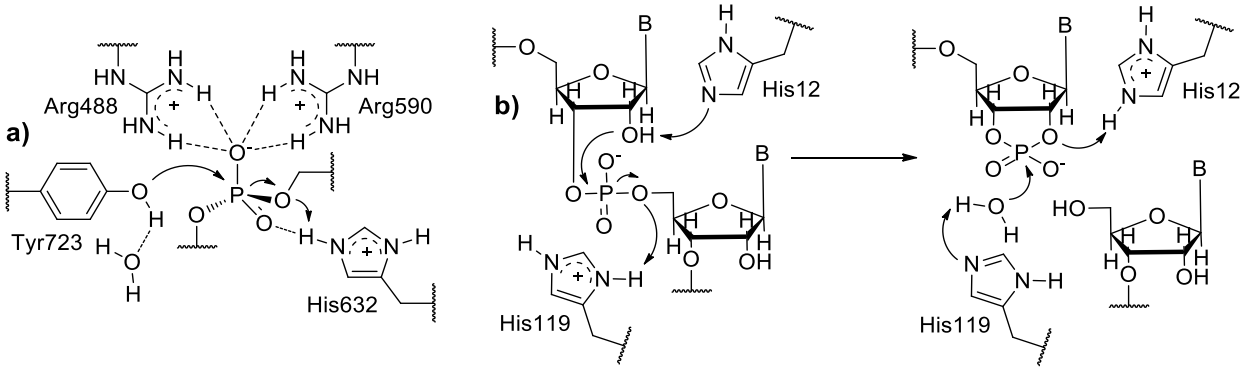


*Figure 8: Schematic representation of the active sites of **a)** DNA topoisomerase I, **b)** RNAse A and corresponding mechanism of action.*

The aforementioned example of phosphodiesterases are only some of the most remarkable of those reported in literature. For many other of these enzymes the mechanism of action has been studied, and in some cases, elucidated. However, the reported cases are sufficient to have a general idea on the basic features required by the active site of an enzyme to be capable of cleaving phosphate diesters bonds, and have been used by supramolecular chemists as reference rules for the creation of artificial catalysts able to perform these processes. The main examples of phosphodiesterase models reported in the literature will be presented in the following paragraph of this chapter.

1.4 Artificial Phosphodiesterases

One of the most important challenges for supramolecular chemists in the last decades has been the creation of artificial catalysts able to mimic the action of natural enzymes[3,4,60]. Phosphodiesterase models[43,61–63] are one of the class of enzyme mimics that have attract the highest interest in this field, in view of their hypothetic future applications in medicine or biotechnology. At the current state of the art, the activity of these compounds is much lower than that of natural enzymes, but the improvement of their catalytic efficiency is still attracting a high interest among researchers for many reasons. First of all, the limited natural abundance, the high substrate selectivity and the scarce stability of enzymes represent a limitation for their use in large scale applications, and the production of less active, but simpler, more stable and versatile surrogates can represent a solution to this problem. Second, the

creations of effective biomimetic systems can allow to obtain deepen elucidations on the mechanism of action of their corresponding natural analogues. Moreover, artificial nucleases can be used as conformational probes to investigate the structure of nucleic acid, in order to obtain refined hydrolytic agents with gradually higher affinity and sequence selectivity for these substrates.

Good phosphodiesterase models should be able to recognize the substrate, promote the conversion of the bound substrate into products, readily release the products and restore the catalytically active form, thus assuring turnover[64–66]. To do that, proper active units have to be placed in proper positions on a certain molecular scaffold. The design of these artificial catalyst plays a crucial role for the effectivity of the system. In fact, the simple mimic of the structural features of the reference active site is not always sufficient to have good catalysts, because even a slight mismatch of the orientation of the reactive groups from the one required to bind the substrate and/or stabilize the transition state could lead to a drastic loss of catalytic activity. It is thus necessary that the molecular scaffold preserves a residual flexibility[67] in order to allow the catalytic units to rearrange for a correct substrate binding and to adapt to its conformational changes along the whole catalytic process. In other words, an enzyme model must provide a *dynamic preorganization*[68] of its active units to be capable of a good phosphodiesterase activity.

The activity of artificial phosphodiesterase, at least in preliminary studies, is generally tested on simple phosphodiester model compounds, often bearing better leaving groups compared to RNA or DNA (Figure **9a** and **9b**, respectively).

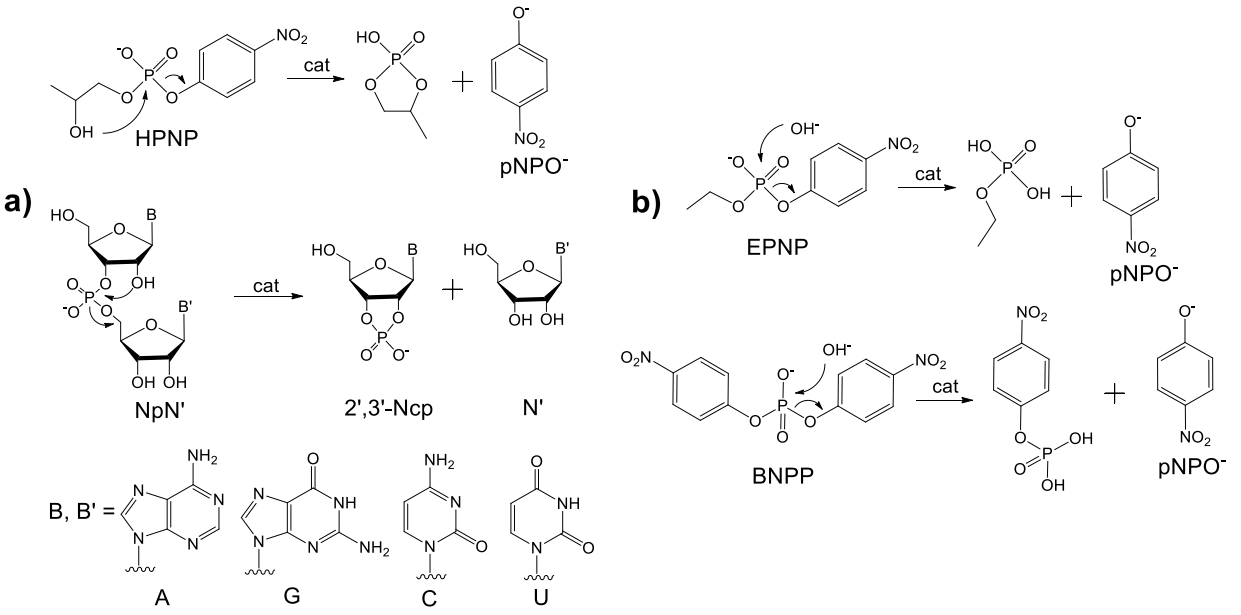


*Figure 9: Generic cleavage mechanism of **a)** RNA and **b)** DNA model compounds generally used to test the activity of artificial phosphodiesterases, indicated as **cat** for simplicity.*

Activated model compounds have the advantage to be easier to handle and to cleave, thus facilitating the performance of the kinetic measurements. However, since these compounds can undergo hydrolysis according to mechanisms that could differ from those followed by nucleic acids, it is necessary to be careful into interpret the results of experiments carried out on them[8,69,70]. For example, enzyme models that are effective in the stabilization of the leaving group can seem poor catalysts in the cleavge of simple phosphate diesters, but can turn to be very effective once tested on natural substrates bearing worse leaving groups. On the other hand, artificial phosphodiesterases that cleave effectively simple phosphate diesters could be sometimes hardly effective on RNA or DNA fragments, because of detrimental steric- or electronic factors generated by the interaction with the more complicated structure of these compounds. Nevertheless, simple phosphate diesters are often very useful to screen the activity of enzyme models and to find good candidates that can be then tested on more complicated substrates or eventually on RNA and DNA fragments or chains.

1.4.1 Metal-based artificial phosphodiesterases

As aforementioned, metal cations are one of the most common cofactors used by natural phosphodiesterases. The main properties that a metal ion should have to effectively promote hydrolysis of the substrate are the hard character (to bind the phosphate oxygens), Lewis acidity (to polarize the phosphate group, the leaving group and eventually water molecules) and fast ligand exchange (to ensure turnover)[8]. Nature selected Ca(II), Mg(II), Fe(II), Fe(III) and Zn(II) to satisfy these requirements[43-46], while chemists prefer to use only Fe(III) and Zn(II) of this list, with the addition of lanthanides, Cu(II), and Co(III).

Probably because of their environmental availability, Ca(II) and Mg(II) are very common in enzymes, but they are not convenient to be used in phosphodiesterase mimics. In aqueous medium, in fact, their ability to provide catalysis is disadvantaged by their reluctance to undergo desolvation and by their poor Lewis acidity, which is further quenched after the coordination to the oxygen ligands required to form stable adducts with appropriate molecular scaffolds. These problems were solved in enzymes in millions of years by evolution but for the production of the corresponding models is more convenient to use lanthanides, which are harder, or transition metal cations, which, being softer ions, can be effectively bound to a proper molecular platform with nitrogen ligands to form complexes with free coordination sites suitable for catalysis and a higher amount of residual Lewis acidity on the metal center. However, also in these cases the coordination of the metal ions causes an important loss of reactivity. To overcome this drawback, it can be helpful to increase the number of metal centers or to insert in the system organic moieties which can participate to the catalytic process, in order to obtain polyfunctional catalysts where several active units cooperate to cleave the substrate.

Despite lanthanide-based artificial catalysts were reported to be capable of moderate[71–73] or in some cases high[74,75] phosphodiesterase activity, they are not attractive for biological applications because of

their toxicity. For these reasons, next paragraphs of this chapter will be focused on the description of the main examples of transition metal-containing enzyme models, with a particular attention for those that use the highly biomimetic Fe(III), Zn(II) and Cu(II) ions.

1.4.2 Mononuclear artificial phosphodiesterases

Very interesting cases of Zn(II)-based phosphodiesterase models are the corresponding complexes formed after coordination to the polyamine ligands **1**-**10** (**Figure 10a**)[76]. These complexes were tested in the cleavage of HPNP (2-hydroxypropyl-*p*-nitrophenylphosphate) and BNPP (bis-*p*-nitrophenylphosphate), which are RNA and DNA model compounds, respectively. For both substrates a metal bound water molecule was supposed to participate to the catalytic process, by directly attacking the phosphate group of BNPP or by deprotonating the β-OH of HPNP, despite in the latter case also a direct coordination of the hydroxyl group of the substrate has not to be excluded in principle. The reactivity of the systems was found to be strongly dependent on the properties of the ligands, in particular on the imposed coordination geometries and on how high is the suppression of Lewis acidity on the Zn(II) cation. In fact, for both substrates, the tridentate cyclic ligands **1-5** were found to form the most active complexes of the present library because they were capable to bind the Zn(II) ion by a facial coordination mode which ensures the presence of free binding sites for the water molecule and the substrates. On the other hand, linear trivalent polymanines **6-8** can hardly provide this type of coordination and their catalytic activity was found to be lower. Moreover, it was found that reactivity of the tested complexes increases with the acidity of the metal-bound nucleophile, which depends on the residual Lewis acidity of the Zn(II) ion. In fact, tetravalent ligands **9-10** were reported to be the worse catalysts of the library because of their simultaneous tendency to coordinate the metal ion with a disadvantageous geometry and to suppress the activation of the metal bound water molecule.

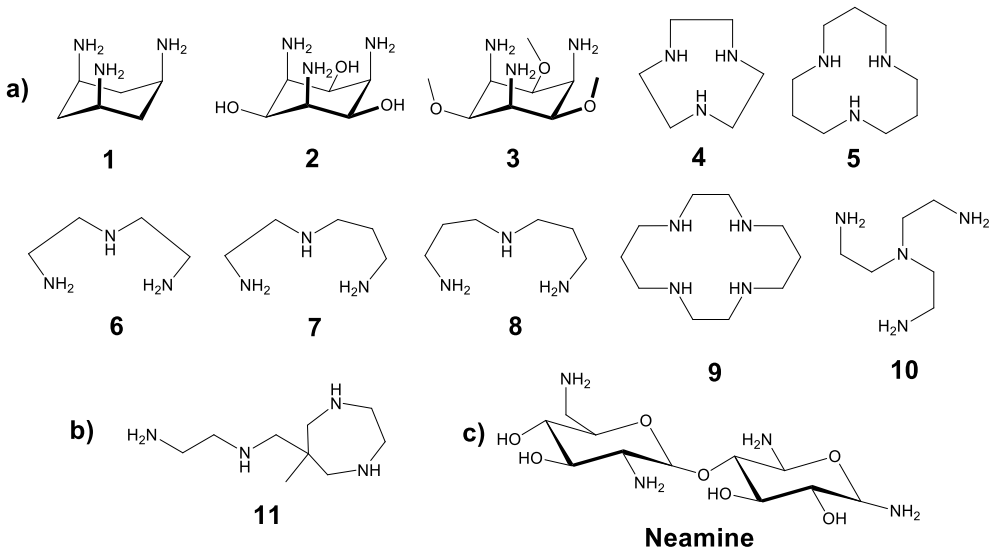


Figure 10: Polyamines ligands for the creation of mononuclear artificial phosphodiesterases.

Cyclic polyamines were successfully used as ligands to generate mononuclear phosphodiesterases also with other metal cations, resulting effective into hydrolyze, in some cases, also DNA sequences. For example, the Co(III) complex of compound **11**[77] and the Cu(II) complexes of **4**[78] and **1**[79] (**Figure 10b**), were found to cleave DNA plasmids with a first order rate constant of $5 \cdot 10^{-5}$ s^{-1} , $5 \cdot 10^{-5}$ s^{-1} and $1.2 \cdot 10^{-3}$ s^{-1} (37°C, pH7.6), respectively. The latter reactivity corresponds to a half-life of 20min for supercoiled DNA in the presence of 75μM Cu(II)-complex and is one of the highest reported in literature. Finally, the adduct between Cu(II) and neamine[80,81] (**Figure 10c**) was found to cleave with a good efficiency plasmid DNA, showing saturation kinetics with a high affinity for the substrate, probably due to the binding of the positively charged amino glycoside to the polyanionic DNA backbone.

1.4.3 Di- and trinuclear artificial phosphodiesterases

In the past years, a wide variety of multinuclear phosphodiesterase models were synthesized, with the aim to enhance the activity of the corresponding mononuclear systems by exploiting the cooperation of two or more metal ions. As stated before, to obtain good enzyme models the chosen molecular scaffold has to preserve a certain degree of flexibility. However, it is also important that the scaffold is rigid enough to fix the metal centers at a distance similar to the one found in the active sites of natural metallo-enzymes, which is known to range from 3 to 6 Å[44].

Good results to reproduce this feature were obtained by spanning proper ligands with (semi)rigid aromatic spacers (**Figure 11**)[82].

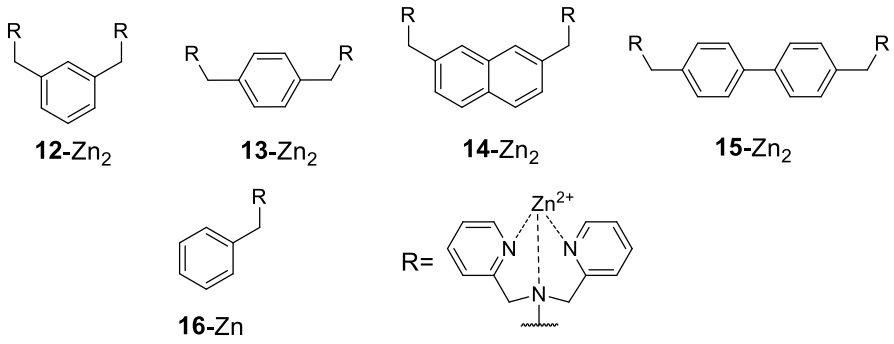


Figure 11: Dinuclear Zn(II) complexes obtained by spanning proper ligands with aromatic spacers and monofunctional control compound.

For example, Zn(II) complexes of compounds **12-15** were studied in the cleavage of the diribonucleoside monophosphate UpU. The catalytic activity was found to be inversely proportional to the distance between the metal centers, with a maximum reactivity for the dimetallic complex **12-** Zn_2, whose molecular scaffold was supposed to guarantee an optimal Zn(II)-Zn(II) distance. For the di-functional complexes of compounds **12-14** was also reported a good degree of cooperativity of the metal cations, resulting more effective with respect to the mono-functional reference compound **16**-Zn, while no cooperation was instead found between the two Zn(II) ions of compound **15-**Zn_2, probably because of

the high metal-metal distance imposed by the spacer. Similarly, Chin et al. developed the naphthalene-based ligand **17**[83] (**Figure 12a**) functionalized with two Cu(II) complexes of 1,4,7-triazacyclonane (TACN, **Figure 10**, compound **4**), which was found to cleave ApA 300 times faster compared to the corresponding mononuclear analog **4**-Cu. According to the kinetic data, a metal bound water molecule was supposed to participate to the catalytic process as general base, by deprotonating the 2'-OH of the substrate.

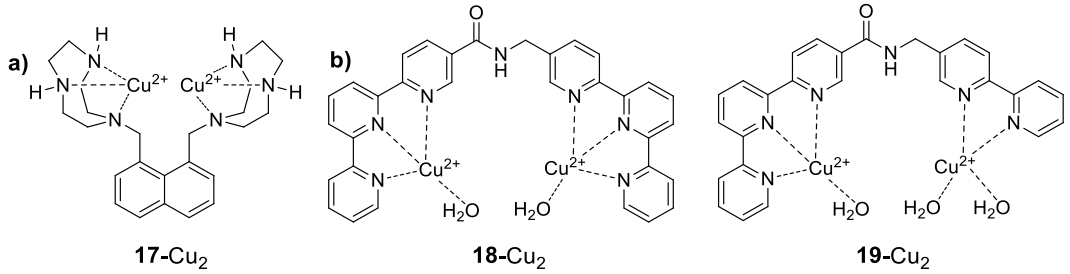


*Figure 12: Dinuclear Cu(II) complexes tested in the cleavage of **a)** ApA and **b)** HPNP.*

An amide spacer was instead used by Hamilton et al. to connect polyaromatic derivatives such as bipyridine and terpyridine, in order to create the corresponding copper complexes **18**-Cu_2 and **19**-Cu_2 (**Figure 12b**)[84]. In the transesterification of the RNA model HPNP both compounds were superior to their mononuclear analogs, and the higher activity of **19**-Cu_2 respect to **18**-Cu_2 was supposed to be due to the lower pK_as of its metal bound water molecules.

Another strategy to keep two metal centers at a correct distance to cooperate in the cleavage of phosphodiester bonds is to insert in a flexible molecular scaffold a phenolic or alcoholic μ-bridging unit.

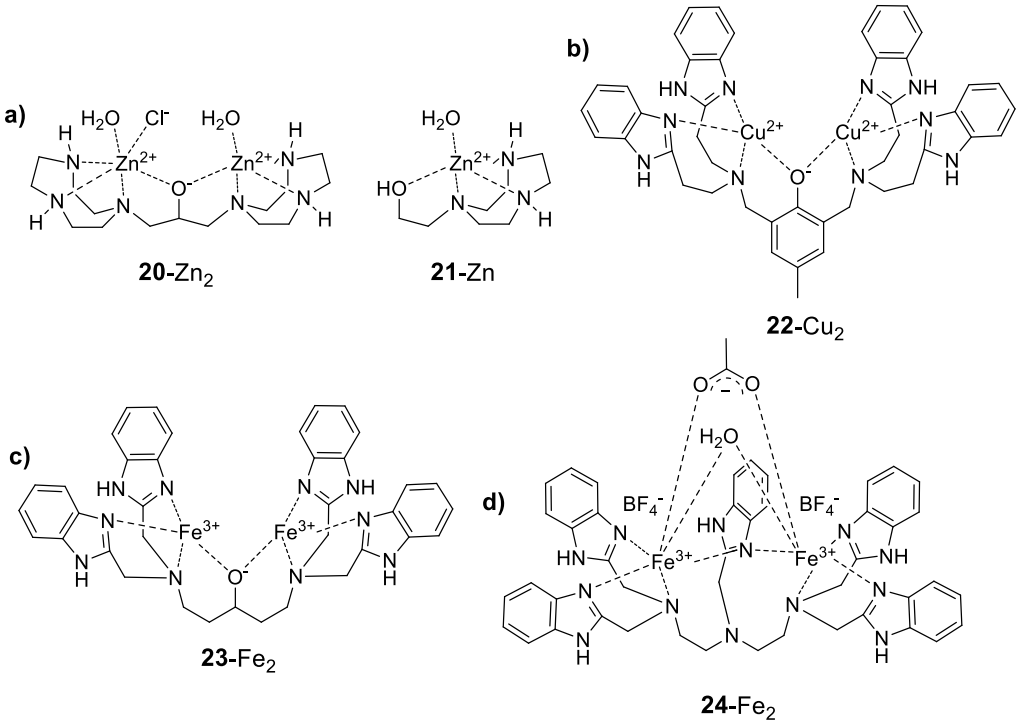


Figure 13: Dinuclear enzyme models exploiting μ-bridging units to keep in proximity their catalytic metal centers.

Ligand **20**[85], for instance, bridges two Zn(II) centers thanks to the action of an alkoxyde anion to form the corresponding dinuclear complex **20**-Zn_2, which was able to cleave HPNP and resulted superior to the reference mononuclear compound **21**-Zn (**Figure 13a**). **20**-Zn_2 and **20**-Cu_2 were also tested in the cleavage of a model for 5'-cap of mRNA[86], resulting both able to cleave effectively this substrate. Compound **22**-Cu_2 (**Figure 13b**) was one of the first Cu(II)-based artificial phosphodiesterase reported in the literature[87] and it was found to cleave very effectively HPNP thanks to a good cooperation of its two metal centers, kept in proximity by coordination to a phenoxyde anion. However, this compound failed into hydrolyze diribonucleoside monophosphates, whose phosphate groups were not accessible to the two copper ions probably because of steric hindrance. Interestingly, the di-Fe(III) complex of the very similar ligand **23** (**Figure 13c**) is known to be extremely effective in the cleavage of plasmid DNA[88]. In fact, in the presence of H_2O_2, **23**-Fe_2 was reported to degrade supercoiled DNA immediately after mixing and in very mild reaction conditions (10 μM complex, 25°C, pH 8.0). Since the DNA fragments produced during the process were found to be consistent with a hydrolytic cleavage mechanism, H_2O_2 was supposed not to act as oxidative agent, but to be instead bound to the metal centers and to attack the substrate as a nucleophile. Another very efficient catalysts was investigated by Liu et al.[89] and it was obtained from ligand **24** (**Figure 13d**), with the aim of producing a mimic of mammalian Purple Acid Phosphatase (**Figure 6a**). The corresponding complex **24**-Fe_2 was reported to be able to cleave plasmid DNA with a rate constant of $2.1 \cdot 10^{-3}$ s^{-1}, which is one of the highest reported in literature for enzyme models that work in the absence of oxygen and corresponds to a half-life of 5 minutes. In this case the role of bridging unit to ensure the proper intermetallic distance was played by either a benzimidazolmethyl group, directly attached to the molecular scaffold, and by a free μ–hydroxyde and a μ–acetate ions. Moreover, the presence of a labile BF_4^- per Fe(III) ion was reported to be crucial to increase the reactivity of the catalyst, by allowing an easy substitution with the substrate on the remaining coordination sites of the metal centers.

The presence of μ–bridging interactions can be also exploited to induce self-assembly of monomeric complexes to form corresponding dinuclear catalysts (**Figure 14**).

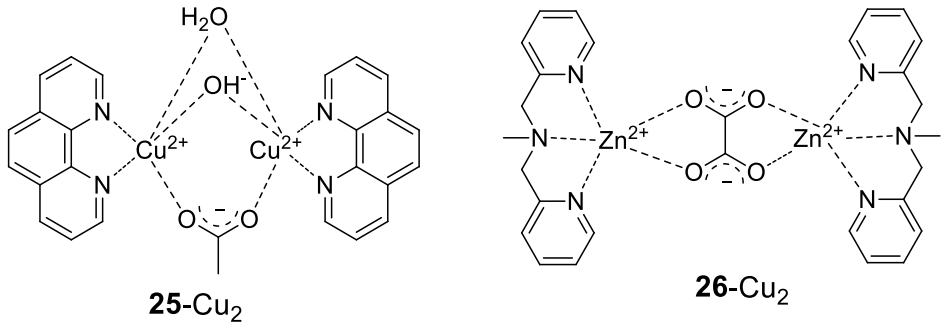


Figure 14: Self-assembly of mononuclear complex by coordination of μ-bridging units to obtain dimetallic artificial phosphodiesterases.

For example, complex **25**-Cu_2 was obtained by bridging two units of its mono-Cu(II) analogs with a water molecule, an hydroxide and an acetate ions, resulting effective in the cleavage of plasmid DNA[90,91] (apparent rate constant k_{obs}=5.2·10^{-4} s^{-1}). On the other hand, this di-Cu(II) compound was found to be not highly stable and for example was reported to co-crystalize with the DNA model compound BNPP in its monomeric form. Similarly, a bridging oxalate anion was used to obtain **26**-Zn_2, which was found to be moderately active as hydrolytic agent for plasmid DNA[92]. No cooperation for the two metal centers was reported, being **26**-Zn_2 only two-folds more reactive than the corresponding mono-Zn(II) complex.
The presence of bridging units in dinuclear artificial phosphodiesterases, however, is not always beneficial for catalysis, because sometimes it can lead to the saturation of coordination positions on the metal centers, thus resulting detrimental in the interaction with the substrate or in the binding of additional water molecules. Cases that support these statements are represented by complexes **27**-Zn_2, **27**-Cu_2 and **28**-Zn_2 (**Figure 15a**).

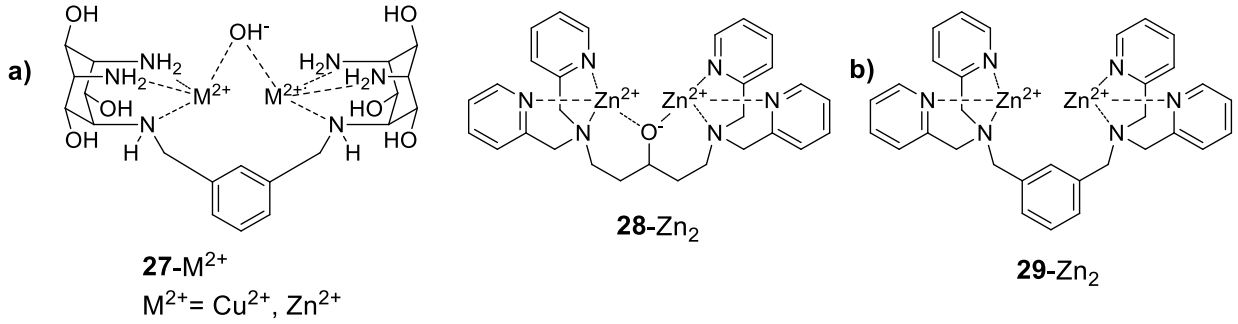


*Figure 15: Dinuclear phosphodiesterases where **a)** the presence of μ-bridging units disadvantage catalysis with respect to **b)** corresponding non-bridged complexes.*

Ligand **27** was synthesized by Mancin et al.[93] according to the good catalytic activity reported for Zn(II) and Cu(II) complexes of compound **2** (**Figure 10**) in the cleavage of simple di- and tri-phosphate esters and also of supercoiled DNA[93]. When tested on the RNA model compound HPNP, both the resulting dimetallic complexes **27**-Cu_2 and **27**-Zn_2 show no cooperation between the two metal cations, and **27**-Cu_2 was even found to be two orders of magnitude less active than its mononuclear control compound **2**-Cu. To explain these results the authors proposed the formation of stable μ-hydroxo bridge(s) between the metal centers which jeopardize the metal ion catalytic role. **28**-Zn_2 was instead tested by Komiyama in the transesterification of a small library of diribonucleoside monophosphate, together with **29**-Zn_2 (**Figure 15b**)[93]. Both the catalysts were effective in cleaving the substrates, but **28**-Zn_2 was reported to be about 3 times less effective with respect to the rather similar complex **29**-Zn_2. The lower reactivity of **28**-Zn_2 was ascribed to the presence of the bridging phenolic group, which allows the two Zn(II) to bind only an additional water molecule to provide general base catalysis, while **29**-Zn_2 has the possibility to coordinate two additional water molecules that can participate to the catalytic process.

Komiyama also compared the activity of **28**-Zn_2 with the trinuclear Zn(II) complex **30**-Zn_3 (**Figure 16a**)[94], finding that the latter was much more effective than the former in the cleavage of diribonucleoside monophosphates. This fact proved the participation of all three metal ions to the catalytic process and highlight this compound as one of the first trimetallic nuclease models. Other examples of trinuclear artificial phosphodiesterases are **31**-Cu_3 and **32**-Cu_3 (**Figure 16b**)[95]. These complexes were tested in the cleavage of a small group of diribonucleoside monophosphates and their activity was compared with a reference dimetallic complex **33**-Cu_2. Only for **31**-Cu_3 was reported a high degree of cooperativity between the three active units because of the higher flexibility of its molecular scaffold compared to **31**-Cu_3, which was in fact found to be a worse catalyst. **31**-Cu_3 had in some cases also a remarkable substrate selectivity, being able to cleave Up(2'-5')U 50-folds faster than the 3'-5' analogue Up(3'-5')U, while the reverse was found for Ap(2'-5')A and Ap(3'-5')A. No substrate selectivity was instead reported for **31**-Cu_3 and **32**-Cu_2.

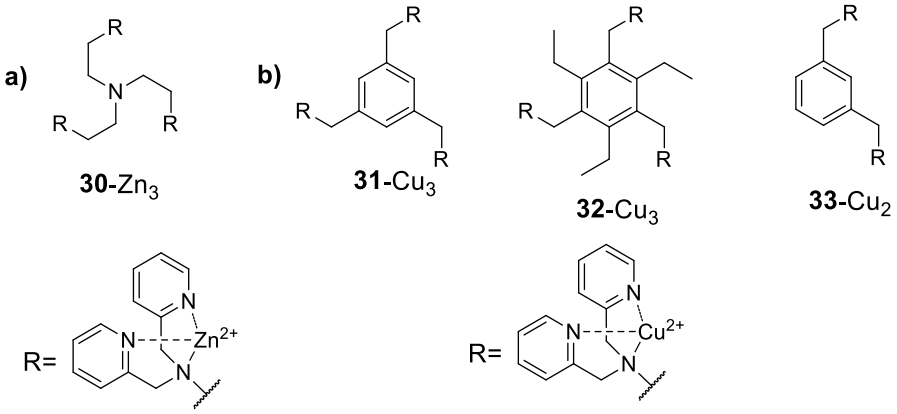


*Figure 16: **a)** Zn(II)-based and **b)** Cu(II) trinuclear phosphodiesterase and corresponding bifunctional control complex.*

1.4.4 Metal ion-organic functionalities cooperation in artificial phosphodiesterases

The high degree of cooperation between metal cation(s) and amino acid residues in the hydrolytic cleavage of phosphate esters observed in many enzymes[43–45] has been a source of inspiration for many chemists, that decided to add active units able to mimic the presence of amino acid side-chains on the molecular scaffold of metal-based artificial phosphodiesterase, in order to increase their catalytic activity.

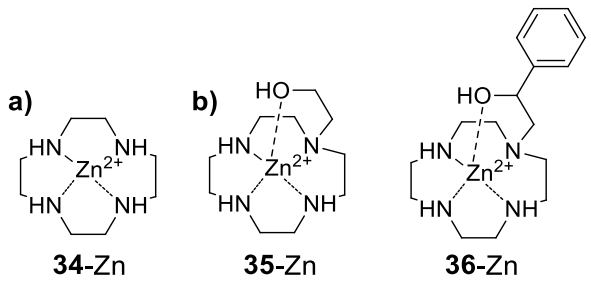


*Figure 17: **a)** Monofunctional Zn(II) complex of ligand **34** and **b)** difunctional analogs bearing different alcoholic pendants.*

One of the earliest example of these hybrid systems was investigated by Kimura et al. as a development of their studies on the Zn(II) complex of tetraazacyclododecane **34** (cyclen, **Figure 17a**)[95]. This complex was found to be effective in the cleavage of many carboxyesters and phosphate di- and triesters and, with the goal to mimic the active site of Alkaline Phosphatase (**Figure 7a**), they thought to insert different alcoholic pendants on the macrocyclic structure and to study their influence on the reactivity of the corresponding derivatives. The resulting **35**-Zn and **36**-Zn complexes[96] (**Figure 17b**) were able to cleave the DNA model compound bis-*p*-nitrophenylphosphate (BNPP) and the alcoholic pendants were found to effectively cooperate with the metal center by enhancing the reaction rate of one order of magnitude compared to the simple metal complex **34**-Zn. Extensive studies on the mechanism of action of compound **36**-Zn pointed out that BNPP undergoes hydrolysis in two steps (**Figure 18**)[96]: the first involving a nucleophilic attack of the alcoholic unit on the phosphate group to form a phosphorylated intermediate and to release a *p*-nitrophenolate anion, while the second step includes a second nucleophilic attack by a water molecule, with the departure of a second equivalent of *p*-nitrophenolate. In these steps, both the alcoholic pendant and the water molecule are activated by coordination to the Zn(II) center. The resulting phosphorylated adduct was found to be highly inert to hydrolysis and for this reason **36**-Zn cannot be considered a catalyst, but, more correctly, a promoter of phosphoryl transfer processes, unless it mimics correctly the action of the corresponding natural enzyme.

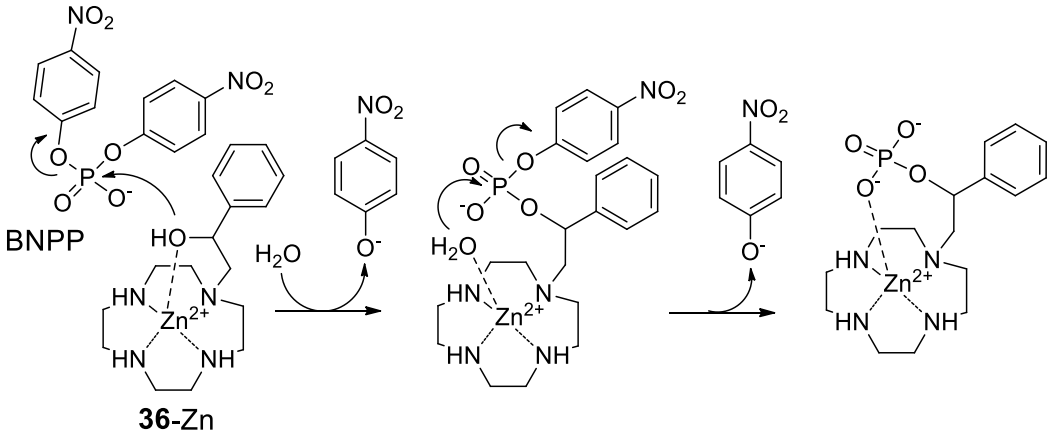


*Figure 18: Mechanism of action of **36**-Zn in the cleavage of BNPP.*

Other hybrid systems where organic functionalities are mixed with metal ions were presented by Anslyn et al. The Zn(II) complexes of ligands **37**, **38**, **39** (**Figure 19**) were tested in the hydrolysis of the ApA diribonucleotide[97], giving pseudo first order rate constant of $2.2 \cdot 10^{-5}$ s^{-1}, $2.5 \cdot 10^{-6}$ s^{-1} and $6.7 \cdot 10^{-9}$ s^{-1}, respectively. These results indicate a good degree of cooperativity between the Zn(II) ion and the ammonium or guanidinium groups of **38**-Zn and **39**-Zn, which were approximately 370- and 3300-folds more active than **37**-Zn. The Zn(II) complex of ligand **39** was found to be the best catalyst of the group, with a rate enhancement for the cleavage reaction of 10^6 times respect to the background hydroxide-

catalyzed reaction. This high reactivity was ascribed to the superior ability of the guanidinium groups to provide beneficial interactions with the substrate by hydrogen bonding and other electrostatic effects.

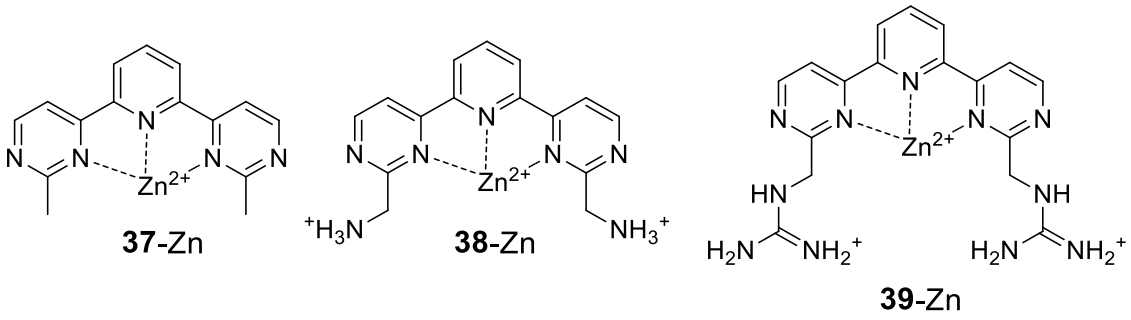


*Figure 19: Zn(II) complex of ligand **37** and corresponding hybrid systems functionalized with an additional ammonium or guanidinium pendant.*

The high effectivity of guanidinium groups in increasing the catalytic activity of phosphodiesterase models is not casual. In fact, as part of the side chain of the amino acid arginine, it plays a leading role as active unit in corresponding natural enzymes[57,58], where it can interact with phosphate esters via two point hydrogen bonding and other pattern of interactions. These interactions are generally used to bind the substrate and to stabilize the transition state, with the synergic participation of metal cations or other amino acid residues. Additionally, arginine side chains are in some cases suggested to act as general acids by protonating the leaving group and thus facilitating its departure. Not surprisingly, thus, this functional group is probably the most used non-metallic active unit to produce artificial phosphodiesterases.

Very interesting examples of metal ions-guanidinium cooperation are shown in **Figure 20a**. In these cases, He and coworkers functionalized a 2,2'-bipyridyl scaffold in 3,3' and 4,4' positions with two methylguanidinium pendants to form ligands **40** and **41**, respectively, and to study the catalytic activity of corresponding Zn(II) or Cu(II) complexes. **40**-Zn and **41**-Zn were tested in the cleavage of BNPP[98], resulting 600- and 300-folds more effective of their analog **42**-Zn that lacks the guanidinium groups. The copper complex of **41** was also reported to hydrolyze supercoiled DNA[99] 10-folds faster than the corresponding control compound **42**-Cu. The plot of the rate constant against the complex concentration gave a Michaelis-Menten profile, thus proving the existence of a binding process with the substrate in the investigated concentrations range. According to the crystal structure of the ligand, the binding of the substrate was suggested to take place thanks to the anchoring of the guanidinium units to the two phosphate groups adjacent to the scissile one, while the cleavage of the substrate was attributed to the nucleophilic attack of a metal-bound hydroxide ion (**Figure 20b**).

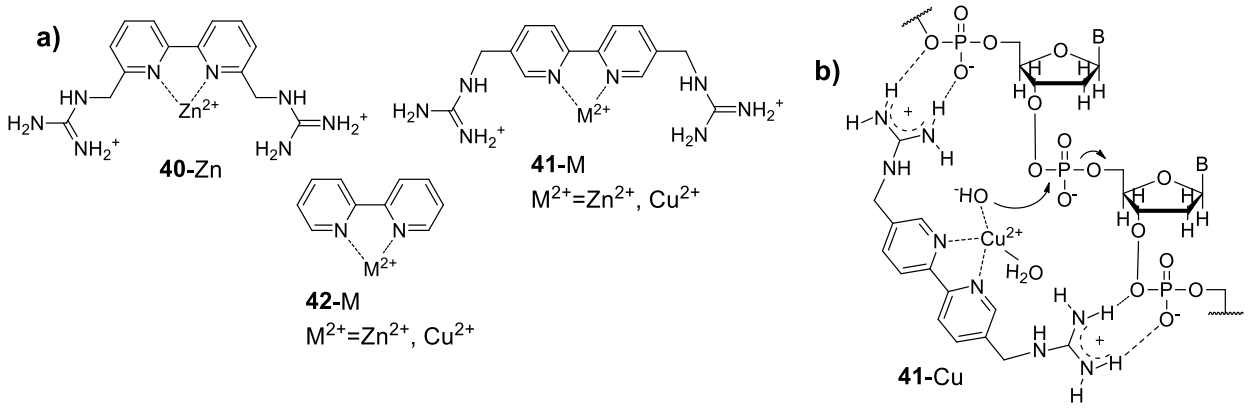


*Figure 20: **a)** Hybrid metal ion- guanidinium-based artificial phosphodiesterases and **b)** mechanism of action of complex **41**-Cu in the cleavage of plasmid DNA.*

Ligands **43** and **44** were synthesized by Graham and Spiccia, by attaching two guanidinium pendants to a triazacyclonane unit (TACN) with different alkyl spacers (**Figura 21a**)[100]. The length of the spacers influenced the catalytic activity of corresponding Cu(II) complexes, in fact **43**-Cu was found to be inactive in the cleavage of both HPNP and BNPP, while **44**-Cu resulted respectively 10- and 40-times faster than the reference TACN-Cu complex in the hydrolysis of the same substrates. The loss of catalytic activity of **43**-Cu was explained by the fact that, as found in its X-ray crystal structure, the guanidinium groups are close enough to the Cu(II) ion to participate to its complexation, thus avoiding a correct interaction with the substrates. Both complexes were instead able to cleave plasmid DNA, thus indicating that, probably, this substrate is able to displace the guanidinium units from the coordination sphere of **43**-Cu. Also in this case **44**-Cu was found to be the most effective catalysts, resulting approximately twice as active as the control complex TACN-Cu, thanks to a moderate cooperativity of the guanidinium groups with the metal center.

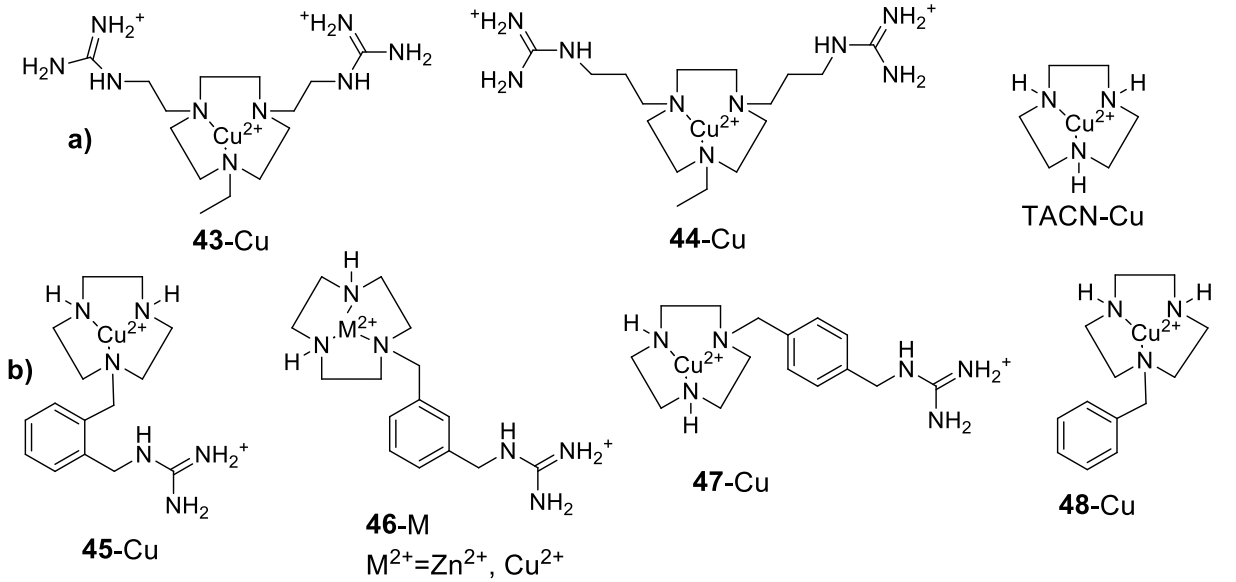


Figure 21: Difunctional artificial phosphodiesterases obtained by linking with different spacers a TACN-M complex (M= Zn(II) or Cu(II)) and one or two guanidinium units and corresponding monofunctional metallo-catalysts.

With the aim of increasing the rigidity of the molecular scaffold and the distance between the catalytic groups, the same researchers disposed these two active units on a xylylene spacer in different orientations, thus obtaining ligands **45**, **46**, **47** (**Figure 21b**)[101]. The related Cu(II) complexes, in aqueous HEPES buffer, pH 7.0, were not significantly more active towards HPNP and BNPP hydrolysis when compared to the non-guanidinylated analogues TACN-Cu and **48**-Cu, but turns to be noticeably faster than both control compounds in the cleavage of plasmid DNA. In particular, **46**-Cu was reported as the best catalyst of this family of compounds and accelerated the cleavage process of 22- and 12-folds compared to TACN-Cu and **48**-Cu, respectively. Interestingly, also Salvio et al. studied the activity of Zn(II) and Cu(II) complexes of ligand **46**[102] and found that, in different experimental conditions (DMSO/H_2O 8/2, pH 9.0 and 9.8), the guanidinium pendant and the metal centers cooperate to cleave HPNP. In fact, in the presence of **46**-Zn and **46**-Cu, the transesterification of the substrate was accelerated by 6- and 40-folds with respect to the monofunctional TACN-Zn(II) and TACN-Cu(II) complexes, respectively.

1.4.5 Metal-free artificial phosphodiesterases

As reported in **Paragraph 1.3.1**, phosphodiesterase enzymes do not always need metal cations as cofactors and can cleave their substrates thanks to the action of the side chains of amino acid residues[54–56]. More in detail, metal free-enzymes use positively charged amino acid residues to interact with the scissile phosphate group, in order to stabilize the negatively charged transition state during the attack of the incoming nucleophile. The role of nucleophile can be played both by an external water molecule and by the alcoholate $-O^-$ or thiolate $-S^-$ groups belonging to the side chain of other amino acids presents in the active sites. Moreover, a general base activation of the nucleophile or a general acid stabilization of the leaving groups can be provided by these important type of functionalities. Accordingly, many researchers focused their efforts into the synthesis of artificial models devoid of metal centers and where all the catalytic functions of the active units have to be performed by organic moieties.

The earliest studies in this field were carried out in the cleavage of RNA sequence or model compounds, because they undergo hydrolysis thanks to the intramolecular attack of a -OH group present in β-position[36] of their backbone. In principle the process can be activated simply by the presence of a general base that promotes the β-OH deprotonation, eventually assisted by the corresponding conjugated acid that can provide electrostatic and/or general acid catalysis. According to this idea, Breslow et al. studied the cleavage of a poly-U RNA sequence by an imidazole buffer (**Figure 22a**)[103]. The plot of the reaction rates versus the pH gave a bell-shaped curve that suggested the participation of both imidazole and imidazolium species in a sequential reaction mechanism[104] that reminds the one observed for bovine Ribonuclease A[56] (**Figure 8b**).

Similar results were reported in a study of Komiyama et al. concerning the cleavage of the ApA diribonucleoside by diamines **49**-**55**[105] (**Figure 22b**). In fact, they found that the only diamines able to cause a noticeable acceleration of the substrate hydrolysis were compounds **50**-**53**, which were also the only ones to be at least partly in their monoprotonated form at neutral pH, together with **49**. These compounds were thus supposed to promote the ApA cleavage by exploiting the intramolecular cooperation of their amino and ammonium groups to deprotonate the ribose 2'-OH and to stabilize the negatively charged phosphate group, respectively. In fact, the control monoamines **56**-**58** and the more basic compounds **54** and **55** tended to be in their fully protonated form in the same experimental conditions, resulting almost catalytically inactive. The lack of reactivity for the monocationic form of **49** was instead ascribed to the marginal basicity of its deprotonated amino group.

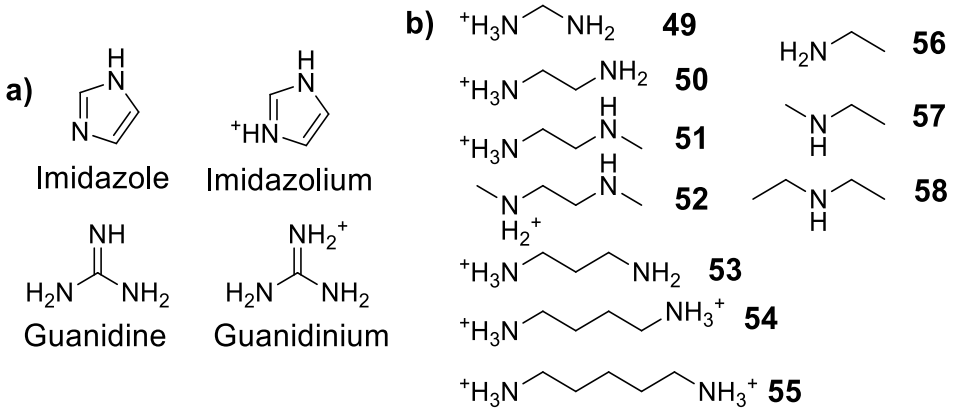


*Figure 22: **a)** Conjugated acid-base couples and **b)** polyamines tested in the cleavage of RNA model compounds.*

The importance of the cooperation of a general base and its conjugate acid in the cleavage of phosphate esters was also reported in the case of the guanidine/guanidinium buffer. A detailed kinetic study of the transesterification of HPNP (DMSO/H_2O 80/20) was performed by Yatsimirsky et al.[106] and it pointed out that, in the presence of a guanidine/guanidinium dyad (**Figure 22a**), the reaction rate is 1000-folds higher compared to the background. On the basis of the experimental data the authors concluded that the buffer cleaves the substrate thanks to the cooperation of both its protonated and deprotonated forms. The protonated form was found to act via electrostatic stabilization of the transition state rather than via proton transfer to the leaving group, at least when, as in this case, a general base catalysis is present and the leaving group is a good one. In the case of hydroxide-catalyzed reactions or in the presence of poor leaving groups the proton transfer can instead occur.

In order to exploit these properties intramolecularly, Salvio et al.[107] synthesized compounds **59**-**61** (**Figure 23a**)**.** All the catalysts were effective in the cleavage of HPNP and the kinetic data, together with the effective molarity values, confirmed that their monoprotonated form was the only catalytically active species and that the guanidine/guanidinium dyad cooperate with a high degree of synergism. The reactivity of the three compounds increase with the rigidity of their molecular scaffold caused by the

steric hindrance produced on the diphenylmethane spacers. In fact, **61** was found to be 2- and 9-times more effective than **60** and **59**, respectively, thanks to the presence of a bulky adamanthyl group that disfavours the rotations around the C-phenyl bonds, thus orienting the two active units in a position suitable for catalysis. Evidences of general acid-base cooperation in the cleavage of HPNP were reported by the same authors also for compound **62** (**Figure 23b**)[102], which was designed with a slightly different approach. In this case a hybrid acid-base couple was formed by attaching to a m-xylylene platform a guanidinium and an imidazole pendant, where the role of general base is expected to be played by the latter. Unless this compound afforded only modest acceleration of the transesterification process compared to the background, the cooperation of the two active units was clearly detectable, being around 6-folds more effective than the control monofunctional compound **63**.

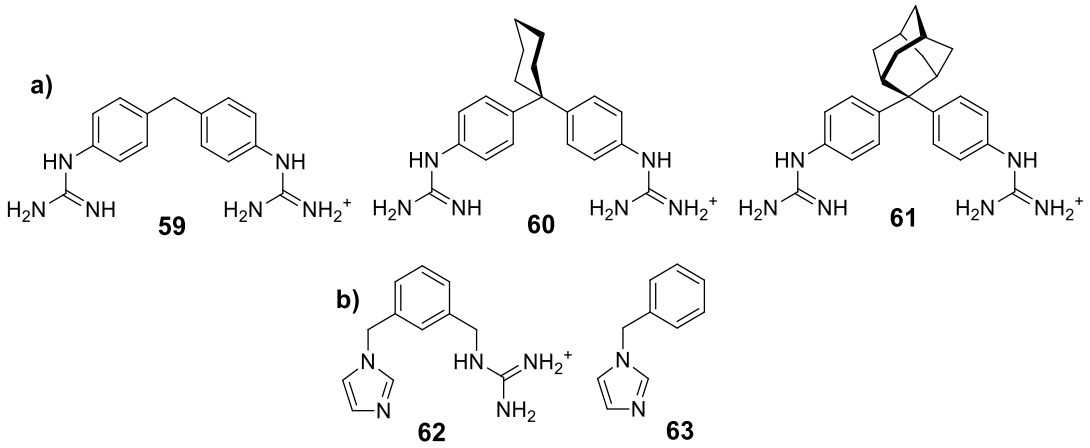


*Figure 23: Artificial phosphodiesterases bearing **a)** a guanidine-guanidinium dyad and **b)** a imidazole-guanidinium dyad as active units with the corresponding monofunctional compound.*

However, the contemporary presence of a general base is not strictly necessary to obtain guanidinium-based phosphodiesterase. According to what has been observed in some enzymes[55,57,59], in fact, the ability of this functionality to provide ion-pair interactions with the phosphate group can be combined to the action of active units that mimic the presence of nucleophilic amino acid residues. Compounds **64**-**67** were synthetized by Göbel et al. according to this approach and were tested in the transphosphorylation of 1,2-phenylcyclophosphate **68** (**Figure 24**). The first catalyst to be tested was **64**[108] and it hydrolyzed the substrate thanks to the nucleophilic attack of its alcoholic pendant 2700-folds faster than the parent compound which miss the guanidinium moiety. In a subsequent study[109], the system was improved by synthesizing the two bisguanidinium derivatives **65** and **66**, which, in the presence of a base, were phosphorylated readily and reversibly, resulting 800- and 75- folds more effective than **64**, respectively. The increase in catalytic activity observed when unsubstituted guanidinium groups are replaced with heterocyclic guanidinium cation was ascribed to their higher acidity, that allow the formation of tighter interactions with the phosphate group and/or a more effective protonation of the leaving group.

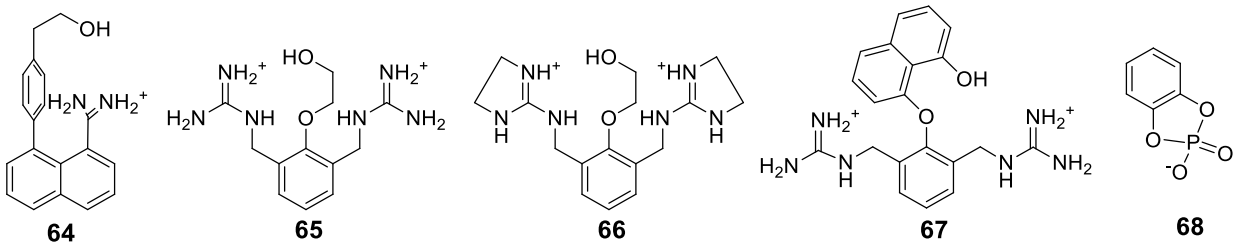


*Figure 24: Guanidinium derivatives exploiting the action of alcoholic or naphtholic pendants in the transphosphorylation of compound **68**, BNNP and plasmid DNA.*

More recently the same group presented compound **67**[110] as a further development of this family of enzyme models. This compound was tested in the cleavage of the BNPP in DMF, with an initial rate of $3.7 \cdot 10^{-6}$ s^{-1} to yield the corresponding phosphorylated analog. HPLC and ^{31}P-NMR studies proved that the reaction takes place via O-phosphorylation of the naphthol moiety rather than N-phosphorylation of a guanidine group. This intermediate then slowly hydrolyzed in DMF/H_2O mixtures to regenerate the starting catalyst and phosphate monoester. The same catalyst was also tested in the cleavage of plasmid DNA and, in HEPES buffer at 37°C, was able to degrade 37% of the substrate after 20 hours of incubation. Appropriate control experiments proved that the DNA degradation is not metal dependent and that it occurs mainly via hydrolytic scission of the phosphate bridges, although pathways that involve naphthoxy radicals cannot be completely excluded.

The activity of compound **69** (**Figure 25**) was instead investigated by Xu and co-workers[111]. This catalyst was able to induce a 10^7-folds acceleration for the cleavage of supercoiled DNA respect to the background, resulting much more effective than the parent non-guanidinium compounds **70** and TACN. Unexpectedly, **63** was around 3 times more effective than its corresponding Zn(II) complex **69**-Zn in the same process. To explain this loss of catalytic activity the authors suggest that, at neutral pH, the guanidinium pendant can coordinate to the metal cation, thus disfavouring the cooperation between the active units of the catalyst.

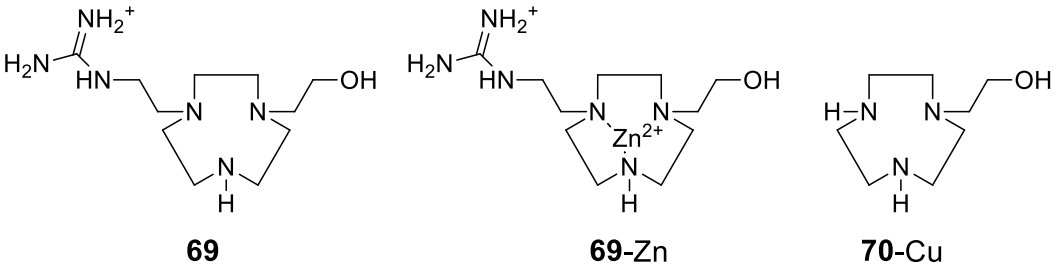


Figure 25: Artificial phosphodiesterases bearing a guanidinium and an alchoholic groups as active units, corresponding Zn(II) complex and parent non-guanidinium monofunctional analog tested in cleavage of supercoiled DNA.

In this paragraph have been described the main strategies used in past years to obtain artificial phosphodiesterases by assembling metal ligands or organic functional groups on a simple molecular

scaffold. Nevertheless, in literature there are several example of enzyme models obtained with more sophisticated architectures[61–63]. For example, these type of catalytic units have been conjugated to DNA or RNA recognition groups like intercalators and complementary oligo(ribo)nucleotides, PNA or peptide sequences, in order to increase the binding to these substrates and to induce site-specific cleavage. Another strategy that has been reported is the functionalization of nanoparticles with spacers bearing at one end metal complexes or guanidinium units, in order to exploit the multivalency effect to increase the reactivity. However, the detailed description of these strategies is out of the topic of this thesis, because the approach used to create this kind of enzyme models was different.

In fact, in the following paragraph will be presented the synthesis and the phosphodiesterase activity of supramolecular catalysts obtained by the arrangement of proper catalytic units on a *cone*-calix[4]arene platform. This molecular scaffold has been reported as very convenient for the creation of artificial catalysts, thanks to its high ability to provide *dynamic preorganization*[68] of the active units which are consequently able to cooperate in the promotion of the studied reaction. Thus, the most important example of calixarene-based artificial phosphodiesterases will be now reviewed.

1.4.6 Calix[4]arene-based artificial phosphodiesterases

The extreme versatility of calix[4]arenes scaffold is at the base of their intensive use in supramolecular catalysis[68,112–114]. In fact, this kind of molecular scaffold can be easily and selectively derivatized, with the possibility to insert, both at the upper and at the lower rims of the macrocycle, a wide variety of functional groups in different reciprocal orientations. Moreover, the opportunity to finely tune the conformational properties of this platform, contemporary preserving a certain degree of flexibility, can be exploited by chemists to achieve effectively the dynamic arrangement of the active sites of enzyme models which is required for catalysis. Additionally, the molecular cavity that constitute the backbone of these family of cyclophane can act as host in the substrate recognition step, having thus a potential active role in increasing the catalytic activity of the system.

Tacking into account the remarkable catalytic activity of many metal-based artificial phosphodiesterases, the first generation of calix[4]arene-based artificial phosphodiesterases was developed by Reinhoudt et al. by preparing the Zn(II) complexes of compounds **71**-**74** (**Figure 26**)[115,116]. **71**-Zn_2 was found to cleave HPNP with turnover (CH_3CN/H_2O 1/1, pH 7, 25°C), giving an acceleration of the process of 23000-folds compared to the background[116]. This catalyst was also reported to be 300-times more active than the corresponding monofunctional complex **72**-Zn, which was in turn 6-folds more effective than the parent simple Zn(II) complex of ligand **75**. According to these data, the high catalytic activity of **71**-Zn_2 was supposed to take origin both from a very effective cooperation between the two metal centers to bind and activate the substrate and from a direct participation of the calix[4]arene cavity to the transesterification process, probably by causing a further lowering of the pKa of a metal-bound water molecule which act as general base once converted into the corresponding

hydroxyl ion. The hypothesis that the residual flexibility of the *cone*-calix[4]arene scaffold could be beneficial to enhance the degree of synergism in the action of the two active units of **71**-Zn_2 was subsequently confirmed by the same authors. In fact, they found that the rigid *flattened cone*-calix[4]arene-based analog **73**-Zn_2, where the mobility of the macrocyclic cavity is suppressed by the presence of a short crown-ether bridge at its lower rim, cleaves HPNP 8-times slower[115].

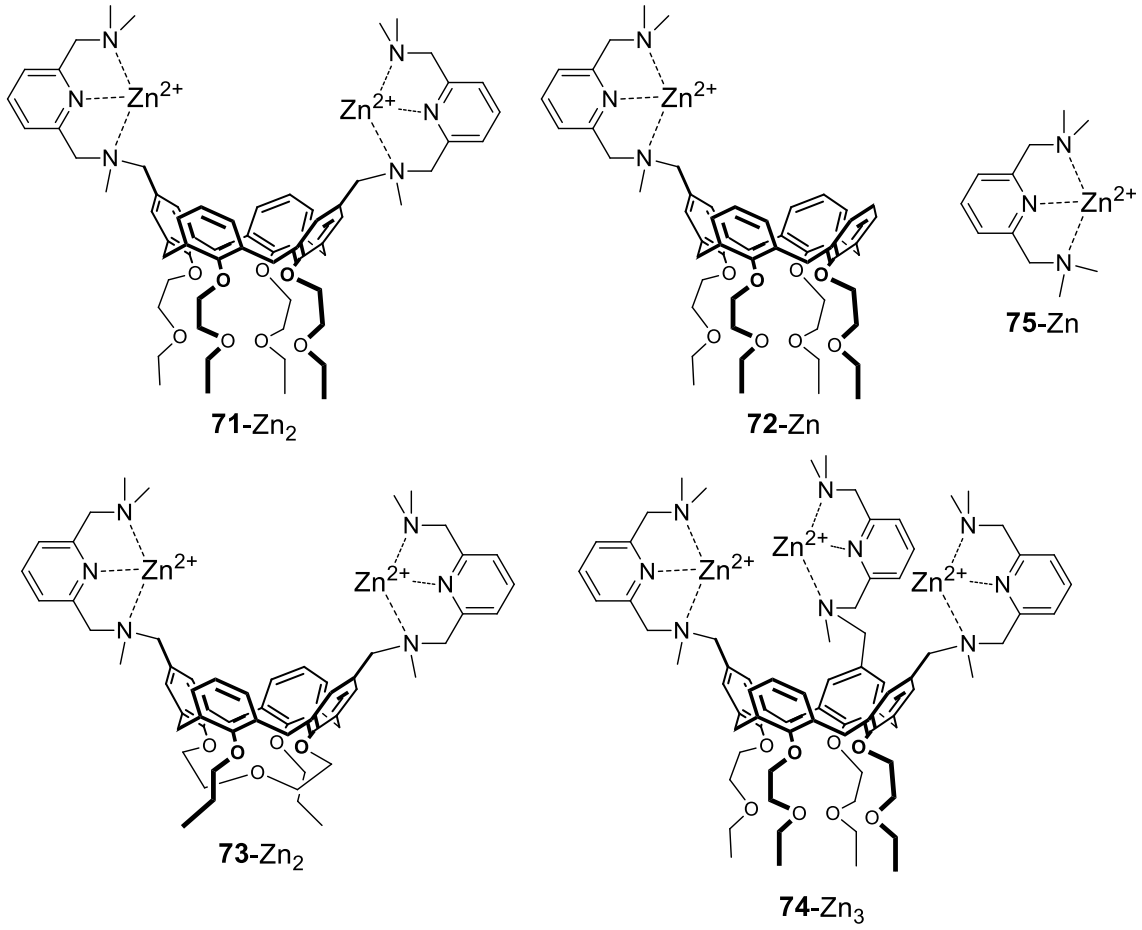


*Figure 26: First generation of calix[4]arene-based artificial phosphodiesterases and simple non-calix[4]arene control complex **75**-Zn.*

A noteworthy phosphodiesterase activity was also reported for the di-Cu(II) complex **76**-Cu_2 (**Figure 27**)[117], which was found to cause the cleavage of HPNP 10000-times faster than the background reaction (EtOH/H_2O 3.5/6.5, pH 6.4). The catalyst was 330-folds more effective than its monometallic control complex **77**-Cu, representing another interesting example of metal-metal cooperation in a dinuclear catalyst. Contrary to **71**-Zn_2, **76**-Cu_2 was also able to cleave the DNA model compound EPNP, whose hydrolysis was accelerated of 27000-folds with respect to the background.

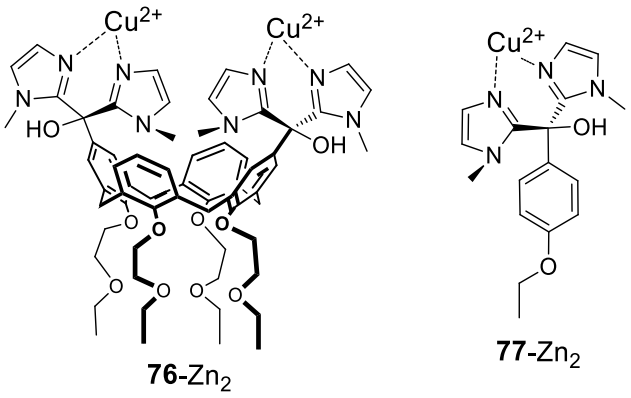


*Figure 27: Dinuclear Cu(II) complex of calix[4]arene **76** tested in the cleavage of HPNP and EPNP and corresponding monofunctional control compound.*

The different activity of the two catalysts was proposed to be due to the different average distance between their metal centers. In fact, in **76**-Cu_2, the two Cu(II) cations were reported to be close enough to provide double Lewis acid activation of the phosphate of EPNP or HPNP, facilitating the participation of a metal-bound water molecule to the cleavage process as nucleophile (**Figure 28a**) or general base, respectively. In **71**-Zn_2, the two metal centers are, instead, too far each from the other and the catalyst can afford only a single Lewis acid activation of the phosphate group, which, in the case of EPNP, is not strong enough to allow the attack of a metal-bound hydroxide on it. The transesterification of HPNP was indeed effectively promoted by this catalyst[116], because its Zn(II) ions were reported to be correctly oriented to provide a contemporary activation of the phosphate group and of the β-OH of the substrate, thus accelerating its cleavage (**Figure 28b**).

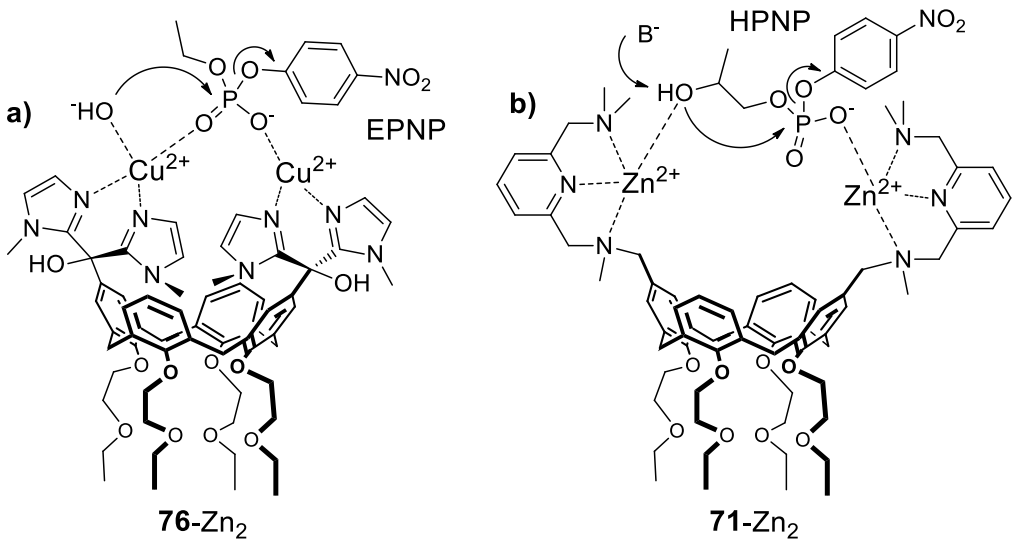


*Figure 28: Mechanism of action of **a) 76**-Cu_2 in the hydrolysis of EPNP and **b)** of **71**-Zn_2 in the transesterification of HPNP.*

Both **71**-Zn_2 and **76**-Cu_2 were tested in the transesterification of different diribonucleoside monophosphates, resulting however almost inactive for these substrates. An interesting increase of

reactivity was observed with the insertion a third metal center to the system, by synthesizing complex **74**-Zn_3[118]. In fact, together with a 3-folds enhancement for the cleavage rate of HPNP with respect to its dinuclear analog **71**-Zn_2 (CH_3CN/H_2O 1/1, pH 7, 25°C), this trifunctional catalyst gave high acceleration to the transesterification of the considered ribonucleoside dimers with catalytic turnover, thanks to the cooperation between its three Zn(II) ions. The maximum rate accelerations were of $72 \cdot 10^5$- and $8.5 \cdot 10^5$-folds compared to the background for GpG and UpU, respectively, which were cleaved 160- and 19-times faster than ApA. Substrate saturation studies pointed out that the selectivity of the catalyst for the two former diribonucleosides rises from a higher binding affinity, which has been ascribed to the coordination of the imidic nitrogen of uridine or guanosine by one of the metal centers.

Despite these calix[4]arene-based catalysts showed a remarkable phosphodiesterase activity, their poor solubility in water is an handicap for their use as biomimetic systems. To overcome this drawback, Ungaro et al. decided to synthesize calix[4]arenes **78**-**80** (**Figure 29a**)[119], decorated at the upper rim with two or three triazacyclododecane units ([12]aneN_3).

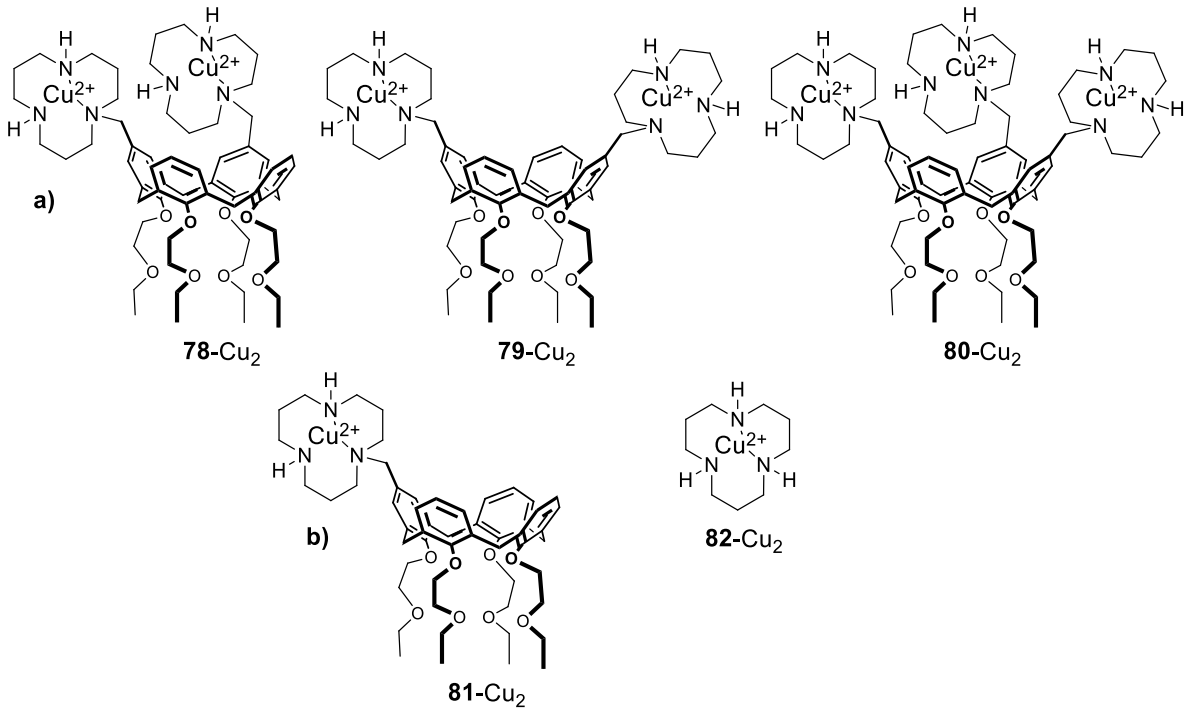


*Figure 29: **a)** Water-soluble Cu(II) complexes of calix[4]arenes **78**-**80** and **b)** parent monofunctional compounds tested as artificial phosphodiesterases.*

The solubility of corresponding Zn(II) complexes was still too low for catalytic measurements, but complexes **78**-Cu_2, **79**-Cu_2 and **80**-Cu_3 were soluble enough to carry out a detailed study of their activity in aqueous medium. The 1,2-vicinal complex **78**-Cu_2 was first tested in the cleavage of HPNP in HEPES buffer (25°C, pH 7) and was able to accelerate of 1090-folds the transesterification process over the buffer-catalyzed reaction. Acceleration of 42- and 73-times respect to the monofunctional compounds **81**-Cu and **82**-Cu (**Figure 29b**), respectively, pointed out once again a high degree of synergism between

the two metal centers and a beneficial influence of the calix[4]arene cavity for catalysis. Interestingly, the 1,3-distal regioisomer **79**-Cu_2 was a poor catalyst for the same substrate and did not exploit any metal-metal cooperation, being only 40- and 2-folds more effective than the background and **81**-Cu, respectively. Also in this case, thus, the catalytic activity of the studied calix[4]arene derivatives was dependent on the geometrical disposition of the active units at the upper rim of the macrocycles. More in details, experimental data for the transesterification of HPNP by **78**-Cu_2 were in agreement with a mechanism of action where the two Cu(II) cations provide double Lewis acid activation of the phosphate group, while a metal-bound hydroxide ion act as general base to deprotonate the β-OH moiety (**Figure 30a**). However, according to CPK models, to afford this pattern of interactions **79**-Cu_2 need to adopt a *pinched cone* conformation where the two [12]aneN_3 units produce a high steric hindrance around the phosphate group which disfavours the attack of the β-hydroxyl group on it, thus causing a loss in catalytic activity of the 1,3-distal complex (**Figure 30b**). Accordingly, the trimetallic complex **80**-Cu_3, always for the cleavage of HPNP, was found to be approximately as effective as **78**-Cu_2**,** confirming that these systems prefer to work by exploiting the action of the active units present at the 1,2-vicinal positions of their upper rim.

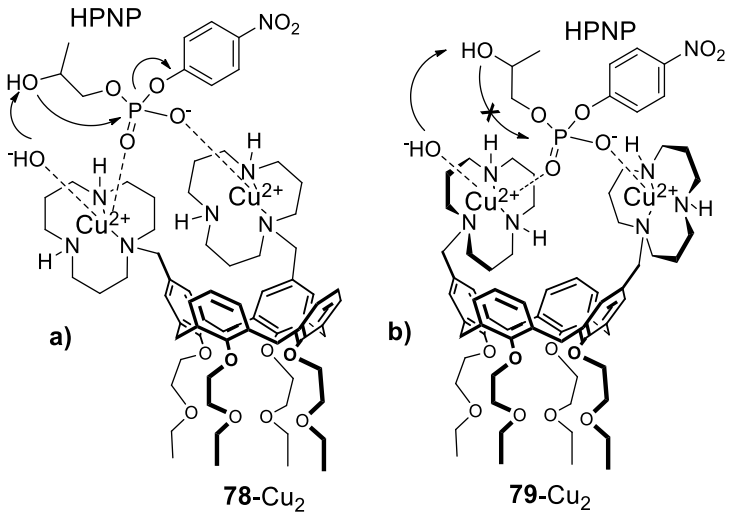


*Figure 30: Mechanism of action of **a) 78**-Cu2 and **b) 79**-Cu2 in the cleavage of HPNP.*

Being the most soluble of this group of Cu(II) complexes, **80**-Cu_3 was also tested as phosphodiesterase agent for a small library of diribonucleoside monophosphates.[119] This complex caused an enhancement of the transesterification rate in the order of 10^4-folds compared to the background for most substrates. As in the case of **74**-Zn_3, the catalyst had a preference for uridine- and guanine-containing dimers and it was able to accelerate the cleavage of UpU and UpG of 5 orders of magnitude. Unexpectedly, **80**-Cu_3 was not active towards CpA, but its inertness towards this substrate was ascribed to the formation of an unproductive adduct with it rather than to a problem of intrinsic reactivity. In fact, the control monofunctional compound **81**-Cu was found to be able to cleave this diribonucleoside. The

transesterification of UpU was also studied in the presence of **78**-Cu_2, whose catalytic activity was similar to the one of its trifunctional analog, thus supporting the idea that in **80**-Cu_3 the third metal center act essentially as innocent spectator.

The Cu(II) complexes of calix[4]arenes **78**-**80** were also able to hydrolyze a series of oligoribonucleotides (ACCAUC, CGCUGA, AGGUUAA, CAGGCC, CCGGCA, ACUAUC) with high efficiency[120], but for these substrates was very difficult to understand the factors that control the reactivity of the catalysts. In fact, if on the one hand a certain tendency for the catalytic efficiency to increase with the number of metal centers was observed and **80**-Cu_3 was in general the best catalyst of the lot, on the other hand, in some cases, the dinuclear complexes were superior to their parent trinuclear analog and, at contrary to what was reported for HPNP, the 1,3-distal derivative **79**-Cu_2 was more effective than the 1,2-vicinal isomer **78**-Cu_2. In marked contrast to the type of substrate-selectivity observed for diribonucleoside monophosphates, all the catalysts had a clear preference for the cleavage of the CpA bonds of the studied oligoribonucleotides. For example, **80**-Cu_3 was able to cleave the CAGGCC hexamer at the CpA bond with an acceleration factor of $5 \cdot 10^5$-folds respect to the background, which is one of the highest values reported in literature for the cleavage of oligoribonucleotides by an artificial metal-based phosphodiesterases. This picture was also confirmed in the reaction of longer oligoribonucleotide (5' GCAAGCACAGACAUCAG 3'), in which the CpA bonds were the only ones to be cleaved by **78**-Cu_2 and **79**-Cu_2, while bonds between other nucleotides did not react at a feasible rate.

The *cone*-calix[4]arene scaffold is an appropriate platform to obtain also effective metal-free artificial phosphodiesterases. This type of enzymes models have been synthesized by the groups of Casnati and Salvio[121] and exploit the action of two to four guanidinium groups as active units.

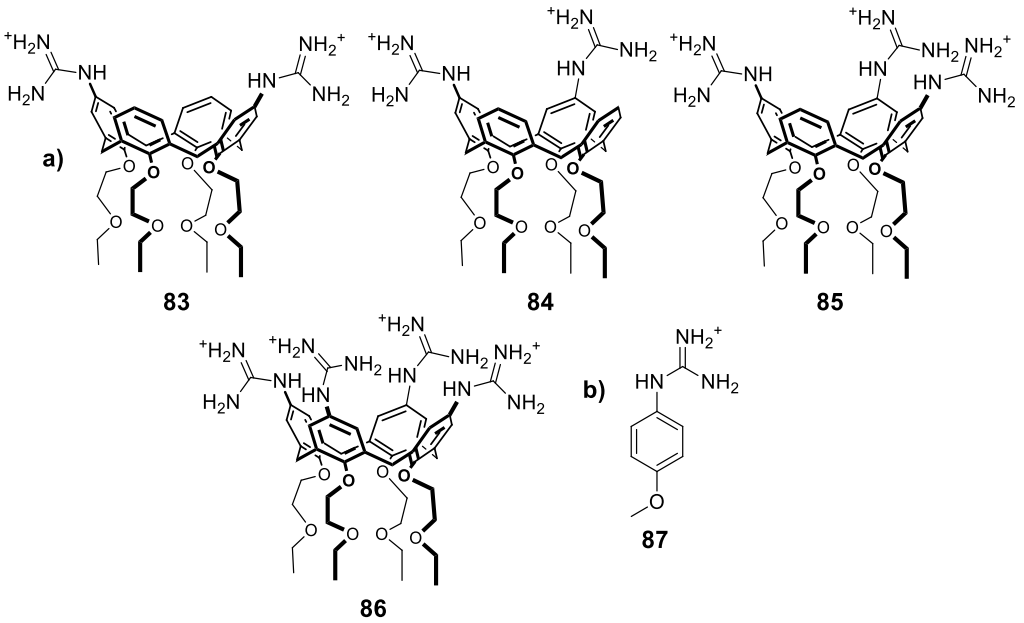
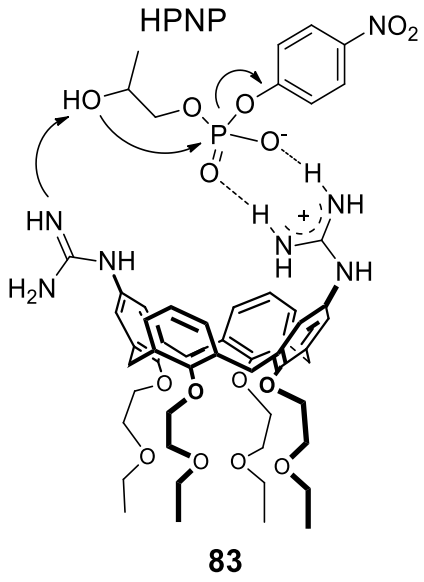


*Figure 31: **a)** Metal-free artificial phosphodiesterase based on guanidino-calix[4]arenes and **b)** monofunctional control compound.*

Compounds **83**-**86** (**Figure 31a**) were first tested in the transesterification of HPNP in DMSO/H_2O 80:20 in the pH range 9-12 which, in this reaction medium, correspond to neutral or moderately basic conditions (pK_w = 18.4 at 25 °C). As reported in literature for other guanidinium-based catalysts, the combination of pH-rate profiles and potentiometric titrations pointed out that these systems require the presence of a guanidine-guanidinium dyad at their upper rim to be active towards the cleavage of the substrate. The best catalyst of this family was the 1,3-distal difunctional calixarene **83**, which accelerated the substrate cleavage of 430-folds compared to the background (DMSO/H_2O 80/20, pH 10.4, 25°C), resulting twice as effective as the corresponding 1,2-vicinal isomer **84**. The catalytic activity of these compounds was proposed to rise from a good degree of cooperation between the two active units, with the guanidine moiety acting as general base and the guanidinium group as an electrophilic activator (**Figure 32**). For example, **83** was found to be three order of magnitude faster in the cleavage of the substrate than a buffer solution of control compound **87** (**Figure 31b**), were the active guanidine-guanidinium dyad is not connected on the same molecular scaffold. The presence of additional guanidinium cations did not cause any enhancement of catalytic activity for the trifunctional calix[4]arene **85**, where the third active unit was thus reported to act as innocent spectator, and led even to a loss of reactivity for the tetrafunctional compound **86**, probably because of overcrowding.



83

*Figure 32: Mechanism of action of **83** in the cleavage of HPNP.*

Calix[4]arenes **83** and **84** were also tested in the cleavage of some diribonucleoside monophosphates[122]. The reaction rate enhancement was in the order of 10^3 – 10^5 folds respect to the background for this set of substrates, with maximum accelerations obtained for guanine- or uridine-containing dimers GpU, GpG and UpU. In contrast with what observed for the transesterification of HPNP, in these cases the 1,2-vicinal compound **84** was in general slightly superior to the 1,3-distal calix[4]arene **83**, although for the CpA and GpA dimers the reverse is true.

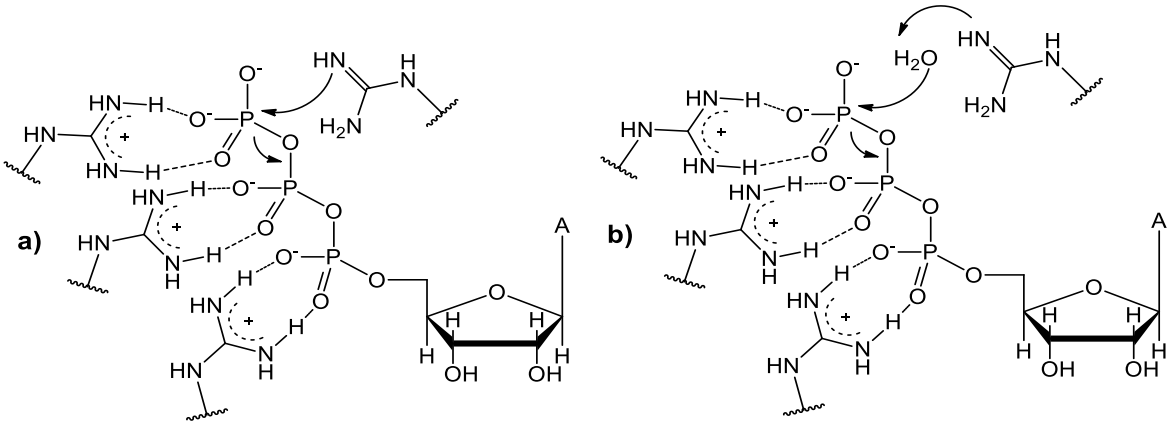


*Figure 33: Cooperation of the active units of calix[4]arene **86** in the cleavage of ATP. The deprotonated guanidine group could provide **a)** direct nucleophilic attack on the terminal phosphate group of the substrate or **b)** general base catalysis by deprotonating an external water molecule.*

In a subsequent study, the whole set of guanidine-calix[4]arenes was tested as cleaving agent for ATP[123]. Compound **86**, when present in solution in its triprotonated form (DMSO/H_2O 80/20, pH 9.8, 80°C), was able to hydrolyze this substrate with turnover, approximately 1000-folds faster than the background reaction. The cleavage process was supposed to occur thanks to the synergic action of the three guanidinium units and the deprotonated guanidine moiety to bind and activate the substrate and to act as a general base or as nucleophile, respectively (**Figure 33a** and **33b**). The contemporary occurrence of a triple electrophilic activation of the phosphate groups and of a nucleophilic attack or activation by a guanidine unit was proved to be crucial, by observing that the catalyst in its triprotonated form was much more effective than in its di- or tetraprotonated species and the tri- or di-functional calix[4]arenes **83**-**85**.

1.5 Conclusions and outline of the research

Phosphodeisterase enzymes are involved in the hydrolytic cleavage of phosphate diester bonds, which constitute the backbone of many biological componds. By exploiting a combination of different strategies this enzymes family is able to induce huge acceleration respect to the background for this type of reaction, resulting one of the most efficient in nature. These observations attracted, in last decades, the attention of supramolecular chemists, that have spent many efforts in the attempt of create arificial systems, called *enzyme models*, which can be able to mimic the action of their natural analogs.

Since the X-ray crystal structure of the active sites of some phosphodiesterase enzymes were reported, it is known that they exploit the synergic action of metal cations and amino acid residues to promote the cleavage of their substrate. Accordingly, corresponding artificial models have been obtained by

functionalizing a proper molecular scaffold with ligands for metal cations or organic functional groups able to mimic the presence of these types of active units.

Although artificial phosphodiesterases are still far to reach the catalytic activity and substrate selectivity of corresponding enzymes, a large variety of system able to cleave very effectively RNA, DNA and model compounds have reported and their efficiency and selectivity is continuously improving. A crucial phase of the building of effective enzyme models is the choice of an appropriate molecular scaffold to be used to preorganize the active units. In fact, in good artificial catalysts the molecular scaffolds have to provide a correct compromise between rigidity and residual flexibility of the system, in order to provide the so-called dynamic preorganization of the reactive groups, that is the ability to contemporary dispose them in an orientation suitable for catalysis and to allow their rearrangement to adapt to the change in conformation that the substrate undergoes during the overall catalytic process.

The *cone*-calix[4]arene cavity has been reported as a platform that can easily afford these particular requirements and, in fact, the synthesis of both metal-based and metal-free artificial phosphodiesterases has been successfully achieved by appropriate functionalization of the upper rim of this family of macrocycles. However, example of calix[4]arene-based phosphodiesterase models where the action of metal cations and organic functionalities is contemporary exploited was never reported before this PhD thesis, despite in many enzymes like Staphilococcal Nuclease and Alkaline Phosphatase this kind of strategy is observed. For this reason a part of my PhD was focused on the investigation of the properties of the hetero-bifunctional calix[4]arene derivatives **88** and **89** (**Figure 34a**)[124,125], whose synthesis and study of catalytic activity of corresponding Zn(II) and Cu(II) complexes will be presented in the next chapter of this thesis.

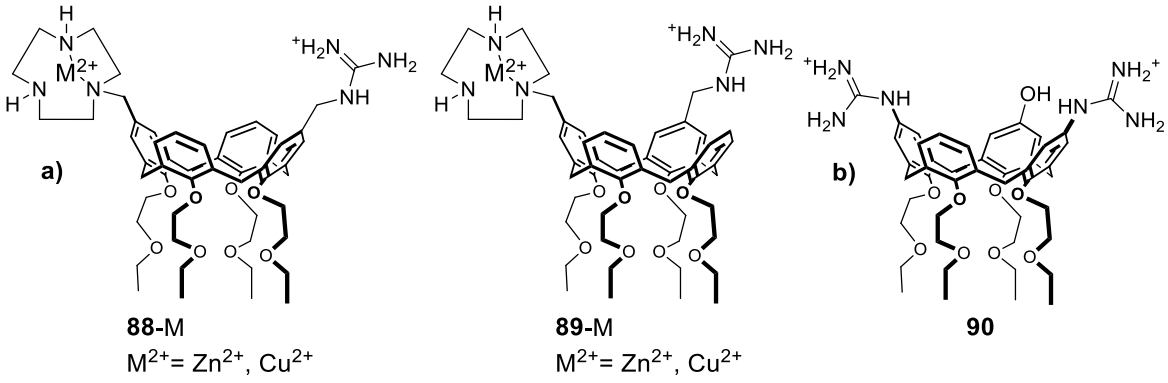


Figure 34: Calix[4]arene based artificial phosphodiesterases whose synthesis and catalytic activity will be described in the following chapters of this thesis.

As mentioned before, the first generation of metal-free artificial phosphodiesterase based on calix[4]arenes exploited the synergic action of guanidine and guanidinium units, to provide general base catalysis and electrophilic activation of the substrate, respectively. In **Chapter 3** of this thesis will be, instead, reported the mechanism of action of the trifunctional diguanidinium-hydroxyl-calix[4]arene **90**

(**Figure 34b**)[126], which was synthesized with the aim of studying the possibility to create a system where electrophilic guanidinium groups can effectively cooperate with nucleophilic functional groups in the cleavage of DNA model compounds, thus creating a mimic of human DNA Topoisomerase I.

1.6 Bibliography

(1) Berg, Jeremy B, Tymoczko, John L, Stryer, L. *Biochemistry, 5th Edition*, 5th ed.; W. H. Freeman: York, N., Ed.; **2002**.

(2) Fersht, A. *Enzyme Structure and Mechanism*, 2nd ed.; W. H. Freeman: New York, Ed.; **1984**.

(3) Dugas, H. *Bioorganic Chemistry*; York, S.-V. N., Ed.; **1988**.

(4) Breslow, R. *Acc. Chem. Res.* **1995**, *28* (3), 146.

(5) Wolfenden, R. *Chem. Rev.* **2006**, *106* (8), 3379.

(6) Rittié, L.; Perbal, B. *J. Cell Commun. Signal.* **2008**, *2* (1–2), 25.

(7) Gleave, M. E.; Monia, B. P. *Nat. Rev. Cancer.* **2005**, *5* (6), 468.

(8) Trawick, B. N.; Daniher, A. T.; Bashkin, J. K. *Chem. Rev.* **1998**, *98* (3), 939.

(9) Fischer, E. *Berichte der Dtsch. Chem. Gesellschaft* **1894**, *27*, 1985.

(10) Armstrong, E. F. *J. Soc. Chem. Ind.* **1930**, *49* (44), 919.

(11) Pauling, L. *Chem. Eng. News Arch.* **1946**, *24* (10), 1375.

(12) Wolfenden, R. *Nature* **1969**, *223* (5207), 704.

(13) Menger, F. M. *Acc. Chem. Res.* **1985**, *18* (5), 128.

(14) Kirby, A. J. *Adv. Phys. Org. Chem.* **1980**, *17*, 183.

(15) Mandolini, L. *Adv. Phys. Org. Chem.* **1986**, *22*, 1.

(16) Page, M. I.; Jencks, W. P. *Proc. Natl. Acad. Sci.* **1971**, *68* (8), 1678.

(17) Villa, J.; Strajbl, M.; Glennon, T. M.; Sham, Y. Y.; Chu, Z. T.; Warshel, A. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (22), 11899.

(18) Lightstone, F. C.; Bruice, T. C. *Bioorg. Chem.* **1998**, *26* (4), 193.

(19) Menger, F. M. *Acc. Chem. Res.* **1993**, *26* (4), 206.

(20) Khanjin, N. A.; Snyder, J. P.; Menger, F. M. *J. Am. Chem. Soc.* **1999**, *121* (50), 11831.

(21) Bruice, T. C.; Lightstone, F. C. *Acc. Chem. Res.* **1999**, *32* (2), 127.

(22) Bruice, T. C.; Benkovic, S. J. *Biochemistry* **2000**, *39* (21), 6267.

(23) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125* (6), 1472.

(24) Menger, F. M. *Biochemistry* **1992**, *31* (23), 5368.

(25) Štrajbl, M.; Shurki, A.; Kato, M.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125* (34), 10228.

(26) Villà, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105* (33), 7887.

(27) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303* (5655), 186.

(28) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106* (8), 3210.

(29) Warshel, a; Papazyan, A. *Proc. Natl. Acad. Sci.* **1996**, *93* (24), 13665.

(30) Olsson, M. H. M.; Warshel, A. *J. Am. Chem. Soc.* **2004**, *126* (46), 15167.

(31) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. *J. Phys. Chem.* **1996**, *100* (31), 12771.

(32) Lim, C.; Truhlar, D. G. *J. Phys. Chem.* **1985**, *89* (1), 5.

(33) Garrett, B. C.; Truhlar, D. G. *J. Phys. Chem.* **1979**, *83* (8), 1052.

(34) Kamerlin, S. C. L.; Sharma, P. K.; Prasad, R. B.; Warshel, A. *Q. Rev. Biophys.* **2013**, *46* (1), 1.

(35) Schroeder, G. K.; Lad, C.; Wyman, P.; Williams, N. H.; Wolfenden, R. *Proc. Natl. Acad. Sci.* **2006**, *103* (11), 4052.

(36) Perreault, D. M.; Anslyn, E. V. *Angew. Chem. Int. Ed. Engl.* **1997**, *36* (5), 433.

(37) Järvinen, P.; Oivanen, M.; Lönnberg, H. *J. Org. Chem.* **1991**, *56* (18), 5396.

(38) Barnes, J. A.; Wilkie, J.; Williams, I. H. *J. Chem. Soc.{,} Faraday Trans.* **1994**, *90* (12), 1709.

(39) Cleland, W. W.; Hengge, A. C. *Chem. Rev.* **2006**, *106* (8), 3252.

(40) Cleland, W. W. *FASEB J.* **1990**, *4* (11), 2899.

(41) Klähn, M.; Rosta, E.; Warshel, A. *J. Am. Chem. Soc.* **2006**, *128* (47), 15310.

(42) Williams, N. H.; Cheung, W.; Chin, J. *J. Am. Chem. Soc.* **1998**, *120* (32), 8079.

(43) Weston, J. *Chem. Rev.* **2005**, *105* (6), 2151.

(44) Wilcox, D. E. *Chem. Rev.* **1996**, *96* (7), 2435.

(45) Jedrzejas, M. J.; Setlow, P. *Chem. Rev.* **2001**, *101* (3), 607.

(46) Jeltsch, A.; Alves, J.; Maass, G.; Pingoud, A. *Febs Lett.* **1992**, *304* (1), 4.

(47) Karlin, K. D. *Science* **1993**, *261* (5122), 701 LP.

(48) Cowan, J. A. *JBIC J. Biol. Inorg. Chem.* **1997**, *2* (2), 168.

(49) Cox, R. S.; Schenk, G.; Mitić, N.; Gahan, L. R.; Hengge, A. C. *J. Am. Chem. Soc.* **2007**, *129* (31), 9550.

(50) Jeltsch, a; Alves, J.; Wolfes, H.; Maass, G.; Pingoud, a. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (18), 8499.

(51) Vipond, I. B.; Baldwin, G. S.; Halford, S. E. *Biochemistry* **1995**, *34* (2), 697.

(52) Groll, D. H.; Jeltsch, A.; Selent, U.; Pingoud, A. *Biochemistry* **1997**, *36* (38), 11389.

(53) Dupureur, C. M. *Metallomics* **2010**, *2* (9), 609.

(54) Kubiak, R. J.; Yue, X.; Hondal, R. J.; Mihai, C.; Tsai, M. D.; Bruzik, K. S. *Biochemistry* **2001**, *40* (18), 5422.

(55) Redinbo, M. R.; Stewart, L.; Kuhn, P.; Champoux, J. J.; Hol, W. G. J. *Science (80).* **1998**, *279* (March), 1504.

(56) Raines, R. T.; Peptide, B. P.; Isomerization, B. **1998**, *2665* (96).

(57) Holtz, K. M.; Kantrowitz, E. R. *FEBS Lett.* **1999**, *462* (1–2), 7.

(58) Cotton, F. A.; Hazen, Edward E., J.; Legg, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76* (6), 2551.

(59) Stewart, L. *Science* **1998**, *279* (5356), 1534.

(60) Kirby, A. J. *Angew. Chem. Int. Ed. Engl.* **1996**, *35* (7), 706.

(61) Mancin, F.; Scrimin, P.; Tecilla, P.; Tonellato, U. *Chem. Commun.* **2005**, 2540.

(62) Mancin, F.; Scrimin, P.; Tecilla, P. *Chem. Commun.* **2012**, *48* (45), 5545.

(63) Salvio, R. *Chem. Eur. J.* **2015**, *21* (31), 10960.

(64) Breslow, R. *Artificial Enzymes*; Wiley-VCH Verlag GmbH & Co. KGaA, Ed.; Weinheim, Germany, **2005**.

(65) Raynal, M.; Ballester, P.; Vidal-Ferran, A.; van Leeuwen, P. W. N. M. *Chem. Soc. Rev.* **2014**, *43* (5), 1660.

(66) Raynal, M.; Ballester, P.; Vidal-Ferran, A.; van Leeuwen, P. W. N. M. *Chem. Soc. Rev.* **2014**, *43* (5), 1734.

(67) Sanders, J. K. M. *Chem. Eur. J.* **1998**, *4* (8), 1378.

(68) Molenveld, P.; Engbersen, J. F. J.; Reinhoudt, D. N. *Chem. Soc. Rev.* **2000**, *29* (2), 75.

(69) Oivanen, M.; Kuusela, S.; Lonnberg, H.; Lönnberg, H. *Chem. Rev.* **1998**, *98* (3), 961.

(70) Menger, F. M.; Ladika, M. *J. Am. Chem. Soc.* **1987**, *109* (10), 3145.

(71) Franklin, S. J. *Curr. Opin. Chem. Biol.* **2001**, *5* (2), 201.

(72) Chand, D. K.; Bharadwaj, P. K.; Schneider, H. J. *Tetrahedron* **2001**, *57* (31), 6727.

(73) Berg, T.; Simeonov, a; Janda, K. D. *J. Comb. Chem.* **1999**, *1* (1), 96.

(74) Tomoyuki, I.; Jun, S.; Makoto, K. *Chem. Lett.* **2000**, *29* (4), 356.

(75) Jun, S.; Tomoyuki, I.; Kenichiro, F.; Makoto, K. *Chem. Lett.* **2000**, *29* (1), 56.

(76) Bonfá, L.; Gatos, M.; Mancin, F.; Tecilla, P.; Tonellato, U. *Inorg. Chem.* **2003**, *42* (12), 3943.

(77) Dixon, N. E.; Geue, R. J.; Lambert, J. N.; Moghaddas, S.; Pearce, D. A.; Sargeson, A. M. *Chem. Commun.* **1996**, No. 11, 1287.

(78) Hegg, E. L.; Burstyn, J. N. *Inorg. Chem.* **1996**, *35* (26), 7474.

(79) Itoh, T.; Hisada, H.; Sumiya, T.; Hosono, M.; Usui, Y.; Fujii, Y. *Chem. Commun.* **1997**, *1* (7), 677.

(80) Sreedhara, A.; Cowan, J. a. *Chem. Commun.* **1998**, No. 16, 1737.

(81) Sreedhara, A.; Freed, J. D.; Cowan, J. A. *J. Am. Chem. Soc.* **2000**, *122* (37), 8814.

(82) Kawahara, S.; Uchimaru, T. *Eur. J. Inorg. Chem.* **2001**, *2001* (9), 2437.

(83) Young, M. J.; Chin, J. *J. Am. Chem. Soc.* **1995**, *117* (42), 10577.

(84) Liu, S.; Hamilton, A. D. *Bioorganic Med. Chem. Lett.* **1997**, *7* (13), 1779.

(85) Iranzo, O.; Kovalevsky, A. Y.; Morrow, J. R.; Richard, J. P. *J. Am. Chem. Soc.* **2003**, *125* (7), 1988.

(86) McCue, K. P.; Morrow, J. R. *Inorg. Chem.* **1999**, *38* (26), 6136.

(87) Wall, M.; Hynes, R. C.; Chin, J. *Angew. Chem. Int. Ed. Engl.* **1993**, *32* (11), 1633.

(88) Schnaith, L. M.; Hanson, R. S.; Que, L. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (2), 569.

(89) Liu, C.; Yu, S.; Li, D.; Liao, Z.; Sun, X.; Xu, H. *Inorg. Chem.* **2002**, *41* (4), 913.

(90) Anbu, S.; Kandaswamy, M.; Kamalraj, S.; Muthumarry, J.; Varghese, B. *Dalt. Trans.* **2011**, *40* (28), 7310.

(91) Roy, M.; Dhar, S.; Maity, B.; Chakravarty, A. R. *Inorganica Chim. Acta* **2011**, *375* (1), 173.

(92) Qian, J.; Wang, L.; Gu, W.; Liu, X.; Tian, J.; Yan, S. *Dalton Trans.* **2011**, *40* (20), 5617.

(93) Mancin, F.; Rampazzo, E.; Tecilla, P.; Tonellato, U. *Eur. J. Org. Chem.* **2004**, 281.

(94) Yashiro, M.; Ishikubo, A.; Komiyama, M. **1997**, No. 3, 83.

(95) Komiyama, M.; Kina, S.; Matsumura, K.; Sumaoka, J.; Tobey, S.; Lynch, V. M.; Anslyn, E. *J. Am. Chem. Soc.* **2002**, *124* (46), 13731.

(96) Kimura, E.; Kodama, Y.; Koike, T.; Shiro, M. *J. Am. Chem. Soc.* **1995**, *117* (32), 8304.

(97) Aït-Haddou, H.; Sumaoka, J.; Wiskur, S. L.; Folmer-Andersen, J. F.; Anslyn, E. V. *Angew. Chem. Int. Ed. Engl.* **2002**, *41* (21), 4013.

(98) He, J.; Sun, J.; Mao, Z. W.; Ji, L. N.; Sun, H. *J. Inorg. Biochem.* **2009**, *103* (5), 851.

(99) An, Y.; Tong, M.-L.; Ji, L.-N.; Mao, Z.-W. *Dalt. Trans.* **2006**, No. 17, 2066.

(100) Tjioe, L.; Joshi, T.; Brugger, J.; Graham, B.; Spiccia, L. *Inorg. Chem.* **2011**, *50* (2), 621.

(101) Baldini, L.; Sansone, F.; Faimani, G.; Massera, C.; Casnati, A.; Ungaro, R. *European J. Org. Chem.* **2008**, No. 5, 869.

(102) Salvio, R.; Cacciapaglia, R.; Mandolini, L. *J. Org. Chem.* **2011**, *76* (13), 5438.

(103) Breslow, R.; Labelle, M. *J. Am. Chem. Soc.* **1986**, *108* (10), 2655.

(104) Anslyn, E.; Breslow, R. *J. Am. Chem. Soc.* **1989**, *111* (12), 4473.

(105) Komiyama, M.; Yoshinari, K. *J. Org. Chem.* **1997**, *62* (7), 2155.

(106) Corona-Martinez, D. O.; Taran, O.; Yatsimirsky, A. K.; Corona-Martínez, D. O.; Taran, O.; Yatsimirsky, A. K. *Org. Biomol. Chem.* **2010**, *8* (4), 873.

(107) Salvio, R.; Mandolini, L.; Savelli, C. *J. Org. Chem.* **2013**, *78* (14), 7259.

(108) Göbel, M. W.; Bats, J. W.; Dürner, G. *Angew. Chem. Int. Ed. Engl.* **1992**, *31* (2), 207.

(109) Muche, M.-S.; Göbel, M. W. *Angew. Chem. Int. Ed. Engl.* **1996**, *35* (18), 2126.

(110) Ullrich, S.; Nazir, Z.; Büsing, A.; Scheffer, U.; Wirth, D.; Bats, J. W.; Dürner, G.; Göbel, M. W. *ChemBioChem* **2011**, *12* (8), 1223.

(111) Sheng, X.; Lu, X. M.; Zhang, J. J.; Chen, Y. T.; Lu, G. Y.; Shao, Y.; Liu, F.; Xu, Q. *J. Org. Chem.* **2007**, *72* (5), 1799.

(112) Rebilly, J.-N.; Reinaud, O. *Supramol. Chem.* **2014**, *26* (August), 454.

(113) Cacciapaglia, R.; Di Stefano, S.; Mandolini, L.; Salvio, R. *Supramol. Chem.* **2013**, *25* (9–11), 537.

(114) Giuliani, M.; Morbioli, I.; Sansone, F.; Casnati, A. *Chem. Commun.* **2015**, *51* (75), 14140.

(115) Molenveld, P.; Stikvoort, W. M. G.; Kooijman, H.; Spek, A. L.; Engbersen, J. F. J.; Reinhoudt, D. N. *J. Org. Chem.* **1999**, *64* (11), 3896.

(116) Molenveld, P.; Kapsabelis, S.; Engbersen, J. F. J.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **1997**, *119* (12), 2948.

(117) Molenveld, P.; Engbersen, J. F. J.; Kooijman, H.; Spek, A. L.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **1998**, *120* (27), 6726.

(118) Molenveld, P.; Engbersen, J. F. J.; Reinhoudt, D. N. *Angew. Chem. Int. Ed. Engl.* **1999**, *38* (21), 3189.

(119) Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Reinhoudt, D. N.; Salvio, R.; Sartori, A.; Ungaro, R. *J. Am. Chem. Soc*. **2006**, 128, 12322.

(120) Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Peracchi, A.; Reinhoudt, D. N.; Salvio, R.; Sartori, A.; Ungaro, R. *J. Am. Chem. Soc.* **2007**, *129* (41), 12512.

(121) Baldini, L.; Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Salvio, R.; Sansone, F.; Ungaro, R. *J. Org. Chem.* **2012**, *77* (7), 3381.

(122) Salvio, R.; Cacciapaglia, R.; Mandolini, L.; Sansone, F.; Casnati, A. *RSC Adv.* **2014**, *4* (65), 34412.

(123) Salvio, R.; Casnati, A.; Mandolini, L.; Sansone, F.; Ungaro, R. *Org. Biomol. Chem.* **2012**, *10* (45), 8941.

(124) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Sansone, F. *J. Org. Chem.* **2015**, *80* (11), 5887.

(125) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Sansone, F.; Mandolini, L.; Casnati, A. *J. Org. Chem.* **2016**, *81* (11), 4728.

(126) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Sansone, F.; Mandolini, L.; Casnati, A. *J. Org. Chem.* **2016**, *81* (19), 9012.

Chapter 2

Guanidium-triazacyclononane calix[4]arenes as catalysts for the cleavage of DNA and RNA model compounds

2.1 Introduction

Phosphodiesterase enzymes are involved in the cleavage of phosphodiester bonds[1–3], which keep together the backbone of important bio(macro)molecules, such as RNA and DNA. Because of the relevance of their biological activity, the production of artificial catalysts able to mimic the action of this family of enzymes has attract much attention over the years, with the final goal to obtain systems that can have applications in biotechnologies and medicine[4,5]. The catalytic activity of these systems, also called enzyme models, is provided by the *active units* which have the task to activate the substrate through electrophilic activation, to afford a nucleophile or a proper general acid-base function and to stabilize the transition state[6–8]. These functions are quite often fulfilled by metal-ion-containing moieties[9–12] and/or functional groups which can be easily protonated/deprotonated in the pH range of operation[12–16]. In a biomimetic approach, quite often these functional groups are those present in the side chains of natural amino acids, since these are known to be exploited by natural enzymes to cleave their substrates (see **Chapter 1, Paragraph 1.3.1**)[11,17–21].

A crucial role to determine the activity of enzyme models is played also by the molecular scaffold on which the active units are placed. The molecular scaffold has the essential task to keep the active sites at a proper distance and with an orientation suitable to afford their cooperation in catalysis. The scaffold should possibly afford a proper *dynamic preorganization*[22] of the active sites, allowing their rearrangement to fit the conformational changes of the substrate along the overall catalytic process[23]. The *cone*-calix[4]arene platform has been reported to be particularly effective into achieve this purpose, thanks to a special compromise between rigidity and residual flexibility of its structure. For example, effective artificial phosphodiesterases able to cleave ATP, DNA or RNA model compounds and oligoribonucleosides have been obtained by decorating the upper rim of this molecular scaffold with guanidinium groups or ligands for metal cations[22,24,25].

Interestingly, the possibility to place these two different types of active units on the same *cone*-calix[4]arene scaffold, in order to exploit their cooperation for the cleavage of phosphodiester bonds, had not been evaluated yet, despite this “mixed active sites” strategy is known to be used by both natural enzymes[11,17,18] and artificial phosphodiesterase models built on simpler platforms[12–14]. Thus, in collaboration with the group of Riccardo Salvio and Luigi Mandolini from the university “La Sapienza” in Rome, we decided to synthetize two new *cone*-calix[4]arene derivatives, bearing at the upper rim a guanidinium group and a triazacyclononane moiety as ligating unit for Zn(II) or Cu(II) ions and to investigate their catalytic activity for the cleavage of different RNA and DNA model compounds. These two different regioisomers, differ for the relative disposition of the active units, which are held in 1,3 distal- or 1,2 vicinal- positions, since in the past it was observed that the distance and relative

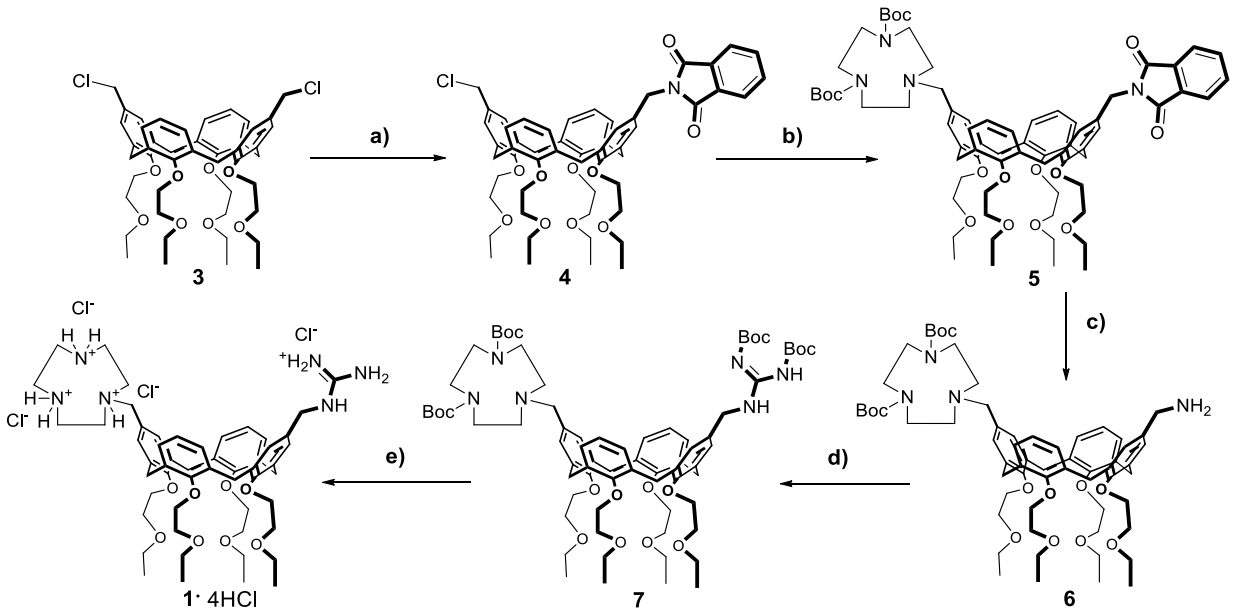
disposition of the active units on the macrocyclic scaffold might result in artificial enzymes with quite a different activity[26–29].

The synthesis, the potentiometric characterization and the investigations on the catalytic activity of these compounds will be presented in the following pages of this chapter. These studies are already published in the literature[30,31].

2.2 Results and discussion

2.2.1 Calix[4]arenes synthesis

Compounds **1** and **2** were synthetized starting from the corresponding 1,3 distal- and 1,2 vicinal dichloromethyl precursors **3**[32] and **8**[33]. In both the cases, the crucial step of the synthetic pathway was the first one, where the desymmetrization of the functional groups present at the upper rim of the studied calix[4]arenes took place. Focusing on the synthesis of **1** (**Scheme 1**), this step was initially carried out by reacting **3** with a stoichiometric amount of potassium phthalimide and dicyclohexyl-18-crown-6 in refluxing toluene, to form the corresponding monochloromethyl-monophthalimidomethyl-derivative **4** as main product (38%) in a mixture containing also the diphthalimido-derivative (16%) and the unreacted starting material (23%).



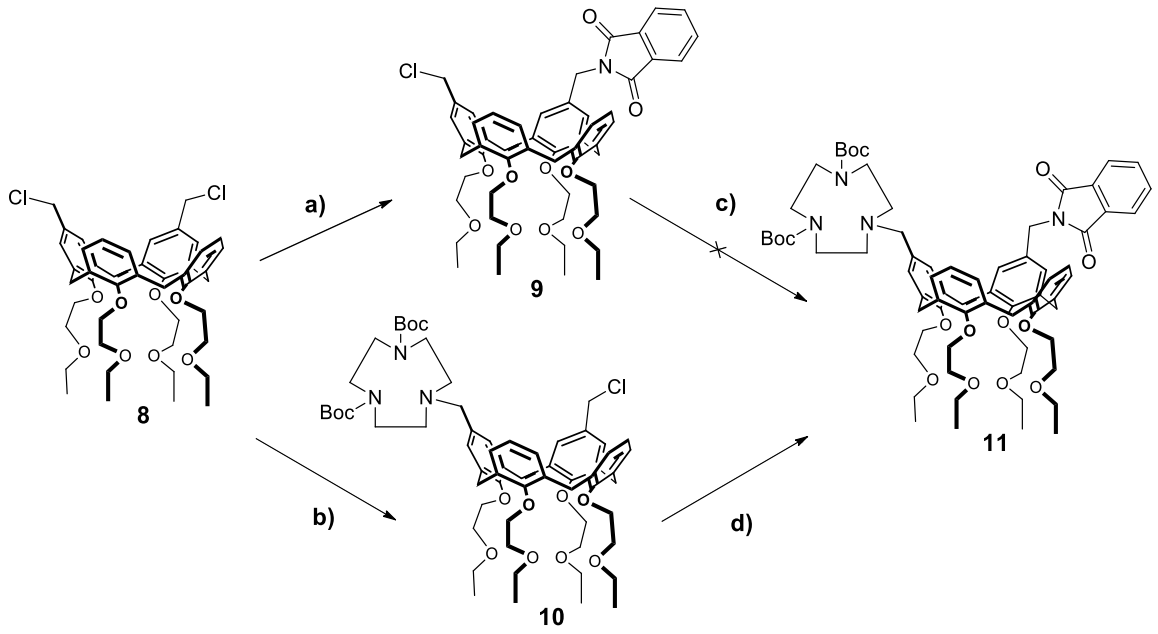
Scheme 1: a) Potassium phthalimide; dry DMF, rt. b) N,N'-Bis-(Boc)-1,4,7-triazacyclononane, K_2CO_3; CH_3CN, rt. c) $N_2H_4 \cdot H_2O$; MeOH, 60 °C. d) N,N'-Bis-(Boc)-N''-triflylguanidine, NEt_3; DCM, rt. e) CF_3COOH (TFA), Et_3SiH (TES); DCM, rt and then 1M aqueous HCl; EtOH, rt.

An improvement of the reaction conditions was then achieved by following a standard protocol for the Gabriel synthesis, that is by performing the same reaction in dry DMF without the use of the crown-

ether, allowing the isolation of the desired compound in larger amount than the expected statistical one (57% yield).

The following steps were then performed easily. A *N,N'*-bis-(Boc)-1,4,7-triazacyclononane (bis-Boc-TACN) group was first inserted on the remaining chloromethyl moiety of **4** to obtain **5** (70% yield), whose phthalimidomethyl group was quantitatively converted in the corresponding aminomethyl unit to yield compound **6.** Subsequently, the insertion of the guanidine group on this latter calix[4]arene was performed by reaction with [*N*,*N*'–bis-(Boc)]triflylguanidine[34] to give **7** (63% yield), from which **1** was obtained as tetrahydrochloride salt by removal of the Boc protecting groups with trifluoroacetic acid (TFA), followed by exchange of the resulting trifluoroacetate counterion with chloride (98% yield).

To obtain compound **2** it was instead necessary to modify the first steps of the just described synthetic strategy (**Scheme 2**).

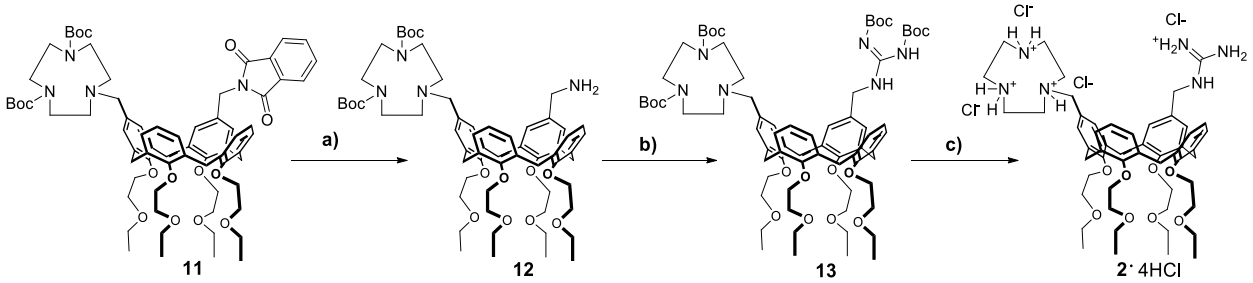


Scheme 2: a) Potassium phthalimide; dry DMF, r.t. b) N,N'-Bis-(Boc)-1,4,7-triazacyclononane, K_2CO_3; CH_3CN, rt. c) N,N'-Bis-(Boc)-1,4,7-triazacyclononane, KI, strong bases; CH_3CN, high temperature. d) Potassium phthalimide, dicyclohexyl-18-crown-6; dry toluene, 110°C.

In fact, on the contrary of what previously reported, after the treatment of the bis-chloromethyl derivative **8** with 1 mol equivalent of potassium phthalimide, which gave **9** in 65% yield, the reaction with bis-Boc-TACN did not afford compound **11** even using strong bases, high temperatures, iodide as catalyst and long reaction times. To explain these results it was suggested that, moving from the 1,3-distal calix[4]arene **3** to the corresponding 1,2-vicinal isomer **8**, the reactivity of the two chloromethyl groups became influenced by steric effects which, if on the one hand seemed to promote a moderate regioselectivity in the synthesis of the corresponding monophthalimido derivative (65% for **9** *vs.* 57% for **4**), on the other hand completely prevented the subsequent insertion of the TACN unit. The low

reactivity of the $–CH_2Cl$ moieties of 1,2-vicinal calix[4]arene derivatives toward the attack of highly hindered nucleophiles was further confirmed once these two reaction steps were reversed, since the introduction of a bis-Boc-TACN group on **8** to give **10** took place in 56% yield, without observing the formation of the disubstituted compound and allowing the recovery of 27% unreacted starting material. These steric effects did not affect negatively the following conversion of **10** into compound **11** by reaction with potassium phthalimide, which was achieved in 90% yield.

The reaming synthetic steps to obtain the target calix[4]arene derivative where then carried out according to the same strategy used for compound **1** (**Scheme 3**).



Scheme 3: a) $N_2H_4·H_2O$; MeOH, 60 °C. b) N,N'-Bis-(Boc)-N''-triflylguanidine, NEt_3; DCM, rt. c) CF_3COOH (TFA), Et_3SiH (TES); DCM, rt and then 1M aqueous HCl; EtOH, rt.

Treatment with hydrated hydrazine of **11** afforded the corresponding aminomethyl derivative **12** (92% yield), which was then guanidinylated by reaction with *N,N'*-bis(Boc)-*N''*-triflylguanidine[34] and converted in the tetrahydrochloride form of compound **2** by removal of the Boc groups and exchange of the triflouroacetate anions with chloride (overall 46% yield).

^{1}H- and ^{13}C NMR spectra of all the intermediates bearing at their upper rim the bis-Boc protected TACN unit (i. e. compounds **5**-**7** and **10**-**13**) showed patterns of signals compatible with the presence in solution of different rotamers, in slow exchange with respect to the NMR timescale. This phenomenon was ascribed to the steric bulkiness afforded by the protected azacrown moiety and it was reported in the literature also for other bis-Boc-TACN functionalized derivatives[35]. After removal of the Boc protecting groups to obtain **1**·4HCl and **2**·4HCl the signals ascribed to these rotamers disappeared and the spectra of the target compounds turned to be compatible with the presence of single chemical species (**Figure 1**).

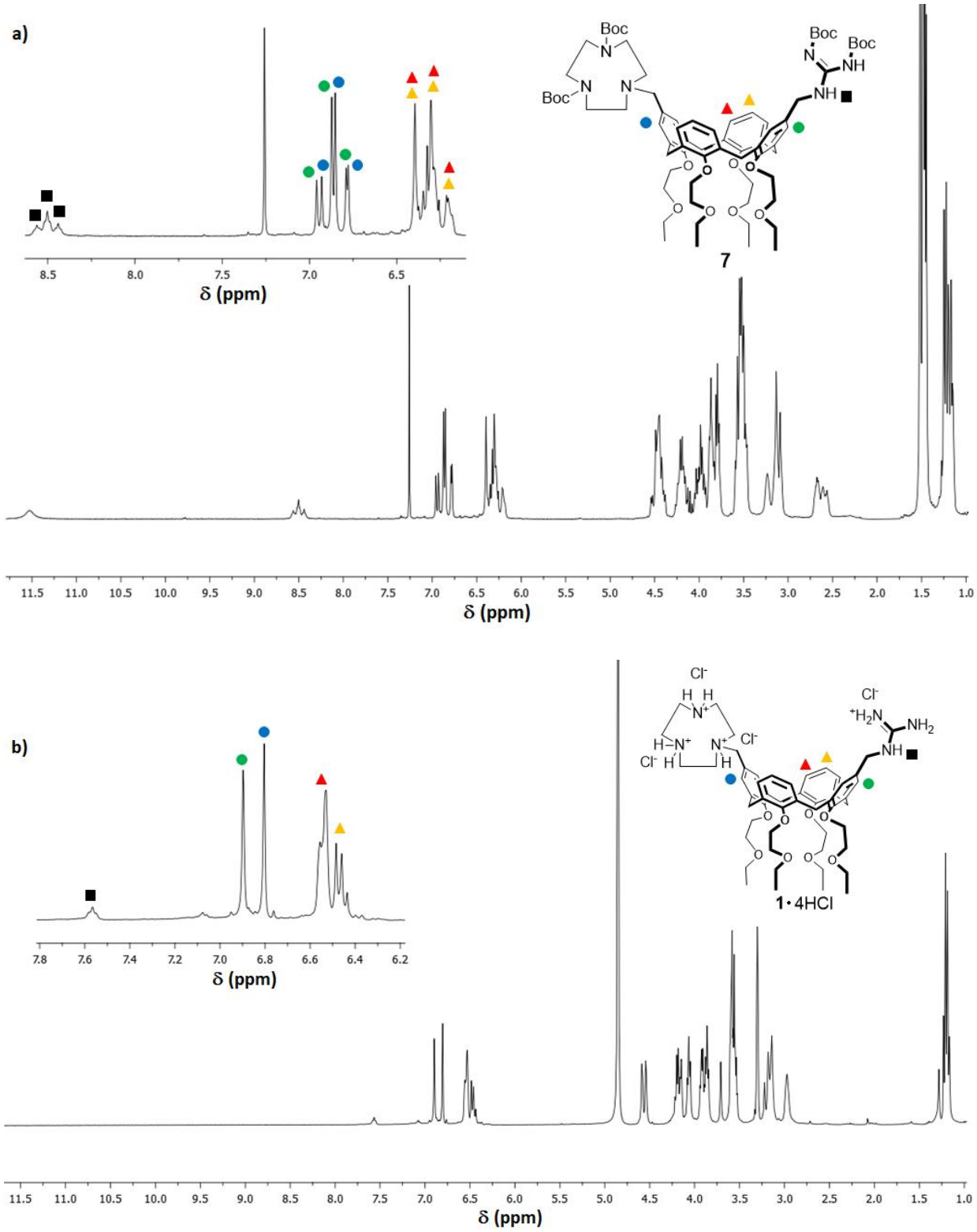


*Figure 1: ^{1}H NMR spectra of **a)** compound **7** (400 MHz, $CDCl_3$) and **b)** compound **1**·4HCl (300 MHz, CD_3OD). The pattern of signals in the upper spectrum was compatible with the presence of different rotamers in slow exchange respect to the NMR timescale. The spectral region comprising the aromatic and the $–CH_2$**NH** guanidine protons was magnified since the phenomenon was particularly evident. After removal of the Boc protecting group from the TACN moiety the lower spectrum turned to be consistent with the presence of a sole chemical species.*

2.2.2 Potentiometric titrations

Potentiometric titration of calix[4]arenes **1**·4HCl and **2**·4HCl were performed to determine their acidity constants and their affinity for the metal cations with whom are expected to form catalytically active complexes. The experiments were carried out in DMSO/H_2O 8/2 (thereafter indicated as simply 80% DMSO), a mixture of solvents where our compounds are well soluble and not associated and that has been demonstrated to be suitable for kinetic studies of hydrolytic reactions[12,36] and for the determination of the pK_a values of different compounds[37,38]. Moreover, the use of this medium is particularly favorable for investigations about the mechanism of action exploited by guanidinium-based catalysts for the cleavage of phosphate esters, since, respect to the use of pure water[36,39], it advantages the establishment of two-point hydrogen bonding interactions to bind and/or activate the substrate[12]. However, for the treatment of the experimental data it is important to remember that in 80% DMSO the autoprotolysis of water is strongly suppressed (pK_w= 18.4)[40], with pH values necessary to reach neutrality equal to 9.2.

The potentiometric titrations, shown in **Figure 2**, were carried out on **1**·4HCl and **2**·4HCl alone and in the presence of a stoichiometric amount of $ZnCl_2$ or $CuCl_2$.

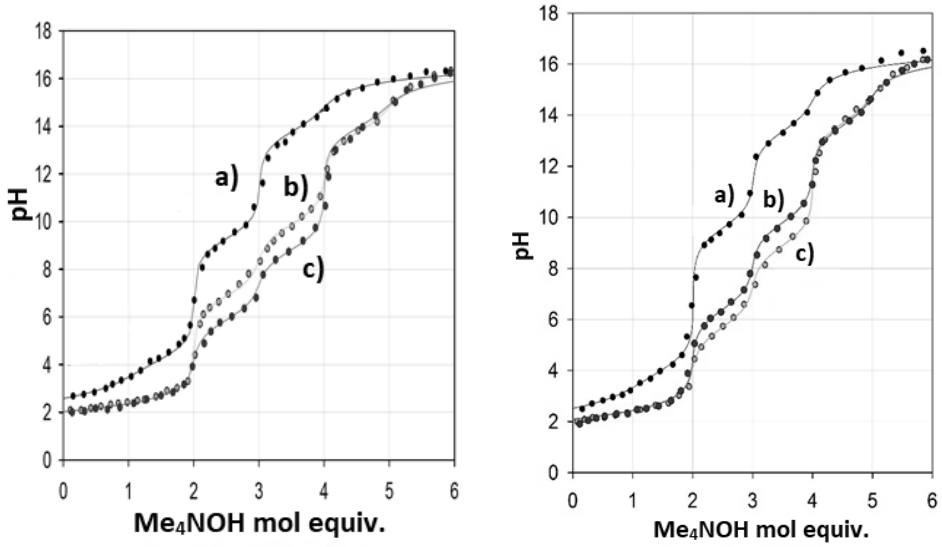


*Figure 2: Titrations curves obtained by addition of increasing aliquots of a solution of Me_4NOH to solutions containing **1**·4HCl and **2**·4HCl (80% DMSO, 25°C, left and right, respectively) in the presence of **a)** no additives, **b)** 1mol equiv. $ZnCl_2$ and **c)** 1mol equiv. $CuCl_2$.*

As expected, in the absence of additives, for both compounds it was detected the presence of four titratable protons, of which the three most acidic were those present on the nitrogen atoms of their TACN groups and the least acidic was the one on their guanidinium moieties. The addition of 1 mol equivalent of Zn(II) or Cu(II) salts strongly modified the shape of the resulting titration curves since it led to a strong increase of the acidity for the protons belonging to the triazacyclonane unit, as a result of the tight binding of the metal cations. In fact, the fitting of the experimental data yielded, respectively,

logK_{Zn}= 7.6 and 6.9 as estimation of the equilibrium constants for the formation of the Zn(II) complexes of the monoprotonated catalysts **1**H$^+$ and **2**H$^+$, while for the coordination of the Cu(II) cations the calculated logK_{Cu} were equal to 9.2 and 8.8, respectively. The pK_a for the deprotonation of the guanidinium group, instead, did not undergo to significant variation[41], thus excluding a participation to the coordination of the metal cations, which could have been detrimental for the catalytic activity of the studied complexes. Moreover, in the presence of both $ZnCl_2$ or $CuCl_2$ it was also observed the appearance of a fifth titratable proton, which was due to the coordination of a water molecule to the metal centers. The pK_a values found for the additional Zn(II)- and Cu(II)-bound water molecules were of 9.9 and 8.8, respectively, in the case of compound **1** and 9.8 and 8.9 in the presence of compound **2**, thus indicting that their deprotonation took place in neutral or slightly basic conditions.

In general, as it can be seen in **Table 1**, the differences in acidity for the same types of titratable protons were negligible for the two regioisomeric calix[4]arenes, pointing out that the acid-base properties of the two compounds were not significantly different. A similar observation can be done for the stability constants of the complexes formed in the presence of Zn(II) or Cu(II), since, as reported just above, the association constant calculated for the coordination of these metal ions by the 1,3-distal derivative **1**H$^+$ were essentially identical to those obtained for the 1,2-vicinal compound **2**H$^+$.

*Table 1: Acidity Constants (pKa) of **1**·4HCl and **2**·4HCl, in the absence and presence of 1 molar equiv of $ZnCl_2$ or $CuCl_2$, (80% DMSO, 25 °C). The pK_a data were obtained from potentiometric titrations with Me_4NOH of 1.0 mM substrate solutions in the presence of 10 mM Me_4NClO_4 with experimental uncertainty = ± 0.1 pK units, unless otherwise stated.[a] Error limit = ±0.3 pK units.[b] Under the same conditions, the pK_a of guanidine·HCl is 13.7 (see **Ref. 41**). [c] From the potentiometric titration: log K_{Zn} = 7.6 ±0.4 and 6.9 ± 0.5, for the binding of Zn(II) to the triazacyclononane moiety in **1**H$^+$ and **2**H$^+$, respectively. [d] From the potentiometric titration: log K_{Cu} = 9.2 ± 0.4 and 8.8 ±0.4, for the binding of Cu(II) to the triazacyclononane moiety in **1**H$^+$ and **2**H$^+$, respectively.*

Tested compound	Additive	pK_{a1}	pK_{a2}	pK_{a3}	pK_{a4}	pK_{a5}
1·4HCl	none	2.2[a]	4.0	9.6	13.4[b]	-
	$ZnCl_2$[c]	< 2	< 2	6.4	9.8	13.6
	$CuCl_2$[d]	< 2	< 2	5.7	8.9	13.7
2·4HCl	none	2.4	4.2	9.3	13.8[b]	-
	$ZnCl_2$[c]	< 2	< 2	6.9	9.9	13.7
	$CuCl_2$[d]	< 2	< 2	5.9	8.8	13.9

2.2.3 Cleavage of HPNP

Both the Zn(II) and the Cu(II) complexes of compounds **1**H$^+$ and **2**H$^+$ were tested as catalysts for the cleavage of different phosphate diesters. The first reaction to be investigated was the transesterification of the RNA model compound 2-hydroxypropyl-*p*-nitrophenyl phosphate (HPNP) in 80% DMSO (**Figure 3a**), at pH values equal to the pK_a of the water molecule bound to the metal center of the catalysts. This choice guaranteed the tight complexation of the metal cations, together with the deprotonation of 50% of the water molecules coordinated to them and the presence the guanidinium unit in its fully protonated form.

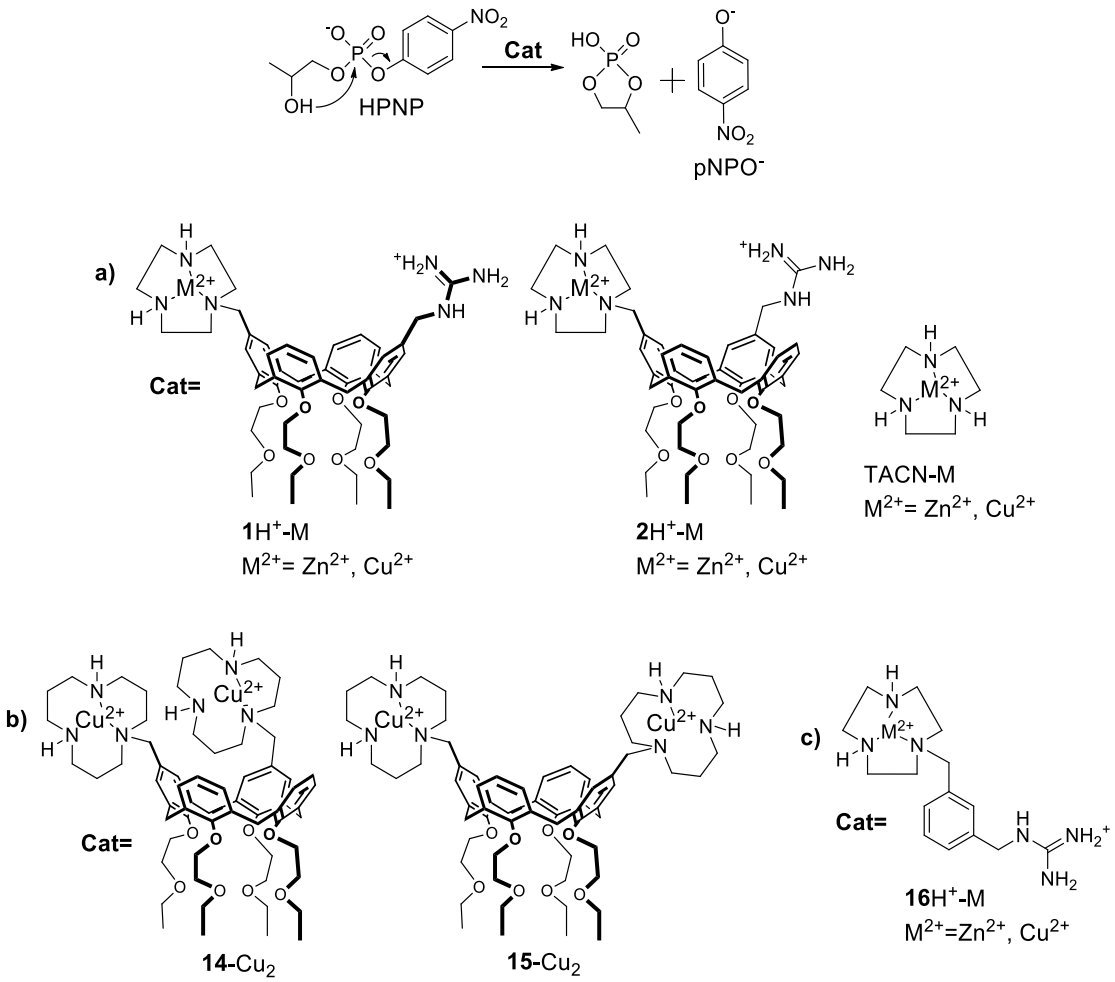


*Figure 3: Different metal complexes used to catalyze the cleavage of the RNA model compound HPNP in **a)** this work of thesis, **b) Ref. 26** and **c) Ref. 41**.*

The pseudo-first-order rate constants for the studied reaction were calculated as $k_{obs}= v_o/[HPNP]$, where v_o was the spectrophotometrically determined initial rate of *p*-nitrophenol (pNPOH) release from the substrate (**Figure 4a**). It is worth to mention that in these experiments the quantity which is directly measured is the absorbance at λ=400 nm due to the liberation of the *p*-nitrophenol/*p*-nitrophenolate dyad in the reaction mixture, since the departing *p*-nitrophenol is released as a mixture of its protonated and deprotonated forms. Calibration curves were built at each tested pH values by measuring the

absorbance of known amounts of *p*-nitrophenol at the same wavelength and in the same reaction medium (**Figure 4b**), in order to extrapolate the total concentration of pNPOH (sum of the concentration of the protonated and the unprotonated forms) released during the cleavage processes, without knowing the molar ratio of each single protonation species.

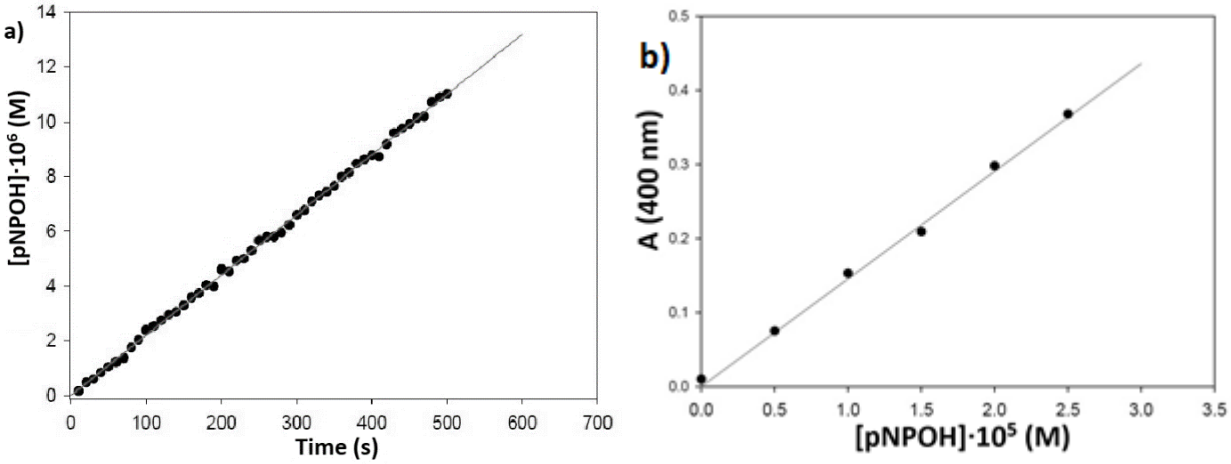


*Figure 4: **a)** Spectrophotometric determination of the initial rate of pNPOH release from a 0.2 mM solution of HPNP in the presence of 0.5 mM **2**H⁺-Zn (80% DMSO, 25°C, pH 9.9). **b)** Calibration curve built at the same pH value to extrapolate the total concentration of pNPOH released during the kinetic investigation. The curve was obtained by plotting the absorbance at 400 nm of solutions containing known aliquots of p-nitrophenol.*

The results of these kinetic investigations were listed in **Table 2**, together with those obtained for the same experiments carried out in the presence of the TACN-Zn or TACN-Cu complexes (TACN-M in **Figure 3a**), used as monofunctional control compounds.

*Table 2: Transesterification of the RNA model HPNP in the presence of metal complexes **1**H⁺-M, **2**H⁺-M, and TACN–M, where M= Zn(II) or Cu(II). The experiments were performed in 0.10 M N,N'-diisopropyl ethanolamine buffer (80% DMSO, 25.0 °C) at 0.20 mM HPNP and 0.50 mM catalyst concentrations.* [a] k_{obs} *was calculated as* v_o*/[HPNP], with error limits on the order of ±10%.* [b] k_{rel} *was calculated as* k_{obs}*(bifunctional catalyst)/*k_{obs}*(TACN–M).* [c] *The background rate constant* (k_{bg}, s^{-1}) *for the hydroxide-catalyzed cleavage is a function of pH and it is given by the following expression*[28]*:* $k_{bg} = 10^{(pH-17.2)}$.

Entry	pH	Catalyst	$10^6 \times k_{obs}$ (s^{-1}) [a]	k_{rel} [b]	k_{obs} / k_{bg} [c]
1	9.9	TACN–Zn	20	1.0	400
2	9.8	**1**H⁺–Zn	110	5.5	2200
3	9.9	**2**H⁺–Zn	73	3.6	1800
4	8.8	TACN–Cu	1.4	1.0	350
5	8.9	**1**H⁺–Cu	1300	930	320000
6	8.8	**2**H⁺–Cu	1800	1600	360000

The best catalytic performances were observed for compounds **1**H^+-Cu and **2**H^+-Cu, which were around two order of magnitude more effective than the corresponding Zn(II) complexes, accelerating the studied transesterification process of $3.2 \cdot 10^5$- and $3.6 \cdot 10^5$-folds respect to the uncatalyzed reaction, respectively. In marked contrast with what found for other calix[4]arene derivatives, the catalytic activity of the tested metal complexes was not significantly affected by the different position of the active units at their upper rim. In fact, after coordination of the same metal ion, the two tested regioisomers gave to the cleavage of the substrate very similar acceleration respect to the background, at contrary of what reported in literature (100% H_2O, 20 mM HEPES buffer, pH 7.0) for the bimetallic Cu(II) complex of the 1,2-vicinal compound **14**, which cleaved HPNP 30-folds faster than the parent 1,3-distal complex **15**-Cu_2 (**Figure 3b**)[26]. The highest activity of catalysts **1**H^+-M and **2**H^+-M respect to their monofunctional control compounds TACN–Zn and TACN–Cu, together with the previous report that guanidinium chloride is not able to enhance the cleavage rate of HPNP in the pH range 9.0–9.8[41], pointed out the existence of cooperation between the TACN-M and the guanidinium catalytic units when supported on the calix[4]arene scaffold. The degree of cooperation was modest for **1**H^+-Zn and **2**H^+-Zn, while was relevant for the corresponding Cu(II) complexes, whose catalytic performances exceed that of the monofunctional control TACN–Cu by a factor of 10^3. These results are in agreement with those reported for the bifunctional catalyst **16**H^+ (**Figure 3c**)[41], bearing the same reactive groups on a simple *m*-xylylene spacer, whose Cu (II) complex **16**H^+-Cu was much more effective than the parent Zn(II)-based analogous **16**H^+-Zn into catalyze the cleavage of HPNP. Moreover, in the same experimental conditions, **16**H^+-Cu promoted this process 25- and 30-folds less effectively than the corresponding supramolecular catalysts **1**H^+-Cu and **2**H^+-Cu, respectively, thus indicating once again the beneficial role played by the calix[4]arene scaffold into increase the degree of synergism between the active units in artificial phosphodiesterases.

Further investigations, performed by using compound **1**H^+-Cu, evidenced a strictly linear dependence of the initial rate of *p*-nitrophenol release respect to the substrate concentration (**Figure 5a**), pointing out that the effect of the catalyst-substrate association on the kinetics of the reaction was negligible in the tested concentration range. This observation means that the energy coming from the interaction between their guanidinium and phosphate groups was not significantly used by the two reacting partners to bind to each other, resulting instead fully translated into catalysis by stabilization of the transition state.

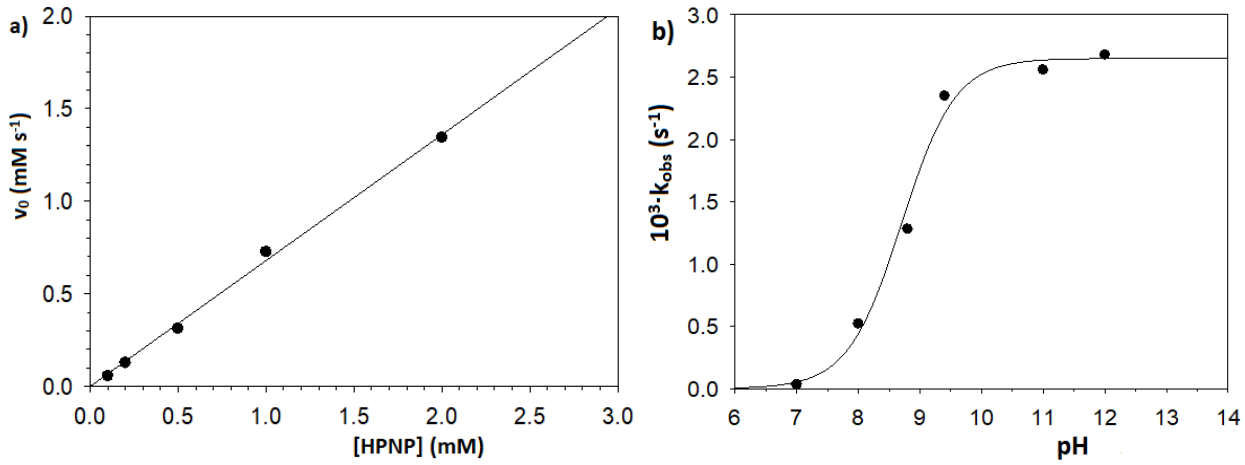


Figure 5: ***a)*** *Plot of initial rates for the cleavage of HPNP catalyzed by 0.50 mM* ***1****H⁺-Cu vs substrate concentration (80% DMSO, pH 25 °C). From the slope of the straight line* $k_{obs} = (1.36 \pm 0.06) \cdot 10^{-3} s^{-1}$. ***b)*** *pH-rate profile for the cleavage of 0.20 mM HPNP catalyzed by 0.50 mM* ***1****H⁺-Cu (80% DMSO, 25 °C). Data points are experimental and the line is the plot of the regression analysis with* ***Eqn. 1****.*

Important results were obtained also from the studies of the catalytic activity of the same Cu(II) complex as function of the pH. The plot of the experimental data gave in fact a sigmoidal shaped curve (**Figure 5b**) which is typical of kinetic titrations in which an unreactive species is transformed into a reactive one at higher pH values. The treatment of the given pH-rate profile with **Eqn. 1**, where k was the second-order rate constant for a reaction catalyzed by the fully deprotonated form of a generic chemical species whose acidity constant was K_a and stoichiometric concentration was $[Cat]_{tot}$, gave $pK_a = 8.7 \pm 0.1$ and $k = 5.3 \pm 0.2$ $M^{-1}s^{-1}$ as best values for the fitting parameters after a least-squares regression procedure.

$$k_{\mathrm{obs}} = \frac{k\,[Cat]_{tot}}{1 + [\mathrm{H}^+]/K_{\mathrm{a}}} \quad \boldsymbol{Eqn.1}$$

The perfect agreement of the pK_a value for the Cu(II)-bound water molecule of **1**H^+-Cu determined by this kinetic investigation (8.7 ± 0.1) with the one obtained from potentiometric titrations (8.8 ± 0.1) clearly indicated that the hydroxo form was the sole active species of the catalyst. To further support this idea, it can be observed that the limiting value of $(2.65 \pm 0.10) \cdot 10^{-1} s^{-1}$ approached by the quantity $k_{obs} = k\,[Cat]_{tot}$ for high pH values (i.e. when $[H^+]/K_a << 1$) was twice as large as the k_{obs} value previously measured at pH 8.9 (**Table 2**, entry 5), where only an half of the water molecules coordinated to the metal center were deprotonated.

According to these data we proposed, for the cleavage of HPNP by the studied calix[4]arene derivatives, the catalytic mechanism reported in **Figure 6**, where the metal-coordinated hydroxide of **1**H^+-Cu acts as a general base to deprotonate the β-OH group of the substrate, while the positively charged guanidinium ion interacts with the phosphate to stabilize the transition state through electrostatic interactions, eventually assisted in this task by the metal center[41].

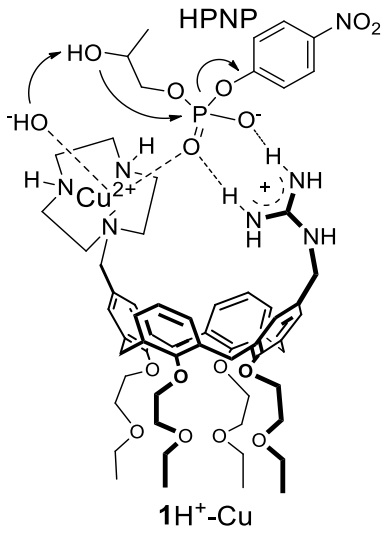


*Figure 6: Proposed mechanism of action of **1**-Cu in the cleavage of HPNP.*

2.2.4 Cleavage of diribonucleoside monophosphates (NpN's)

Regardless to whether the cleavage of phosphate diesters occurs via an associative (A_N+D_N) or a concerted (A_ND_N) mechanism (see **Chapter 1, Paragraph 1.3.1**)[10], it is currently accepted that the replacement of a good leaving group with a poor one leads the transition state to bear a higher pentavalent phosphorane dianion character. Thus, it has been argued that the cleavage of substrates less activated than HPNP by the given calix[4]arene catalysts could undergo to larger rate enhancements respect to the background, thanks to a stronger electrophilic/electrostatic interaction between the altered form of the phosphate group and the neighboring guanidinium unit and/or metal center of the catalysts. Accordingly, a set of investigations was performed to study the transesterification of four different diribonucleoside monophosphate (NpN'), namely GpU, CpA, UpU and GpA, in the presence of compounds **1**H$^+$-Cu and **2**H$^+$-Cu (**Figure 7**).

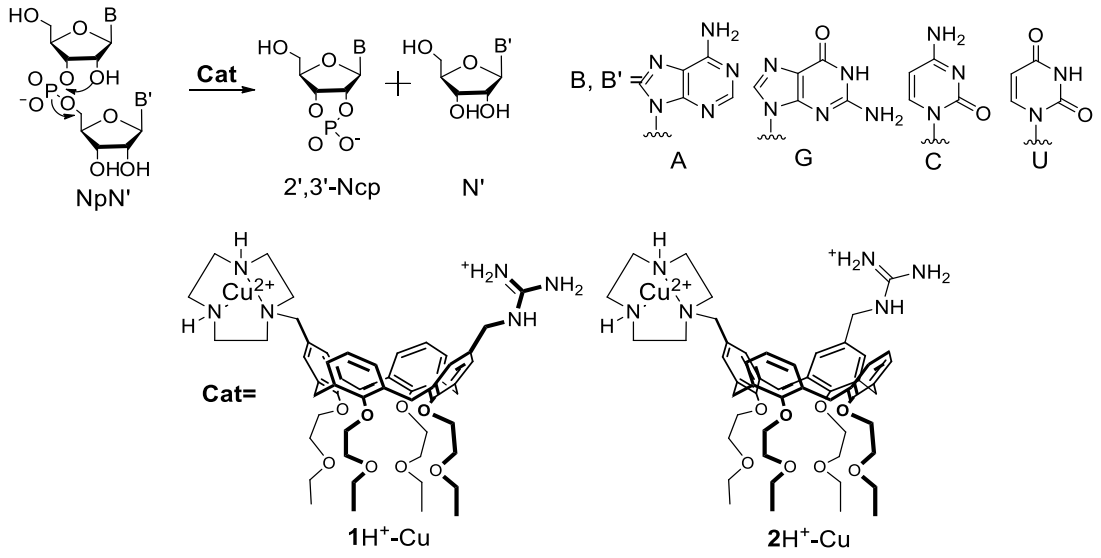


*Figure 7: Generic mechanism for the cleavage of diribonucleoside monophophosphates (NpN's) by **1**H$^+$-Cu and **2**H$^+$-Cu.*

The experiments were carried out in the same conditions tested for HPNP, but increasing the temperature from 25°C to 50°C, since, bearing a worse leaving group, NpN's are less prone to transesterification than HPNP. The reaction progress was monitored by HPLC analysis of periodically withdrawn samples of the reaction mixtures[29,42,43] (**Figure 8**) and pseudo-first-order rate constants k_{obs} were calculated as $k_{obs} = v_o/[NpN']$, where v_o was the initial rate of nucleoside *N*' formation.

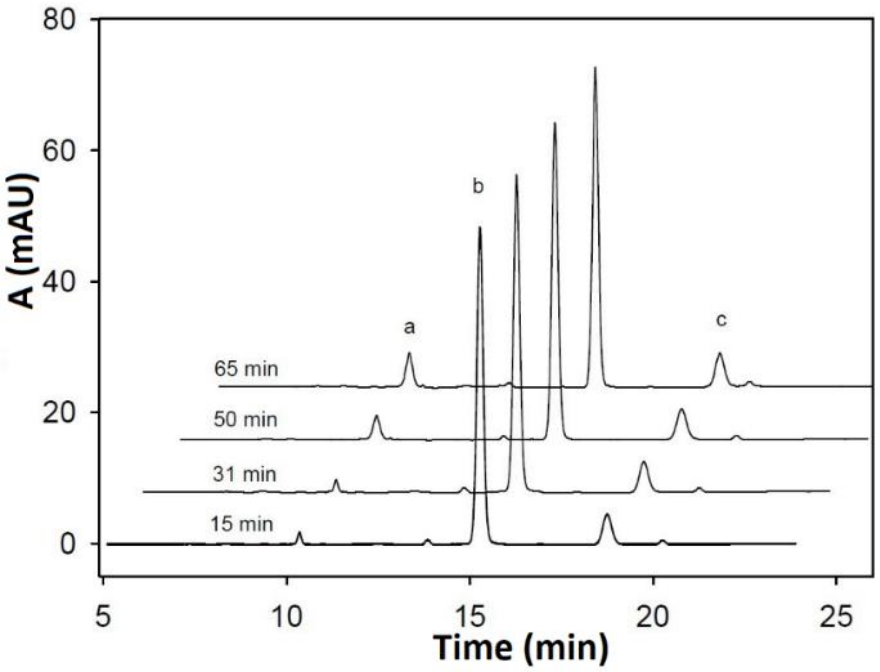


*Figure 8: HPLC monitoring of the cleavage of UpU dimer into 2',3'-Ucp and U by **1**H⁺-Cu. The reaction progress was studied by analyzing periodically withdrawn samples from the reaction mixture (0.1 mM diribonucleoside monophosphate, 0.5 mM catalyst, 80% DMSO, 50°C, pH 9.9). The peaks corresponding to uridine, the starting diribonucleoside monophosphate and the internal standard (p-hydroxybenzoic acid) are indicated as **a**, **b** and **c**, respectively.*

The experimental data (**Table 3**) show that both compounds were able to cleave all the tested substrates, giving an increase of their cleavage rates ranging from 4 to 7 order of magnitude respect to the uncatalyzed reaction. The catalytic efficiency was quite similar within the two regioisomers only for the transesterification of GpU and UpU. On the other side, **1**H⁺-Cu cleave GpA much faster than CpA, while the opposite is true for reaction catalyzed by **2**H⁺-Cu. More in details, GpA and CpA were cleaved 160-folds faster and 40-folds slower by the 1,3-distal Cu(II) complex than by its 1,2-vicinal isomer, respectively, indicating that the phosphodiesterase activity was strongly dependent on the identity of both the substrate and the catalyst which composed the catalytic system. In fact, being the rate constant for the spontaneous cleavage (k_{bg}) similar for each diribonucleoside monophosphates, it seemed unlikely that the different transesterification rates enhancements observed within the various substrates could be originated by their different intrinsic reactivity.

Table 3: Cleavage of diribonucleoside 3',5'-monophosphates NpN′ *by and* **2**H^+–Cu **1**H^+–Cu. *The reaction mixtures were composed of 0.50 mM catalysts, 0.050 mM* NpN′, *0.10 M* N,N′*-diisopropyl ethanolamine buffer in 80% DMSO (pH 8.9, 50°C). Error limits were on the order of ±10%.* [a] *The* k_{bg} *values (s^{-1}), were calculated from data measured in the presence of 1.0 mM* Me_4NOH *in 80% DMSO, 50.0 °C, and extrapolated to pH 8.9.* k_{bg} *(GpU)= 3.1×10^{-11} s^{-1}(from* ***Ref. 29****),* k_{bg} *(CpA)= 1.8×10^{-11} s^{-1} (from* ***Ref. 29****);* k_{bg} *(UpU) = 2.6×10^{-11} s^{-1} (from* ***Ref. 43****);* k_{bg} *(GpA)= 1.5×10^{-11} s^{-1} (from this work).*

	1H^+-Cu		**2H^+-Cu**	
	$10^6 \cdot k_{obs}$ (s^{-1})	**k_{obs}/k_{bg}**	**$10^6 \cdot k_{obs}$ (s^{-1})**	**k_{obs}/k_{bg} [a]**
GpU	85	3.4×10^6	93	3.0×10^6
CpA	1.2	8.6×10^4	48	2.7×10^6
UpU	26	1.3×10^6	15	5.8×10^5
GpA	160	1.3×10^7	1.0	6.7×10^4

A convenient comparison of the cleavage activity experienced by this type of substrates was obtained (**Table 4**) by dividing the k_{obs}/k_{bg} values listed in **Table 3** for NpN' ribonucleosides with the those indicated in **Table 2** for HPNP in the presence of both **1**H^+-Cu and **2**H^+-Cu, under the assumption that the data obtained for HPNP at 25°C give a reasonable estimation of the data that would have been measured at 50°C.

Table 4: Comparison of rate accelerations over background induced by **1**H^+*–Cu and* **2**H^+*–Cu in the cleavage of NpN' and HPNP substrates. The* $(k_{obs}/k_{bg})_{HPNP}$ *and* $(k_{obs}/k_{bg})_{NpN'}$ *values are taken from* ***Table 2*** *and* ***Table 3****, respectively.*

Catalyst /substrate pairs	$\frac{(k_{obs}/k_{bg})_{NpN'}}{(k_{obs}/k_{bg})_{HPNP}}$
1H^+–Cu / GpA	41
1H^+–Cu / GpU	11
2H^+–Cu / GpU	8.3
2H^+–Cu / CpA	7.5
1H^+–Cu / UpU	4.1
2H^+–Cu / UpU	1.6
1H^+–Cu / CpA	0.27
2H^+–Cu / GpA	0.19

Most of the diribonucleosides were cleaved more efficiently than HPNP by the tested Cu(II)-complexes, thanks to the exploitation of a larger electrostatic stabilization of their transition states. The effect was more evident for the most reactive catalyst-diribonucleoside pairs, with the best results obtained for the GpA which was cleaved 41-folds faster than HPNP by **1**H^+-Cu, while became gradually less important for the other substrates. In two cases, namely for the transesterification of CpA by **1**H^+-Cu and of GpA by **2**H^+-Cu, the rate accelerations were instead lower than those observed for HPNP, indicating that the benefits arising from a higher stabilization of the transition state was overcome by other unfavorable factors, probably of steric origin. Thus, once again, it can be observed that the structural features of both the catalysts and the substrates can have an important role to determine the efficiency of the catalytic transesterification of diribonucleoside monophosphates.

The comparison of the phosphodiesterase activity of **1**H^+-Cu and **2**H^+-Cu in the cleavage of the tested NpN' ribonucleosides with the ones reported for other artificial catalysts is quite difficult since, often, the investigations have been performed under different experimental conditions and reliable estimation of the k_{bg} values, necessary for the calculations of the rate enhancements respect to the background, are not available. To the best of our knowledge, the only investigations which were carried out in conditions similar enough to allow a consistent comparison with the kinetic data in our possession are referred to catalysts based on *cone*-calix[4]arenes[29] or nanostructured supports functionalized with the guanidine-guanidinium pair(s) [42,43]. Also in these cases the catalytic efficiency was strongly dependent on the identity of the various catalyst-substrate combination, but, in general the two calix[4]arene derivatives presented in this chapter are the best catalysts for the studied substrates, both in terms of absolute cleavage rates and of acceleration respect to the background in the given solvent mixture, temperature and concentration ranges.

2.2.5 Cleavage of BNPP

Both the Zn(II)- and Cu(II) complexes of **1**-H^+ and **2**-H^+ were also tested as cleaving agents of the DNA model compound bis(*p*-nitrophenyl)-phosphate (BNPP) (**Figure 9a**), together with their corresponding monofunctional compounds TACN-Zn and TACN-Cu. For consistency with the previously presented investigations on the transesterification of HPNP, the experiments were carried out in 80% DMSO at pH values coincident with the pK_a of the metal bound water molecule present on each catalyst. The reaction was carried out at the temperature of 50°C in order to counterbalance the lower reactivity of the substrate compared to HPNP, as a consequence of the absence of the intramolecular nucleophilic β-OH group.

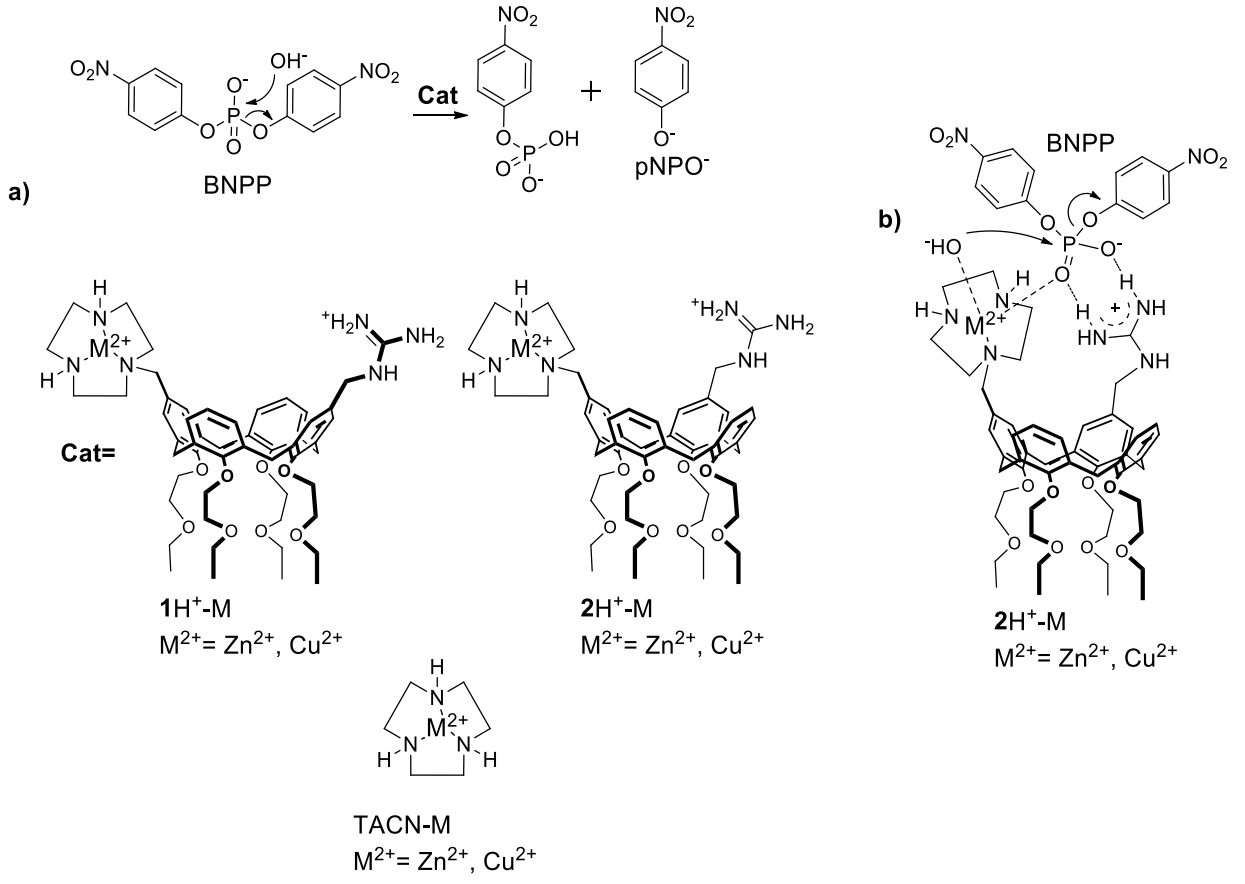


Figure 9: ***a)*** *Artificial bifunctional catalysts* ***1***H^+*-M and* ***2***H^+*-M and corresponding monofunctional compounds TACN-M (M= Zn(II) and Cu(II)) tested in the cleavage of the DNA model compound BNPP and* ***b)*** *proposed mechanism of action for the cleavage of the substrate by the Zn(II) and Cu(II) complexes of* ***2***H^+.

Pseudo-first-order rate constants k_{obs} were calculated from the initial rate values of the *p*-NPOH liberation (spectrophotometric detection) and are listed in **Table 5**, together with rate enhancements over background values (k_{obs}/k_{bg}). The ability to hydrolyze BNPP was remarkable for all the bifunctional catalysts, with rate enhancements respect to the background process ranging from $4.9\cdot10^4$- to $1.1\cdot10^6$-folds. A large degree of cooperation between the active units was also observed since the increase of catalytic activity respect to the corresponding control compounds was comprised between 110- and 1400-folds for **1**H^+-Zn and **2**H^+-Cu, respectively. The high degree of synergism observed for the reactive units of the Zn(II) complexes, together with the marked superiority of the 1,2-vicinal calix[4]arene over the regioisomeric 1,3-distal compounds, are in contrast with what found in similar investigations on HPNP (see **Table 2**). Moreover, in terms of specific rate constants, the catalytic activity of the Zn(II)-binding calix[4]arenes was comparable or even higher than those of the corresponding Cu(II) complexes (see entry 2 *vs* 5 or entry 3 *vs* 6 of **Table 5** for a comparison), despite **1**H^+-Cu and **2**H^+-Cu were more efficient in terms of rate enhancement respect to the background hydroxide-catalyzed reaction, since they were tested at lower pH values.

*Table 5: Cleavage of the DNA model BNPP in the presence of metal complexes **1**-H⁺-M, **2**-H⁺-M, and TACN-M (M = Zn(II), Cu(II)). The experiments were carried out in the presence of 0.20 mM BNPP, 0.50 mM catalyst, 0.10 M* N,N′*-diisopropyl ethanolamine buffer (80% DMSO, 50.0 °C).* [a] *In the presence of 0.50 mM guanidinium chloride (pH range 8.8–9.9), no detectable liberation of* p-nitrophenol *was observed within 12 h.* [b] k_{obs} *were calculated as* v_o*/[HPNP]; error limits on the order of ±10%.* [c] k_{rel} *were calculated as* k_{obs}*(bifunctional catalyst)/*k_{obs}*(TACN-M).* [d] k_{bg} *data from extrapolation to the given pH value of the hydroxide catalyzed cleavage of BNPP were measured in the presence of 1.0 mM* Me_4NOH ($k_{obs}=5.4 \times 10^{-4}\ s^{-1}$). *The following* k_{bg} *values (s^{-1}) were obtained (pH value in parenthesis):* 1.7×10^{-9} *(9.9),* 1.4×10^{-9} *(9.8),* 1.7×10^{-10} *(8.9),* 1.4×10^{-10} *(8.8).*

Entry	pH	Tested catalyst [a]	$10^6 \cdot k_{obs}$ (s^{-1}) [b]	k_{rel} [c]	(k_{obs}/k_{bg}) [d]
1	9.9	TACN–Zn	0.74	1.0	430
2	9.9	**1**H⁺–Zn	84	110	49000
3	9.8	**2**H⁺–Zn	180	240	130000
4	8.8	TACN–Cu	0.14	1.0	1100
5	8.8	**1**H⁺–Cu	20	140	140000
6	8.9	**2**H⁺–Cu	190	1400	1100000

It can be also pointed out that, on the contrary of what it could be expected by simple considerations of the electronic activation of the substrates[44], BNPP was cleaved approximately with the same efficiency of HPNP by **1**H⁺-Cu and **2**H⁺-Cu and even faster by the corresponding Zn(II)-complexes. In the hydrolysis of BNPP, in fact, the two negative charges present in the phosphorane-like transition state, formed during the attack of the incoming nucleophile, can be delocalized on two electron withdrawing *p*-nitrophenyl moieties, while HPNP bears only one of these groups (**Figure 10**).

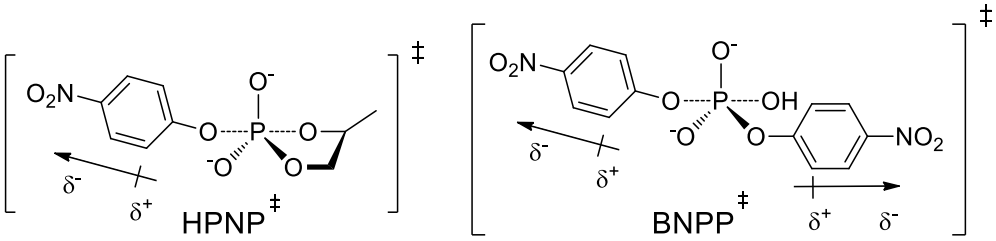


*Figure 10: Schematic representation of the dianionic phosphorane-like transition state of HPNP (HPNP‡) and BNPP (BNPP‡) during the cleavage by **1**-H⁺-M and **2**-H⁺-M (M=Zn(II), Cu(II)). In the case of BNPP‡ the negative charge density on the altered form of the phosphate group is lower with respect to HPNP‡, because of the presence of two electron withdrawing p-nitrophenyl units instead of one.*

Thus, in the former substrate, the lower electron density in the altered form of the scissile phosphate would be suggested to disfavor the establishment of electrostatic/electrophilic interaction with the guanidinium unit and/or the metal center of the catalysts, thus resulting in a lower stabilization of the transition state and consequently in lower acceleration of the studied reaction, when compared to the background. However, for the metal ion-catalyzed cleavage of BNPP, these detrimental factors could be balanced by the entropic gain generated thanks to the conversion of an otherwise intermolecular

reaction (**Figure 9a**) into an intramolecular nucleophilic attack of a metal-bound hydroxide ion (**Figure 9b**)[45,46,47]. Additionally, it has not to be excluded that the participation of the β-OH group could determine a lower reactivity of HPNP by imposing particular steric and geometrical requirements to allow transesterification to occur. Moreover, it is possible that the importance of these steric and geometrical requirements depends on whether between Zn(II) and Cu(II) is the metal center coordinated by the catalysts, suggesting that they are at the origin of the lower catalytic activity of **1**-H^+-Zn and **2**-H^+-Zn respect to the parent Cu(II) complexes into cleave HPNP. In the case of BNPP, it can be also suggested that analogous steric and geometric effects were responsible of the different catalytic efficiency reported for the tested calix[4]arene as function of the orientation of the two neighboring active units at their upper rim, resulting in a superiority of the 1,2-vicinal regioisomer which was not observed in the case of HPNP.

According to the analysis of the kinetic data in our possession, the mechanism of action that we propose for the cleavage of BNPP was analogous to that previously proposed for HPNP (**Figure 9b** and **6**, respectively), with the difference that, while the guanidinium group and the Zn(II) or Cu(II) ion cooperate to provide electrostatic/electrophilic activation of the substrate, the metal-bound hydroxide act as nucleophile to attack the phosphate group instead of as general base to deprotonate the lacking β-OH unit.

2.3 Conclusions

In this chapter the synthesis and characterization of the novel bifunctional calix[4]arenes **1** and **2**, decorated at the upper rim with a guanidinium and a TACN unit in 1,3 distal or 1,2-vicinal position, respectively, have been described. These derivatives were subsequently tested as artificial phosphodiesterases in the cleavage of different DNA and RNA model compounds. The metal complexes **1**H^+-M and **2**H^+-M, formed after coordination of Zn(II) or Cu(II) metal cations by the monoprotonated form of the catalysts, were able to cleave efficiently all the tested phosphate diesters, despite the catalytic performances were dependent on the identity of the studied calix[4]arene-substrate pair.

For the transesterification of the RNA model compound HPNP the best results were obtained with the Cu(II)-complexes of the two catalysts, which were significantly more efficient respect to the corresponding Zn(II)-adducts and gave an acceleration of the process of 5 order of magnitudes respect to the background. Also the extent of the cooperation between the two active units was much higher in the Cu(II) derivatives, suggesting that the different features of the coordination spheres of the two metal centers could give different steric requirements that influence the activity of the corresponding catalysts. No significant variations of the catalytic efficiency were instead found by comparing the results obtained within the two studied regioisomers after coordination of the same metal cation. The

combination of the kinetic and potentiometric data clearly evidenced the participation of a metal-bound water molecule to the cleavage of HPNP, acting as generic base for the deprotonation of the β-OH group of the substrate and facilitating its intramolecular nucleophilic attack on the phosphate group. Quite interestingly, a significant role was demonstrated to be played by the guanidinium unit and the metal center of the catalysts, which lower the activation energy of the phosphodiester cleavage by electrostatic stabilization of the negatively charged transition state.

Another set of kinetic investigations involved compounds **1**H^+-Cu and **2**H^+-Cu and four different diribonucleoside monophosphates. The observed rate acceleration respect to the hydroxide catalyzed reaction were as high as 10^4-10^7-folds and strongly dependent on the structure of the catalyst-substrate pair, with diribonucleoside which were cleaved more efficiently by the 1,2-vicinal Cu(II) complex and others by the 1,3-distal regioisomer. Being the intrinsic reactivity of the given substrates similar, it was proposed that the reported differences in catalytic efficiency arose from the different steric and geometrical requirements of the each catalyst-substrate pairs. Most of the catalyst-diribonucleoside pairs experienced cleavage accelerations higher than those given by the same calix[4]arenes on HPNP, thanks to the operation of a larger electrostatic stabilization of the transition states, whose dianionic phosphorane-like character was increased by the presence of a worse leaving group on the phosphate. However, for the CpA/**1**H^+-Cu and GpA/**2**H^+-Cu pairs the reverse was observed, probably because of the presence of detrimental steric effects that overcame the benefits deriving from a higher stabilization of the transition state.

Compounds **1**H^+-M and **2**H^+-M (M= Zn(II), Cu(II)) were also tested in the catalytic hydrolysis of the DNA model compound BNPP. Independently of whether a Zn(II) or a Cu(II) ion was complexed, the 1,2-vicinal calix[4]arene derivative was superior than its 1,3-distal isomer, with a maximum rate acceleration respect to the uncatalyzed reaction of $1.1 \cdot 10^6$-folds reported for **2**H^+-Cu. Considering the two isomers separately, the catalytic activity of their Zn(II) and Cu(II) complexes was similar in terms of specific cleavage rates, despite for the latter ones highest acceleration over the background were reported, since the investigations were performed at lower pH values. Similar, and particularly relevant, was also the extent of cooperation between the catalytic units, being the tested calix[4]arene complexes 2-3 order of magnitude more active than the corresponding monofunctional TACN-M complexes (M=Zn(II), Cu(II)). According to the experimental data, it was proposed that the cleavage of the substrate took place via nucleophilic attack on the phosphate of a hydroxide ion bound to the Zn(II) or Cu(II) cations of the studied catalysts, assisted by both the metal center and the guanidinium moiety, which provided electrostatic stabilization of the transition state. BNPP was cleaved as efficiently as HPNP by the tested catalysts, despite the presence of better leaving groups on its phosphate moiety would suggested a lower electrostatic stabilization of the transition state respect to the latter substrate. To explain these finding it was proposed that the conversion of the intermolecular hydroxide nucleophilic attack of the

uncatalyzed reaction into a calix[4]arene-catalyzed intramolecular process was particularly beneficial to promote the hydrolysis of the BNPP substrate.

2.4 Experimental section

2.4.1 General information

All moisture sensitive reactions were carried out under a nitrogen atmosphere, using previously oven-dried glassware. All dry solvents were prepared according to standard procedures, distilled before use and stored over 3 or 4 Å molecular sieves. Analytical TLC were performed using prepared plates of silica gel (Merck 60 F-254) and then revealed with UV light or with ninhydrin (5% in EtOH) when free amino groups were present on the studied compounds. Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates. mQ water were used in the preparation of 80% DMSO used in kinetic and potentiometric experiments. HPNP[48], was prepared according to literature procedures. All other solvents and reagents were commercial samples and used as such. Commercial samples of *N*p*N*' and their aqueous solutions were stored at -20 °C.

Care was taken when handling tetramethylammonium perchlorate because it is potentially explosive[49]. No accident occurred in the course of the present work.

^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^{1}H nucleus at 300 MHz and 400 MHz, respectively, and of ^{13}C nucleus at 75 MHz and 100 MHz, respectively) and partially deuterated solvents were used as internal standards to calculate the chemical shifts (δ values in ppm). All ^{13}C NMR spectra were performed with proton decoupling. High resolution electrospray ionization (ESI) mass analysis were performed with a Waters single-quadrupole spectrometer in positive mode using MeOH or CH_3CN as solvents. Melting points were determined on an Electrothermal apparatus in closed capillaries. Potentiometric titrations were performed by an automatic titrator equipped with a pH electrode. Spectrophotometric measurements of *p*-nitrophenol liberation were carried out at 400 nm on a double beam spectrophotometer. HPLC analyses were performed on a liquid chromatograph equipped with a UV-vis detector operating at 254 nm.

2.4.2 Potentiometric titrations

Potentiometric titrations were carried out according to a previously reported procedure[28]. A freshly prepared solution of 50 or 100 mM Me_4NOH in 80% DMSO was added in small increments under an argon atmosphere a 1 mM solution (5mL, 80% DMSO, 25°C) of **1**·4HCl or **2**·4HCl, using 10 mM Me_4NClO_4 as ionic strength buffer, in the absence or presence of 1 mol equiv of $ZnCl_2$ or $CuCl_2$. The electrode calibration was performed as reported in literature[41]. Elaboration of the titration plots was carried out with the software HYPERQUAD 2000[50,51].

2.4.3 Kinetic Measurements

Metal complexes were formed in situ by addition of the calculated stoichiometric amount of a concentrated aqueous solution of the metal salt ($ZnCl_2$ or $CuCl_2$) to the reaction mixture. The solutions were incubated for 1h before the start of the kinetic run by fast addition of a small volume of the substrate solution.

2.4.4 HPNP and BNPP

Observed rate constants were obtained by UV–vis monitoring of *p*-nitrophenol liberation at 400 nm (initial rate method). Experiments were carried out in the presence of 0.50 mM precatalysts **1**·4HCl or **2**·4HCl, 10 mM Me_4NClO_4, 0.50 mM $CuCl_2$ or $ZnCl_2$, and 0.10 M N,N' diisopropyl ethanolamine buffer, (80% DMSO, 25°C for HPNP, 50°C for BNPP). The total concentration of released *p*-nitrophenol was extrapolated from calibration curves built at the required pH values by measuring the absorbance at 400 nm of known aliquots of pNPOH dissolved in the same medium. The pH of the solutions was adjusted with a 50 or 100 mM solution of $HClO_4$ in 80% DMSO.

The pseudo first order rate constants for the spontaneous background reaction (k_{bg}) of the cleavage of BNPP were obtained by extrapolation to the required pH values of the pseudo first order constant measured in the presence of 1.0 mM Me_4NOH (pH 15.4), 10 mM Me_4NClO_4, (80% DMSO, 50.0 °C).

2.4.5 Diribonucleoside monophosphates

HPLC monitoring of nucleoside liberation (initial rate method) was carried out on solutions with the following composition: 0.50 mM precatalysts **1**·4HCl or **2**·4HCl, 10 mM Me_4NClO_4, 0.50 mM $CuCl_2$, 0.10 mM NpN', 0.10 M N,N'-diisopropyl ethanolamine buffer (80% DMSO, pH 8.9, 50.0°C); the pH of the solution was adjusted with a 50 or 100 mM solution of $HClO_4$ in 80% DMSO. Aliquots of the reaction mixture (60–80 μL) were withdrawn at appropriate time intervals and quenched with the same volume of a 15 mM solution of $HClO_4$ in 80% DMSO. After addition of *p*-hydroxybenzoic acid (internal standard) in 80% DMSO, the solution was filtered and analyzed by HPLC on a Supelcosil C-18 DB column (25 cm × 4.6 mm I.D., particle size 5 μm), using the following eluent: H_2O (0.1% trifluoroacetic acid)/CH_3CN, gradient from 100:0 to 85:15 in 28 min, flow 0.9 mL/min. The pseudo-first order specific rate for background reaction (k_{bg}) of GpA was extrapolated at pH 8.8, from the initial rate of hydroxide catalyzed nucleoside liberation in the presence of 1.0 mM Me_4NOH (pH 15.4), 10 mM Me_4NClO_4, 80% DMSO, 50.0 °C.

2.4.6 Calix[4]arene synthesis

Compounds **1**·4HCl and **2**·4HCl were prepared starting from the 5,17-bis(chloromethyl)-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene **3** and the 5,11-bis(chloromethyl)-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene **8**, respectively, in turn synthetized according to literature procedures[32,33].

5-(N-phthalimidomethyl)-17-(chloromethyl)-25,26,27,28-tetrakis(2-ethoxy-ethoxy)calix[4]arene (4).

To a solution of **3** (0.24 g, 0.30 mmol) in dry DMF (20 mL), potassium phthalimide (0.061 g, 0.30 mmol) was added. The reaction mixture was stirred overnight at r.t., then was quenched by adding distilled water (20 mL) and extracted with AcOEt (2 x 20 mL). The organic layers were combined, washed with brine (3 x 20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude was purified by flash chromatography (hexane/AcOEt 6:4) to give **4** as a white solid foam (0.16 g, 0.170 mmol; 57% yield): mp 68-72°C. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 7.87-7.82 (m, 2H, Ar*H* phthalimide); 7.70-7.69 (m, 2H, Ar*H* phthalimide); 6.80 (s, 2H, Ar*H*-meta-CH_2phthalimide); 6.77 (s, 2H, Ar*H*-meta-CH_2Cl); 6.51 (m, 6H, Ar*H*-meta and Ar*H*-para); 4.54 (s, 2H, *CH_2*phthalimide); 4.51 (d, *J* =13,4 Hz, 2H, Ar*CH_2*Ar ax.); 4.44 (d, *J* =13,4 Hz, 2H, Ar*CH_2*Ar ax); 4.40 (s, 4H, *CH_2*Cl); 4.15 (t, *J* =5,6, 4H, ArO*CH_2*CH_2O); 4.12-4.03 (m, 4H, ArO*CH_2*CH_2O); 3.85-3.78 (m, 8H, ArOCH_2*CH_2*O); 3.56-3.45 (m, 8H, O*CH_2*CH_3); 3.16 (d, *J* =13,4 Hz, 2H, Ar*CH_2*Ar eq.); 3.11 (d, *J* =13,4 Hz, 2H, Ar*CH_2*Ar eq.); 1.28-1.11 (m, 12H, CH_2*CH_3*). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 168.1, 157.1, 153.5, 155.8, 135.9, 135.6, 134.4, 134.2, 133.9, 132.2, 130.9, 129.7, 128.7, 128.2, 128.0, 123.2, 122.5, 73.3, 73.1, 72.9, 69.7, 69.6, 66.4, 66.3, 46.8, 41.3, 30.8, 30.7, 15.3. HR ES-MS: *m/z* Calcd for $C_{54}H_{62}O_{10}ClNNa$ [(**4**+Na)$^+$] 942.39599, found 942.39545; *m/z* Calcd for $C_{54}H_{66}ClN_2O_{10}$, [(**4**+NH_4)$^+$] 937.44060, found 937.44005.

5-(N-phthalimidomethyl)-17-[N,N'-bis(Boc)-1,4,7-triazacyclo-non-1-yl]methyl-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (5).

To a solution of **4** (0.105 g, 0.114 mmol) in dry CH_3CN (4 mL), 1,4-bis(*Boc*)-1,4,7-triazacyclononane (0.055 g, 0.17 mmol) and K_2CO_3 (0.021 g, 0.15 mmol) were added. The reaction mixture was stirred for 4 days at room temperature, during which additional 1,4-bis-(*Boc*)-1,4,7-triazacyclononane (2 x 0.015 g, 0.046 mmol) was added. The reaction was then quenched by evaporating the solvent under reduced pressure. The residue was taken with DCM (20 mL) and the organic phase was washed with a saturated solution of $NaHCO_3$ (20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude was purified by flash chromatography (hexane/AcOEt 6:4 – AcOEt/MeOH 7:3) to give **5** as a colorless oil (0.11 g; 0.080 mmol; 70% yield) and recover unreacted 1,4-bis-(*Boc*)-1,4,7-triazacyclononane (0.066 g , 0.055 mmol; 48% yield): ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 7.83-7.82 (m, 2H, Ar*H* phthalimide); 7.70-7.67 (m, 2H, Ar*H* phthalimide); 7.08, 6.99, 6.97 (3 s, 2H, Ar*H*-meta-CH_2phthtalimide); 6.89, 6.85, 6.75 (3 s, 2H, Ar*H*-meta-CH_2TACN); 6.46, 6.33-6.30 (2 m, 6H, Ar*H*-meta and *ArH*-para); 4.76, 4.72, 4.68, 4.58 (4 s, 2H, *CH_2*phthtalimide); 4.55-4.40 (m, 4H, Ar*CH_2*Ar ax.); 4.18-3.92 (m, 8H, ArO*CH_2*CH_2O); 3.86-3.76 (m, 8H, ArOCH_2*CH_2*O); 3.57-3.45 (m, 14H, O*CH_2*CH_3, *CH_2*TACN and N'(Boc)*CH_2CH_2*N''(Boc)); 3.23-3.08 (m, 8H, Ar*CH_2*Ar eq., N'(Boc)*CH_2*CH_2N and N''(Boc)*CH_2*CH_2N); 2.36 (m, 4H, N'(Boc)CH_2*CH_2*N and N''(Boc)CH_2*CH_2*N); 1,47-1,44 (m, 18H, OC*$(CH_3)_3$*); 1.28-1.11 (m, 12H, CH_2*CH_3*). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 168.1, 157.6, 155.8-155.3, 136.2-135.7, 133.9, 132.5-132.0, 129.8, 129.2, 127.8, 123.3, 122.4, 73.7, 73.3, 72.7,

69.6, 66.4, 60.7, 52.6, 52.4, 50.3, 49.9, 49.4, 41.6, 30.8, 28.6, 15.3. HR ES-MS: *m/z* Calcd for $C_{70}H_{93}N_4O_{14}$ [(**5**+H)$^+$] 1213.66883, found 1213.66828; *m/z* Calcd for $C_{70}H_{92}N_4O_{14}Na$, [(**5**+Na)$^+$] 1235.65078, found 1235.65453.

5-(Aminomethyl)-17-[N,N'-bis(Boc)-1,4,7-triazacyclonon-1-yl]-methyl-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (6).

To a solution of **5** (0.043 g, 0.035 mmol) in dry MeOH (4 mL), $N_2H_4{\cdot}H_2O$ (44 µL, 1.42 mmol) was added. The reaction mixture was heated to 60°C and stirred for 3 h, then was quenched by evaporating the solvent under reduced pressure. The residue was taken with DCM (20 mL) and the organic phase was washed with a solution of NaOH 1M (20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. **6** was obtained as a colorless oil (0.038 g, 0.035 mmol; quantitative yield), pure enough to avoid further purifications. ^{1}H NMR (300 MHz, $CDCl_3$) *δ* (ppm): ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 6.83, 6.75, 6.67 (3 s, 2H, Ar*H*-meta-CH_2NH_2); 6.79, 6.69, 6.64 (3s, 2H, Ar*H*-meta-CH_2TACN); 6.53, 6.44, 6.36 (3 m, 6H, Ar*H*-meta and *ArH*-para); 4.48-4.43 (m, 4H, Ar*CH*$_2$Ar ax.); 4.17-4.01 (m, 8H, ArO*CH*$_2$CH$_2$O); 3.86-3.81 (m, 8H, ArOCH$_2$*CH*$_2$O); 3.69-3.41 (m, 16H, O*CH*$_2$CH$_3$, *CH*$_2$TACN, *CH*$_2$NH$_2$ and N'(Boc)*CH*$_2$*CH*$_2$N''(Boc)); 3.13-3.08 (m, 8H, Ar*CH*$_2$Ar eq. and N'(Boc)*CH*$_2$CH$_2$N e N''(Boc)*CH*$_2$CH$_2$N); 2.60-2.46 (m, 4H, N'(Boc)CH$_2$*CH*$_2$N e N''(Boc)CH$_2$*CH*$_2$N); 1.47-1.44 (m, 18H, OC*(CH*$_3$*)*$_3$); 1.28-1.17 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (75 MHz, $CDCl_3$) *δ* (ppm): 157.1, 155.9, 155.8, 155.7, 155.6, 155.5, 155.3, 136.7, 135.9-134.0, 133.2, 133.0, 132.8, 129.2, 128.9, 128.5, 128.1, 127.9, 127.8, 127.8, 127.1, 127.0, 122.3, 79.4, 79.3, 79.3, 73.5-72.6, 69.7, 69.6, 66.4-66.2, 60.5, 53.6, 53.4, 50.3, 49.9, 49.6, 49.3, 46.1, 30.9, 28.6, 15.3. HR ES-MS: *m/z* Calcd for $C_{62}H_{91}N_4O_{12}$ [(**6**+H)$^+$] 1083.66335, found 1083.66638; *m/z* Calcd for $C_{62}H_{90}N_4O_{12}Na$, [(**6**+Na)$^+$] 1105.64530, found 1105.64758.

5–N-[N',N''–bis(Boc)guanidine]methyl-17-[4,7-bis(Boc)-1,4,7-triazacyclonon-1-yl]methyl-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (7).

To a solution of **6** (0.040 g, 0.037 mmol) in dry DCM (4 mL), *N*,*N*'-bis(*Boc*)-*N*''-triflylguanidine (0.017 g, 0.044 mmol) and triethylamine (5 µL, 0.037 mmol) were added. The reaction mixture was stirred overnight at room temperature, then was quenched by adding distilled water (20 mL) and was vigorously stirred for an additional 30 min. The aqueous phase was separated and washed with DCM (2 x 20 mL), then the combined organic phases were dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude was purified by preparative TLC plates (hexane/AcOEt 6.5:3.5) to give **7** as a light yellow oil (0.033 g, 0.025 mmol; 68% yield). ^{1}H NMR (400 MHz, $CDCl_3$) *δ* (ppm): ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 11.52 (s, 1H, *NH*Boc); 8.56, 8.50, 8.44 (3 br. t, *J*=4,8 Hz, CH$_2$*NH*guanidinium); 6.96, 6.87, 6.79 (3 s, 2H, Ar*H*-meta-CH$_2$guanidinium); 6.93, 6.85, 6.78 (3 s, 2H, Ar*H*-meta-CH$_2$TACN); 6.40-6.20 (m, 6H, Ar*H*-meta and *ArH*-para); 4.52-4.39 (m, 6H, Ar*CH*$_2$Ar ax. and *CH*$_2$guanidinium); 4.21-4.10 (m, 4H, ArO*CH*$_2$CH$_2$O); 4.03-3.93 (m, 4H, ArO*CH*$_2$CH$_2$O); 3.87-3.81 (m, 8H, ArOCH$_2$*CH*$_2$O); 3.56-3.46 (m, 14H,

O*CH*$_2$CH$_3$, *CH*$_2$TACN and N'(Boc)*CH*$_2$*CH*$_2$N''(Boc)); 3.23-3.09 (m, 8H, Ar*CH*Ar eq, N'(Boc)*CH*$_2$CH$_2$N and N''(Boc)*CH*$_2$CH$_2$N); 2.65-2.55 (m, 4H, N'(Boc)CH$_2$*CH*$_2$N and N''(Boc)CH$_2$*CH*$_2$N); 1.51-1.44 (m, 36H, OC*(CH*$_3$*)*$_3$); 1.26-1.22 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (75 MHz, CDCl$_3$) δ (ppm): 163.6, 157.7, 156.9, 156.5, 155.9, 155.7, 155.5, 155.1, 154.8, 136.7-135.3, 133.9-133.0, 130.5, 129.6, 129.3, 128.8, 128.0, 127.8, 122.3, 83.0, 79.5-79.2, 73.5-72.3, 69.7, 69.6, 66.4-66.2, 60.5, 53.9-53.2, 51.0-49.0, 44.2, 30.9, 28.6, 28.3, 28.1, 15.3. HR ES-MS: *m/z* Calcd for $C_{73}H_{109}N_6O_{16}$ [(**7**+H)$^+$] 1325.79001, found 1325.78946; *m/z* Calcd for $C_{73}H_{108}N_6O_{16}Na$ [(**7**+Na)$^+$] 1347.77195, found 1347.77140.

5–N-guanidinomethyl-17-N-[1,4,7-triazacyclonon-1-yl]methyl-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene, Tetrahydrochloride (1·4HCl).

In a mixture of DCM/TFA/TES 95:2.5:2.5 (10 mL), **7** (0.040 g, 0.030 mmol) was dissolved. The reaction mixture was stirred over night at room temperature and quenched by evaporating the solvent under reduced pressure. The residue was first taken in a solution of HCl 1 M in EtOH (3mL) and vigorously stirred for 30 min three times, to exchange the TFA anion, then was recrystallized from DCM with hexane to give **1**·4HCl as a light gray solid (0.032 g, 0.030 mmol; 98% yield): mp 123-130 °C. ^{1}H NMR (300 MHz, CD$_3$OD) δ (ppm): 7.56 (t, *J*=5,0Hz, 1H, CH$_2$*NH*guanidinium); 7.08 (s, 1H, *NH*$_2$guanidinium); 6.89 (s, 2H, Ar*H*-meta-CH$_2$guanidinium); 6.80 (s, 2H, Ar*H*-meta-CH$_2$TACN); 6.53 (d, *J*=6,6 Hz, 4H, Ar*H*-meta); 6.49 (t, *J*=6,6 Hz, 4H, Ar*H*-para); 4.59 (d, 4H, Ar*CH*$_2$Ar ax.); 4.19-4.15 (m, 6H, ArO*CH*$_2$CH$_2$O e *CH*$_2$guanidinium); 4.08-4.05 (m, 4H, ArO*CH*$_2$CH$_2$O); 3.92-3.85 (m, 8H, ArOCH$_2$*CH*$_2$O); 3.70 (s, 2H, *CH*$_2$TACN); 3.60-3.54 (m, 12H, O*CH*$_2$CH$_3$ and N'H*CH*$_2$*CH*$_2$N''H); 3.22-2.97 (m, 8H, Ar*CH*$_2$Ar eq., N'H*CH*$_2$CH$_2$NH and N''H*CH*$_2$CH$_2$NH); 2.69 (m, 4H, N'HCH$_2$*CH*$_2$N and N''HCH$_2$*CH*$_2$N); 1.23-1.16 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (75 MHz, CD$_3$OD) δ (ppm): 157.1, 156.9, 156.8, 155.8, 136.1, 135.8, 134.2, 134.1, 130.2, 129.4, 129.4, 127.9, 127.8, 127.4, 122.0, 73.3, 72.1, 69.9, 69.7, 69.6, 66.0, 58.8, 47.4, 44.8, 43.3, 42.1, 30.5, 30.4, 14.3. ES-MS: *m/z* 925.9 [(**2**+H)$^+$]. HR ES-MS: *m/z* Calcd for $C_{53}H_{77}N_6O_8$ [(**2**+H)$^+$] 925.58029, found 925.58091.

5-(N-phthalimidomethyl)-11-chloromethyl-25,26,27,28-tetrakis(2- ethoxyethoxy) calix[4]arene (9).

To a solution of **8** (0.255 g, 0.31 mmol) in dry toluene (8 mL), potassium phthalimide (0.058 g, 0.31 mmol) and dicyclohexyl-18-crown-6 (0.117 g, 0.31 mmol) were added. The reaction mixture was stirred at 110 °C under nitrogen atmosphere. After 3 h, 25 mL of distilled water and 10 mL of toluene were added and the mixture was stirred for 20 min. The separated organic phase was washed with distilled water (2 x 20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The residue was then submitted to flash chromatography (hexane/AcOEt 6/4). Pure compound **9** was isolated as a white foam after evaporation of the solvent (0.185 g, 0.201 mmol; 65% yield): 7.86-7.84 (m, 2H, Ar*H* phthalimide); 7.71-7.67 (m, 2H, Ar*H* phthalimide); 6.85 (d, 1H, J= 2.0 Hz, Ar*H*-meta-CH$_2$phthalimide); 6.83 (d, 1H, J= 2.0 Hz, Ar*H*-meta-CH$_2$phthalimide); 6.80-6.78 (m, 2H, Ar*H*-meta); 6.71 (t, 1H, J= 7.6 Hz,

Ar*H*-para); 6.40-6.36 (m, 5H, Ar*H*-meta-CH_2Cl, Ar*H*-meta and Ar*H*-para); 4.63 (s, 2H, *CH_2*phthalimide); 4.49 (d, *J* =13,6 Hz, 2H, Ar*CH_2*Ar ax.); 4.45 (d, *J* =13,6 Hz, 2H, Ar*CH_2*Ar ax.); 4.12 (m, 6H, *CH_2*Cl and ArO*CH_2*CH_2O); 4.04-3.98 (m, 4H, ArO*CH_2*CH_2O); 3.85-3.78 (m, 8H, ArOCH_2*CH_2*O); 3.56-3.45 (m, 8H, O*CH_2*CH_3); 3.15-3.07 (m, 4H, Ar*CH_2*Ar eq.); 1.25-1.11 (m, 12H, CH_2*CH_3*). ^{13}C NMR (100 MHz, $CDCl_3$) δ (ppm): 168.1, 157.1, 153.5, 155.8, 135.9, 135.6, 134.4, 134.2, 133.9, 132.2, 130.9, 129.7, 128.7, 123.2, 122.5, 73.3, 73.1, 72.9, 69.7, 69.6, 66.4, 66.3, 46.8, 41.3, 30.8, 30.7, 15.3. HR ES-MS: m/z Calcd for $C_{54}H_{62}ClNO_{10}Na$, [(**7**+Na)$^+$] 942.39544, found 942.39552.

5-Chloromethyl-11-[N,N'-bis(Boc)-1,4,7-triazacyclonon-1-ylmethyl]-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (10).

To a solution of **8** (0.084 g, 0.10 mmol) in dry CH_3CN (5 mL), 1,4-bis(*Boc*)-1,4,7-triazacyclononane (0.034 g, 0.10 mmol) and K_2CO_3 (0.014 g, 0.10 mmol) were added. The reaction mixture was stirred at room temperature for 6 days. In the course of reaction additional amounts of 1,4-bis-(*Boc*)-1,4,7-triazacyclononane (2 x 0.010 g, 0.030 mmol) were added. The reaction was then quenched by evaporating the solvent under reduced pressure. The residue was dissolved with DCM (20 mL) and the organic phase was washed with distilled water (2 x 20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude material was purified by flash chromatography (hexane/AcOEt/DCM 7:2:1 – hexane/AcOEt 7:3) to recover unreacted **8** (0.022 g , 0.027 mmol; 27% yield) and give **10** as a colorless oil (0.064 g, 0.058 mmol; 56% yield): ^{1}H NMR (300 MHz, $CDCl_3$) *δ* (ppm): 6.82-6.78, 6.71-6.65 (2 m, 5H, Ar*H*-meta-CH_2TACN, Ar*H*-meta and Ar*H*-para); 6.60 (t, 1H, J= 7.4 Hz Ar*H*-para); 6.51-6.40, 6.39-6.38 (2 m, 4H, Ar*H*-meta-CH_2Cl and Ar*H*-meta); 4.51 (br. d, 2H, *J* =12.6 Hz, Ar*CH_2*Ar ax.); 4.48 (br. d, 2H, *J* =13.2 Hz, Ar*CH_2*Ar ax.); 4.32-4.05 (m, 10H, *CH_2*Cl and ArO*CH_2*CH_2O); 3.84-3.82 (m, 8H, ArOCH_2*CH_2*O); 3.55-3.44 (m, 14H, O*CH_2*CH_3, *CH_2*TACN and N'(Boc)*CH_2*CH_2*N''(Boc)); 3.15-3.08 (m, 8H, ArCH_2Ar eq. and N'(Boc)*CH_2*CH_2N and N''(Boc)*CH_2*CH_2N); 2.59-2.37 (m, 4H, N'(Boc)CH_2*CH_2*N and N''(Boc)CH_2*CH_2*N); 1.49 (s, 9H, OC*$(CH_3)_3$*); 1.45 (s, 9H, OC*$(CH_3)_3$*); 1.23-1.17 (m, 12H, CH_2*CH_3*). ^{13}C NMR (100 MHz, $CDCl_3$) *δ* (ppm): 156.8-155.2, 135.7-133.9, 133.3-133.1, 128.7-127.8, 122.2, 79.4, 79.3, 73.4, 73.2, 72.9, 69.7, 66.4, 66.3, 60.2, 59.8, 53.6-52.9, 50.8-49.0, 46.6, 30.8, 28.6, 28.5, 15.3. HR ES-MS: *m/z* Calcd for $C_{62}H_{89}O_{12}N_3Cl$ [(**3**+H)$^+$] 1102.61293, found 1102.61236.

5-(N-Phthalimidomethyl)-11-[N,N'-bis(Boc)-1,4,7-triazacyclo-non-1-ylmethyl]-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (11).

To a solution of **10** (0.060 g, 0.055 mmol) in dry DMF (8 mL), potassium phthalimide (0.011 g, 0.060 mmol) was added. The reaction mixture was stirred overnight at room temperature and quenched by evaporating the solvent under reduced pressure. The crude material was dissolved in AcOEt (20 mL) and the organic phase was washed with a saturated solution of NaCl (3 x 20 mL), dried over anhydrous Na_2SO_4, and evaporated under reduced pressure. Compound **11** was obtained as a pale white oil (0.060

g, 0.049 mmol; 90% yield), pure enough to avoid further purifications: ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 7.83-7.81 (m, 2H, Ar*H* phthalimide), 7.71-7.67 (m, 2H, Ar*H* phthalimide); 7.04, 6.93, 6.88, 6.78, 6.72, 6.63, 6.54, 6.45, 6.32 (9 m, 10H, Ar*H*-meta-CH_2phthtalimide, Ar*H*-meta-CH_2TACN, Ar*H*-meta, Ar*H*-para); 4.65-4.38 (m, 6H, *CH_2*phthtalimide, Ar*CH_2*Ar ax.); 4.23-3.97 (m, 8H, ArO*CH_2*CH_2O); 3.90-3.76 (m, 8H, ArOCH_2*CH_2*O); 3.64-3.34 (m, 14H, O*CH_2*CH_3, *CH_2*TACN and N'(Boc)*CH_2CH_2*N''(Boc)); 3.26-3.06 (m, 6H, Ar*CH_2*Ar eq., N'(Boc)*CH_2*CH_2N); 3.00-3.70 (m, 2H, N''(Boc)*CH_2*CH_2N); 2.49-2.29 (m, 4H, N'(Boc)CH_2*CH_2*N and N''(Boc)CH_2*CH_2*N); 1.50 (s, 9H, OC*$(CH_3)_3$*); 1.44 (s, 9H, OC*$(CH_3)_3$*); 1.22-1.10 (m, 12H, CH_2*CH_3*). ^{13}C NMR (100 MHz, $CDCl_3$) δ (ppm): 167.7, 156.9-155.7, 135.9, 135.5, 134.2, 133.8, 133.1, 129.9, 128.7-127.9, 123.1, 122.1, 79.2, 73.3, 72.9, 69.6, 66.3, 60.0, 53.4, 50.3, 50.4-49.1, 41.8, 30.8, 28.6, 15.2. HR ES-MS: *m/z* Calcd for $C_{70}H_{93}N_4O_{14}$ [(**4**+H)$^+$] 1213.66828, found 1213.66847; *m/z* Calcd for $C_{70}H_{92}N_4O_{14}Na$, [(**4**+Na)$^+$] 1235.65022, found 1235.65075.

5-(Aminomethyl)-11-[N,N'-bis(Boc)-1,4,7-triazacyclonon-1-yl-methyl]-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (12).

To a solution of **11** (0.034 g, 0.028 mmol) in dry MeOH (4 mL), $N_2H_4{\cdot}H_2O$ (35 µL, 0.72 mmol) was added. The reaction mixture was heated to 65°C and stirred for 3 hours, then was quenched by evaporating the solvent under reduced pressure. The residue was dissolved with DCM (20 mL) and the organic phase was washed with a solution of NaOH 1M (20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. Compound **12** was obtained as a colorless oil (0.028 g, 0.026 mmol; 92% yield). Because of the instability of the amino group[52] the compound was not fully characterized and used as such in the following reaction. HR ES-MS: *m/z* Calcd for $C_{62}H_{91}N_4O_{12}$ [(**5**+H)$^+$] 1083.66280, found 1083.66238.

5–N-[N',N''–Bis(Boc)guanidine]methyl-11-[4,7-bis(Boc)-1,4,7-triazacyclonon-1-ylmethyl]-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (13).

To a solution of **12** (0.028 g, 0.026 mmol) in dry DCM (5 mL), *N*,*N*'-bis(*Boc*)-*N*''-triflylguanidine (0.015 g, 0.039 mmol) and triethylamine (5.4 µL, 0.039 mmol) were added. The reaction mixture was stirred overnight at room temperature, then quenched by adding distilled water (20 mL) and vigorously stirred for an additional 30 min. The aqueous phase was separated and washed with DCM (2 x 20 mL), then the combined organic phases were dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude material was purified by preparative TLC plates (hexane/AcOEt/DCM 6:3:1) to give **13** as a yellow oil (0.020 g, 0.018 mmol; 58% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 11.55 (s, 1H, *NH*Boc); 8.38, 8.31, 8.25 (3 br. t, 1H, *J*=4.8 Hz, CH_2*NH*guanidinium); 6.77-6.72 (m, 1H, Ar*H*-para); 6.66-6.61 (m, 4H, Ar*H*-meta-CH_2guanidinium and Ar*H*-para); 6.58-6.52 (m, 4H, ArH-meta-CH_2TACN and Ar*H*-meta); 6.49-6.3 (m, 1H, Ar*H*-para); 4.52-4.45 (m, 4H, Ar*CH_2*Ar ax.); 4.37, 4.29, 4.24 (3 t, 2H, *J*=4Hz, *CH_2*guanidinium);

4.22-4.04 (m, 8H, ArO*CH*$_2$CH$_2$O); 3.91-3.84 (m, 8H, ArOCH$_2$*CH*$_2$O); 3.72-3.40 (m, 14H, O*CH*$_2$CH$_3$, *CH*$_2$TACN and N'(Boc)*CH*$_2$*CH*$_2$N''(Boc)); 3.32-3.08 (m, 6H, Ar*CH*Ar eq and N'(Boc)*CH*$_2$CH$_2$N); 2.98-2.78 (m, 2H, N''(Boc)*CH*$_2$CH$_2$N); 2.53-2.25 (m, 4H, N'(Boc)CH$_2$*CH*$_2$N and N''(Boc)CH$_2$*CH*$_2$N); 1.54-1.46 (m, 36H, OC*(CH*$_3$*)*$_3$); 1.24-1.20 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (100 MHz, $CDCl_3$) δ (ppm): 163.6, 156.2, 155.9, 155.8, 155.7, 155.2, 153.1, 150.8, 135.7-134.1, 133.1, 130.4, 128.3, 128.2, 127.6, 127.4, 122.2, 82.9, 79.3, 73.5-72.8, 69.5, 66.3, 60.2, 53.4-52.9, 50.5-49.2, 45.0, 30.9, 28.7, 28.6, 28.3, 28.1, 15.3. HR ES-MS: *m/z* Calcd for $C_{73}H_{109}N_6O_{16}$ [(**6**+H)$^+$] 1325.78946, found 1325.78946; *m/z* Calcd for $C_{73}H_{108}N_6O_{16}Na$ [(**6**+Na)$^+$] 1347.77140, found 1347.77186.

5–N-Guanidinomethyl-11-N-[1,4,7-triazacyclononane-1-methyl]-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene, Tetrahydrochloride (2·4HCl).

In a mixture of DCM/TFA/TES 95:2.5:2.5 (10 mL), **13** (0.020 g, 0.015 mmol) was dissolved. The reaction mixture was stirred overnight at room temperature and quenched by evaporating the solvent under reduced pressure. The residue was first treated with a solution of 1 M HCl in EtOH (3mL), and vigorously stirred for 30 min three times, to exchange the TFA anion, then recrystallized from DCM/hexane to give **2**·4HCl as a waxy white solid (0.012 g, 0.013 mmol; 86% yield): ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 6.92-6.87 (m, 4H, Ar*H*-meta-CH$_2$guanidinium and Ar*H*-meta-CH$_2$TACN); 6.77-6.72 (m, 2H, Ar*H*-para); 6.64-6.57 (m, 4H, Ar*H*-meta); 4.62 (d, 2H, *J*=12.8 Hz, Ar*CH*$_2$Ar ax.); 4.61 (d, 1H, *J*=13.2 Hz, Ar*CH*$_2$Ar ax.); 4.60 (d, 1H, *J*=12.8 Hz, Ar*CH*$_2$Ar ax.); 4.27-4.22 (m, 6H, *CH*$_2$guanidinium and ArO*CH*$_2$CH$_2$O); 4.08-4.07 (m, 4H, ArO*CH*$_2$CH$_2$O); 3.99-3.96 (m, 4H, ArOCH$_2$*CH*$_2$O); 3.90-3.85 (m, 4H, ArOCH$_2$*CH*$_2$O); 3.65-3.56 (m, 14H, *CH*$_2$TACN, O*CH*$_2$CH$_3$ and N'H*CH*$_2$*CH*$_2$N''H); 3.25-3.18 (m, 4H, Ar*CH*$_2$Ar eq.); 3.03-3.00 (m, 4H, N'H*CH*$_2$CH$_2$NH and N''H*CH*$_2$CH$_2$NH); 2.55 (m, 4H, N'HCH$_2$*CH*$_2$N and N''HCH$_2$*CH*$_2$N); 1.27-1.20 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 157.1, 156.8, 156.7, 155.8, 136.3, 136.2, 135.7, 135.6, 134.8, 134.6, 134.5, 134.2, 130.0, 129.9, 129.6, 128.3, 128.2, 128.0, 127.9, 127.8, 127.5, 127.4, 122.0, 73.8, 73.7, 73.0, 72.8, 69.9, 69.8, 69.7, 69.6, 66.0, 58.2, 47.4, 44.6, 43.2, 42.1, 30.4, 14.3. HR ES-MS: *m/z* Calcd for $C_{53}H_{77}N_6O_8$ [(**1**+H)$^+$] 925.57974, found 925.57982.

2.5 Bibliography

(1) Berg, Jeremy B, Tymoczko, John L, Stryer, L. *Biochemistry, 5th Edition*, 5th ed.; W. H. Freeman: York, N., Ed.; **2002**.

(2) Fersht, A. *Enzyme Structure and Mechanism*, 2nd ed.; W. H. Freeman: New York, Ed.; **1984**.

(3) Rittié, L.; Perbal, B. *J. Cell Commun. Signal.* **2008**, *2* (1–2), 25.

(4) Gleave, M. E.; Monia, B. P. *Nat Rev Cancer* **2005**, *5* (6), 468.

(5) Trawick, B. N.; Daniher, A. T.; Bashkin, J. K. *Chem. Rev.* **1998**, *98* (3), 939.

(6) Breslow, R. *Artificial Enzymes*; Wiley-VCH Verlag GmbH & Co. KGaA, Ed.; Weinheim, Germany, **2005**.

(7) Raynal, M.; Ballester, P.; Vidal-Ferran, A.; van Leeuwen, P. W. N. M. *Chem. Soc. Rev.* **2014**, *43* (5), 1660.

(8) Raynal, M.; Ballester, P.; Vidal-Ferran, A.; van Leeuwen, P. W. N. M. *Chem. Soc. Rev.* **2014**, *43* (5), 1734.

(9) Mancin, F.; Scrimin, P.; Tecilla, P.; Tonellato, U. *Chem. Commun.* **2005**, No. 20, 2540.

(10) Mancin, F.; Scrimin, P.; Tecilla, P. *Chem. Commun.* **2012**, *48* (45), 5545.

(11) Weston, J. *Chem. Rev.* **2005**, *105* (6), 2151.

(12) Salvio, R. *Chem. Eur. J.* **2015**, *21* (31), 10960.

(13) Tjioe, L.; Joshi, T.; Brugger, J.; Graham, B.; Spiccia, L. *Inorg. Chem.* **2011**, *50* (2), 621.

(14) Kimura, E.; Kodama, Y.; Koike, T.; Shiro, M. *J. Am. Chem. Soc.* **1995**, *117* (32), 8304.

(15) Muche, M.-S.; Göbel, M. W. *Angew. Chem. Int. Ed. Engl.* **1996**, *35* (18), 2126.

(16) Scheffer, U.; Strick, A.; Ludwig, V.; Peter, S.; Kalden, E.; Göbel, M. W. *J. Am. Chem. Soc.* **2005**, *127* (7), 2211.

(17) Holtz, K. M.; Kantrowitz, E. R. *FEBS Lett.* **1999**, *462* (1–2), 7.

(18) Cotton, F. A.; Hazen, Edward E., J.; Legg, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76* (6), 2551.

(19) Redinbo, M. R.; Stewart, L.; Kuhn, P.; Champoux, J. J.; Hol, W. G. J. *Science (80).* **1998**, *279* (March), 1504.

(20) Stewart, L. *Science (80).* **1998**, *279* (5356), 1534.

(21) Raines, R. T.; Peptide, B. P.; Isomerization, B. **1998**, *2665* (96).

(22) Molenveld, P.; Engbersen, J. F. J.; Reinhoudt, D. N. *Chem. Soc. Rev.* **2000**, *29* (2), 75.

(23) Sanders, J. K. M. *Chem. Eur. J.* **1998**, *4* (8), 1378.

(24) Cacciapaglia, R.; Di Stefano, S.; Mandolini, L.; Salvio, R. *Supramol. Chem.* **2013**, *25* (9–11), 537.

(25) Giuliani, M.; Morbioli, I.; Sansone, F.; Casnati, A. *Chem. Commun.* **2015**, *51* (75), 14140.

(26) Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Reinhoudt, D. N.; Salvio, R.; Sartori, A.; Ungaro, R. *J. Am. Chem. Soc*. **2006**, 128, 12322.

(27) Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Peracchi, A.; Reinhoudt, D. N.; Salvio, R.; Sartori, A.; Ungaro, R. *J. Am. Chem. Soc.* **2007**, *129* (41), 12512.

(28) Baldini, L.; Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Salvio, R.; Sansone, F.; Ungaro, R. *J. Org. Chem.* **2012**, *77* (7), 3381.

(29) Salvio, R.; Cacciapaglia, R.; Mandolini, L.; Sansone, F.; Casnati, A. *RSC Adv.* **2014**, *4* (65), 34412.

(30) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Sansone, F. *J. Org. Chem.* **2015**, *80* (11), 5887.

(31) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Sansone, F.; Mandolini, L.; Casnati, A. *J. Org. Chem.* **2016**, *81* (11), 4728.

(32) Arduini, A.; Fanni, S.; Manfredi, G.; Pochini, A.; Ungaro, R.; Sicun, A. R.; Ugozzoli, F. *J. Org. Chem.* **1995**, *60* (5), 1448.

(33) Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Reinhoudt, D. N.; Salvio, R.; Sartori, A.; Ungaro, R. *J. Org. Chem.* **2005**, *70* (2), 624.

(34) Feichtinger, K.; Sings, H. L.; Baker, T. J.; Matthews, K.; Goodman, M. *J. Org. Chem.* **1998**, *63* (23), 8432.

(35) Veiga, A. X.; Arenz, S.; Erdélyi, M. *Synth.* **2013**, *45* (6), 777.

(36) Corona-Martínez, D. O.; Taran, O.; Yatsimirsky, A. K. *Org. Biomol. Chem.* **2010**, *8* (4), 873.

(37) Georgieva, M.; Velinov, G.; Budevsky, O. *Anal. Chim. Acta* **1977**, *90* (Supplement C), 83.

(38) Wróbel, R.; Chmurzy, L. **2000**, *405*, 303.

(39) Ariga, K.; Anslyn, E. V. *J. Org. Chem.* **1992**, *57* (2), 417.

(40) Baughman, E. H.; Kreevoy, M. M. *J. Phys. Chem.* **1974**, *78* (4), 421.

(41) Salvio, R.; Cacciapaglia, R.; Mandolini, L. *J. Org. Chem.* **2011**, *76* (13), 5438.

(42) Savelli, C.; Salvio, R. *Chem. Eur. J.* **2015**, *21* (15), 5856.

(43) Salvio, R.; Cincotti, A. *RSC Adv.* **2014**, *4* (54), 28678.

(44) Hengge, A. C.; Cleland, W. W. *J. Org. Chem.* **1991**, *56* (5), 1972.

(45) Aguilar-Pérez, F.; Gómez-Tagle, P.; Collado-Fregoso, E.; Yatsimirsky, A. K. *Inorg. Chem.* **2006**, *45* (23), 9502.

(46) Ruiz Kubli, M.; Yatsimirsky, A. K. *Inorganica Chim. Acta* **2016**, *440*, 9.

(47) Livieri, M.; Mancin, F.; Saielli, G.; Chin, J.; Tonellato, U. *Chem. Eur. J.* **2007**, *13* (8), 2246.

(48) Brown, D. M.; Usher, D. A. *J. Chem. Soc.* **1965**, No. 0, 6558.

(49) Luxon, S. G. *Hazards in the Chemical Laboratory*, 5th ed.; Chemistry, T. R. S. of, Ed.; Cambridge, UK, 1992.

(50) Gans, P.; Sabatini, A.; Vacca, A. *Talanta* **1996**, *43* (10), 1739.

(51) Alderighi, L.; Gans, P.; Ienco, A.; Peters, D.; Sabatini, A.; Vacca, A. *Coord. Chem. Rev.* **1999**, *184* (1), 311.

(52) Baldini, L.; Melegari, M.; Bagnacani, V.; Casnati, A.; Dalcanale, E.; Sansone, F.; Ungaro, R. *J. Org. Chem.* **2011**, *76* (10), 3720.

Chapter 3

A calix[4]arene – based DNA Topoisomerase I mimic for the promotion of phosphoryl transfer processes

3.1 Introduction

The presence of arginine residues in the active site of phosphodiesterase enzymes is often crucial to allow the accomplishment of their biological role[1–4]. The guanidinium group at the end of this amino acid side chain has been in fact reported to assist the action of metal cations[1,2] or other amino acid residues[3,4] in the cleavage of the target substrates, by providing stabilization of the transition state through electrostatic interactions with the oxygen atoms of the scissile phosphate group. Accordingly, both metal-free[5–8] and metal-containing[5,9–12] artificial catalysts, showing remarkable phosphodiesterase activity, have been developed by inserting guanidinium units on different types of molecular scaffolds, including the *cone*-calix[4]arene macrocyclic cavity[13–17]. However, still little explored is the possibility to create models for the human DNA Topoisomerase I[18], an enzyme able to cleave single strands of DNA by exploiting the nucleophilic attack of a tyrosine residue on specific phosphate groups of the sequence, together with the establishment of electrostatic interactions between two neighboring arginine side chains and the phosphate group (**Chapter 1**, **Paragraph 1.3.1**)[3,4]. Thus, with the aim to mimic the catalytic triad present in the active site of this enzyme, it was decided to synthetize compound **1**H·2HCl, by functionalizing the upper rim of a *cone*-calix[4]arene with a phenolic hydroxyl group and two guanidinium units. This new artificial phosphodiesterase was then tested in the cleavage of the DNA model compound BNPP (bis-*p*-nitrophenyl phosphate) and the results of the corresponding kinetic investigations, which were the object of also of a joint publication with the group of Salvio and Mandolini in Rome[19], will be reported in this chapter.

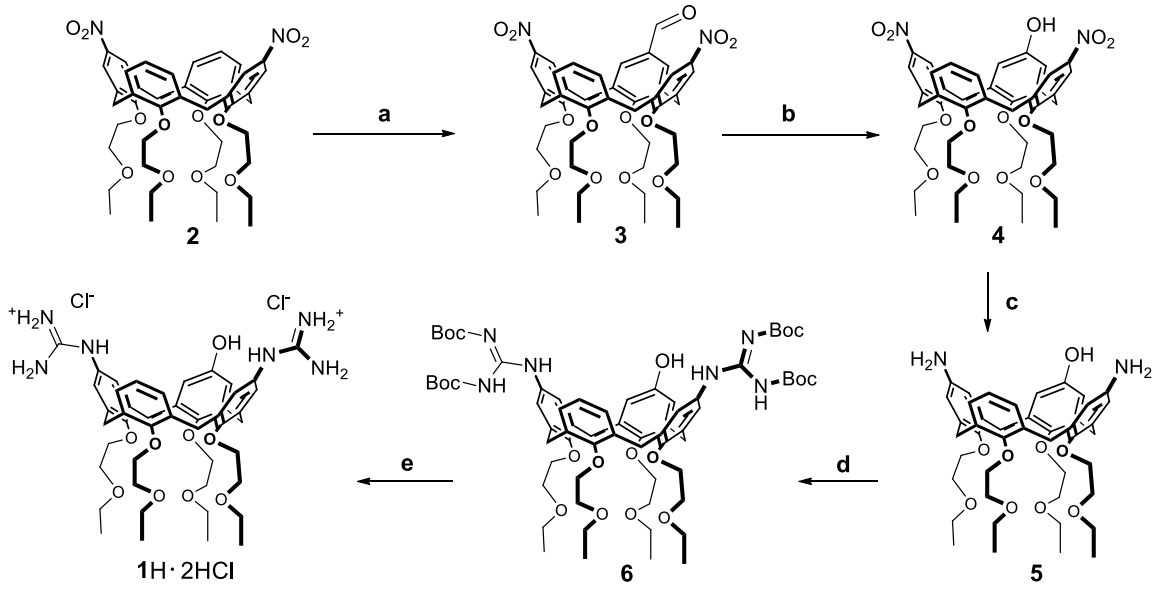
3.2 Results and discussion

3.2.1 Calix[4]arene synthesis

Compound **1**H·2HCl was obtained starting from the dinitrocalix[4]arene **2**[20], according to the synthetic pathway reported in **Scheme 1**.

The first reaction step consisted in the insertion of a formyl moiety on **2** to give **3**, by modifying a protocol already known in literature to perform Gross formylation on calix[4]arene derivatives[21]. The desired product was isolated from a mixture containing also the diformyl compound, in an amount exceeding the expected statistical distribution (59% yield). However, the yield value obtained for the synthesis of **3** was lower than the 84% one reported for the obtainment, under similar conditions, of an analogous dinitro-monoformyl derivative of a tetrapropoxy-calix[4]arene[22]. To explain these results, it was suggested that the four ethoxyethyl chains at the lower rim of **2** can complex the Sn(IV) ions used to provide Lewis acid catalysis, thus decreasing the reactivity toward electrophilic aromatic substitution at its upper rim. Subsequently, compound **3** was submitted to Baeyer-Villiger oxidation[23] to convert its formyl unit into the corresponding formate ester, which was readily hydrolyzed under basic conditions

to give the desired dinitro-hydroxy calix[4]arene **4** (81% overall yield from **3**). Then, the two nitro groups of **4** were quantitatively reduced to obtain **5**[24], in turn converted in compound **6** (42% yield) by guanidinylation of the amino groups at the upper rim with [N,N'–bis(Boc)]thiourea and $HgCl_2$, as already reported in literature for analogous calix[4]arene derivatives[13,15]. Target compound **1**H·2HCl was finally obtained by removal of the Boc protecting groups and exchange of the trifluoroacetate with chloride as counterions of the guanidinium units (93% yield)[16,17].



Scheme 1: a) $SnCl_4$, Cl_2CHOCH_3; dry $CHCl_3$, -10° C. b) m-CPBA; DCM and then NaOH 2M; MeOH-H_2O 4:1. c) $NaBH_4$, $NiCl_2 \cdot 6H_2O$; MeOH, 65° C. d) [N,N'–bis(Boc)]thiourea, $HgCl_2$, NEt_3; dry DMF. e) TFA, TES; DCM and then HCl 1M; EtOH.

HR-MS, ^{1}H- and ^{13}C NMR spectra are fully consistent with the structure of all the synthetized calix[4]arene derivatives. To demonstrate the presence of ABAH trifunctionalized compounds, in the ^{1}H NMR spectra, the pattern of signals observed for the Ar*CH*$_2$Ar and Ar*H* resonances are particularly diagnostic (see **Figure 1** for the case of **1**H·2HCl). In fact, both the axial and the equatorial protons of the aromatic methylene bridges give rise to two doublets of equal intensity, according to the substitution pattern created at the upper rim. Moreover, for all the compounds excluding **3**, the signals of the aromatic protons adjacent to the two A substituents in 1,3 distal-position (A= nitro, amino or guanidinium groups depending on the considered calix[4]arene derivative) are split in two singlets or two doublets when the resolution was sufficient to see the meta J coupling, since they were differently shielded by the electron density of the B- (B= -OH) and H-functionalized vicinal aromatic rings.

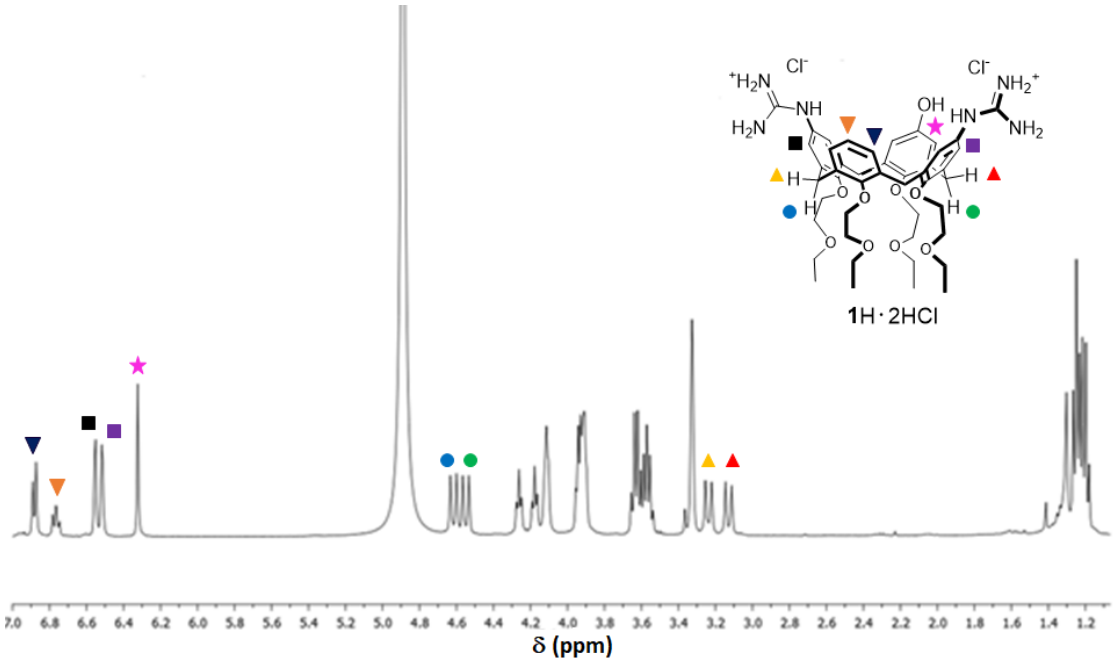


*Figure 1: ^{1}H NMR spectrum of compound **1**H·2HCl (400 MHz, CD_3OD). The pattern of signals belonging to the $ArCH_2Ar$ and to the aromatic protons are diagnostic to demonstrate the occurrence of a ABAH trifunctional substitution at the upper rim.*

3.2.2 Acid-base titrations

Potentiometric and UV-VIS titrations were performed on **1**H·2HCl, which will be hereafter indicated as $(\mathbf{1}H_3)^{2+}$ for simplicity, to investigate its acid-base properties before its use as artificial phosphodiesterase. The experiments were carried out in DMSO/water 8/2 (80% DMSO) which, as already reported in **Paragraph 2.2.2**, is a solvent mixture suitable to carry out the pK_a determination[25,26] and kinetic studies for the hydrolytic cleavage of phosphate diesters[5,13–17,27]. For a reliable treatment of the experimental data it is important to remember that in this medium the acidity constant for the autoprotolysis of water (K_w) is increased to 18.4[28], thus implying that neutrality is reached in solutions with pH equal to 9.2.

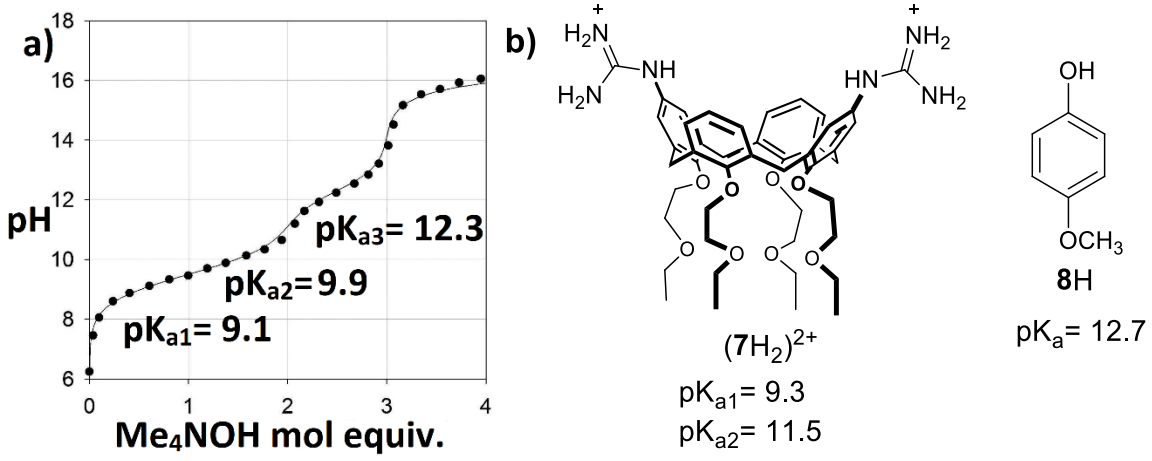


*Figure 2: **a)** Titration of 2 mM $(\mathbf{1}H_3)^{2+}$ with Me_4NOH (80% DMSO, 25 °C, 10 mM Me_4NClO_4) and **b)** pK_a values determined for model compounds $(\mathbf{7}H_2)^{2+}$ (**Ref. 13**) and **8**H under the same experimental conditions.*

Potentiometric titrations (**Figure 2a**) pointed out, as expected, the presence of three titratable protons in the tested calix[4]arene, whose identification was not trivial without the support of other investigations. In fact, the comparison of the pK_a values reported for its parent compound (**7**H_2)$^{2+}$ (**Figure 2b**)[13] seems to indicate that the most acidic proton of (**1**H_3)$^{2+}$ belongs to one of the two guanidinium units, whose acidity were suggested to be not significantly affected by the presence of an additional OH-group at the upper rim. However, the possibility that the same pK_a has to be assigned to the phenolic hydroxyl was not excluded in principle, since it is expected to be much more acidic than that of the model compound **8**H (**Figure 2b**), thanks to the electrostatic stabilization of the resulting phenolate anion afforded by the neighbouring guanidinium dyad. To clarify this question, it was decided to perform a UV-VIS titration in which the absorbance change at λ= 320 nm as function of the increasing concentration of Me_4NOH was used to monitor the deprotonation of the phenolic -OH moiety of both (**1**H_3)$^{2+}$ and **8**H (**Figure 3a** and **3b**, respectively).

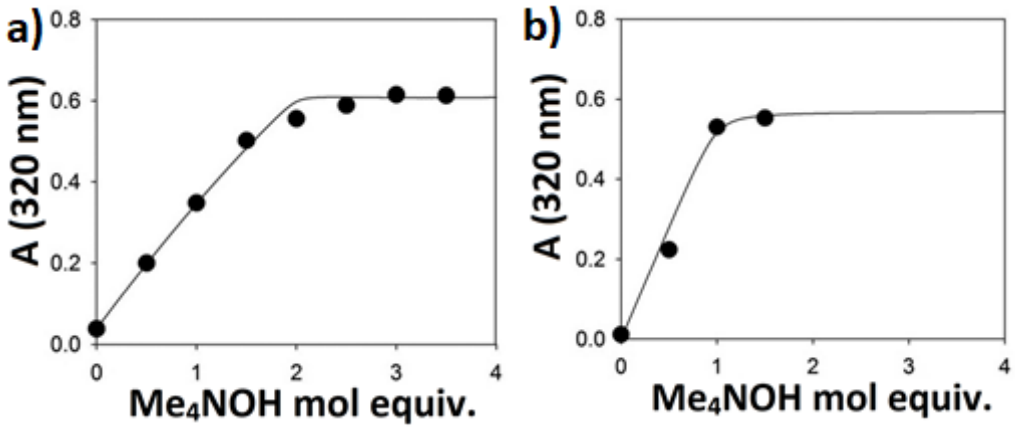


*Figure 3: UV-Vis titration of **a)** 0.2 mM calix[4]arene (**1**H_3)$^{2+}$ and **b)** 0.2mM **8**H with Me_4NOH (DMSO 80%, 25°C).*

The experiment pointed out that after the addition of 1mol equiv. of base only 55% of the phenolic hydroxyl of (**1**H_3)$^{2+}$ is converted into the corresponding phenolate form, contrarily with what observed for its monomeric control compound (**8**H), which is instead fully deprotonated under the same condition. The proton removal process was completed only in the presence of a second equivalent of Me_4NOH, thus indicating that the OH-unit and one of the two guanidinium groups at the upper rim participate with similar contributions to the first two deprotonation steps of (**1**H_3)$^{2+}$ (pK_{a1}= 9.1 and pK_{a1}= 9.9). As expected, these observations indicated also that the phenolate anion acts as an innocent spectator in the neutralization of the least acidic proton (pK_{a3}= 12.3) by a third equivalent of base, which involves only, or very nearly so, the remaining guanidinium moiety.

Accordingly, the combination of the potentiometric and UV-VIS data allows to identify the various protonation species of compound **1** as function of the pH, which were visualized in the distribution diagram calculated on the basis of the given pK_a values (**Figure 4a**). More in details, it is possible to state without ambiguity that the bis- and monoprotonated form of the tested calix[4]arene (**1**H_2)$^+$ and **1**H are

represented, respectively, by almost equal amounts of the two tautomers A and B and by the zwitterionic derivative C shown in **Figure 4b**.

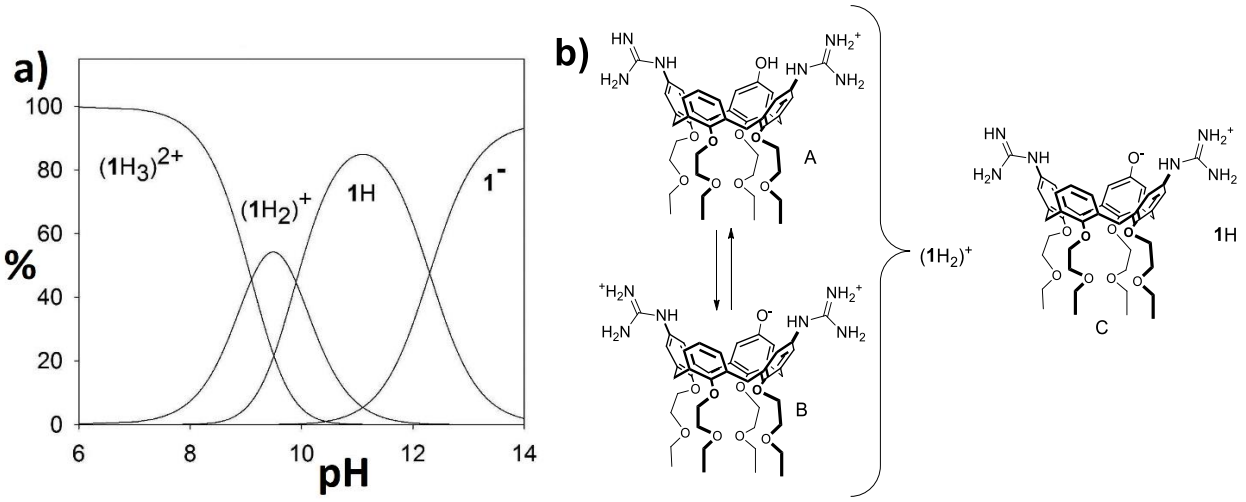


*Figure 4: **a)** Distribution diagram of the species of compound **1** as a function of pH and **b)** structure of its bis- and monoprotonated forms $(\mathbf{1}H_2)^+$ and **1**H.*

3.2.3 BNPP cleavage

3.2.3.1 Initial rate experiments

Kinetic investigations were performed for the cleavage of the DNA model compound BNPP (bis-*p*-nitrophenyl phosphate) by $(\mathbf{1}H_3)^{2+}$ in 80% DMSO (**Figure 5**), buffered in the pH range 8.5-12.0 by the presence of *N,N'*-diisopropyl ethanolamine-perchlorate salt.

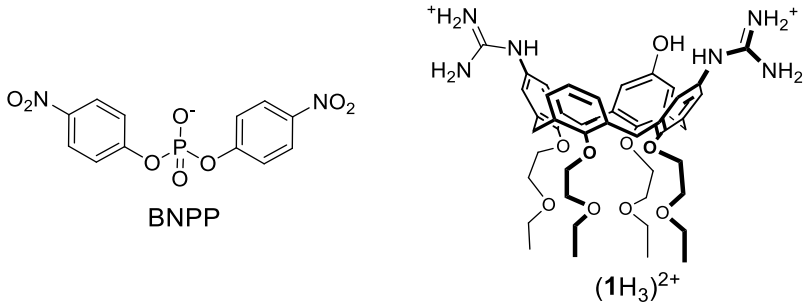


Figure 5: DNA model compound used as substrate for the human DNA Topoisomerase mimic $(\mathbf{1}H_3)^{2+}$.

The pseudo-first-order rate values for the cleavage of BNPP were calculated as $k_{obs} = v_0/[BNPP]$, where v_0 is the spectrophotometrically determined initial rate of *p*-nitrophenol (pNPOH) liberation from the substrate. The total concentration of released pNPOH was extrapolated from calibration curves obtained, at the given pH values, by plotting the absorbance at 400 nm of solutions at increasing *p*-nitrophenol concentration (see **Paragraph 2.2.3**, **Figure 3b**). The experimental data (**Table 1**, entries 1-7), pointed out that, in the given pH range, the tested compound is able to accelerate the studied process from 10^5- to 10^2-folds, with respect to the background reaction. The catalytic activity of $(\mathbf{1}H_3)^{2+}$ is significantly higher than that reported for the control compounds $(\mathbf{7}H_2)^{2+}$ and **8**H in the cleavage of the

same substrate (**Table 1**, entries 8 and 9, respectively), thus indicating a good degree of cooperation between the three catalytic units present at the upper rim. In fact, at the given pH values, the monofunctional model compound **8**H is not able to induce significant acceleration over the background for the hydrolysis of BNPP, while the bis-guanidinium calix[4]arene (**7**H_2)$^{2+}$, lacking the penolic -OH group, is found to catalyze this reaction 42-folds slower than the studied trifunctional derivative.

*Table 1: Cleavage of BNPP in the presence of different additives, (80% DMSO, 50 °C). The reaction mixtures were composed of 0.20 mM BNPP, 1.0 mM additive, 0.10 M N,N'-diisopropyl ethanolamine buffer, 10 mM Me_4NClO_4 solutions. [a] k_{obs} were calculated as v_o/[BNPP], with maximum error = ±10%. [b] The background rate constant (k_{bg}, s^{-1}) for the hydroxide-catalyzed cleavage has been calculated by the expression $k_{bg} = 10^{(pH-18.67)}$ (see **Ref. 16**). [c] Corrected for the contribution of the background reaction OH^- + BNPP: $k_{bg} = 1.1 \times 10^{-6} \cdot s^{-1}$ at pH 12.7.*

Entry	Tested additive	pH	$10^6 \cdot k_{obs}$ (s^{-1}) [a]	k_{obs} / k_{bg} [b]
1	**1**H·2HCl	8.5	8.8	$1.3 \cdot 10^5$
2		9.0	23	$1.1 \cdot 10^5$
3		9.5	44	$6.5 \cdot 10^4$
4		10.0	42	$1.9 \cdot 10^4$
5		10.5	35	$5.1 \cdot 10^3$
6		11.1	26	$9.6 \cdot 10^2$
7		12.0	19	$8.8 \cdot 10$
8	**7**·2HCl	10.4	0.64	$1.2 \cdot 10^2$
9	**9**H	12.7	0.50 [c]	

Other information about the extent of the cooperation between the reactive groups of (**1**H_3)$^{2+}$ were obtained by studying the plot of the k_{obs} values reported in **Table 1** as function of the pH. In fact, the corresponding plot of the data point resulted in a nonsymmetrical bell-shaped profile, consistent with the presence of more than one catalytically active species in the tested conditions (**Figure 6a**). According to the distribution diagram in **Figure 4a**, it was suggested that the active species of the studied calix[4]arene are its bis- and monoprotonated forms (**1**H_2)$^+$ and **1**H, whose pH dependent concentrations were calculated from standard equations for a triprotic acid (**Eqn. 1** and **Eqn.2**, respectively), where C_{tot} is the initial concentration of (**1**H_3)$^{2+}$ and K_{a1}, K_{a2} and K_{a3} its experimental acidity constants.

$$[(\mathbf{1}H_2)^+] = \frac{C_{tot}[H^+]^2 K_{a1}}{[H^+]^3 + [H^+]^2 K_{a1} + [H^+] K_{a1} K_{a2} + K_{a1} K_{a2} K_{a3}} \quad \textbf{Eqn. 1}$$

$$[\mathbf{1}H] = \frac{C_{tot}[H^+] K_{a1} K_{a2}}{[H^+]^3 + [H^+]^2 K_{a1} + [H^+] K_{a1} K_{a2} + K_{a1} K_{a2} K_{a3}} \quad \textbf{Eqn. 2}$$

The resulting concentration values were then inserted as known quantities into **Eqn. 3** to fit the given pH-rate profile, using k_1 and k_2 as adjustable parameters to estimate the individual contribution of the two species to the overall catalytic activity of compound $(\mathbf{1}H_3)^{2+}$.

$$k_{obs} = k_1[(\mathbf{1}H_2)^+] + k_2[\mathbf{1}H] \quad \textbf{Eqn. 3}$$

A non-linear least-squares treatment of the data points gave as best fit values $k_1 = (6.2 \pm 0.4)\cdot10^{-2}$ $M^{-1}s^{-1}$ and $k_2 = (2.8 \pm 0.2)\cdot10^{-2}$ $M^{-1}s^{-1}$, indicating that the bis-protonated form of the tested calix[4]arene $(\mathbf{1}H_2)^+$ is twice as active than its neutral species **1**H. Taking into account that nearly a half of $(\mathbf{1}H_2)^+$ is in the catalytically-active-phenolate form, it turned out that tautomer B (**Figure 4b**), featuring two positively charged guanidinium units, is about four times more effective than **1**H, where only a guanidinium group is protonated. Thus, it appears clear that the participation of both the guanidinium units to the catalytic process is crucial to improve the phosphodiesterase activity of the tested compound.

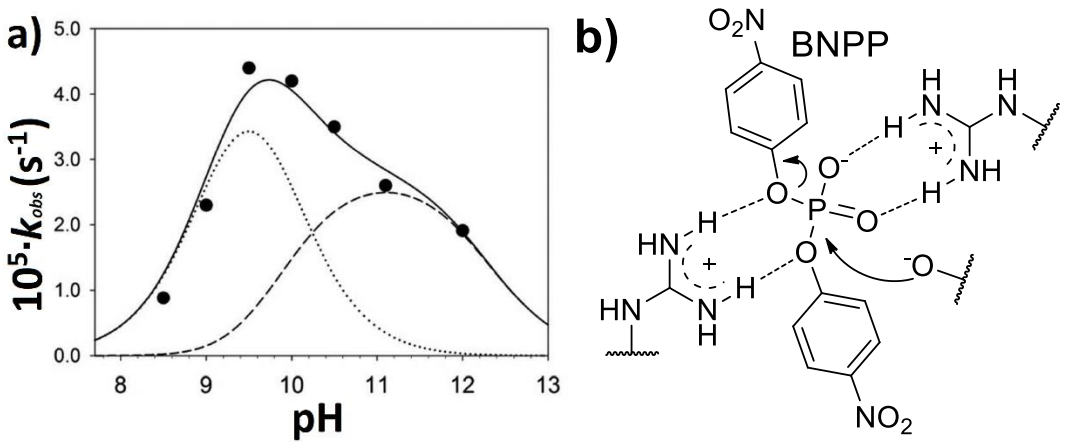


*Figure 6: **a)** pH–rate profiles for the cleavage of 0.20 mM BNPP in the presence of 1.0 mM $(\mathbf{1}H_3)^{2+}$, 0.10 M N,N-diisopropyl ethanolamine buffer, 10 mM Me_4NClO_4, (data points from **Table 1**). The solid line is the plot of **Eqn. 3** with best fit values of k_1 and k_2. The dotted and dashed lines are the plots of the individual contribution of the species $(\mathbf{1}H_2)^+$ and **1**H, respectively, to the overall reactivity. **b)** Suggested mechanism of BNPP cleavage promoted by $(\mathbf{1}H_2)^+$ involving two guanidinium units as electrophilic activators and a phenolate moiety acting as a nucleophile.*

The experimental data in our possess are hence consistent with the suggestion that the three catalytic units of the studied calix[4]arene act synergistically into the cleavage of BNPP, which presumably occurs by nucleophilic attack of the deprotonated phenolic hydroxyl moiety on the phosphate group of the substrate, assisted by the electrostatic stabilization of the negatively charged transition state by at least one of the two guanidinium units (**Figure 6b**).

3.2.3.2 Time course experiments

The release of *p*-nitrophenol (pNPOH) from BNPP was monitored by UV-VIS spectrometry in time-course experiments under substrate-excess conditions, with the aim to obtain more information about the ability of compound $(\mathbf{1}H_2)^+$ in the promotion of phosphoryl transfer processes. The experiments were carried out in the presence of 0.1 mM calix[4]arene and 5.0 mM BNPP, buffering the reaction mixture at

pH 9.5, in order to guarantee the maximum absolute rate of the substrate cleavage (see **Table 1**, entry 3).

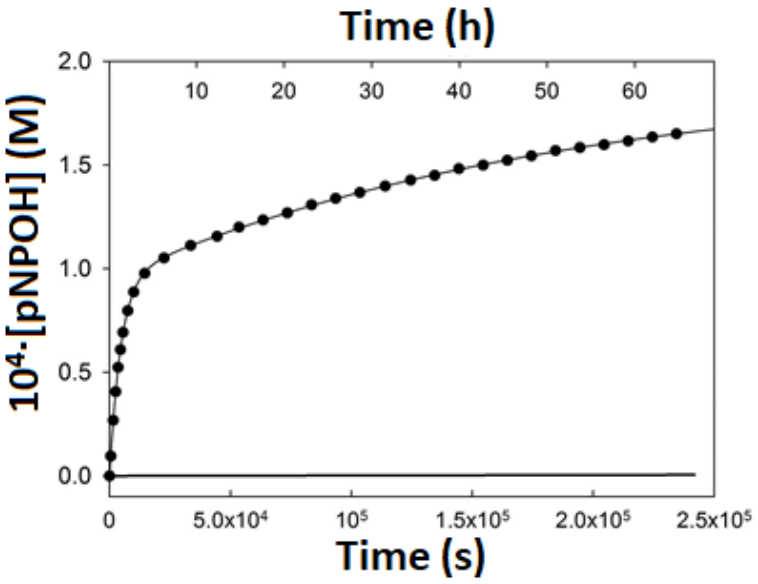


*Figure 7: Liberation of pNPOH in a 5.0 mM solution of BNPP upon addition of 0.10 mM calixarene (**1**H$_2$)$^+$ (10 mM Me$_4$NClO$_4$, 0.10 M N,N-diisopropyl ethanolamine buffer, pH 9.5; 80% DMSO, 50° C). Data points are experimental and the solid line is the plot of **Eqn. 4** with best fit parameters k' = 2.0 × 10^{-4} s^{-1}, and k ''= 5.4 × 10^{-6} s^{-1}. The solid line at the bottom corresponds to background pNPOH liberation calculated at pH 9.5.*

The pNPOH release took place with an initial burst followed by a much slower phase (**Figure 7**), suggesting that the cleavage of BNPP was occurring according to a kinetic scheme composed by two consecutive irreversible first-order reactions[29]. A prediction of the time-dependent concentration of liberated pNPOH was obtained by using **Equation 4**, yielded after adaptation of standard integrated equations to the investigated reaction system, where C_{tot} is the total concentration of the calix[4]arene derivative, k' = k[BNPP] and $\tau = (k'+k'')^{-1}$.

$$[\mathrm{pNPOH}]=C_{\mathrm{tot}}\,\tau\,k'\left[\tau\,k'\left(1-e^{-\frac{t}{\tau}}\right)+(1-e^{-k''t})\right] \quad \textbf{Eqn. 4}$$

This equation can also be rearranged into a simpler form to give **Eqn. 5** once the first exponential term dies out and the second exponential decay becomes apparent.

$$[\mathrm{pNPOH}]=C_{\mathrm{tot}}\,\tau^2k'^2+C_{\mathrm{tot}}\left(1-e^{-k''t}\right) \quad \textbf{Eqn. 5}$$

A nonlinear least-squares procedure to fit the experimental data to **Eqn 4** gave k' = (2.0 ± 0.2)·10^{-4} s^{-1} and k''= (5.4 ± 0.2)·10^{-6} s^{-1} as best values for the fitting parameters. Interestingly, the pseudo-first-order rate constants k' and the k_{obs} measured at the same pH but with different reactant concentrations (**Table 1**, entry 3) corresponds to second order rate constants of 4.0·10^{-2} M^{-1} s^{-1} and of 4.4·10^{-2} M^{-1} s^{-1}, respectively. The good agreement of the two values indicates that, as previously reported in other cases[13,16,30], the association between (**1**H$_2$)$^+$ and the substrate is too weak to affect the kinetics, since the

rate constants are concentration-independent quantities. In other words, in the tested conditions, the studied calix[4]arene works under sub-saturating conditions and it selectively use all of the available binding energy arising from the interaction of its guanidinium unit(s) with the scissile phosphate group to stabilize the transition state, rather than to interact with the reactant state.

These additional experimental data are thus in agreement with a reaction sequence featuring a fast step of phosphoryl transfer from BNPP to the nucleophilic phenolate moiety of (**1**H_2)$^+$, with the release of the first molar equivalent of pNPOH to form **9**, which subsequently slowly converts into **10** by liberation of a second equivalent of pNPOH, presumably thanks to the assistance of the neighbouring guanidine/guanidinium catalytic dyad (**Figure 8**)[7,6,30].

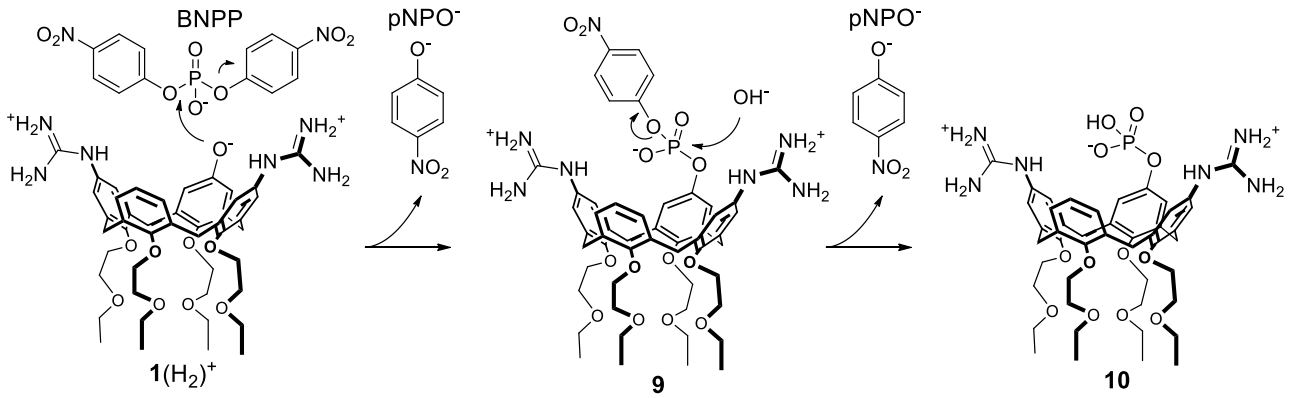


*Figure 8: Sequential two-steps mechanism proposed for the cleavage of BNPP by (**1**H_2)$^+$.*

The occurrence of this reactive pathway was further supported by ES-MS analysis of aliquots of the reaction mixtures after different times of incubation. In fact, the spectrum of a sample withdrawn after 15 min showed the formation of the O-phosphorylated derivative **9**, whose peak was detected with higher intensity in an aliquot withdrawn after 3 days, together with the m/z signal belonging to calix[4]arene **10** (**Figure 9a** and **9b**, respectively). On the other hand, the complete dephosphorylation of this latter intermediate was not observed in the explored time range, indicating that the regeneration of the starting compound (**1**H_2)$^+$ was not occurring. Thus, turnover is not ensured and it is consequently more appropriate to describe the tested calix[4]arene as a promoter of phosphoryl transfer processes rather than as a catalysts.

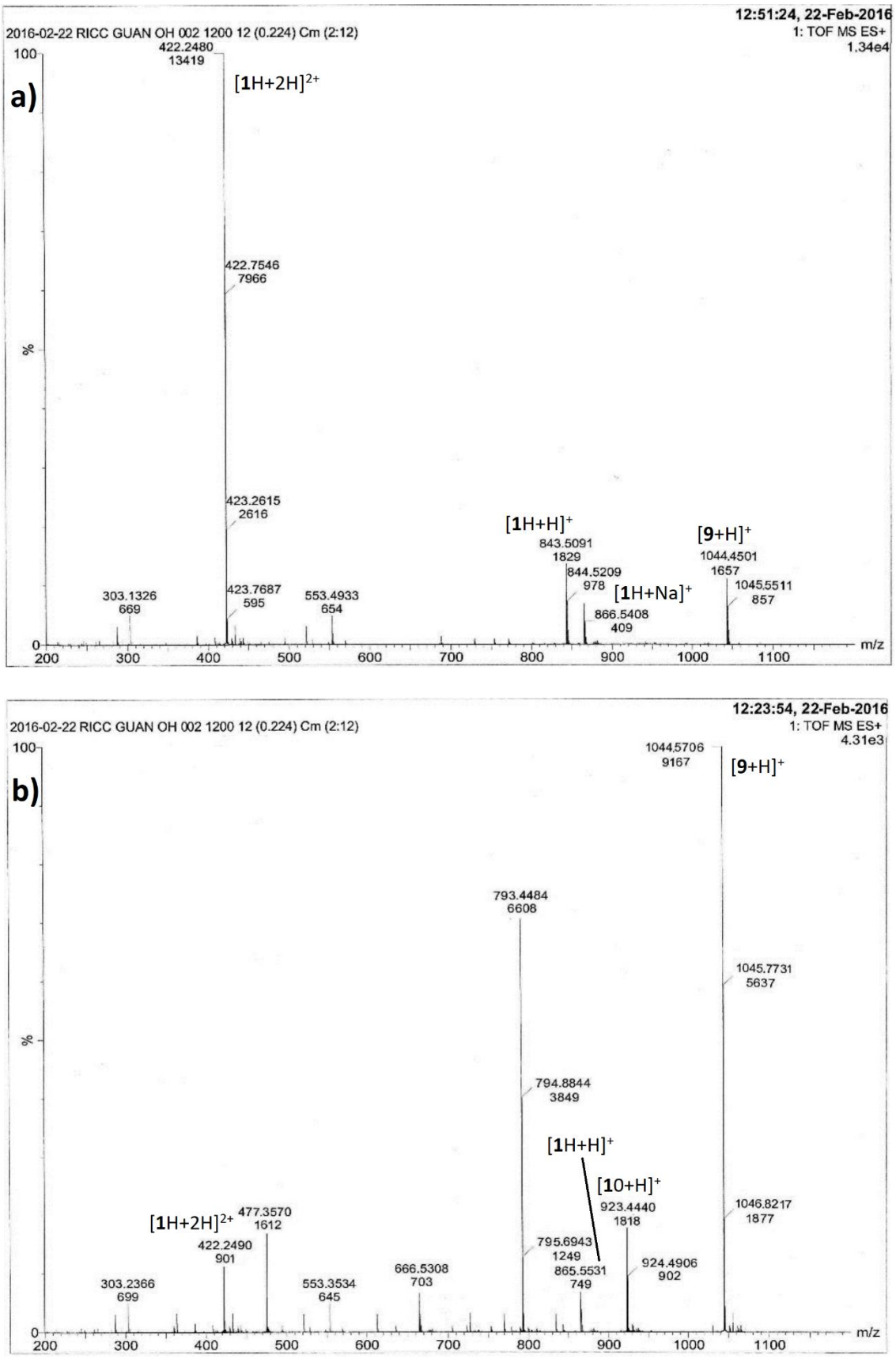


*Figure 9: ESI-MS spectra after **a)** 15 min and **b)** 3 days of the reaction mixture of the experiment time-course experiments for the cleavage of BNPP. The analysis confirms the presence of the* p*-nitrophenyl phosphoryl derivative **9** (m/z=1044) and of the phosphoryl derivative of the calix[4]arene **11** (m/z=923), according to the mechanism in **Figure 8**.*

3.3 Conclusions

In this chapter it has been presented the synthesis and the investigations of the catalytic activity of the trifunctional calix[4]arene (**1**H_3)$^{2+}$, functionalized at the upper rim with a phenolic –OH and two guanidinium units to mimic the catalytic triad at the active site of the human Topoisomerase I enzyme.
The studied compound was found to be effective in the cleavage of the DNA model compound BNPP, whose hydrolysis was accelerated respect to the background by 2 to 5 orders of magnitude, under weakly acid or moderately basic conditions (pH 8.5-12 in 80% DMSO). The superiority of (**1**H_3)$^{2+}$ over its control compounds (**7**H_2)$^{2+}$ and **8**H, lacking respectively the phenolic hydroxyl and the guanidinium dyad, indicates the presence of cooperation between the three catalytic units at the upper rim. Moreover, the data collected from kinetic investigations about the pH-dependent initial rate liberation of pNPOH from the substrate, together with the results given by UV-VIS and potentiometric titrations, clearly pointed out that the presence of two catalytically active species for the tested calix[4]arene derivative, namely its bis- and monoprotonated forms of (**1**H_2)$^+$ and **1**H. More in details, it was found that the tautomeric form B (**Figure 4b**) is four time more active than the neutral calix[4]arene **1**H in the cleavage of the substrate, presumably thanks to the double electrostatic stabilization of the transition state afforded by the two positively charged guanidinium units, to assist the nucleophilic attack of the deprotonated phenolate on the scissile phosphate group.
Time-course kinetic investigations, in combination with ES-MS analysis of periodically withdrawn aliquots of the reaction mixtures, evidenced that the cleavage of BNPP by (**1**H_2)$^+$ occurs in two steps, with the overall release of two mol equiv. of pNPOH. The first step, involving a phosphoryl transfer process to the phenolate moiety of (**1**H_2)$^+$ to form **9**, was found to be much faster than the second, in which compound **10** is given by liberation of the second equivalent of pNPOH form the resulting phosphorylated intermediate, probably thanks to the electrophilic activation provided by the neighbouring guanidinium dyad.
In the investigated timescale, compound **10** was inert toward dephosphorylation, thus impairing the turnover of the starting active species (**1**H_2)$^+$. The low efficiency of this step prevents the complete mimicking of the catalytic cycle of DNA topoisomerase I and it has to be further improved to develop more effective mimics of this enzyme.

3.4 Experimental section

3.4.1 General information

Synthesis of compounds **3** and **6** were carried out under nitrogen atmosphere, using previously oven-dried glassware. All dry solvents were prepared according to standard procedures, distilled before use and stored over 3 or 4 Å molecular sieves. Analytical TLC were performed using prepared plates of silica

gel (Merck 60 F-254) and then revealed with UV light. 2,4-dinitrophenylhydrazine (0.1 M in H_2SO_4/H_2O/EtOH 1:2:4), $FeCl_3$ (1% in MeOH/H_2O 1:1) and Ninhydrin (5% in EtOH) TLC stains where also used to confirm the presence of, respectively, free formyl, penolic hydroxyl or amino groups on the studied compounds. Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates. DMSO, purged 30 min with Argon, and mQ water were used in the preparation of 80% DMSO used for kinetic and potentiometric experiments. The bis(hydrochloride) diguanidino derivative **7**·2HCl, was prepared according to literature procedures[13]. All other solvents and reagents were commercial samples and used as such.

Care was taken when handling tetramethylammonium perchlorate because it is potentially explosive[31]. No accident occurred in the course of the present work.

^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^{1}H nucleus at 300 MHz and 400 MHz, respectively, and of ^{13}C nucleus at 75 MHz and 100 MHz, respectively) and partially deuterated solvents were used as internal standards to calculate the chemical shifts (δ values in ppm). All ^{13}C NMR spectra were performed with proton decoupling. High resolution electrospray ionization (ESI) mass analysis were performed with a Waters single-quadrupole spectrometer in positive mode using MeOH or CH_3CN as solvents. Melting points were determined on an Electrothermal apparatus in closed capillaries. Potentiometric titrations were performed by an automatic titrator equipped with a pH electrode. Spectrophotometric measurements of *p*-nitrophenol liberation were carried out at 400 nm on a double beam spectrophotometer. HPLC analyses were performed on a liquid chromatograph equipped with a UV-vis detector operating at 254 nm.

3.4.2 Acid-base titrations

Potentiometric titrations were carried out according to a previously reported procedure[13]. A freshly prepared solution of 50 or 100 mM Me_4NOH in 80% DMSO was added in small increments under an argon atmosphere a 2.0 - 3.0 mM solution (5mL, 80% DMSO, 25°C) of **1**H·2HCl or **8**H, using 10 mM Me_4NClO_4 as ionic strength buffer. The electrode calibration was performed as reported in literature[9]. Elaboration of the titration plots was carried out with the software HYPERQUAD 2000[32,33]. UV-Vis spectrophotometric titrations of a 0.2 mM solutions **1**H·2HCl or **8**H were carried out analogously under argon atmosphere, by incremental addition of Me_4NOH (DMSO 80%, 25 °C).

3.4.3 Kinetic measurements

The liberation of pNPOH from the substrate was monitored by UV-VIS spectrometry at 400 nm. Initial rate experiments were performed on reaction mixtures composed of 0.20 mM BNPP, 1.0 mM additive ((**1**H_3)$^{2+}$ or (**7**H_2)$^{2+}$ or **8**H), 0.10 M *N*,*N*-diisopropyl ethanolamine buffer, 10 mM Me_4NClO_4 (80% DMSO, 50.0 °C). The total concentration of released *p*-nitrophenol was extrapolated from calibration curves built at the required pH values by measuring the absorbance at 400 nm of known aliquots of pNPOH

dissolved in the same medium. The pH of the solution was adjusted at the selected pH value with a 50−100 mM solution of $HClO_4$ in 80% DMSO.

Time-course experiment was carried out in the presence of 5.0 mM BNPP, 0.1 mM $(\mathbf{1}H_3)^{2+}$, 10 mM Me_4NClO_4, 0.10 M *N*,*N*-diisopropyl ethanolamine buffer solution (80% DMSO, 50.0 °C). The pH of the solution was adjusted to pH 9.5 by addition of the proper volume of the 50 – 100 mM solution of $HClO_4$ in 80% DMSO. Non-linear least-square fit of experimental data to **Eqn. 4** was carried out with the software SigmaPlot 12.0 (Systat Software, Inc.).

3.4.4 Calix[4]arene synthesis

Compound **1**H·2HCl was prepared starting from the 5,17-dinitro-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene **2**, in turn synthetized according to literature procedures[20].

5,17-Dinitro-11-formyl-25,26,27,28-tetrakis(2-ethoxyethoxy) calix[4]arene (3).

To a solution of **2** (0.22 g, 0.27 mmol) in dry $CHCl_3$ (20 mL) cooled at -10°C, $SnCl_4$ (1.6 mL, 13.92 mmol) and Cl_2CHOCH_3 (1.26 mL, 13.81 mmol) were added. The reaction mixture was stirred for 2 hours at -10°C, then it was quenched with distilled water (30 mL) and vigorously stirred for additional 30 minutes. The organic layer was washed with a saturated solution of $NaHCO_3$ (2 x 20 mL) and brine (20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude material was purified by flash chromatography (hexane/AcOEt 5.5:4.5 – hexane/AcOEt 1:1) to give **3** as a white solid (0.13 g, 0.16 mmol; 59% yield): mp 65–67 °C. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 9.61 (s, 1H, ArC*H*O), 7.75 (s, 4H, Ar*H*-meta-NO_2); 7.03 (s, 2H, Ar*H*-meta-CHO); 6.48 (m, 3H, Ar*H*-meta and Ar*H*-para); 4.70 (d, 2H, *J* =12.6 Hz, Ar*CH*$_2$Ar ax.); 4.61 (d, 2H, *J* =13.8 Hz, Ar*CH*$_2$Ar ax.); 4.35-4.26 (m, 4H, ArO*CH*$_2$CH$_2$O); 4.17 (t, 2H, *J*=4.8 Hz, ArO*CH*$_2$CH$_2$O); 4.07 (t, 2H, *J*=4.8 Hz, ArO*CH*$_2$CH$_2$O); 3.85-3.76 (m, 8H, ArOCH$_2$*CH*$_2$O); 3.55-3.47 (m, 8H, O*CH*$_2$CH$_3$); 3.38 (d, 2H, *J*=12.6 Hz, Ar*CH*$_2$Ar eq.); 3.31 (d, 2H, *J*=13.8 Hz, Ar*CH*$_2$Ar eq.); 1.23-1.15 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (100 MHz, $CDCl_3$) δ (ppm): 191.1, 162.7, 161.2, 155.6, 142.6, 137.1, 136.3, 134.8, 133.3, 131.6, 130.2, 128.5, 124.2, 123.6, 123.1, 74.1, 74.0, 73.7, 69.8, 69.6, 69.5, 66.6, 66.5, 66.4, 30.9, 15.3. HR ES-MS: *m/z* Calcd for $C_{45}H_{55}O_{13}N_2$ [(**3**+H)$^+$] 831.36987, found 831.36950.

5,17-Dinitro-11-hydroxy-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (4).

To a solution of **3** (0.13 g, 0.16 mmol) in DCM (15 mL), *m*-CPBA (0.19 g, 1.10 mmol) was added. The reaction mixture was stirred for 5 days at room temperature, then was quenched with a solution of $NaHSO_3$ 0.2 M and vigorously stirred for additional 30 minutes. The organic layer was washed with brine (2 x 20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The resulting material was taken with a solution of NaOH 2 M in MeOH:H_2O 4:1 (5 mL) and was stirred for 2 hours at room temperature, then the mixture was concentrated by evaporation of the MeOH under reduced pressure. The remaining aqueous layer was neutralized with a solution of 1M HCl and extracted with DCM (2 x 15 mL), then the combined organic layers were dried on anhydrous Na_2SO_4 and evaporated

under reduced pressure. The residue was triturated in refluxing hexane overnight, for three times, to give **4** as a pale yellow oil (0.11 g, 0.13 mmol; 81% yield): ^{1}H NMR (300 MHz, CD_3OD) δ (ppm): 7.38 (d, 2H, *J*=2.7 Hz, Ar*H*-meta-NO_2); 7.34 (d, 2H, *J*=2.7 Hz, Ar*H*-meta-NO_2); 6.91 (dd, 2H, J_1=7.6 Hz, J_2=2.1 Hz, Ar*H*-meta); 6.86 (td, 1H, J_1=7.6 Hz, J_2=2.1 Hz, Ar*H*-para); 6.36 (s, 2H, Ar*H*-meta-OH); 4.67 (d, 2H, *J*=13.8 Hz, Ar*CH*$_2$Ar ax.); 4.60 (d, 2H, *J*=13.8 Hz, Ar*CH*$_2$Ar ax.); 4.20 (t, 6H, *J*=5.1 Hz, ArO*CH*$_2$CH$_2$O); 4.14 (t, 2H, *J*=5.4 Hz, ArO*CH*$_2$CH$_2$O); 3.90 (t, 8H, *J*=5.4 Hz, ArOCH$_2$*CH*$_2$O); 3.60-3.55 (m, 8H, O*CH*$_2$CH$_3$); 3.33 (d, 2H, *J*=13.8 Hz, Ar*CH*$_2$Ar eq.); 3.22 (d, 2H, *J*=13.8 Hz, Ar*CH*$_2$Ar eq.); 1.24-1.17 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (75 MHz, CD_3OD) δ (ppm): 161.6, 156.7, 152.0, 149.7, 142.4, 136.5, 136.4, 135.3, 134.7, 128.7, 122.8, 122.7, 115.1, 74.1, 73.1, 73.0, 69.7, 69.6, 60.1, 66.0, 65.9, 30.6, 30.5, 14.3. HR ES-MS: *m/z* Calcd for $C_{44}H_{55}N_2O_{13}$ [(**4**+H)$^+$] 819.36987, found 819.36987; *m/z* Calcd for $C_{44}H_{54}N_2O_{13}Na$ [(**4**+Na)$^+$] 841.35181, found 841.35284.

5,17-Diamino-11-hydroxy-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (5).

To a solution of **4** (0.053 g, 0.065 mmol) in MeOH (10 mL), $NiCl_2 \cdot 6H_2O$ (0.062 g, 0.072 mmol) and $NaBH_4$ (0.025 g, 0.66 mmol) were added. The reaction mixture was stirred for 3 hours at room temperature, then was quenched with a solution of 1 M HCl (15 mL) and the pH raised to 8-9 with a solution of 1 M NaOH. The resulting mixture was extracted with AcOEt (3 x 15 mL) and the combined organic layers were washed with distilled water (2 x 20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. Compound **5** was obtained as an orange to brown oil (0.049 g, 0.065 mmol; quantitative yield), pure enough to avoid further purifications. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 7.10 (br. d, 2H, *J*=7.6 Hz, Ar*H*-meta); 6.92 (br. t, 1H, *J*=7.6 Hz, Ar*H*-para); 6.56 (s, 2H, Ar*H*-meta-OH); 5.99 (s, 2H, Ar*H*-meta-NH_2); 5.89 (s, 2H, Ar*H*-meta-NH_2); 4.58 (br. d, 2H, *J*=12.8 Hz, Ar*CH*$_2$Ar ax.); 4.52 (br. d, 2H, *J*=12.8 Hz, Ar*CH*$_2$Ar ax.); 4.26 (m, 2H, ArO*CH*$_2$CH$_2$O); 4.17 (m, 2H, ArO*CH*$_2$CH$_2$O); 3.94-3.83 (m, 12H, ArO*CH*$_2$CH$_2$O and ArOCH$_2$*CH*$_2$O); 3.63-3.55 (m, 8H, O*CH*$_2$CH$_3$); 3.17 (br. d, 2H, *J*=12.8 Hz, Ar*CH*$_2$Ar eq.); 3.06 (br. d, 2H, *J*=12.8 Hz, Ar*CH*$_2$Ar eq.); 1.26-1.19 (m, 12H, CH$_2$*CH*$_3$). ^{13}C NMR (75 MHz, CD_3OD) δ (ppm): 151.3, 140.7, 135.5, 135.4, 134.9, 127.9, 116.2, 114.3, 73.1, 69.6, 66.0, 30.6, 30.5, 14.3. HR ES-MS: *m/z* Calcd for $C_{44}H_{59}N_2O_9$ [(**5**+H)$^+$] 759.42151, found 759.42188.

5,17–Bis-[*N*,*N*'–bis(*Boc*)guanidine]-11-hydroxy-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene (6).

To a solution of **5** (0.049 g, 0.065 mmol) and triethylamine (72 µL, 0.039 mmol) in dry DMF (5 mL), *N*,*N*'-bis-(*Boc*)-thiourea (0.015 g, 0.039 mmol) and $HgCl_2$ (0.070 g, 0.26 mmol) were added. The reaction mixture was stirred for 2 days at room temperature, then was quenched by adding AcOEt (15 mL) and the precipitated HgS was filtered off. The filtrate was washed with brine (3 x 20 mL) and the organic layer was dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. The crude material was purified by flash chromatography (hexane/AcOEt 7:3) to give **6** as a colorless oil (0.034 g, 0.027 mmol; 42% yield). ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 11.62 (br. s, 2H, *NH*Boc); 10.18 (s, 2H, Ar*NH*guanidinium);

7.25 (s, 4H, Ar*H*-meta-guanidinium); 6.38 (t, 1H, *J*=6.9 Hz, Ar*H*-para); 6.28 (d, 2H, *J*=6.9 Hz, Ar*H*-meta); 5.68 (s, 2H, Ar*H*-meta-OH); 4.49 (d, 2H, *J*=12.6 Hz, ArCH_2Ar ax.); 4.42 (d, 2H, *J*=13.5 Hz, ArCH_2Ar ax.); 4.24 (t, 4H, *J*=6.6 Hz, ArOCH_2CH_2O); 3.90-3.84 (m, 8H, ArOCH_2CH_2O and ArOCH$_2CH_2$O); 3.78-3.72 (m, 4H, ArOCH$_2CH_2$O); 3.60-3.48 (m, 8H, OCH_2CH_3); 3.15 (d, 2H, *J*=12.6 Hz, ArCH_2Ar eq.); 3.09 (d, 2H, *J*=13.5 Hz, ArCH_2Ar eq.); 1.60-1.49 (m, 36H, OC$(CH_3)_3$); 1.25-1.15 (m, 12H, CH_2CH_3). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 163.6, 155.5, 155.0, 153.9, 153.4, 150.7, 137.1, 137.0, 133.9, 133.0, 127.8, 123.6, 123.3, 122.5, 114.7, 83.6, 79.5, 74.0, 72.5, 69.6, 66.5, 66.2, 30.9, 30.8, 29.7, 28.2, 28.1, 15.4, 15.3. HR ES-MS: *m/z* Calcd for $C_{66}H_{95}N_6O_{17}$ [(**6**+H)$^+$] 1243.67482, found 1243.67444; *m/z* Calcd for $C_{61}H_{87}N_6O_{15}$ [(**6**-Boc+H)$^+$] 1143.62239, found 1143.62122.

5,17–Diguanidine-11-hydroxy-25,26,27,28-tetrakis(2-ethoxyethoxy)calix[4]arene, bis-hydrochloride (1·2HCl).

In a mixture of DCM/TFA/TES 95:2.5:2.5 (10 mL), **6** (0.034 g, 0.027 mmol) was dissolved. The reaction mixture was stirred overnight at room temperature and quenched by removal of the solvent under reduced pressure. The residue was taken into a 1 M HCl EtOH solution (3 mL), vigorously stirred for 30 min and then the solvent removed under reduced pressure. This procedure was repeated three times to exchange the CF_3COO^- anion to chloride. Then **1**·2HCl was obtained as a colorless oil (0.021 g, 0.025 mmol; 93% yield), pure enough to avoid further purifications. ^{1}H NMR (400 MHz, CD_3OD) δ (ppm): 6.89 (d, 2H, *J*=7.2 Hz, Ar*H*-meta); 6.77 (t, 1H, *J*=7.2 Hz, Ar*H*-para); 6.55 (s, 2H, Ar*H*-meta-guanidinium); 6.51 (s, 2H, Ar*H*-meta-guanidinium); 6.32 (s, 2H, Ar*H*-meta-OH); 4.63 (d, 2H, *J*=13.2 Hz, ArCH_2Ar ax.); 4.57 (d, 2H, *J*=13.2 Hz, ArCH_2Ar ax.); 4.26 (t, 2H, *J*=5.6 Hz, ArOCH_2CH_2O); 4.18 (t, 2H, *J*=5.2 Hz, ArOCH_2CH_2O); 4.12 (m, 4H, ArOCH_2CH_2O); 3.94-3.91 (m, 8H, ArOCH$_2CH_2$O); 3.67-3.56 (m, 8H, OCH_2CH_3); 3.26 (d, 2H, *J*=13.6 Hz, ArCH_2Ar eq.); 3.15 (d, 2H, *J*=12.8 Hz, ArCH_2Ar eq.); 1.27-1.18 (m, 12H, CH_2CH_3). ^{13}C NMR (100 MHz, CD_3OD) δ (ppm): 156.6, 155.5, 151.8, 149.6, 136.6, 136.5, 135.5, 135.0, 128.4, 128.1, 124.9, 124.8, 122.5, 114.8, 73.9, 73.0, 72.9, 69.9, 69.8, 69.7, 66.2, 66.1, 66.0, 30.5, 30.4, 14.3 HR ES-MS: *m/z* Calcd for $C_{46}H_{63}N_6O_9$ [(**1**+H)$^+$] 843.46510, found 843.46593.

3.5 Bibliography

(1) Holtz, K. M.; Kantrowitz, E. R. *FEBS Lett.* **1999**, *462* (1–2), 7.

(2) Cotton, F. A.; Hazen, Edward E., J.; Legg, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76* (6), 2551.

(3) Redinbo, M. R.; Stewart, L.; Kuhn, P.; Champoux, J. J.; Hol, W. G. J. *Science (80).* **1998**, *279* (March), 1504.

(4) Stewart, L. *Science (80).* **1998**, *279* (5356), 1534.

(5) Salvio, R. *Chem. Eur. J.* **2015**, *21* (31), 10960.

(6) Salvio, R.; Cincotti, A. *RSC Adv.* **2014**, *4* (54), 28678.

(7) Savelli, C.; Salvio, R. *Chem. Eur. J.* **2015**, *21* (15), 5856.

(8) Sheng, X.; Lu, X. M.; Zhang, J. J.; Chen, Y. T.; Lu, G. Y.; Shao, Y.; Liu, F.; Xu, Q. *J. Org. Chem.* **2007**, *72* (5), 1799.

(9) Salvio, R.; Cacciapaglia, R.; Mandolini, L. *J. Org. Chem.* **2011**, *76* (13), 5438.

(10) Tjioe, L.; Joshi, T.; Brugger, J.; Graham, B.; Spiccia, L. *Inorg. Chem.* **2011**, *50* (2), 621.

(11) An, Y.; Tong, M.-L.; Ji, L.-N.; Mao, Z.-W. *Dalt. Trans.* **2006**, No. 17, 2066.

(12) He, J.; Sun, J.; Mao, Z. W.; Ji, L. N.; Sun, H. *J. Inorg. Biochem.* **2009**, *103* (5), 851.

(13) Baldini, L.; Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Salvio, R.; Sansone, F.; Ungaro, R. *J. Org. Chem.* **2012**, *77* (7), 3381.

(14) Salvio, R.; Cacciapaglia, R.; Mandolini, L.; Sansone, F.; Casnati, A. *RSC Adv.* **2014**, *4* (65), 34412.

(15) Salvio, R.; Casnati, A.; Mandolini, L.; Sansone, F.; Ungaro, R. *Org. Biomol. Chem.* **2012**, *10* (45), 8941.

(16) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Sansone, F.; Mandolini, L.; Casnati, A. *J. Org. Chem.* **2016**, *81* (11), 4728.

(17) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Casnati, A.; Mandolini, L.; Sansone, F. *J. Org. Chem.* **2015**, *80* (11), 5887.

(18) Ullrich, S.; Nazir, Z.; Büsing, A.; Scheffer, U.; Wirth, D.; Bats, J. W.; Dürner, G.; Göbel, M. W. *ChemBioChem* **2011**, *12* (8), 1223.

(19) Salvio, R.; Volpi, S.; Cacciapaglia, R.; Sansone, F.; Mandolini, L.; Casnati, A. *J. Org. Chem.* **2016**, *81* (19), 9012.

(20) van Loon, J.-D.; Heida, J. F.; Verboom, W.; Reinhoudt, D. N. *Recl. des Trav. Chim. des Pays-Bas* **1992**, *111* (7–8), 353.

(21) Arduini, A.; Fanni, S.; Manfredi, G.; Pochini, A.; Ungaro, R.; Sicun, A. R.; Ugozzoli, F. *J. Org. Chem.* **1995**, *60* (5), 1448.

(22) Struck, O.; Verboom, W.; Smeets, W. J. J.; Spek, A. L.; Reinhoudt, D. N. *J. Chem. Soc. Perkin Trans. 2* **1997**, *5* (2), 223.

(23) Morita, Y.; Agawa, T.; Nomura, E.; Taniguchi, H. *J. Org. Chem.* **1992**, *57* (13), 3658.

(24) Kametani, T.; Umezawa, O. *Japanese J. Crop Sci.* **1981**, *50* (4), 476.

(25) Georgieva, M.; Velinov, G.; Budevsky, O. *Anal. Chim. Acta* **1977**, *90* (Supplement C), 83.

(26) Wróbel, R.; Chmurzy, L. **2000**, *405*, 303.

(27) Corona-Martínez, D. O.; Taran, O.; Yatsimirsky, A. K. *Org. Biomol. Chem.* **2010**, *8* (4), 873.

(28) Baughman, E. H.; Kreevoy, M. M. *J. Phys. Chem.* **1974**, *78* (4), 421.

(29) Frost, A. A; Pearson, R. G. *Kinetics and Mechanism*, 2nd ed.; John Wiley & Sons, Ed.; New York, 1961.

(30) Salvio, R.; Mandolini, L.; Savelli, C. *J. Org. Chem.* **2013**, *78* (14), 7259.

(31) Luxon, S. G. *Hazards in the Chemical Laboratory*, 5th ed.; Chemistry, T. R. S. of, Ed.; Cambridge, UK, 1992.

(32) Gans, P.; Sabatini, A.; Vacca, A. *Talanta* **1996**, *43* (10), 1739.

(33) Alderighi, L.; Gans, P.; Ienco, A.; Peters, D.; Sabatini, A.; Vacca, A. *Coord. Chem. Rev.* **1999**, *184* (1), 311.

Chapter 4

Strategies for the study of protein-ligand interactions and for the development of supramolecular ligands

4.1 Introduction

Proteins[1] are large biomolecules that perform a wide variety of functions in living organisms. They are constituted by one or more polypeptide sequences which impart them a sophisticated and unique 3D structure, called *protein folding*[2], and determine their specific biological activity. Proteins are known to act as enzymes and antibodies in living organisms, by catalyzing several types of biological transformations[3] and by supervising the activation of the immune response[4], respectively. They are also involved in cell signaling[5,6], working as selective receptors present on the cell surface to govern cell proliferation, growth and differentiation[7], and participate to the transport or storage processes of various types of molecules through the cell membrane. Receptorial roles are also played in infections, where pathogens use proteins as recognition tools to tether the target cell membrane[8] and in cancer, where they might favor the proliferation and spreading of metastases[9,10]. Moreover, proteins can also have several important mechanical and structural roles, being constitutive parts of connective tissues, cell cytoskeletons, animal shells and muscles[6,11,12].

A common feature of all proteins is that their functions are driven by recognition processes of specific counterparts, which can be both macromolecules, such as nucleic acids, polysaccharides and other proteins or small molecules, like oxygen, metal cations, peptides and other types of ligands[13–15]. In order to understand in detail how these processes take place, it is crucial to achieve the most complete characterization possible of the mechanism of the interactions which involve the given protein-ligand pair, both under a structural[14,16] and under an energetic[17,18] point of view. These information can give to researchers precious insights for the creation of artificial molecules that can tightly bind to a target protein in order to influence its action, thus affording a great contribution for the discovery, the design and the development of new drugs[19]. In fact, many drugs available on the market, sometimes also being mimics of their natural substrates, exploit interactions with target proteins in order to interfere or modulate their action. These interactions can inhibit protein aggregation[20] or specific protein-protein interaction[21] which have a pathological effect on the human body or, at contrary, can promote, manage and stabilize oligomerization processes[22,23] with a beneficial physiological effect.

In the first part of this chapter, the main features of the kinetics and energetics that drive protein-ligand interactions will be briefly reported, followed by a description of some of the most important experimental techniques that can be used to obtain both a structural and a thermodynamic characterization of these processes. In the second part will be also reported the main results and progresses in the design of small molecules as ligands for target proteins. Of course, the amount of literature regarding this field of research is huge and an exhaustive review of this subject is out of the scopes of this thesis. Hence, in that part of the chapter, only the most important examples of calix[4]arene-based ligands will be surveyed, with a short description of the strategies followed to make

this family of macrocycles able to interact with the target protein and, in some cases, to influence its biological activity.

4.2 Mechanism of protein-ligand interactions

To achieve a deep understanding of the protein-ligand interaction process a detailed study of the physico-chemical factors that govern the formation of the corresponding complex is required. For this reason, a basic description and/or rationalization of the kinetic and thermodynamic concepts and of the binding driving forces behind this kind of phenomenon will be now reported.

4.2.1 Kinetics of protein-ligand binding

The rate of association of a generic protein P with a ligand molecule L to form the parent protein-ligand complex PL can be described according to **Scheme 1**[19]:

$$P + L \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} PL \qquad \textbf{Scheme 1}$$

where k_1 and k_{-1} are the rate constants for the association and the dissociation reaction of PL, respectively. At the equilibrium the rate for the binding process equals the one for the reverse unbinding event and the kinetic law for the overall process can be written as follow in **Eqn. 1**, whit [*P*], [*L*] and [*PL*] indicating the molar concentration of the protein, of the ligand and of the protein-ligand complex, respectively.

$$k_1[P][L] = k_{-1}[PL] \quad \textbf{Eqn. 1}$$

Consequently, both the equilibrium constant for the association (K_a) and the dissociation (K_d) of PL can be expressed as function of the corresponding kinetic rate constant (**Eqn. 2**).

$$K_a = \frac{1}{K_d} = \frac{[PL]}{[P][L]} = \frac{k_1}{k_{-1}} \quad \textbf{Eqn. 2}$$

In most cases, tight protein ligand binding rises from processes in which the formation of the complex is much faster than the backward dissociation reaction.

4.2.2 Thermodynamics of protein-ligand binding

Under a thermodynamic point of view, the protein-ligand binding process can be described by considering a system composed by the solvent and by the two binding partners as solutes. The resulting binding event is generated by the balance of the various interactions that take place among the protein, the ligand and the solvent and the occurrence of these interactions is dictated by the laws of thermodynamics[17,18].

In analogy with any other spontaneous processes, the protein-ligand binding takes place when the variation of its standard free energy $\Delta G_b°$ is negative, and the magnitude of this change is related to the equilibrium constant K_a for the association of the protein-ligand complex. The higher is the K_a value the

most negative will be ΔG_b and the most favorite will be the binding event. For a generic binding reaction, $\Delta G_b°$ is composed of an enthalpic contribution $\Delta H_b°$, which is a measure of the total heat exchange caused by the interactions involving the components of the system, and of an entropic contribution $\Delta S_b°$, which indicates the distribution of heat among the whole thermodynamic system and is commonly reported as an index of its degree of disorder. These concepts can be easily visualized in the well-known definition of binding Gibbs free energy[24] which is reported in **Eqn. 3**.

$$\Delta G_b^\circ = -RTlnK_a = \Delta H_b^\circ - T\Delta S_b^\circ \quad \mathbf{Eqn.3}$$

In a non-strict sense $\Delta H_b°$ can be also indicated as the energy balance deriving from the formation of new noncovalent interactions between the protein and ligand and from the disruption of preexisting interactions between the solvent and both the ligand and the protein[24,25]. According to this principle, when this balance is favorable for binding $\Delta H_b°$ has negative values and the process is exothermic, while when $\Delta H_b°$ is positive the protein-ligand association is endothermic and not spontaneous. At contrary, positive and negative $\Delta S_b°$ values will advantage and disadvantage binding, respectively. The net entropic variation $\Delta S_b°$ associated with the protein-ligand binding process can be expressed as a sum of three terms[25,26]: $\Delta S_{solv}°$ which is the entropy change caused by the solvent release from both the ligand and the protein, $\Delta S_{conf}°$ which indicates the change in conformational freedom of both the two entities upon binding and $\Delta S_{r/t}°$ which is related to the loss of rotational and translational degrees of freedom that the protein and the ligand experience during their complexation (**Eqn. 4**).

$$\Delta S_b^\circ = \Delta S_{solv}^\circ + \Delta S_{conf}^\circ + \Delta S_{r/t}^\circ \quad \mathbf{Eqn.4}$$

$\Delta S_{solv}°$ is always largely positive and thus contributes favorably to the protein-ligand association, while $\Delta S_{r/t}°$ is always negative and disadvantages the occurrence of binding[25]. $\Delta S_{conf}°$, instead, can afford both a negative and a positive contribution to the total binding entropy because the conformational freedom of the ligand and the protein can both increase and decrease when they bind to each other[27]. Hence, since the sign and the magnitude of ΔG_b depends on two contributions, the driving forces that govern the protein-ligand association can be different. More in details, when the change of binding free energy is mainly due to enthalpic factors it is said that the process is enthalpy-driven, while when it is dominated by the entropic ones the phenomenon is indicated as entropy-driven.

It has been observed that the contributions of ΔH_b and ΔS_b to ΔG_b are closely related[28,29]. For example, when the formation of the protein-ligand complex derives from the formation of multiple interactions between the associating entities a highly negative enthalpy change will be recorded, but at the same time, a decrease of entropy of the system will take place because of the reduction of mobility of both of the binding partners. On the hand, when the protein-ligand binding is entropy driven, for instance because of the release of a high number of solvent molecules from the surface of both the protein and the ligand, a positive binding enthalpy is generally expected, due to the disruption of a high number of

pre-existing interactions. The complementary change of entropy and enthalpy often observed during the protein-ligand binding event is called *enthalpy-entropy compensation*. The observation of this phenomenon in several thermodynamic binding studies[30,31], together with the results of some theoretical models[32,33] leads to suggest that it could be a general feature of protein-ligand interaction and of other association events, despite its origin has not been elucidated with certainty. In fact, a large variety of factors, such as the properties of the solvent[29,34,35], the structure and mobility of both the ligand and the protein and the nature of the molecular forces involved in the binding process[28,32,36], have been reported to influence the enthalpy-entropy compensation and its detailed analysis is generally difficult.

In this paragraph an introduction on the basic thermodynamic concepts that govern the protein-ligand binding have been reported. A quantitative dissection of the enthalpic and entropic contributions to the overall free energy of these processes is crucial, because it can provide precious insights for the improvement of the design of ligands able to interact the most effectively as possible with a target protein.

In the next paragraph it will be shown how these general notions can be used for a detailed description of the multivalency effect, which is one of the main tools that can be exploited by both synthetic[15,37] and natural ligands[38] to bind a protein with high affinity.

4.2.3 Multivalency effect

Multivalency[38] is one of most powerful instruments exploited by nature to achieve strong but yet reversible binding between two or more entities. Considering the case of proteins, the non-covalent interactions that they use to recognize their counterparts, *i.e.* hydrogen bonding, π-π stacking, cation- π, hydrophobic, ionic, and van der Waals forces, are too weak to promote strong and selective binding processes, if considered singularly. Thus, to achieve the formation of stable complexes, they exploit the contemporary use of several of these interactions in their binding site, where the contact with the given binding partner takes place. To further increase the strength and the specificity of the forming complexes, multivalent proteins present on their surfaces multiple identical binding sites, each of them used to establish the same non-covalent interactions with an equal number of binding units belonging to the ligand. The power of this phenomenon, generally indicated as *multivalent effect*, is that the resulting overall affinity of proteins for their ligands is higher than the one that would have been obtained by simply adding the binding energies arising from the interaction between each single receptor/binding unit pair.

The generic binding of a *i*-meric protein by an *i*-valent ligand can be described as a sequential process that starts with an intermolecular interaction followed by *i*-1 intramolecular ones, leading to the

formation of the corresponding ligand-receptor complex. A general model for the description of such a system has been introduced by Jenks[39] by exploiting the principle of additivity of free energies (**Eqn. 5**).

$$\Delta G^{\circ}_{multi} = i\Delta G^{\circ}_{mono} + i\Delta G^{\circ}_{interaction} \quad \mathbf{Eqn.\,5}$$

In this model the total free energy associated with the process ΔG°_{multi} is separated in two contributions, one including the sum of the binding energy ΔG°_{mono}, corresponding to *i* monofunctional interactions between the protein and the ligand, and the other representing the balance of the favourable and unfavourable energetic factors $\Delta G^{\circ}_{interaction}$ which derive from the occurrence of tethering. In other words, $\Delta G^{\circ}_{interaction}$ is a correction term introduced to describe the deviation of the overall free energy from the one expected by summing *i* intermolecular monofunctional binding events. When the multivalent effect is beneficial for protein-ligand interaction this term can be extremely negative, thus resulting in a high free energy gain.

However, a detailed thermodynamic analysis of the $\Delta G^{\circ}_{interaction}$ term can be difficult and Kitov et al. proposed an alternative model for the description of multivalency[40]. According to the limiting conditions that only one multivalent ligand can bind to a target oligomeric receptor at a time, thus excluding the formation of aggregates, and that all the binding units of the ligand and the binding sites of the protein are identical and act independently, **Eqn. 5** can be in fact rearranged as follow:

$$\Delta G^{\circ}_{i} = \Delta G^{\circ}_{inter} + (i-1)\Delta G^{\circ}_{intra} - RTln\Omega_{i} \quad \mathbf{Eqn.\,6}$$

where $\Delta G^{\circ}_{inter}=\Delta G^{\circ}_{mono}$ and $(i\text{-}1)\Delta G^{\circ}_{intra}=(i\text{-}1)\Delta G^{\circ}_{mono} + \Delta G^{\circ}_{interaction}$ are the contributions to the binding free energy derived from intermolecular and intramolecular interactions, respectively, and $RT\ln\Omega_i$ is a statistical term consistent with the Boltzmann-Gibbs definition of entropy[40].

The convenience of **Eqn. 6** is that it allows to perform a dissection of the different energetic factors that give rise to the multivalent effect. The enthalpic and part of the entropic factors are in fact combined in its first and second terms, which represent the gain in binding free energy due to presence of additional intramolecular interactions once the initial intermolecular one has occurred. The maximum number of these additional interactions determine the highest value possible for this term, which is approached asymptotically with the increasing of binding units of the ligand. The third term, also indicated by the author as *avidity entropy* $\Delta S^{\circ}_{avidity,}$ is instead related to the probability of the occurrence of a generic binding event and incorporate the statistical entropic contributions of the multivalent effect. The magnitude of the avidity entropy is always positive and increase non-linearly with the number of both the binding units of ligand and the binging sites of the receptor, becoming a determining contribution for high valency values.

Good estimation of the extent of the multivalency effect can be also obtained without rigorous treatment of the energetics associated with the binding process, by considering simple parameters. One of the most used of them is the *enhancement factor* β[37], which was introduced by Whitesides et al. and was defined as the ratio between the binding constant K_{multi} and K_{mono} obtained for the binding of,

respectively, a multivalent ligand and its corresponding monovalent analog to the same target receptor (**Eqn. 7**).

$$\beta = \frac{K_{multi}}{K_{mono}} \quad \textbf{Eqn. 7}$$

A parameter that can be obtained with a similar approach is the *relative potency rp* (**Eqn. 8**), calculated as the ratio between the IC50 values of a monovalent compound ($IC50_{mono}$) and of its parent multivalent ligand ($IC50_{multi}$), where the IC50 indicates the concentration required to inhibit a certain process by 50% and is frequently used to establish the power of inhibitors for target proteins.

$$rp = \frac{IC_{50_{mono}}}{IC_{50_{multi}}} \quad \textbf{Eqn. 8}$$

These parameters are useful because allow us to obtain an estimation for the magnitude of the multivalent effect also when the number of effective interactions is unknown, because in general the higher the β or *rp* values will be, the more beneficial will be multivalency for the binding process, despite no insight about the cooperativity and on the topology of the ligating groups will be obtained. More detailed information can be instead deduced when the valency of the ligand is known. In fact, for a *i*-valent ligand, by dividing β or *rp* by *i*, it is possible to calculate how many time a single ligating unit is more effective in binding (or inhibiting) target protein, when compared to the same binding unit in the corresponding monovalent compound. In other words, the resulting β/i or *rp/i* values easily allow to compare the binding ability of ligands with different valency, topology or structure.

4.3 Experimental techniques for the study of protein-ligand interactions

The various features of protein-ligand interactions can be studied according to several experimental techniques. X-ray crystallography, nuclear magnetic resonance (NMR), Laue X-ray diffraction, small-angle X-ray scattering, transmission electron microscopy (TEM), scanning electron microscopy (SEM) and cryo-electron microscopy (cryo-EM, recently awarded with the Nobel Price[41]) furnish (near-)atomic-resolution of the structure of protein ligand-complexes and related larger aggregates, while isothermal titration calorimetry (ITC) and differential scanning calorimetry (DSC) are powerful techniques to obtain a detailed thermodynamic characterization of the corresponding binding processes[19]. The combination of the data obtained from these techniques can provide a detailed understanding of the structural and energetic characteristics of the given protein-ligand binding events, providing precious insights on the strategy to exploit these events for practical applications.

In the next paragraphs the attention will be focused on the introduction and the description of ITC, NMR and X-ray crystallography, which are the techniques used in this work of thesis to study the binding of synthetic ligands to different model proteins.

4.3.1 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC)[42,43] is one of the standard techniques used to study intermolecular interaction[44], by measuring the heat exchange that takes place during these processes. In a single experiment it is able to provide quantitative thermodynamic information, such as the binding affinity, the binding enthalpy and the binding stoichiometry, which can be directly obtained from the plot of experimental data. The high sensitivity of the technique for measuring small heat of reaction, togetter with the large practical window for reliable binding affinity, make ITC suitable for the study of the non-covalent interactions which drive most of the biological processes, including the formation of protein-ligand complexes.

The design of ITC instrument is quite simple (**Figure 1**)[45]. It consists of two cells, the reference cell and the sample cell, whose temperature is constantly measured and kept constant by heaters connected to a power compensating system. For a typical study of protein-ligand binding the sample cell is filled with a solution of the protein and is titrated with precise aliquots of a concentrate solution of the ligand, although the reverse setup of the experiment can also be used. In the reference cell is instead put a solution of buffer and other additives lacking the protein. During the titration the interactions between the ligand and the protein cause a heat release or absorption, depending on the exothermic or endothermic nature of the process, leading to a change in temperature of the sample cell with respect to the reference cell, which is counterbalanced by modulating the feedback power applied by the instrument.

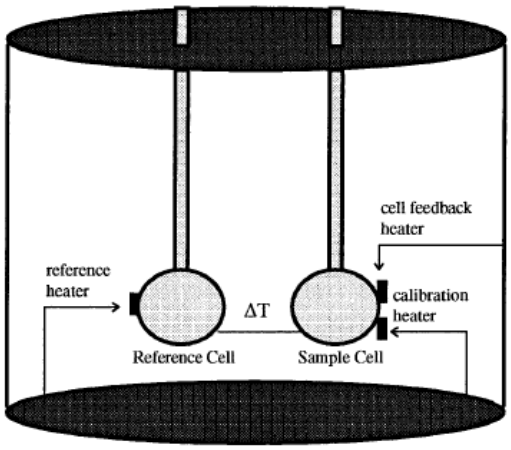


*Figure 1: Schematic representation of an ITC instrument (Figure taken from **Ref. 45**).*

The raw data obtained from ITC are the heat exchanges associated with the addition of each aliquot of ligand per unit of time (**Figure 2a**). The area of the corresponding peaks, proportional to the fraction of protein complexed at each addition in the sample cell, decreases gradually until saturation of the binging sites is reached and only heat of dilution will be measured. The most common representation of these data is the so-called *Wiseman isotherm*[46], where the molar heat of reaction detected for each

addition of ligand is plotted versus the ratio between the ligand and the protein concentration [L]/[M]. The resulting binding curve has generally a typical sigmoidal shape (**Figure 2b**), from which it is possible to extract directly the stoichiometry, the enthalpy and the equilibrium constant of binding as fitting parameters, after a proper nonlinear analysis.

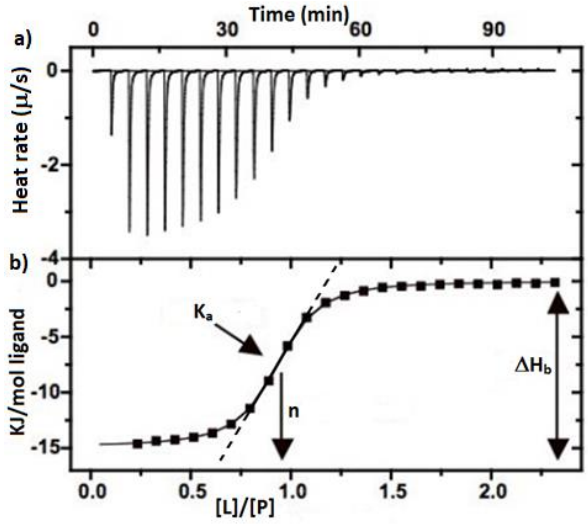


Figure 2: ***a)*** *Raw data and* ***b)*** *typical binding curve derived from a generic ITC experiment, from which it is possible to extract estimations of n, ΔH_b and K_a by simple graphical analysis.*

In fact, according to the expression of the association constant (K_a) for the formation of the given protein-ligand complex PL and to the mass conservation of both the protein and ligand, the total heat exchanged (Q_{tot}) during the binding process can be expressed as follow (**Eqn. 9**):

$$Q_{tot} = nV_0[PL]\Delta H_b =$$

$$= nV_0\Delta H_b \frac{[P]_{tot} + [L]_{tot} + \frac{1}{K_a} - \sqrt{\left([P]_{tot} + [L]_{tot} + \frac{1}{K_a}\right)^2 - 4[P]_{tot}[L]_{tot}}}{2} \quad \boldsymbol{Eqn.\,9}$$

where n is the stoichiometry of the complex, V_0 the volume of the sample cell, ΔH_b the enthalpy of binding, and $[P]_{tot}$ and $[L]_{tot}$ are the total concentrations of the protein and the ligand, respectively. Moreover, further information about the energetics of the protein-ligand binding can be extracted indirectly by considering the connection between the association constant and the binding free energy (**Eqn. 3**), thus obtaining a complete thermodynamic characterization of the process.

The convenience of the Wiseman isotherm is that an estimation of n, ΔH_b and K_a can be obtained also by simple graphical analysis of the experimental data [47]. In fact, the binding enthalpy can be extrapolated as the difference between the initial and the final heat exchanged during the titration, while the stoichiometry and the association constant will be deduced by the inflection point and the slope of the linear part of the titration curve, respectively (**Figure 2b**).

Other important information that can be obtained from the Wiseman isotherm are included in the dimensionless parameter *c* (**Eqn. 10**)[46], expressed as multiplication of the total protein concentration, the stoichiometry of binding and the association constant or, as reported more commonly in biochemistry, the dissociation constant K_d.

$$c = n[P]K_a = \frac{n[P]}{K_d} \quad \boldsymbol{Eqn.10}$$

This parameter gives crucial guidelines for the setup of an ITC experiment, because it can be used to predict the form of the resulting binding isotherm. As shown in **Figure 3**, in fact, in the limiting case of $c<1$ the experimental titration curve will approach linearity and, despite the association constant can be determined, it will not be possible to calculate the binding enthalpy, while the reverse will happen for $c>10000$, where a rectangular isotherm will be obtained. For intermediate *c* values, and in particular for $10<c<1000$, will be instead obtained a reliable curve with a well-defined inflection point, allowing an accurate thermodynamic analysis of the binding process.

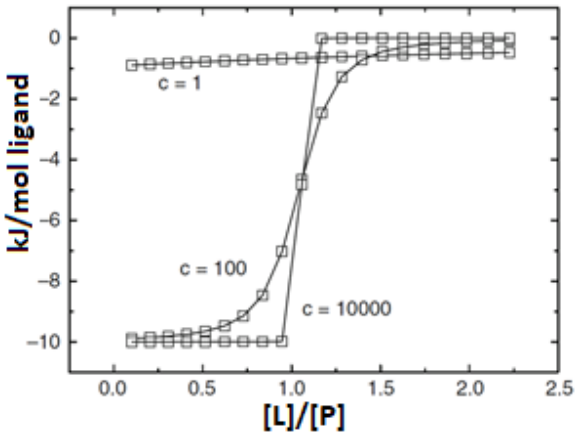


Figure 3: Shape of the Wiseman isotherm for different c values.

Since *n* and K_a are fixed parameters, it is necessary to play with concentration of the protein to obtain optimal ITC experimental data[42,45]. However, this purpose it is not always possible to reach. In fact, for very tight binding events (K_a~10^8-10^9 M^{-1}) the required concentrations of protein will be too low to generate detectable heat of reaction during the ligand additions and, at contrary, for weak binding processes (K_a ~10^2-10^3 M^{-1}) prohibitively high reactants concentrations will be necessary, yielding calorimetric signals out of the instrumental range or leading to a not practically feasible preparation of the sample.

Looking at **Eqn. 9** it appears clear that another parameter that modulate the form of the Wiseman isotherm is the enthalpy of binding ΔH_b, from which the amount of reaction heath developed during the binding is directly proportional. Thus, for low affinity ligand a high enthalpic contribution to the overall energetics of the process could be determining to have measurable reaction heat, while the opposite

can happen for processes involving high affinity ligand, whose reaction heat could exceed the upper detection limit of the instrument.

Nevertheless, a window of 5-6 order of magnitudes for measurable $K_a s$ is one of the largest for the available techniques used to study protein-ligand interaction and, to obtain suitable experimental data, corresponds to the concentration range commonly used to investigate this kind of process (10^{-9} M to 10^{-3} M). Moreover, according to literature, for ligands with affinity values that approach both the lower and the higher instrumental detection limits it is possible to obtain reliable ITC data by performing displacement titrations, where an additional ligand with intermediate affinity is present[43,48–50]. Finally, it is necessary to remember that the heat exchange recorded by the instrument during the titration is the total heat exchange that takes place once the ligand enters in the sample cell, including the contributions related to dilution and to interactions different from protein-ligand binding. For a correct evaluation of the heat of reaction deriving from binding it is thus necessary to perform always control experiments where the heat of dilution of both the ligand and the protein are measured and then subtracted from the raw data. However, in some cases, this process can be much complicated by the occurrence of undesired processes, such as self-aggregation of both the reactants, that provides additional heat of reaction to the system.

Also considering the disadvantages of this technique, from this short description it appears clear that ITC is one of the most powerful techniques to study protein ligand interactions, in particular when most of the attention is focused on the thermodynamics of the binding.

4.3.2 Nuclear magnetic resonance – ^{15}N HSQC experiment

Nuclear magnetic resonance (NMR)[51] is one of the most important experimental techniques for the study of protein-protein[52] or protein-ligand[53,54] interactions. Currently are in fact available different experiments[55] from which it is possible to get detailed information about these binding processes, under both the structural and the thermodynamic point of view.

One of the most used experiments to obtain these information exploits the ^{1}H-^{15}N heteronuclear single quantum correlation and is commonly named ^{15}N HSQC. In this experiment the pulse sequence applied by the instrument on the sample results in 2D-spectra where the cross-peaks indicate the N-H couplings present in the studied protein. This is a powerful technique for the structural characterization of proteins in solution, since the cross-peaks related to the N-H amide bonds of the backbone fall in a characteristic window of chemical shifts (approximately from 100 to 140 ppm for ^{15}N and from 6 to 11 ppm for ^{1}H, respectively) and, within this window, each cross peak will occupy a unique position. In fact, because of the folding of the protein, every amide bond of the backbone experiences a particular chemical environment and, in the corresponding ^{15}N HSQC spectrum, they will give rise to resonances with a precise position, allowing the assignment of the amino acid residue to which are referred. In this way each amino acid of the backbone will be identified by a single signal of the 2D-spectrum, with the

exception of proline, since as tertiary amide lacks the N-H amide proton, and of asparagine and glutamine, whose lateral amide NH groups will give rise to an additional cross peak.

The typical strategy to extract information on the binding of a ligand to a protein consists in performing an NMR titration, where increasing aliquots of ligand are added to a solution of a ^{15}N-labelled protein and the shifts experienced by the N-H correlations in the 2D spectra of the ^{15}N HSQC are monitored. Depending on the kinetics of the protein-ligand binding process, the appearance of the resulting spectra will be different[56,57,58].

The presence of the ligand, in fact, will cause a chemical shift perturbation (CSP) for the resonances belonging to the amino acid residues involved in the interaction. In the most common case of fast binding processes, when compared to the instrument timescale, the observed cross peaks are the average between those belonging to the bound and the unbound form of the protein. The change in position of these resonances will be therefore proportional to the fraction of bound protein present in solution and will reach a plateau when the binding site(s) will be saturated (**Figure 4**). The CSPs provide structural information about the protein-ligand complexation, because only the cross-peaks of the residues involved in the interaction will undergo significant changes of position. However, it is necessary to specify not all the resonances that experience shifts belong to amino acid residues directly in contact with the ligand, because the occurrence of the binding often causes a certain change in conformation of the protein that can also influence the chemical environment around other residues. As mentioned before, during the titration, the observed CSPs will be proportional to the amount of added ligand and will reach a plateau when it is present in excess[57]. By exploiting proper non-linear regression models to fit the resulting saturation curves on both the ^{1}H and the ^{15}N dimensions, it will be possible to extract an estimation of the association constant for the binding process, thus obtaining also thermodynamic information about the studied protein-ligand interactions.

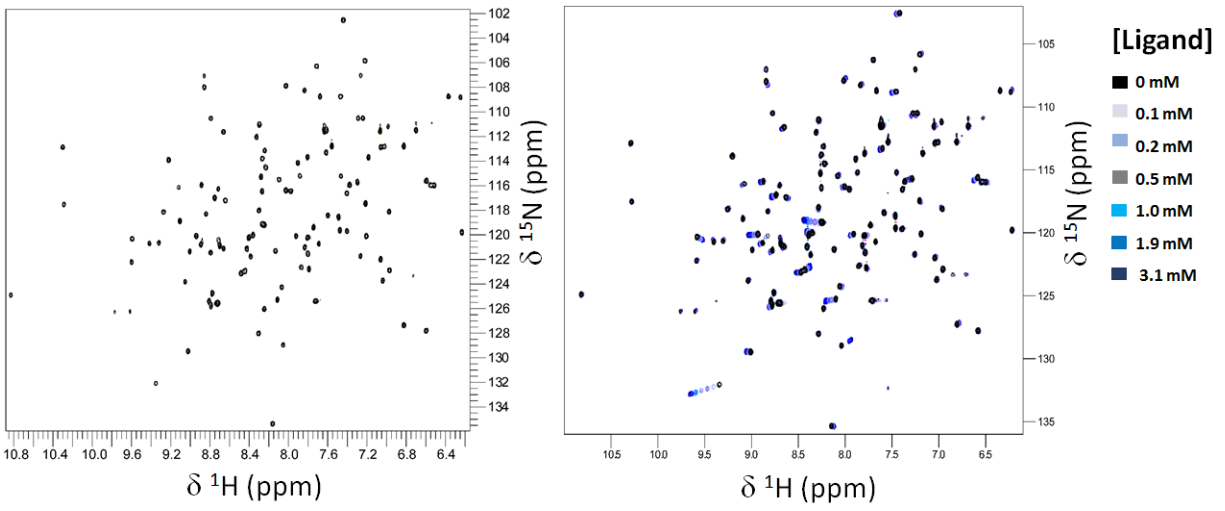


*Figure 4: **a)** Generic example of ^{15}N HSQC spectrum of a protein and **b)** chemical shift perturbations of some of its amide resonances in the presence of increasing amount of a ligand, for a protein-ligand binding process faster than the NMR timescale.*

When the protein-ligand complex formation is slow compared to the NMR timescale the signals belonging to both the free and the bound form of the protein will be contemporary present in the ^{15}N HSQC spectra, with an increase and a decrease of the volumes of corresponding cross peaks which will be recorded along the titration, respectively[57,58]. Thus, the parameter to monitor for the determination of the association constant will be the change in volume of the cross-peaks belonging to the residues involved in the process against the increasing concentration of the ligand. An intermediate situation is instead observed for binding processes with a rate similar to the timescale of the experiment[52]. In these cases, in fact, the resonances involved in the interaction with the ligand give rise to weak and broad signals in the spectra, which become weaker and broader during its addition, resulting suitable for a structural analysis of the binding process, but not for the calculation of the association constant.

For all the three just descripted cases it is important to remember that the results of the ^{15}N HSQC experiments strongly depends on the experimental conditions in which they are carried out, such as pH, temperature and concentration range of reactants[57]. Hence, to have reliable experimental data it is important to keep these parameters constant among the overall titration and, if the purpose is to compare the activity of different ligands, among all the different experiments that will be carried out.

The possibility to obtain contemporary structural and thermodynamic information about protein-ligand interaction from the same experiment makes ^{15}N HSQC titration one of the gold standard for the study of these processes in solution. The fact that also very weak ligands are generally able to impart measurable CSPs into the spectrum of target protein[58] extents the window for the estimation of corresponding dissociation constant in the mM range, which are values hardly detectable with other common experimental techniques (*i.e.* ITC). On the other hand, the low sensitivity of this technique forces to work in a quite high range of protein concentration, that is from the micro- to the millimolar range. Moreover, the interpretation of ^{15}N HSQC spectra can become extremely difficult for large proteins (MW>30 KDa) because of the overcrowding of the signals and of the broadening of their line width due to fast transverse relaxation rates (T_2). These drawbacks are gradually going to be resolved thanks to the introduction on the market of cryogenically-cooled probes (which have their electronics cooled to 20-25K) and the invention of TROSY (Transverse Relaxation Optimized SpectroscopY) experiments. The cryoprobe technology[59–61], in fact, allow to greatly reduce the background thermal noise on the raw data recorded by the instruments, leading to an improvement of the signal to noise ratio and consequently to an enhancement of the sensitivity of this experimental technique. TROSY experiments[62–64] permit instead to obtain a substantial increase of the resolution of corresponding 2D-spectra and thus to study the interaction between ligands and large proteins[65], despite the intrinsic overcrowding of signals still makes the assignment of an exceedingly high number of cross peaks complicated in many cases.

4.3.3 X-ray crystallography

Although NMR spectroscopy allow to obtain many insights on the protein-ligand interaction in solution, X-ray diffraction of protein-ligand co-crystals remains the most reliable technique for the structural characterization of these processes in the solid state[66,67,68]. The combination of data from these two techniques is probably the most effective approach to obtain a detailed understanding of the factors that drive the formation of supramolecular assemblies and can provide important guidelines for the design and development of new drugs.

Regardless of the fact that the aim of the work is to obtain crystals of a small molecule or of a protein-ligand complex, the fundamental stages of the crystallization process are:

- nucleation, where a proper amount of molecules aggregates to form the so-called critical nucleus, which afford the surfaces required for crystal growth.
- growth, where additional molecules diffuse from the bulk solution to the surface of critical nuclei and stick on them to form the crystal lattice, by exploiting specific and highly directional interactions.
- end of growth, which takes place when the concentration of the protein in the bulk solution decreases under the saturation level or when the growing lattice is destabilized by deformation-induced strain or by the presence of impurities.

The protein-ligand co-crystallization process can be described as a particular case of protein crystallization in which the presence of the ligand is one of the variables that have to be optimized to obtain the target co-crystals, together with pH, temperature, concentration of the protein, of the precipitant agents and of other additives[69–71].

Thus, a convenient representation of this process can be obtained according to bi-dimensional phase diagrams used to describe standard protein crystallization (**Figure 5**)[67,68,72], in which the stable states of the protein (liquid, crystalline, amorphous solid) are represented as function of only two of these parameters at time, where one is generally the initial protein concentration [P].

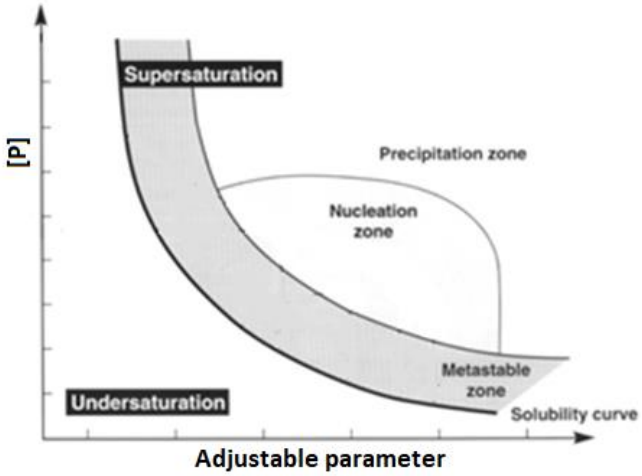


*Figure 5: Schematic illustration of a protein crystallization phase diagram (Figure adapted from **Ref. 72**).*

The regions of the diagrams that can be identified are the undersaturation zone, which is delimited by the values of [P] and of the other adjustable variable where the protein is soluble in the studied medium, and a supersaturation zone. The latter region can be further divided in three parts: the precipitation zone where the supersaturation level is too high and amorphous aggregates are obtained; the nucleation one, where the formation of the crystal around the critical nuclei begins and crystal growth starts to occur; and the metastable zone, where only growth can slowly take place because nucleation is kinetically inhibited. To obtain optimal conditions for crystallization, the various parameters that govern the process should be tuned in order to drive the system in a part of the nucleation zone as close as possible to the border with the metastable one (**Figure 5**). In this way, since the protein concentration will gradually decrease during the process, once that nucleation has occurred the system will pass readily in the metastable zone, where growth occurs slowly and the resulting crystal will be of higher quality.

Despite extensive studies to understand the physico-chemical principles that rule protein crystallization have been carried out, a full comprehension of this phenomenon has not been achieved yet and the obtainment of protein crystals suitable for X-ray diffraction is still a hard challenge. Further complications appear when the purpose is to crystallize the target protein together with its ligand[71], because the presence of the binding partner represent an additional parameter to be optimized. However, according to the great interest attracted by this topic, several methods for protein-ligand co-crystallization have been developed over the years.

Among them, the most commonly used is by far the vapor-diffusion experiment[68,73,72], which has been exploited also in this work of thesis and hence will be now quickly described. In this technique an aqueous droplet containing the protein, the ligand, the buffer and the precipitating agents (generally inorganic salts or polyethylene glycols (PEGs) of high molecular weight) is equilibrated against a reservoir solution, composed by the same buffer and precipitant agents but at a higher concentration. Because of the different amount of buffer and precipitants contained in them, slow diffusion of water from the crystallization mixture to the reservoir solution will take place, causing an increase of the protein and ligand concentrations and, in some cases, their co-crystallization. These experiments are carried out in appropriate plates containing hermetically sealed wells, where the relative position of the two solutions can be different. In the so-called hanging drop design the crystallization mixture is put on the cover slides that close the wells and is suspended over the reservoir solution, while in the sitting drop design it is placed on proper support (**Figure 6a** and **6b**, respectively). Under a practical point of view[70], the sitting drop experiments are easier to be prepared, but sometimes are disadvantageous because the crystals can adhere on the support where they are located, resulting hard to be collected[67,68,70]. On the other hand, the set-up of the hanging drop experiments is more difficult, but they give often better results since crystals do not stick on the cover slide and the higher distance[74] from

the reservoir solution guarantee a slower evaporation of the crystallization droplet and thus a higher quality of the crystals.

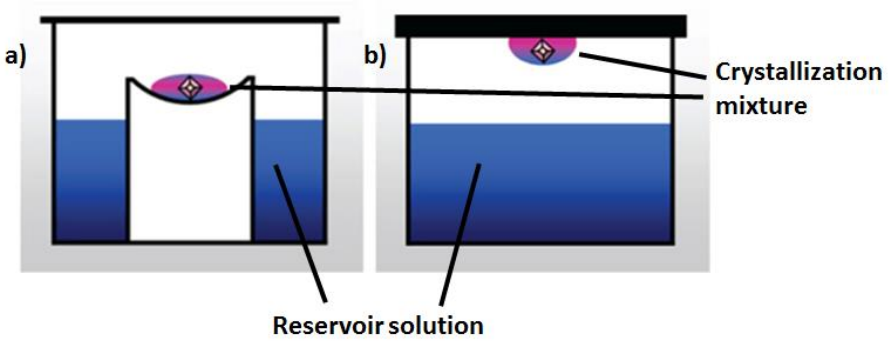


*Figure 6: Schematic representation of vapor diffusion experiments in **a)** the sitting drop and **b)** the hanging drop configurations.*

The most important features of the vapor diffusion method, in both the possible configurations, are the straightforward preparation of the experiments and the low amount of protein and ligand required, that offer the opportunity to explore systematically a large variety of different experimental conditions[69,70,74]. On the market are in fact available multi-well plates (**Figure 7a** and **7b**) in which very small droplets (with a typical volume of few μL) containing protein and ligand in different concentration can be equilibrated against reservoir of diverse compositions, which can be prepared manually or taken from commercial crystallization kits. Moreover, for both the sitting drop and in the hanging drop configurations, robotic technologies to automatize the set-up of the experiments have been developed (**Figure 7c**), allowing to further decrease the amount of time and raw materials (droplets volume of less than 1μL) required to carry out the crystallization tests[75].

These facts have a crucial significance since, currently, protocols for the systematic co-crystallization of protein-ligand complexes have not been established yet, and the possibility to perform quickly and easily detailed screening of the various parameters that influences these phenomena can be determining to succeed in the achievement of their occurrence.

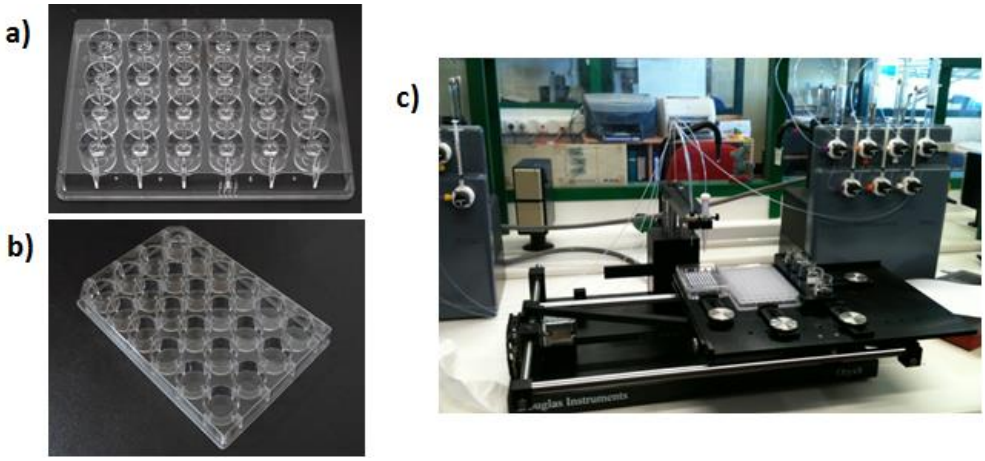


*Figure 7: 24-well plates commonly used to set up manual **a)** sitting drop and **b)** hanging drop vapor diffusion experiments. On the right is reported a photo of **c)** a robot for automatized hanging drop vapor diffusion screening.*

4.4 Calix[4]arenes as ligands for proteins

Over the years, the special properties of calix[n]arenes has been exploited to develop effective and selective receptors for the recognition of many different guest molecules[76–78]. Moreover, since the synthesis of water-soluble derivatives of these macrocycles has been reported[79], their use as ligands for biological macromolecules, such as enzymes[15,80,81], proteins, nucleic acids and lypopolysaccharides, has been investigated, in order to test their potential application as anti-tumor agents, effector of the immune response, gene- or drug-delivery systems and anti-adhesion agents against viruses, bacteria and bacterial toxins. In these fields, the possibility to exploit the different non-covalent interactions (i. e. hydrogen bonding, π-π stacking, cation- π, hydrophobic, ionic, and Van der Waals forces) provided by both the lipophilic cavity[82] of these macrocycles and the wide variety of functional groups[83] that can be easily inserted at both its upper and lower rim is a powerful tool to achieve the desired binding events. Additionally, the possibility to tune the valency and the conformation of the calix[n]arene molecular scaffold is particularly useful for the creation of multivalent ligands[81] whose number and orientation[84,85] of the binding units can be suitable to fit the surfaces of target macromolecule.

The amount of literature about calix[n]arene-based ligand for bio-macromolecules[15,80,81,86–88] is huge and a complete review of this subject is out of the aim of this paragraph. In the following pages, thus, will reported a general description of the features that allow these compounds to be effective ligands for proteins and, in some cases, to influence their biological activity, with a particular attention for calix[4]arene derivatives, since they are the sub-family of these macrocycles on which this thesis has been focused.

Cone-calix[4]arene derivatives, for example, has been reported to be suitable for the appropriate orientation at their upper rim of binding units able to interact with the so-called *hot-spots*[89,90], that is groups of amino acid surrounding the binding sites of target proteins. This property has been exploited by Hamilton et al. for the synthesis of compounds **1** and **2** (**Figure 8**)[91,92], decorated at the upper rim with four polypeptide loops. Taking inspiration by the action of antibodies, these derivatives were designed to bind the target proteins by recognizing a large region of their surfaces thanks the simultaneous exploitation of both hydrophobic and electrostatic multiple interactions.

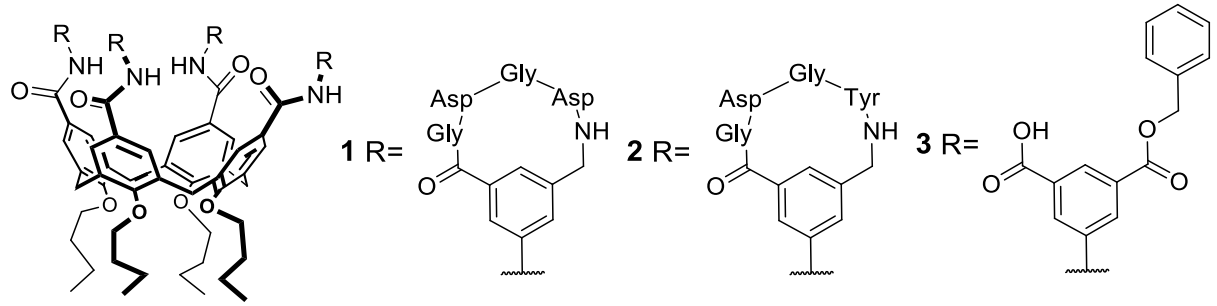


Figure 8: Cone-calix[4]arene derivatives tested as ligands for cythochrome c.

The first protein-ligand system to be investigated was composed by **1** and cytochrome *c* (cyt *c*)[91]. In cyt *c* the catalytic heme group is surrounded by a lipophilic region containing a positively charged belt which can be covered by **1** to form a highly stable 1:1 complex (K_a~10^8 M^{-1}). The binding process was proposed to take place thanks to the presence, at the upper rim of the calix[4]arene, of four cyclic GlyAspGlyAsp loops containing hydrophobic and negatively charged amino acid residues, which provide the complementary interactions necessary to recognize the target region of the protein. Moreover, also the 3-aminobenzoyl moieties included in each peptide sequences was suggested to participate to the recognition process, by affording additional hydrophobic contacts with the protein surface. Thanks to the occurrence of the binding with **1**, the reduction of the Fe(III) ion, contained in the heme group of the protein, by ascorbate was inhibited, such as the interaction with its catalytic partner cyt *c* peroxidase[93]. Similarly, when tested as ligand for α-chymotrypsinase, **1** was reported to bind the target protein with a dissociation constant in the nanomolar range and to avoid the formation of complexes between the target protein and soybean trypsin[94]. Compound **2** was instead tested by the same authors as inhibitor of the binding to the corresponding receptors of growth factors involved in angiogenesis and tumor progression. More in details, it was studied the ability of this calix[4]arene to prevent the association between the platelet derived growth factor (PDGF) and its membrane receptor (PDGFR)[92], which is a crucial event for the maintenance of blood vessels in oncogenesis. As cytochrome *c* and α-chymotrypsinase, PDGF has on its surface a hydrophobic region containing a positively charged belt and, as expected, compound **2** was a strong inhibitor of the adhesion process (IC_{50}=250nM). This fact seems to highly disadvantage the growth of different human tumors implanted in nude mice.

Inhibitor properties were also reported for calix[4]arene **3**[95], in which the alternation of negatively charged and hydrophobic groups was obtained by functionalizing its upper rim with monobenzylester isophthalic acid moieties and, for this reason, it has been presented as simplified version of compounds **1** and **2**. The inhibitory action was played against both PDGF and the vascular endothelial growth factor VEGF (IC_{50}= 190 and 480 nM, respectively), from which the development of tumors depends on, being both involved in angiogenesis. Also in this case, the in vivo treatment of mice with this compound lead to a suppression of implanted human tumors.

C*one*-calix[4]arenes can also provide a single-point recognition of studied target proteins, by exploiting their aromatic cavity, assisted by the action of proper functional groups at its upper rim, to host lipophilic or positively charged amino acid residues. Examples of this kind of protein-ligand interactions have been reported for the tetrasulfonato calix[4]arene **4** (**Figure 9a**). This compound[96] is soluble in water and buffer solution and, at neutral pH, adopts a rigid cone conformation in these media, which is reinforced by a network of hydrogen bonds that involves the phenolic and phenolate groups at its lower rim. The ability of **4** to bind cyt *c* was demonstrated by Crowley et al.[97], according to the results of NMR titrations and to the resolution of the X-ray crystal structure of the corresponding protein-ligand

complex. More in details, the X-ray diffraction data showed that this compound is able to recognize three lysines by exploiting, in each cases, the inclusion of the alkyl part of the side chains into its hydrophobic cavity and the interaction of the ammonium group toward the four sulfonate units at its upper rim (**Figure 9b**). The three molecules of **4** were located at the interface of a cytochrome *c* dimer, whose aggregation was supposed to be promoted by the burial of the target lysine residues. In fact, additional CH-π and van der Waals contacts, formed between apolar amino acid side chains of the opposite protein monomer and, respectively, the external part of the calix[4]arene cavity and the phenolic/phenolate moieties at the lower rim, were supposed to allow the oligomerization of cyt *c* during the co-crystallization process (*molecular glue effect*).

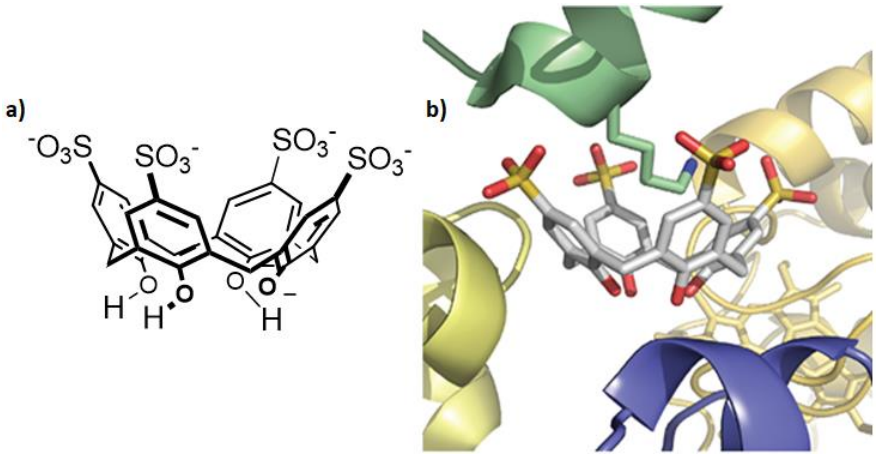


*Figure 9: **a)** sulfonato-calix[4]arene tested as ligand for cytochrome c and **b)** inclusion of Lys4 of target protein in the crystal structure of the cyt c-**4** complex (figure taken from **Ref. 97**).*

Similar results were reported for the lysozyme-**4** system, after the resolution of the X-ray crystal structure of the corresponding protein-ligand complex[98]. In this case the calix[4]arene derivative **4** recognized selectively the C-terminal Arg128 instead of lysine residues, but the feature of the binding were very similar to those observed for the **4**-cyt *c* couple, since the side chain of the bound amino acid sits in the macrocyclic cavity and the guanidinium group established salt bridge interactions with the four sulfonate groups. The selectivity observed for Arg128 versus other arginine residues was ascribed to the fact that it is the most exposed residue from the protein surface and thus the most accessible for the ligand. Also in this case the ligand was reported to act as molecular glue, by inducing the crystallization of the target protein as tetramer. A change in selectivity was reported when, in a subsequent study, bis-methylation of the lysine side chains of lysozime was carried out[99]. In this case, in fact, **4** turned to bind preferentially to the dimethylated Lys116, which is the most accessible between the six lysine residues present in the protein backbone. The results of NMR and X-ray diffraction experiments pointed out that, on the contrary of what reported for the previous crystal structures, the dimethyl ammonium head group penetrates in the cavity of the calix[4]arene to establish a cation-π interaction, while the remaining part of the lysine side chain stays outside the cavity. The selectivity of **4**

for dimethylated lysine side chains has been reported as a mimic of the action of proteins involved in gene regulation and signaling pathways, such as chromodomains.

The cone-calix[4]arene scaffold has been reported to work also as stopper of lipophilic pockets deriving from the quaternary structure of proteins. Noteworthy example of this function has been described by De Mendoza et al., in a study where compounds **5**-**9** (**Figure 10**) were tested as inhibitors of the activity of voltage-dependent potassium channels[100].

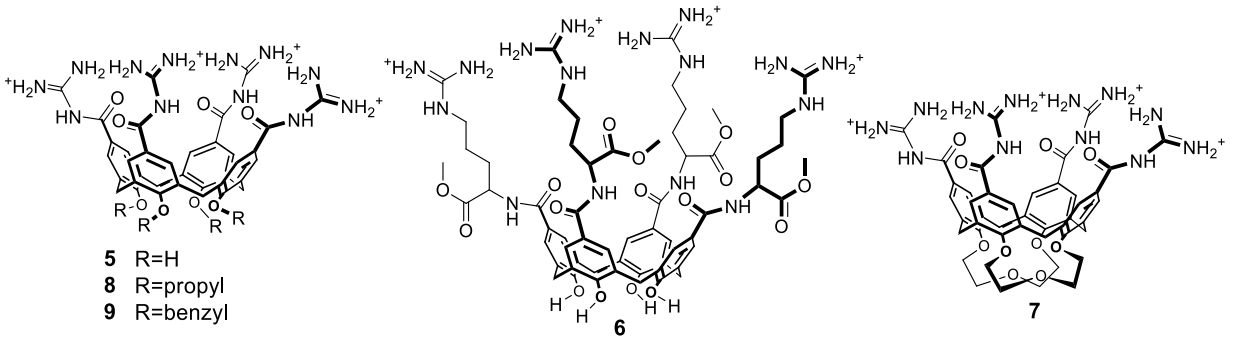


Figure 10: Cone-calix[4]arene derivatives tested as stopper of voltage-dependent ionic channels.

The idea behind this study was that the cup-like shape of the tested compounds would be suitable to fit the portals of these channels, while the guanidinium groups at their upper rim would be exploited to bind the carboxylates moieties present at the channel entry. The tested compounds were able to inhibit reversibly the action of the Shaker potassium channels and the higher inhibitor activity was reported for **7**, whose *rigid-cone* structure, imparted by the presence of crown ether-like motifs at its lower rim, allowed a better fit to the target channel surface. However, the presence of substituents at the lower rim of tested calix[4]arene was not always reported as beneficial for their biological activity. For example, compound **8** and **9** were found to cause an important, but irreversible, decrease of ionic current for the channels expressing oocytes and this drawback was attributed to the amphiphilic nature of the ligands, provided by their propyl and benzyl groups at lower rim, respectively. In fact, both calix[4]arene derivatives were supposed to play disruptive detergent action against the cell membranes containing the target ionic channels. Control experiments, performed by using carboxylate- or ammonium groups-bearing calix[4]arenes and a monofunctional model compound of **5**, pointed out no appreciable activity for the tested derivatives, thus confirming the importance of both the macrocyclic cavity and the positively charged guanidinium groups to obtain effective inhibitors.

Compound **10** and **11**, synthetized with a similar design by the same authors and shown in **Figure 11**, were instead found to restore the physiological activity of a mutated form of the tetrameric p53 transcription factor[101].

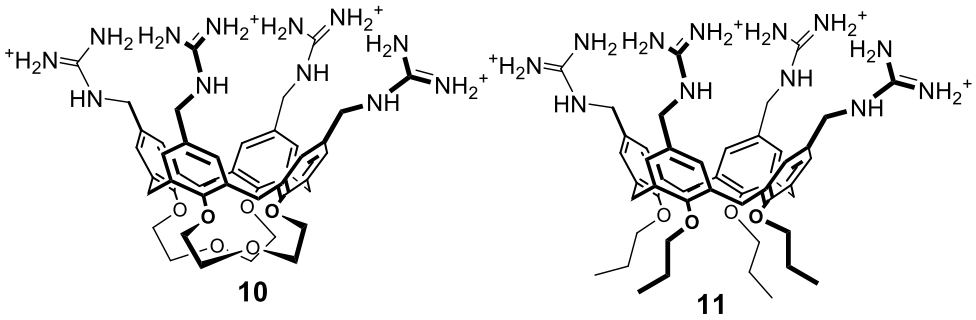


Figure 11: Calix[4]arene derivatives used to restore the activity of p53 mutated form.

The wild type form of p53 has an important anti-tumor activity in the human body, acting as "genome guardian" when its association as tetramer takes place correctly. A crucial role in this process is covered by Arg337, which recognize corresponding Asp residues on the opposite protein units, thus stabilizing the resulting oligomer. Unfortunately, p53 can be sometimes observed in a mutated form, where Arg337 is replaced by a histidine residue which, being not fully protonated at physiological pH, lead to the formation of unstable tetramers, whose anti-tumor activity is strongly inhibited. A common feature of the tetramers, formed upon association of both the wild type and the mutated form of p53, is the presence of a Glu336 and Glu339 couple on the extremity of each of their four protein monomers, which provides a squared distribution of negative charges at the two bases of each oligomeric protein. DSC and circular dichroism experiments indicated this set of negatively charged residues as effective anchoring point for the four guanidinium groups of **10**, which was proposed to approach each extremity of the tetrameric mutated protein by its lower rim side. The stability of the resulting tetramer was strongly enhanced in the presence of **10** at 400 mM concentration and the formation of complex containing two calix[4]arene molecules and the associated protein was confirmed by ESI mass spectrometry. As expected, the stability of the wild-type tetrameric protein was not affected by the presence of the ligand.

Compound **11**, alkylated with simple propyl chains at the lower rim, was even more effective than its more rigid analogue **10** into keep together the four units of mutated p53[102]. In fact, DSC studies showed that a 25mM concentration of this compound is sufficient to restore the oligomerization of the target protein and this gain of activity was assigned to the different feature of its molecular scaffold. More in details, it was supposed that the residual flexibility of **11** allowed a better rearrangement of its guanidinium groups at the upper rim to fit the target Glu residues and that the propyl chains at the lower rim can penetrate the lipophilic pocket of the protein to establish additional hydrophobic interactions.

Other important examples of protein recognition by calix[4]arene-based ligands are those concerning the interaction with carbohydrate binding proteins[15,80,81], also called lectins. Lectins are involved in many physiological and pathological process, in which they act as receptors for specific binding partners by

exploiting interactions with multiple glycoside moieties, disposed on the considered ligands in specific reciprocal orientation. This kind of binding process can be described as a particular case of multivalency, which is commonly called *glycoside cluster effect*[103].

Several calix[n]arene derivatives of different valency and conformation have been functionalized with saccharide units in order to interact with lectins and influence the biological processes in which they are involved. Focusing on calix[4]arenes, relevant example of these ligands, also called *glycocalixarenes*, are represented by compounds **12**-**14** (**Figure 12**)[104,105], blocked in different macrocyclic geometries, and functionalized at the lower rim with different glycoside units.

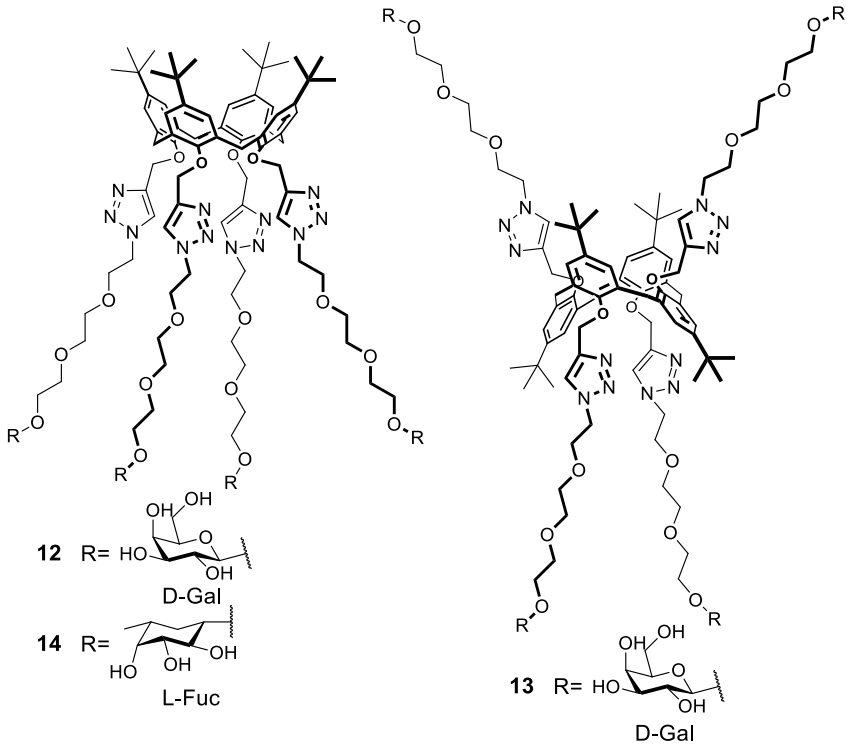


Figure 12: Glycocalix[4]arenes of different conformation tested as inhibitors of PA-IL and/or PA-IIL lectines.

The target chosen for ligands **12** and **13**, decorated with four β-galactoside units, was the tetrameric galactose-binding lectins PA-IL from *Pseudomonas aeruginosa*[105] (a bacterium responsible of chronic respiratory diseases) and their inhibitor activity was found to be dependent on their conformational features. In fact, ITC titrations pointed out the superiority of the 1,3-alternate-calix[4]arene **12** over the corresponding conical isomer **13** into bind target proteins, with dissociation constants of 176 nM and 420 nM and stoichiometry for the resulting complexes of 1:4 and 1:3, respectively. Thus, the stoichiometry of the complexes indicates that the lower binding affinity of **13** rises from the fact that only three of its four binding units can interact with PA-IL. Since it is known that the binding sites of this protein are located two by two in the minor faces of its parallelepipedon-like structure, to explain these findings, it was proposed that **13** provides an orientation of its four galactoside groups which is suitable to bridge two protein tetramers at time, whereas for the cone-calix[4]arene derivative **14** this geometry

of complexation is disadvantaged, being its binding groups are all disposed at the same side of the macrocyclic scaffold.

These suggestions were formulated according to docking calculations and then confirmed by AFM analysis of samples containing proper amount of PA-IL and **13**. Similar investigations were performed by other authors in an analog study, where **13** and the parent compound **14,** functionalized with four L-fucose units instead of four galactoside groups, were tested respectively as inhibitors PA-IL of PA-IIL[104], which is another important virulence factor expressed by *Pseudomonas aeruginosa*. The experiments, carried out *in vivo* on mice, pointed out a block of bacterial proliferation upon treatment with both the calix[4]arene-based ligands, thus confirming the inhibition of the action of the target lectins. At contrary to what observed for the **13**-PA-IL system, it was proposed that **14** interact with PA-IIL by binding four different protein tetramers instead of bridging two of them at time. In fact, despite the structures of PA-IL and PA-IIL are very similar, it is known that in the latter protein aggregate the binding sites for fucose are farer and, for this reason, an interaction with two of them by same the calix[4]arene derivative was supposed to be unlikely.

Other studies focused on the recognition of lectins were performed by using as ligands calix[4]-, calix[6]- and calix[8]arene functionalized at their upper rim with galactose or lactose units[106]. The studied lectins were gal-1, gal-3 and gal-4, belonging to different subfamilies of galectins, which are proteins able to bind selectively galactose moieties on cell surfaces, resulting involved in the development of tumors and in the progression of metastases. Focusing on the results obtained for the calix[4]arene-based ligand and, among them, on the more active L-lactosylated derivatives, the highest inhibitor activity was reported for compounds **15** and **16** (**Figure 13**).

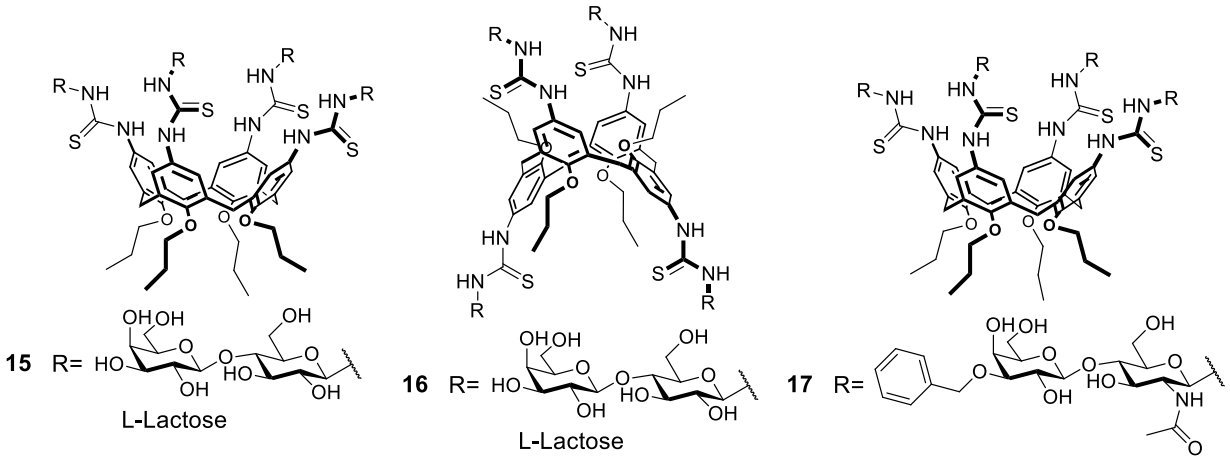


Figure 13: Lactosylcalix[4]arenes tested as anti-adhesive agents for different subfamilies of galectins.

Compounds **15** and **16** were found to inhibit very effectively the adhesion of, respectively, gal-4 and gal-1 to human pancreatic carcinoma cells. Interestingly, for gal-1 a high ligand selectivity was observed, depending on the conformation of tested compounds. In fact, at contrary to what just mentioned for the *1,3-alternate* calix-[4]arene derivative **16**, the corresponding *cone*- isomer **15** has been reported as

an extremely poor inhibitor towards the target protein. The opposite was observed in the case of gal-3, whose galectin adhesion to colon adenocarcinoma cells, was strongly suppressed by **15**, but almost not affected by the presence of **16**. The selective inhibitor properties observed for **16** have an important medical relevance because, under appropriate experimental conditions, gal-1 has been reported to play a beneficial role into induce programmed death of some tumor cell lines, whereas gal-3 was found to disadvantage this action. Hence, ligands able to block the activity of only the latter protein, without influencing the activity of the former, look promising for the development of effective anti-tumoral drugs.

Compound **17** (**Figure 13**)[107] was subsequently obtained as a development of the aforementioned *cone*-calix[4]arene, by inserting at the upper rim an N-acetyl-lactosamine (LacNAc) moiety, bearing a *m*-methoxybenzyl group at the galactose 3' position of each lactosyl unit. This modification has been suggested according to the X-ray crystal structures of complexes including lectins and this LacNAc derivative, that pointed out the possible use of additional interactions with both the aromatic and the amide groups of the sugar during the binding. Thanks to this structural adjustment, the selectivity of **17** towards the two target proteins was improved respect to the one observed for **16**, since this new compound resulted to inhibit the adhesion of gal-3 to target tumor cells better than its precursor, while was completely inactive as ligand for gal-1.

4.5 Conclusions and outline of the research

Proteins are involved in a plethora of physiological and pathological processes which are driven by the recognition of specific counterparts. The detailed understanding of the structural and thermodynamic features of these processes is crucial to extract the information on the properties that synthetic ligands could have to effectively bind target proteins and to influence their biological activity, possibly with a beneficial effect under a medicinal point of view.

In fact, the rational design of ligands as inhibitors for certain proteins is a central point for the development of new drugs, while the possibility to restore the activity of other types of proteins, by using synthetic molecules able to reinforce and stabilize their oligomeric structure, has been reported as a promising approach for the same purpose.

According to what observed in nature, the exploitation of the multivalency effect, that is the gain of overall binding affinity that can be achieved by utilizing several identical binding units to establish non-covalent interactions with the target binding site, has been reported as an effective strategy for the synthesis of different artificial ligands for proteins. A family of ligands that effectively takes advantage of this effect is based on the use of calix[n]arenes as molecular scaffolds, since these macrocycles gives the possibility to insert easily several functional groups able to contemporary interact on the target protein surfaces at both the rims of their cavity.

Among this kind of derivatives, calix[4]arenes are worth of a particular attention since they can be blocked in different conformations which could induce a sharp selectivity even towards very similar target proteins, thanks to the different orientation of the binding groups imparted by the aromatic scaffold. Of course, the way in which the recognition of target proteins is achieved depends also on the nature of the substituent present on the calix[4]arene scaffold, which cooperate with the aromatic cavity of these compounds to establish different modes of binding. For example, as described in the previous pages of this chapter, Crowley et al. reported that the sulfonato-calix[4]arene **4** is able to recognize single positively charged amino acid residues on the surface of both cytochrome c and of lysozyme, thanks to the synergic action of the negatively charged groups at its upper rim and of the macrocyclic cavity, which can host a part of the bound amino acid side chains[97–99]. Moreover, it has been found that this compound is able to induce the crystallization of target proteins as oligomers.

On the basis on these observations we decided, together with the same group of research, to study the interaction between cyt *c* and compound **18** (**Figure 14a**), blocked in a rigid-cone conformation by two crown-ether like motifs at the lower rim and bearing four C-terminal L-alanine moieties at the upper rim as binding units for the positively charged regions of the protein surfaces.

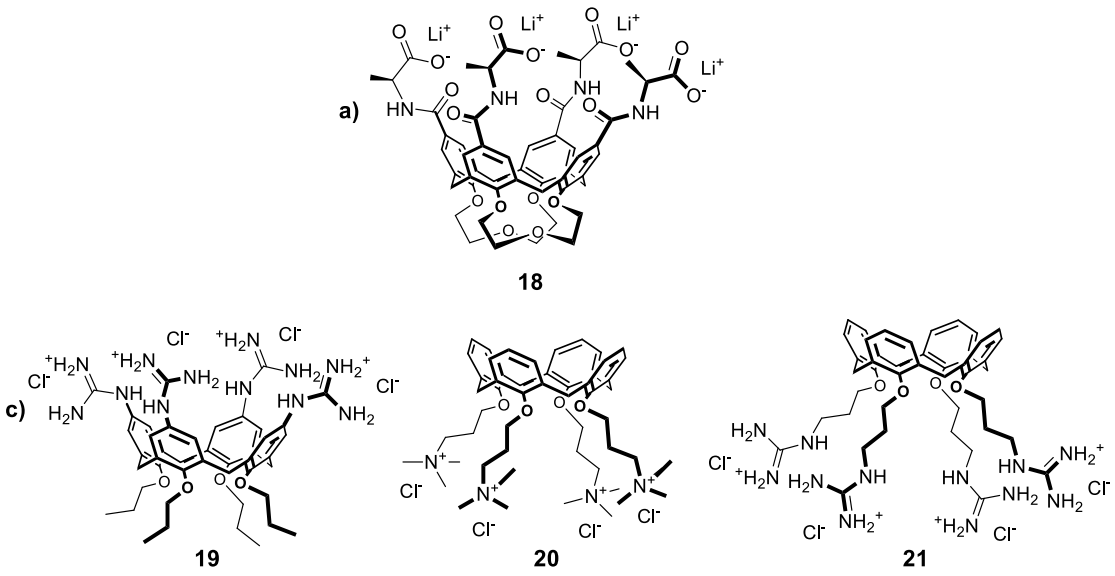


*Figure 14: **a)** Tetraalanino- and **b)** tetraguanidinium-calix[4]arene derivatives tested as ligand for cytochrome c and GB 1, respectively.*

The binding features of this system were investigated by NMR titrations and the results of these studies will be described in the first part of the following chapter of this thesis.

The second part of the same chapter will be instead focused on the description of ITC investigations for the protein-ligand systems constituted by the B1 immunoglobulin-binding domain of the streptococcal protein G (GB1) and compounds **19**–**21** (**Figure 14b**).These experiments were carried out to strengthen, with detailed thermodynamic data, the results obtained in a previous work[108], where the binding processes involving these protein-ligand couples were already studied by NMR titrations.

4.6 Bibliography

(1) Berg, Jeremy B, Tymoczko, John L, Stryer, L. *Biochemistry, 5th Edition*, 5th ed.; W. H. Freeman: York, N., Ed.; **2002**.

(2) Gething, M.-J.; Sambrook, J. *Nature* **1992**, *355* (6355), 33.

(3) Fersht, A. *Enzyme Structure and Mechanism*, 2nd ed.; W. H. Freeman: New York, Ed.; **1984**.

(4) Medzhitov, R. *Nature* **2007**, *449* (7164), 819.

(5) Pawson, T.; Nash, P. **2000**, No. 416, 1027.

(6) Kim, S. H.; Turnbull, J.; Guimond, S. *J. Endocrinol.* **2011**, *209* (2), 139.

(7) Norton, J. D. *J Cell Sci* **2000**, *113 (Pt 2*, 3897.

(8) Källström, H.; Liszewski, M. K.; Atkinson, J. P.; Jonsson, a B. *Mol. Microbiol.* **1997**, *25* (4), 639.

(9) Morin, P. J. *Cancer Res.* **2005**, *65* (21), 9603.

(10) Zhou, M.; Barrette, T. R.; Kumar-Sinha, C.; Ghoshk, D.; Varambally, S.; Dhanasekaran, S. M.; Zhou, M.; Barrette, T. R.; Kumar-Sinha, C.; Sanda, M. G.; Ghosh, D.; Pienta, K. J.; Sewalt, R. G. A. B.; Otte, A. P.; Rubin, M. A.; Chinnaiyan, A. M. *Nature* **2002**, *419* (OCTOBER), 388.

(11) Friedl, P.; Wolf, K. *J. Cell Biol.* **2010**, *188* (1), 11.

(12) Glickman, N. S.; Yelon, D. *Semin. Cell Dev. Biol.* **2002**, *13* (6), 507.

(13) Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100* (7), 2479.

(14) Stites, W. E. *Chem. Rev.* **1997**, *97* (5), 1233.

(15) Giuliani, M.; Morbioli, I.; Sansone, F.; Casnati, A. *Chem. Commun.* **2015**, *51* (75), 14140.

(16) Chothia, C.; Janin, J. *Nature* **1975**, *256* (5520), 705.

(17) Perozzo, R.; Folkers, G.; Scapozza, L. *J. Recept. Signal Transduct.* **2004**, *24* (1–2), 1.

(18) Janin, J. *Prog. Biophys. Mol. Biol.* **1995**, *64* (2–3), 145.

(19) Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q. *Int. J. Mol. Sci.* **2016**, *17* (2), 144.

(20) Yin, H.; Hamilton, A. D. *Angew. Chem. Int. Ed. Engl.* **2005**, *44* (27), 4130.

(21) Sheng, C.; Dong, G.; Miao, Z.; Zhang, W.; Wang, W. *Chem. Soc. Rev.* **2015**, *44* (22), 8238.

(22) Qureshi, S. a; Kim, R. M.; Konteatis, Z.; Biazzo, D. E.; Motamedi, H.; Rodrigues, R.; Boice, J. a; Calaycay, J. R.; Bednarek, M. a; Griffin, P.; Gao, Y. D.; Chapman, K.; Mark, D. F. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (21), 12156.

(23) Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Angew. Chem. Int. Ed. Engl.* **2006**, *45* (15), 2348.

(24) Olsson, T. S. G.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E. *J. Mol. Biol.* **2008**, *384* (4), 1002.

(25) Liu, S.; Ji, X.; Tao, Y.; Tan, D.; Zhang, K.-Q.; Fu, Y.-X. *Protein Eng.* **2009**, 207.

(26) Amzel, L. M. In *Energetics of Biological Macromolecules, Part C*; Methods in Enzymology; Academic Press, 2000; Vol. 323, pp 167–177.

(27) MacRaild, C. A.; Daranas, A. H.; Bronowska, A.; Homans, S. W. *J. Mol. Biol.* **2007**, *368* (3), 822.

(28) Dunitz, J. D. *Chem. Biol.* **1995**, *2* (11), 709.

(29) Lumry, R.; Rajender, S. *Biopolymers* **1970**, *9* (10), 1125.

(30) Eftink, M. R.; Anusiem, A. C.; Biltonen, R. L. *Biochemistry* **1983**, *22* (16), 3884.

(31) Olsson, T. S. G.; Ladbury, J. E.; Pitt, W. R.; Williams, M. A. *Protein Sci.* **2011**, *20* (9), 1607.

(32) Qian, H. *J. Chem. Phys.* **1998**, *109* (22), 10015.

(33) Chang, C.-E.; Gilson, M. K. *J. Am. Chem. Soc.* **2004**, *126* (40), 13156.

(34) Breiten, B.; Lockett, M. R.; Sherman, W.; Fujita, S.; Al-Sayah, M.; Lange, H.; Bowers, C. M.; Heroux, A.; Krilov, G.; Whitesides, G. M. *J. Am. Chem. Soc.* **2013**, *135* (41), 15579.

(35) Grunwald, E.; Steel, C. *J. Am. Chem. Soc.* **1995**, *117* (21), 5687.

(36) Ryde, U. *Medchemcomm* **2014**, *5*, 1324.

(37) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem. Int. Ed. Engl.* **1998**, *37* (20), 2754.

(38) Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Koksch, B.; Dernedde, J.; Graf, C.; Knapp, E. W.; Haag, R. *Angew. Chem. Int. Ed. Engl.* **2012**, *51* (42), 10472.

(39) Jencks, W. P. *Proc. Natl. Acad. Sci.* **1981**, *78* (7), 4046.

(40) Kitov, P. I.; Bundle, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 16271.

(41) Callaway, E. *Nature* **2017**, 167.

(42) Freyer, M. W.; Lewis, E. A. In *Biophysical Tools for Biologists, Volume One: In Vitro Techniques*; Methods in Cell Biology; Academic Press, 2008; Vol. 84, pp 79–113.

(43) Velázquez-Campoy, A.; Ohtaka, H.; Nezami, A.; Muzammil, S.; Freire, E. In *Current Protocols in Cell Biology*; John Wiley & Sons, Inc., 2001.

(44) Falconer, R. J. *J. Mol. Recognit.* **2016**, No. May, 504.

(45) Pierce, M. M.; Raman, C. S.; Nall, B. T. *Methods* **1999**, *221*, 213.

(46) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. *Anal. Biochem.* **1989**, *179* (1), 131.

(47) Indyk, L.; Fisher, H. F. In *Energetics of Biological Macromolecules Part B*; Methods in Enzymology; Academic Press, 1998; Vol. 295, pp 350–364.

(48) Zhang, Y. L.; Zhang, Z. Y. *Anal. Biochem.* **1998**, *261* (2), 139.

(49) Bradshaw, J. M.; Mitaxov, V.; Waksman, G. *J. Mol. Biol.* **1999**, *293* (4), 971.

(50) Sigurskjold, B. W. *Anal. Biochem.* **2000**, *277* (2), 260.

(51) Ferentz, A. E.; Wagner, G. *Q. Rev. Biophys.* **2000**, *33* (1), 29−65.

(52) Zuiderweg, E. R. P. *Biochemistry* **2002**, *41* (1), 1.

(53) Takeuchi, K.; Wagner, G. *Curr. Opin. Struct. Biol.* **2006**, *16* (1), 109.

(54) Clarkson, J.; Campbell, I. D. *Biochem. Soc. Trans.* **2003**, *31* (Pt 5), 1006.

(55) Meyer, B.; Peters, T. *Angew. Chem. Int. Ed. Engl.* **2003**, *42* (8), 864.

(56) Fielding, L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51* (4), 219.

(57) Williamson, M. P. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1.

(58) Lian, L. Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *71*, 59.

(59) Goger, M. J.; McDonnell, J. M.; Cowburn, D. *Spectroscopy* **2003**, *17* (2–3), 161.

(60) Shapiro, M. J., Wareing, J. R. *Curr. Opin. Drug Discov. Devel.* **1999**, *2* (4), 396.

(61) Ban, D.; Gossert, A. D.; Giller, K.; Becker, S.; Griesinger, C.; Lee, D. *J. Magn. Reson.* **2012**, *221*, 1.

(62) Zhu, G.; Yao, X. *Prog. Nucl. Magn. Reson. Spectrosc.* **2008**, *52* (1), 49.

(63) Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (23), 13585.

(64) Wu, K. **1997**, *94* (November), 12366.

(65) Skinner, A. L.; Laurence, J. S. *J. Pharm. Sci.* **2008**, *97* (11), 4670.

(66) McPherson, A. *Eur. J. Biochem.* **1990**, *189* (1), 1.

(67) Gavira, J. A. *Arch. Biochem. Biophys.* **2016**, *602*, 3.

(68) Krauss, I. R.; Merlino, A.; Vergara, A.; Sica, F. *Int. J. Mol. Sci.* **2013**, *14* (6), 11643.

(69) Luft, J. R.; Collins, R. J.; Fehrman, N. A.; Lauricella, A. M.; Veatch, C. K.; DeTitta, G. T. *J. Struct. Biol.* **2003**, *142* (1), 170.

(70) Bolanos-Garcia, V. M.; Chayen, N. E. *Prog. Biophys. Mol. Biol.* **2009**, *101* (1–3), 3.

(71) Hassell, A. M.; An, G.; Bledsoe, R. K.; Bynum, J. M.; Carter, H. L.; Deng, S. J. J.; Gampe, R. T.; Grisard, T. E.; Madauss, K. P.; Nolte, R. T.; Rocque, W. J.; Wang, L.; Weaver, K. L.; Williams, S. P.; Wisely, G. B.; Xu, R.; Shewchuk, L. M. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *63* (1), 72.

(72) Chayen, N. E. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1998**, *54* (1), 8.

(73) Benvenuti, M.; Mangani, S. *Nat. Protoc.* **2007**, *2* (7), 1633.

(74) Cudney, R.; Patel, S.; Weisgraber, K.; Newhouse, Y.; McPherson, A. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50* (4), 414.

(75) Korczyńska, J.; Hu, T. C.; Smith, D. K.; Jenkins, J.; Lewis, R.; Edwards, T.; Brzozowski, A. M. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2007**, *63* (9), 1009.

(76) Matthews, S. E.; Beer, P. D. *Supramol. Chem.* **2005**, *17* (6), 411.

(77) Casnati, A. *Chem. Commun. (Camb).* **2013**, *49*, 6827.

(78) Arduin, A.; Pochini, A.; Secchi, A.; Ugozzoli, F. In *Calixarenes 2001*; Asfari, Z., Böhmer, V., Harrowfield, J., Vicens, J., Saadioui, M., Eds.; Springer Netherlands: Dordrecht, 2001; pp 457–475.

(79) Casnati, A.; Sciotto, D.; Arena, G. In *Calixarenes 2001*; Asfari, Z., Böhmer, V., Harrowfield, J., Vicens, J., Saadioui, M., Eds.; Springer Netherlands: Dordrecht, 2001; pp 440–456.

(80) Kim, Taisum; Nimse, S. B. *Chem Soc Rev* **2013**, *42*, 366.

(81) Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. **2010**, 2715.

(82) Danylyuk, O.; Suwinska, K. *Chem. Commun.* **2009**, No. 39, 5799.

(83) Casnati, A.; Sansone, F.; Ungaro, R. *Acc. Chem. Res.* **2003**, *36* (4), 246.

(84) Synthesis, S.; Studies, N. M. R.; Duynhoven, J. P. M. Van; Janssen, R. G.; Verboom, W.; Franken, S. M.; Casnati, A.; Pochini, A.; Ungaro, R.; De Mendoza, J.; Nieto, P. M.; Prados, P.; Reinboudt, D. N. **1994**, No. 15, 5814.

(85) Groenen, L. C.; Ruël, B. H. M.; Casnati, A.; Timmerman, P.; Verboom, W.; Harkema, S.; Pochini, A.; Ungaro, R.; Reinhoudt, D. N. *Tetrahedron Lett.* **1991**, *32* (23), 2675.

(86) Perret, F.; Coleman, A. W. *Chem. Commun.* **2011**, *47* (26), 7303.

(87) Peters, M. S.; Schrader, T. In *Calixarenes and Beyond*; Neri, P., Sessler, J. L., Wang, M.-X., Eds.; Springer International Publishing: Cham, 2016; pp 627–669.

(88) Zadmard, R.; Alavijeh, N. S. *RSC Adv.* **2014**, *4*, 41529.

(89) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280* (1), 1.

(90) Clackson, T.; Wells, J. A. *Science (80).* **1995**, *267* (5196), 383 LP.

(91) Hamuro, Y.; Calama, M. C.; Park, H. S.; Hamilton, A. D. *Angew. Chem. Int. Ed. Engl.* **1997**, *36* (23), 2680.

(92) Sebti, S. M.; Hamilton, A. D. **2000**, 6566.

(93) Wei, Y.; McLendon, G. L.; Hamilton, A. D.; Case, M. A.; Purring, C. B.; Lin, Q.; Park, H. S.; Lee, C. S.; Yu, T. *Chem. Commun. (Camb).* **2001**, No. 17, 1580.

(94) Park, H. S.; Lin, Q.; Hamilton, A. D. **1999**, No. 10, 8.

(95) Sun, J.; Wang, D.-A.; Jain, R. K.; Carie, A.; Paquette, S.; Ennis, E.; Blaskovich, M. a; Baldini, L.; Coppola, D.; Hamilton, A. D.; Sebti, S. M. *Oncogene* **2005**, *24* (29), 4701.

(96) Shinkai, S.; Araki, K. *J. Am. Chem. Soc.* **1990**, No. 112, 9053.

(97) McGovern, R. E.; Fernandes, H.; Khan, A. R.; Power, N. P.; Crowley, P. B. *Nat Chem* **2012**, *4* (7), 527.

(98) McGovern, R. E.; McCarthy, A. A.; Crowley, P. B. *Chem. Commun.* **2014**, *50* (72), 10412.

(99) McGovern, R. E.; Snarr, B. D.; Lyons, J. A.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. *Chem Sci* **2015**, *6* (1), 442.

(100) Martos, V.; Bell, S. C.; Santos, E.; Isacoff, E. Y.; Trauner, D.; De Mendoza, J. **2009**.

(101) Gordo, S.; Martos, V.; Santos, E.; Menéndez, M.; Bo, C.; Giralt, E.; de Mendoza, J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16426.

(102) Gordo, S.; Martos, V.; Vilaseca, M.; Menéndez, M.; De Mendoza, J.; Giralt, E. *Chem. - An Asian J.* **2011**, *6* (6), 1463.

(103) Lee, R. T.; Lee, Y. C. *Glycoconj. J.* **2000**, *17* (7–9), 543.

(104) Boukerb, A. M.; Rousset, A.; Galanos, N.; Méar, J. B.; Thépaut, M.; Grandjean, T.; Gillon, E.; Cecioni, S.; Abderrahmen, C.; Faure, K.; Redelberger, D.; Kipnis, E.; Dessein, R.; Havet, S.; Darblade, B.; Matthews, S. E.; De Bentzmann, S.; Guéry, B.; Cournoyer, B.; Imberty, A.; Vidal, S. *J. Med. Chem.* **2014**, *57* (24), 10275.

(105) Cecioni, S.; Lalor, R.; Blanchard, B.; Praly, J. P.; Imberty, A.; Matthews, S. E.; Vidal, S. *Chem. Eur. J.* **2009**, *15* (47), 13232.

(106) André, S.; Sansone, F.; Kaltner, H.; Casnati, A.; Kopitz, J.; Gabius, H. J.; Ungaro, R. *ChemBioChem* **2008**, *9* (10), 1649.

(107) André, S.; Grandjean, C.; Gautier, F.-M.; Bernardi, S.; Sansone, F.; Gabius, H.-J.; Ungaro, R. *Chem. Commun.* **2011**, *47* (21), 6126.

(108) Giuliani, M. *New multivalent calixarene-based ligands for protein recognition and inhibition*; PhD Thesis, Parma University, **2015**.

Chapter 5

Calix[4]arene-based ligands for model proteins

5.1 Introduction

Recognition events involving proteins and specific binding partners, which can be both other bio-macromolecules or small molecule ligands, are at the basis of several biological processes[1–3]. The development of artificial ligands able to promote, modulate or inhibit these processes, with the final goal to create bioactive compounds that can have an application in medicine, has attracted much attention over the years[4–6]. A detailed characterization of the studied protein-ligand interactions events is the first step to achieve this purpose[7–10], since it allows to obtain crucial information about the features that synthetic molecules could have to bind target proteins, by mimicking the action of their natural analogues. As described in the previous chapter, the calix[4]arene scaffold has been effectively used for the creation of ligands able to recognize the surface of different proteins and to influence their physico-chemical properties or biological activity[3,11,12]. According to these observations we decided to test some calix[4]arene derivatives as ligands for different model proteins.

The characterization of the binding processes involving these protein-ligand systems was carried out during a five month-placement period spent in Galway at the National University of Ireland, under the supervision of Dr. Peter Crowley, and will be presented in the following pages.

5.2 Studied proteins

5.2.1 Cytochrome c

Cytochrome *c* (Cyt *c*)[13] (**Figure 1a**) is a small, monomeric, highly water soluble, heme-containing protein found in the inner mitochondrial membrane of several organisms, ranging from yeasts to humans. Among these organisms the primary sequence of this protein is highly conserved and consist in around 100 amino acid residues, for a molecular weight of approximately 12KDa[14].

Cyt *c* plays different biological roles[15], which are driven by the occurrence of its interaction with other proteins or small molecules. For example, its interactions within complex III (Coenzyme Q – Cyt *c* reductase) and complex IV (Cyt *c* oxidase) of the electron transport chain in mitochondria are known to result in the transfer of one electron from the former to the latter protein[16], which constitutes an important stage of the cellular respiration. Cyt *c* is also known to work as intermediate in cell apoptosis[17,18]. In this case the binding of this protein to the IP3 receptor[19], which takes place in the cytosol after that appropriate stimuli cause its release from the mitochondria, has been reported to activate a cascade of events that results in the disruption of the considered cell. Moreover, cyt *c* has also been found to play enzymatic roles, by catalyzing different redox reaction such as the removal of superoxide and hydrogen peroxide from mitochondria or the oxidation of different electron donor molecules[20,21].

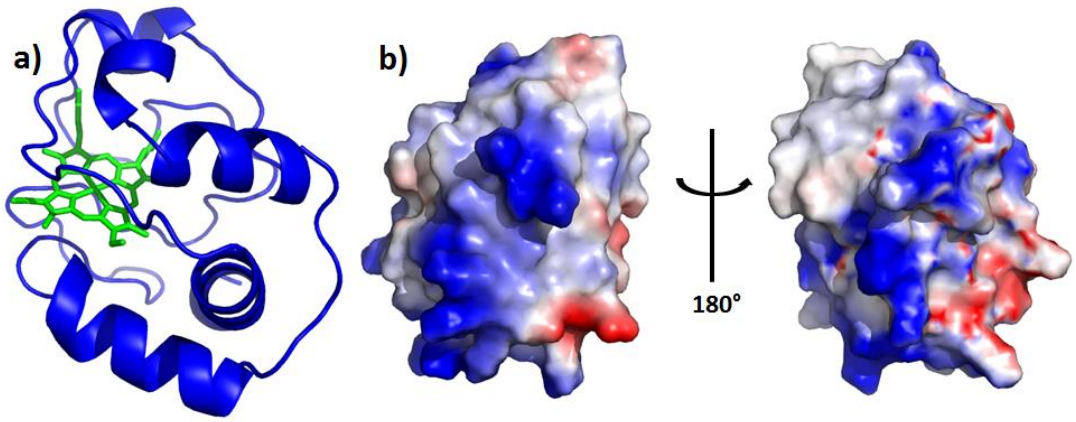


Figure 1: ***a)*** *Cartoon representation of cytochrome c were the backbone of the protein is shown in blue and the heme group is colored in green.* ***b)*** *Electrostatic potential of the protein surfaces with blue and red areas corresponding to positively and negatively charged regions, respectively.*

The presence of standard protocols for its over-expression from bacteria and its subsequent purification[22,23], together with the large availability of data about its structure[24,25] and its interactions with different binding partners[26–30], renders this protein suitable for different practical application, such as the development of electron transfer devices for the construction of biosensors[31–33] and its usage as standard for different laboratory analysis. A noteworthy example of the latter type of applications is the exploitation of cyt *c* as model protein for the study of the interaction with new artificial ligands which are supposed to influence its properties. This usage is particularly appropriated when the tested ligand bears negatively charged groups on their molecular scaffold, since the protein, with a pI around 10[34], own large positively charged regions on its surfaces (**Figure 1b**) because of the presence of several Arg and Lys side chains in its primary sequence (in particular, for cyt *c* from *Saccharomyces cerevisiae*, which is commonly used for these investigations, the number of these amino acid residues is 8 and 11, respectively).

Among the different ligands shown to be able to interact with cyt c, those synthetized by functionalization of the *cone*-calix[4]arene[28,29,35] molecular scaffold were reported to provide a remarkable binding ability and, in this chapter, it will be presented the characterization of the new calix[4]arene-based ligand system.

5.2.2 B1 Immunoglobulin-binding domain of streptococcal protein G

Protein G from *Group G Streptococcus* is a large membrane protein which is believed to help these bacteria to elude the immune response during infection processes, by interacting with immunoglobulin (Ig) and other proteins[36]. Its B1 domain (**Figure 2a**), commonly indicated as GB1, is one of those that are directly involved in these binding events and, because of its important biological role, it has been the object of several investigations that have afforded detailed characterization of its structure[37–39] and mechanism of action[40–42].

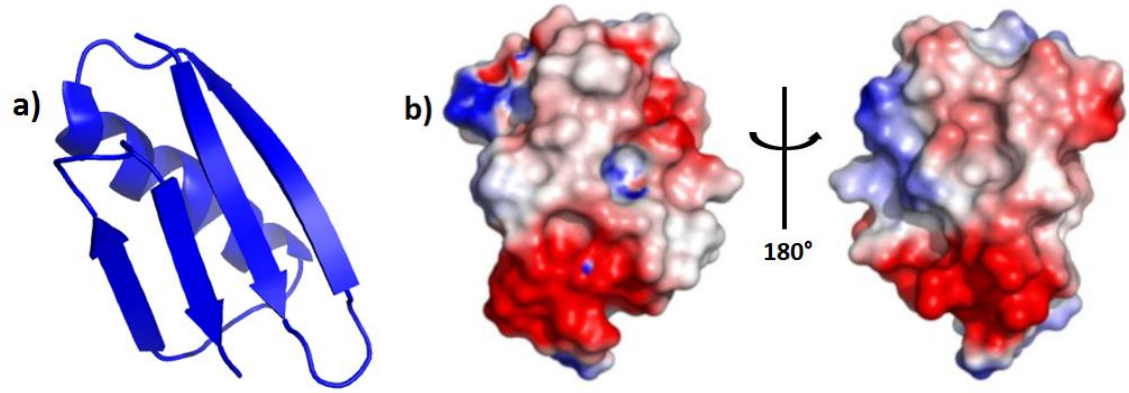


Figure 2: ***a)*** *Cartoon representation of GB1 and* ***b)*** *electrostatic potential of the protein surfaces with blue and red areas corresponding to positively and negatively charged regions, respectively (Figure taken from* ***Ref. 49****).*

GB1 is a small protein composed of 56 amino acids residues which are arranged in an α-helix and two hairpins in its secondary structure. It is highly rich in Asp or Glu amino acid residues (10 in total) and, with a pI of 4.0[43], its surfaces are mainly negatively charged at neutral pH. Its ability to tightly bind other proteins is often exploited in medicine for the separation of different types of antibodies[44] and, thanks to its high water solubility, its remarkable thermal stability[45] (T_m= 87°C) and its straightforward production and purification[46,47], it is also an excellent model protein for folding studies or other structural investigations. Despite wild-type GB1 is a pretty stable protein, for laboratory application it is normally preferred the usage of the GB1-QDD form containing the T2Q, N8D and N37D mutations that further improve its robustness under drastic conditions[48]. These modifications widen the range of pH in which the protein is stable leaving essentially unchanged its structure, unless the insertion of two additional Asp residues increase the negative charge density on its surfaces (**Figure 2b**). This feature renders GB1-QDD extremely suitable for the interaction with positively charged ligands, such as some calix[4]arene derivatives that will be presented below[49]. Since the structural similarity between GB1-QDD and wild-type GB1 allows to consider the two proteins as equivalent, in the rest of the chapter this mutated form will be indicated simply as GB1.

5.3 The ligands

The tetraalanino-calix[4]arene **1**[50] (**Figure 3a**) was tested several years ago as receptor for different ammonium salts, native amino acids and corresponding methyl esters hydrochloride, resulting able to bind most of the tested guest molecules with association constants (K_a) in the order of 10^2-10^3 M^{-1}. Within all the tested partners, the highest binding affinity was reported for amino acid methyl esters bearing apolar residues. It was proposed that this calix[4]arene derivative could bind the ammonium groups of the guest amino acids by establishing electrostatic or dipole/ion interactions with the -C(O)NH- or carboxylate moieties belonging to its alanine residues. Combined with these interactions, it

was supposed that also hydrophobic effects, originated by the inclusion of the guest lipophilic residues in the calix[4]arene aromatic cavity could be effective to bind target substrates. Calixarene **1** was therefore considered promising as ligand for cyt *c*, also because it has been reported that a negatively charged calixarene such as the tetrasulfonato calix[4]arene **2** (**Figure 3b**) can recognize lysine residues on the surfaces of this protein exploiting the simultaneous formation of salt bridges with the ammonium group and the inclusion of the alkyl part of the amino acid side chain in the calixarene cavity[29]. Moreover, as reported in **Chapter 4**, calix[4]arene-based ligands bearing polypeptide loops at their upper rim were already reported in the literature to bind tightly to the same model protein[28,35]. Thus, it was considered reasonable to explore the binding properties of **1** toward cyt c and the results of these investigations will be reported in following pages.

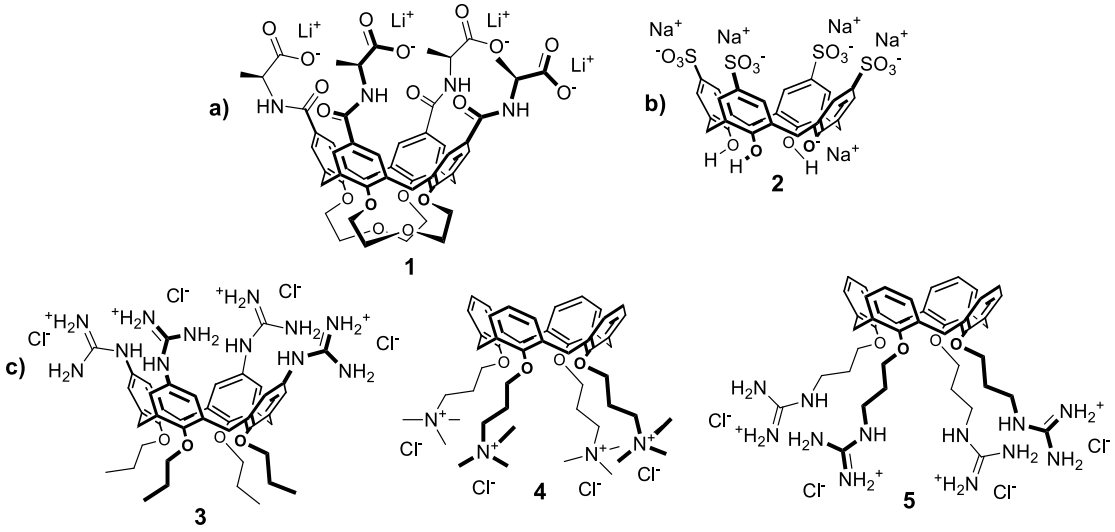


*Figure 3: **a)** Calix[4]arene derivatives tested as ligands for cyt c and **b)** tetrasulfonato calix[4]arene known to interact effectively with the same protein by recognizing some Lys residues on its surfaces. In part **c)** of the figure are instead shown three compounds tested as ligands for GB1.*

Compounds **3**-**5** (**Figure 3c**) were instead tested as ligands for GB1 in a previous work of thesis, with the aim to explore the ability of positively charged guanidinium-calix[4]arenes to bind proper model proteins. In this previous work, the binding between the protein and ligands **3-5** were studied by NMR titrations[49], which yielded interesting results. In fact, if on the one hand for compound **3** it was not possible to obtain information about the interaction with the target protein since it caused its precipitation in the tested experimental conditions, on the other hand **4** and **5** were found to bind GB1, with dissociation constant of 800 μM and 80 μM, respectively. In both cases, the maps of chemical shift perturbations seemed to point out that the protein-ligand recognition mechanism takes places probably thanks to the operation of both electrostatic and hydrophobic interactions with negatively charged and lipophilic amino acid side residues, respectively.

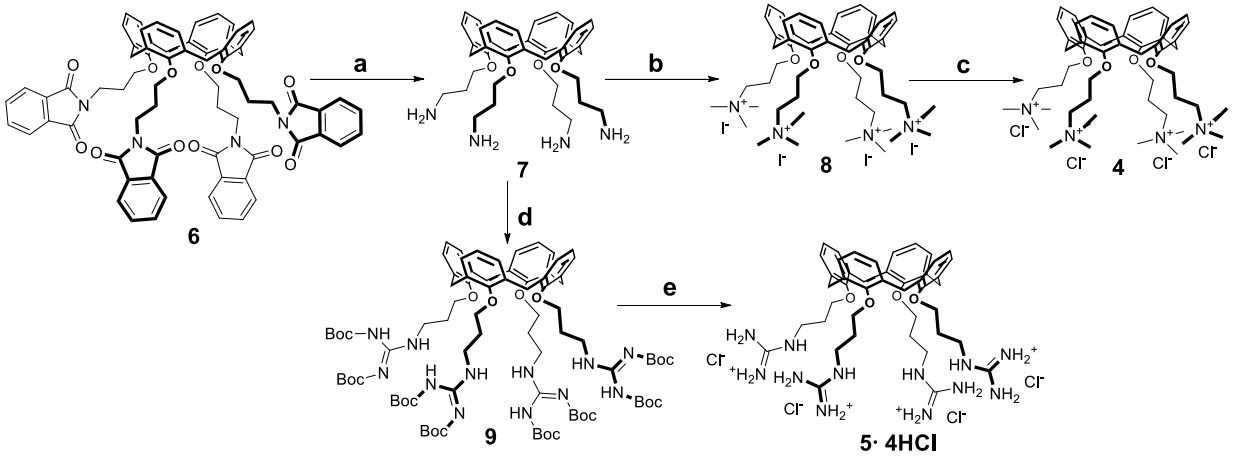
In order to obtain further detailed information on the contributions that these two types of interactions could give to the overall binding energy of the given processes it was decided to study the same protein to ligand binding by ITC experiments, whose set-up and results will be presented in the next paragraphs.

5.4 Results and discussion

5.4.1 Synthesis of the ligands

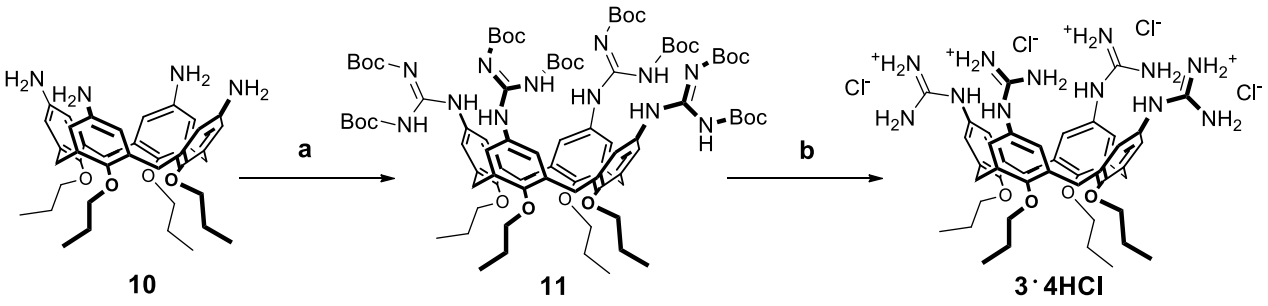
A 50 mg batch of calixarene **1** was already available in our laboratories from previous studies as molecular host. The calix[4]arene derivatives **3**-**5** were instead obtained according to protocols already present in literature and developed in our research group[51–53]. More in details, **4** and **5** were synthetized starting from the tetrakis[3-(phthalimido)propoxy]- intermediate **6**, which was converted into the corresponding tetrakis[3-aminopropoxy]calix[4]arene **7** by reaction with hydrated hydrazine in refluxing EtOH in 74% yield (**Scheme 1**). This reaction was performed under nitrogen atmosphere and the work up was carried quickly in order to expose the product to the air for the shortest time possible, to avoid the formation of stable carbamate-byproducts upon adsorption of CO_2 from the atmosphere[54]. Subsequently, an aliquot of **7** was reacted with CH_3I to yield the parent tetrakis[3-(trimethylammonium)propoxy]- salt **8** (quantitative yield), whose iodide counterions were exchanged with the more hydrophilic chloride anions by treatment with DOWEX11 resin to obtain compound **4** (75% yield).

The remaining part of **7** was instead transformed into the calix[4]arene derivative **9** after reaction with [N, N'–bis-(Boc)]triflylguanidine[55], which was then deprotected from the Boc groups in 1,4-dioxane with 37% HCl in 1,4 dioxane to yield the target tetrakis(guanidiniumproxy)calix[4]arene **5** as tetrahydrochloride salt (quantitative yield for both steps).



Scheme 1: a) $N_2H_4{\cdot}H_2O$; MeOH, reflux. b) CH_3I, $KHCO_3$; dry MeOH, N_2. c) DOWEX11 resin; H_2O, 30 min. d) [N, N'–bis-(Boc)]triflylguanidine, NEt_3; dry DCM. e) HCl 37%, TES; 1,4-dioxane.

Compound **3** was obtained according to the two step-synthetic pathway shown in **Scheme 2**, starting from its tetrapropoxy-*p*-aminocalix[4]arene precursor **10**. This calix[4]arene derivative was coupled with [N, N'–bis-(Boc)]thiourea in the presence of $HgCl_2$ to obtain compound **11**, whose guanidine moieties were freed from the Boc protecting groups by reaction with 37% HCl, giving compound **3** in quantitative yield.



Scheme 2: a) [N, N'–bis-(Boc)]Thiourea, $HgCl_2$, NEt_3; dry DMF. b) HCl 37%, TES; 1,4-dioxane.

For all the aforementioned calix[4]arene derivatives the NMR and ESI-MS spectra, as well as the found physico-chemical properties, were consistent with those expected according to their structure and with the data reported literature[51–53], thus confirming the successful synthesis of target compounds.

5.4.2 Characterization of the cyt c-1 binding

5.4.2.1 NMR studies

Investigations about the interaction between cyt *c* and compound **1** were performed by NMR titrations, where increasing aliquots of a 25 mM aqueous solution of the calix[4]arene derivative were added to a sample containing the ^{15}N labelled protein at 0.1 mM concentration (20mM KH_2PO_4 buffer, 50mM NaCl, pH 6). The occurrence of the protein-ligand interaction was monitored by ^{15}N HSQC experiments, where the chemical shift perturbations (CSPs) of the resonances belonging to the amide NH groups of the cyt *c* backbone were observed during the titrations[56–58]. In order to avoid the observation of CSPs due to oxidation of the protein, cyt *c* was kept in its reduced form by addition of sodium ascorbate (1 mM) to the samples. The experiments pointed out that the positions of most of the cross peaks appearing in the ^{15}N HSQC spectra remain almost unchanged during the addition of the ligand (**Figure 4a**). In fact, only for two small groups of resonances around both Lys4 and the two neighboring Lys87 and Lys 89 residues were reported significant CSPs, thus suggesting the presence of two binding sites for calix[4]arene **1** on the protein surface. Moreover, the increasing magnitude of the observed perturbations as function of the concentration of **1** (**Figure 4b**) indicated that the bound and the unbound form of cyt *c* were in fast exchange compared to the NMR timescale[56,58].

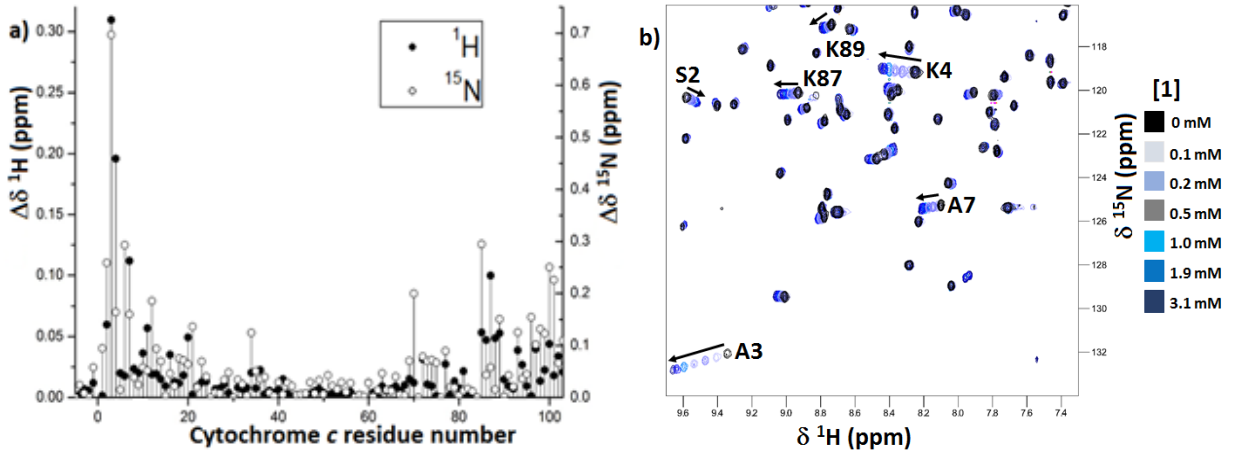


Figure 4: ***a)*** *Plot of chemical shift perturbations (CSPs) of ^{1}H signals (black circles) and ^{15}N peaks (white circles) measured for cytochrome c backbone amides in the presence of a 30-folds excess of* ***1****. Cytochrome c residues are numbered from –5 to 103. Blanks correspond to proline residues 25, 30, 71 and 76, and unassigned Gly84.* ***b)*** *Spectral region from overlaid ^{1}H-^{15}N HSQC spectra of 0.1 mM pure cytochrome c (black) and cytochrome c in the presence of 0.1–3.1 mM* ***1*** *(blue scale). The arrows indicate the direction of the CSPs experienced by some amide resonances during the titration.*

These results were consistent with what reported in literature for the binding of the same protein with compound **2**. In fact, the co-crystal structure of cyt c and compound **2** shows Lys4 and Lys89 to be two of the three anchoring points exploited by the ligand to bind cyt *c*[29]. NMR experiments alone were not sufficient to distinguish which, between Lys87 and Lys89, was indeed most involved in the interaction with **1** because of the close proximity of the two amino acid residues. Therefore, both Lys87 and Lys89 have to be considered as possible binding sites for this ligand. No evidence of interaction with calix[4]arene **1** was recorded for Lys22, which is the third lysine residue bound by compound **2** in the solid state. The position of the corresponding amide resonance, in fact, did not experience appreciable perturbations during the titrations indicating that probably, in solution, this amino acid residue does not interact so efficiently with **1**. However, despite the NMR data pointed out a similar ability of ligands **1** and **2** to recognize the same binding sites on the protein surfaces, it is necessary to be careful before claiming that they exploit also the same mode of binding to the protein residues. In fact, ligand **1** and **2**, although sharing the same *cone*-conformation of the calix[4]arene cavity and a quite similar total charge (4^- and 5^- respectively), are quite different for the polar groups adorning the upper rim. Ligand **1** has a considerable larger distance of the negatively charged carboxylic groups from the cavity compared to the sulfonate moieties of ligand **2** and presents four additional H-bonding acceptor/donating groups (the amide –CONH at the upper rim). Moreover, on the contrary to what found for compound **2**, the conformational mobility, the steric hindrance and the possibility to give rise to intramolecular H-bonds of the four alanine residues at upper rim, could impair calix[4]arene **1** to host the alkyl part of the lysine residues in the cavity. An alternative and possible mode of binding for target lysines by host **1**, for instance, could be through H-bonds between the $–NH_3^+$ head group of the guest residue and the amide

C=O groups of the calix[4]arene. These interactions could possibly be also assisted by CH-π interactions of the methylene groups of Lys with the calix[4]arene cavity. In the case of ligand **1**, it cannot be even excluded that additional non-covalent interactions between the remaining –CONH- and -COO^- groups of the calix[4]arene alanine moieties and adjacent polar residues present on the protein surface could play an important role in binding. All these alternative modes of binding appear reasonable since they were already observed in solution for the tetramethyl ester[59] of calixarene **1** and their occurrence cannot be excluded in principle, especially in the absence of a crystal structure of the complex between cyt c and **1** that would provide further elucidation on the mechanism that this calix[4]arene derivative exploits to bind the target lysine side chains. More recently, Crowley et al.[60] could also highlighted a rather interesting interaction, at the solid state, of the tetrasulfonato calix[4]arene **2** with the dimethylated lysine residues on the surfaces of lysozime. The dimethylammonium head groups of two lysines were buried in the calixarene cavity, interacting through cation/π interactions. Such interactions should however be excluded in the absence of methylated lysines since the highly hydrophilic and polar NH_3^+ head group has a clear preference to bind to charged (-COO^- /-SO_3^-) or polar (-CO-) groups rather than to enter the lipophilic cavity of a calixarene.

It was also noted that in the cyt c – calixarene **1** binding process, the amide resonance experiencing the highest perturbation ($\Delta\delta$= 0.31 ppm for $^1H^N$) was that belonging to Ala3. Although some of the observed chemical shift perturbations might simply originate from the conformational rearrangement of cyt *c* associated with the binding of lysine residues, this considerably high value of CSP might suggest more specific interactions. In previous studies **1** was reported to have a fair affinity for alanine O-methylester hydrochloride[50], but it seems unlikely that this amino acid side chain is selectively bound by **1** in place of the adjacent residue of Lys4. Moreover, it was also noted that the binding ability of **1** towards amino acids O-methylester hydrochlorides was strongly suppressed at basic pH values, thus indicating that is the α-NH_2 group of the guest molecule in its protonated form (-NH_3^+) to allow the recognition event to take place. Consequently, since in cyt *c* the α-amino group of Ala3 is part of the amide backbone and thus not free to interact as –NH_3^+ with **1**, the high CSP observed for the corresponding H-^{15}N resonance should be originated by the formation of additional H-bonds between the alanine –NHCO group of the protein and the alanine moieties on the calix[4]arene structure, after that the recognition of the neighboring Lys4 has taken place.

An overall view of the regions of the protein surfaces most involved in the interaction with the tested calix[4]arene derivative can be visualized in **Figure 5**, in which the amino acid residues whose amide resonances were significantly perturbed at the end of the performed titrations ($\Delta\delta \geq$ 0.04 ppm or 0.4 ppm for $^1H^N$ and ^{15}N, respectively) are colored in blue. As expected, the presence of the ligand exerted an appreciable influence only on a few resonances located around Lys4 and Lys 87/Lys 89, which have

been instead highlighted in green, thus confirming that the interaction with the protein took place in small regions of its surfaces which contain lysine residues.

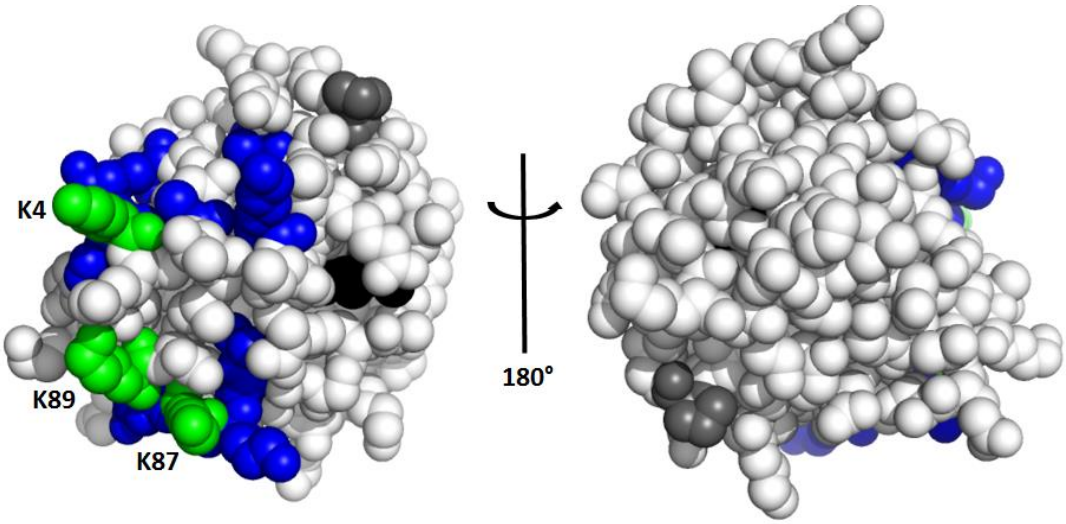


*Figure 5: Space-filling representation of cyt c highlighting the **1** binding surfaces. Residues for which the amide resonance had a significant chemical shift perturbation ($\Delta\delta\,^{1}H^{N} \geq 0.04$ or $^{15}N \geq 0.4$ ppm) are in blue. The heme edge is black and prolines are grey. Lys4, Lys87 and Lys89 are colored in green.*

Together with this important information on the possible structure of the cyt *c*-**1** complex, the NMR titration data also allowed to calculate the binding affinities of the system. In fact, by plotting the CSPs observed in the ^{1}H dimension for the two small clusters of the amide resonances around Lys4 and Lys 87/Lys 89 as function of the ligand concentration[56–58], the hyperbolic curves of **Figure 6** were obtained. A non-linear regression of these experimental data with a two independent 1:1 binding models yielded K_d gave values of ~0.4 mM and ~1.2 mM for the two groups of resonances, respectively, thus supporting the hypothesis that the protein has at least two binding sites for **1** on its surfaces. These K_d values were similar to those of ~0.8 mM and ~1.6 mM reported in literature for the cyt *c*-**2** system having as a target the clusters of amino acid residues of Lys4 and Lys 89, respectively.

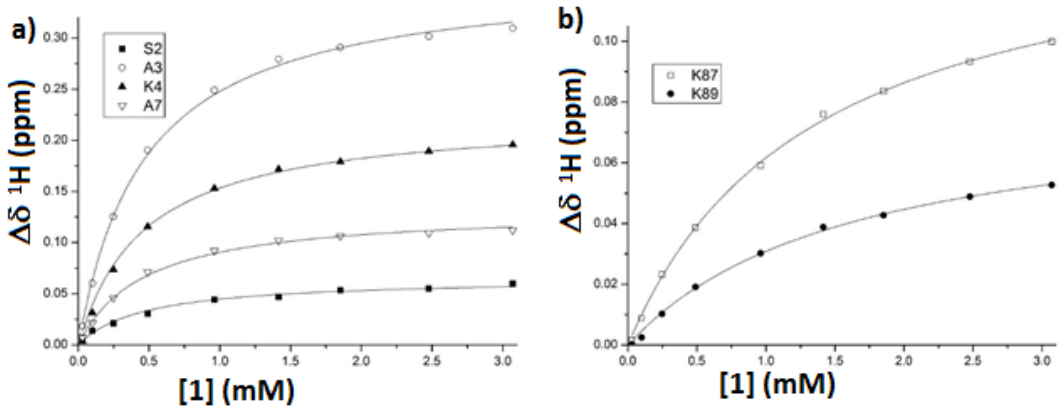


*Figure 6: Binding isotherms for the interaction of **1** with cyt c. The increasing CSPs of two groups of amide resonance around **a)** Lys4 and **b)** Lys87 were fitted independently with a 1:1 binding model as function of the ligand concentration.*

5.4.2.2 ITC studies

Attempts to achieve a full thermodynamic characterization for the binding process involving cyt *c* and compound **1** were performed by running ITC experiments[7,61,62]. In the absence of a standard protocol to set up experiments for this specific protein-ligand system, it was decided to place cyt *c* in the sample cell at 50 μM concentration, since this value is included in the 10-100 μM range normally used for this technique[63]. Accordingly, also taking into account that during the NMR titrations the saturation of the protein binding sites was achieved after the addition of a 30-folds excess of **1**, a 2 mM concentration was chosen for the ligand stock solution. All the experiments were carried out by sequential addition of known aliquots of titrant (24 x 10 μL) and, for consistency, it was decided to perform them in the same experimental conditions used for the NMR studies (30°C, pH 6, 20mM KH_2PO_4 buffer, 50mM NaCl). However, it was not possible to work with the reduced form of cyt *c,* since a buffer into protein control titration, carried out in the presence of 1mM sodium ascorbate in both solutions, yielded highly deformed baseline and remarkable peaks for the heat of dilution.

After the oxidation of the protein with potassium ferricyanide, which was then eliminated from the sample by dialysis against the buffer, the same experiments gave instead heat of dilution comparable to those observed by injecting buffer into buffer (**Figure 7b** and **7a**, respectively), thus indicating that the sodium ascorbate was not tolerated in the reaction medium.

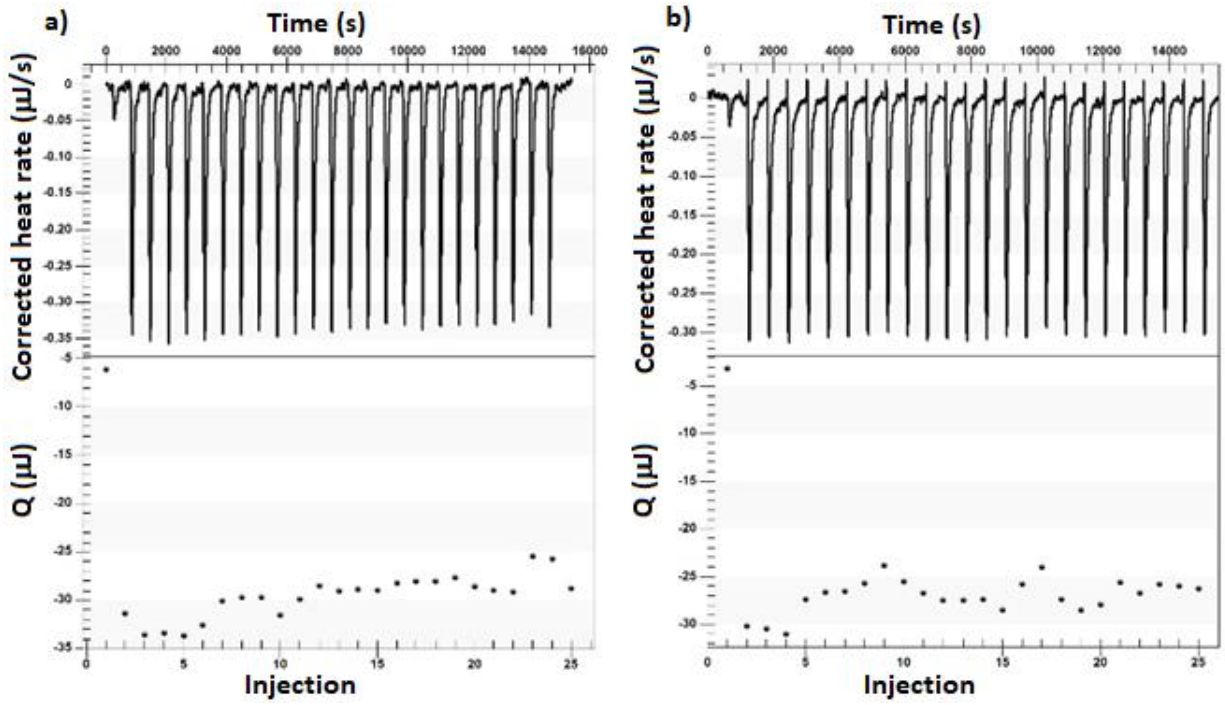


*Figure 7: Raw heat rate corrected for the baseline drifting and corresponding peak areas (top and bottom, respectively) measured for each injection of **a)** buffer into buffer and **b)** buffer into 50 μM oxidized cyt c control experiments (24 x 10 μL injections, 30°C, pH 6, 20mM KH_2PO_4 buffer, 50mM NaCl).*

The use of cyt *c* in different oxidation states in NMR and ITC titrations was not considered problematic to obtain comparable information from the these two techniques since, being the ligand supposed to recognize single amino acid residues on the protein surfaces, the conformation changes undertaken by the protein upon oxidation[24] should not affect significantly the mode of binding.

Also the ligand solution required some treatment before it could be used in ITC studies. In fact, when the stock solution was simply prepared by dissolving compound **1** in the buffer, the titration for the determination of the heats of dilution gave rather broad peaks (**Figure 8a**), thus suggesting the occurrence of undesired by-processes. Considering that the tested calix[4]arene was available as tetralithium salt[50], the detection of additional heat could be ascribed to the exchange of the four alanine counterions with the K^+ ions coming from the buffer. This problem was solved (**Figure 8b**) by adding a step of dialysis against the buffer during the preparation of the solution of **1**, that avoided this cation exchange process to take place in the sample cell.

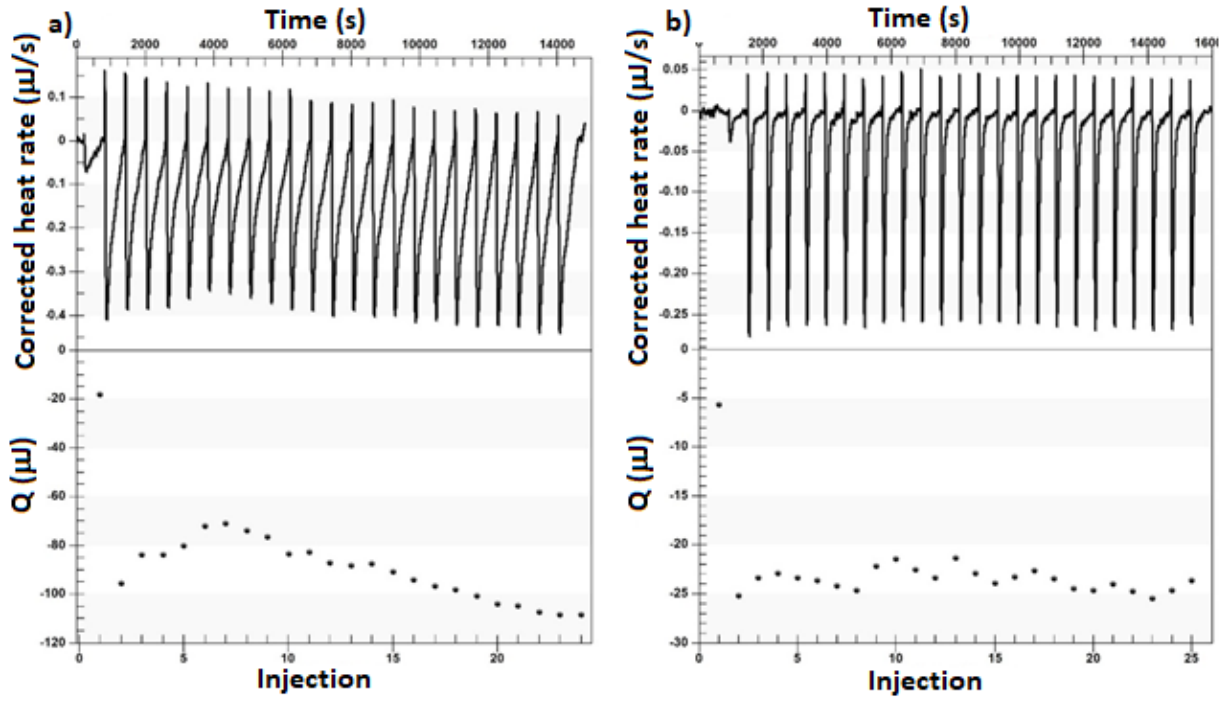


*Figure 8: Heat rate corrected for the baseline drifting and corresponding peak areas (top and bottom graphs, respectively) measured for two solutions of **1** prepared by **a)** simply dissolving the ligand in the buffer and **b)** dialyzing the ligand solution in buffer against the buffer.*

After this set-up of the ideal operation conditions, a ligand into protein titration was performed, yielding the experimental data shown in **Figure 9a**.

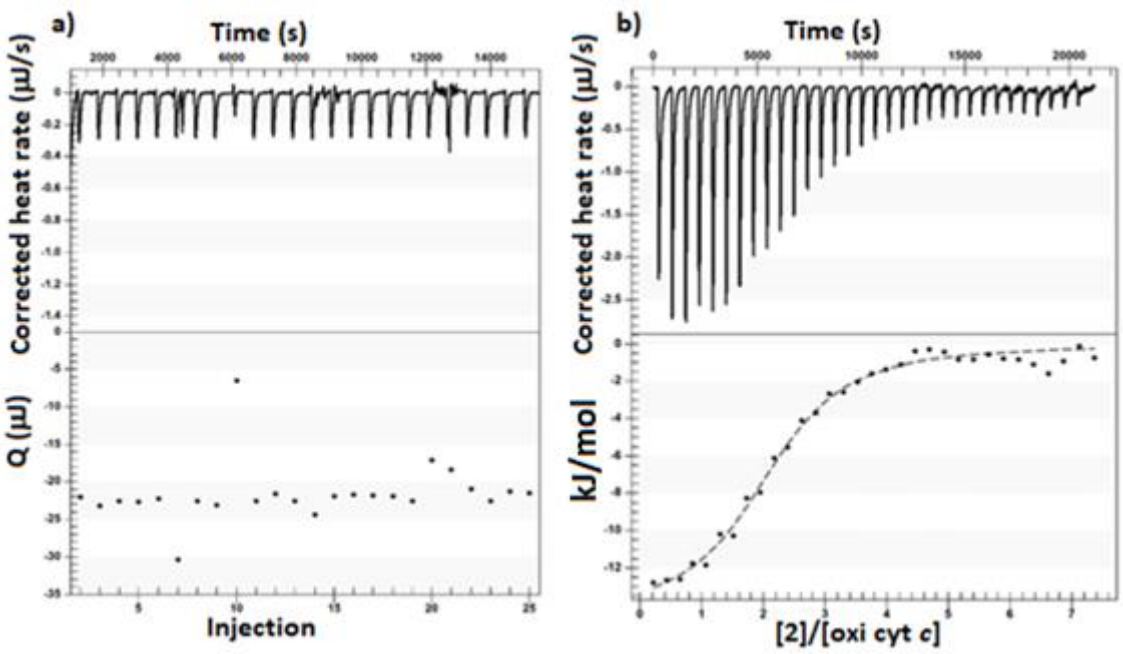


Figure 9: ***a)*** *Heat rate corrected for the baseline drifting and corresponding peak areas (top and bottom graphs, respectively) measured for each injections of a solution of ligand* ***1*** *into oxidized cyt c (ITC experiment: 2mM ligand, 50 μM protein, 24x10 μL injections, 30°C, pH 6, 20mM KH_2PO_4 buffer, 50mM NaCl).* ***b)*** *Heat rate corrected for the baseline drifting (top) and corresponding binding isotherm (bottom) measured for a solution of ligand* ***2*** *into oxidized cyt c (ITC experiment 4mM ligand, 100 μM protein, 48x5 μL injections, 30°C, pH 6, 20mM KH_2PO_4 buffer, 50mM NaCl).*

The detection of the heat of reaction developed at each ligand injection was disturbed by a high level of noise in the baseline, which was subsequently found to be caused by a problem of the stirring system of the instrument. Notwithstanding, most of the resulting peaks had a regular shape and it was possible to point out that the heat release associated with the studied cyt c – ligand **1** interactions was negligible in the tested experimental conditions. This preliminary result does not appear so surprising since, as mentioned in **Chapter 4**, when estimations of the stoichiometry n and of the dissociation constant K_d are available, to obtain an optimal ITC binding isotherm the initial protein concentration [*P*] should be chosen in order to afford values of the *c parameter* between 10 and 1000 (**Eqn. 1**)[62,64,65].

$$c = n[P]K_a = \frac{n[P]}{K_d} \quad \textbf{Eqn. 1}$$

In the case of the cyt c – **1** complex, considering the values of n and K_d (n=2, K_d ~0.4 mM) obtained by NMR studies and the value of [*P*] = 50 μM, **Eqn. 1** gives *c* ~ 0.25, indicating that at least a 40-folds increase in the protein concentration is necessary, which would mean to reach a cyt *c* concentration of 2 mM in the sample cell and a ligand concentration of ~80 mM in the syringe. However, these conditions are prohibitive under a practical point of view, since require the production of a high amount of both the protein and the ligand and could induce the occurrence of aggregation phenomena involving the binding partners, which are known to be detrimental to get useful data from this experimental

technique. Therefore, it should be concluded that the interactions between the cyt *c* and **1** are too weak to allow the detection of the corresponding heat of complexation. On the other hand, a titration of cyt *c* with compound **2**, carried out by the group of Crowley[66] under similar conditions (4 mM ligand into 100 μM oxidized protein, 48x5 mL injections, 20mM KH_2PO_4, 50mM NaCl, pH6, 25°C, unpublished data), resulted in a perfect sigmoidal binding isotherm (**Figure 9b**), whose regression analysis gave, as average, a K_d value for the two binding sites of 22 μM. This value was 50-folds lower than the one determined by NMR[29], indicating that this latter technique was overestimating the dissociation constant, as reported in literature for other binding events[67,68]. Considering that the affinities of calix[4]arenes **1** and **2** for cyt c were suggested to be similar (by NMR), it seems thus possible that a moderate increase of the protein concentration could have been sufficient to obtain manageable data even in our case. This hypothesis could unfortunately not be experimentally confirmed because problems to the stirring system of the ITC instrument prevented us to perform further calorimetric titrations at higher protein concentrations.

Interestingly, the titration on the cyt c-**2** system pointed out that the enthalpy and the entropy of binding were similar in magnitude (ΔH= -14.53 KJ/mol and $-T\Delta S$= -12.03 KJ/mol), indicating that the ligand exploits both electrostatic and hydrophobic interactions to bind the protein surfaces, which are known to give rise to favorable enthalpic and entropic contributions to the overall binding free energy[7,69,70], respectively. These calorimetric results are also in nice agreement with what observed in the crystal structure of the corresponding complex[29], where the target lysine residues are recognized by the ligand thanks the formation of salt bridges between its sulfonic units and the ammonium head group at the amino acid side chain and thanks to the inclusion of the alkyl part of the same side chain in its cavity. In the case of the binding process involving ligand **1** and cyt *c*, of course, these kind of information cannot be inferred from the collected data, but some considerations may be done on this subject. In fact, despite the formation of the adducts of cyt c with ligands **1** and **2** was investigated at similar protein concentrations and associated with comparable K_d values, the ITC investigations of these processes yielded different results. Being the heat exchange generated by protein-ligand interactions directly proportional to the binding enthalpy (**Chapter 4**)[62,64,65], it can be suggested that the enthalpic contributions to the overall free energy was lower for the cyt *c*-**1** system, thus affording further complications to the detection of useful heat of reaction. As a balance, it can be supposed that to allow the binding to take place the ligand exploit a consistent entropic contribution. Also in this case, however, the absence of substantial ITC data did not allow us to confirm our hypothesis and, at the moment, it can be only reported that in the tested experimental conditions this technique is not suitable to obtain insights on the thermodynamics of the studied binding process.

5.4.2.3 Crystallization tests

Co-crystallization experiments[10,71,72] were carried out for the cyt *c*-**1** system, in order to try to get further information about their reciprocal interactions from the solid state structure of the corresponding

complex. According to the procedure reported in the literature for the crystallization of the same protein with compound **2**[29], the first attempts were done by following the manual hanging drop vapor diffusion experiments. The protein and ligand concentrations (1.7 mM and 17 mM, respectively) and the composition of the reservoir solution (20% PEG 8000, 50 mM NaCl, 50 mM sodium cacodylate buffer, pH 6.0) used were similar to those adopted for the system cyt *c* - **2**. However, these tests yielded only completely clear drops even after one month of equilibration and efforts to induce crystallization by increasing the ligand concentration and/or by changing the percentage and the molecular weight of the PEG used as precipitant were unsuccessful.

It was then decided to explore different experimental conditions by running a robot screening over a commercially available kit of 96 reservoir solutions, according to the sitting drop vapor diffusion method. The experiment, carried out keeping the protein concentration at 1.7 mM, evidenced that the presence of 10 or 20 mM ligand caused the formation of red jelly-like precipitates (**Figure 10a**) in drops incubated over 22 different reservoir solutions, 11 of which containing phosphate citrate or sodium citrate buffer together with different additives.

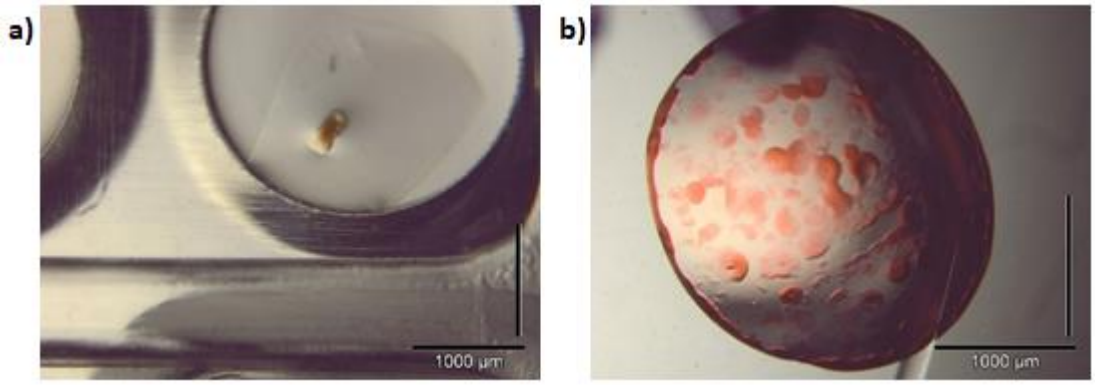


*Figure 10: Gel formation observed during crystallization tests for the cyt c - **1** complex performed by **a)** a robot sitting drop- and **b)** manual hanging drop vapor diffusion method (1.7 mM protein, 10 mM ligand concentrations for both experiments). The crystallization mixtures were equilibrated over reservoir solutions composed of, respectively, 20% PEG 6000, 1M LiCl, 100mM sodium citrate buffer (pH 4) and 25% PEG 8000, 50mM sodium citrate buffer (pH 5).*

In all cases the control drops, where the solution of the ligand was replaced by the same volume of water, were devoid of any precipitates, demonstrating that the formation of the observed red gels was induced by **1**. Since this phenomenon was easily replicated in manual hanging drop tests where the same protein and ligand concentrations were tested and the reservoir solution was composed by 50mM citrate buffer (pH 5) and PEG 8000 as only precipitant agent (**Figure 10b**), it was considered useful to perform some cycles of optimization of the experimental conditions, in order to prompt the system to afford crystalline precipitates.

These investigations (**Table 1**) pointed out that the jellification process was promoted between pH 4 and 5 and in the presence of 20-30% PEG 8000 or PEG 3350 and at both 10 or 20 mM **1**, while it was prevented at pH 6 or by increasing the buffer concentration at 100 mM. The use of phosphate citrate buffer (pH 4.2) to set up the same experiments led instead to harsher jellification or even to precipitation of the protein when high PEG percentage (30% PEG 3350) were employed. For both the tested buffer, the lowering of the **1** concentration at 1.7 mM caused a milder gel formation in the studied drops, while a further decrease of the concentration for both the binding partners to 0.1 or 0.5 mM caused the precipitation of the protein even in the control drops for the all the tested PEGs percentages.

*Table 1: Results obtained from manual hanging drop experiments for the crystallization of the cyt c-**1** system, performed by testing different compositions of the reservoir solution and ligand concentrations. The drops were prepared by combining 1 mL volumes of the protein (1.7 mM), the calix[4]arene derivative and the reservoir solution. Control drops were obtained by replacing the ligand with 1 μL of water.*

[1]	**Buffers**											
	50 mM sodium citrate (pH 5)						**50mM phosphate citrate (pH 4.2)**					
	PEG 8000			**PEG 3350**			**PEG 8000**			**PEG 3350**		
	10%	**20%**	**30%**	**10%**	**20%**	**30%**	**10%**	**20%**	**30%**	**10%**	**20%**	**30%**
0 mM	Clear	Clear	Clear	Clear	Clear	Clear	Clear	Clear	Clear	Clear	Clear	Gel
1.7mM	Clear	Gel (mild)	Gel (mild)	Clear	Gel (mild)	Gel (mild)	Clear	Gel	Gel	Clear	Gel (strong)	Prec.
10mM	Clear	Gel	Gel	Clear	Gel	Gel	Clear	Gel	Gel	Clear	Gel (strong)	Prec.
20mM	Clear	Gel	Gel	Clear	Gel	Gel	Clear	Gel	Gel	Clear	Gel (strong)	Prec.

For none of the above described tests the conversion of the formed gels or amorphous precipitates in crystals was achieved even after several weeks of incubation and it was thus decided to study the influence that the presence of different additives in the reservoir solution could have on this process. More in details (**Table 2**), it was investigated the effect of the addition of different salts at 50 mM or 1 M concentration to a reservoir solution containing 50 mM sodium citrate buffer and 25% PEG 8000, chosen

as most promising composition according to what observed in the previous experiments. The concentration of both the ligand and the protein, instead, were initially fixed at 1.7 mM.

*Table 2: Results obtained from manual hanging drop experiments for the crystallization of the cyt c-**1** system, performed by testing different salts as additives for a reservoir solution composed of 50 mM sodium citrate (pH 5) and 25% PEG 8000. The drops were prepared by combining 1 mL volumes of the protein (1.7 mM), the calix[4]arene derivative (1.7 or 50 mM) at and the reservoir solution. Control drops were obtained by replacing the ligand with 1 μL of water.*

[1]	**Tested salts**										
	None	**Na_2SO_4**		**Li_2SO_4**		**NH_4OAc**		**NaOAc**		**NaCl**	
	-	**50mM**	**1M**	**50mM**	**1M**	**50mM**	**1M**	**50mM**	**1M**	**50mM**	**1M**
0 mM	Clear	Gel	Phase sep.	Gel	Phase sep.	Phase sep.	Clear	Phase sep.	Prec.	Clear	Phase sep
1.7mM	Gel	Gel	Phase sep.	Gel	Phase sep.	Phase sep.	Clear	Phase sep.	Prec.	Clear	Phase sep.
50mM	Strong gel	Gel	Gel	Strong gel	Phase sep.	Gel	Clear	Phase sep.	Prec.	Clear	Phase sep

In general, the addition of the salts created too harsh conditions, as it can be confirmed by the fact that in almost all control drops their presence caused phase separation or jellification of the protein. The only exceptions were represented by the solutions containing NH_4OAc or NaCl, where a salting-in and a salting-out effect were observed moving by from 50 mM to 1M concentrations, respectively. These experiments yielded clear drops also in the presence of the ligand, but they looked interesting since suggested that an intermediate concentration of these two salts or their combination with the other tested ones may be appropriate to optimize the crystallization conditions. Moreover, the same tests were also performed in the presence of 50 mM of **1**, since the consequences of its presence in large excess respect to cyt c was not evaluated before. The increase of the ligand concentration did not lead to the formation of crystalline precipitates in any experiment, but in some cases it seemed to be beneficial, since it led to jellification processes instead of phase separations. Despite many efforts have been spent to achieve this purpose, every attempts to find proper crystallization conditions for the studied protein-ligand complex failed. These observations are in contrast with what reported in literature for the cyt *c*-**2** system used initially as reference, for which rough co-crystals were obtained in preliminary studies by simply adding to the reservoir solution PEG 8000 as sole precipitant[29].

This marked difference in the properties of the two protein-ligand systems has to be probably ascribed to the different structural features of the two ligands and in particular of the functional groups that they bear at both the upper and the lower rim. In fact, the four alanine residues of **1**, when compared to the sulfonate moieties of **2**, are much more flexible because of the presence of three rotationally mobile bonds (*i. e.* between the aryl units and the –C(O)NH- groups and between the chiral carbon and both the C- and N- terminal carbonyl moieties). This feature could disadvantage ligand **1** to fix the binding units according to a spatial disposition suitable to crystallize with cyt *c*, thus rendering trickier the obtainment of the desired co-crystals. Moreover, the presence of the crown-ether motifs at the lower rim, instead of the four free phenolic –OH units, could disfavor the same compound to form additional interactions (i. e. CH-π and van der Waals contacts) with other protein monomers once that the main recognition event of the target lysine residues has occurred at upper rim. Ligand **2** has been reported to exploit these kind of interactions to crystallize with different target proteins by inducing their self-assembly as oligomers (*molecular glue* effect)[29,73], but, in the case of compound **1**, the different structural and electronic features of the substituent at the lower rim could prevent the formation of these further contacts, disadvantaging its co-crystallization with cyt *c* to take place.

5.4.3 Characterization of the binding of GB 1 by the guanidinium calix[4]arenes 3-5

5.4.3.1 ITC studies

The binding processes involving GB 1 and compounds **3**-**5** have been already studied in a previous PhD thesis work by ^{15}N-HSQC NMR titrations[49]. In the case of ligands **4** and **5**, these investigations were suitable to obtain a preliminary estimation of their binding affinity for the target protein and of the regions of its surfaces in which the recognition events were localized. On the other hand, for compound **3** it was not possible to get useful information since it induced the immediate precipitation of the protein from the given solution, despite this observation was considered as a good indication that the interactions between the two binding partners were significantly strong. In the present thesis work, it was therefore decided to study the binding processes for GB1 and calix[4]arenes **4** and **5** by ITC, to achieve a deepen knowledge of the driving forces that govern the complexation event and to possibly extract useful suggestions about the mode in which the interactions with the protein were taking place. For the GB 1-**3** system it was instead suggested that the lower reactant concentrations normally required by this technique could have been helpful to prevent protein precipitation and thus to demonstrate the occurrence of the proposed protein-ligand binding process.

Initially it was decided to carry out the titration experiments according to the standard configuration for ITC[7,61,62], that is by placing the protein in the sample cell and the ligand at higher concentration in the syringe, and to use conditions the most similar possible to those reported for the NMR studies (30°C, pH 6, 20mM KH_2PO_4 buffer, 50mM NaCl). On the one hand these choices were suspected to be risky because all the three tested ligands have a pronounced amphiphilic structure that, in the selected

experimental conditions, could be suitable to induce aggregation processes able to prevent the collection of useful ITC data[7,62]. For example, evidence of aggregate formation in water were reported in literature for calix[4]arene derivatives analogous to ligand **3**, being functionalized at the upper rim with ammonium[74] or tetraalkylammonium[75] groups and at the lower rim with long alkyl chains (≥ C_6H_{13}). However, for compound **3**, it was demonstrated by DOSY experiments that it is fundamentally in a monomeric form, at least at concentration lower than 1 mM[76]. For compounds **4** and **5** the possible aggregation in D_2O was instead studied by monitoring the increase of the line-width in their ^1H NMR spectra as function of their concentration and of that of different added salts[53]. For 1 mM aqueous solutions of the calix[4]arene **4**, the addition of moderate amount of inorganic salts (10 mM NaCl, 2 mM $MgCl_2$) did not caused a significant aggregate formation, suggesting that it is also present as monomer at this or lower concentrations. For calixarene **5**, on the other side, a broadening of the NMR spectra were instead observed in the presence of higher concentrations of the same additives (250 mM NaCl, 50 mM $MgCl_2$), indicating that aggregates can form for this compound even though to a moderate extent and under high saline concentrations.

Thus, it seemed worthy at least to verify the feasibility of the planned titrations by performing a ligand into buffer control experiment with a 2mM solution of **4**. This preliminary test, however, gave completely unreliable results (**Figure 11a**).

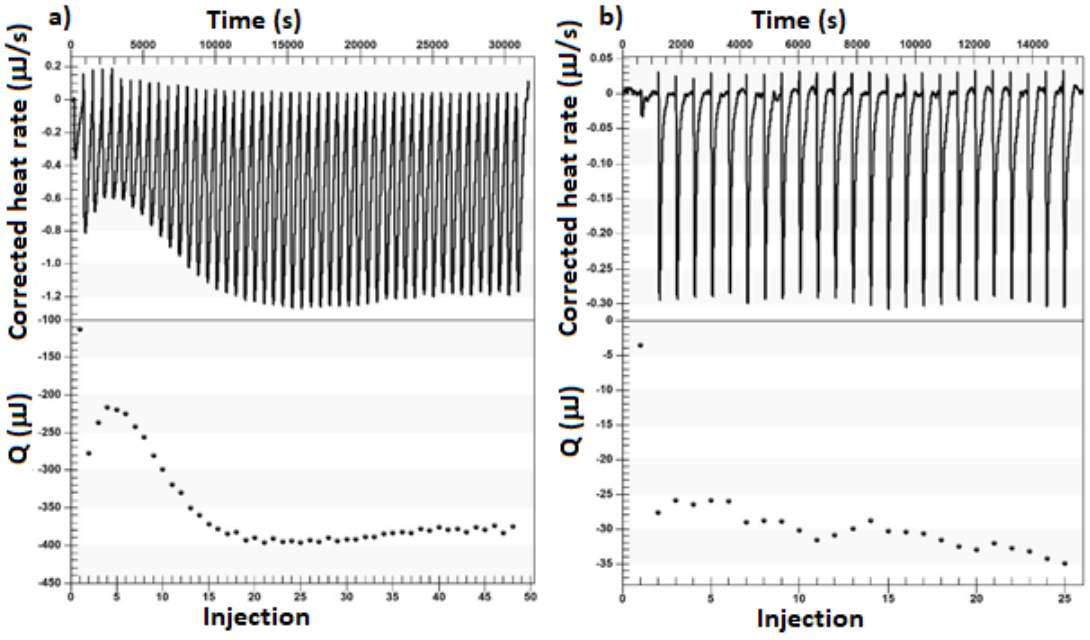


*Figure 11: Raw heat rate corrected for the baseline drifting and corresponding peak areas (top and bottom, respectively) measured for each injection of **a)** 2mM solution of **4** into buffer and **b)** buffer into 50 µM solution of **5** control experiments (50 x 5 µL and 24 x 10 µL injections respectively, 30°C, pH 6, 20mM KH_2PO_4 buffer, 50mM NaCl).*

In fact, the heats of dilution detected for each ligand injections were huge and the shape of the corresponding peaks was highly broadened, as it can be observed in ITC experiments highly disturbed by

undesired aggregation processes. Since after centrifugation of the solution recovered at the end of the titration was not collected any precipitate, the detection of those exceedingly large heats of dilution was ascribed to the inclusion of calixarene **4** in aggregates still soluble in the studied medium. Considering that their occurrence was not observed by NMR even though at slightly lower saline concentrions[53], it is likely to suppose that the fraction of **4** involved in this phenomenon is low, but high enough to affect negatively the results given by a very sensitive technique such as ITC. According to these observations, it was then decided to evaluate the possibility to improve the results of the planned ITC experiments by inverting their configuration[62], that is by placing the ligands in the sample cell and the protein in the syringe. In this way the initial concentration of calix[4]arene is much smaller (50 μM vs. 2 mM) and even decreases during the titration. These conditions seemed rather promising since disadvantage the aggregation processes of the ligands, although the risk that the contemporary increase of the protein concentration would advantage its aggregation is on the other hand present.

At the tested concentration values, namely 50 μM for compounds **3**-**5** and 1 mM for GB 1, the undesired aggregation phenomena seemed to be prevented. In fact, the heats of dilution measured during buffer into ligand control experiments for ligands **4** and **5** (see **Figure 11b** for the case of compound **5**) were analogous to those observed for buffer into buffer model test (**Figure 7a**), while were only slightly higher and associated with a sharp shape of the corresponding peaks in the case of the protein into buffer blank titration (**Figure 12a**). Unfortunately, compound **3** was not soluble in the buffer even in this low concentration range, thus impairing the collection of data about its interactions with GB 1 also by ITC.

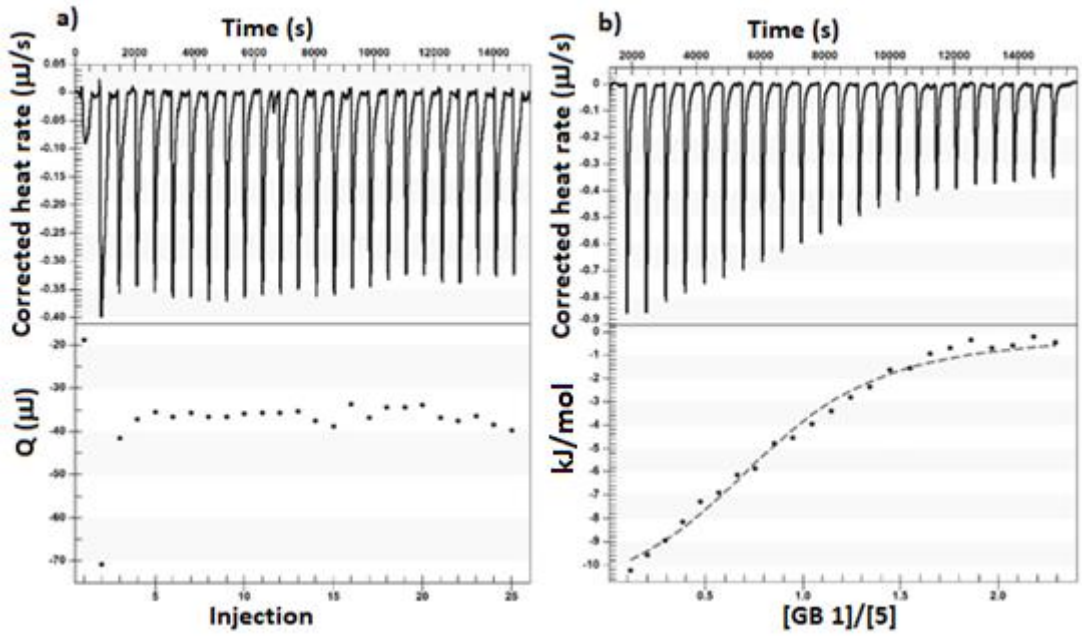


*Figure 12: **a)** Raw heat rate corrected for the baseline drifting and corresponding peak areas (top and bottom, respectively) measured for each injection of a 1 mM GB 1 into buffer control experiment. **b)** Raw heat rate corrected for the baseline drifting and corresponding binding isotherm (top and bottom, respectively) detected for a 0.5 mM GB 1 into 50 μM solution of **5** titration (24 x 10 μL injections, 30°C, pH 6, 20mM KH_2PO_4 buffer, 50mM NaCl for both experiments).*

At the end of this long refinement of the experimental conditions, a tests of protein into ligand titration were finally carried out. The first ligand to be tested was **5** since it was known, according to NMR estimation of K_d, to be able to bind more tightly GB 1 than **4** and thus the detection of measurable heats of reaction for its interaction with the protein was considered easier. To limit as much as possible the waste of the protein, its concentration was decreased at 0.5 mM, while that of **5** was kept at 50 μM.

The heats of reaction developed for each addition of GB 1 were sufficient to give to the plots of heat rate versus the protein/ligand molar ratio the typical sigmoidal shape expected for an ITC binding isotherm. It was therefore possible to fit the experimental data according to standard binding equations[65] (**Figure 12b**). Also in this case the K_d obtained as fitting parameter was lower than that given by NMR titration in analogous conditions (8.9 μM instead of ~80 μM, respectively), in agreement with what reported in **Paragraph 5.4.2.2** for the interactions involving cyt c and ligand **2** or in literature for the inclusion of different amino acids by the same compound[67,68]. Regarding to the thermodynamic contributions that gave rise to the studied binding event, the treatment of the experimental data gave ΔH= -11.85 KJ/mol and -TΔS= -17.44 KJ/mol as values for the enthalpic and entropic terms to the overall free energy of association, respectively. The favorable influence of both the energetic contributions pointed out that the ligand bound to the protein by establishing both electrostatic and hydrophobic interactions on its surfaces[7,69,70], with the latter appearing slightly more important than the former. These results support the NMR data previously collected to describe the same binding event[49], since from those investigations seemed likely that both polar and lipophilic side chains of GB 1 were able to interact with the ligand.

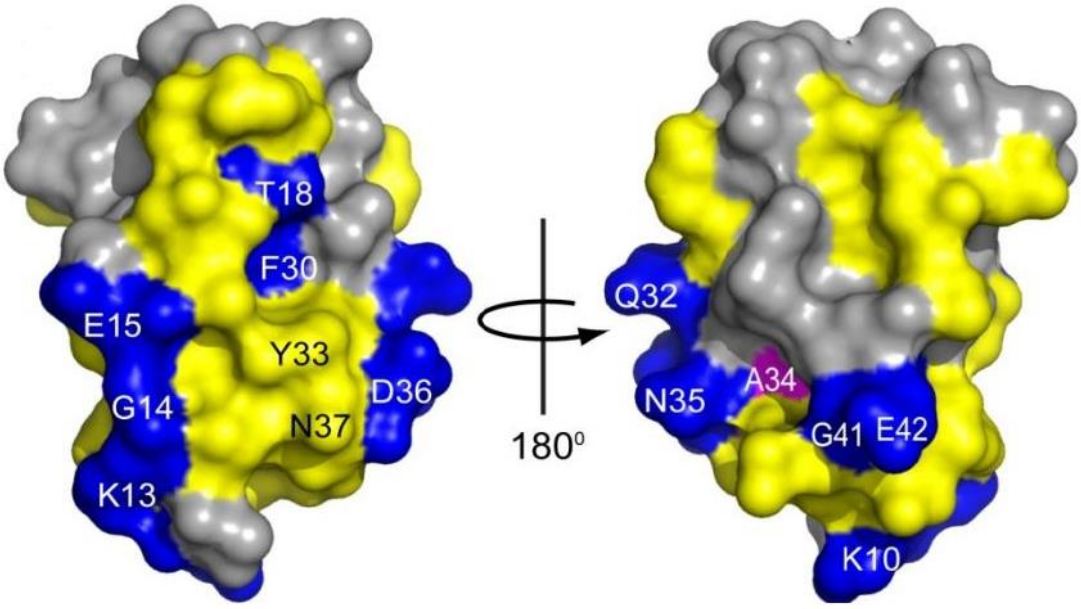


*Figure 13: Chemical-shift perturbation map of GB1 in the presence of 10 equivalents of **5**. Residues for which the amide resonances experienced large ($\Delta\delta^1H^N \geq 0.2$ or $^{15}N \geq 2$ ppm), or small ($\Delta\delta^1H^N \geq 0.05$ or $^{15}N \geq 0.5$ ppm) shifts are colored blue and yellow, respectively. Resonances that broadened beyond detection are magenta (Figure taken from **Ref. 49**).*

More in details (**Figure 13**), the highest CSPs were concentrated in regions of the protein surfaces with predominant negative electrostatic potentials, that means around Asp 36, Glu15 and Glu42, thus indicating that probably interactions between the guanidinium pendants at the lower rim of **5** and the carboxylic lateral groups of those amino acids were occurring. The guanidinium units of compound **5** could also originate additional H-bonds by interacting with the polar side chains of Gln32, Asn35, whose amide resonances experienced significant CSP at the end of the given titrations. However, also the peptide skeleton resonances belonging to aromatic or even positively charged amino acid residues like Thr18, Phe30, and Lys13 were highly perturbed, pointing out that the establishments of π-π and/or CH-π contacts between portion of the corresponding side chains and the rest of the backbone of the ligand (i. e. the alkyl part of the propoxy guanidinium groups or both the internal and the external sides of the macrocyclic cavity) could contribute to the recognition process and/or cause modification in the neighboring peptide bonds. The information about the binding stoichiometry suggested by the two techniques are instead in partial disagreement, since the *n*= 0.85 value obtained by ITC suggested the coexistence of 1:1 and 1:2 protein-ligand complexes in solution[77], while the NMR investigations are more consistent with the idea that the protein bind only a ligand molecule at a time. Nevertheless, the ITC data were not sufficient to provide a definitive elucidation of the mode exploited by this calix[4]arene to bind to the target protein, although they confirmed the combination of different types of interactions and gave insights on the binding stoichiometry. Moreover, the results of the NMR investigations are quite difficult to be interpreted since, despite a certain preference for some regions was reported, the ligand does not seem to be particularly selective for specific parts of the GB 1 surfaces. Additionally, for some residues, together with a consistent CSP, a variation of the line width for the corresponding amide resonances was observed, with cross peaks which disappeared and reappeared during the titrations, indicating that maybe different binding events are contemporary tacking place[56,78]. Accordingly, it can be thus suggested that the tested calix[4]arene recognize GB 1 by laying on its surfaces to take advantage from the formation of different types of interactions, thus explaining the fact that a high number of amino acid residues of different nature seem to be affected by the binding event. The possibility that a second ligand molecule can bind to the protein could instead explain the observation of different phenomena affecting some amide resonances during the NMR titrations, since the two recognition processes could take place according to different mechanism.

Attempts to get a comparison with the data obtained from analogous investigations for the GB 1-**4** system failed. Experiments carried out by injecting a 1 mM solution of protein into a 50 μM solution of ligand **4** gave negligible heat of reaction and no further insights about the thermodynamics of this binding event could be collected. These results are not completely surprising since, as reported after NMR studies, the K_d of this ligand is around 10-folds lower than the one measured for compound **5** and, similarly to the case described above, the studied reaction was not expected to be highly

exothermic[62,64,65]. Considering that the binding affinity observed by ITC should be at least one order of magnitude larger than that obtained from an NMR titration[67,68], it might be possible that a moderate increase of the ligand in the calorimeter sample cell could give measurable heats of reaction. However, this titration, together with a second one where an enhancement of the protein concentration, although programmed could not be performed because of technical problems with the ITC instrument.

5.5 Conclusions

^{15}N HSQC NMR and ITC experiments were used to characterize binding processes involving different proteins and calix[4]arene-based ligands.

The first protein-ligand system to be investigated was composed by cyt *c* and compound **1**. Useful information about their reciprocal interaction have been extracted only by NMR titrations. In these experiments the ligand was found to recognize selectively two regions of the protein surfaces including only a small number of amino acid residues. Among them, Lys4 and one of the neighboring Lys87 and Lys89 were proposed to act as anchoring point for **1** on the protein surfaces. This is in agreement with what observed in the crystal structure reported in literature for the complex composed by the same protein and compound **2**, where it could be evidenced that the residues of both these lysines are included into the calix[4]arene cavity. Contrary to what observed by NMR and in the crystal structure of *cyt c* and ligand **2,** where a third molecule of **2** is also bound to Lys22, the performed NMR studies with ligand **1** did not evidence significant interaction with this amino acid residue. The increasing CSP of the amide resonances involved in the binding process as function of the increasing concentration of the ligand indicated that the studied recognition event was fast in the NMR time scale. By fitting the corresponding plot of experimental data with an appropriate binding model it was possible to obtain an estimation of the binding affinities associated with this phenomenon. The perturbations experienced by the two clusters of resonances which were more affected by the presence of **1** were suitable to be treated with independent 1:1 binding models, yielding K_d values of ~0.4 mM and ~1.2 mM, respectively. Since the K_d given by analogous NMR studies for the binding of compound **2** to the same regions of cyt *c* surfaces were as high as ~0.8 mM and ~1.6 mM, the performed experiments pointed out that the ability of the two ligands in the association with the target protein was similar. Attempts to get more information about the thermodynamics of the studied binding process by ITC were unsuccessful, since resulted in the development of negligible heat of reaction, probably because of both the weakness of the occurring interactions and the modesty of the enthalpic contribution to the overall binding energy. A slight enhancement of the concentrations for both the binding partners is proposed as possible solution of these problems. Unfortunately, a technical problem with the ITC instrumentation prevented the demonstration of this hypothesis. In analogy with the examples found in the literature, ligand **1** was

supposed to recognize target lysine residues according to two possible binding mechanisms, both of them including the contemporary establishment of electrostatic interactions with the ammonium group of the bound side chain and hydrophobic contacts with the alkyl part of it. To obtain further elucidation about this subject, tests to achieve co-crystallization of the studied protein-ligand complex were performed. However, despite a robot screen of different reservoir solutions highlighted the ability of the ligand to induce the jellification of the protein in the presence of sodium citrate or phosphate buffer and different additives, all the subsequent tests to optimize manually the experimental conditions did not yield the desired crystals. Considering that compound **2** was reported to interact similarly with cyt *c* but also to co-crystallize relatively easily with it, it has been proposed that the difficulties faced into obtain the same results with ligand **1** rises from the different structural features of the two calix[4]arenes derivatives and in particular from the higher flexibility of the four alanine groups that the latter ligand bear at the upper rim.

In this chapter were also presented ITC studies of the binding between GB 1 and compounds **3**-**5**, performed with the aim to obtain thermodynamic data to be combined with the information given by NMR titrations, carried out in a previous thesis work. Useful titrations were performed by placing the ligands in the sample cell instead than in the syringe, in order to decrease as much as possible their concentration and to avoid their aggregation in the reaction medium. The only protein-calix[4]arene system suitable to be studied by ITC was that formed by GB 1 and ligand **5**, since ligand **3** was not soluble enough in the used buffer and in the tested experimental conditions, while compound **4** did not produce significant heats of reaction once in contact with the protein, probably because of weak binding. The K_d value of **5** and GB 1, obtained by ITC as fitting parameter of the resulting binding isotherm is 8.9 μM, while the value determined previously by NMR results is ~80 μM, nearly one order of magnitude larger. As reported with other systems in literature, NMR seems to slightly overestimated K_d values compared to ITC. ITC also suggest the existence, besides the 1:1 species, also of a 1:2 protein to ligand complex. The analysis of the different thermodynamic contributions to the binding free energy pointed out that both the binding enthalpy and the binding entropy are favorable to binding and similar in magnitude, thus indicating that both electrostatic and hydrophobic interactions seem to be important for the recognition process. In the absence of a crystal structure for the considered complex it is quite difficult to clarify definitely the binding mechanism followed by the ligand to approach the target protein, but, according to both the NMR and ITC data in our hands and to previous examples found in the literature, some hypothesis can be proposed. More in details, it was suggested that dipole-ion or electrostatic interactions could involve the guanidinium pendant groups at the lower rim of **5** and the polar or negatively charged amino acid residues on the protein surfaces (i. e. Asp36, Glu15, Glu42, Gln32 and Asn35), while hydrophobic, C-H/π or cation/π interactions could take place between the outer or

inner aromatic surface of the calix[4]arene scaffold and aliphatic/aromatic groups (i. e. Thr18, Phe30) and maybe also positively charged amino acid side chains (Lys13).

5.6 Experimental section

5.6.1 General Information

All moisture sensitive reactions were carried out under nitrogen atmosphere, using previously oven-dried glassware. All dry solvents were prepared according to standard procedures, distilled before use and stored over 3 or 4 Å molecular sieves. Most of the solvents and reagents were commercial samples and used as such. Analytical TLC were performed using prepared plates of silica gel (Merck 60 F-254) and then revealed with UV light or with ninhydrin (5% in EtOH) when free amino groups were present on the studied compounds. Merck silica gel 60 (70-230 mesh) was used for flash chromatography and for preparative TLC plates.

^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV300 and Bruker AV400 spectrometers (observation of ^{1}H nucleus at 300 MHz and 400 MHz, respectively, and of ^{13}C nucleus at 75 MHz and 100 MHz, respectively) and partially deuterated solvents were used as internal standards to calculate the chemical shifts (δ values in ppm). All ^{13}C NMR spectra were performed with proton decoupling. For ^{1}H NMR spectra recorded in D_2O at temperatures higher than 25°C the correction of chemical shifts was performed using the expression $\delta = 5.060 - 0.0122\times T(°C) + (2.11\times10^{-5})\times T^2(°C)$ to determine the resonance frequency of water protons[79]. Electrospray ionization (ESI) mass analysis were performed with a Waters single-quadrupole spectrometer in positive mode with MeOH or CH_3CN as solvents. Melting points were determined on an Electrothermal apparatus in closed capillaries.

5.6.2 Protein production

All the reagents were from Sigma-Aldrich. Both unlabeled and ^{15}N-labelled proteins were over-expressed in Escherichia coli BL21 (DE3) and purified according to literature methods[22,24,46]. The purity and concentration of the protein were estimated by using 15% SDS polyacrylamide gel electrophoresis (SDS-PAGE) and UV-vis spectroscopy (Perkin Elmer Lambda 35). Luria-Bertani (LB) medium was prepared by dissolving in 1L of water 10 g N-Z amine, 10 g Yeast extract, 5 g NaCl and then by adjusting pH to 7. LB Agar plates were prepared according to same procedure, but with the addition of 15 g Agar and 2 mM $MgSO_4$. ^{15}N labelled minimal medium (MM) was produced with 50 mM Na_2HPO_4, 50 mM KH_2PO_4, 2g/L D-glucose, 1 g/L $(^{15}NH_4)_2SO_4$, 20 mM citrate/20 mM succinate pH 7.0, 20 mL/L 50x5052, 30 mg/L Thiamine), 1 µL/mL Carbenicillin (75 mg/mL), where 50x5052 contains 25 g glycerol, 7 mL water, 2.5 g glucose, 10 g lactose monohydrate.

5.6.3 Protein over-expression

E. coli BL21, optimized for IPTG-inducible recombinant protein over-expression, was transformed with the plasmid encoding for the chosen protein (see paragraphs below).

1 µl of DNA plasmid was added to 50 µl of BL21 competent cells, incubated for 30 min on ice and then treated by heat shock for exactly 45 s at 42 °C. To allow the bacteria to recover, 250 µl LB was added and the sample was incubated for 45 min at 37 °C. Two aliquots of 100 and 150 µL of the resulting mixture were then plated on a LB Carbenicillin plate, which was incubated overnight at 37 °C. The next day a pre-culture was set-up by picking a single colony and inoculating 5 ml of LB medium containing 2 mM $MgSO_4$ and 1 mM Carbenicillin for selection. The pre-culture was then incubated overnight at 30 °C. The next day LB medium containing again $MgSO_4$ and Carbenicillin was inoculated with the pre-culture. The cells were grown at 37°C until an OD value of 0.6-0.8 at 600 nm was reached in (1h 30 min-2h). Protein over-expression was induced by addition of IPTG (Isopropyl β-D-1-thiogalactopyranoside) to a final concentration of 1 mM. After 4 h of over-expression at 37 °C the cells were harvested by centrifugation for 20 min at 4500 rpm. The pellet was then re-suspended in 5 mM of lysis buffer (10 mM TRIS-HCl, 1 mM EDTA, pH 7.5) and frozen at -20 °C, for at least one night to be ready for the further purification steps.

For the preparation of ^{15}N-proteins when the OD value at 600 nm of 0.6-0.8 the cells were harvested by centrifugation for 20 min at 4500 rpm and the pellet was re-suspended in minimal medium. After 30 min of shaking (300 rpm) at 37 °C the culture was split in aliquots of 250 mL volumes in 2 L Erlenmeyer flasks. Protein overexpression was then induced similarly by adding IPTG to a final concentration of 1 mM.

5.6.4 SDS-PAGE

A total of 20 µl cell/protein extract was mixed with SDS buffer and boiled at 95°C for 5min. The purity of denatured, reduced samples were tested on 15% SDS polyacrylamide gel electrophoresis (80 minutes, 140 V). As a protein size marker 5 µl of EZ-Run™ Pre-Stained Rec Protein Ladder (Fisher BioReagents) were loaded. Coomassie Blue staining was performed as reported in literature[46].

5.6.5 Cyt c purification

^{15}N labelled and unlabeled samples of saccharomyces cerevisae cytochrome c C102T were over-expressed according to the standard procedures reported above, by transforming E. coli BL21 (DE3) with a pBTR1 plasmid[22,24].

After the over expression steps, the protein was purified as reported in literature[22]. A cell pellet, previously re-suspended during cell harvesting and frozen, was placed in a beaker of water to thaw and 1 mL of 1 mg/mL DNase was then added. The resulting sample was sonicated for a total of 10 min (5x2 min cycles using manual pulses of 5 s on and 2 s off) at 8 microns, allowing the mixture to cool on ice.

Then, the cell membranes and debris were removed by centrifugation (4°C, 11000 rpm, 15 min) and the supernatant was transferred into a beaker where $(NH_4)_2SO_4$, previously pulverized in a mortar, was added slowly to make a 2 M solution in the cell lysate/extract to cause precipitation of hardly water soluble proteins. After that the solution went cloudy, it was centrifuged (4°C, 11000 rpm, 15 min) and the supernatant was dialyzed overnight against 20 mM KH_2PO_4 buffer at 4°C.

The next day the sample was centrifuged (4°C at 11000 rpm for 15 minutes) and further purified by sequential steps of Fast Protein Liquid Chromatography (FPLC). The first had been performed on a carboxymethyl (CM) ion-exchange column, equilibrated with buffer A (20 mM KH_2PO_4 and 100 mM NaCl, pH 6). The elution of neutral or negatively charged proteins was induced by adding 10% buffer B (20 mM KH_2PO_4, 1 M NaCl, pH 6) to the eluent, while the increase of the fraction of buffer B to 23% allowed cyt *c* to come out. The same purification procedure was repeated by using the same buffers and gradients at pH 8, after concentration and reduction of the protein with 1mM sodium ascorbate. Size exclusion chromatography (SEC) was used as a further polishing step, separating proteins on the basis of their size. The column was equilibrated with the same Buffer A as above (pH 6) and a < 500 µL sample was carefully loaded onto the column. Cyt c eluted at 80 mL based on its molecular weight. The fractions of interest were collected, concentrated and exchanged to H_2O. Protein concentration was confirmed by UV-visible spectroscopy according to the Lambert-Beer law, using an extinction coefficient of ε=28.5 $mM^{-1}cm^{-1}$ at λ=550 nm[22] after reduction of the protein with 1mM sodium ascorbate.

5.6.6 GB1 purification

Protein over-expression was carried out according to standard protocol found in literature[46]. The GB1K10H plasmid was transformed onto *E. coli* BL21 (DE3) and the protein purification followed standard procedures[22,47]. The re-suspended cells from the culture were heated to 80°C and then directly cooled on ice. They were centrifuged (10°C, 11000 rpm, 10 min) and then poured into 15 ml of buffer A (20 mM TRIS/HCl pH 7.5).

Subsequently the protein underwent to FPLC purification steps onto anion-exchange resin. A diethylaminoethyl sepharose (DEAE) column was equilibrated in buffer A (20 mM TRIS HCl, pH 7.5) and, after that all the positively charged proteins came off, a gradient from 0 to 50% of buffer B (20 mM TRIS HCl, 1 M NaCl, pH 7.5) was applied to elute GB 1-QDD. Fractions containing the protein were pooled, concentrated and loaded into a second DEAE column, performed with a manual gradient of buffer B (four steps from 0% to 10%, 25%, 50% B). After concentration, the sample was further purified onto an SEC column (20 mM KPi, 50 mM NaCl pH 6.0). The presence of GB 1 and its purity were checked with SDS-PAGE and NMR (for ^{15}N-labelled samples), while its concentration was determined via UV spectroscopy using as extinction coefficient[80] ε_{280} = 99.7 $mM^{-1}cm^{-1}$.

5.6.7 NMR titrations

Typical samples for NMR titrations contained 0.1 mM ^{15}N cyt c in the presence of 20 mM KH_2PO_4 buffer, 50 mM NaCl, 10 % D2O at pH 6, for a total volume of 0.5 mL. The experiments were performed at 30°C by adding 0.6-15 μL aliquots of a 25 mM stock solution of ligand, previously dissolved in water at pH 6. Over the course of the titrations the protein was diluted by 1.14-folds and the samples were corrected to pH 6 ± 0.05 for each addition of **1**. Two-dimensional ^{1}H-^{15}N HSQC experiments were obtained a 600 MHz Varian NMR spectrometer equipped with a HCN cold probe, with spectral widths of 16 and 40 ppm for ^{1}H and ^{15}N, respectively. Ligand-induced chemical shift perturbations (CSPs) were analyzed on Biopack and NMRpipe[81] as $\Delta\delta$ with respect to the spectrum of pure cyt *c*[24]. Titrations were repeated three times to ensure reproducibility and spectral changes due to ligand binding were assessed in CCPN. Binding isotherms were obtained by plotting the magnitude of the chemical shift change (Δδ) as a function of the concentration of **1**. The data were fitted in Origin by non-linear least squares regression to a one-site binding model, with Δδ and [**1**] as the dependent and independent variables, respectively, and the dissociation constant (K_d) and the maximum chemical-shift change ($\Delta\delta_{max}$) as fitting parameters[56–58]. A global data analysis was performed in which the curves were fit simultaneously to a single K_d value, while $\Delta\delta_{max}$ was varied for each resonance. The calculated K_ds were identical whether or not tacking in account the final protein dilution.

5.6.8 ITC titrations

All the proteins and ligands stock solutions were prepared in an identical buffer, composed by 20 mM KH_2PO_4 buffer, 50 mM NaCl at pH 6, centrifuged and degassed before to be placed in the sample cell or in the syringe. The samples containing cyt c, GB 1 and **1** were dialyzed over-night against the same buffer before their use and then their concentrations were checked by UV-visible spectroscopy (ε_{550} = 28.5 $mM^{-1}cm^{-1}$, ε_{280} = 99.7 $mM^{-1}cm^{-1}$ and ε_{270} = 78.1 $mM^{-1}cm^{-1}$, respectively[22,80]) and corrected to the target values, while the ones containing ligands **3**, **4** and **5** were prepared by directly diluting 80 mM, 50 mM and 25 mM water solutions, respectively. The titrations were performed at 30°C using a Nano ITC (TA Instruments), after centrifugation (4500 rpm, 10 min) and degassing (586 mm Hg, 8 min) of all the samples. Depending on the features of studied systems, one of the ligands or of the proteins were placed in a 250 μL syringe and added every 600 s in known aliquots (24 x 10 μL, 250 rpm) to the sample cell containing the corresponding binding partner.

For all the protein ligand systems, control experiments in which both the binding partners were titrated alone against the buffer gave negligible heats of dilution, excluding the case **4** at 2 mM concentration, probably because of its involvement in aggregation processes[53] (see **Paragraph 5.4.3.1**). In the case of the cyt c-**1** system, the experiments were performed by placing the ligand in the syringe and the protein in the sample cell at 2mM and 50 μM concentration, respectively, as provided by standard protocols for

this techniques[7,62]. To avoid the occurrence of the aforementioned aggregation processes, for the systems composed of GB 1 and ligands **3**-**5** it was decided to reverse the experimental set up, that is to place the protein in the syringe and the calix[4]arene derivatives in the sample cell, in order to decrease as much as possible the concentration of the latter binding partner. The experiments were carried out by incremental addition of 1 or 0.5 mM solution of GB 1 to 50 μM **4** and **5**, respectively, while no investigations were performed on compound **3** since it was not soluble in the buffer even at this concentration range. Among all the performed titration, only those involving the GB 1-**5** system yielded significant heats of reaction in the tested conditions and the corresponding experimental data were fitted on NanoAnalyze according to standard procedures[65]. The stoichiometry (*n*) and the enthalpy of binding (ΔH), together with the dissociation constant (K_d) for the studied process were obtained as fitting parameters directly from the analysis of binding isotherm, while other thermodynamic parameters (i. e. the changes in entropy ΔS) were calculated from the equation $\Delta G= \Delta H-T\Delta S= RTlnK_a$ where T is the absolute temperature and R=8.314 $Jmol^{-1}K^{-1}$. Two independent titrations were performed for each ligand tested.

5.6.9 Co-crystallization tests

The hanging- and sitting drop vapour diffusion methods[10,71,72] were used for crystallization tests of the cyt *c*-**1** complex at 20°C. Drops for the former experimental configuration were prepared manually in 24-well plates, by mixing 1 μL volumes of the reduced protein, the ligand and the reservoir solution. Control drops were obtained by replacing the solution of the ligand with 1 μL of water. For the latter set up was the drops were instead prepared with an OryxNano robot from Douglas Instrumets with same strategy, but by decreasing the volumes of the combined solutions to 0.3 μL.

Preliminary hanging drops experiments were carried out according to literature procedure[29] by mixing 1.7 mM cyt c, 17 mM **1** and using a reservoir solution composed of 20% PEG 8000, 50 mM NaCl and 50 mM sodium cacodylate buffer (pH 6.0), but yielded only clear drops. Subsequently, a robot screening of 96 reservoir solution from JenaBioscience was performed and in 22 wells was observed the formation of red gels using analogous protein and ligand concentration (1.7 mM and 10-20 mM, respectively), with the control drops that remained devoid of any precipitate. The same jellification process was reproduced manually without varying the concentrations of the binding partners, with a reservoir solution containing 50 mM sodium citrate buffer (pH 4-5) and 25% PEG 8000 as sole precipitant. Several round of optimization to convert the observed gels in crystals were performed by changing the percentage and type precipitant (10-20-30% PEG 3350 or PEG 8000), by increasing the concentration and the pH of the buffer (100 mM and 6, respectively) or by decreasing the ligand and protein concentration (1.7 and 0.1-0.5 mM, respectively), but did not yielded the desired results. No

crystals were obtained also by repeating the same experiments with 50 mM phosphate citrate buffer at pH 4.2. The presence of Na_2SO_4, Li_2SO_4, NaCl, NH_4OAc, NaOAc at 50 mM or 1 M concentrations as additives for a reservoir solution originally composed of 50 mM sodium citrate buffer (pH 5) and 25 % PEG 8000 was then studied as further step of refinement of the experimental conditions. Once mixed to equal volumes of 1.7 mM cyt *c* and 1.7 or 50 mM **1** in manual hanging drop experiments, also these latter reservoir solutions were not suitable to induce the formation of the desired co-crystals.

5.6.10 Synthesis of the ligands

A batch of ligand **1** was already available in our labs[50] and was used as such after a check of its purity by ^{1}H NMR. Compounds **3**-**5** were instead synthetized according to procedures found in literature, starting from aliquots of 25,26,27,28-Tetrakis[3-phtalimidopropoxy]calix[4]arene **6**[51] or 5,11,17,23-Tetraamino-25,26,27,28-tetrapropoxycalix[4]arene **10**[52] as precursors. All the compounds show the same physico-chemical properties already reported in literature[51–53].

25,26,27,28-Tetrakis[3-aminopropoxy]calix[4]arene (7)

To a solution of **6** (0.9 g, 0.77 mmol) in absolute EtOH (50 mL), $N_2H_4{\cdot}H_2O$ (7.5 mL, 145.6 mmol) was added. The mixture was stirred overnight under reflux for 24 h and then was quenched by evaporating the solvent under reduced pressure. The residue was taken with DCM (20 mL) and washed with distilled water (20 mL). The aqueous layer was extracted with DCM (2 x 20 mL), while the organic layers were combined, washed with brine (2 x 20 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure. **7** was obtained as white solid, pure enough to avoid further purifications (0.37 g, 0.57 mmol; 74% yield): ^{1}H-NMR (CDCl3, 300 MHz) δ (ppm): 6.60-6.57 (m, 12 H, Ar*H* meta and Ar*H* para); 4.41 (d, 4H, *J* = 13.3 Hz, Ar*CH*$_2$Ar ax.); 3.95 (t, 8H, *J* = 7.1 Hz, OC*H*$_2$CH$_2$CH$_2$); 3.19 (d, *J* = 13.3 Hz, 4H, Ar*CH*$_2$Ar eq.); 2.89 (t, *J* = 7.1 Hz, 8H, OCH$_2$C*H*$_2$CH$_2$); 2.89 (quin, *J* = 7.1 Hz, 8 H, OCH$_2$CH$_2$C*H*$_2$); 1.84 (br s, 8H, CH$_2$CH$_2$*NH*$_2$).

25,26,27,28-Tetrakis[3-(trimethylammonium)-propoxy]calix[4]arene, Tetraiodide (8)

To a solution of **7** (0.18 g, 0.28 mmol) in dry MeOH (5 mL) $KHCO_3$ (0.14 g, 1.38 mmol) and CH_3I (0.25 mL, 4.0 mmol) were added. The reaction mixture was stirred overnight at r.t. and then was quenched by evaporating the solvent under reduced pressure. The residue was taken with a 9/1 mixture of MeOH/DCM (10 ml), the insoluble inorganic salts were filtered off and the solvents were removed from the filtrate under reduced pressure. The crude was triturated in Et_2O for 3 h to give **8** as yellowish solid (0.38 g, 0.28 mmol; quantitative yield): ^{1}H-NMR (300 MHz, D_2O) δ (ppm): 6.70 (d, 8H, J = 6.6 Hz, Ar*H* meta); 6.61 (t, 4H, J = 6.6 Hz, Ar*H* para); 4.25 (d, 4H, J = 13.5 Hz, Ar*CH*$_2$Ar ax.); 4.16 (t, 8H, J = 7.0 Hz, O*CH*$_2$CH$_2$CH$_2$); 3.39-3.24 (m, 12H, Ar*CH*$_2$Ar eq. and OCH$_2$CH$_2$*CH*$_2$); 3.01 (s, 36H, N(*CH*$_3$)$_3$); 2.19 (br quint., 8H, J= 7.0 Hz, OCH$_2$*CH*$_2$CH$_2$).

25,26,27,28-Tetrakis[3-(trimethylammonium)-propoxy]calix[4]arene, Tetrachloride (4)

To a solution of **8** (0.38 g, 0.28 mmol) in distilled water (10 mL) DOWEX11 resin (5 mL in volume) was added. The reaction mixture was stirred for 30 min, then the resin was filtered off, washed with distilled water (3 x 10 mL) and the filtrate was lyophilized under reduced pressure to give **4** as white solid (0.21 g, 0.21 mmol; 75% yield), pure enough to avoid further purifications: ^{1}H-NMR (300 MHz, D_2O) δ (ppm): 6.65 (d, 8H, J = 6.2 Hz, Ar*H* meta); 6.60 (t, 4H, J = 6.2 Hz, Ar*H* para); 4.23 (d, 4H, J = 13.5 Hz, Ar*CH*$_2$Ar ax.); 4.09 (t, 8H, J = 6.9 Hz, O*CH*$_2$CH$_2$CH$_2$); 3.31 (d, 4H, J = 13.5 Hz, Ar*CH*$_2$Ar eq.); 3.20-3.14 (m, 8H, OCH$_2$CH$_2$*CH*$_2$); 2.97 (s, 36H, N(CH$_3$)$_3$); 2.14 (br quint., 8H, J= 6.9 Hz, OCH$_2$*CH*$_2$CH$_2$).

25,26,27,28-Tetrakis[3-(N,N'-bis-(Boc)-guanidine)-propoxy]calix[4]arene (9)

To a solution of **7** (0.18 g, 0.28 mmol) in dry DCM (7 ml), [N, N'–bis-(Boc)]triflylguanidine (0.48 g, 1.22 mmol) and NEt_3 (0.15 mL, 1.08 mmol) were added. The reaction mixture was stirred overnight at r.t., then was quenched with distilled water (10 mL) and was vigorously stirred for additional 30 min. The organic phase was separated, washed with distilled water (3 x 10 mL), dried over anhydrous Na_2SO_4 and evaporated under reduced pressure to give **9** as white solid (0.47 g, 0.28 mmol; quantitative yield): ^{1}H-NMR (300 MHz, CDCl3) δ (ppm): 11.51 (s, 4H, *NH*Boc); 8.41 (t, 4H, J = 5.1 Hz, CH$_2$CH$_2$CH$_2$*NH*); 6.61 (m, 12H, Ar*H* meta and Ar*H* para); 4.41 (d, 4H, J= 13.2 Hz, Ar*CH*$_2$Ar ax.); 3.99 (t, 8H, J = 7.2 Hz, O*CH*$_2$CH$_2$CH$_2$); 3.60-3.56 (m, 8H, OCH$_2$CH$_2$*CH*$_2$); 3.22 (d, 4H, J= 13.2 Hz, Ar*CH*$_2$Ar eq.); 2.22 (quint, 8H, J = 7.2 Hz, OCH$_2$*CH*$_2$CH$_2$); 1.50 (s, 36H OC*(CH$_3$)$_3$*); 1.48 (s, 36H, OC*(CH$_3$)$_3$*).

25,26,27,28-Tetrakis(3-guanidiniumpropoxy)calix[4]-arene, Tetrachloride (5)

To a solution of **9** in 1,4-dioxane (0.47 g, 0.28 mmol) 37% HCl (2 mL, 23.5 mmol) and TES (0.5 mL, 3.1 mmol) were added. The reaction mixture was stirred overnight at r. t. and then quenched by evaporating the solvent under reduced pressure to obtain **5** as white solid (0.27 g, 0.28 mmol; quantitative yield), pure enough to avoid further purifications: ^{1}H-NMR (400 MHz, D_2O) δ (ppm): 6.80 (d, 8H, J = 8.0 Hz, Ar*H* meta); 6.59 (t, 4H, J = 8.0 Hz, Ar*H* para); 4.28 (d, 4H, J = 13.3 Hz, Ar*CH*$_2$Ar ax.); 3.93 (t, 8H, J = 7.0 Hz, O*CH*$_2$CH$_2$CH$_2$); 3.25-3.17 (m, 12H, Ar*CH*$_2$Ar ax. and OCH$_2$CH$_2$*CH*$_2$); 2.09 (quint, 8H, J = 7.0 Hz, OCH$_2$*CH*$_2$CH$_2$).

5,11,17,23-Tetra[N, N'-bis-(Boc)-guanidino]-25,26,27,28-tetrapropoxycalix[4]arene (11)

To a solution of **10** (0.41 g, 0.63 mmol) in dry DMF (5 ml), [N, N'–bis-(Boc)]Thiourea (0.86 g, 3.1 mmol), $HgCl_2$ (1.36 g, 5.0 mmol) and NEt_3 (1.6 mL, 11.2 mmol) were added. The reaction mixture was stirred overnight at r.t., then was quenched by adding AcOEt (10 mL) and the insoluble salts were filtered off. The filtrate was washed with brine (3 x 10 mL) and evaporated under reduced pressure. The crude was purified by flash chromatography (Et_2O/hexane 3/7-Et_2O/hexane 6/4) to give **11** as white solid (0.68 g, 0.42 mmol; 67% yield): ^{1}H-NMR (300 MHz, $CDCl_3$) δ (ppm): 11.59 (s, 4H, *NH*Boc); 9.82 (s, 4H, Ar*NH*); 6.90

(s, 8H, Ar*H* meta); 4.41 (d, 4H, J = 12.9 Hz, Ar*CH*$_2$Ar ax.); 3.81 (t, 8H, J = 7.4 Hz, O*CH*$_2$CH$_2$CH$_2$); 3.18 (d, 4H, J = 12.9 Hz, Ar*CH*$_2$Ar eq.); 1.86-188 (m, 8H, OCH$_2$*CH*$_2$CH$_2$); 1.48 (s, 36H OC*(CH*$_3$*)*$_3$); 1.44 (s, 36H, OC*(CH*$_3$*)*$_3$); 0.96 (t, 12H, J = 7.4 Hz, OCH$_2$CH$_2$*CH*$_2$).

5,11,17,23-Tetraguanidinium-25,26,27,28-tetrapropoxycalix[4]arene, Tetrachloride (3)

To a solution of **10** (0.67 g, 0.42 mmol) in 1,4-dioxane (10 ml), 37% HCl (1 mL, 33.6 mmol) and TES (0.81 mL, 5.0 mmol) were added. The reaction mixture was stirred overnight at r. t. and then quenched by evaporating the solvent under reduced pressure to obtain **3** as white solid (0.41 g, 0.42 mmol; quantitative yield), pure enough to avoid further purifications: ^{1}H-NMR (400 MHz, CDCl$_3$) δ (ppm): 6.65 (s, 8H, Ar*H* meta); 4.45 (d, 4H, J = 13.5 Hz, Ar*CH*$_2$Ar ax.); 3.87 (t, 8H, J = 7.4 Hz, O*CH*$_2$CH$_2$CH$_2$); 3.22 (d, 4H, J = 13.5 Hz, Ar*CH*$_2$Ar eq.); 1.89 (sext., 8H, J= 7.4 Hz, OCH$_2$*CH*$_2$CH$_2$); 0.90 (t, 12H, J = 7.4 Hz, OCH$_2$CH$_2$*CH*$_2$).

5.7 Bibliography

(1) Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100* (7), 2479.

(2) Stites, W. E. *Chem. Rev.* **1997**, *97* (5), 1233.

(3) Giuliani, M.; Morbioli, I.; Sansone, F.; Casnati, A. *Chem. Commun.* **2015**, *51* (75), 14140.

(4) Yin, H.; Hamilton, A. D. *Angew. Chem. Int. Ed. Engl.* **2005**, *44* (27), 4130.

(5) Sheng, C.; Dong, G.; Miao, Z.; Zhang, W.; Wang, W. *Chem. Soc. Rev.* **2015**, *44* (22), 8238.

(6) Kiessling, L. L.; Gestwicki, J. E.; Strong, L. E. *Angew. Chem. Int. Ed. Engl.* **2006**, *45* (15), 2348.

(7) Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q. *Int. J. Mol. Sci.* **2016**, *17* (2), 144.

(8) Meyer, B.; Peters, T. *Angew. Chem. Int. Ed. Engl.* **2003**, *42* (8), 864.

(9) Clarkson, J.; Campbell, I. D. *Biochem. Soc. Trans.* **2003**, *31* (Pt 5), 1006.

(10) Hassell, A. M.; An, G.; Bledsoe, R. K.; Bynum, J. M.; Carter, H. L.; Deng, S. J. J.; Gampe, R. T.; Grisard, T. E.; Madauss, K. P.; Nolte, R. T.; Rocque, W. J.; Wang, L.; Weaver, K. L.; Williams, S. P.; Wisely, G. B.; Xu, R.; Shewchuk, L. M. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *63* (1), 72.

(11) Kim, Taisum; Nimse, S. B. *Chem Soc Rev* **2013**, *42*, 366.

(12) Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. **2010**, 2715.

(13) Berg, Jeremy B, Tymoczko, John L, Stryer, L. *Biochemistry, 5th Edition*, 5th ed.; W. H. Freeman: York, N., Ed.; 2002.

(14) Schagger, H.; Cramer, W. A.; Vonjagow, G. *Anal. Biochem.* **1994**, *217* (2), 220.

(15) Hüttemann, M.; Pecina, P.; Rainbolt, M.; Sanderson, T. H.; Valerian, E.; Samavati, L.; Doan, J. W.; Lee, I. *Mitochondrion* **2012**, *11* (3), 369.

(16) Hüttemann, M.; Lee, I.; Samavati, L.; Yu, H.; Doan, J. W. *Biochim. Biophys. Acta - Mol. Cell Res.* **2007**, *1773* (12), 1701.

(17) Matsuno-Yagi, A. *J. Biol. Chem.* **1996**, *271* (35), 21629.

(18) Liu, X.; Kim, C. N.; Yang, J.; Jemmerson, R.; Wang, X. *Cell* **1996**, *86* (1), 147.

(19) Beręsewicz, M.; Kowalczyk, J. E.; Zabłocka, B. *Neurochem. Int.* **2006**, *48* (6), 568.

(20) Zhi-Bo Wang, Min Li, Yungang Zhao, J.-X. X. *Protein Pept. Lett.* **2003**, *10* (3).

(21) Pereverzev, M. O.; Vygodina, T. V; Konstantinov, a a; Skulachev, V. P. *Biochem. Soc. Trans.* **2003**, *31*, 1312.

(22) Crowley, P. B.; Chow, E.; Papkovskaia, T. *ChemBioChem* **2011**, *12* (7), 1043.

(23) Gorban, A. M. S.; Izzeldin, O. M. *Biochem. Mol. Biol. Int.* **1999**, *47* (5), 899.

(24) Volkov, A. N.; Vanwetswinkel, S.; de Water, K.; van Nuland, N. A. J. *J. Biomol. NMR* **2012**, *52* (3), 245.

(25) Louie, G. V; Brayer, G. D. *J. Mol. Biol.* **1990**, *214* (2), 527.

(26) Worrall, J. A.; Kolczak, U.; Canters, G. W.; Ubbink, M. *Biochemistry* **2001**, *40* (24), 7069.

(27) Lange, C.; Hunte, C. *Proc. Natl. Acad. Sci.* **2002**, *99* (5), 2800.

(28) Lin, Q.; Park, H. S.; Hamuro, Y.; Lee, C. S.; Hamilton, A. D. *Pept. Sci.* **1998**, *47* (4), 285.

(29) McGovern, R. E.; Fernandes, H.; Khan, A. R.; Power, N. P.; Crowley, P. B. *Nat Chem* **2012**, *4* (7), 527.

(30) Crowley, P. B.; Ganji, P.; Ibrahim, H. *ChemBioChem* **2008**, *9* (7), 1029.

(31) McGovern, R. E.; Feifel, S. C.; Lisdat, F.; Crowley, P. B. *Angew. Chem. Int. Ed. Engl.* **2015**, *54* (21), 6356.

(32) Zadmard, R.; Schrader, T. *J. Am. Chem. Soc.* **2005**, *127* (3), 904.

(33) Beissenhirtz, M. K.; Scheller, F. W.; Stöcklein, W. F. M.; Kurth, D. G.; Möhwald, H.; Lisdat, F. *Angew. Chem. Int. Ed. Engl.* **2004**, *43* (33), 4357.

(34) Gelder, B. F. Van; Slater, E. C. *Biochim. Biophys. Acta* **1962**, *58* (3), 593.

(35) Hamuro, Y.; Calama, M. C.; Park, H. S.; Hamilton, A. D. *Angew. Chem. Int. Ed. Engl.* **1997**, *36* (23), 2680.

(36) Sjobringso, U.; Bjorcksli, L.; Kasternll, W. *J. Biol. Chem.* **1991**, *266* (1), 399.

(37) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253* (5020), 657.

(38) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33* (15), 4721.

(39) Orban, J.; Alexander, P.; Bryan, P. *Biochemistry* **1992**, *31* (14), 3604.

(40) Margeta-Mitrovic, M.; Jan, Y. N.; Jan, L. Y. *Proc. Natl. Acad. Sci.* **2001**, *98* (25), 14649.

(41) Sjöbring, U.; Trojnar, J.; Grubb, A.; Akerström, B.; Björck, L. *J. Immunol.* **1989**, *143* (9), 2948 LP.

(42) Sjöbring, U.; Björck, L.; Kastern, W. *J. Biol. Chem.* **1991**, *266* (1), 399.

(43) Alexander, P.; Orban, J.; Bryan, P.; Fahnestock, S.; Lee, T. *Biochemistry* **1992**, *31* (14), 3597.

(44) Akerström, B.; Brodin, T.; Reis, K.; Björck, L. *J. Immunol.* **1985**, *135* (4), 2589 LP.

(45) Sturtevant, J. M. *Annu. Rev. Phys. Chem.* **1987**, *38* (1), 463.

(46) Studier, W. F. *Protein Expr. Purif.* **2005**, *41* (1), 207.

(47) Lindman, S.; Linse, S.; Mulder, F. A. A.; André, I. *Biochemistry* **2006**, *45* (47), 13993.

(48) Lindman, S.; Xue, W.-F.; Szczepankiewicz, O.; Bauer, M. C.; Nilsson, H.; Linse, S. *Biophys. J.* **2006**, *90* (8), 2911.

(49) Giuliani, M. *New multivalent calixarene-based ligands for protein recognition and inhibition*; PhD Thesis, Parma University, **2015**.

(50) Sansone, F.; Barboso, S.; Casnati, A.; Sciotto, D.; Ungaro, R. *Tetrahedron Lett.* **1999**, *40* (25), 4741.

(51) Bagnacani, V.; Sansone, F.; Donofrio, G.; Baldini, L.; Casnati, A.; Ungaro, R. *Org. Lett.* **2008**, *10* (18), 3953.

(52) Dudic, M.; Colombo, A.; Sansone, F.; Casnati, A.; Donofrio, G.; Ungaro, R. *Tetrahedron* **2004**, *60* (50), 11613.

(53) Bagnacani, V.; Franceschi, V.; Fantuzzi, L.; Casnati, A.; Donofrio, G.; Sansone, F.; Ungaro, R. *Bioconjug. Chem.* **2012**, *23* (5), 993.

(54) Baldini, L.; Melegari, M.; Bagnacani, V.; Casnati, A.; Dalcanale, E.; Sansone, F.; Ungaro, R. *J. Org. Chem.* **2011**, *76* (10), 3720.

(55) Feichtinger, K.; Sings, H. L.; Baker, T. J.; Matthews, K.; Goodman, M. *J. Org. Chem.* **1998**, *63* (23), 8432.

(56) Williamson, M. P. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1.

(57) Fielding, L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51* (4), 219.

(58) Lian, L. Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *71*, 59.

(59) Casnati, A.; Sansone, F.; Ungaro, R. *Acc. Chem. Res.* **2003**, *36* (4), 246.

(60) McGovern, R. E.; Snarr, B. D.; Lyons, J. A.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. *Chem Sci* **2015**, *6* (1), 442.

(61) Freyer, M. W.; Lewis, E. A. In *Biophysical Tools for Biologists, Volume One: In Vitro Techniques*; Methods in Cell Biology; Academic Press, 2008; Vol. 84, pp 79–113.

(62) Pierce, M. M.; Raman, C. S.; Nall, B. T. *Methods* **1999**, *221*, 213.

(63) Titration, I.; Itc, C. *Life Sci.* **2000**, No. 10121042406, 1.

(64) Velazquez-Campoy, A. *J. Therm. Anal. Calorim.* **2015**, *122* (3), 1477.

(65) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. *Anal. Biochem.* **1989**, *179* (1), 131.

(66) Doolan, A.; Rennie, M.; Crowley, P. B. *Unpublished data*; 2016.

(67) Buschmann, H. J.; Mutihac, L.; Schollmeyer, E. *J. Incl. Phenom. Macrocycl. Chem.* **2003**, *46* (3–4), 133.

(68) Volkov, A. N.; Bashir, Q.; Worrall, J. A. R.; Ullmann, G. M.; Ubbink, M. *J. Am. Chem. Soc.* **2010**, *132* (33), 11487.

(69) Olsson, T. S. G.; Williams, M. A.; Pitt, W. R.; Ladbury, J. E. *J. Mol. Biol.* **2008**, *384* (4), 1002.

(70) Amzel, L. M. In *Energetics of Biological Macromolecules, Part C*; Methods in Enzymology; Academic Press, 2000; Vol. 323, pp 167–177.

(71) McPherson, A. *Eur. J. Biochem.* **1990**, *189* (1), 1.

(72) Krauss, I. R.; Merlino, A.; Vergara, A.; Sica, F. *Int. J. Mol. Sci.* **2013**, *14* (6), 11643.

(73) McGovern, R. E.; McCarthy, A. A.; Crowley, P. B. *Chem. Commun.* **2014**, *50* (72), 10412.

(74) Shinkai, S.; Arimura, T.; Araki, K.; Kawabata, H.; Satoh, H.; Tsubaki, T.; Manabe, O.; Sunamoto, J. *J. Chem. Soc. Perkin Trans. 1* **1989**, 2039.

(75) Rodik, R. V.; Klymchenko, A. S.; Jain, N.; Miroshnichenko, S. I.; Richert, L.; Kalchenko, V. I.; Mély, Y. *Chem. Eur. J.* **2011**, *17* (20), 5526.

(76) Sansone, F.; Dudič, M.; Donofrio, G.; Rivetti, C.; Baldini, L.; Casnati, A.; Cellai, S.; Ungaro, R. *J. Am. Chem. Soc.* **2006**, *128* (45), 14528.

(77) Cecioni, S.; Lalor, R.; Blanchard, B.; Praly, J. P.; Imberty, A.; Matthews, S. E.; Vidal, S. *Chem. Eur. J.* **2009**, *15* (47), 13232.

(78) Gordo, S.; Martos, V.; Santos, E.; Menéndez, M.; Bo, C.; Giralt, E.; de Mendoza, J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16426.

(79) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62* (21), 7512.

(80) Byeon, I.-J. L.; Louis, J. M.; Gronenborn, A. M. *J. Mol. Biol.* **2003**, *333* (1), 141.

(81) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6* (3), 277.

Acknowledgments

At the end of this work many peoples have to be mentioned in this section. First of all, I would like to thanks Prof. Alessandro Casnati and Prof. Francesco Sansone, who supervised me along these three years, always leaving me free to perform independently my research. Your experience, ideas and advices were fundamental to improve my skills and to render very stimulating my last cycle of studies. An additional remark is due to Alessandro Casnati for the efforts spent in the correction of this thesis.
Thanks also to Riccardo Salvio from "Università La Sapienza" of Rome, for his collaboration in the development of the phosphodiesterase mimics presented in this work and in particular for the kinetic measurements reported in **Chapter 2** and **3**.

I would like to say thank you to Dr. Peter Crowley, for hosting me in his group of research during my five-months placement in Galway, where I had the occasion to deepen my knowledge on the "protein's world". It is necessary to acknowledge Aishling Doolan and Martin Rennie for the ITC data about the cytochrome *c*-ligand **2** system and for the assistance with protein production, but more importantly for the friendly atmosphere that they were able to create in the lab, together with Sreenu, Jimi, Ciara, Kiefer and Antoine. It has been a great pleasure to work with you all.

Now, it is time for the informal acknowledgments, so I can skip to Italian: grazie Augusto per le mille serate passate al Carrol's e al Roisin Dubh e per la gita malata che abbiamo fatto a Castelderg quest'estate. Fai il bravo a Galway che a marzo ti aspetto a Parma per i festeggiamenti!!!
Grazie ai ragazzi dei lab. 192-193: Fede (sì, fai parte della lista, ma non ti montare troppo la testa), Ilaria, Marta, Mag, Davide, Lola, Sergio, Vap, Ste e a tutti quelli che ora dimentico di nominare, ma che hanno condiviso con me una parte di questo cammino. Come ben sapete, i calix sentono l'umore di chi li sintetizza e voi avete contribuito, a parte qualche logico momento di frustrazione, a rendermi sempre positivo e ottimista.
Ora è il momento dei ragazzi de La Villa. Lollo, Tia, Leti, Niky, Lucca, Cillo, Albi, Ciuich 1, Ciuich 2, Agne, Kommander, Lollino, Carol, Lucky, Tagghi, Dis, Dano, Davidino, Serena: che facciate parte dello "zoccolo duro" o che siate dei "nuovi arrivati", che ci vediamo spesso o una volta all'anno, è sempre un piacere stare insieme a voi. Vi voglio bene e, ora che questa tesi è finita, ricominceremo a vederci più assiduamente al bancone.
Maria Chiara (credo di non averti mai chiamata così, ma siamo pur sempre in una tesi sperimentale), fai parte del lab. 192 anche te, ma meriti una menzione a parte. Grazie per avermi sopportato in questi tre mesi di scrittura, ora tornerò ad essere più simpatico. Ci conosciamo da pochi mesi, eppure ci sarebbero

già tantissime cose da scrivere, ma non è questa la sede più adatta. Che dire, sei stata l'ultima persona che ho conosciuto durante questo dottorato… evidentemente le cose più belle succedono alla fine. Grazie di tutto.

Infine grazie alla mia famiglia, per aver sempre assecondato le mie scelte ed avermi sempre sostenuto durante il mio lungo percorso di studi.

Stefano