Tutor/s

Dr. Jordi Poater
*Departament de Química Inorgànica i Orgànica*

Grau de Química

# Treball Final de Grau

**Interactions in hydrogen-bonded macrocycles**
**Interaccions en macrocicles per pont d'hidrogen**

Montserrat Estellers Moreno

*June 2022*

UNIVERSITAT DE BARCELONA

B:KC Barcelona Knowledge Campus
Campus d'Excel·lència Internacional

*Science never solves a problem without creating 10 more.*

George Bernard Shaw

Primer de tot, agrair a en Jordi Poater, el meu tutor del treball de fi de grau, per tota la seva ajuda durant els últims mesos. També agrair als meus amics de la universitat per haver fet d'aquests últims anys una experiència inoblidable, i per últim i no menys important, a la meva família per tota la seva ajuda i recolzament.

# REPORT

# IDENTIFICATION AND REFLECTION ON THE SUSTAINABLE DEVELOPMENT GOALS (ODS)

The Sustainable Development Goals (ODS) are a set of global objectives aimed to achieve a sustainable health at all levels going from planetary biosphere to local community. The objectives of these ODS are to end poverty, to protect the planet and to ensure peace and prosperity for all human beings. The United Nations organization (UN), aimed at achieving the millennium development goals, created the 2030 agenda for humanity containing the 17 different ODS collecting the development of mankind under the banner of sustainability.

All 17 ODS can be grouped into five broad areas known as the 5P's which interact with each other to ensure sustainable development. The 5P's are planet, people, prosperity, peace, and partnership. Higher education institutions have a key role on implementing these sustainable development goals by offering lifelong learning adapted to challenges of the 21st century.

The collaborative and open research of supramolecular macrocycles is extensive, and these compounds have been reported to have several applications in different sectors going from pharmaceutical, used as drug enhancing excipients to the chemical industry with several uses in photocatalysis and the synthesis of new materials which contribute to an economy growth. This has all been possible by the previous studies on interactions in supramolecular macrocycles because, without all this previous research, it would not have been possible to know enough about these systems and introduce them in the industry innovation and health sectors, which are key on the prosperity of our society.

So, since all the studies collected on this bibliographic research conclude that the successful synthesis of supramolecular macrocycles is possible, it can be said that this work can help to achieve the ODS by transferring knowledge given by higher educational institutions capable of being helpful in different sectors of our society (health, economic, industry innovation) leading to the achievement of the prosperity of humankind according to the sustainable development goals (ODS).

# CONTENTS

# 1. SUMMARY

**Hydrogen bonded macrocycles** are a particular type of supramolecular structures built from several monomers which are bonded by **hydrogen bonds**. These macrocycles can be composed by different numbers of building blocks creating cyclic structures of different sizes. The dimension of these macrocycles will rely on the geometrical requirements of the self-assembly imposed by the nature of the monomers involved.

Since the synthesis of hydrogen bonded macrocycles is based on non-covalent interactions, its self-assembly is ruled by a **thermodynamic controlled equilibrium** that will determinate if the monomers self-assemble into the ring closed structure or the open oligomeric one. So, for the macrocycle to be synthetised, the factors ruling this thermodynamic equilibrium must be studied.

For this equilibrium to be shifted into the macrocycle architecture, certain conditions regarding the structure of the monomers and the characteristics of their binding sites must be met. If these characteristics are satisfied, the overall stability of the cyclic structure will be higher than the open form and therefore, the self-assembly will be in the macrocycle architecture.

The stabilizer factor of the self-assembly into the ring closed structure is known as **chelate cooperativity,** which can be mathematically defined by the **effective molarity (EM)**, a value quantifying how favourable the intramolecular interaction is over the intermolecular one, or in other words, how favourable the self-assembly into the cyclic structure is over the open oligomeric one.

So, the key on obtaining these hydrogen bonded macrocycles relies on molecular preorganization by using building blocks capable of enhancing EM values because its monomeric structure and binding sites characteristics can display large chelate cooperativities.

**Keywords**: Hydrogen bonded macrocycles, hydrogen bonds, thermodynamically controlled equilibrium, chelate cooperativity, effective molarity (EM).

# 2. RESUM

Els **macrocicles basats en enllaços per ponts d'hidrogen** són un tipus particular d'estructures supramoleculars construïdes a partir de la unió de diferents monòmers mitjançant **enllaços d'hidrogen**. Aquests macrocicles poden estar composats per diferents nombres de monòmers creant estructures cícliques de diferents mides. La dimensió d'aquests dependrà dels requisits geomètrics imposats per la naturalesa dels monòmers implicats.

Donat que la síntesis de macrocicles basats en enllaços per pont d'hidrogen té la seva base en interaccions no covalents, el seu autoassemblatge està controlat per un **equilibri termodinàmic** que determinarà si els monòmers s'enllacen formant una estructura oberta o una de tancada. Per tant, per poder sintetitzar amb èxit aquests macrocicles serà necessari estudiar els factors que afecten a l'equilibri termodinàmic.

Per tal de desplaçar l'equilibri cap a la formació de l'estructura cíclica, s'han de complir certes condicions relatives tant a l'estructura com a les característiques dels enllaços. Si les condicions són favorables, l'estabilitat global de l'estructura cíclica serà major que la de la forma oberta i, per tant, l'autoassemblatge dels monòmers serà cap a la forma tancada.

El factor estabilitzador de l'estructura cíclica per sobre de l'oberta és conegut com a **cooperativitat del quelat**, que pot ser definida matemàticament amb la **molaritat efectiva (EM)**, un valor capaç de quantificar com de favorable és la interacció intramolecular per sobre la intermolecular, o en altres paraules, com de favorable és la formació de l'estructura cíclica per sobre de l'oberta.

Per tant, la clau per obtenir estructures macrocícliques recau en la preorganització molecular mitjançant l'ús de monòmers capaços de fer augmentar al màxim el valor d'EM degut tant a que la seva estructura monomèrica com les característiques dels seus punts d'enllaç són capaces d'aconseguir una gran cooperativitat en el quelat.

**Paraules clau**: Macrocicles per ponts d'hidrogen, enllaços d'hidrogen, equilibri termodinàmic, cooperativitat del quelat, molaritat efectiva (EM)

# 3. INTRODUCTION

An emerging topic of supramolecular chemistry relies on the formation of well-defined, discrete self-assembled structures based on noncovalent interactions. These discrete well defined self-assembled structures are called macrocycles.

Macrocycles are supramolecular ring closed structures formed by several monomers bonded together by noncovalent interactions. The focus of this work will be on the macrocycles self-assembled by hydrogens bonds.

These macrocycles can be formed by different number of monomers creating cyclic structures of different sizes. The size of these hydrogen bonded macrocycles is dictated by geometric requirements of the monomers that compose them. Below, there is a representation of two different sized macrocycles based on six and four monomers nicknamed rosette and cyclic tetramer, respectively (Figure 1).

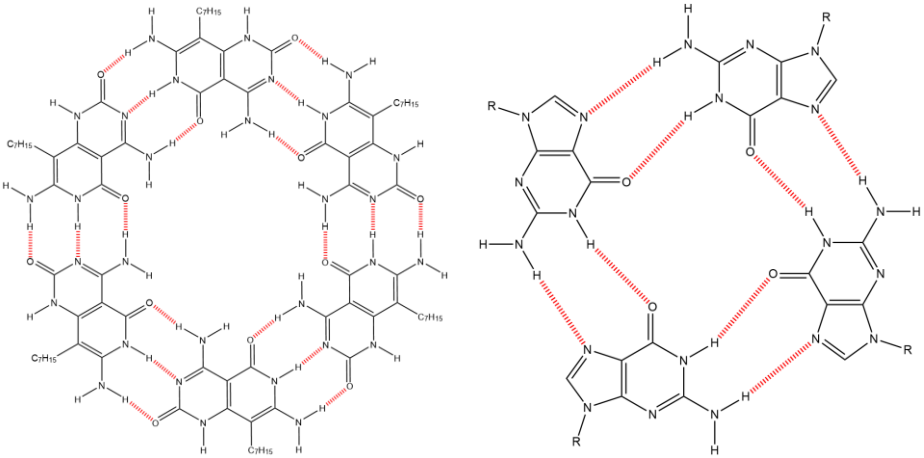


Figure 1. Self-assembly of two different monomers into two different sized macrocycles.

Macrocycles are synthesized by what is called supramolecular synthesis which relies on the principle of self-assembly that can be defined as the creation of a well-ordered architecture based on noncovalent interactions. Since the nature of these type of interactions is weak, these kinds of architectures are formed by an equilibrium controlled by several thermodynamic factors. (1)

In order to synthetise self-assembled structures, it is crucial to understand how noncovalent interactions work both individually and as a group, as well as to understand which factors will affect thermodynamically the process in order to obtain high yields on the synthesis of cyclic assemblies.

When a monomer with more than one binding site is used, it is expected that it will self-assemble forming an equilibrium between an open structure (linear) and a closed one (cyclic). (1) Since the goal of supramolecular chemistry is achieving well-defined, discrete supramolecular structures, macrocycles are usually the focus because its size can be controlled by geometric requirements and binding sites.

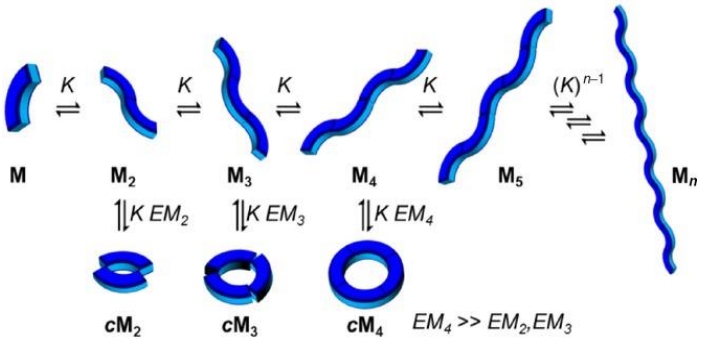



Figure 2. Thermodynamically controlled equilibrium on the self-assembly
Image from González-Rodríguez et al, ref. 1

So, in order to achieve these desired macrocycles, the question that must be answered is: how can the equilibrium be shifted between the open structure and the closed one? In this bibliographic study it will be discussed how to shift said equilibrium to form cyclic structures over linear ones by the analysis of the thermodynamic equilibrium and the binding interactions between the monomeric structures.

The key of obtaining ring closed structures relies on high chelate cooperativity, which can be defined as the cooperative effect that increases the thermodynamic stability of the ring over the open structure. In fact, there are other cooperative effects, but chelate cooperativity is the one

responsible for the "all or nothing" processes found in synthetic systems that produce discrete, well-defined supramolecular architectures, which means that chelate cooperativity will define if the cycle is completely assembled or totally dissociated into its monomers. (2)

Chelate cooperativity is quantified by the effective molarity (EM) which affords an estimate of how favourable the intramolecular binding interaction is over the intermolecular one. (1) In other words, to synthesize supramolecular macrocycles quantitatively it is necessary to know how to achieve the highest values of EM.

To be able to do that, it is important to understand what factors will affect chelate cooperativity. EM can be mathematically defined by the ratio between the equilibrium constant leading to a cyclic system ($K_{intra}$) and the one leading to the open one following equation 1:

$$EM = {K_{intra}}/{K_{inter}} \qquad \text{Eq 1}$$

Where $K_{intra}$ represents the equilibrium constant of the formation of the chelated system. (1) However, the magnitude of EM can also be regulated by enthalpic and entropic contributions associated to the cycle formation, as equation 2 represents: (2)

$$EM = e - (\Delta H^{o}_{intra} - \Delta H^{o}_{inter})/RT \cdot e(\Delta S^{o}_{intra} - \Delta S^{o}_{inter})/R \qquad \text{Eq 2}$$

The enthalpic component is dominated, amongst other, by electrostatic interactions that may occur in the cyclic assembly but, since these effects are rare and difficult to foresee, it is considered that the principal factor increasing enthalpic penalization would be the creation of strained macrocycles. If the macrocycle is unstrained, the enthalpic contribution to EM becomes negligible and its value depends on entropic contributions. It has been demonstrated that monomers with a preorganized structure that self-assemble into unstrained rings are most suited to produce higher EM than those that have more strained cycles. (1,2)

The entropic component is key to obtain either high or low EM values. In other words, the entropy factor has much more power over the cyclization process than the enthalpic contribution. The entropy of the formation of the cyclic systems depends on one hand on the symmetry and the number of components of the cycle since EM value tends to decrease when the self-assembly is composed by a relatively large number of molecules, and on the other hand, on the number of degrees of freedom (conformational and rotational) lost upon cyclization. (2)

As said before, this bibliographic research will be based on the study of the synthesis of supramolecular macrocycles based on non-covalent interactions, specifically, those formed by hydrogen bonds. Being that the point of this work, it is necessary to briefly discuss how hydrogen

bonds work in noncovalent synthesis as well as how to prevent entropic and enthalpic penalizations upon the self-assembled macrocycles.

## 3.1. HYDROGEN BONDS IN SELF-ASSEMBLY

Hydrogen bonds are attractive noncovalent interactions between two electronegative atoms, one that acts like the acceptor (A) because it has lone electron pairs, and the other one that is covalently bonded to an acidic proton that acts as a donor (D). The strength of this interaction, quantified by the association constant of the bond (Kass), will depend strongly on the solvent since it can compete with the formation of hydrogen bonds depending on its polarity.

Hydrogen bonds have taken special interest in the self-assembly of supramolecular macrocycles because of its high directionality and specificity. These noncovalent interactions are very abundant in biological systems being one of them the Watson-Crick base pairing on the DNA and RNA genetic structures [3].

DNA and RNA nucleosides such as guanine(G)-cytosine(C), adenine(A)-thymine(T) or uracil (U)-thymine(T) in RNA, are complementary bases because they can recognise one another and form hydrogen bonds creating a structure known as the double helix. It is important to highlight that the complementary G-C bases have an ADD-DAA hydrogen motif, whereas A-T or U-T present a DA-AD motif (Figure 3).

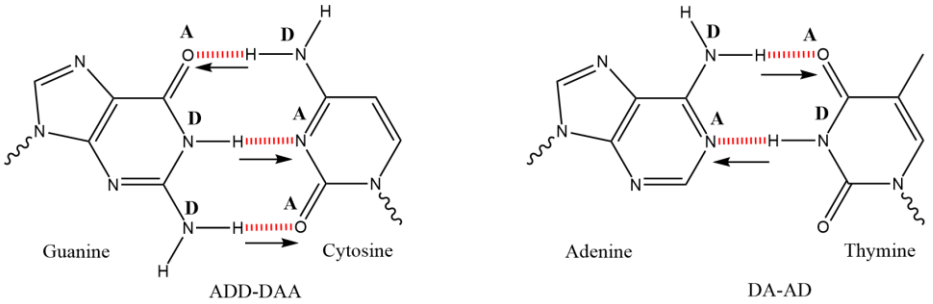



Figure 3. Hydrogen bond pattern between DNA complementary bases

The synthesis of supramolecular macrocycles is highly oriented on using DNA nucleosides derivates as building blocks to create new supramolecular structures with special applications. What needs to be understood is that Watson-Crick bases without any modifications can only form hydrogen bonds with one face of the molecule and that is why the self-assembly is into the double

helix. To synthesize supramolecular macrocycles, both faces of the monomer unit must be able to form hydrogen bonds, so the point is to synthesize DNA analogues equipped with two hydrogen bonding faces and use them as building blocks.

By the comparison of association constants (Kass) of the triply hydrogen bonded complex guanine-cytosine ($10^4$-$10^5$ $M^{-1}$) and the double bonded complex adenine/uracil-thymine (10-$10^2$ $M^{-1}$), it can be stablished that three hydrogen bonded associations are more stable because they have a larger association constant (4). The problem comes when the comparison between different complexes formed by three hydrogen bonds is made.

According to the study made by Jorgensen and co-workers (4), there is a diminution of the association constant in chloroform between a guanine-cytosine (1. G-C) and uracil-2,6-diaminopyridine (2. U-DAP) dimers being the Kass for G-C around $10^4$-$10^5$ $M^{-1}$ and the Kass for U-DAP around $10^2$-10 $M^{-1}$, meaning that there is a difference of 2-3 magnitude orders (Figure 4).

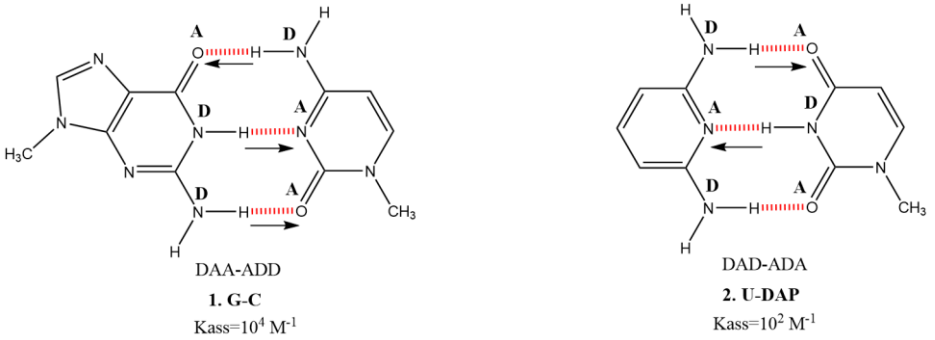



Figure 4. G-C and U-DAP dimeric structure

The reason of this difference is not immediately obvious since both have the same number of hydrogen bonds, but according to Monte Carlo molecular modelling and molecular dynamics simulations, the G-C dimer, which has a (ADD-DAA) hydrogen pattern, is more stable than the U-DAP dimer that has a (DAD-ADA) hydrogen array (4).

Since the primary hydrogen bonds are the same, there must be an explanation about the energetic difference between the two complexes. This difference was justified by the consideration of secondary interactions between the positively and negatively polarized atoms involved in the proximal hydrogen bonds (5).

Following the definition of hydrogen bonds seen before, the acceptor atom being an oxygen or a nitrogen will be partially negatively charged because of its lone electron pair, and the donor, being formally the hydrogen, will be partly positively charged in the formation of the hydrogen bond. This will create the secondary interaction between contiguous atoms that can be attractive or repulsive.

Understanding that and knowing the hydrogen bonding arrays of both complexes, the G-C dimer will have two attractive and two secondary interactions while the U-DAP complex will have four repulsive secondary interactions (Figure 5). That justifies the lower stability seen by Jorgensen and co-workers of the U-DAP dimer over the G-C one.

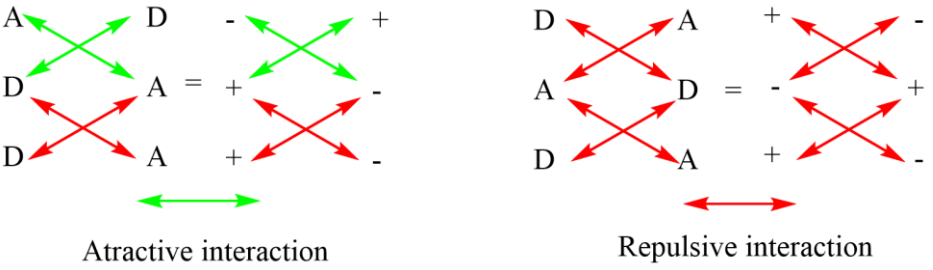



Figure 5. Secondary interactions

Jorgensen study about the effect of secondary interactions in hydrogen bond strength has been challenged by many other research groups over the years and, even though it has been stablished that there is an oversimplification in the nature of the hydrogen bond since they treat it just as interacting point charges and the study does not take under consideration orbital interactions, the general conclusion is that it remains a useful principle for designing multipoint hydrogen bonds interactions since the predictions of binding strengths are in line with experimental results. [6]

Knowing how secondary interactions work, one can reach the conclusion that even a (ADD-DAA) hydrogen array has high stability, a pattern of (DDD-AAA) should be even more stable because the four secondary interactions would be attractive. The association constants of DDD-AAA complexes in solvents like chloroform typically exceed $10^5$ $M^{-1}$ [7], while the one for ADD-DAA array has a value of $10^4$ – $10^5$ $M^{-1}$ [4]. For a system of DDD-AAA a remarkable value of $10^{10}$ $M^{-1}$ has been achieved [7] making this hydrogen bonding array extremely relevant but yet poorly explored due to the fact that the synthesis of these monomers involve low-yielding steps, especially for the AAA part. [7]

Anyway, thanks to the Jorgensen and Prenata studies [4], the trend on the Kass between triple hydrogen-bonds can be understood: DDD-AAA > ADD-DAA > DAD-ADA [8]. Since the Kass is a measure of how the equilibrium is shifted between the formation or the dissociation of the hydrogen bond, and that this constant depends on the symmetry of the referred bonding pattern, it is easy to foresee that the symmetry will also have an impact on achieving high EM values on the synthesis of hydrogen bonded macrocycles.

## 3.2 THERMODINAMIC CONSIDERATIONS OF THE SELF-ASSEMBLY

As it has been commented before in the introduction, the self-assembly of a supramolecular macrocycle will be ruled by an equilibrium between an open form and a closed one controlled by thermodynamic factors. It has also been commented what affects the entropic and the enthalpic factors of the cyclization, but what has not been discussed is how the monomeric structure itself can have an impact on the equilibrium.

In the last section, the pattern of the hydrogen bond array has been discussed since it is believed that it can affect chelate cooperativity leading to the conclusion that a proper use of hydrogen bonding patterns would help to enhance the value of EM. However, it is yet to be discussed if just by using an appropriate bonding pattern would be enough to shift the equilibrium completely into the cyclic form since it is unknown how that affects the entropic factor which is, in fact, the key to obtain the desired macrocycles.

Just to recap, chelate cooperativity is enhanced on one hand when no strain is generated upon cyclization because in such case the enthalpic factor becomes negligible and, on the other, when the smallest number of degrees of freedom (conformational and torsional) are lost during the formation of the cycle so, the key to quantitively obtain an unstrained macrocycle with a high EM value relies in the molecular preorganization of the building blocks.

One of the most basic considerations is that there will be a competition on the formation of hydrogen bonds between the acceptor and donor atoms in the complementary bases composing the monomers and the solvent molecules. To achieve the formation of the self-assembly, this competition must be won by the complementary bases and, to do that, the solvent used on the reaction must be as less polar as possible. That means that, when the design of the monomeric structure is being considered, the introduction of substituents in the building block capable of enhancing its solubility in non-polar media is important.

Having said that, a good example of why preorganization of the monomer is so important is the study made by Sessler and co-workers [2] where they studied the dimerization constant of the molecule 3 (Figure 6). This monomer uses guanine and cytosine as complementary bases, which means that it has an ADD-DAA pattern with a high association constant. Based on this, the dimerization of monomer 3 should be successful in a nonpolar media, but that is not the case.

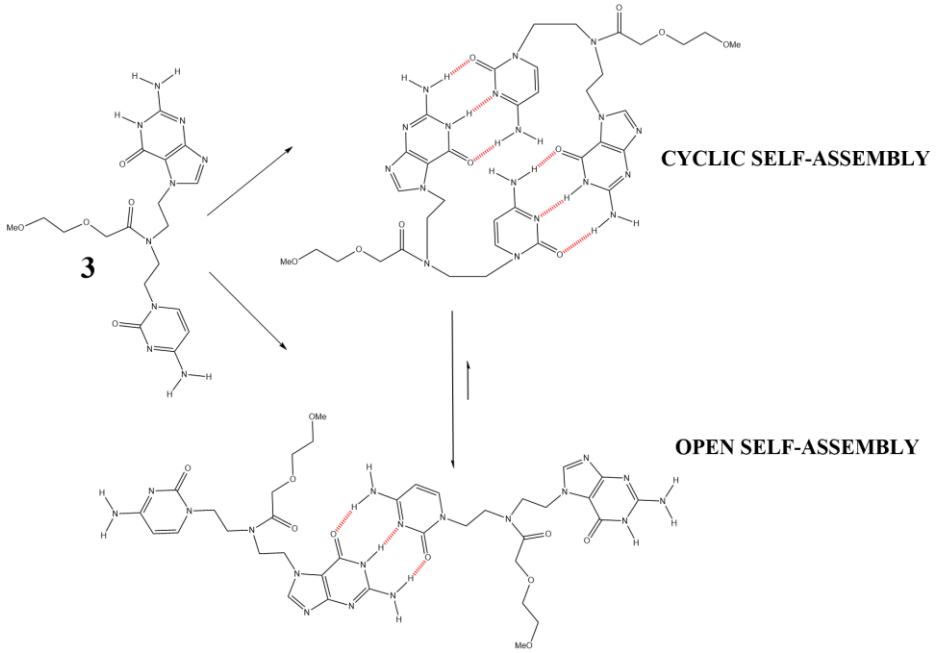



Figure 6. Self-assembly of monomer 3

Monomer 3 rather self-assemble into the open structure than the closed cyclic one because of the number of sigma bonds in its structure. Sigma bonds are directly proportional to the number of degrees of freedom the molecule will have since these types of bonds are able to rotate, a possibility that would be lost in the formation of the cycle. As said before, the loss of freedom degrees upon association creates a large entropic penalization decreasing chelate cooperativity and stabilizing the self-assembly into the open structure.

There is also the enthalpic penalization that comes from the strain generated in the cyclic structure as well as the lack of substituents able to increase the solubility of the complex in non-

polar media. Both penalizations in enthalpic and entropic values and the lack of solubility are what provoke the thermodynamically controlled equilibrium to be shifted into the open structure. (2)

The same research group, seeing the bad results of the first monomer, decided to study the dimerization of monomer 4 (Figure 7). This new monomer can only form two hydrogen bonds but, compared to monomer 3, it has a much more rigid structure with only 4 routable pi-conjugated bonds. Also, the complementary bases have a substituent that increases the solubility of the complex in non-polar solvents. (2)

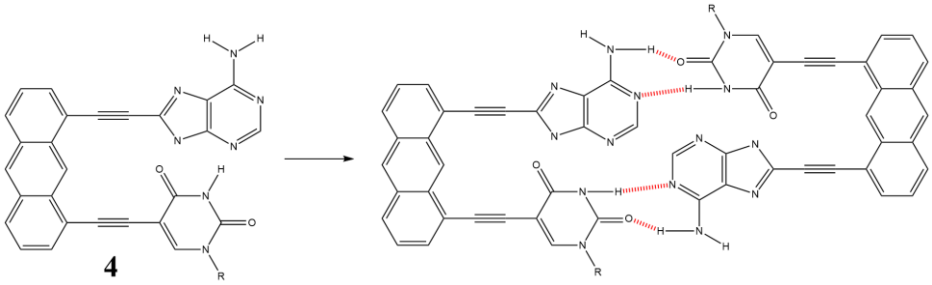


Figure 7. Self-assembly of monomer 4

Sessler and co-workers studied the dimerization process by adding volume fractions of DMSO, a polar solvent, into the solution of $CDCl_3$ and then following its course by $^1H$ NMR. They observed that the dimer complex based on monomer 4 could handle more fractions of DMSO than the one from 3 before the dissociation of the dimer into its monomers, meaning that the dimer based on monomer 4 is thermodynamically more stable than the one from monomer 3. (2)

This outcome stablishes that the rigidity of the monomeric structure is way more important in the stabilization of the closed dimeric structure over the open one than the number of hydrogen bonds between both monomers in the dimer. Therefore, the rigidity of the structure must be what influences the entropic value of the self-assembly, which, as mentioned before, is the key on obtaining high EM values.

However, this research group went even further, and they studied how the change of the spacer connecting the two complementary bases could affect the dimerization constant because of the creation of strain. To do that they compared monomer 4 and 5 (Figure 7 and 8 respectively). (2)

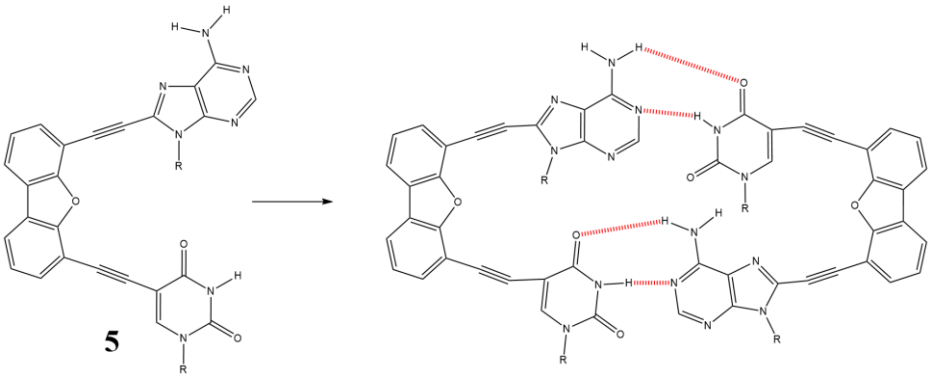



Figure 8. Self-assembly of monomer 5

As can be seen in the design of monomer 5, they used a spacer that orientates the complementary bases into a wide angle, forcing the complex to bend to form the hydrogen bonds while, in monomer 4, the spacer orients the bases into parallel planes allowing the closed dimer to be formed without strain.

This is corroborated experimentally by the relative stability measure of both dimers in fractions of DMSO where, by following the addition of the polar solvent by NMR, showed that while the dissociation of monomer 4 occurs when there is a 60% fraction of DMSO, only a 25% fraction was necessary to dissemble monomer 5. Moreover, they also observed that the dimer based on monomer 4 dissociated slowly while the one from 5 did it quickly in an NMR timescale, which meant that the kinetic behaviour was also different, concluding that dimer of monomer 4 is thermodynamically and kinetically more stable than the one from 5. [(2)]

The conclusion that should be extracted from these examples is that the rigidity of the monomer will be key to obtain high EM, not only because it will ensure the minimization of the loss of freedom degrees upon cyclization but also because it will guarantee the correct orientation of the hydrogen bonding faces, being both factors key on the obtention of the cyclic structure over the open one.

# 4. OBJECTIVES

The aim of this work is, via bibliographic research and using experimental examples, to show what factors, regarding the interactions and the structure, can affect the thermodynamically controlled equilibrium between the formation of a macrocycle and an open architecture generated during the self-assembly of a monomeric structure with two different binding sites.

The focus will be on how this equilibrium can be shifted to the macrocyclic form over the open architecture and on how this can be achieved by using monomers capable of enhancing chelate cooperativity for different reasons.

To do that, a literature survey on the synthesis of cyclic tetramers will be carried out by studying on one hand how different hydrogen bonding arrays (ADD-DAA, DAD-ADA) affect chelate cooperativity and, on the other, how different changes in the monomeric structure can induce variations in the EM value due to different reasons.

# 5. METHODS

All the information used in this bibliographic study has been found either in Web of Science, a data base of scientific knowledge by Clarivate Analytics, or in Google Scholar.

The research started by the three references enclosed in this research project call, based on previously studied guanine quartets and a particular rosette (six-membered macrocycle). First, some of the references cited by this initial manuscript were checked and next, the following step of the research procedure was to search on Web of Science the words “hydrogen bonded macrocycles” using the “article” filter. By doing that, a basic understanding of supramolecular macrocycles and its self-assembly was achieved.

After this first step, the research was oriented on what monomers were able to self-assemble into these cyclic structures and that is when several papers regarding the importance of monomeric structure and hydrogen bonding symmetry were found. Following that line of work, the studies on the self-assembly of cyclic tetramers achieving high chelate cooperativities used in the results and discussion section were encountered.

# 6. RESULTS AND DISCUSSION

During the introduction, the key concepts concerning the self-assembly of supramolecular macrocycles have been explained. Hydrogen bonds have been described as well as studied in the formation of dimers to see their relative strengths by the comparison of their association constants. It has also been commented that these hydrogen bonded macrocycles are formed via an equilibrium that is controlled by thermodynamic factors closely related to the monomeric structure as well as the characteristics of the binding sites.

Now, what remains to be studied is how these hydrogen bonds interactions and monomeric structure can affect chelate cooperativity and shift the thermodynamical controlled equilibrium towards the formation of ring closed macrocycles.

To study the effect of monomeric structure and binding characteristics on the self-assembly, this section will begin with a small introduction on the first ever synthesized monomers based on DNA and RNA complementary bases capable of self-assembling into stable macrocycles. After that, the presentation of a monomer capable of forming cyclic tetramers with incredible thermodynamic stability will be discussed followed by the study of how the modification of this first monomer can affect chelate cooperativity.

In the introduction, it is said that the ideal monomer is the one that can self-assemble into an unstrained macrocycle without losing any degrees of freedom in the process. The formation of the unstrained macrocycle comes from the correct orientation of the hydrogen bonding faces in the right angle while the responsible for the loss of freedom degrees upon cyclization is the number of rotatable bonds in the structure.

So, to begin, it is important to discuss the first ever successfully synthesized and characterised macrocycle based on DNA complementary bases. This success was accomplished by the groups of Lehn and Mascal. They pioneered the synthesis of monomers based on DNA hybrids nick-named Janus-type bases. These molecules incorporate both hydrogen bonding arrays of guanine and cytosine in one molecule without using any spacer to connect them achieving with that, a monomer with two hydrogen bonding faces (Figure 9).

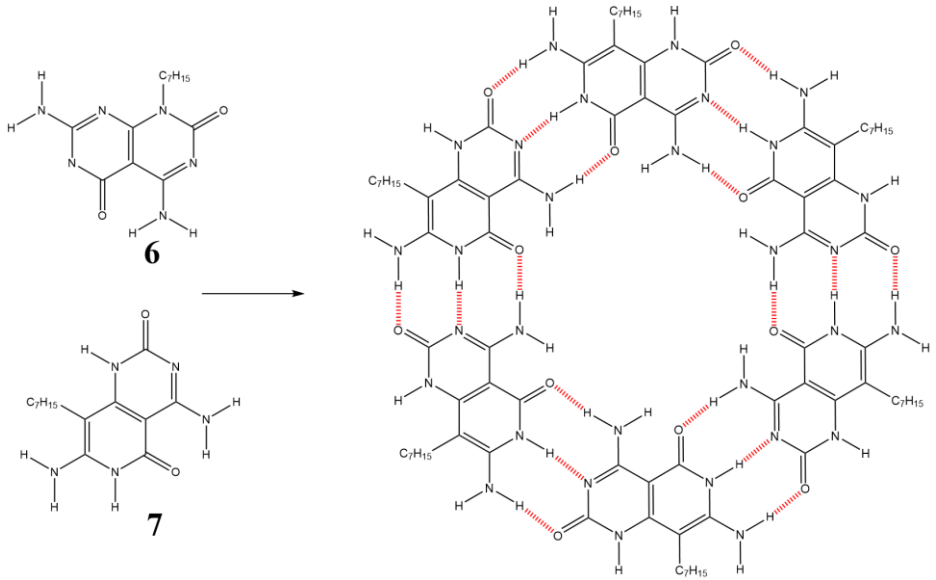



Figure 9. Janus-type rosette (six membered macrocycle)

The geometry of this hybrid situates both hydrogen bonding faces in a 120º of each other forcing the monomers to self-assemble into a hexagonal rosette. It is important to highlight that the high chelate cooperativities monomers 6 and 7 display is due to the ADD-DAA hydrogen bonding array as well as its monomeric rigidity. This allows the rosette to be self-assembled in polar solvents which is remarkable since the competition for the formation of hydrogen bonds will be high. [9]

Another factor allowing molecules 6 and 7 to be successfully self-assembled into hexagonal macrocycles is their perfectly defined topology, which means that the geometry of the monomer eliminates the possibility of a different cyclic structure to be formed other than the rosette. This factor, as well the lack of freedom degrees lost upon self-assembly and the use of ADD-DDA bonding pattern, are key on the high chelate cooperativity monomers 6 and 7 present.

In top of their rigidity, there is no strain generated upon cyclization since the hydrogen bonding faces are perfectly oriented making the enthalpic contribution negligible. All these factors may suggest that, on paper, these Janus-type bases may be the ideal monomer.

Reached this point and seeing that it was in fact possible to obtain stable macrocyclic structures by using rigid monomers with a favourable hydrogen bonding array, the question that

needed to be answered was if this Janus-type structure could be chemically programmed to self-assemble into different sized macrocycles.

Perrin and co-workers were the first ones to successfully achieve that. They used the same idea of the Janus-type base but with the introduction of a pyrrole ring as the central connector between the two complementary bases. This spacer forced the hydrogen bonding faces to be oriented into a 90º angle from each other compelling the self-assembly into a cyclic tetramer named G^C quartet (Figure 10). [10]

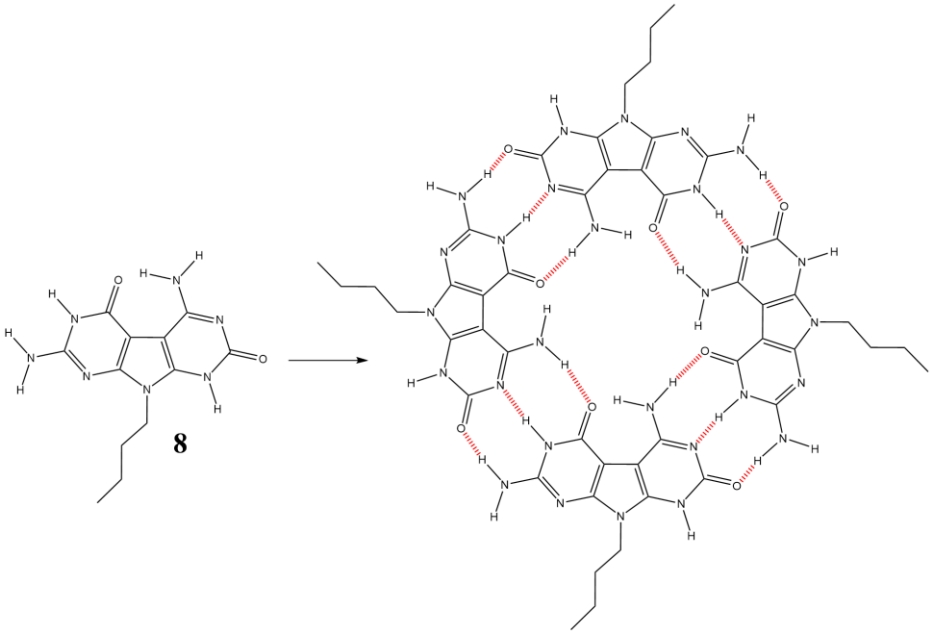



Figure 10. G^C quartet

To characterize the self-assembly, the equilibrium was studied by mass spectrometry, $^{1}H$ NMR spectroscopy, and diffusion-ordered spectroscopy (DOSY). All these characterization techniques provided conclusive evidence of the formation of the cyclic tetramer.

By adding the pyrrole spacer, the group of Perrin discovered that, with a small modification of the monomeric structure, a different sized macrocycle could be obtained without adding any extra degrees of freedom that would penalize the cyclization.

Perrin also investigated a Janus type molecule based on adenine and thymine hybrids [11], but the results were not satisfactory since the molecules they used as building blocks were not soluble in non-polar solvents and therefore the experiment had to be carried out in DMSO.

Parting by the fact that the hybrids, which were based on a DAD-DAD hydrogen bonding array, had a smaller association constant (Kass) than the ADD-DAA pattern, self-association in a polar environment was not observed, meaning that Janus type molecules based on DAD-ADA patterns are less useful to synthesize cyclic structures.

These experiments carried out by the groups of Lehn, Mascal and Perrin laid the groundwork for future researchers to achieve the synthesis of different types of macrocycles with high chelate cooperativity. It also allowed the study of the factors concerning the different noncovalent interactions and type of monomeric structures that would enhance the EM value and therefore improve the synthesis of supramolecular macrocycles.

Having stablished this basis, a series of experimental studies will be presented on how and why the obtention of a cyclic tetramer based on the DNA complementary bases guanine and cytosine can be achieved. It will also be shown, using this same monomer, if via modifications on the monomeric structure either by the addition of substituents or by varying the central connector between the complementary bases, the EM of the self-assembly can be increased.

Below, there is a table containing the experimental calculated thermodynamic and kinetic values for the monomeric structures in different solvents of different polarity that will be explained in this section. These values will be key to understand how both the interactions and monomeric structure affect the enthalpic and entropic value and therefore the effective molarity.

| Monomer | Solvent | Kass ($M^{-1}$) | $\Delta H$ (KJ $mol^{-1}$) | $\Delta S$ ($Jmol^{-1}$ $K^{-1}$) | EM (M) | Ref |
|---|---|---|---|---|---|---|
| GC | DMF | 5.7 +/- 0.3 | -155 | -425 | 218 | 8, 13 |
| | THF | 1.5 +/- 0.1 x $10^3$ | -225[a], -98.7[b] | -465[a], -32.3[b] | 180 | 8,13[a],14[b] |
| | $CHCl_3$ | 2.8 +/- 0.3 x $10^4$ | ----- | ----- | 730-910 | 8, 13 |
| | 1:1 $CDCl_3$/DMSO-$D_6$ | | -142 | -387 | ----- | 8, 13 |
| iGiC | DMF | 6.1 +/- 0.8 | ----- | ----- | 246 | 8 |
| | THF | 1.7 +/- 0.6 x $10^3$ | ----- | ----- | 294-463 | 8 |
| | $CHCl_3$ | 3.2 +/- 0.5 x $10^4$ | ----- | ----- | 314 | 8 |
| AU | DMF | ----- | ----- | ----- | ----- | ----- |
| | THF | ----- | ----- | ----- | ----- | ----- |
| | $CHCl_3$ | 2.5 +/- 0.4 x $10^2$ | ----- | ----- | 0.10 | 8 |
| $GC_{AR}$ | DMF-$D_7$ | ----- | -86 | -190 | ----- | 13 |
| | THF | ----- | -196 | -347 | ----- | 13 |
| | 1:1 $CDCl_3$/DMSO-$D_6$ | ----- | -101 | -240 | ----- | 13 |
| GC2 | DMF | 5.7 | -166.3 | --558.8 | ----- | 14 |
| | THF | 1.5 x $10^3$ | -91.9 | -66.3 | 2.4 | 14 |
| | $CHCl_3$ | 2.8 x $10^4$ | ----- | ----- | 1.1 x $10^1$ | 14 |
| GC3 | DMF | 5.7 | ----- | ----- | ----- | 14 |
| | THF | 1.5 x $10^3$ | -95.8 | -87.6 | 1.6 x $10^{-11}$ | 14 |
| | $CHCl_3$ | 2.8 x $10^4$ | ----- | ----- | 4.9 x $10^{-1}$ | 14 |
| GC4 | DMF | 5.7 | ----- | ----- | ----- | 14 |
| | THF | 1.5 x $10^3$ | ----- | ----- | ----- | 14 |
| | $CHCl_3$ | 2.8 x $10^4$ | ----- | ----- | 3.1 x $10^{-2}$ | 14 |
| GC5 | DMF | 5.7 | ----- | ----- | ----- | 14 |
| | THF | 1.5 x $10^3$ | -101.6 | -159.8 | 1.2 x $10^{-3}$ | 14 |
| | $CHCl_3$ | 2.8 x $10^4$ | ----- | ----- | 2.2 x $10^{-3}$ | 14 |

Table 1. Thermodynamic and kinetic parameters calculated for all the monomers that will be seen in this section.

(a), (b) Different $\Delta H$, $\Delta S$ values obtained from the same research group on different studies of the same monomer.

## 6.1 MACROCYCLE SYNTHESIS WITH HIGH CHELATE COOPERATIVITY

The group of González-Rodríguez studied several self-assembled systems based on hydrogen bonds to be able to stablish a strategy on how to design a building block based on Watson-Crick complementary bases guanine and cytosine that could self-assembly into a stable cyclic tetramer with the highest EM value as possible. To do that, they studied the self-assembly of the following monomer via experimental techniques such as $^1$H NMR (Figure 11). [12]

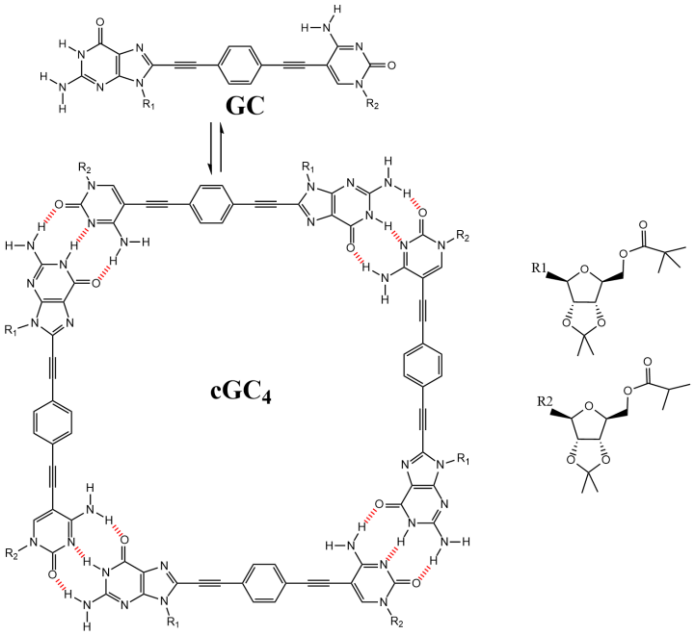



Figure 11. $cGC_4$ cyclic tetramer

Since they used the same complementary bases, the main difference between this monomer and the one used by Perrin and co-workers is the spacer. Perrin used a pyrrole ring to separate the complementary guanine cytosine bases and González-Rodríguez group chose to use a linear, $\pi$-conjugated, p-diethynyl benzene. This monomer also has a bulky lipophilic group in both guanine and cytosine that will make the structure more soluble in non-polar media.

González-Rodríguez and co-workers were able to prove that this monomer self-assembled into a cyclic tetramer in solution via experimental analysis as well as to portray a study of the thermodynamics and kinetics of the system. They also were able to quantify the EM value via experimental work. To study the thermodynamics and kinetics of the system, they designed

different experiments to see how easy the cyclic tetramer would be dissociated in different conditions.

The first one consisted of the addition of increasing portions of DMSO, a polar solvent, in order to see which mixture of DMSO/$CDCl_3$ could the cyclic tetramer survive before the dissociation into its monomers. They found out that the cyclic tetramer could resist large amounts of DMSO (until 80% v/v) which is quite notable for a hydrogen bonded structure. The fact that the cyclic tetramer survived into such high polar solvent fractions means that this macrocycle must have a large chelate effect keeping the cyclic tetramer from its dissociation.

They also observed that, when the dissociation occurred, the exchange between the cyclic structure and the monomer seemed to be very slow in the NMR timescale, which meant that the cyclic tetramer had a remarkably high kinetic stability. They also observed an "all or nothing" behaviour, meaning that no other supramolecular species were detected other than the cyclic tetramer. This same study performed with DMF instead of DMSO proved that the macrocycle does not dissociate into its monomers even when the solvent is only DMF.

They also studied the stability of the cycle by the addition of a competitor for the hydrogen bonding. They added increasing amounts of cytosine that would compete with the bonds already formed by the initial $cGC_4$ tetramer to see how strong the intramolecular hydrogen bonds were compared to the intermolecular ones that could be formed by the addition of extra cytosine.

Since the addition of cytosine into the system would compete to form hydrogen bonds, it is logical to think that, as the concentration of free cytosine increased, the weaker the intramolecular interactions would be resulting in a gradual dissociation of the cycle. The conclusion of this experiment was that the cyclic tetramer formed in solution was stable because its dissociation did not occur until large concentrations of cytosine were added to the system.

González-Rodríguez's group was also able to experimentally obtain the EM value as well as the association constant and thermodynamic values in different solvents. The values obtained for these parameters can be seen in Table 1. Thanks to the exhaustive analysis made of the self-assembly process of this building-block, it has been once again proved that the rigidity is key to successfully obtain the macrocycle.

Subsequent work of this group was based first on the understanding of why this monomer self-assembled into such a stable macrocycle and then, on how to increase EM value by making small modifications of the monomeric structure.

## 6.2 ROLE OF H-BONDING PATTERN IN CHELATE COOPERATIVITY

As it has already been discussed in the introduction, the strength of the hydrogen bond can be quantified by the association constant which is a measure of the equilibrium between the formation and the dissociation of hydrogen bonds. Since this value is affected by secondary interactions according to Jorgensen's study, [4] the following trend on the value of Kass can be stablished: DDD-AAA>ADD-DAA>DAD-ADA. [8]

Following this trend and knowing how the secondary interactions affect the binding strength, it is logical to think that the symmetry of the hydrogen face used by the complementary base in the monomer will have an impact on chelate cooperativity, therefore affecting EM value by increasing it if the Kass is large or decreasing it if the constant is small.

In order to study the role of the hydrogen bonding symmetry in the cyclization process, the group of González-Rodríguez studied and compared the self-assembly of three different monomers, all of them containing a rigid central block capable of reducing entropic penalizations with lipophilic bulky groups enhancing solubility. [8]

The difference between these three monomers were the complementary bases used. Two of them GC (studied in the last example) and iGiC were based on guanine-cytosine complementary bases, so they had an ADD-DAA hydrogen bonding pattern. The third one, AU, is based on adenine and uracil bases. In order to make this monomer competitive, the adenine edge was chemically modified by the addition of an amino group, making the possibility of forming a third hydrogen bond a reality. Since it is based on adenine and uracil, the hydrogen bonding pattern of this monomer was DAD-ADA. (Figure12)

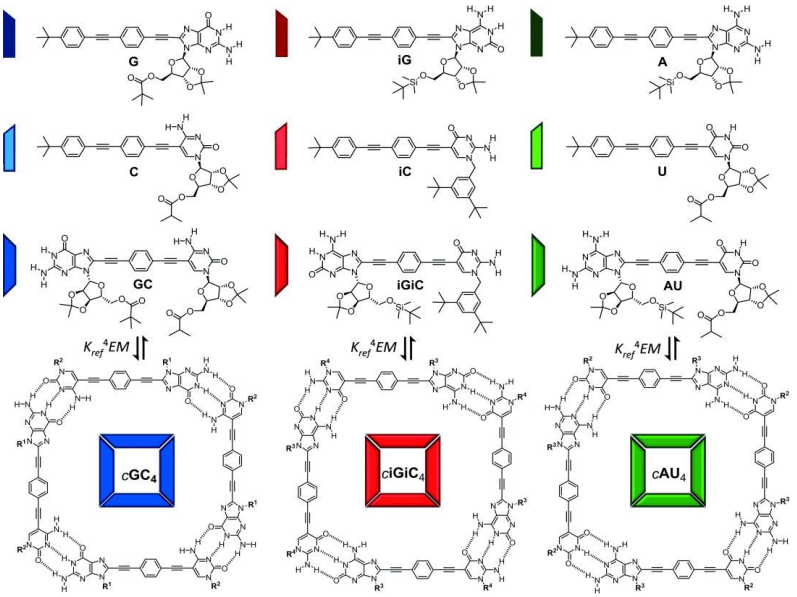



Figure 12. $cGC_4$, $ciGiC_4$ and $cAU_4$ cyclic tetramers
Image from González-Rodríguez et al, ref 8

It is easy to preview which monomer should self-assemble into a more stable cyclic tetramer by their hydrogen bonding arrays. Based on the conclusion reached by Jorgensen [4], ADD-DAA hydrogen bonding pattern will have a higher Kass meaning that the constant of the cyclotetramerization process and therefore EM value should be higher. Both EM and Kass values obtained for all three monomers in different solvents can be seen in Table 1.

By a closer study of Table 1 values, González-Rodríguez group noticed that while the data obtained for Kass are what could be expected for all three monomers, the EM value for $cAU_4$ was way smaller than what they predicted compared to the one's obtained for $cGC_4$ and $ciGiC_4$. That observation led them to think that maybe, there would be an unpredicted penalization factor other than the lower association constant making the EM value decrease.

To check the strength of the hydrogen bonds keeping the cycle together, they performed an experiment consisting of the gradual addition of the corresponding complementary nucleoside to a solution containing each cyclic tetramer. To carry out the experiment they used different solvents of different polarities.

González-Rodríguez group discovered that, in order to dissociate the cyclic tetramers $cGC_4$ and $ciG\text{-}iC_4$, were needed 60 equivalents of C and iC respectively, while the $cAU_4$ cycle was fully dissociated after the addition of about 3 equivalents of U regardless the solvent employed, which would affect the strength of the hydrogen bond between complementary bases. These results meant that the cyclotetramer of AU has a much weaker chelate effect than the other two.

In order to explain why GC and iGiC complexes form more stable cycles than AU ones, it is necessary to comprehend the geometrical requirements of the self-assembly. First of all, it is crucial to understand that, for all three monomers, the formation of the cycle is only possible when the building blocks involved in the formation of the hydrogen bonds are oriented in a 90º angle from each other.

However, there is another key factor, which is the conformational degrees of the monomers. In order for the monomers to self-assemble into the cyclic tetramer not only they must be oriented in a 90º angle, but they also must be found in the syn-conformation and, since all three monomers have rotatable bonds, all of them can be found in the anti-conformation as well, which would lead to the self-assembly into the open structure (Figure 13).

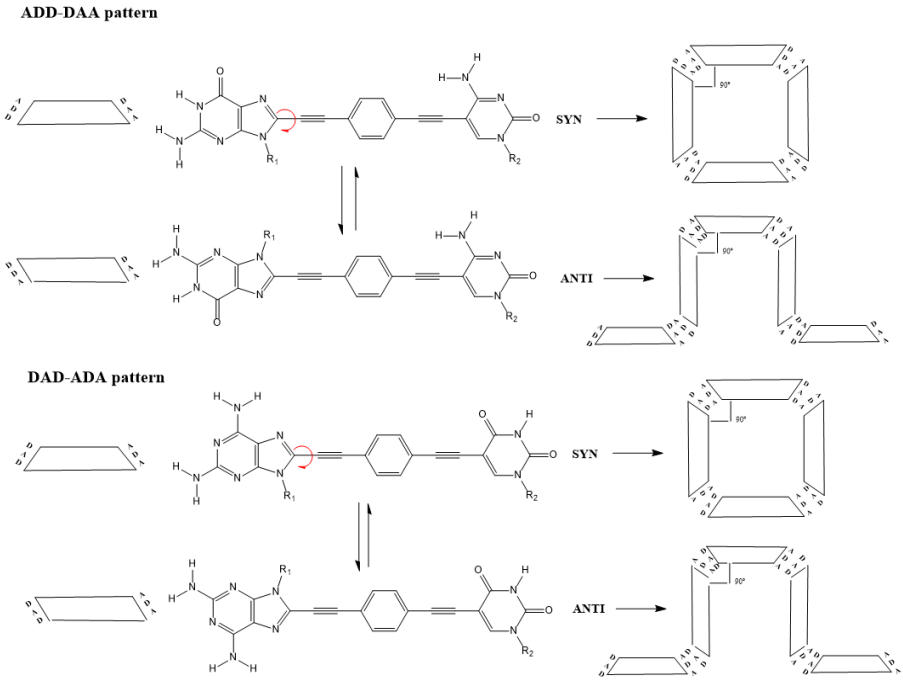



Figure 13. Geometry of ADD-DAA and DAD-ADA pattern. Self-assembly in a 90º

This is, in fact, a degree of freedom that will be lost upon cyclization for all of three complexes, so now the question remained to be answered is: if the three monomers have the same entropic penalization to form the cycle, why is the EM for $cAU_4$ way smaller than the one seen for $cGC_4$ and $ciGiC_4$? To answer this question, the hydrogen bonding pattern must be examined.

The difference between guanine-cytosine and adenine-uracil monomers comes from the fact that, when AU monomer is found in the anti-conformation, it can also self-assemble into the open structure when the angle between both monomers is 210º because of the symmetry of its hydrogen bonding array. At difference, GC and iGiC do not have this possibility since they do not have the symmetric hydrogen bonding motif (Figure 14).

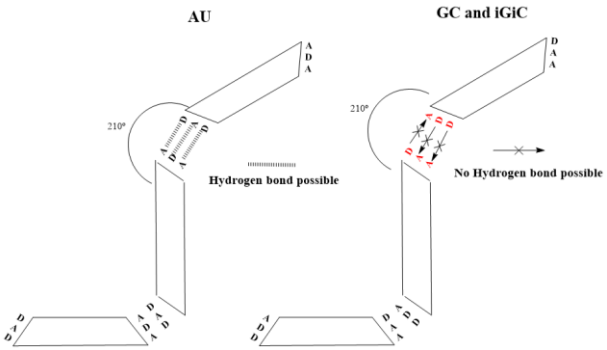



Figure 14. Self-assembly in a 210º

This extra degree of freedom that must be lost upon cyclization that AU has but the other two monomers do not, is the reason why the EM value for $cAU_4$ is way smaller than expected. This extra degree of freedom is what enhances the entropic penalization in the self-assembly of AU.

The conclusion that must be extracted from this example is that DAD-ADA hydrogen bonding arrays will, on one hand, form weaker bonds because of the repulsive secondary interactions which will be reflected on a lower association constant affecting the overall strength of the cyclic assembly, and on the other hand, generate an extra degree of freedom by the introduction of different biding modes creating an extra entropic penalization that will affect EM by decreasing its value.

So, in order to generate hydrogen bonded macrocycles quantitatively, monomers with an unsymmetric bonding array such as those based on guanine-cytosine (ADD-DAA) are recommended to be used over the ones with DAD-ADA arrays.

Nevertheless, there is still one hydrogen-bonding pattern that must be studied, DDD-AAA. In principle, one may think that, since this complementary monomers with this bonding pattern have been reported to have a Kass around $10^{10}$ [7], the self-assembly will form stable macrocycles, but following the conclusions from the last paragraph, this should be discussed.

González Rodríguez study [8] reached the conclusion that, even if a cyclic tetramer based on purine-pyridine couple could be synthetised with a DDD-AAA hydrogen pattern, which is impossible with the current knowledge, the conclusions could be extrapolated. This tetramer would have a high Kass, but a small value of EM, like $cAU_4$, for the very same reasons, being those the possibility of binding with two different angles.

With the last two examples portrayed experimentally by González-Rodríguez and co-workers [12, 8], it has been demonstrated that, both monomeric structure and binding interaction geometry are important to obtain unstrained macrocycles self-assembled with the minimal loss of freedom degrees possible.

## 6.3 EFFECT OF THE POLARITY OF THE SOLVENT

Having studied the importance of the rigidity as well as the role of symmetry in the obtention of macrocycles, the next step should be focused on answering the second part of the question presented before, which is how to increase the EM value. To do that, González-Rodríguez group though the introduction of a substituent in an atom involved on the formation of the hydrogen bond capable of enhancing its donor or acceptor characteristics, would increase the strength of the bond enhancing with it chelate cooperativity and obtaining higher EM values. To enhance chelate cooperativity, this research group introduced a para-electron poor aryl substituent (p-nitrobenzene) in the guanosine amino nitrogen acting as a donor. [13]

The objective of this modification is easy to understand. Following what has already been commented on the introduction, the hydrogen bond is formed between a donor, that is partially positively charged and an acceptor that is partially negatively charged. Therefore, the introduction of an electron poor aryl group as a substituent in the donor would increase its ability to form bonds since said substituent would attract electrons making the hydrogen bonded to the nitrogen to be more positively charged and, therefore more willing to bond with the lone electron pairs of the acceptor atom in the cytosine edge, since its acidity is increased.

There is one predictable problem though. The bond between the donor, being the nitrogen, and the guanine is a sigma bond, meaning that it can rotate. The fact that this bond can rotate is irrelevant when the nitrogen is not substituted because this nitrogen is bonded to two hydrogens that can interact with the acceptor equivalently. But, in this case the rotation of this sigma bond will create two different conformations that can affect the value of EM since the hydrogen of the donor atom involved in the formation of the bond needs to be properly oriented to interact with the oxygen acting as an acceptor in the cytosine.

In order to see if the substituent made the self-assembly more stable, this research group studied the known GC monomer and compared its thermodynamic values with the electron poor substituted one, $GC_{Ar}$, in non-polar and polar solvents. The values obtained can also be seen in Table 1.

González-Rodríguez group did not provide the EM values, but they did calculate the tetramerization constant ($K_T$) that, according to equation 1 must be directly proportional to the value of EM, since Kass should be the same for both monomers as they are using the same hydrogen motifs. So, considering that $K_T$ obtained for $cGC_{Ar4}$ and $cGC_4$ are of the same order of magnitude for all solvents but for THF, which is only 1 order of magnitude bigger for $cGC_{Ar4}$ than for $cGC_4$ ($10^{16}$ and $10^{15}$ $M^{-3}$ respectively), it can be concluded that $cGC_{Ar4}$ is not significantly more stable than $cGC_4$.

However, what is interesting from this example is that it was expected for $GC_{Ar}$ to have higher $\Delta H$ values than GC since the hydrogen involved in the formation of the bond is more acidic due to the electron attracting substituent in the nitrogen, which would mean a stronger hydrogen bond but that is not what the collected data in Table 1 reflected. Instead, the difference in the tetramerization constants between the cyclic tetramers must come from the decrease of the $\Delta S$ absolute values (Observe (a) values for GC monomer).

To justify these unexpected lower numbers, the research group thought that, if the hydrogen could form stronger bonds with the cytosine in these substituted monomers it must also bind strongly to the solvent molecules as well. And that is in fact the effect what would attenuate the increase on the $\Delta H$ value for $cGC_{Ar4}$.

The difference in the entropic term comes from solvation effects which increase with the polarity of the solvent. When a monomer with the capacity to form hydrogen bonds is introduced in a polar media, there will be a competition for the formation of these bonds between the solvent and the other monomers creating the solvation effect.

This solvation effect is stronger in the monomeric state since it has more donor and acceptor atoms available for the formation of hydrogen bonds, and that introduces some order in the solvation sphere. Since cyclization blocks some of these positions, the solvation sphere becomes more disordered when the monomers self-assemble. The dissociation of the solvation sphere produces a small decrease in the entropy upon association.

Recapitulating what has been said before, the presence of the aryl group leads to different conformations in the monomeric state because of the rotation of the sigma bond. The two possible conformations are presented below, but in order to self-assemble, the monomer has to be in conformation A since it is more stable for several factors (Figure 15).

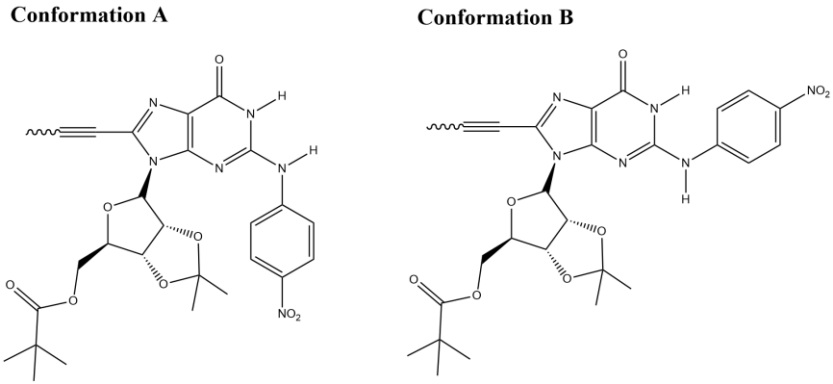



Figure 15. Two possible conformations based on the rotation of the amino sigma bond

What happens is that conformation A produces a certain sterically crowded region around the formation of the hydrogen bond that hinders solvation and, since the alteration of the solvation sphere induces the entropy to decrease, the substituted monomers have lower absolute values for $\Delta S$ making the overall association constant bigger.

So, the conclusion that must be extracted is that the introduction of an electron poor substituent in a donor atom involved in the formation of the hydrogen bond will, in fact, increase the EM value, but the increase of said value will not be significant since the tetramerization constants are about the same order. Also, the small increase in EM values does not come from a stronger interaction between the donor's and the acceptor's atom but from a change of the solvation sphere that results in smaller entropy values which make EM increase. This means that the solvation effect can play in favour and help to shift the equilibrium into the ring closed species.

## 6.4 INFLUENCE OF THE CENTRAL CONNECTOR

All the examples presented by González-Rodríguez group have proven experimentally that the monomeric structure and the nature of the complementary bases are key to obtain high chelate cooperativity, but there is still one factor that needs to be addressed.

If one compares the monomer used by Perrin and co-workers at the beginning of this section and the one from González-Rodríguez, it is easy to see that both are based on the same complementary bases, and both self-assemble into cyclic tetramers, but the only difference is the spacer used and therefore the size of the resulting cycle.

In order to see the effect that the spacer separating the complementary bases has on chelate cooperativity, or said in other words, how the size of the cycle affects EM value, González-Rodríguez group synthetised a series of monomers with guanosine and cytosine as complementary bases at the edge connected by spacers of increasing lengths resulting in cyclic tetramers of different sizes (Figure 16). [14]

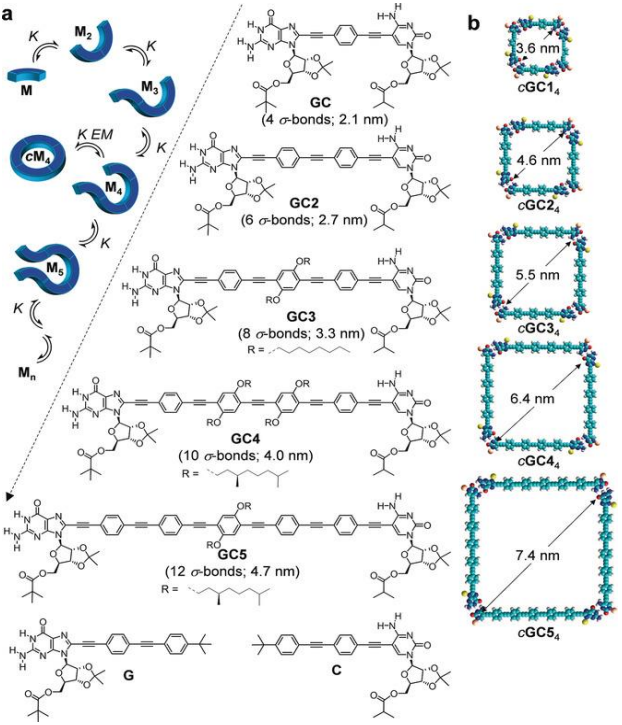



Figure 16. Self-assembly of monomers of increasing length
Image from González-Rodríguez et al, ref 14

It is logical to think that, if the size of the spacer is larger, the more sigma bonds it will have and therefore the more degrees of freedom (conformational and torsional) would be lost upon cyclization increasing its entropic penalties. In fact, that is exactly what they proved experimentally, but this research group went even further and tried to stablish a relationship between the number of sigma bonds that a monomer can have before the entropic penalties shift the equilibrium into the oligomeric structure by analysing EM values of each cycle in different solvents.

As it can be seen in the Table 1, EM value gets lower as the size of the cycle gets larger because the entropy value (observe (b) value for GC monomer) increases with the length of the spacer. Since the interaction by hydrogen bonds is the same in all the cycles, the weaker chelate cooperativity must be produced by entropic penalizations as it was predicted, since enthalpy does not suffer major changes.

They noticed that both ln(EM) and $\Delta S$ followed a linear relationship with the number of sigma bonds between sp and $sp^2$ carbons in the spacer while $\Delta H$ remained pretty much constant for all the assemblies (Figure 17 a and b respectively).

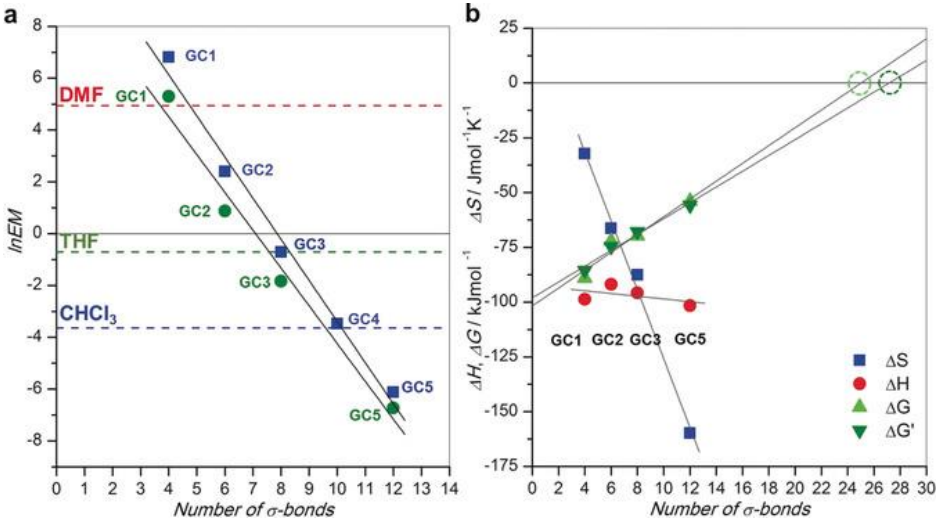



Figure 17. 17a) Graphic comparing ln EM vs number of σ-bonds.
17b) Graphic comparing thermodynamic values vs number of σ-bonds
Images from González-Rodríguez et al, ref 14

They also noticed that the relationship between the number of sigma bonds and the decrease of the EM value could be explained from two different points of view. One of them was by following the rule Ecrolani defined as the condition to form a macrocycle quantitively at a given concentration. Ecrolani rule is defined by the next equation where n is the number of the monomers in the cycle (Equation 3).

$$K_{ref} \cdot EM \geq 185 \cdot n \qquad \text{Eq 3}$$

On Figure 17a, there is a representation of ln(EM) vs the number of sigma bonds for all five cycles. The dashed lines show the limit above which this Ecrolani condition is met for the three main solvents studied here. In DMF (extremely polar) the only monomer that checks these conditions is GC, in THF GC, $GC_2$ and $GC_3$ can self-assembly quantitively and all but $GC_5$ form cyclic tetramers on $CHCl_3$. Monomer $GC_5$ do not meet Ecrolani's condition in any of the three studied solvents at a given concentration.

On Figure 17b, there is a representation of $\Delta H$, $\Delta S$, $\Delta G$ values vs the number of sigma bonds of each monomer when the solvent is THF. With these data, it is possible to estimate for which monomer the cyclization process becomes endergonic ($\Delta G^o > 0$), meaning the length at which the self-assembly into the open structure is preferable over the one into a cyclic tetramer regardless the concentration.

The conclusion of Figure 17b is that it is possible to know the maximum number of sigma bonds a monomeric structure can contain before its self-assembly is shifted into the open structure by the extrapolation of $\Delta G^o > 0$. According to the figure, the maximum number of sigma bonds a molecule could contain in its structure is 26.

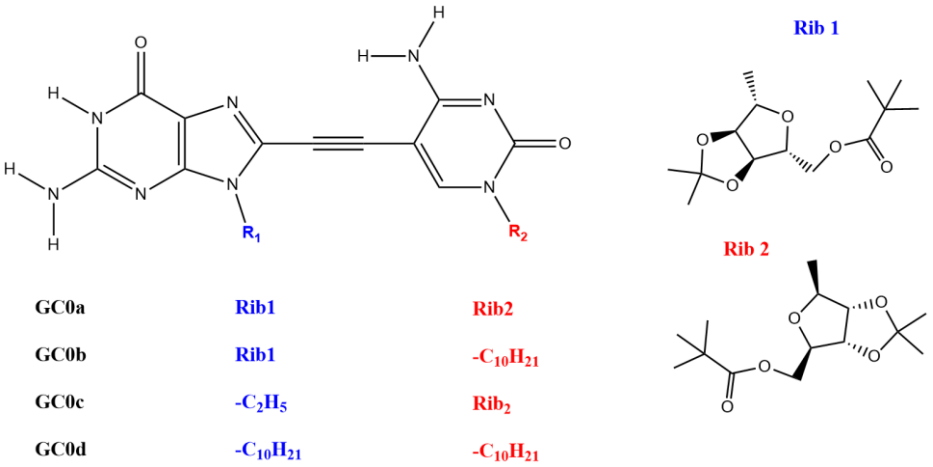
The problem with this analysis is that it strongly relies on the solvent employed, so this number is not a solid answer for every system. For example, if the solvent used was $CHCl_3$ the number of sigma bonds necessary to shift the equilibrium into the open structure would be larger because the value of Kass would increase and therefore the $K_T$ value would do the same.

It is necessary to highlight that the analysis made by González-Rodríguez of this monomer is only applicable to this monomeric structure and this specific binding interaction. They also did not take under consideration the alkyl chains in $GC_3$, $GC_4$ and $GC_5$ that needed to be introduced for solubility purposes. The length and position of this alkyl chains could influence the freedom of movement of the sigma bonds and introduce solvation, conformational and steric effects.

Anyway, the conclusions extracted from this study on how the length of the spacer can affect chelate cooperativity can be extrapolated to many supramolecular structures since the principle the study of González-Rodríguez is based on is that, as the length of the spacer gets larger, more sigma bonds will be needed in the formation of the monomeric structure and therefore more rotational and conformational degrees of freedom will be lost upon cyclization.

Further work of this group was aimed at the modification of this central spacer in order to try and reduce either the number of sigma bonds in the central connector by using shorter structures or by the addition of substituents in the spacer to block the rotation of the sigma bonds.

To evaluate how the chelate cooperativity would be affected using a shorter central spacer, González-Rodríguez group synthetised a batch of four new monomers using cytosine and guanine as complementary bases all connected by an ethynylene spacer. The difference between these new monomers is the substituents used on guanine and cytosine (Figure 18). [15]




| | | |
|---|---|---|
| **GC0a** | **Rib1** | **Rib2** |
| **GC0b** | **Rib1** | **-$C_{10}H_{21}$** |
| **GC0c** | **-$C_2H_5$** | **$Rib_2$** |
| **GC0d** | **-$C_{10}H_{21}$** | **-$C_{10}H_{21}$** |

Figure 18. guanine-cytosine monomers with ethynylene as the central connector

In theory, the use of a shorter spacer must enhance chelate cooperativity because the structure of the spacer will contain less sigma bonds and, therefore the number of degrees of freedom that would be lost upon association should be decreased, but another kind of penalizations in the self-assembly into the cyclic tetramer may appear.

One should consider two principal factors that could penalize the cyclization of these shorter monomers. First one is steric effect. Since the self-assembly into the cyclic tetramer requires the monomer to be in the syn-conformation, the bulky lipophilic substituents present in both edges of the monomer would create this steric hindrance because both of them will be oriented in the same space frame.

Another consideration is that, since the monomeric structure is short, the diameter of the cyclic tetramer would be small, meaning that the monomers conforming the cyclic assembly would be closer in space and that could generate unexpected intramolecular interactions in the cycle.

The first monomer González-Rodríguez group synthetised was GC0a, but they soon realized that it was not soluble and capable of forming gels in a variety of solvents. This behaviour was completely different from the ones observed for the larger monomers synthesized in the last example. This different behaviour seemed to indicate that the self-assembly into a cyclic tetramer did not take place, meaning that the anti-conformation for this monomer must be more stable for steric reasons or for intramolecular interactions between the lipophilic bulky groups.

In order to study the effect of these possible intramolecular interactions between the lipophilic groups and its steric effect, González-Rodríguez team proceeded to synthesize the other monomers seen in Figure 18. In these new monomers they used one or two long alkyl chains instead of the lipophilic groups, since these alkyl chains could release some steric hindrance as well as to not generate unexpected intramolecular interactions between each other.

On one hand, GC0b presented the same behaviour as GC0a forming gels suggesting high population on the anti-conformation, but on the other hand, GC0c and GC0d, even though they were not very soluble, they did not form gels on any solvent which meant that while GC0b self-assembled into the open structure GC0c and GC0d formed cyclic tetramers.

In order to compare the stability of the self-assembly of these new shorter monomers with the cyclic tetramer synthesized using monomer GC, González-Rodríguez group carried out an experiment where they added increasing amounts of DMSO into different solutions containing GC0a, GC0b, GC0c, GC0d and GC to see which fraction of this polar solvent could provoke the dissociation of the self-assembly.

As it was expected for the observations made before, the addition of DMSO into the solutions containing GC0a and GC0b resulted in large aggregates when the solvent was mainly non-polar while its dissociation was observed when DMSO concentration increased.

On the contrary, the results of this same experiment on the alkyl substituted monomers GC0c and GC0d concluded that a slow exchange was observed on an NMR scale when the fraction of DMSO increased. This observation meant that the cyclic assemblies formed by the self-assembly of these two monomers were kinetically stable.

Anyway, when comparing the amount of DMSO required for the dissociation between these shorter monomers and GC, this trend was observed: GC>>>GC0d>GC0c which indicates a clear lower cycle stability for the shorter monomers.

With the aim to understand the behaviour of these shorter monomers, González-Rodríguez group performed a series of calculations and they concluded that for GC0a, the anti-conformer is in fact more stable than the syn one required for cyclotetramerization. This difference on the stability is the reason why GC0a self-assembles into a linear oligomer. Regarding the other monomers, GC0b, GC0c, GC0d the results of these calculations were that the syn and the anti-conformation were equally stable.

González-Rodríguez group wanted to find out the reason of the difference in the stability between the syn and the anti-conformations on GC0a monomer. First, they studied if the steric hindrance was the reason of this difference, but they found out that this was not the problem.

The discard of this possible destabilisation factor for the syn conformation led the group to believe that the problem may be something else. Further inspection of GC0a and GC0b structures revealed the existence of the possibility for intramolecular interactions between the bulky lipophilic substituents and the donor atom in the cytosine (Figure 19).

**Intramolecular interacion for the anti-conformation for GC0a and GC0b**

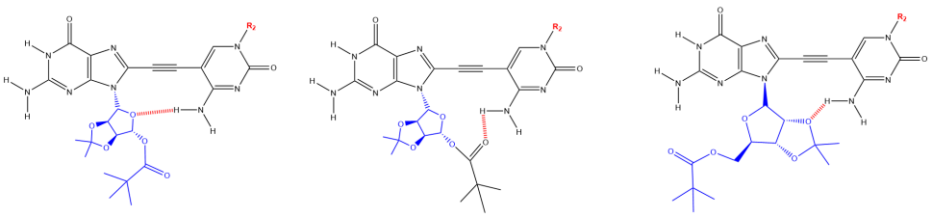


Figure 19. Stabilizing intramolecular interactions for the anti-conformation for GC0a and GC0b

In order for these intramolecular interactions to occur, the monomer must be arranged in the anti-conformation and, since the existence of these interactions results in a considerable hindrance on the rotation of the sigma bonds, the anti-conformer is predominant and therefore, these macrocycles self-assemble into a linear oligomer.

Having discovered the existence of these intramolecular interactions between the lipophilic substituents and the donor atom on the cytosine edge of the monomer makes it easier to understand why Gc0d self-assembles into the most stable cyclic tetramer from this batch of new monomers since it is substituted by two alkyl chains incapable of that kind of interaction.

Now, the question one may be asking is, why GC0c self-assembles into a cyclic tetramer while GC0b forms open structures if they both have a lipophilic group and an alkyl chain as substituents? The answer to that relies on an important characteristic of the hydrogen bonds, which is proximity. While there is a hydrogen available close in space to form the hydrogen bond for GC0b, that is not the case for GC0c and, therefore the anti-conformation does not have this extra stability concluding into the formation of the cyclic tetramer.

So, to recap, even though it is true that shortening the monomers length minimizes the sigma bonds in its structure which would mean that there would be fewer degrees of freedom present in the monomer and, therefore the entropic penalizations would decrease, it must also be considered that other problems like steric hindrance and intramolecular interactions between the monomers and the substituents may appear increasing the stability of the open structure over the closed one.

Reached this point, González-Rodríguez group thought that, maybe, the answer to this problem would be not the decrease of the number of sigma bonds in the structure but to use the steric hindrance on their favour to block the rotation of these bonds in the syn-conformation by the addition of substituents in the spacers.

With this intention, González-Rodríguez group synthetised two different monomers containing a biphenyl unit substituted either by four hydrogens (GC'-H) or by four methyl groups (GC'-Me) as a central connector (Figure 20). These substituents will create a steric hindrance around the central sigma bond of the monomer that will block at different extents its ability to rotate. [15]

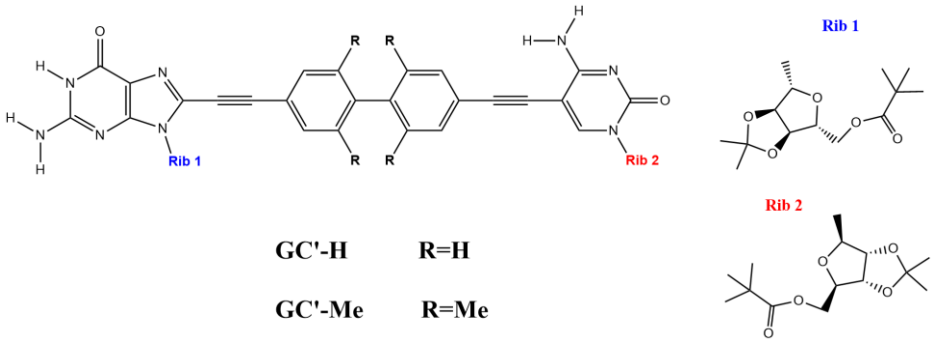



Figure 20. Monomeric structures with blocked rotation of the central σ-bond

In order to understand this example, it is important to highlight that, for the cyclic tetramer to be formed, both nucleobases must be in the same plane and, since biphenyl molecule cannot adopt a perfectly planar conjugation because of steric effects between both ortho substituted rings, some problems may arise in the cyclization. So, if the steric hindrance produced by these ortho substituted rings is increased, the internal π-conjugation of the system will not make it possible for the monomer to self-assemble into the cyclic tetramer.

Therefore, the study of the self-assembly of these two monomers will be useful to investigate the influence of blocked sigma bonds and internal π-conjugation on chelate cooperativity.

In order to study the stability and the nature of the self-assemblies, González-Rodríguez group monitored the aggregation of increasing fractions of DMSO by $^1$H NMR to see which fraction of the polar solvent is needed to disassociate the macrocycle. To understand how the rotation and the π-conjugation affects chelate cooperativity, they compared the self-assembly of these two monomers with GC and GC2 (Figure 16) containing four and six sigma bonds respectively.

GC'-H monomer showed an all or nothing behaviour plus a slow exchange in NMR time scale in its dissociation, which meant that the hydrogen substituted monomer formed kinetically and thermodynamic cyclic tetramers. They proceeded to compare the stability of cGC'-H$_4$ with cGC$_4$ and cGC2$_4$ and they stablished this trend: cGC$_4$> GC2$_4$> cGC'-H$_4$. The results of the same study in GC'-Me monomer were completely different since amongst other experimental proves, GC'-Me afforded viscous solutions characteristic of the polymerization.

To understand why the GC'-H$_4$ self-assembles into less stable cycles than the one seen before in this work and why GC'-Me forms open oligomeric architectures it is necessary to

recapitulate what has been said before about how internal π-conjugation affects the self-assembly.

In order to form the cyclic tetramer, it is required for the guanine and cytosine nucleobases to be in the same plane. This is a problem that was not needed to be considered on the other examples used in this work since all the monomers studied to this point contained central connectors that could adopt a perfectly planar conjugation, but that is not the case for GC'-H and GC'-Me monomers because of the biphenyl spacer. (Figure 21).

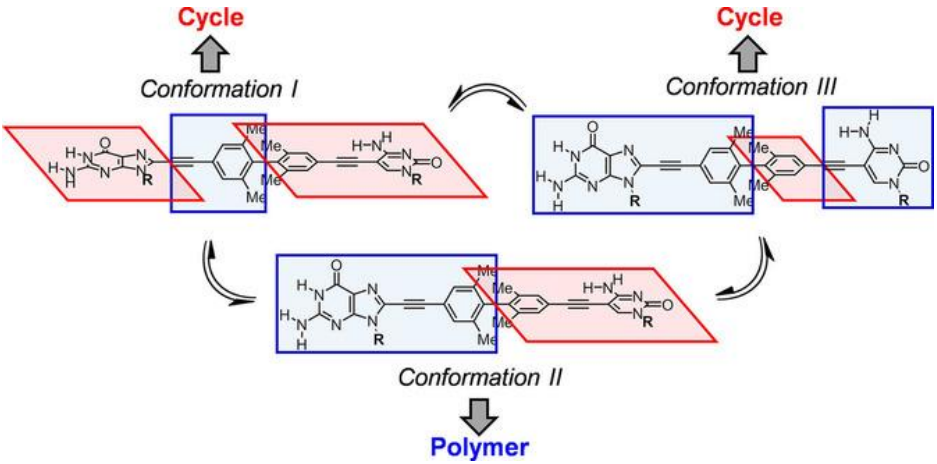



Figure 21. Different conformations due to the non-planar behaviour of the central connector
Image from González Rodríguez et al, ref 15

The figure above shows how the central sigma bond can rotate and form different possible π-conjugated conformations. The question now is, since both GC'-H and GC'-Me can be found in these conformations, why does the GC'H self-assemble into a cyclic tetramer while GC'-Me forms polymeric structures? The answer to this question relies on steric effects.

Both monomers present steric hindrance around the rotation of the central sigma block, but since GC'-Me is substituted by four methyl groups, its steric hindrance will be bigger than the hydrogen substituted one and, therefore the biphenyl group is obligated to orient both complementary bases in orthogonal planes (blue and red planes in the figure). Anyway, if the steric hindrance is lower, this biphenyl group may be able to place both complementary bases in the same plane, allowing the self-assembly into the cyclic tetramer.

Since experimental studies concluded that GC'-Me formed polymeric structures, this monomer must prefer to be found in conformation II forming the polymeric structure. To confirm

the hypothesis, González-Rodríguez team performed several calculations of the stability of each conformation for GC'-Me concluding that conformation II is indeed more stable than conformations I and III and, therefore justifying the self-assembly into the polymeric architecture.

No computational studies were made to see the energy of the different conformations for the GC'-H monomer, but, since steric hindrance is lower for this monomer it is logical to think that conformations I and III are more populated than for GC'-Me monomer, which led the hydrogen bonding molecule to be able to self-assemble into the cyclic structure.

The study of these two monomers has brought a new concern which is the importance of the planarity of the monomeric structure. As seen in this example, if the central connector used due to its $\pi$-conjugation cannot be found in the correct conformation because it cannot place the complementary bases on the same plane, the cyclic tetramer will not be the result of the self-assembly of the monomers.

# 7. CONCLUSIONS

All along this work, it has been proven many times how fragile the thermodynamic controlled equilibrium between the formation of a macrocycle and the open structure of a supramolecular system is. There have been many examples on how this equilibrium can be shifted to one side or the other for small details on monomeric structure, hydrogen bonding array, $\pi$-conjugation of the system, steric effects, or intramolecular unexpected interactions, amongst others.

After this study on cyclic tetramers, it can be concluded that there are two principal factors affecting **chelate cooperativity**. The first one is the **rigidity of the monomeric structure** and, the second one, the **hydrogen bonding array**. They both are key to shift the equilibrium into the macrocyclic form with high **EM** values.

Regarding the monomeric structure, the principal concern is the **number of degrees of freedom** since, the bigger this number is, less possible the self-assembly into the cyclic structure becomes. The number of degrees of freedom is directly proportional to the number of sigma bonds so, when **molecular preorganization** is taking place, a rigid $\pi$-conjugated system is preferred. It is also important to consider the **length of the molecule**. If the monomer is **too large**, the entropic penalization will shift the equilibrium into the open structure, but, if it is **too short**, other effects like **steric hindrance and intramolecular interactions** between substituents and monomers may jeopardize the formation of the cyclic tetramer. It is also important to consider the **$\pi$-conjugation** of the system since it can add difficulties to the formation of the cyclic assembly. Regarding this matter, **planarity is key,** so it is crucial to use a monomeric structure able to adopt a perfectly planar conformation placing both complementary bases on the same plane.

Concerning hydrogen bonding interactions, there are two main factors to consider. The first one must be the polarity of the media since there will be a competitive effect in the formation of hydrogen bonds, so generally, it is better to use **non-polar solvents.** The second one is related with the **symmetry of the hydrogen bonding pattern** since it will be key on obtaining high **chelate cooperativity**. There is a trend on Kass value, influenced by the symmetry of hydrogen bonding arrays, that has been stablished regarding secondary interactions **DDD-AAA>ADD-DAA>DAD-ADA**. It has also been demonstrated that the hydrogen bonding motif do not only has an influence on the overall strength of the system but also on the number of ways the monomers can self-assemble depending on its symmetry, affecting the overall stability of the self-assembly into the hydrogen bonded macrocycle.

# 8. REFERENCES AND NOTES

**1.** Mayoral, M. J., Bilbao, N., & González-Rodríguez, D. (2016). Hydrogen-Bonded Macrocyclic Supramolecular Systems in Solution and on Surfaces. *ChemistryOpen*, *5*(1), 10–32. https://doi.org/10.1002/open.201500171

**2.** Aparicio, F., Mayoral, M. J., Montoro-García, C., & González-Rodríguez, D. (2019). Guidelines for the assembly of hydrogen-bonded macrocycles. *Chemical Communications*, *55*(51), 7277–7299. https://doi.org/10.1039/c9cc03166a

**3.** Sivakova, S., & Rowan, S. J. (2005). Nucleobases as supramolecular motifs. *Chemical Society Reviews*, *34*(1), 9–21. https://doi.org/10.1039/b304608g

**4.** Jorgensen, W. L., & Pranata, J. (2008). Additional examples of imidcs and 1,6-diamidopyridines binding with K,'s of 125-150 M"1 are available: (a). In *Proc. Natl. Acad. Sci. O.S.A* (Vol. 112, Issue 2). https://pubs.acs.org/sharingguidelines

**5.** Schramm, U. (n.d.). *The Cooperativity Concept Gulliver constrained by a multitude of weak ªbondsº. Illustration by*.

**6.** Karas, L. J., Wu, C. H., Das, R., & Wu, J. I. C. (2020). Hydrogen bond design principles. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *10*(6). https://doi.org/10.1002/wcms.1477

**7.** Papmeyer, M., Vuilleumier, C. A., Pavan, G. M., Zhurov, K. O., & Severin, K. (2016). Molecularly Defined Nanostructures Based on a Novel AAA-DDD Triple Hydrogen-Bonding Motif. *Angewandte Chemie*, *128*(5), 1717–1721. https://doi.org/10.1002/ange.201510423

**8.** Montoro-García, C., Camacho-García, J., López-Pérez, A. M., Mayoral, M. J., Bilbao, N., & González-Rodríguez, D. (2016). Role of the Symmetry of Multipoint Hydrogen Bonding on Chelate Cooperativity in Supramolecular Macrocyclization Processes. *Angewandte Chemie*, *128*(1), 231–235. https://doi.org/10.1002/ange.201508854

**9.** Mascal, M., Hext, N. M., Warmuth, R., Arnall-Culliford, J. R., Moore, M. H., & Turkenburg, J. P. (1999). The G-C DNA base hybrid: Synthesis, self-organization and structural analysis. *Journal of Organic Chemistry*, *64*(23), 8479–8484. https://doi.org/10.1021/jo990719t

**10.** Asadi, A., Patrick, B. O., & Perrin, D. M. (2008). G∧C quartet - A DNA-inspired Janus-GC heterocycle: Synthesis, structural analysis, and self-organization. *Journal of the American Chemical Society*, *130*(39), 12860–12861. https://doi.org/10.1021/ja8047128

**11.** Asadi, A., Patrick, B. O., & Perrin, D. M. (2007). Janus-AT bases: Synthesis, self-assembly, and solid state structures. *Journal of Organic Chemistry*, *72*(2), 466–475. https://doi.org/10.1021/jo061840k

**12.** Montoro-García, C., Camacho-García, J., López-Pérez, A. M., Bilbao, N., Romero-Pérez, S., Mayoral, M. J., & González-Rodríguez, D. (2015). High-Fidelity Noncovalent Synthesis of Hydrogen-Bonded Macrocyclic Assemblies. *Angewandte Chemie*, *127*(23), 6884–6888. https://doi.org/10.1002/ange.201501321

**13.** Romero-Pérez, S., Camacho-García, J., Montoro-García, C., López-Pérez, A. M., Sanz, A., Mayoral, M. J., & González-Rodríguez, D. (2015). G-arylated hydrogen-bonded cyclic tetramer assemblies with remarkable thermodynamic and kinetic stability. *Organic Letters*, *17*(11), 2664–2667. https://doi.org/10.1021/acs.orglett.5b01042

**14.** Montoro-García, C., Mayoral, M. J., Chamorro, R., & González-Rodríguez, D. (2017). How Large Can We Build a Cyclic Assembly? Impact of Ring Size on Chelate Cooperativity in Noncovalent Macrocyclizations. *Angewandte Chemie*, *129*(49), 15855–15859. https://doi.org/10.1002/ange.201709563

**15.** Montoro-García, C., Bilbao, N., Tsagri, I. M., Zaccaria, F., Mayoral, M. J., Fonseca Guerra, C., & González-Rodríguez, D. (2018). Impact of Conformational Effects on the Ring–Chain Equilibrium of Hydrogen-Bonded Dinucleosides. *Chemistry - A European Journal*, *24*(46), 11983–11991. https://doi.org/10.1002/chem.201801704