GDAŃSK UNIVERSITY OF TECHNOLOGY

The author of the doctoral dissertation: Michał Kosno
Scientific discipline: Chemical Science

DOCTORAL DISSERTATION

Title of doctoral dissertation: Spectroscopic studies on physicochemical properties of selected unsymmetrical bisacridine derivatives and NMR analysis of their interactions with the model sequence Pu22 aided by molecular dynamics.

Title of doctoral dissertation (in Polish): Spektroskopowe badania fizykochemiczne wybranych niesymetrycznych pochodnych bisakrydyn oraz analiza NMR ich oddziaływań z sekwencją modelową Pu22 wspomagana metodami dynamiki molekularnej.

Supervisor *signature*	Second supervisor *signature*
Professor Zofia Mazerska, PhD, DSc, Eng.	<Title, degree, first name and surname>
Auxiliary supervisor *signature*	Cosupervisor *signature*
Tomasz Laskowski, PhD, Eng.	<Title, degree, first name and surname>

Gdańsk, year 2023

Dedication

I dedicate this dissertation to my fiancée Marta and to my parents.

Acknowledgements

I want to thank my supervisor, Professor Zofia Mazerska, for her time, understanding and support.

I would like to express my deepest appreciation to my auxiliary supervisor Dr. Tomasz Laskowski and to Professor Jan Mazerski for their help, scientific support, valuable advice and knowledge they shared with me.

I am also sincerely grateful to Dr. Julia Borzyszkowska-Bukowska for her invaluable help and maintaininga pleasant atmosphere at work.

Abstract

Compounds derived from acridine have been studied for many years at the Department of Pharmaceutical Technology and Biochemistry of the Gdansk University of Technology. In recent years, new promising derivatives have appeared, belonging to the unsymmetrical bisacridines (UAs) family with high anticancer activity. Both their physicochemical properties and their mechanism of action at the molecular level have not been thoroughly analyzed so far.

Four derivatives were selected for the study, termed as: C-2028, C-2041, C-2045 and C-2053. The first aim of this work was to determine the protonation forms in which the studied compounds exist at different pH. The second aim was the analysis of aggregation process at different conditions, which is important from the point of view of further in vitro and in vivo studies. Methods such as UV-Vis and NMR spectroscopy were used for this purpose. The obtained spectral data were subjected to a thorough chemometric analysis, using techniques such as principal components analysis, multiple regression and numerical optimization.

Previous studies have shown that a potential molecular target for unsymmetrical bisacridines may be the promoter sequence of the C-Myc proto-oncogene, known as Pu22. Therefore, the next goal of the presented work was to determine the interactions of the studied compounds with the above mentioned sequence using advanced NMR spectroscopy and molecular dynamics (MD) methods. These studies were narrowed down to the two most promising derivatives: C-2045 and C-2053. In order to determine the exact molecular structure of the obtained complexes, a series of one- and two-dimensional NMR spectra of both investigated compounds and their monomeric derivatives (C-1311 and C-1748) were recorded, which allowed to build the initial structures of the studied complexes. In the next step, they were subjected to an equilibrium MD simulation. In the final stage of the research, the nature and strength of the interactions between the studied compounds and the selected G-quadruplex were estimated using the umbrella sampling method, which allowed the determination of the free energy profiles of the obtained complexes.

The results obtained allowed to determine the way in which the unsymmetrical bisacridines interact with the C-Myc promoter sequence - their potential molecular

target. In addition, their basic physicochemical properties have been determined, which will allow further research on the compounds described above. In this work, a simple and reliable method for the determination of physicochemical constants has been developed and presented, which yields well credible results, even in the case of compounds with a complex structure having many sites capable of protonation.

Streszczenie

W Katedrze Technologii Leków i Biochemii Politechniki Gdańskiej od lat bada się związki będące pochodnymi akrydyny. W ostatnich latach pojawiły się nowe, obiecujące pochodne należące do rodziny niesymetrycznych bisakrydyn (UAs), wykazujące wysoką aktywność przeciwnowotworową. Zarówno ich właściwości fizykochemiczne jak i mechanizm działania na poziomie molekularnym nie zostały wcześniej wnikliwie przeanalizowane.

Do badań zostały wyselekcjonowane cztery pochodne oznaczone jako: C-2028, C-2041, C-2045 oraz C-2053. Niniejsza praca miała na celu określenie form protonowania w jakich istnieją badane związki w różnych wartościach pH. Drugim celem była analiza procesu agregacji w różnych warunkach – istotnych z punktu widzenia dalszych badań *in-vitro* oraz *in-vivo*. W tym celu posłużono się metodami spektroskopii UV-Vis oraz NMR. Otrzymane dane spektralne zostały poddane wnikliwej analizie chemometrycznej, która korzystała m.in. z analizy głównych składowych, regresji wielokrotnej oraz optymalizacji numerycznej.

Wcześniej badania wykazały, że potencjalnym celem molekularnym niesymetrycznych bisakrydyn może być sekwencja promotorowa protoonkogenu C-Myc, znana jako Pu22. Zatem, kolejnym celem pracy było określenie sposobu oddziaływania badanych związków z wyżej wspomnianą sekwencją stosując zaawansowane metody spektroskopii NMR oraz metody dynamiki molekularnej (MD). Badania zostały zawężone do dwóch najbardziej obiecujących pochodnych: C-2045 i C-2053. W celu ustalenia dokładnej struktury molekularnej otrzymanego kompleksu zarejestrowano szereg jedno i dwuwymiarowych widm NMR, zarówno badanych związków, jak i ich monomerycznych pochodnych (C-1311 i C-1748), co pozwoliło na zbudowanie wstępnych struktur badanych kompleksów. W następnym kroku zostały one poddane równowagowej symulacji metodami MD. W końcowej fazie badań natura i siła oddziaływań pomiędzy badanymi związkami a wybranym G-kwadrupleksem została oszacowana metodą próbkowania parasolowego. Pozwoliło to na wyznaczenie profili energii swobodnej otrzymanych kompleksów.

Otrzymane wyniki wskazały sposób, w jaki niesymetryczne bisakrydyny oddziałują z sekwencją C-Myc, która stanowi ich potencjalny cel molekularny. Ponadto, zostały

określone ich podstawowe właściwości fizykochemiczne, które umożliwiają dalsze badania opisanych powyżej związków. W niniejszej pracy została także opracowana oraz zaprezentowana prosta i wiarygodna metoda wyznaczania stałych fizykochemicznych, która sprawdza się nawet w przypadku związków o złożonej strukturze, posiadających wiele miejsc zdolnych do protonowania.

CONTENTS

LIST OF ABBREVIATIONS

ACQ	–	data acquisition period
COSY	–	COrrelation SpectroscopY
DNA	–	DeoxyriboNucleic Acid
G4	–	G-quadruplex
HMBC	–	Heteronuclear Multiple Bond Coherence
HSQC	–	Heteronuclear Single Quantum Correlation
MD	–	Molecular Dynamics
NMR	–	Nuclear Magnetic Resonance
NOE	–	Nuclear Overhauser Enhancement
NOESY	–	Nuclear Overhauser Effect SpectroscopY
PCA	–	Principal Component Analysis
pKa	–	Proton dissociation Constant
ppm	–	Parts Per Million
RD	–	Relaxation Delay
ROESY	–	Rotating-frame nuclear Overhauser Effect SpectroscopY

SVD	–	Singular Value Decomposition
T	–	Delay
TOCSY	–	Total Correlation SpectroscopY
UAs	–	Unsymmetrical Bisacridines
US	–	Umbrella Sampling
UV-Vis	–	UltraViolet-Visible spectroscopy
WHAM	–	Weighted Histogram Analysis Method
ξ	–	The distance between the centers of mass determined for the upper and lower G4 tetrads and the imidazoacridinone rings of the upper and lower ligand, respectively

1 THEORETICAL PART

1.1 INTRODUCTION

Acridines have been widely investigated for many years. Some of their analogues (1-nitroacridine derivatives) exhibit high antitumor activity. Their heteroaromatic ring systems are not planar and possess peculiar electron density distribution. One of the most known compounds belonging to this family is Ledakrin, which is the first officially registered Polish anticancer drug. Ledakrin was developed in the Department of Pharmaceutical Technology and Biochemistry at Gdańsk University of Technology[1]. This drug is based on a 1-nitroacridine ring, with an aminoalkyl chain in the 9 position and is characterized by high cytostatic effects on e.g. HeLa, Walker 256 or NK lymphocytic leukemia LI210. Unfortunately, it is no longer recommended for usage because of the highly cardiotoxic effect. Other group of compounds, extensively investigated at the Gdańsk University of Technology, are acridinones. Compounds belonging to this particular family are characterized by a planar, at least three-membered aromatic ring system. Interestingly, the maximal deviation from planarity is not higher than 0.02 Å[2]. The most successful acridinone-based compound is Symadex (C-1311), which possesses an additional five-membered imidazole ring and an aminoalkyl chain (Figure 1). It exhibits high antitumor activity against many different tumor cell lines, including several breast and colon cancer[3]. This agent was not officially registered, but it is a drug which has been recommended for second phase clinical trials.

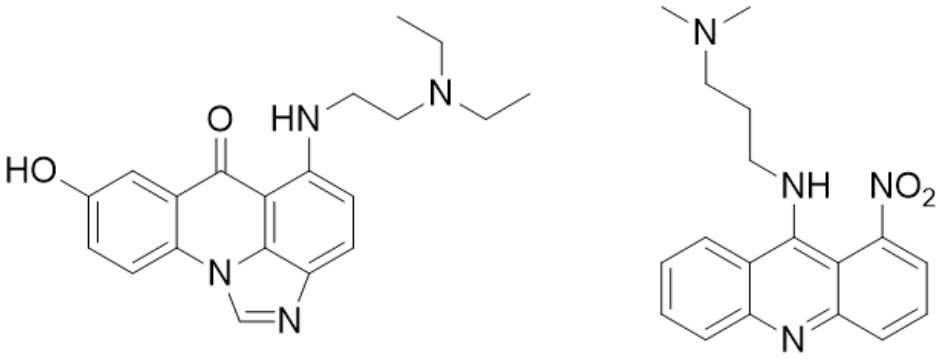


Fig.1. Symadex (left)[3], Ledakrin (right)[1].

Both of them, Symadex and Ledakrin, mainly act via interacting with DNA, because of the presence of a highly conjugated aromatic system and a quite long aminoalkyl chain. Ledakrin requires prior metabolic activation, which includes reduction of the nitro group. One of the intermediate products of this process is reactive enough to form a covalent link between the different strands of DNA[4]. C-1311 has a different mechanism of action. It is a compound based on a planar four-membered flat aromatic ring. Flat fragment of C-1311 inserts between two bases of DNA. The aminoalkyl chain is arranged in the minor groove[5].

Success of those both compounds prompted researchers to further investigations, which resulted in development of a new class of compounds, named unsymmetrical bisacridines (UAs). At first glance, it seems to be a straightforward combination of the two compounds. However, it quickly turned out that such compounds are characterized by a number of unique properties. The first studies indicated a tremendous biological activity, which is much higher than it could be expected. Preliminary studies also showed a lack of significant interactions with helical forms of DNA. Moreover, there are some reports indicating that UAs are poorly metabolized. All of those may suggest a fundamentally different mechanism of action. It has been hypothesized that UAs may act by stabilizing G-quadruplexes. However, due to the large variability in the G-quadruplex structure, it is difficult to determine the type of binding. Indeed, preliminary studies on the stabilization of guanine tetrads in DNA, which were based on the determination of the melting points, initially supported this hypothesis. Despite the fact that UAs are completely different from other compounds, some of their properties could have been predicted by the chemical structure. Taking into account all the previous research, it can be expected that the imidazoacridinone ring, which is exceptionally flat, could act as an intercalator[5,6], while the aminoalkyl chain with the entire 1-nitroacridine ring could be considered as a highly flexible side chain. Furthermore, the nitro group could be suspected of DNA cross-linking after prior metabolic activation. Considering metabolism, substituents of both rings may also play a role. Previous research proves that the hydroxyl group at 8 position of the imidazoacridinone ring is prone to metabolism by conjugation with glucuronic acid[7], whereas the methyl group at 4 position of the 1-nitroacridine affects susceptibility to conjugation with glutathione (metabolic activation or inactivation)[8,9].

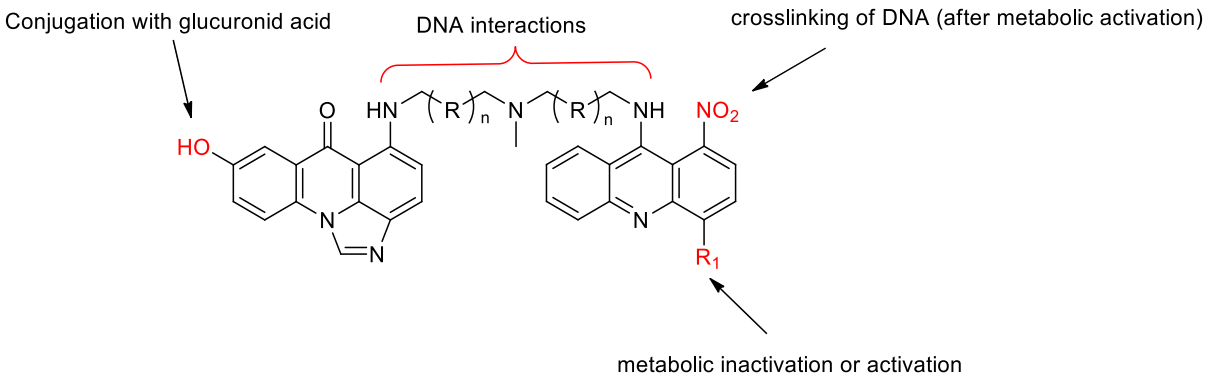


Fig. 2. Proposed structure-activity relationship (SAR) study of UAs.

Unsymmetrical bisacridines were not widely investigated so far, especially from the physicochemical point of view, so I decided to tackle this problem and try to reveal their secrets, which were shrouded in mystery so far.

1.2 Noncovalent ligand/DNA interactions

Interactions between a ligand and a DNA sequence are often ambiguous and complex in nature. However, they are usually a combination of many different types of interactions, that are much simpler. The following types of interactions are worth mentioning:

- electrostatic interactions[10],
- groove binding[11],
- classical intercalation[12],
- threading intercalation[13],
- nonclassical intercalation[14].

Electrostatic interactions

DNA requires the presence of positively charged metal ions to maintain stability in aqueous solution. This is known as counterion condensation[15]. The mechanism of ligand binding to the DNA molecule associated with counterion condensation is that the positively charged ligand atom neutralizes the negative charge carried by the phosphate residues, resulting in the release of counterions into the solution. Such release causes changes in the entropy of the system, which is the main physicochemical driving force of this process. Compounds, that interact with DNA

by intercalation, often have a positively charged atom built into their structure to facilitate approach to the DNA fragment[16].

Groove binding

Groove binding is one of the most commonly observed types of ligand/DNA interactions. This interaction is possible due to the fact that there is a negative charge in the DNA grooves. If the ligand is a sufficiently long and flexible molecule with positively charged fragments, it is able to geometrically adapt to the shape of the DNA and bind within the groove[17]. Examples of compounds interacting with DNA grooves are presented in Figure 3[18-23].

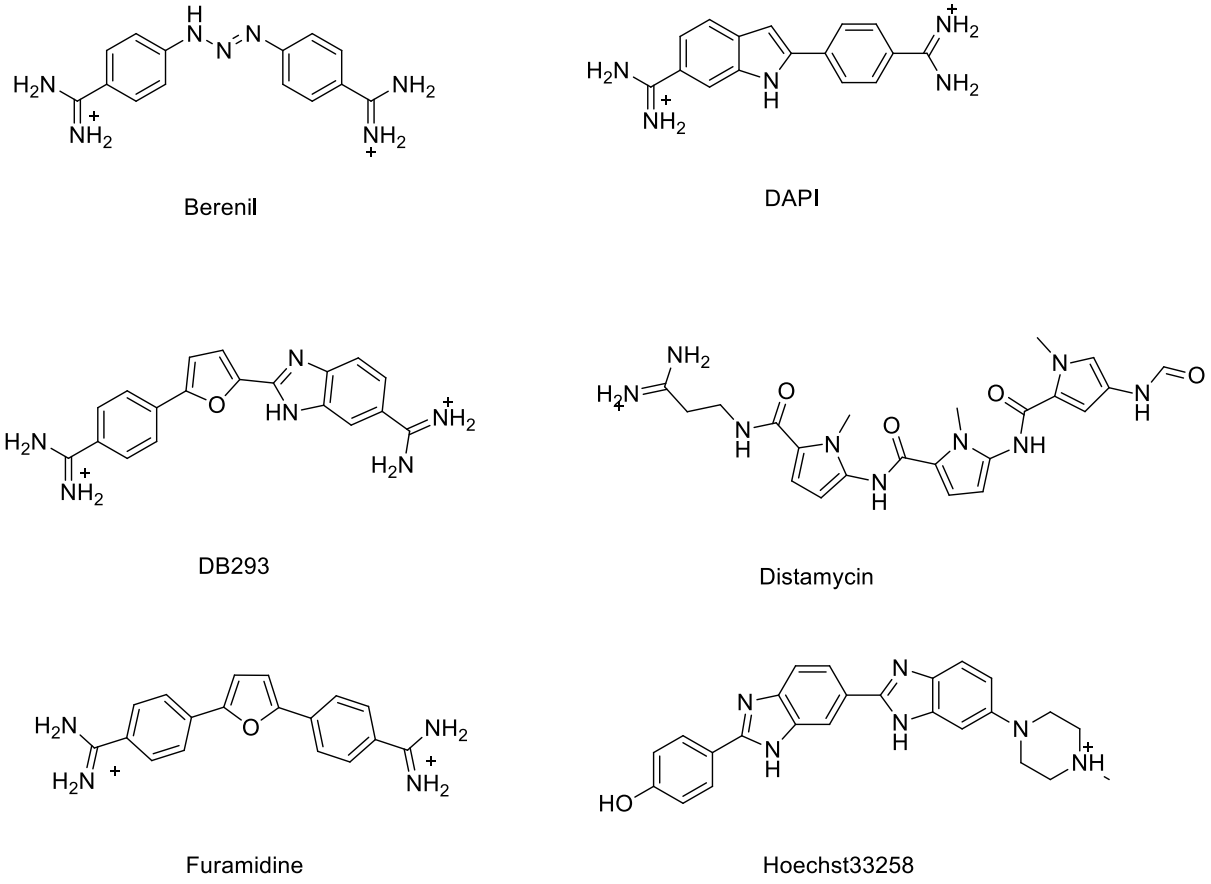


Fig. 3. Structures of compounds binding to minor groove[18-19].

Classical intercalation

Intercalation is the process of reversible binding of a ligand to DNA by inserting a flat fragment of a molecule between DNA base pairs. In this type of interaction, the

hydrogen bonds between DNA bases are not broken or weakened. In addition, the presence of an additional flat element in the DNA causes the base pairs to move slightly apart and therefore the potential adjacent intercalation site is blocked. This is known as the nearest neighbor exclusion principle. A typical intercalating compound contains in its structure a fused aromatic ring, with a positive charge located on one of the ring atoms or in the side chain, if present. Such a charge allows the ligand to approach the DNA molecule through electrostatic interactions. The phenomenon of counterionic condensation is used here. Thus, in the first stage, the driving force of this process are electrostatic interactions and favorable for complex formation changes in entropy. In the second step, hydrophobic interactions within DNA base pairs play the key role[16,20]. Examples of intercalators are presented in Figure 4[6,25-29].

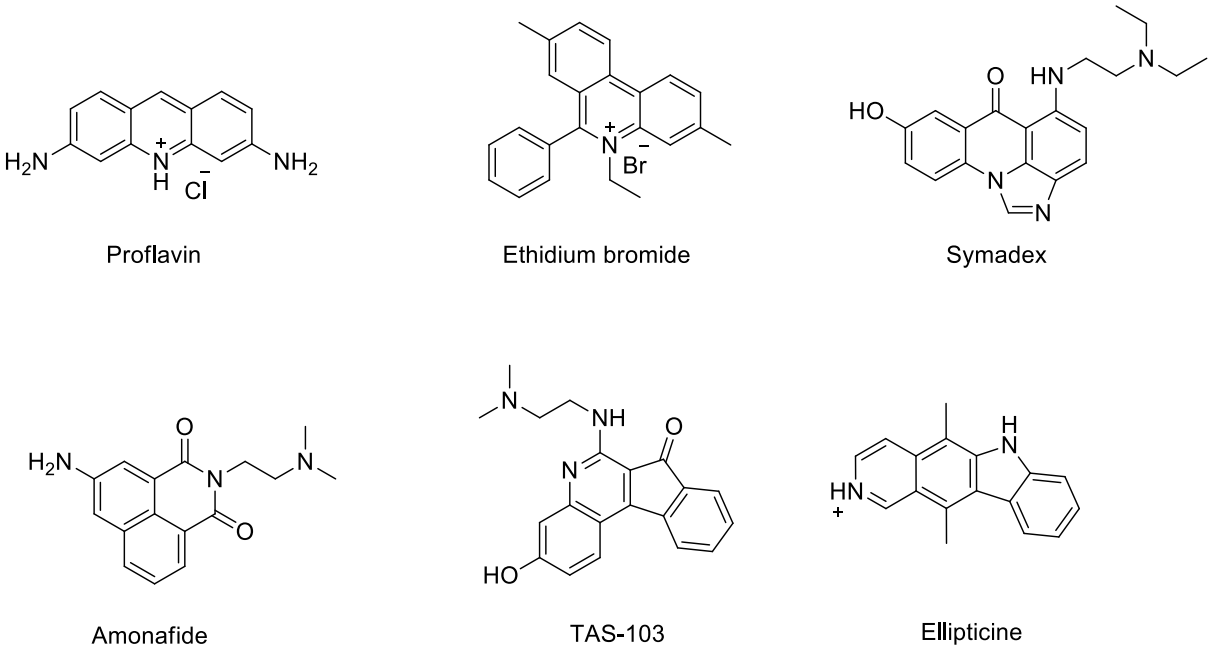


Fig. 4. Examples of intercalating agents[21,22,6,23,24,25].

Intercalation complexes are particularly interesting due to their complex nature. Their stability may be due to the interactions between the frontier molecular orbitals, especially between the lowest occupied ligand orbital (LUMO) and the highest occupied orbital of the nearest purine base (HOMO)[26].

Threading intercalation

Another interesting method of intercalation is the threading intercalation. A ligand must meet a number of structural requirements in order to interact with DNA in this way. Such a compound must consist of two relatively large subunits, connected by a linker. The kinetics of this process is extremely slow, due to the fact that the ligand must "squeeze" through the DNA base pairs, causing them to be severely deformed[27]. The main idea behind threading intercalators is presented in Figure 5[28].

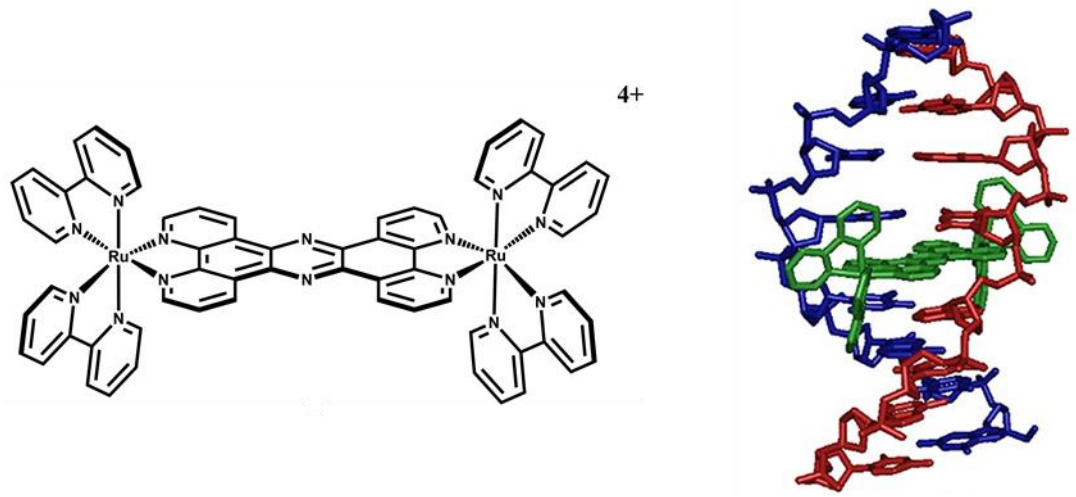
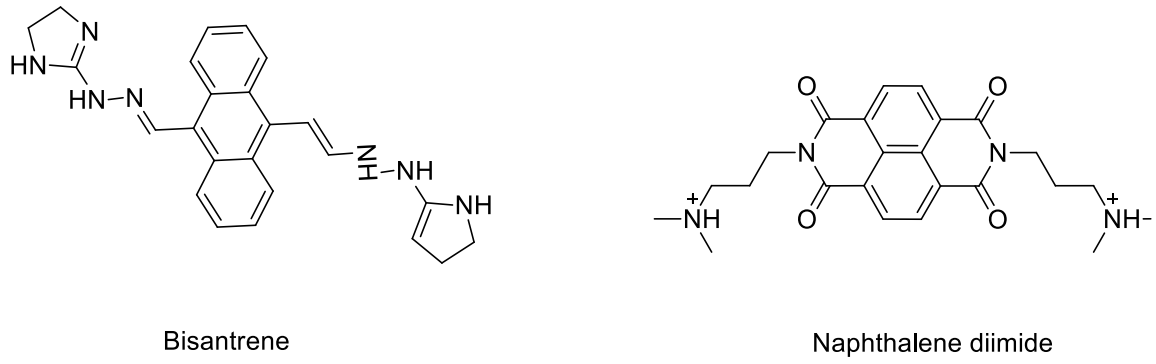


Fig. 5. Polypyridylruthenium coordination complexes (left) and its intercalation complex with DNA[28].

Examples of another threading intercalators are presented below in Figure 6[33,34].



Bisantrene Naphthalene diimide

Fig. 6. Examples of threading intercalators[29,30].

Compounds that interact with DNA in more complex ways, that cannot be described by a single binding mode, are called non-classical intercalators. They often combine many simpler interactions with DNA, such as classical intercalation or minor groove interactions. The figure below (Figure 7) shows some of these compounds[35,36].

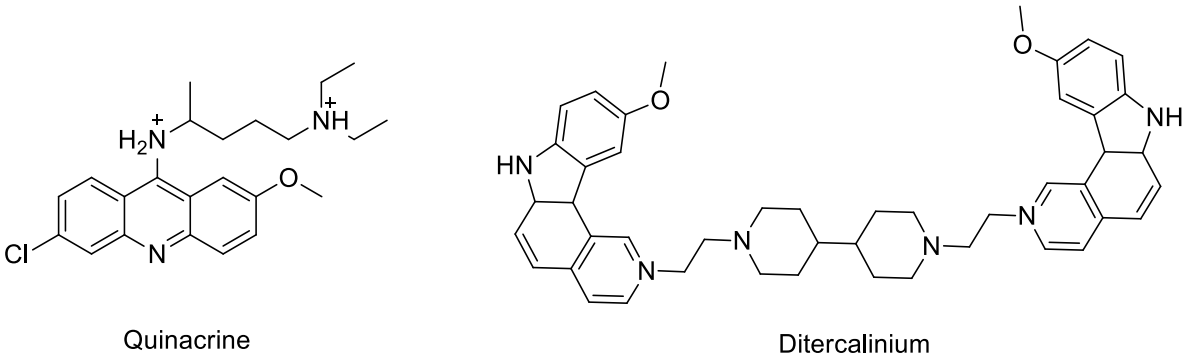


Fig. 7. Examples of non-classical intercalators[31,32].

1.3 G-quadruplexes

1.3.1 General information

G-quadruplexes have been known for over a century, but have not been extensively studied earlier[33]. The interest suddenly increased, when such structures were discovered in a human genome. Something, that was initially considered as a structural curiosity, turned out to be crucial for the life itself[34,35]. DNA G-quadruplexes are composed of guanidine rich oligonucleotides, assembled into guanine tetrads. Such structures usually form four-column topological structures, which are held together by Hoogsteen hydrogen bonding. There are over a dozen possible structures of DNA G-quadruplexes, which turn them into useful biological tools. These structures were found in many various genome regions, such as promotor regions (e.g. C-Myc, K-RAS) or telomeric sequences. They are involved in many critical cellular processes, including replication, translation, gene transcription and genomic stability in general[36]. However, the role of G-quadrupelexes is not fully elucidated. The best known structures of this type are telomeric G-quadruplexes. The presence of G-quadruplexes inhibits the activity of telomerase, which is responsible for maintaining a constant length of telomeres. G-quadruplexes consist of tetrads, which are built of four guanosines. These structures are stabilized by cations – and the stabilization is stronger while the cation better fits into the space between the tetrads. Most often they are sodium and potassium ions, sometimes also ammonium ions. There are eight hydrogen bonds in each G-tetrad, and each guanosine residue is both a donor and an acceptor of two hydrogen bonds (Figure 8).

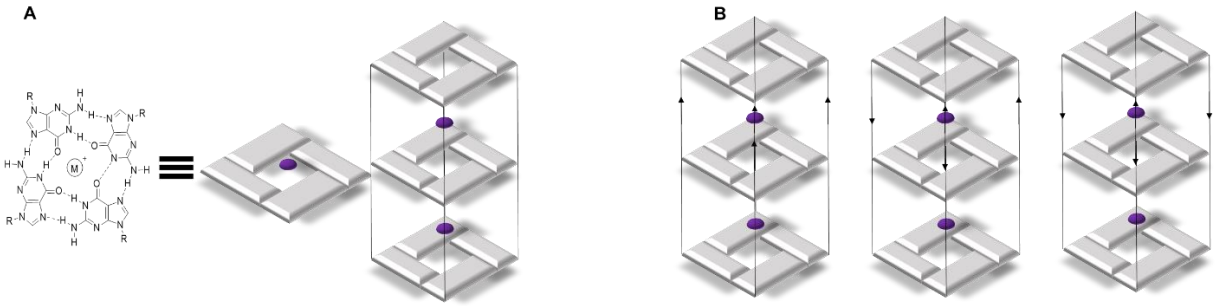


Fig. 8. Chemical structure of G-tetrad stabilized by Hoogsteen baseparing and central cation and stacking interaction (A). A variety of structures due to the direction of the strands (B).

Due to their secondary structure, G-quadruplexes can be divided into two key elements: the core and the loops. The core is formed of tetrads, which are held together by hydrogen bonds, π-π interactions, and backbone strands. The loops connect the backbone strands with each other[37]. The four corners of each of the tetrads form columns, that are particularly important for the determination of the structure of a G-quadruplex molecule using nuclear magnetic resonance spectroscopy. According to the length, base-sequence and configuration of the loops, G-quadruplexes exhibit various topologies, which determine their stability and availability for several types of ligands.

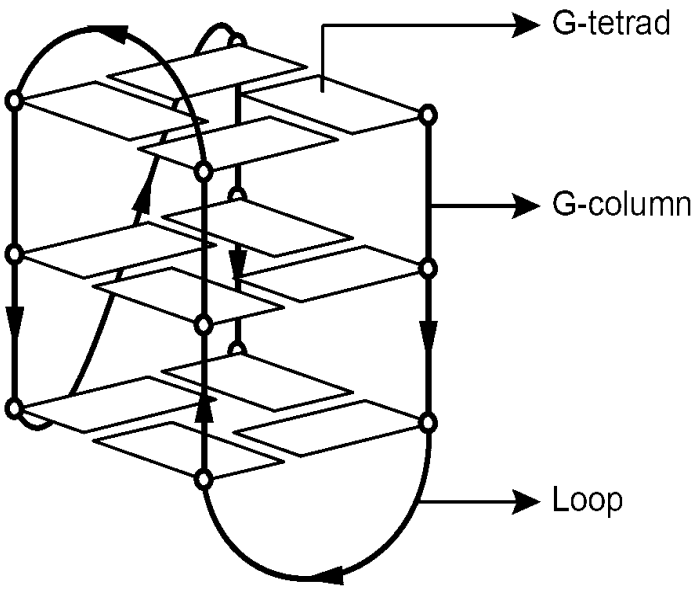


Fig. 9. Key elements of G-quadruplex: G-tetrad, G-column and Loop[37].

1.3.2 Interactions with G-quadruplexes

Due to their unique structure, G-quadruplexes interact with ligands slightly different than helical DNA fragments. Ligands can bind to the four grooves of a G-quadruplexes, or interact with the lower or upper tetrad. The presence of tetrads and numerous loops makes intercalation interactions rather unlikely. The most possible binding modes are presented in Figure 10[42].

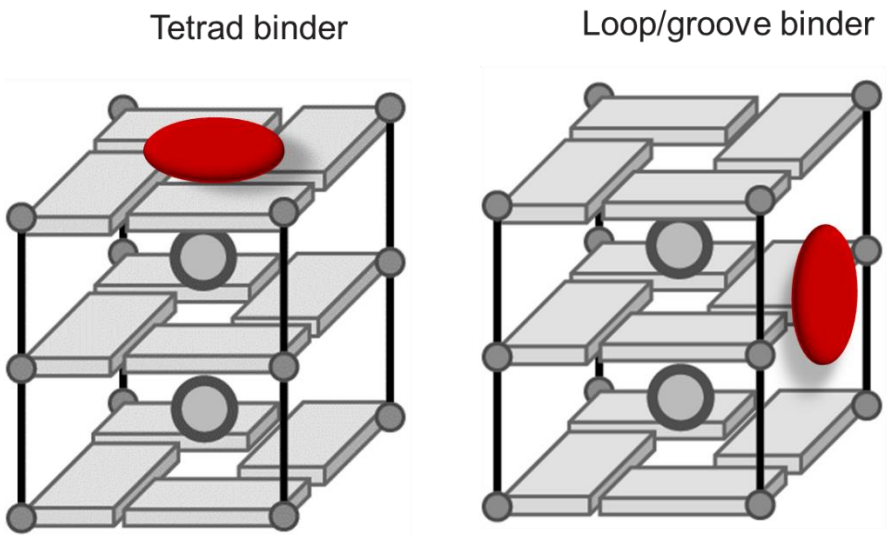


Fig. 10. G-quadruplex binding modes of ligands[38].

Compounds that interact in this way are depicted below (Figure 11).

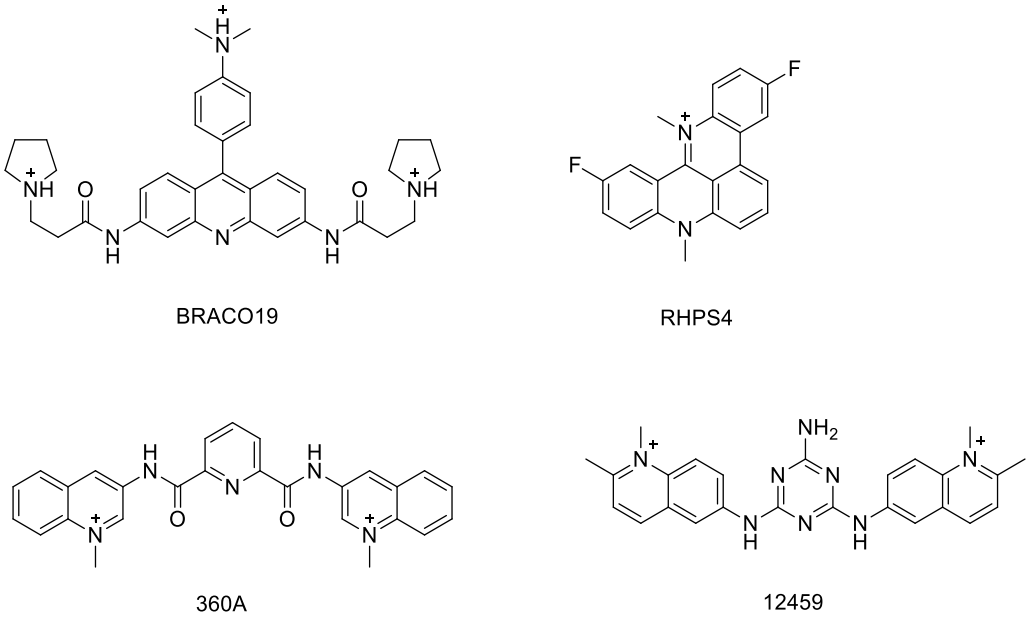


Fig. 11. Examples of compounds interacting with G-quadruplexes[38].

1.3.3 Promoter of the C-Myc proto-oncogene G-quadruplex as a molecular target for potential anti-cancer drugs

Myc is an important family of genes involved in transcription process and is responsible for regulation of many physiological processes, like protein synthesis, cell cycle control, cell adhesion or apoptosis[39]. C-Myc is an important part of Myc family, that encodes the c-MYC transcription factor, which was first time identified as the cellular homolog of the viral oncogene (c-myc) of the avian myelocytomatosis retrovirus[40]. According to current knowledge, a large number of cancers may be caused by overexpression of C-Myc[41]. Promotor region of C-Myc can fold into G-quadruplex structure, hence these particular structures are often taken into consideration as a molecular target for small ligands. Examples of the compounds that bind to the promoter region of the C-Myc proto-oncogene are:

- quinolones[42],
- quindolines[43],
- qlkaloids[44],
- metal complexes[45].

Quinolones

As an example of quinolones, CX-3543 molecule (Figure 12) can be presented[42]. This agent is derived from a fluoroquinolone and targets parallel-stranded G-quadruplexes only, i.e., those containing parallel G-columns. CX-3543 also selectively binds non-covalently to the c-Myc fragment. It is currently in the second phase of clinical trials with a proven antineoplastic effect and works by inhibiting the synthesis of ribosomes, which lead to cell apoptosis[46].

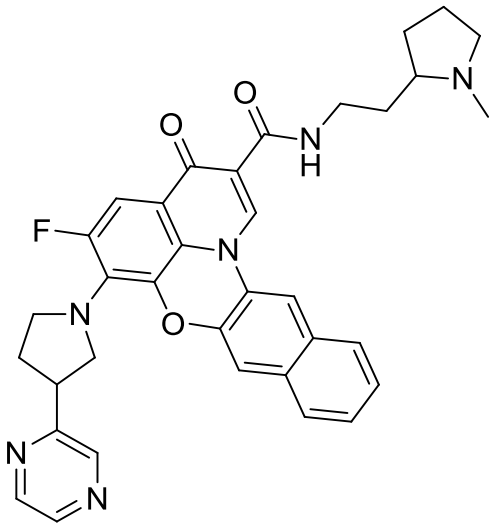


Fig. 12. Chemical structure of CX-3543 (quarfloxin)[46].

Quindolines

A good example of a quindoline derivative is the ligand described by Dai et al.[47], which was obtained as a synthetic derivative of a naturally occurring compound called cryptolepine. This derivative is interesting because it inhibits c-MYC expression in H2p G2 hepatocellular carcinoma and forms a well-defined complex with a G-quadruplex located in the c-Myc promoter region, named mycPu22. The structure formed in this way resembles a sandwich, because the two ligand molecules arrange themselves respectively below and above the lower and upper planes. The three-dimensional structure of the complex was resolved using advanced NMR techniques, supported by molecular modelling methods (Figure 13)[51].

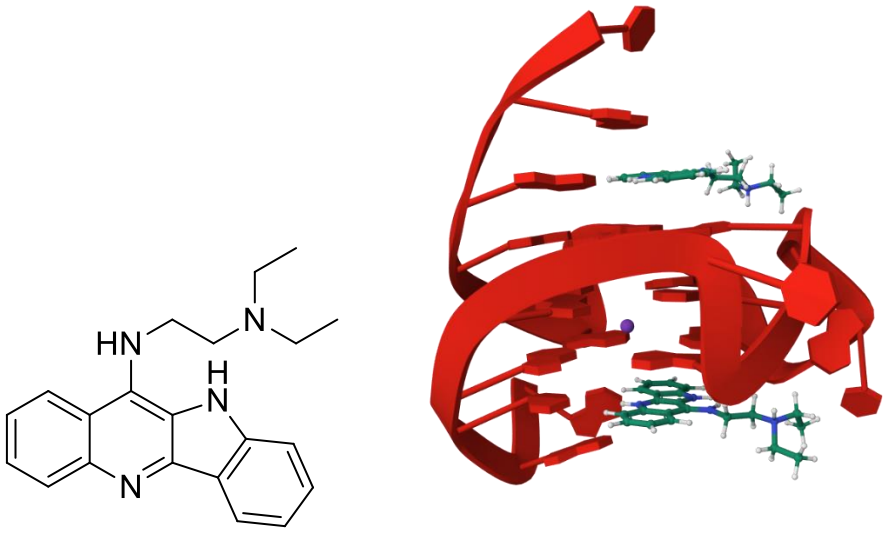


Fig. 13. Quindoline derivatives (left), DNA-ligand complex formed in ration 2:1 (right). DNA is marked in red, ligand is shown as a stick representation[47].

Other quindoline derivatives also interact with c-Myc G-quadruplexes, which was proven by Ou and colleagues[48]. The side chain plays the most important role in the nature of these interactions, as its arrangement determines binding affinity and selectivity. As an example, a compound with the name SYUIQ-05 shows a greater preference for binding to the c-Myc G-quadruplex than to the telomeric one (Figure 14)[52].

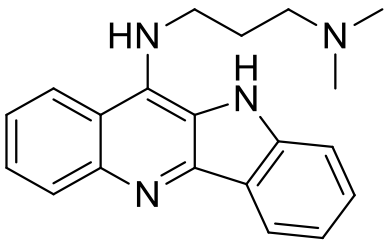


Fig.14. Chemical structure of SYUIQ-05[52].

Alkaloids

Alkaloids have long been used in traditional medicine, and some have also been shown to be ligands that may interact with G-quadruplexes. Alkaloids interacting with DNA have been described by Ji[49]. In this paper, three of the eleven compounds studied, named San, Pal, and Beb, respectively, were shown to increase the stability of G-quadruplexes (Figure 15)[53].

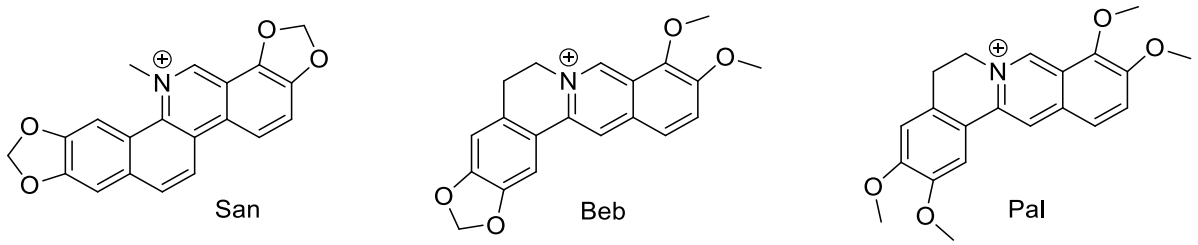


Fig. 15. Alkaloids with the highest G-quadruplex binding ability[49].

These compounds interact with G-quadruplex DNA mainly through intercalation. The alkaloids shown above were found to be the most potent because they have a structure of multi-membered flat rings with a centrally located positively charged nitrogen atom. An interesting property of San, Pal and Beb is their ability to induce the formation of G-quadruplexes (Hum24 sequence). They also interact with DNA G4 in ways other than intercalation. They have the ability to bind in the major groove as well as at the ends of the sequence[49].

Metal complexes

Compounds containing positively charged metal ions tend to interact very well with DNA in general. This is supported by the fact that DNA side chains are negatively charged, so electrostatic interactions play a major role[50]. There are many compounds containing metal ions build into their structures. The most common are platinum atoms (cisplatin and its derivatives)[51,52]. However, there are also compounds containing titanium, ruthenium, or gallium atoms[53]. Examples of these kind of agents are presented in Figure 16[57,58].

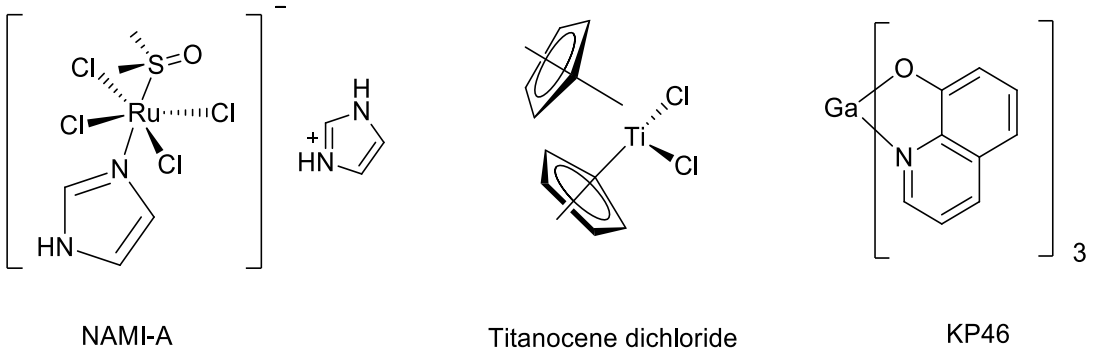


Fig. 16. Examples of compounds containing a metal ion in their structure with potential anticancer activity[53,54].

Such compounds are rarely used, mainly because of their poor water solubility and stability. Their interaction with DNA is also often difficult to determine. Titanocene dichloride is suspected to interact with DNA phosphate residues and with one of the nitrogen atoms of DNA bases[53]. Gallium(III) compounds most likely have an antiproliferative effect by inhibiting the ribonucleotide reductase[55].

1.4 Nuclear Magnetic Resonance (NMR)

NMR spectroscopy is indeed an absorption spectroscopy. It means that the obtained spectrum is a plot of the amount of quantum of electromagnetic radiation energy, absorbed by a sample, as a function of the frequency of this radiation. In NMR spectroscopy, the coupling phenomenon is essential. There are two basic types of coupling: spin-spin coupling and dipolar coupling. The latter is of particular interest. It allows to reveal the spatial proximity of two or more protons. This spectroscopy is a good tool for studying physicochemical properties, such as acid-base equilibrium[56], self-association[57] or resolving DNA and protein structures. Exchangeable protons with fast or slow exchange ratio can also be "visible" during NMR experiments. The exchangeable proton signals may be easily found in the NOESY spectrum, due to the presence of the NOE signals from the proton of the studied compound to the water resonance. Similarly, NOESY can be used to find correlation signals between two different molecules[58]. This kind of spectroscopy is widely employed for study of DNA/ligand interactions. There are many other techniques, which can be applied for such an analysis. A brief overview of the types of NMR spectroscopy is provided below.

1.4.1 COSY (correlation spectroscopy) and TOCSY (total correlation spectroscopy)

Correlation spectroscopy is the simplest 2D NMR spectroscopy method, which is able to reveal mainly proton – proton correlations. In most cases, geminal and vicinal couplings are visible. Long range couplings are rare. The COSY experiment is relatively easy to perform (straightforward pulse sequence given below[59]) and interpretation of the results is not very difficult[60].

$$-RD-\ 90°-t_1-90°-\ ACQ(t_2)- \qquad [1]$$

RD – Relaxation Delay,

t – Delay,

ACQ – Data acquisition period.

To simplify the spectrum and to obtain the cross peaks in pure absorption mode, an enhanced COSY called Double Quantum Filtered COSY (DQF-COSY) is used. This experiment is performed with the following pulse sequence[59]:

$$-RD-90°-t_1-90°(\emptyset 1)-\Delta-90°(\emptyset 2)-ACQ(t_2)- \qquad [2]$$

TOCSY is a useful modification of this experiment, which allows to reveal a chain of couplings over a specific spin system. The advantage of TOSCY spectrum is a great number of recorded correlation signals, which in many cases makes the assignment of resonances to the corresponding protons easier[59,61]. TOCSY is performed with pulse sequence[59]:

$$-RD-\ 90°-t_1-[spin\ lock]-\ ACQ(t_2)- \qquad [3]$$

1.4.2 NOESY - nuclear Overhauser effect spectroscopy

NOESY is a powerful tool to study three-dimensional structure of small chemical compounds, as well as biomacromolecules. Nuclear Overhauser effect (NOE) is an interaction of two spins through space, which derives from cross-relaxation phenomenon. With this method, it is possible to find protons which are close to each other in space. Such a feature allows to establish three-dimensional structures, both of small and relatively large chemical entities. Particularly useful is the fact, that there is strong and precisely described correlation between NOE signal intensity and the distance between two nuclei[62]:

$$a = k\sum_{ij} r_{ij}^{-6} \qquad (1)$$

a – intensity of a NOE signal,

k – constant (correlated to the reference distance),

r_{ij} – distance between two protons.

This experiment is carried out by pulse sequence[59]:

$$-RD-\ 90°-t_1-90°-\ t_m-90°-ACQ(t_2)- \qquad [4]$$

NOESY also can be used to find exchangeable protons in the structure. Besides that the exchangeable proton signals are often broadened and have an ambiguous integration, a NOE effect between the compound's protons and solvent's labile protons can be observed. Fortunately, it is easy to find and very reliable. However, one needs to be cautious. "Visibility" of such a proton is closely related to the dynamics of chemical exchange. Only slow exchanging and fast exchanging protons are registrable

in the spectrum. If the proton is in an indirect exchange (comparable to the NMR timescale), there will be no signals visible. To reveal whether there are protons in an indirect exchange present within a structure, additional experiments involving different temperatures should be performed[63].

The molecular tumbling rate and the NMR observation frequency is closely related to the magnitude and sign of a proton-proton NOE. Small molecules generally generate positive NOEs, while the same effect is usually negative for large entities. Molecules with a mass of about 1 kDa may not produce NOE signals at all. This is where ROESY comes in handy. ROESY is very similar to NOESY, but NOE signals are generated in a rotated frame. The sequence of pulses is analogous to that used in the case of TOCSY.

1.4.3 HSQC and HMBC

Sometimes, ^{13}C NMR data is essential for structural studies. For instance, it may be necessary to determine which proton is bound to which carbon. It is possible with heteronuclear single quantum coherence spectroscopy (HSQC), with pulse sequence given below.

$$-{}^{1}H - \ RD - \ 90°x - \tau - 180°y - \tau - 90°y - 180° - 90° - \tau - 180° - \tau - \ ACQ - \quad [5]$$

$$^{13}C - 180°y - 90°y - \frac{t_1}{2} - \frac{t_1}{2} - 90° - 180° - ACQ - \quad [6]$$

The main idea, standing behind hetero-correlation spectroscopy, involves transferring magnetization from the proton nucleus to the heteroatom nucleus. Most often it concerns protons and carbons. The spectrum consists of two axes, with the x axis representing ^{1}H nuclei and the y axis representing ^{13}C atoms. This spectroscopy reveals a coupling solely between a proton and a carbon via one bond. If such information is not satisfactory, there is another kind of NMR experiment, named heteronuclear multiple bond correlation spectroscopy (HMBC). The most interesting difference between HSQC and HMBC is that HBMC allows to “trace” a coupling via 2 or even (rarely) 4 bonds. It is a very useful technique to unambiguously perform an appropriate assignment. However, one needs to be cautious again, since the presence of a signal is an evidence that there is a coupling, but when there is no signal, it means that the proton may, or may not, be there. This complication

arises from the Karplus relationship, which strongly affects the correlation intensity, hence for a dihedral angle equal to 90 degrees the coupling constant is close to zero[64].

1.5 UV-Vis spectroscopy

Despite the fact that the NMR technique is indisputably useful and adequate method to investigate the physicochemical properties of chemical entities, it is rather not very handy in the case of determining various physicochemical constants, especially at low concentrations. Likewise, not everyone has access to such an advanced and costly apparatus. Fortunately, UV-Vis spectroscopy comes to the rescue. It is also a fairly simple and fast method that uses a relatively cheap spectrophotometer. Importantly, the amount of information that can be obtained by this method is quite large. All of that makes UV-Vis spectroscopy an ideal method to investigate physicochemical properties, such as proton dissociation[65], aggregation and general behaviour in the solution[6966]. The combination of the UV-Vis spectroscopy with advanced numerical methods[67] is especially useful.

Electronic spectroscopy provides information about the structure of chemical entities by studying electronic transitions between ground and excited states. Depending on the type of electrons present in a chemical compound, different types of electronic transition are possible. Absorption of ultraviolet and visible radiation causes electrons to be excited and transferred from the ground state to the excited state. As a result of this process, electrons are transferred from the bonding orbital to the antibonding orbital. Possible electronic transition are listed below:

- $\sigma \rightarrow \sigma^*$,
- $n \rightarrow \sigma^*$,
- $\pi \rightarrow \pi^*$,
- $n \rightarrow \pi^*$.

The $\sigma \rightarrow \sigma^*$ transition causes the transfer of electrons a from the σ-bonding orbital to the σ*-antibonding orbital. It is a transition that requires high energy due to the strength of the σ bond, which is manifested by the absorption of radiation in a very short range of electromagnetic waves. This transition is typical for saturated hydrocarbons[68].

More interesting is the $n \rightarrow \sigma^*$ transition. In this case, n electrons (a single pair of electrons) are transferred to the σ* antibonding orbital. This transition does not require as much energy as the $\sigma \rightarrow \sigma^*$ transition, so it occurs for the longer

wavelength. This transition is characteristic of compounds containing oxygen, nitrogen, sulfur, halogen atoms or other atoms with a lone pair of electrons in their structure[69]. For this transition, the wavelength at which radiation absorption occurs is closely related to the electronegativity of the atoms. This is because with increasing electronegativity, more energy is required to excite the electrons[70].

One of the most interesting electron transitions is the $\pi \rightarrow \pi^*$ transition. This transition occurs in compounds which have π electrons in their structure, such as aromatic compounds, ketones, aldehydes, alkenes and many others, and involves transfer of electrons from bonding π orbital to antibonding π^* orbital[71].

The transition with the lowest energy is the $n \rightarrow \pi^*$ transition. The electrons of the lone electron pair of n orbital become excited and move to the antibonding π* orbital[71].

One might consider a chemical compound, in which many electronic transitions can occur simultaneously. If there are transitions: $\sigma \rightarrow \sigma^*$, $n \rightarrow \sigma^*$, $\pi \rightarrow \pi^*$, $n \rightarrow \pi^*$, the compound will absorb energy over a relatively wide range of wavelengths. According to the above, the ideal UV-Vis spectrum should look like the one depicted in Figure 17[72].

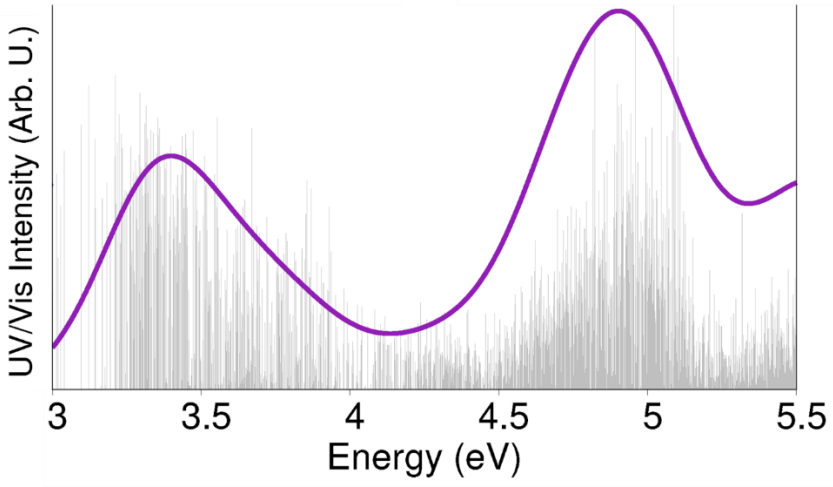


Fig. 17. Stick plot and spectrum of flavonoid. Sticks represents intensity of electronic transition at given energy (wavelength) whereas curve is a real spectrum[72].

Sticks are broadened due to the uncertainty principle, measurements errors and many other reasons. The actual spectrum is a linear combination of all of these broadened signals and can be expressed as a sum of Voigt curves (convolution of a Cauchy-

Lorentz distribution and a Gaussian distribution)[73]. The resulting envelope can be further analysed using advanced chemometric techniques, which are described in a chapter entitled *Techniques applied in the ligand/DNA interactions studies*.

2 OBJECTIVE AND SCOPE OF WORK

The aim of the present study was **to investigate the interactions of selected antitumor unsymmetrical bisacridine derivatives with the C-Myc protooncogenes' promotor region using advanced NMR techniques, enhanced by molecular dynamic methods**. This sequence plays important role in human cells and is often taken into consideration as the first molecular target when testing new chemotherapeutics. Moreover, C-Myc is highly overexpressed in tumor cells. Previous study indicates that selected UAs strongly interact with C-Myc protooncogenes' promotor regions. In addition, unlike C-2028 and C-2041, C-2045 and C-2053 exhibited much higher cytotoxicity against cancer cells than against normal cells, as demonstrated in previous studies[74]. Taking all the data into account, derivatives C-2045 and C-2053 were chosen to study their interactions with selected G-quadruplex. Studied agents, i.e., unsymmetrical bisacridines (UAs), are based on two different acridine aromatic systems, which are linked by diamino-alkylamino chain. Such a complex structure can exhibit the ability to self-associate in relation to various pH, concentration and ionic strengths. Moreover, several pH-dependent forms of the compounds are also possible. There are many papers describing quantum chemical structure of acridines, which show how many various variants can be considered[75,76].

In the present work on the development of anticancer drugs, knowledge about the individual compound forms is essentially important, mainly due to the various pH values in tumor microenvironment and the different pH values compared to normal cells. This kind of dysregulation can promote metastasis, or can even make cancer more invasive. In the case of cancer cells, the pH of intracellular fluid is equal to 7.4 and the pH of the extracellular microenvironment may vary from 6.6 to 6.9[77]. For normal cells, these values are slightly different: the pH of the intracellular fluid is about 7.2, while the extracellular fluid is 7.4[81]. Interestingly, UAs compounds probably exist in at least two different forms in the pH range typical for cancer cells, while they exist in different forms in the pH range characteristic for normal cells. All above reveals how important it is to know the three-dimensional structure, as well as physicochemical properties of the studied compounds. It is a key to further research. Due to the differences in protonation state, UA compounds can interact differently with various

cellular components. In addition, when helical DNA, G-quadruplexes or RNA are taken into consideration as main molecular targets, it is crucial to know the electronic structure of the studied compounds, due to the negatively charged parts of nucleic acids. Lack of this knowledge can lead to misinterpretation of research results.

Considering all the above, the undertaken research should be properly planned. The first step was to build numerical models of the studied compounds in such a way, that the models were as close as possible to the actual UA structures in aqueous solutions. Therefore, a complex spectroscopic analysis supported by advanced chemometric techniques was performed.

In summary, the scope of the work can be divided into the following subtasks:

- **Physicochemical properties of UAs:**
- determination of proton dissociation constants,
- discovering protonation states of studied UAs,
- UAs self-association study.
- **Interaction of UAs with G-quadruplexes:**
- recording NMR spectra of non-complexed G-quadruplex and signal assignments,
- recording NMR spectra of G-quadruplex/UAs complexes and signal assignments,
- building molecular models of resulting complexes,
- study on complexes using molecular dynamic methods.

3 MATERIALS AND METHODOLOGY

3.1 Materials

Hydrochloride salts of:

- **C-2028**: (9'-{N-[(imidazo [4,5,1-de]-acridin-6-on-5-yl)aminopropyl]-N--methylaminopropylamino}-10 -- nitroacridine),
- **C-2041**: (1-[3-(imidazo [4,5,1-de]-acridin-6-on-5-yl)aminopropyl]-4-[3'-(10 --nitroacridin-1- yl)-aminopropyl]piperazine),
- **C-2045**: (9'-{N-[(8- hydroxyimidazo[4,5,1-de]-acridin-6-on-5-yl)aminopropyl]-N--methylaminopropylamino}-40 -methyl-10 -nitroacridine),
- **C-2053**: (9'-{N-[(imidazo [4,5,1-de]-acridin-6-on-5-yl)aminopropyl]-N--methylaminopropylamino}-40 -methyl-10 -nitroacridine),
- **C-1311**: (5-Diethylaminoethylamino-8-hydroxyimidazoacridinone),
- **Ledakrin**: (1-Nitro-9-(3-dimethylaminopropylamino)-acridine),
- **Structural analogue of the C1311**: 5-((3-((3-aminopropyl)(methyl)amino)propyl)amino)-6H_imidazo[4,5,1-de]acridin--6-one

were studied. All unsymmetrical bisacridines (UAs) were synthesized in the Department of Pharmaceutical Technology and Biochemistry, Gdańsk University of Technology.

Other chemical entities

- formic acid,
- HEPES,
- hydrochloric acid,
- methanol,
- PIPES,

- potassium cacodylate,
- sodium acetate,
- sodium hydroxide,
- sodium tetraborate,
- sulphuric acid,

Oligonucleotide

- Mutated 22-mer of the G-quadruplex sequence from human C-Myc oncogene promotor: 5'-TGA GGG T GGG TA GGG T GGG TAA-3', designated as Pu22.

3.2 Techniques applied in the ligand/DNA interactions studies

In order to solve the problems defined in the objectives and scopes of work, I applied several research techniques. Their theoretical basis is described below.

3.2.1 UV-Vis Absorption Spectroscopy

A single absorption spectrum recorded at *n* wavelengths can be thought of as an *n* element column vector.

$$\mathbf{w} = [w_1\ w_2\ w_3\ \ldots\ w_n] \qquad (2)$$

An ordered set of *m* such spectra can be treated as a two-dimensional matrix **W** with *m* columns and *n* rows:

$$\mathbf{W} = [\mathbf{w}_1\ \mathbf{w}_2\ \mathbf{w}_3\ \ldots\ \mathbf{w}_m] = \mathbf{W}_{nm} \qquad (3)$$

According to the principle of additivity, the absorbance of the *k* component sample at wavelength λ can be described as:

$$w(\lambda) = \sum_{i=1}^{k} c_i\, w_i(\lambda) \qquad (4)$$

c_i – concentration of *i*-th component of the sample,
$w_i(\lambda)$ – absorbance of *i*-th component at wavelength λ.

3.2.1.1 Chemometric approach in spectra analysis

Chemometrics is especially useful in the analysis of spectral data. Set of spectra can be organised column-wise into an absorbance matrix, **W**. From now on, each column of **W** shall be treated as a vector $\mathbf{w}_i$ in the *m*-dimensional vector space. According to the Lambert-Beer law, the absorption is linearly dependent on concentration. This means that the absorbance matrix **W** can be represented as a product of the two matrices: the concentration matrix **C** and the matrix of basic spectra **B**. In order to decompose matrix **W** into matrices that contain the concentration and spectra of the pure components, respectively, the absorbance matrix must be verified and properly prepared. Verifying the data includes analysis of: mean and standard deviation values for each column, as well as correlation coefficients analysis of each pair of columns. The curve showing the mean values and standard deviations of each vector should be reasonably smooth and there should be no sharp spikes in the correlation coefficients plot. All spectra (vectors) that produce a spike on one of the curves mentioned above should be removed from the data set.

First of all, the data need to be prepared for further analysis, which means that the dataset has to be placed at the origin of the coordinate system. It can be done by subtraction the mean value of each column vector (for column-wise representation) from each absorbance value (5).

$$x'_{ij} = x_{ij} - \overline{x_j} \qquad (5)$$

x'_{ij} – centred data,

x_{ij} – raw data set,

$\overline{x_j}$ – mean values of j-th column.

Sometimes, to simplify the analysis, the data are additionally variance scaled to reach the standard deviation of each column equal to 1. It can be done using an equation given below. This approach combines data centering and variational scaling, which is called autoscaling (standardization) (6).

$$z_{ij} = \frac{x_{ij} - \overline{x_j}}{s_j} \qquad (6)$$

z_{ij} – autoscaled data matrix,

x_{ij} – initial value of the variable,

$\bar{x_j}$ – mean value of the j-th variable,

s_j - standard deviation of the j-th variable.

An example of different scaling is given below (Figure 18).

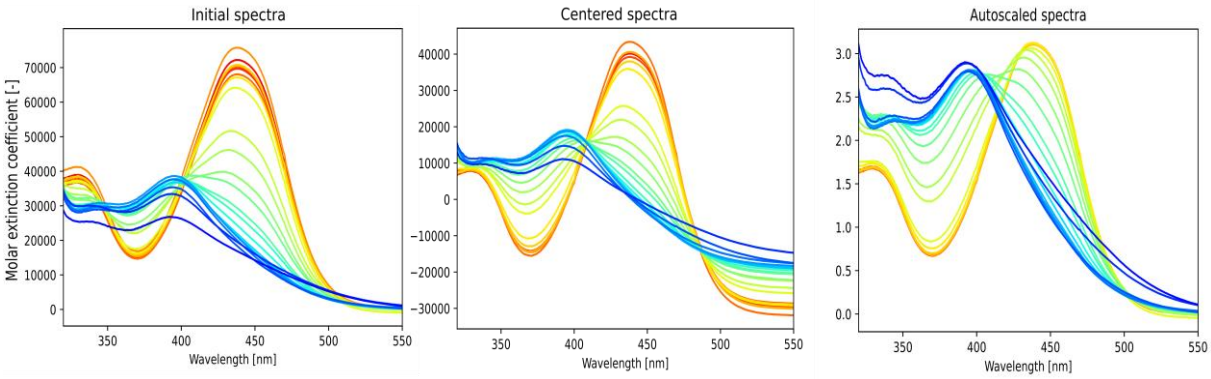


Fig. 18. Data scaling (no-scaled spectra – left, centered spectra – in the middle, autoscaled spectra – right.

Figure 18 shows the effect of centering and autoscaling on the spectra. In the case of centering, the set of spectra is shifted relatively to the y-axis so that 0 passes through its center, but spectra intensity remains unaltered by this operation. During autoscaling, division by standard deviation additionally scales spectra intensities and they become easier to compare.

Both types of scaling ensure that all spectral vectors are anchored to the origin. However, it is important to be aware that, after standardization, all vectors have a length equal to 1, so the ends of the vectors are placed on an arc. Conversely, after centering, all vectors have an original length and are placed on a plane. In general, for the *n*-dimensional vector, its end is located on the hypersphere (for autoscaled vectors) or on the hyperplane (for centered vectors).

Using centered data is preferable because of omitting spherical coordinate systems during further analysis. An easiest way to illustrate this concept is presented below (Figure 19):

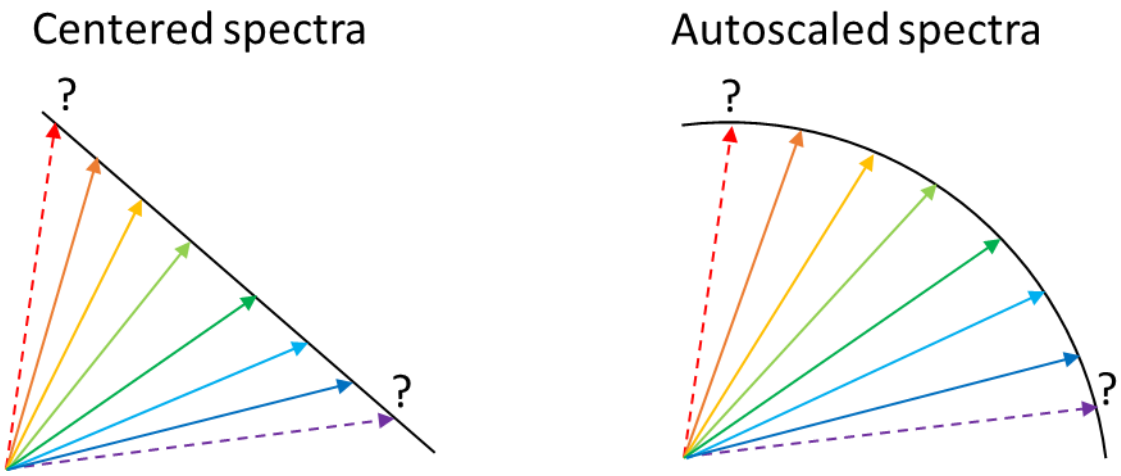


Fig. 19. Centered spectra (left), autoscaled spectra (right).

It is now easy to see that the ends of the centered vectors align on a line, while the ends of the autoscaled vectors trace out a curve. In this dissertation, only centered spectra were used, to avoid curved coordinate systems and to make it possible to find an analogy to the ternary plot.

Sometimes, especially if the spectra are collected in low concentration, or in case one is dealing with CD spectra (performing low number of scans to save time), there is a need to smooth the spectra. This is crucial, because some chemometric techniques treat the noisy spectrum as an additional spectral form, since it cannot be recreated by other base spectra (vectors) and it becomes a base vector itself. There are a number of data smoothing algorithms in signal analysis techniques. In the chemometric analysis of UV-VIS spectra, the one commonly used is so-called Savitzky–Golay filter (SG)[78], which employs low-degree polynomials to approximate the experimental curve (spectrum). The biggest advantage of this approach is the fact that SG is shift-invariant. It means that there is no shift in the smoothed spectrum in respect to the initial one.

The basic technique for the numerical analysis of a set of spectra is Principal Component Analysis (PCA). Principal component analysis is a procedure, which performs an orthogonal transformation of the system of correlated variables, X, into

new variables, Y. Y are a linear combination of the initial variables, X. The main idea behind principal component analysis is to describe the variability of *n* points in p-dimensional space of a given feature by introducing a new, linear and orthogonal coordinate system. This is possible by solving the so-called the eigen problem of the correlation/covariance matrix:

$$Rv = \lambda v \qquad (7)$$

R – m*m square matrix (correlation matrix if data was autoscaled or covariance matrix if data was centred),

v - n-element vector (eigenvector),

λ – eigenvalues corresponding to a particular eigenvector.

The first step in PCA is the determination of the eigenvectors and corresponding eigenvalues, according to the equation (7).

What the 'principal components' actually are and what is their true meaning? One could consider an exemplary dataset, consisting of two column matrix (columns contain explanatory variables). Hence, it is possible to make a projection of the variables onto the column vectors (Figure 20).

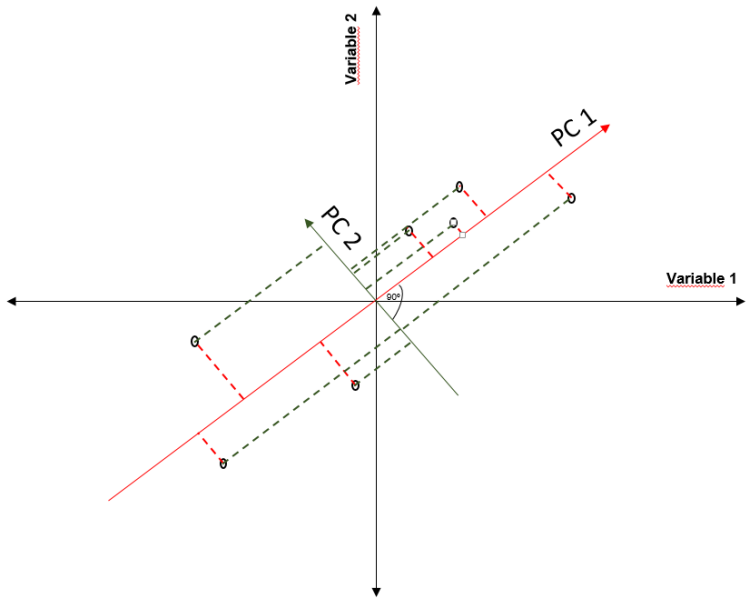


Fig. 20. Two dimensional illustration of the principal component analysis.

Eigenvalues for PC 1 and PC 2 are the sum of their squared distances, correspondingly. If each eigenvalue is divided by *n*-1, the result is the variance of the corresponding specific PC. In fact, PCA is often performed via another algorithm than presented above, which is named singular value decomposition (SVD)[79]. This approach is more effective because of the higher numerical stability[83].

PCA of the UV-Vis spectra of a given molecular system (chemical compound) over specific experimental conditions (pH values, concentrations, temperature, etc.) is typically used to determine number of spectral forms and their molecular fractions. The studied molecular system can still be represented as a single point in vector space, and this point traces out the trajectory according to the experimental conditions, as it happens in classical phase space.

It is very convenient to think of spectra as vectors placed in vector space span on the basic vectors. Continuing this line of thought, it might be stated that one is sampling a vector space in such a way that spectra are collected under varying conditions, which can be represented as points in a space, which is in full analogy to the configuration space (as in dynamical system theory). For instance, one can imagine that there is a set of spectra collected for an acridine derivative, with two nitrogen atoms able to protonate, at varying pH values. It may be expected that there are three 'spectral pure' chemical entities, for example neutral (A), bearing one positive charge (A+) and the third with two positive charges (A2+). Hence, it might be stated that all of the collected spectra are linear combinations of three basic spectra (vectors), which is indeed the case. Going further, while visualizing all the measured spectra as a set of individual points, this space can be treated as a kind of configurational space, which will be named *molar fraction space* from now on. In this space, the entire state of system (protonation state of chemical compounds under certain conditions) is represented by single points, while each axis of the coordinate system expresses a position along specific protonation state. Evolution of the system takes place according to the given parameter (pH, concentration etc.).

However, how points placed somewhere in molar fraction space may be transformed into actual molar fractions? Here, an analogy to the ternary plot may be used, but generalised to *N* dimensions. It means, for example, that for five pure spectral forms,

the system corresponds to the five-dimensional analogue of a triangle and all the mathematical formulas are in full analogy.

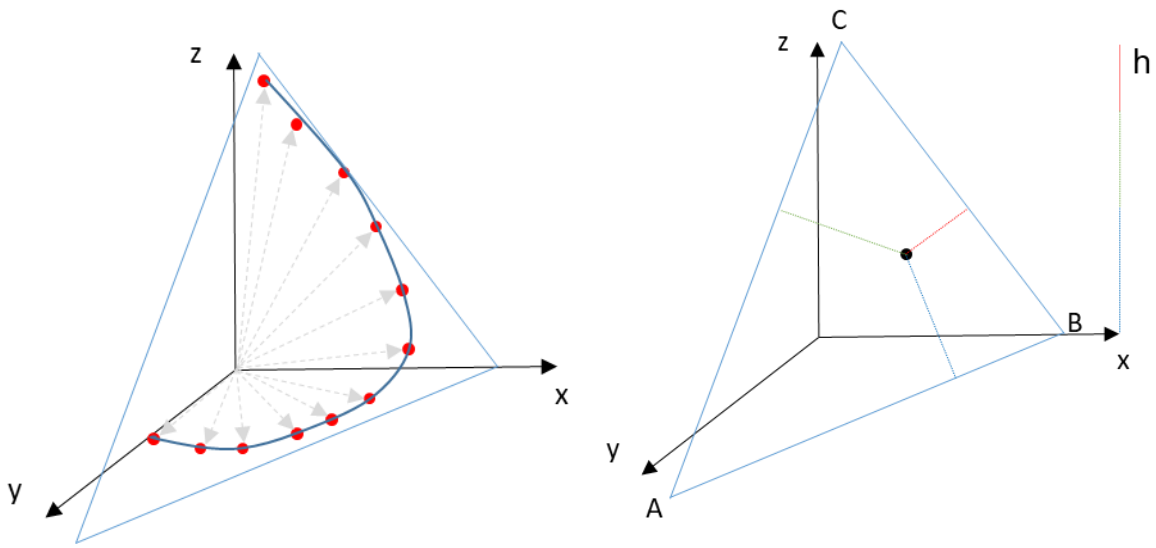


Fig. 21. Projection of the spectra points onto two dimensional space (left) and illustrating of the altitude method derived from Viviani's theorem[80] (right). Coordinates: x, y, z – base vectors of pure spectral forms, A, B, C – pure spectral forms, red points – system states under given condition, gray dotted arrows – various position of vectors tracing out system states.

An example for a three-dimensional system is presented in Figure 21. The sum of the perpendicular components, i.e., the shortest distances from a given point to each of the three sides (or *N* sides in general) are equal to the height of the triangle (or *N* dimensional simplex). The ratio of each component to the height is indeed a molar fraction. In order to discuss a higher dimensional space, another example for a four-dimensional system is presented in Figure 22.

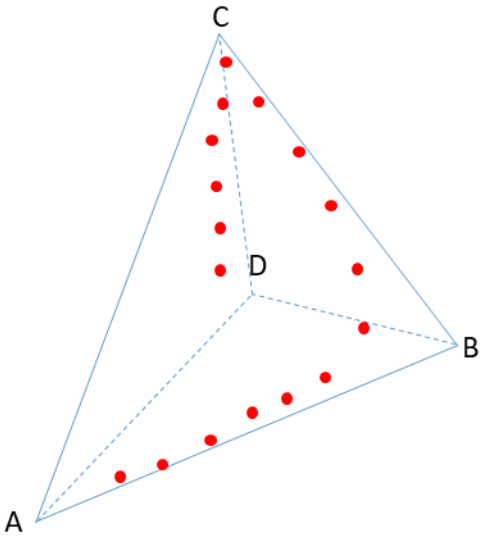


Fig. 22. Projection of the spectra onto two dimensional space (tetrahedral projection for a four-dimensional object).

Once again, each point can be translated into molar fraction using perpendicular components, but this time there is no interest in components perpendicular to each side, but components perpendicular to a plane. In the previous example, a projection of the points onto the *N*-1 dimensional plane – triangle – was used. Herein, the space contains additional dimension, which is impossible to show on a plane, but fortunately it is possible to project an object from four dimensional space onto three dimensional space. Going further, it is also possible to find a projection of a three dimensional object onto a two dimensional plane. As a result, a tetrahedron is obtained. This approach can be easily generalized to *N* dimensions.

$$d = \frac{|wx+b|}{||w||} \qquad (8)$$

$$wx + B = 0 \qquad (9)$$

While examining higher dimensional problems, triangles or tetrahedrons are no longer applicable. One has to deal with an *N*-dimensional simplex, which is a generalized version of a triangle. Fortunately, all of the computations are analogous, hence it doesn't matter how high the dimensionality of the problem actually is.

3.2.1.2 Selection of physicochemical and biophysical models

In scientific research, there is often a need to fit the obtained results to a theoretical model. For example, in studies on self-associations of the chemical agents, it is possible to determine a mode of aggregation by finding the best fit to a model curve, corresponding to a specific self-association model. However, how to find the best fit? Herein, the numerical optimization method comes to the aid. The basic problem is to define an appropriate, objective function and to select an algorithm that allows for a quick determination of its minimum. The best way to find an objective function extremum of the unknown model function (while only its dimensionality and approximate shape derived from sampling its surface is known) is the Nelder-Mead simplex algorithm[81], which employs a direct search method and is useful for solving the unconstrained optimization problems.

$$\min f(x) \quad (10)$$

$$f : \mathbb{R}^n \rightarrow \mathbb{R} \quad (11)$$

f – objective function,
n – number of dimensions.

This method is based on comparison of the values of the objective function at *n*+1 vertices of the simplex (geometric figure in *n* dimensions with *n*+1 vertices) and moving it towards the extreme. A typical objective function is the sum of squared differences between the experimental values and the model outputs.

3.2.2 Protonation equilibrium and self-association

3.2.2.1 Protonation

The equilibrium of a protonation/deprotonation state is:

$$AH \overset{K_a}{\leftrightarrow} A^- + H^+ \quad (12)$$

From now, all concentrations are written without square brackets. For equation given above, K_a is equal to:

$$K_a = \frac{H^+A^-}{AH} = 10^{-pK_a} \quad (13)$$

The molar fractions of undissociated protons are:

$$u_H = \frac{AH}{A^- + AH} \quad (14)$$

Concentration of protons can be expressed as:

$$H^+ = 10^{-pH} \quad (15)$$

For multiprotic compound with different dissociation constants, equation (14) can be expressed as:

$$u_H = \frac{\frac{AH}{A^-}}{1+\frac{AH}{A^-}} = \frac{\frac{H^+}{K_a}}{1+\frac{H^+}{K_a}} \quad (16)$$

Based on equations (12) and (15), the equation above can be rewritten as:

$$u_H = \frac{10^{pK_a - pH}}{1+10^{pK_a - pH}} \quad (17)$$

Now, assuming that there is a chemical agent which can be described with two equilibrium constants, K_1 and K_2:

$$H^+ + A^{2-} \overset{K_1}{\leftrightarrow} AH^- \quad (18)$$

$$H^+ + AH^- \overset{K_2}{\leftrightarrow} AH_2 \quad (19)$$

For equation above, equilibrium constants are:

$$K_1 = \frac{AH}{H^+A} \quad (20)$$

$$K_2 = \frac{AH_2}{H^+AH} \quad (21)$$

Now, the molar fraction of undissociated protons are:

$$u_H = \frac{AH+2AH_2}{A+AH+AH_2} = \frac{\frac{AH}{A}+2\frac{AH_2}{A}}{1+\frac{AH}{A}+\frac{AH_2}{A}} \quad (22)$$

From equations (20) and (21), one can calculate:

$$\frac{AH}{A} = K_1 H^+ \qquad (23)$$

and

$$\frac{AH_2}{A} = \frac{AH_2}{AH}\frac{AH}{A} = K_2 K_1 (H^+)^2 \qquad (24)$$

Hence, equation (22) can be rewritten as:

$$u_H = \frac{K_1 H^+ + 2K_1 K_2 (H^+)^2}{1+K_1 H^+ + 2K_1 K_2 (H^+)^2} \qquad (25)$$

The above can be generalized for *n* K constants and written in short as:

$$\gamma = \frac{\prod_{i=1}^{n} K_i}{\prod_{i=1}^{n} H^i} \qquad (26)$$

$$u_1 = \frac{1}{1+\gamma} \qquad (27)$$

$$u_n = u_1 \cdot \gamma \qquad n = \{1,4\} \quad (28)$$

3.2.2.2 Self-association

As the concentration of the chemical agents increases, they can aggregate. In the simplest case, the compounds start to form dimers. The equation of this process is as follows:

$$2M \Leftrightarrow D \qquad (29)$$

With dimerization constants:

$$K_D = \frac{D}{M^2} \qquad (30)$$

Concentration of a dimer can be expressed as:

$$D = K_D M^2 \qquad (31)$$

The next possible model is trimerization, available in two variants. One step trimerization is very unlikely from a statistical and thermodynamic point of view. Two steps process seems to be more reliable, in which the third agent attaches itself to the previously formed dimer. However, a more interesting situation occurs while considering unlimited self-association. There are two possibilities: the first, when all aggregation constants are equal; and the second one, when a particular aggregation constant differs. In the first case, a concentration of monomer can be calculated from the equation given below:

$$c = M\frac{\partial}{\partial M}\left(\frac{M}{1-K_A M}\right) = M\frac{1(1-K_A M)-M(-K_A)}{(1-K_A M)^2} = M\frac{1-K_A M+K_A M)}{(1-K_A M)^2} = \frac{M}{(1-K_A M)^2} \qquad (32)$$

If the aggregation constants are different, a concentration of a monomer can be calculated as follow:

$$c = M\frac{\partial z}{\partial M} = M + \frac{2K_D M^2}{(1-K_A K_D M^2)^2} \qquad (33)$$

M – monomer concentration,
K_A – aggregation constant,
K_D – dimerization constant.

3.2.3 1D and 2D NMR studies of G-quadruplex structures

3.2.3.1 Proton labelling

Nuclear magnetic resonance (NMR) is a powerful tool for the determination of the three-dimensional structure of G-quadruplexes. Two-dimensional experiments are of particular importance, but one-dimensional techniques are also very useful. The most widely employed spectrum is NOESY, mainly because it gives the opportunity to study the proximity of protons in space. Figure 23 presents proton labelling in DNA:

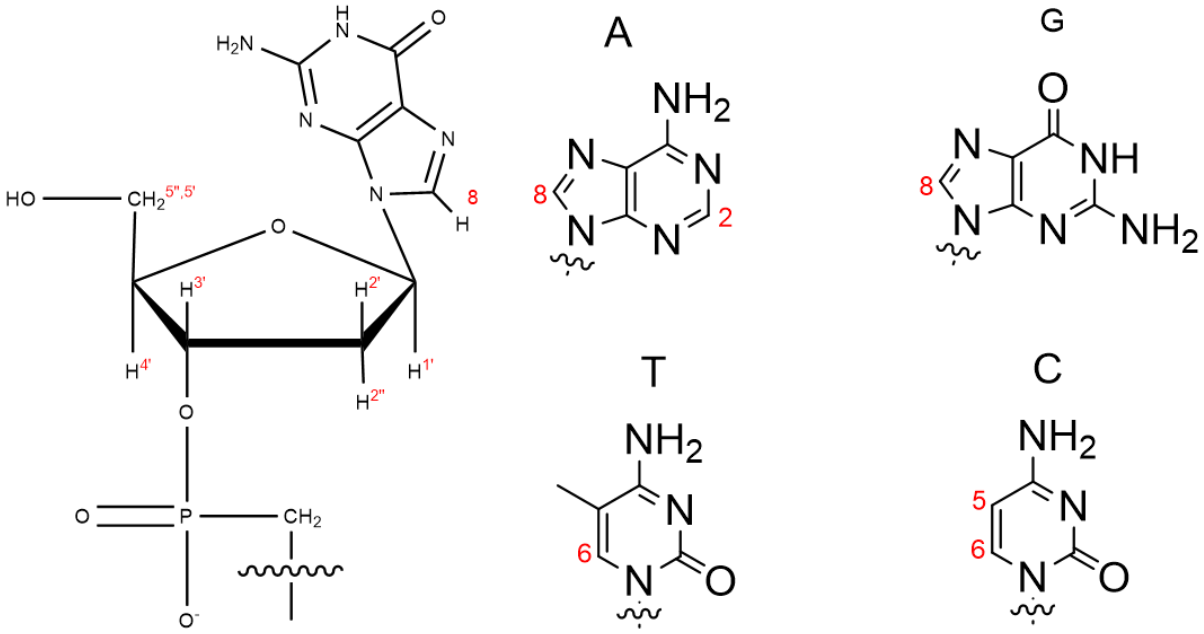


Fig. 23. Proton labelling for DNA bases.

3.2.3.2 Rules for proton assignments

The most important part of the NOE spectrum is the region containing correlation signals between protons located on 8 carbon atoms of guanines and $H^{1'}$ protons of their deoxyribose. The $H^{1'}$ proton of a sugar moiety exhibits dipolar coupling to the guanine proton of its own base and to the guanine proton of the next (3' direction) tetrad (Figure 24).

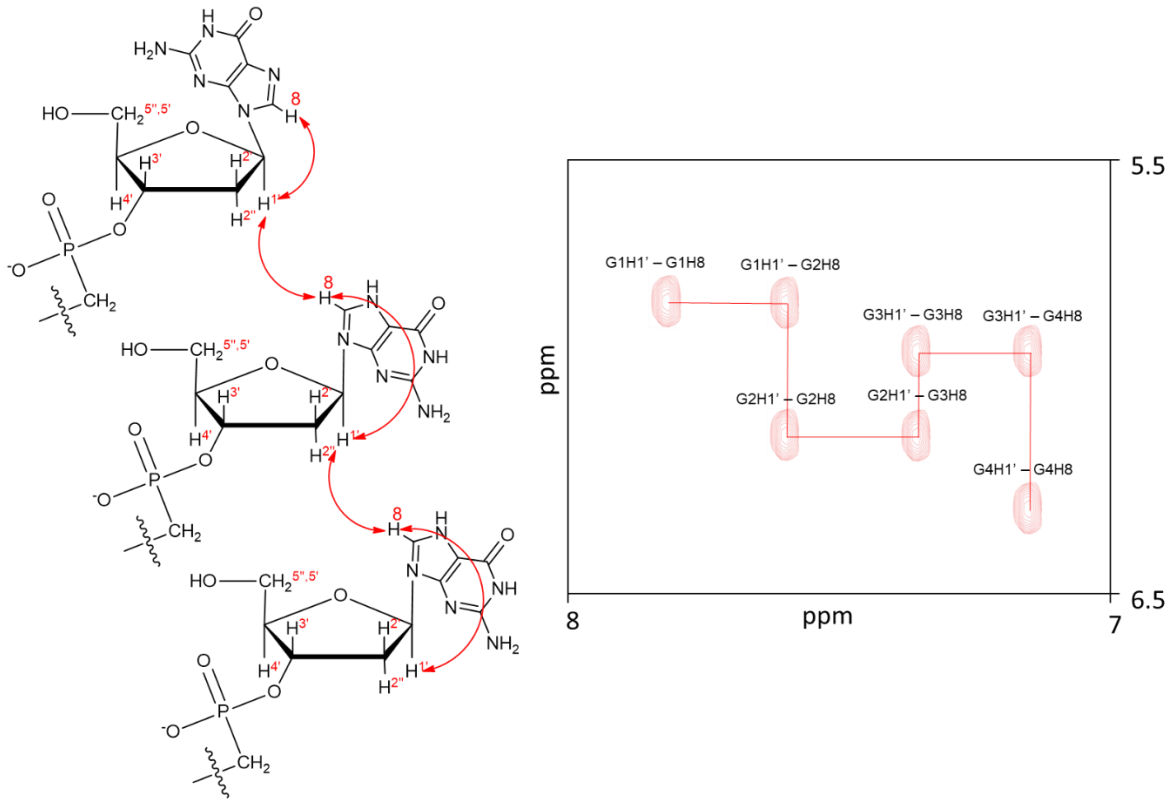


Fig. 24. Rules NOE proton assignment pathway for G-quadruplex (proton coupling – left, signals that correspond to protons – right).

H1' protons usually display dipolar couplings to all other sugar protons of deoxyribose, which is shown in Figure 25.

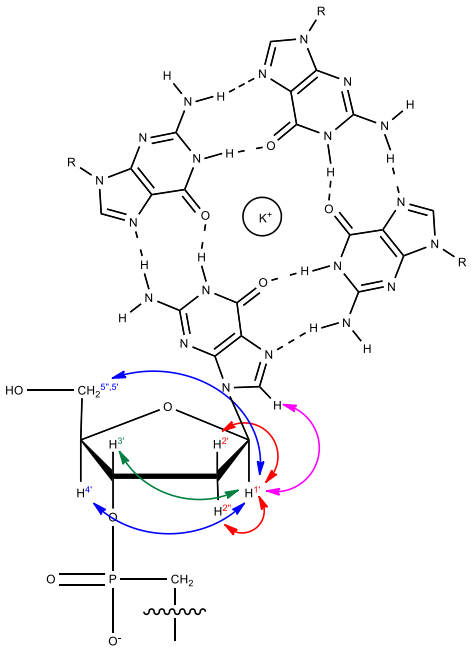


Fig. 25. Proton coupling between deoxyribose protons.

The best way to distinguish guanine protons from other DNA bases (adenine, thymine, cytosine) is the HSQC (Heteronuclear Single Quantum Coherence) spectrum, particularly the region presented in the figure below (Figure 26).

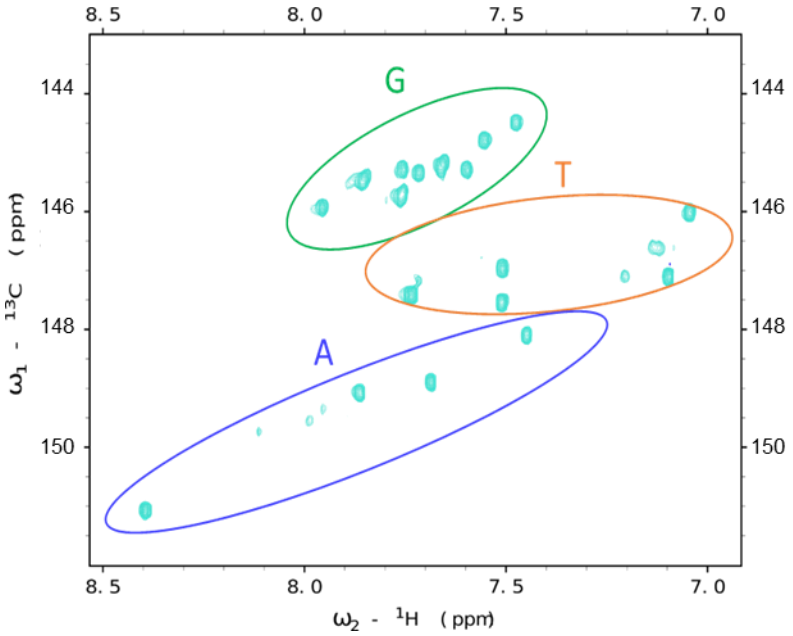


Fig. 26. DNA base protons clustering in the HSQC spectrum.

Guanine signals tend to group in the upper part of the presented fragment, thymines are placed slightly lower and adenine signals are in the lowest part of the aromatic region.

3.2.4 Molecular dynamics for ligand/DNA complexes

Molecular dynamics is usually employed to investigate the behavior of the studied complexes in aqueous environment. This approach involves using Newton's equations of motion to represent the evolution of a system over time. The entire process can be divided into several basic steps:

a) ligand parametrization exploiting quantum chemistry calculations,
b) preparation of an initial structure of the complex,
c) energy minimization,
d) molecular dynamics.

From now on, to simplify considerations discussed herein, an approach derived from differential geometry will be employed.

An initial state of a system (for instance, a given UAs/DNA complex) containing a large number of atoms can be represented by single point, placed somewhere on 3N dimensional configuration space (while considering only the velocity components) or on 6N dimensional phase space (while considering both velocities and momentum). The main purpose of molecular dynamics is to sample this phase space, in order to observe the behavior of the studied complex over time. This concept is presented in Figure 27.

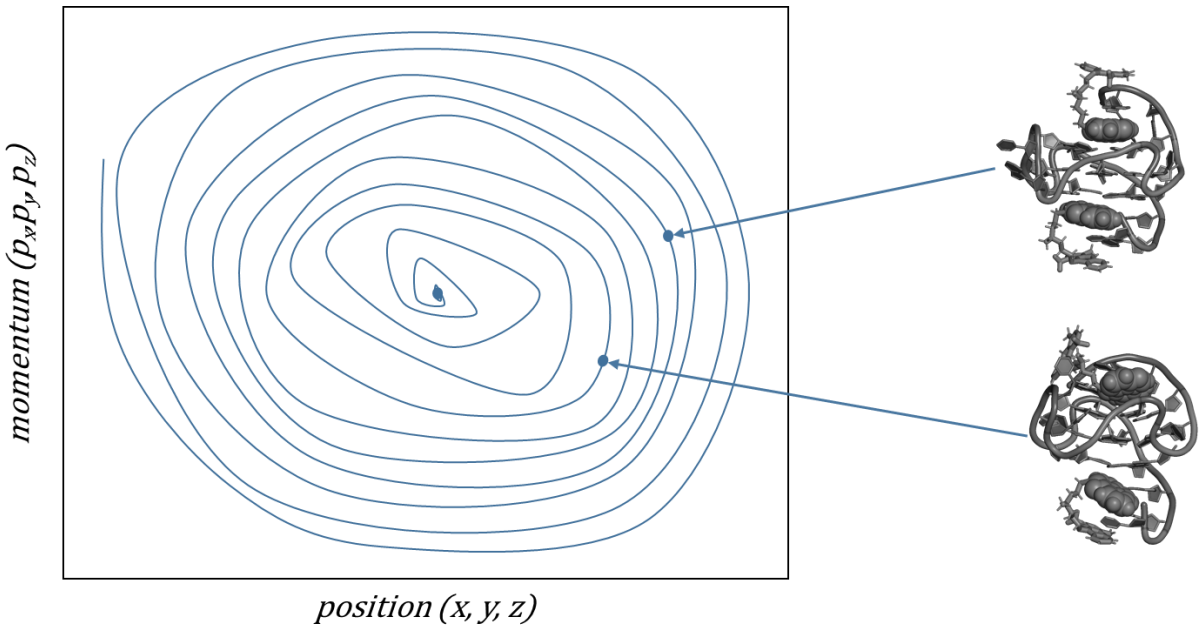


Fig. 27. Example of phase trajectory (with focal point) traced out by ligand/DNA complex.

Phase trajectory is a parametrical curve, given by the integration of the Hamiltonian equations (describing evolution of a studied system over time):

$$\frac{\partial H}{q} = \dot{p},\ \frac{\partial H}{\partial p} -= \dot{q} \qquad (34)$$

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V \qquad (35)$$

p – generalized position,

q – generalized momentum,

m – mass of particles,

V – potential energy,

H – Classical Hamiltonian.

The structures of a studied complex (points tracing out by vector state), resulting from solving the Hamilton equation presented above, can be clustered after simulation to obtain the most representative structures.

Umbrella sampling

When analyzing intermolecular interactions, one of the most important elements of the description of the phenomenon is quantification of the free energy of the interactions between them. Many methods allow the estimation of free energy values; one of them is the umbrella sampling (US) method described below. This energy, denoted as *G*, can be determined by the following relationship:

$$G = -\frac{1}{\beta} lnQ \qquad (36)$$

Where *Q* can be expressed as:

$$Q = \int \exp[-\beta E(r)] d^N r \qquad (37)$$

$$\beta = \frac{1}{k_B T}$$

Q – canonical partition function,

k_B – Boltzmann's constant,

T – absolute temperature,

N – number of degrees of system freedom,

r – coordinates of a system.

To calculate the free energy of a system using the above-mentioned relationship, extensive sampling of the phase space is required. Hence, a new parameter has to be introduced, named the reaction coordinate (ξ). This parameter corresponds

to thermodynamic states of a system. In the first step, the probability of system distribution along ξ needs to be established, which can be expressed as:

$$Q(\xi) = \frac{\int \delta[\xi(r) - \xi] \exp[-\beta E(r)] d^N r}{\int \exp[-\beta E(r)] d^N r} \quad (38)$$

Based on that equation, free energy of a system along ξ can be calculated as:

$$G(\xi) = -\frac{1}{\beta} ln Q(\xi) \quad (39)$$

However, in practice, G cannot be determined from these equations directly. To overcome this problem the following formula is used:

$$P(\xi) = \lim_{t \to \infty} \frac{1}{t} \int_0^t \rho[\xi(t')] dt' \quad (40)$$

P(ξ) – time average,

t – time,

ρ – number of occurrences in a given state.

According to the above equation, the free energy of a system can be determined by counting the occurrences of the system states along the reaction coordinate. To ensure that the phase space is sampled at a sufficiently high level, several parallel simulations are performed simultaneously by applying an external potential (usually harmonic), in order to force the system to pass through energetically unfavorable states. The data obtained in this way are further analyzed using the weighted histogram analysis method (WHAM)[82].

4 Research approaches

4.1 Molar fractions of spectral forms

Stock UAs solutions (0.01 M) were prepared in deionized water and then diluted to 10^{-3} M or 10^{-4} M concentration if necessary.

4.1.1 Determination of pKa

Series of mixtures containing UAs (2×10^{-5} M) and appropriate buffer solutions or hydrochloric acid (0.2 M) with following pH values: 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6.1, 6.3, 6.5, 6.7, 6.9, 7.1, 7.3, 7.5, 7.7, 7.9, 8.1, 9, 10 and 11 were prepared. In the range from 1.5 to 3.5 the solution of hydrochloric acid was applied. For pH range from 4 to 11, the most suitable buffer solutions were used as follows: acetic buffer, PIPES buffer, HEPES buffer and borate buffer. In the next step, the sets of UV-Vis spectra were collected. For all pH-dependent measurements the optical pathway of a quartz cuvette was equal 1 cm.

In all the cases the spectra were centered. In the next step, the transformed sets of spectra (W) were used to perform the principal component analysis of the W matrix. The correct way to perform singular value decomposition (SVD) for centered spectra is decomposition via covariance matrix:

$$cov(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \qquad (41)$$

x,*y* – variables,

n – number of wavelengths.

The covariance matrix W was decomposed into singular values (according to the SVD algorithm), which allowed to determine eigenvalues and eigenvectors (U).

The contributions of the principal components were calculated using the matrix equation given below (42):

$$U = [V^T V]^{-1} V^T W \quad (42)$$

Conversely, the coordinates of the individuals in the space of principal components are calculated by:

$$V = WU[UU^T]^{-1} \quad (43)$$

V – the coordinates of the individuals in the space of principal components,

W – centred spectrum,

U – contributions of the principal components.

Matrices U and V were used to calculate the components of the spectra. If the matrix of the components of the spectra is denoted by C, it can be calculated as:

$$C = VU^T \quad (44)$$

Based on C, reconstructed spectra (D) were determined and then residual spectra (R) were calculated by:

$$R = W - D \quad (45)$$

The spectra are placed in a vector space, which is spanned by the principal components. These vectors are placed into a hyperplane with the number of dimensions equal to the number of pure spectral forms. This means that the first step is to determine the number of spectral forms. It can be done by the analysis of the eigenvalues and residual spectra. It is assumed that the rank of residual spectra at which the systematic changes of these spectra disappear defines the number of significant principal components (spectral forms). Subsequently, the points corresponding to the pure spectral forms need to be found over the principal component space.

Calculation of the X matrix allows to solve equation (46), which results in the B matrix:

$$W = BX \quad (46)$$

W – initial data matrix,

B – matrix of pure spectral forms,

X – estimation of the molar fraction matrix.

Values of the U matrix were optimized using the penalty function and using Nelder-Meada algorithm. Optimized values are given a penalty for violation of the constraints. It means that a value needs to be added until a minimum of the optimized function (FC) is reached.

$$penalty = \sum\sum x_{ij}^2 \qquad (47)$$

$$reward = \sum_i var(x_{ij}) \qquad (48)$$

$$FC = penalty - reward \qquad (49)$$

Based on the chosen number of significant eigenvectors, *n* principal component analysis was performed once again to extract (n-1)-dimensional hyperplane from n-dimensional space. Points from the X matrix were placed into a four-dimensional tetrahedron analogue. Corners of the given entities were found manually and then optimized using the simplex method (via Nelder-Meada algorithm). Molar fractions were obtained basing on contributions of the principal components and were calculated using coefficients of the plane equations (50, 51). A suitable plane was established using the multiple linear regression method by finding the best fit of the N-dimensional planes to the vertices of the simplex.

$$\sum_i d_i = h \qquad i \in \{R\} \qquad (50)$$

$$u_i = \frac{d_i}{h} \qquad (51)$$

h – height of a five dimensional simplex,

d_i – distance from u_i to the opposite plane of the simplex,

u_i – molar fractions.

4.1.2 Self-association study

Series of the mixtures containing UAs and HEPES buffer solutions (0.1 M) with following pH values: 2.5, 6.5, 7.4, 8.5 were prepared. The following UAs concentrations were analyzed: 10^{-6}, 2×10^{-6}, 5×10^{-6}, 10^{-5}, 2×10^{-5}, 5×10^{-5}, 10^{-4}, 2×10^{-4}, 5×10^{-4},

10^{-3} M. To determine whether UAs undergo `aggregation under conditions applied during enzymatic reactions, an appropriate sets of UAs UV-Vis spectra at various temperatures (4ºC, 22ºC and 35ºC) were collected. The optical pathways of a quartz cuvettes varied from 0.1 cm to 1 cm.

In the case of self-association, molar fractions were calculated in exactly the same way as for pKa determination.

4.2 Selection of the theoretical model and its parameters

4.2.1 Acid-base balance approach

Dissociation constants (pKa) were determined based on the molar fractions using an extended version of the Henderson-Hasselbalch equation (52 – 54).

Based on previously determined pKa values, theoretical molar fractions were calculated and compared with those determined experimentally.

$$\gamma = \frac{\prod_{i=1}^{n} K_i}{\prod_{i=1}^{n} H^i} \qquad (52)$$

$$u_1 = \frac{1}{1+\gamma} \qquad (53)$$

$$u_n = u_1 \cdot \gamma \qquad n = \{1,4\} \quad (54)$$

K – Dissociation constant,

u_n – Molar fractions.

In the next step, pKa values were optimized using the simplex optimization (via Nelder-Meada algorithm) and then were validated using *leave-one-out* cross-validation.

4.2.2 Self-association analysis

Dissociation constants (pKa) were determined based on the molar fractions, using an appropriate model of self-association. Three different models involving dimerization, unlimited self-association and micellar aggregation were considered. Mathematical formulas of employed models are presented in 3.2.2.2 Data were fitted to the models

using least squares method (with simplex optimization) and the best models were identified based on comparative analysis of residual sums of squares. In the next step, dissociation constants were validated using *leave-one-out* cross-validation.

All chemometric calculations were performed in the R environment (version 4.0.2) and python (version 3.8.10).

4.3 Identification of chemical structures of spectral forms

Series of mixtures containing UAs (1 mM) and an appropriate buffer solution or acid (0.01 M) with the following pH values: 1, 5, 7, 8, 11 were prepared. For pH = 1, hydrochloric or sulfuric acid was used. The solutions with pH equal to 5, 7 and 8 were prepared using a cacodylate buffer solution. Borate buffer was used for pH = 11. All mixtures contained 10% D_2O. The mixture for the TOCSY experiment was prepared in pure D_2O. Subsequently, the sets of 1D and 2D NMR (TOCSY) were recorded on a 700 MHz Bruker Avance III spectrometer equipped with a QCl CryoProbe (In collaboration with Institute of Bioorganic Chemistry, Polish Academy of Sciences). All spectra were proceeded using the Topspin software (version 4.1.3).

4.3.1 NMR study of the UAs complexes with G-quadruplexes

Sample preparations

G-quadruplexes were obtained from Metabion, GmbH (Germany) and additionally purified from triethylamine acetate by multiple spinning down, using Amicon centrifugal filters with 2 ml volume and 3 kDa cut-off. Purified G-quadruplex sequences were dissolved in 10 mM potassium cacodylic buffer of pH 5.0, containing 10 mM KCl. Subsequently, 20 µl D_2O was added. Total volume of sample was 200 µl. The series of spectra were collected (NOESY, TOCSY and HSQC). Afterwards, DNA samples were titrated with C-2045 or C-2053, dissolved in pure water.

NMR experiments

Spectra were collected on a 700 MHz Avance III spectrometer (Bruker) equipped with a QCl CryoProbe (in collaboration with the Institute of Bioorganic Chemistry, Polish

Academy of Sciences, Poznań, Poland). All spectra were processed with the Bruker TOPSPIN 4.0 software and NMRFAM-SPARKY suite[83].

4.3.2 Molecular modelling

Quantum chemical parametrization

The numerical models of C-2045 and C-2053 were built in the Avogadro software. Initial geometries were optimized using the MMFF94 force field. Then, the models were refined using GAUSSIAN09 software. Optimal geometry was found with Minnesota hybrid functionals (MN12SX) at the 6-31G* level of theory. In the next step, electronic charges were computed using a second-order Møller–Plesset perturbation theory (MP2). In parallel, basic parameters from CHARMM36 Generalized Force Field were derived using a computational tool available on *paramchem.org*. Models prepared in this way were later used in all the simulations[24,23,6].

Preparation of the simulation systems

Numerical model of G-quadruplex was taken from Protein Data Bank (PDB), model number 2L7V[47]. Based on the crucial correlation signals between ligand and DNA protons, found in the NOESY spectra, an appropriate ligand (C-2045 or C-2053) was placed manually in such a way that the protons of the ligand were located in space near the corresponding protons of the DNA. These models of the complexes were built using VMD v1.9.3. software.

Simulation procedures

Resulting G-quadruplex/ligand complexes were placed in cubic boxes with the size of 1.2 Å. Subsequently, all built systems were minimized in vacuum to find the initial minimum of potential energy function. The simulation boxes were filled with water (spc216 model). The energy was again minimized, and in the next step, the environment was neutralized by adding an appropriate amount of potassium and chlorine ions (two potassium ions were kept in the center of G-quadruplexes during all computations). From a statistical thermodynamic point of view, all simulations were performed in the NPT ensemble, i.e., maintaining constant pressure and temperature. For the molecular dynamics, the CHARMM36 force field and the GROMACS package equipped with the PLUMED plug-in were used. To maintain the continuity

of the simulated system, periodic boundary conditions were applied in all three dimensions. The temperature of the systems during the simulation was 310 K while the pressure was kept at 1 bar using the Berendsen barostat. The equations of motion were integrated using Verlet's algorithm with a time step of 2 fs[6,5,23].

Umbrella sampling

The Hamiltonian Replica Exchange Umbrella Sampling method was employed to obtain free energy profiles of DNA/ligand interactions. As a reaction coordinate, the distance ξ was chosen, defined as the distance between the centers of mass determined for the upper and lower G4 tetrads and the imidazoacridinone rings of the upper and lower ligand, respectively.

The initial structures were selected from the trajectories of additional 100 ns simulations, in which the reaction coordinate (DNA/ligand ξ distance) was increased from 0.4 to 1.5 nm, due to the applied harmonic potential with a force constant of 2500 kJ/mol·Å2.

In each case, the reaction coordinate was divided into 12 equal windows, spaced every 0.1 nm. Each window was then simulated for 1000 ns with the application of Parrinello-Rahman barostat at a temperature of 318 K. The weighted histogram method (WHAM) was used to obtain the free energy profiles[84,85].

5 RESULTS

5.1 Acid-base equilibrium of UAs

Considering that there are hardly any reported results regarding the physicochemical properties of UAs, this research began with investigation of some of them. The first feature taken into account was the proton dissociation constants. It is crucial to know the protonation state of a potential drug under certain conditions. Firstly, this knowledge is important since the compounds express different protonation forms under various pH conditions, what is crucial for their solubility under various physiological conditions. Moreover, considering the compounds' interactions with DNA, their protonation level also plays a crucial function. UV-Vis spectroscopy was chosen as the most suitable method for the studies on drug protonation and dissociation. All UAs are aromatic compounds, so they are characterized by high absorption in the wide range of wavelengths, also in the visible region. The absorption values above 350 nm are especially interesting, because the electromagnetic radiation in that region is less energetic than in the UV region. This allows to avoid unnecessary noise and some other problems with further data analysis.

As the first step, UAs spectra sets were recorded under the conditions described in section 4.1. (Figure 28).

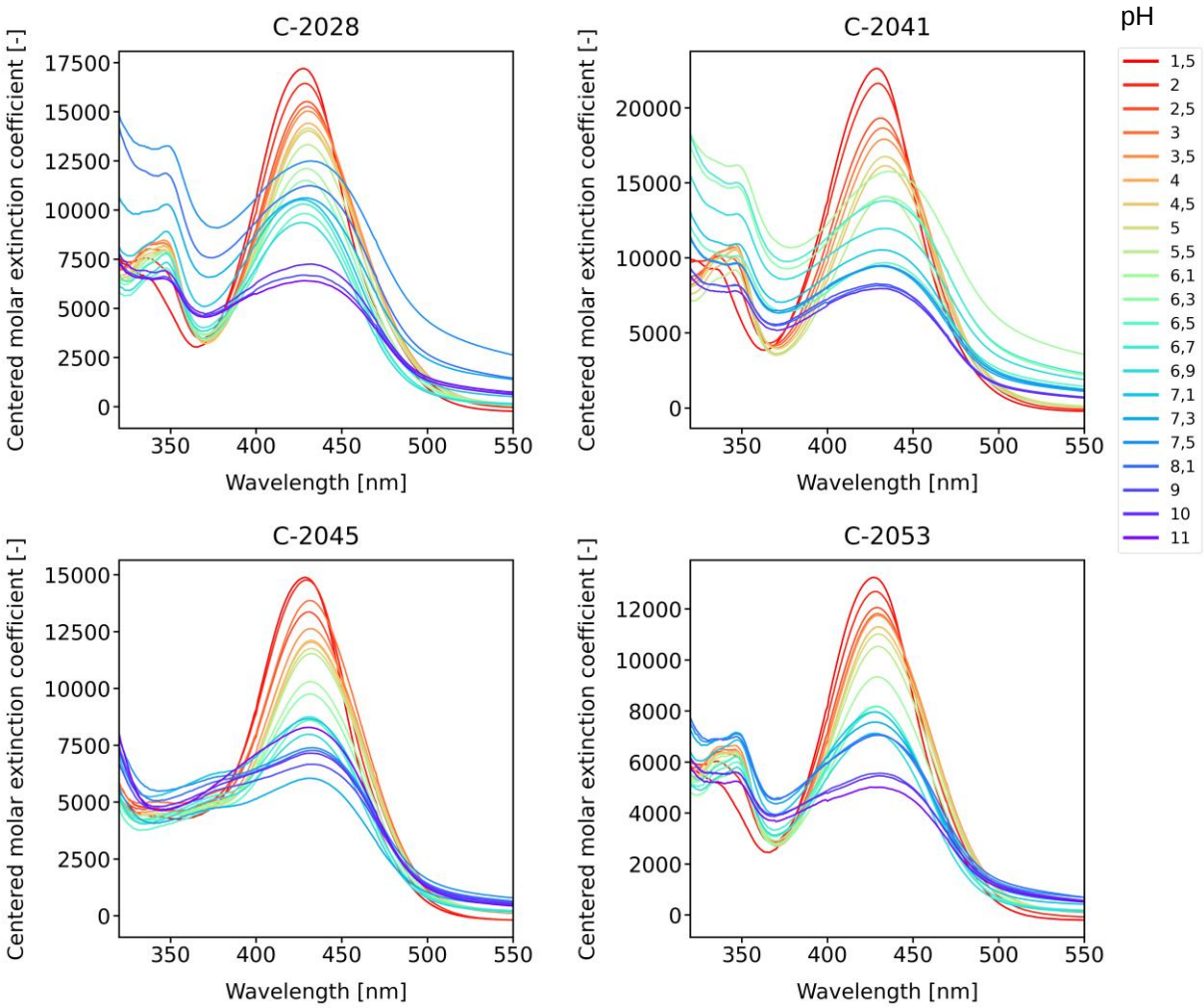


Fig. 28. Spectra in the pH range from 1.5 to 3.5 were recorded in hydrochloric acid. Spectra of the UAs in pH range from 4 to 11 recorded in an appropriate buffer solution as follows: acetate buffer, HEPES, PIPES and boric buffer.

As can be seen, the spectra above (initial spectra) are characterized by significant changes in the shape, corresponding to the increase of pH values. It indicates the differentiation in electron density distribution over the system, which probably is mainly related to the proton detachment. These data were then analyzed numerically. Firstly, the obtained spectra were prepared for further analysis. For this purpose, all spectra sets were presented as a matrix, containing wavelengths in rows and particular spectra at different pH values in columns. All columns were centered, with the mean value equal to 0. Centering of spectra operation allows to investigate changes in shape only, yet not in intensity (Figure 29).

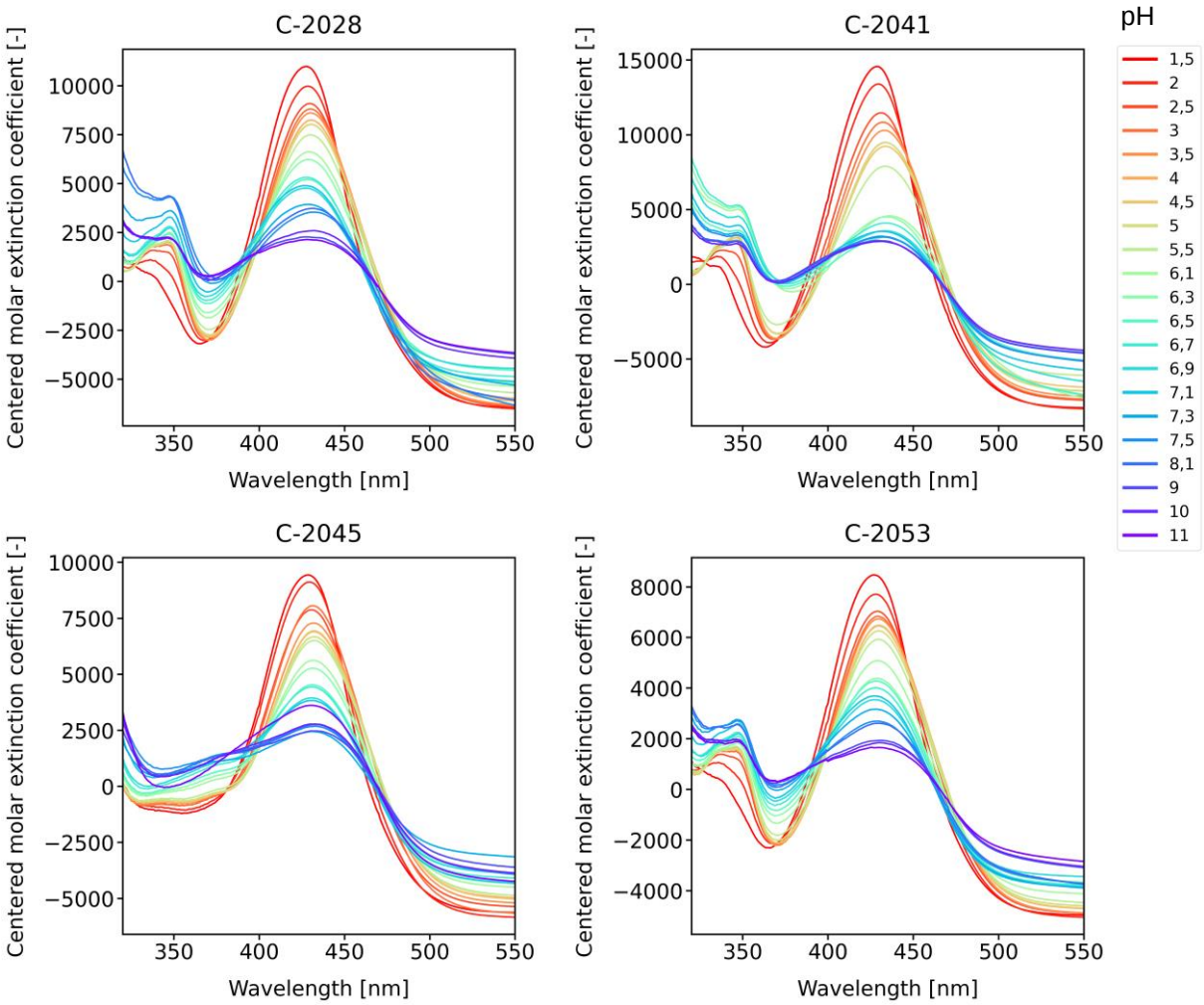


Fig. 29. Spectra in the pH range from 1.5 to 3.5 were recorded in hydrochloric acid. Centred spectra of the UAs in pH range from 4 to 11 recorded in an appropriate buffer solution as follows: acetate buffer, HEPES, PIPES and boric buffer.

The quality of the collected data has been assessed on the basis of the correlation coefficients. Each set of spectra was numerically decomposed into eigenvalues and eigenvectors, which were further used as the first approximation of the molar fractions of a particular spectral form. The number of valid spectra was chosen mainly based on residual spectra analysis (Figures 30, 32, 34, 36).

5.1.1 C-2028 protonation forms

To decide how many eigenvectors should be taken into equation (46), residual spectra analysis was performed (Figure 30). The number of pure spectral forms corresponds to the different protonation forms of UAs. The first and the second rank spectra have a regular shape and quite high intensity. The third and fourth rank spectra have a lower intensity, but the regularity in their shape is visible. The fifth rank spectrum has very

low intensity, but it still has the character of a spectrum and it is not as noisy as the last one. Another symptom, that the fifth principal component is important, are the intersections of all spectra at several wavelengths. However, the systematic changes vanished at the sixth rank spectrum, which indicates that six pure spectral forms can be present in the studied solutions. An attempt to analyze the six principal components was taken but it failed, so finally five principal components were chosen.

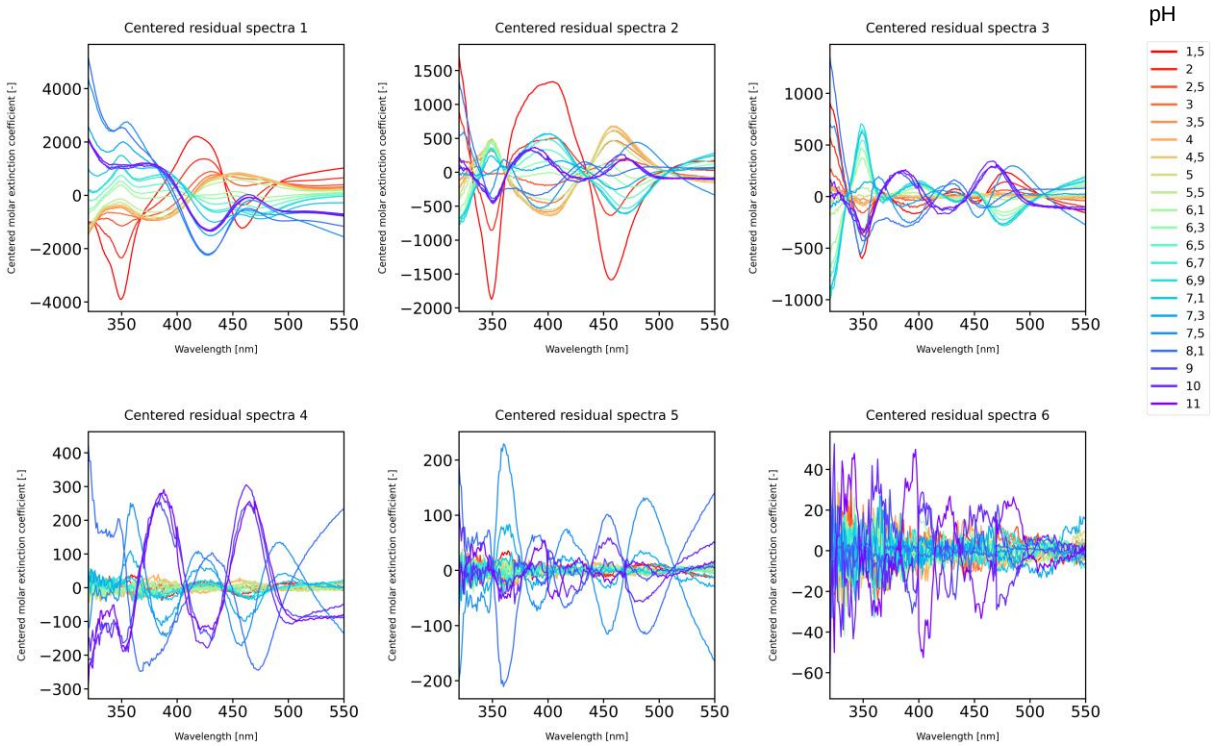


Fig. 30. Residual spectra of C-2028 ($2x10^{-5}$ M); pH range from 1.5 to 11 (ionic strength ~ 0.02 M).

Taking into account the selected eigenvectors, the molar fractions, spectral pure components, and pKa values were calculated (Figure 31). Points obtained from an experiment differed slightly from those obtained from the calculated pKa values. Differences arose from numerical inaccuracy and experimental errors. Potential aggregation might also play a role. Molar fractions were used to determine the pure component spectra. Interestingly, it turned out that the spectra of the pure components were similar to each other. The maxima for each band were almost the same. They only differentiated in the intensity and the band width. In the case of the first band, of a maximum at 350 nm, the hyperchromic effect was observed as the pH increased. Conversely, for the band with a maximum at 430 nm, a hypochromic effect was visible.

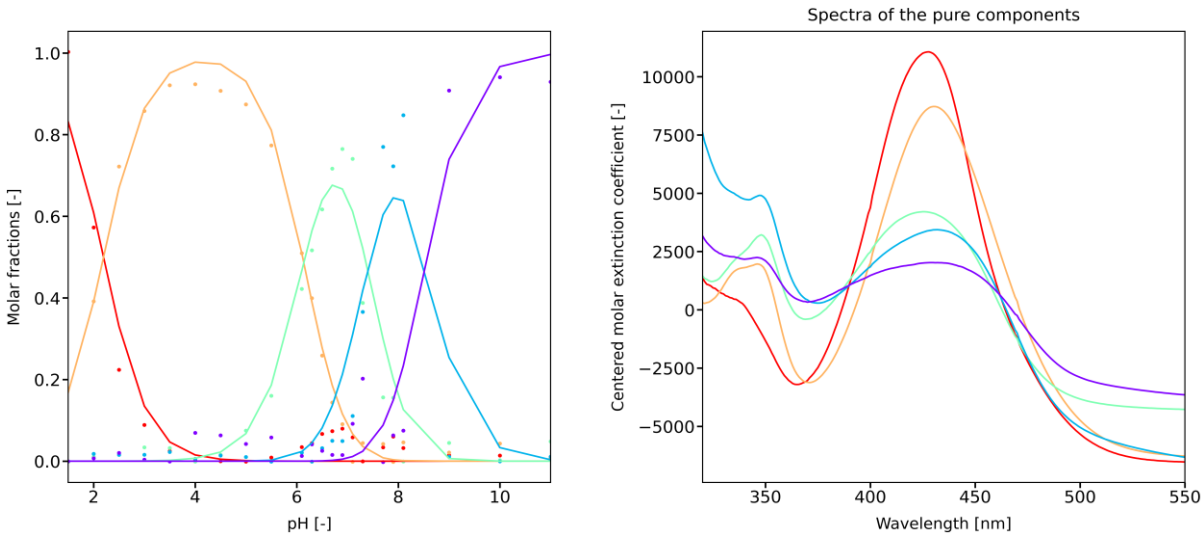


Fig. 31. C-2028 molar fractions (dotted lines – optimal molar fractions calculated based on the data; solid line – the best fit of the theoretical model to the data) - left, spectra of the pure components – right. Each color represents a different spectral form. Red corresponds to the form at the lowest pH and purple corresponds to the form at the highest pH.

5.1.2 C-2041 protonation forms

The results clearly indicated that the four eigenvectors are valid. The fifth rank spectra are very noisy and there are no points of intersection. Hence, these spectra presumably consist only of noise and measurement errors (Figure 32).

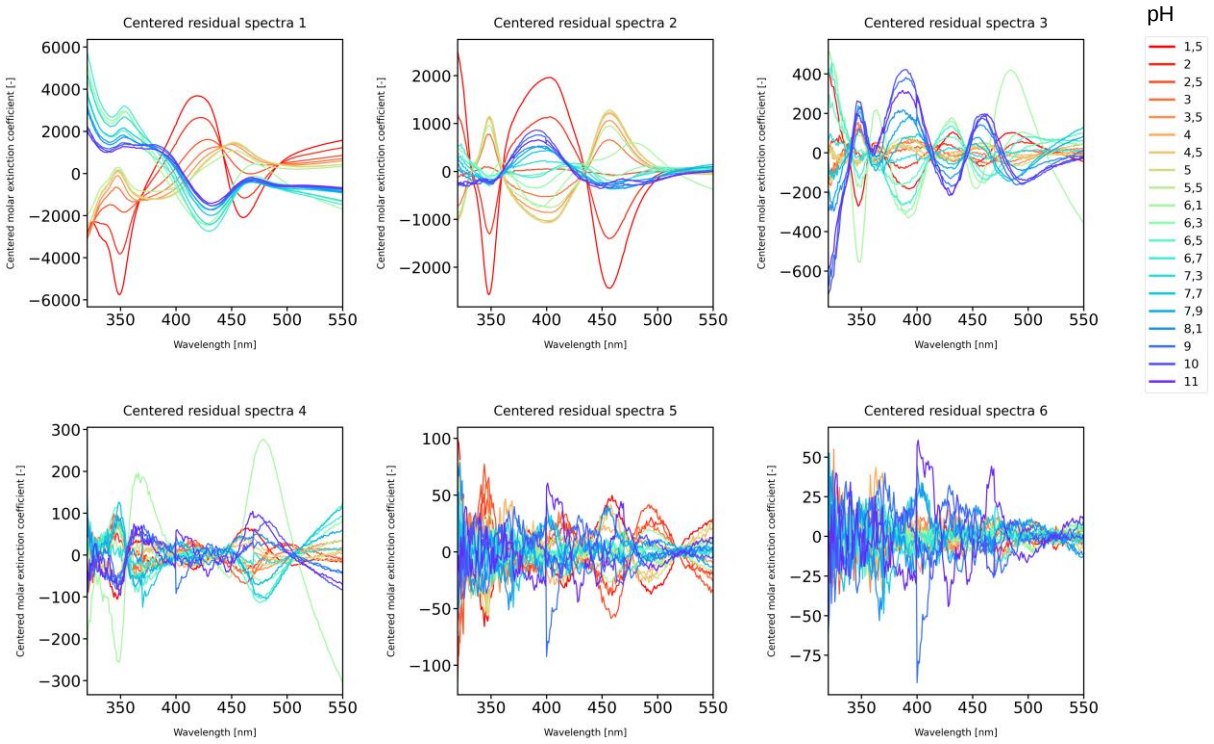


Fig. 32. Residual spectra of C-2041 ($2x10^{-5}$ M); pH range from 1.5 to 11 (ionic strength ~ 0.02 M).

C-2041 compound seems to be a special one due to the fact that the piperazine ring is present in the linker. This is probably the reason why the pure component spectra, as well as pKa values, are so different in comparison to C-2028 and to other UAs. Based on chemical knowledge, four dissociation constants may be expected, but in this case only three pKa values are “visible”. If two spectra are too similar to each other, performing a reliable chemometric analysis is troublesome. The compound C-2041, which has two nitrogen atoms in the piperazine ring, serves as a good example of such a conundrum. The protonation of the aforementioned two nitrogen atoms is indistinguishable in the UV-Vis spectrum. Considering three dissociation constants, the theoretical model is fitted properly (Figure 33).

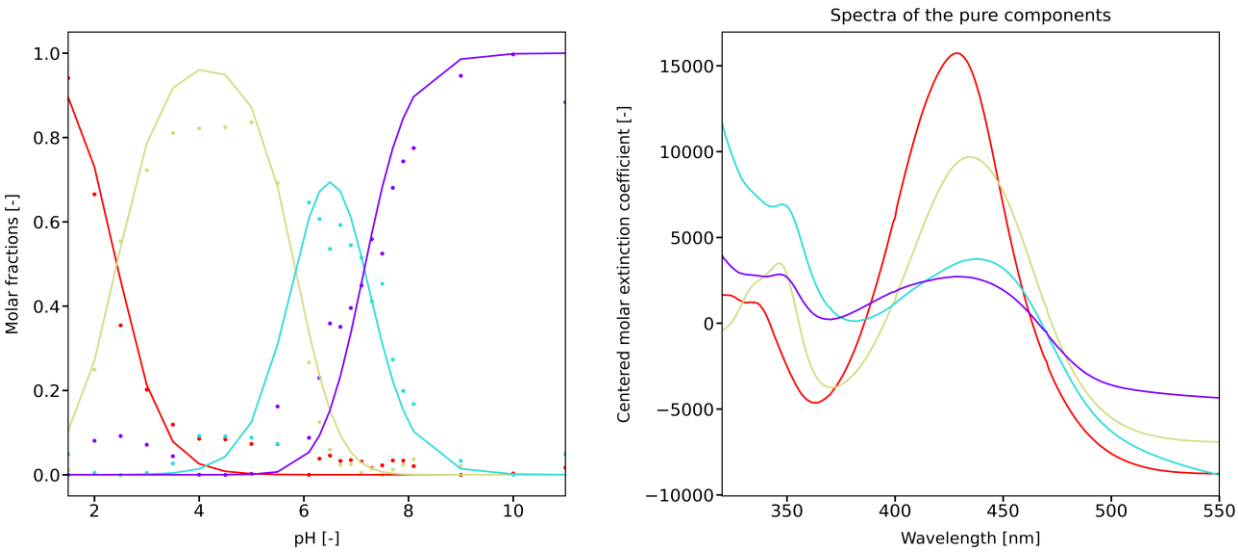


Fig. 33. C-2041 molar fractions (dotted lines – optimal molar fractions calculated based on the data; solid line – the best fit of the theoretical model to the data) - left, spectra of the pure components – right. Each color represents a different spectral form. Red corresponds to the form at the lowest pH and purple corresponds to the form at the highest pH.

5.1.3 C-2045 protonation forms

Residual spectra analysis once again clearly indicated that four eigenvalues were valid (Figure 34). All the spectra have reasonable intensity and regularity. The regularity almost vanished at the fifth rank spectrum, which is noisy and has low intensity. However, there are points of intersection in the fourth rank residual spectra, which suggest that also fifth rank spectra should be taken into consideration. This is enough for the acceptance of the fifth eigenvector. Again, it is easy to see that the last spectra are all noisy and hardly exhibit any regularity. Further analysis confirmed that such a choice was appropriate.

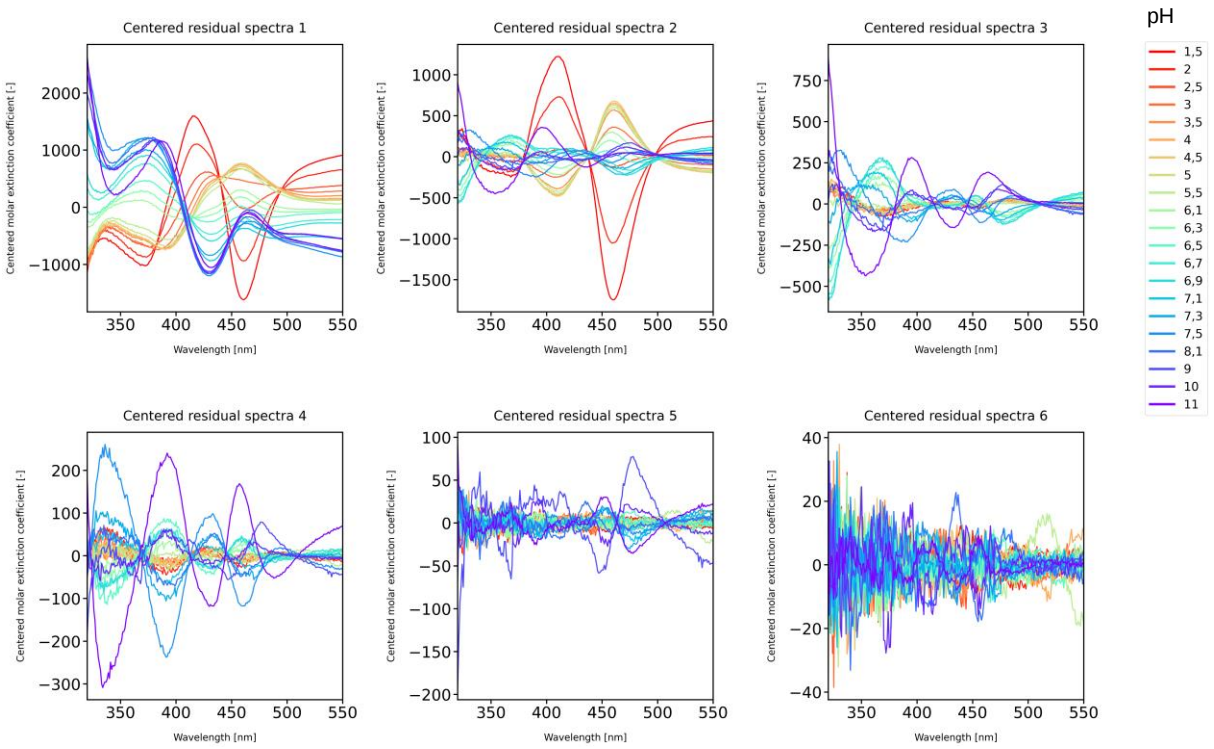


Fig. 34. Residual spectra of C-2045 ($2x10^{-5}$ M); pH range from 1.5 to 11 (ionic strength ~ 0.02 M).

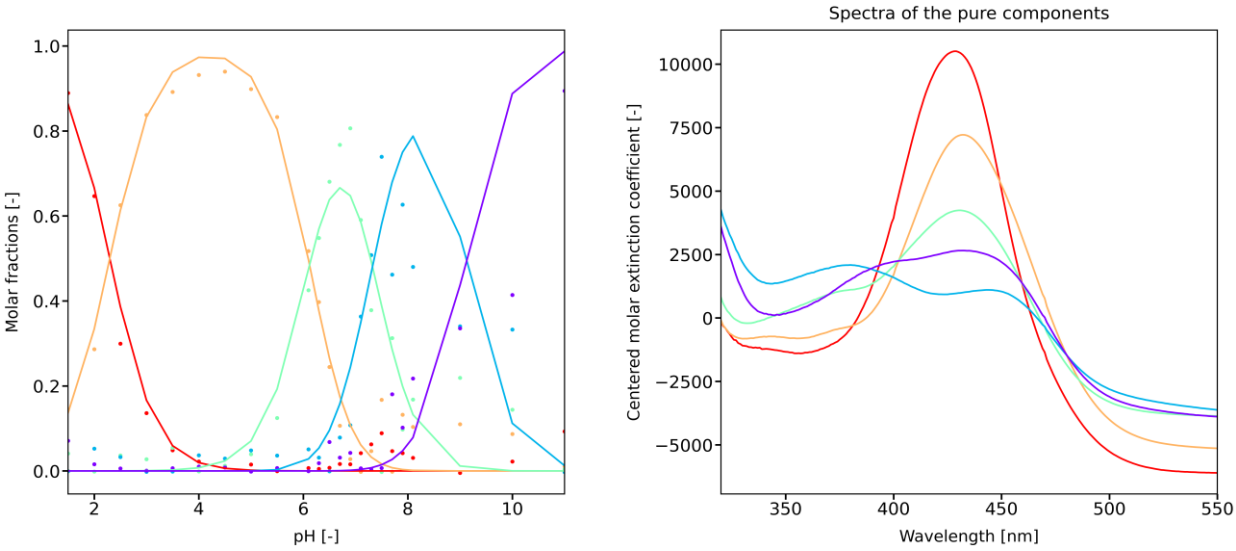


Fig. 35. C-2045 molar fractions (dotted lines – optimal molar fractions calculated based on the data; solid line – the best fit of the theoretical model to the data) - left, spectra of the pure – right. Each color represents a different spectral form. Red corresponds to the form at the lowest pH and purple corresponds to the form at the highest pH.

5.1.4 C-2053 protonation forms

The residual spectra analysis for C-2053 again showed that five eigenvalues are valid (Figure 36). The fifth residual spectrum is very noisy and it is difficult to see any regularity in it. The intensity is also very low. However, the fourth rank spectra of the C-2053 have points of intersection, which are visible despite the strong noise. Considering all above, five eigenvectors were taken for further analysis.

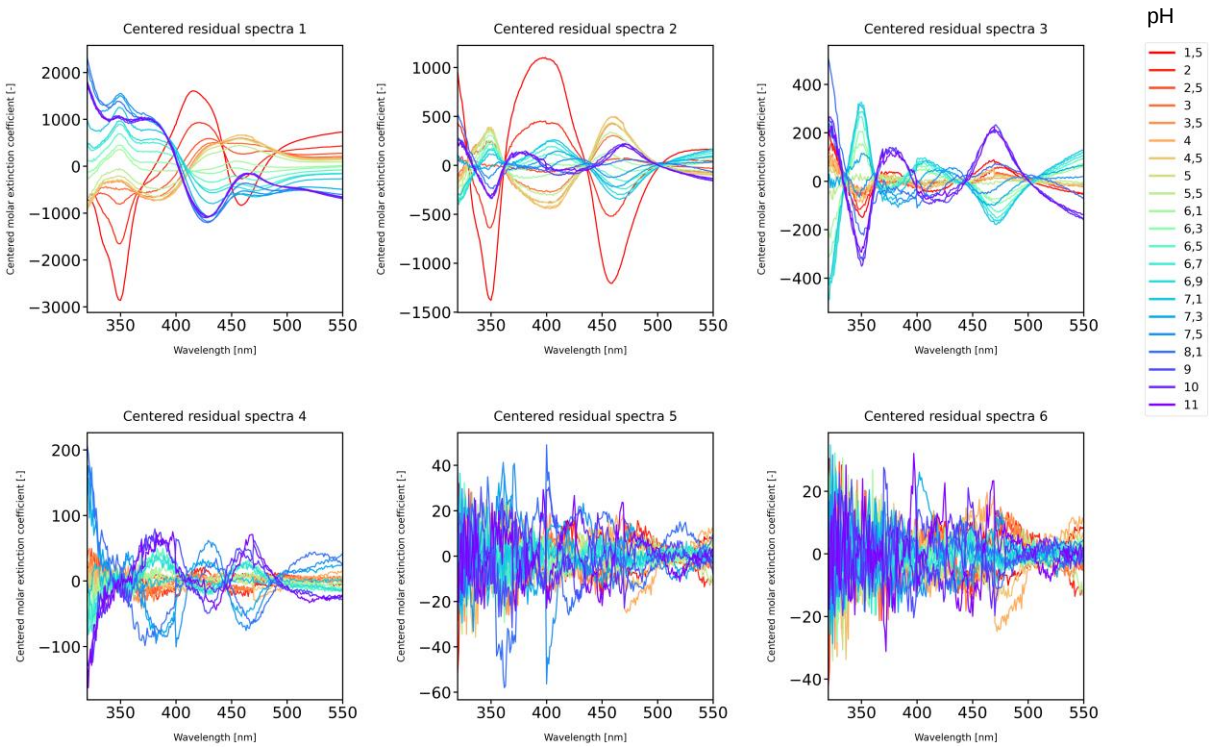


Fig. 36. Residual spectra of C-2053 ($2x10^{-5}$ M); pH range from 1.5 to 11 (ionic strength ~ 0.02 M).

Molar fractions and spectra of the pure components are very similar to those obtained for C-2028. The only difference between C-2053 and C-2028 is that the latter lacks methyl group in the 4th position of the 1-nitroacridine ring, hence one should not be confused by their spectral similarity (Figure 37).

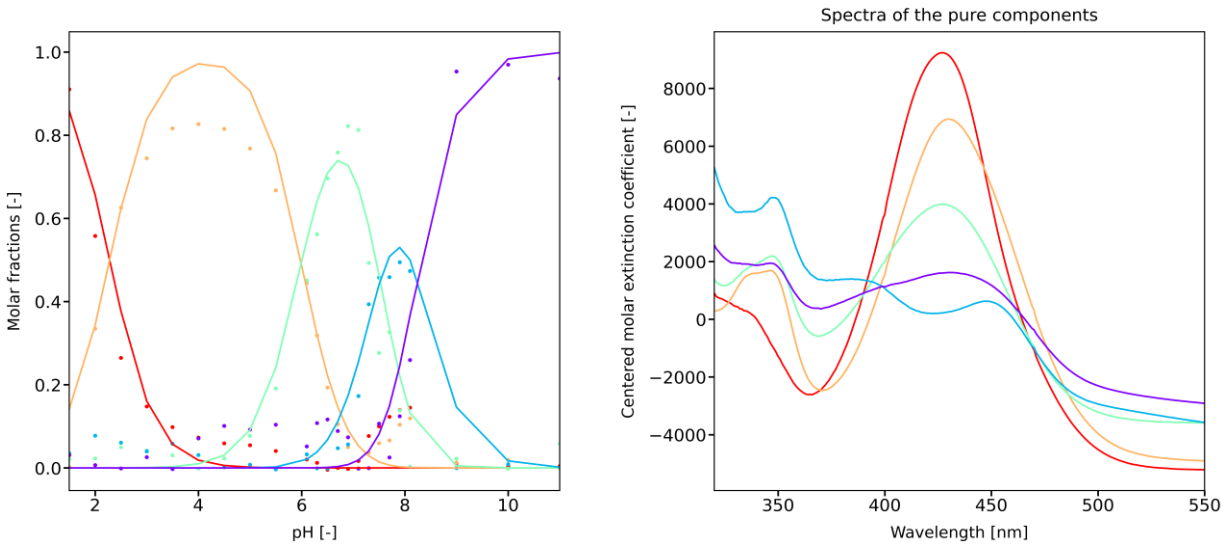


Fig. 37. C-2053 molar fractions (dotted lines – optimal molar fractions calculated based on the data; solid line – the best fit of the theoretical model to the data) - left, spectra of the pure components – right. Each color represents a different spectral form. Red corresponds to the form at the lowest pH and purple corresponds to the form at the highest pH.

5.1.5 Dissociation constants of UAs structural elements

To initially assess the order of the proton detachment for UAs, an additional experiment involving UAs analogues was conducted. Three compounds, which essentially are the monoacridine fragments of UAs, were chosen: 1) C-1311, 2) 8-dehydroxy imidazoacridinone structural analogue of 1311 and 3) Ledakrin (Figure 38). C-1311 and its 8-dehydroxy analogue have allowed to find the pH range in which the hydroxyl group of imidazoacridinone moiety dissociates. Ledakrin was selected to find which UAs pKa value corresponds to 1-nitroacridine ring.

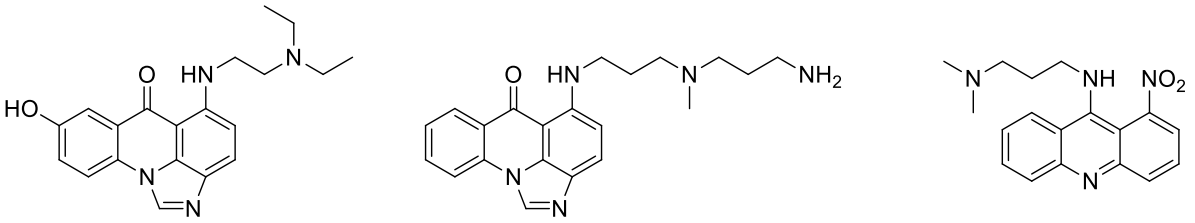


Fig. 38. C-1311 (left), ledakrin (right) and structural analogue of C1311 (in the middle).

C-1311

C-1311 differs only in the aminoalkyl chain in comparison to the imidazoacridinone part of C-2045. Dissociation constants were determined according to the same

methodology as it was previously described for UAs. The results are presented below (Figure 39, 40).

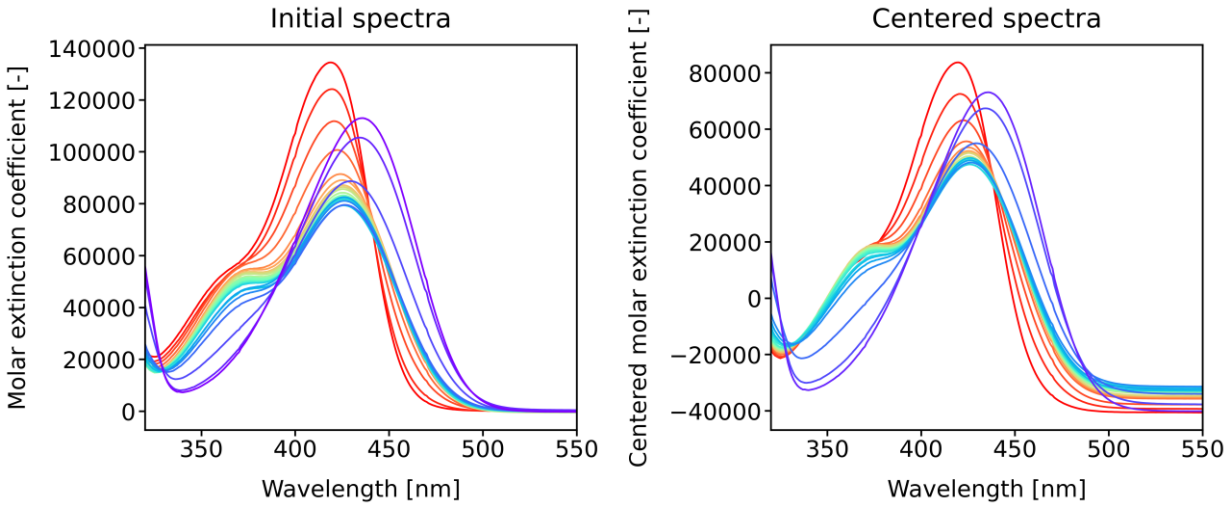


Fig. 39. The initial spectra of the C-1311 (left) and the centered spectra of the C-1311 (right); conditions: (C-1311 concentration = $2x10^{-5}$ M), recorded in hydrochloric acid in the pH range from 1.5 to 3.5. Spectra in the pH range from 4 to 11 were recorded in an appropriate buffer solution as follows: acetate buffer, HEPES, PIPES and boric buffer (ionic strength ~ 0.02 M).

The spectra presented above display at least one isosbestic point, which shows that there are more than one spectral forms in the solutions. In order to precisely determine the pKa values, PCA was performed, which resulted in the residual spectra (Figure 40).

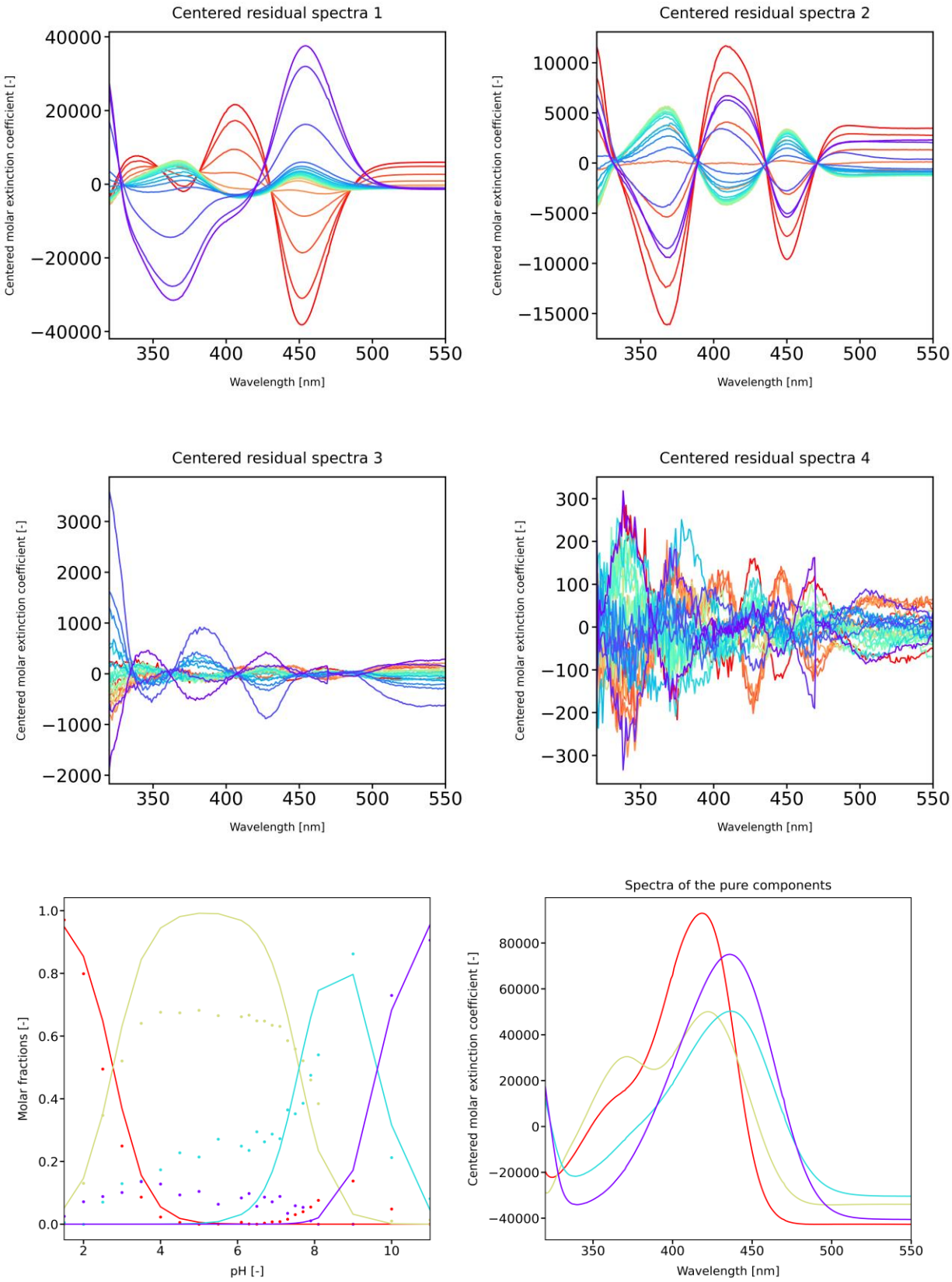


Fig. 40. Residual spectra of C-1311 ($2x10^{-5}$ M); pH range from 1.5 to 11 (ionic strength ~ 0.02 M). – upper; molar fractions and spectra of the pure components – lower. Each color represents a different spectral form. Red corresponds to the form at the lowest pH and purple corresponds to the form at the highest pH.

Residual spectra analysis indicated that there are four valid eigenvectors for C-1311, because the regularities vanish at fourth rank residual spectrum. Dissociation constants were determined according to the methodology described in paragraph 4.2.1. Results suggest that the first pKa value of all UAs is associated with the imidazoacridinone ring. The pKa equal to 9.675 probably corresponds to the dissociation of the hydroxyl group, due to the lack of this value for the analogue of 1311 described below. Moreover, pKa value of a pure phenol is ~10, which is coherent with the obtained results. The value equal to 7.597 can be assigned to the imidazole part of the ring mainly by elimination. Dissociation of the proton corresponding to the tertiary amine in the side chain probably does not affect the aromatic ring of C-1311 strongly enough to impact the UV-Vis spectrum. Conversely, in the case of UAs, the tertiary amine can be visible because of the presence of the second aromatic ring.

Structural analogue of C-1311

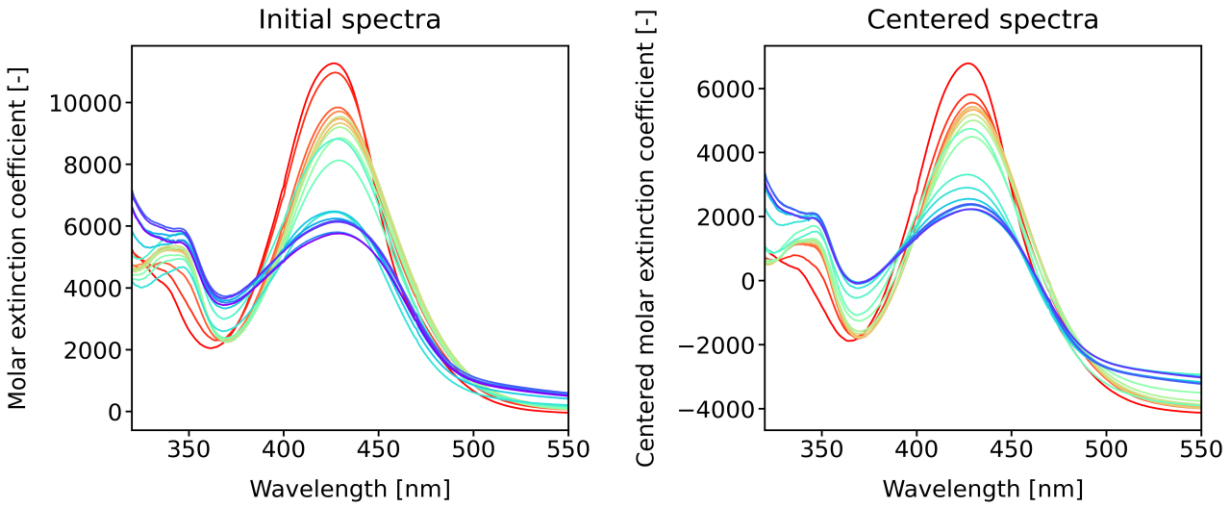


Fig. 41. The initial spectra of the structural analogue of C-1311 (left) and the centered spectra of the structural analogue of C-1311 (right); conditions: concentration = $2x10^{-5}$ M, recorded in hydrochloric acid in the pH range from 1.5 to 3.5. Spectra in the pH range from 4 to 11 were recorded in an appropriate buffer solution as follows: acetate buffer, HEPES, PIPES and boric buffer (ionic strength ~ 0.02 M).

At least one isosbestic point is visible in the spectra above, which indicates at least two spectral forms (Figure 41). To precisely determine the pKa values, PCA was performed according to the procedure presented above, which resulted in the residual spectra (Figure 42).

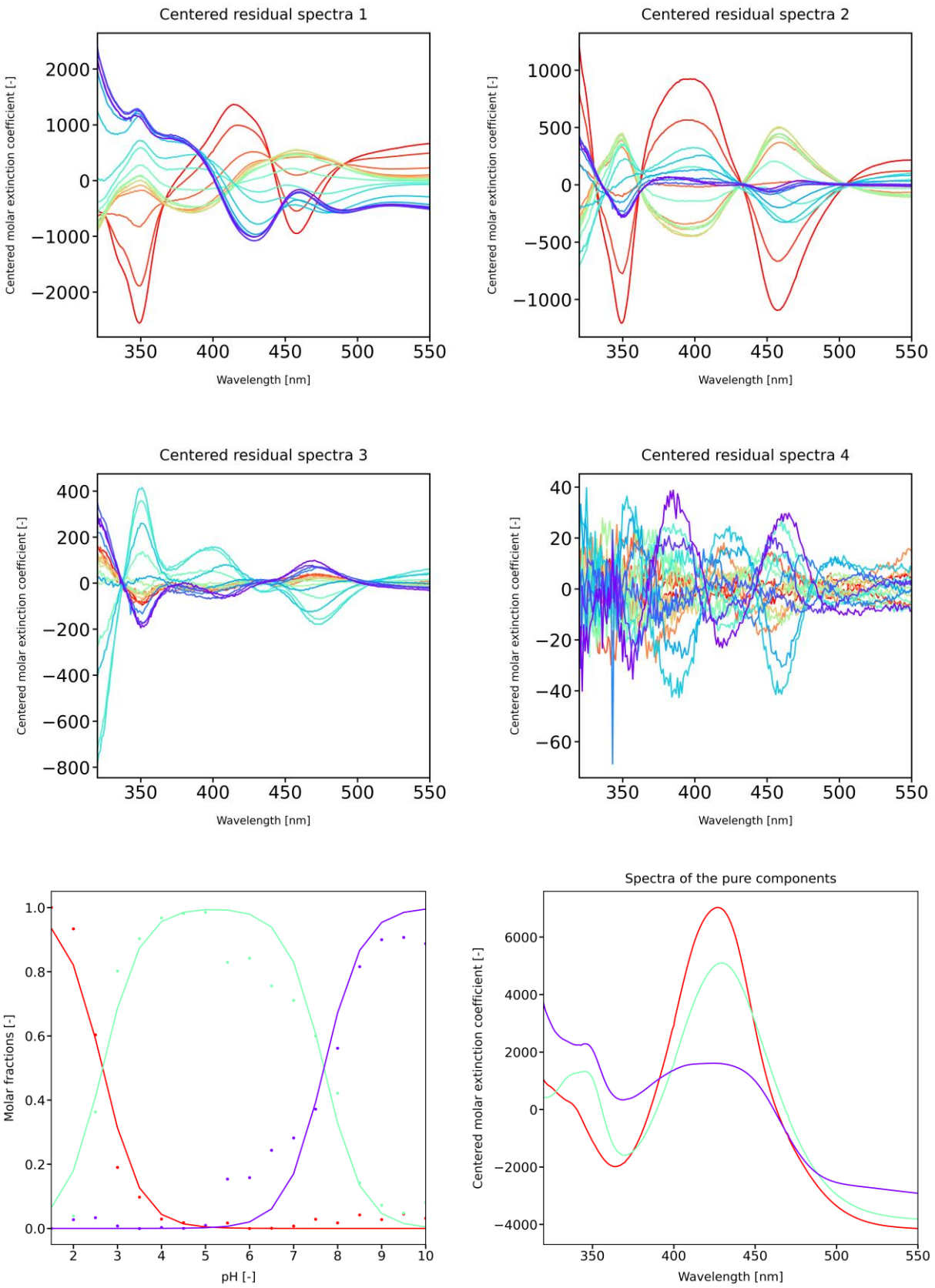


Fig. 42. Residual spectra of imidazoacridinone analogue of C-2028 ($2x10^{-5}$ M); pH range from 1.5 to 11 (ionic strength ~ 0.02 M) – upper; molar fractions and spectra of the pure components – lower. Each color represents to a different spectral form. Red corresponds to the form at the lowest pH and purple corresponds to the form at the highest pH.

From the figure above, it can be seen that the first four residual spectra should also be taken into account because regularities vanish at fourth rank residual spectra. However, an attempt to use the fourth component revealed that the changes seen in the fourth residual spectra are artifacts. The main difference in the structure of C-1311 and the presented analogue is the absence of the hydroxyl group, which means that there is no pKa value around 9.5. Based on the above analysis, it may be decided that the pKa value of the hydroxyl group in C-2045 is around 9.5, due to the similarity of its imidazoacridinone moiety to C-1311.

Ledakrin

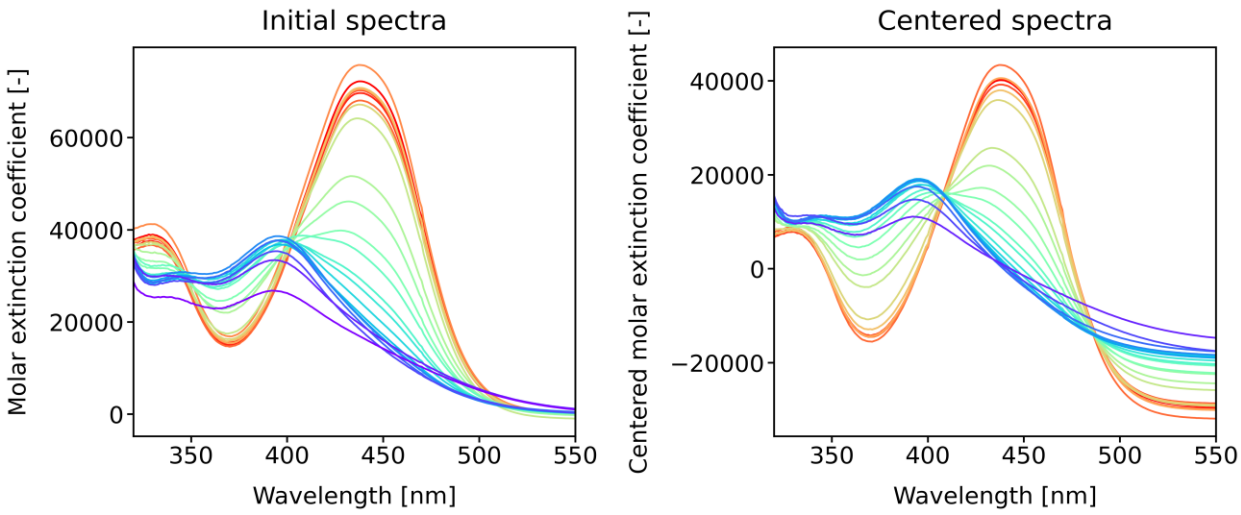


Fig. 43. The initial spectra of the ledakrin (left) and the centred spectra of the ledakrin (right); conditions: (ledakrin concentration = $2x10^{-5}$ M), recorded in hydrochloric acid in the pH range from 1.5 to 3.5. Spectra in the pH range from 4 to 11 were recorded in an appropriate buffer solution as follows: acetate buffer, HEPES, PIPES and boric buffer (ionic strength ~ 0.02 M).

UV-Vis spectra of Ledakrin indicate at least two spectral forms (Figure 43). Once again, for precise determination of the pKa values, PCA was performed (Figure 44).

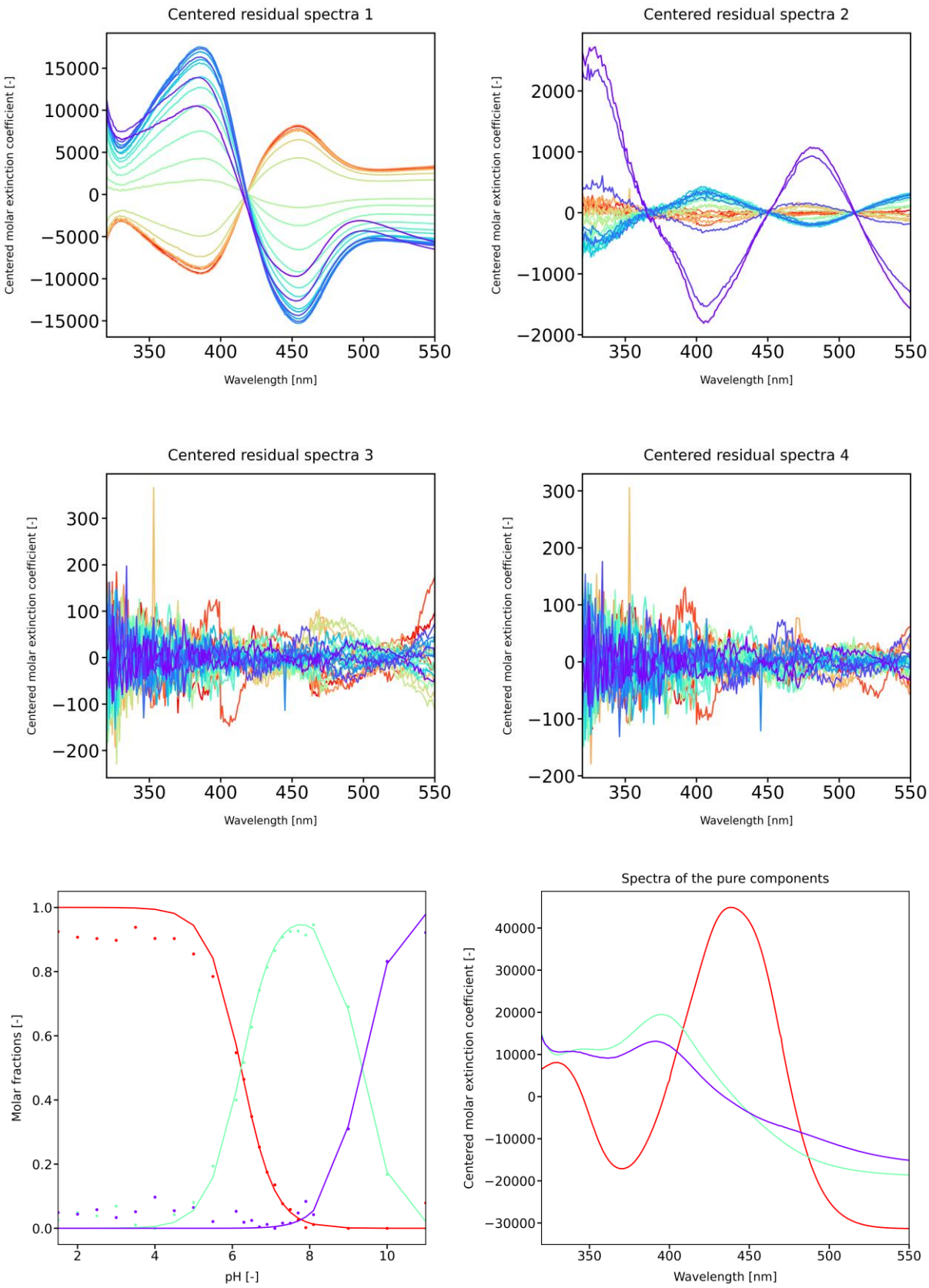


Fig. 44. Residual spectra of Ledakrin ($2x10^{-5}$ M); pH range from 1.5 to 11 (ionic strength ~ 0.02 M) – upper; molar fractions and spectra of the pure components - lower. Each color represents to a different spectral form. Red corresponds to the form at the lowest pH and purple corresponds to the form at the highest pH.

Analysis of Ledakrin residual spectra suggested that there is rather one pKa value than two. The third eigenvector was taken into account to simplify the computations, since it is virtually impossible to solve the equation (46) for two components. The third spectral form appears to be false, because at high pH Ledakrin becomes insoluble and the change in the spectrum shape is a result of the dispersion effects. Presumably, there is only one pKa value equal to 6.227 arising from the spectra (Figure 44). The tertiary amine group dissociation in the side chain is probably not seen for the same reason as in the case of C-1311.

5.1.6 NMR study on UAs protonation forms

The results presented above have revealed the number of dissociation constants and corresponding pKa values. However, the proton dissociation order still remained a partial mystery. Some predictions were done by the comparative analysis of structural analogues. Nevertheless, for this purpose, the most reliable experiment is nuclear magnetic resonance (NMR) analysis. The first step was the localization of the signals corresponding to the exchangeable protons, which was done by total correlation spectroscopy (TOCSY). Subsequently, a set of one dimensional NMR experiments for all UAs at four selected pH values was registered. The lowest possible and the highest possible pH were chosen to study fully protonated and fully deprotonated UAs structures. The pH values equal to 5 and 7 were selected based on the previously performed UV-Vis analysis, as these values correspond to the compositions of the proper mixtures, in which one of the spectral forms exists in majority.

The exchangeable protons were identified using TOCSY spectrum. The aim of the experiment was to find peaks showing couplings between UAs' protons and water. The figure below (Figure 45) shows that there are three UAs protons, which are coupled with H_2O. This result indicates that there are at least three exchangeable protons. Other protons might remain "invisible" due to chemical exchange ratio.

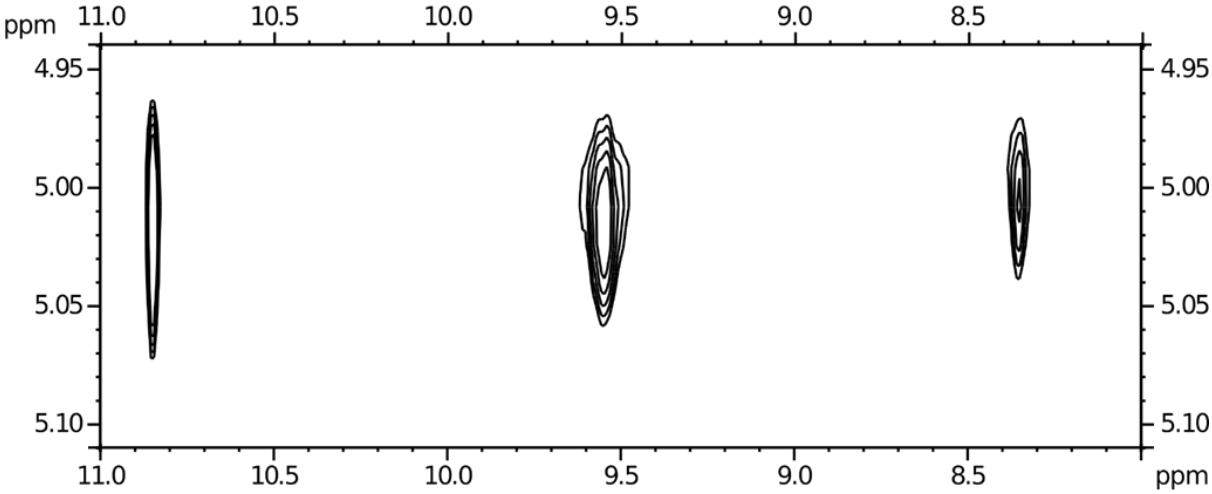


Fig. 45. TOCSY spectrum of the C-2045 (pH = 5, temp. = 5ºC, H_2O/D_2O) collected in 10 mM cacodylate buffer.

The next step was the identification of the exchangeable protons. It has been done by collecting TOCSY spectrum at the lowest possible pH at 5ºC. Figure 45 displays that a signal around 8.35 ppm shows correlation to the protons of low chemical shift. These protons relate to the methylene groups in the side chain of C-2045, so the proton corresponding to the signal around 8.35 is connected to the aniline group in the imidazoacridinone ring. Interestingly, this proton has quite high chemical shift, presumably because of the hydrogen bond formed between the aniline group and the ketone function embedded in the imidazoacridinone ring. The signal around 9.55 arises from the proton attached to the tertiary amine group in the side chain because there is only one correlation signal with the peak of integration equal to 3. Besides, there are no other correlations, which confirms the aforementioned assignment. The signal at almost 11 ppm is presumably connected to the proton attached to the 1-nitroacridine fragment (i.e., to the nitrogen atom built into the acridine ring). It is the only one proton which cannot display any correlation signals at the TOCSY spectrum, because there are no other neighboring protons within the structure (Figure 46). Moreover, ROESY spectrum recorded at pH 1 displayed a dipolar coupling between this proton and the methyl group at the position 4 of the nitroacridine moiety, which ultimately proved the discussed assignment (Figure 47).

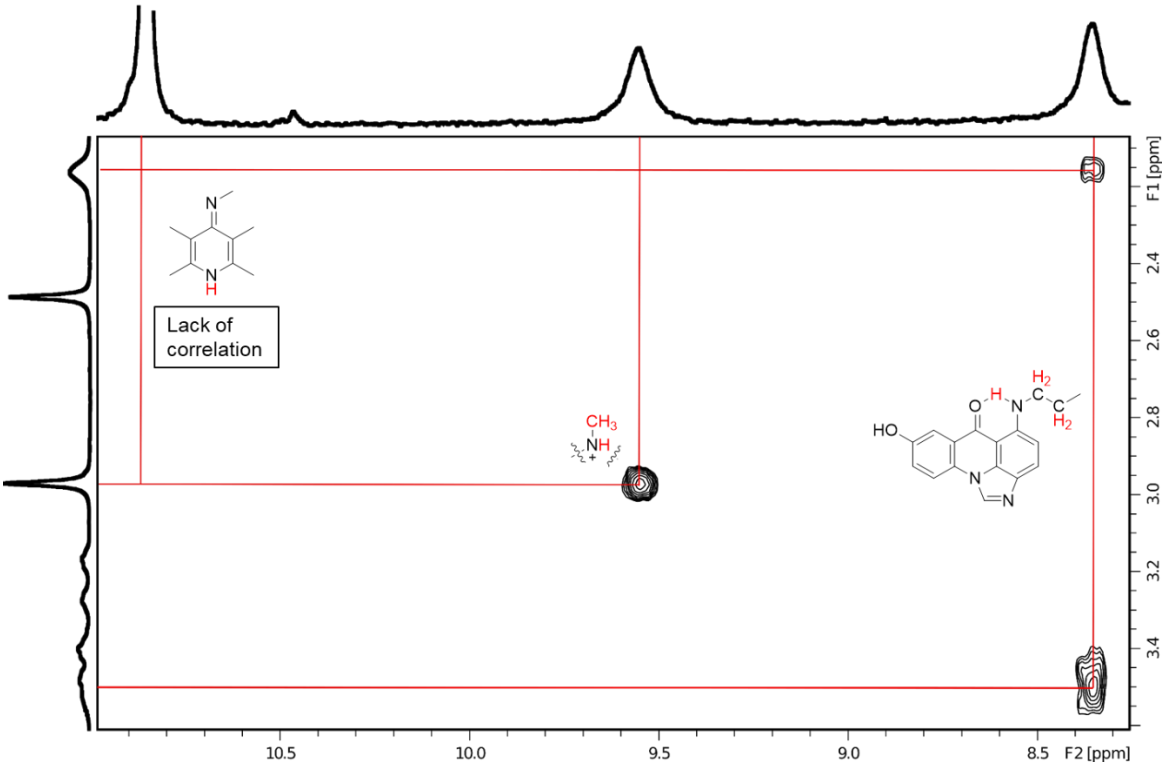


Fig. 46. TOCSY spectrum of the C-2045 (pH = 4.5, temp. = 5ºC, H_2O/D_2O) collected in 10 mM H_2SO_4.

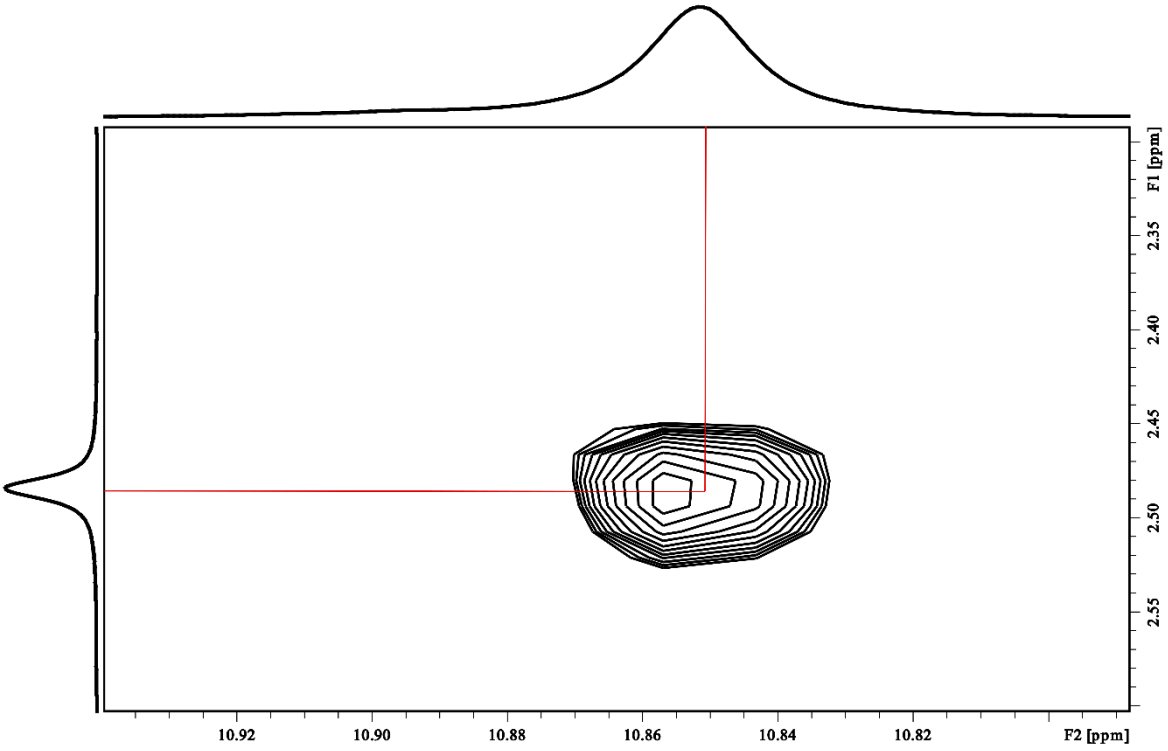


Fig. 47. Fragment of the ROESY spectrum of the C-2045 (pH = 1, temp. = 5ºC, H_2O/D_2O) collected in 10 mM H_2SO_4.

Spectrum recorded at pH = 1 revealed two signals in the imine region which disappeared at pH = 7. Based on that observation and on the previously described TOCSY spectra. The pKa values ranging from 5.8 to 6.14 were assigned to the nitrogen atom of the 1-nitroacridine ring. Noteworthy is the fact, that the amino group at position 9 of the 1-nitro-9-aminoacridine ring is not protonated at lower pH values, but it exists as an imine nitrogen. Protons bound to the amino group of imidazoacridinone ring and to the imidazole part of this ring are not visible in NMR spectra (Figure 48).

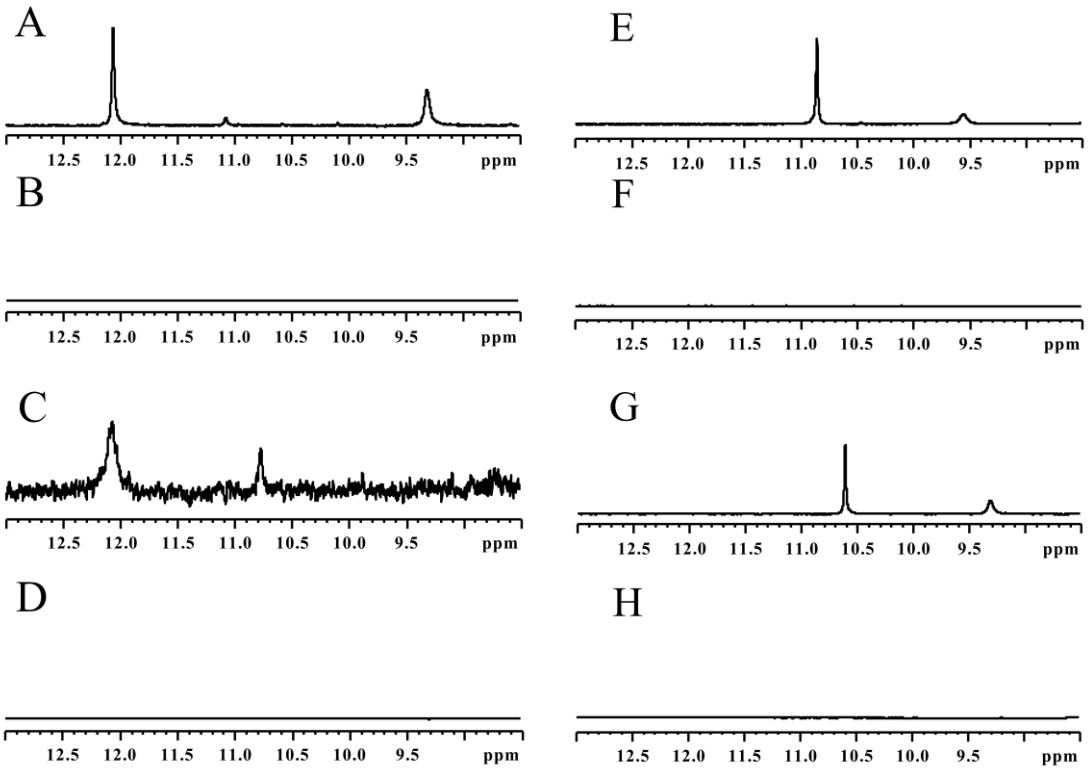
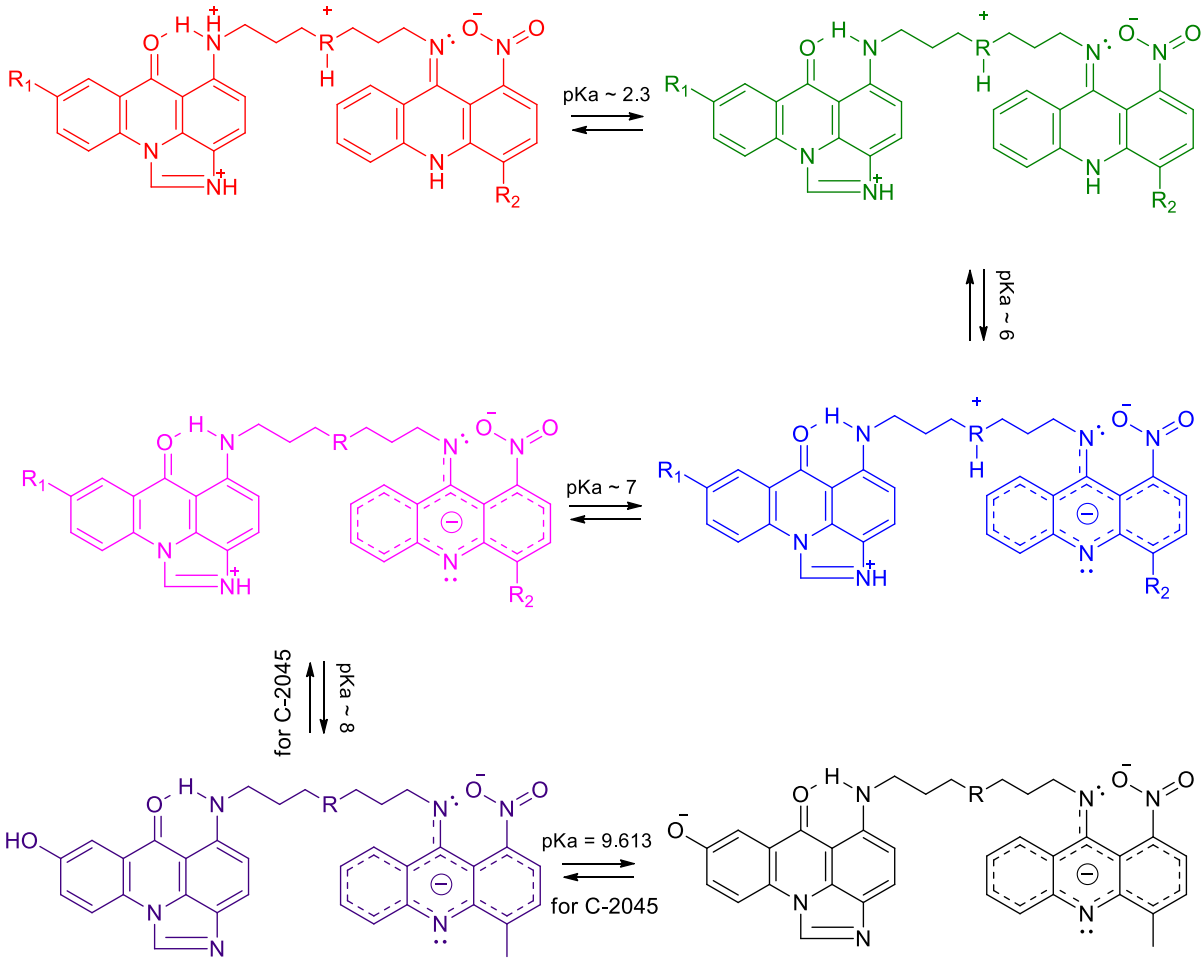


Fig. 48. 1H NMR spectra of UAs: C-2028 at pH = 1 (A) and pH = 7 (B), C-2041 at pH = 1 (C) and pH = 7 (D), C-2045 at pH = 1 (E) and pH = 7 (F), C-2053 at pH = 1 (G) and pH = 7 (H). Temperature = 5 ◦C; H2O/D2O collected in sulfuric acid (pH = 1) and in 10^{-2} M cacodylate buffer (pH equal to 5 and 7).

5.1.7 Protonation structures of UAs

NMR studies have allowed to resolve only two protons, which correspond to pKa values around 6 and around 7. However, considering the results obtained from

NMR studies and comparative analysis of structural analogues, following pH-dependent structures of the UAs were proposed (Figure 49).



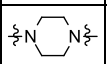
	C-2028	C-2041	C-2045	C-2053
R				
R_1	-	-	OH	-
R_2	-	-	CH_3	CH_3

Fig. 49. Proposed structure of UAs different forms depending on various pH conditions and corresponding them approximate pKa values. C-2045 is used as an example.

Comparative analysis of C-1311 and the structural analog of C-1311 has indicated that pKa values around 2.3 are related to the amino group of imidazoacridinone ring of both compounds. NMR spectra analysis strongly suggested that the next proton to dissociate at pH of approximately 6 is the proton bound to the acridine nitrogen of nitroacridine moiety. This process changes the entire ring from cyclic amidine

to aromatic ring with delocalized negative charge, which increases the susceptibility of the whole compound to hydrolysis (data not shown). Next pKa values around 7 are associated with hydrogen atom from tertiary amino group build into aminoalkyl linker. These pKa values are surprisingly low. However, an NMR signal related to that proton was identified at pH = 1, vanishing at pH = 7. The two last pKa values, corresponding to the imidazole part of imidazoacridinone ring and hydroxyl group for C-2045 were identified based of comparative analysis of C-1311 and structural analogue of C-1311.

5.2 Self-association of UAs

UAs, due to their partially planar fragments and the presence of two aromatic rings may be suspected to undergo self-association phenomenon in water solutions. Moreover, among acridine based compounds there are many examples of self-associating agents reported[86,87]. To investigate the aggregation of UAs, UV-Vis study followed by chemometric analysis was performed.

5.2.1 Self-association at pH 2.5

Residual spectra analysis of all UAs at pH = 2.5 have revealed that only two eigenvectors were valid, which resulted in two pure aggregation forms under given conditions. The best-fitting model, corresponding to the obtained results, is dimerization. C-2045 has the strongest tendency to self-associate, while C-2041 has the weakest tendency. Interestingly, C-2028 and C-2053 dimerize in a similar way. Considering the chemical structures, presumably an additional hydroxyl group of C-2045 increases susceptibility to self-associate. As long as the UAs are not substituted with groups that substantially impact the ring electron density (C-2028 and C-2053), the self-association is similar. Piperazine ring in aminoalkyl linker of C-2041 strongly decreased the ability to dimerize (Figure 50).

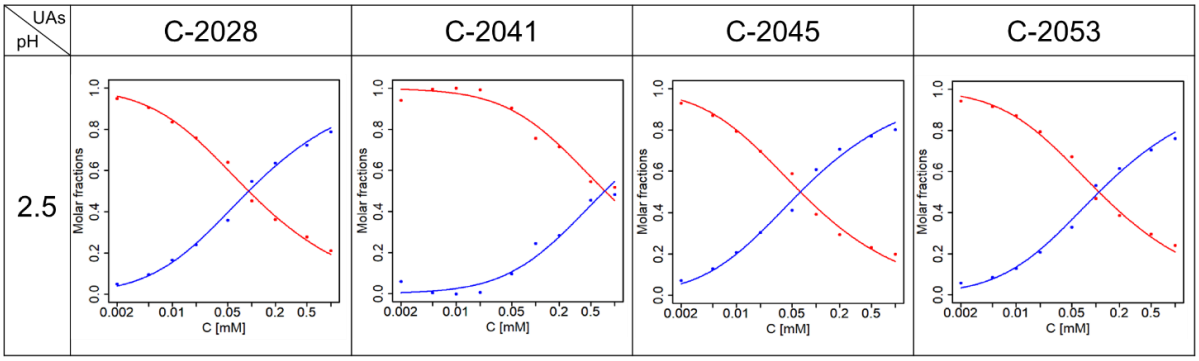


Fig. 50. UA molar fractions for UAs studied at pH = 2.5 and 35 °C. Points - optimal molar fractions; curves - the best fit of the theoretical model of dimerization to the data. Red color - monomer, blue color - dimer.

5.2.2 Self-association at pH 3.5

Self-association is more complex at pH = 3.5 and takes the form of unlimited aggregation. This process is also stronger. Interestingly, C-2041 aggregates with the greatest ratio (Figure 51).

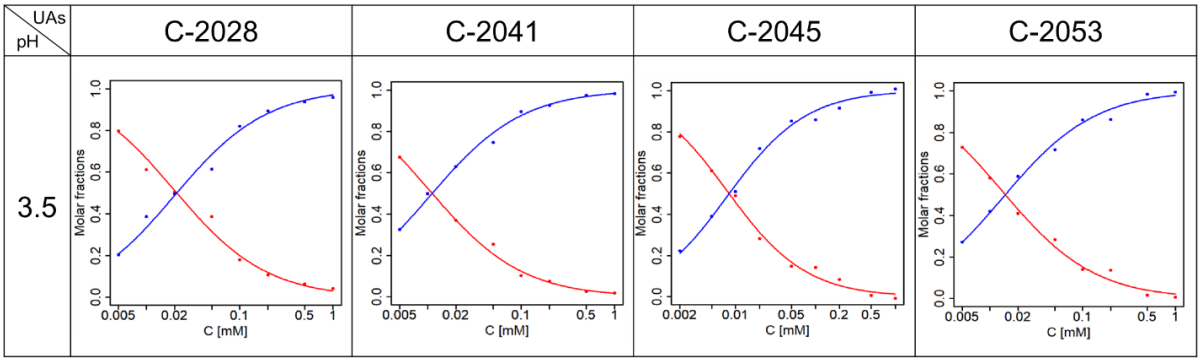


Fig. 51. Molar fractions for UAs studied at pH = 3.5 and 35 °C. Points - optimal molar fractions; curves - the best fit of the theoretical model of dimerization to the data. Red color - monomer, blue color - dimer.

5.2.3 Self-association at higher pH values

At pH 6.5 and above, higher order aggregation of UAs occurs and the results do not fit any simple self-association model (Figure 52).

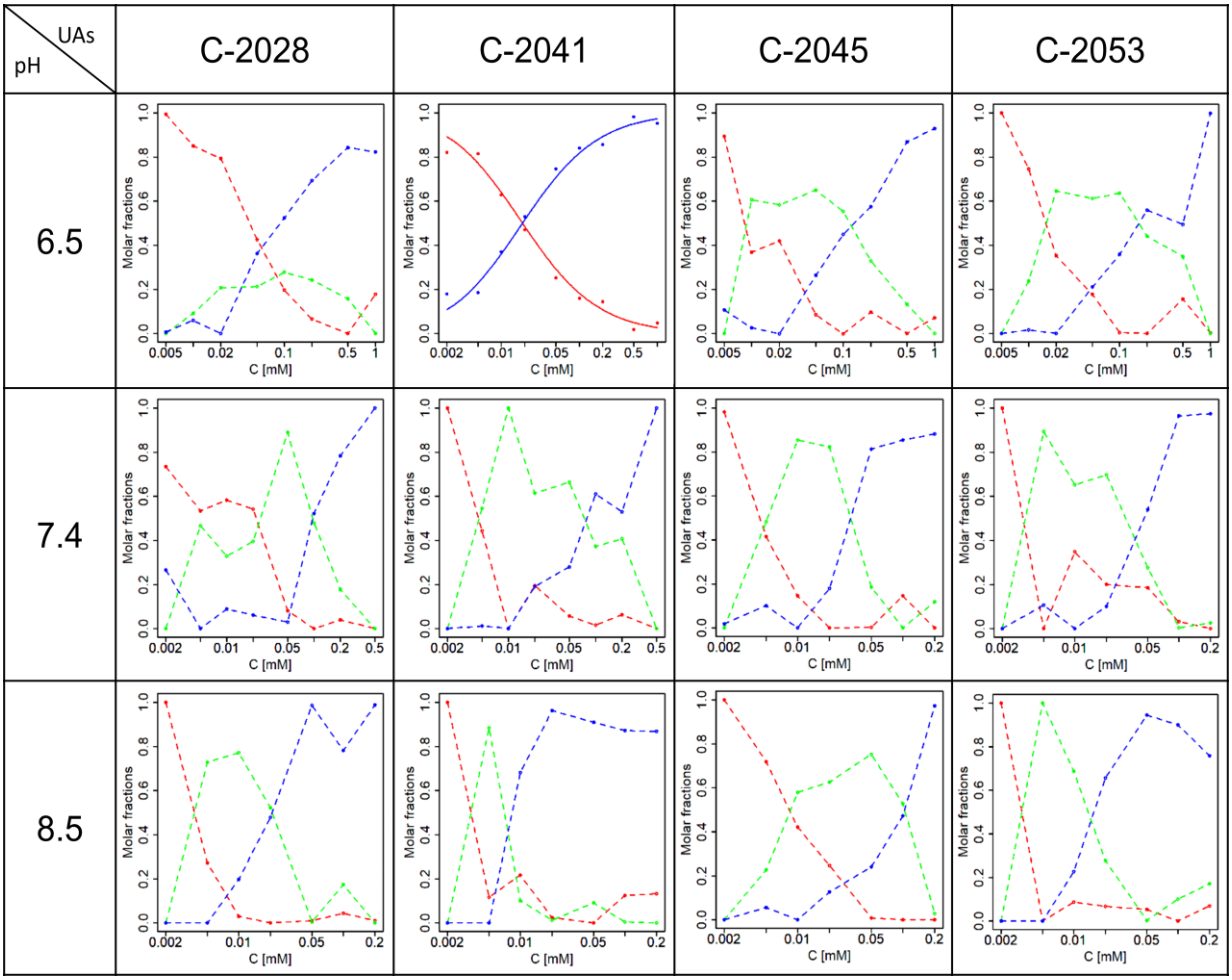


Fig. 52. Molar fractions of UAs at pH = 6.5, 7.4, 8.5 at 35 °C. Dotted curves - molar fractions of individual self-associates, solid curves - the best fit of the theoretical model. Red color - monomer, green color - first order aggregate (presumably a dimer), blue color - higher order aggregate.

The table below summarizes the determined self-association constants.

Table 1. Self-association of UAs – summary.

pH	2.5		3.5		6.5		7.4		8.5	
UAs	**K**	**STD**	**K**	**STD**	**K**	**STD**	**K**	**STD**	**K**	**STD**
C-2028	**10.89***	**0.02**	**26.94***	**0.033**	**~20**		**~40, ~10**		**~220, ~50**	
C-2041	**1.33**	**0.04**	**50.77***	**0.02**	**33.25**	**0.037**	**~200, ~13**		**~250, ~143**	
C-2045	**15.72***	**0.026**	**~63**		**~125, ~40**		**~200, ~40**		**~125, ~10**	
C-2053	**9.14***	**0.026**	**~50**		**~60, ~5**		**~250, ~20**		**~400, ~60**	

5.3 C-2045 and C-2053 with Pu22 sequence – model of interactions

Studies of small-molecule ligand's interactions with DNA are complicated and often require sophisticated methods to reveal the nature of the studied system. All of the studies over such complexes involved several steps:

1. recording of 2D NMR spectra of the DNA sequence (reference spectra);
2. recording of 2D NMR spectra of the studied complexes;
3. assigning signals to protons;
4. building three-dimensional models of the investigated systems;
5. studying their properties using molecular dynamics techniques.

In this study, based on the results of previous studies, the Pu22 sequence, which is a fragment of the promoter region of the C-Myc proto-oncogene, was used for further research.

5.3.1 Reference spectrum of Pu22

The three-dimensional structure of Pu22 is already known and the numbering of the nitrogenous bases was taken from the publication of Dai et al[47].

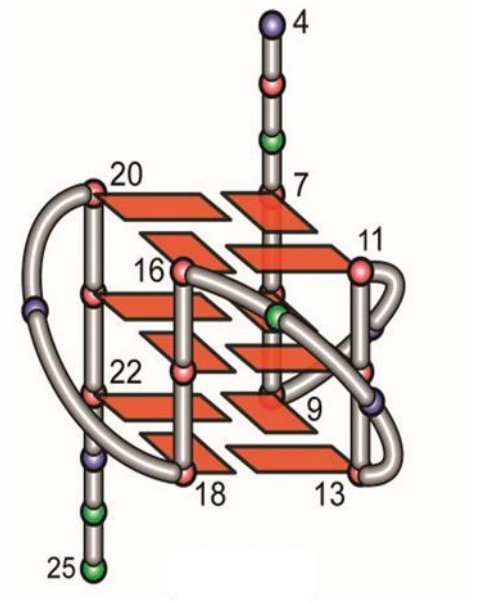


Fig. 53. Labeling of DNA bases and protons[47].

The spectrum below displays assignments of the imino protons of free (uncomplexed) Pu22 (Figure 54).

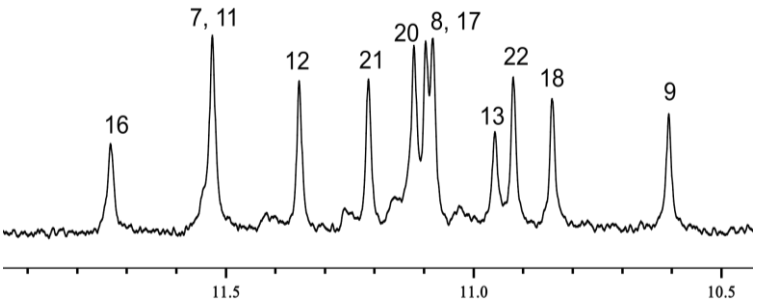


Fig. 54. Assignments of the imino protons of Pu22. at pH = 5.0, temperature = 50ºC; H2O/D2O collected in 10^{-2} M cacodylate buffer with 10 mM KCl.

Figure 55 shows the NOESY walk for the Pu22 sequence. The signals were assigned to protons according to the principle that the H1' deoxyribose sugar proton exhibits dipolar coupling to proton H6 or H8 of 'its own' nitrogenous base and the 3'-adjacent nucleotide in most cases, which was described in detail in chapter 3.2.3.

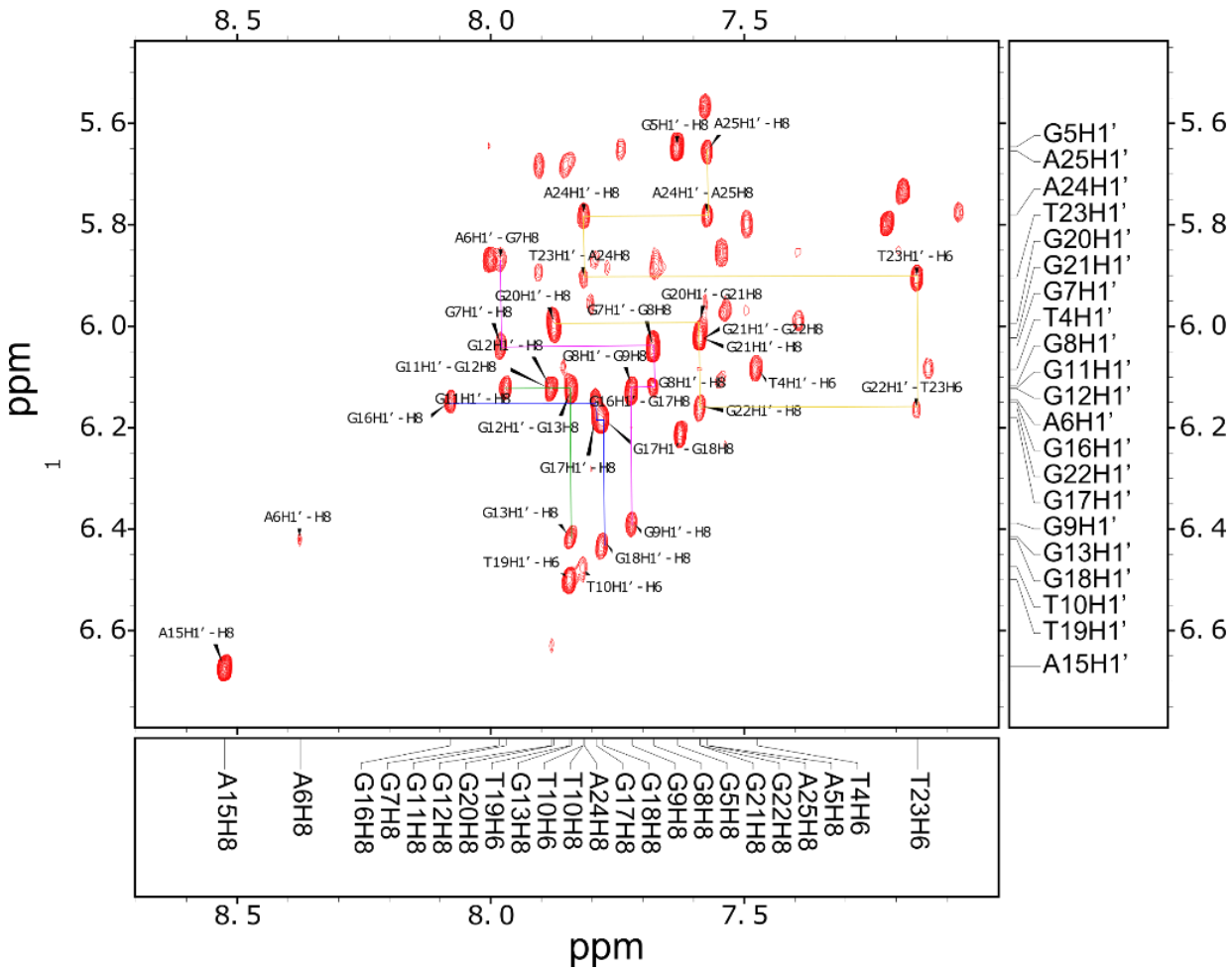


Fig. 55. The expanded H8/H6-H1' region of the 2D-NOESY spectrum of the Pu22.

In the next step, the NOESY walk allowed the assignment of the signals to the remaining deoxyribose protons. The following figures present signals of H1' /H2" -H2' and H2 " -H2'/ H3' deoxyribose protons (Figure 56 and 57).

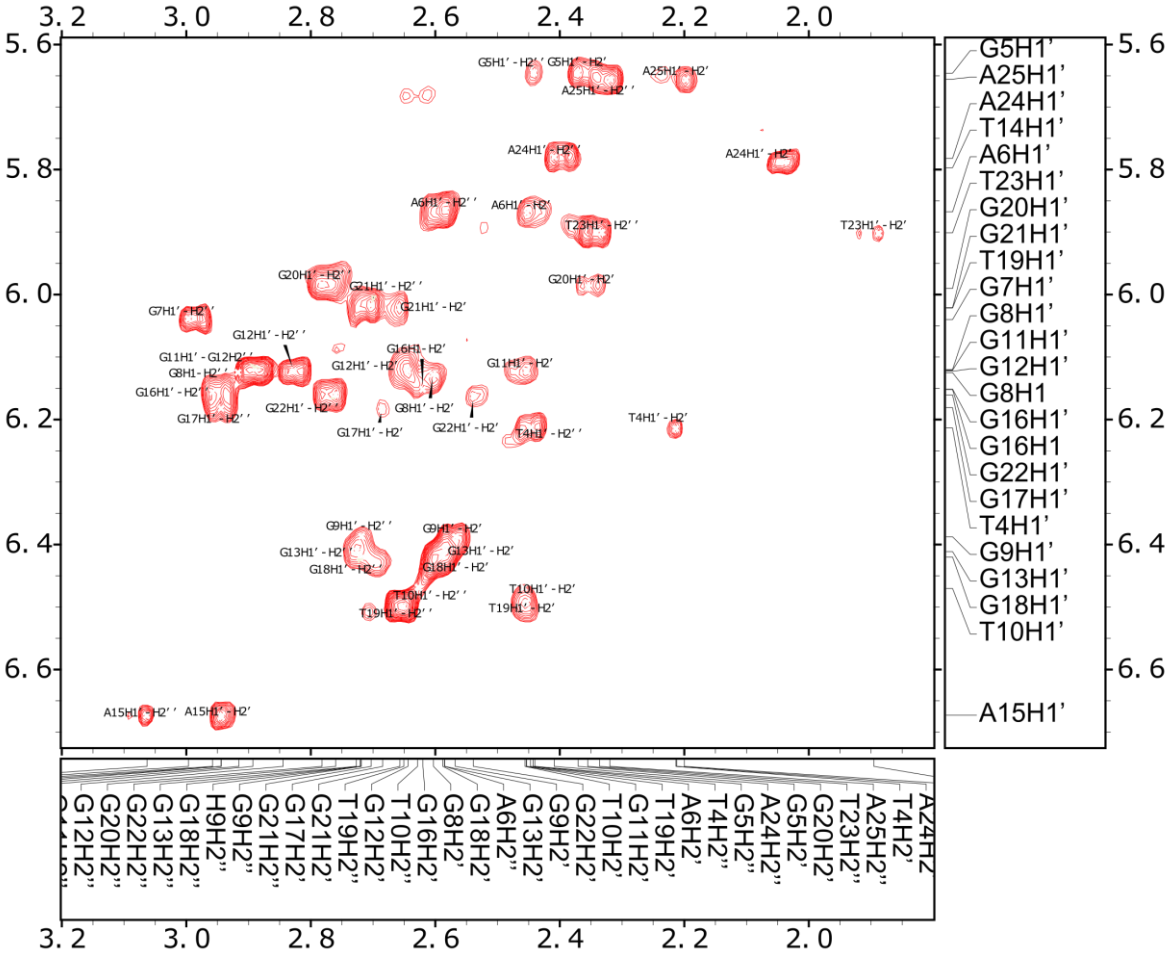


Fig. 56. The expanded H1' /H2''-H2' region of the 2D-NOESY spectrum of the Pu22.

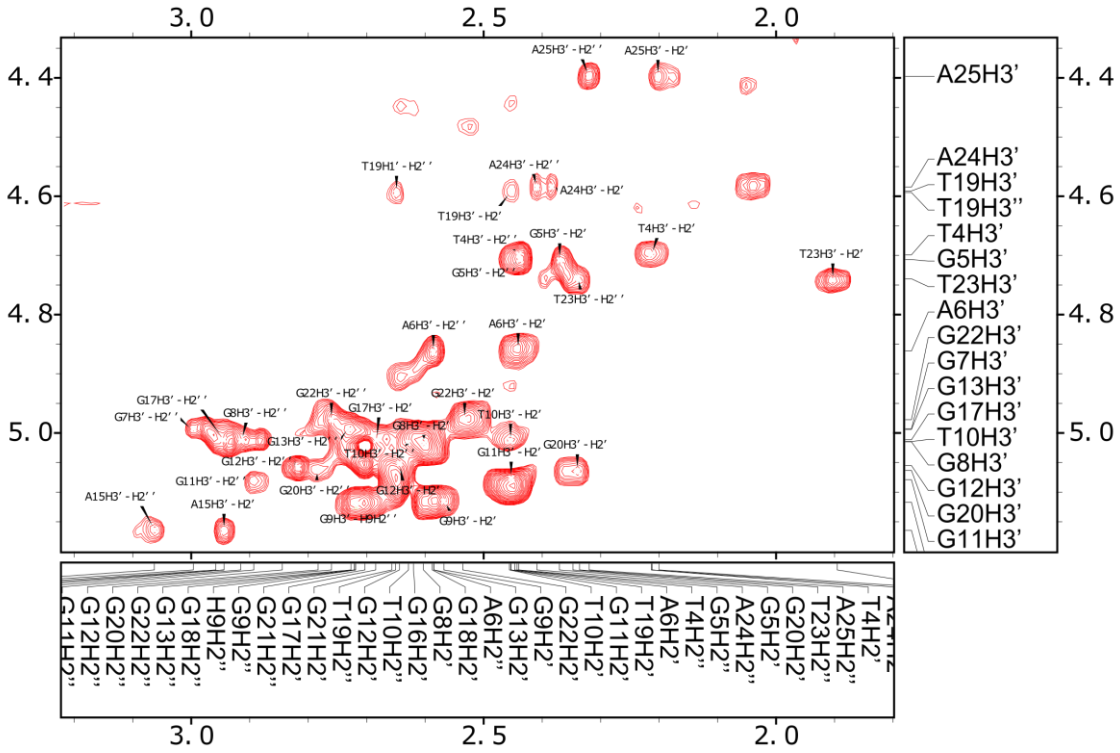


Fig. 57. The expanded H2¨-H2/ H3'' region of the 2D-NOESY spectrum of the Pu22.

5.3.2 Studies on the formation of the complexes - titration of Pu22 with acridine derivatives.

The first step to unravel the nature of the interaction of UAs with Pu22 was to collect 1D NMR spectra with various ligand ratios. The results have strongly indicated that both C-2045 and C-2053 interact with Pu22 with a DNA:ligand 1:2 mol/mol ratio (Figure 58).

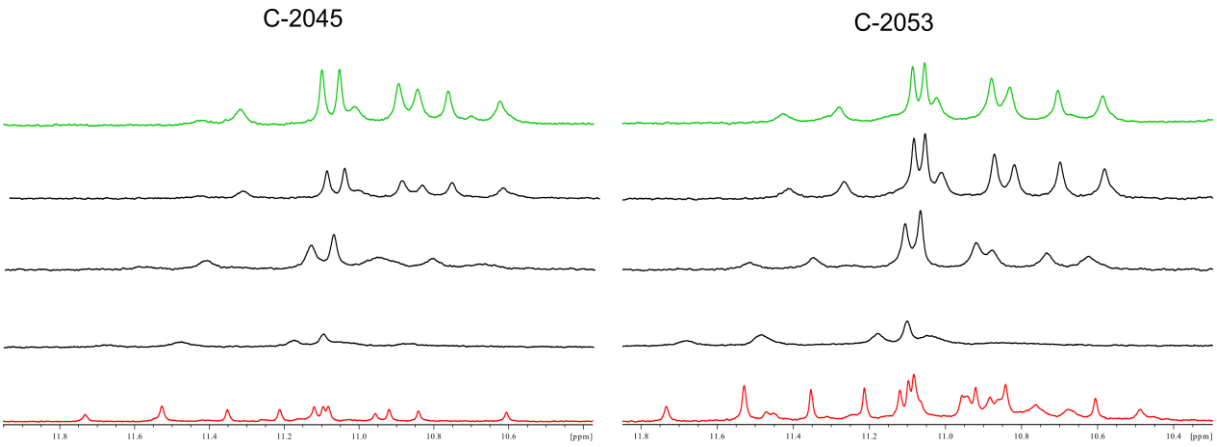


Fig. 58. Spectrum of C-2045/Pu22 complex (left), spectrum of C-2053/Pu22 (right). The first spectrum is the spectrum of the pure DNA, while the next one is the spectrum of the complex, recorded every 0.5 eq at pH = 5.0, temperature = 50ºC; H2O/D2O collected in 10^{-2} M cacodylate buffer with 10 mM KCl.

In the spectra above, it can be seen that the signals of the imino protons shift as the titration progresses. The largest shifts are observed for the signals corresponding to the upper and lower tetrad imino protons. The signals corresponding to the protons of the middle tetrad are barely shifted at all. Only their width notably changes during the titration. Such changes are observed for both C-2045 and C-2053. These findings suggested that the interaction is specific and strong. Based on that results, the NOESY spectra were performed. Despite the fact that the spectra were recorded under various pH and temperature conditions, no UAs resonance signals were visible in the NOESY spectra. To overcome this problem, the fact that UAs are formed by two aromatic rings that are not electron-coupled to each other, has been exploited. Each ring can be approximated by a monomeric acridine analogue, whose aromatic systems are identical to the UA rings. For this purpose, two monomeric acridines named C-1311 and C-1748 were employed for further studies (Figure 59).[24]

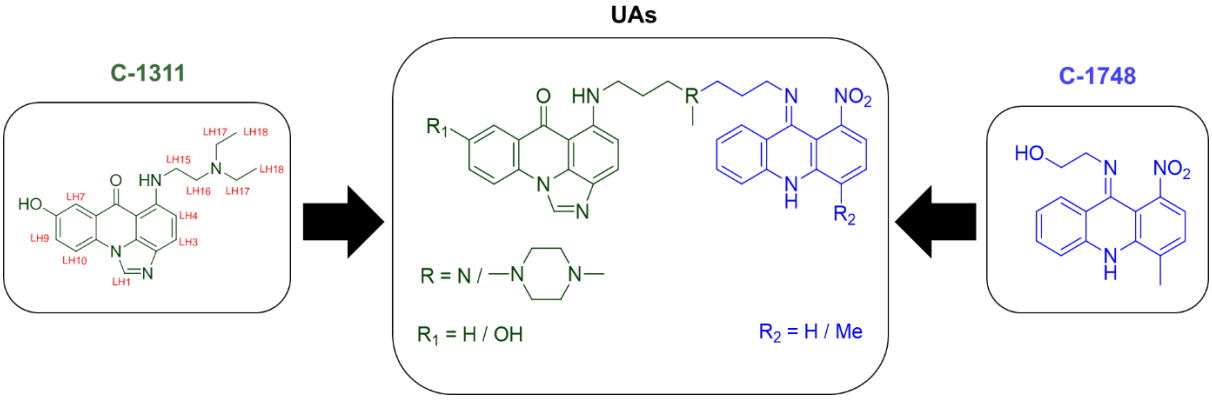


Fig. 59. Unsymmetrical bisacridine analogues: C1311 – with proton labelings (left), C-1748 (right) and general structure of UAs (in the middle).

Figure below shows the 1D NMR spectra of C-1748 and C-1311 complexed with Pu22. It is easy to see that the spectra of Pu22 and the complex consisting of Pu22 and C-1748 in a ratio of 2:1 are practically indistinguishable, whereas the spectra of its complex with C1311 are very different (Figure 60).

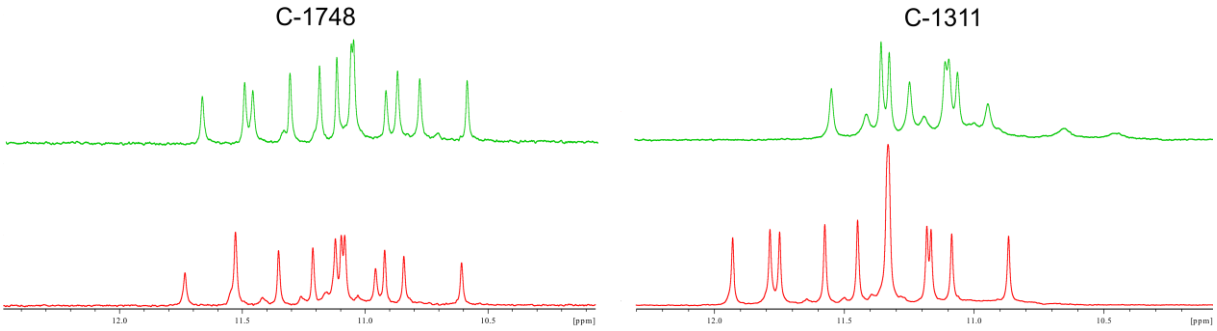


Fig. 60. NMR spectra of C-1748 (left) and C1311 (right). The red spectrum is the spectrum of pure DNA, while the green spectrum is the spectrum of the 2:1 complex. The spectra were recorded at pH = 5.0, temperature = 50ºC; H2O/D2O and collected in 10^{-2} M cacodylate buffer with 10 mM KCl.

An intricate analysis of the spectra presented above allowed to conclude that C-1311 forms a complex with Pu22 in such a way that one ligand molecule interacts with the upper tetrad, while the other interacts with the lower tetrad (DNA:ligand 1:2 mol/mol). The interactions of C-1748 with Pu22 seem to be weak and unspecific, as the compound barely impacts the imino resonances of the studied G-quadruplex. Based on these observations, complexes consisting of Pu22 sequences and compounds C-2045 and C-2053, respectively, were constructed.

5.3.3 NOESY spectrum of Pu22/C-1311 complex

The figure below shows a fragment of the NOESY spectrum containing signals of aromatic protons and signals of two ligand protons. The ligand protons are very distinctive and are clearly distinguished by their size (Figure 61, 62 and 63).

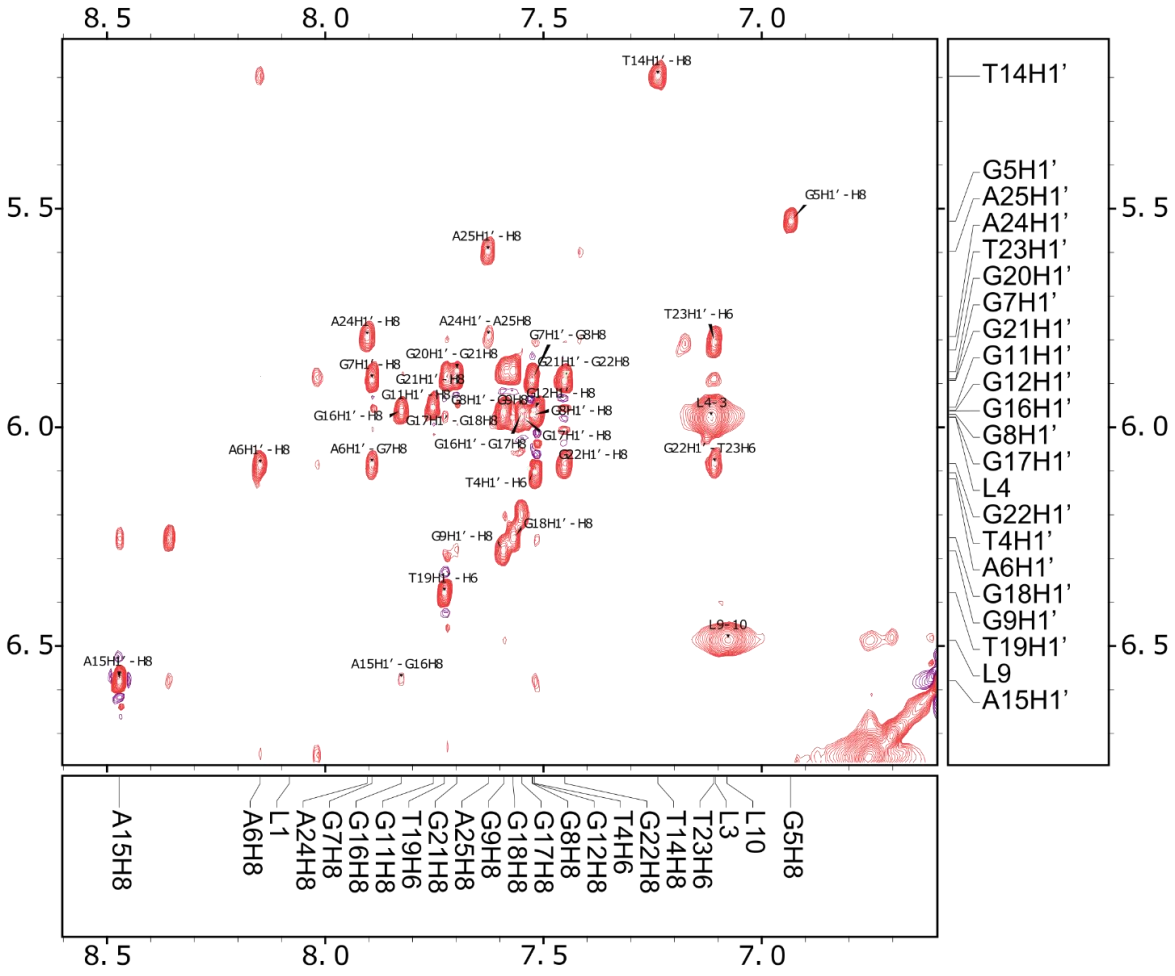


Fig. 61. The expanded H8/H6-H1' region of the 2D-NOESY spectrum of the Pu22/C-1311 complex. L4-3 and L9-10 are ligand correlation signals.

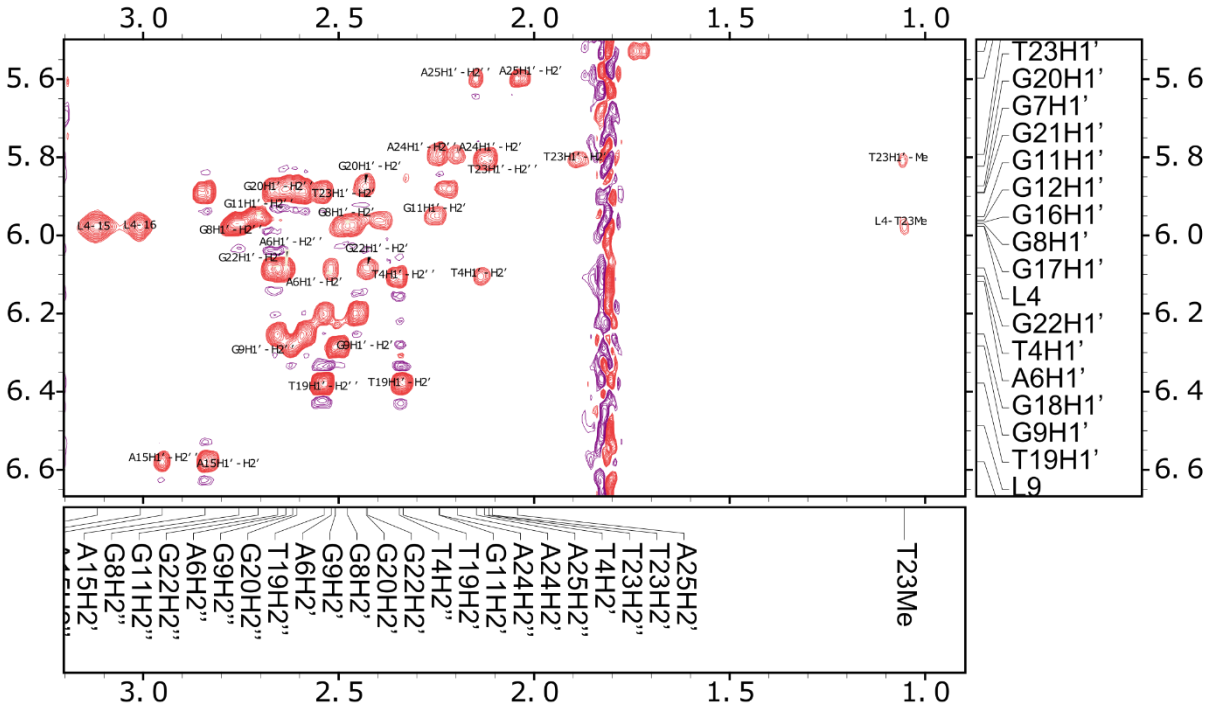


Fig. 62. The expanded H1'/H2''-H2' region of the 2D-NOESY spectrum of the Pu22/C-1311 complex. L4-15 and L4-16 are ligand correlation signals.

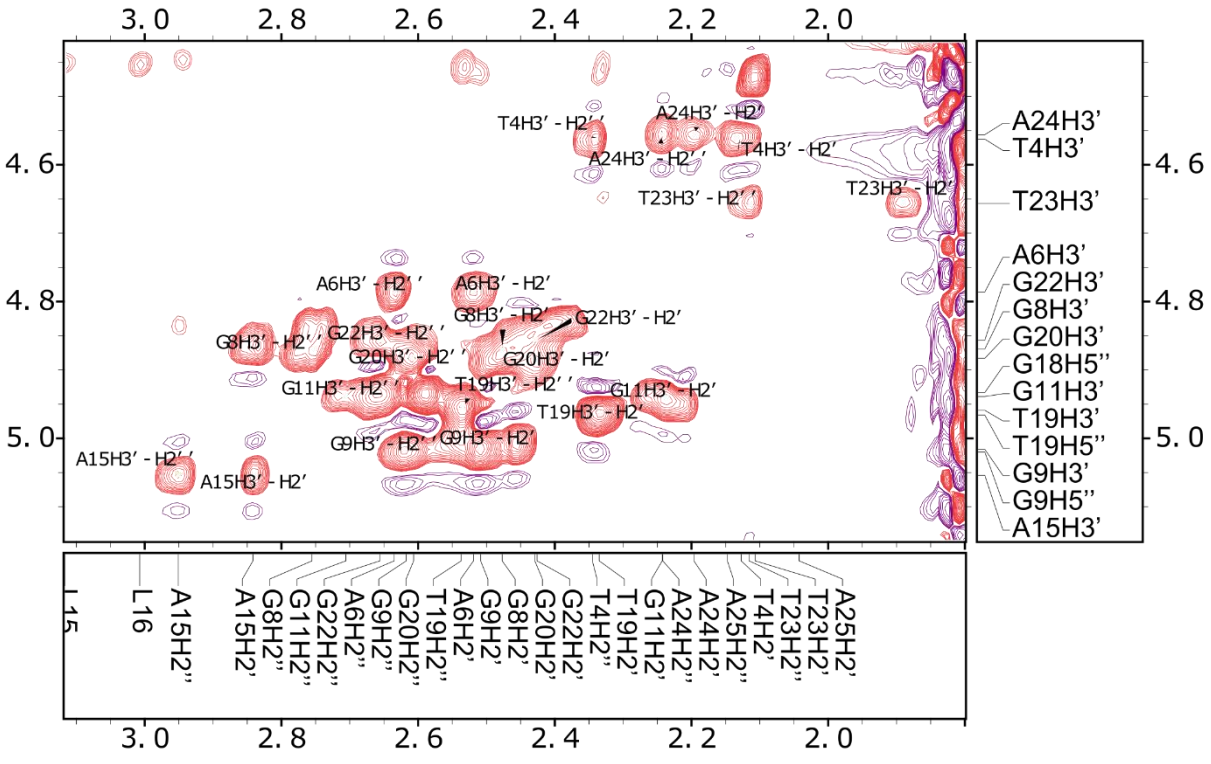


Fig. 63. The expanded H2''-H2'/ H3' region of the 2D-NOESY spectrum of the Pu22/C-1311 complex.

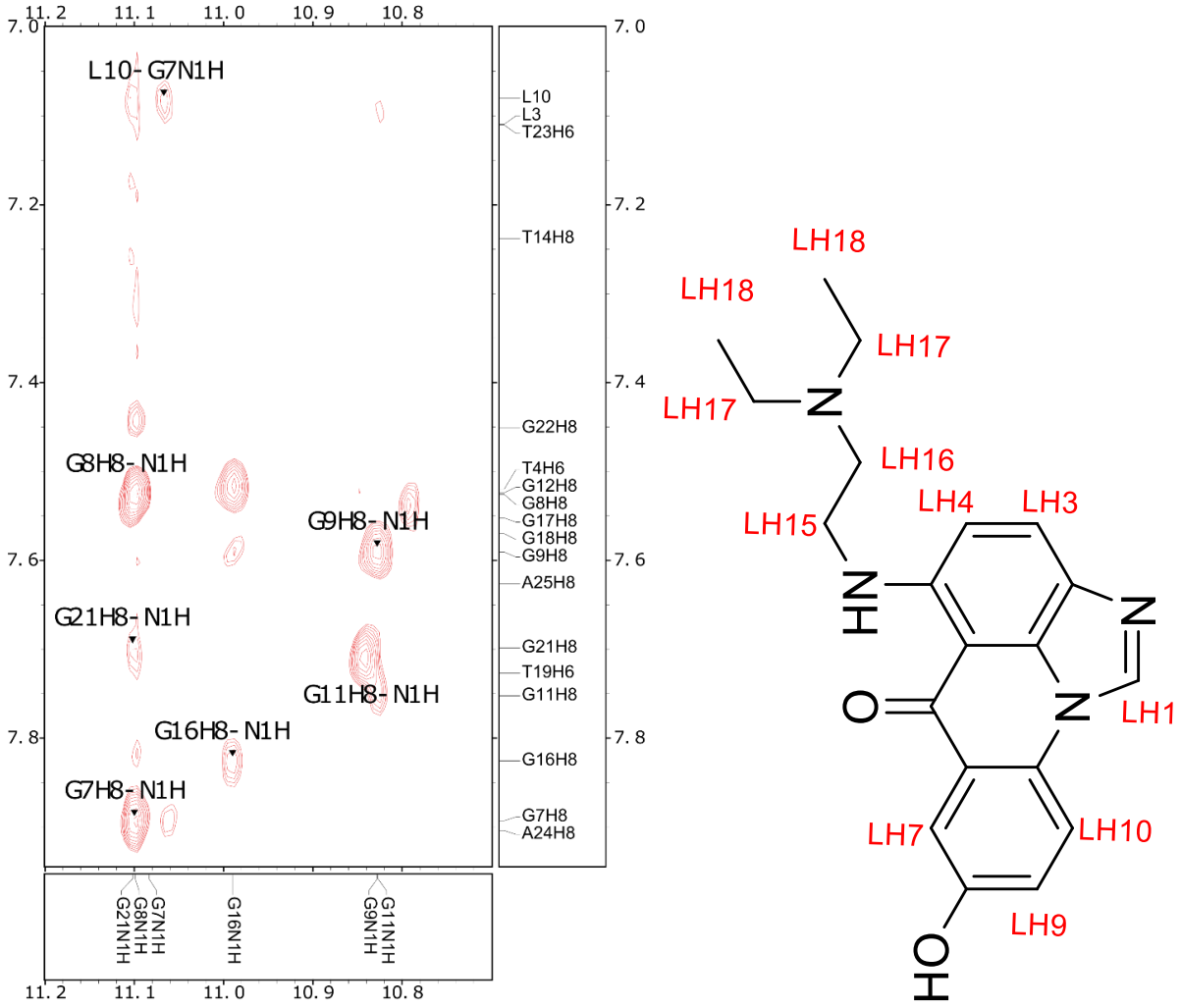


Fig. 64. The expanded imino region of the 2D-NOESY spectrum of the Pu22/C-1311 complex (left) and ligand proton labeling (right).

List of the observed Pu22/C-1311 correlation signals is presented below:

Table 2. Pu22 – C1311 NOE signals.

DNA protons	Ligand protons
T23Me	L15
T23Me	L16
T23Me	L4
T23H2''/H2'	L10
G7N1H	L10

Signals from Table 2 show which DNA and ligand atoms are in spatial proximity. They enabled the construction of the initial structures of the complexes, which were further investigated by molecular dynamics methods.

5.3.4 Umbrella Sampling

Based on 1 μs MD simulations with the application of the US method, free energy profiles depicting the interactions between UAs and upper/lower guanine tetrads for the complexes formed by C-2045 and C-2053 with Pu22, respectively, were obtained. For each of the complexes, free energies for the upper (3' end) and lower (5' end) tetrad were determined separately. A comparison of the free energy profiles was presented below (Figure 65).

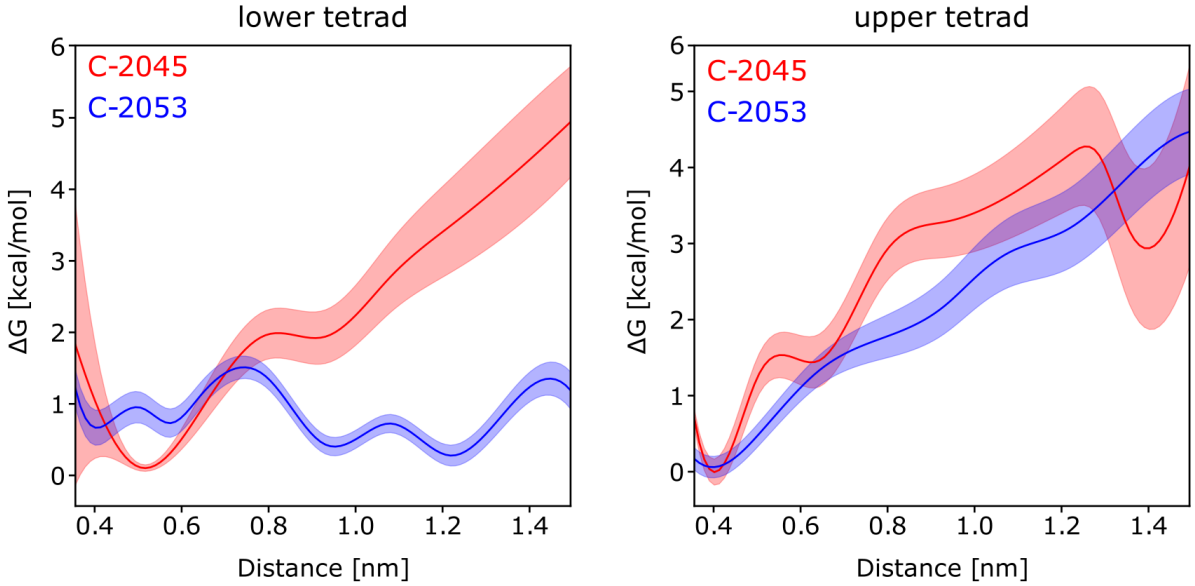


Fig. 65. Free energy for the lower (left) and upper tetrad (right).

The figure above shows that the interactions with the upper tetrad are more specific than with the lower tetrad in the case of both ligands. However, the most specific interaction among the studied compounds (narrowest energy minimum) is observed for C-2045. In the case of the lower tetrad, the interactions of both ligands are much less specific. A well-defined energy minimum was displayed only for compound C-2045. In the case of C-2053, there is no clear energy minimum.

The figure below shows the free energy profiles with their corresponding, representative structures, obtained via clustering of the US trajectories (Figure 66).

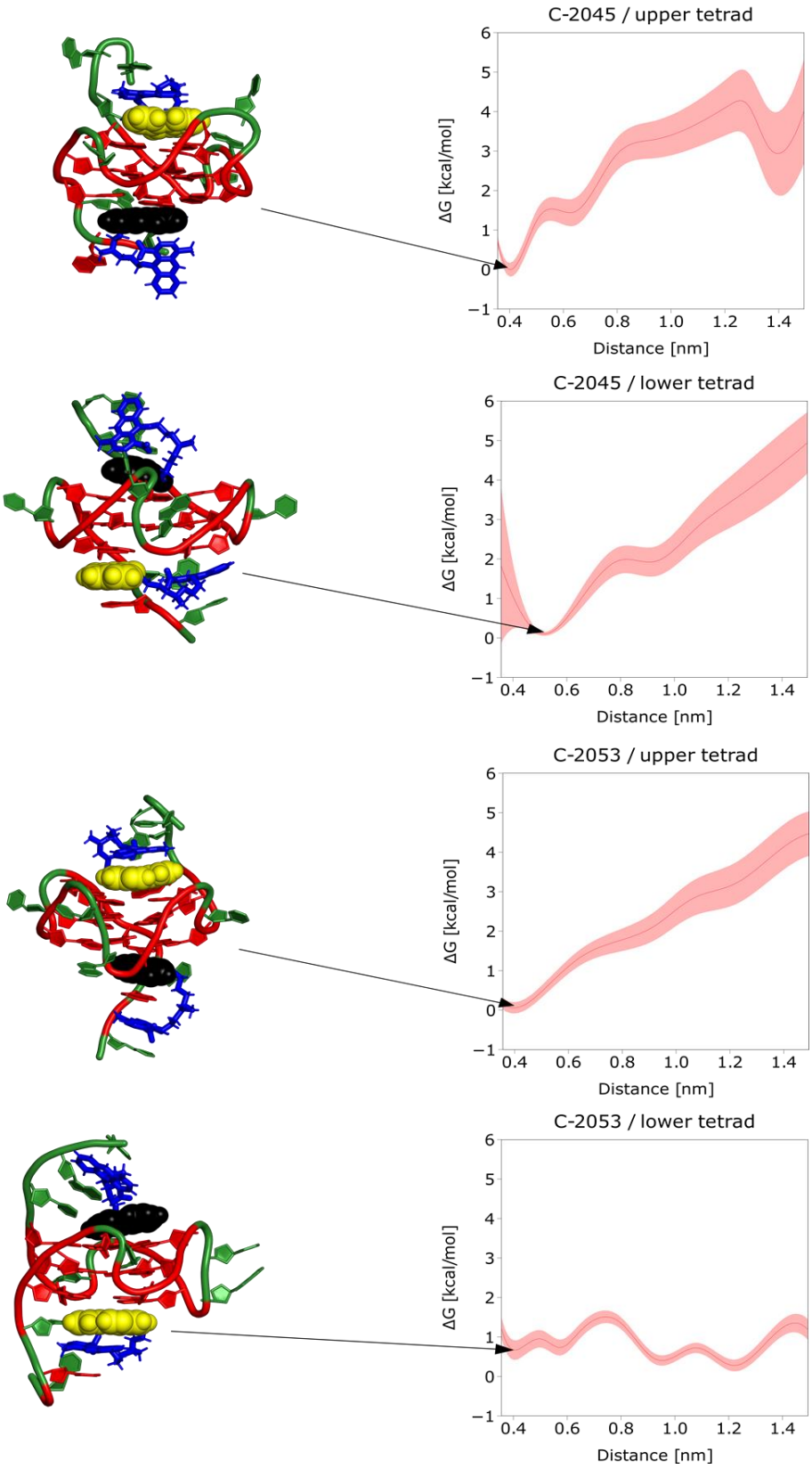


Fig. 66. Free energy profiles for complexes of Pu22 with C-2045 and C-2053, respectively. Guanosines are depicted in red while cytosines, thymidines and adenosines are marked in green. An aminoalkyl linker together with the 1-nitroacridine ring is shown in blue. The imidazoacridone ring corresponding to a particular energy minimum is marked in yellow, while the other is depicted in black.

The figure above clearly shows that both studied compounds interact with Pu22 by imidazoacridinone ring, while the linker and 1-nitroacridine ring play a role of a side chain.

6 Discussion

The first part of the dissertation presents an in-depth analysis of the acid-base equilibria and self-association of compounds belonging to the family of unsymmetrical bisacridines (UAs). These physicochemical properties of acridine derivatives have been studied previously to some extent. However, the complex approach used in this research is innovative, while most of the previous work was based on simpler and less reliable methodology[88,89,90]. Determination of pKa values and the assignment of the corresponding protonable heteroatoms, combined with self-association studies, have shown how many forms the UAs can exist in. The complex nature of the compounds under study has an enormous impact on their interactions with DNA, proteins and other biological structures[91,92], which sheds a new light on their potential biological mechanisms of action. The second stage of the research involved investigating the way that UAs interact with the model Pu22 sequence, using advanced methods of nuclear magnetic resonance spectroscopy. A thorough analysis of one-dimensional proton spectra, as well as two-dimensional homo- and hetero--correlation spectra allowed for a preliminary determination of the drug-DNA interaction mode. Molecular dynamics made it possible to find the most energetically favorable states of the system, which were presented as molecular structures resulting from cluster analysis. Furthermore, to evaluate the effect of the additional hydroxyl group at the 8-position of the imidazoacridinone ring, free energy profiles were determined for the ligands (C-2045 and C-2053) interacting with the upper and lower tetrad of the G-quadruplex, respectively. To determine the free energy of the system, umbrella sampling method was applied[93,94,95].

6.1 Protonation equilibrium

UAs exist in five independent forms in the pH range from 1.5 to 11. Under physiological pH (from 6 to 8), there are mainly two or even three various protonation forms. Each of them presumably would interact with DNA in a different way, which could have a significant impact on cell viability. The Table 3 below presents pKa values determined for all UAs studied.

Table 3. pKa of UAs – summary.

Result / UA	pK_{a1}	STD	pK_{a2}	STD	pK_{a3}	STD	pK_{a4}	STD
C-2028	2.194	0.005	6.139	0.001	7.397	0.005	8.546	0.02
C-2041	2.435	0.004	5.843	0.005	7.158	0.004	-	-
C-2045	2.301	0.003	6.120	0.002	7.327	0.006	9.137	0.04
C-2053	2.285	0.005	5.997	0.002	7.522	0.003	8.235	0.004

The pKa values of all UAs are quite similar. Only C-2041 expressed two pKa values of notably lower values, whereas the first pKa value is higher than the others. The element, which distinguishes this compound from others, is the additional piperazine ring in the linking chain.

Table 4. pKa of UAs analogues – summary.

Result / Comp	pK_{a1}	pK_{a2}	pK_{a3}
C-1311	2.767	7.597	9.675
Structural analogue of C-1311	2.662	7.687	-
Ledakrin	-	6,227	-

The pKa values for monomeric structural analogues are gathered in Table 4. The determined pKa values, which were around 2.3, 7.2 and pKa_3 around 9 are associated with C-1311, which correspond to imidazoacridinone part of the studied UAs dimers. Moreover, NMR spectra allowed to prove that pKa_2 values equal to 6.139, 5.843, 6.120, 5.997 for C-2028, C-2041, C-2045 and C-2053, correspondingly (Table 3), relate to pKa_2 of 1-nitroacridine ring of Ledakrin (nitrogen atom of acridine ring), presented in Table 4. The amino groups at the position 9 of 1-nitroacridine ring is not protonated at all, it exists mainly as the imine form at the pH lower than 1-nitroacridine pKa value (Ledakrin, Table 4). Dissociation of this proton makes the whole aromatic

ring negatively charged and, as a result, the imine group at position 9 susceptible to the hydrolysis (data not shown).

The higher pKa values from 7.158 to 7.522 (pKa_3) (pKa_4 for C-2028 and C-2053) are proposed to the tertiary amine group in the side chain or to the nitrogen atom in the imidazole part of the imidazoacridinone ring (not visible in NMR spectra). However, the proton attached to the tertiary amine group is no longer visible in NMR at pH equal to 7, which allows to assume that the highest pKa for C-2028, C-2041 and C-2053 is associated with the nitrogen atom in the imidazole part of the imidazoacridinone ring. The spectral form corresponding to the pKa of C-2045 imidazole part was not visible. Therefore, the pKa_4 for C-2045 was assigned to the hydroxyl group at position 8 of the imidazoacridinone ring. The proposed pH-depending UAs structures are presented in paragraph 5.1.7.

6.2 Self-association

UAs also undergo aggregation process, which is dependent on the pH of a solution. The aggregation ratio is stronger at higher pH values. All studied compounds exist in monomeric form only at very low concentrations (~1uM). At higher concentrations, they begin to self-associate, yet the mode of self-association depends on the pH value. At the lowest pH, when all nitrogen atoms are protonated, self-association occurs gradually, with the first step being dimerization. At pH 3.5, all studied compounds are less protonated and self-association is notably more chaotic. This means that it occurs in an unlimited mode. At physiological pH and higher, the compounds' protonation state is so low that self-association occurs in two stages, yet it is impossible to determine the nature of this process.

Self-association data are critical for all other chemical and biological studies, since the examined compounds are drug candidates. Compounds existing in multimeric form may display problems with cell penetration, as well as with an access to enzyme active sites or binding to DNA. In addition, they may strongly bind to proteins[96].

6.3 Interactions of UAs with Pu22

The one-dimensional NMR spectra have clearly showed that both studied compounds (C-2045 and C-2053) form well-defined complexes with Pu22 in a 2:1 (ligand:DNA)

ratio. However, their complex physicochemical properties; particularly their ability to self-associate at relatively low concentrations, cause NOE signals to disappear even when ligand and DNA protons are relatively close together, due to undefined and constantly changing chemical neighborhood. This effect can be clearly seen by comparing one-dimensional spectra with two-dimensional NOESY spectra. Despite the fact that some of the imino protons are strongly shielded or deshielded by the proximity of the ligand, their NOE signals remain invisible. In the case of both C-2045 and C-2053, even the NOE signals between the protons of a ligand itself are missing. The use of monomeric analogues of bisacridines, codenamed C-1311 and C-1748, allowed to prove that the imidazoacridinone ring indeed interacts with the upper and lower tetrad of Pu22, while the position and orientation of the 1-nitroacridine fragment of UAs remains undetermined. Moreover, analysis of the free energy profiles, resulting from umbrella sampling, indicates that the presence of the hydroxyl group at the position 8 of the imidazoacridinone part of UAs provides a significant contribution to studied ligand/DNA interactions. The difference in interactions between the upper and lower tetrads is interesting. One side chain of the quadruplex (5'-TGA) contains a guanine moiety, unlike the second ending (TAA-3'), which incorporates an adenine base. This may indicate that the interaction with guanine may be preferable to other bases. This preference may be due to the HOMO/LUMO frontier orbitals interaction between one of the aromatic UAs systems and guanine[97].

The high energy and the presence of local energy minima at relatively large distances are a consequence of the complex structure of the ligands. The more distant minima correspond to structures, in which the 1-nitroacridine ring unspecifically interacts with the tetrads, or where both rings interact with the side chains of the G-quadruplex. Examples of the structures corresponding to less preferable conformations are presented in Figure 67.

Fig. 67. Less favorable structures at more distant minima. Guanosines are depicted in red while cytosines, thymidines and adenosines are marked in green. An aminoalkyl linker together with the 1-nitroacridine ring is shown in blue. The imidazoacridone ring corresponding to a particular energy minimum is marked in yellow, while the other is depicted in black.

The position of the aminoalkyl chain with 1-nitroacridine moiety could be estimated using metadynamics. However, this approach requires a much deeper analysis of the conformational spaces of the complexes, and the obtained results would probably not contribute significantly to the explanation of the mechanism of action of the studied compounds at a molecular level.

The interactions of UAs with Pu22 G-quadruplex described in this work are not unique by their occurrence, yet they excel with their nature and complexity. This is a direct consequence of the dimeric structures of C-2045 and C-2053. One could expect that such molecules might not yield specific interactions with DNA G-quadruplexes, yet the presented studies have proven otherwise. Nevertheless, the literature shows that Pu22 is targeted by many other agents. Among others, the interaction of the Pu22 sequence with the compound designated as DC-34, along with the interaction of the mutated version of this sequence designated as Myc2345 with the PEQ compound were investigated. In both cases, the interaction involved non-covalent binding to the upper and lower tetrads of a G-quadruplex[98,99].

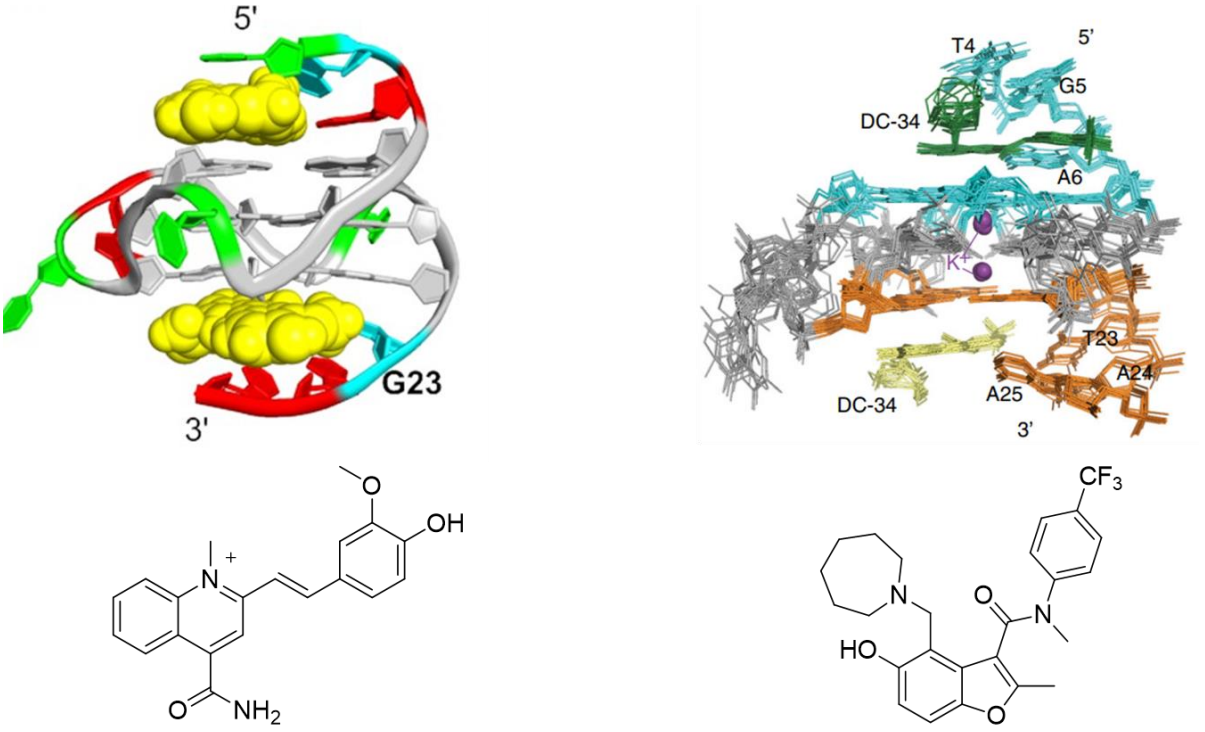


Fig. 68. Structure of the PEQ/Myc2345 complex with molecular structure of PEQ (left) and structure of the DC-34/MYC G4 complex with molecular structure of DC-34 (right)[98,98].

Many other complex structures formed by different G-quadruplexes and ligands have also been described. A complex formation usually results from ligand's interactions with the outer tetrads, but sometimes other interactions occur, e.g. intercalation[100].

7 Conclusions

To conclude, UAs exist in five independent forms in the pH range from 1.5 to 11. Under physiological pH (from 6 to 8), there are mainly two or even three various protonation forms. Each of them probably can interact with DNA in a different way, which can have a significant influence on cell viability. The pKa values of studied UAs are quite similar. Only C-2041 has two pKa values clearly lower, whereas the first pKa value is higher than the others. Except the fact, that UAs possess an ability to adopt a wide variety of protonation states, they also exhibit a strong tendency to self-associate, which is highly pH dependent. Nevertheless, the flat structure of the imidazoacridinone rings causes them to form well-defined complexes with DNA fragments capable of forming G-quadruplexes, which may be the basis of their mechanism of action at the molecular level. The interaction of the planar imidazoacridone aromatic system with the investigated Pu22 sequence is typical. Noteworthy is the fact, that both ligands interact more strongly and specifically with the upper tetrad, whose side chain does not contain guanine in its sequence. The analogous interaction with the lower tetrad is weaker and less specific. The fact, that the 5'-ending side chain has guanine in its structure, suggests that the interaction with this base is preferred. In addition, it can be seen that the hydroxyl group at the 8-position of the imidazoacridone ring significantly affects the distribution of its electron density and makes the interactions of C-2045 with both tetrads somewhat more specific than those of C-2053. Possibly, a more accurate description of these interactions could be obtained using quantum chemistry approach.

BIBLIOGRAPHY

1. WHO Chronicle. International Non-proprietary Names for Pharmaceutical Substances. Suppl. *WHO Chron.* **30**, 1–18 (1976).
2. Kumar, R., Sharma, S. & Prasad, D. Acridones: A Relatively Lesser Explored Heterocycle for Multifactorial Diseases. *Key Heterocycle Cores Des. Multitargeting Mol.* 53–132 (2018).
3. Wiśniewska, A., Chrapkowska, A., Kot-Wasik, A., Konopa, J. & Mazerska, Z. Metabolic transformations of antitumor imidazoacridinone, C-1311, with microsomal fractions of rat and human liver. (2007).
4. Pawlak, J. W., Pawlak, K. & Konopa, J. The mode of action of cytotoxic and antitumor 1-nitroacridines. II. In vivo enzyme-mediated covalent binding of a 1-nitroacridine derivative, ledakrin or nitracrine, with DNA and other macromolecules of mammalian or bacterial cells. *Chem. Biol. Interact.* **43**, 151–173 (1983).
5. Laskowski, T., Czub, J., Sowiński, P. & Mazerski, J. Intercalation complex of imidazoacridinone C-1311, a potential anticancer drug, with DNA helix d(CGATCG)2: stereostructural studies by 2D NMR spectroscopy. *J. Biomol. Struct. Dyn.* **34**, 653–663 (2016).
6. Laskowski, T., Borzyszkowska, J., Grynda, J. & Mazerski, J. C-1311 (Symadex), a potential anti-cancer drug, intercalates into DNA between A and G moieties. NMR-derived and MD-refined stereostructure of the d(GAGGCCTC)2:C-1311 complex. *J. Mol. Struct.* **1141**, 357–367 (2017).
7. Mieszkowska, A. *et al.* Metabolic Profiles of New Unsymmetrical Bisacridine Antitumor Agents in Electrochemical and Enzymatic Noncellular Systems and in Tumor Cells. *Pharmaceuticals* **14**, 317 (2021).
8. Potęga, A., Kosno, M. & Mazerska, Z. Novel insights into conjugation of antitumor-active unsymmetrical bisacridine C-2028 with glutathione: Characteristics of non-enzymatic and glutathione S-transferase-mediated reactions: Conjugation of antitumor-active C-2028 with glutathione. *J. Pharm. Anal.* **11**, 791–798 (2021).
9. Potęga, A. Glutathione-mediated conjugation of anticancer drugs: an overview

of reaction mechanisms and biological significance for drug detoxification and bioactivation. *MOLECULES* **27**, 5252 (2022).

10. Murade, C. U. & Shubeita, G. T. A fluorescent reporter on electrostatic DNA-ligand interactions. *Biomed. Opt. Express* **13**, 159 (2022).

11. Shaikh, S. A., Ahmed, S. R. & Jayaram, B. A molecular thermodynamic view of DNA-drug interactions: A case study of 25 minor-groove binders. *Arch. Biochem. Biophys.* **429**, 81–99 (2004).

12. Soni, A., Khurana, P., Singh, T. & Jayaram, B. A DNA intercalation methodology for an efficient prediction of ligand binding pose and energetics. *Bioinformatics* **33**, 1488–1496 (2017).

13. Li, M., Lincoln, P. & Andersson, J. Slow threading intercalation of monomeric Ru(II) complexes with 10,13-diarylsubstituted dppz ligands. *J. Phys. Chem. B* **115**, 7923–7931 (2011).

14. Strekowski, L. *et al.* A non-classical intercalation model for a bleomycin amplifier. *Anticancer. Drug Des.* **2**, 387–398 (1988).

15. Manning, G. S. Counterion Condensation on Charged Spheres, Cylinders, and Planes†. *J. Phys. Chem. B* **111**, 8554–8559 (2007).

16. Strekowski, L. & Wilson, B. Noncovalent interactions with DNA: an overview. *Mutat. Res.* **623**, 3–13 (2007).

17. Khan, G. S., Shah, A., Zia-Ur-Rehman & Barker, D. Chemistry of DNA minor groove binding agents. *J. Photochem. Photobiol. B Biol.* **115**, 105–118 (2012).

18. Pearl, L. H., Skelly, J. V., Hudson, B. D. & Neidle, S. The crystal structure of the DNA-binding drug berenil: molecular modelling studies of berenil-DNA complexes. *Nucleic Acids Res.* **15**, 3469 (1987).

19. Bazhulina, N. P. *et al.* Binding of Hoechst 33258 and its derivatives to DNA. *J. Biomol. Struct. Dyn.* **26**, 701–718 (2009).

20. Mukherjee, A. & Sasikala, W. D. Drug-DNA intercalation: from discovery to the molecular mechanism. *Adv. Protein Chem. Struct. Biol.* **92**, 1–62 (2013).

21. Sasikala, W. D. & Mukherjee, A. Molecular mechanism of direct proflavine-DNA intercalation: Evidence for drug-induced minimum base-stacking penalty

pathway. *J. Phys. Chem. B* **116**, 12208–12212 (2012).

22. Vardevanyan, P. O., Antonyan, A. P., Parsadanyan, M. A., Davtyan, H. G. & Karapetyan, A. T. The binding of ethidium bromide with DNA: interaction with single- and double-stranded structures. *Exp. Mol. Med.* **35**, 527–533 (2003).

23. Laskowski, T., Czub, J., Sowiński, P. & Mazerski, J. Intercalation complex of imidazoacridinone C-1311, a potential anticancer drug, with DNA helix d(CGATCG)2: stereostructural studies by 2D NMR spectroscopy. *J. Biomol. Struct. Dyn.* **34**, 653–663 (2016).

24. Laskowski, T. *et al.* The interactions of monomeric acridines and unsymmetrical bisacridines (UAs) with DNA duplexes: an insight provided by NMR and MD studies. *Sci. Rep.* **13**, 3431 (2023).

25. Avendaño, C. & Menéndez, J. C. DNA Intercalators and Topoisomerase Inhibitors. *Med. Chem. Anticancer Drugs* 199–228 (2008).

26. Arshad, N. *et al.* Investigations on Anticancer Potentials by DNA Binding and Cytotoxicity Studies for Newly Synthesized and Characterized Imidazolidine and Thiazolidine-Based Isatin Derivatives. *Molecules* **27**, (2022).

27. Nordell, P. & Lincoln, P. Mechanism of DNA threading intercalation of binuclear Ru complexes: uni- or bimolecular pathways depending on ligand structure and binding density. *J. Am. Chem. Soc.* **127**, 9670–9671 (2005).

28. Fairbanks, S. D., Robertson, C. C., Keene, F. R., Thomas, J. A. & Williamson, M. P. Structural Investigation into the Threading Intercalation of a Chiral Dinuclear Ruthenium(II) Polypyridyl Complex through a B-DNA Oligonucleotide. *J. Am. Chem. Soc.* **141**, 4644–4652 (2019).

29. Lown, J. W., Hanstock, C. C., Bradley, B. D. & Scraba, D. G. Interactions of the antitumor agents mitoxantrone and bisantrene with deoxyribonucleic acids studied by electron microscopy. *Mol. Pharmacol.* **25**, 178–184 (1984).

30. Tanious, F. A., Wilson, W. D. & Yen, S. F. Kinetic and equilibrium analysis of a threading intercalation mode: DNA sequence and ion effects. *Biochemistry* **30**, 1813–1819 (1991).

31. Yoshikawa, M., Kubota, Y. The interaction of DNA with quinacrine and its analogs studied by absorption and circular dichroism spectroscopy. *Nucleic*

Acids Symp Ser **29**, 87–88 (1993).

32. Gao, Q. *et al.* Drug-induced DNA repair: X-ray structure of a DNA-ditercalinium complex. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 2422 (1991).

33. Ivar, B. Untersuchungen über die Guanylsäure. *Hoppe. Seylers. Z. Physiol. Chem.* **32**, 201–213 (1901).

34. Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366 (1988).

35. JR, W., MK, R. & TR, C. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **59**, 871–880 (1989).

36. Fry, M. Tetraplex DNA and its interacting proteins. *Front. Biosci.* **12**, 4336–4351 (2007).

37. Winnerdy, F. R. & Phan, A. T. Quadruplex structure and diversity. *Annu. Rep. Med. Chem.* **54**, 45–73 (2020).

38. Müller, S. & Rodriguez, R. G-quadruplex interacting small molecules and drugs: From bench toward bedside. *Expert Rev. Clin. Pharmacol.* **7**, 663–679 (2014).

39. Dang, C. V. MYC on the path to cancer. *Cell* **149**, 22–35 (2012).

40. Pelengaris, S., Khan, M. & Evan, G. c-MYC: more than just a matter of life and death. *Nat. Rev. Cancer* **2**, 764–776 (2002).

41. Madden, S. K., de Araujo, A. D., Gerhardt, M., Fairlie, D. P. & Mason, J. M. Taking the Myc out of cancer: toward therapeutic strategies to directly inhibit c-Myc. *Mol. Cancer 2021 201* **20**, 1–18 (2021).

42. Li, Q., Xiang, J.-F., Zhang, H. & Tang, Y.-L. Searching Drug-Like Anti-cancer Compound(s) Based on G-Quadruplex Ligands. *Curr. Pharm. Des.* **18**, 1973–1983 (2012).

43. Liu, H.-Y. *et al.* New Disubstituted Quindoline Derivatives Inhibiting Burkitt's Lymphoma Cell Proliferation by Impeding c-MYC Transcription. (2017).

44. Chen, B.-J., Wu, Y.-L., Tanaka, Y. & Zhang, W. Small Molecules Targeting c-Myc Oncogene: Promising Anti-Cancer Therapeutics. *Int. J. Biol. Sci* **10**, (2014).

45. Wang, J. *et al.* A Pt(ii)–Dip complex stabilizes parallel c-myc G-quadruplex.

Chem. Commun. **49**, 4758–4760 (2013).

46. Shalaby, T. *et al.* G-Quadruplexes as Potential Therapeutic Targets for Embryonal Tumors. *Molecules* **18**, 12500 (2013).

47. Dai, J., Carver, M., Hurley, L. H. & Yang, D. Solution Structure of a 2:1 QuindolineÀc-MYC G-Quadruplex: Insights into G-Quadruplex-Interactive Small Molecule Drug Design. *J. Am. Chem. Soc* **133**, 17673–17680 (2011).

48. Ou, T.-M. *et al.* Stabilization of G-Quadruplex DNA and Down-Regulation of Oncogene c-myc by Quindoline Derivatives. *J. Med. Chem.* **50**, 1465–1474 (2007).

49. Xiaohui, J. *et al.* The Interaction of Telomeric DNA and C-myc22 G-Quadruplex with 11 Natural Alkaloids. *Nucleic Acid Ther.* **22**, 127–136 (2012).

50. Strekowski, L. & Wilson, B. Noncovalent interactions with DNA: an overview. *Mutat. Res.* **623**, 3–13 (2007).

51. Dasari, S. & Bernard Tchounwou, P. Cisplatin in cancer therapy: molecular mechanisms of action. *Eur. J. Pharmacol.* **740**, 364 (2014).

52. Florea, A. M. & Büsselberg, D. Cisplatin as an Anti-Tumor Drug: Cellular Mechanisms of Activity, Drug Resistance and Induced Side Effects. *Cancers (Basel).* **3**, 1351–1371 (2011).

53. Turel, I. & Kljun, J. Interactions of metal ions with DNA, its constituents and derivatives, which may be relevant for anticancer research. *Curr. Top. Med. Chem.* **11**, 2661–2687 (2011).

54. Rademaker-Lakhai, J. M., Van Den Bongard, D., Pluim, D., Beijnen, J. H. & Schellens, J. H. M. A Phase I and pharmacological study with imidazolium-trans-DMSO-imidazole-tetrachlororuthenate, a novel ruthenium anticancer agent. *Clin. Cancer Res.* **10**, 3717–3727 (2004).

55. Arion, V. B., Jakupec, M. A., Galanski, M. S., Unfried, P. & Keppler, B. K. Synthesis, structure, spectroscopic and in vitro antitumour studies of a novel gallium(III) complex with 2-acetylpyridine 4N-dimethylthiosemicarbazone. *J. Inorg. Biochem.* **91**, 298–305 (2002).

56. Sayer, T. L. & Rabenstein, D. L. Nuclear magnetic resonance studies of the

acid–base chemistry of amino acids and peptides. III. Determination of the microscopic and macroscopic acid dissociation constants of α,ω-diaminocarboxylic acids. *Can. J. Chem.* **54**, 3392–3400 (2011).

57. Pérez, A. *et al.* NMR spectroscopic study of the self-aggregation of 3-hexen-1,5-diyne derivatives. *Chem. - A Eur. J.* **19**, 10271–10279 (2013).

58. Fattori, J., Rodrigues, F. H. S., Pontes, J. G. M., Paula Espíndola, A. & Tasic, L. Monitoring Intermolecular and Intramolecular Interactions by NMR Spectroscopy. *Appl. NMR Spectrosc.* **3**, 180–266 (2015).

59. Lindon, J. C. NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY TECHNIQUES | Multidimensional Proton. *Encycl. Anal. Sci. Second Ed.* 350–357 (2005).

60. Agarwal, N., Nair, M. S., Mazumder, A. & Poluri, K. M. Characterization of Nanomaterials Using Nuclear Magnetic Resonance Spectroscopy. *Charact. Nanomater. Adv. Key Technol.* 61–102 (2018).

61. Szakács, Z. & Sánta, Z. NMR Methodological Overview. *Hum. Asp. Sci. Think. NMR Spectrosc. Mass Spectrom.* 257–289 (2015).

62. Butts, C. P. *et al.* Interproton distance determinations by NOE - Surprising accuracy and precision in a rigid organic molecule. *Org. Biomol. Chem.* **9**, 177–184 (2011).

63. Lindon, J. C. NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY TECHNIQUES | Multidimensional Proton. *Encycl. Anal. Sci. Second Ed.* 350–357 (2005).

64. Werner, M. H., Gupta, V., Lambert, L. J. & Nagata, T. Uniform 13C/15N-Labeling of DNA by Tandem Repeat Amplification. *Methods Enzymol.* **338**, 283–304 (2002).

65. MM, P., A, J., A, K., R, M. & SY, C. Determination of pK(a) of felodipine using UV-Visible spectroscopy. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **115**, 887–890 (2013).

66. A, H. & E, B. UV-visible Absorption Study of the Self-association of Non-ionic Chromonic Triphenylenes TP6EOnM (n = 2, 3, 4) in Dilute Aqueous Solutions: Impact of Chain Length on Aggregation. *Chimia (Aarau).* **69**, 520–523 (2015).

67. Kirsanov, D., Rudnitskaya, A., Legin, A. & Babain, V. UV–Vis spectroscopy with chemometric data treatment: an option for on-line control in nuclear industry. *J. Radioanal. Nucl. Chem. 2017 3123* **312**, 461–470 (2017).

68. Raymonda, J. W., Simpson, W. T. & Simpson, W. T. Experimental and Theoretical Study of Sigma-Bond Electronic Transitions in Alkanes. Experimental and Theoretical Study of Sigma-Bond Electronic Transitions in Alkanes. *J. Chem. Phys.* **47**, 430–448 (1967).

69. Ying, J. F., Leung, ; K T & Leung, K. T. Nonoptical excited state spectroscopy of CHF 2 Cl: Characterization of nondipole n→σ* valence transitions by angle-resolved electron energy loss spectroscopy. Nonoptical excited state spectroscopy of CHF 2 CI: Characterization of nondipole n-+ u* valence . *J. Chem. Phys* **100**, 1011–1020 (1994).

70. Appleton, A. L. *et al.* Effects of electronegative substitution on the optical and electronic properties of acenes and diazaacenes. (2010).

71. Akash, M. S. H. & Rehman, K. Ultraviolet-Visible (UV-VIS) Spectroscopy. *Essentials Pharm. Anal.* 29–56 (2020).

72. Skoko, S., Ambrosetti, M., Giovannini, T. & Cappelli, C. Simulating Absorption Spectra of Flavonoids in Aqueous Solution: A Polarizable QM/MM Study. *Mol. 2020, Vol. 25, Page 5853* **25**, 5853 (2020).

73. Al-Jalali, M. A., Aljghami, I. F. & Mahzia, Y. M. Voigt deconvolution method and its applications to pure oxygen absorption spectrum at 1270 nm band. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **157**, 34–40 (2016).

74. Pawłowska, M., Kulesza, J. & Augustin, E. c-Myc Protein Level Affected by Unsymmetrical Bisacridines Influences Apoptosis and Senescence Induced in HCT116 Colorectal and H460 Lung Cancer Cells. *Int. J. Mol. Sci.* **23**, 3061 (2022).

75. Oliveira, H. P. M., Camargo, A. J., De Macedo, L. G. M., Gehlen, M. H. & Da Silva, A. B. F. A quantum chemical and photophysical study of acridine-9-N-methacrylamide. *J. Mol. Struct. THEOCHEM* **674**, 213–225 (2004).

76. El-Sheshtawy, H. S., Assran, A. S. & AbouBaker, A. M. Synthesis, Structural Characterization, Spectroscopic Properties, and Theoretical Investigations

of Aminoacridine Derivatives. *Polycycl. Aromat. Compd.* **39**, 1–13 (2019).

77. Omidi, Y. & Barar, J. Dysregulated pH in Tumor Microenvironment Checkmates Cancer Therapy. **3**, 149–162 (2013).

78. Press, W. H. & Teukolsky, S. A. Savitzky-Golay Smoothing Filters. *Comput. Phys.* **4**, 669–672 (1990).

79. Van Loan, C. F. Generalizing the Singular Value Decomposition. *SIAM J. Numer. Anal.* **13**, 76–83 (2006).

80. Abboud, E. The College Mathematics Journal Viviani's Theorem and Its Extension. *Coll. Math. J.* **41**, 1–16 (2017).

81. Gao, F. & Han, L. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Comput. Optim. Appl.* **51**, 259–277 (2012).

82. Johannes K. Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 932–942 (2011).

83. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

84. Zhou, J. K., Yang, D. Y. & Sheu, S. Y. The molecular mechanism of ligand unbinding from the human telomeric G-quadruplex by steered molecular dynamics and umbrella sampling simulations. *Phys. Chem. Chem. Phys.* **17**, 12857–12869 (2015).

85. Haider, S. Computational Methods to Study G-Quadruplex–Ligand Complexes. *J. Indian Inst. Sci.* **98**, 325–339 (2018).

86. Evstigneev, M. P., Evstigneev, V. P. & Davies, D. B. 1H NMR determination of the self-association of an acridine homodimer and its complexation with ethidium bromide in aqueous solution. *J. Mol. Struct.* **784**, 162–168 (2006).

87. Blears, D. J. & Danyluk, S. S. A Nuclear Magnetic Resonance Investigation of the Aggregation of Acridine Orange in Aqueous Solution. *J. Am. Chem. Soc.* **89**, 21–26 (2002).

88. Martínez, C. H. R. & Dardonville, C. Rapid Determination oflonization Constants (pKa) by UV SpectroscopyUsing 96-Well Microtiter Plates. *ACS Med. Chem. Lett.* **4**, 142 (2013).

89. Reijenga, J., van Hoof, A., van Loon, A. & Teunissen, B. Development of Methods for the Determination of pKa Values. *Anal. Chem. Insights* **8**, 53 (2013).

90. Shenderovich, I. G. The Partner Does Matter: The Structure of Heteroaggregates of Acridine Orange in Water. *Molecules* **24**, (2019).

91. Wanat, K. Biological barriers, and the influence of protein binding on the passage of drugs across them. *Mol. Biol. Rep.* **47**, 3221–3231 (2020).

92. Manallack, D. T. The pKa Distribution of Drugs: Application to Drug Discovery. *Perspect. Medicin. Chem.* **1**, 25 (2007).

93. Moraca, F. *et al.* Ligand binding to telomeric G-quadruplex DNA investigated by funnel-metadynamics simulations. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E2136–E2145 (2017).

94. Nayis, A., Liebl, K., Frost, C. V. & Zacharias, M. Targeting Telomeres: Molecular Dynamics and Free Energy Simulation of Gold-Carbene Binding to DNA. *Biophys. J.* **120**, 101 (2021).

95. Zhou, J. K., Yang, D. Y. & Sheu, S. Y. The molecular mechanism of ligand unbinding from the human telomeric G-quadruplex by steered molecular dynamics and umbrella sampling simulations. *Phys. Chem. Chem. Phys.* **17**, 12857–12869 (2015).

96. Ishida, T., Horiike, K., Tojo, H. & Nozaki, M. Interaction of protein with a self-associating ligand. Deviation from a hyperbolic binding curve and the appearance of apparent co-operativity in the Scatchard plot. *J. Theor. Biol.* **130**, 49–66 (1988).

97. Patterson, S. E., Coxon, J. M. & Strekowski, L. Intercalation of ethidium and analogues with nucleic acids: a molecular orbital study. *Bioorg. Med. Chem.* **5**, 277–281 (1997).

98. Calabrese, D. R. *et al.* Chemical and structural studies provide a mechanistic basis for recognition of the MYC G-quadruplex. *Nat. Commun. 2018 91* **9**, 1–15 (2018).

99. Dickerhoff, J., Dai, J. & Yang, D. Structural recognition of the MYC promoter G-quadruplex by a quinoline derivative: insights into molecular targeting of

parallel G-quadruplexes. *Nucleic Acids Res.* **49**, 5905–5915 (2021).

100. Criscuolo, A. *et al.* Insights into the Small Molecule Targeting of Biologically Relevant G-Quadruplexes: An Overview of NMR and Crystal Structures. *Pharmaceutics* **14**, 2361 (2022).

TABLES

SCIENTIFIC ACHIEVEMENTS

Journal articles

1. A. Potęga, D. Rafalska, D. Kazimierczyk, **M. Kosno**, A. Pawłowicz, W. Andrałojć, E. Paluszkiewicz & T. Laskowski: In vitro enzyme kinetics and NMR-based product elucidation for glutathione S-conjugation of the anticancer unsymmetrical bisacridine C-2028 in liver microsomes and cytosol: major role of glutathione S-transferase M1-1 isoenzyme; *MOLECULES* (submitted on Sep 01, 2023)

2. T. Laskowski, **M. Kosno**, W. Andrałojć, J. E. Frackowiak, J. Borzyszkowska--Bukowska, P. Szczeblewski, N. Radoń, M. Świerżewska, A. Woźny, E. Paluszkiewicz, Z. Mazerska: The interactions of monomeric acridines and unsymmetrical bisacridines (UAs) with DNA duplexes: an insight provided by NMR and MD studies; *Scientific Reports* 2023.

3. P. Szczeblewski, M. Wróblewski, J. Borzyszkowska-Bukowska, T. Bairamova, J. Górska, T. Laskowski, A. Samulewicz, **M. Kosno**, Ł. Sobiech, J. Polit, Kukula--Koch: The role of centrifugal partition chromatography in the removal of β-asarone from Acorus calamus essential oil; *Scientific Reports* 2022.

4. **M. Kosno**, T. Laskowski, J. E. Frackowiak, A. Potęga, A. Kurdyn, W. Andrałojć, J. Borzyszkowska-Bukowska, K Szwarc-Karabyka, Z. Mazerska: Extensive chemometric analysis exploring the acid-base equilibrium and self-association of unsymmetrical bisacridines reveals characteristics possibly relevant for their anticancer activity; *Molecules* – 2022.

5. A. Potęga, **M. Kosno**, Z. Mazerska: Novel insights into conjugation of antitumor-active unsymmetrical bisacridine C-2028 with glutathione: characteristics of non-enzymatic and glutathione S-transferase-mediated reactions; J*ournal of Pharmaceutical Analysis* - 2021.

6. A. Mieszkowska, A. Nowicka, A. Kowalczyk, A. Potęga, M. Pawłowska, **M. Kosno**, E. Augustin Z. Mazerska: Metabolic Profiles of New Unsymmetrical Bisacridine Antitumor Agents in Electrochemical and Enzymatic Noncellular Systems and in Tumor Cells; *Pharmaceuticals* - 2021.

7. E. Wagner-Wysiecka, P.Szulc, E. luboch, J. Chojnacki, K. Szwarc-Karabyka, N. Łukasik, M. Murawski, **M. Kosno**: Photochemical Rearrangement of a 19-Membered Azoxybenzocrown: Products and their Properties; *ChemPlusChem* - 2020.

Conferences

1. 10th Conversatory on Medicinal Chemistry; Polish Society of Medical Chemistry; Lublin, Poland; 03 – 05. 09. 2021 Unique physicochemical properties of the unsymmetrical bisacridines and their importance.

2. 26th The Biochemistry Global Summit; Portuguese Biochemical Society; Lisbon, Portugal 09 – 14. 07. 2022; Preliminary studies on the interactions between antitumor unsymmetrical bisacridines, UAs, and DNA using monomer acridine derivatives as model compounds.

3. Chemometrics and metrology in analytics; Zakopane, Poland; 16 – 18. 10. 2022; Application of chemometric methods in self-association study of bisacridine anticancer compounds.

4. 11th Conversatory on Medicinal Chemistry; Polish Society of Medical Chemistry; Lublin, Poland; 14 – 05. 16. 2023 Interactions of antitumor unsymmetrical bisacridines, UAs, with C-Myc protooncogene promoter studied by NMR spectroscopy, enhanced by MD techniques.