Structure elucidation of artificial, self-assembling squalene-conjugates and β-peptides

PhD thesis

Dr. Dóra Bogdán

Semmelweis University
Doctoral School of Pharmaceutical Sciences

Supervisor:	Dr. István Mándity, PhD
Consultant:	Dr. Tamás Gáti, PhD

Official reviewers:	Dr. Gyula Batta, DSc., PhD
	Dr. Márta Mazák-Kraszni, PhD

Head of the final examination committee: Dr. Klebovich Imre, DSc., PhD
Members of the final examination committee: Dr. Tóth Gábor, DSc., PhD
Dr. Hosztafi Sándor, PhD

Budapest
2019

1. Introduction

Ecdysteroids are defined as compounds structurally related to ecdysone. Later, real edcysteroids and ecdysteroid-like compounds were distinguished. Compounds belong to the group of real ecdysteroids, if the A/B ring connection is *cis* in the steroid moiety (cyclopentano-perhydrophenantrene skeleton) and if the 14α-hydroxy-7-ene-6-on chromophores are also present; the classification does not consider pharmacological properties. Ecdysteroid-like compounds do not fulfil the abovementioned structural requirements, those compounds have only structural relationship to ecdysone.

In my PhD thesis the investigated ecdysteroids are the synthetically modified derivatives of posterone. Posterone is a C21 ecdysteroid, an acetyl group is connected to C17 in β-position, the skeleton is similar to β-androstane. Anellation of A and B rings is *cis*.

Some investigated ecdysteroid compounds were oxime derivatives. Oximes and oxime-ethers are $R_2C=NOH$ / $R_2C=OR$ compounds, which are derived from the condensation of aldehydes or ketones with hydroxylamine or its derivatives. Oximes, which are formed from aldehydes, are aldoximes, which are synthesized from ketones, are ketoximes. Two configurational isomers are possible around the C=N double bond: *E* and *Z* isomers, also known as *syn* and *anti* isomers in case of aldoximes.

By the formation of nano-assemblies, several feature of a drug molecule could be optimized: controlled liberation and distribution, enhancement of absorption or protection against degradation.

Squalenoylated drugs are prodrugs. In our previous work, we applied two different types of linkers: one only with CH_2 units and another with S-S disulfide bridge in the middle of this unit. The squalene part is responsible for self-association and protection of the bioactive molecule, the linker part plays role in intracellular liberation. If the connection is pH sensitive between linker and drug molecule, this feature could be used for targeted liberation in cancer cells, in which there is acidic environment. The use of the linker with disulfide bond was found rather advantageous to release the drug molecules from a prodrug form mainly in the inner environment of cancer cells. Increased concentration of glutathione cause a cleavage in the linker, then the drug molecule liberates during the formation of a thiolactone ring.

Various supramolecular systems could be built from squalene-conjugates: micelles/liposomes, lamellar bilayers, cubic, hexagonal or amorphous structures.

Foldamers are biomimetic polymers, which fold into conformally ordered structure with the help of noncovalent interactions. One of the most important group of peptidomimetic foldamers is the group of β-peptides. β-Peptides are built from β-amino acids, in which amino

group is bonded to the β-carbon. The α- and β-carbon atoms could be variously substituted, high structural diversity is available and this effects on conformal behaviour. β-Peptides have also high conformal stability, the same degree of order could be built from shorter sequences than from α-peptides.

Helices, sheets from filaments and loops are possible secondary structures. Different helical systems could be formed: H14, H12, H10/12 and H14/16 helices based on the number of members in the pseudoring by H-bonds.

The *trans*-ACHC (*trans*-2-aminocyclohexanecarboxylic acid) is one of the most often used building block for H14 helix formation. The formation of an organized helical structure is possible for ACHC tetramers. In case of *cis*-ACHC oligomers with heterochiral alternating stereochemical pattern, H10/12 helices are formed. Self-association was detected for this β-peptides in polar solvents.

In case of *diexo*-ABHEC (*diexo*-3-aminobicyclo[2.2.1]hept-5-ene-2-carboxylic acid) a filament-like secondary structure is found in sequences built from 2*S*,3*R* or 2*R*,3*S* isomers based on NH^i – $C_\beta H^{i+1}$ and NH^i – $C_\beta H^{i-1}$ NOE interactions. In alternating heterochiral sequences a circle-like fold was shown with H-bonds between NH^i – $C=O^i$ atoms.

For dimers of oxanorbonene derivatives – *diexo*-3-amino-7-oxabicyclo[2.2.1]hept-5-ene-2-carboxylic acid (*diexo*-AOBHEC) – H8 helical conformation was detected.

2. Aims

Our aim was the structure elucidation of self-organising systems and to investigate the ordered conformation mainly by NMR spectroscopy.

The first step was to make the assignment of the investigated compounds. The structures were identified and the relative configuration of the chiral carbon atoms were determined.

Complete ^{1}H and ^{13}C assignment were made for squalenoylated ecdysteroids and free ecdysteroids, the diastereopic hydrogens were also differentiated.

It was a challenging task to make the complete assignment for the squalene-linker side chain. The repeating structural elements are in a very similar chemical environment, therefore the chemical shifts will be very close to each other, hard to differentiate them; in addition to this, the steroid signals also overlap with these signals in the spectrum. Our aim was to elaborate a strategy for this problem, and to use methods with higher resolution.

In case of ecdysteroids we have an additional task in structure elucidation. Some molecules were oximes, therefore we had to identify the *E* and *Z* isomers. The configuration of A/B ring connection in the steroid skeleton is also had to be determined.

The next task was to investigate the self-organisation. In case of squalene-conjugates, nano-sized particles are expected in aqueous systems. For this purpose also NMR spectroscopic methods were used.

Backbone protons were identified in β-peptides. Our aim was to investigate the self-organisation in three different solvents, in H_2O/D_2O 90:10 mixture, CD_3OH and DMSO-d_6. In this study helical and unordered/less stable ordered structures were identified by NMR spectroscopy and CD method.

3. Methods

3.1 NMR spectroscopy

In case of ecdysteroids and squalene conjugates 1-5 mg of the samples were dissolved in 0.6 ml deuterated chloroform. β-Peptide NMR samples were prepared in 4 mM concentration in 0.5 ml CD_3OH, DMSO-d_6 and water (H_2O/D_2O 90:10 v/v%). CD_3OD was used to determine NH-ND exchange. For nanoassemblies of squalene-conjugates a stock solution was made in deuterated acetone, this solution was added to deuterium oxide, then acetone was removed under vacuum, the resulting sample was measured in a Shigemi-tube. The diffusion of the monomer of squalene-conjugate was investigated in CD_3OD. All samples were transferred into a 5 mm NMR tube (except for squalene-conjugate nanoassemblies).

NMR measurements were taken on Bruker 950/239 MHz, 800/200 MHz and a 500/125 MHz spectrometers equipped with cryoprobe and a 400/100 MHz spectrometer with room temperature probe. All pulse programs were applied from the Bruker and Varian software library.

In case of ecdysteroids and their derivatives for 1D measurement, 64K data points were used to yield the FID. Spectral widths for the 950 MHz ^1H spectra were set to 7600 Hz, whereas in case of other spectrometers to 6500 Hz, 64K data points and 32 scans were used. For ^{13}C measurements, a sweep width of 48000 Hz, 128K data points 64 scans; for DEPTQ a sweep width of 48000 Hz, 64K data points and 64 scans were applied. For 2D measurement, in case HSQC the following parameters were used: sweep width in F2 dimension: 7000 Hz or 7600 Hz (950/239 MHz); 2K x 128 or 2K x 4K (950/239 MHz) data points (t_2 x t_1). HMBC measurements were acquired with 4000 Hz or 7600 Hz (950/239 MHz) F2 sweep width and 2K x 256 or 16K x 8K (950/239 MHz) data points (t_2 x t_1). In the band-selective HSQC and HMBC experiments data points were 2K x 128 (t_2 x t_1), sweep width was used in the selected region. Selective 1D ROESY was taken with 300 ms mixing time, 64 scans and 32K data points. Selective 1D TOCSY was acquired with 128K data points, 7600 Hz sweep width and 8 scans. In DOSY measurements F2 sweep width set to 7500 Hz, 128 scans (for nano-assemblies) or 16 scans (for monomer and calibration) and 32K x 16 (t_2 x t_1) data points were used. The gradient length was 1.5 — 2 ms, the diffusion time 100 – 200 ms. Calibration was made on methanol-D_2O mixture.

In case of β-peptides 64 scans, 12000 Hz sweep width and 64K data points were the parameters for ^1H measurements. 2D-TOCSY experiments were taken with 80 or 120 ms mixing time, 32 scans and 2K x 256 (t_2 x t_1) data points. 2D ROESY experiments were

acquired with 300 or 400 ms mixing time, 32 scans and 2K x 256 (t_2 x t_1) data points. For solvent suppression gradient water suppression techniques were used: p3919gp (water suppression using 3-9-19 pulse sequence with gradients) and zgesp (water suppression using excitation sculpting with gradients.

TopSpin 3.5 software was used for workup. Chemical shifts are given on the δ-scale and are referenced to the solvents: CDCl$_3$: ^{1}H: δ= 7.27 ppm; ^{13}C: δ=77.00 ppm; DMSO-d_6: ^{1}H: δ= 2.50 ppm; CD$_3$OD/CD$_3$OH: ^{1}H: δ= 3.31 ppm; D$_2$O: ^{1}H: δ= 4.79 ppm. Most ^{1}H assignments were accomplished using general knowledge of chemical shift dispersion with the aid of the proton-proton coupling pattern (^{1}H NMR spectra). The NMR signals of the products were assigned by comprehensive one- and two-dimensional NMR methods using widely accepted strategies.

3.2 ECD spectroscopy

The concentration of the sample solutions was 1 mM in CH$_3$OH and H$_2$O. 0.1 cell was used for measurements.

CD spectra were measured on a Jasco J715 dichrograph at 298K. Three spectra were accumulated for each sample.

CD curves were corrected by the spectral contribution of the blank solvent. Jasco Spectra Manager software was applied for workup of spectra.

4. Results

4.1 Structure elucidation of ecdysteroids and ecdysteroid-squalene-conjugates

- In my PhD thesis two ecdysteroids and their squalene-conjugates with two-two different linker regions were studied.
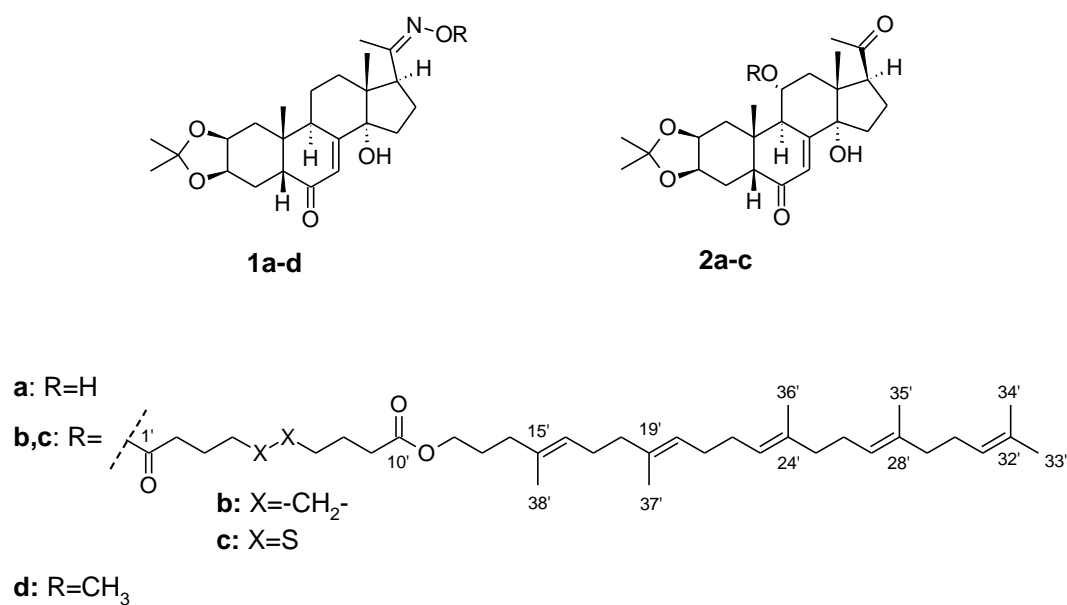

Figure 1: Structures of the investigated compounds

- A complete ^{1}H and ^{13}C assignment was made for all compounds, diastereotopic hydrogens and groups were differentiated.

- In case of squalene-conjugates derived from the same ecdysteroid, minimal differences were shown for the signals of the same ^{1}H and ^{13}C nuclei in the steroid skeleton. By comparing the chemical shifts of the squalene moiety, the data were also in good agreement.

- **1a** is known in the literature, the *E* isomer is identified by X-ray crystallography, its NMR assignment is written in solution in deuterated chloroform. Differences were found in our NMR assignment. C-11 and C-16 were misassigned in the literature. In our HMBC spectrum these ^{13}C signals could be unequivocally differentiated. This misassignation was corrected and the NMR assignment was completed by the identification of diastereotopic hydrogens.

- In our first experiments, a spectrometer at 500/125 MHz was used. Unfortunately, mainly in case of squalene signals the resolution was not enough for the unequivocal

assignment. By the use of band-selective and ultrahigh field measurements the digital resolution could be enhanced up to 19-times.
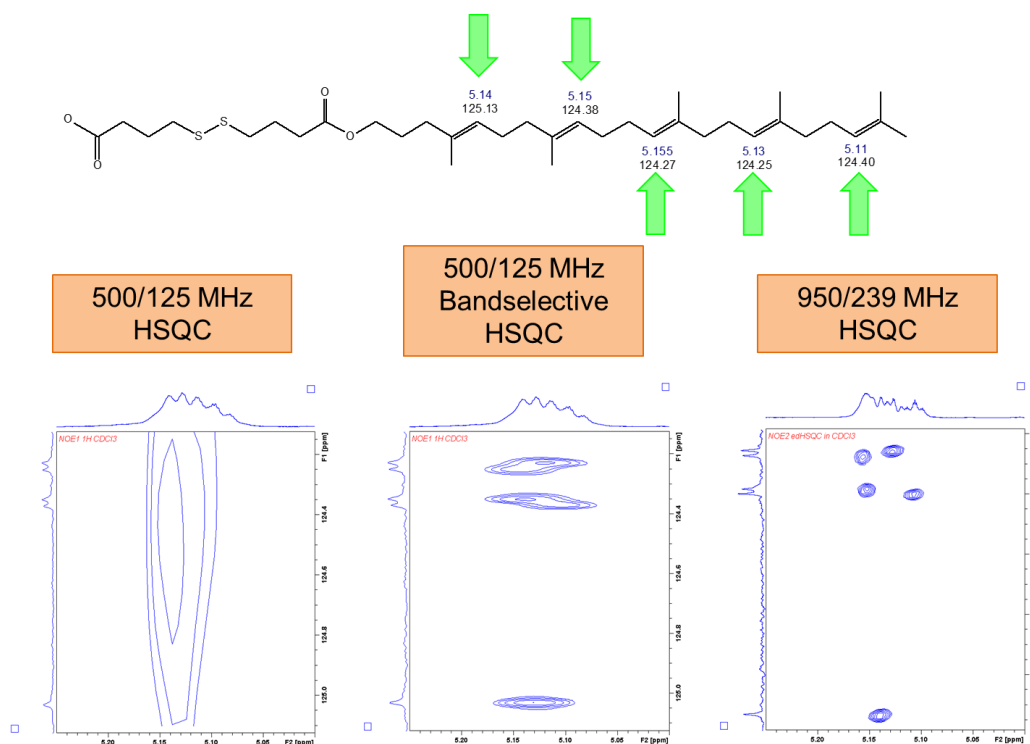

Figure 2: Differentation of squalene sp^2 CH groups in the HSQC spectrum

- Methylene hydrogens in the steroid skeleton are diastereotopic, their chemical shifts are different. In order to determine α or β position selective 1D ROESY measurements were applied.
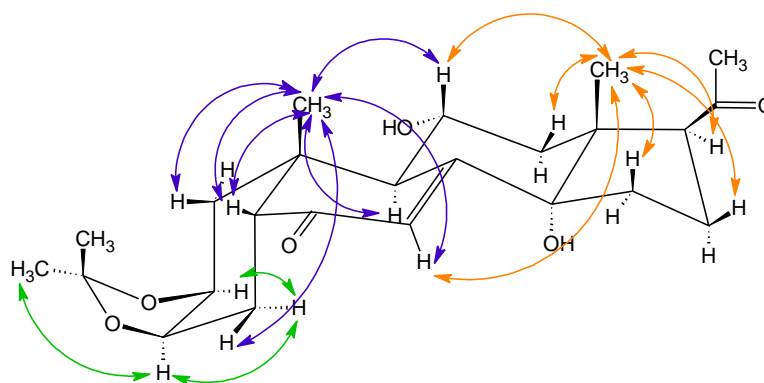

*Figure 3: **2a** most important ROESY interactions by irradiation of H₃-18 (orange) H₃-19 (blue) and H-2 + H-3 (green).*

- No remarkable NOE sterical interaction could be observed between the broad OH and the H_3-21 hydrogen atoms, not even when DMSO-d_6 was used as solvent. As only one signal set was present in the spectrum we had no information about $\Delta\delta$*syn-anti* parameter of C-21. The oxime-methylether was investigated to overcome this problem. Now the $NOCH_3 \leftrightarrow CH_3$-21 NOE interaction could be measured and two signal sets appeared in the spectrum in 95% - 5% ratio. Both ROESY, and investigation of $\Delta\delta$*syn-anti* parameter confirmed the presence of *E* isomer. Chemical shift of C-21 was 16.02 ppm in the *E* isomer and 20.26 ppm in the *Z* isomer, therefore a $\Delta\delta$*syn-anti* parameter of 4.24 ppm, is in good agreement with the literature. In the solution of **1a-c** only one signal set is observed with these chemical shifts of C-21: 15.53 ppm, 17.48 ppm és 17.44 ppm. These data are correlated to the chemical of *E* isomer, these compounds were determined as *E* isomers.

- Squalene-conjugates in deuterium oxide were detected as large associates based on NMR spectra. Two methods were used for the calculation of the number of monomers in the supramolecular assembly. The two results differ (1000 vs. 2600 number of monomers, 9.28 nm hydrodynamic radius), this discrepancy could be because we handled the assembly and the monomer as spherical particles. Other squalene-conjugates known in the literature have much bigger diameter (200 – 300 nm) based on DLS and TEM measurements. These bigger particles are supposed to be built from smaller subparticles, and we probably determined these subparticles by DOSY NMR spectroscopy.

4.2. Structure elucidation of β-peptides

- Six different β-peptides were investigated, these compounds were built from *trans*-ACHC units and other amino acids as third building block.
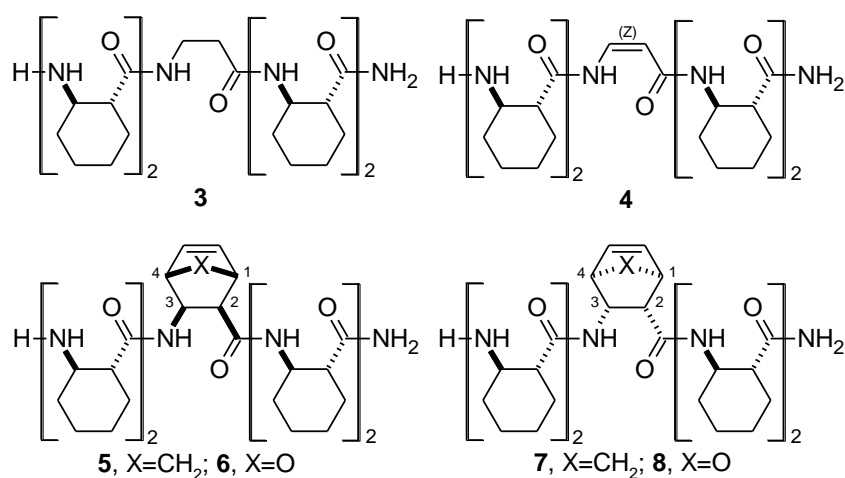

Figure 4: The investigated β-peptides

- The NH signals were found with good dispersion in the ^{1}H spectrum, which is characteristic for ordered secondary structures. The NH, $C_\alpha H$ and $C_\beta H$ hydrogen signals were in the expected chemical shift regions.

- Self-organization was investigated in details for the six β-peptide pentamers. For **3**, **5** and **6** characteristic NOE-correlations for ordered conformation were shown in ROESY spectra. The $NH^i - C_\alpha H^{i+2}$ and $C_\alpha H^i - C_\beta H^{i+2}$ interactions were identified, of which are characteristic for H14 helix.
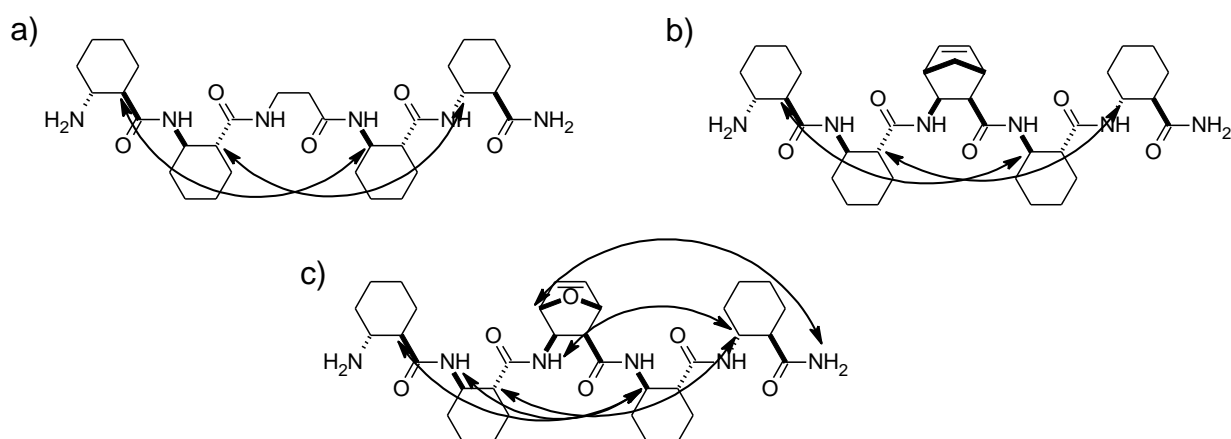

*Figure 5: NOE-interactions in **3** (a), **5** (b) and **6** (c); 4mM; 298K*

- In case of **4**, **7** and **8** only weak cross-peaks were found in methanol, therefore less stable ordered secondary structure could be observed based on ROESY NMR measurements. In water and DMSO-d_6 no correlations were found typical to ordered conformation, thus the helical structure is not favourable in these solvents.

- Amide NH-ND exchange investigations are in good agreement with ROESY NMR measurements. For **7** an immediate proton exchange was observed, indicating an unfolded structure. For **4** and **8** proton exchange occurred in ca. 1 hour, demonstrating weak structure stabilizing H-bonds and poorly shielded amide protons. Significantly longer proton exchange times (10 hours, over 1 day) were shown for **3**, **5** and **6** structures, these results may suggest highly shielded NH protons from the solvent and stable H-bonds which indicate a helical secondary structure.
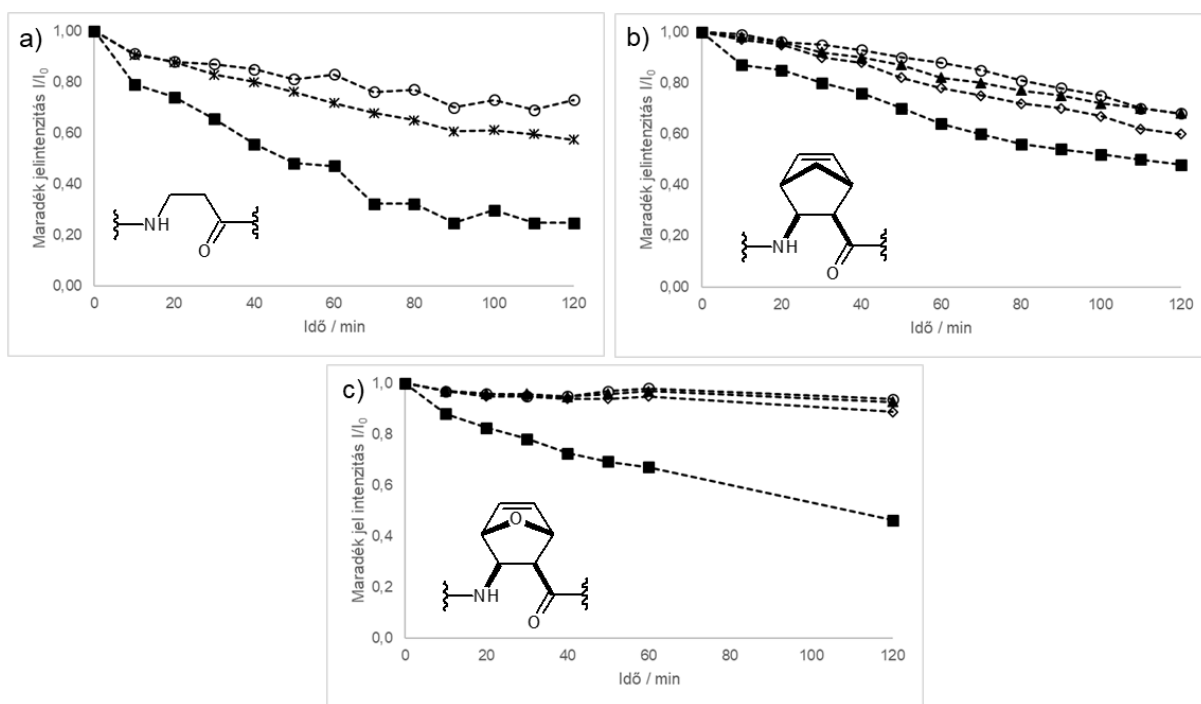

*Figure 6: NH-ND exchange diagram for **3** (a), **5** (b) and **6** (c); 4 mM; 298K*
*■: NH₂; ○:NH₃; ▲ : NH₄; ◊: NH₅, *: NH₄+NH₅*

- Electronic circular dichroism spectra were recorded in water and methanol at room temperature, for **3**, **5** and **6** the intensive CD bands were characteristic for H14 helix. The decreased intensities for **4**, **7** and **8** suggest the reduced secondary structure stability. The quite identical absorption maxima and minima in water and methanol proved the absence of solvent-driven structural changes.
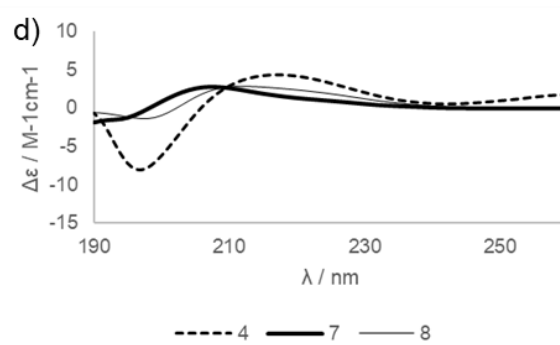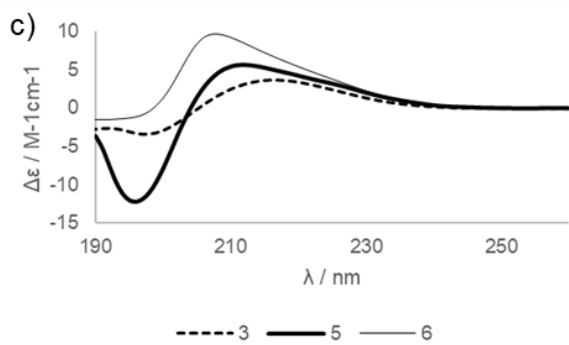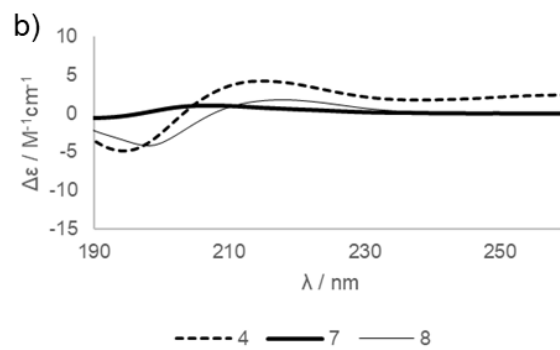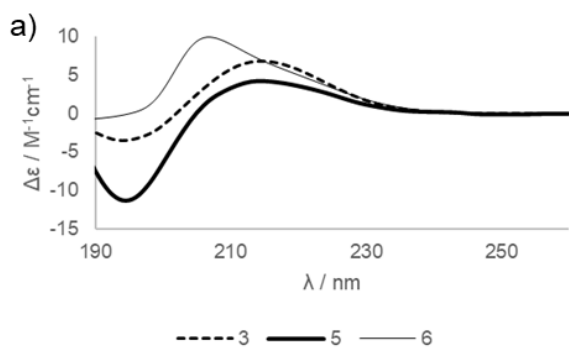
Figure 7: ECD spectra of the investigates β-peptides; 1mM; 298K
a), b) in water
c), d) in methanol

5. Conclusions

In case of ecdysteroids and their squalene-conjugates minimal differences were detected in the chemical shifts of the same steroid moiety or the squalene-linker-side chain. The complete assignment of newly synthesized squalene-conjugates could be faster accomplished based on our results. A new $\Delta\delta$*syn-anti* parameter of C_α atom was determined for an oxime compound and a combined synthetic-NMR spectroscopic method was described in order to identify the configuration.

The formation of nanoassemblies was verified in deuterium oxide. The investigation of secondary structure could not be performed because of formation of nanoassemblies. A nanoparticle with ca. 1000 monomers was defined by DOSY measurements. This type of nanoassembly has not been determined yet by NMR spectroscopy, our results highlighted a method for investigating squalene-conjugate particles.

In case of β-peptides the built-in of a different β-amino acid into the *trans*-ACHC sequence cause a definitive change in conformation. The flexible β-alanine did not influence the formation of H14 helix from *trans*-ACHC. On the other hand, the unsaturated *Z*-dehydro-β-alanine prohibited the ordered secondary structure, a possible explanation is the conjugation between the C=C double bond and neighbouring amide bonds. Our investigations on cyclic amino acids showed an effect of configuration on helical structure stability. When [1*R*,2*R*,3*S*,4*S*]-*diexo*-ABHEC and [1*R*,2*S*,3*R*,4*S*]-*diexo*-AOBHEC were used as third building block H14 helix was found. But by the use of the enantiomeric pairs of these amino acids, as [1*S*,2*S*,3*R*,4*R*]-*diexo*-ABHEC and [1*S*,2*R*,3*S*,4*R*]-*diexo*-AOBHEC, no stable ordered secondary structure was found. Our results showed, that some types of β-amino acids can stabilize the H14 helix of *trans*-ACHC, some other building blocks can interrupt this ordered conformation. Comparing ABHEC and AOBHEC units, no effect of *O*-atom was found. In designing new peptides this information could be used in favour of the required secondary structure.

List of publications

I. Publications related to the present thesis

1. **D. Bogdán**, R. Haessner, M. Vágvölgyi, D. Passarella, A. Hunyadi, T. Gáti, G. Tóth. Stereochemistry and complete ^{1}H and ^{13}C NMR signal assignment of C-20-oxime derivatives of posterone 2,3-acetonide in solution state, Magn. Reson. Chem. *56* (2018) 859-866
   DOI: 10.1002/mrc.4750
   IF: 1.78

2. I. Nekkaa[1], **D. Bogdán[1]**, T. Gáti, Sz. Béni, T. Juhász, M. Palkó, G. Paragi, G. K. Tóth, F. Fülöp, I. M. Mándity. Flow-chemistry enabled efficient synthesis of β-peptides: backbone topology vs. helix formation, ChemComm. *55* (2019) 3061-3064, DOI: 10.1039/C8CC10147G
   IF: 6.29
   [1]: Authors contributed equically in this work

II. Publications not related to the present thesis

1. M. Atia, **D. Bogdán**, M. Brügger, N. Haider, P. Mátyus. Remarkable regioselectivities in the course of the synthesis of two new Luotonin A derivatives, Tetrahedron *73* (2017) 3231-3239

2. Á. Horváth, A. Menghis, B. Botz, É. Borbély, Á. Kemény, V. Tékus, J. Zs. Csepregi, A. Mócsai, T. Juhász, R. Zákány, **D. Bogdán**, P. Mátyus, J. Keeble, E. Pintér & Zs. Helyes. Analgesic and Anti-Inflammatory Effects of the Novel Semicarbazide-Sensitive Amine-Oxidase Inhibitor SzV-1287 in Chronic Arthritis Models of the Mouse, Nature Scientific Reports *7* (2017) 39863

3. R. Meleddu, S. Distinto, A. Corona, G. Bianco, V. Cannas, F. Esposito, A. Artese, S. Alcaro, P. Matyus, **D. Bogdan**, F. Cottiglia, E. Tramontano, E. Maccioni. (3Z)-3-(2-[4-(aryl)-1,3-thiazol-2-yl]hydrazin-1-ylidene)-2,3-dihydro-1H-indol-2-one derivatives as dual inhibitors of HIV-1 reverse transcriptase, Eur. J. Med. Chem. *93* (2015) 452-460