Synthesis of proteophenes that can be utilized as fluorescent ligands for biological targets

Linnea Björk

Examiner
Professor K. Peter R. Nilsson

Supervisors
Linda Lantz
Marcus Bäck
Synthesis of proteophenes that can be utilized as fluorescent ligands for biological targets

Abstract
Small fluorescent probes are important tools when studying protein aggregates involved in different neurodegenerative diseases, such as Alzheimer’s disease. Luminescent conjugated oligothiophenes have been developed and shown to be excellent ligands when studying morphology among amyloids, due to their conjugated thiophene backbone that provides them with unique photophysical properties. This kind of probes are being developed successively to enhance the specificity of their biological targets. In this project, luminescent conjugated oligothiophenes functionalized with amino acids, so called proteophenes, have been synthesized to investigate their optical properties. Since amino acids are chiral molecules, the possibility of induced chirality to the thiophene backbone was examined, as well as the proteophenes ability to work as amyloidospecific ligands for the study of protein aggregates. The synthesis of four different proteophenes are presented in this report, along with analysis results of their photophysical properties.

Keyword
Fluorescent probes, Neuredegenerative diseases, Luminescent Conjugated Oligothiophenes, Protein misfolding, Aggregates, Amyloids, Alzheimer’s disease, Aβ-plaque
Abstract

Small fluorescent probes are important tools when studying protein aggregates involved in different neurodegenerative diseases, such as Alzheimer’s disease. Luminescent conjugated oligothiophenes have been developed and shown to be excellent ligands when studying morphology among amyloids, due to their conjugated thiophene backbone that provides them with unique photophysical properties. This kind of probes are being developed successively to enhance the specificity of their biological targets. In this project, luminescent conjugated oligothiophenes functionalized with amino acids, so called proteophenes, have been synthesized to investigate their optical properties. Since amino acids are chiral molecules, the possibility of induced chirality to the thiophene backbone was examined, as well as the proteophenes ability to work as amyloidospecific ligands for the study of protein aggregates. The synthesis of four different proteophenes are presented in this report, along with analysis results of their photophysical properties.
Abbreviations

AcN: Acetonitrile
AD: Alzheimer disease
ADI: Alzheimer’s Disease International
Aβ: Amyloid beta
BF₃·OEt₂: Boron trifluoride diethyl etherate
BTD: 2,1,3-benzothiadiazole
CD: Circular dichroism
DCM: Dichloromethane
DIPEA: N,N-Diisopropylethylamine
DMF: Dimethylformamide
E/EtOAc: Ethyl acetate
FC: Flash column chromatography
H: Heptane
H₂SO₄: Sulfuric acid
HCl: Hydrochloric acid
HOAc: Acetic acid
HPLC: High Performance Liquid Chromatography
K₂CO₃: Potassium carbonate
LCO: Luminescent conjugated oligothiophene
LiOH: Lithium hydroxide
Lys: Lysine
MeOH: Methanol
MgSO₄: Magnesium sulfate
MS: Mass spectrometry
NaCl: Sodium chloride
NaHCO₃: Sodium bicarbonate
NaOH: Sodium hydroxide
NBS: N-bromo succinimide
NFTs: Neurofibrillary tangles
NH₄OAc: Ammonium acetate
NMR: Nuclear magnetic resonance
PBS: Phosphate-buffered Saline
T: Toluene
TBTA: t-butyl-2,2,2-trichloroacetimidate
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin-layer chromatography</td>
</tr>
<tr>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
</tbody>
</table>
Table of contents

1 Introduction ....................................................................................................................... 1
  1.1 Neurodegenerative diseases ....................................................................................... 1
      Protein misfolding and amyloid disease ........................................................................ 1
  1.2 Luminescent conjugated oligothiophenes ................................................................. 2
      Proteophenes ................................................................................................................. 3
  1.3 Photophysical properties ......................................................................................... 4
      Fluorescence ................................................................................................................. 4
      Circular dichroism ......................................................................................................... 5
  1.4 Aim ............................................................................................................................ 5
  1.5 Synthetic approach and target molecules ................................................................. 5

2 Results and discussion .................................................................................................. 9
  2.1 Synthesis of compound 6 ......................................................................................... 9
  2.2 Synthesis of target compounds 9, 10, 11 and 12 .................................................... 12
  2.3 Photophysical properties ......................................................................................... 16

3 Conclusions and future perspectives ............................................................................. 20
Acknowledgements .......................................................................................................... 21
References ........................................................................................................................ 22
Experimental details ......................................................................................................... 24
1 Introduction

This chapter of the thesis encompasses the background of the project, describing neurodegenerative diseases and their challenges, as well as an introduction to Alzheimer’s disease (AD) and other protein misfolding diseases and why studying aggregated proteins are of great importance. The background of fluorescent molecules as amyloidospecific ligands and especially luminescent conjugated oligothiophenes (LCOs) are described since LCOs together with amino acids are the synthetic focus of this project. The synthesis of new molecules that can be utilized as fluorescent ligands for biological targets are of high importance since new findings can help in the fight against many neurodegenerative diseases.

The theory behind the photophysical properties of the target molecules is also included in this section, which is ended with the synthetic route for reaching the target molecules.

1.1 Neurodegenerative diseases

The brain is a fascinating organ, consisting of up to over 80 billion neurons, being the “control centre” of human beings. The brain contains much of what makes us who we are, storing all our memories and knowledge, defining us as individuals. Therefore, it does not come as a surprise that diseases disturbing the function of our brains are of great interest when it comes to research concerning society and health around the world. Since the brain is the most complex organ in our bodies there are great challenges ahead when facing the diseases affecting it. Especially neurodegenerative diseases, diseases causing neuronal cell death in the central nervous system, have shown to be problematic. The characteristics of these disorders are the degradation of cells in the central nervous system, linked to neuronal malfunction and the loss of specific brain functions, which have extreme consequences on health, giving rise to both physical and mental disabilities.\(^1\)

One of the main pathways involved in neurodegenerative diseases is the misfolding, aggregation, and accumulation of proteins in the brain. When a peptide or a protein fails to adopt or sustain its native functional structure it can cause pathogenesis, which are commonly referred to as protein misfolding diseases.\(^2\)

Protein misfolding and amyloid disease

Proteins play a huge role in all living organisms where they participate in almost all biochemical reactions necessary for life. Proteins are consisting of a mixture of different amino acids that are connected via amide bonds, also called peptide bonds. Proteins have different levels of structural conformation, from the primary structure with its specific set of amino acids, to the secondary structure where the protein folds in to \(\alpha\)-helices and \(\beta\)-strands, before they fully fold in to their tertiary monomeric structure and eventually, if possible, builds up into a complex of numerous tertiary proteins forming the quaternary structure. Despite a lot of regulatory processes and protein quality control during the folding procedure the proteins are sometimes misfolded in to wrong conformations during the many steps of the folding process. This can cause the protein to lose its natural function, which in turn can cause the protein to aggregate and accumulate, forming possibly harmful aggregates called amyloids.\(^3,4\)

Amyloids are highly organized fibrils with a typical cross-\(\beta\) structure, a structure characteristic for amyloid fibrils regardless of the precursor protein. The accumulation of harmful misfolded proteins is
associated with many diseases, especially neurodegenerative diseases such as Alzheimer’s disease, Parkinson’s disease, Huntington’s disease, and many more. In AD, mainly two different types of aggregated proteins are found: extracellular neuritic plaque of amyloid β-peptide (Aβ) and intracellular neurofibrillary tangles (NFTs) of hyperphosphorylated tau protein.\(^3\)

Alzheimer’s disease is a very common disease causing dementia and is believed to account for 50-70% of all cases of dementia today. According to Alzheimer’s Disease International, ADI, a new person is affected by the disease every three seconds. Advanced age is the main risk factor for developing dementia and the majority of the affected are 65 years old or older. Among the many symptoms for AD, loss of short-term memory, cognitive impairments and eventually dementia, are included. The molecular pathways of AD are still not fully understood, but some pathological hallmarks have been found as previously mentioned. Therefore, studying aggregated proteins are of great importance for the understanding of many neurodegenerative diseases.\(^1\)

Amyloid-specific ligands have long been used to study protein aggregation. The first molecular dyes used for the detection of amyloids were Congo red (Figure 1, A) and Thioflavin T (Figure 1, B). These molecular dyes interact with the repetitive pattern of the cross-β structure in the amyloid aggregates. The identification of the aggregates is based on the changes in fluorescence of Thioflavin T, and the changes in absorbance of Congo red, upon binding.\(^3,5\)

These molecules, and their derivatives, are still used today as tools for detecting protein aggregates. However, there are some downsides with these amyloid ligands since they have limited ability to distinguish between different morphotypes among the aggregates. The development of new amyloid-specific ligands is an ongoing process and recently a novel class of luminescent conjugated oligothiophenes have been shown to have increased specificity for amyloid aggregates.\(^2,3\)

![Figure 1. Structures of the molecular dyes A) Congo red and B) Thioflavin T.](image)

### 1.2 Luminescent conjugated oligothiophenes

Thiophenes are aromatic heterocycles containing a sulfur atom. The thiophenes can undergo electrophilic aromatic substitution reactions at both its α- and β-positions, where the α-position is preferred because of its higher reactivity. This makes way for the thiophenes to link together at the α-positions, creating polymers or oligomers that are highly conjugated. Because of the conjugation, delocalized π-electrons can move through the alternating single- and double bonds in the structure,
which makes the conjugated thiophenes excellent chromophores as they can absorb and emit light. The luminescence of the molecules can be used as a detection method when interacting with different targets and is one of the reasons why LCOs are being developed as amyloid ligands.\textsuperscript{3}

Another interesting feature of LCOs is the rotational freedom around the single bonds linking the thiophene monomers together. This flexibility allows the structure to twist out of its plane, changing the total effective conjugation. When the structure is twisted, the conjugation is shortened which can be illustrated by the shifted colour of the emitted fluorescence signal. The twisted conformation gives emission at shorter wavelengths compared to the planar conjugation. This quality can be useful when studying the conformational changes of LCOs upon binding to different target structures. It also enables one LCO to optically distinguish between different amyloid morphotypes in a sample.\textsuperscript{3,6}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Conformational twist of the thiophene backbone changes the effective conjugation length which is demonstrated by change of colour of the emitted fluorescence signal.\textsuperscript{6}}
\end{figure}

By alternating the length of the thiophene backbone, together with different functionalized side chains at diverse positions, the LCOs can display different properties. The changes can affect the solubility and biocompatibility, as well as the selectivity for different targets. Even small variations in the structure can have a great impact on the interaction with the target.\textsuperscript{7}

By achieving these alterations in the structure, heterogeneous populations of aggregates can be assigned. Previous studies have shown that LCOs consisting of at least five thiophene units with carboxyl groups along the conjugated system could spectrally distinguish between A\textsubscript{β}-plaques and tau neurofibrillary tangles in brain tissue with AD pathology. It has also been shown that a mixture of LCOs can be used to monitor age-related structural changes of A\textsubscript{β} aggregates.\textsuperscript{8,9}

**Proteophenes**

The colour shift of the emitted fluorescent light observed as LCOs undergo changes in conformation has been utilized as a tool for optical identification of different morphologies among protein aggregates. LCOs display numerous optical properties in this sense, although they are not optically active. However, associating LCOs to proteins and peptides can generate changes in the thiophene backbone conformation, giving rise to modifications in optical properties. The linking of LCOs to chiral molecules can induce chirality to the ligand, which makes them optically active.\textsuperscript{10}
With the induced chirality, especially in conjugated oligomers such as oligothiophenes, many opportunities follow. For example, to be used as optically active conjugated polymeric materials with semiconducting and optoelectronic properties.\textsuperscript{11} Oligothiophenes with an optically active substituent at distinct positions along the thiophene backbone normally exhibit optical activity in the $\pi-\pi^*$ transition region, derived from the main-chain chirality when the oligothiophenes are forming supramolecular, $\pi$-stacked self-assembled aggregates in a poor solvent or at low temperature.\textsuperscript{12}

Amino acids are a group of organic substances present in all living organisms, since they are the building blocks of proteins and peptides. They are defined as molecules containing both an amine and a carboxylic acid in their chemical structure, along with a carbon-based side chain specific for each amino acid. There are twenty different amino acids occurring in proteins, building up the protein’s primary structure. Although all twenty amino acids have different chemical properties, they have one thing in common (except from the simplest amino acid glycine) – they are chiral. This means they can exist in two mirror images of each other, as enantiomers. Amino acids can either be in the L- or D-form, where L-amino acids are the ones present in naturally occurring proteins. The only physical difference between the enantiomers is that they rotate plane-polarized light in opposite directions, whilst in Nature they behave totally different and the D-enantiomer would not be recognized by the body. Among the twenty amino acids are tyrosine (Tyr) and glutamic acid (Glu), which are the ones in focus in this thesis. Tyrosine (Figure 3, A) is an aromatic amino acid, containing a hydroxyl group on the aromatic ring. Glutamic acid (Figure 3, B) is an amino acid with an acidic side chain. Functionalizing LCOs with these amino acids may induce chirality to the molecule which is one of the things to be investigated in this project. LCOs linked to amino acids are here called proteophenes.\textsuperscript{13}

![Figure 3](image-url). Structure of amino acids A) L-Tyrosine, and B) L-Glutamic acid

### 1.3 Photophysical properties

The target molecules ability to act as chromophores, as well as the possibility of induced chirality to the thiophene backbone are two interesting features of these compounds. To investigate this matter more carefully the theory behind these photophysical phenomena and how they are measured needs to be explained.

#### Fluorescence

Fluorescence has been mentioned already in this report as a tool for studying conformational changes of molecules in different environments. Fluorescence can be described as a phenomenon where light interacts with molecules. A molecule has different energy states, where electrons can excite to or relax from. When electrons in a molecule are excited by absorption of photons, the electrons move to a higher energy level. As the electrons relax down to the lower level again some of the energy gained is released as light, while some of the energy has been consumed by other processes. This means that the molecule absorbs light of higher energy (lower wavelength) and emits light of lower energy (higher
wavelength). This difference is called Stokes shift. The different energy levels and how the electrons excite can be visualised by a Jablonski energy diagram (Figure 4).\textsuperscript{14,15}

![Jablonski energy diagram of fluorescence illustrating absorption of light, vibrational relaxation from higher to lower state of excitation and fluorescence to ground state.\textsuperscript{16}](image)

**Figure 4.** Jablonski energy diagram of fluorescence illustrating absorption of light, vibrational relaxation from higher to lower state of excitation and fluorescence to ground state.\textsuperscript{16}

**Circular dichroism**

Circular dichroism (CD) is a technique that measures the difference in absorbance between right and left circularly polarized light. This method can only be applied to optically active chiral molecules and is based on the ability for chiral molecules to absorb left- (L) and right- (R) handed circularly polarized light to varying extent. This phenomenon is called dichroism. When two plane-polarized waves with the same intensity and wavelength meet in a plane perpendicular to each other circular polarized light is created. This light is formed into a circle, which is either turned clockwise or counter-clockwise giving rise to L- or R-circularly polarized light. If the molecule being studied is chiral one of the two versions of polarized light is absorbed more than the other. The CD instrument measures this differential absorption, yielding a spectrum of the molecule where dichroism is plotted against wavelength. This enables studies of the molecular structure, which is why CD is most commonly used as a method to study the secondary structure of proteins. In this study however, we are using CD measurements to study whether the LCOs have developed chirality after attaching the chiral amino acids, or not.\textsuperscript{17}

**1.4 Aim**

In this project, the aim was to synthesize LCOs functionalized with amino acids, so called proteophenes, for potential use as amyloidospecific ligands. The proteophenes were also to be tested for their photophysical properties to investigate whether they are suitable as fluorescent probes.

**1.5 Synthetic approach and target molecules**

Research on synthesizing LCOs has a long history in the Nilsson group at Linköping University, where the synthetic route for these types of molecules has been developed and optimized along the way. As the research has advanced, so have the structures of the molecules, as well as the methods used to synthesize them. A general synthetic route for synthesizing LCOs has so been developed,
containing repetitive bromination- and Suzuki coupling reactions. Depending on whether the LCO is to be further functionalized or not, protection- and deprotection reactions are also necessary. In this thesis, pentamic LCOs with amino acid side chains were to be synthesized, why amide coupling reactions are also included in the synthetic route. Constructing the symmetric thiophene pentamer 6 (Figure 5), where the acidic side chains are protected with orthogonal protecting groups, is the first target molecule since this molecule can be further functionalized with different amino acids in the different positions.3

Figure 5. Structure of the thiophene pentamer 6, blue arrows showing the orthogonal protecting groups.

The synthesis (Scheme 1) starts with 3-thiopheneacetic acid 1 being methylated via Fischer esterification using methanol in excess, together with catalytic amounts sulfuric acid. The formed methyl ester is further brominated using N-bromo succinimide (NBS) in equivalent amount, resulting in methyl-2-bromo-3-thiopheneacetic acid 2. Compound 2 is Suzuki cross coupled with 5-carboxythiophene-2-boronic acid using the palladium catalyst PEPPSI-IPr in a 3 mol % equivalent ratio, together with potassium carbonate as base, yielding the asymmetric dimer 3. The carboxylic acid on the thiophene dimer is then tert-butylated with tert-butyl-2,2,2-trichloroacetimidate (TBTBTA) and catalytic amounts of boron trifluoride diethyl etherate (BF3.OEt2) to give the protected thiophene dimer 4. Bromination of compound 4 is performed using the same conditions as previously mentioned, yielding compound 5 which is then reacted in a Suzuki cross coupling using this time 2,5-thiophenebisboronic acid to result in the symmetric thiophene pentamer 6.

Scheme 1. Synthesis of the symmetric thiophene pentamer 6. General conditions and reagents: i) 1. MeOH, H2SO4, 80°C; 2. DMF, NBS, -15°C to r.t; ii) PEPPSI-IPr, K2CO3, 1,4-dioxane:MeOH (4:1), 80°C; iii) 1,4-dioxane, t-butyl-2,2,2-trichloroacetimidate, BF3.OEt2; iv) DMF, NBS, -15°C to r.t.
Once the thiophene pentamer 6 is formed it can be further functionalized with amino acids. The amino acids chosen in this thesis is L-tyrosine and L-glutamic acid. Previous studies on LCOs functionalized with these amino acids have been made, although only in either the α- or the β-position on the thiophene backbone. In this thesis, both α- and β-positions are subjects of functionalization, hence the carboxylic moieties being protected with different groups. The groups can thereby undergo deprotection selectively, enabling possible separate coupling with two different sets of amino acids. The target molecules are functionalized with Tyr and Glu, either with the same amino acid on all positions, or alternated on α- and β-positions. The amino acids used in the amide coupling reactions in the synthetic route are protected with tert-butyl groups, why the tert-butylated β-positions of the thiophene pentamer are deprotected and coupled first, allowing deprotection of all tert-butyl groups on the amino acids in the last step of the synthesis.

The synthesis continues (Scheme 2) by deprotecting the tert-butyl esters on thiophene pentamer 6 using trifluoroacetic acid (TFA). The deprotected pentamer is then reacted with H-Tyr-OtBu in an amide coupling reaction, using the specialized activating reagent HATU along with N,N-diisopropylethylamine (DIPEA) as base, yielding the tert-butyl protected tyrosine coupled thiophene pentamer 7. Next, the methyl esters on compound 7 is deprotected using lithium hydroxide (LiOH) and subsequently reacted in another amide coupling reaction using this time H-Glu(OtBu)-OtBu, affording the tert-butyl protected Tyr-Glu coupled thiophene pentamer 8. The final step of the synthesis allows for all tert-butyl groups to be deprotected at once, using TFA, resulting in the target compound Tyr-Glu 9.

**Scheme 2.** Synthesis of target compound 9. General conditions and reagents: i) 1. DCM:TFA 4:1, 2. H-Tyr-OtBu, HATU, DIPEA, DMF; ii) 1. LiOH (1M, aq.), 1,4-dioxane, H2O, 2. H-Glu(OtBu)-OtBu x HCl, HATU, DIPEA, DMF; iii) DCM:TFA 4:1.
The same procedure is followed, using the two diverse amino acids in couplings to different positions of the thiophene pentamer, yielding target molecules 10, 11 and 12 (Figure 6).

**Figure 6.** Structures of the target proteophenes Tyr-Glu 9, Glu-Glu 10, Tyr-Tyr 11 and Glu-Tyr 12.
2 Results and discussion

In this section of the thesis, the results from the synthesis of the target compounds as well as the results from the analyses of the photophysical properties will be reported and discussed. More detailed experimental methods are reported in the experimental section in the report.

2.1 Synthesis of compound 6

The first target molecule was compound 6, since it could be used further to synthesize the target molecules 9, 10, 11 and 12.

The commercially available 3-thiopheneacetic acid 1 was chosen as starting material, protecting the carboxylic acid with a methyl ester via Fischer esterification, using methanol and sulfuric acid under heat (80°C). The yield of the first step in the synthesis was 95%. The formed thiophene methyl ester was then brominated at the α-position via electrophilic aromatic substitution, using NBS in equivalent amount, with dry DMF as solvent, giving compound 2 in a 79% yield.

![Figure 7. Proposed mechanism for the bromination of methyl-2-bromo-3-thiopheneacetate 2.](image)

The bromination (Figure 7) is somewhat regioselective for the α-position since the formed intermediate carbocation is stabilized by the methyl ester substituent on the β-position in the thiophene ring. NBS is used in a 1:1 equivalent ratio to decrease the risk of dibromination. Anyhow, there were some side products formed, such as dibrominated product and bromination in the other α-position. To separate the isomers, versa flash chromatography was used. The reaction was proceeded under reduced temperature (at -15°C), and with a polar solvent, which facilitates electrophilic aromatic substitution. The polar solvent stabilizes the transition state during the reaction. The lower temperature also decreases the risk of side products forming.18

The formed methyl-2-bromo-3-thiopheneacetate 2 was then reacted with 5-carboxythiophene-2-boronic acid in equivalent amount via a palladium catalyzed Suzuki cross coupling reaction. The catalyst used was the commercially available PEPPSI-IPr, in a 3 mol % ratio, and with potassium carbonate as base. Solvents were 1,4-dioxane:methanol 4:1. The reaction was heated to 80°C yielding crude asymmetric thiophene dimer 3.
The mechanism of the Suzuki coupling reaction (Figure 8) starts with oxidative addition of the brominated thiophene monomer 2 to the activated PEPPSI-IPr palladium (0) complex, forming a palladium (II) intermediate. The oxidative addition is the rate determining step. The next step is transmetallation which the base facilitates by activation of the boronic acid. During the transmetallation the thiophene unit \( R^2 \) on the boronic acid is transferred from the boron group to the palladium complex. In the last step of the reaction the coupled product 3 is formed through reductive elimination, which also regenerates the palladium (0) catalyst.\(^{19, 20}\)
Next, compound 3 was reacted without any further purification. *Tert*-Butyl 2,2,2-trichloroacetimidate (TBTA) was used as a source of *tert*-butyl groups, in a 2 mol equivalent ratio, together with the Lewis acid boron trifluoride diethyl etherate (BF₃OEt₂). Dry 1,4-dioxane was used as solvent, resulting in a 63% yield of the *tert*-butylated product 4.

![Figure 9. Proposed mechanism for the *tert*-butyl protection reaction, resulting in compound 4.](image)

The mechanism proposes (Figure 9) that the reaction starts with a carbocation being formed *in situ*, when the oxygen in TBTA interacts with the Lewis acid. The reactive *tert*-butyl carbocation is then subject for nucleophilic attack from the negatively charged oxygen on the free carboxylate at the end of the thiophene dimer.²¹

The formed *tert*-butylated dimer 4 is brominated at the remaining α-position according to the same procedure as previously described, using NBS and DMF, giving the brominated thiophene dimer 5 in a 99% yield. In this bromination step the formation of side products is less since there is just the one α-position left to react at. Compound 5 is then coupled to 2,5-thiophenebisboronic acid in a 2.5:1 mol ratio, via Suzuki cross coupling under the same conditions as earlier mentioned. This yielded the symmetric thiophene pentamer, target compound 6 in 86%.

**Table 1.** Summary of yields for the synthesis of compound 6.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>99</td>
</tr>
<tr>
<td>6</td>
<td>86</td>
</tr>
<tr>
<td>Total yield of Compound 6 (%)</td>
<td>42</td>
</tr>
</tbody>
</table>
2.2 Synthesis of target compounds 9, 10, 11 and 12

As mentioned, compound 6 was used as precursor molecule to synthesize the target proteophenes 9, 10, 11 and 12.

The first step of this synthesis was deprotection of the tert-butyl groups on compound 6, which was performed using trifluoroacetic acid (TFA) in dichloromethane (DCM), ratio 1:4.

![Diagram](image)

**Figure 10.** Proposed mechanism for the deprotection of tert-butyl groups using trifluoroacetic acid.

*Ter*-butyl esters are selectively deprotected under acidic conditions, while the methyl esters are deprotected under basic conditions.\(^{22}\) The deprotection starts by the *ter*-butyl ester being protonated, causing the loss of *ter*-butyl as a carbocation. The formed carbocation can either be trapped by for example methanol or deprotonate to form isobutylene gas.\(^ {23}\)

Without further purification, the deprotected pentamer was reacted with H-Tyr-OtBu in a 3 mol excess, via an amide coupling reaction using the commercially available activation reagent HATU and the steric base DIPEA. The solvent used was DMF and the reaction was carried out in room temperature resulting in 82% of the tyrosine functionalized thiophene pentamer 7.
HATU is used in amide coupling reactions to transform the carboxylic acid into an active ester. The base DIPEA, also known as Hünig’s base, is a sterically hindered base which makes it excellent for this kind of reaction since it is a poor nucleophile and does not compete with the amino acid in the coupling step. The reaction is well carried out in a polar aprotic solvent, why DMF was chosen. The mechanism starts with DIPEA deprotonating the carboxylic acid making it more reactive. The carboxylate ion formed is then attacking HATU to form an unstable salt which is immediately attacked by the formed oxygen ion in the remaining HATU complex, releasing tetramethylurea. The activated ester is now formed and the nucleophilic amine on the amino acid can attack through an addition-elimination reaction, kicking out the better leaving group, the cyclic nitrogen system. A seven membered cyclic transition state is thought to form during the nucleophilic addition, which stabilizes the forming amide through a neighboring group effect from the nitrogen in the pyridine ring.\(^\text{24}\)

The methyl esters in compound 7 were in the next step hydrolyzed, using lithium hydroxide (LiOH) in aqueous solution, with 1,4-dioxane as solvent. LiOH is a good base for this kind of hydrolysis since it shows little racemization at the chiral centre in the amino acid.\(^\text{25}\)
The resulting carboxylates were subsequently reacted with H-Glu(OtBu)-OtBu in an amide coupling reaction, under the same conditions as previously described, affording the tert-butyl protected Tyr-Glu functionalized thiophene pentamer 8 in 58% yield. Finally, all tert-butyl groups were deprotected using TFA, as described before, giving the target compound Tyr-Glu 9.

The same procedure was repeated, affording target molecules Glu-Glu 10, Tyr-Tyr 11 and Glu-Tyr 12. However, when synthesizing Glu-Glu 10, all protecting groups on the pentamer 6 were removed at once, using TFA followed by LiOH, enabling coupling with H-Glu(OtBu)-OtBu at all four positions in one reaction step. This worked out fine, affording slightly crude tert-butyl protected Glu-Glu. When all tert-butyl groups were to be removed in the final step, HPLC-MS analysis showed some side products had formed with an extra mass of 14 g/mol. One possible explanation discussed was that methanol, that had been used to quench the deprotection reaction, could interact with the free acids, possibly through Fischer esterification. If a methyl group had been added to the free acid that would correspond to the peaks with +14 g/mol shown on the HPLC-MS analysis. Therefore, the side products were dissolved in 1,4-dioxane together with LiOH (1M, aq.) to investigate if the possibly added methyl group(s) could be hydrolysed. This did not work out, hence the low yield of 10% of target compound Glu-Glu 10.

When synthesizing compound Glu-Tyr 12, problems occurred during the methyl ester hydrolysis. HPLC analysis showed that also tert-butyl groups had been removed by treatment with LiOH. This in turn gave rise to side products in the coupling reaction with H-Tyr-OtBu. Peaks with masses +163 g/mol was shown on HPLC-MS analysis after coupling, which corresponds to the mass of extra coupled H-Tyr-OtBu, that could have occurred where tert-butyl groups were accidentally missing. This gave the low yield of 13% in the second last step of the synthetic route. The hydrolysis problem could perhaps be avoided by using a different hydrolysis reaction than the one used. One reaction discussed was the Krapcho demethylation which is an S_N2 reaction involving a halide ion as the nucleophile. LiCl in DMF could be used under warm conditions, possibly by microwave irradiation. This method would possibly be more selective for the methyl esters, although the high temperature could be a risk factor for epimerization of the amino acid chiral centres. 26
Figure 12. Possibly formed side product in the Glu-Tyr coupling reaction. The unwanted coupling could have taken place at any of the positions with tert-butyl protected carboxylic acids being removed from the glutamic acids.

Table 2. Summary of yields for the synthesis of target compounds, using compound 6 as precursor molecule.

<table>
<thead>
<tr>
<th>Synthesis step (from Compound 6)</th>
<th>Yield (%) Tyr-Glu 9</th>
<th>Yield (%) Glu-Glu 10</th>
<th>Yield (%) Tyr-Tyr 11</th>
<th>Yield (%) Glu-Tyr 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>82</td>
<td>-</td>
<td>82</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>58</td>
<td>109</td>
<td>51</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>119</td>
<td>10</td>
<td>49</td>
<td>65</td>
</tr>
<tr>
<td>Total yield (%)</td>
<td>57</td>
<td>10</td>
<td>20</td>
<td>8</td>
</tr>
</tbody>
</table>
2.3 Photophysical properties

![Absorption and emission spectra for Tyr-Glu, Glu-Glu, Tyr-Tyr, and Glu-Tyr.](image)

Figure 13. Absorption and emission spectra for the target compounds dissolved in PBS buffer. 30 µM solutions of the ligands were used for the measurements. Absorption- and emission spectra were collected between 340-850 nm with excitation at 420 nm.

Table 3. Summary of absorption- and emission maxima for the different proteophenes, in PBS buffer and when bound to Aβ-aggregates in cryosections from transgenic mice. Stokes shifts are also reported.

<table>
<thead>
<tr>
<th>Proteophene</th>
<th>$\text{Abs}_{\text{max}}$ (nm)</th>
<th>$\text{Em}_{\text{max}}$ (nm)</th>
<th>Stoke shift (nm)</th>
<th>$\text{Em}_{\text{max}}$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PBS</td>
<td>PBS</td>
<td>PBS</td>
<td>Aβ aggregate</td>
</tr>
<tr>
<td>Tyr-Glu</td>
<td>419</td>
<td>546</td>
<td>127</td>
<td>512, 538</td>
</tr>
<tr>
<td>Glu-Glu</td>
<td>415</td>
<td>545</td>
<td>130</td>
<td>512, 538</td>
</tr>
<tr>
<td>Tyr-Tyr</td>
<td>433</td>
<td>548</td>
<td>115</td>
<td>512, 538</td>
</tr>
<tr>
<td>Glu-Tyr</td>
<td>424</td>
<td>547</td>
<td>123</td>
<td>512, 538</td>
</tr>
</tbody>
</table>
Figure 14. Absorption and emission spectra for the target molecules at different pH, dissolved in HCl and NaOH (20 mM solutions).
The results from the absorption- and fluorescence measurements at different pH indicates that the molecules form aggregated structures under acidic conditions. This can be shown by the lower intensity in the fluorescence spectra, together with the slight red shift of the wavelengths. The fluorescence signal is quenched in lower pH. One plausible explanation to this could be that the molecules stack on top of each other through π-π-interactions forming some type of clusters. The only molecule not showing this trend is Glu-Tyr 12. As can be seen in figure 14, in the bottom right corner, Glu-Tyr 12 does not show as distinct red shift as the rest of the molecules.

![Figure 15. CD spectra of the target molecules in HCl or NaOH (20 mM solutions). 30 µM of the ligands were used in the measurements.](image)

The CD measurements showed the highest induced chirality in Tyr-Tyr 11, both under acidic and basic conditions. In basic conditions however, the chirality induced ranged from very low to none in all molecules. This also indicates that the compounds at lower pH forms aggregated supramolecular structures were the induced helicity forms because of interchain π-π interactions. Glu-Tyr 12 does not show any induced chirality under either of these conditions. This, together with the fact that Glu-Tyr 12 did not show any red shift in fluorescence signal under acidic conditions, indicates that it does not form the same kind of aggregated structure as the other molecules seems to do.
Figure 16. Fluorescence images and emission spectra of the ligands bound to Aβ-aggregates in APPPS1 transgenic mice with AD pathology. 3 µM of the ligands in PBS was used in the staining experiment.

In agreement with previous studies, also pentameric LCOs functionalized with amino acids in α- and β-positions, so called proteophenes, displayed selective binding towards Aβ plaques as can be seen in figure 16. Glu-Glu 10 showed specifically interesting results since it stained the aggregates with high selectivity, not showing much interaction with the background media, even in rather high concentration (3 µM). In contrast to Glu-Glu 10, Tyr-Tyr 11 interacted significantly more with the background, possibly due to hydrophobic interactions from the phenol moiety.
3 Conclusions and future perspectives

The aim of the project to synthesise proteophenes for potential use as fluorescent ligands for biological targets was achieved. The proteophenes were also tested for their photophysical properties and showed potential as fluorescent probes. Especially target molecule Glu-Glu 10 showed interesting features, why analogues of this molecule would be particularly interesting to study further. For example, by altering the molecule’s backbone, inserting a 2,1,3-benzothiadiazole (BTD) in the middle of the structure, might change the photophysical properties of the molecule. The BTD building block, together with the thiophene backbone, usually gives fluorescence at higher wavelengths and could therefore be used to further investigate different morphologies among aggregates.

Also, more analyses of the proteophenes photophysical properties would be necessary to investigate other areas of use for the molecules, for example as optically active conjugated polymeric materials with semiconducting and optoelectronic properties.

Other alterations in the structures of the proteophenes that would be interesting to try would be to synthesize the compounds enantiomers. In this study, the L-form of the amino acids were used, why the amino acids in their D-form could be used to confirm and further investigate the induced chirality of the molecules.

Optimization of the synthetic steps were made through-out the project, however there are still more to investigate. As mentioned previously, different methods for hydrolysing methyl esters were discussed but due to time limitations they were not tested. If there had been more time, the Krapcho demethylation reaction would have been tested to see if higher selectivity for the methyl esters could be achieved.

As one can tell, there are still a lot to explore and develop when it comes to conjugated molecules being utilized as fluorescent ligands and it is therefore a very intriguing and exciting area for a synthetic organic chemist to be working in.
Acknowledgements

First and foremost, I would like to thank my examiner professor Peter Nilsson for giving me the opportunity to work with this project. I am looking forward to future projects in your research group.

Secondly, I would like to give thanks to my supervisor Linda Lantz for always encouraging me and guiding me in the lab. Also, a thank you to my other supervisor Marcus Bäck for guidance and help throughout the project. Both of you are excellent chemists and certainly nice company in the lab.

Last, but not least, I want to thank my friend and opponent Jonathan Thornell for carefully dissecting my report and for being a fair opponent.
References

4. Gomes, C.M. Protein Misfolding Diseases. (Humana Press, 2019)


Experimental details

General methods
All chemicals and solvents used were obtained from commercial sources. TLC was carried out on 0.25 mm precoated silica-gel plates (Merck 60 F254) using UV-light (λ=254 nm, 366 nm), staining with ethanol/sulfuric acid/p-anisaldehyde/acetic acid 90:3:2:1 for visualization. Organic extracts were dried over anhydrous magnesium sulfate, filtered and concentrated in vacuo. Flash column chromatography (FC) was performed using silica gel (Merck Grade 9385, high purity grade, pore size 60 Å, 230-400 mesh particle size) with different organic solvents as mobile phase. Analytical HPLC was performed on a Waters system equipped with a Waters 1525 gradient pump, 2998 Photodiode Array Detector, 2424 Evaporative Light Scattering Detector, SQD 2 Mass Detector and an Xbridge® C18 column (4.6 x 50 mm, 2.5 μm), using water:acetonitrile 95:5 with 10 mM NH₄OAc, and acetonitrile:water:10 with 10 mM NH₄OAc as mobile phase. Preparative HPLC was carried out on the same Waters system, with an XSELECT™ CSH™ Prep Phenyl-Hexyl column (19 x 250 mm, 5 μm). Versa flash chromatography was performed using a VersaPak™ C18 Cartridge column (40 x 150 mm), with acetonitrile:water as mobile phase. NMR-spectra were recorded on a Varian 300 MHz instrument and chemical shifts were defined with the solvent residual peak as reference.

Optical characterization of the proteophenes
Target molecules 9, 11 and 12 were dissolved in de-ionized water together with LiOH (1M) to a final concentration of 1.5 mM, target molecule 10 was dissolved in de-ionized water to a final concentration of 1.5 mM. All stock solutions were diluted further to a final concentration of 30 μM with PBS buffer. Absorption-, excitation- and emission spectra were collected between 340-850 nm using a Tecan Infinite M1000 Pro microplate reader (Tecan, Männedorf, Switzerland) with excitation at 420 nm. CD spectra were recorded with a Chirascan (Applied Photophysics, Leatherhead, U.K.) using a 1 cm quartz cell. 

Proteophene staining of cryosections from transgenic mice APPPS1 with AD pathology
Cryosections were fixed in absolute ethanol for 10 min and rehydrated in water followed by PBS buffer. Stock solutions of the target compounds (1.5 mM) were further diluted to a final concentration of 3 μM in PBS and applied to the brain sections, which were then incubated in room temperature. The sections were mounted with a fluorescent mounting medium. Spectral images of the stained tissue sections were collected on an inverted Zeiss (Axio Observer. Z1) LSM 780 microscope equipped with a 32 channel QUASAR GaAsP spectral array detector. For all imaging a Plan-Apochromat 20x/1.3 DIC objective lens was used. Excitation was done by excitation with an argon laser at 458 nm. Emission spectra were collected between 416-687 nm.

Methyl 2-(thiophen-3-yl)acetate
3-thiophene acetic acid 1 (5.00 g, 35.2 mmol) was dissolved in MeOH (100 mL). A catalytic amount of H₂SO₄ (18 M) was added drop wise, and the solution was heated at 80°C overnight. The solution was neutralized using NaHCO₃, extracted with DCM, washed with H₂O and saturated NaCl (aq.), dried over MgSO₄, filtered and concentrated. The product was purified using FC (H/E 10:1), giving a 95% yield (5.19 g) of the product. 1H NMR (300 MHz, CDCl₃) δ 7.30 (dd, J=2.9 Hz, 4.7 Hz, 1H), 7.16 (dd, J=1.2 Hz, 2.9 Hz, 1H), 7.05 (dd, J=1.2 Hz, 4.7 Hz, 1H), 3.70 (s, 3H), 3.66 (s, 2H). 13C NMR (75 MHz, CDCl₃) δ 171.7, 133.7, 128.6, 125.9, 123.0, 52.2, 35.8.

Methyl 2-(2-bromothiophen-3-yl)acetate (2)
The thiophene methyl ester (5.29 g, 33.9 mmol) was dissolved in dry DMF (15 mL) before NBS (6.03 g, 33.9 mmol), also dissolved in dry DMF (10 mL), was added drop wise at -15°C. The reaction was allowed to reach room temperature and stirred overnight. The solution was poured over H₂O, extracted with DCM, washed with water and saturated NaCl (aq.), dried over MgSO₄, filtered and...
concentrated. The product was purified using FC with a gradient mobile phase system of H/E (40:1, 10:1). A second purification with FC, this time using a mobile phase of T/E (80:1), gave 79% (6.31 g) yield of methyl-2-bromo-3-thiophenecarboxylate 2. To separate the isomers, versal flash chromatography was performed (AcN/H₂O 40:60). ¹H NMR (300 MHz, CDCl₃) δ 7.24 (d, J=5.9 Hz, 1H), 6.93 (d, J=5.9 Hz, 1H), 3.72 (s, 3H), 3.64 (s, 2H). ¹³C NMR (75 MHz, CDCl₃) δ 170.6, 133.4, 128.6, 125.8, 111.6, 52.2, 34.8.

3'- (2-methoxy-2-oxoethyl)-[2,2'-bithiophene]-5-carboxylic acid (3)

Methyl-2-bromo-3-thiophenecarboxylate 2 (1.19 g, 5.07 mmol) was dissolved in 1,4-dioxane:MeOH (4:1, 50 mL), together with K₂CO₃ (2.10 g, 15.2 mmol) and 5-carboxythiophene-2-boronic acid (872 mg, 5.07 mmol), before PEPPSI-IPr (103 mg, 3 mol %) was added. The mixture was heated to 80°C and stirred for 2 hours. pH was adjusted using HCl, before the solution was extracted with EtOAc, washed with H₂O and saturated NaCl (aq.), dried over MgSO₄ and filtered. Evaporation of the solvents gave crude thiophene dimer carboxylic acid 3 (1.56 g), which was used without further purification in the next step.

_Tert-butyl 3'- (2-methoxy-2-oxoethyl)-[2,2'-bithiophene]-5-carboxylate (4)_

The crude thiophene dimer 3 (1.43 g, 5.07 mmol) was dissolved in dry 1,4-dioxane (12 mL) before a solution of TBTA (2.20 g, 10.1 mmol) in dry 1,4-dioxane (12 mL) was added. BF₃OEt₂ (192 µL, 1.51 mmol) was added and the mixture was stirred for 2 hours. Additional amounts of TBTA (1.1 g) and BF₃OEt₂ (96 µL) were added and the mixture was stirred overnight. An additional amount of TBTA (1.1 g) was added and the solution was stirred for 3 hours, extracted with EtOAc, washed with saturated NaHCO₃ and H₂O, dried over MgSO₄, filtered and concentrated. The product was purified by FC (H:E 10:1), to give 63% yield (1.08 g) of the tert-butylated product 4. ¹H NMR (300 MHz, CDCl₃) δ 7.66 (d, J=3.5 Hz, 1H), 7.28 (d, J=5.3 Hz, 1H), 7.13 (d, J=3.5 Hz, 1H), 7.06 (d, J=5.3 Hz, 1H), 3.78 (s, 2H), 3.72 (s, 3H), 1.58 (s, 9H). ¹³C NMR (75 MHz, CDCl₃) δ 171.2, 161.4, 141.4, 135.7, 133.3, 132.7, 131.3, 130.6, 127.1, 125.6, 82.1, 52.4, 34.7, 28.3.

_Tert-butyl 5'-bromo-3'-(2-methoxy-2-oxoethyl)-[2,2'-bithiophene]-5-carboxylate (5)_

Dimer 4 (64.7 mg, 0.19 mmol) was dissolved in dry DMF (0.5 mL), before a solution of NBS (34 mg, 0.19 mmol) in dry DMF (0.125 mL), was added at -15°C. The mixture was stirred overnight. An additional amount of NBS (10 mg) was added and the solution was stirred for 3 hours, poured over H₂O, extracted with DCM, washed with H₂O and saturated NaCl (aq.), dried over MgSO₄ and filtered. Concentration of the solution gave crude monobrominated compound 5 (79 mg). ¹H NMR (300 MHz, CDCl₃) δ 7.64 (d, J=4.1 Hz, 1H), 7.08 (d, J=4.1 Hz, 1H), 7.03 (s, 1H), 3.72 (s, 3H), 3.70 (s, 2H), 1.57 (s, 9H). ¹³C NMR (75 MHz, CDCl₃) δ 170.7, 161.2, 139.9, 136.4, 133.3, 133.2, 131.9, 127.5, 112.6, 82.3, 52.5, 34.5, 28.3.

Compound 6

Dimer 5 (591 mg, 1.42 mmol) was dissolved in 1,4-dioxane:MeOH (4:1, 11 mL), together with K₂CO₃ (430 mg, 3.11 mmol) and 2,5-thiophenecarboxylic acid (97.2 mg, 0.57 mmol), before PEPPSI-IPr (11.5 mg, 3 mol %) was added at 80°C and the reaction was stirred for 2 hours. An additional amount of 2,5-thiophenecarboxylic acid (49 mg) was added and the solution was left to stir for 2 hours, neutralized with HOAc, extracted with EtOAc, washed with H₂O and saturated NaCl (aq.), dried over MgSO₄, filtered and concentrated. The product was purified by FC (T/E, 40:1), giving 86% (368 mg).
of compound 6. $^1$H NMR (300 MHz, CDCl$_3$) $\delta$ 7.67 (d, $J=4.1$ Hz, 2H), 7.16 (d, $J=4.1$ Hz, 2H), 7.14 (s, 2H), 7.10 (s, 2H), 3.78 (s, 4H), 3.76 (s, 6H), 1.59 (s, 18H). $^{13}$C NMR (75 MHz, CDCl$_3$) $\delta$ 170.7, 161.3, 140.9, 133.4, 132.1, 128.4, 127.2, 127.0, 125.1, 82.2, 52.5, 34.9, 28.4.

**Compound 7**

Compound 6 (50.0 mg, 0.07 mmol) was dissolved in DCM (2.5 mL), before TFA (0.6 mL) was added and the solution was stirred for 3.5 hours, before the solvents were co-evaporated with toluene. The deprotected pentamer (42.6 mg, 0.07 mmol) was dissolved in DMF (2.5 mL) together with H-Tyr-OtBu (47.1 mg, 0.20 mmol) before DIPEA (51.3 mg, 0.40 mmol) and HATU (75.4 mg, 0.20 mmol) were added and the mixture was stirred over night at room temperature. The mixture was diluted with EtOAc, washed with H$_2$O, dried over MgSO$_4$, filtered and concentrated. Preparative HPLC was used to purify the product, yielding 82% (58.4 mg) of compound 7. $^1$H NMR (300 MHz, CD$_3$CO) $\delta$ 7.75 (d, $J=8.2$ Hz, 2H), 7.71 (d, $J=4.1$ Hz, 2H), 7.28 (s, 2H), 7.23 (s, 2H), 7.22 (d, $J=4.1$ Hz, 2H), 7.14 (d, $J=8.8$ Hz, 4H), 6.78 (d, $J=8.8$ Hz, 4H), 4.73-4.71 (m, 2H), 3.83 (s, 4H), 3.69 (s, 6H), 3.15-3.03 (m, 4H), 1.44 (s, 18H).

**Compound 8**

Compound 7 (15.5 mg, 0.0143 mmol) was dissolved in 1,4-dioxane (1 mL) before LiOH (aq. 1M) (42.9 µL, 0.429 mmol) was added and the reaction was stirred overnight. An additional amount of LiOH (aq. 1M) (11 µL) was added before the solution was neutralised with HCl (1M) and the solvents evaporated. Hydrolysed compound 7 (15.1 mg, 0.0143 mmol) was dissolved in DMF (0.6 mL) together with H-Glu(OtBu)-Otbu x HCl (12.7 mg, 0.0429 mmol), before HATU (16.3 mg, 0.0429 mmol) and DIPEA (15.0 µL, 0.0859 mmol) were added and the mixture was stirred for 25 minutes. The solution was diluted using EtOAc, washed with H$_2$O, dried over MgSO$_4$, filtered and concentrated. FC (DCM/MeOH (5%)) was used to purify the product. A second purification with preparative HPLC was performed, giving 58% (12.8 mg) of compound 8. $^1$H NMR (300 MHz, CD$_3$CO) $\delta$ 7.74 (d, $J=8.2$ Hz, 2H), 7.72 (d, $J=4.1$ Hz, 2H), 7.61 (d, $J=8.2$ Hz, 2H), 7.40 (s, 2H), 7.38 (d, $J=4.1$ Hz, 2H), 7.25 (s, 2H), 7.14 (d, $J=8.2$ Hz, 4H), 6.77 (d, $J=8.2$ Hz, 4H), 4.73-4.71 (m, 2H), 4.44-4.42 (m, 2H), 3.74 (s, 4H), 3.11-3.02 (m, 4H), 2.37-2.32 (m, 4H), 2.09-2.03 (m, 2H), 1.95-1.85 (m, 2H), 1.42 (s, 54H). $^{13}$C NMR (75 MHz, CD$_3$CO) $\delta$ 172.4, 171.8, 171.7, 170.0, 161.7, 157.1, 140.3, 136.6, 136.3, 135.3, 132.0, 131.2, 129.7, 128.9, 128.7, 128.0, 126.1, 116.0, 81.9, 80.6, 55.9, 53.3, 37.6, 37.1, 32.2, 29.4, 27.9.
Compound 9

Compound 8 (0.00832 mmol, 12.8 mg) was dissolved in DCM (2 mL) before TFA (0.5 mL) was added and the mixture was stirred for 4 hours, quenched with MeOH, co-evaporated with toluene, giving 11.9 mg of the product. $^1$H NMR (300 MHz, (CD$_3$)OD) $\delta$ 7.62 (d, $J$=4.1 Hz, 2H), 7.25 (s, 2H), 7.23 (d, $J$=3.5 Hz, 2H), 7.16 (s, 2H), 7.10 (d, $J$=8.8 Hz, 4H), 6.71 (d, $J$=8.8 Hz, 4H), 4.52-4.47 (m, 2H), 3.73 (s, 4H), 3.26-3.21 (m, 2H), 2.96-2.98 (m, 2H), 2.46-2.44 (m, 4H), 2.25-2.20 (m, 2H), 2.05-1.97 (m, 2H). $^{13}$C NMR (75 MHz, (CD$_3$)OD) $\delta$ 176.3, 175.1, 174.9, 172.7, 163.8, 157.3, 141.3, 139.3, 137.4, 137.1, 134.9, 132.4, 131.3, 130.8, 129.3, 128.5, 128.1, 126.2, 116.3, 55.9, 53.3, 37.5, 37.1, 31.3, 27.8.

Tert-butyl protected compound 10

Compound 6 (41 mg, 0.0542 mmol) was dissolved in DCM (2 mL) before TFA (0.5 mL) was added drop wise and the solution was stirred for 3 hours, solvents evaporated and co-evaporated with toluene and DCM. The deprotected compound 6 (17.5 mg, 0.0271 mmol) was dissolved in 1,4-dioxane (2 mL) before LiOH (aq. 1M) (0.163 mmol, 163 µL) was added at 50°C. An additional amount LiOH (aq. 1M) (81 µL) was added and the mixture was stirred additionally. The solution was neutralised using HCl. To the carboxylic acid compound 6 (0.0271 mmol, 16.7 mg) in DMF (2 mL), H-Glu(OzBu)-OzBu x HCl (0.163 mmol, 48.2 mg) was added, before HATU (0.163 mmol, 62.0 mg) and DIPEA (0.326 mmol, 56.7 µL) were added and the solution was stirred at room temperature. An additional amount of DIPEA (14 µL) was added and the solution was stirred for 2 hours, diluted with EtOAc, washed with H$_2$O and saturated NaCl (aq.), dried over MgSO$_4$, filtered and concentrated. The product was purified with FC (DCM/MeOH(5%)) to give 109% (46.9 mg) of the tert-butyl protected compound 10. $^1$H NMR (300 MHz, (CD$_3$)$_2$CO) $\delta$ 7.86 (d, $J$=8.2 Hz, 2H), 7.77 (d, $J$=3.5 Hz, 2H), 7.61 (d, $J$=8.2 Hz, 2H), 7.42-7.40 (m, 4H), 7.25 (s, 2H), 4.57-4.55 (m, 2H), 4.44-4.43 (m, 2H), 3.75 (s, 4H), 2.45-1.86 (m, 16H), 1.45 (s, 72H). $^{13}$C NMR (75 MHz, (CD$_3$)$_2$CO) $\delta$ 172.6, 172.4, 171.8, 169.9, 161.9, 140.4, 140.3, 136.7, 136.3, 135.3, 132.0, 129.8, 128.7, 128.0, 126.1, 82.0, 81.9, 80.7, 80.6, 53.6, 53.3, 53.0, 38.8, 37.1, 32.4, 32.2, 30.6, 28.2, 27.7.

Compound 10

Tert-butyl protected compound 10 (46.9 mg, 0.0296 mmol) was dissolved in DCM (2 mL) before TFA (0.5 mL) was added dropwise and the mixture was stirred for 3 hours. An additional amount TFA (0.25 mL) was added. When the reaction was finished, it was quenched using MeOH and co-evaporated with toluene and DCM. The reaction was purified using preparative HPLC to give 10% (3.2 mg) of compound 10. $^1$H NMR (300 MHz, (CD$_3$)$_2$CO) $\delta$ 7.76 (d, $J$=4.1 Hz, 2H), 7.32-7.31 (m, $J$=4.1 Hz, 4H), 7.24 (s, 2H), 4.44-4.40 (m, 2H), 4.29-4.26 (m, 2H), 3.79 (s, 4H), 2.40-2.36 (m, 4H), 2.31-2.25 (m, 6H), 2.14-2.11 (m, 4H), 1.95-1.91 (m, 2H).
**Tert-butyl protected compound 11**

Compound 7 (18 mg, 0.0166 mmol) was dissolved in 1,4-dioxane (1 mL) before LiOH (aq. 1M)(0.0498 mmol, 49.8 μL) was added and the mixture was stirred for 3 hours. An additional amount of LiOH (aq. 1M) (40 μL) was added and the solution was stirred additionally, neutralised using HCl (1M) and concentrated. The hydrolysed compound 7 (0.0166 mmol, 17.5 mg) was dissolved in DMF (1 mL) together with H-Tyr-OtBu (0.0498 mmol, 11.8 mg) before DIPEA (0.0996 mmol, 17.3 μL) and HATU (0.0498 mmol, 18.9 mg) were added and the reaction was stirred at room temperature for 35 minutes. The reaction solution was diluted with EtOAc, washed with H2O, dried over MgSO4, filtered and concentrated. The product was purified using preparative HPLC, giving 51% (12.7 mg) of tert-butylated compound 11. 1H NMR (300 MHz, (CD3)OD) δ 7.60 (d, J=4.1 Hz, 2H), 7.12-7.08 (m, 10H), 7.01 (d, J=8.2 Hz, 4H), 6.73-6.68 (m, 8H), 4.68-4.65 (m, 2H), 4.55-4.52 (m, 2H), 3.63 (s, 4H), 3.11-2.90 (m, 8H), 1.41 (s, 36H). 13C NMR (75 MHz, (CD3)OD) δ 172.8, 172.4, 172.2, 163.3, 157.4, 141.3, 139.2, 137.5, 137.0, 134.7, 132.4, 131.4, 131.3, 130.8, 129.0, 128.7, 128.4, 128.0, 126.3, 116.4, 116.3, 83.2, 83.1, 56.8, 56.3, 37.7, 37.1, 28.3.

**Compound 11**

Tert-butyl protected compound 11 (12.7 mg, 0.0085 mmol) was dissolved in DCM (2 mL) before TFA (0.5 mL) was added dropwise and the solution was stirred for 3 hours, co-evaporated with toluene and DCM. The mixture was purified using preparative HPLC to give 49% (5.3 mg) of compound 11. 1H NMR (300 MHz, (CD3)OD) δ 7.57 (d, J=4.1 Hz, 2H), 7.12-7.10 (m, 4H), 7.09-7.06 (m, 4H), 7.05 (d, J=4.1 Hz, 2H), 7.02-6.99 (m, 4H), 6.71-6.66 (m, 8H), 4.77-4.71 (m, 2H), 4.63-4.62 (m, 2H), 3.61 (s, 4H), 3.25-2.86 (m, 8H).

**Glu(OtBu)-OtBu coupled pentamer**

Compound 6 (51.6 mg, 0.0682 mmol) was dissolved in DCM (2 mL) before TFA (0.5 mL) was added dropwise and the reaction was stirred for 3 hours, quenched with MeOH and co-evaporated with toluene and DCM. The deprotected compound 6 (44.0 mg, 0.00682 mmol) was dissolved in DMF (2 mL) together with H-Glu(OtBu)-OtBu x HCl (60.5 mg, 0.205 mmol) before HATU (77.8 mg, 0.205 mmol) and DIPEA (71.3 μL, 0.409 mmol) were added and the reaction was stirred for 1 hour, diluted with EtOAc, washed with H2O and saturated NaCl (aq.), dried over MgSO4, filtered and concentrated. The product was purified using FC (T/E, 2:1) to give 95% (72.9 mg) of Glu(OtBu)-OtBu coupled pentamer. 1H NMR (300 MHz, (CD3)CO) δ 7.87 (d, J=7.6 Hz, 2H), 7.77 (d, J=4.1 Hz, 2H), 7.28 (s, 2H), 7.25 (d, J=4.1 Hz, 2H), 7.24 (s, 2H), 4.54-4.52 (m, 2H), 3.81 (s, 4H), 3.67 (s, 6H), 2.42-2.37 (m, 4H), 2.15-2.11 (m, 2H), 2.01-1.93 (m, 2H), 1.43 (s, 36H). 13C NMR (75 MHz, (CD3)CO) δ 172.6, 171.9, 171.2, 161.9, 140.6, 140.1, 136.4, 133.7, 132.3, 129.7, 128.7, 127.9, 126.3, 82.0, 80.7, 53.8, 52.4, 35.1, 32.4, 30.6, 28.3, 27.7.
**Tert-butyl protected compound 12**

Glu(OtBu)-OtBu coupled pentamer (0.0223 mmol, 25.1 mg) was dissolved in 1,4-dioxane (2 mL) before LiOH (aq., 1M) (0.0668 mmol, 66.8 µL) was added and the reaction was stirred for 2 hours. An additional amount LiOH (aq. 1M) (66.8 µL) was added and the mixture was stirred overnight, neutralised using HCl (1M). The hydrolysed Glu(OtBu)-OtBu coupled pentamer (0.0223 mmol, 24.5 mg) was dissolved in DMF together with H-Tyr-OtBu (0.0668 mmol, 15.9 mg) before HATU (0.0668 mmol, 25.4 mg) and DIPEA (0.0668 mmol, 11 µL) were added and the reaction was stirred for 3 hours. Additional amounts of H-Tyr-OtBu (11 mg), HATU (17 mg) and DIPEA (7 µL) were added and the mixture was stirred additionally, diluted with EtOAc, washed with H2O and Saturated NaCl (aq.), dried over MgSO4, filtered and concentrated. The product was purified by FC (DCM:MeOH (5%)) to give 13% (4.3 mg) of tert-butyl protected compound 12. 1H NMR (300 MHz, (CD3)2CO) δ 7.85 (d, J=8.2 Hz, 2H), 7.78 (d, J=3.5 Hz, 2H), 7.40-7.35 (m, 4H), 7.34 (s, 2H), 7.16 (d, J=3.5 Hz, 2H), 6.90-6.87 (m, 4H), 6.55-6.52 (m, 4H), 4.59-4.56 (m, 4H), 3.82 (s, 2H), 3.77 (s, 2H), 3.00-2.96 (m, 2H), 2.85-2.78 (m, 2H), 2.38-2.33 (m, 4H), 2.21-2.05 (m, 4H).

**Compound 12**

Tert-butyl protected compound 12 (9.8 mg, 0.00638 mmol) was dissolved in DCM (2 mL) before TFA (0.5 mL) was added dropwise and the solution was stirred for 2 hours, co-evaporated with toluene and DCM. The product was purified by preparative HPLC to give 65% (5 mg) of compound 12. The product was dissolved in LiOH (aq, 1M) to make it a salt. 1H NMR (300 MHz, (D2)O) δ 7.75 (d, J=3.5 Hz, 2H), 7.37 (s, 2H), 7.27 (s, 2H), 7.16 (d, J=3.5 Hz, 2H), 6.90-6.87 (m, 4H), 6.55-6.52 (m, 4H), 4.41-4.32 (m, 4H), 3.82 (s, 2H), 3.77 (s, 2H), 3.00-2.96 (m, 2H), 2.85-2.78 (m, 2H), 2.38-2.33 (m, 4H), 2.21-2.05 (m, 4H).
Appendix

**Figure 1.** $^1$H NMR for methyl 2-((thiophen-3-yl)acetate.

**Figure 2.** $^{13}$C NMR for methyl 2-((thiophen-3-yl)acetate.
Figure 3. $^1$H NMR for methyl 2-(2-bromothiophen-3-yl)acetate (2).

Figure 4. $^{13}$C NMR for methyl 2-(2-bromothiophen-3-yl)acetate (2)
Figure 5. $^1$H NMR for tert-butyl 3’-(2-methoxy-2-oxoethyl)-[2,2’-bithiophene]-5-carboxylate (4)

Figure 6. $^{13}$C NMR for tert-butyl 3’-(2-methoxy-2-oxoethyl)-[2,2’-bithiophene]-5-carboxylate (4)
Figure 7. $^1$H NMR for tert-butyl 5'-bromo-3'-((2-methoxy-2-oxoethyl)-[2,2'-bithiophene]-5-carboxylate (5)

Figure 8. $^{13}$C NMR for tert-butyl 5'-bromo-3'-((2-methoxy-2-oxoethyl)-[2,2'-bithiophene]-5-carboxylate (5)
Figure 9. $^1$H NMR for Compound 6

Figure 10. $^{13}$C NMR for Compound 6
Figure 11. $^1$H NMR for Compound 7

Figure 12. $^{13}$C NMR for Compound 7
Figure 13. $^1$H NMR for Compound 8

Figure 14. $^{13}$C NMR for Compound 8
Figure 15. $^1$H NMR for Compound 9

Figure 16. $^{13}$C NMR for Compound 9
Figure 17. $^1$H NMR for tert-butyl protected compound 10

Figure 18. $^{13}$C NMR for tert-butyl protected compound 10
Figure 19. $^1$H NMR for compound 10
Figure 20. $^1$H NMR for tert-butyl protected compound 11

Figure 21. $^{13}$C NMR for tert-butyl protected compound 11
Figure 22. $^1$H NMR for compound 11
Figure 23. $^1$H NMR for Glu-diOtBu coupled pentamer

Figure 24. $^{13}$C NMR for Glu-diOtBu coupled pentamer
Figure 25. $^1$H NMR for $t$-butyl protected compound 12

Figure 26. $^1$H NMR for compound 12