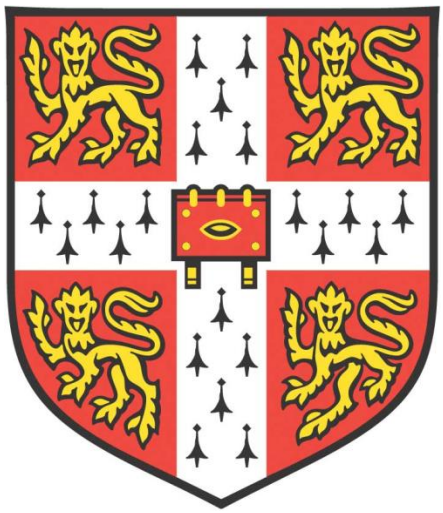
Self-assembly and aggregation of glucagon-like peptide 1 and its analogues



Eva Přáda

Peterhouse

Yusuf Hamied Department of Chemistry

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

March 2023

DECLARATION

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It is not substantially the same as any that I have submitted for a degree or diploma or other qualification at the University of Cambridge or any other University, and no part has already been or is concurrently being submitted for any degree, diploma or other qualification. In accordance with the Physics and Chemistry Degree Committee guidelines, this thesis does not exceed the 60,000-word limit

Eva Přáda

Cambridge, March 2023

SUMMARY

Self-assembly and aggregation of glucagon-like peptide 1 and its analogues

Eva Přáda

Aggregation and physical instability of peptide-based drugs poses a great challenge to the pharmaceutical industry. Glucagon-like peptide 1 (GLP-1) is a hormone that is used in the treatment of type-2 diabetes. However, GLP-1 has a short half-life *in vivo* and it is prone to aggregate which complicates its pharmaceutical usage. Strategies to overcome the short half-life *in vivo* include numerous chemical modifications of the native peptide. The focus of this Thesis is on the effect of two sets of chemical modification strategies, lipidation and C-terminal amidation, on the physical stability of the peptide.

The first part of this work combines experimental and computational approaches to better understand the molecular basis of the aggregation of GLP-1 and its C-terminally amidated variant, GLP-1-Am. In particular, the off-pathway aggregation of GLP-1 and GLP-1-Am into disordered low-molecular weight oligomers is described. This process competes with the amyloid formation pathway and the addition of pre-formed off-pathway oligomers slightly slows down the fibrillation rate. Energy Landscape Theory was employed to investigate and rationalize the conformational behaviour and aggregation propensity of GLP-1 in different protonation states. Under all conditions studied, the GLP-1 energy landscape possesses a multi-funnel character with a variety of structurally different ensembles with low energy, which is a typical feature of intrinsically disordered proteins and aggregating systems. It is also shown that β-structure-containing conformations are more energetically favoured at acidic pH compared to neutral pH conditions, which agrees with a greater propensity of GLP-1 for aggregation at acidic pH which was observed experimentally.

The second part of this Thesis focuses on the self-assembly and aggregation of lipidated analogues of GLP-1. Four lipidated GLP-1 analogues, which varied in the position of lipidation, and one additional analogue differing by the nature of the lipid moiety, were

studied to establish the effect of the lipidation site and the lipid moiety on the physical stability of the peptide. The lipidation was shown to induce formation of large stable oligomers (i.e. > 7 monomeric units). The aggregation mechanism and kinetics were shown to be highly dependent on the lipidation position and the nature of the lipid moiety. Moreover, the aggregation kinetics of lipidated analogues were rarely observed to follow a classical nucleation-elongation mechanism but were rather likely to consist of more complex processes. Aggregates with a high content of β-sheet were formed by all analogues studied, however, they were distinct in their tertiary structure and aggregate morphology.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Prof Sophie Jackson for giving me an opportunity to work in her group and for her continuous support, advice and guidance during my PhD.

A big thank you belongs to all members of the Jackson group for their friendly support. In particular, I would like to thank Irina Edu, Katie Gibson, Xinyang Li, Jack Barber and Frederik Becher for their helpful input into various parts of this project. I would also like to thank Dr. Konstantin Röder and Dr. Alasdair Keith for helping me to set up (and keep running) the computational part of this project and for answering all my naive questions.

My gratitude also belongs to Prof. Petr Bouř, Dr. Tao Wu, and Dr. Monika Krupová from the Institute of Organic Chemistry and Biochemistry of the CAS in Prague for allowing me to perform the IR/VCD measurements in their lab and for their help and guidance during the experiments.

I am extremely grateful to Peterhouse for funding my PhD and, therefore, for making this research possible.

Last but definitely not least, I would like to thank my husband, my parents and my grandma for their support throughout my life and my studies.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

λ_{max}	fluorescence emission maximum
a.u.	arbitrary units
A_{280}	absorbance at the wavelength of 280 nm
ACN	acetonitrile
AFM	atomic force microscopy
Aib	α-aminoisobutyric acid
ANS	8-anilinonaphthalene-1-sulfonic acid
APR	aggregation-prone region
AUC	analytical ultracentrifugation
cac	critical aggregation concentration
CD	circular dichroism (spectroscopy)
ddH_2O	distilled deionized water
DOPS	1,2-dioleoyl-sn-glycero-3-phospho-L-serine
DPP-4	dipeptidyl-peptidase IV (enzyme)
DPS	discrete path sampling
DSSP	Define Secondary Structure of Proteins (algorithm)
EDTA	ethylenediaminetetraacetic acid
FEL	free energy landscape
FPLC	fast protein liquid chromatography
GLP-1	glucagon-like peptide 1
GLP-1-Am	glucagon-like peptide 1 amide
HPLC	high-performance liquid chromatography
IEF	isoelectric focusing
IM	ion mobility

IR	infra-red (spectroscopy)
IS	ionic strength
K_{av}	gel phase distribution parameter
LC	liquid chromatography
L-BFGS	limited-memory Broyden-Fletcher-Golgfarb-Shano algorithm
MRE_{222}	mean residue ellipticity at 222 nm (units of circular dichroism)
MS	mass spectrometry
*M*W	molecular weight
MWCO	molecular weight cut-off
PEL	potential energy landscape
PES	polyethersulfone
pI	isoelectric point
PVDF	polyvinylidene fluoride (membrane)
RMS	root mean square (force)
rpm	revolutions per minute
R_{St}	Stokes radius
SEC	size-exclusion chromatography
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
ThT	thioflavin T
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
UV-Vis	ultraviolet-visible (spectroscopy)
VCD	vibrational circular dichroism (spectroscopy)

CHAPTER 1
INTRODUCTION

1.1 Therapeutic peptides

In 1922, the field of peptide therapeutics started with the first medical use of insulin, which was extracted from bovine and porcine pancreases. This discovery revolutionized the treatment of type I diabetes. Another milestone in the field was the development of solid-phase peptide synthesis in 1963[1]. However, the peptide therapeutics field only started to significantly grow 20 years later in the 1980s with the approval of synthetic gonadotropin-releasing hormones leuprolide and goserelin. In 2021, there were around 80 peptide-based drugs on the global market, more than 150 peptides in clinical development and another 400–600 peptides in preclinical studies[2].

To overcome rapid enzymatic cleavage and, therefore, short half-lives of peptide-based therapeutics *in vivo*, multiple strategies have been employed. These strategies include: insertion of unnatural amino acids or D-amino acids, N-terminal capping, cyclization, truncation, peptide stapling[2] or targeting peptides directly to cells or tissues[3]. In the latter case, nano-carrier based delivery systems are being developed which have the potential to replace the traditional subcutaneous injection in the near future[4].

Due to their size, native peptides are rapidly cleared from plasma, therefore, protection against enzymatic cleavage alone is often insufficient to increase the half-life of a drug *in vivo*. Since the mechanism of renal clearance is size-dependent[5,6], strategies to increase the molecular weight of peptides have been developed. Such strategies include lipidation, conjugation to larger proteins and PEGylation[6,7]. Steric hindrance from these bulky moieties additionally protects against proteolytic degradation[6,7]. Selective conjugation to plasma proteins has been employed in the development of two glucagon-like peptide 1 (GLP-1) analogues – albinoglutide and dulaglutide. For those analogues, the half-life of the peptide was extended to allow once-weekly injections[8,9]. Albinoglutide consists of two tandem copies of modified and truncated variants of the GLP-1 peptide which are fused to human

albumin[10]. In dulaglutide (Figure 1.1) each of the two modified GLP-1 analogues is covalently linked by a flexible peptide linker to human immunoglobulin G4 heavy chain fragment. The two fragments (with covalent linked GLP-1) are joined through two disulfide bridges. A PEGylation strategy, in which polyethylene glycol (PEG) moiety is attached to peptides/proteins, has been successfully employed for multiple proteins (e.g., PEG-bovine adenosine deaminase and PEG-α-interferon). Several other PEGylated peptides and proteins are currently in clinical trials[11].

Lipidation is another common strategy to improve the stability of peptide drugs *in vivo*[12–14]. Lipidation is also known to extend the systemic circulation half-life of the peptide molecule by allowing reversible binding to serum albumin[15]. This strategy has been applied to GLP-1 to produce liraglutide and semaglutide which are commercially available lipidated forms of GLP-1 analogues[16], as well as other lipidated GLP-1 analogues that are currently under clinical evaluation[17,18]. Lipidation of peptides may often negatively affect peptide solubility, however, these limitations can be overcome by a co-formulation with excipients that increase solubility, such as hydroxypropyl-β-cyclodextrin[19].

In the case of liraglutide and the four lipidated GLP-1 analogues studied in this Thesis, a palmitic acid chain is linked to a lysine in the peptide *via* a γ-glutamic acid spacer[17]. Lipidation was shown to enhance the potency (both *in vivo* and *in vitro*) of the peptide compared to a non-lipidated version, and the γ-glutamic acid linker was found to be superior to other spacers[20].

1.2 Glucagon-like peptide 1

Glucagon-like peptide 1 (GLP-1) is a 31-amino-acid incretin hormone regulating the glucose level in the blood[21] and its derivatives are widely used in the treatment of type 2 diabetes[3,16,22–26]. It is secreted by endocrine L cells in the gut in response to food intake. The main effect of GLP-1 is regulation of insulin exocytosis from pancreatic β-cells. GLP-1 also suppresses appetite and reduces glucagon secretion[27]. Due to these effects, it is used in the treatment of type-2 diabetes mellitus and it also seems to be a promising candidate for treatment of obesity[28,29]. Moreover, positive neuroprotection effects of GLP-1 have also been observed[30].

In this work, the 31-residue GLP-1(7-37) version and its C-terminally amidated and lipidated analogues were studied. Further in the text, the residues in GLP-1(7-37) are numbered from 1 to 31 (not 7 to 37), starting from N-terminal His1.

Early structure-activity studies suggested that the N-terminal residues His1, Gly4, Phe6, Thr7 and Asp9 are essential for interaction with the GLP-1 receptor[31]. Similarly, Phe22 and Ile23 residues are also important for receptor binding as they stabilize the secondary structure of GLP-1[31].

Generally, naturally occurring peptide hormones have a short plasma half-life that limits their therapeutic usage[32] and GLP-1 is not an exception. *In vivo*, this peptide is degraded by the enzyme dipeptidyl-peptidase IV (DPP-4), which cleaves off the two N-terminal amino acids of GLP-1, which results in a biologically inactive form, GLP-1-(3-31)[33]. Additionally, neutral endopeptidase has a minor contribution to GLP-1 degradation by cleaving at six sites within the central and C-terminal GLP-1 region[34]; Asp9-Val10, Ser12-Tyr13, Tyr13-Leu14, Glu21-Phe22, Phe22-Ile23 and Trp25-Leu26. Since DPP-4 cleavage between Ala2 and Glu3 is the major cause of the short plasma half-life, a substitution of Ala2 is a common strategy to reduce fast GLP-1 degradation. Substituting Ala2 with α-aminoisobutyric acid (Aib) was found to completely prevent the GLP-1 cleavage by DPP-4[35]. In two other commercially-available GLP-1 analogues, albiglutide and dulaglutide, Ala2 is substituted with Gly for the same reason. Moreover, production of this therapeutic peptide is complicated by its intrinsic propensity to aggregate. GLP-1 was shown to aggregate under a wide range of conditions resulting in the formation of amyloid fibrils[36–38].

To overcome the rapid *in vivo* elimination of therapeutic GLP-1 analogues, various half-life extension strategies have been successfully employed including sequential modification, attachment of fatty-acid to peptide, fusion with human serum albumin, fusion with the fragment crystallizable region of a monoclonal antibody, sustained drug delivery systems, and PEGylation[16]. Sequential modifications of GLP-1 enhanced DPP-4 resistance and improved GLP-1 receptor activation potency. Using this strategy, twice-daily and once-daily analogues, exenatide and lixisenatide (Figure 1.1), respectively, have been developed[16]. Amino acid substitutions were shown to be effective in preventing proteolytic degradation, however, these modified GLP-1 analogues are still rapidly removed by renal clearance. Thus, other strategies have been employed to prolong peptide circulation in plasma. An attachment of a fatty acid to GLP-1 led to its slower absorption and mediated its albumin binding in plasma[39]. This approach was employed in the development of liraglutide and later

semaglutide (Figure 1.1). *In vitro* studies showed that while liraglutide could be a substrate for proteolytic degradation by both DPP-4 and neutral endopeptidase, its degradation was much slower compared to non-modified GLP-1, possibly due to the steric hindrance provided by the fatty acid and the reversible binding of liraglutide to albumin[40]. Compared to liraglutide, semaglutide has an additional amino acid substitution Ala2→Aib2, a longer stearic diacid (C18) and a longer optimized spacer connecting the lipid with the peptide backbone. These modifications in semaglutide result in the *in vivo* half-life of 7 days, allowing for once-weekly administration of the drug[22]. More complicated strategies that led to approval of GLP-1 therapeutical analogues were fusion with recombinant human serum albumin or an antibody fragment crystallizable domain **(**Figure 1.1).

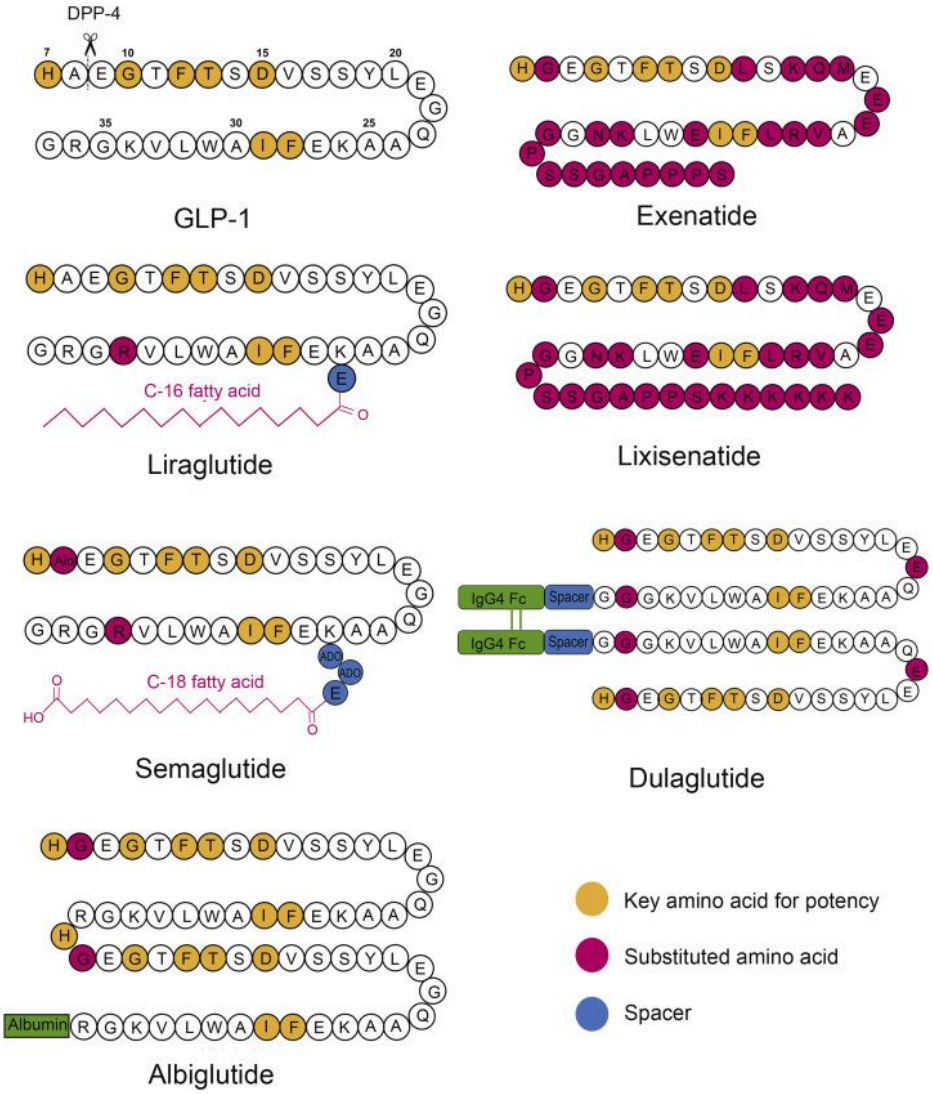


Figure 1.1: GLP-1 and its commercially-available analogues. Figure reproduced with permission from Yu *et al.* (2018)[16].

1.3 Aggregation of peptides and proteins

Aggregation of peptides and proteins is a widely studied phenomenon not only because of its great biological importance[41–45] but also for the development of novel materials[46,47], stability of peptide-based drugs[16,23,25,48–51] or therapeutic antibodies[52–54]. The term *aggregation* refers to self-association of multiple peptide/protein chains that are held together by non-covalent interactions. However, in some cases, aggregation may refer also to self-assemblies linked with covalent bonds[55]. Aggregation can be affected by many factors including peptide/protein concentration[36], pH[56], trace impurities, metal ions[57,58], salts[59–62], ionic strength[63–68], temperature[69], agitation and shear forces[70,71], and many other factors[72].

In general, protein or peptide aggregation is a self-assembly process that is often accompanied by changes in secondary structure. However, a rigorous definition of aggregation in this respect is lacking and terms “aggregation” and “self-assembly” are sometimes used interchangeably[73,74]. In this Thesis, the term “aggregation” is used for a self-assembly process which includes changes in secondary structure of the peptide, which are often induced by prolonged incubation. In contrast, the term “self-assembly” in this Thesis refers to the formation of higher molecular weight species which is not accompanied by a change in secondary structure of the peptide. Aggregation, as described in this Thesis, often leads to the formation of aggregates of various morphologies.

Probably the most studied form of aggregation is the formation of amyloid fibrils (fibrillation), mainly because of its link to neurodegenerative disorders. It is generally accepted in the field that amyloid fibrils are formed by a “nucleation-elongation” mechanism (Figure 1.2A)[44,75]. In this mechanism, monomers assemble into oligomeric species which can form nuclei that further rapidly elongate to form amyloid fibrils. In the latter stages, the process has a self-catalytic loop due to “secondary nucleation processes” on the surface of existing fibrils and/or fibril fragmentation followed by further elongation[76–79]. These secondary processes allow new aggregates to be formed through processes that involve existing aggregates. However, recently there have been an increasing number of observations of off-pathway oligomers being populated which affects the rate of fibril formation (Figure 1.2B)[36,80–82]. Off-pathway oligomers likely cannot form amyloid fibrils directly, instead, their formation competes with fibrillation. Nevertheless, off-pathway oligomers may be able to interconvert back to the monomer or to on-pathway oligomers and

thus, eventually convert into fibrils. This additional process affects the overall kinetics of the fibril formation[83].

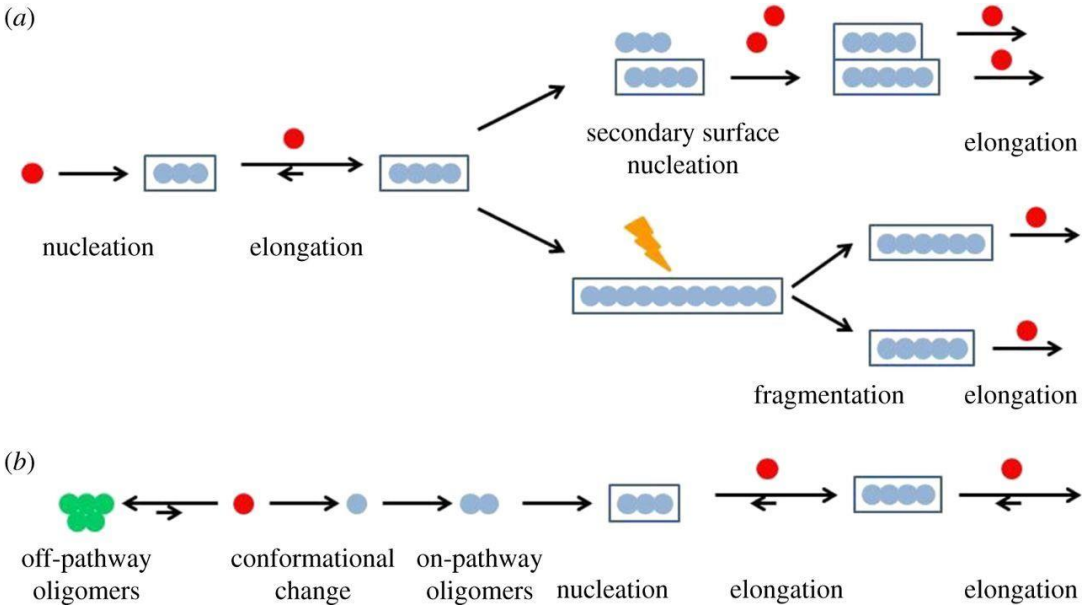


Figure 1.2: Amyloid fibril formation mechanisms. In the classical nucleation-elongation mechanism (A), peptide/protein monomers (marked as red circles) assemble first into oligomers and then nuclei that are further elongated into amyloid fibrils. The process can be self-catalysed by fibril fragmentation and secondary surface nucleation. (B) Additionally, there are mechanisms that involve off-pathway oligomers that may interconvert back to the monomer or to on-pathway oligomers and, therefore, affect the overall kinetics of the reaction. The figure was reproduced with permission from Zapadka *et al.*, *Interface Focus*, *7*, 2017[72].

A typical kinetic profile of fibrillation that follows a nucleation-elongation mechanism (Figure 1.2A) is depicted in Figure 1.3. The profile is typically of a sigmoidal shape, with three phases: a lag phase, a growth phase and a plateau phase (Figure 1.3A); the lag time and half-time of the fibrillation process generally decreases with increasing peptide/protein concentration[78,84], Figure 1.3B and C. The sigmoidal kinetic curves for the mass concentration of fibrils as a function of time are the result of fibril growth through nucleation, elongation and fragmentation[85].

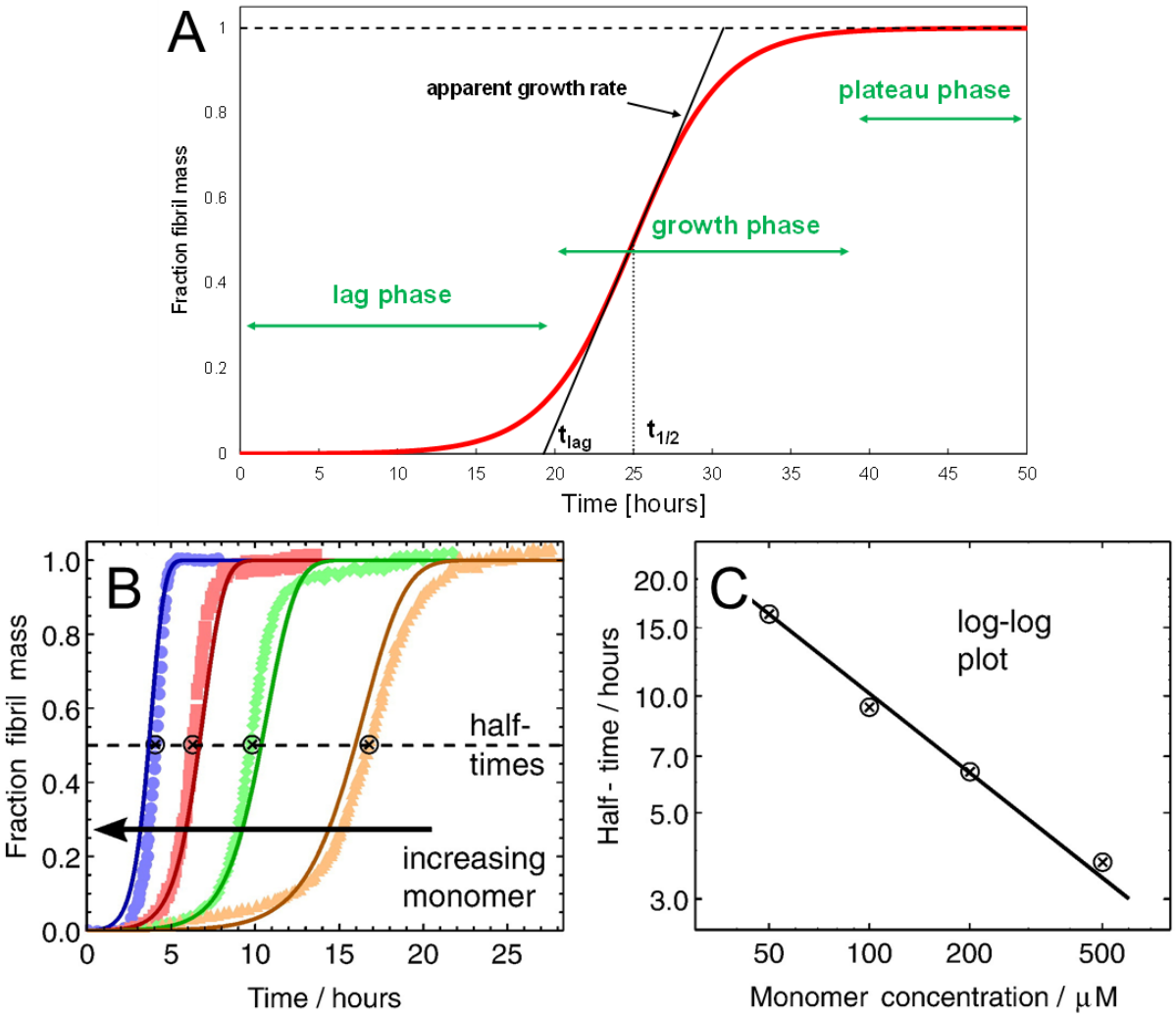


Figure 1.3: Typical fibrillation kinetic profiles corresponding to a nucleation-elongation mechanism. (A) A sigmoidal profile of a typical fibrillation reaction with lag, growth and plateau phases highlighted along with the kinetic parameters that can be measured. Figure adapted from Brichtová (2019)[86]. (B) Global fitting of fibrillation curves at different monomer concentrations of a WW domain[87]. The log-log plot of the kinetic data from B, half-time of the reaction *vs* monomer concentration. Panels B and C were adapted with permission from Cohen *et al.* (2012)[84].

The formation of amyloid fibrils, which follows the classical nucleation-elongation mechanism can be expressed in the form of the integrated rate laws for the mass $[M]$ and number $[P]$ concentrations of fibrils as a function of time t[78,85,88]:

$$\frac{d[P]}{dt} = k_n[m]^{n_c} + k_2[m]^{n_2}[M] + k_-[M] \quad (1.1)$$

$$\frac{d[M]}{dt} = 2k_+[m][P], \quad (1.2)$$

where $[m]$ is the monomer concentration, k_n is a rate constant for homogeneous formation of nuclei in solution, k_2 is a rate constant for heterogenous nuclei formation on the surface of existing fibrils, k_- is a fragmentation rate constant, k_+ is an elongation rate constant, n_c and

n_2 stand for the reaction orders of primary and secondary nucleation with respect to monomer concentration. The different terms in Eq. 1.1 represent the elementary microscopic processes. The first term ($k_n[m]^{n_c}$) accounts for the primary nucleation, the second term ($k_2[m]^{n_2}[M]$) represents secondary nucleation on the surface of existing fibrils and the third term ($k_-[M]$) describes the fibril fragmentation.

To include the saturation of secondary nucleation, Eq. 1.1 can be expressed as:

$$\frac{d[P]}{dt} = k_n[m]^{n_c} + k_2 \frac{[m]^{n_2}}{1 + [m]^{n_2}/K_M}[M] + k_-[M], \quad (1.3)$$

where K_M is the Michaelis constant given by $K_M = (k_b + \bar{k}_2)/k_f$; where k_b and $\bar{k}_2$ in the expression for the Michaelis constant stand for the rate constants for dissociation into n_2 monomers and a free site or into a new nucleus and a free site, respectively, k_f represents a rate constant for n_2 monomers binding to free fibril site[76,78,88].

Information about the elementary microscopic processes can be obtained from experimental kinetic measurements. A framework based on the global fitting of the data from experimental kinetic assays to integrated rate laws (Eq. 1.1, 1.2, 1.3) was developed. Currently, this framework can be used in the free online web-based platform, AmyloFit[88,89]. This approach starts from the kinetic experiments, in which the fibril mass is measured as a function of time, and the approach requires the fibrillation kinetic traces of multiple concentrations of the analyzed aggregating protein/peptide. The software is suitable for the systems in which a purified protein/peptide aggregate into linear aggregates, such as amyloid fibrils, following a nucleation-elongation mechanism.

Nevertheless, this global fitting-based framework could not be applied to the systems studied in this Thesis due to several reasons. In Chapter 3, the aggregation of GLP-1 and GLP-1-Am at neutral pH is studied. Under these conditions, GLP-1 and GLP-1-Am were shown to undergo multiple types of aggregation processes resulting in amyloid fibrils but also stable low-molecular weight oligomers of disordered structure. The formation of low-molecular weight oligomers not only competes with the formation of amyloid fibrils but the low-molecular weight oligomers were also shown to partly inhibit the fibrillation process. Due to these observations (competing aggregation pathways, partial inhibition of fibrillation) in GLP-1 and GLP-1-Am aggregation, it was not possible to fit the data to models used in the AmyloFit framework. In Chapter 5, the aggregation of lipidated analogues of GLP-1-Am is investigated. The lipidated GLP-1-Am analogues do not follow the classical nucleation-elongation mechanism. Their kinetic traces indicate the presence of

multiple underlying aggregation mechanisms and the morphology of aggregates is not always fibrillar. In addition, shown in Chapter 4, in solution the lipidated analogues exist mostly in the form of stable higher-molecular weight oligomers and the free monomers are present only in trace amounts. These features make the aggregation of lipidated GLP-1-Am analogues unsuitable for analysis in the AmyloFit framework.

1.4 Oligomeric states in peptide and protein aggregation

In recent years, protein and peptide oligomeric states present during the process of amyloid formation have received enormous attention as they have been considered to be the most detrimental species in protein misfolding diseases[90]. However, protein and peptide oligomers can have a very versatile character in many aspects such as their size, structure, toxicity, stability or their role in the aggregation process. The term "oligomer" usually represents an assembly of nanometer dimensions which comprises of a small number of molecules of smaller size[73].

Due to the link with protein misfolding diseases and also the safety control of protein and peptide biotherapeutics formulation, potential cytotoxicity and immunogenicity of oligomeric species has been extensively investigated. Several studies showed a correlation between oligomer toxicity and the oligomer size, such that toxicity generally decreases with an increase in oligomer size[90–92]. However, the most critical factor determining toxicity of oligomeric species was shown to be the presence of solvent-exposed hydrophobic residues on the oligomer surface[91,93,94]. This is particularly true if these exposed hydrophobic patches interact with the lipid bilayer of biological membranes[95,96]. On the other hand, the toxicity of oligomers appears to be rather independent of their secondary structure[93]. Although a high β-sheet content is often present in misfolded protein oligomers[97], it is rather a stabilizing element, not an essential prerequisite for toxicity[93].

With respect to the kinetic role of oligomers in the process of amyloid formation, the oligomers are termed as "on-pathway", if they are intermediates that react onward to form fibrils, or "off-pathway", if they are produced in side-reactions[36,65,66,80,98,99]. Nevertheless, some studies suggested that in many cases, this binary definition is not sufficient to describe all oligomers observed in aggregation reactions[83], e.g., only a small fraction of "on-pathway" oligomers is eventually transformed into amyloid fibrils since most of these oligomers

dissociate back into monomers[100]. The definition of "off-pathway" oligomer is unambiguous for stable oligomers whose formation competes with the fibrillation process or even partly inhibits fibrillation. The existence of these off-pathway oligomers have been described for GLP-1 and its C-terminally amidated variant (this work, Chapter 3) but also for SOD1[101] (stable off-pathway trimers) or Aβ peptide[102] (stable dimer). The inhibitory effect of off-pathway oligomers may lead to significant changes in the fibrillation mechanism. For some proteins and peptides, such as GLP-1[36] or human calcitonin[80], an inverse relationship of increasing fibrillation lag time with increasing peptide/protein concentration was observed. This switch in fibrillation kinetics was in both cases attributed to the formation and inhibitory effects of off-pathway oligomers.

The stability of peptide and protein oligomers ranges from transient intermediates to stable final aggregation products[99]. Also, the structure of oligomeric species is highly variable usually with larger species being more structured. Larger and more structured oligomers usually contain a high degree of β-sheet, which stabilize the structure[93,103]. On the other hand, many stable, yet highly disordered oligomers have been observed as well [101,102,104–111]. Additionally, it has been shown that some small aromatic molecules are able to drive the aggregation to stable disordered oligomers instead of amyloid fibrils[99,112–114]; e.g., for two amyloidogenic peptides/proteins amyloid-β and α-synuclein, it was demonstrated that the small aromatic molecule, polyphenol (-)–epigallocatechin gallate, inhibits formation of amyloid fibrils and unstructured non-toxic oligomers are formed instead[115].

1.5 Self-assembly and aggregation of lipidated peptides

Lipidated analogues of GLP-1 studied in this Thesis can generally be classified as peptide amphiphiles since they contain a largely hydrophilic peptidic part conjugated to a hydrophobic lipid chain. For peptide amphiphiles, the balance of many attractive and repulsive forces determines the morphology, size, structure of their assemblies and aggregates, and a deeper understanding of these contributions is important for the future research and applications in this field[116]. To a large degree, the hydrophobic elements drive the self-assembly and aggregation which may result in supramolecular structures of various morphologies including micelles, bilayers, amyloid fibrils, nanotubes or vesicles[117–119], Figure 1.4. A typical structure of a micelle has a hydrophobic inner core surrounded by a

hydrophilic water-exposed outer layer. Micelles can be spherical or they can have disc or worm-like structures[49,120]. For many systems, the self-assembly is peptide concentration-dependent and these amphiphilic molecules self-assemble above a certain concentration threshold, which is known as the critical aggregation concentration (cac)[49].

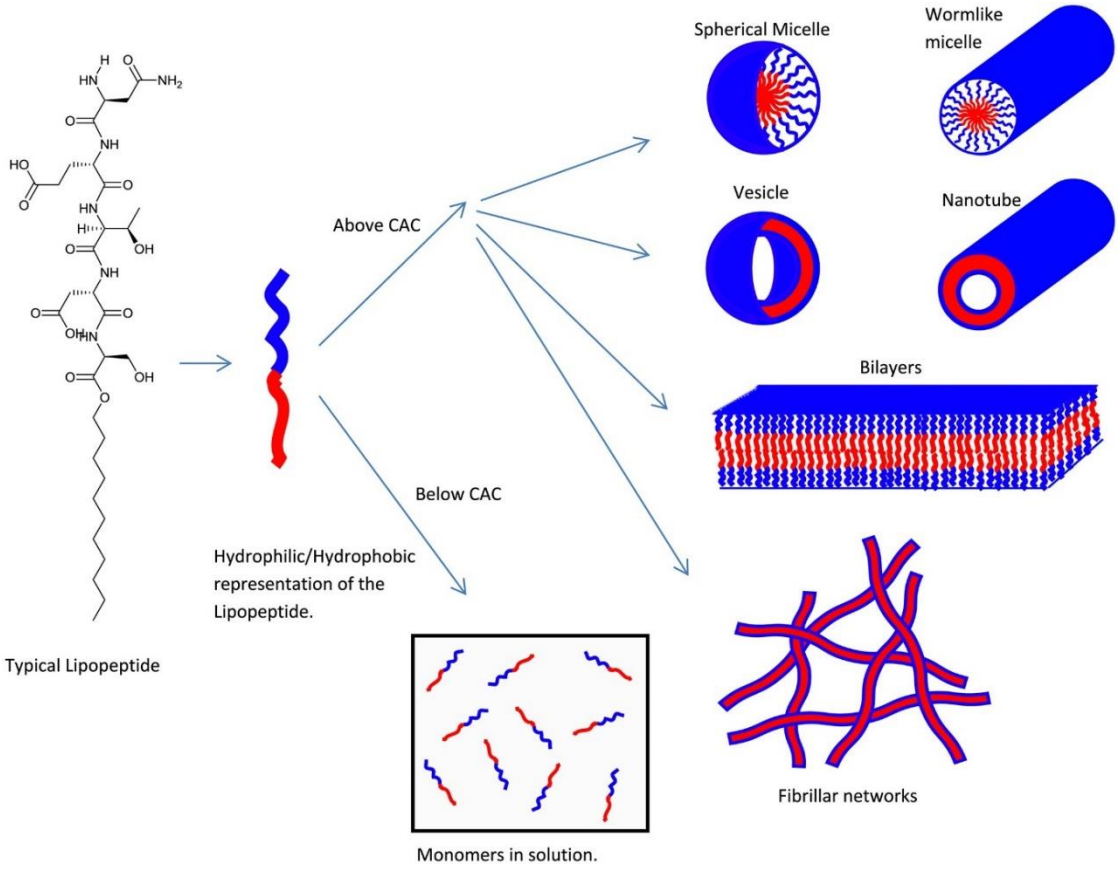


Figure 1.4: Self-assembly and aggregation of lipopeptides. The hydrophobic parts of the amphiphile are represented by red colour, hydrophilic parts are highlighted in blue. Reproduced from Hutchinson *et al.* (2017)[121].

It was previously shown that alkylation of amylin, a peptide that is prone to amyloid formation, can selectively regulate the morphology of observed aggregates or suppress their formation based on the length of the alkyl chain[122]. A similar effect was observed for lipidated exendin-4-derived dual peptide agonists[123]. Another study showed that self-assembly of toll-like receptor agonist lipopeptides into either spherical micelles or flexible worm-like micelles is dependent on the number of lipid chains attached to the peptide[124]. Several studies also showed that the morphology of self-assembled species/aggregates can be regulated by pH, ionic strength or temperature[17,125–127].

Conjugation of a hydrophobic lipid chain to a peptide often leads to the stabilization or even induction of secondary structure[128,129]. This is due to the fact that these amphiphilic molecules tend to self-assemble/oligomerize rapidly, which increases the local concentration

of the peptide and increase in intramolecular non-covalent bonds which can help to stabilize specific secondary structure[116]. This effect has mostly been demonstrated for α-helical peptides, e.g., for collagen-based peptides, it was established that a series of dialkyl chains enhances the thermal stability of the α-helical conformation, compared to non-lipidated peptides[130].

1.6 Structure of amyloid fibrils

The amyloid structure is undoubtedly one of the most frequently occurring structural motifs observed in self-assembled aggregates of peptides, proteins and peptide and protein amphiphiles. The amyloid structure was first observed over 85 years ago as a unique pattern in X-ray diffraction due to cross-β sheet reflections[131]. Amyloid fibrils are structurally defined as fibrillar aggregates with a cross-β structure. The central spine of the fibril is formed by an extended β-sheet structure[132]. The cross-β structure has characteristic distances between the individual polypeptide chains forming the structure[133] – main chain spacing of 4.7 to 4.8 Å (Figure 1.5A) and side chain spacing of approximately 10 Å (Figure 1.5B). The main chain spacing, i.e., the distance between two hydrogen-bonded β-strands within the same β-sheet, is relatively conserved between different polypeptide sequences. The side chain spacing, i.e., the packing distance between two parallel β-sheets, shows greater variability with distances from 8.8 to 14.6 Å being observed[132]. Amyloids also bind Congo Red dye with resulting green birefringence under cross-polarized light[134]. Many peptides and proteins, if not nearly all[135], are capable of amyloid fibril formation under certain conditions despite the fact that they do not share any similarities in terms of sequence or structure.

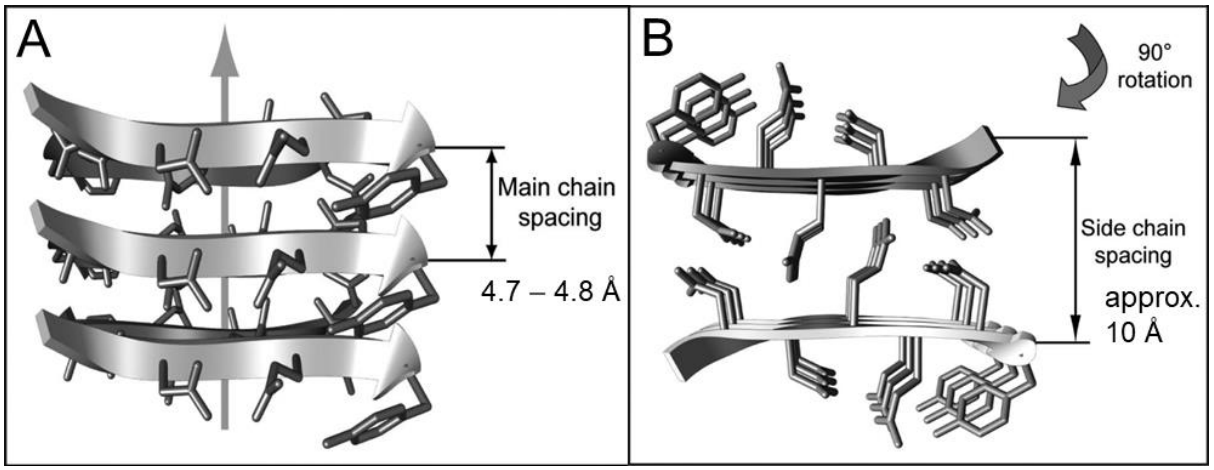


Figure 1.5: Cross-β-sheet structure of amyloid fibrils. Side (A) and top (B) view of the cross-β core structure of the yeast protein Sup35 fibril obtained by X-ray diffraction (PBD ID: 1YJP)[136] illustrating the main chain and side chain spacing and their dimensions. Figure adapted from Fändrich *et al.* (2009)[132].

Despite the common cross-β sheet structure found in the amyloid core, there are many different possible ways in which to pack protofilaments which result in different fibril morphologies even for the same protein/peptide. This is known as structural polymorphism and it has been observed for fibrils formed both *in vivo* and *in vitro*[132,137–139]. Fibril morphology is influenced by various environmental factors, e.g. pH, temperature or ionic strength[132]. However, ensembles of fibrils with different morphological types from the same peptide/protein may coexist within the same sample. Different fibril morphologies originate from three main causes: firstly, fibrils may consist of a different number of protofilaments, secondly, the mutual orientation of protofilaments may differ, third, the protofilament substructure may differ. Figure 1.6 depicts different structures of amyloid fibrils presented in a view down the fibril axis, which have been determined using cryo-EM[140]. This figure nicely illustrates differences in number and mutual orientation of protofilaments within different fibril structures.

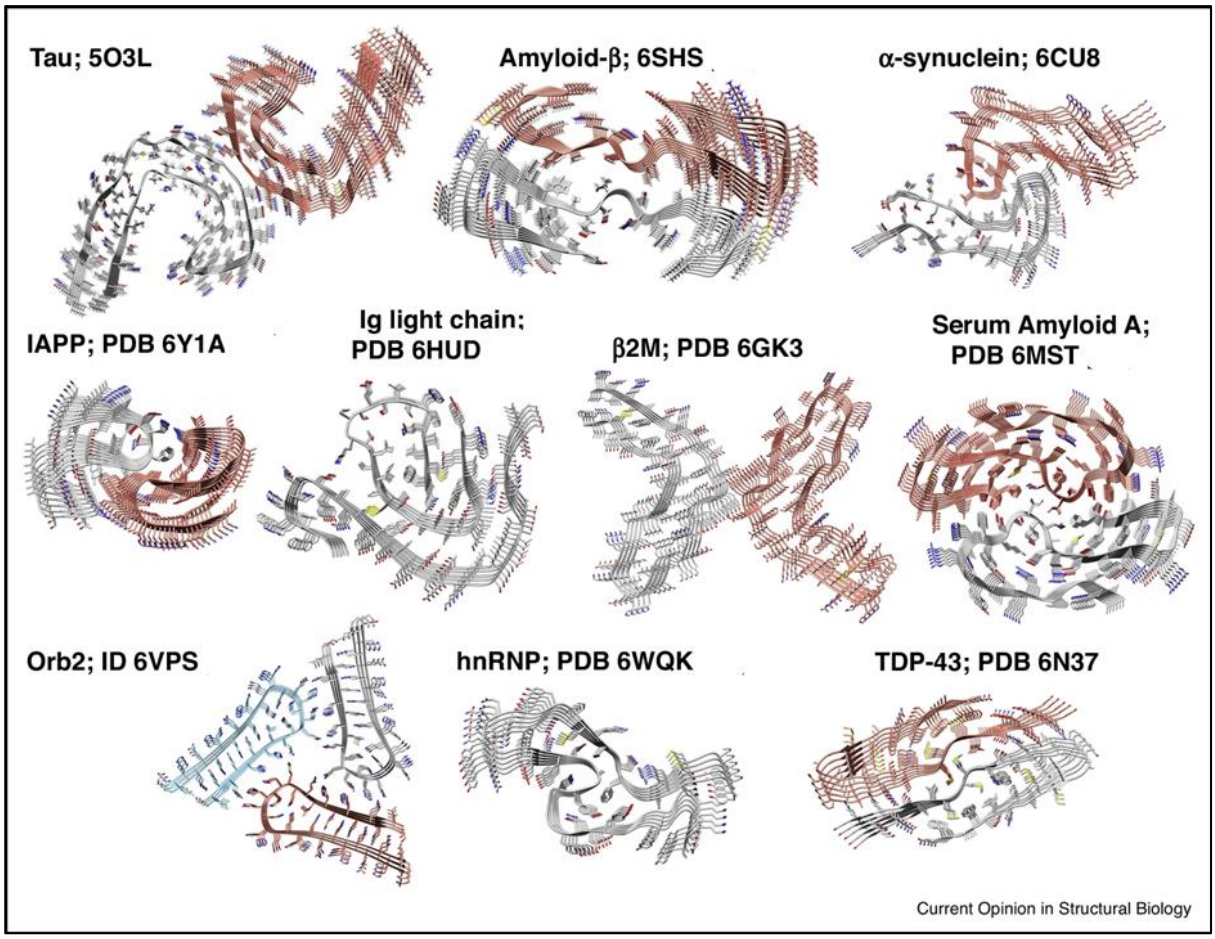


Figure 1.6: Amyloid fibrils: examples of different structures obtained by cryo-EM. Structures of amyloid fibrils of tau[141], amyloid-β[142], α-synuclein[143], IAPP[144], Ig light chain[145], β2M[146], serum amyloid A[147], Orb2[148], hnRNP[149] and TDP-43[150] as determined by cryo-EM. Fibrils are visualized in a view down the fibril axis. The backbone of each protofilament is coloured differently. Figure reproduced with permission from Ragonis-Bachar, P., & Landau, M. (2021). *Current Opinion in Structural Biology*, 68, 184-193.

Various approaches and methods have been developed to predict which regions of proteins/peptides have an intrinsic propensity to form amyloid – APR (aggregation-prone region). The propensity to aggregate is affected by both amino acid sequence and secondary and tertiary structure of the protein. The prediction methods can be roughly divided into two categories based on the type of input, which can be either a protein sequence or a PDB structure. Table 1.1 provides a summary of the most frequently used methods and highlights the results of the APRs predicted by each for GLP-1.

Sequence input methods		
Method name	Basic principle	Result on GLP-1
Tango[151]	statistical mechanics algorithm, based on physico-chemical principles of β-sheet formation	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
Aggrescan[152]	amino acid scoring derived from *in vivo* mutation experiments of the amyloid-β peptide	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
Cordax[153]	machine learning based method using hexapeptide amyloid core structures	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
FoldAmyloid[154]	amino acid scoring based on packing density and hydrogen bonds observed in spatial structures of proteins	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
Waltz[155]	algorithm based on position-specific properties of amyloidogenic hexapeptides	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
MetAmyl[156]	consensus method, combines 11 prediction algorithms;	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG

	PASTA, SALSA, AGGRESCAN, PAFIG, FoldAmyloid (5 tables), Tango and Waltz	
Structure input methods		
Method name	Basic principle	Result on GLP-1
AGGRESCAN 3D 2.0[157]	combines protein structure and amino acid scoring from AGGRESCAN method	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG
NetCSSP[158]	predicts contact-dependent secondary structure propensities using artificial neural network-based algorithms	HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG

Table 1.1: Overview of the most widely used aggregation prediction algorithms and their results on GLP-1. Table was adapted from Housmans *et al*. (2021)[159]. For structure-based methods, a crystal structure of GLP-1 bound to its receptor was obtained from the Protein Databank, ID: 5OTU[160,161], and used as an input. Parts of GLP-1 sequence highlighted in red indicate β-sheet segments.

1.6.1 Non-cross-β-sheet fibril morphologies

Nowadays, there are several examples of non-cross-β-sheet fibril morphologies which have mainly been observed for cytotoxic microbial functional amyloids or natural antimicrobial peptides. In these cases, the peptides form a cross-α fibril morphology[162–164]. In cross-α fibrils, individual peptide chains are in an α-helical conformation within the fibril core. The α-helices are positioned perpendicular to the fibril axis and they are stacked into sheets that are parallel to the fibril axis and which interact through the hydrophobic faces of the α-helices, analogously to β-sheets in cross-β structures, Figure 1.7.

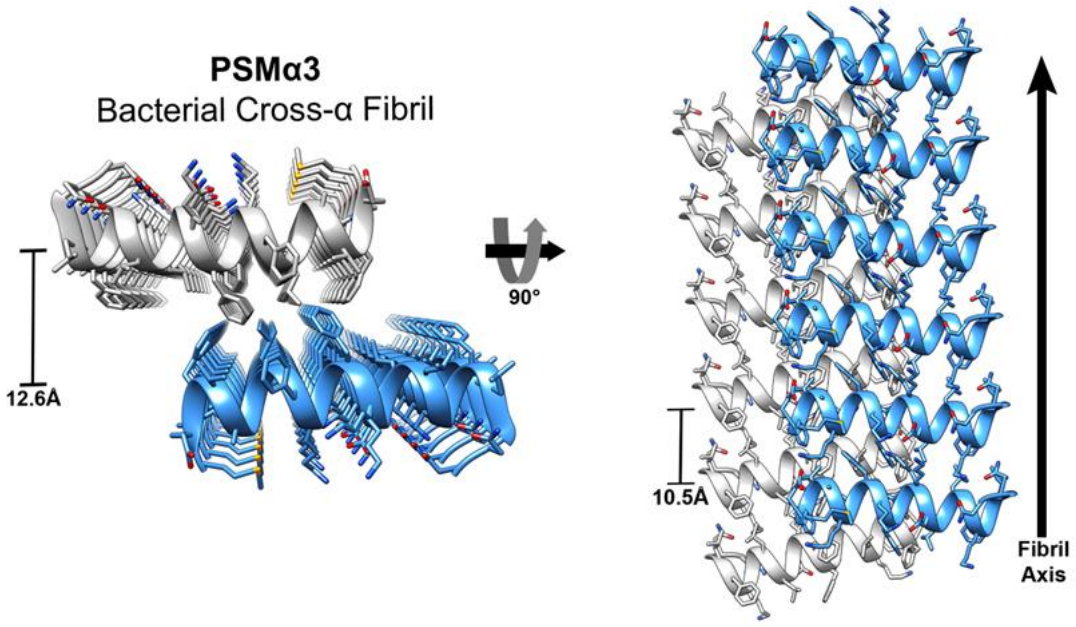


Figure 1.7: Cross-α fibril architecture of PSMα3 bacterial peptide. Figure adapted from Salinas *et al*. (2021)[165].

A cross-α fibril structure has been reported for the 22-residue phenol-soluble modulin α3 (PSMα3), which is secreted by *Staphylococcus aureus*[164]. Another example of a cross-α architecture was reported for the amphibian antimicrobial peptide uperin 3.5. Interestingly, uperin 3.5 exhibits a switch in secondary structure between the cross-α and cross-β fibril morphology depending on the presence or absence of lipids[165]. This has been suggested to be a storage mechanism for this cytotoxic peptide[165]. In both cases, the cross-α fibril morphology is connected with a high cytotoxicity of the peptide/fibrils and both uperin 3.5 and PSMα3 fibrils can be classified as functional fibrils[163,166].

1.7 Techniques to study aggregation and amyloid fibril formation

1.7.1 Morphological characterization

Morphology of amyloid fibrils and other larger (> *circa* 10 nm) protein aggregates is usually studied using advanced microscopic techniques such as transmission electron microscopy (TEM) and atomic force microscopy (AFM). In TEM, samples of protein aggregates are deposited on carbon-coated copper grids and negatively stained with a solution containing a heavy metal to increase the contrast. TEM images at high resolution can provide information

on the morphology of fibrils, their homogeneity and dimensions. In AFM, amyloid fibrils/aggregates can be visualized in aqueous solutions, under physiological-like conditions. AFM images provide information about fibril length, width and height. This technique also allows measurement of force and elastic properties of fibrils[167–170].

Nowadays, high-resolution techniques, which can provide information about both structure and morphology of fibrils, are available. These are mainly cryo-electron microscopy, solid-state NMR and X-ray diffraction. Due to the recent advances in solid-state NMR and particularly in cryo-electron microscopy, many new structures, close to atomic resolution, have been elucidated[140,171]. Solid-state NMR can be used to determine dihedral angles and distances between atoms in the subunits of fibrils. This information can be used to build atomic models of fibrils or their subunits, which can later be fitter into lower-resolution electron microscopy models[172–174]. X-ray diffraction has been used for structure determination of small amyloidogenic peptides, although these peptides may not crystalise in the fibril form, the interactions elucidated in the peptide crystal may still give insight into the amyloid structure which is formed by the same peptide[175,176]. However, the field has been recently revolutionized by advances and increasing resolving power of cryo-EM[177]. These advances have been driven by the technological progress in microscope and electron detection equipment and improvements in image processing software[178]. The recently solved cryo-EM structures of fibrils have revealed the differences between pathological and functional amyloids, which can be related to the fibril stability and reversibility of their formation[140].

1.7.2 Structural characterization

1.7.2.1 Circular dichroism

Circular dichroism (CD) is defined as the differential absorption of left- and right-circularly polarized light. CD is used in the analysis of a wide range of chiral optically active molecules, which absorb light preferentially in one direction. In proteins and peptides, the chromophores of interest include the peptide bond, which absorbs below 240 nm; aromatic amino acid side chains with absorption between 260–320 nm; and disulphide bonds which have a weak, broad absorption band at around 260 nm. Information about the secondary structure of proteins can be obtained from the CD spectra in the far-UV region (180–260 nm), in which the peptide bonds, i.e., the protein backbone, absorb. Usually, a β-sheet

structure is characterized by a maximum at *circa* 196 nm and a single broad minimum at around 218 nm. α-helices have two characteristic minima at 208 and 222 nm[167,179]. A single minimum in the spectrum at *circa* 198 nm is attributed to disordered conformations. Deconvolution algorithms can be employed to quantify the contribution of the different secondary structure elements to the signal. In this Thesis, two algorithms – Contin-LL method with dataset 3 from DichroWeb server[180–182] and BeStSel method[183–187] – are used to estimate the different elements of secondary structure. The Contin-LL method uses ridge regression techniques and incorporates variable selection[188] which filters the datasets so that only the closest matching spectra to the analyzed spectrum are used in the analysis[189]. The "dataset 3" from DichroWeb server contains 37 reference spectra of different soluble globular proteins in the wavelength range from 185 to 240 nm, whose structures are known. The BeStSel method was developed and optimized specifically for β-sheet-rich proteins including aggregates with high a content of β-sheet[183]. It is based on the SP175 reference spectra set of 71 soluble globular proteins[190] which is enriched with three additional proteins – native β_2-microglobulin, amyloid fibrils of K3 peptide fragment of β_2-microglobulin, and fibrils of Alzheimer's amyloid-β (1–42) peptide. The BeStSel fitting algorithm is based on linear least-square regression methods. However, all estimates of secondary structure taken from either of these algorithms have to be treated with caution. Results using these methods should primarily be used as indicators of changes in secondary structure rather than absolute values[167]. Moreover, the available algorithms are optimized mostly for soluble proteins and the presence of aggregates in the sample may lead to distortions and artifacts, and therefore, lead to inaccuracies in the absolute determination of secondary structure content[167]

The signal in the near-UV region of the CD spectrum (250–300 nm) arises from aromatic amino acid side chains and depends on their mobility and environment. The presence of a near-UV CD signal reflects restricted movement of aromatic side chains which are fixed within a chiral environment, e.g., in a protein core, in an oligomeric structure, or some other species such as an aggregate. The absence of any signal in this region denotes that the rotation of the side chains of aromatic residues is not restricted[191,192].

1.7.2.2 Infrared spectroscopy and vibrational circular dichroism

IR spectra of peptides and proteins are dominated by the amide I signal (mainly carbonyl C=O stretch) and they are routinely used for secondary-structure determination. An amide I stretch occurs between 1600 and 1700 cm^{-1} and the exact wavenumber of this absorption band maxima depends on the secondary structure of the protein backbone. Information about

the changes in secondary structure can be also obtained from the amide II absorption maximum (between 1400 and 1500 cm^{-1}), which correspond to N-H in-plane bending and C-N stretching[193].

Amyloid fibrils typically display a maximum of amide I band between 1611 and 1630 cm^{-1}, whereas for a "native" β-sheet in globular proteins, the amide I maximum lies between 1630 and approx. 1643 cm^{-1}. For proteins with disordered and α-helical secondary structure, the amide I maximum occurs approx. in the range of 1637–1645 cm^{-1} and 1645–1662 cm^{-1}, respectively[194]. Due to the significant interference of the H-O-H bending mode of water with the amide I band, IR experiments with proteins are often done in the solid state (dried or lyophilized samples) or in deuterated buffers, since the D-O-D bending mode does not overlap with the amide I band.

Vibrational circular dichroism (VCD) is defined as the difference in the absorbance of left- minus right-circularly polarized light for a molecule undergoing a vibrational transition[195]. In VCD, the stereoselectivity of CD complements the structural specificity of IR. VCD is nowadays increasingly employed in aggregation studies due to the signal enhancement and splitting often observed for amyloid fibrils. Previous studies have demonstrated the use of VCD as a unique and sensitive probe of peptide and protein fibrillation dynamics in solution[196–199]. When amyloid fibrils are present in the solution, a VCD spectrum shows a large enhancement in the amide I and amide II vibrational regions which is induced by a highly regular cross-β sheet structure [200–204]. Most likely, this effect originates from the long-range vibrational coupling of inter-β-sheet contacts as well as possible braiding of multiple β-sheet protofilaments[200,202]. VCD can, therefore, be used for monitoring a fibrillation process and structural characterization of fibrils formed. Based on the sign pattern of the enhanced spectra, the different morphologies of fibrils, or the overall organization of the macro-assembly can be distinguished [138,139,197,205,206].

1.7.2.3 Intrinsic tryptophan fluorescence

The intrinsic fluorescence signal of tryptophan and tyrosine side chains has been extensively employed to study protein conformational changes. Tryptophan and tyrosine fluorescence emission spectra are dependent on the local environment of these residues. A steady-state fluorescence spectrum is obtained by excitation of a sample at a fixed wavelength followed by recording the emission fluorescence intensity over a specific wavelength range. The maximum of tryptophan fluorescence occurs between 315 and 360 nm after sample

excitation at 280 nm, where the tryptophan side chain displays an absorption maximum. The wavelength of the maximum fluorescence intensity (λ_{max}) varies with the polarity of the local environment of the tryptophan side chain and therefore with the extent of side chain burial in the protein structure/aggregate. For example, tryptophan in a hydrophobic pocket, e.g., in the core of an aggregate, λ_{max} can be as low as 315 nm, whereas tryptophan in a position that is fully exposed to the solvent exhibits a λ_{max} of approx. 355–360 nm[207,208].

The fluorescence of tyrosine side chain is less sensitive to environmental changes compared to tryptophan, however, the shift of fluorescence maxima between exposed and fully buried side chain is still significant[167]. Tyrosine fluorescence exhibits its maximum on average at around 305 nm, after sample excitation at 260 nm.

1.7.3 Monitoring the kinetics of aggregation

1.7.3.1 Thioflavin T assay

Thioflavin T (ThT) binding assays have been widely used to monitor amyloid formation. ThT is known for its capability to bind β-sheet-rich structures, such as amyloid fibrils[209]. When ThT is bound to a β-sheet structure, it has an enhanced fluorescence emission at around 480 nm upon excitation at 448 nm[210]. Due to these features, ThT has become "a gold standard" for identifying and monitoring the formation of amyloid fibrils both *in vivo* and *in vitro*[167,211]. Nevertheless, the specificity and selectivity of ThT for amyloid fibrils remains under question – multiple examples of non-amyloid proteins or aggregates with suitable cavities have been reported to bind to ThT and increase its fluorescence[212–216]. In addition, ThT fluorescence intensity is not a quantitative parameter since it is dependent on both fibril morphology and electrostatic charge on the surface of fibrils[217], moreover, amyloid fibrils which do not cause ThT fluorescence enhancement have also been reported[217,218].

1.7.3.2 8-anilinonaphthalene-1-sulfonic acid assay

8-anilinonaphthalene-1-sulfonic acid (ANS) can be employed as a molecular probe to monitor early conformational changes associated with aggregation. ANS has been used in folding and aggregation studies due to its ability to detect exposed hydrophobic regions of proteins/peptides[94,219]. Using an excitation wavelength of 355 nm, ANS is barely fluorescent in polar environments, however, in a hydrophobic environment, e.g., upon interaction with

exposed hydrophobic patches of proteins/peptides, its fluorescence increases with a fluorescence emission maximum at around 500 nm.

1.7.4 Monitoring of oligomeric populations using size-exclusion chromatography and analytical ultracentrifugation

Size-exclusion chromatography (SEC) can be used as an analytical technique to detect the amount of aggregates or stable oligomeric species in a sample. SEC separates the molecules based on their hydrodynamic radii as they pass through the SEC resin packed in the column. Molecules in SEC do not directly bind to the chromatography resin but they interact with pores in the resin and the extent of this interaction affects their elution volume. Smaller molecules have greater interaction with the pores of the resin (as they are able to enter the pores), and therefore they are delayed as they pass through the column, which results in greater elution volumes compared to larger molecules.

Another technique which can analyse the size distribution of species in a sample is analytical ultracentrifugation. There are two main techniques in analytical ultracentrifugation – sedimentation equilibrium and sedimentation velocity. In sedimentation equilibrium experiments, lower centrifugal forces are used compared to sedimentation velocity experiments. Under these conditions, sedimentation is balanced by diffusional transport and sedimentation equilibrium is achieved when the net transport vanishes throughout the solution[220]. In sedimentation velocity experiments, the radial concentration distributions (boundaries) of solutes during transport through a centrifugation field are monitored over time. At the start of the experiment, the protein concentration distribution along the radial position of the cell is uniform. At various sedimentation times, at constant speed and temperature, the boundary moves down the cell as the solute is depleted from the sample meniscus. The shift of the sedimentation boundary, in the sedimentation curves recorded over time, is indicated by a red arrow in Figure 1.8A. During the experiment, individual sedimentation curves at different time points are recorded (Figure 1.8A), fitted to a suitable model (Figure 1.8A and B) and based on this model, the sedimentation coefficient distribution is predicted (Figure 1.8C).

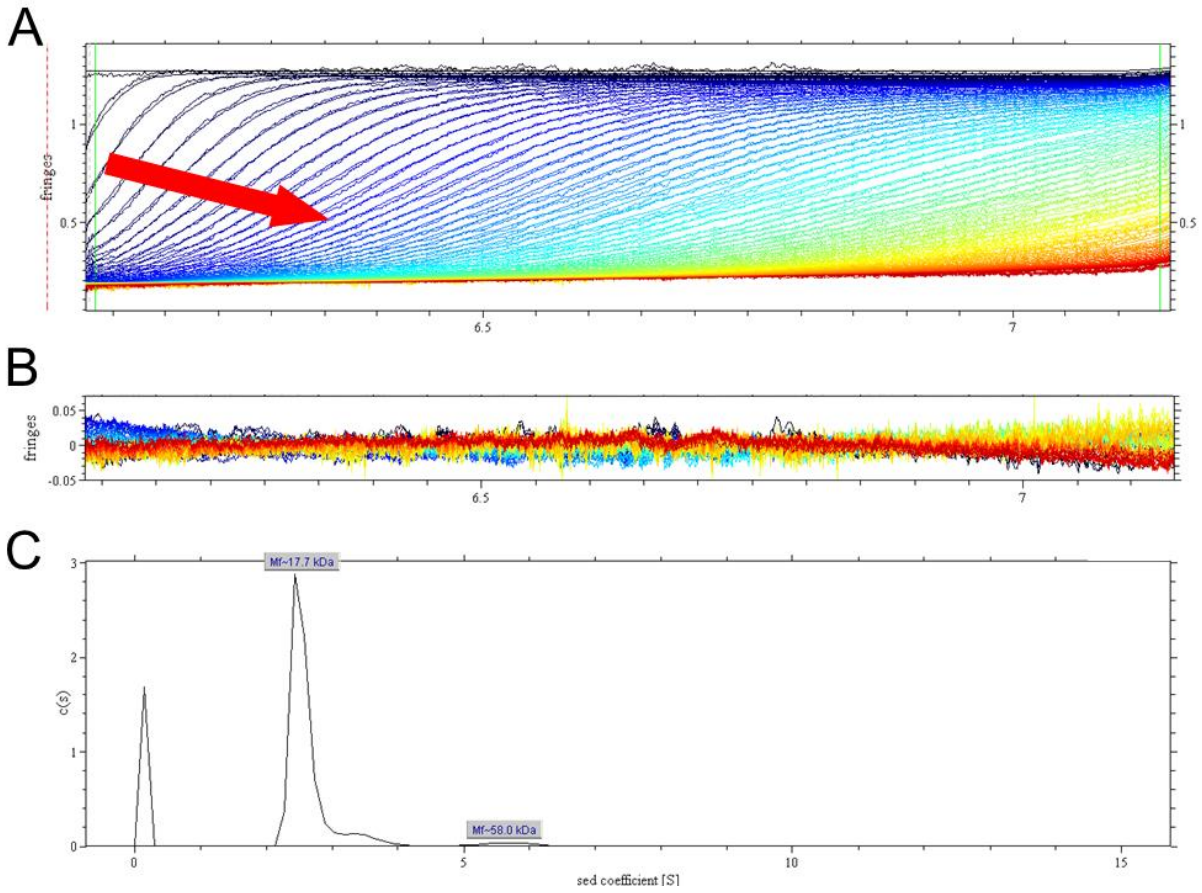


Figure 1.8: A sedimentation velocity experiment analysis. The analysis of sedimentation curves was performed in the Sedfit program[220–222]. (A) Sedimentation curves were fitted to the c(s) distribution model. The shift of the sedimentation boundary over a time course of the experiment is illustrated by the red arrow. (B) The residual plot shows the goodness of the fit. Based on the model, the sedimentation coefficient distribution is calculated (C).

1.7.4.1 Principle of sedimentation velocity

In sedimentation velocity experiments, the centrifugal force leads to the depletion of macromolecules at the meniscus and the formation of a concentration boundary which moves towards the bottom of the cell during the experiment, Figure 1.8A. The sedimentation coefficient, s, is given by the Svedberg equation:

$$s = \frac{u}{\omega^2 r} = \frac{M(1-\bar{\nu}\rho)}{N_A f} = \frac{MD(1-\bar{\nu}\rho)}{RT} \tag{1.4}$$

where u is the observed radial velocity of the macromolecule, ω is the angular velocity of the rotor, r is the radial position, $\omega^2 r$ is the centrifugal field, M is the molar mass, $\bar{\nu}$ is the partial specific volume, ρ is the density of the solvent, N_A is Avogadro's number, f is the frictional coefficient, D is the diffusion coefficient, and R is the gas constant. The s-values are reported in Svedberg [S] units.

The Stokes equation can be used to determine the frictional coefficient of a spherical particle (f_0):

$$f_0 = 6\pi\eta R_0 \tag{1.5}$$

where η is the viscosity of the solution and R_0 is the radius of the sphere. By substituting R_0 with[223]:

$$R_0 = \left(\frac{3M\bar{\nu}}{4\pi N_A}\right)^{\frac{1}{3}} \tag{1.6}$$

the relationship for the sedimentation coefficient of a spherical particle (s_{sphere}) can be obtained in the form:

$$s_{sphere} = \frac{M(1-\bar{\nu}\rho)}{N_A 6\pi\eta \left(\frac{3M\bar{\nu}}{4\pi N_A}\right)^{1/3}} \tag{1.7}$$

A correction of the experimental s-values to a standard state of water at 20 °C is necessary for comparison of data obtained by different instruments and under different conditions[220]. The correction is calculated using:

$$s_{20,w} = s_{T,B}\left(\frac{\eta_{T,B}}{\eta_{20,w}}\right)\frac{(1-\bar{\nu}\rho)_{20,w}}{(1-\bar{\nu}\rho)_{T,B}} \tag{1.8}$$

where *T* and *B* stands for the experimental temperature and buffer conditions and *20,w* indicate standard conditions. The ratio of the maximum *s*-value to the observed *s*-value, $s_{sphere}/s_{20,w}$, is equal to the ratio of the experimental frictional coefficient to the minimum frictional coefficient (f/f_0), which is an indication of the maximum shape asymmetry from a sphere.

The output of the sedimentation velocity experiments are concentration profiles in the radial direction as a function of time. Nowadays, the models for the experimental data are based directly on the transport equation, also known as the Lamm equation:

$$\frac{\partial\chi(r,t)}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[rD\frac{\partial\chi(r,t)}{\partial r} - s\omega^2 r^2\chi(r,t)\right] \tag{1.9}$$

This equation describes the evolution of the concentration distribution of macromolecular species χ as a function of time and radial position under the influence of sedimentation and diffusion[224].

1.7.4.2 Comparison of analytical ultracentrifugation and size-exclusion chromatography

Sedimentation velocity experiments are more demanding on equipment, instrumentation, time and data processing compared to size-exclusion chromatography. However, it also has several advantages; the samples do not undergo dilution or interaction (e.g., with the column matrix) during the experiment, therefore, less stable oligomeric species, which may otherwise dissociate upon dilution or matrix interaction, can be detected. In addition, sedimentation velocity can be applied to a wide range of particle sizes and therefore the mass of the analysed species can be predicted with a greater accuracy since the sedimentation rates depend on the 2/3 power of the molar mass compared to size-exclusion chromatography which depends more on the Stokes radius and the separation is based on the 1/3 power of the mass[220].

1.8 Previous work in the Jackson group on GLP-1 aggregation

Aggregation of GLP-1(7-37) and its C-terminally amidated variant, GLP-1-Am, has been previously studied in the Jackson group[36,218,225]. It was discovered that both peptides form amyloid fibrils under a wide range of conditions, however, the exact aggregation propensity and kinetics is pH-dependent. In Zapadka *et al.*[36], the fibrillation of GLP-1 was studied as a function of peptide concentration between pH 7.5 and 8.2, Figure 1.9. At pH 8.2, the fibrillation kinetics are consistent with a nucleation-elongation mechanism, Figure 1.9E and F. However, at pH 7.5, the lag time and half-time of fibrillation increase with increasing peptide concentration (Figure 1.9A and B), which is the opposite of what is expected in a nucleation-elongation mechanism. This observation was attributed to the formation of off-pathway species and an initial unimolecular step, in which monomer converts to a different monomeric form that can form oligomers on-pathway to fibril formation.

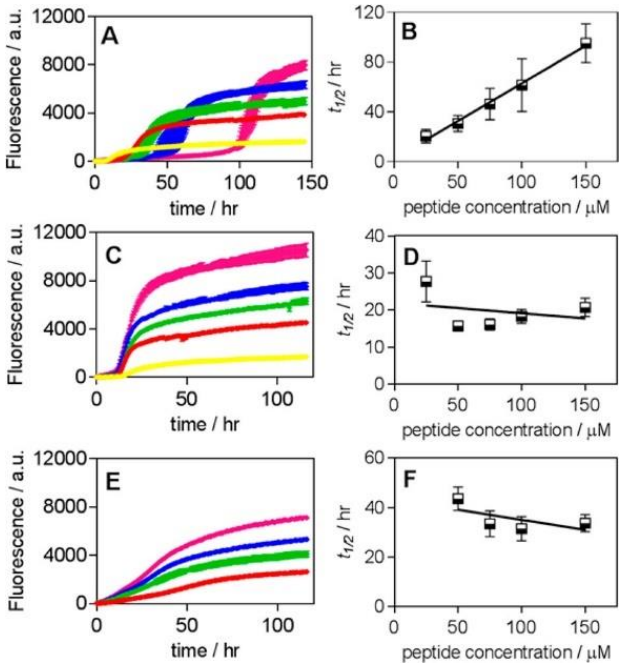


Figure 1.9: Fibrillation kinetics of GLP-1 as a function of peptide concentration at pH 7.5, 7.7 and 8.2. GLP-1 fibrillation kinetics was monitored using ThT assay at pH 7.5 (A, B), pH 7.7 (C, D), and pH 8.2 (E, F). Panels A, C, and E show typical ThT fluorescence traces. Each peptide concentration was run in triplicate. Panels D, E, and F show the dependence of the $t_{1/2}$ on the concentration of GLP-1. In panels B, D and F, the error bars correspond to the standard deviations of the kinetic parameters determined from three independent experiments each of which was run in triplicate. GLP-1 concentration colour scale for panels A, C, E: pink – 150 μM; blue – 100 μM; green – 75 μM; red – 50 μM; yellow – 25 μM. Reprinted with permission from *J. Am. Chem. Soc.* 2016, 138, 50, 16259-16265. Copyright 2023 American Chemical Society.

In Becher (2020)[225], the fibrillation of GLP-1 and GLP-1-Am was investigated over a wider range of pH conditions. The propensity for fibrillation was shown to be highly pH sensitive, Figure 1.10A and B. Moreover, it was shown that both GLP-1 and GLP-1-Am have the lowest fibrillation propensity around their isoelectric point, Figure 1.10C, which is 5.8 for GLP-1 and 6.8 for its C-terminally amidated variant.

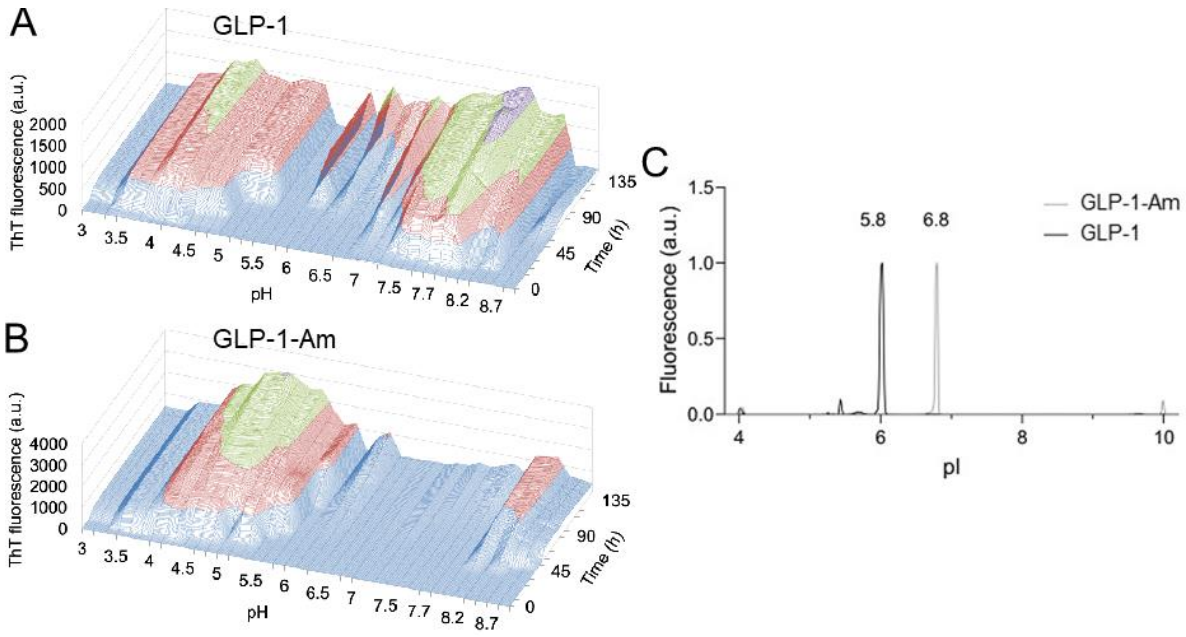


Figure 1.10: Aggregation profiles of GLP-1 and GLP-1-Am over a pH range 3–8.7 and isoelectric point determination. Aggregation of 75 µM GLP-1 (A) and 75 µM GLP-1-Am (B) in a pH range from pH 3.0 to 8.7 at 37 °C was monitored by a thioflavin T assay for 135 hours. The x-axis shows the pH value, z-axis time and the y-axis ThT fluorescence with a signal colour scale: 0–500 blue, 500–1000 red, 1000–1500 green, 1500–2000 purple. An isoelectric point of GLP-1 and GLP-1-Am was determined using imaged capillary isoelectric focusing (C). Figure adapter from Becher (2020)[225].

In my M.Phil Thesis[86], the aggregation of GLP-1-Am at neutral pH conditions, i.e., in the proximity of its isoelectric point, was investigated in greater detail. It was shown that the fibrillation under these conditions is slow with lag times > 50 hours. Additionally, apart from formation of amyloid fibrils, the formation of stable low-molecular weight oligomers (Figure 1.11) was observed over time. These oligomers remained in the solution even after depletion of the monomeric peptide and coexisted in solution with amyloid fibrils.

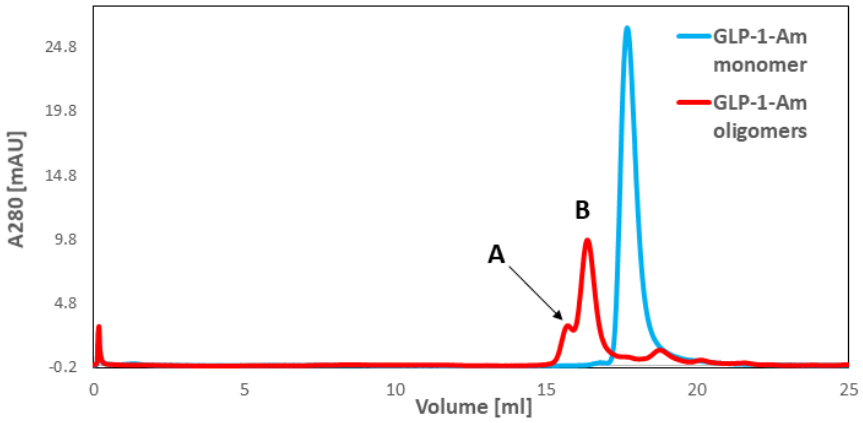


Figure 1.11: Size-exclusion chromatography elution profiles of the GLP-1-Am monomer and oligomers. The sample was analysed prior to the incubation (blue line) and after 16 days (red line) of incubation at 37 °C with continuous shaking at 80 rpm. After 16-day sample incubation two main oligomeric species A and B were observed. Elution conditions: 25 mM sodium phosphate buffer at pH 8, flow rate 0.75 mL min^{-1}, Superose 12 10/300. Figure adapted from Brichtová (2019)[86].

It was also shown that these low-molecular weight oligomers are distinct from both monomeric peptide and amyloid fibrils in terms of their secondary structure, Figure 1.12. Using far-UV CD, the low-molecular weight oligomers were determined to have a high content of disordered structure, Figure 1.12C. Chapter 3 of this Thesis, expands the investigation of formation and stability of these oligomers and focuses on their effect on fibrillation kinetics.

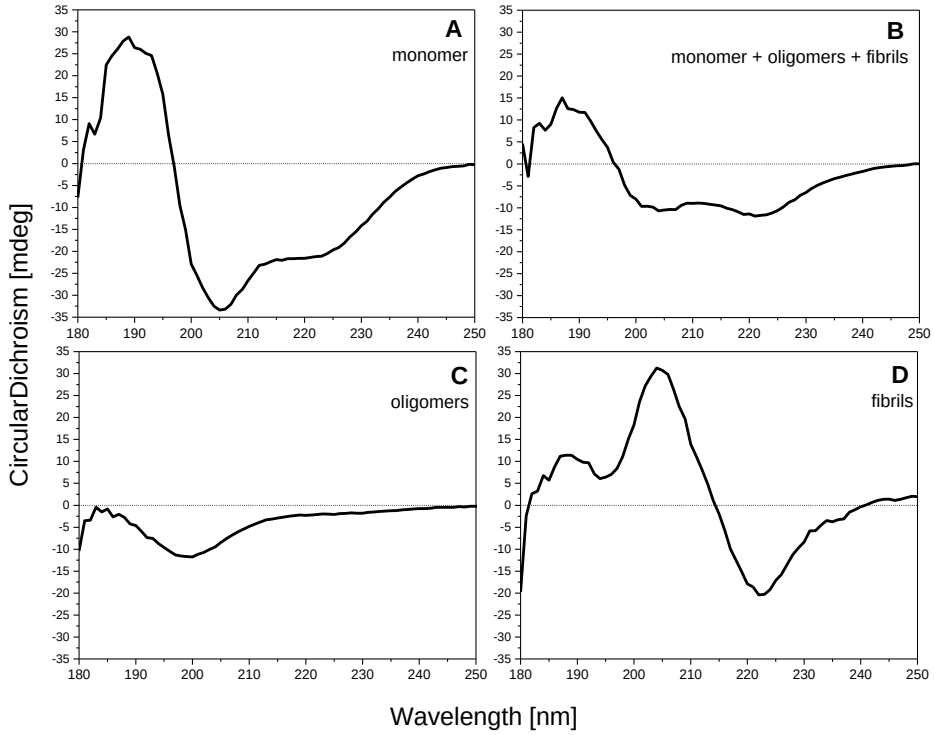


Figure 1.12: Far-UV CD spectra of different species present in the aggregating mixture of GLP-1-Am. All samples were measured in 25 mM phosphate at pH 8 in a 0.1 cm pathlength cuvette. (A) GLP-1-Am monomer at 85 μM concentration, a freshly prepared sample. (B) GLP- 1-Am (85 μM) after 9 days incubation at 37 °C with constant shaking at 80 rpm. (C) GLP-1-Am oligomers isolated from sample B using SEC. (D) Fibrils isolated (by centrifugation) from a sample prepared analogously to sample B. All spectra were recorded at 25 °C. Figure adapted from Brichtová (2019)[86].

1.9 The scope of this Thesis

This Thesis is focused on the investigation of the physical stability of both non-lipidated and lipidated variants of GLP-1. The research part of this Thesis is organized into four chapters (Chapter 3 to Chapter 6). In Chapter 3, the aggregation processes of GLP-1 and its C-terminally amidated variant, GLP-1-Am, at around neutral pH are investigated with emphasis on the effect of additional off-pathway aggregation processes, which compete with amyloid formation. In Chapter 4, the computational approach, Energy Landscape Theory, is employed in order to rationalize differences in fibrillation propensities of GLP-1 at acidic and neutral pH values. Chapters 5 and 6 compare the self-assembly and aggregation behaviour of four lipidated GLP-1 analogues, varying in the position of lipidation, and one additional analogue differing by the nature of the lipid moiety in order to establish the effect of both lipidation site and the lipid moiety on the physical stability of the peptide.

CHAPTER 2
MATERIALS AND METHODS

2.1 Peptides

Non-lipidated peptides

GLP-1(7-37), H-HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG-OH, molecular weight 3355 Da, was purchased from GenScript in the form of an acetate salt with 98.1 % purity. The residual content of the TFA salt was 0.94 % as determined by GenScript.

GLP-1-Am(7-37), H-HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG-NH_2, molecular weight 3355 Da. Two different batches were used in the experiments. Batch 1 was purchased from Bachem in the form of an acetate salt with 96.7 % purity. The residual content of the trifluoroacetic acid TFA counterion was not specified. Batch 2 was purchased from GenScript in the form of an acetate salt with 99.2 % purity. The residual content of the TFA salt was 0.02 % as determined by GenScript.

βAsp_GLP-1-Am, H-HAEGTFTS{Beta-Asp}VSSYLEGQAAKEFIAWLVKGRG-NH_2, molecular weight 3355 Da, was purchased from GenScript in the form of an acetate salt with 92.5 % purity.

IPP2_non-lipidated, H-HAEGTFTSDVSSYLEGQAAREFIAWRVRGRG-NH_2, molecular weight 3454 Da, was purchased from GenScript in the form of an acetate salt with 95.4 % purity. The residual content of the TFA salt was 0.42 % as determined by GenScript.

GLP-1-Am_Aib, H-H{2-Me-Ala}EGTFTSDVSSYLEGQAAKEFIAWLVKGRG-NH_2, molecular weight 3369 Da, was purchased from GenScript in the form of an acetate salt with 96.7 % purity. The residual content of the TFA salt was 0.74 % as determined by GenScript.

Lipidated peptides

IPP2, H-HK(γ-Glu-palmitoyl)EGTFTSDVSSYLEGQAAREFIAWRVRGRG-NH_2, was purchased from Bachem in the form of an acetate salt with 96 % purity. The molecular weight of IPP2 is 3878 Da.

IPP3, H-HAEGTFTSDVSK(γ-Glu-palmitoyl)YLEGQAAREFIAWLVRGRG-NH_2, was purchased from Bachem in the form of an acetate salt with 95.6 % purity. The molecular weight of IPP3 is 3819 Da.

IPP4, H-HAEGTFTSDVSSYLEGK(γ-Glu-palmitoyl)AAREFIAWLVRGRG-NH_2, was purchased from Bachem in the form of an acetate salt with 96.3% purity. The molecular weight of IPP4 is 3778 Da.

IPP5, a C-terminally amidated liraglutide analogue with 31 residues with the following sequence: H-HAEGTFTSDVSSYLEGQAAK(γ-Glu-palmitoyl)EFIAWLVRGRG-NH_2; was purchased from Peptides International in the form of an acetate salt with > 96% purity. The molecular weight of IPP5 is 3750 Da.

IPP10, a C-terminally amidated semaglutide analogue, with the following sequence: H-H(Aib)EGTFTSDVSSYLEGQAAK(PEG2-PEG2-γ-Glu-distearoyl)EFIAWLVRGRG-NH_2; was supplied by Peptides International in the form of an acetate salt with approx. 96 % purity. The molecular weight of IPP10 is 4113 Da.

All peptides and lipidated analogues were produced using solid-phase peptide synthesis and purified using HPLC. All peptides and lipidated peptides were stored in the form of lyophilized peptide powder at −20 °C and were used without further purification.

2.2 Buffers

All buffers were prepared using the respective acids and salts of analytical grade quality purchased from Sigma-Aldrich or Fisher Chemicals, using distilled deionized water (at a resistivity of 18.2 MΩ cm at 25 °C) and volumetric glassware, and filtered through a 0.22 µm pore size membrane (PES membranes, Millex). The pH of buffers was checked using a SevenCompact pH meter S210 (Mettler Toledo). Buffers were prepared gravimetrically without further pH adjustment.

The amount of acid and salt required for a buffer at a particular pH were calculated using the Henderson-Hasselbalch equation and the conservation of molar amount:

$$\mathrm{pH} = \mathrm{p}K_\mathrm{a} + \log \frac{[\mathrm{A}^-]}{[\mathrm{HA}]} \quad (2.1)$$

$$[\mathrm{A}^-] + [\mathrm{HA}] = c_\mathrm{buffer} \quad (2.2)$$

where [A $^-$] is the concentration of the basic buffer component, [HA] is the concentration of the acidic buffer component (dissociation in solution was neglected), and c_buffer is the total buffer concentration.

The deuterated versions of buffers for IR and VCD measurements were prepared analogously using deuterium oxide (D_2O) purchased from Sigma Aldrich with an isotopic purity of 99.9 atom % D.

Composition of the main buffers for aggregation assays and biophysical measurements:

Sodium phosphate buffers over the pH range from 6 to 8 were prepared at 25 mM concentration by dissolving the corresponding amount of $NaH_2PO_4 \cdot 2\ H_2O$ (MW: 156.01 g mol^{-1}) and Na_2HPO_4 (MW: 141.96 g mol^{-1}) in distilled deionized water. The amounts of acidic and basic component required was calculated for each pH value using Equations 2.1 and 2.2 and a pK_a value of 6.95.

Sodium phosphate buffer, 25 mM at pH 3 was prepared by dissolving 0.346 g of $NaH_2PO_4 \cdot 2\ H_2O$ (MW: 156.01 g mol $^{-1}$) and 19.12 µL of H_3PO_4 (MW: 197.99 g mol $^{-1}$, ≥99.0%, ρ = 1.685 g mL^{-1}) in 100 mL of distilled deionized water. The pK_a value used was 2.1.

Citrate buffer, 25 mM at pH 3 was prepared by dissolving 2.540 g of citric acid and 2.523 g of monosodium citrate in 1 L of distilled deionized water. The pK_a value used was 3.05.

Tris buffer, 25 mM at pH 8.5 at 37 °C was prepared by dissolving 0.631 g of Trizma® hydrochloride (MW: 157.6 g mol $^{-1}$) and 2.544 g of Trizma® base (MW: 121.14 g mol $^{-1}$) (both SigmaAldrich) in 1 L of distilled deionized water. The pK_a of Tris is 7.78 (37 °C).

2.3 Other chemicals

Thioflavin T (ThT) was purchased from Sigma-Aldrich. The dye content in the powder was ≥65%. A stock solution was prepared in distilled deionized water at 5 mM concentration as determined gravimetrically. The ThT stock was filtered through a 0.22 μm filter (PES membrane, Millex) and stored at -20 °C.

Deuterium oxide (D_2O) was purchased from Sigma Aldrich in an isotopic purity 99.9 atom % D.

Pyrene was purchased from Sigma-Aldrich at 98 % purity. The stock solution of pyrene was prepared in 75 % (v/v) methanol in water solution and the concentration of pyrene in the stock was gravimetrically determined to be 197.6 μM.

8-Anilinonaphthalene-1-sulfonic acid (ANS) was purchased from Sigma-Aldrich in ≥ 97 % purity. The stock solution was prepared in distilled deionized water at 9.84 mM dye concentration, filtered through a 0.22 μm filter (PES membrane, Millex), and stored at -20 ˚C. The concentration of the dye in the stock solution was determined spectrophotometrically using an extinction coefficient of at 350 nm, ε_{350} = 5000 cm^{-1} M^{-1}.

2.4 Peptide sample preparation and determination of peptide concentration

Peptide powder was dissolved in the appropriate buffer. After peptide dissolution, the sample was filtered through a 0.22 μm filter (PES membranes, Millex). The concentration of the peptide was determined spectrophotometrically on a Cary 60 UV-Vis spectrophotometer (Agilent Technologies) using the Beer-Lambert Law and a theoretical extinction coefficient of 6990 M^{-1} cm^{-1} at 280 nm. The extinction coefficients of peptides were calculated using the ExPASy Bioinformatics Resource Portal[226] using:

$$\varepsilon_{280} = 5500\, n_{\mathrm{Trp}} + 1490\, n_{\mathrm{Tyr}} + 125\, n_{\mathrm{S\text{-}S}} \qquad (2.3)$$

where ε_{280} represents the extinction coefficient at 280 nm, n_{Trp} is the number of tryptophan residues in the peptide sequence, n_{Tyr} is the number of tyrosine residues in the peptide sequence and $n_{\mathrm{S\text{-}S}}$ is the number of disulfide bridges in the peptide sequence. Absorbance measurements were performed by scanning the sample between 200 and 400 nm and subtracting the absorption spectrum of the buffer measured in the same cuvette over the same

wavelength range. If no significant light scattering was observed (i.e., the absorbance at 320 nm was close to zero), the peptide concentration was calculated directly using A_{280} without further correction:

$$c = \frac{A_{280}}{\varepsilon_{280} \cdot l} \tag{2.4}$$

When light scattering of the sample was observed as a significant absorbance at 320 nm (A_{320}), the concentration was calculated using the scattering-corrected expression:

$$c = \frac{A_{280} - 1.929\, A_{330}}{\varepsilon_{280} \cdot l} \tag{2.5}$$

In both equations (Eq. 2.4 and 2.5), c is the concentration in mol L^{-1} and l is the pathlength of the cuvette in cm, here $l = 1$.

Preparation of aggregated/fibrillar samples

Samples were incubated in 1.5 mL plastic microcentrifuge tubes (STARLAB) sealed and wrapped in aluminium foil to protect from sunlight. Incubation/aggregation was performed in an Incubator Shaker (Innova® 43) at 37 °C with 180 rpm agitation.

2.5 Determination of peptide solubility

The solubility of non-lipidated GLP-1-Am and its lipidated analogues was tested over a pH range from 2.5 to 8.5. The peptide solubility was tested in the individual buffers differing by 0.5 on the pH scale (i.e., pH 2.5, 3.0, 3.5, 4.0, …). 500 µL of buffer (25 mM phosphate, citrate or Tris) of a corresponding pH was added to *circa* 0.5 mg of lyophilized peptide powder, gently mixed and left for *circa* 5 minutes at room temperature before the solution was filtered through a 0.22 µm filter (Millex, PVDF Membrane). Subsequently, the concentration of the peptide was measured on a spectrophotometer (Cary 60 UV-Vis, Agilent Technologies) using the absorption at 280 nm.

2.6 Thioflavin T binding assay

Kinetics of peptide fibrillation was probed by Thioflavin T binding assays using a FLUOstar Omega microplate reader (BMG Labtech). Peptide samples at a given concentration were

incubated at 37 °C with 50 μM ThT. Peptide samples with ThT were pipetted into a 96-well half-area plate (Corning 3881) and sealed with tape (Costar Thermowell) to prevent samples from evaporating. The total volume of sample in a well was 100 μL (non-lipidated analogues) or 120 μL (lipidated analogues). Bottom reading of the plate was performed every 30 minutes with 5 minutes of shaking prior to each reading (orbital shaker mode at 600 rpm). ThT binding to fibrils and other species was monitored by recording the fluorescence emission at 482 nm after excitation filter at 448 nm. Fluorescence was measured at a gain of 500 with 8 flashes per well.

2.7 8-Anilinonaphthalene-1-sulfonic acid binding assay

The exposure of hydrophobic patches in species populated during peptide aggregation was probed using 8-anilino-1-naphthalenesulfonic acid (ANS) fluorescent dye. Samples were prepared in the wells of a 96-well half-area plate (Corning 3881) by mixing the peptide samples with ANS to a total volume of 120 μL, in which the final concentration of the ANS dye was 250 μM. To prevent evaporation of the samples, the plate was sealed with tape (Costar Thermowell).

The fluorescence measurements were performed using a FLUOstar Omega (BMG Labtech) plate reader, with an excitation filter at 355 nm and an emission filter at 482 nm, at a gain of 500 and 8 flashes per well. The plate was incubated at 37 ˚C and readings were taken through the bottom of the wells every 35 minutes, after 5 minutes of shaking at 600 rpm, over 6 days. All samples were measured in triplicate.

2.8 Calcein-release assay

The assay was performed with 50 μM of calcein-encapsulated 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS) vesicles of an average diameter of 125 nm (calcein concentration inside vesicles >70 mM). The assay buffer was 20 mM HEPES, 120 mM NaCl, 0.8 mM EDTA, pH 8.0. Low-molecular weight oligomers of GLP-1-Am were isolated using size-exclusion chromatography (SEC) from aged sample of 200 μM peptide incubated in 25 mM sodium phosphate at pH 8 for 7 days. Isolated oligomers were concentrated using Amicon Ultra Centrifugal Filter Unit with a 3 kDa MWCO. The fluorescence emission signal at

520 nm was recorded every 5 minutes upon excitation at 484 nm. Samples were incubated at 37 °C with agitation during the assay. Every sample was measured in triplicate. Fluorescence kinetic measurements were carried out using a FLUOstar Omega microplate reader (BMG Labtech). Samples of calcein-encapsulated vesicle with 4 % (w/w) of GLP-1-Am oligomers were pipetted into a 96-well half-area plate (Corning 3694) and sealed with transparent tape to prevent samples from evaporation. The total volume of each sample in a well was 100 μL. Fluorescence was measured at a gain of 800 with 8 flashes per well. At the end of the assay 5 % (v/v) Triton X100 to each well which causes instant disruption of vesicles. The fluorescence signal was then normalized, taking the fluorescence value after the addition of 5 % (v/v) Triton X100 (complete disruption of vesicles) as 1.

2.9 Pyrene fluorescence

The critical micelle/aggregation concentration and/or the presence of micelle-like oligomers in the peptide sample was probed based on changes in pyrene fluorescence, which increases in intensity when the pyrene is in hydrophobic environment, e.g., inside a micelle. Pyrene stock solution (197.6 μM) was added to the peptide sample such that the final concentration of pyrene in the sample was 1 μM.

Fluorescence spectra were recorded using a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies). Samples were measured in a 120 μL quartz cuvette (Hellma Analytics). The excitation wavelength was 339 nm and the emission spectra were collected between 360 nm and 460 nm in 1 nm steps. The emission and excitation band passes were 5 nm and the voltage on the photomultiplier tube was 600 V. Measurements were carried out at room temperature.

2.10 Intrinsic tryptophan fluorescence

Intrinsic tryptophan fluorescence spectra were measured on a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies). Spectra were obtained using an excitation wavelength of 280 nm and emission spectra were recorded between 300 and 400 nm with a step of 1 nm. Emission and excitation band passes of 10 nm, and a voltage on the photomultiplier tube of 550 V were used. Samples were measured in a 120 μL quartz cuvette (Hellma Analytics). Measurements were carried out at room temperature.

2.11 Circular dichroism

Circular dichroism spectra were measured on a Chirascan CD spectrometer (Applied Photophysics). Far-UV CD spectra over the range 180–250 nm were measured in a 1 mm pathlength cuvette and the measurement was performed with a 1 nm step size and with a 1 nm spectral bandwidth. Near-UV CD spectra over the range 250–350 nm were measured with a 0.5 nm step size and 1 nm spectral bandwidth in a 2 mm pathlength cuvette. The resulting spectrum was obtained as an average of three scans and the spectrum of the pure buffer was subtracted. All measurements were performed at room temperature. The CD machine units (ellipticity – signal expressed in mdeg) were converted to molar ellipticity $[\theta]_{\mathrm{molar}}$ using the following equation:

$$[\theta]_{\mathrm{molar}} = \frac{m^0}{10 \cdot l \cdot c} \tag{2.6}$$

where $[\theta]_{\mathrm{molar}}$ is the molar ellipticity (with units deg cm^2 dmol^{-1}), m^0 is the CD signal in mdeg (machine units), l is the cuvette pathlength in cm, and c is the sample concentration in mol L^{-1}.

Far-UV CD spectra were analysed using DichroWeb, Contin-LL method, dataset 3[180–182] and BeStSel method[183–187] to estimate the percentage of different secondary structure in the sample. In DichroWeb server the "closest matching solution with all proteins" fit was used.

In addition, α-helical content for soluble (i.e., non-aggregated) samples was estimated using mean residue ellipticity value at 222 nm (MRE_{222}), which was calculated as follows:

$$MRE_{222} = \frac{{m^0}_{222}}{10 \cdot l \cdot c \cdot n} \tag{2.7}$$

where ${m^0}_{222}$ is the CD signal in mdeg (machine units) at 222 nm, l is the cuvette pathlength in cm, c is the sample concentration in mol L^{-1}, and n is the number of amino acid residues. α-helical content was estimated using a method based on a liner interpolation between experimentally determined MRE_{222} values for purely α-helical and purely coiled protein[227–229]. α-helical content is then calculated as:

$$\%\ helicity = 100 \cdot \left(1 + \frac{(MRE_{222} - MRE_{helix})}{(MRE_{coil} - MRE_{222})}\right)^{-1} \tag{2.8}$$

where MRE_{222} is the observed ellipticity at 222 nm, MRE_{helix} is the value for purely α-helical structure (-35 791 deg cm^2 $dmol^{-1}$, at 25°C), and MRE_{coil} is the value for purely coiled structure (-725 deg cm^2 $dmol^{-1}$, at 25°C).

2.12 Infrared spectroscopy and vibrational circular dichroism

Infrared (IR) and vibrational circular dichroism (VCD) spectra were measured on a ChiralIR instrument (Biotools), using BaF_2 windows, 50 μm optical pathlength, 8 cm^{-1} resolution, and 10 blocks of 2048 scans. Samples of an aggregated GLP-1-Am and its lipidated variants were measured at a concentration of *circa* 4 mg mL^{-1} (*circa* 1.1–1.2 mM) in deuterated 25 mM phosphate buffer at pD 3, 7.5 or 8. All measurements were performed at room temperature. In each case, deuterated phosphate buffer spectra (of corresponding pD) were subtracted as a baseline. Both IR and VCD spectra were normalized to approximately equal the highest amide I absorption. For the samples with an enhanced VCD signal, the VCD/IR intensity ratio was calculated.

Infrared spectra of solid samples were recorded on a FT-IR Perkin-Elmer Spectrum One spectrometer, by loading the lyophilised peptide powder onto the diamond crystal of the apparatus. The spectra were collected using 20 scans, averaged and normalised such that the maximum absorbance represents 100 %. The amide I region (1600–1700 cm^{-1}) was analysed to obtain information about the secondary structure. The second derivative of amide I and II region were plotted to obtained more detailed comparison of the individual spectra. The second derivative spectra were smoothed using Savitzky-Golay filter (polynomial order: 2; points of window 20). The aged samples were flash frozen in liquid nitrogen, freeze dried overnight under vacuum in a CoolSafe LaboGene freeze dryer prior to the IR measurements.

2.13 Transmission electron microscopy

Samples were imaged using a Thermo Scientific Talos F200X G2 Transmission Electron Microscope with an acceleration voltage of 200 kV. 2 μL of the sample was loaded onto carbon coated 300 mesh copper grid (EMResolutions or Agar Scientific), which was glow discharged using a Quorum Technologies GloQube system prior to sample application. The

sample was dried by blotting, then negatively stained with 2 µL of 2 % (w/w) uranyl acetate solution for 15–30 seconds and dried again.

Size analysis of oligomeric species of lipidated GLP-1-Am analogues from TEM images

The diameters of IPP3 and IPP5 oligomers imaged at 5 µM and 25 µM concentration by TEM were measured using TEM Imaging & Analysis (TIA) software. For each peptide, 12 different TEM images were analysed (including both 5 and 25 µM concentrations) with 6 data points (oligomer diameters) taken from each image. In total, 72 data points were analysed for each oligomeric distribution and the data were plotted as histograms.

An average ($\bar{x}$) with a standard error of the mean, $\sigma(\bar{x})$, and a standard deviation, $\sigma(x_i)$ for each distribution were calculated as follows:

$$\bar{x} = \frac{\sum_i x_i}{N} \tag{2.9}$$

$$\sigma_{(\bar{x})} = \frac{\sigma_{(x_i)}}{\sqrt{N}} \tag{2.10}$$

$$\sigma_{(x_i)} = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{N - 1}} \tag{2.11}$$

where x_i is the oligomer diameter and N is number of data points.

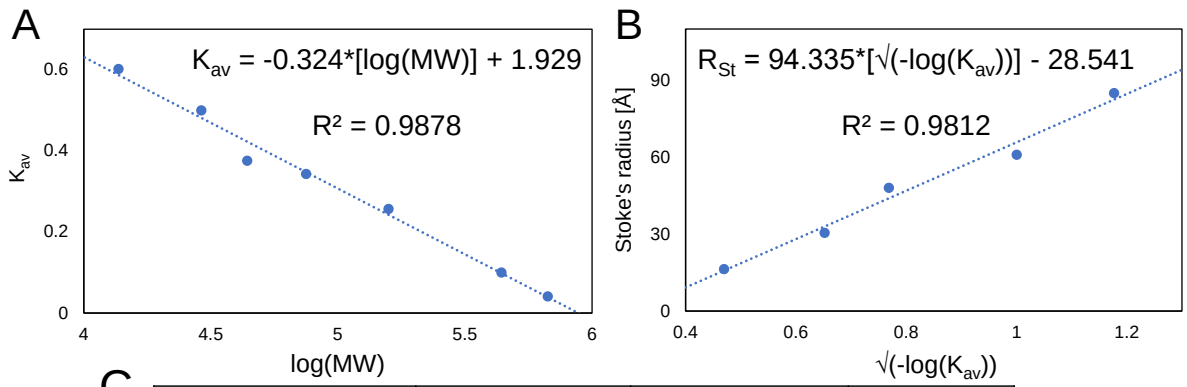
2.14 Size-exclusion chromatography

Analytical size-exclusion chromatography was performed on an ÄKTA FPLC system (GE Healthcare), using a Superdex 200 Increase 10/300 column, Superdex75 10/300 and Superose12 10/300 (all GE Healthcare). Samples were loaded using 200 or 500 µL loops. Prior to loading, the samples were filtered through a 0.22 µm filter (Millex, PVDF Membrane). All samples were eluted at a flow rate of 0.75 mL min^{-1} at room temperature and UV absorbance detection at 280 nm through a 0.5 cm flow cell was used.

For each column and set of conditions, globular protein standards (GE Healthcare) were used for column calibration, Figure 2.1 and Figure 2.2. The gel phase distribution parameter, K_{av}, of each protein standard was calculated as

$$K_{av} = \frac{V_e - V_0}{V_c - V_0} \tag{2.12}$$

where V_e is an elution volume of a protein standard, V_0 is the void volume of the column (determined by elution volume of Blue Dextran 2000) and V_c is the total volume of the column, 24 mL for all columns used. The logarithm of the molecular weight of each protein standard was plotted against its K_{av}. A linear regression of this plot was used to determine the molecular weight of any eluting species observed in the peptide experiments. For Stokes radius (R_{St}), a calibration plot of $\sqrt{-\log(K_{av})}$ against R_{St} and a linear regression was used.



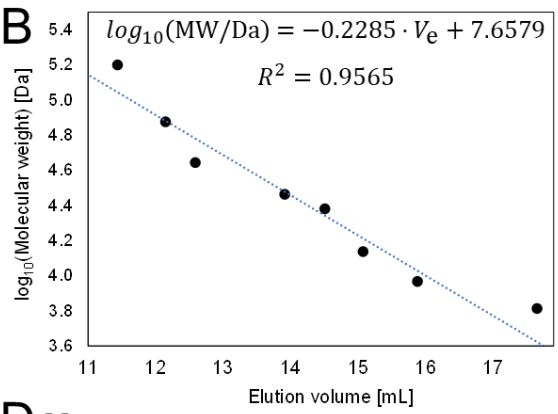
C

Standard	M_W [Da]	Stoke's radius [Å]	$V_{elution}$ [ml]
Blue Dextran 2000	N/A (void volume)	N/A	8.1
Aldolase	158 000	48.1	12.2
Ovalbumin	44 000	30.5	14.1
Ferritin	440 000	61	9.7
Conalbumin	75 000		13.6
Thyroglobulin	669 000	85	8.8
Ribonuclease A	13 700	16.4	17.7
Carbonic Anhydrase	29 000		16.0

Figure 2.1: Calibration of the Superdex 200 Increase 10/300 column. Calibrations on the molecular weight (A) and on the Stokes radius (B) were performed in 25 mM sodium phosphate buffer at pH 7.5 using a set of protein standards (C) at room temperature. Elution was performed with a flow rate 0.75 mL min^{-1} and 100 µL of approximately 100 µM of each protein was injected using a 200 µL injection loop. The blue dotted line is the best fit of the data to a linear regression and the parameters obtained are shown in the equations displayed in the plots (A, B).

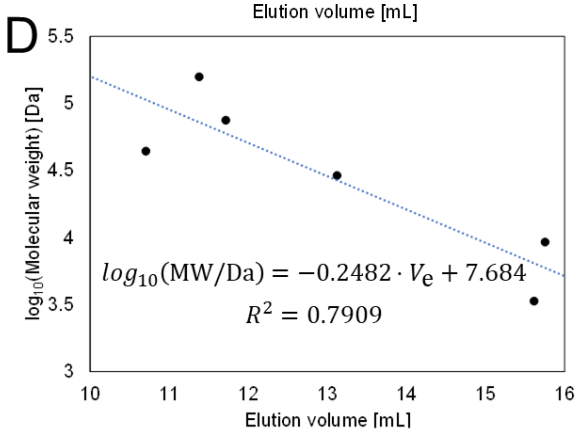
A

Protein standard	MW [Da]
aldolase	158,000
conalbumin	75,000
ovalbumin	44,000
carbonic anhydrase	29,000
trypsinogen	24,000
ribonuclease A	13,700
chymotrypsin inhibitor 2	9,265
GLP-1-Am monomer	3,355



C

Protein standard	MW [Da]
aldolase	158,000
conalbumin	75,000
ovalbumin	44,000
carbonic anhydrase	29,000
chymotrypsin inhibitor 2	9,265
GLP-1 monomer	3,355



E

Protein standard	MW [Da]
ovalbumin	44,000
carbonic anhydrase	29,000
trypsinogen	24,000
ribonuclease A	13,700
chymotrypsin inhibitor 2	9,265
GLP-1-Am monomer	3,355

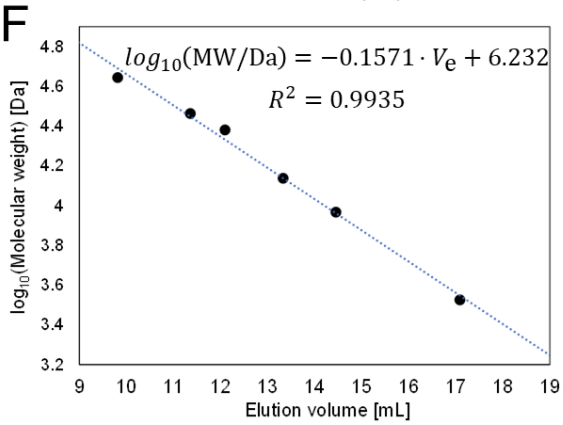


Figure 2.2: Calibrations of the Superose 12 10/300 and the Superdex 75 10/300 columns. Calibration and set of protein standards for calibration of the Superose 12 10/300 in 25 mM sodium phosphate buffer at pH 8 (A, B). Calibration and set of protein standards for calibration of the Superose 12 10/300 in 10 mM sodium phosphate buffer at pH 7 (C, D). Calibration and set of protein standards for calibration of the Superdex 75 10/300 in 25 mM sodium phosphate buffer at pH 8 (E, F). Elution was performed with a flow rate 0.75 mL min^{-1} and 100 µL of approximately 100 µM of each protein was injected using a 200 µL injection loop. The protein standards elution volumes were plotted against the logarithm of their molecular weight and the correlation is described by a linear regression fit (blue dotted line) and its square of the correlation coefficient, where V_e is the protein elution volume measured from the centre of a peak. Calibration data in panels B and F are taken from Brichtová (2019)[86].

2.15 Analytical centrifugation – sedimentation velocity

Sedimentation velocity experiments were performed using a Beckman Optima XL-I Analytical Ultracentrifuge equipped with An-60 Ti rotor. Samples of 85 μM peptide concentration were either freshly prepared before the measurement or incubated for 7 days at 37 °C with 180 rpm agitation. After a 2-hour temperature equilibration of samples in the centrifuge to 20 °C, the experiment was performed with centrifugation at 50 000 rpm. The interference sedimentation curves were collected as 300 scans (approx. 24 h run time) for non-lipidated analogues or as 150 scans (approx. 12 h run time) for lipidated analogues and fitted to a continuous c(*s*) distribution model implemented in the Sedfit program[220–222]. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$). The molecular weight and relative amount of each detected species was calculated using the Sedfit program.

2.16 Liquid chromatography-mass spectrometry

Liquid chromatography-mass spectrometry (LC-MS) analysis was performed on a Waters' Xevo G2-S bench top QTOF using electrospray ionization in the positive mode. The peptides at around 10 μM concentration in 25 mM Tris buffer at pH 8.5 were run according to the LC method indicated in Table 2.1 with total ion detection.

LC method			
Flow rate	0.2 mL min^{-1}		
Run time	11.55 min		
Sample injection	6 µL		
Solvent system			
A	H_2O + 0.1% formic acid		
B	75 % acetonitrile, 25 % H_2O, 0.075 % formic acid		
Gradient			
Time [min]		A [%]	B [%]
0.00		95	5
2.00		80	20
9.00		50	50
9.21		0	100
11.55		95	5

Table 2.1: LC method parameters.

2.18 Isoelectric focusing electrophoresis

The isoelectric point (pI) of aged and fresh samples of GLP-1-Am, βAsp_GLP-1-Am and GLP-1 was determined by isoelectric focusing (IEF) gel Novex™ pH 3-10 run in buffers from Novex™ pH 3–10 IEF Buffer Kit at a constant voltage of 200 V for 1 hour. Peptide samples were prepared in a sample buffer (Novex™ pH 3-10 IEF Buffer Kit) and immediately loaded onto the gel. A protein marker for isoelectric focusing SERVA Liquid mix 3–10 was run in parallel for calibration. Gels were stained with InstantBlue (Expedeon).

2.19 Computational methods – Energy landscapes

Energy landscape calculations were performed using GMIN[230], OPTIM[231] and PATHSAMPLE[232] programs together with the AMBER program suite[233–236]. All calculations were performed using the AMBER ff14SB force field[237] that was symmetrized in the same way[238,239] as the ff99SB force field[240]. An implicit generalized Born solvation model (igb = 2)[241] using infinite interaction cutoff and the Debye−Hückel approximation for salt (0.1 M)[242] was employed. Starting structures were either generated from sequence by antechamber or obtained from the Protein Databank (ID: 5OTU)[160,161] and modified to correspond to the GLP-1(7-37) sequence. Basin-hopping global optimization (60 000 steps) in a GMIN program[230] was used on starting structures to obtain low energy conformers. 200 lowest energy minima from each basin hopping search were used to initiate exploration of the energy landscape.

Discrete path sampling[243,244] (DPS) was employed to create kinetic transition networks[245,246] of local minima and transition states that connect them. DPS was performed using OPTIM[231] and PATHSAMPLE[232] programs. The doubly-nudged elastic band algorithm[247–249] was used to generate transition state candidates, and hybrid eigenvector-following algorithm[250] to converge them. Approximate steepest-descent paths were used to obtain the minima directly connected to these transition states. A customised L-BFGS approach[251,252] with an RMS force convergence criterion of 10^{-6} kcal mol^{-1} was employed. The databases were refined by finding alternative pathways (SHORCUT scheme[253,254]), removing kinetic traps (UNTRAP scheme[254]) and local sampling (CONNECTUNC scheme[255]). A more detailed description of the sampling methods and of the building of the kinetic transitions network is available in Chapter 4.

Free-energy disconnectivity graphs were constructed for T = 310 K. Secondary structure content and the radius of gyration were calculated using the cpptraj program[256] using DSSP algorithm[257]. Kinetic transition networks were visualized in the form of disconnectivity trees[258] and peptide geometries were visualized using the PyMOL molecular visualization system[259].

CHAPTER 3
AGGREGATION OF NON-LIPIDATED GLP-1 ANALOGUES: FORMATION OF STABLE LOW-MOLECULAR WEIGHT OLIGOMERS OF GLP-1 AND GLP-1-AM

Some of the results presented in this Chapter are published as:

Přáda Brichtová E., Krupová M., Bouř P., Lindo V., Gomes dos Santos A., Jackson S. E. Glucagon-like peptide 1 aggregates into low-molecular-weight oligomers off-pathway to fibrillation, *Biophysical Journal* (2023), DOI: 10.1016/j.bpj.2023.04.027

All the experiments presented in this Chapter were carried out by myself unless otherwise stated.

3.1 Introduction

GLP-1 is of great interest to the pharmaceutical industry as a therapeutic peptide for the treatment of type 2 diabetes. Several GLP-1 analogues are already commercially available[16,23]. However, previous studies have reported the self-association of GLP-1 into amyloid fibrils over a wide range of conditions[36,260,261], which hinders its pharmaceutical use.

The intrinsic physical instability of GLP-1 remains a challenge in drug processing and formulation. This peptide hormone is also one of the few aggregation-prone peptides/proteins that exhibits unusual fibrillation kinetics where, under specific conditions, the lag time increases with increasing peptide concentration[36]. This behaviour has been also reported for human calcitonin[80], ribosomal protein S6[262], immunoglobulin light chain[263] and liraglutide[81]. The unusual dependence of lag time on the peptide concentration has been explained by various off-pathway, self-assembly processes that affect the rates of fibrillation, although details of this off-pathway reaction remain unknown. Despite the increasing evidence for off-pathway oligomers/aggregates, there is still a need to elucidate their structural properties, stabilities and the mechanisms by which such off-pathway species form and affect the fibrillation kinetics.

3.2 Aims

The main objectives of this Chapter were to isolate and characterize the putative off-pathway oligomers formed by GLP-1 and GLP-1-Am, to understand their structures, why they form and how they affect amyloid formation. Size-exclusion chromatography (SEC) and sedimentation velocity experiments were employed to detect and isolate (SEC only) stable, low-molecular weight oligomers of GLP-1 and GLP-1-Am. A variety of spectroscopic methods were used to probe and monitor aggregation, as well as characterize the oligomeric species observed, including ThT assays, circular dichroism, FT-IR and vibrational circular dichroism.

3.3 Results

3.3.1 GLP-1 and its C-terminally amidated analogue show pH-dependent fibrillation propensity

Using ThT assays, the fibrillation behaviour of GLP-1 and GLP-1-Am was investigated at acidic and neutral/slightly basic pH conditions, Figure 3.1A and B. GLP-1 was incubated with ThT at pH 3.5, 7 and 8 at 37 °C with agitation over 6 days (Figure 3.1A), whilst GLP-1-Am was incubated with ThT at pH 3, 7 and 8 at 37 °C with agitation over 6 days (Figure 3.1B). The fibrillation of both GLP-1 and GLP-1-Am at acidic pH show sigmoidal profiles with distinct lag, elongation and growth, and plateau phases. In contrast, GLP-1 at pH 7 and 8 (in 25 mM sodium phosphate) did not shown any increase in ThT fluorescence over the timescale of the assay. Interestingly, over the timescales studied, GLP-1-Am in 25 mM sodium phosphate at pH 7 and 8 shows a linear-like increase in ThT fluorescence, lacking the distinct lag or elongation and growth phases, observed at acidic pH. The fibrillation behaviour of GLP-1-Am at pH 3 and 8 was also systematically studied over a range of peptide concentrations, Figure 3.1C and D. At pH 3, the kinetics of fibrillation of GLP-1-Am are consistent with a nucleation-elongation mechanism, in which increasing the peptide concentration leads to a decrease in the lag time. In contrast, although the ThT curves at pH 8 could not be fitted to a standard sigmoidal function and the fact that the rate of fibrillation is somewhat variable under these conditions, it does not appear to vary significantly with respect to peptide concentration.

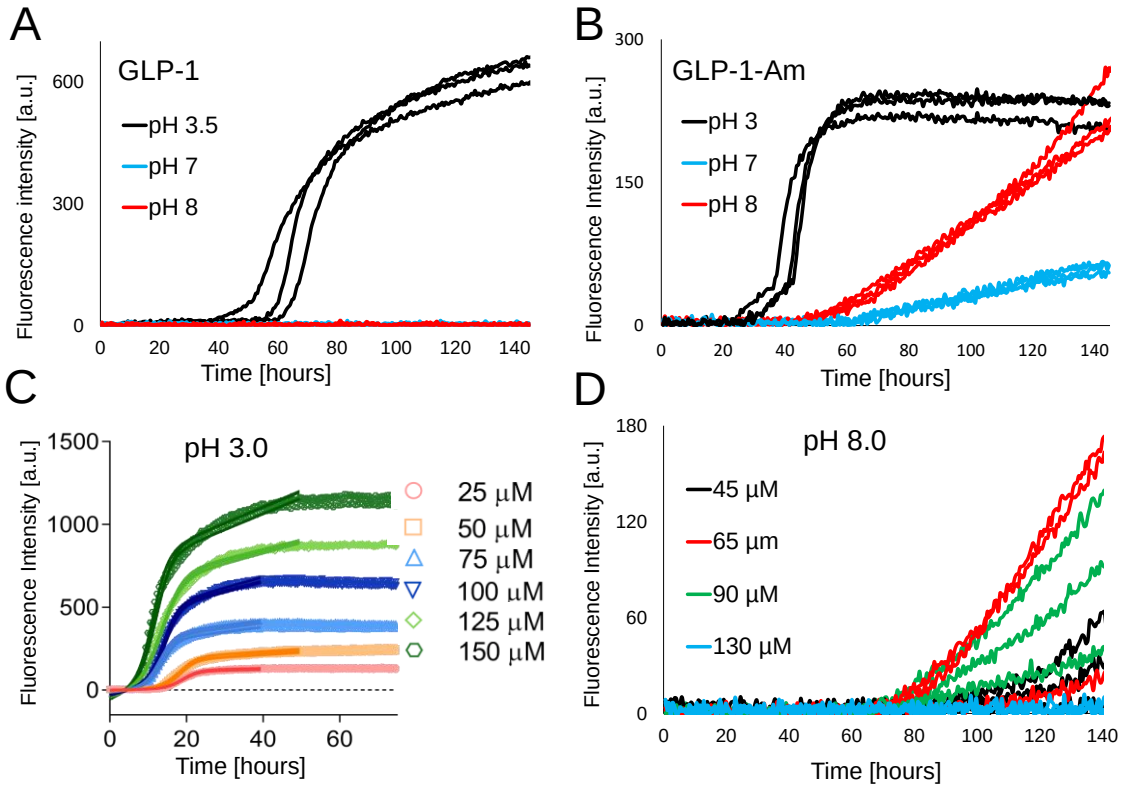


Figure 3.1: pH-dependent fibrillation of GLP-1 and GLP-1-Am monitored by thioflavin T assay. The fibrillation of GLP-1 and GLP-1-Am was monitored using a ThT assay. All conditions were performed in triplicate. (A) GLP-1 at 43 µM concentration was incubated in: pH 3.5 (25 mM citrate, shown in black), pH 7 (10 mM sodium phosphate, shown in blue), and pH 8 (10 mM sodium phosphate, shown in red), at 37 °C with agitation over 6 days. (B) GLP-1-Am at 85 µM concentration was incubated in: pH 3 (25 mM citrate, shown in black), pH 7 (25 mM sodium phosphate, shown in blue), and pH 8 (25 mM sodium phosphate, shown in red), at 37 °C with agitation over 6 days. (C) Fibrillation of GLP-1-Am in 25 mM citrate at pH 3 for different peptide concentrations from 25 to 150 µM was monitored over 80 hours at 37 °C. The concentration-dependence data of GLP-1-Am at pH 3 were collected by Dr. Frederik Becher. (D) Fibrillation of GLP-1-Am in 25 mM sodium phosphate at pH 8 for different peptide concentrations from 45 to 130 µM was monitored over 6 days in 37 °C.

After 6-days of incubation at 37 °C, samples of GLP-1 and GLP-1-Am at pH 3 and 7 (for GLP-1) or 8 (for GLP-1-Am) were imaged using transmission electron microscopy (TEM), Figure 3.2. These conditions correspond to the end point of the ThT assay (Figure 3.1). At pH 3, a dense network of fibrils was formed for both GLP-1 and GLP-1-Am. These fibrils (Figure 3.2A and C) are several micrometers long and have a diameter of about 15 to 18 nm. Fibrils of GLP-1 and GLP-1-Am were also observed at pH 7 or 8 (Figure 3.2B and D) but, under these conditions, they did not form a dense network but rather isolated clusters containing several fibrils. They are shorter compared to those formed at pH 3 and it is difficult to distinguish the diameter of individual fibrils due to their clustering.

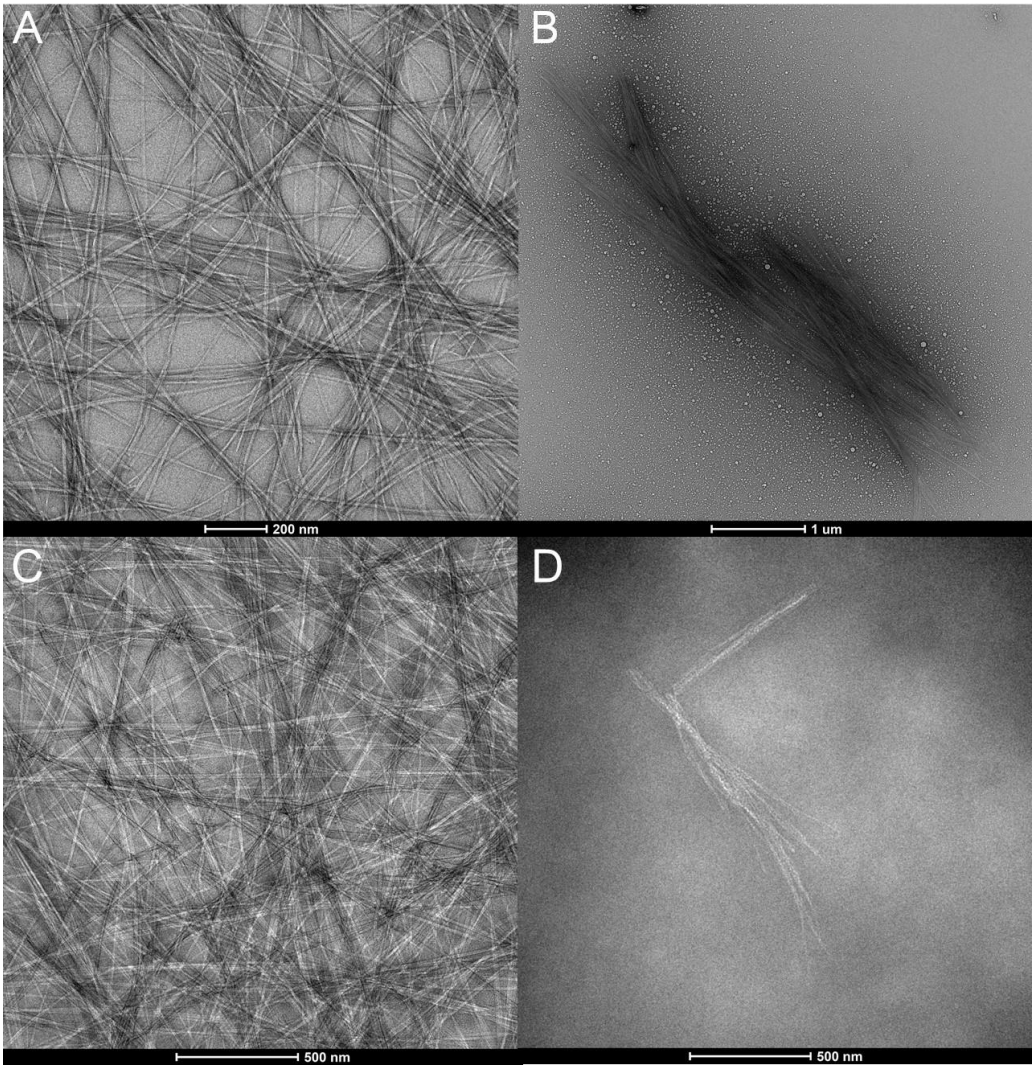


Figure 3.2: TEM images of GLP-1 and GLP-1-Am fibrils formed at acidic and around neutral pH. GLP-1 samples at 85 µM concentration were incubated either in 25 mM citrate at pH 3 (A) or 25 mM phosphate at pH 7 (B) at 37 °C with agitation for 6 days prior to the imaging. GLP-1-Am samples at 85 µM concentration were incubated either in 25 mM citrate at pH 3 (C) or 25 mM phosphate at pH 8 (D) at 37 °C with agitation for 6 days prior to the imaging. Aged samples were applied onto a carbon coated copper grid and negatively stained using 2 % uranyl acetate.

Fourier-transform infrared absorption (FT-IR) and vibrational circular dichroism (VCD) spectroscopy were also used to characterize the GLP-1-Am sample after aggregation. VCD is nowadays increasingly employed in aggregation studies due to the signal enhancement and splitting often observed for amyloid fibrils. It mainly occurs in the amide I region and the enhancement is induced by a highly regular cross-β sheet structure[200–204], more details on use of technique in amyloid research is given in Chapter 1, Section 1.6.2.2. Both FT-IR and VCD spectra were simultaneously recorded for each sample in the solution of a deuterated buffer of a corresponding pD. Figure 3.3 shows the spectra of samples of GLP-1-Am aggregated/fibrillated under acidic and basic conditions after 8-days of incubation in 25 mM deuterated sodium phosphate buffer at pD 3 and 8. In the FT-IR spectrum at pD 8,

the broad amide I band with a maximum at 1643 cm^{-1} corresponds to a disordered structure. In contrast, an enhanced VCD signal was observed ($\Delta A/A \approx 10^{-3}$, at approx. 1647 cm^{-1}), indicating the formation of amyloid fibrils. These, seemingly contradictory results, can be explained by the presence of multiple species in the sample. Contributions from all the species are visible in the resulting IR spectrum, while the VCD spectrum is dominated by the enhanced VCD signal of the amyloid fibrils – the VCD signal from non-fibrillar species is expected to be about two orders of magnitude lower. Nevertheless, the species with the disordered structure are likely to be in excess of the fibrils under these conditions and therefore dominate the FT-IR spectrum. A more characteristic "amyloid-like fibril" FT-IR (amide I signal split – 1612/1681 cm^{-1}) and VCD spectrum (even higher intensity between approx. 1590 and 1687 cm^{-1}) was measured at pD 3. Thus, a low pD favors a higher content of β-structured amyloid fibrils compared to pD 8, consistent with what was observed in the ThT assays. Due to the availability of the instrument, IR and VCD experiment was performed only with GLP-1-Am.

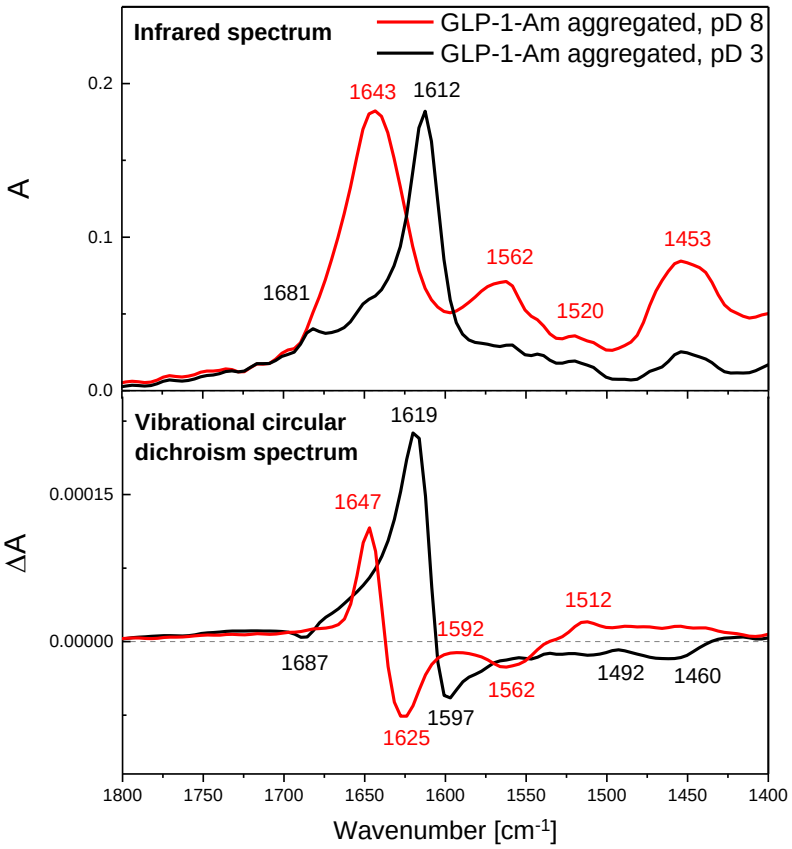


Figure 3.3: FT-IR and VCD spectra of aggregated GLP 1 Am at pD 3 and pD 8. Samples of 1.2 mM GLP-1-Am in 25 mM deuterated sodium phosphate buffer at pD 3 and pD 8 were incubated for 8 days at 37 °C with 180 rpm agitation. The FT-IR and VCD spectra have been normalized to the amide I absorption.

3.3.2 GLP-1 and GLP-1-Am form stable low-molecular weight oligomers

The aggregation of GLP-1 and GLP-1-Am was also studied using size-exclusion chromatography at pH 7 and 8. These studies were not extended to pH 3 because it was not possible to calibrate the SEC column under these conditions. GLP-1 and GLP-1-Am samples at concentrations of 43, 85 or 150 μM were incubated in an aqueous buffer at pH 7.0 or 8.0 at 37 °C with agitation. At various time points, aliquots were taken from the aggregating reaction, filtered using a 0.22 μm pore size membrane filter and applied onto a Superose 12 10/300 or a Superdex 75 10/300 size-exclusion column to determine the distribution of monomers and oligomers at each time point. The filtration step removed any amyloid fibrils formed, leaving only soluble monomer and oligomers in solution. The first datapoint recorded within an hour after the preparation of a fresh sample is referred in text as "time zero". For both GLP-1 and GLP-1-Am, oligomers with a larger hydrodynamic radius than the monomer were observed to form slowly during the incubation period and did not deplete even when there was substantial aggregation into amyloid fibrils, Figure 3.4. For both peptides, the formation of the oligomeric species was accompanied by monomer depletion. However, it should be noted that, under some conditions, the monomer depletion is also caused by the fibrillation of the peptide and not just oligomer formation. The intensity of the peaks corresponding to oligomers increased during the incubation period and did not decrease even at time points where significant amyloid fibril formation had occurred, suggesting that these oligomers are not capable of further aggregation/fibrillation or dissociation back into the monomeric state. After 10 days of incubation with shaking at 37 °C for GLP-1 at pH 7 (Figure 3.4A) and 16 days of incubation and shaking at 37 °C for GLP-1-Am at pH 8 (Figure 3.4C), the low-molecular weight oligomers represented around 40 % of the initial sample concentration, Table 3.1. A detailed quantitative analysis of the relative amounts of oligomers formed during the incubation is given in Table 3.1. The increase in the relative percentage of GLP-1 and GLP-1-Am oligomers over time are shown in Figure 3.5. An exponential function was used to fit the increase in low-molecular weight oligomers of GLP-1 with time (Figure 3.5A), however, only a single datapoint suggests a non-linear evolution of the oligomers and more datapoints are needed to confirm the non-linear behaviour and to obtain a more reliable fit. The increase in GLP-1-Am oligomers observed was gradual over the timescale studied and a linear regression fit was used for the datapoints, Figure 3.5B.

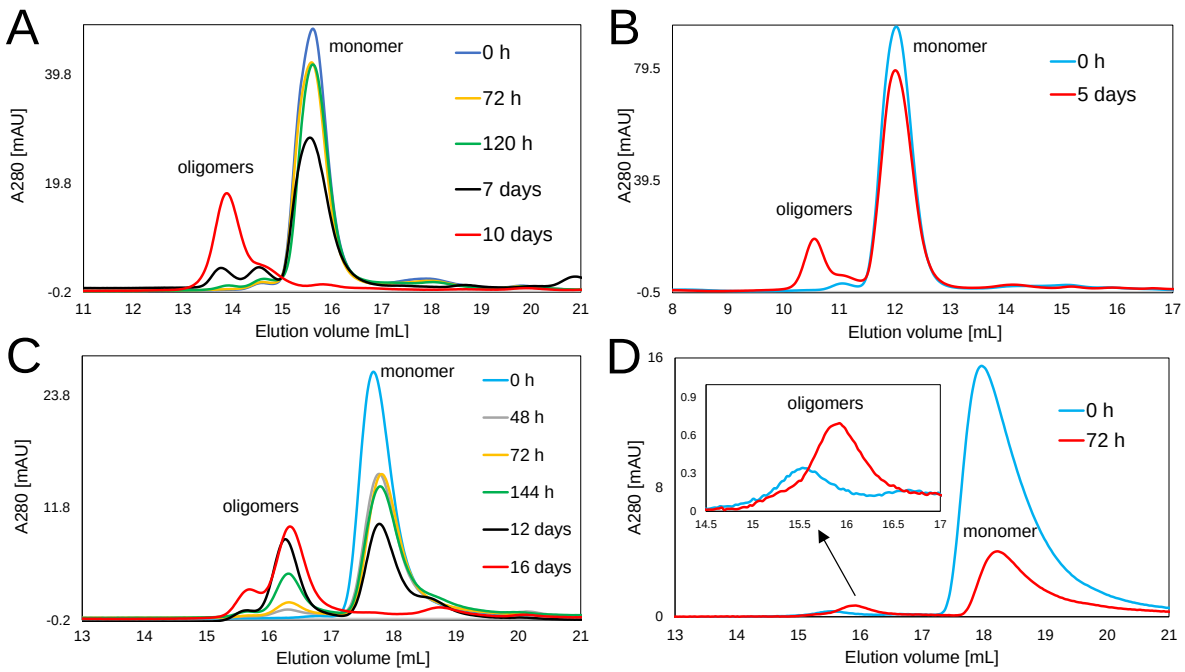


Figure 3.4: Formation of low-molecular weight oligomers of GLP-1 and GLP-1-Am observed using size-exclusion chromatography. (A) Size-exclusion chromatography of a sample of 85 µM GLP-1 incubated in 10 mM sodium phosphate buffer at pH 7 at 37 °C with 180 rpm agitation over 10 days. (B) Size-exclusion chromatography of a sample of 150 µM GLP-1 in 10 mM Tris buffer at pH 8 at time zero and after a 5-day incubation at 37 °C with 180 rpm agitation. (C) Size-exclusion chromatography of a sample of 85 µM GLP-1-Am incubated in 25 mM sodium phosphate buffer at pH 8 at 37 °C with 180 rpm agitation over 16 days. The data in panel C has already been published in Brichtová, 2019[86]. (D) Size-exclusion chromatography of a sample of 43 µM GLP-1-Am in 25 mM sodium phosphate buffer, pH 7, at time zero and after 72-hour incubation at 37 °C with 180 rpm agitation. The samples were analysed at different time points using a Superose 12 10/300 (in A, C and D) or a Superdex 75 10/300 (in B) size-exclusion columns, where the elution buffer was the same as the incubation buffer for each sample. The difference in elution volumes of monomeric GLP-1 and GLP-1-Am is most likely due to their slightly different conformations and differences in the ionic strength and pH of elution buffers used in the experiments.

GLP-1 at pH 7, Figure 3.4A			
Time	Content [%]		
[days]	monomer	oligomers	other
0	99	1	0
3	85	1	14
5	82	3	15
7	59	7	34
10	2	41	57

GLP-1 at pH 8, Figure 3.4B			
Time	Content [%]		
[days]	monomer	oligomers	other
0	98	2	0
5	81	15	4

GLP-1-Am at pH 8, Figure 3.4C			
Time	Content [%]		
[days]	monomer	oligomers	other
0	100	0	0
2	61	2	37
3	59	5	36
6	56	14	30
12	39	25	36
16	0	38	62

GLP-1-Am at pH 7, Figure 3.4D			
Time	Content [%]		
[days]	monomer	oligomers	other
0	100	0	0
3	20	2	78

Table 3.1: A quantitative estimate of relative amounts of GLP-1 and GLP-1-Am monomer and oligomers present at different time points determined using SEC. Values in this Table are obtained from the analysis of chromatograms shown in Figure 3.4. A peak integration was performed in the Unicorn software of the ÄKTA instrument. Content of "other" species, i.e., different from monomers and low-molecular weight oligomer, was determined from the difference of the total peak(s) area at 0 days and the total peak(s) area at a given time point. This determination assumes that at 0-day time point 100 % of peptide total content is injected into the column, i.e., larger aggregates or fibrils are not yet formed, at later time points "other" (i.e., larger aggregated species) are formed and those are removed by filtration before the injection onto SEC column. For each peptide/condition, aliquots from the same sample are taken at different time points.

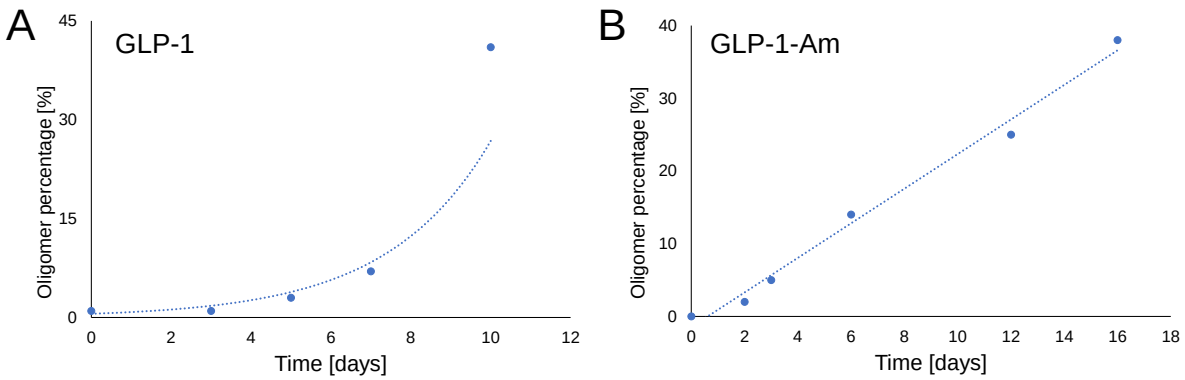


Figure 3.5: Increase of the relative amounts of GLP-1 and GLP-1-Am oligomers as monitored by SEC. The plots are based on the relative percentage of GLP-1 and GLP-1-Am oligomers at different time points given in **Table 3.1**. GLP-1 at 85 µM peptide concentration was incubated at 37 °C with agitation in 10 mM phosphate at pH 7 over 10 days (A). GLP-1-Am at 85 µM concentration was incubated at 37 °C with agitation in 25 mM phosphate at pH 7 over 16 days (B). The data in (A) are fitted to an exponential function whilst in (B) the data are fitted to a linear function (in both cases the best fit is shown by the blue dotted line.

The size of the observed oligomeric species was estimated using calibration curves of the two SEC columns, Superose 12 10/300 and Superdex 75 10/300 (details in Chapter 2). Based on the calibration curves, the size of the oligomers for GLP-1-Am was determined to be in the range from 6.7 to 13 kDa, approximately dimers to tetramers. For GLP-1, the size of oligomers was estimated to be in the range of trimers to pentamers. Oligomerization of GLP-1-Am was reproducible for peptide batches supplied by different manufacturers, Figure 3.6. For GLP-1, only one batch was used.

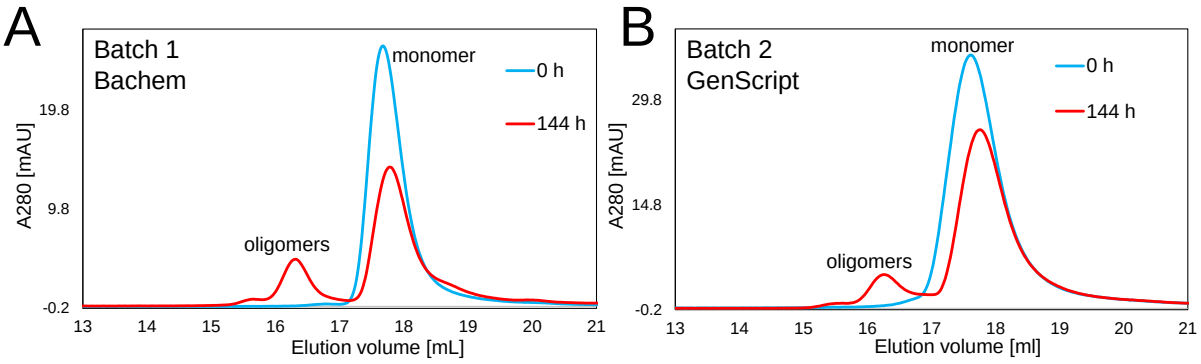


Figure 3.6: Batch-to-batch reproducibility of GLP-1-Am oligomerization. 85 µM GLP-1-Am from Bachem (A) or GenScript (B) was incubated in 25 mM sodium phosphate buffer, pH 8 at 37 °C with agitation. Size-exclusion chromatograms of samples of both peptide batches were analysed on a Superose 12 10/300 column directly after preparation and after 144 h of incubation.

The formation of these low-molecular weight oligomers of GLP-1 and GLP-1-Am was observed in multiple buffers in the pH range 7–8.5, including sodium phosphate, Tris or ammonium bicarbonate buffer, Figure 3.7 and Figure 3.8. Both GLP-1 and GLP-1-Am show a higher tendency to form these oligomers in lower ionic strength buffers, i.e., 10 mM sodium phosphate buffer, or 10 and 25 mM Tris buffer, than in buffers with higher ionic strengths (>100 mM) – Figure 3.7 and Figure 3.8.

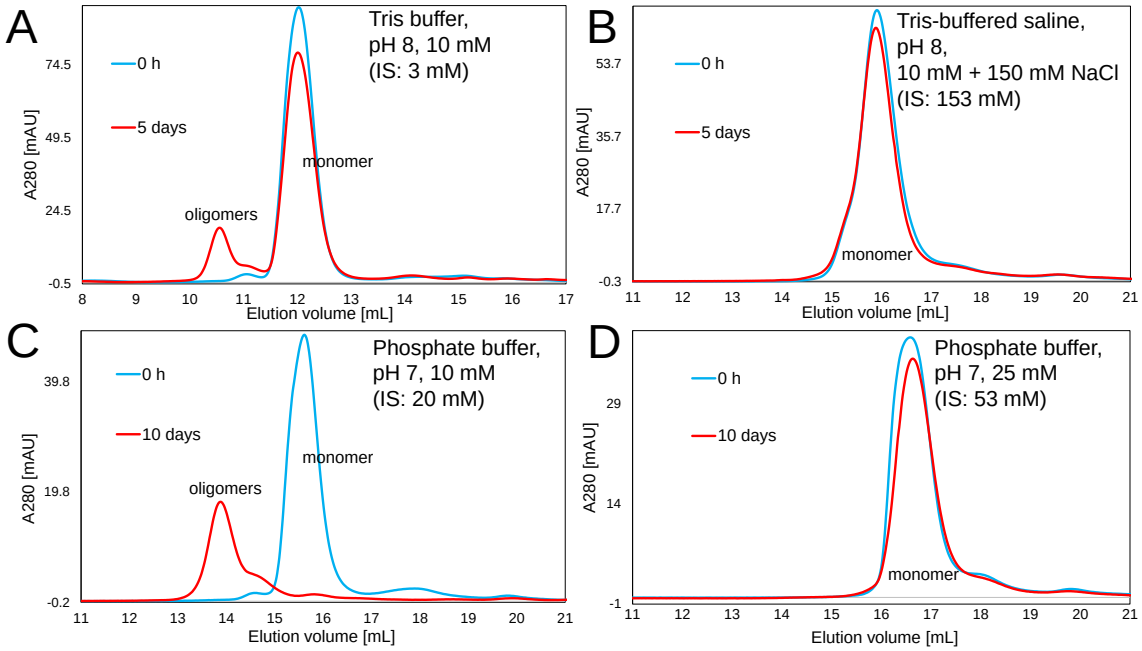


Figure 3.7: Formation of low-molecular weight oligomers of GLP-1 in low and high ionic strength (IS) buffers monitored using SEC. (A) 150 µM GLP-1 in 10 mM Tris buffer at pH 8 (IS = 3 mM) at time zero and after a 5-day incubation (at 37 °C with shaking) were analysed using a Superdex 75 10/300 size-exclusion column. (B) 150 µM GLP-1 in Tris-buffered saline (10 mM Tris +150 mM NaCl) at pH 8 (IS = 153 mM) at time zero and after a 5-day incubation (at 37 °C with shaking) were analysed using a Superdex 75 10/300 size-exclusion column. (C) 85 µM GLP-1 in 10 mM sodium phosphate buffer at pH 7 (IS = 20 mM) at time zero and after 10-day incubation (at 37 °C with shaking) were analysed using a Superose 12 10/300 size-exclusion column. (D) 85 µM GLP-1 in 25 mM sodium phosphate buffer at pH 7 (53 mM ionic strength) at time zero and after a 10-day incubation (at 37 °C with shaking) were analysed using a Superose 12 10/300 size-exclusion column.

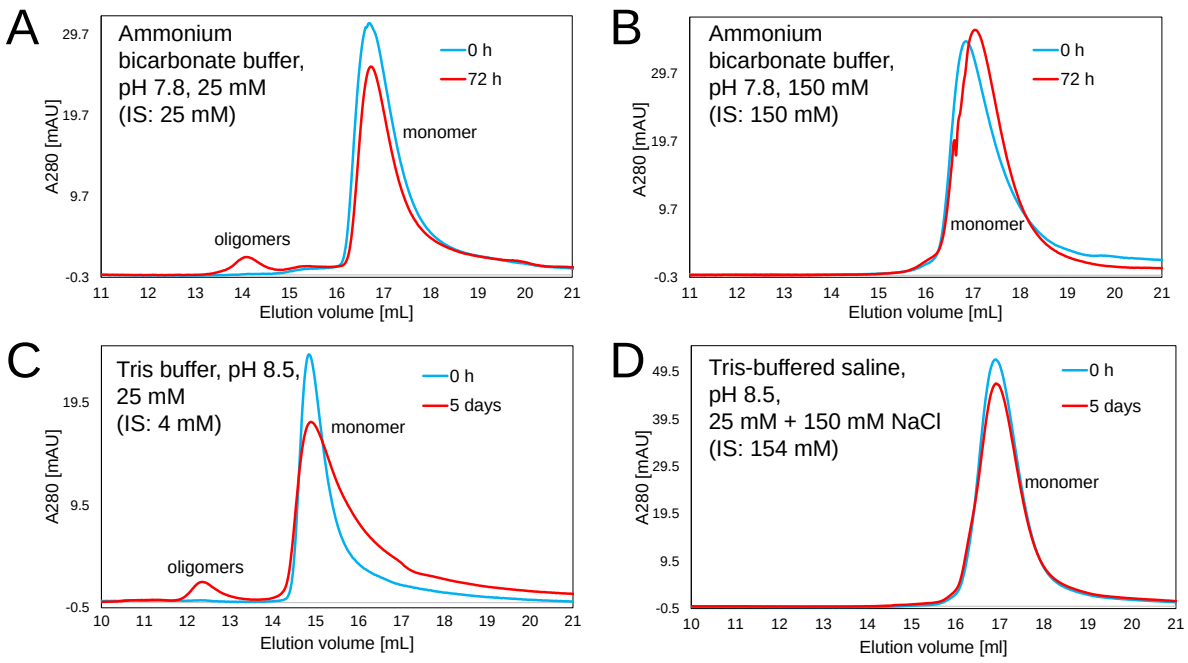


Figure 3.8: Formation of low molecular weight oligomers of GLP-1-Am in low and high ionic strength (IS) buffers. (A) 100 µM GLP-1-Am in 25 mM ammonium bicarbonate at pH 7.8 (IS = 25 mM) was analysed at time zero and after 72-hour incubation at 37 °C with agitation. (B) 100 µM GLP-1-Am in 150 mM ammonium bicarbonate at pH 7.8 (IS = 150 mM) was analysed at time zero and after 72-hour incubation at 37 °C with agitation. (C) 120 µM GLP-1-Am in 25 mM Tris buffer at pH 8.5 (IS = 4 mM) was analysed at time zero and after 5-day incubation at 37 °C with agitation. (D) 120 µM GLP-1-Am in Tris-buffered saline at pH 8.5 (IS = 154 mM, 25 mM Tris and 150 mM NaCl) was analysed at time zero and after 5-day incubation at 37 °C with agitation. All samples were analysed using a Superdex 75 10/300 size exclusion column.

Additionally, the changes in oligomer distribution of GLP-1 and GLP-1-Am over time were studied using velocity sedimentation experiments. GLP-1 and GLP-1-Am samples of 85 µM concentration were either freshly prepared before the measurement or incubated for 7 days at 37 °C with 180 rpm agitation. The experiment was performed with centrifugation at 50,000 rpm and sedimentation curves were collected over *circa* 24 hours for each sample. The data were fitted to a continuous c(*s*) distribution model in a Sedfit program[221]. Apart from a peak corresponding to the non-sedimented monomer with a sedimentation coefficient close to 0, both peptides formed low-molecular weight oligomers, with a high content of dimer (over 40 % in fresh samples) and lower amounts of large oligomer (from 4-mer to 13-mer), Figure 3.9. The mass and abundance of each of the detected oligomeric species for both fresh and aged samples, as determined by the Sedfit program, are given in Table 3.2. In contrast to the size-exclusion chromatography, during the sedimentation velocity experiment, the samples do not undergo additional dilution and do not interact with the column matrix. Therefore, oligomeric species which do not have sufficient stability to be detected by SEC (mainly oligomers formed in freshly prepared samples) can be detected in

sedimentation velocity experiments. Changes in oligomer distribution over time were apparent for both GLP-1 (Figure 3.9A) and GLP-1-Am (Figure 3.9B). GLP-1 showed changes in the ratio of monomeric and oligomeric peaks and GLP-1-Am formed new oligomeric species over time. Discrepancies in the amounts of oligomers detected by SEC (Table 3.1) and sedimentation velocity (Table 3.2) are likely to be due to the fact that it is possible to detect less stable transient oligomers in sedimentation velocity experiments, e.g., two types of dimers may be present in the sample – the less stable transient dimer, which is detected only by sedimentation velocity, and the stable dimer, which can be detected in both sedimentation velocity and SEC experiments. In addition, different timescales were used in the SEC and sedimentation velocity experiments which may also contribute to the differences observed.

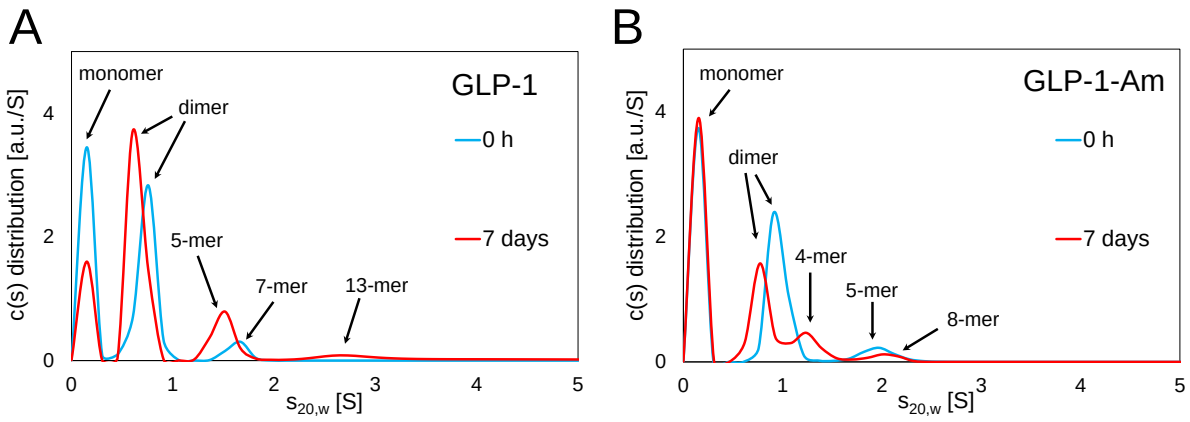


Figure 3.9: Formation of low-molecular weight oligomers of GLP-1 and GLP-1-Am observed by sedimentation velocity. GLP-1 (A) and GLP-1-Am (B) at 85 µM were prepared in 25 mM phosphate at pH 8 prior to the experiment or incubated at 37 °C with 180 rpm agitation for 7 days. After a 2-hour temperature equilibration to 20 °C, the experiment was performed with centrifugation at 50,000 rpm using a Beckman Optima XL-I Analytical Ultracentrifuge. The interference sedimentation curves were collected as 300 scans (approx. 24 h run time) and fitted to a continuous c(*s*) distribution model implemented in the Sedfit program. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$). The size of oligomeric species was calculated using the Sedfit program.

GLP-1 fresh samples				**GLP-1-Am fresh samples**		
$s_{20,w}$ [S]	MW [kDa]	Content [%]		$s_{20,w}$ [S]	MW [kDa]	Content [%]
0.15 0.75 1.67	N/A* (monomer) 7.39 (dimer) 23.9 (7-mer)	42 45 10		0.15 0.91 1.97	N/A* (monomer) 6.09 (dimer) 18.0 (5-mer)	46 46 8
GLP-1 7-day incubation				GLP-1-Am 7-day incubation		
$s_{20,w}$ [S]	MW [kDa]	Content [%]		$s_{20,w}$ [S]	MW [kDa]	Content [%]
0.15 0.60 1.52 2.73	N/A* (monomer) 4.73 (dimer) 16.7 (5-mer) 43.6 (13-mer)	20 58 16 6		0.15 0.77 1.24 2.01	N/A* (monomer) 7.49 (dimer) 14.5 (4-mer) 27.7 (8-mer)	48 34 13 4

Table 3.2: Quantitative estimation of mass and abundance of species detected in sedimentation velocity experiments. The data correspond to the sedimentation profiles shown in Figure 3.9. The molecular weight and relative abundance of detected species were calculated in the Sedfit program. * It was not possible to calculate the mass of the GLP-1/GLP-1-Am monomer in the Sedfit program due to its limited sedimentation under the conditions used. The molecular weight of the GLP-1/GLP-1-Am monomer is approximately 3.35 kDa.

3.3.3 Structural characterization of the low-molecular weight oligomers

As was shown in my Masters project[86], low-molecular weight oligomers, detected during aggregation of GLP-1-Am, have a highly disordered structure (Chapter 1, Section 1.7). Low-molecular weight oligomers of GLP-1 are structurally characterized in a similar manner in this Section.

Spectra of non-aggregated peptide were recorded immediately after dissolving the lyophilized peptide powder. The low-molecular weight oligomers formed slowly over time

were either separated using SEC or obtained after filtering out the fibrils from the aggregated sample using a 0.22 µm membrane, and their spectra were recorded directly afterwards.

For a freshly prepared sample of 85 µM GLP-1 sample at pH 7, the far-UV circular dichroism (CD) spectrum showed a contribution from α-helical and β-structure as well as disordered regions, Figure 3.10A. The far-UV CD spectra of isolated oligomeric species (Figure 3.10A) formed by GLP-1 at pH 7 showed a prevalence of disordered regions. Intrinsic tryptophan fluorescence spectra were also measured for both samples. The fluorescence maximum (λ_{max}) is at 354 nm for a freshly prepared GLP-1 sample and at 357 nm for the sample containing only low-molecular weight oligomers, Figure 3.10B. Trp25 in the freshly prepared sample has a slightly lower λ_{max} than in the low-molecular weight oligomers (354 nm *vs* 357 nm) suggesting that the Trp side chain more solvated after oligomerization which is consistent with the observations from the far-UV CD spectra that the peptide is largely disordered under these conditions.

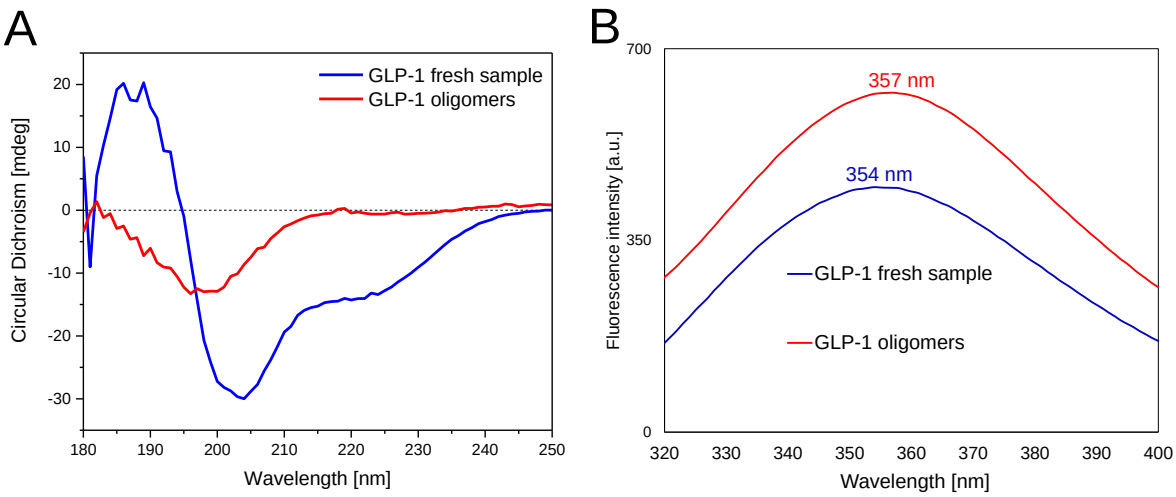


Figure 3.10: Far-UV CD and intrinsic tryptophan fluorescence spectra of freshly prepared GLP-1 and low-molecular weight oligomers of GLP-1. Far-UV CD (A) and intrinsic tryptophan fluorescence (B) spectra of the monomeric GLP-1 were measured for a freshly prepared peptide sample which was dissolved in 10 mM sodium phosphate buffer at pH 7. Spectra of GLP-1 oligomers were measured after filtration of a sample that had undergone a 10-day incubation at 37 °C with agitation, conditions in which all monomer was depleted as shown by SEC, Figure 3.4A.

In addition, the low-molecular weight oligomers of GLP-1 exhibit light scattering revealed by an increased in absorbance over the range 310-350 nm in the UV-Vis absorption spectrum in contrast to the freshly prepared GLP-1 sample (Figure 3.11). This is a further indication of aggregation[264]. Considering the small size of these oligomers, they show a surprisingly high level of light scattering as observed in the UV-Vis absorption spectra, which may suggest their further coagulation in the solution.

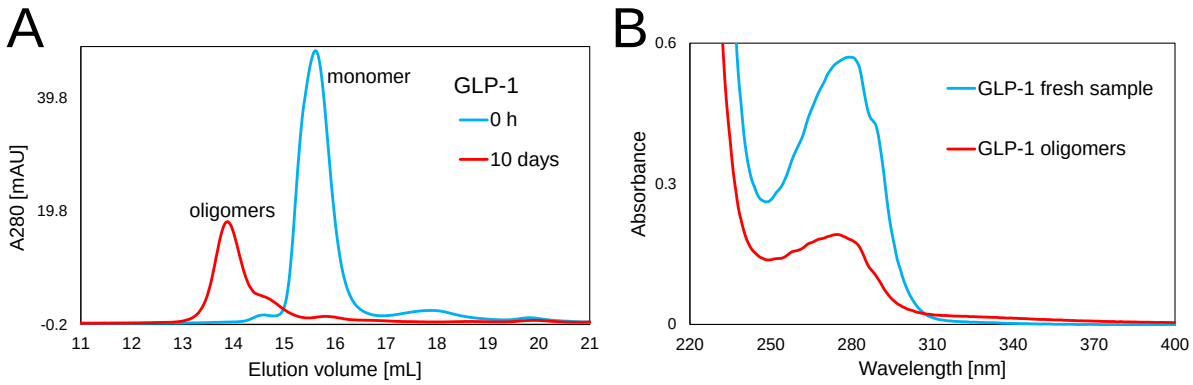


Figure 3.11: SEC chromatograms and absorption spectra of freshly prepared GLP-1 and low-molecular weight oligomers of GLP-1. (A) GLP-1 at 85 µM peptide concentration was incubated in 10 mM sodium phosphate buffer at pH 7, at 37 °C with agitation. This sample was analysed using SEC (Superose 12 10/300) at different time points. Panel A shows the elution profile corresponding to a freshly prepared GLP-1 sample along with the elution profile of oligomers of GLP-1 which were formed after 10 days of sample incubation. (B) UV absorption spectra of fresh sample and oligomer as described in (A). The oligomeric GLP-1 shows detectable light scattering as apparent from an increased absorption at around 320 nm. The concentration of oligomers is lower than the initial concentration of the peptide due to the competing aggregation/fibrillation processes that resulted in the formation of large insoluble aggregates which were removed by filtration prior to the analysis.

3.3.4 Membrane disruption experiments

The membrane disruption potential of low-molecular weight oligomers of GLP-1-Am was tested using a calcein-release assay. The assay was performed with 50 µM of calcein-encapsulated 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS) vesicles of an average diameter of 125 nm (calcein concentration inside vesicles >70 mM) at pH 8. Low-molecular weight oligomers of GLP-1-Am were isolated using SEC from an aged sample of 200 µM peptide incubated in 25 mM phosphate at pH 8 for 7 days. Isolated oligomers were concentrated using an Amicon Ultra Centrifugal Filter Unit with a 3 kDa MWCO. At the end of the assay 5 % (v/v) Triton X100 to each well which causes instant disruption of the vesicles. The fluorescence signal was then normalized, taking the fluorescence value after the addition of 5 % (v/v) Triton X100 (complete disruption of vesicles) as 1.

Interestingly, the calcein-release caused by disruption of about 35 % of the total amount of vesicles was detected when 4 % (w/w) of freshly dissolved GLP-1-Am sample was incubated with the vesicles, Figure 3.12. For the samples with 4% (w/w) of pre-formed isolated GLP-1-Am oligomers, no release of calcein from vesicles was detected over the timescale of the assay, Figure 3.12, which indicates that the low-molecular weight oligomers are not capable of disruption of the liposome membrane. This rather surprising observation

may suggest some affinity of native GLP-1-Am for membranes which is lost after its misfolding which accompanies the oligomerization.

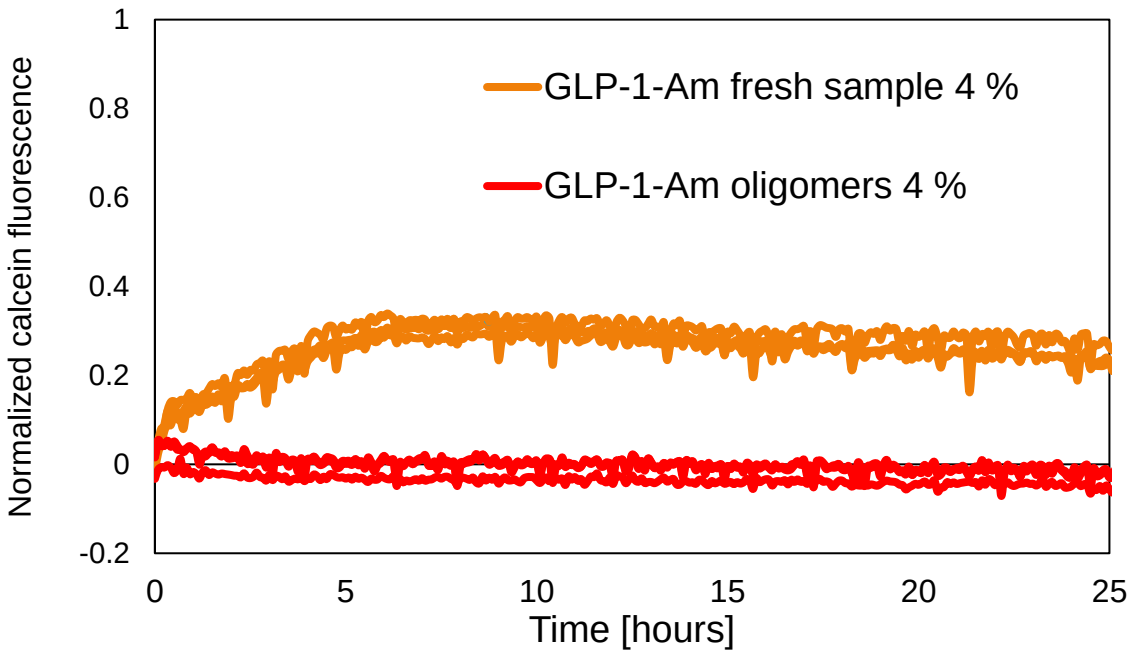


Figure 3.12: Calcein-release assay with 4 % of either freshly dissolved sample of GLP-1 or low-molecular weight oligomers of GLP-1-Am. The assay was performed with 50 µM 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS) vesicles at pH 8 in 37 °C over 25 hours. The fluorescence signal at 520 nm was recorded every 5 minutes upon 484 nm excitation.

3.3.5 Stability of low-molecular weight oligomers

The stability of the low-molecular weight oligomers of GLP-1 and GLP-1-Am was investigated using SEC. Samples of GLP-1 and GLP-1-Am were incubated at pH 8 for 4/5 days until a sufficient population of low-molecular oligomers was formed. SEC fractions containing only low-molecular weight oligomers were subsequently reinjected onto the column, Figure 3.13. It was shown that the low-molecular weight oligomers of both GLP-1 (Figure 3.13A) and GLP-1-Am (Figure 3.13B) are stable upon their reinjection, i.e., they do not dissociate back into the monomer, although their concentration in the separated fraction is significantly lower than in the original sample. In my previous work during my Master's degree, the oligomers were also shown to be stable with respect to many stressors such as high temperature (90 °C), agitation or low percentages of isopropanol[86].

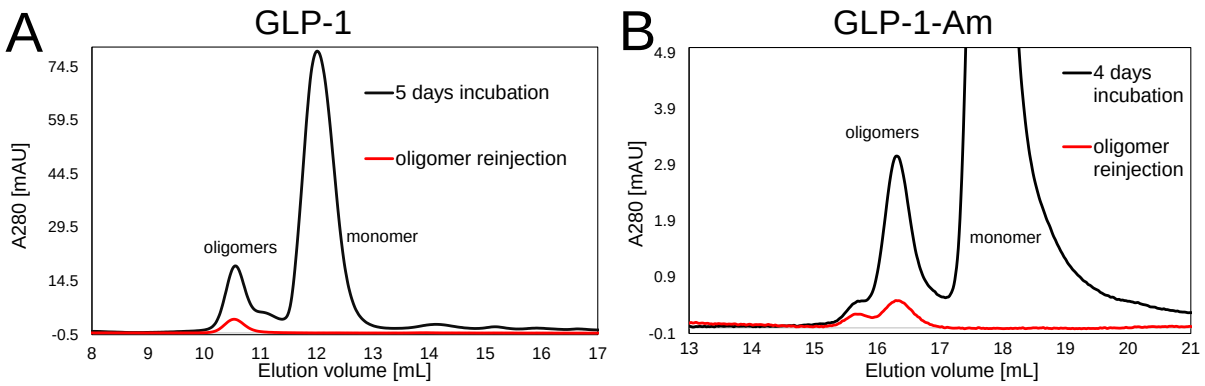


Figure 3.13: Stability of low-molecular weight oligomers of GLP-1 and GLP-1-Am upon reinjection onto a size-exclusion column. (A) 150 µM GLP-1 was analysed after a 5-day incubation at 37 °C with agitation in 10 mM Tris buffer at pH 8, oligomeric fractions were isolated and immediately reinjected onto the Superdex 75 10/300 size-exclusion column. (B) 85 µM GLP-1-Am was analysed after a 4-day incubation at 37 °C with agitation in 25 mM sodium phosphate buffer at pH 8, oligomeric fractions were isolated and immediately reinjected onto a Superose 12 10/300 size-exclusion column.

The stability of these unstructured oligomers, in addition to their slow rate of formation, suggests that they might be stabilized by covalent links. LC-MS was used to investigate this hypothesis. The LC-MS analysis of GLP-1-Am low-molecular weight oligomers, using electrospray ionization in the positive mode, showed the MS spectrum of GLP-1-Am oligomers to be identical to the spectrum of the GLP-1-Am monomer, Figure 3.14B and D. The total ion chromatogram of GLP-1-Am monomer showed a single sharp peak corresponding to the peptide elution (Figure 3.14A) whereas the chromatogram of oligomeric sample showed multiple coeluting peaks (Figure 3.14C). Nevertheless, the coeluting peaks corresponded to the same MS spectrum, consistent with a GLP-1-Am monomer, in both molecular weight and isotopic distribution, Figure 3.15. Therefore, the LC-MS analysis of GLP-1-Am oligomers did not identify any covalent links between peptide chains leading to oligomer formation that might have occurred during the prolonged incubation at 37 °C.

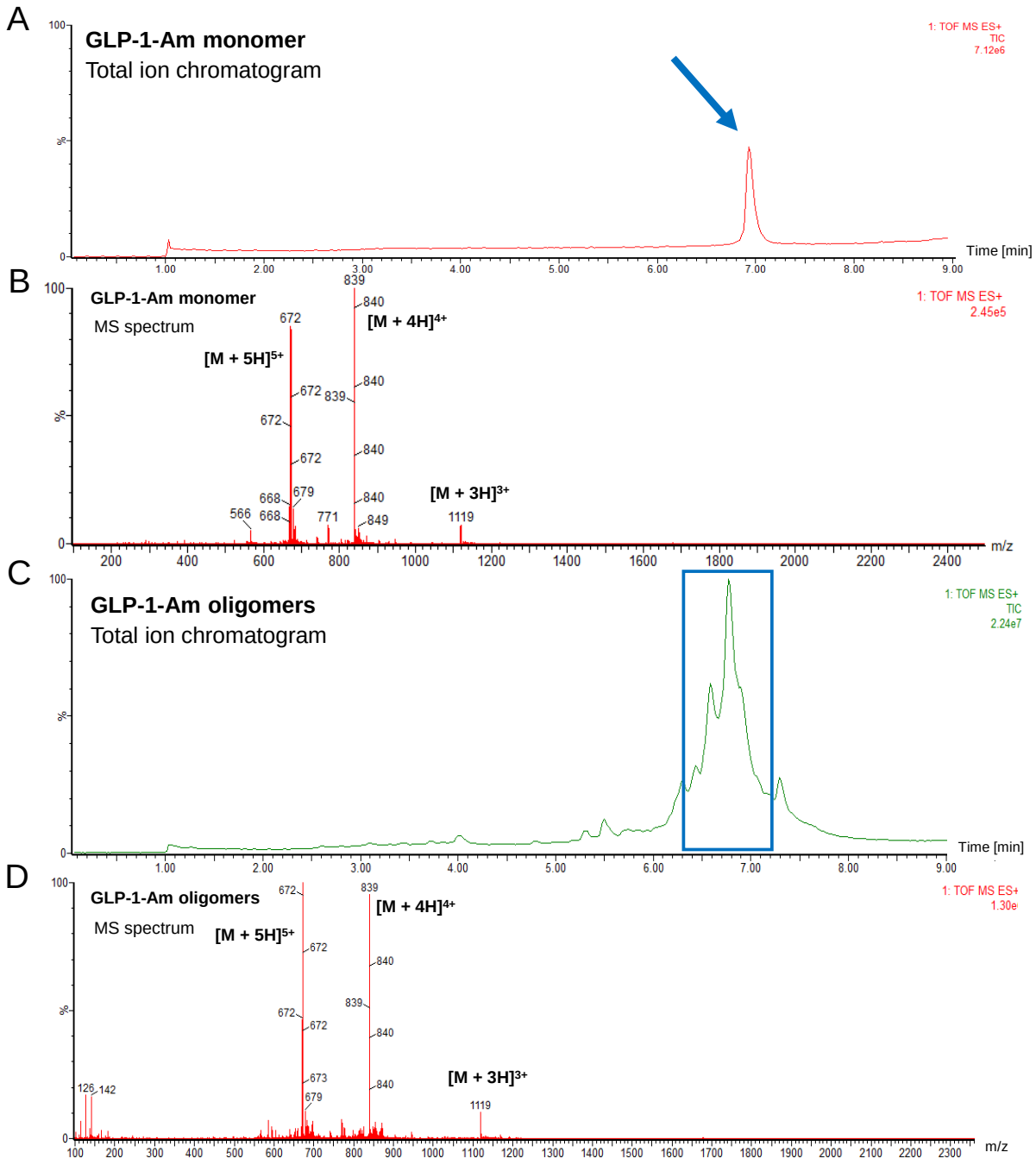


Figure 3.14: LC-MS analysis of monomeric and oligomeric GLP-1-Am. Total ion chromatogram of monomeric GLP-1-Am – the blue arrow indicates the peptide peak on the total ion chromatogram (A), corresponding mass spectrum is shown in (B). LC-MS analysis of GLP-1-Am oligomers (C, D). The total ion chromatogram shows multiple overlapping peaks (C). All the peaks in the area marked with the blue rectangle correspond to the MS spectrum in (D) which is consistent with the spectrum of the GLP-1-Am monomer. Total ion chromatogram reflects a sum of intensities of all mass spectral peaks at a given time point (i.e., plotted against the time).

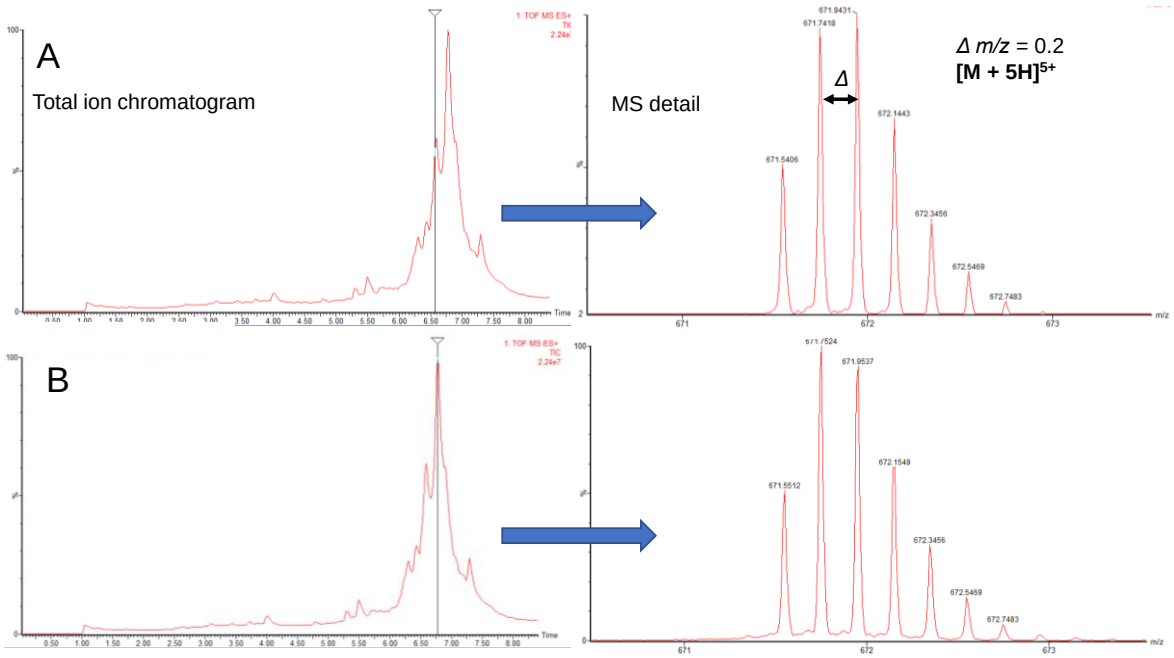


Figure 3.15: LC-MS analysis of GLP-1-Am oligomers – isotopic distribution. The isotopic distribution of the peak with *m/z* around 672 was analysed at two different time points in the total ion chromatogram (A and B). Both isotopic distributions correspond exactly to the five times charged state of the GLP-1-Am monomer.

Due to the slow formation and high stability of the observed oligomers, SEC and sedimentation velocity experiments were also performed with GLP-1-Am containing a β-aspartic at position 9, βAsp_GLP-1-Am, to exclude the possibility that the observed species originate from isomerization of aspartic acid in the peptide. The isomerization of Asp9 did not have any effect on the elution volume of the monomer and only a minor effect was observed in the sedimentation velocity experiment compared with a freshly prepared sample of GLP-1-Am, Figure 3.16. Therefore, the observed low-molecular weight oligomer species are not likely to result from the slow isomerization of Asp9.

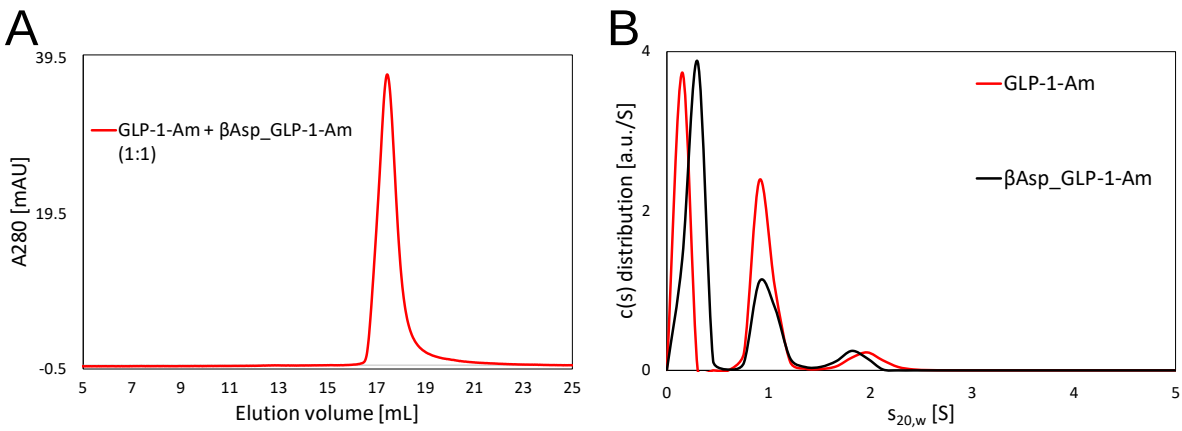


Figure 3.16: Size-exclusion chromatography and sedimentation velocity of freshly prepared samples of GLP-1-Am and βAsp_GLP-1-Am. (A) Size-exclusion chromatography of a freshly prepared mixed sample of GLP-1-Am and βAsp_GLP-1-Am in a 1:1 ratio at a total concentration of 85 µM in 25 mM phosphate, pH 8. Size-exclusion chromatography was performed using a Superose12 10/300 column. (B) Sedimentation velocity of freshly prepared samples of GLP-1-Am and βAsp_GLP-1-Am at 85 µM in 25 mM phosphate pH 8. After a 2-hour temperature equilibration to 20 °C, the experiment was performed at 50,000 rpm. The interference sedimentation curves were collected as 300 scans (approx. 24 h run time) and fitted to a continuous c(*s*) distribution model implemented in a Sedfit program[221]. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$).

Other chemical modifications which may be induced by prolonged incubation at 37 °C are deamidation and/or isomerization of Gln17. Although glutamine deamidation occurs with extremely slow rates (half-lives ranging from 600 to 20,000 days)[265–267], the possibility of deamidation in the aged sample of GLP-1-Am was tested using an isoelectric focusing gel. As shown in Figure 3.17, 85 µM GLP-1-Am incubated in 25 mM phosphate at pH 8 for 8 days shows a single band corresponding to the pI of a freshly prepared sample of GLP-1-Am. Therefore, the observed low-molecular weight oligomeric species are not likely to be formed as a result of deamidation.

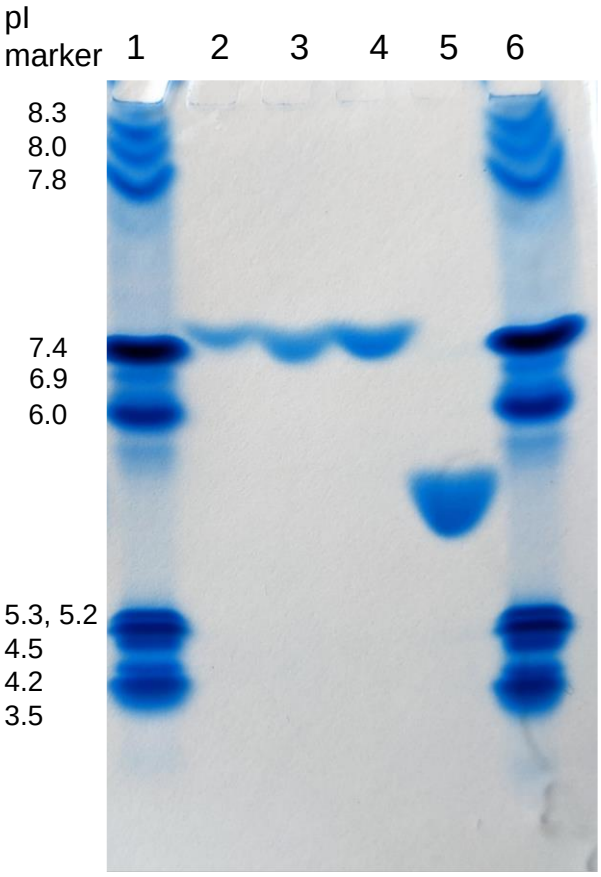


Figure 3.17: Isoelectric focusing gel electrophoresis of fresh and aged samples of GLP-1-Am, βAsp_GLP-1-Am and GLP-1. Lanes 1 and 6: pI protein standard markers, lane 2: GLP-1-Am incubated at 85 µM peptide concentration in 25 mM phosphate, pH 8, at 37 °C with agitation for 8 days, lane 3: GLP-1-Am (fresh sample), lane 4: βAsp_GLP-1-Am (fresh sample), lane 5: GLP-1 (fresh sample). Values of the isoelectric point (pI) standards are given on the left side of the gel image. The pI values of samples in lanes 2, 3, 4 are around 7.4; The pI of a sample in lane 5 is in the range 5.5–5.8.

3.3.5 Effect of low-molecular weight oligomers of GLP-1-Am on fibrillation kinetics

To assess the effect of low-molecular weight oligomers on the fibrillation kinetics, a ThT assay was performed in which pre-formed purified oligomers of GLP-1-Am were added to the GLP-1-Am solution at the beginning of the aggregation reaction, Figure 3.18. Two different concentrations of GLP-1-Am were used. The addition of pre-formed oligomeric species was shown to have a slight inhibitory effect on the fibrillation process. For 50 µM GLP-1-Am (Figure 3.18A), the addition of oligomers affected mainly the apparent growth phase but not the lag time of the reaction, indicating that the oligomers are likely to inhibit/slow down the secondary nucleation processes that occur in this phase of fibril formation. In the case of 25 µM GLP-1-Am (Figure 3.18B), the addition of pre-formed oligomers also delayed the lag time of the reaction.

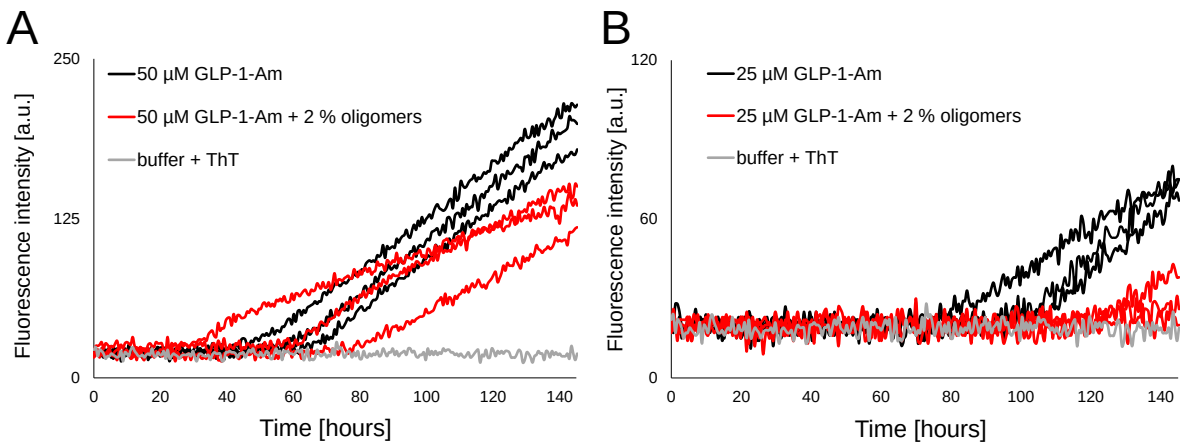


Figure 3.18: Thioflavin T assay of GLP-1-Am with the addition of pre-formed oligomers. Samples of 50 (A) and 25 µM (B) GLP-1-Am in 25 mM phosphate at pH 8 with and without the addition of 2 % (w/w) of pre-formed purified GLP-1-Am oligomers were incubated with 50 µM thioflavin T at 37 °C with agitation for 145 hours. All samples were run in triplicate. GLP-1-Am oligomers were pre-incubated in 25 mM phosphate pH 8 for 5 days and then isolated using SEC.

3.4 Discussion

As was previously reported, GLP-1 forms amyloid fibrils under a wide range of conditions with differing fibrillation rates which depend critically on pH and peptide concentration[36,225]. Under some conditions, there is evidence that GLP-1 not only forms on-pathway oligomers, which can convert directly into amyloid fibrils, but that it can also populate oligomeric species that are off-pathway products, and which thus reduce the rate of formation of amyloid fibrils[36]. This Chapter shows that a C-terminally amidated variant, GLP-1-Am, also fibrillates under similar conditions to GLP-1 and also populates off-pathway oligomers, as shown by the fact that the lag time associated with its aggregation kinetics does not decrease with increasing peptide concentration at pH 8, Figure 3.1D. A detailed comparison of the fibrillation of GLP-1 and GLP-1-Am has been published[225]. The main focus of this investigation was to isolate and characterize the putative off-pathway oligomers for GLP-1 and GLP-1-Am, in order to understand their structures, why they form and how they affect the self-assembly process.

3.4.1 GLP-1 and GLP-1-Am form oligomeric species that compete with fibrillation

In this Chapter, the formation of low-molecular weight oligomers of both GLP-1 and GLP-1-Am in the pH range 7 to 8.5 is reported. Using SEC, distinct well-resolved oligomeric

peaks were detected after a prolonged incubation at 37 °C in different aqueous buffers, Figure 3.4, Figure 3.7, Figure 3.8. The formation of low-molecular weight oligomers was reproducible for both GLP-1 and GLP-1-Am. In addition, for GLP-1-Am these results were reproduced for two different peptide batches produced by different manufacturers (Figure 3.6), for GLP-1 only one batch of the peptide was used. The formation of these low-molecular weight oligomers during incubation at 37 °C was also confirmed by sedimentation velocity experiments, Figure 3.9. The increase in the intensity of the oligomeric peaks occurs over days showing that their rate of formation is rather slow, indicating that this oligomerization process is not self-propagated/catalyzed, in contrast to fibrillation[76]. Importantly, these oligomers remained in solution even after depletion of the peptide monomer, under conditions where a considerable amount of peptide was converted into a fibrillar species, Figure 3.4C. This suggests that the oligomers are not easily capable of further aggregation into amyloid fibrils or of dissociation back into monomers. Most likely, they represent a stable form of aggregate and, therefore, their formation competes with fibrillation.

3.4.2 GLP-1 and GLP-1-Am form oligomers ranging from dimers to pentamers

Using a set of protein calibration standards on two size-exclusion columns Superose 12 10/300 and Superdex 75 10/300, the size of the “low-molecular weight” oligomers of GLP-1-Am was estimated to be in the range of dimers to tetramers, whilst the GLP-1 oligomers were estimated to be in the size range from trimers to pentamers, Figure 3.4. For small peptides such as GLP-1, it is beyond the scope of this technique to determine the size more precisely, because the calibration of the size-exclusion column is based on the separation of globular proteins according to their hydrodynamic radii. This could cause inaccuracies in the size estimates since GLP-1 and GLP-1-Am monomers and oligomers are unlikely to be fully globular. The presence of GLP-1 and GLP-1-Am oligomers from a size range of dimers to pentamers (with a small amount of larger species) was also confirmed in sedimentation velocity experiments, Figure 3.9.

3.4.3 Oligomers of GLP-1 and GLP-1-Am have highly disordered structures

As was shown in my previous Master's research[86], low-molecular weight oligomers of GLP-1-Am have a highly disordered character. In this Chapter, low-molecular weight oligomers of GLP-1 are also characterized to be highly disordered. Far-UV CD spectra reveal that freshly prepared, solution of GLP-1 has regions of α-helical and β-structure, whereas the low-molecular weight oligomeric species show significantly less of these secondary structure motifs, and a prevalence of disordered regions, Figure 3.10A. Highly disordered oligomers have also been reported in the aggregation of Aβ40 and Aβ42[109], transthyretin[105] and α-synuclein[96,104]. Some studies on the cytotoxicity of α-synuclein[268] and Aβ40[269] oligomers suggested that higher structural complexity and a high content of β-sheet is needed to disrupt the cell membrane. Indeed, the highly disordered low-molecular weight oligomers described in this study did not show any membrane disruption potential when assessed using calcein-release assay with DOPS vesicles, Figure 3.12. However, as reported elsewhere[94,270], the secondary structure of an oligomer cannot always be linked with cytotoxicity without experimental assessment, since the presence of hydrophobic patches on the surface of an oligomer may induce cytotoxicity.

3.4.4 The low-molecular weight oligomers of GLP-1 and GLP-1-Am are stable but not covalent species

The fact that the low-molecular-weight oligomers of both GLP-1 and GLP-1-Am were detected as fully resolved peaks in SEC experiments highlights their stability, as they did not dissociate during the ten-fold dilution of the sample as it ran through the SEC column. N.B., detection of peptide or protein oligomers using SEC is not always possible without oligomers first being cross-linked to enhance their stability[82]. These results, along with the timescales on which these oligomers are formed (1–15 days), suggest that the stability might be a direct result of a covalent modification of the peptide resulting in essentially cross-linked species. The GLP-1 sequence does not contain any cysteines, therefore disulfide bridges cannot form, but potential dimerization might have occurred through tyrosine or tryptophan side chains, although this is less likely. Nevertheless, LC-MS did not detect any chemical modification of this type within the oligomerised peptide. It is, therefore, highly likely that despite the unusual stability and disordered conformation of the low-molecular weight oligomers of

GLP-1 and GLP-1-Am, they are non-covalent in character. Non-covalent and non-crosslinked oligomers of self-assembly-prone peptides and proteins that of a similar stability with respect to time and dilution have been previously observed for α-synuclein[104], transthyretin[105], SOD1[101], cystatin[106,107], prion protein[108] and amyloid-β peptide[102,109,110] and β-2 microglobulin where Cu^{2+} ions drive oligomer formation but are not essential for its stability[111].

Given these results, it is interesting to speculate on the non-covalent forces that maintain the peptides in their oligomeric state. Although the hydrophobic effect is frequently of importance in self-assembly reactions, there is no burial of the hydrophobic side chain of Trp25 in the oligomers of GLP-1 and GLP-1-Am, suggesting that it may not play a dominant role here (of course, the burial of other hydrophobic sides chains in the peptide cannot be excluded). In contrast, there is some evidence that electrostatic interactions play a role, as demonstrated by the fact that the oligomers form preferentially in low ionic strength buffers. The fact that very disordered polypeptide chains can form complexes with high stability has been shown recently in the case of the C-terminal tails of histone H1[271,272]. These results further establish that fixed structure is not essential for maintaining a complex – in the case of H1 a hetero-oligomer with the nuclear chaperone prothymosin-α, in the case of GLP-1 and GLP-1-Am, described here, homo-oligomers.

3.4.5 Low-molecular weight oligomers of GLP-1 and GLP-1-Am are likely off-pathway states

As was shown in this Chapter, low-molecular weight oligomers of GLP-1 and GLP-1-Am are very resistant to further association into fibrils, suggesting that they are not intermediates *en route* to fibrils, i.e., on-pathway oligomers, but they are rather off-pathway species formed by a process in competition with fibrillation. Moreover, as shown in the ThT assay with pre-formed low-molecular weight oligomers of GLP-1-Am, the addition of oligomers slightly slows down the fibrillation kinetics, Figure 3.18. A similar inhibitory effect on amyloid formation has recently been reported for D76N-β2-microglobulin dimers[273]. This, in part, helps to rationalize the unusual dependence that has been observed between aggregation lag times and peptide concentration for GLP-1, where under specific conditions the lag times increase with increasing GLP-1 concentration[36]. For such unusual kinetics, a mechanism containing a unimolecular step has been proposed which can explain the inverse concentration-fibrillation trend[80]. In this mechanism, the original “fibrillation-inactive”

monomer has to convert into a "fibrillation-active" monomer, which can further form on-pathway oligomers and amyloid fibrils[36,80]. "Fibrillation-inactive" monomers have also been proposed to inhibit/slow the fibrillation process by interacting with on-pathway oligomers[80]. We hypothesize that the observed stable low-molecular weight oligomers of GLP-1 and GLP-1-Am are either formed by "fibrillation-inactive" monomer self-assembly or by the conversion of on-pathway oligomers upon interaction with the "fibrillation-inactive" monomer or by a combination of these two mechanisms, Figure 3.19A. As was suggested by the ThT assay with pre-formed isolated oligomers (Figure 3.18), the low-molecular weight oligomers may also partly inhibit the secondary fibrillation processes and therefore slow the formation of amyloid fibrils, Figure 3.19B.

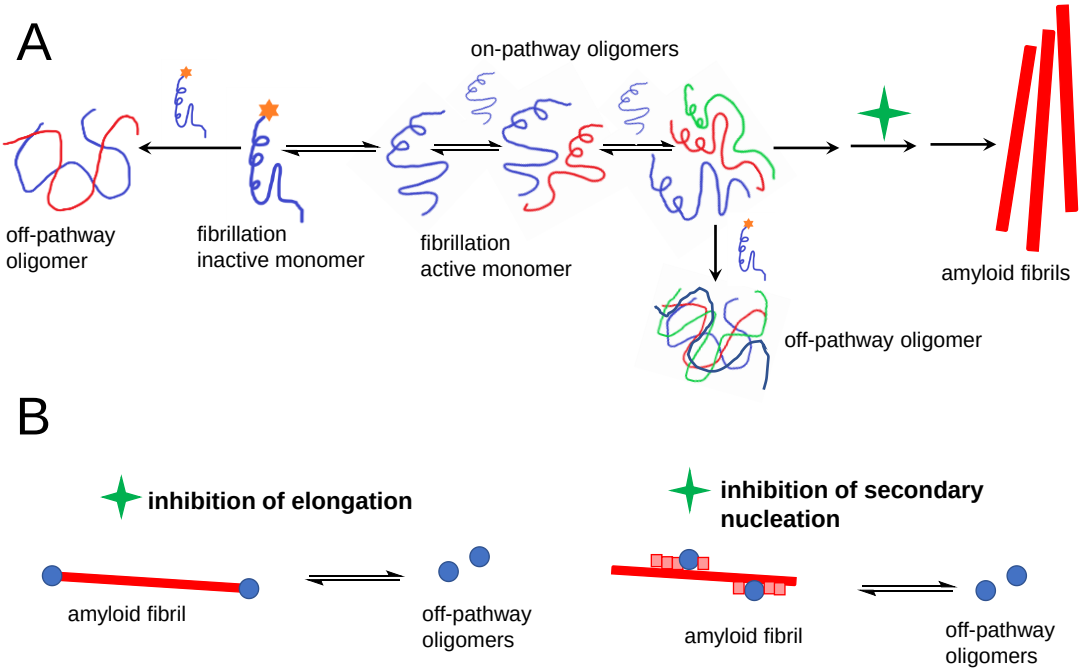


Figure 3.19: Possible mechanisms for the formation of off-pathway low-molecular weight oligomers and their inhibition of secondary nucleation processes and fibrillation. "Fibrillation-inactive" monomers (indicated by a star) either self-assemble into an off-pathway oligomer or convert into the "fibrillation-active" monomer in a unimolecular step and/or the off-pathway oligomer is formed upon interaction of the on-pathway oligomer with the "fibrillation-inactive" monomer (A). The green star in the scheme A refers to the processes illustrated in B. Off-pathway oligomers may also reversibly inhibit elongation and/or secondary nucleation processes and slow the overall kinetics of amyloid formation (B).

To sum up, it was shown that GLP-1 and GLP-1-Am, which are prone to amyloid fibril formation over a wide range of conditions, both form low-molecular weight oligomers at pH values between 7 and 8. These oligomers are not easily capable of further self-assembly into larger oligomers or amyloid fibrils or dissociation back into monomers. Moreover, the addition of pre-formed oligomers to an aggregation reaction of monomer slightly slows

down the fibrillation kinetics. Therefore, they are likely products of an alternative self-assembly pathway that competes with fibril formation. The size of the oligomers was estimated to be from dimer to pentamer, and they were also shown to have a highly disordered structure. Surprisingly, these oligomers are stable with respect to time, temperature, and even a low percentage of isopropanol[86]. However, no covalent modification of oligomers was revealed by LC-MS. This observation of stable oligomeric species which are "off-pathway" to the fibrillation process further illustrates the diversity and complexity of the aggregation landscape of peptides in which on-pathway oligomers, protofibrils and fibrils are not the only stable detectable species. It also sheds some light on the unusual dependence of GLP-1 aggregation kinetics on peptide concentration, which has been previously observed under some conditions[36].

CHAPTER 4
ENERGY LANDSCAPES OF MONOMERIC GLP-1

Calculations in this Chapter were done in collaboration with the Wales group (Yusuf Hamied Department of Chemistry, Cambridge) using an approach developed in their group. Guidance and help with the workflow of the calculations was provided by Dr. Konstantin Röder and Dr. Alasdair Keith, who also initiated the sampling of databases of monomeric GLP-1 at protonation states corresponding to pH 4 and 7.5. My contribution to this project included constructing the database at protonation state corresponding to pH 3, final refinement of all databases and their analysis.

4.1 Background

As described in the Introduction and previous Chapters, GLP-1 forms β-sheet-rich amyloid fibrils over a wide range of pH values[36,38,260,261]. In these cases, the monomeric peptide first self-associates to form oligomers which go on to form a critical nucleus. This nucleus further elongates into an amyloid fibril which can then self-catalyse fibrillation through secondary nucleation mechanisms. Both the kinetics and mechanism of GLP-1 aggregation are known to be highly pH-dependent and different pH values favour the formation of different oligomeric states[36]. Monomers play a key role, both in the initial formation of oligomers and primary nuclei, as well as in elongation, so understanding more about the monomeric state at different pH values and how this influences the aggregation kinetics/mechanism is vital. In this Chapter, a computational potential and the free energy landscape framework is applied to study monomeric GLP-1(7-37).

The construction of energy landscape networks using computational methods provides a powerful way to study the structural features and global dynamics of biomolecular systems[274]. The energy landscape consists of stationary points – local minima and transition states. A local minimum is geometrically characterized by having all non-zero normal mode frequencies real, whereas the transition states have exactly one imaginary normal mode frequency. The energy landscapes in this Chapter are visualized using disconnectivity trees[275,276] (Figure 4.1). Disconnectivity trees are constructed from the databases of stationary points and their vertical axis represents either potential or free energy.

An energy landscape dominated by a single funnel leads to a single structural ensemble as shown in Figure 4.1A, e.g., a native state in the example of protein folding. These energy landscapes, which are said to have minimal frustration, are characteristic of the systems without competing off-pathway products and with only one stable, folded structure. Peptides and intrinsically disordered proteins show multiple distinct structural ensembles in equilibrium with each other[277]. In the energy landscape framework, these systems are represented as frustrated energy landscapes possessing groups of low energy minima of different structural motifs which are separated by high energy barriers[278,279]. These groups of low energy minima are visualized as separate funnels, Figure 4.1B.

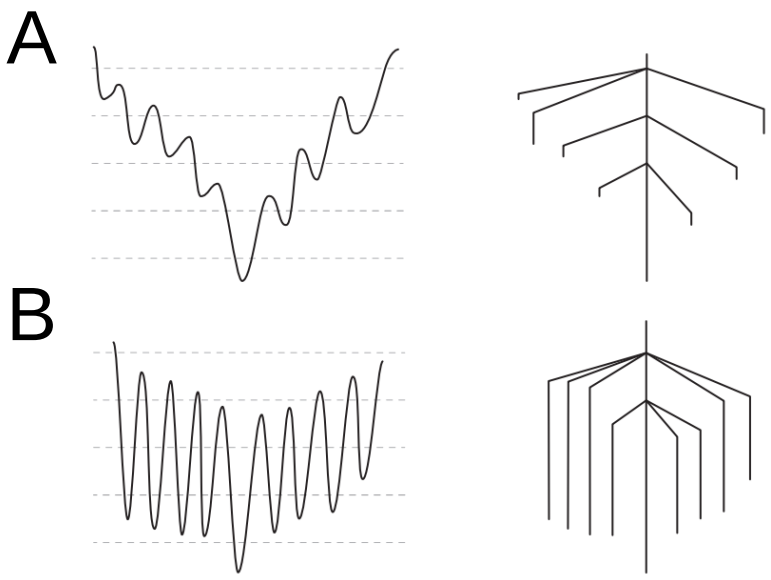


Figure 4.1: One-dimensional representation of distinct types of potential energy landscapes. Adapted from Wales (2003)[280]. The energy landscape in A is dominated by a single funnel. This type of energy landscape is typical of well-behaved folded proteins. The energy landscape in B shows a multi-funnel character and contains multiple deep funnels separated by energy barriers. This landscape is characteristic of intrinsically disordered proteins or short peptides with high conformational flexibility. Left and right side of the figure in both A and B shows different graphical representation of the same energy landscape (left – folding funnel scheme; right – disconnectivity tree scheme).

4.2 Aims

In this Chapter, the energy landscape framework method is employed to study the conformational propensities and flexibility of monomeric GLP-1 with different protonation states which correspond to different pH values. The aim of this simplified approach is to establish whether there are any correlations between the energy landscapes of monomeric GLP-1 and its propensity to form cross-β-sheet-rich fibrils over a wide pH range.

4.3 Building the database – workflow

The databases of stationary points, i.e. the potential energy landscapes, were constructed using GMIN[230], OPTIM[231] and PATHSAMPLE[232] programs with an interface to the molecular dynamics package AMBER[235,236]. All calculations were performed with an ff14SB force field[237], an implicit generalized Born solvent model (igb = 8)[241] and the Debye-Hückel approximation for salt (0.1 M)[242]. These parameters were chosen based on the previous study of amyloid-β aggregation[255].

The workflow employed in constructing each energy landscape is illustrated in Figure 4.2. Four initial input structures of GLP-1 were generated to seed four independent runs of the basin hopping algorithm[281–284] to enable a better scan through the conformational space of the peptide. Three starting structures were generated from the peptide sequence by the antechamber program and two of these were run through the first round of basin-hopping algorithm comprising of 60 000 steps, the fourth structure was obtained from the Protein Databank (ID: 5OTU)[160,161] and modified to correspond to the GLP-1(7-37) sequence. This resulted in four initial structures (*de novo* generated extended chain, α-helical and two β-structure conformations) which were adapted for each protonation state. Each input structure was used as the starting point in a basin-hopping run comprising of 60 000 steps. The sampling of each of four sub-databases was based on 200 of the lowest energy structures found by the basin-hopping algorithm. Discrete path sampling[244,285] was then utilized to further grow the database, as well as systematically identify transition states connecting the minima. The sub-databases were expanded to contain at least 12 000 local minima and then connected into one large database *via* discrete path sampling, creating a large kinetic transition network[245,246] of local minima connected by the transition states.

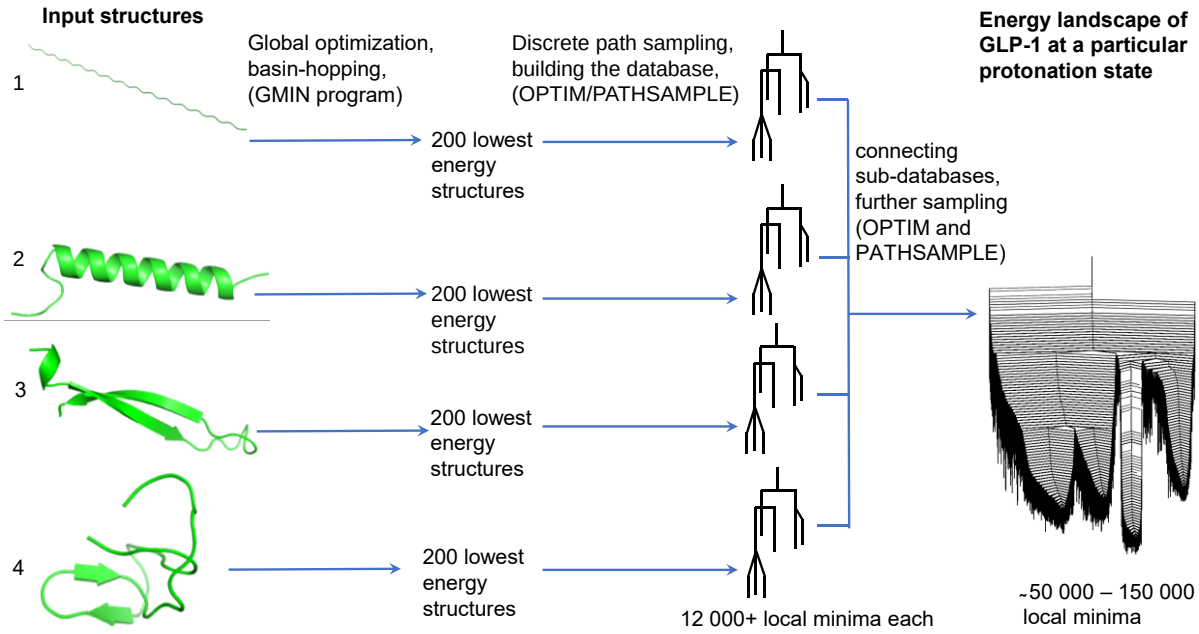


Figure 4.2: Schematic illustration of the construction of a potential energy landscape for GLP-1. Four separate basin-hopping runs were initiated with four input structures of GLP-1 (linear chain, α-helix and two different β-sheet structures). After 60 000 steps of the basin-hopping algorithm, 200 of the lowest energy structures were selected and used as a basis to form a transition kinetic network. Using discrete path sampling, the local minima were connected through the transition states and the database was further sampled. When each sub-database contained at least 12 000 local minima, the sub-databases were connected into one large database representing the potential energy landscape of one protonation state of GLP-1.

4.4 Energy landscapes of monomeric GLP-1 at different protonation states

To explore the conformational behaviour of monomeric GLP-1 at different pH values and determine the relative stabilities of different ensembles of structures, the energy landscapes of GLP-1 monomer at different protonation states were investigated.

In total, three separate energy landscapes were generated. The protonation states studied approximately correspond to pH 3, 4 and 7.5, see Table 4.1. The protonation states corresponding to pH 3 and 4 differ by the protonation state of the Asp9 side chain. The third protonation state, corresponding to approx. pH 7.5, further differs by the protonation state of the His1 side chain. The potential energy landscape of each protonation state was constructed as described in the previous section.

Approximate pH	Sequence
3	HIP ALA GLH GLY THR PHE THR SER ASH VAL SER SER TYR LEU GLH GLY GLN ALA ALA LYS GLH PHE ILE ALA TRP LEU VAL LYS GLY ARG GLY
4	HIP ALA GLH GLY THR PHE THR SER ASP VAL SER SER TYR LEU GLH GLY GLN ALA ALA LYS GLH PHE ILE ALA TRP LEU VAL LYS GLY ARG GLY
7.5	HIE ALA GLU GLY THR PHE THR SER ASP VAL SER SER TYR LEU GLU GLY GLN ALA ALA LYS GLU PHE ILE ALA TRP LEU VAL LYS GLY ARG GLY

Table 4.1: Protonation states of GLP-1. Three-letter abbreviations of residues indicating the protonation state following the convention developed by the AMBER package. Here, histidine is represented by HIP if its side chain is protonated (i.e. hydrogen atoms are on both delta and epsilon nitrogen of the imidazole ring). HIE represents the histidine side chain in its deprotonated state where a hydrogen is present only on the epsilon nitrogen. GLU represents the deprotonated form of glutamic acid whereas GLH stands for the protonated form of the carboxylate group. Similarly, ASP stands for the deprotonated side chain, whereas ASH represent the protonated form. Changes in the protonation state compared to that at pH 3 are highlighted in red.

4.4.1 Potential energy landscapes: secondary and tertiary structures

The disconnectivity graphs for the potential energy landscapes (PELs) of three different protonation states of monomeric GLP-1 are depicted in Figure 4.3 and the number of stationary points for each PEL is given in Table 4.2. All three energy landscapes exhibit a multi-funnel character which is typical of intrinsically disordered proteins or short peptides which can adopt multiple conformations[274,278,286–288]. A representative structure of the ensemble at the bottom of each funnel is shown. In all protonation states, the lowest energy minima contain predominantly α-helical structures. At pH 4, the α-helical structure is "broken" forming two shorter helices, whereas for protonation states corresponding to pH 3 and 7.5, Figure 4.3A and C, the deepest and lowest-energy funnel belongs to an ensemble of structures with a single long helix,

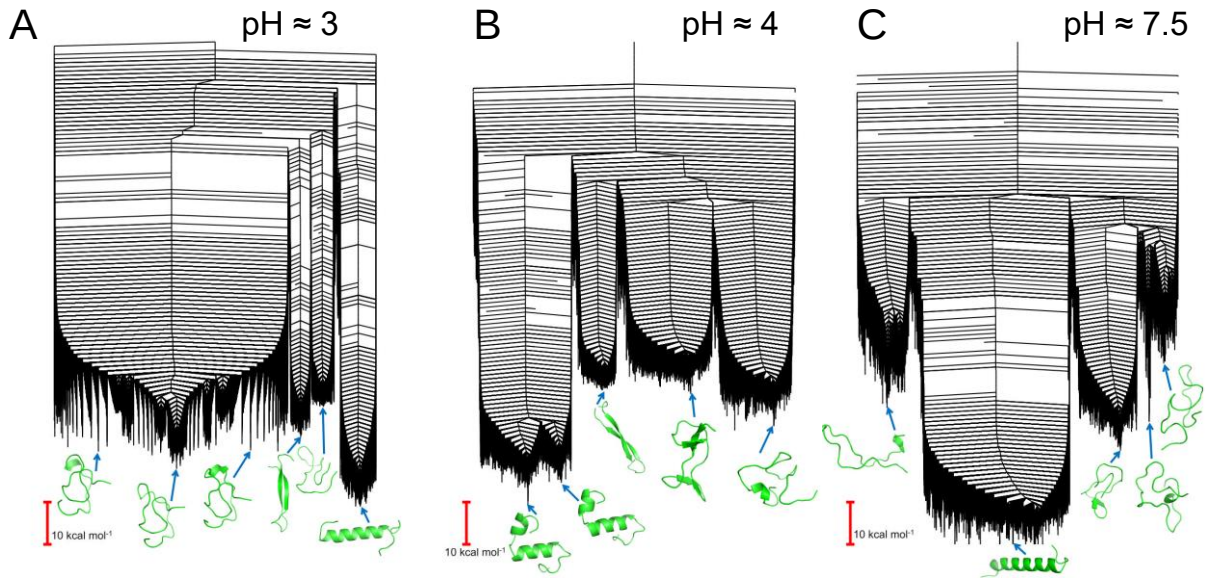


Figure 4.3: Potential energy landscapes of GLP-1 monomer at different protonation states. A representative structure corresponding to the ensemble of structures in each minimum are shown below each funnel.

PE landscape	Number of minima	Number of transition states
A (pH ≈ 3)	145 113	195 147
B (pH ≈ 4)	70 133	99 884
C (pH ≈ 7.5)	51 093	68 717

Table 4.2: Number of stationary points in GLP-1 potential energy landscapes.

The lowest energy α-helical structures at each pH value are visualized in a greater level of detail in Figure 4.4. The residues found in the α-helical region are highlighted within the GLP-1 sequence in grey. The α-helical content in the structure was analysed by the cpptraj program[256] and the percentages of α-helix in each structure shown in Figure 4.4A, B and C were 52, 48 and 68 %, respectively.

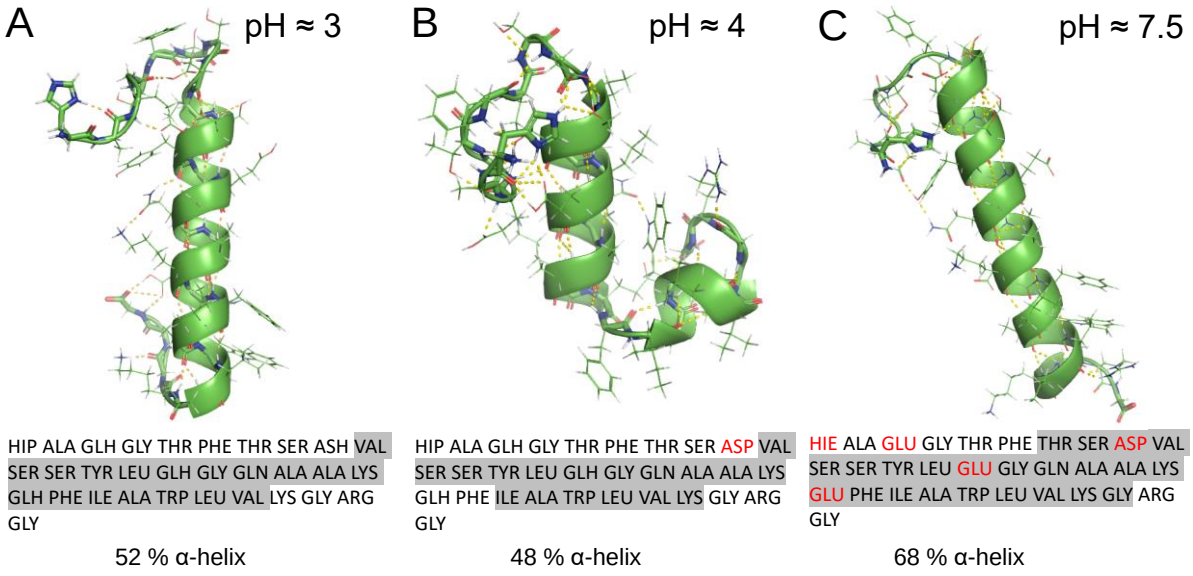


Figure 4.4: Lowest energy α-helical structures of monomeric GLP-1 at different protonation states. The structures lying in the bottom of the α-helical funnel at protonation states corresponding to pH 3 (A), 4 (B) and 7.5 (C) are visualised. Residues in the α-helical regions of the peptide are highlighted in grey. Side chains that differ in their protonation states are highlighted in red. Percentages of α-helical content of each conformation were calculated using cpptraj program[256].

Interestingly, only the energy landscape corresponding to the protonation state at around pH 4 shows a "broken" α-helical conformation consisting of two separate helices, first one in the region Val10–Lys20 and the second between Ile23 and Lys28. These two helices are separated by a short non-helical region containing Glu21 and Phe22, Figure 4.4. The only difference between the protonation state of GLP-1 corresponding to pH 3 and 4 is the protonation state of the Asp9 side chain, which becomes deprotonated at the higher pH. This deprotonation changes the hydrogen-bonding pattern in the N-terminal region of α-helical GLP-1 conformation as the deprotonated side chain of Asp9 forms a hydrogen bond with the side chain of His1 under these conditions. This is likely to contribute towards the greater stability of the more compact structure observed in the N-terminal region which interacts with the surface of the first helical region (Val10–Lys20).

Figure 4.5 shows the side chain-main chain hydrogen bonds which are formed within the "broken" helix structure and which are thought to contribute to the stability of the lowest energy structure formed at pH 4. There are two main low-energy structural ensembles of a "broken" α-helix – Figure 4.5A and C. In the first ensemble (Figure 4.5A), a hydrogen bond between Tyr13 side chain and the carbonyl group of Arg30 is formed. The free energy landscape at a temperature of 310 K in Figure 4.5B shows a distance between the hydrogen of the hydroxy group of Tyr13 side chain and carbonyl oxygen of Arg30. As apparent from the free energy landscape, conformations with a short distance between these two atoms,

which enable the formation of a hydrogen bond, are abundant and occupy the bottom parts of the α-helical funnel. The second ensemble consists of structures forming a hydrogen bond between the hydrogen attached to the nitrogen in the side chain of Trp25 and carbonyl oxygen of Leu14, Figure 4.5C. The free energy landscape in Figure 4.5D shows that conformations containing this hydrogen bond are not as abundant as those with the Tyr13–Arg30 hydrogen bond but are nevertheless present among the lowest energy minima. Due to the similarity in energies of both ensembles represented by structures in Figure 4.5A and C, both populations are likely to co-exist.

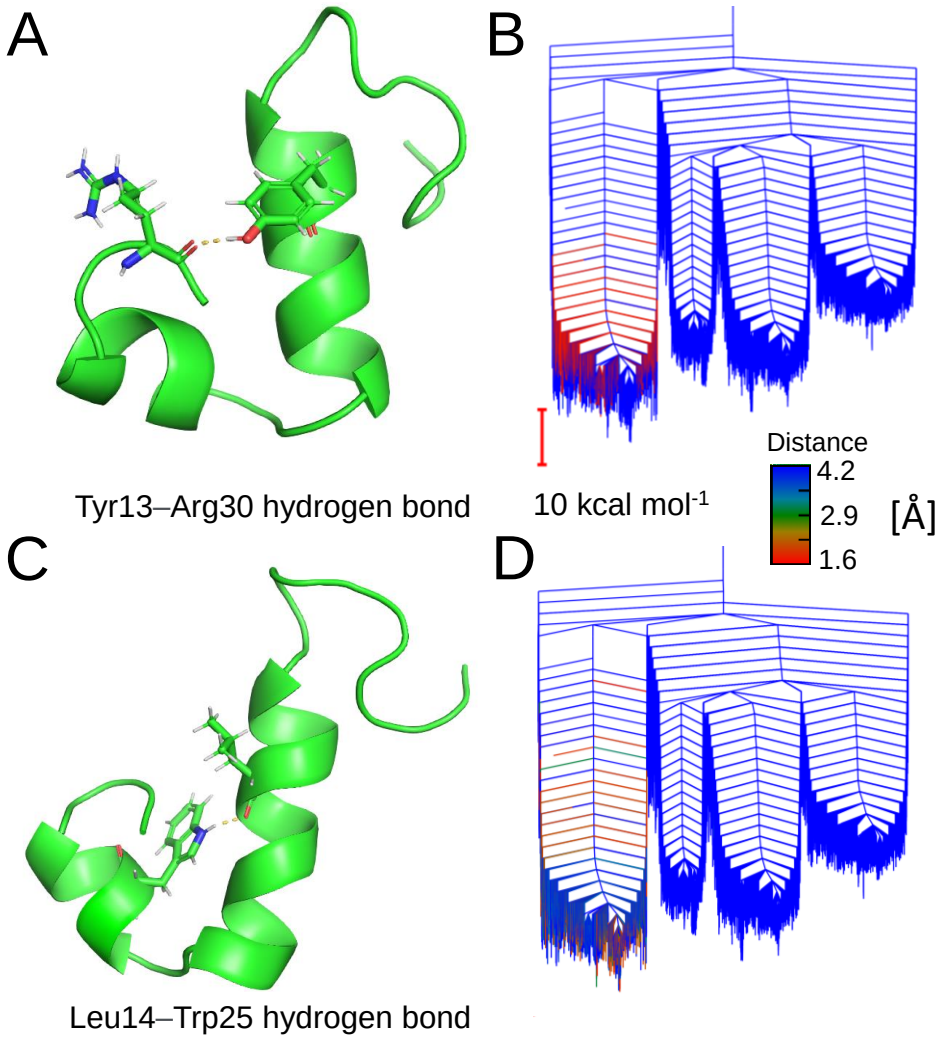


Figure 4.5: Hydrogen bonds stabilizing a broken α-helical conformation of GLP-1 monomer at pH 4. The distance between the atoms forming the hydrogen bond in Tyr13–Arg30 (A) is visualized in the free energy landscape (B) (T = 310 K) using a colour scale. The distance between the atoms forming the hydrogen bond in Leu14–Trp25 (C) is visualized in the free energy landscape (D) (T = 310 K) using the same colour scale. The scale bar of 10 kcal mol is relevant to both energy landscapes in B and D.

Structural changes in GLP-1 between pH 3, 4 and 7.5 have also been observed experimentally. Figure 4.6A shows the far-UV CD spectra of freshly prepared samples of

100 µM GLP-1 at pH 3, 4 and 7.5. Whereas at pH 3 and 7.5, the far-UV CD spectra of GLP-1 are similar, there is a visible difference in the α-helical content at pH 4 as shown by the change in far-UV CD signal at 222 nm. This observation correlates with the "broken" α-helical conformation predicted by the energy landscape at pH 4, Figure 4.4B. In addition to the information on secondary structure provided by far-UV CD, the experimentally determined fluorescence emission spectra of the single tryptophan residue provides information on tertiary structure at different pH values, Figure 4.6B. The wavelength at which the fluorescence is at a maximum (λ_{max}) reports on the local environment of the tryptophan residue – the λ_{max} is blue-shifted the more hydrophobic its local environment. At pH 3 and 7.5, λ_{max} is at 355 nm, whereas at pH 4, λ_{max} is slightly blue-shifted towards 352 nm. This is in agreement with the structures observed in the energy landscape at pH 4, Figure 4.3, where, due to the break in the α-helix, the GLP-1 conformation is more compact and Trp25 side chain interacts with other structured regions and can therefore be considered partially buried (Figure 4.4).

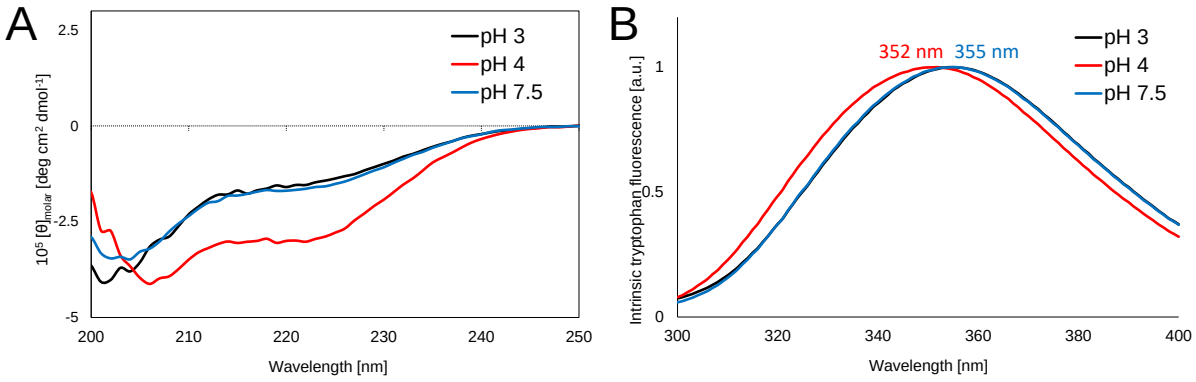


Figure 4.6: Far-UV CD and intrinsic tryptophan fluorescence of freshly prepared samples of GLP-1 at pH 3, 4 and 7.5. Samples were freshly prepared at 100 µM peptide concentration in 25 mM citrate at pH 3 and 4 or 25 mM phosphate at pH 7.5. Fluorescence spectra (B) were recorded after an excitation at 280 nm. Both experiments were performed by Jack Barber.

4.4.3 Free energy landscapes

The free energy landscapes (FELs) were derived from the potential energy landscape using the harmonic superposition approach[289] to calculate the vibrational contribution to the partition function. The free energy surface is temperature dependent, the energy landscapes in this Chapter were computed at a temperature of 310 K. Figure 4.7 shows the free energy landscape (at T = 310 K) of GLP-1 monomer protonation states corresponding to the PELs in Figure 4.4. The total number of stationary points (minima and transition states) of each FEL is given in Table 4.3. The distinct individual funnels from PELs are also preserved in

the FELs. Consistent with the PELs, the deepest (i.e., the lowest energy) funnels contain α-helical structures. In the energy landscape corresponding to pH 7.5, the difference in energy between the bottom of the α-helical funnel and the bottom of other funnels is around 20 kcal mol^{-1}, whereas for FELs at pH 3 and 4 this difference is lower than 10 kcal mol^{-1}. However, the funnels on these two FELs are still separated by high-energy barriers. The difference in the relative stabilities of minima may correlate with the aggregation propensities of GLP-1 determined at different pH values as discussed later in the text.

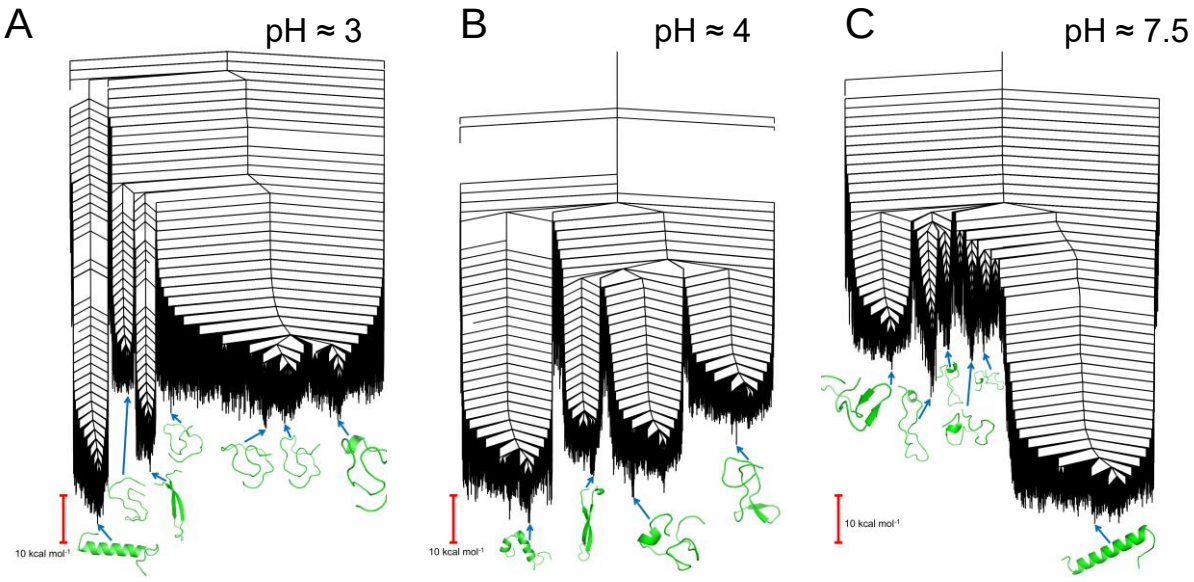


Figure 4.7: Free energy landscapes of monomeric GLP-1 in different protonation states. Free energy landscapes constructed for T = 310 K are derived from the potential energy landscapes shown in Figure 4.3. A representative structure of each of the minima is shown below each funnel.

FE landscape	number of minima	Number of transition states
A (pH ≈ 3)	143 847	183 643
B (pH ≈ 4)	69 496	95 080
C (pH ≈ 7.5)	50 490	63 924

Table 4.3: Number of stationary points in GLP-1 free energy landscapes.

Figure 4.8 summarizes some of the properties of structures obtained throughout the FELs including radius of gyration, α-helical and β-sheet content. The radius of gyration, a measure of peptide compactness, is shown in Figure 4.8A, B and C, whilst the secondary structure is depicted in Figure 4.8D, E, F (α-helix) and G, H, I (β-sheet). The funnels containing the least compact structures are the α-helical funnels at pH ≈ 3 and 7.5, and β-sheet-rich funnels. The most collapsed structures are in the random coil funnels. As expected, the α-helical funnel at pH 4 contains more compact structures due to the "breakage" in the α-helix and more compact N-terminal region of GLP-1. The FELs of protonation states which correspond to acidic pH values (pH 3 and 4) also contain funnels with high β-sheet content, whereas overall the FEL at pH 7.5 contains a lower percentage of β-sheet containing structures, which are present only in the high-energy funnels. Additionally, the FEL corresponding to a neutral pH contains a funnel with the highest overall content of α-helix observed in any of the FELs.

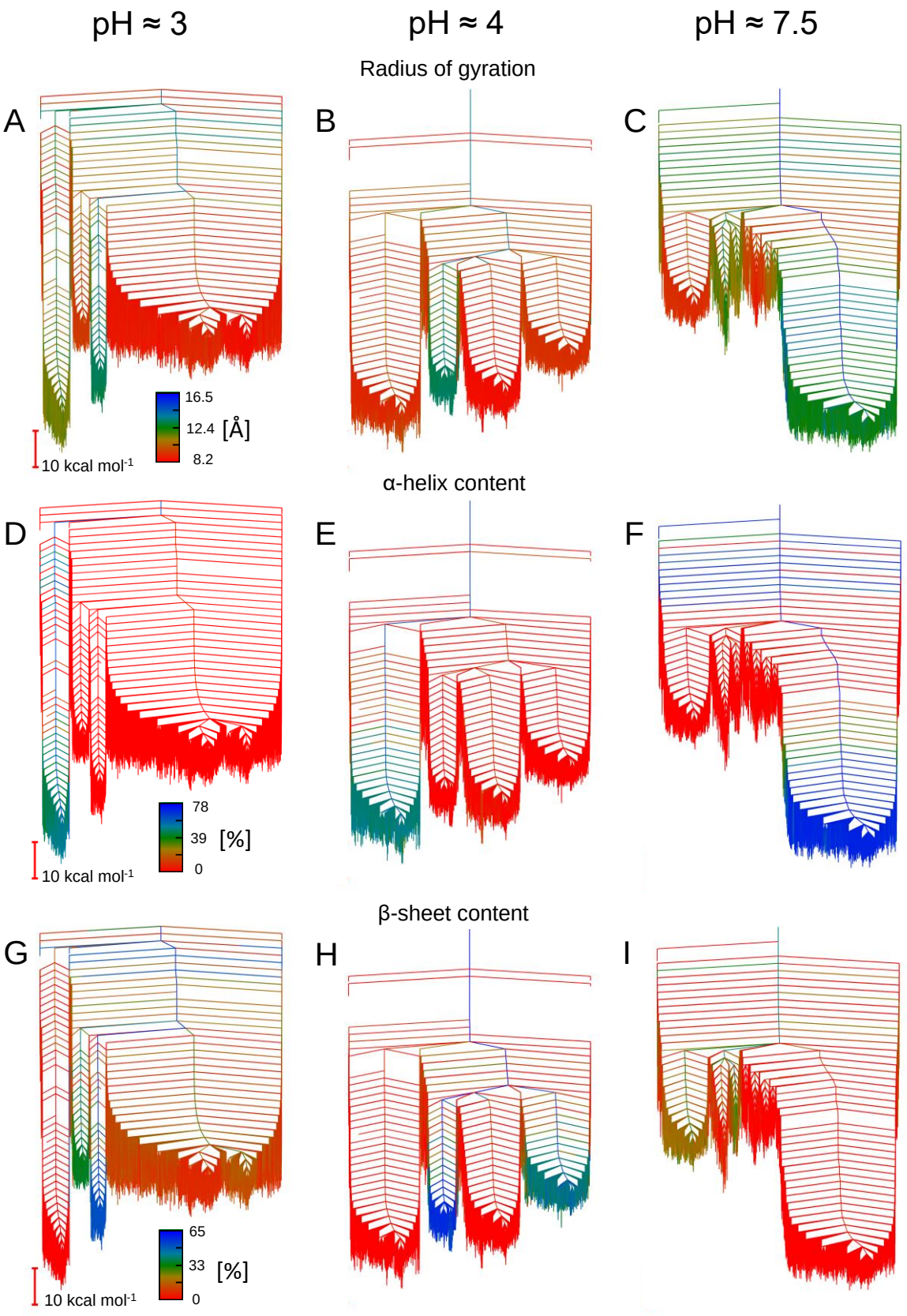


Figure 4.8: Analysis of the radius of gyration, α-helical and β-sheet content of structure in the free energy landscapes of monomeric GLP-1. pH 3 (A, D and G), 4 (B, E and H) and 7.5. (C, F and I). FELs were constructed for T = 310 K. The structural parameters of the stationary points in the FELs were obtained using the cpptraj program[256].

To correlate the trends observed in the FELs of monomeric GLP-1 at different protonation states with the aggregation propensity of GLP-1, a thioflavin T assay was performed to probe the rate of formation of β-sheet-rich aggregates. The assays, in which samples were incubated at 37 °C with agitation for 6 days, were conducted with 25–100 μM GLP-1 at pH 3, 4 and 7.5, Figure 4.9. GLP-1 shows a greater propensity for aggregation at pH 3 and 4 compared to pH 7.5 which is apparent from the shorter lag times observed in the sigmoidal aggregation curves. These experimental observations correlate with the trends observed in the FELs as more β-sheet-rich structures are present in the energy landscapes corresponding to acidic pH values. In contrast, the FEL at a protonation state corresponding to approx. pH 7.5, where aggregation is slower, shows funnels with lower content of β-sheet. Since the initial input structures employed in the basin-hopping have an identical backbone for all protonation states, the lower content of β-sheet in the final FEL at pH ≈ 7.5 points towards a lower stability of β-sheet in this protonation state. Moreover, in the FELs calculated at protonation states corresponding to acidic pH values, the β-sheet funnels are closer in energy to the α-helical funnel than they are in the FEL determined at pH ≈ 7.5.

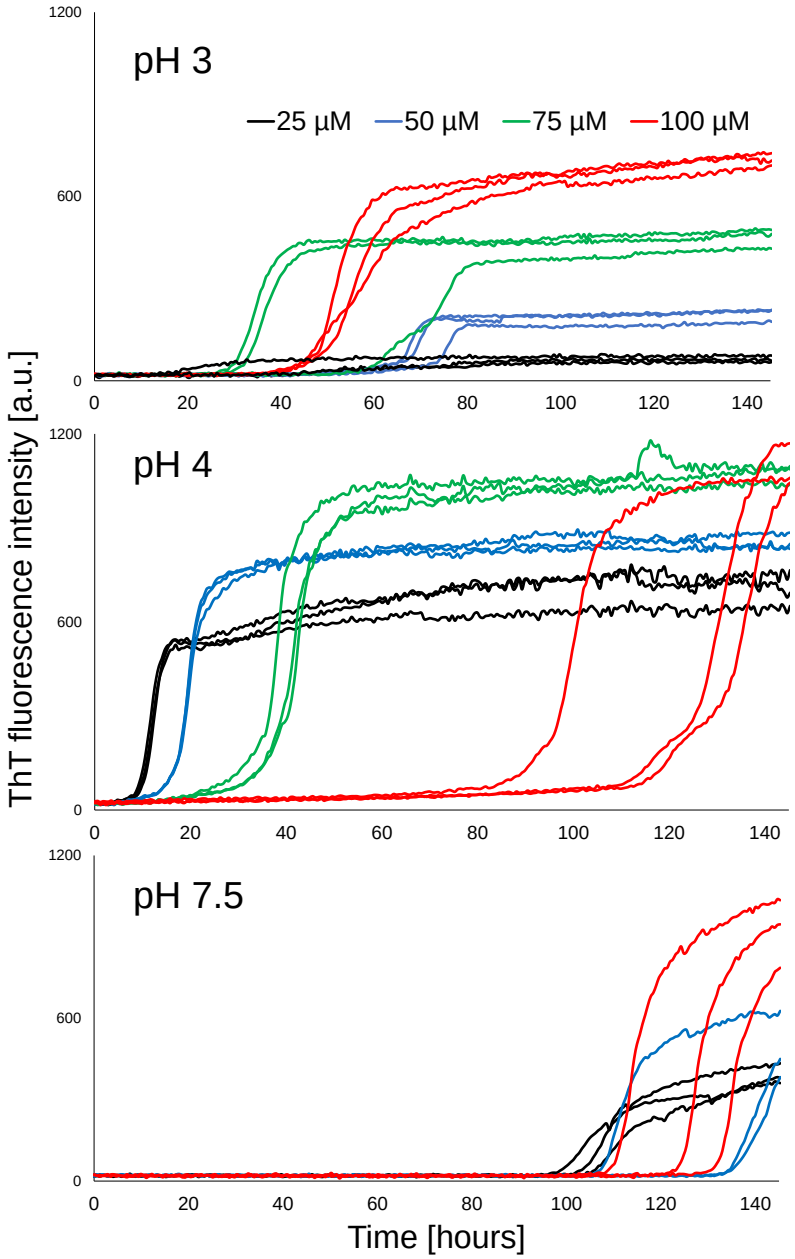


Figure 4.9: Thioflavin T assays of GLP-1 at pH 3, 4 and 7.5. GLP-1 samples were prepared in 25 mM citrate at pH 3 and 4 and in 25 mM phosphate at pH 7.5. Samples were incubated at 37 °C with agitation for 6 days. Fluorescence was recorded every 30 minus at 482 nm after excitation at 448 nm. The experiment was performed by Jack Barber.

4.5 Discussion

The potential and free energy landscapes of monomeric GLP-1 at different protonation states were investigated. The energy landscape framework was employed to rationalize the experimentally observed changes in secondary structure and aggregation propensity at different pH values.

4.5.1 Lowest energy funnels correspond to α-helical structure

In all protonation states of GLP-1 monomer studied, the lowest energy funnel (in both PELs and FELs) corresponded to α-helical structures. Using the cpptraj program[256], the α-helical content for the structures at the bottom of the lowest-energy funnel were calculated to be 52, 48 and 68 % for pH 3, 4, and 7.5, respectively. This α-helical content is significantly higher compared to experimental predictions from the far-UV circular dichroism in which the α-helical content was in a range of 10 to 33 %, Table 4.4. The overprediction of α-helicity in the energy landscape may be due to a bias introduced by the highly helical input structure and/or the force field used (ff14SB). It is interesting to note that the force field was previously reported to overpredict the α-helical regions and α-helical stability in intrinsically disordered proteins and peptides[290–292]. Additionally, the α-helical structures represent only one, albeit the deepest, funnel on the multi-funnel energy landscape, thus the contributions of other structural ensembles to the average structural properties of GLP-1 are not considered.

	BeStSel			**Dichroweb**		
	pH 3	pH 4	pH 7.5	pH 3	pH 4	pH 7.5
α-helix [%]	10	27	13	17	33	21
β-sheet [%]	20	16	19	24	10	20
Turns [%]	18	17	17	24	22	25
Disordered [%]	52	40	51	35	35	34

Table 4.4: α-helical content of GLP-1 determined from experimental far-UV circular dichroism spectra. The corresponding far-UV CD spectra are shown in **Figure 4.6**A. The content of individual secondary structure elements was predicted using DichroWeb, Contin LL method, dataset 3[180–182] and BeStSel method[183–187]. Experimental far-UV CD spectra were obtained by Jack Barber.

The experimental far-UV CD spectra reveal an increase in α-helical structure at pH 4 compared to pH 3 and 7.5, Table 4.4. Although the predicted content of α-helix from the experimental data differs based on the analysis software employed, this trend is clear in both BeStSel and Dichroweb analysis. Using the energy landscape approach, the change in α-helical conformations at pH 4 was also reflected. In this case, the helical funnel at pH ≈ 4, contains more compact structures with two shorter helical regions, Figure 4.4B. In contrast, the energy landscapes of GLP-1 at pH ≈ 3 and 7.5, the helical funnel is occupied by

structures with a single, long α-helix. The only change between the protonation states corresponding to pH ≈ 3 and 4 is the protonation state of the side chain of Asp9. Interestingly, this residue does not seem to be directly involved in the stabilization of the "broken" α-helical structure, however, the deprotonated Asp9 side chain forms a hydrogen bond with His1 side chain which stabilizes the more compact structure in the N-terminal region of the peptide and therefore can influence overall stability of the monomer conformation.

4.5.2 Free energy landscapes correlate with the aggregation propensity

As reported previously, the aggregation of GLP-1 is pH-dependent and even subtle changes in pH can lead to a change in the aggregation kinetics and even the aggregation mechanism[36]. Using a thioflavin T assay, it was experimentally shown that aggregation, specifically amyloid fibril formation, is faster at acidic pH compared to neutral pH, Figure 4.9. In agreement with these observations, the FELs of monomeric GLP-1 at pH ≈ 3 and 4 showed funnels with higher β-sheet content that were closer in energy to the deepest α-helical funnel, Figure 4.8G and H. In the FEL corresponding to the GLP-1 protonation at pH 7.5, less β-sheet-rich structures are observed and the funnels with some content of β-sheet structure are much less energetically favourable, Figure 4.8I.

The β-sheet-rich funnels observed under acidic conditions are still separated by high energy barriers, therefore, any α→β structural transition probably does not occur readily in monomeric GLP-1. One possible consequence of this is that such a conformational change which is necessary to get the β-sheet structure found in amyloid fibers may occur in an oligomer. However, the greater stability of monomeric β-sheet-rich structure may still indicate better feasibility of α→β structure conversion even if the conversion itself occurs in the oligomer.

4.5.3 Limitations of the approach

The energy landscapes framework provides a comprehensive way to study biomolecular energy landscapes. This framework provides the insight into the structure, thermodynamic properties and global dynamics of the system[293]. However, this approach still faces some limitations in capturing the energy landscape of complex biomolecular systems. First of these is locating the global minimum on the energy landscape. Although compared to conventional molecular dynamics, the energy landscape framework offers the possibility of a much more extended conformational search due to the basin-hopping algorithm, it may still not be sufficient to access all structures/state. This is especially true in the case of large and complex biomolecules, if the initial structural information and starting points are limited. In this case, the system may still become trapped in local minima and never find the global minimum. This is likely the case with the α-helical ensembles of GLP-1 monomer. As one of the initial input structures, corresponding to the structure of GLP-1 bound to its receptor, was highly helical, the local minima after the basin-hopping run still showed a high content of α-helix compared to the experimental far-UV CD results, Table 4.4. The limited conformational search is likely to be caused by high-energy transition structures which are rejected by the algorithm.

The other limitation is given by the nature of the system – GLP-1 in solution. As was shown experimentally (Figure 4.10), GLP-1 in solution is not only in a monomeric form but a large percentage oligomerize into dimers and even higher-order oligomers. Thus, the energy landscape of monomeric GLP-1 is only an approximation and cannot reflect all the features of the system.

Despite of all the above-mentioned limitations, the observed trends in PELs and FELs may still be important and reflect what is really going on in solution as the energy landscapes of different protonation states of GLP-1 were constructed based on the same structural inputs, force fields and other parameters.

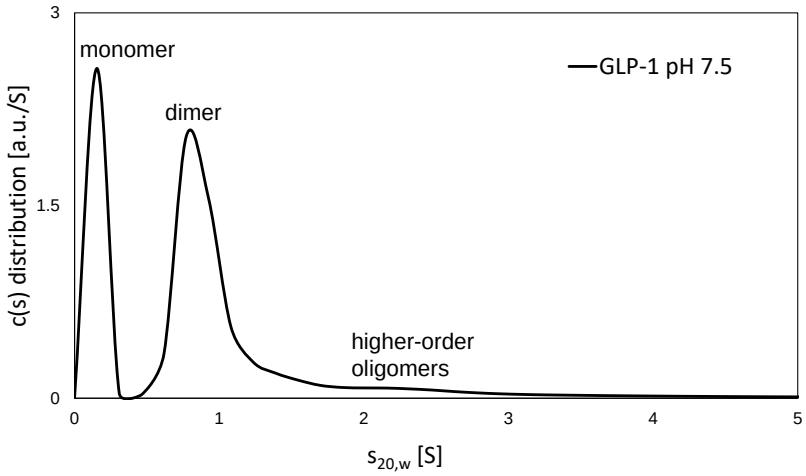


Figure 4.10: Sedimentation velocity experiment for freshly prepared GLP-1 sample at pH 7.5. The sample was freshly prepared at 85 µM peptide in 25 mM phosphate at pH 7.5. The interference sedimentation curves were fitted to a continuous c(*s*) distribution model implemented in the Sedfit program[220,221]. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$).

CHAPTER 5
EFFECT OF LIPIDATION ON THE BIOPHYSICAL PROPERTIES OF GLP-1-AM

This Chapter describes the changes in biophysical properties of GLP-1-Am analogues induced by lipidation in solid samples or in the samples freshly dissolved in aqueous buffer. The longer-term physical stability and aggregation processes of lipidated GLP-1-Am analogues are studied in Chapter 6.

5.1 Introduction

Many peptide-based biopharmaceuticals suffer from short half-lives *in vivo*[294]. For instance, native GLP-1 has a half-life *in vivo* of *circa* 2 minutes, which is caused by the rapid degradation by the dipeptidyl peptidase-4 enzyme[295,296]. Several types of modifications of peptide-based biopharmaceuticals have been developed to increase stability and proteolytic resistance of these molecules and extend their half-life *in vivo*[16]. These half-life extension strategies include peptide stapling[297], incorporation of non-natural amino-acid residues[298,299], covalent conjugation to larger proteins[300,301], covalent attachment of a polyethylene glycol chain (PEGylation)[302] or fatty acid chain (lipidation)[294,303,304] or a combination of these approaches[22].

Peptide lipidation was proven to be an effective strategy with several marketed peptide-based drugs such as liraglutide[305], semaglutide[22], long-acting insulin detemir[306] or somapacitan[307]. Mechanisms underlying the half-life extension of lipidated therapeutic peptides *in vivo* include an increased tendency to self-assemble into larger oligomers as well

as binding to human serum albumin. These mechanisms reduce the enzymatic degradation of therapeutic peptides and slow their renal clearance.[3,16,39,49,294,303,308]

5.2 Aims

This Chapter describes studies aimed at understanding the effect of lipidation on the biophysical properties of GLP-1 analogues. Specifically, the effect of lipidation on peptide solubility, secondary structure and self-assembly behaviour was investigated. Characterization of the biophysical properties and physical stability of peptide-based drugs in solution is a critical step in the drug-development process as it is related not only to drug activity *in vivo*, but also the physical stability of the drug which has to be considered during the manufacturing, storage and formulation of the drug. Four lipidated GLP-1 analogues differing only in the position of lipidation in the peptide sequence were studied to assess the role of position of lipidation on the peptide properties.

Results showed that the biophysical properties and the tendency for rapid formation of stable oligomeric species of the studied analogues closely correlate with the longer-term physical stability and propensity for aggregation.

In this Thesis, the term “aggregation” is used for a self-assembly process which includes changes in secondary structure of the peptide, which are often induced by prolonged incubation. In contrast, the term “self-assembly” in this Thesis refers to the formation of higher molecular weight species which is not accompanied by a change in secondary structure of the peptide.

5.3 Lipidated variants of GLP-1-Am

GLP-1-Am is a C-terminally amidated variant of GLP-1(7−37). Lipidated analogues of GLP-1-Am studied in this Chapter differ mainly by their position of lipidation − Figure 5.1. The IPP2 analogue is lipidated close to the N-terminus at position 2, whereas IPP3, IPP4 and IPP5 are lipidated in the central region of the peptide, at positions 12, 17 and 20, respectively. Lipidated analogues also contain lysine-to-arginine substitutions at positions 20 and 28. In addition, IPP2 has another substitution leucine-to-arginine at position 26. Due to very different biophysical properties and rapid aggregation, IPP2 analogue is studied in more detail in the following Chapter 6.

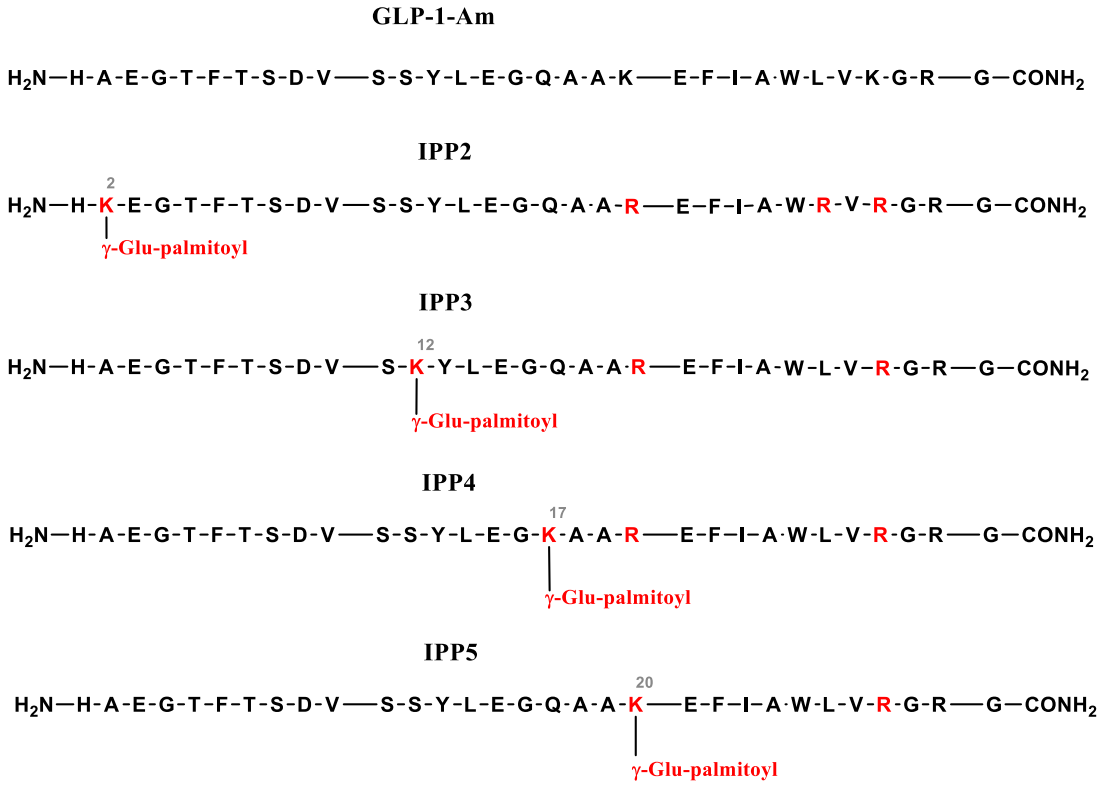


Figure 5.1: Sequences of GLP-1-Am and its lipidated analogues. All sequences have a free N-terminus and a C-terminal amidation. For lipidated variants, the position of the lipidation site is numbered and highlighted in red. Additionally, the substitution sites relative to GLP-1-Am are also highlighted in red.

All lipidated analogues have the same lipidation – a palmitic acid is attached through the γ-glutamic acid to a lysine side chain. The detailed structure of the lipid and the linker attaching it to GLP-1-Am is depicted in Figure 5.2.

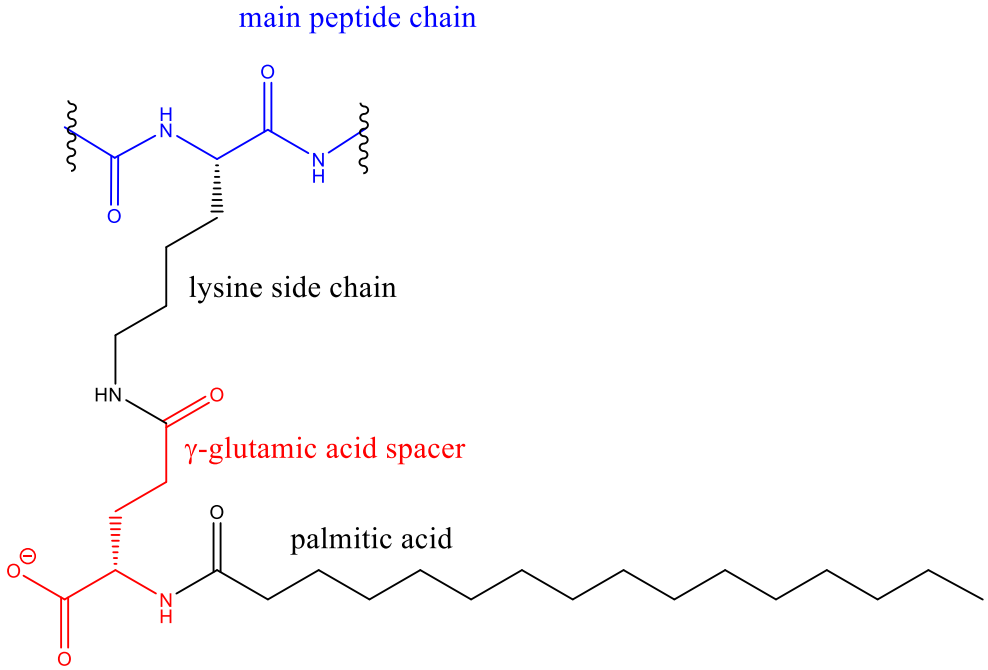


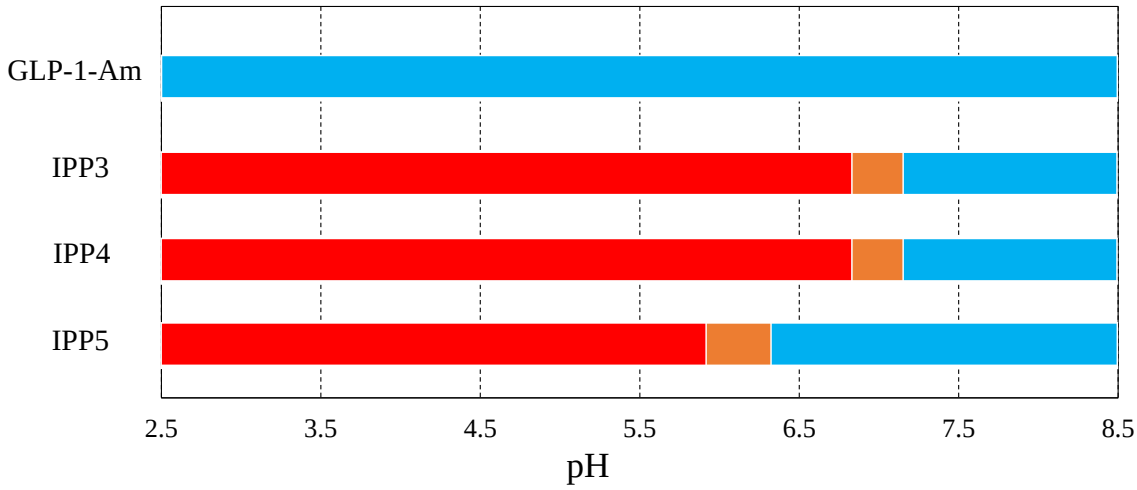
Figure 5.2: Detailed structure of the lipidation and attachment site. Palmitic acid (black) is attached to the peptide chain (blue) through the γ-glutamic acid spacer (red) and lysine side chain (black).

5.4 Lipidation limits the solubility of GLP-1-Am variants

The solubility of non-lipidated GLP-1-Am and its lipidated analogues was tested over a pH range from 2.5 to 8.5. A solution of buffer (25 mM phosphate, citrate or Tris) of a corresponding pH was added to the lyophilized peptide powder, gently mixed and left for *circa* 5 minutes at room temperature before the solution was filtered through a 0.22 µm filter (Millex, PVDF Membrane). Subsequently, the concentration of the peptide was measured using absorption at 280 nm. Figure 5.3 illustrates the solubility of each lipidated analogue and non-lipidated GLP-1-Am at different pH values – pH conditions in which a peptide was soluble are indicated in blue (peptide concentration > 50 µM), pH values with limited solubility in orange and pH conditions in which the peptide was insoluble in red (peptide concentration < 1 µM). A "limited solubility" stands for the peptide concentration between 1 and 50 µM.

Whereas non-lipidated GLP-1-Am is soluble at all pH values tested, the solubility of lipidated analogues is pH-limited. IPP3, IPP4, and IPP5 which are soluble at neutral and

basic pH values. IPP5 has the widest pH range of solubility of all the tested lipidated analogues – it is soluble from pH 6 to pH 8.5.



Red: insoluble
Orange: limited solubility (maximum concentration < 50 µM)
Blue: soluble

Figure 5.3: pH-dependent solubility of lipidated GLP-1-Am analogues. The solubility in a buffer of a corresponding pH was determined from an absorption measurement after sample filtration through a 0.22 µm membrane filter. Concentrations were determined spectrophotometrically. The peptide under specific conditions is defined as "soluble" if it the concentrations above 50 µM can be achieved. The peptide is defined as "insoluble" if it is not possible to attain concentrations above 1 µM. "Limited solubility" stands for concentrations from 1 to 50 µM.

To rationalize the solubility of the lipidated analogues with their net charge, the isoelectric point (pI) of each lipidated analogue was determined using an isoelectric focusing gel pH 3–10 (Invitrogen, Thermo Fisher Scientific) – Figure 5.4A. All isoelectric points were in the range of 5–6, with the pI of IPP5 being the lowest at around 5.5, followed by IPP3 with pI ≈ 5.7 and IPP4 with pI ≈ 5.9. For IPP4 and IPP5, pI values were also measured using capillary isoelectric focusing (Figure 5.4B) and similar results obtained – 6.1 for IPP4 and ≈ 5.5 for IPP5.[225] It is interesting to note that, IPP3, IPP4 and IPP5 are soluble only when their net charge is negative, i.e., at a pH above their pI values. The widest solubility range of IPP5, correlates with it having the lowest pI (5.5) of all the measured analogues.

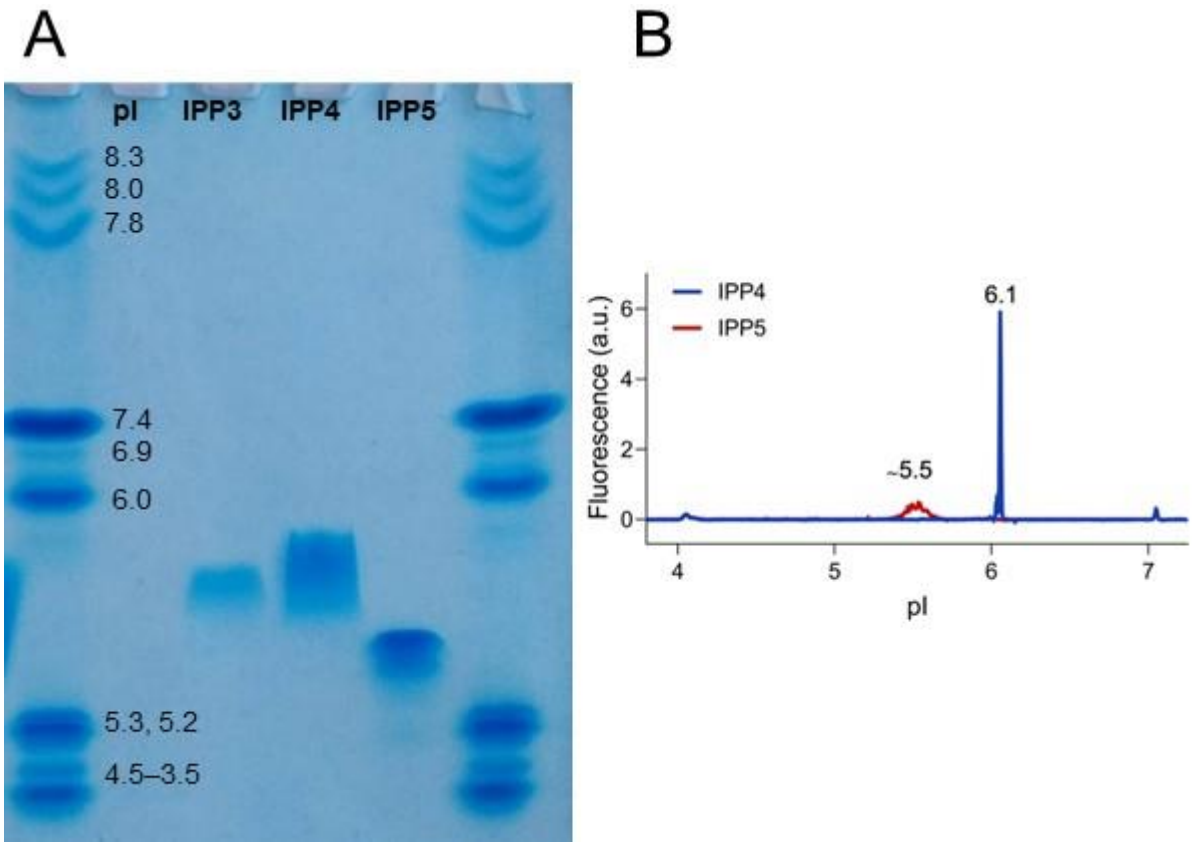


Figure 5.4: Determination of the isoelectric point values (pI) of lipidated GLP-1-Am analogues. pI of lipidated analogues was determined using an isoelectric focusing gel pH 3–10 (Invitrogen, Thermo Fisher Scientific) (A), pI of IPP4 and IPP5 was also determined using imaged capillary isoelectric focusing (B). Capillary isoelectric focusing experiments were performed by Frederik Becher at Astra Zeneca.[225]

5.5 Effect of lipidation on peptide structure and oligomerization in solution

The effect of lipidation on the structure and oligomerization of GLP-1-Am was investigated mainly at pH values close to physiological conditions. Experiments in this Section were performed at pH 7.5 in 25 mM sodium phosphate (ionic strength 65 mM) unless specified otherwise.

This Section focuses on two questions – firstly, to what extent is the secondary structure of GLP-1-Am peptide chain affected by the attachment of palmitic acid via a γ-glutamic acid linker and secondly, what effect has the lipidation on peptide self-assembly behaviour.

5.5.1 Effect of lipidation on secondary structure (non-aged samples)

First, the effect of the lipidation itself and its position on the secondary structure of the peptide was analysed for solid samples of the lyophilized peptide powder. Infrared spectroscopy in the amide I and amide II regions (range between *circa* 1350 and 1800 cm^{-1}) was employed for this purpose. Figure 5.5A depicts attenuated total reflection (ATR) FT-IR spectra of the lyophilized powder of GLP-1-Am and its lipidated analogues IPP3, IPP4 and IPP5. In Figure 5.5B, the second derivate of the original spectra was plotted to identify the peak frequencies of characteristic components and thus to allow more detailed qualitative comparison[309]. The FT-IR spectra show rather smaller effect of lipidation on the peptide secondary structure in the solid phase. The maximum of amide I absorption lies in the range from 1647 to 1654 cm which corresponds to α-helical and/or disordered (random coil) structure[193,310]. The second derivative of the spectra (Figure 5.5B) indicates a dominant contribution of α-helix at 1655 cm^{-1} with a minor contribution of disordered regions at 1640 cm^{-1}, which are present mainly for GLP-1-Am and IPP4.

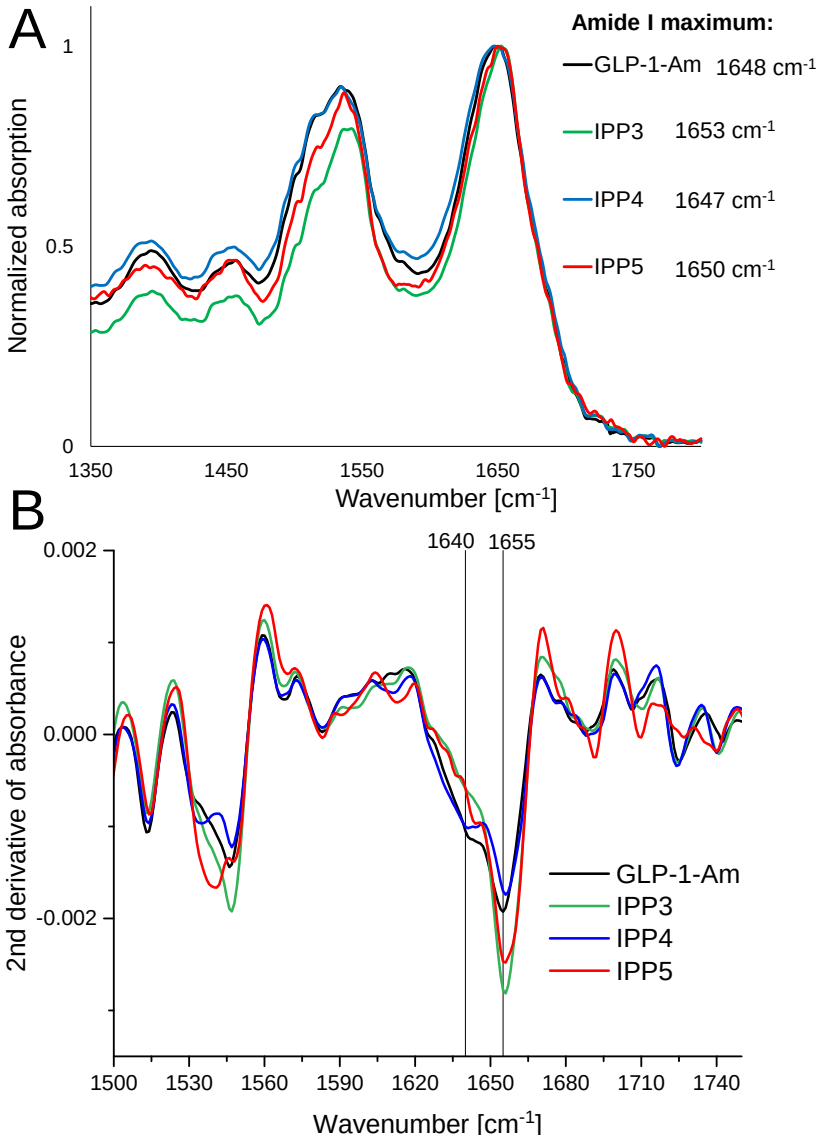


Figure 5.5: FT-IR ATR spectra and their second derivative of the amide I and amide II regions of GLP-1-Am and lipidated analogues in a powder state. FT-IR spectra of amide I and II regions (A) and their second derivatives (B). The figure legend in A highlights the amide I band maximum for each spectrum. Spectra (in A) have been normalized to the highest amide I absorption. In panel B, two vertical lines at 1640 and 1655 cm^{-1} were added to highlight the signal at corresponding frequencies.

To investigate the effect of lipidation on the secondary structure of the peptide chain for samples in aqueous solution, far-UV circular dichroism (CD) spectra were recorded and analyzed. Far-UV CD spectra of freshly prepared samples (85 µM) of all lipidated analogues and non-lipidated GLP-1-Am were measured in 25 mM phosphate, pH 7.5 (Figure 5.6). Table 5.1 summarizes the approximate percentages of different secondary structures of GLP-1-Am and its lipidated analogues as predicted from the far-UV CD spectra using three different algorithms DichroWeb, BeStSel method and estimation of α-helical content from mean residue ellipticity at 222 nm. The first two algorithms are based on the reference dataset of the far-UV CD spectra of known secondary structure (details in Section 1.7.2.1).

The latter algorithm is based on the experimentally determined values of mean residue ellipticity at 222 nm for fully α-helical and fully coiled protein (details in Section 2.11). The predicted percentages of secondary structure elements are not fully consistent between different algorithms (Table 5.1). BeStSel method and algorithm for estimation of α-helical content showed an increase in the percentage of α-helix upon lipidation. None of the algorithms employed was optimized for lipidated peptides, given values should be, therefore, taken with great caution. Here, the secondary structure estimates should be used for the indication of changes in secondary structure during aggregation (e.g., comparison to Section 6.2.1) rather than absolute values.

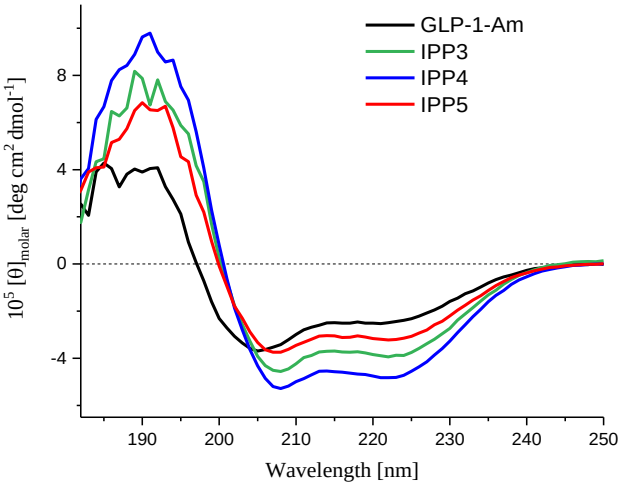


Figure 5.6: Far-UV CD spectra of freshly prepared samples of lipidated analogues of GLP-1-Am at pH 7.5. All samples were prepared at 85 µM peptide in 25 mM phosphate, pH 7.5 and measured in a 0.1 cm pathlength cuvette. The far-UV CD signal was converted into concentration independent molar ellipticity, $[\theta]_{molar}$, units, see Chapter 2, Section 2.11, Eq. 2.6.

Dichroweb				
	α-helix [%]	β-sheet [%]	turns [%]	disordered [%]
GLP-1-Am	37	15	18	30
IPP3	39	15	20	26
IPP4	47	12	17	25
IPP5	34	18	20	28
BeStSel method				
	α-helix [%]	β-sheet [%]	turns [%]	disordered [%]
GLP-1-Am	20	20	16	44
IPP3	40	16	13	31
IPP4	42	13	10	35
IPP5	48	0	33	19
α-helical content estimation based on MRE_{222}				
	α-helix [%]			
GLP-1-Am	21			
IPP3	34			
IPP4	42			
IPP5	28			

Table 5.1: Percentages of secondary structure of GLP-1-Am and its lipidated analogues predicted from far-UV CD spectra. The content of individual secondary structure elements was predicted using three different algorithms: DichroWeb, BeStSel, and α-helical content estimation based on the mean residue ellipticity at 222 nm (MRE_{222})[180–182]. Details of the employed algorithms are given in Chapter 2, Section 2.11.

Intrinsic tryptophan fluorescence emission spectra and circular dichroism in the near UV region were also measured to monitor the changes in the local environment of the single tryptophan (position 25) and the local environment of other aromatic side chains in the peptide sequence upon lipidation. Intrinsic tryptophan fluorescence spectra were recorded after excitation at 280 nm. The position of fluorescence emission maxima (λ_{max}) reflects the degree of tryptophan exposure to the aqueous solvent. The lower λ_{max}, the more buried and

less accessible to the solvent the tryptophan side chain is. Figure 5.7A shows normalized intrinsic tryptophan fluorescence emission spectra of lipidated analogues and non-lipidated GLP-1-Am. Intrinsic tryptophan fluorescence emission maxima of all lipidated analogues are lower in wavelength than λ_{max} of non-lipidated GLP-1-Am which indicates that the tryptophan residue in lipidated analogues is less exposed to the solvent compared to non-lipidated GLP-1-Am. This may be caused by the transient association of the lipid with the tryptophan side chain. In addition, as shown later in this Chapter (Figure 5.8, Figure 5.9), lipidation promotes the peptide self-assembly into larger oligomers. The observed shift in λ_{max} of lipidated analogues may also originate from the formation of these higher-order structures. IPP4 shows the greatest blue shift, 337 nm *vs* 354 nm (for non-lipidated GLP-1-Am) and shows the greatest change in secondary structure caused by lipidation as observed by far-UV CD (Figure 5.6). Figure 5.7B shows the near-UV CD spectra of GLP-1-Am and its lipidated analogues. For peptides and proteins, CD in the near-UV region reflects the restricted movement of aromatic side chains illustrating that aromatic residues are fixed within a chiral environment, e.g., in a protein core, in an oligomeric structure, or in the aggregate. The absence of any signal in this region denotes that the rotation of the side chains of aromatic residues is not restricted (i.e., the residues are not fixed in chiral environment). As apparent in Figure 5.7B, freshly prepared samples of GLP-1-Am or its lipidated analogues do not have a significant CD signal in this region. Only in the case of IPP4, is there a slight indication of a positive maximum at around 275 nm.

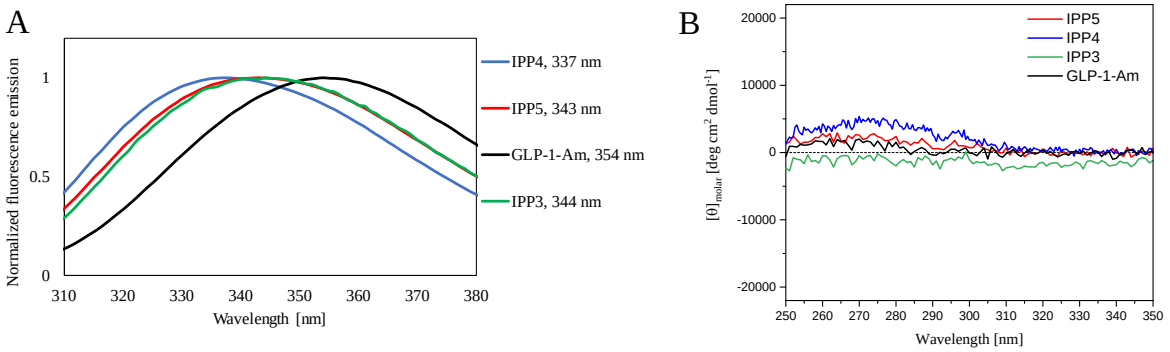


Figure 5.7: Intrinsic tryptophan fluorescence and near-UV CD spectra of freshly prepared samples of GLP-1-Am and its lipidated analogues at pH 7.5. Both intrinsic tryptophan fluorescence and near-UV CD spectra were measured at 85 µM peptide in 25 mM phosphate, pH 7.5. Intrinsic tryptophan fluorescence spectra (A) were normalized due to the different voltages across the photomultiplier tube which was used in the measurements, the λ_{max} values for the different peptides are shown on the right of the panel. Near-UV CD spectra (B) were measured in a 0.2 cm pathlength cuvette.

5.5.2 Effect of lipidation on oligomerization behaviour

Size-exclusion chromatography (SEC) and analytical ultracentrifugation (AUC) were employed to study the self-assembly of lipidated GLP-1-Am analogues. SEC separates species in solution based on their size (i.e., hydrodynamic radius). Large species do not enter the pores in the column matrix, therefore, they elute in the void volume of the column. Molecules with partial access to the pores of the matrix are separated and elute from the column in order of decreasing size. In contrast to SEC, in analytical ultracentifugation the sample is not in contact with a column matrix. Also, the sedimentation coefficient of a particle depends on the (2/3) power of the molecular weight, wheras SEC separation depends on the Stokes radius which is proportional to the (1/3) power of the molecular weight[220,311].

Figure 5.8 shows the elution profile of freshly prepared samples of IPP3, IPP4, IPP5 and GLP-1-Am at 135 µM peptide in 25 mM phosphate, pH 7.5. Samples were filtered through the 0.22 µm membrane filter prior to the SEC analysis to remove any larger insoluble aggregates. Lipidated analogues were analysed using a Superdex200 Increase 10/300 size-exclusion column with a resolution range from 10 kDa to 600 kDa (for globular proteins). Size-exclusion chromatography of all the lipidated GLP-1-Am analogues clearly revealed the presence of oligomeric species. Briefly, the IPP5 chromatogram shows multiple populations of different oligomeric species as apparent from well-resolved peaks at elution volumes around 13 and 15 mL, and a smaller peak eluting at around 18 mL. For comparison, non-lipidated GLP-1-Am, with a molecular weight of *circa* 3.36 kDa, elutes as a monomeric peak and has an elution volume of around 19.2 mL. However, the molecular weight of the GLP-1-Am monomer is already far below the recommended separation range of a Superdex200 Increase 10/300 column. In contrast to IPP5, IPP3 and IPP4 chromatograms show only single peaks which are, in both cases, broader than the peaks observed for IPP5.

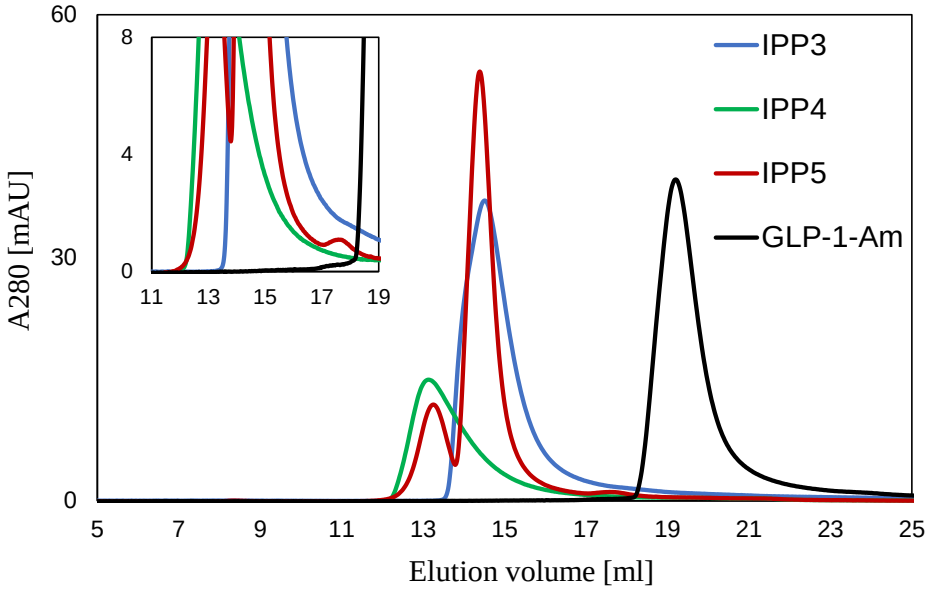


Figure 5.8: Size-exclusion chromatograms of freshly prepared samples of lipidated GLP-1-Am analogues and non-lipidated GLP-1-Am at pH 7.5. All samples were analysed at 135 µM peptide in 25 mM phosphate, pH 7.5 and filtered using a 0.22 µm membrane filter prior to the injection onto the Superdex200 Increase 10/300 column.

To estimate the size of the species observed in the SEC experiments, a column calibration was performed using a standard set of globular proteins. These proteins, with known molecular weights and Stokes radii, were dissolved in 25 mM phosphate, pH 7.5, injected onto the Superdex 200 Increase 10/300 column and run under the same buffer conditions as the GLP-1 analogues. Details about SEC column calibration are given in Chapter 2, Section 2.14, Figure 2.1.

Estimates of the molecular weight and Stokes radius of the observed species as determined from the SEC calibration are given in the first part of Table 5.2. However, it should be noted that this determination has certain limitations since the commercially available SEC calibration kits contain only globular proteins whose hydrodynamic properties may differ from those of the lipidated peptide oligomers observed.

A velocity sedimentation measurement using an analytical ultracentrifuge was employed to gain complementary information about the oligomeric populations present in solution for non-lipidated and lipidated GLP-1-Am analogues. Freshly prepared samples of 85 µM GLP-1-Am and its lipidated analogues IPP3, IPP4 and IPP5 in 25 mM phosphate, pH 7.5, were used. The interference sedimentation curves were collected as 300 scans for GLP-1-Am or 150 scans for lipidated analogues (approx. 24 h or 12 h run time, respectively) and fitted to a continuous c(*s*) distribution model implemented in a Sedfit program[220,221]. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$). Figure 5.9 shows sedimentation velocity plots from freshly prepared samples of IPP3, IPP4, IPP5 and non-lipidated GLP-1-Am. All peptides coexist in solution in multiple oligomeric states. The peak close to 0 S is likely be the monomeric peptide which is too small to sediment. Whereas for non-lipidated GLP-1-Am, the peptide is mostly in monomeric and dimeric form with small percentage of larger (*circa* hexameric) oligomers (Table 5.2), the lipidated analogues have a tendency to form the species of higher sedimentation coefficients corresponding to higher molecular weight. IPP5 was shown to coexist in mainly two oligomeric states – 8-mer and 13-mer – as determined from the sedimentation plot analysis. Sedimentation plots of IPP4 and IPP3 showed less distinct peaks. For IPP3, two interweaving peaks corresponding to *circa* 9-mer and 18-mer were determined. In the case of IPP4, the distribution of sedimentation coefficients showed a broad unresolved peak over the range of *circa* 1 – 9.5 S and two small peaks at 0 and 0.7 S. This distribution indicates a presence of trace amounts of small species (monomers, dimers) and a broad range of larger oligomers which are likely to be rapidly interconverting.

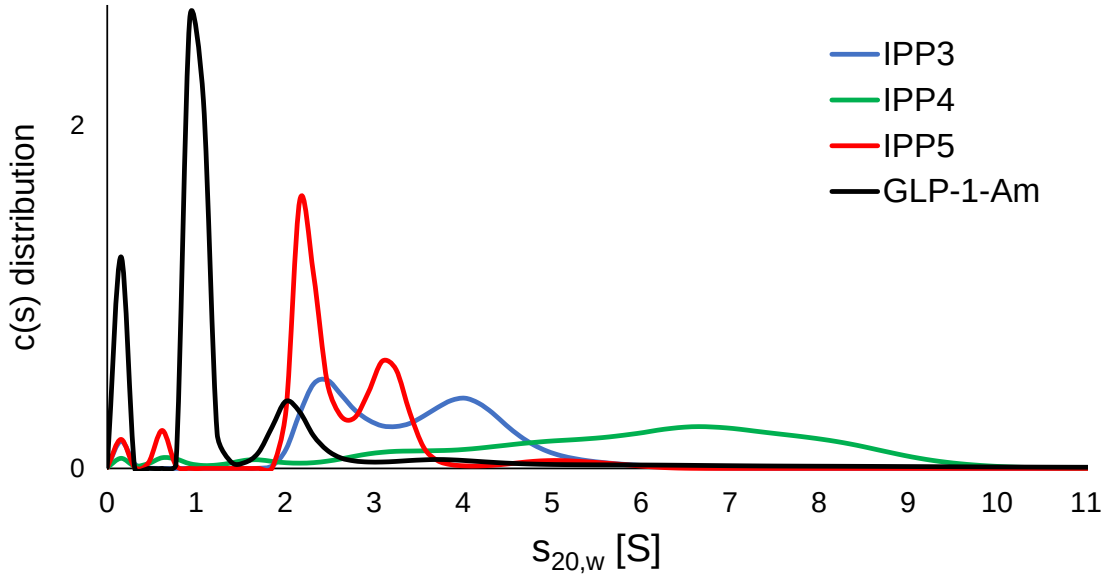


Figure 5.9: Velocity sedimentation plots of freshly prepared samples of GLP-1-Am and its lipidated analogues at pH 7.5. Samples were freshly prepared at 85 µM peptide in 25 mM phosphate at pH 7.5. After a 2-hour temperature equilibration of samples in the centrifuge to 20 °C, the experiment was performed with centrifugation at 50 000 rpm. The interference sedimentation curves were fitted to a continuous c(*s*) distribution model implemented in Sedfit[220,221]. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$).

Table 5.2 provides a summary of estimated masses and Stokes radii of detected oligomeric species for GLP-1-Am and its lipidated analogues IPP3, IPP4 and IPP5 obtained using SEC and sedimentation velocity.

Size-exclusion chromatography					
Peptide	Concentration [μM]	Peak elution volume(s) [mL]	MW estimate [kDa]	Estimated number of monomeric units in oligomer	Predicted Stokes radius [nm]
GLP-1-Am*	135	19.2	6.2	1-2	0.9
IPP3	135	14.5	50.6	13	3.1
IPP4	135	13.2	90.8	24	3.8
IPP5	135	13.2 14.4	88.9 53.2	24 14	3.7 3.1
Sedimentation velocity – analytical ultracentrifugation					
Peptide	Concentration [μM]	Sedimentation coefficient (s) $s_{20,w}$ [S]	MW estimate [kDa]	Estimated number of monomeric units in oligomer	Predicted Stokes radius [nm]
GLP-1-Am	85	1.00 2.11	6.58 (62 %) 20.1 (19 %)	2 6	1.6 2.3
IPP3	85	2.64 4.16	33.7 (47 %) 66.7 (53 %)	9 18	2.8 3.5
IPP4	85	0.74 1.73 6.16	4.95 17.9 120	1 5 32	1.1 2.5 4.4
IPP5	85	0.61 2.34 3.19 5.15	4.15 31.4 (61 %) 50.0 (33 %) 103 (5 %)	1 8 13 28	1.5 2.9 3.4 4.4

Table 5.2: Summary of the oligomeric population distribution in freshly prepared samples of GLP-1-Am, IPP3, IPP4 and IPP5 at pH 7.5. The molecular weight of the species, number of monomeric units in the oligomer and Stokes radius of the species were obtained from size-exclusion chromatography (SEC) and sedimentation velocity experiments. SEC parameters were calculated using the equations obtained from the Superdex 200 Increase 10/300 column calibration, **Figure 2.1**. Sedimentation velocity parameters were obtained directly from the Sedfit program[220,221] using a continuous c*(s)* distribution model. All frictional coefficients in the sedimentation velocity analysis were between 1.3 and 1.4. Percentages in brackets in the MW estimation column (AUC measurements), are estimates of the amount of peptide in a particular peak (determined only for easily distinguishable peaks). All experiments were performed in 25 mM phosphate at pH 7.5. *The molecular weight of GLP-1-Am is below the resolution range of the Superdex 200 Increase 10/300 size-exclusion column.

Peptide concentration-dependent oligomer populations of IPP4

Using the size-exclusion chromatography and sedimentation velocity experiments, peptide concentration-dependent changes in the population of IPP4 oligomers were observed. This phenomenon was observed for IPP4 only at micromolar concentrations, the populations of oligomer species formed by IPP3 or IPP5 did not show a similar dependence on peptide concentration.

Figure 5.10A shows SEC chromatograms of freshly prepared solutions of IPP4 over a range of peptide concentrations from 10 to 100 µM in 25 mM phosphate, pH 7.5. Figure 5.10B shows far-UV CD spectra of the same samples. Far-UV CD spectra did not show any significant changes in the secondary structure of IPP4 over the range of peptide concentrations studied. However, as apparent from the SEC chromatograms shown in Figure 5.10A, populations of stable oligomeric species varied with the peptide concentration. At low peptide concentrations (10 and 25 µM), two peaks eluting at around 10 mL and 21 mL were observed. Surprisingly, as the peptide concentration increased, a new broad peak appeared with a maximum elution volume at around 13–14 mL. Figure 5.8 shows the SEC chromatogram of IPP4 at an even higher concentration of 135 µM, and this confirms the results observed at lower peptide concentrations – the peaks at around 10 mL and 21 mL completely disappeared under these conditions giving rise to a single broad peak with an elution volume of 13–14 mL. This behaviour was reproducible using different IPP4 batches made by different manufacturers (Appendix A, Figure A1).

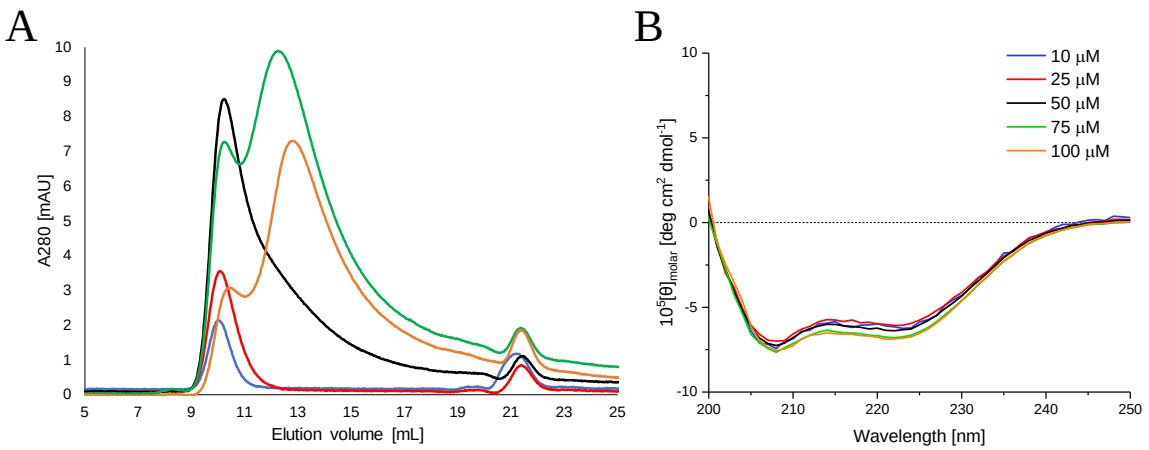


Figure 5.10: Peptide concentration-dependent distribution of IPP4 oligomeric species at pH 7.5. (A) SEC chromatograms of freshly prepared IPP4 samples at 10, 25, 50, 75, and 100 µM in 25 mM phosphate, pH 7.5. Samples were analysed using a Superdex200 Increase 10/300 column. (B) Far-UV CD spectra of freshly prepared IPP4 samples at 10, 25, 50, 75, and 100 µM in 25 mM phosphate, pH 7.5. All samples were measured in a 0.1 cm pathlength cuvette. The CD signal was converted into molar ellipticity, $[\theta]_{molar}$, units, see Chapter 2 for details. The legend located in the top right corner applies to both panels A and B. An IPP4 batch obtained from Bachem was used for these experiments.

Sedimentation velocity was employed as a complementary experiment to study the peptide concentration-dependent distribution of oligomer species of IPP4. Freshly prepared solutions of IPP4 at various concentrations (from 10 to 135 µM) in 25 mM phosphate at pH 7.5 were centrifuged at 50 000 rpm for *circa* 12 hours. The interference sedimentation curves were recorded and fitted to a continuous c(*s*) distribution model implemented in the Sedfit program[220,221] providing a sedimentation coefficient distribution plot, Figure 5.11. Figure 5.11 shows the shift of the sedimentation coefficient distributions towards higher values with increasing peptide concentration. In addition, with increasing peptide concentration, the peaks corresponding to individual oligomer populations merge together to form a single broad peak. These observations indicate the formation of new large oligomeric species which are stable only at higher peptide concentrations and which are in rapid equilibrium with the smaller oligomeric species present in the sample. Table 5.3 provides a summary of the species observed at different IPP4 concentrations.

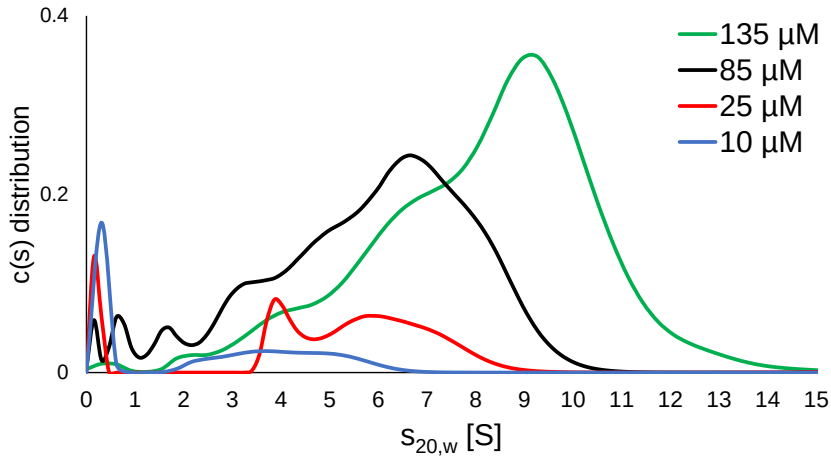


Figure 5.11: A sedimentation plot of freshly prepared samples of IPP4 at different peptide concentrations. Samples of different concentration (from 10 to 135 µM) were prepared in 25 mM phosphate at pH 7.5 prior to the experiment. After 2-hour temperature equilibration to 20 °C, the experiment was performed with centrifugation at 50 000 rpm using Beckman Optima XL-I Analytical Ultracentrifuge. The interference sedimentation curves were collected as 150 scans (approx. 12 h run) and fitted to a continuous c(*s*) distribution model. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$).

Sedimentation velocity: IPP4 at pH 7.5				
Concentration [μM]	Sedimentation coefficient $s_{20,w}$ [S]	MW estimation [kDa]	Estimated number of monomeric units in oligomer	Frictional coefficient
10	4.08	66.7 (60 %)	17–18	1.31
25	4.10 6.35	66.8 (26 %) 128 (65 %)	17–18 34	1.31
85	0.74 1.73 6.16	4.95 17.9 120	1 4–5 32	1.29
135	8.20	181	48	1.27

Table 5.3: Sedimentation coefficients and calculated molecular weights of oligomeric species of IPP4 at different peptide concentrations at pH 7.5. Sedimentation coefficients correspond to the distributions in Figure 5.11. Molecular weights were calculated using the Sedfit program[220,221] using a continuous c(s) distribution model. The frictional coefficients were optimized to provide the best fit to the model. Percentages in brackets in the MW estimation column give the proportion of a peptide in an oligomeric state (determined only for easily distinguishable peaks).

To further probe the peptide concentration-dependent behaviour of IPP4 at pH 7.5, a pyrene fluorescence assay was also employed. Pyrene fluorescence is sensitive to the solvent environment, and it is, therefore, also used as a probe in protein/peptide self-assembly studies.[312,313] Pyrene fluorescence intensity can reflect its local hydrophobic environment and it has been routinely used to determine the critical aggregation concentration of peptides and peptide amphiphiles.[17,80,124,126] IPP4 at different concentrations, ranging from 1 to 100 μM, was tested with 1 μM pyrene in 25 mM phosphate, pH 7.5. After excitation at 339 nm, pyrene fluorescence emission at 373 nm was measured. In addition to the fluorescence intensity at 373 nm, the ratio of the intensity corresponding to the pyrene vibronic bands I (at 373 nm) and III (at 384 nm) was also sometimes plotted[80]. In Figure 5.12, pyrene fluorescence intensity at 373 nm is plotted against IPP4 concentration on a logarithmic scale. The fluorescence intensity shows a large change at around 18 μM IPP4. This change can be attributed to the transition between oligomers formed at "low" and "high"

peptide concentration. The sharp increase in pyrene fluorescence in the high peptide concentration oligomer regime is likely to be due to the larger size of the oligomers formed (see Figure 5.11) and therefore greater hydrophobic clusters present in the high concentration oligomers which provide better hydrophobic environment for pyrene. It is also interesting to note that this concentration is close to that observed for the formation of the oligomer detected using SEC at peptide concentrations between 25 and 50 µM, Figure 5.10. It should be noted that one must take into the account the additional dilution of the sample during the chromatography, which will result in a decrease from the initial concentration injected onto the column.

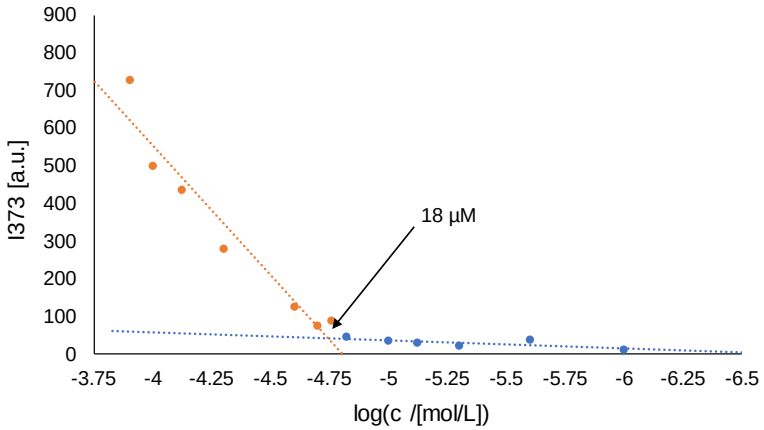


Figure 5.12: Measurement of the critical aggregation concentration of IPP4: dependence of pyrene fluorescence on IPP4 concentration. Fluorescence intensity at 372 nm was recorded after sample excitation at 339 nm for a range of IPP4 concentrations (from 1 to 100 µM) with 1 µM pyrene. All samples were freshly prepared in 25 mM phosphate, pH 7.5. Concentrations are plotted as log(c). The intersection of blue and orange line defines the "critical aggregation concentration" which corresponds to approx. 18 µM.

Transmission electron microscopy (TEM) was employed to image the oligomeric species present in freshly prepared samples of IPP3, IPP4, and IPP5 at pH 7.5. Freshly prepared samples of 5 and 25 µM were applied onto a copper grid and negatively stained using 2 % uranyl acetate solution. Figure 5.13 depicts representative TEM images of the early oligomers formed in IPP3, IPP4 and IPP5 samples at two different peptide concentrations. Individual oligomeric species were distinguishable only for IPP3 and IPP4. In the case of IPP5, it was not possible to observe individual oligomers, which may be due to their heterogeneity, shown by SEC and AUC (Figure 5.8 and Figure 5.9), and a high tendency to cluster during drying on the TEM grid.

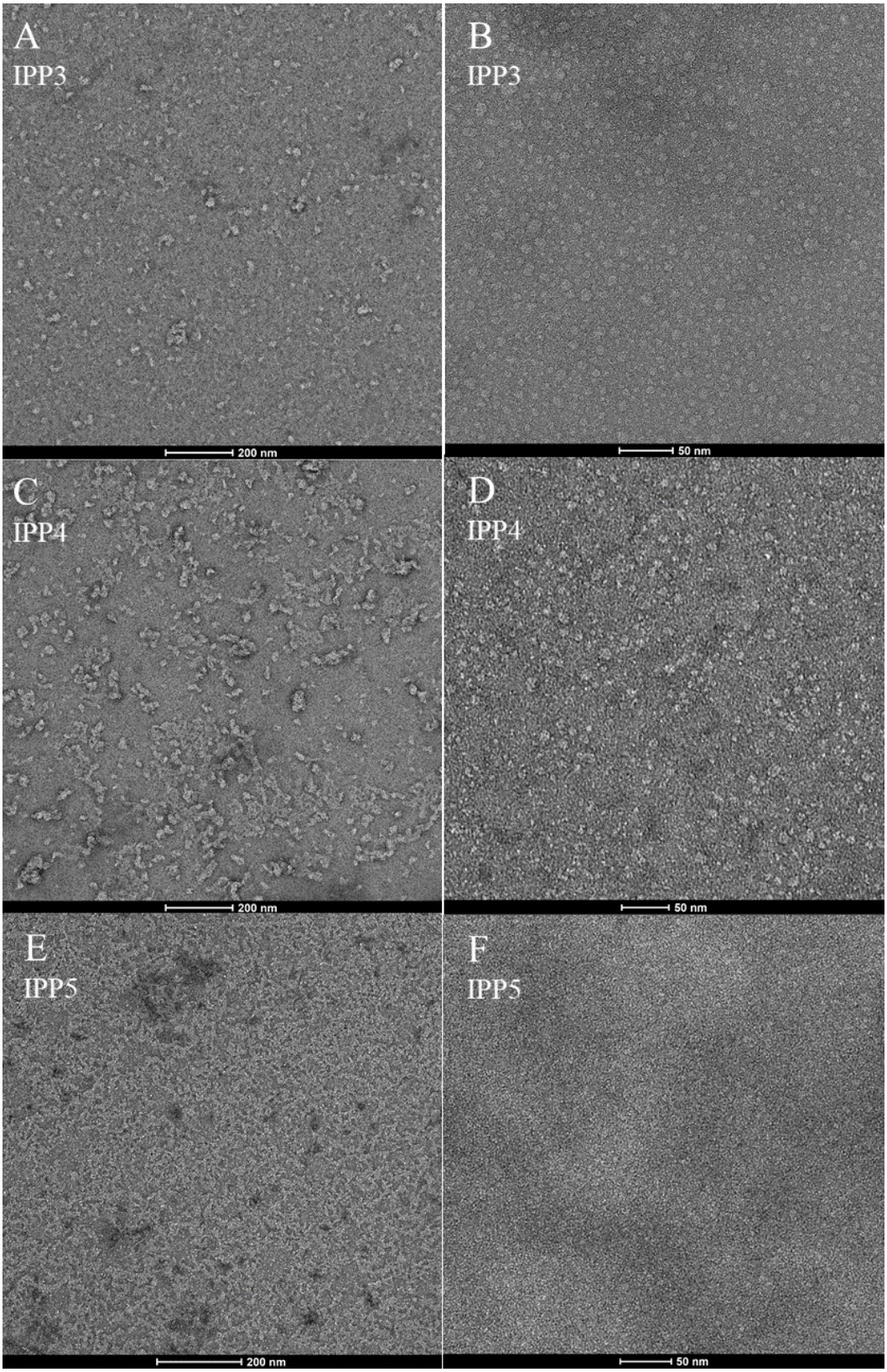


Figure 5.13: TEM images of freshly prepared samples of lipidated GLP-1 Am analogues at pH 7.5. Freshly prepared samples at 25 µM (A, C, E) or 5 µM (B, D, F) were applied onto carbon coated copper grids and negatively stained using 2 % uranyl acetate. Images were taken at 36000 x (A, C, E) or 94000 x (B, D, F) magnification.

Histograms displaying the distribution in the diameters of oligomeric species were constructed for IPP3 and IPP4 – Figure 5.14. An average, a standard error of the mean and a standard deviation were calculated for each distribution (see Chapter 2 for details). For IPP3 species, the average particle diameter with the standard error of the mean was (20.5 ± 0.7) nm and the standard deviation was 5.8 nm. If the same analysis was applied to IPP4, the average diameter with the standard error of the mean was 30.1 ± 1.5 nm and the standard deviation was 12.8 nm, however, IPP4 histogram indicate that two separate subpopulations – one with a diameter of around 10 nm and the other with a diameter of around 37 nm. This observation is likely due to concentration-dependent oligomeric distribution as the TEM images were taken at 5 and 25 µM peptide concentrations which correspond to "low" and "high" peptide concentration regime as indicated previously by the pyrene fluorescence assay, Figure 5.12.

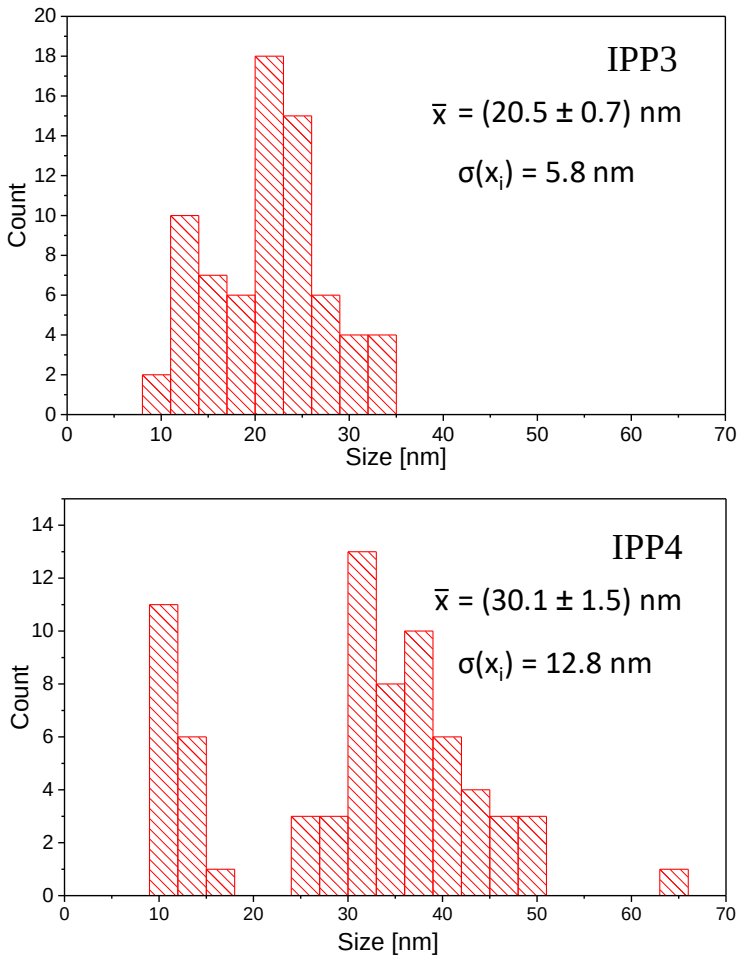


Figure 5.14: Distribution of the diameters of oligomeric species of IPP3 and IPP4 observed by TEM. Each histogram was constructed from 72 data points (12 different images were analysed, 6 data points from each TEM image). An average ($\bar{x}$) with a standard error of the mean and a standard deviation ($\sigma(x_i)$) were calculated for each distribution, see Chapter 2 for details.

5.6 pH-dependent behaviour of lipidated analogues

As shown earlier in this Chapter, the solubility of lipidated analogues is affected by pH – IPP3, IPP4 and IPP5 analogues are soluble only at neutral and basic pH values. The previous Section describes results obtained at pH 7.5, this Section illustrates the behaviour of lipidated GLP-1-Am analogues at other pH values.

5.6.1 IPP5: pH-dependent oligomer populations

Previously published studies on liraglutide[81,314] reported a pH-induced transformation of oligomers formed in solution. Two oligomeric subpopulations which changed in their ratios over a pH range from 6.4 to 8.1 were observed. IPP5 is structurally very similar to Liraglutide, the only difference is the additional C-terminal amidation in IPP5. This Section describes the pH-dependent behaviour of IPP5.

IPP5 behaviour was investigated over a pH range from 6.5 to 7.5 using size-exclusion chromatography and sedimentation velocity experiments. For SEC experiments, IPP5 samples were freshly prepared and filtered through a 0.22 μm filter to remove any larger aggregates prior to injection onto the column. All IPP5 samples were prepared at 170 μM in 25 mM phosphate buffer of a corresponding pH. The same buffer was used for sample preparation and as an elution buffer. Figure 5.15A shows the SEC chromatogram of IPP5 at pH 6.5, 7.0 and 7.5. Depending upon the pH, the three chromatograms show major peaks which are labelled 1, 2, and 3, and minor, low intensity peaks, which are labelled 4 and 5. Peak 3, which is the largest of all species observed is only seen at pH 6.5. The ratio of the heights of peaks 1 and 2 varies with pH. At pH 7.5, peak 2 is the dominant, whilst at pH 7, the ratio of peaks 1 and 2 increases but peak 2 is still larger, however, at pH 6.5, peak 1 has a higher intensity than peak 2. In addition, at pH 6.5, the area under the curve for peaks 1 and 2 has decreased, these species converting into the higher molecular weight species observed (peak 3). This suggests a pH-dependent equilibrium of coexisting oligomers.

Similar results were also observed using velocity sedimentation experiments (Figure 5.15B). IPP5 samples (85 μM) were freshly prepared in 25 mM phosphate buffer of a corresponding pH. Sedimentation plots show species that broadly follow the oligomeric distribution

observed using SEC. Peaks which are likely to be the same species are labelled with the same number in both plots (Figure 5.15A and B).

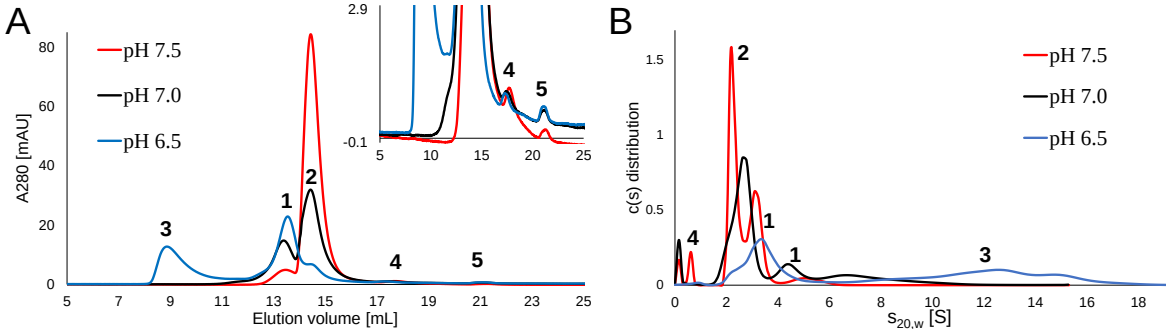


Figure 5.15: pH-dependent distribution of IPP5 oligomers in freshly prepared samples. (A) SEC chromatograms of freshly prepared IPP5 samples at 170 µM concentration in 25 mM phosphate buffer of a corresponding pH were measured using a Superdex200 Increase 10/300 column. (B) Sedimentation velocity experiments were performed with freshly prepared IPP5 samples at 85 µM concentration in 25 mM phosphate buffer of a corresponding pH. Size-exclusion experiments were performed by Xinyang Li.

Using SEC, it was possible to isolate peaks 1, 2 and 3 and further characterize them using circular dichroism, intrinsic tryptophan fluorescence and a thioflavin T binding assay. The oligomeric species in peak 3 were also visualized using TEM. Oligomeric peaks were isolated in concentrations ranging from 5 to 10 µM and further analysed within two hours from isolation. It was shown, by reinjecting the sample back onto a SEC column that the species are sufficiently stable over these timescales, for characterization to be possible (Chapter 6, Figure 6.8). Figure 5.16A shows far-UV CD spectra of isolated oligomeric peaks 1, 2 and 3 and estimated percentages of α-helical and β-sheet secondary structure. Oligomer 1 (labeling consistent with peak labeling in the SEC plot – Figure 5.15A) shows significantly higher β-sheet content and lower α-helical content compared to oligomer 2. Oligomer 3, the higher-molecular weight oligomer, which was detected only at pH 6.5, also shows a high content of β-sheet structure but the decrease in α-helical content is smaller compared to oligomer 1. Intrinsic tryptophan fluorescence spectra (Figure 5.16B) reflect the local environment of tryptophan residue (position 25) in the oligomer structure. Oligomer 2 shows the lowest λ_{max} (334 nm) of all the analysed IPP5 oligomers which implies the tryptophan side chain is buried in the oligomer. In contrast, oligomer 1 has tryptophan λ_{max} at 345 nm which implies greater extent for tryptophan exposure to the solvent, therefore, tryptophan side chain is likely to be in a less structured region or on the surface of the oligomer. For oligomer 3, λ_{max} is *circa* in between the values for oligomer 1 and 2, at 340 nm. A thioflavin T binding assay was employed to test if any of the species had "amyloid-like" features and extended β-sheets (Figure 5.16C). ThT fluorescence from 460-600 nm was measured upon

excitation at 448 nm. As apparent from Figure 5.16C, all three IPP5 oligomers show a certain level of ThT binding, however, fluorescence intensity is greater for oligomer 1 and 3, which were also shown to be β-sheet-rich using far-UV CD. It was also possible to image oligomer 3 using TEM, Figure 5.16D shows *circa* 50 nm long "worm-like" oligomers. Oligomers 1 and 2 were too small to be visualized using TEM.

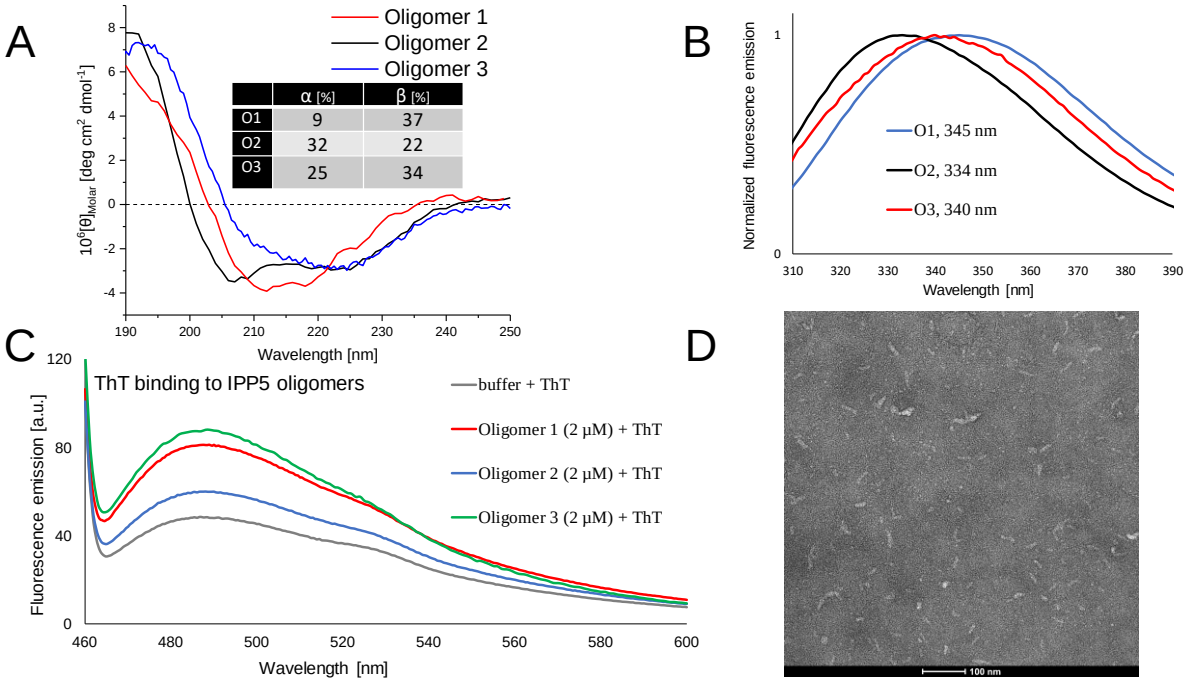


Figure 5.16: Characterization of IPP5 oligomers. IPP5 oligomers (1, 2 and 3) were isolated using SEC and further characterized using far-UV CD (A), intrinsic tryptophan fluorescence (B), and thioflavin T binding experiment (C). Oligomers 1 and 2 were isolated and analysed in 25 mM phosphate at pH 7.5 at 25 °C. Oligomer 3 was prepared and analysed in 25 mM phosphate at pH 6.5 at 25 °C. High-molecular weight oligomer 3 was also imaged using TEM. Far-UV CD spectra of oligomers 1 and 2 were recorded by Xinyang Li.

5.6.2 IPP4: Toroidal assemblies formation at the solubility threshold

As shown earlier in this Chapter, IPP3 and IPP4 show similar solubility profiles, both analogues are soluble above pH 7, but have limited solubilities (< 50 µM) at pH 7 which is termed the "solubility threshold" (Figure 5.3). Their behaviour at this solubility threshold at pH 7 is investigated in this Section.

Samples of 40 µM IPP4 and IPP3 were freshly prepared in 25 mM phosphate, pH 7, filtered through a 0.22 µm membrane filter and imaged using TEM. Whereas the IPP4 sample, Figure 5.17A, shows a fairly uniform population of spherical assemblies with diameters in

the range of 12–27 nm, IPP3 under the same conditions forms amorphous aggregates which are quite variable in size as shown in Figure 5.17B. Due to the observed diameter of these species, it is unlikely that they are single-layer micelles as the diameter of those is estimated to be around 5 nm given the size of the peptide[315,316]. Interestingly, the behaviour of IPP4 around pH 7 seems to be unique compared to all the other lipidated GLP-1-Am analogues studied and also compared to other pH values. The formation of these regular spherical assemblies was also reflected in an increased level of scattering between 300 and 350 nm in the absorption spectrum of IPP4 at pH 7, in comparison with the level of light scattering seen in the absorption spectrum for IPP4 at pH 7.5, Appendix A, Figure A2.

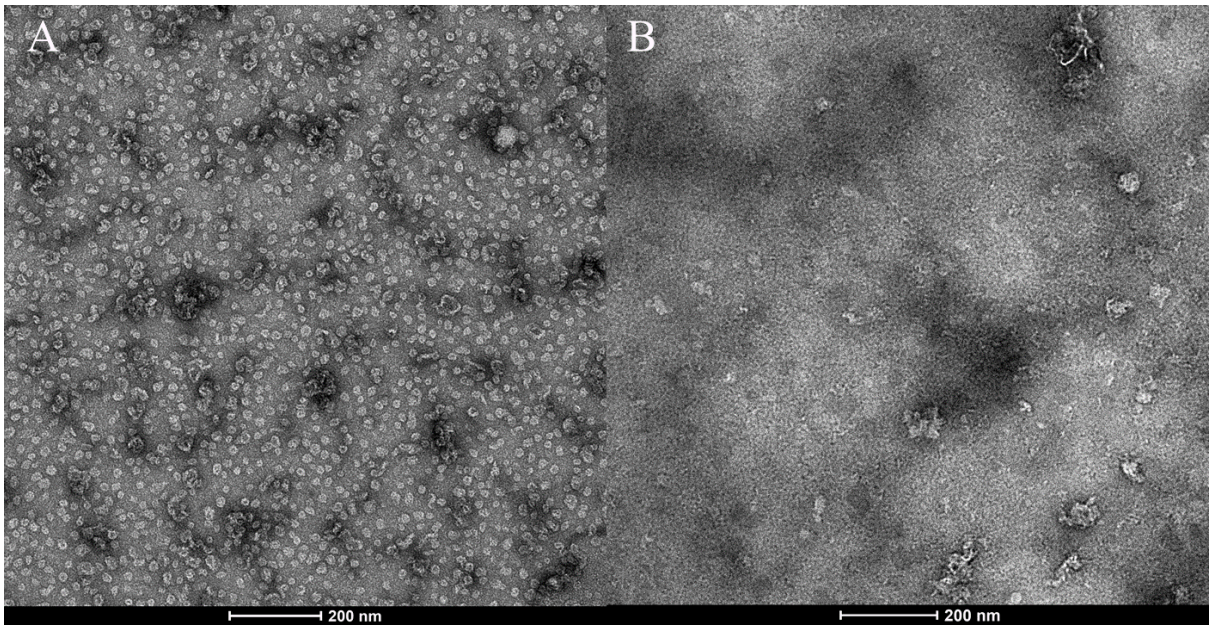


Figure 5.17: TEM images of freshly prepared IPP4 and IPP3 samples at pH 7. Samples of 40 µM IPP4 (A) and 35 µM IPP3 (B) were freshly prepared in 25 mM phosphate, pH 7, filtered through a 0.22 µm membrane filter, applied onto a carbon coated copper grid and stained using 2 % uranyl acetate.

Freshly prepared samples of IPP4 and IPP3 at pH 7 were also studied using SEC. Figure 5.18 shows the SEC chromatograms of IPP4 and IPP3 at 10, 20 and 30 µM in 25 mM phosphate, pH 7. IPP4 chromatograms show only a single peak at around 8 mL (Figure 5.18A) which indicates the formation of a single species which is independent of peptide concentration under the conditions used. Using the calibration curve for Superdex200 Increase 10/300 column (Chapter 2, Figure 2.1), the diameter of eluted species was estimated to be 21 nm which fits into the size range observed by TEM. IPP3 chromatograms show a very low intensity peak at around 18 mL (Figure 5.18B), indicating a species of low molecular weight somewhere in the range from monomer to tetramer. The low intensity

signal of IPP3 samples is likely to be caused by the peptide aggregates in solution remaining in the filter at the top of the SEC column.

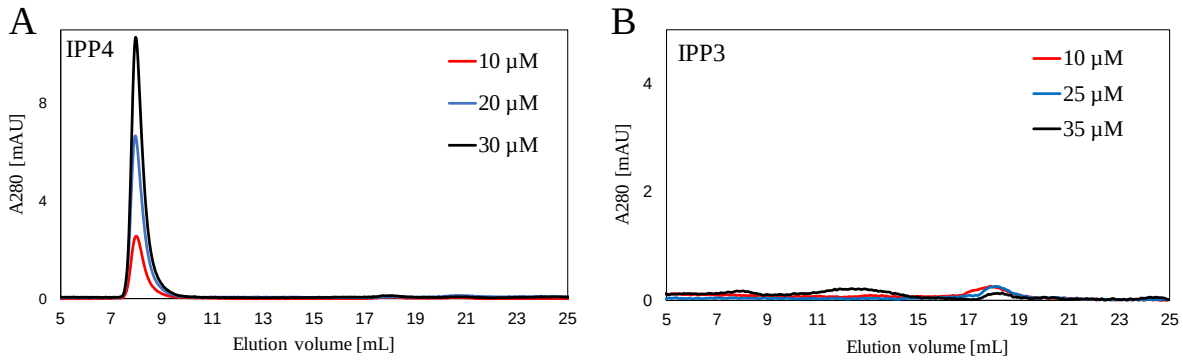


Figure 5.18: Size-exclusion chromatograms of IPP4 and IPP3 at pH 7. Freshly prepared IPP4 (A) and IPP3 (B) samples in 25 mM phosphate, pH 7 were analysed using a Superdex200 Increase 10/300 SEC column.

The self-assembly behaviour of IPP4 and IPP3 were also investigated using a pyrene assay. Fluorescence emission spectra of 50 µM IPP4 and 30 µM IPP3 (lower peptide concentration due to the lower solubility) with 1 µM pyrene, over a pH range from 7 to 8, were recorded, (Figure 5.19). IPP4 shows a significantly higher pyrene fluorescence intensity at pH 7 than at other pH values (Figure 5.19A). IPP3 samples show about the same pyrene fluorescence intensity at all pH values (Figure 5.19B). This experiment was reproducible and independent on the ionic strength or buffer used (Appendix A, Figure A3). This observation corresponds with the formation of spherical assemblies only at pH 7 and only in the case of IPP4 and indicates that the spherical assemblies increase the pyrene fluorescence to a greater degree than other oligomers. For IPP4, the pyrene assay was also performed over a range of peptide concentrations from 2 to 50 µM at pH 7, however, the analysis of the fluorescence emission intensity at 373 nm did not show any critical concentration for the formation of these spherical assemblies in this concentration range (Appendix A, Figure A4). Therefore, this suggests that the critical micellar concentration IPP4 at pH 7 is below 2 µM and, therefore, that the spherical assemblies are stable in all the concentrations used in this study.

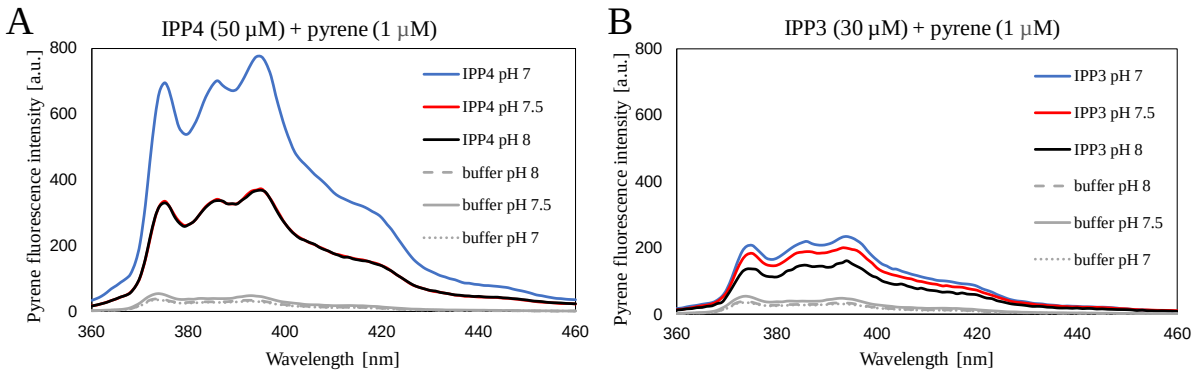


Figure 5.19: Pyrene fluorescence spectra of 1 µM pyrene with IPP4 and IPP3 over a pH range from 7 to 8. Fluorescence emission spectra of 50 µM IPP4 and 30 µM IPP3 (lower peptide concentration due to the lower solubility) with 1 µM pyrene in 25 mM phosphate at pH from 7 to 8 at 25 °C were recorded after excitation at 339 nm.

As observed by negatively stained TEM, IPP4 assemblies have a spherical shape with a hollow cavity in the middle (Figure 5.20) resembling a toroid. A sample prepared under the same conditions (40 µM peptide concentration in 25 mM phosphate at pH 7) was also imaged using cryo-EM to eliminate the effect of staining and the drying on grid process. Figure 5.21 shows that the morphology of assemblies imaged using the cryo-EM technique is similar to assemblies which were imaged using negative staining. Therefore, the staining and drying process is likely not to cause significant disruption of the morphology of the observed IPP4 assemblies.

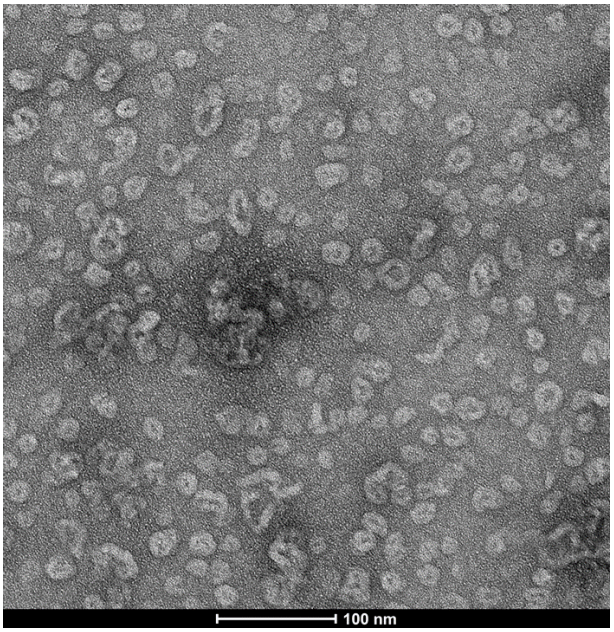


Figure 5.20: Detailed TEM image of IPP4 assemblies formed at pH 7. 40 µM IPP4 was freshly prepared in 25 mM phosphate, pH 7, filtered through a 0.22 µm membrane filter, applied onto a carbon coated copper grid and stained using 2 % uranyl acetate.

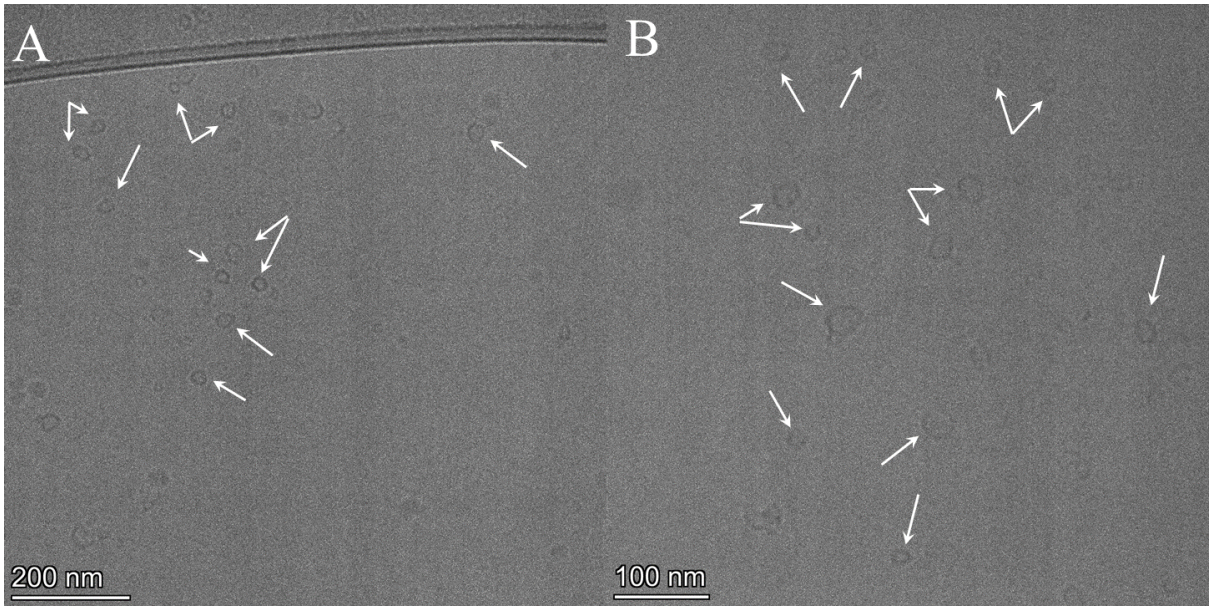


Figure 5.21: Cryo-EM images of IPP4 toroidal assemblies formed at pH 7. 40 µM IPP4 was freshly prepared in 25 mM phosphate, pH 7, filtered through a 0.22 µm membrane filter, applied onto a lacey carbon TEM copper grid and plunge-frozen in the liquid ethane. The cryo-EM samples were prepared and imaged with the help of Dr. Heather Greer.

The peptide conformation within these assemblies was also characterized by far-UV CD and intrinsic tryptophan fluorescence. The far-UV CD spectrum of freshly prepared IPP4 at pH 7 has two minima at around 222 nm and 210 nm (Figure 5.22A). The secondary structure prediction shows a high content of α-helix in the assembly – 43 %. The high content of α-helix has been already reported in a micelle-like structure of a lipidated gut-derived peptide hormone.[17] Intrinsic tryptophan fluorescence spectrum of freshly prepared IPP4 at pH 7 shows a maximum at 337 nm implying that tryptophan residue is already quite buried in the assembly structure (Figure 5.22B).

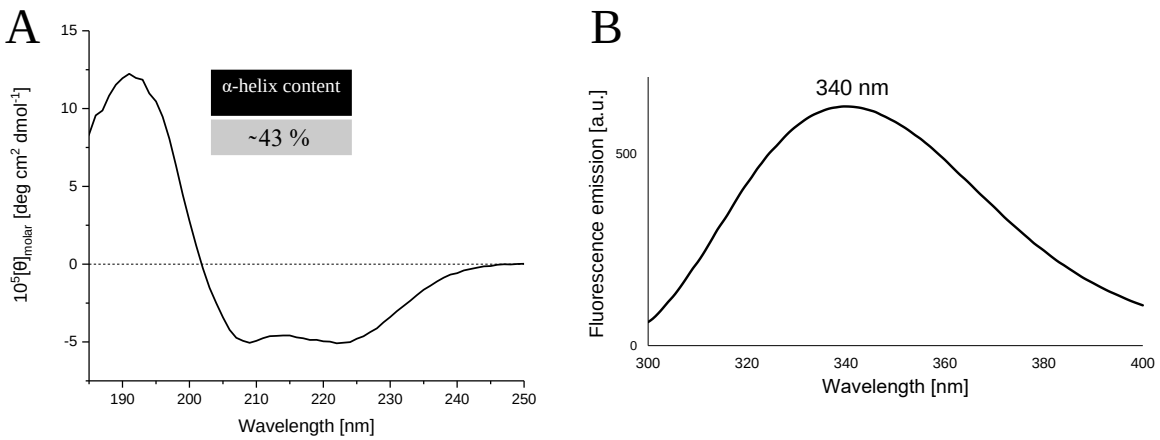


Figure 5.22: Far-UV CD and intrinsic tryptophan fluorescence characterization of assemblies formed by IPP4 at pH 7. 40 µM IPP4 was freshly prepared in 25 mM phosphate, pH 7 and characterized using far-UV CD (A) and intrinsic tryptophan fluorescence spectra (B). Far-UV CD spectra were measured in a 0.1 cm pathlength cuvette. CD signal was converted into molar ellipticity, $[\theta]_{molar}$, units, see Eq. 2.6 (Chapter 2). Total α-helical content (in A) was calculated using Eq. 2.7 and Eq. 2.8 as indicated in Section 2.11 (Chapter 2).

5.7 Discussion

The results presented in this Chapter provide an insight into the effect of lipidation on C-terminally amidated GLP-1. The behaviour of three analogues, IPP3, IPP4 and IPP5, which differ by the position of lipidation in the peptide sequence, was investigated. All analogues studied in this Chapter are lipidated by palmitic acid which is attached to a lysine size chain via a γ-glutamic acid linker, Figure 5.2. The focus of the Chapter is on the behaviour of freshly prepared samples of IPP3, IPP4 and IPP5, in particular, changes in solubility, structural changes caused by the lipidation and self-assembly behaviour of individual lipidated analogues.

5.7.1 Lipidation of GLP-1-Am: General effects on the biophysical properties of the peptide

As shown in Figure 5.3, the attachment of a fatty acid moiety limits the solubility of GLP-1-Am analogues. Whereas non-lipidated GLP-1-Am was soluble over entire pH range tested (from pH 2.5 to 8.5), the solubility of IPP3, IPP4 and IPP5 is restricted to neutral and basic pH. IPP5 has the widest solubility range (from pH 6 higher) of all the lipidated analogues studied which correlates with it having the lowest isoelectric point of 5.5, Figure 5.4. The solubility threshold of lipidated analogues IPP5, IPP4 and IPP3 at around pH 6 and

7 and may therefore be related to the deprotonation of the N-terminal histidine as its deprotonation is likely to occur over this pH range.

The attachment of the palmitic acid via γ-glutamic acid linker caused rather smaller changes in the secondary structure of the peptide backbone. This is apparent in the IR spectra of the lyophilized peptide samples (Figure 5.5) and in far-UV CD spectra of freshly dissolved sample in 25 mM phosphate at pH 7.5. The α-helical content in the peptide seems to slightly increase upon lipidation, Table 5.1, however, the result from several secondary structure predicting algorithms were not fully conclusive. To conclude, neither the palmitic acid lipidation nor the introduction of additional negative charge in the linker disrupt the interactions stabilizing the α-helix in the peptide regardless of lipidation position.

5.7.2 Self-assembly of IPP3

IPP3 is lipidated at Lys12, is soluble at pH > 7 and has an isoelectric point at around 5.7. In general, IPP3 behaves similarly to IPP4, however, lipidation at position 12 does not cause an increase in the α-helicity of the peptide as was observed for IPP4 (which is lipidated at position 17), Table 5.1. In the size-exclusion chromatography, freshly prepared sample of IPP3 eluted as a single broad peak, corresponding to an oligomer consisting of approx. 13 monomers. However, the sedimentation velocity experiment revealed two partly overlapping peaks which correspond to a 9-mer and 18-mer. These observations indicate an existence of two oligomeric species in rapid equilibrium, Figure 5.23. Sedimentation velocity experiments are more likely to better reflect oligomeric species in the rapid equlibrium[222] compared to the size-exclusion chromatography, in which the rapidly interconverting species elute in a single broad peak which theoretical mass is somewhere in between the masses of the interconverting species[317–319].

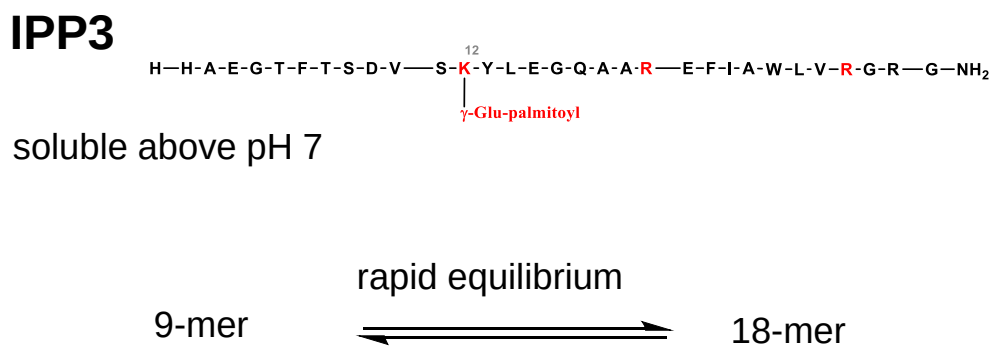


Figure 5.23: IPP3 – self-assembly scheme.

5.7.3 Self-assembly of IPP4

IPP4 analogue is lipidated at Lys17 position and it has an isoelectric point at around 6.1. The analogue is soluble above pH 7. At pH 7, the solubility of IPP4 is limited with the highest achievable IPP4 concentration at around 50 µM. Under these conditions, IPP4 forms regular toroidal assemblies, Figure 5.24. The morphology of IPP4 at pH 7 was studied using negatively stained TEM and cryo-EM, Figure 5.20 and Figure 5.21, both showing a uniform population of assemblies of toroidal morphology. The uniformity of their population was further shown using SEC (Figure 5.18A). The formation of IPP4 assemblies is not concentration-dependent over the concentration range studied (2–50 µM), Appendix A, Figure A4. Far-UV CD spectra of IPP4 assemblies revealed high percentage of α-helix – 43 %, Figure 5.22. The formation of micelles with α-helical structure has been previously reported for lipidated analogues of gut-derived peptide hormone $PYY_{3–36}$[17]. Other relevant examples of formation of spherical or ellipsoidal micelles include the cyclic lipopeptide Daptomycin[315] and TM5, an antimicrobial peptoid with a long aliphatic chain[320]. The hydrophobic core of IPP4 assemblies is likely to be formed by the palmitic acid aliphatic chains on the lipid whereas charged hydrophilic parts of the peptide chain are likely to be on the outer surface of the assembly, however the diameter of observed assemblies is too large to correspond to single-layer micelles, for which the diameter would be expected to be lower than 5 nm.

At higher pH values (pH 7.5 and above), IPP4 assembles to different (non-toroidal) oligomeric species. This pH-related transition of IPP4 oligomers is clearly reflected in the pyrene assay, Figure 5.19A, in which pyrene exhibits significantly more intense fluorescence emission in the presence of toroidal assemblies (at pH 7) than in the presence of non-toroidal oligomers (at pH 7.5) for the same IPP4 concentration. This is likely due to the penetration of pyrene into the highly hydrophobic core of the assembly. In contrast to pH 7, at pH 7.5, oligomer formation and distribution is peptide-concentration dependent. At low IPP4 concentrations, a single oligomeric species (18-mer) in equilibrium with monomer is present, Figure 5.10 and Figure 5.11. At higher IPP4 concentrations, a range of rapidly interconverting oligomers are formed. In addition, the average oligomer size increases with increasing peptide concentration, Figure 5.11. These concentration-induced changes in oligomer distribution are not accompanied by a change in peptide secondary structure as indicated by far-UV CD, Figure 5.10B. Using the pyrene assay, a threshold IPP4 concentration of 18 µM was identified, Figure 5.12. This value may be analogous to “critical

aggregation concentration" [17,321] or "critical oligomer concentration"[65] values reported for other self-assembling systems. Whereas, formation of a single stable oligomeric species, as observed for concentrations < 18 µM, can slow down/inhibit higher-order self-assembly and aggregation, a distribution of multiple oligomeric species can promote further self-assembly and aggregation over time.

A summary of IPP4 self-assembly behaviour is shown in Figure 5.24. The observed pH-transition, from toroidal assemblies at pH 7 to oligomeric species at pH 7.5, is likely to be associated with deprotonation of either the N-terminus or N-terminal histidine side chain as those groups have pK_a values in the right pH range[36]. Therefore, the N-terminal histidine residue is likely to play an important role in solubility and self-assembly properties of the peptide.

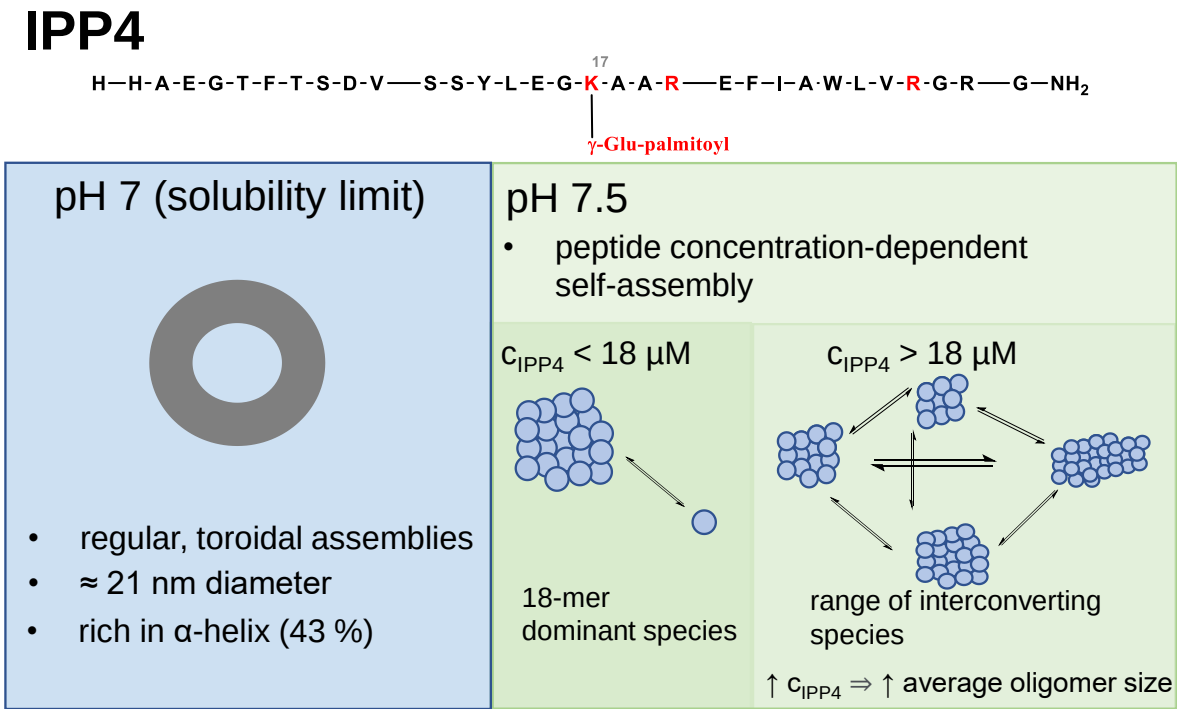


Figure 5.24: IPP4 – self-assembly scheme. The peptide monomer is represented as a blue circle.

5.7.4 Self-assembly of IPP5

The IPP5 analogue, lipidated at Lys20, has the widest pH-range of solubility of all lipidated analogues studied. It is soluble above pH 6 which correlates with its pI value which is around 5.5, the lowest of all lipidated analogues studied. Apart from a C-terminal amidation, the chemical structure of IPP5 is the same as the commercially available GLP-1 analogue

liraglutide, sold under the brand names Victoza or Saxenda, and used for type 2 diabetes treatment[16].

Similarly to liraglutide[81,314], IPP5 was shown to populate two major oligomeric species in solution, Figure 5.8, Figure 5.9. These species were resolved as distinct peaks in both size-exclusion chromatography and sedimentation velocity experiments, suggesting that their interconversion is slow. Sedimentation velocity experiments estimate the size of these two species to be approximately an 8-mer and 13-mer, Table 5.2. The prediction of oligomer size using size-exclusion chromatography is likely to be overestimated as it is dependent on the column calibration in which globular proteins were used and which may differ from lipidated oligomers in many parameters. However, due to limitations of both the size-determination methods used, it is possible that the larger of the observed oligomeric species is formed by dimerization of the smaller oligomer (i.e., heptamer self-association into 14-mer). Both size-exclusion chromatography and sedimentation velocity experiments also showed the presence of lower amounts of smaller oligomers and/or monomer, Figure 5.15. The oligomeric distribution of IPP5 also shows a pH dependence with the larger oligomer being more favourable at lower pH values.

Interestingly, two major oligomers of IPP5 are structurally distinct. The smaller oligomer (≈ 8-mer) has a higher content of α-helix compared to the larger oligomer (≈ 13-mer) which shows a prevalence of β-sheet structure, Figure 5.16. This α-helix to β-sheet transition in the oligomeric species is probably the first step leading to further aggregation of the analogue, which is studied in more depth in Chapter 6. The observed self-assembly properties of IPP5 are summarized in Figure 5.25.

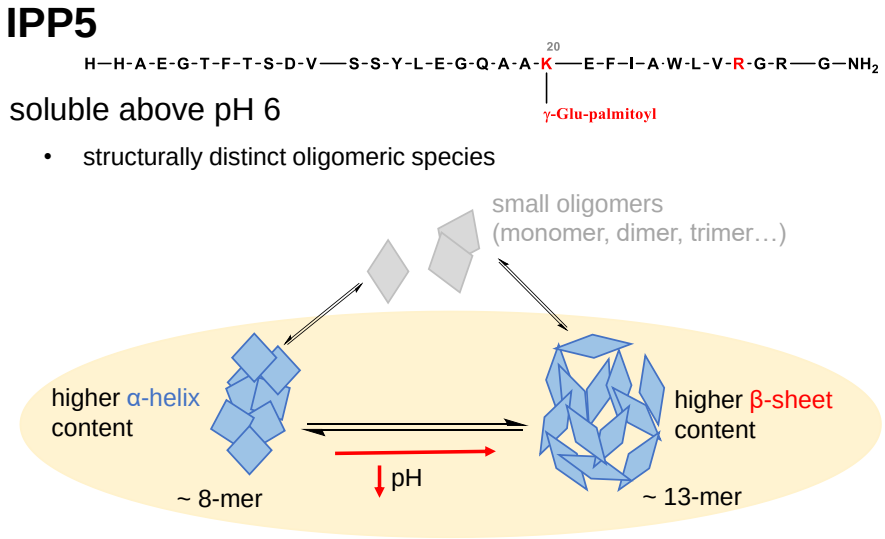


Figure 5.25: IPP5 – self-assembly scheme.

5.7.5 Summary and general characteristics of self-assembly of lipidated GLP-1-Am analogues

In this Chapter, it was demonstrated that the lipidation of GLP-1-Am generally triggers the formation of larger and more stable oligomers compared to non-lipidated GLP-1-Am. However, it was also shown that the size, stability, and distribution of the population of the oligomeric species formed can be regulated by the position of lipidation. This was shown using size-exclusion chromatography (SEC) and sedimentation velocity experiments (Figure 5.8, Figure 5.9, Table 5.2). IPP3, IPP4 and IPP5 were found to be present in solution mainly in the form of larger oligomers, which range from 8 to 32 monomeric units. These larger oligomeric states exhibit higher stability than the oligomeric species of non-lipidated GLP-1-Am and were detectable using SEC. Consistent with these observations are the intrinsic Trp fluorescence spectra (Figure 5.7A), which show that the tryptophan side chain is buried or partially buried within the oligomers. These observations indicate that the lipid chains contribute towards the stabilization of the oligomeric states. It is likely that the lipid chains are in close proximity in the hydrophobic core of the oligomer and that their interactions contribute to the stability of the oligomer as was previously reported in molecular dynamics studies of the self-assembly of liraglutide and semaglutide-[15,322]. The position of lipidation can affect not only the size and stability of the oligomeric species but also their structural characteristics, as was shown for the IPP5 analogue. For IPP5, two types of stable and well-resolved oligomers were observed, differing in the ratio of α-helical and β-sheet structures of the peptides in the oligomer. A similar structural transition in oligomeric species was previously reported in liraglutide studies[81,314].

CHAPTER 6 AGGREGATION BEHAVIOUR OF LIPIDATED GLP-1-AM VARIANTS

This Chapter is focused on the aggregation of lipidated analogues of GLP-1-Am which occurs on longer timescales (hours to days). It follows directly on from the previous Chapter which describes self-assembly processes shortly after the dissolution of the lyophilized peptide in an aqueous buffer.

6.1 Introduction and aims of the Chapter

Lipidated peptides have a highly amphiphilic character and therefore a high propensity for self-assembly and aggregation[121]. The morphology of the resulting aggregates ranges from amyloid fibrils with highly regular secondary structure to amorphous aggregates. Upon aggregation, which is frequently associated with changes in the secondary structure of a peptide or protein, the peptide/protein usually loses its biological activity. In most cases, the presence of aggregates in biopharmaceutical formulations is non-desirable as it not only lowers the amount of active drug in the formulation but may also present a cytotoxicity risk[260,323]. On the other hand, the nanostructures formed by the aggregation of peptides and proteins are of a great interest in the field of biomaterial development[324,325] and as long-acting, slow release formulations[326,327]. This Chapter deals with the long-term physical stability and aggregation of lipidated analogues of GLP-1.

Using a range of biophysical techniques, the kinetics of formation, structure and morphology of aggregates of four lipidated GLP-1 analogues varying in the position of the lipidation

(IPP2, IPP3, IPP4 and IPP5), as well as the nature of the lipid moiety (IPP10), were studied to establish the effect of the lipidation site and the lipid moiety on the physical stability of the peptide. This Chapter focuses on the structural changes of each GLP-1 analogue during the aggregation process and highlights common features in the aggregation of lipidated GLP-1 peptides.

The Chapter begins with comparison of the structure and morphology of aggregates of three lipidated GLP-1-Am analogues – IPP3, IPP4 and IPP5, which are formed over days at pH 7.5 (Section 6.2.1). The following section, 6.2.2, focuses on monitoring the aggregation process which led to the formation of the aggregates studied in Section 6.2.1. The Chapter continues with results highlighting the pH-dependent aggregation behaviour of individual analogues: pH-sensitive morphology of IPP3 aggregates and aggregation mechanism (Section 6.3.1) and aggregation of IPP4 toroidal assemblies into amorphous aggregates at pH 7.0 (Section 6.3.2). In Section 6.3.3, a GLP-1-Am analogue lipidated in the N-terminal region, IPP2, is investigated since its biophysical stability differs greatly from the rest of the analogues studied. In the final section, 6.4, the effect of the nature of the attached lipid on physical stability is assessed by comparing two lipidated analogues (IPP10 and IPP5) which differ solely by the lipid moiety and the linker at the same lipidation site.

6.2 Aggregation of lipidated GLP-1-Am analogues at pH 7.5

6.2.1 Structure and morphology of aggregates

To investigate the changes in secondary structure of aged samples of lipidated GLP-1-Am analogues, far-UV CD spectra were measured after 8 days of incubation at 37 °C with 180 rpm agitation, Figure 6.1. Significant changes in the spectra compared with those of freshly prepared samples (Chapter 5, Figure 5.6) was observed for all lipidated analogues studied in this Section (IPP3, IPP4 and IPP5). Whereas, for freshly prepared samples, the position of lipidation does not have any major effect on the secondary structure of the peptide (Chapter 5, Figure 5.6), the far-UV CD spectra of aged samples are clearly distinct from each other. These observations indicate that the lipidated analogues studied undergo structural changes during their incubation at 37 °C and the secondary structure of resulting aggregates is affected by the position of lipidation.

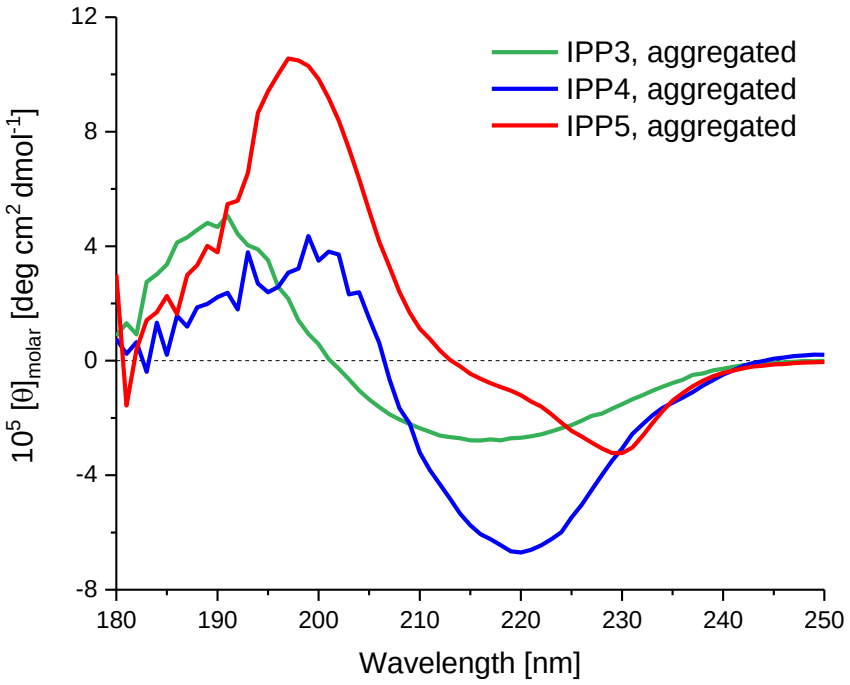


Figure 6.1: Far-UV CD spectra of lipidated analogues of GLP-1-Am after 8 days of aggregation. IPP3, IPP4 and IPP5 were incubated at 85 µM concentration in 25 mM phosphate at pH 7.5 for 8 days with agitation. Samples were measured in a 0.1 cm pathlength cuvette. The CD signal was converted into concentration independent molar ellipticity units.

Table 6.1 shows the percentages of different elements of secondary structure in aggregated samples of lipidated GLP-1-Am analogues as estimated using DichroWeb, Contin-LL method, dataset 3[180–182] and the BeStSel method[183–187]. Whereas the dataset used by the DichroWeb program contains only non-aggregating globular proteins, the BeStSel method dataset also contains the spectra of amyloid fibrils[183]. Therefore, the BeStSel method should be more suitable for the analysis of far-UV CD spectra of aggregating/fibrillating protein and peptide samples. However, CD spectra of amyloid fibrils and other type of aggregates may also be affected by the supramolecular chirality of these systems or light-scattering effects causing spectral artefacts which may complicate the analysis of secondary structure[328–331]. Nevertheless, an estimate of the dominant secondary structures in an aggregate can still be obtained. All analogues studied in this Section (IPP3, IPP4 and IPP5) show a decrease in α-helical structure and an increase in β-sheet content upon aggregation, Table 6.1 (t_{end} structure) and Table 5.1 in Chapter 5 (secondary structure of freshly prepared samples). Both far-UV CD analysis methods showed a higher percentage of disordered regions in aggregated IPP3 and IPP4, whereas for IPP5 the percentage of disordered regions

is significantly lower. For the IPP5 analogue, around 20 % of α-helical structure is maintained in the aggregated state.

DichroWeb, Contin LL method, dataset 3					
	α-helix [%]	β-sheet [%]		turns [%]	disordered [%]
IPP3	21	29		18	32
IPP4	12	30		15	43
IPP5	25	48		10	17
BeStSel					
	α-helix [%]	β-sheet total [%]	antiparallel/parallel β-sheet [%]/[%]	turns [%]	disordered [%]
IPP3	10	37	37/0	12	41
IPP4	8	43	29/14	10	39
IPP5	20	45	45/0	19	16

Table 6.1: Estimates of the amount of different elements of secondary structure in aggregated samples of lipidated GLP-1-Am analogues as predicted from far-UV CD spectra. The content of individual secondary structure elements was predicted using DichroWeb, Contin LL method, dataset 3[180–182] and BeStSel method[183–187]. Far-UV spectra shown in **Figure 6.1**.

To probe the tertiary structure of aged samples of aggregated IPP3, IPP4 and IPP5, intrinsic tryptophan fluorescence emission spectra and near-UV circular dichroism were measured. Figure 6.2A shows normalized intrinsic tryptophan fluorescence spectra of aggregated IPP3, IPP4 and IPP5, which reflects the local environment of the single tryptophan residue present (Trp25). All lipidated analogues studied in this Section show a blue shift in the wavelength of fluorescence intensity maximum, λ_{max}, compared to the spectra of freshly prepared samples, Chapter 5, Figure 5.7A. IPP5 has the greatest blue shift of 14 nm (343 nm → 329 nm) upon aggregation which indicates a significant increase in the hydrophobicity of the environment surrounding the tryptophan, suggesting that IPP5 forms a compact aggregate in which Trp25 is buried. The λ_{max} of IPP3 also shifts considerably from 344 to 332 nm (12 nm) whereas IPP4 shows only a minor shift from 337 to 335 nm.

In near-UV CD spectra of proteins and peptides, the signal arises from aromatic amino-acid side chains which are “locked” in their local environment and can, therefore, be used to

provide further information on the environment of aromatic side chains[167,332]. Figure 6.2B shows the near-UV CD spectra of aged, aggregated samples of IPP3, IPP4 and IPP5. All spectra show a non-zero CD amplitude, in contrast with the near-UV CD spectra of freshly prepared samples which did not show any CD signal in this region (Chapter 5, Figure 5.7B). The appearance of the signal indicates that the aromatic side chains are fixed in the structure of the aggregate formed. Interestingly, the spectra of aggregated IPP3 and IPP4 are almost a mirror image of each other suggesting that the contribution to the signal arises from the same side chains having an "opposite" supramolecular chirality[333,334]. IPP5 has a different near-UV CD spectrum which may suggest a different supramolecular structure of its aggregate compared to IPP3 and IPP4.

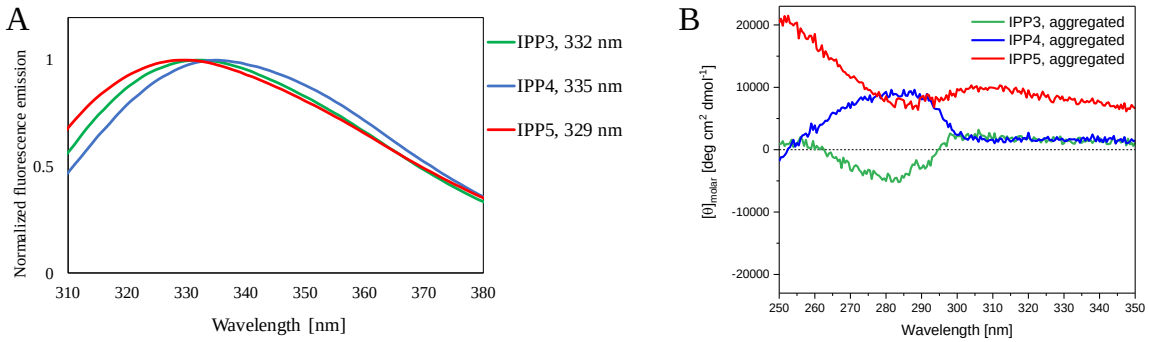


Figure 6.2: Intrinsic tryptophan fluorescence and near-UV CD spectra of aged samples of GLP-1-Am lipidated analogues. Both intrinsic tryptophan fluorescence and near-UV CD spectra were measured at 85 µM peptide in 25 mM phosphate, pH 7.5 after 8-day incubation at 37 °C with agitation. Intrinsic tryptophan fluorescence spectra (A) were normalized due to the different voltages across the photomultiplier tube which were used in the measurements. Near-UV CD spectra (B) were measured in a 0.2 cm pathlength cuvette.

To further investigate the nature, chirality and morphology of the aggregates formed, infrared spectroscopy and its chiroptical extension, vibrational circular dichroism, were employed. Due to the availability of the instrument (international collaboration), only two lipidated analogues, IPP4 and IPP5, were characterized by this technique.

IR and VCD spectra of aggregated IPP4 and IPP5 in 25 mM deuterated sodium phosphate buffer at pD 7.5 were recorded after a 7-day incubation at 37 °C with 180 rpm agitation, Figure 6.3. The IR spectra of IPP4 and IPP5 are very similar, and both possess a splitting at wavenumbers of 1643(1640)/1616 cm^{-1} which is typical of a high content of β-sheet structure. However, only the VCD spectrum of IPP5 shows the "fibril-like" enhanced intensity with a VCD/IR ratio of 1×10^{-3}. This observation illustrates a different higher-order structure of aggregates (IPP4 *vs* IPP5) with a high content of β-sheet structure.

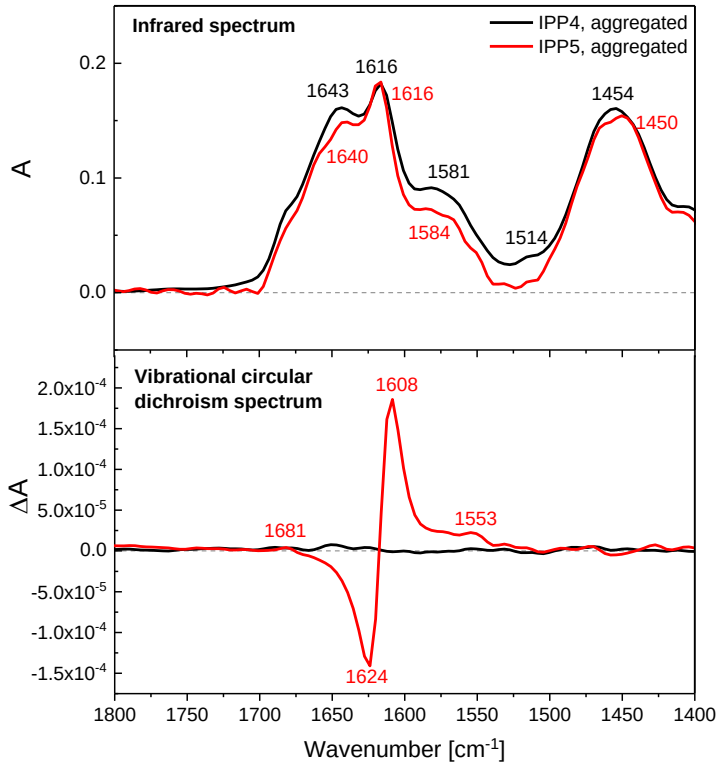


Figure 6.3: Infrared and vibrational circular dichroism spectra of aged, aggregated samples of IPP4 and IPP5. Spectra were measured in 25 mM deuterated sodium phosphate buffer at pD 7.5 after a 7-day incubation at 37 °C with 180 rpm agitation. The concentration of IPP4 and IPP5 was 4 mg mL^{-1} (1.06 mM and 1.07 mM, respectively). The infra-red absorbance spectra and vibrational circular dichroism spectra have been normalized such that the highest amide I absorption peaks are equal intensity.

The aggregates formed by IPP3, IPP4 and IPP5 after 8 days of incubation at 37 °C with 180 rpm agitation were imaged using negative-stain transmission electron microscopy (TEM), Figure 6.4. Two different, final peptide concentrations were imaged – 5 and 25 µM, the 5 µM samples were prepared by dilution of the 25 µM samples prior to application onto the TEM grid. IPP5 is the only analogue, which is observed to form long, rigid mature fibrils (Figure 6.4E and F). In contrast, IPP3 forms thread-like structures, however, these are short, curly, more flexible and tend to assemble into clusters (Figure 6.4A and B). TEM images of aggregated IPP4 show some short fragments with fibril-like structures as well as small irregular spherical and elliptical oligomers/aggregates (Figure 6.4C and D). Table 6.2 summarizes the dimensions and morphological features of species observed by TEM.

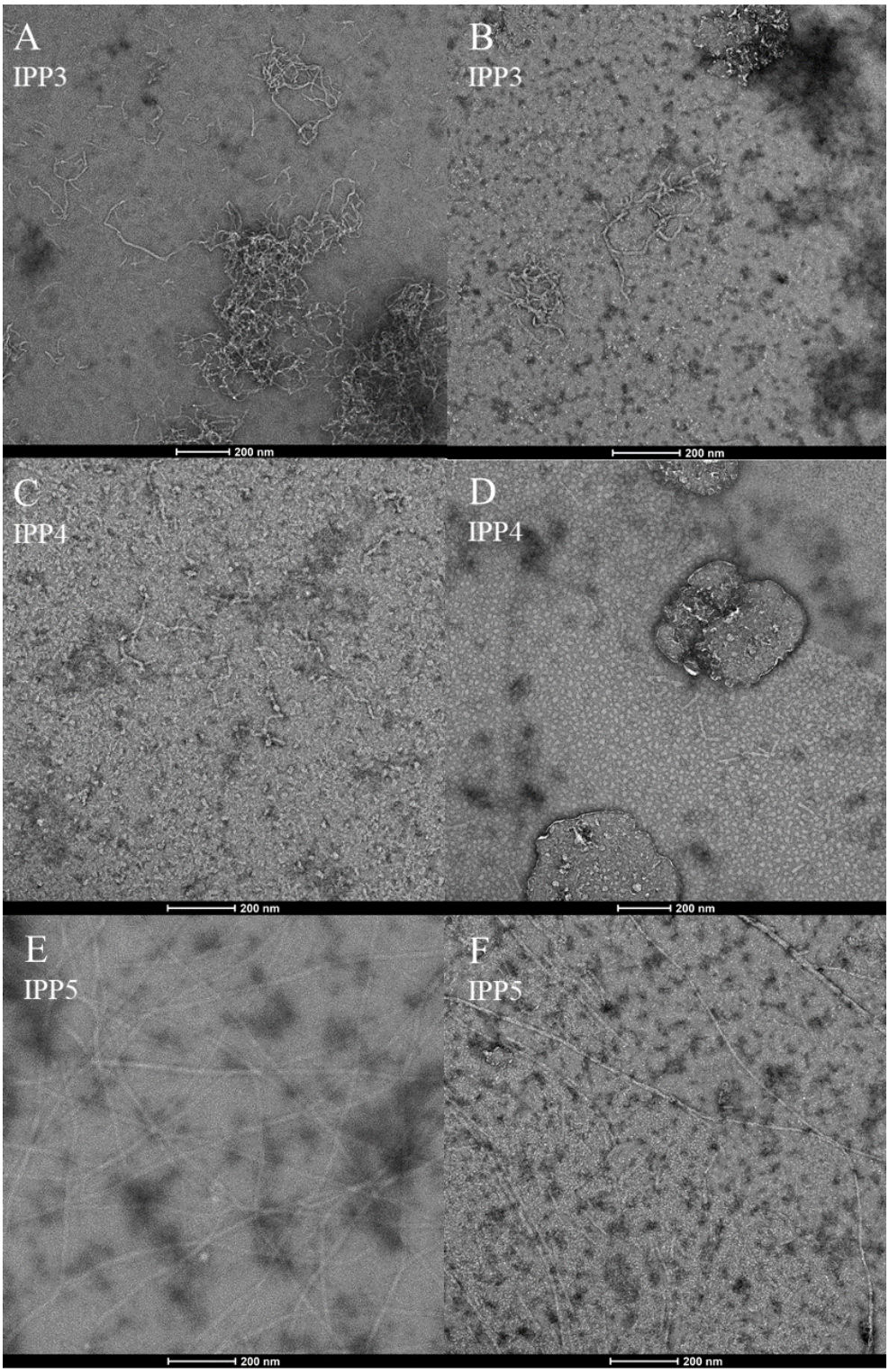


Figure 6.4: TEM images of aged, aggregated samples of IPP3, IPP4 and IPP5 at pH 7.5. Samples at 25 µM concentration were incubated at 37 °C in 25 mM phosphate, pH 7.5, with agitation for 8 days. Aged samples were applied onto a carbon coated copper grid at either 25 µM (A, C, E) or 5 µM (B, D, F) concentration, 5 µM samples were prepared by a five-fold dilution of 25 µM samples prior to the application onto the copper grid. Samples were negatively stained using 2 % uranyl acetate.

	Morphology of an aggregate	**Dimensions of the aggregate species**
IPP3	Short, curly fibrils with a tendency to cluster	fibril length: 150–500 nm fibril width: 9–14 nm
IPP4	irregular aggregates with occasional curly fibril fragments	Curly fibril fragments: length: 75–320 nm width: 11–16 nm Other aggregates: spherical or ellipsoidal species with *circa* 20–40 nm diameter
IPP5	Long, rigid fibrils	fibril length: 3–6 μm fibril width: 10–12 nm

Table 6.2: Morphological features of IPP3, IPP4 and IPP5 aggregates observed by TEM. To obtain the dimensions of each aggregated species, at least 25 data points from at least five TEM images were measured.

6.2.2 Aggregation kinetics

The previous section described the structure and morphology of aggregates formed after aging of the peptides at 37 °C for 7-8 days with agitation. This section is focused on the kinetics of the aggregation process. Fluorescence dye-binding assays were used in combination with size-exclusion chromatography and negative-stain transmission electron microscopy to monitor the aggregation process of the lipidated GLP-1-Am analogues.

A ThT-binding assay was used to monitor the aggregation kinetics of lipidated GLP-1-Am analogues. Multiple concentrations of IPP3, IPP4 and IPP5 at pH 7.5 were tested and an increase in ThT fluorescence over 6 days was observed for all of them (Figure 6.5). However, in no case were sigmoidal-shaped curves typical of a nucleation-elongation mechanism observed, and it was, therefore, not possible to fit any of the data to equations describing a nucleation-elongation model[209]. In addition, IPP3, IPP4 and IPP5 aggregation curves all lack a flat lag phase, also typical of standard nucleation-elongation fibrillation mechanisms[335]. The aggregation kinetics of IPP3 show two distinct phases in which the ThT intensity increases, Figure 6.5A. The gradient of the first phase of IPP3 aggregation, between 0 and

circa 50 hours, increases with peptide concentration. At lower IPP3 concentrations (25 and 50 µM, purple and blue lines in Figure 6.5A), the two phases are very distinct. The start of the second phase may occur as the concentration of some species formed as part of the first phase accumulates and reaches some critical concentration.

The aggregation kinetics of IPP4 show a gradual increase in ThT fluorescence over time, Figure 6.5B, which cannot be fitted to a standard sigmoidal function. The aggregation kinetics of IPP5, Figure 6.5C, show a non-zero ThT fluorescence intensity at the start of the assay and a rapid increase in fluorescence intensity in the first 10 hours. In this case, after a maximum ThT intensity has been reached, the fluorescence then decreases slowly. These curves can be explained by the presence of oligomeric species at the earliest time points, which can bind to and increase the fluorescence of ThT, and which were reported and characterized in Chapter 5 (Figure 5.8 and Figure 5.9). A smaller oligomer which is formed rapidly after the sample is dissolved binds ThT resulting in the initial, non-zero signal. This then converts slowly into a larger oligomer, which has a greater effect on the ThT fluorescence. The final slow decay in ThT fluorescence intensity may be caused by photobleaching or by the fact that fibrils formed in the later stages of the assay bind ThT less than oligomeric intermediates which prevail in earlier stages.

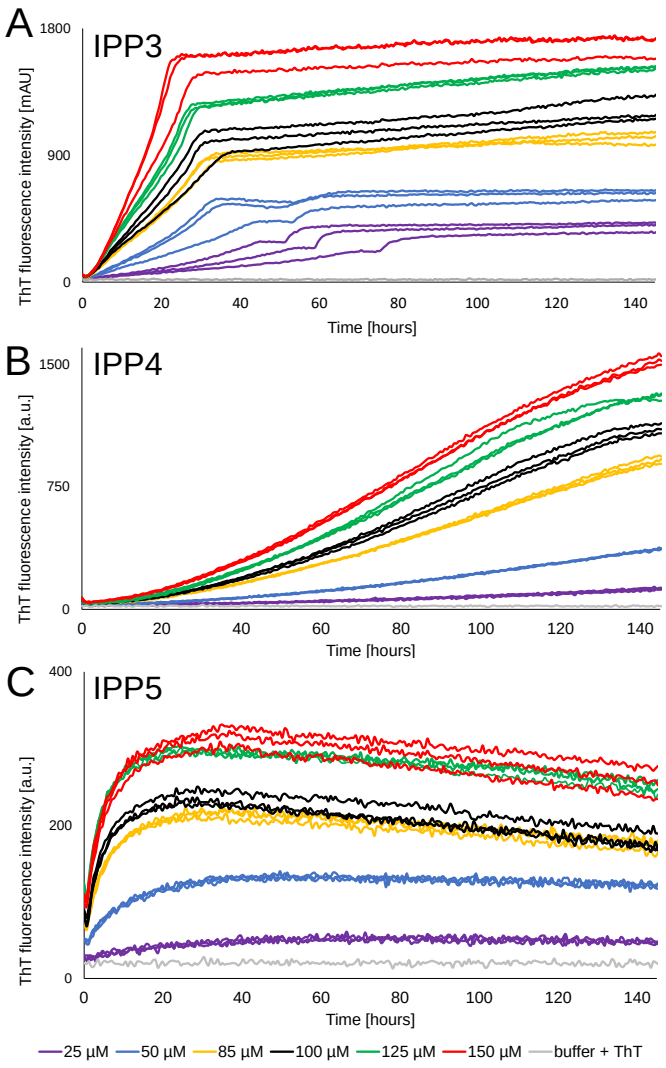


Figure 6.5: Thioflavin T assays – aggregation kinetics of IPP3, IPP4 and IPP5. Different concentrations of IPP3 (A), IPP4 (B) or IPP5 (C) were incubated in 25 mM phosphate, pH 7.5, with 50 µM thioflavin T (ThT) at 37 °C with agitation over 145 hours. ThT fluorescence emission was recorded at 482 nm, after excitation at 448 nm, every 30 minutes. Each sample was measured in triplicate in the same plate.

In addition, the aggregation of lipidated analogues was followed by a fluorescence-based assay employing 8-anilinonaphthalene-1-sulfonic acid (ANS) as a molecular probe. Using an excitation wavelength of 355 nm, ANS is barely fluorescent in polar environments, however, in a hydrophobic environment its fluorescence increases with a fluorescence emission maximum at around 500 nm. ANS has been used in folding and aggregation studies due to its ability to detect exposed hydrophobic regions of proteins/peptides[94,219].

An ANS assay was employed to gain complementary information about the aggregation kinetics of lipidated analogues. Samples of different peptide concentrations, ranging from 10 to 150 µM, were incubated with 250 µM ANS in 25 mM phosphate at pH 7.5 in a plate reader at 37 °C with 180 rpm agitation. All ANS fluorescence curves show a sharp decrease

in intensity during the first hour of incubation, Figure 6.6. As shown in Appendix B, Figure B1 and B2, this is caused by an initial temperature equilibration (25 °C → 37 °C) since the samples were prepared at room temperature and then transferred to the plate reader thermostatted at 37 °C. Consistent with the ThT assays, the aggregation kinetics of IPP3 followed by ANS show two phases which are more distinct at lower peptide concentrations. IPP4 samples show an increase in fluorescence intensity over time in both ANS and ThT assays, Figure 6.5B and Figure 6.6B. However, whereas the ANS assay shows an increase in fluorescence during first 60 hours followed by a plateau, the ThT assay of IPP4 shows a gradual increase in fluorescence intensity without reaching the plateau phase even at 145 hours. These results suggest that different processes are being probed by ANS and ThT, with ANS being more sensitive to the exposure of hydrophobic patches, whereas ThT may primarily probe the increase in β-sheet structure. ANS curves of IPP5 aggregation do not show any significant change in fluorescence intensity in contrast to the ThT curves, this may be explained by a greater sensitivity of ANS to pre-fibrillar oligomers which may be formed rapidly and therefore cause high fluorescence intensity from the start of the assay[81,314,336,337]. In addition, all ANS curves start from non-zero values, which indicates binding of ANS to pre-fibrillar oligomers and/or to the hydrophobic lipid moiety itself.

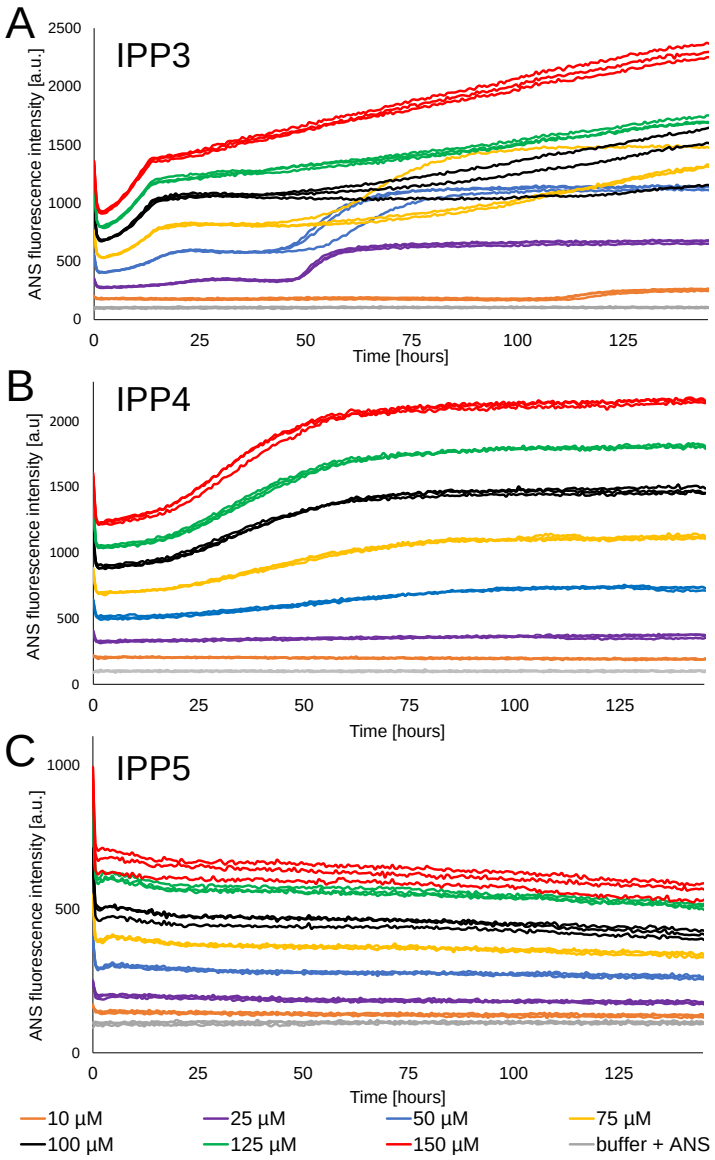


Figure 6.6: 8-Anilinonaphthalene-1-sulfonic acid (ANS) assays – aggregation kinetics of IPP3, IPP4 and IPP5. Different concentrations of IPP3 (A), IPP4 (B) or IPP5 (C) were incubated in 25 mM phosphate, pH 7.5, with 250 µM ANS at 37 °C with agitation over 145 hours. ANS fluorescence was recorded at 482 nm, after an excitation at 355 nm, every 30 minutes. Each sample was measured in a triplicate in the same plate.

The aggregation of lipidated GLP-1-Am analogues was also monitored using size-exclusion chromatography (SEC). IPP3, IPP4 and IPP5 at 135 µM concentration in 25 mM phosphate at pH 7.5 were analysed using a Superdex 200 Increase 10/300 size-exclusion column which has a separation range from 10 to 600 kDa. Samples were analysed immediately after the peptide was dissolved in buffer and at different time points after incubation at 37 °C with 180 rpm agitation. All samples were filtered through 0.22 µm membrane filter prior to their injection onto the column. Figure 6.7 shows the chromatograms for IPP3, IPP4 and IPP5. Species eluted in the void volume of the size-exclusion column (at around 7.5 mL), which are larger than the upper limit of the separation range of the column, were also imaged by

negative-stain transmission electron microscopy (Figure 6.7B, D and F). Over time, all the lipidated analogues show a decrease in the peak intensity corresponding to oligomeric species present from the start, which is compensated by an increase in the peak in the void volume. In addition, the results for IPP4 show a new oligomeric species formed which elutes at around 10 mL, which may represent a metastable intermediate state populated *en route* to the formation of large aggregates. This observation may also be related to the concentration-dependence of IPP4 oligomeric population as shown earlier in Chapter 5.

TEM images of void volume fractions (Figure 6.7B, D, F) show the formation of short fibril-like species. The void-volume fraction of IPP3 shows flexible thread-like structures which have a high tendency to cluster (Figure 6.7B), IPP4 formed only short fibrils (Figure 6.7D). In contrast to IPP3 and IPP4, the void-volume fraction of IPP5 contains longer protofibrils of a uniform width which do not have a tendency for clustering – Figure 6.7F. It is assumed that these can further elongate into the long, rigid fibrils observed after an 8-day incubation which are of a similar width (Figure 6.4E and F).

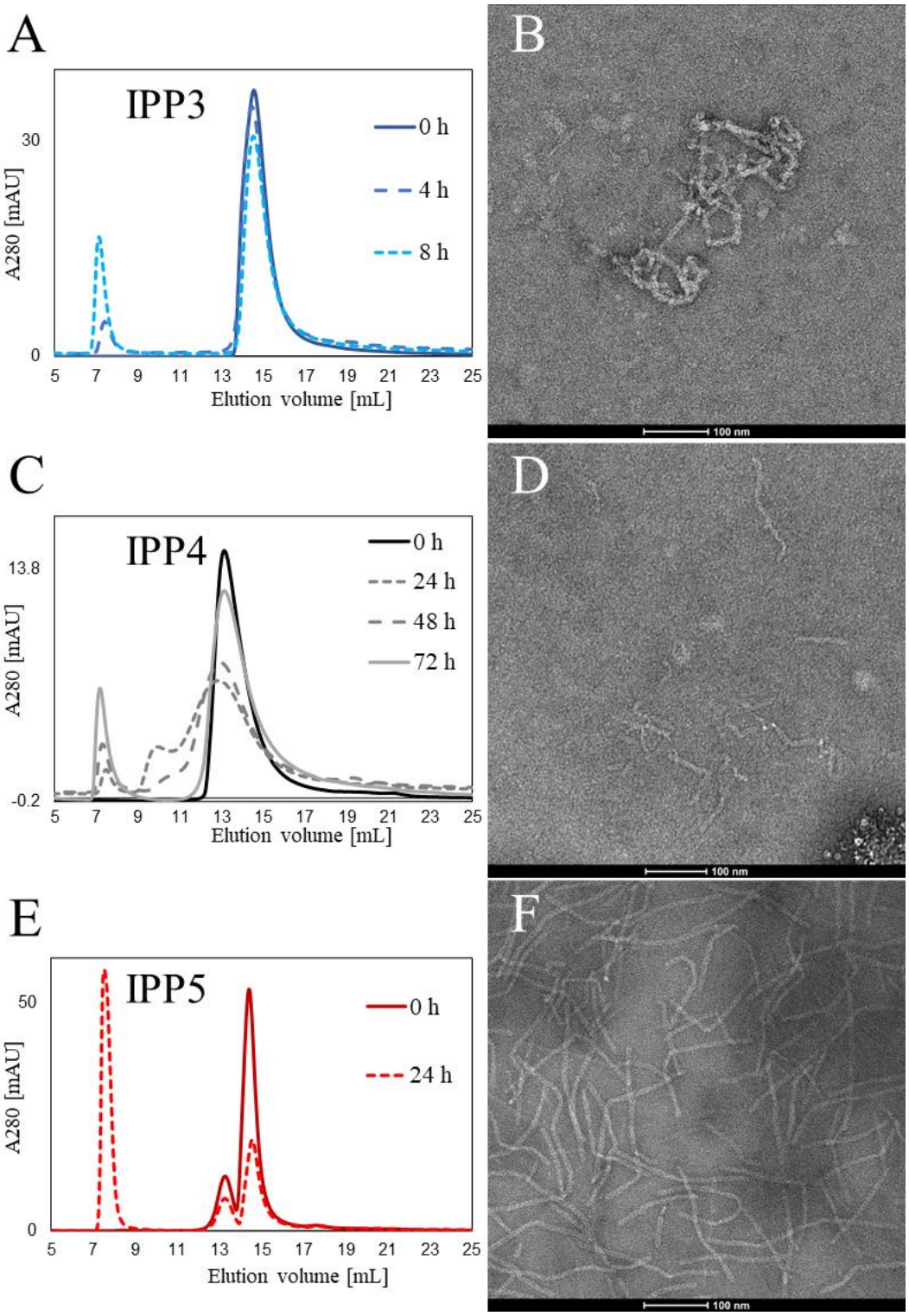


Figure 6.7: IPP3, IPP4 and IPP5 aggregation observed by size-exclusion chromatography and transmission electron microscopy. Size-exclusion chromatography experiments were performed in 25 mM phosphate, pH 7.5, using a Superdex200 Increase column at room temperature. Both freshly prepared and aged samples (incubated at 37 °C with agitation) were filtered through 0.22 µm membrane filter prior to injection onto the column. For each sample, a fraction in the void volume was isolated, applied onto a carbon-coated copper grid, stained with 2% uranyl acetate, and imaged using transmission electron microscopy (B, D, F).

Size-exclusion chromatography was also employed to monitor the interconversion between individual oligomeric populations of IPP5 over time, providing further insight into the overall aggregation mechanism. A freshly prepared sample of IPP5 at pH 7.5 was injected onto a Superdex200 Increase 10/300 column. The two oligomeric peaks with the highest intensity were isolated – Figure 6.8A, red and blue area. Half of each isolated fraction was immediately reinjected onto the column; the rest of the sample was left at room temperature for 24 hours and was subsequently reinjected onto the column – Figure 6.8B and C. As apparent from these figures, oligomer 1 (isolated red fraction) was stable when immediately reinjected onto the column, i.e., the isolated fraction did not re-equilibrate back into the initial oligomeric distribution of injected sample, moreover, oligomer 1 was stable even after the fraction was left at room temperature for 24 prior to its reinjection. In contrast, oligomer 2 (isolated blue fraction) partly interconverted into oligomer 1 immediately. The proportion of oligomer 1 in the fraction seems to slightly increase after 24 hours at room temperature. These observations indicate that oligomer 1 is more stable than oligomer 2, but oligomer 2 is formed faster. In addition, trace amounts of smaller oligomers eluting between 17 and 19 mL were detected using SEC. Oligomer 1, which has a larger hydrodynamic radius than oligomer 2 and was found to have a higher percentage of β-sheet structure (Chapter 5, Figure 5.16), is likely to be an on-pathway intermediate to fibril formation (Figure 6.8D).

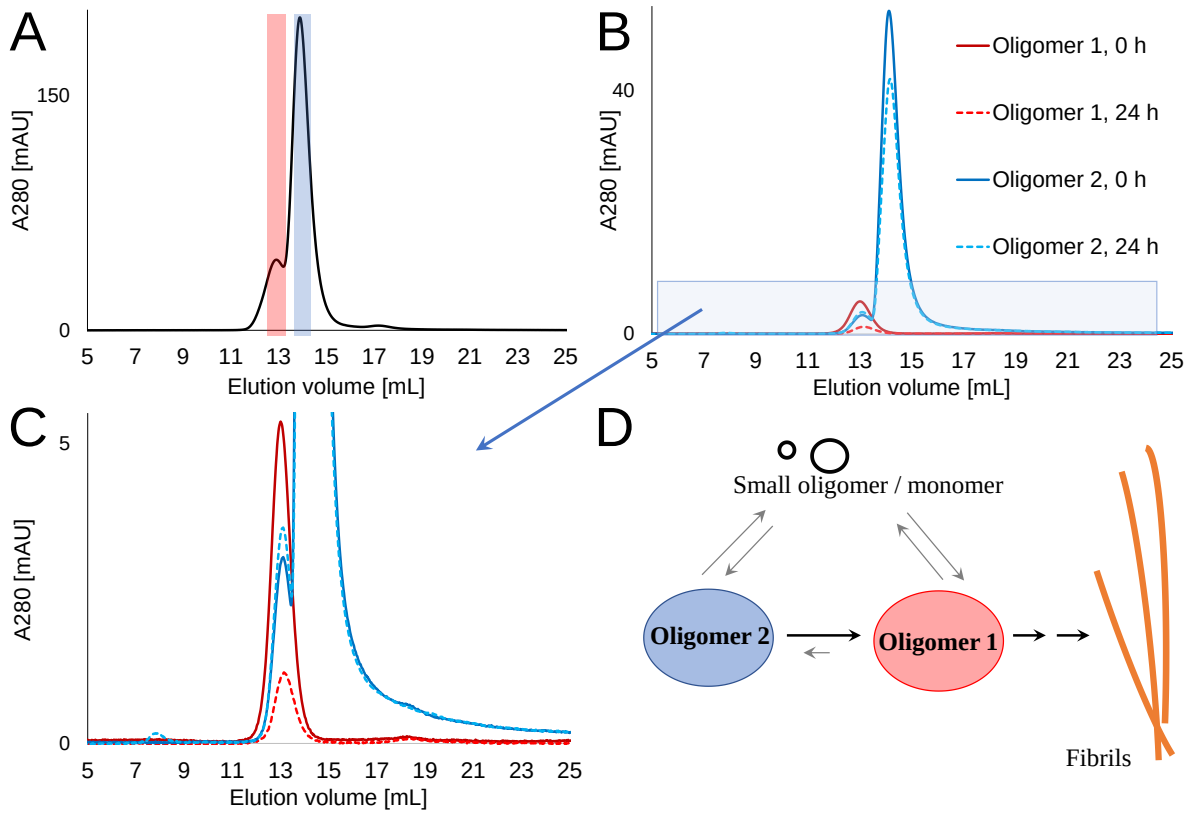


Figure 6.8: Stability of IPP5 oligomers investigated using size-exclusion chromatography. Size-exclusion chromatography experiments were performed in 25 mM phosphate, pH 7.5, using a Superdex200 Increase column at room temperature. Samples were filtered through 0.22 µm membrane filter prior to injection onto the column. Initial concentration of IPP5 was 200 µM (chromatogram in A). Isolated fractions (from A) were either immediately reinjected onto the column or left for 24 hours at room temperature before reinjection (B). Panel C shows a magnification of the area in light blue rectangle from (B). A scheme of IPP5 oligomeric species and their mutual conversion is depicted in (D).

Additionally, the same experiment investigating the oligomeric distribution and oligomer stability was performed at pH 6.5. First, a freshly prepared sample of 170 µM IPP5 was injected onto a Superdex200 Increase column. An IPP5 sample of the same concentration was incubated at 37 °C with 180 rpm agitation for 24 hours, filtered through 0.22 µm membrane filter, and injected onto the column. As shown in Figure 6.9A, the oligomeric distribution in the freshly prepared sample at pH 6.5 differs from pH 7.5 mainly by the presence of a broad peak eluting between 8 and 11 mL. After 24 hours of incubation this broad peak disappears but an intense peak elutes in the void volume of the column (at around 7.5 mL) indicating the formation of a new species with a higher hydrodynamic radius. The void volume peak was collected, and imaged using negative-stain TEM (Figure 6.9B) showing the presence of relatively short thread-like structures, most likely amyloid fibrils/protofibrils with similar widths to fibrils/protofibrils observed at pH 7.5.

The broad peak, observed in the chromatogram of a freshly prepared sample of IPP5 at pH 6.5, was isolated (Figure 6.9C – fraction indicated by a green rectangle) and reinjected onto the column either immediately after its separation or after the sample was left at room temperature for 24 hours, Figure 6.9D. As apparent from Figure 6.9D, in the sample which was reinjected immediately after its isolation, the peak position remains the same. However, in samples that were reinjected after 24 hours, a new oligomeric distribution emerges, consisting of the short fibrillar species eluting in the void volume and species eluting at around 14 mL. The peak eluting at around 14 mL is also present in the chromatogram of a freshly prepared sample – Figure 6.9A and C. This observation indicates a slow re-equilibration of the isolated species either to fibrils and/or smaller oligomers which were previously observed in the mixture. The TEM image of the fraction containing a broad peak (Figure 6.9C, green rectangle area) shows the presence of pre-fibrillar worm-like oligomers with a high content of β-structure as predicted from its far-UV CD spectrum, Figure 6.10.

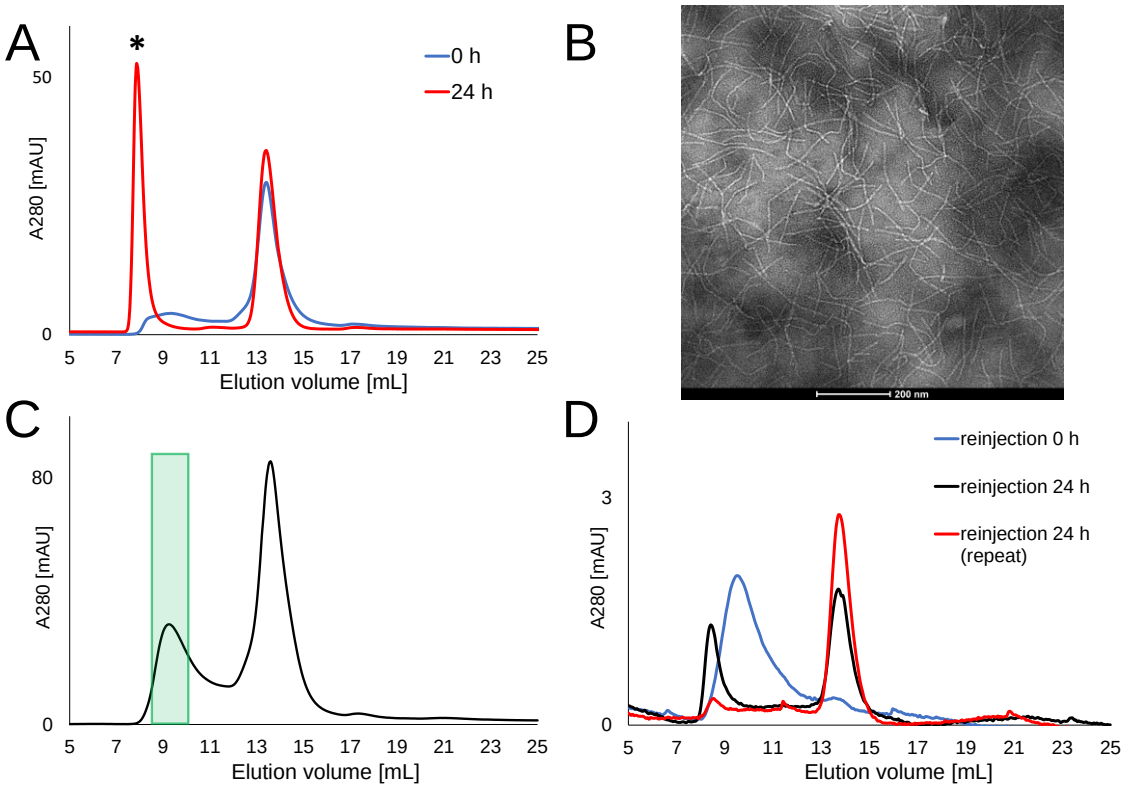
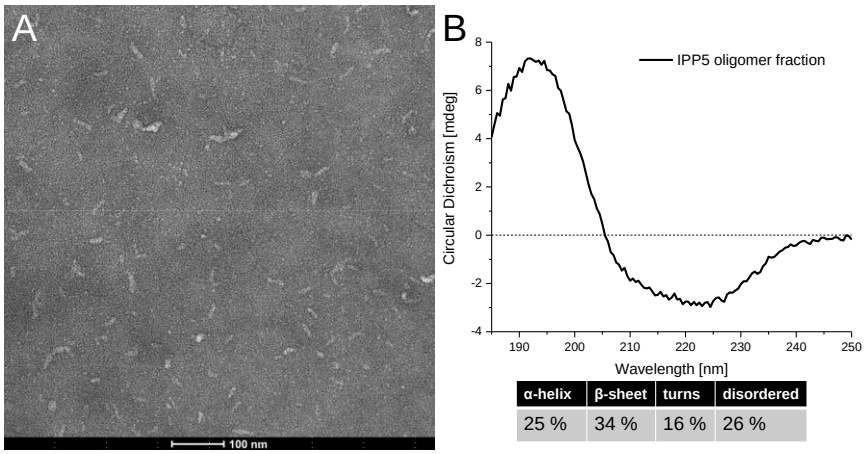


Figure 6.9: Oligomeric distribution and aggregation of IPP5 at pH 6.5. Size-exclusion chromatography experiments were performed with 170 µM IPP5 in 25 mM phosphate, pH 6.5, using a Superdex200 Increase column. Samples were filtered through 0.22 µm membrane filter prior to the injection onto the column. The void volume peak (A, red line, peak marked with asterisk) at around 7.5 mL was isolated, applied onto a carbon-coated copper grid, stained with 2% uranyl acetate and imaged using transmission electron microscopy (B). Isolated fractions (C, green area) were pooled and either immediately reinjected onto the column or left for 24 hours at room temperature before reinjection (D).



α-helix	β-sheet	turns	disordered
25 %	34 %	16 %	26 %

Figure 6.10: Morphological and structural characterization of IPP5 oligomers formed at pH 6.5. IPP5 oligomer formed in 25 mM phosphate at pH 6.5 was isolated using SEC (**Figure 6.9**C, green area). The isolated sample of 5 µM peptide was imaged using negatively stained TEM (A) and structurally characterized using far-UV CD (B). The sample was measured in a 0.1 cm pathlength cuvette. An estimate of the secondary structure was made using DichroWeb, Contin LL method, dataset 3[180–182].

6.3 pH-dependent aggregation of lipidated GLP-1-Am analogues

6.3.1 pH affects the aggregation mechanism and aggregate morphology of IPP3

As shown in Section 6.2.2, at pH 7.5, the IPP3 analogue has rather unusual aggregation kinetics with distinct phases, Figure 6.5A. The aggregation of IPP3 was studied at different pH values to further explore this behaviour.

The aggregation of different concentrations of IPP3 was monitored using ThT assays and was performed at pH 7.5, 8.0 (both 25 mM phosphate) and 8.5 (25 mM Tris), Figure 6.11. Aggregation profiles at pH 7.5 and 8.0 show similar characteristics – two phases are observed which have different initial slopes, the aggregation being faster at pH 8.0 compared to pH 7.5. At pH 8.5, the aggregation kinetics changes. The transition between the two different phases observed at pH 7.5 and 8.0 is no longer distinguishable, especially at higher peptide concentrations. In addition, at higher peptide concentrations, the ThT fluorescence shows a slight decrease after 110 hours, which may be due to dye degradation, which is faster under basic pH conditions[338], and/or photobleaching.

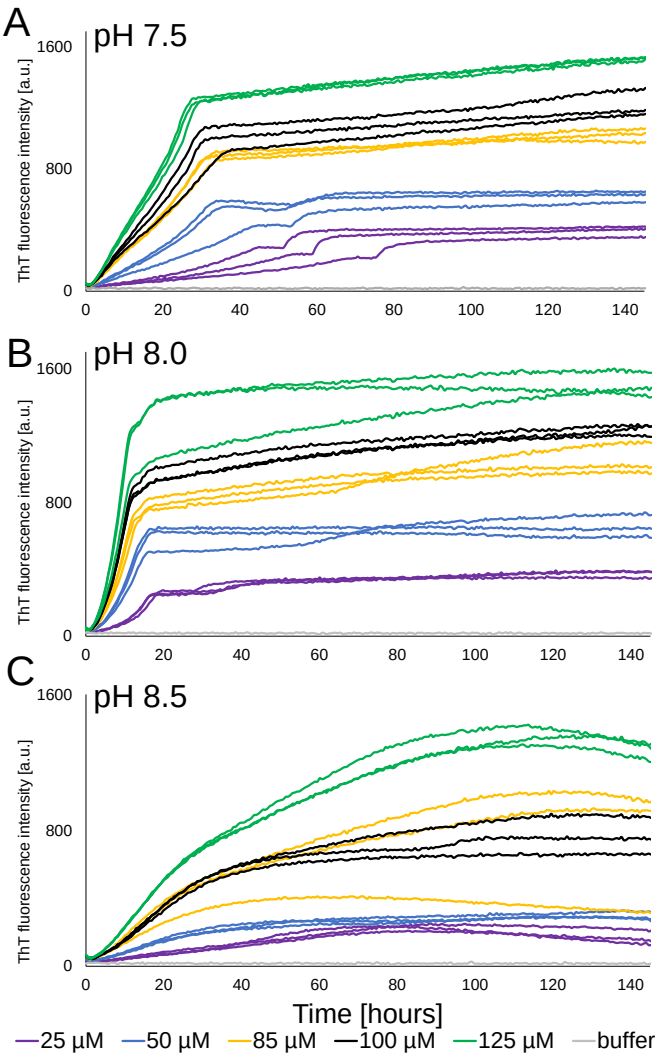


Figure 6.11: Aggregation of IPP3 at different pH values monitored by thioflavin T assays. A range of IPP3 concentrations were incubated at 37 °C with agitation over 6 days and monitored by ThT fluorescence. Samples were prepared in 25 mM phosphate at pH 7.5 (A) and 8.0 (B) and 25 mM Tris at pH 8.5 (C). The ThT fluorescence emission was recorded at 482 nm, after excitation at 448 nm, every 30 minutes. Each sample was measured in triplicate within the same plate.

The morphology of aggregates at the end point of the ThT assays was imaged using negative-stain TEM. A new ThT assay was performed using 25 µM IPP3 at pH 7.0, 7.5 and 8.5, Figure 6.12. IPP3 shows a limited solubility at around pH 7 with the highest achievable concentration of approx. 30 µM. At pH 7.0, at the end of the 6-day assay, IPP3 had formed compact aggregates of an amorphous shape which may consist of short fibril fragments. At pH 7.5, short, curved fibrils with a tendency to form clusters were observed. At 8.5, longer fibrils with a lower tendency to cluster were formed.

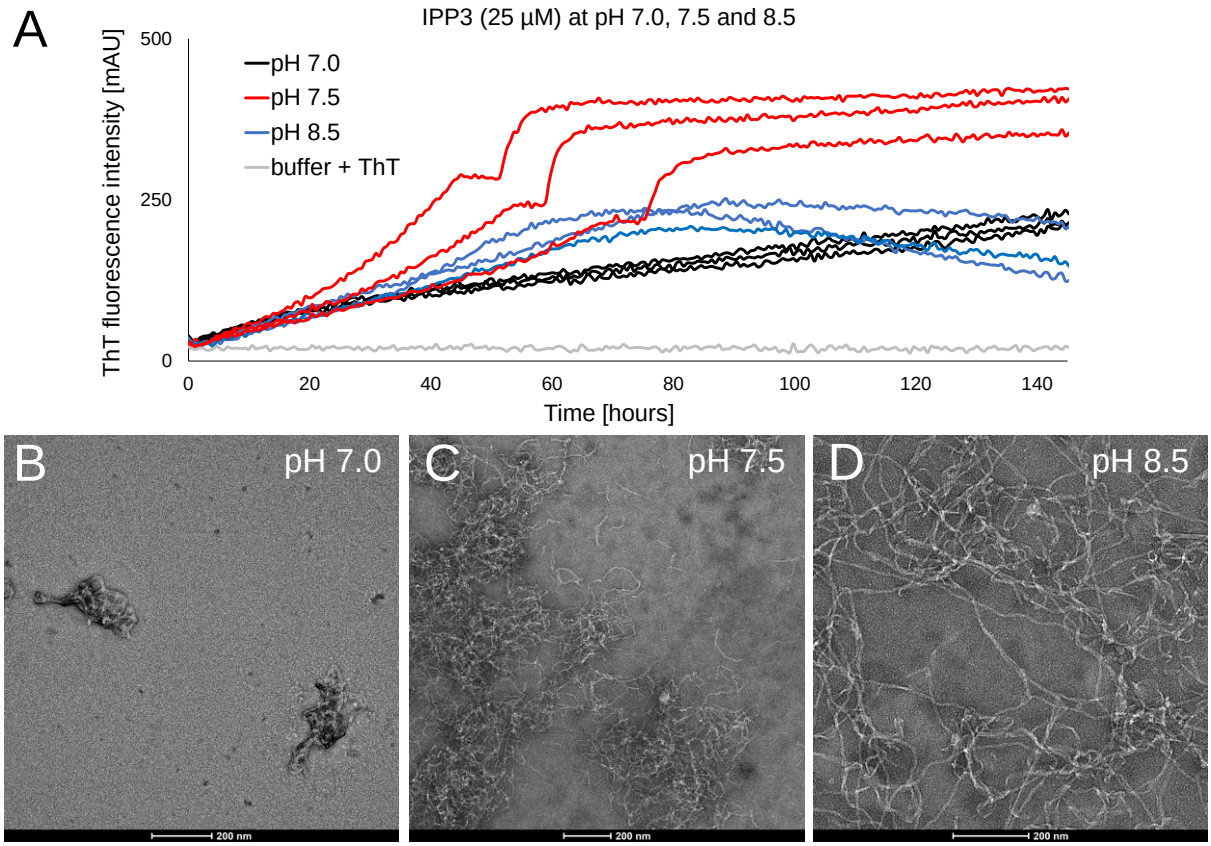


Figure 6.12: IPP3 aggregation kinetics and aggregate morphology at different pH values. The aggregation of 25 µM IPP3 in 25 mM phosphate at pH 7.0 and 7.5 and 25 mM Tris at pH 8.5 was monitored by ThT assays (A). Samples with 50 µM ThT dye were incubated at 37 °C with agitation for 145 hours. ThT fluorescence was recorded at 482 nm, after excitation at 448 nm, every 30 minutes. Each sample was measured in triplicate within the same plate. At the end of the assay, the samples were imaged using negative-stain TEM (B, C, D).

6.3.2 IPP4 at its solubility threshold at pH 7: Aggregation of toroidal assemblies into amorphous aggregates

IPP4 was shown to form regular toroidal assemblies at pH 7 – Chapter 5, Section 5.6.2. This behaviour occurs at a pH where the peptide has a limited solubility, only concentrations lower than 50 µM can be attained. In this Section, the aggregation behaviour of these toroidal assemblies over time is described.

A sample of 40 µM IPP4 in 25 mM phosphate at pH 7 was imaged using TEM shortly after sample preparation (20 minutes) and after 8 days of incubation at 37 °C with 180 rpm agitation – Figure 6.13. These results reveal that the smaller regular toroidal assemblies (green arrow, Figure 6.13A & B) slowly convert to larger irregular species (Figure 6.13B, red arrow), in addition to forming amorphous aggregates (Figure 6.13B, blue arrow). The

diameter of larger irregular species is between 70 and 90 nm whereas the diameter of small toroidal assemblies is in the range of 12–27 nm as determined from TEM images.

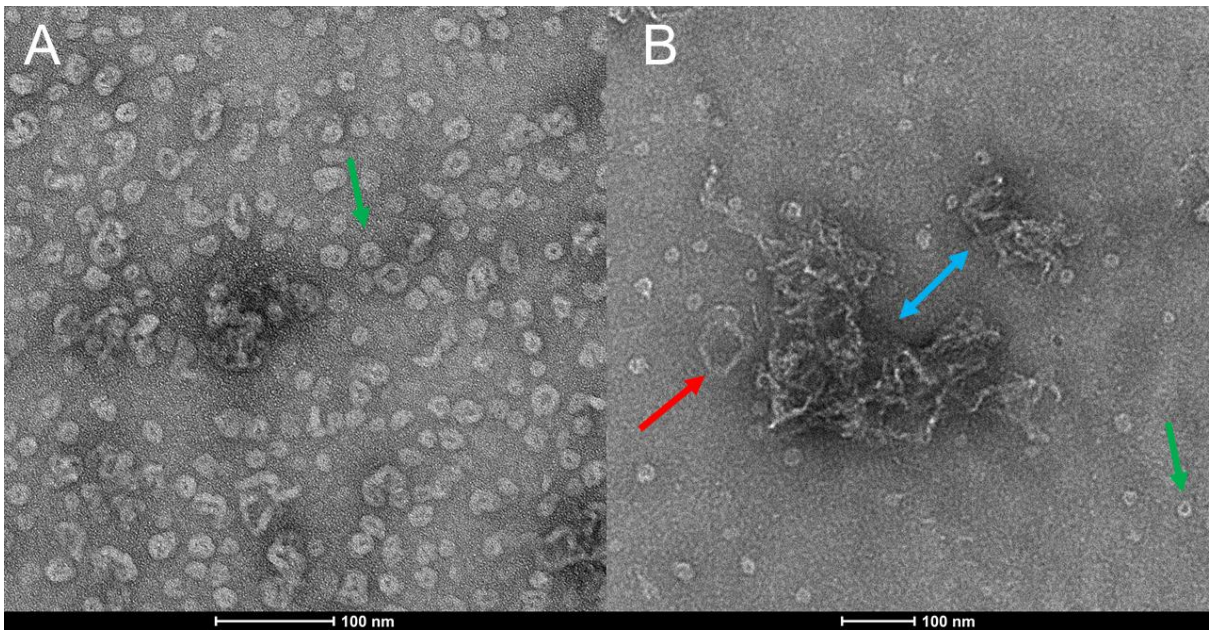


Figure 6.13: Aggregation of IPP4 toroidal assemblies imaged using transmission electron microscopy. Freshly prepared sample of 40 µM IPP4 in 25 mM phosphate at pH 7.0 (A) and the same sample incubated for 8 days at 37 °C with agitation (B) were applied onto a carbon-coated copper grid, stained with 2 % uranyl acetate and imaged using transmission electron microscopy.

To obtain information about the changes in secondary and tertiary structure that occur during this reaction, far-UV CD and intrinsic tryptophan fluorescence were measured, Figure 6.14. The analysis of far-UV CD spectra shows a slight decrease in α-helical content after the 8-day incubation, 43 % → 32 %, which correlates with an increase in β-sheet structure (10 % → 24 %) and disordered regions (21 % → 26 %), suggesting that α-helical structure in the smaller, regular toroidal assemblies is lost upon formation of the larger, irregular species and/or the short, curly fibrils, which have more β–sheet and disordered secondary structure.

As shown in Figure 6.14B, the λ_{max} of the tryptophan fluorescence did not significantly change during the incubation. This indicates that the tryptophan side chain has similar degrees of burial in all the different species observed.

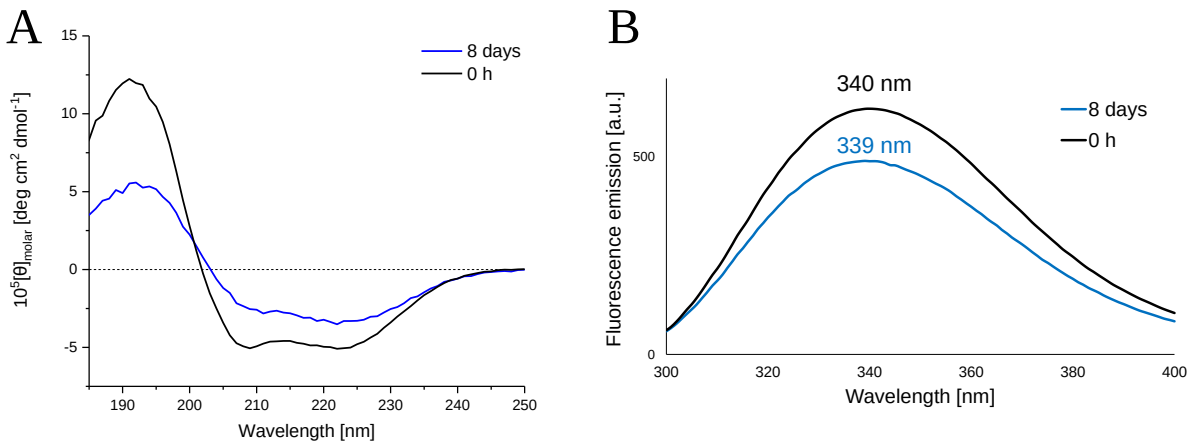


Figure 6.14: Aggregation of IPP4 toroidal assemblies monitored by far-UV circular dichroism and intrinsic tryptophan fluorescence. Far-UV CD (A) and intrinsic tryptophan fluorescence (B) spectra were measured for a freshly prepared sample of 40 µM IPP4 in 25 mM phosphate at pH 7.0 and for the same sample aged for 8 days at 37 °C with agitation.

The reaction was also monitored using ANS and ThT assays. Figure 6.15A shows the results of an ANS binding assay with 50 µM IPP4 at different pH values over 6 days at 37 °C with agitation. The ANS curve obtained at pH 7 differs from those observed at higher pH values in both its fluorescence intensity and shape. Two distinct kinetic phases are apparent at pH 7, which may correspond to i) the conversion of the initial smaller toroidal assemblies into larger hollow species and ii) the subsequent aggregation of these species into amorphous aggregates. Figure 6.15B, shows the aggregation kinetics of 50 µM IPP4 at pH 7.0 and 7.5 as probed by ThT fluorescence. Unlike in the ANS assay (Figure 6.15A), only a single phase is observed in the ThT assays with a slow increase of ThT fluorescence emission over time. This observation suggests an increase in β-sheet structure over time.

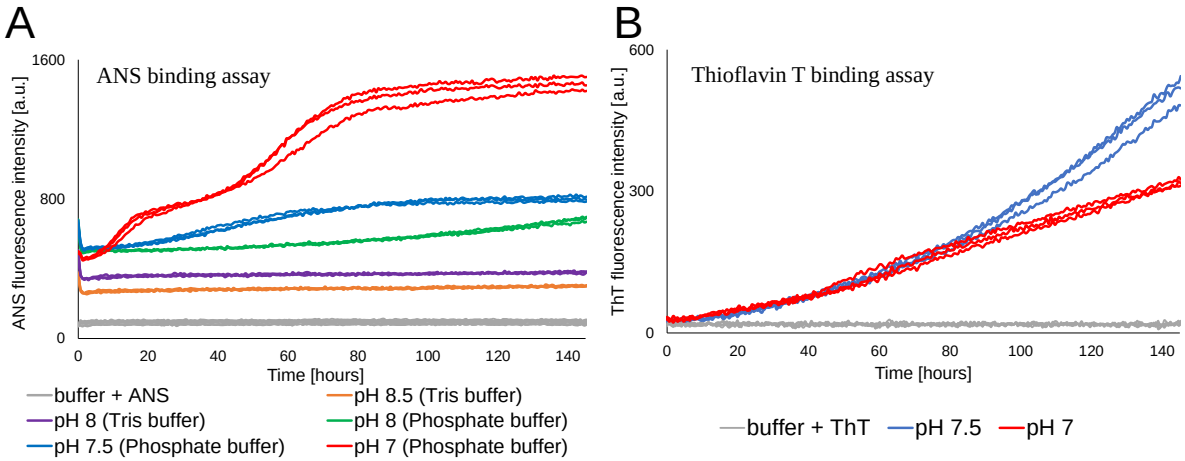


Figure 6.15: Kinetics of IPP4 aggregation – ANS and ThT binding assays. Assays were performed in 25 mM phosphate or Tris of a corresponding pH with either 250 µM ANS (A) or 50 µM ThT (B). Samples of 50 µM IPP4 at pH 7.0, 7.5, 8.0 and 8.5 were incubated for 145 hours at 37 °C with agitation and readings were taken every 30 minutes. Each sample was measured in triplicate within the same plate.

The aggregation of IPP4 toroidal assemblies formed at pH 7 was also monitored using size-exclusion chromatography. Figure 6.16 shows the chromatograms of IPP4 in 25 mM phosphate, pH 7, at different time points after incubation at 37 °C. The UV trace of a freshly prepared sample (in black) shows a single intense peak eluting at around 9.5 mL corresponding to a homogeneous population of small toroidal assemblies. Over time, a second peak eluting in the void volume of the column appears which indicates the formation of larger species, presumably the larger hollow species and/or amorphous aggregates observed by TEM. After 72 hours of incubation at 37 °C, the ratio of species eluting at 9.5 mL and in the void volume was approximately equal indicating the existence of multiple species in the sample at this time.

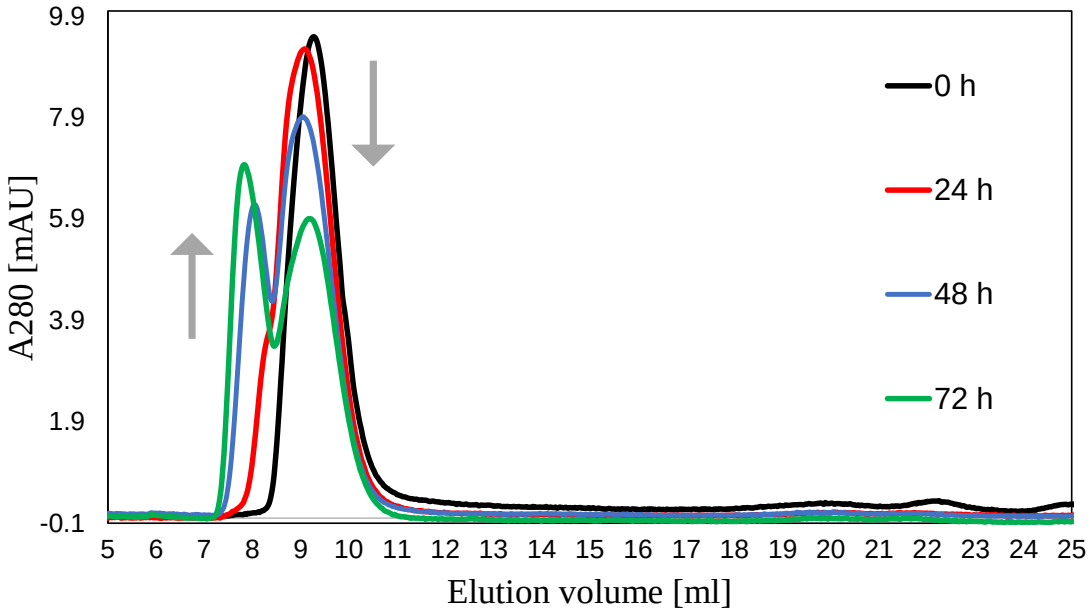


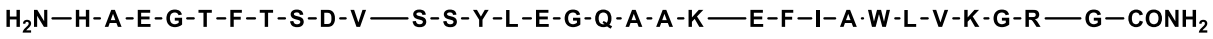
Figure 6.16: Aggregation of small, regular IPP4 toroidal assemblies into larger species monitored by size-exclusion chromatography. Samples of 40 µM IPP4 in 25 mM phosphate at pH 7.0 were injected at different time points, after their incubation at 37 °C with agitation, onto a Superdex200 Increase column using a 200 µL injection loop. All samples were filtered through a 0.22 µm membrane filter prior to injection onto the column.

6.3.3 IPP2 analogue – rapid formation of amyloid fibrils

The experiments with the lipidated IPP2 analogue described in this Section were performed by a Part III student, Irina-Alexandra Edu, under my supervision and assistance. The following data have also been published in a Part III Project Report[339].

The IPP2 analogue of GLP-1-Am contains the same lipidation moiety as the other analogues studied in this Chapter, i.e., palmitic acid attached to a lysine residue via a γ-glutamic acid linker. However, the position of lipidation within the peptide backbone differs from the other analogues studied. IPP2 is lipidated close to its N-terminus at position 2. In addition, this analogue contains an extra leucine-to-arginine substitution at position 26 (Figure 6.17). The IPP2 analogue also greatly differs in the pH range over which it is soluble. Whereas IPP3, IPP4 and IPP5 are soluble at neutral and basic pH values (Chapter 5, Figure 5.3), IPP2 is soluble only at around pH 3 and lower.

GLP-1-Am



IPP2

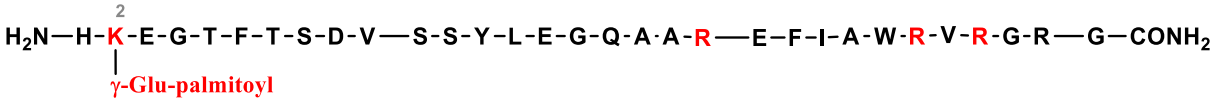


Figure 6.17: Sequences of GLP-1-Am and IPP2. Amino acid substitutions and a lipidation site are highlighted in red.

A set of biophysical experiments was performed to characterize the properties and aggregation behaviour of IPP2 at pH 3. Using FT-IR spectroscopy, IPP2 was characterized as i) a lyophilized powder, ii) 20 minutes after its dissolution in 25 mM phosphate at pH 3 and subsequent filtration through 0.22 µm membrane filter, and iii) after 8 days of incubation at pH 3 in 37 °C with agitation, Figure 6.18A. Samples prepared or incubated in aqueous buffer were subsequently flash-frozen and lyophilized before the IR spectrum was recorded. The FT-IR absorption spectrum was recorded between 1370 and 1770 cm^{-1} covering amide I and amide II regions. In the IR spectrum of the IPP2 lyophilized powder, the maximum of amide I absorption band lies at around 1655 cm^{-1} which is typical of α-helical structure. In contrast, after the sample is dissolved in buffer, the amide I maximum is shifted towards *circa* 1616 cm^{-1}, an indication of β-sheet structure. The IR spectrum of the sample which was flash-frozen and lyophilized shortly after its dissolution in the buffer has a shoulder at around 1655 cm^{-1} which reflects some residual α-helical structure. This shoulder is missing in the sample recorded after an 8-day incubated suggesting all the α-helical structure has been lost.

The high β-sheet content in the two aqueous samples was also confirmed by far-UV CD spectra which were recorded using the same conditions and which showed broad minima at around 215–220 nm. Samples were also imaged using TEM. Short, curly (proto)fibrils were observed directly after dissolution of the lyophilised powder in aqueous buffer, Figure 6.18C, whereas samples incubated for 8-days show longer and more rigid fibrils and additional spherical aggregates, Figure 6.18D.

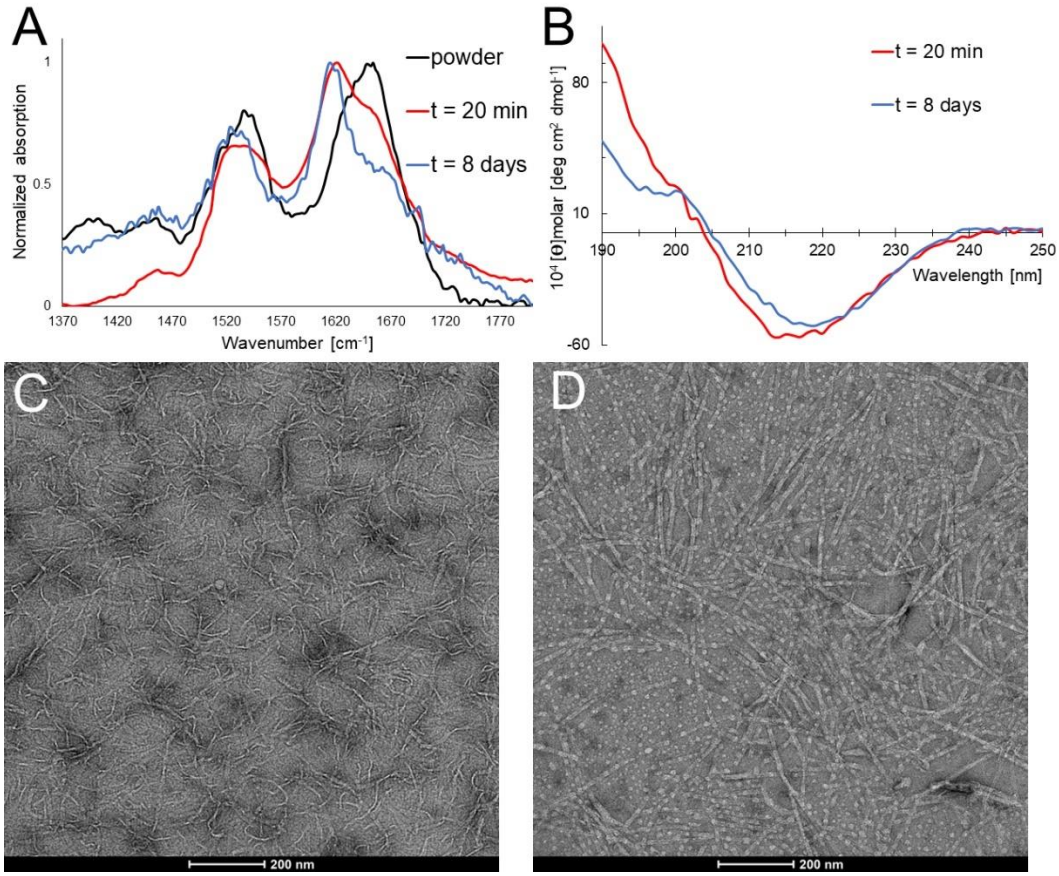


Figure 6.18: Structure and morphology of species formed by IPP2. (A) FT-IR spectra of solid lyophilized samples of IPP2. Spectra of the original lyophilised IPP2 as well as samples that were flash-frozen and lyophilized either shortly after dissolution in the buffer or after 8-days of incubation at 37 °C. (B) Far-UV CD spectra of IPP2 samples at 25 µM in 25 mM phosphate at pH 3. Far-UV CD spectra were measured in a 0.1 cm pathlength cuvette, and the CD signal was converted into concentration-independent molar ellipticity, $[\theta]_{molar}$, units, see Eq. 2.6 (Chapter 2). (C, D) TEM images of uranyl acetate-stained samples. Samples of 25 µM IPP2 in 25 mM phosphate at pH 3 were applied onto carbon-coated copper grids either *circa* 20 minutes after sample preparation (C) or after 8-days of incubation at 37 °C with agitation (D).

The aggregation kinetics of IPP2 were monitored using ThT and ANS binding assays. Both assays were performed in 25 mM phosphate buffer at pH 3 and included samples of IPP2 at multiple concentrations. The ThT assay, Figure 6.19A, shows low levels of ThT fluorescence which do not change over time. In contrast, the ANS assay, Figure 6.19B, shows high ANS fluorescence which does not change over time. The ANS assays show a steep increase in fluorescence within the first hour, which is due to temperature re-equilibration as the samples were prepared at room temperature and the assay was performed at 37 °C.

Neither ThT or ANS binding assays, in this case, reflect the changes during the maturation process of fibrils occurring under these conditions, as proven by TEM and partly also IR and CD spectroscopy, Figure 6.18. The different results, in particular the lack of ThT

fluorescence, may be due to electrostatic interactions between the dye and the aggregate[217]. At pH 3, ThT has a positive charge whereas ANS is negative, causing either repulsive or attractive interactions with the positively charged peptide/peptide aggregates, respectively.

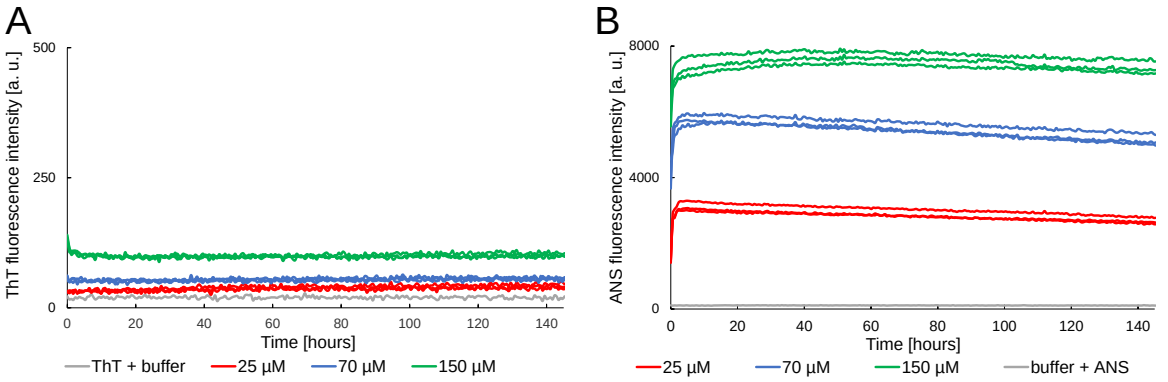


Figure 6.19: ThT and ANS binding assays of IPP2 aggregation. Assays were performed in 25 mM phosphate at pH 3 with either 50 µM ThT (A) or 250 µM ANS dye (B). Three different concentrations of IPP2 were monitored in both assays – 25, 70 and 150 µM. Samples were incubated for 145 hours at 37 °C with agitation and readings were taken every 30 minutes. Each sample was measured in triplicate in the same plate.

6.3.3.1 Effect of Lys26 to Arg26 substitution on peptide properties

To deconvolute the contribution of lipidation and amino-acid substitutions in the peptide sequence to the properties and aggregation of the IPP2 analogue, a non-lipidated peptide with the same sequence as IPP2 was investigated and was found to have a solubility maximum at around pH 3. Above pH 3, the solubility of the peptide is limited with the highest achievable peptide concentration at around 50 µM (in a pH range from 3 to 8.5), Figure 6.20.

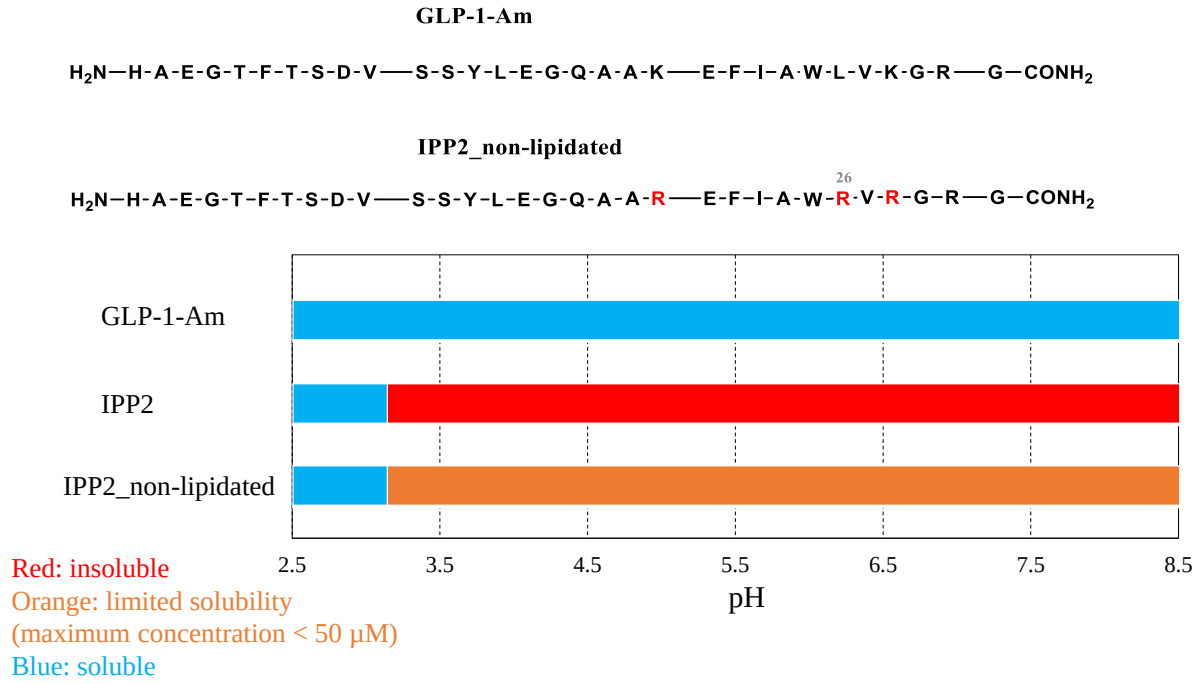


Figure 6.20: Sequence and solubility chart of non-lipidated peptide with peptide sequence identical to IPP2. The peptide solubility was determined spectrophotometrically. Samples were filtered through a 0.22 µm membrane filter prior to the absorbance measurement. The peptide under specific conditions is defined as "soluble" if a concentration above 50 µM can be achieved. The peptide is defined as "insoluble" if it is not possible to attain concentrations above 1 µM. "Limited solubility" stands for concentrations from 1 to 50 µM.

The aggregation of the non-lipidated IPP2 analogue was monitored by ThT assays following the same method as IPP2. Samples of freshly prepared non-lipidated IPP2 in 25 mM phosphate at pH 3 (25, 70, 85, 150 µM) and 7.5 (25 µM) were used. At pH 3, the non-lipidated IPP2 analogue shows only a low propensity for aggregation with only the 150 µM sample showing some aggregation after *circa* 100 hours, Figure 6.21. In contrast, at pH 7.5, the 25 µM sample shows a considerably higher aggregation propensity. Moreover, the aggregated peptide was observed to precipitate forming white flakes under these conditions.

In conclusion, the lack of a lipid attachment, the fact that Lys2 in the non-lipidated is charged but in the lipidated peptide it is neutral, as well as potentially the Lys to Arg substitution at position 26 (although it is noted that there is no change in the charge in this case – just the charge distribution) affect the solubility of the peptide. However, the amino-acid substitutions in the IPP2 sequence, are not likely to be the main cause of the rapid aggregation observed at pH 3 (Figure 6.18).

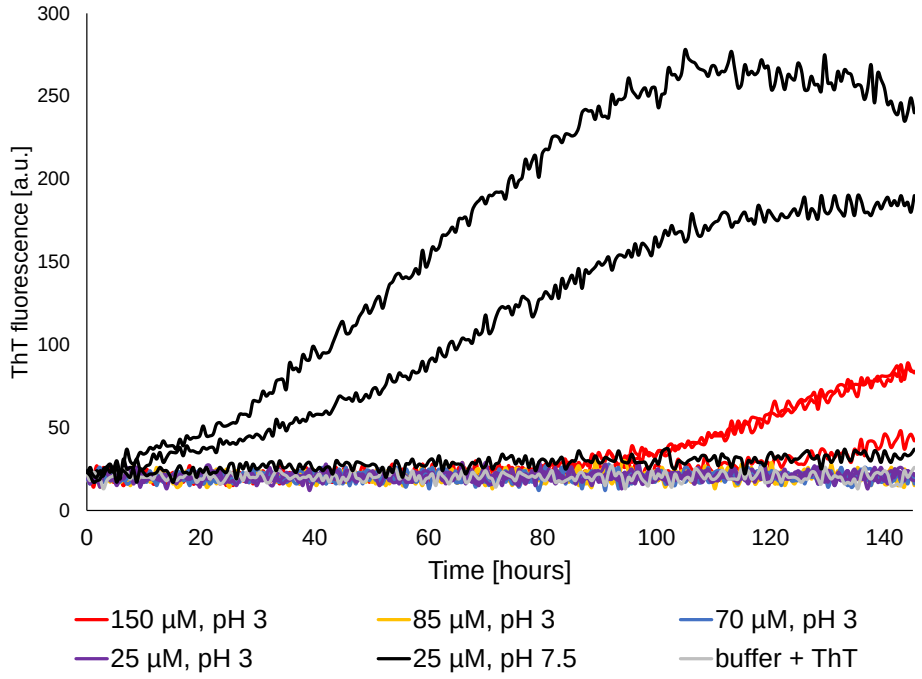


Figure 6.21: Aggregation of non-lipidated IPP2 peptide monitored by ThT assay. The aggregation of non-lipidated IPP2 analogue in 25 mM phosphate, pH 3 (25, 70, 85 and 150 µM) and 7.5 (25 µM), was monitored by ThT assay over 8 days at 37 °C with agitation. ThT fluorescence was recorded at 482 nm, after excitation at 448 nm, every 30 minutes. Each sample was measured in triplicate within the same plate.

6.4 IPP10 analogue – effect of lipidation moiety on peptide properties: comparison with the IPP5 analogue

The experiments with the lipidated IPP10 analogue described in this Section (excluding FT-IR and sedimentation velocity) were performed by a Part III student, Irina-Alexandra Edu, under my supervision and assistance. The following data have also been published in a Part III Project Report[339].

This Section is focused on the comparison of the previously studied IPP5 analogue with the IPP10 analogue, which, apart from substituting Ala2 with 2-aminoisobutyric acid (Aib), differs only by the type of lipidation moiety and its linker, Figure 6.22. IPP10 has a higher net charge compared to IPP5 due to the additional carboxylate group in the lipidation moiety. Apart from C-terminal amidation, the IPP10 analogue is the same as the commercially-available drug – semaglutide[22].

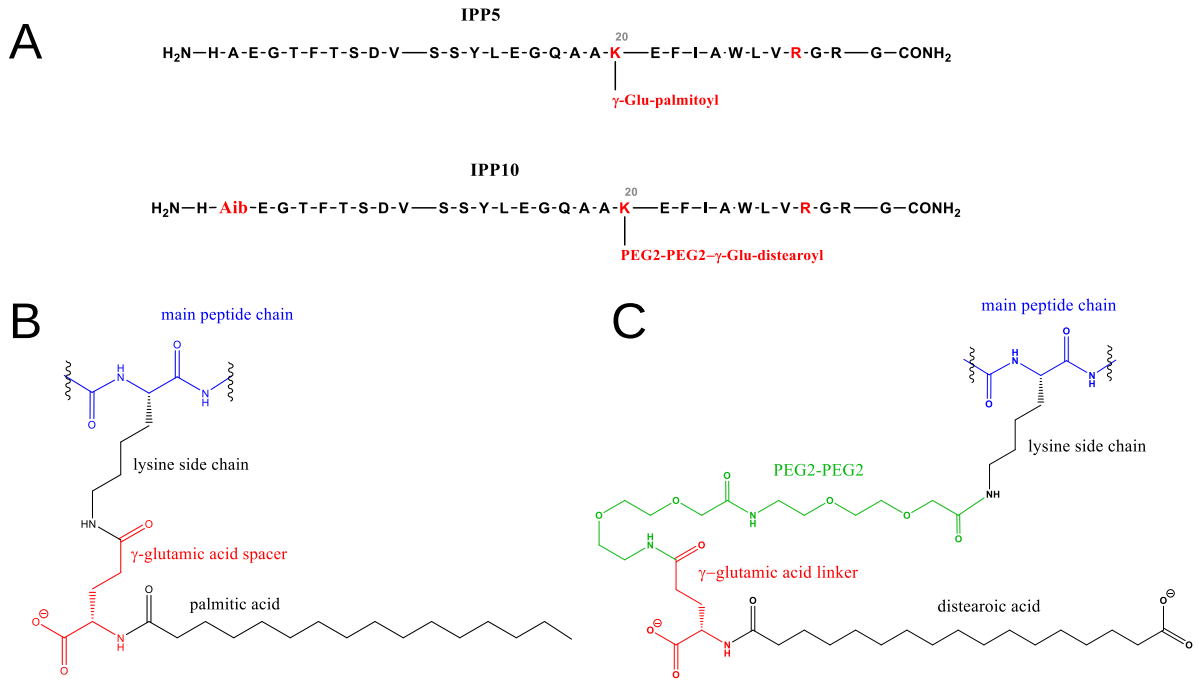
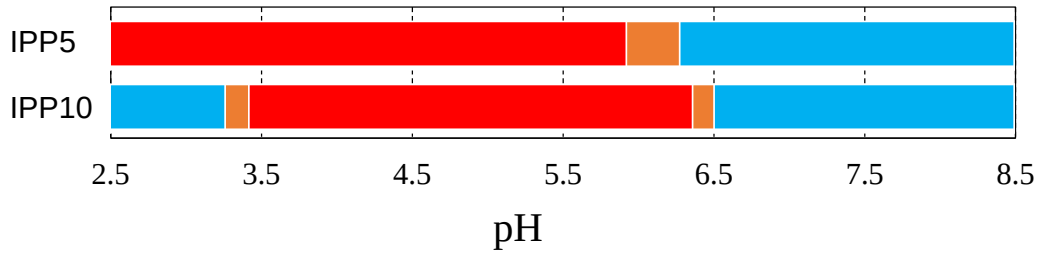


Figure 6.22: Comparison of IPP5 and IPP10 sequence, lipidation sites and lipidation moieties. (A) IPP5 and IPP10 sequences – residues which differ in the sequence from non-lipidated GLP-1-Am are highlighted in red. "Aib" at position 2 in the sequence of IPP10 stands for 2-aminoisobutyric acid. (B) Detail of the linker and lipid in the IPP5 analogue. (C) Detail of the linker and the lipid in the IPP10 analogue.

Solubility of both IPP5 and IPP10 was tested in a pH range from 2.5 to 8.5, Figure 6.23. A solution of buffer (25 mM phosphate, citrate or Tris) of a corresponding pH was added to the lyophilized peptide powder, gently mixed and left for *circa* 5 minutes at room temperature before the solution was filtered through a 0.22 µm filter. Subsequently, the concentration of the peptide was measured on a spectrophotometer using absorption at 280 nm.

Both IPP5 and IPP10 are soluble at pH values from 6.5 and higher – this solubility window is slightly wider for IPP5 which starts to be soluble at around pH 6. IPP10 has an additional solubility window at around pH 3, which most likely is due to the extra negative charge at the lipidation site due to the linker used (in comparison to IPP5).



Red: insoluble
Orange: limited solubility (maximum concentration < 50 µM)
Blue: soluble

Figure 6.23: The comparison of IPP5 and IPP10 solubility over a pH range from 2.5 to 8.5. The solubility in a buffer of a corresponding pH was determined from an absorption measurement after sample filtration through a 0.22 µm membrane filter. Concentrations were determined spectrophotometrically. The peptide under specific conditions is defined as "soluble" if concentrations above 50 µM can be achieved. The peptide is defined as "insoluble" if it is not possible to attain concentrations above 1 µM. "Limited solubility" stands for concentrations from 1 to 50 µM.

The secondary structure of both peptides was studied using FT-IR and far-UV CD, Figure 6.24. First, solid samples of the lyophilized peptide powder were analysed using FT-IR. Figure 6.24A shows the FT-IR spectra in the amide I and amide II region of lyophilized powders of IPP5 and IPP10. For both peptides, the maximum of amide I peak is located at around 1650 cm^{-1} indicating α-helical and/or disordered (random coil) structure[193,310]. Overall, the FT-IR spectra did not show any major differences in secondary structure between IPP5 and IPP10. Far-UV CD spectra were recorded for 85 µM samples freshly prepared in 25 mM phosphate at pH 7.5, Figure 6.24B. The far-UV CD spectra of fresh samples in solution also did not show any major differences in secondary structure between these two peptides.

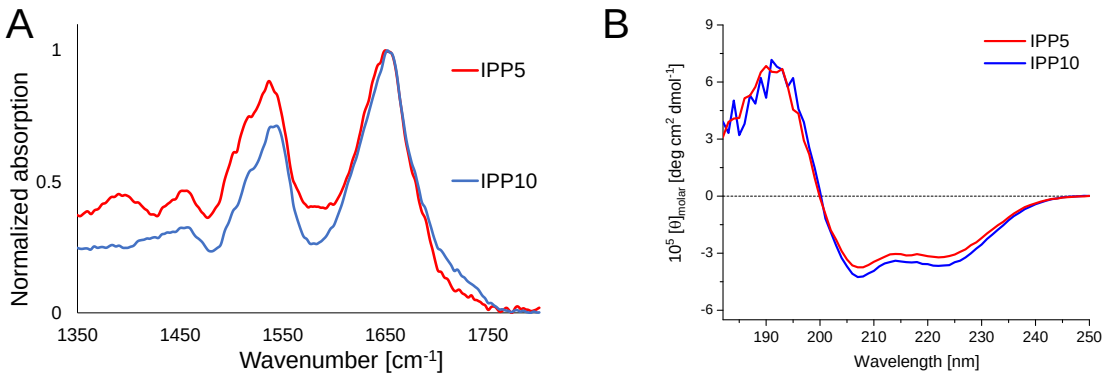


Figure 6.24: IPP5 and IPP10: FT-IR spectra of peptide powders and far-UV CD spectra of freshly prepared samples at pH 7.5. FT-IR spectra (A) have been normalized to the highest amide I absorption. Far-UV CD spectra (B) were recorded for freshly prepared 85 µM IPP5 and IPP10 in 25 mM phosphate at pH 7.5. The CD signal was converted into molar ellipticity.

The distribution of oligomeric species in freshly prepared samples at pH 7.5 was monitored using size-exclusion chromatography and sedimentation velocity experiments, Figure 6.25. Samples for size-exclusion chromatography were freshly prepared at 135 µM concentration in 25 mM phosphate at pH 7.5 (Figure 6.25A) and samples for sedimentation velocity experiments (Figure 6.25B) were prepared at 85 µM concentration under the same conditions. Differences in oligomer distribution between IPP5 and IPP10 were observed by both techniques. In size exclusion chromatography, IPP5 elutes in two main peaks (studied in detail in Chapter 5, Section 5.6.1) whereas the IPP10 chromatogram contains only a single peak. Consistent with this, a single dominant species was observed in sedimentation velocity experiments which had a sedimentation coefficient at around 2.5 S.

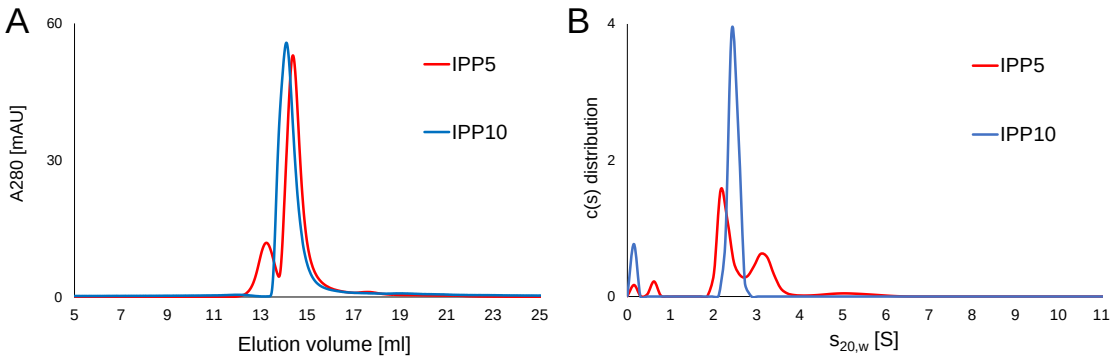


Figure 6.25: Size-exclusion chromatography and sedimentation velocity plots of freshly prepared samples of IPP5 and IPP10 at pH 7.5. (A) SEC chromatograms of freshly prepared IPP5 and IPP10 samples at 135 µM concentration in 25 mM phosphate at pH 7.5 were measured using a Superdex200 Increase 10/300 column. (B) Sedimentation velocity experiments were performed with freshly prepared IPP5 and IPP10 samples at 85 µM concentration in 25 mM phosphate at pH 7.5. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$).

The aggregation behaviour of IPP5 and IPP10 was tested using ThT assays, Figure 6.26. Different concentrations of either IPP5 or IPP10 in 25 mM phosphate at pH 7.5 were

incubated over 6 days at 37 ° with agitations. IPP5 samples show a fast increase in fluorescence intensity whereas in the case of IPP10 the fluorescence stayed mostly low and constant. As was previously shown in this Chapter (Figure 6.4E and F), IPP5 aggregates into fibrillar structures. For IPP10, the far-UV CD spectra did not show any changes between the freshly prepared sample and the sample after 8-day incubation at 37 °C with agitation; TEM images of the sample after 8-day incubation also did not reveal any visible aggregated species – these data are shown in Appendix B, Figure B3.

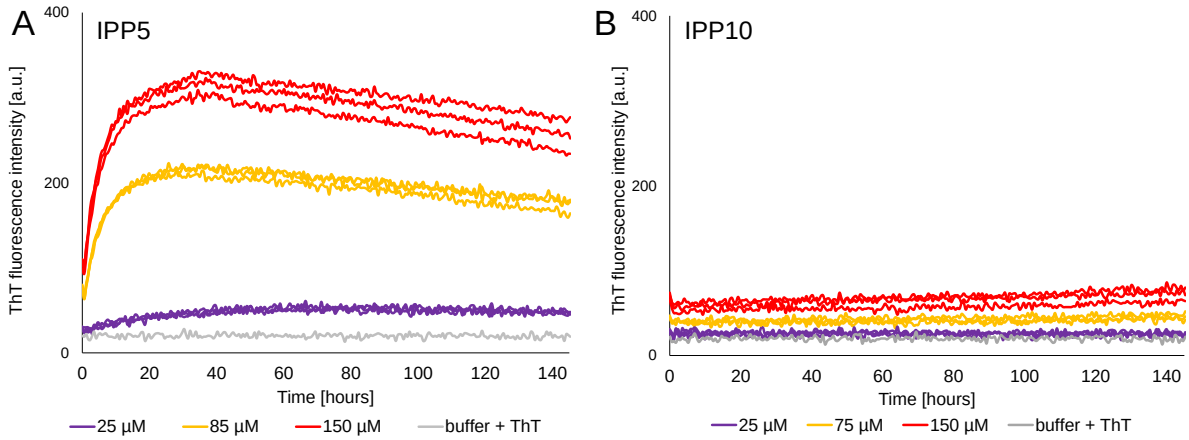


Figure 6.26: ThT assay of IPP5 and IPP10 at pH 7.5. (A) ThT assay of IPP5. (B) ThT assay of IPP10. Assays were performed in 25 mM phosphate at pH 7.5 with 50 µM ThT. Samples were incubated for 145 hours at 37 °C with agitation and readings were taken every 30 minutes. Each sample was measured in triplicate in the same plate.

The nature of the linker and the lipid moiety is the most significant difference between IPP5 and IPP10 analogues. However, a minor change in amino-acid sequence at position 2 (substitution of Ala with Aib) may also affect the physical stability of the peptide. The oligomeric distribution and aggregation behaviour of non-lipidated GLP-1-Am and GLP-1-Am_Aib, which differ only by this substitution at position 2, was assessed using sedimentation velocity experiments and ThT assay (Figure 6.27 and Figure 6.28). Noticeable differences were observed in both experiments – Aib substitution induced a change in the distribution of oligomers and lowered the fibrillation propensity compared to GLP-1-Am.

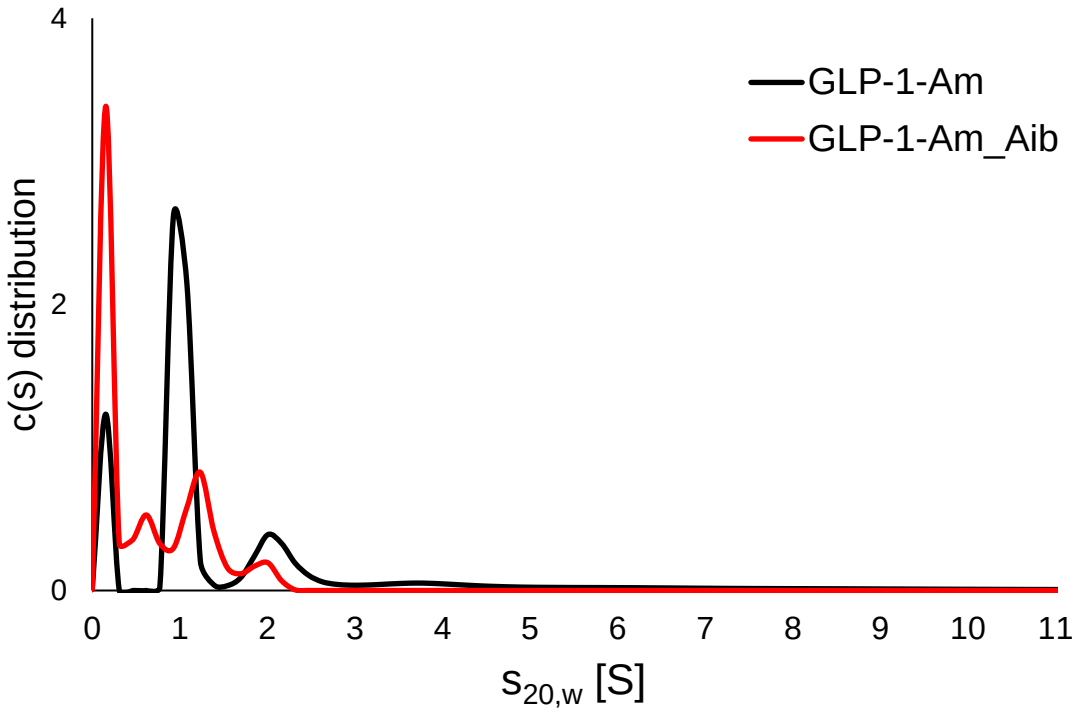


Figure 6.27: Sedimentation velocity plot of GLP 1 Am and GLP 1 Am_Aib at pH 7.5. Samples were freshly prepared at 85 µM peptide in 25 mM phosphate at pH 7.5. After a 2-hour temperature equilibration of samples in the centrifuge to 20 °C, the experiment was performed with centrifugation at 50 000 rpm. The interference sedimentation curves were fitted to a continuous c(s) distribution model implemented in a Sedfit program. The sedimentation coefficient was corrected for the standard state of water at 20 °C ($s_{20,w}$).

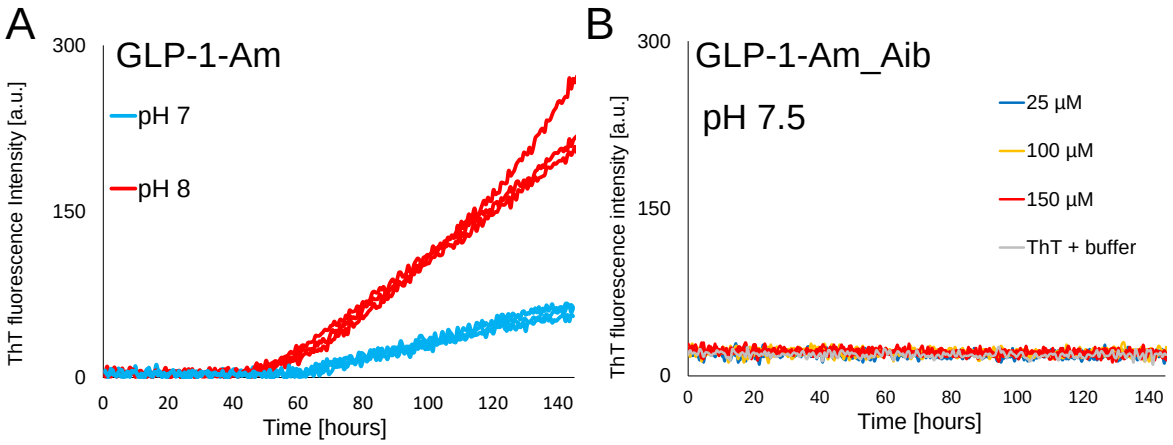


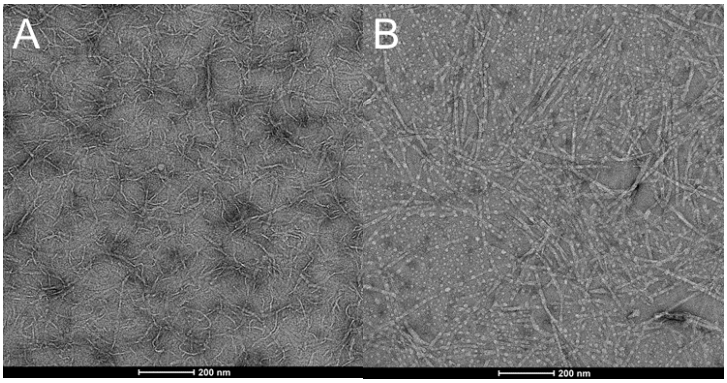
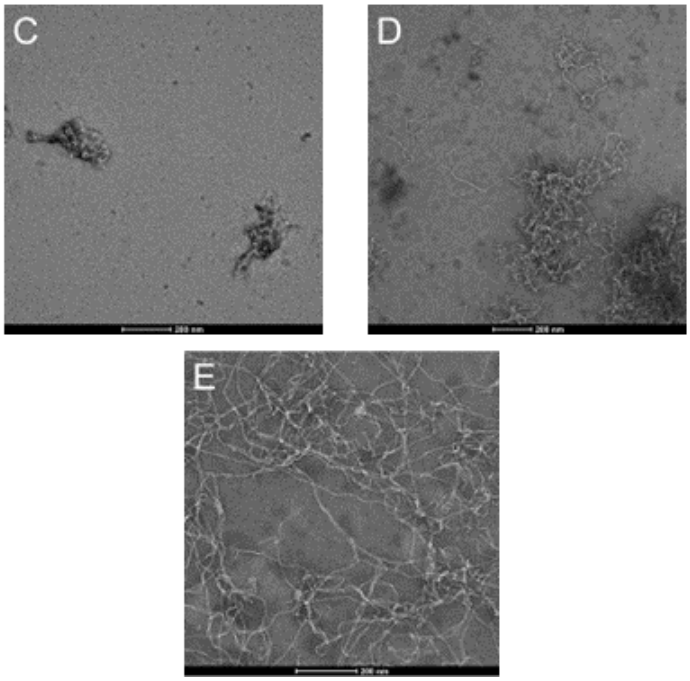
Figure 6.28: ThT assays of GLP 1 Am and GLP 1 Am_Aib at neutral pH. (A) GLP-1-Am at 85 µM concentration was incubated in: pH 7 (25 mM sodium phosphate, shown in blue), and pH 8 (25 mM sodium phosphate, shown in red), at 37 °C with agitation over 6 days. (B) GLP-1-Am_Aib samples at 25, 100 and 150 µM concentration were incubated at pH 7.5 in 25 mM phosphate at 37 °C with agitation over 6 days. All samples were measured in triplicate.

6.5 Discussion

This Chapter describes the aggregation processes which five lipidated GLP-1 analogues undergo. Apart from IPP2 where the aggregation into fibrils is rapid and IPP10 which has greater physical stability and aggregate on even longer timescales, the other three analogues undergo progressive aggregation when the sample is aged over 8 days at 37 °C with agitation. In addition, the aggregation described in this Chapter is accompanied by changes in secondary structure of the peptide in contrast with the "self-assembly" processes which were characterized in Chapter 5, and which were observed in freshly prepared samples.

6.5.1 Morphological classification of species observed during aggregation

The common characteristic of all the species observed during and after the aggregation of all the lipidated GLP-1 analogues studied is an increase in β-sheet structure, Table 6.1. However, a range of aggregate morphologies was observed depending on the lipidation position and pH. Table 6.3 provides an overview of the different types of species and their morphologies for each analogue. It should be noted that because the aggregation of IPP2 is rapid, TEM images and description of species of both a freshly prepared solution and an aged sample are shown. In all other cases, only images of the aged samples (either after 24 hours or 6-8 days) are shown and described.

TEM images of aggregates	**Morphology description**
IPP2	
	A) short curly protofibrils formed immediately after IPP2 dissolution at pH 3 B) fibrils and spherical aggregates of IPP2 formed after 8 days at pH 3
IPP3	
	C) compact aggregates of IPP3 consisting of short fibrillar species observed after 6 days at pH 7.0 D) short fibrils or thread-like aggregates of IPP3 with a tendency for clustering observed after 6 days at pH 7.5 E) long flexible fibrils of IPP3 observed after 6 days at pH 8.5

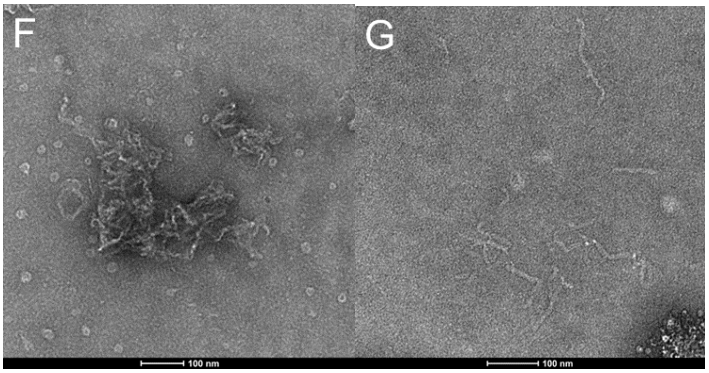
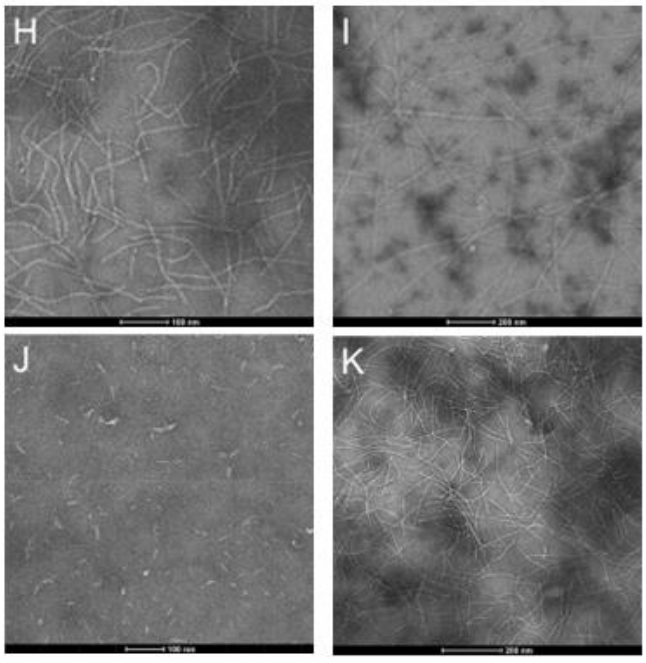
<table>
<tr><td colspan="2">IPP4</td></tr>
<tr><td> 
</td><td>F) aggregates formed by IPP4 toroidal assembly aggregation at pH 7, along with some of the small toroidal assemblies, which are the dominant species immediately after dissolution.
G) short fibrils formed at pH 7.5 after 8 days</td></tr>
<tr><td colspan="2">IPP5</td></tr>
<tr><td> 
</td><td>H) protofibrils formed at pH 7.5 after 24 hours and separated using SEC
I) mature rigid fibrils of IPP5 formed at pH 7.5 after 8 days
J) worm-like oligomers observed after peptide dissolution at pH 6.5; oligomeric fraction separated using SEC
K) protofibrils formed at pH 6.5 after 24 hours and separated using SEC</td></tr>
</table>

Table 6.3: Observed morphologies of aggregated species of lipidated GLP-1 analogues.

6.5.2 Aggregation of lipidated GLP-1-Am analogues – general trends

The highly amphiphilic character of the lipidated peptides promotes their rapid self-assembly into oligomeric species, Chapter 5, such that they form a variety of different oligomeric species (from heptamers to much larger oligomers) in freshly prepared solutions. In addition, all the lipidated peptides undergo further self-association (in this case, we use the term aggregation to distinguish it from the rapid self-association described in Chapter 5) to form several different aggregated states, results described in this Chapter. The resulting aggregates showed morphological diversity from amyloid-like fibrils with differing properties to amorphous aggregates, however, they all showed higher β-sheet content compared to non-aggregated peptides, Table 6.1 and Table 5.1. Nevertheless, the far-UV CD spectra of age samples varied dependent on the position of lipidation in the peptide sequence, Table 6.1. Whereas IPP3 and IPP4 aggregates contained a higher content of disordered structure, around 40 %, IPP5 aggregates showed only around 16 % of disordered regions. During the aggregation, all the analogues lose a large proportion of α-helical structure. For the IPP5 analogue, however, about 20 % of α-helical structure was maintained even in completely aggregated samples which is a significantly larger proportion compared to other analogues. Together with the fact that IPP5 forms mature and rigid fibrils, this suggest that IPP5 fibrils contain a compact β-sheet core, however, the flanking regions, which are not in the amyloid core, are still partly α-helical. IPP3 and IPP4 aggregates containing a high percentage of disordered regions did not form long rigid fibrils as observed for IPP5, but rather clusters of short fibrils, Table 6.3. Lipidation at position 12 (IPP3) or 17 (IPP4) may disfavour fibril maturation. In contrast, IPP5 is lipidated at position 20 which does not seem to disrupt the fibril structure suggesting the C-terminal end of the peptide may not form part of the fibrillar core. It is interesting to note, that the other peptide lipidated far from the centre of the peptide is IPP2 (lipidated at position 2) and it too is capable of the formation of long, rigid fibrils.

The aggregation kinetics of lipidated GLP-1 analogues do not show classical sigmoidal aggregation curves when monitored by ThT assay. Such sigmoidal growth profiles are characteristic of many systems which form amyloid fibrils and results from a nucleation-elongation mechanism in which there is a lag phase which is needed for sufficient fibril nuclei to form before rapid growth occurs[84,335,340]. In the case of lipidated GLP-1 analogues, the lag phase was missing for all the concentrations studied, Figure 6.5. This observation may be explained by the population of oligomers which are formed rapidly upon

peptide dissolution for all the analogues studied. Due to the high local peptide concentration, the structural conversion of these oligomeric species and subsequent aggregation may occur faster compared to the systems which form lower amounts or smaller oligomeric species. The presence of these oligomeric species in solution may also affect the aggregation kinetics due to changes caused in the mechanisms of secondary growth processes (i.e., surface-catalyzed nucleation, elongation after fragmentation of fibrils) which are monomer-dependent as they include monomeric species which elongate/form new fibrils[76,341].

6.5.3 Aggregation of IPP3

At neutral pH, the aggregation of IPP3 results in clusters of short amyloid fibrils, Table 6.3. Compared to the IPP5 analogue, fibrils of IPP3 are visibly shorter, less rigid and they tend to assemble into larger clusters, Table 6.2. We hypothesize that the lipid side chain of IPP3 attached at position 12 is likely to sterically disfavor fibril maturation and therefore formation of long rigid fibrils in contrast to IPP5 which is lipidated at position 20. At basic pH values, the morphology of the IPP3 aggregates changes and the fibrils become longer albeit still less rigid compared with IPP5 fibrils. This observation may be related to the change of electrostatic interactions, specifically deprotonation of tyrosine at position 13 which may cause an additional repulsive interaction with the negative charge on γ-glutamic acid linker which is in the proximity of Tyr13. Although the pK_a of a tyrosine side chain is around 10, a partial deprotonation starts to occur at around pH 8 and the pK_a may be lowered due to the environment of the side chain.

It is also possible that multiple aggregate morphologies are present on the same energy landscape with a certain morphology being kinetically trapped and the other thermodynamically favoured. It has been shown for a lipidated peptide with the sequence V3A3K3 conjugated to a 16-carbon alkyl chain at the N-terminus that at low intermolecular repulsion, i.e., low peptide concentration and low ionic strength, long fibrils with β-sheets are favoured and monodisperse short fibrils represent a metastable state. At high intermolecular repulsion regime, short fibrils represent the thermodynamically favoured product[342].

The aggregation of IPP3 was monitored using a thioflavin T assay. At pH 7.5, an unusual growth profile with two distinct aggregation phases was obtained, Figure 6.29. Two concentrations of IPP3, 125 µM and 25 µM, representing the “high and low concentration

regime" are shown. The multiple phases observed must relate to multiple aggregation processes which can be either sequential (Figure 6.29a) or can occur in parallel (Figure 6.29b). In the sequential case, the second process occurs after the completion of the first process as the "product" of the first process is also the "reactant" in the second process. Another possibility is parallel pathways involving two processes, the processes can be dependent or independent of each other as illustrated in Figure 6.29b. Both processes are likely to have different kinetics and the final aggregation profile may represent a sum of the two processes.

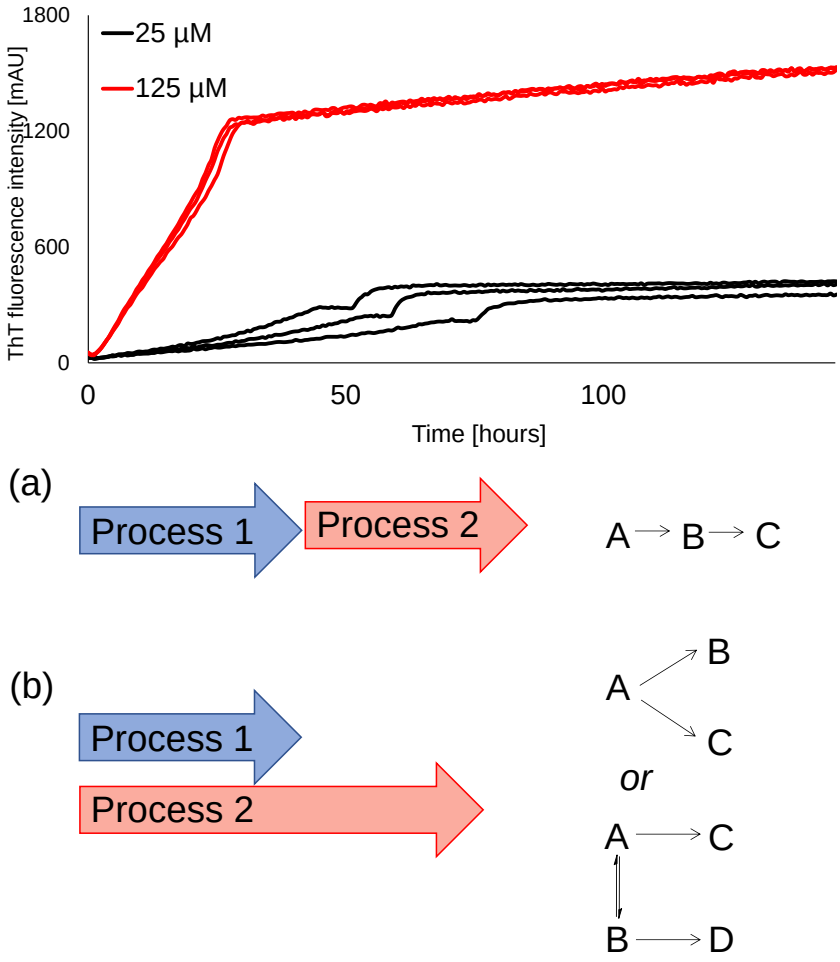


Figure 6.29: Aggregation profile of IPP3 at pH 7.5 and suggested kinetic schemes. IPP3 aggregation was monitored for samples incubated at 37 °C in 25 mM phosphate (pH 7.5) over 6 days using a ThT assay. Two concentrations of IPP3 are plotted – 125 µM representing the "high concentration regime" and 25 µM representing the "low concentration regime". Each concentration was measured in triplicate, run in the same plate. Aggregation kinetics observed likely result from multiple aggregation processes. Possible kinetic schemes illustrating different mechanisms are depicted in (a) reactions in series and (b) parallel pathways.

6.5.4 Aggregation of IPP4

As previously shown (Figure 6.4, Figure 6.7, Table 6.3), at pH 7.5, IPP4 aggregates into short fibril-like species, however, these species do not appear to be capable of maturation into longer, more rigid fibrils. This may be an effect of the lipidation at position 17, which favours the formations of shorter and less ordered species.

IPP4 was also observed to form amorphous aggregates at pH 7. These amorphous aggregates result from the aggregation of regular toroidal species which are formed rapidly under these conditions, Figure 6.13. The amorphous aggregates contain a higher percentage of β-structure compared to the original toroidal assemblies, Figure 6.14. The process of aggregate formation shows kinetics which are distinct from classical nucleation-elongation mechanisms characteristic of amyloid fibril formation, Figure 6.15. A thioflavin T assay showed a gradual increase of β-sheet-rich aggregate, however, an ANS assay, which reflects the exposure of hydrophobic patches, showed two phases reflecting that different processes are occurring. In this case, the second phase likely corresponds to the formation of a species that has a higher content of exposed hydrophobic patches (or an exposed hydrophobic patch that can bind ANS more tightly) with greater β-structure than the metastable species formed in the first phase. Most likely, the metastable species are small and larger spherical species and the thermodynamically stable product is a β-structure-rich amorphous aggregate.

As apparent from the ThT assay performed at pH 7.5 (Figure 6.5B), IPP4 has a slow gradual increase in an aggregated species capable of binding ThT and the process has not reached equilibrium over the timescale of the assay. In the ANS assay, the plateau phase was reached after *circa* 50 to 100 hours depending on the concentration of IPP4 used (Figure 6.6B), which suggests that the exposure of the hydrophobic patches precedes the formation of β-structure capable of binding ThT. The relatively slow conversion of IPP4 into the final aggregated form will also affect spectroscopic characterization of the IPP4 aggregate after 8 days of incubation (Figure 6.1, Figure 6.2, Figure 6.3) as these methods provide a bulk characterization which averages contributions of all species in the sample.

IPP4 aggregates formed at pH 7.0 and 7.5 are morphologically distinct (amorphous aggregates *vs* short fibril species, Table 6.3). The differences observed in this pH range are likely due to the deprotonation of the N-terminal amino group and/or the side chain of His1. Thus, establishing that interactions in the N-terminus are likely to play a key role in determining the morphology of the aggregate formed. This hypothesis is in agreement with

the previously observed shift in aggregation kinetics of non-lipidated GLP-1 in the pH range from 7.5 to 8.2, where the protonation of N-terminus and His1 side chain changes[36].

In general, lipidated peptides seem to be more diverse in the morphologies of aggregates (from micelles to fibrils and hydrogels) compared to non-lipidated peptides and proteins[17,124,126,315,322,343]. Some of the numerous examples of morphological diversity of aggregates and self-assembled species (reversible self-assembly) of lipidated peptides and proteins include: glutamic acid-rich lipopeptide, which is capable of self-assembly into nanofibers or micelles in blood serum in the pH range between 7.0 and 7.4[344]. Another example is a peptide conjugate consisting of heptapeptide sequence for which the formation of either spherical micelles or nanofibers is regulated by temperature[345].

6.5.5 Aggregation of IPP5

At neutral pH, IPP5 aggregates into long, rigid fibrils (Figure 6.4E and F). This is consistent with the fibril formation observed for liraglutide[81,346] which is identical in structure apart from C-terminal amidation. In contrast to the IPP3 and IPP4 analogues, the lipidation of IPP5 at position 20 favours the formation of long rigid fibrils. There are several possibilities with regards to the orientation of the lipid chain in the fibril. It may be buried inside the core of the fibril, helping to stabilize the fibrillar state by burial of hydrophobic surface area, or pointing out of the fibril core, which may minimize steric hindrance which might otherwise occur in the compact fibril core. The latter option is more likely as the lipid chain would otherwise interfere with the peptide backbone hydrogen bonding which forces the monomer to pack into the β-sheet core in the fibril.

The ThT assay revealed fast formation of β-structure, albeit the final ThT fluorescence intensity is considerably less intense compared to the other analogues studied, Figure 6.5. The steep increase in fluorescence in the first 10 hours can be explained by ThT binding to the β-structure-rich oligomer, which is formed upon the peptide dissolution (Figure 5.16), and to newly formed amyloid fibrils. The ANS assay does not detect any major change in the level of hydrophobic patch exposure (Figure 6.6C) which may be related to the high compactness of the fibrils formed and a low percentage of disordered regions (in which hydrophobic patches may be exposed).

As was observed in the assessment of the stability of IPP5 oligomers (Figure 6.8), the larger oligomer containing the higher percentage of β-structure does not convert back into the

smaller and more α-helical oligomer. However, the separated fraction of the smaller oligomeric species re-equilibrates into the original oligomeric population containing both smaller and larger oligomers. This observation suggests that the larger oligomer with higher β-structure is a metastable intermediate in the aggregation process and may serve as a nucleus for fibril elongation, Figure 6.8D.

6.5.6 Rapid aggregation of IPP2

The IPP2 analogue, which is lipidated close to its N-terminus at position 2, shows considerably faster aggregation compared to all the other analogues studied. Another feature specific to this analogue is its very limited solubility – IPP2 is soluble only at around pH 3 and lower. The pH-limited solubility of IPP2 is partly due to an additional amino-acid substitution, Lys26→Arg26, which on its own limits the solubility of the peptide, as was shown for the non-lipidated variant of IPP2, Figure 6.20. However, this substitution is not likely to be responsible for the rapid aggregation observed for this analogue (Figure 6.18) as the non-lipidated peptide with the analogous sequence to IPP2 shows only a very low aggregation propensity when incubated at 37 °C with agitation for 6 days, Figure 6.21.

Upon the dissolution of lyophilised IPP2 in buffer, a rapid structural reorganization, from α-helix to β-sheet, was observed (Figure 6.18). Neither ThT nor ANS assay detected any further changes over 6 days (Figure 6.19) which is probably due to the completion of the α-helix → β-sheet conversion before the first timepoint of the assay was measured. A slight difference in the morphology of fibrillar species was observed for the samples imaged shortly after the IPP2 dissolution in buffer and after 8 days of incubation at 37 °C, Figure 6.18C and D, which suggests that the fibrillar species formed rapidly undergo further maturation, which is not accompanied by any further change in secondary structure.

6.5.7 Effect of the nature of lipidation moiety on aggregation: comparison of the physical stability of IPP5 and IPP10

The physical stability of two analogues, IPP5 and IPP10, differing in the nature of the lipid and linker attached was compared in Section 6.4. Both analogues are lipidated at position 20 and apart from the C-terminal amidation, IPP5 and IPP10 are analogous to commercially available approved diabetes drugs liraglutide and semaglutide.

Using FT-IR and far-UV CD (Figure 6.24), it was shown that the nature of lipidation does not affect the secondary structure of the peptide either in the solid (powder) state or in an aqueous buffer at pH 7.5. Similarly, the position of lipidation was shown to have only a minor effect on the secondary structure of the peptide, Chapter 5, Figure 5.6. However, the nature of the lipidation affects the oligomeric species and their distribution that IPP5 and IPP10 adopt in solution at pH 7.5, Figure 6.25. The nature of the lipid and the linker connecting the lipid with the peptide chain is therefore likely to influence the stability of oligomeric species. This effect may be sterically driven and/or driven by electrostatic interactions since the linker used in IPP10 contains an additional negative charge.

It is interesting to note that the long-term stability of IPP5 and IPP10 differ significantly. Whereas, for IPP5 the fibrillation occurs within several hours of incubation at 37 °C, IPP10 did not show any visible formation of fibrils or other aggregates over this timescale. The long-term stability of IPP10 is consistent with previously published stability studies of semaglutide, in which the semaglutide aggregation at neutral pH was detected only after several weeks of incubation at 37 °C[322]. Results from another study comparing semaglutide and liraglutide analogues suggested a lower tendency for further self-assembly of semaglutide at higher concentrations, again attributed to the additional negative charge in the linker compared to liraglutide[347].

IPP5 and IPP10 (as well as liraglutide and semaglutide) also differ from each other by a Ala → Aib substitution at position 2, which is present in IPP10 (and semaglutide). Although this change is minor, it also affects the size and distribution of oligomeric species and, to some extent, the aggregation propensity (see a comparison of non-lipidated GLP-1-Am and GLP-1-Am_Aib, Figure 6.27 and Figure 6.28). Oligomers observed for lipidated analogues are not the same as those formed by non-lipidated variants. However, the data obtained for non-lipidated variants further highlight the importance of the N-terminus in the self-assembly and aggregation processes and also points to the fact that the effect of this substitution is probably not negligible for the physical stability of the analogue.

CHAPTER 7
CONCLUSIONS AND FUTURE OUTLOOK

7.1 Conclusions

In this Thesis, the physical stability, i.e., the self-association and aggregation of the therapeutic peptide hormone GLP-1, its amidated (GLP-1-Am) and lipidated analogues (IPP2, IPP3, IPP4, IPP5 and IPP10) were studied. The first part of the Thesis is focused on non-lipidated GLP-1 and GLP-1-Am. Chapter 3 describes the fibrillation of these peptides under different conditions and focusses on the formation of low-molecular weight oligomers of GLP-1 and GLP-1-Am at near physiological conditions. Chapter 4 employed complementary computational approaches to determine the energy landscape of monomeric GLP-1 at acidic and neutral pH in order to understand the intrinsic structural preferences of the peptide under different conditions. The second part of the Thesis (Chapters 5 and 6) is focused on the self-assembly and aggregation of lipidated GLP-1-Am analogues. In this part, the effects of the lipidation itself on peptide aggregation, as well the effect of the lipidation position in the GLP-1 sequence and the nature of the lipid moiety on the peptide physical stability were studied.

In Chapter 3, stable low-molecular weight oligomers of the GLP-1 and its C-terminally amidated analogue, GLP-1-Am, were detected, isolated and characterized. GLP-1 and GLP-1-Am, which are prone to amyloid fibril formation over a wide range of conditions, both form low-molecular weight oligomers at pH values between 7 and 8. These oligomers are not easily capable of further self-assembly into larger oligomers or amyloid fibrils or dissociation back into monomers. Moreover, the addition of pre-formed oligomers to an aggregation reaction of monomer slightly slows down the rate of fibrillation. Therefore, these oligomers are likely the products of an alternative self-assembly pathway that competes with fibril formation. The size of the oligomers was estimated to be from dimer to

pentamer, and they were also shown to have a highly disordered structure. Surprisingly, these oligomers are stable with respect to concentration, time, and other stressors. However, LC-MS analysis of low-molecular weight oligomers did not show any covalent modification of the monomeric peptide within the oligomer. This observation of stable oligomeric species which are "off-pathway" to the fibrillation process further illustrates the diversity and complexity of the aggregation landscape of peptides in which on-pathway oligomers, protofibrils and fibrils are not the only stable detectable species. It also sheds some light on the unusual dependence of GLP-1 aggregation kinetics on peptide concentration, which has been previously observed under some conditions[36]. The observation and characterization of these species is of great importance since GLP-1 and its analogues are used as therapeutic drugs. Therefore, a careful analysis of all (even non-covalent) species formed during drug manufacturing and storage is essential.

In Chapter 4, Energy Landscape Theory was employed to investigate and rationalize the conformational behaviour and aggregation propensity of GLP-1 in different protonation states. Under all conditions studied, the energy landscapes possess a multi-funnel character with a variety of structurally different ensembles with low energy, a typical feature in the energy landscapes of intrinsically disordered proteins and aggregating systems[255,278,279,286]. High-energy barriers between the individual funnels imply that the probability of observing individual structural ensembles in experiments may not correspond to equilibrium and may be determined, in part, by kinetics[255]. α-helical conformational ensembles were the lowest energy for all conditions investigated. However, it was shown that at acidic pH values, β-structure-containing conformations are lower in energy, i.e., more energetically favoured, compared to neutral pH. These observations are consistent with the experimental data which, in general, shows a greater propensity for aggregation at acidic pH. Another interesting observation, which was also confirmed experimentally, is that a structural change occurs within the α-helix between pH 3 and 4. At a protonation state corresponding to pH 3, as well as the protonation state at pH ≈ 7.5, a conformational ensemble containing a single long α-helix was observed, whereas at pH 4, the α-helix was broken in the middle resulting in two shorter α-helices. Interestingly, the only change between pH 3 and 4 is the protonation of the side chain of Asp9 which is not directly involved in the helix break, i.e., it is not near the site at which the helix is disrupted. However, deprotonation of Asp9 stabilizes the more compact structures in the N-terminal region of GLP-1 and therefore can influence overall stability of the monomer conformation. Although the overall content of α-helical secondary structure in all energy landscapes was likely overpredicted by the method (as was discussed

in Chapter 4), the change in α-helical structure between pH 3 and 4 was also observed experimentally using far-UV CD. This provides validation that the computational methods capture important features of the conformational ensemble of GLP-1 at different pH values. These results highlight the potential of the method to help rationalize experimental results, providing detailed structural information that it is challenging to obtain experimentally, as well as predict the aggregation/misfolding propensity of peptides.

Chapters 5 and 6 are focused on the self-assembly and aggregation of lipidated analogues of GLP-1. Four lipidated GLP-1 analogues, varying in the position of lipidation and one additional analogue differing in the nature of the lipid moiety, were studied to establish the effect of the lipidation site and the lipid moiety on the physical stability of the peptide. Lipidation was shown to have only a minor influence on the secondary structure of GLP-1-Am. However, it greatly limits the peptide solubility, which depends upon the position of lipidation. Using size-exclusion chromatography and sedimentation velocity experiments, it was determined that lipidated GLP-1-Am analogues are mainly present in solution in the form of larger oligomers (ranging approx. from 7 to 25 monomeric units). The size and heterogeneity of these oligomeric species varies considerably with both the position of lipidation site in the peptide sequence and type of the lipidation moiety. The oligomeric populations in freshly dissolved peptide samples may already serve as an indication of its stability. For example, the IPP10 analogue exists in a single oligomeric state in solution directly after preparation at pH 7.5, and it was also shown to have the greatest stability over the timescale studied. Additionally, for IPP4, populations of oligomeric species were shown to be peptide concentration-dependent and pH-dependent for the IPP5 analogue.

In the case of IPP4, an interesting toroidal assembly to oligomer transition was observed in the pH range from 7.0 to 7.5 which is likely due to deprotonation of the N-terminus and/or the His1 side chain. At pH 7.0, IPP4 showed limited solubility and the formation of uniform toroidal species with an average diameter of about 20 nm. These toroidal assemblies have a high content of α-helix and they are stable over several hours. When incubated at 37 °C with agitation, these assemblies further aggregate into larger toroidal species and later into amorphous aggregates. This process is accompanied with an increase in β-structure. This work, together with other studies[17,315,320,348], shows that micelle-like species and toroidal assemblies of self-assembling peptides and lipopeptides are another stable type of species on the aggregation landscape.

The aggregation occurring on longer timescales (days) was shown to be highly dependent on lipidation position and the nature of the lipid moiety. All aggregates formed during incubation at 37 °C had a high content of β-structure as opposed to non-aggregated peptides, however, the morphology and tertiary structure of the resulting aggregates was found to vary depending on the position of lipidation. The aggregation kinetics of lipidated analogues were rarely observed to follow that of a classical nucleation-elongation mechanism leading to the formation of amyloid fibrils[90]. In contrast, the aggregation of lipidated GLP-1-Am analogues was found to consist of more complex processes which may involve competing pathways. In addition, a number of different morphologies of aggregates were observed, not all of them fibrillar in nature, suggesting that for lipidated peptides the fibrillar state is not always the most favourable and/or there is a competition between multiple stable aggregate types. The aggregation of lipidated analogues studied is also affected by the initial population of oligomeric species (size, structure and stability) and the fact that only a minor fraction of peptide exists in the solution as monomers.

In Chapter 6, the stability of the IPP2 analogue, lipidated near the N-terminus, was investigated. In contrast to the other lipidated analogues studied in this Thesis, IPP2 is soluble only at pH 3 and rapidly (seconds to minutes) aggregates into short fibrils. This observation shows that lipidating GLP-1-Am in the N-terminal region is not desirable for either its solubility or stability.

The influence of the nature of the lipid moiety on analogue stability was investigated in Chapter 6 by comparison of the IPP5 and IPP10 analogues, which are structurally similar to commercially available liraglutide and semaglutide GLP-1 analogues, respectively. They differ only in the linker and the lipid moiety attached and not the site of lipidation. It was shown that modifying the linker and the lipid moiety, both affected the initial oligomers formed, as well as the overall stability of the analogue, which is increased for IPP10.

Results presented in this Ph.D. Thesis contribute to the understanding of the physical stability and related aggregation processes of lipidated peptide pharmaceutics. Not only do the results provide insight into the mechanisms of aggregation of this family of peptides but also important information that may be useful for drug manufacture, formulation and storage.

7.2 Future outlook

The field of peptide and protein aggregation is vast and it is rapidly developing since aggregation phenomena are of importance not only in protein misfolding diseases but also during the production of pharmaceutical proteins/peptides and in the field of biomaterials. Although much is already known, there are still significant gaps mainly in understanding of "non-standard" aggregation processes which either do not lead to the formation of amyloid fibrils or in which different types of stable aggregates are formed.

The work presented in this Thesis directly opens up several new research questions. Firstly, a more detailed investigation on the inhibition of aggregation by the formation of competing low-molecular weight aggregates is needed to better understand these processes. Additionally, the hypothesis of growth-incompetent monomers[80] and unimolecular step conversion[36] needs experimental verification. Nowadays, there are methods which enable one to distinguish between different structural ensembles of monomers in the sample, such as ion-mobility mass spectrometry, and this technique may, therefore, be able to distinguish between the populations of "growth-incompetent" and "growth-competent" monomers. Ion-mobility mass spectrometry allows a separation of charged molecules based on their size, mass, and charge, which is related to the overall 3-dimensional structure of the ion. This technique is becoming increasingly employed in biomolecular interaction studies including the early stages peptide aggregation[349–352].

Many questions remain around the physical stability of lipidated peptides which are getting increasing attention in the pharmaceutical industry[12,39,294,303,308,353–355] due to their improved stability, longer half-life *in vivo*, and the possibility of increased drug permeability and bioavailability, and slow-release formulation[303,347,356]. Tools to predict the aggregation propensity of lipopeptides/lipoproteins from their structure, similar to those already available for non-lipidated peptides/proteins[151–153,357], are needed. In order to develop such predictors, further studies on different lipidated peptides/proteins are needed to create a database from which general rules can be extracted. Computational methods applied in Chapter 4, could also be applicable to lipidated analogues and their oligomeric states. These studies would provide further insight into the structure and stability of these oligomeric states and might be able to estimate the self-assembly and aggregation propensities of the system by comparing the relative stabilities of the oligomeric species. However, applying the Energy Landscape Theory to larger oligomeric states without the prior knowledge of their

structure is also very demanding with respect to both computational resources and time and therefore novel methods for the prediction of aggregation propensities of lipopeptides and lipoproteins are needed.

Of interest would be more detailed studies of the micellar states formed by lipopeptides. It would be interesting to have greater structural insight into such micellar states. Specifically for toroidal assemblies observed for IPP4, further experiments, such as staining with peptide-specific fluorescence dyes and imaging using fluorescence microscopy, would provide more detailed morphological information about localization of the peptidic part of the peptide within the assembly. The design of micellar formulations is an interesting option for the development of slow-release drug formulations.

Additionally, better techniques are needed for monitoring rapid aggregation processes, such as observed for IPP2. Processes occurring on the time scale of seconds and lower are difficult to capture by conventional techniques. Here, the microfluidic devices offer a promising perspective. Since it was shown that modelling the molecular diffusion of individual molecules in microchannels can be used to rapidly determine the size of the protein particles and their aggregates under native conditions[358,359].

Unanswered questions also remain regarding the arrangement of lipidated peptides in oligomers and other aggregated species. Understanding the arrangement of lipid moieties in oligomers, micelles, fibrils and other aggregates could help with aggregation predictions and could improve the design of potential therapeutic candidates. High-resolution techniques, e.g., cryo-electron microscopy, solid-state NMR or X-ray diffraction (for aligned fibrillar samples or crystallizable aggregates) could help with answering these questions.

7.3 Future directions in the development of lipidated GLP-1 analogues

The research presented in this Thesis highlights several findings which may play an important role in the future development of lipidated GLP-1 analogues. Firstly, lipidation can greatly impact the solubility of GLP-1; in most cases, the solubility of the lipidated analogue was lower compared to non-lipidated variant. The analogue solubility was dependent not only on the types of the lipid and the linker moiety but also on the position of the lipidation in the peptide sequence. The analogue solubility, therefore, varies with the mutual arrangement of the lipid (+ linker) and the peptide moiety, even if their structures in both analogues are identical. These observations may be valuable for the development of computational solubility predictors of (not only) peptide and lipid conjugates.

The poorest solubility and physical stability were reported for the lipidation site close to the N-terminus (Section 6.3.3). Moreover, the lipidation in the proximity of the N-terminus may also interfere with the biological activity of GLP-1[27]. GLP-1 analogues lipidated in the central region of the peptide, approximately from position 10 to position 20, showed a much better solubility profile and physical stability[19,22,40].

For all lipidated GLP-1 analogues studied in this Thesis, the lipidation triggered the formation of larger and more stable oligomers compared non-lipidated variants. However, the nature of the oligomers formed varied greatly in size, stability, structure and the number of distinct oligomeric species, depending on the position of the lipidation site and the nature of the linker and lipid moieties. By comparing the oligomerization and aggregation behaviour of IPP10 and IPP5 (Section 6.4), it was demonstrated that the formation of a single stable oligomer in the solution decreases the propensity to further aggregation, whereas formation of multiple oligomeric species may, in contrast, speed up further aggregation of the analogue. These aspects should be considered in further stability and aggregation studies. It is also desirable for future research to focus on the development of tools which enable *in silico* predictions of oligomeric populations in solution.

In further development of lipidated GLP-1 analogues an emphasis should also be put on the compatibility with the current oral delivery trends (oral delivery devices, drug encapsulation…)[360]. The oral delivery option has been already successfully implemented for semaglutide[16] and the pharmaceutical industry is nowadays increasingly focusing on overcoming the obstacles in the oral delivery of other peptide-based pharmaceuticals[361].

APPENDIX A
ADDITIONAL DATA TO CHAPTER 5: EFFECT OF LIPIDATION ON BIOPHYSICAL PROPERTIES OF GLP-1-AM

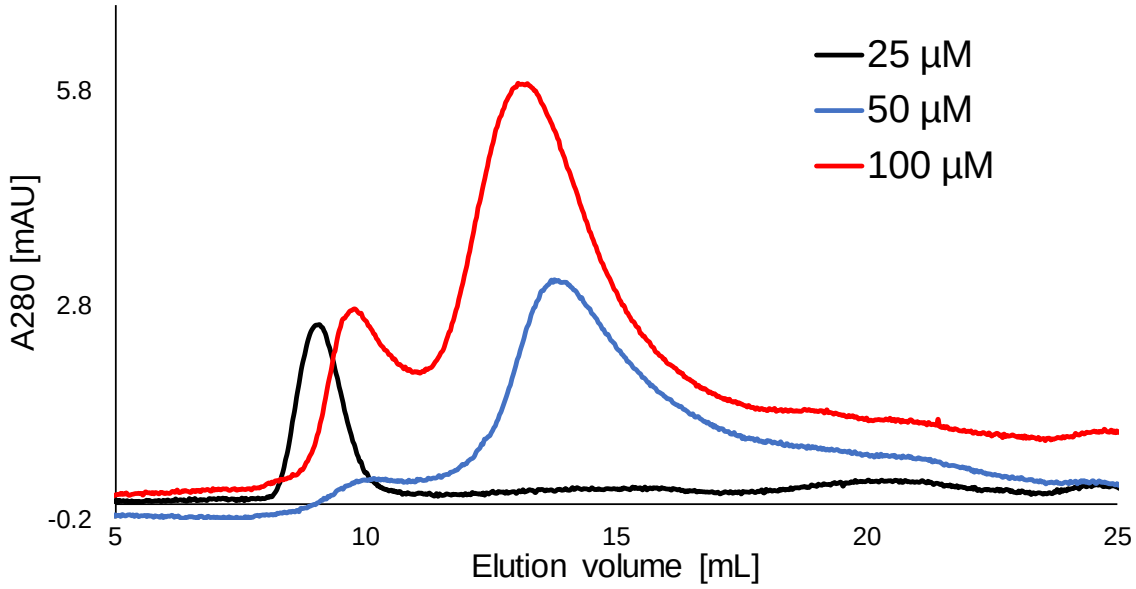


Figure A1: Peptide concentration-dependent distribution of IPP4 oligomeric species at pH 7.5 – Peptides International batch. SEC chromatograms of freshly prepared IPP4 samples at 25, 50 and 100 µM concentration in 25 mM phosphate, pH 7.5. Samples were analysed using a Superdex200 Increase 10/300 column and a 0.75 mL min^{-1} flow rate. *Circa* 200 µL of each sample was injected onto the column using a 200 µL injection loop. IPP4 batch from Peptides International was used for this experiment.

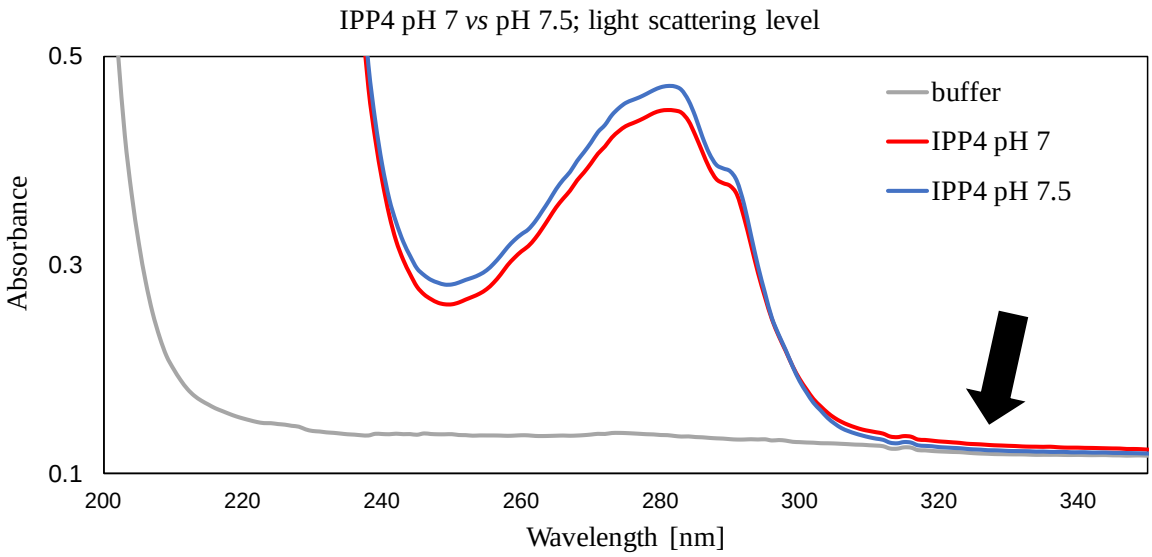


Figure A2: Absorption spectra of freshly prepared IPP4 samples in 25 mM phosphate at pH 7 and 7.5. The increased level of light scattering due to larger particles (toroidal assemblies) in solution is highlighted by a black arrow.

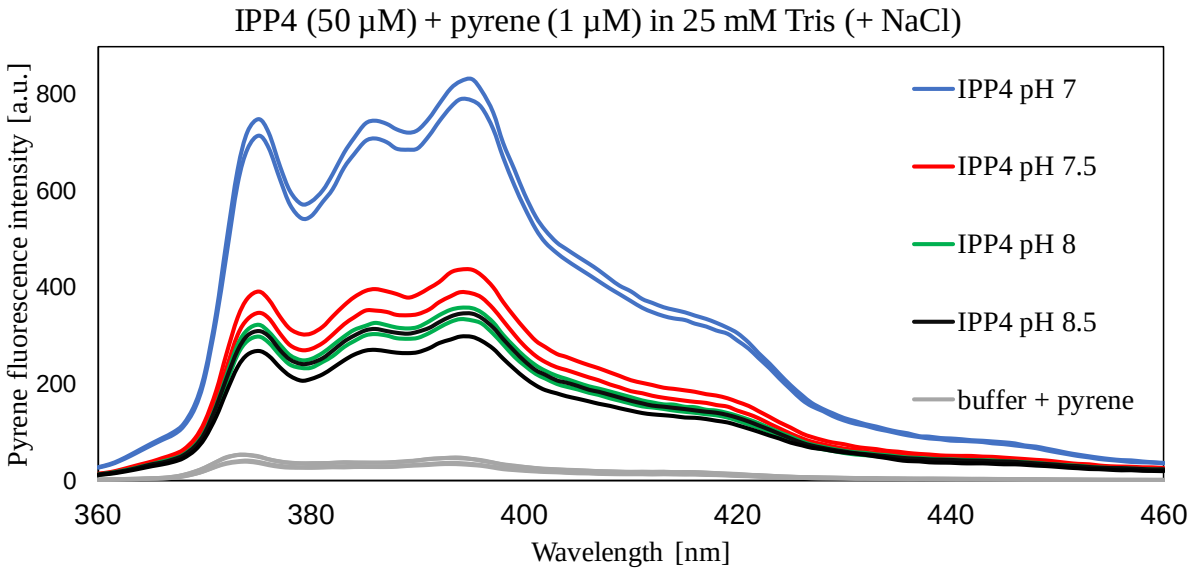


Figure A3: Fluorescence spectra of 1 µM pyrene with and 50 µM IPP4 at a pH range from 7 to 8.5. All samples were freshly prepared in 25 mM Tris buffer adjusted by NaCl (if needed) to keep the ionic strength constant (IS = 22 mM). Each sample was prepared and measured in a duplicate. Fluorescence emission spectra were recorded after excitation at 339 nm.

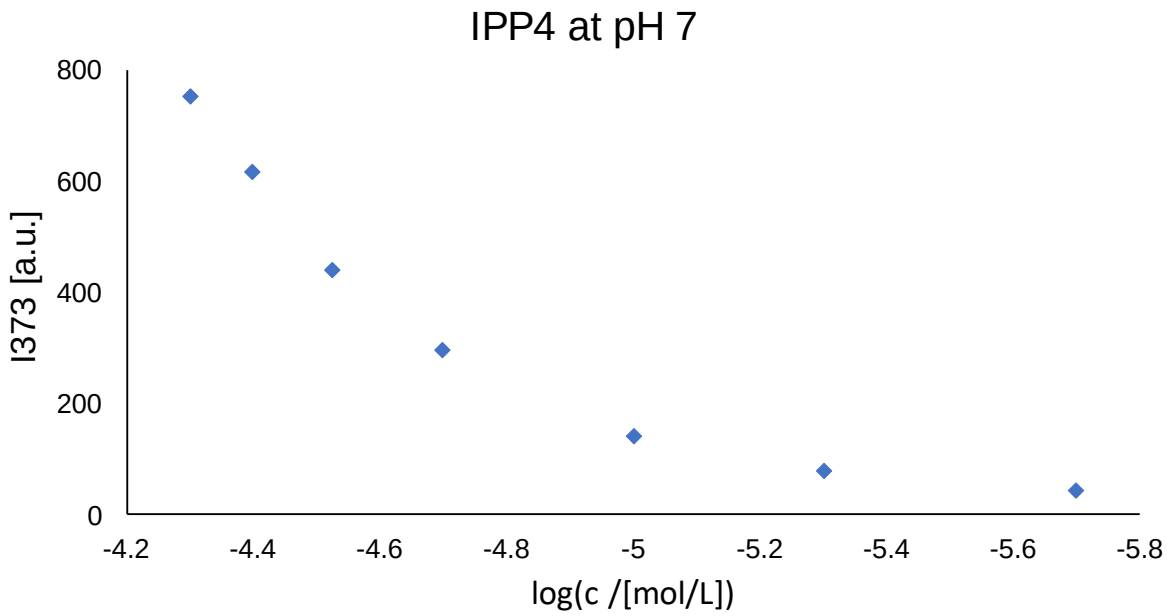


Figure A4: Dependence of pyrene fluorescence on IPP4 concentration at pH 7. Fluorescence intensity at 372 nm was recorded after the sample excitation at 339 nm for a range of IPP4 concentrations (from 2 to 50 µM) with 1 µM pyrene. All samples were freshly prepared in 25 mM phosphate, pH 7. Concentrations are plotted as log(c). Here, the emission intensity did not show any apparent transition indicating that the "critical aggregation concentration" lies outside of the peptide concentration range tested.

APPENDIX B
ADDITIONAL DATA TO CHAPTER 6: AGGREGATION BEHAVIOUR OF LIPIDATED GLP-1-AM VARIANTS

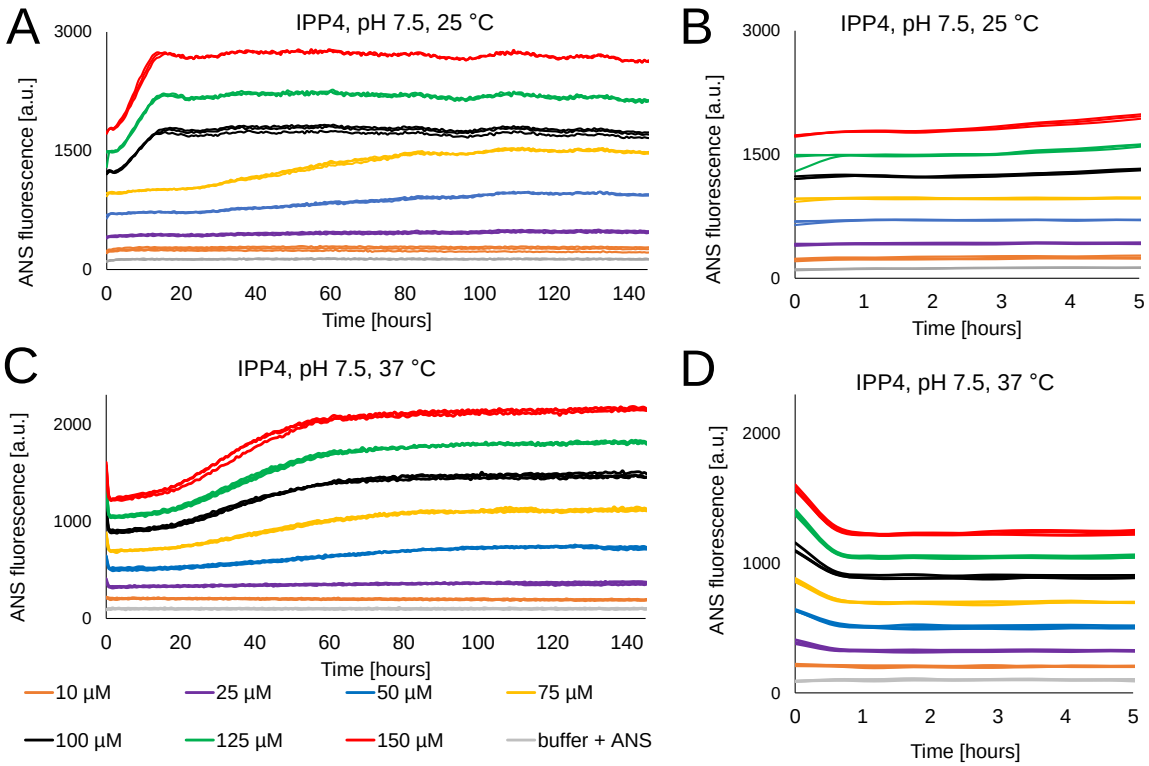


Figure B1: ANS fluorescence assay of IPP4 at pH 7.5 – signal drop due to temperature equilibration. An aggregation of IPP4 (range of concentrations) at pH 7.5 in 25 mM phosphate was monitored over 6 days using an ANS assay. The same assay was performed at 25 °C (A) and 37 °C (C). Panels B and D show a detail of the first 5 hours of assays in A and C, respectively. As apparent from B and D comparison, the initial drop in first 30 minutes observed for the assay performed at 37 °C is very likely to be caused by initial temperature equilibration as the samples were prepared at room temperature and then transferred into 37 °C for the assay. The assay, which was performed in room temperature, and therefore the samples did not undergo any temperature change, did not show this sharp initial drop in ANS fluorescence. An ANS fluorescence emission was recorded at 482 nm, after an excitation at 355 nm, every 30 minutes. Each sample was measured in a triplicate.

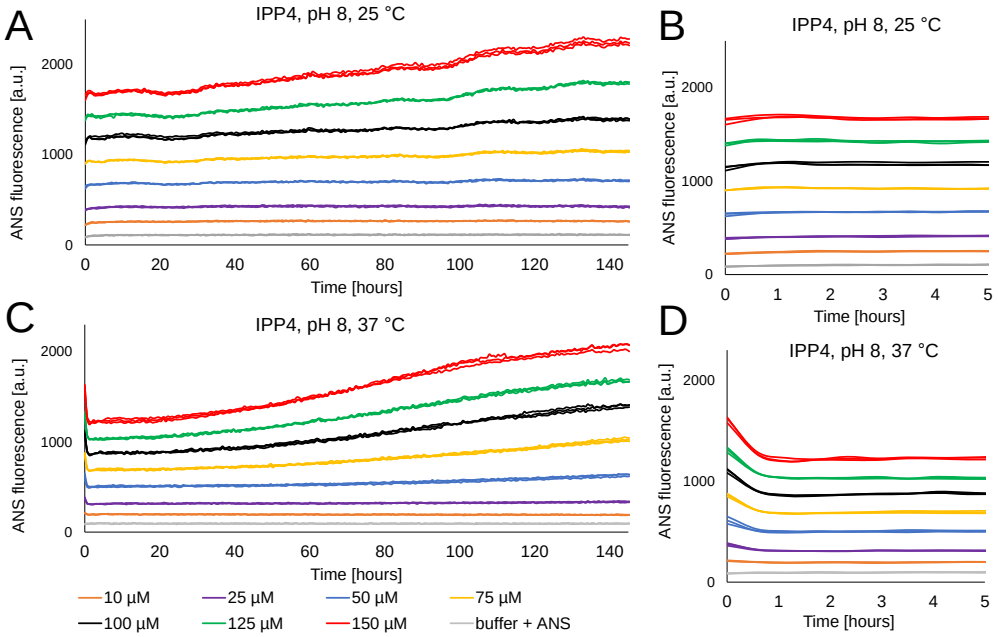


Figure B2: ANS fluorescence assay of IPP4 at pH 8 – signal drop due to temperature equilibration. An aggregation of IPP4 (range of concentrations) at pH 8 in 25 mM phosphate was monitored over 6 days using an ANS assay. The same assay was performed at 25 °C (A) and 37 °C (C). Panels B and D show a detail of the first 5 hours of assays in A and C, respectively. As apparent from B and D comparison, the initial drop in first 30 minutes observed for the assay performed at 37 °C is very likely to be caused by initial temperature equilibration as the samples were prepared at room temperature and then transferred into 37 °C for the assay. The assay, which was performed in room temperature, and therefore the samples did not undergo any temperature change, did not show this sharp initial drop in ANS fluorescence. An ANS fluorescence emission was recorded at 482 nm, after an excitation at 355 nm, every 30 minutes. Each sample was measured in a triplicate.

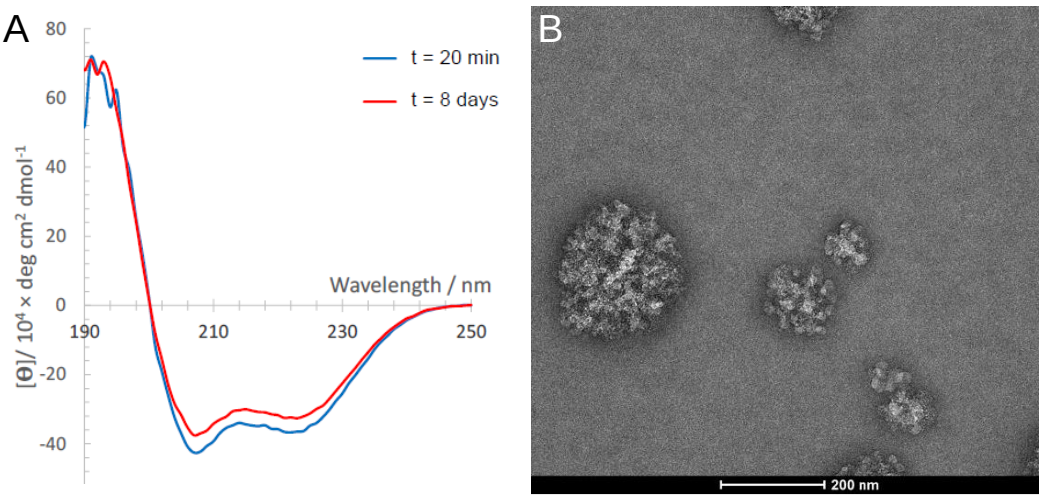


Figure B3: Aging of IPP10 samples at pH 7.5. (A) Far-UV CD spectra of freshly prepared and aged sample of 85 µM IPP10 in 25 mM phosphate at pH 7.5. The aged sample was incubated for 8 days at 37 °C with agitation. CD signal was converted into molar ellipticity. (B) TEM image of aged sample, 25 µM IPP10, after 8-day incubation at 37 °C with agitation in 25 mM phosphate at pH 7.5. The sample was negatively stained using 2 % uranyl acetate. Spherical objects in the image are likely to originate from the phosphate buffer. Both experiments were performed by Irina Edu.

REFERENCES

(1) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154.

(2) Muttenthaler, M.; King, G. F.; Adams, D. J.; Alewood, P. F. Trends in Peptide Drug Discovery. *Nat. Rev. Drug Discov.* **2021**, *20* (4), 309–325. https://doi.org/10.1038/s41573-020-00135-8.

(3) Gilroy, C. A.; Luginbuhl, K. M.; Chilkoti, A. Controlled Release of Biologics for the Treatment of Type 2 Diabetes. *J. Control. Release* **2016**, *240*, 151–164. https://doi.org/10.1016/j.jconrel.2015.12.002.

(4) Sharma, G.; Sharma, A. R.; Nam, J. S.; Doss, G. P. C.; Lee, S. S.; Chakraborty, C. Nanoparticle Based Insulin Delivery System: The next Generation Efficient Therapy for Type 1 Diabetes. *J. Nanobiotechnology* **2015**, *13* (1), 1–13. https://doi.org/10.1186/s12951-015-0136-y.

(5) Maack, T.; Johnson, V.; Kau, S. T.; Figueiredo, J.; Sigulem, D. Renal Filtration, Transport, and Metabolism of Low-Molecular-Weight Proteins: A Review. *Kidney Int.* **1979**, *16* (3), 251–270.

(6) Pollaro, L.; Heinis, C. Strategies to Prolong the Plasma Residence Time of Peptidedrugs. *Medchemcomm* **2010**, *1* (5), 319–324. https://doi.org/10.1039/C0MD00111B.

(7) Kolate, A.; Baradia, D.; Patil, S.; Vhora, I.; Kore, G.; Misra, A. PEG — A Versatile Conjugating Ligand for Drugs and Drug Delivery Systems. *J. Control. Release* **2014**, *192*, 67–81. https://doi.org/10.1016/j.jconrel.2014.06.046.

(8) Baggio, L. L.; Huang, Q.; Brown, T. J.; Drucker, D. J. A Recombinant Human Glucagon-Like Peptide (GLP)-1–Albumin Protein (Albugon) Mimics Peptidergic Activation of GLP-1 Receptor–Dependent Pathways Coupled With Satiety, Gastrointestinal Motility, and Glucose Homeostasis. *Diabetes* **2004**, *53* (9), 2492–2500. https://doi.org/10.2337/diabetes.53.9.2492.

(9) Matthews, J. E.; Stewart, M. W.; De Boever, E. H.; Dobbins, R. L.; Hodge, R. J.; Walker, S. E.; Holland, M. C.; Bush, M. A.; Group, A. S. Pharmacodynamics, Pharmacokinetics, Safety, and Tolerability of Albiglutide, a Long-Acting Glucagon-like Peptide-1 Mimetic, in Patients with Type 2 Diabetes. *J. Clin. Endocrinol. Metab.* **2008**, *93* (12), 4810–4817.

(10) Glaesner, W.; Mark Vick, A.; Millican, R.; Ellis, B.; Tschang, S.; Tian, Y.; Bokvist, K.; Brenner, M.; Koester, A.; Porksen, N.; et al. Engineering and Characterization of the Long-Acting Glucagon-like Peptide-1 Analogue LY2189265, an Fc Fusion Protein. *Diabetes. Metab. Res. Rev.* **2010**, *26* (4), 287–296. https://doi.org/10.1002/dmrr.1080.

(11) Park, E. J.; Choi, J.; Lee, K. C.; Na, D. H. Emerging PEGylated Non-Biologic Drugs. *Expert Opin. Emerg. Drugs* **2019**, *24* (2), 107–119. https://doi.org/10.1080/14728214.2019.1604684.

(12) Zhang, L.; Bulaj, G. Converting Peptides into Drug Leads by Lipidation. *Curr. Med. Chem.* **2012**, *19* (11), 1602–1618. https://doi.org/10.2174/092986712799945003.

(13) Bellmann-Sickert, K.; Elling, C. E.; Madsen, A. N.; Little, P. B.; Lundgren, K.; Gerlach, L.-O.; Bergmann, R.; Holst, B.; Schwartz, T. W.; Beck-Sickinger, A. G. Long-Acting Lipidated Analogue of Human Pancreatic Polypeptide Is Slowly Released into Circulation. *J. Med. Chem.* **2011**, *54* (8), 2658–2667. https://doi.org/10.1021/jm101357e.

(14) Simerska, P.; Moyle, P. M.; Toth, I. Modern Lipid-, Carbohydrate-, and Peptide-Based Delivery Systems for Peptide, Vaccine, and Gene Products. *Med. Res. Rev.* **2011**, *31* (4), 520–547. https://doi.org/10.1002/med.20191.

(15) Frederiksen, T. M.; Sønderby, P.; Ryberg, L. A.; Harris, P.; Bukrinski, J. T.; Scharff-Poulsen, A. M.; Elf-Lind, M. N.; Peters, G. H. Oligomerization of a Glucagon-like Peptide 1 Analog: Bridging Experiment and Simulations. *Biophys. J.* **2015**, *109* (6), 1202–1213. https://doi.org/10.1016/j.bpj.2015.07.051.

(16) Yu, M.; Benjamin, M. M.; Srinivasan, S.; Morin, E. E.; Shishatskaya, E. I.; Schwendeman, S. P.; Schwendeman, A. Battle of GLP-1 Delivery Technologies. *Adv. Drug Deliv. Rev.* **2018**, *130*, 113–130. https://doi.org/10.1016/j.addr.2018.07.009.

(17) Hutchinson, J. A.; Burholt, S.; Hamley, I. W.; Lundback, A. K.; Uddin, S.; Gomes Dos Santos, A.; Reza, M.; Seitsonen, J.; Ruokolainen, J. The Effect of Lipidation on the Self-Assembly of the Gut-Derived Peptide Hormone PYY 3-36. *Bioconjug. Chem.* **2018**, *29* (7), 2296–2308. https://doi.org/10.1021/acs.bioconjchem.8b00286.

(18) Ambery, P.; Parker, V. E.; Stumvoll, M.; Posch, M. G.; Heise, T.; Plum-Moerschel, L.; Tsai, L.-F.; Robertson, D.; Jain, M.; Petrone, M.; et al. MEDI0382, a GLP-1 and Glucagon Receptor Dual Agonist, in Obese or Overweight Patients with Type 2 Diabetes: A Randomised, Controlled, Double-Blind, Ascending Dose and Phase 2a Study. *Lancet* **2018**, *391* (10140), 2607–2618. https://doi.org/10.1016/S0140-6736(18)30726-8.

(19) Lainé, A. L.; Houvenagel, S.; Broo, A.; Jones, I.; Goodman, J.; Corkill, D.; Rose, J.; Coward, S.; Sandinge, A. S.; Petrone, M.; et al. Developing an Injectable Co-Formulation of Two Antidiabetic Drugs: Excipient Impact on Peptide Aggregation and Pharmacokinetic Properties. *Int. J. Pharm.* **2020**, *576* (October 2019), 119019. https://doi.org/10.1016/j.ijpharm.2020.119019.

(20) Madsen, K.; Knudsen, L. B.; Agersoe, H.; Nielsen, P. F.; Thøgersen, H.; Wilken, M.; Johansen, N. L. Structure−Activity and Protraction Relationship of Long-Acting Glucagon-like Peptide-1 Derivatives: Importance of Fatty Acid Length, Polarity, and Bulkiness. *J. Med. Chem.* **2007**, *50* (24), 6126–6132. https://doi.org/10.1021/jm070861j.

(21) Holst, J. J. The Physiology of Glucagon-like Peptide 1. *Physiological Reviews*. October 2007, pp 1409–1439. https://doi.org/10.1152/physrev.00034.2006.

(22) Lau, J.; Bloch, P.; Schäffer, L.; Pettersson, I.; Spetzler, J.; Kofoed, J.; Madsen, K.; Knudsen, L. B.; McGuire, J.; Steensgaard, D. B.; et al. Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide. *J. Med. Chem.* **2015**, *58* (18), 7370–7380. https://doi.org/10.1021/acs.jmedchem.5b00726.

(23) Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discov. Today* **2015**, *20* (1), 122–128. https://doi.org/10.1016/j.drudis.2014.10.003.

(24) Lau, J. L.; Dunn, M. K. Therapeutic Peptides: Historical Perspectives, Current Development Trends, and Future Directions. *Bioorg. Med. Chem.* **2018**, *26* (10), 2700–2707. https://doi.org/10.1016/j.bmc.2017.06.052.

(25) Kaspar, A. A.; Reichert, J. M. Future Directions for Peptide Therapeutics Development. *Drug Discov. Today* **2013**, *18* (17–18), 807–817. https://doi.org/10.1016/j.drudis.2013.05.011.

(26) D'Addio, S. M.; Bothe, J. R.; Neri, C.; Walsh, P. L.; Zhang, J.; Pierson, E.; Mao, Y.; Gindy, M.; Leone, A.; Templeton, A. C. New and Evolving Techniques for the Characterization of Peptide Therapeutics. *J. Pharm. Sci.* **2016**, *105* (10), 2989–3006. https://doi.org/10.1016/j.xphs.2016.06.011.

(27) Drucker, D. J. The Biology of Incretin Hormones. *Cell Metab.* **2006**, *3* (3), 153–165. https://doi.org/10.1016/j.cmet.2006.01.004.

(28) Ahrén, B. GLP-1 for Type 2 Diabetes. *Exp. Cell Res.* **2011**, *317* (9), 1239–1245. https://doi.org/10.1016/j.yexcr.2011.01.010.

(29) Madsbad, S. The Role of Glucagon-like Peptide-1 Impairment in Obesity and Potential Therapeutic Implications. *Diabetes, Obes. Metab.* **2014**, *16* (1), 9–21. https://doi.org/10.1111/dom.12119.

(30) Briyal, S.; Shah, S.; Gulati, A. Neuroprotective and Anti-Apoptotic Effects of Liraglutide in the Rat Brain Following Focal Cerebral Ischemia. *Neuroscience* **2014**, *281*, 269–281. https://doi.org/10.1016/j.neuroscience.2014.09.064.

(31) Adelhorst, K.; Hedegaard, B. B.; Knudsen, L. B.; Kirk, O. Structure-Activity Studies of Glucagon-like Peptide-1. *J. Biol. Chem.* **1994**, *269* (9), 6275–6278. https://doi.org/10.1016/S0021-9258(17)37366-0.

(32) Uddin, S. *Development of Biopharmaceutical Drug-Device Products*; Jameel, F., Skoug, J. W., Nesbitt, R. R., Eds.; AAPS Advances in the Pharmaceutical Sciences Series; Springer International Publishing: Cham, 2020; Vol. 35. https://doi.org/10.1007/978-3-030-31415-6.

(33) Yi, J.; Warunek, D.; Craft, D. Degradation and Stabilization of Peptide Hormones in Human Blood Specimens. *PLoS One* **2015**, *10* (7), e0134427. https://doi.org/10.1371/journal.pone.0134427.

(34) Deacon, C. F. Circulation and Degradation of GIP and GLP-1. *Horm. Metab. Res.* **2004**, *36* (11/12), 761–765. https://doi.org/10.1055/s-2004-826160.

(35) Sebokova, E.; Christ, A. D.; Wang, H.; Sewing, S.; Dong, J. Z.; Taylor, J.; Cawthorne, M. A.; Culler, M. D. Taspoglutide, an Analog of Human Glucagon-Like Peptide-1 with Enhanced Stability and in Vivo Potency. *Endocrinology* **2010**, *151* (6), 2474–2482. https://doi.org/10.1210/en.2009-1459.

(36) Zapadka, K. L.; Becher, F. J.; Uddin, S.; Varley, P. G.; Bishop, S.; Gomes dos Santos, A. L.; Jackson, S. E. A pH-Induced Switch in Human Glucagon-like Peptide-1 Aggregation Kinetics. *J. Am. Chem. Soc.* **2016**, *138* (50), 16259–16265. https://doi.org/10.1021/jacs.6b05025.

(37) Chang, X.; Keller, D.; O'Donoghue, S. I.; Led, J. J. NMR Studies of the Aggregation of Glucagon-like Peptide-1: Formation of a Symmetric Helical Dimer. *FEBS Lett.* **2002**, *515* (1–3), 165–170. https://doi.org/10.1016/S0014-5793(02)02466-3.

(38) Egbu, R.; van der Walle, C. F.; Brocchini, S.; Williams, G. R. Inhibiting the Fibrillation of a GLP-1-like Peptide. *Int. J. Pharm.* **2020**, *574*, 118923. https://doi.org/10.1016/j.ijpharm.2019.118923.

(39) Wang, Y.; Lomakin, A.; Kanai, S.; Alex, R.; Belli, S.; Donzelli, M.; Benedek, G. B. The Molecular Basis for the Prolonged Blood Circulation of Lipidated Incretin Peptides: Peptide Oligomerization or Binding to Serum Albumin? *J. Control. Release* **2016**, *241*, 25–33. https://doi.org/10.1016/j.jconrel.2016.08.035.

(40) Malm-Erjefält, M.; Bjørnsdottir, I.; Vanggaard, J.; Helleberg, H.; Larsen, U.; Oosterhuis, B.; van Lier, J. J.; Zdravkovic, M.; Olsen, A. K. Metabolism and Excretion of the Once-Daily Human Glucagon-Like Peptide-1 Analog Liraglutide in Healthy Male Subjects and Its In Vitro Degradation by Dipeptidyl Peptidase IV and Neutral Endopeptidase. *Drug Metab. Dispos.* **2010**, *38* (11), 1944–1953. https://doi.org/10.1124/dmd.110.034066.

(41) Paslawski, W.; Lorenzen, N.; Otzen, D. E. *Protein Amyloid Aggregation*; Eliezer, D., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2016; Vol. 1345. https://doi.org/10.1007/978-1-4939-2978-8.

(42) Aguzzi, A.; Calella, A. M. Prions: Protein Aggregation and Infectious Diseases. *Physiol. Rev.* **2009**, *89* (4), 1105–1152. https://doi.org/10.1152/physrev.00006.2009.

(43) Benilova, I.; Karran, E.; De Strooper, B. The Toxic Aβ Oligomer and Alzheimer's Disease: An Emperor in Need of Clothes. *Nat. Neurosci.* **2012**, *15* (3), 349–357. https://doi.org/10.1038/nn.3028.

(44) Dobson, C. M.; Knowles, T. P. J.; Vendruscolo, M. The Amyloid Phenomenon and Its Significance in Biology and Medicine. *Cold Spring Harb. Perspect. Biol.* **2019**, a033878. https://doi.org/10.1101/cshperspect.a033878.

(45) Thal, D. R.; Walter, J.; Saido, T. C.; Fändrich, M. Neuropathology and Biochemistry of Aβ and Its Aggregates in Alzheimer's Disease. *Acta Neuropathologica*. February 23, 2015, pp 167–182. https://doi.org/10.1007/s00401-014-1375-y.

(46) Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W. Functional Amyloid Formation within Mammalian Tissue. *PLoS Biol.* **2005**, *4* (1), e6. https://doi.org/10.1371/journal.pbio.0040006.

(47) Fowler, D. M.; Koulov, A. V.; Balch, W. E.; Kelly, J. W. Functional Amyloid – from Bacteria to Humans. *Trends Biochem. Sci.* **2007**, *32* (5), 217–224. https://doi.org/10.1016/j.tibs.2007.03.003.

(48) Wang, W. Instability, Stabilization, and Formulation of Liquid Protein Pharmaceuticals. *Int. J. Pharm.* **1999**, *185* (2), 129–188. https://doi.org/10.1016/S0378-5173(99)00152-0.

(49) Hutchinson, J. A.; Burholt, S.; Hamley, I. W. Peptide Hormones and Lipopeptides: From Self-Assembly to Therapeutic Applications. *J. Pept. Sci.* **2017**, *23* (2), 82–94. https://doi.org/10.1002/psc.2954.

(50) Manning, M. C.; Chou, D. K.; Murphy, B. M.; Payne, R. W.; Katayama, D. S. Stability of Protein Pharmaceuticals: An Update. *Pharm. Res.* **2010**, *27* (4), 544–575. https://doi.org/10.1007/s11095-009-0045-6.

(51) Evers, A.; Pfeiffer-Marek, S.; Bossart, M.; Heubel, C.; Stock, U.; Tiwari, G.; Gebauer, B.; Elshorst, B.; Pfenninger, A.; Lukasczyk, U.; et al. Peptide Optimization at the Drug Discovery-Development Interface: Tailoring of Physicochemical Properties Toward Specific Formulation Requirements. *J. Pharm. Sci.* **2019**, *108* (4), 1404–1414. https://doi.org/10.1016/j.xphs.2018.11.043.

(52) Jones, L. S.; Randolph, T. W.; Kohnert, U.; Papadimitriou, A.; Winter, G.; Hagmann, M.; Manning, M. C.; Carpenter, J. F. The Effects of Tween 20 and Sucrose on the Stability of Anti-L-Selectin during Lyophilization and Reconstitution. *J. Pharm. Sci.* **2001**, *90* (10), 1466–1477. https://doi.org/10.1002/jps.1098.

(53) Kameoka, D.; Masuzaki, E.; Ueda, T.; Imoto, T. Effect of Buffer Species on the Unfolding and the Aggregation of Humanized IgG. *J. Biochem.* **2007**, *142* (3), 383–391. https://doi.org/10.1093/jb/mvm145.

(54) Austerberry, J. I.; Thistlethwaite, A.; Fisher, K.; Golovanov, A. P.; Pluen, A.; Esfandiary, R.; Van Der Walle, C. F.; Warwicker, J.; Derrick, J. P.; Curtis, R. Arginine to Lysine Mutations Increase the Aggregation Stability of a Single-Chain Variable Fragment through Unfolded-State Interactions. *Biochemistry* **2019**, *58* (32), 3413–3421. https://doi.org/10.1021/acs.biochem.9b00367.

(55) Cromwell, M. E. M.; Hilario, E.; Jacobson, F. Protein Aggregation and Bioprocessing. *AAPS J.* **2006**, *8* (3), E572–E579. https://doi.org/10.1208/aapsj080366.

(56) Schmittschmitt, J. P.; Scholtz, J. M. The Role of Protein Stability, Solubility, and Net Charge in Amyloid Fibril Formation. *Protein Sci.* **2009**, *12* (10), 2374–2378. https://doi.org/10.1110/ps.03152903.

(57) Asthana, S.; Bhattacharyya, D.; Kumari, S.; Nayak, P. S.; Saleem, M.; Bhunia, A.; Jha, S. Interaction with Zinc Oxide Nanoparticle Kinetically Traps α-Synuclein Fibrillation into off-Pathway Non-Toxic Intermediates. *Int. J. Biol. Macromol.* **2020**, *150*, 68–79. https://doi.org/10.1016/j.ijbiomac.2020.01.269.

(58) Uversky, V. N.; Li, J.; Fink, A. L. Metal-Triggered Structural Transformations, Aggregation, and Fibrillation of Human α-Synuclein. *J. Biol. Chem.* **2001**, *276* (47), 44284–44296. https://doi.org/10.1074/jbc.M105343200.

(59) Raibekas, A. A.; Bures, E. J.; Siska, C. C.; Kohno, T.; Latypov, R. F.; Kerwin, B. A. Anion Binding and Controlled Aggregation of Human Interleukin-1 Receptor Antagonist. *Biochemistry* **2005**, *44* (29), 9871–9879. https://doi.org/10.1021/bi050388g.

(60) Bottomley, S. P.; Tew, D. J. The Citrate Ion Increases the Conformational Stability of A1-Antitrypsin. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **2000**, *1481* (1), 11–17. https://doi.org/10.1016/S0167-4838(00)00118-7.

(61) Katayama, D. S.; Nayar, R.; Chou, D. K.; Valente, J. J.; Cooper, J.; Henry, C. S.; Vander Velde, D. G.; Villarete, L.; Liu, C. P.; Manning, M. C. Effect of Buffer Species on the Thermally Induced Aggregation of Interferon-Tau. *J. Pharm. Sci.* **2006**, *95* (6), 1212–1226. https://doi.org/10.1002/jps.20471.

(62) Fayos, R.; Pons, M.; Millet, O. On the Origin of the Thermostabilization of Proteins Induced by Sodium Phosphate. *J. Am. Chem. Soc.* **2005**, *127* (27), 9690–9691. https://doi.org/10.1021/ja051352e.

(63) Marek, P. J.; Patsalo, V.; Green, D. F.; Raleigh, D. P. Ionic Strength Effects on Amyloid Formation by Amylin Are a Complicated Interplay among Debye Screening, Ion Selectivity, and Hofmeister Effects. *Biochemistry* **2012**, *51* (43), 8478–8490. https://doi.org/10.1021/bi300574r.

(64) Lo Nostro, P.; Ninham, B. W. Hofmeister Phenomena: An Update on Ion Specificity in Biology. *Chem. Rev.* **2012**, *112* (4), 2286–2322. https://doi.org/10.1021/cr200271j.

(65) Hasecke, F.; Miti, T.; Perez, C.; Barton, J.; Schölzel, D.; Gremer, L.; Grüning, C. S. R.; Matthews, G.; Meisl, G.; Knowles, T. P. J.; et al. Origin of Metastable Oligomers and Their Effects on Amyloid Fibril Self-Assembly. *Chem. Sci.* **2018**, *9* (27), 5937–5948. https://doi.org/10.1039/C8SC01479E.

(66) Miti, T.; Mulaj, M.; Schmit, J. D.; Muschol, M. Stable, Metastable, and Kinetically Trapped Amyloid Aggregate Phases. *Biomacromolecules* **2015**, *16* (1), 326–335. https://doi.org/10.1021/bm501521r.

(67) Klement, K.; Wieligmann, K.; Meinhardt, J.; Hortschansky, P.; Richter, W.; Fändrich, M. Effect of Different Salt Ions on the Propensity of Aggregation and on the Structure of Alzheimer's Aβ(1-40) Amyloid Fibrils. *J. Mol. Biol.* **2007**, *373* (5), 1321–1333. https://doi.org/10.1016/j.jmb.2007.08.068.

(68) Jain, S.; Udgaonkar, J. B. Salt-Induced Modulation of the Pathway of Amyloid Fibril Formation by the Mouse Prion Protein. *Biochemistry* **2010**, *49* (35), 7615–7624. https://doi.org/10.1021/bi100745j.

(69) Wang, W.; Roberts, C. J. Non-Arrhenius Protein Aggregation. *AAPS J.* **2013**, *15* (3), 840–851. https://doi.org/10.1208/s12248-013-9485-3.

(70) Batzli, K. M.; Love, B. J. Agitation of Amyloid Proteins to Speed Aggregation Measured by ThT Fluorescence: A Call for Standardization. *Mater. Sci. Eng. C* **2015**, *48*, 359–364. https://doi.org/10.1016/j.msec.2014.09.015.

(71) Hamilton-Brown, P.; Bekard, I.; Ducker, W. A.; Dunstan, D. E. How Does Shear Affect Aβ Fibrillogenesis? *J. Phys. Chem. B* **2008**, *112* (51), 16249–16252. https://doi.org/10.1021/jp805257n.

(72) Zapadka, K. L.; Becher, F. J.; Gomes dos Santos, A. L.; Jackson, S. E. Factors Affecting the Physical Stability (Aggregation) of Peptide Therapeutics. *Interface Focus* **2017**, *7* (6), 20170030. https://doi.org/10.1098/rsfs.2017.0030.

(73) Narhi, L. O.; Schmit, J.; Bechtold-Peters, K.; Sharma, D. Classification of Protein Aggregates. *J. Pharm. Sci.* **2012**, *101* (2), 493–498. https://doi.org/10.1002/jps.22790.

(74) Morris, A. M.; Watzky, M. A.; Finke, R. G. Protein Aggregation Kinetics, Mechanism, and Curve-Fitting: A Review of the Literature. *Biochim. Biophys. Acta - Proteins Proteomics* **2009**, *1794* (3), 375–397. https://doi.org/10.1016/j.bbapap.2008.10.016.

(75) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **2017**, *86* (1), 27–68. https://doi.org/10.1146/annurev-biochem-061516-045115.

(76) Törnquist, M.; Michaels, T. C. T.; Sanagavarapu, K.; Yang, X.; Meisl, G.; Cohen, S. I. A.; Knowles, T. P. J.; Linse, S. Secondary Nucleation in Amyloid Formation. *Chem. Commun.* **2018**, *54* (63), 8667–8684. https://doi.org/10.1039/c8cc02204f.

(77) Ferrone, F. Analysis of Protein Aggregation Kinetics. *Methods Enzymol.* **1999**, *309*, 256–274. https://doi.org/10.1016/S0076-6879(99)09019-9.

(78) Meisl, G.; Yang, X.; Hellstrand, E.; Frohm, B.; Kirkegaard, J. B.; Cohen, S. I. A.; Dobson, C. M.; Linse, S.; Knowles, T. P. J. Differences in Nucleation Behavior Underlie the Contrasting Aggregation Kinetics of the Aβ40 and Aβ42 Peptides. *Proc. Natl. Acad. Sci.* **2014**, *111* (26), 9384–9389. https://doi.org/10.1073/pnas.1401564111.

(79) Dear, A. J.; Meisl, G.; Michaels, T. C. T.; Zimmermann, M. R.; Linse, S.; Knowles, T. P. J. The Catalytic Nature of Protein Aggregation. *J. Chem. Phys.* **2020**, *152* (4). https://doi.org/10.1063/1.5133635.

(80) Kamgar-Parsi, K.; Hong, L.; Naito, A.; Brooks, C. L.; Ramamoorthy, A. Growth-Incompetent Monomers of Human Calcitonin Lead to a Noncanonical Direct Relationship between Peptide Concentration and Aggregation Lag Time. *J. Biol. Chem.* **2017**, *292* (36), 14963–14976. https://doi.org/10.1074/jbc.M117.791236.

(81) Bothe, J. R.; Andrews, A.; Smith, K. J.; Joyce, L. A.; Krishnamachari, Y.; Kashi, S. Peptide Oligomerization Memory Effects and Their Impact on the Physical Stability of the GLP-1 Agonist Liraglutide. *Mol. Pharm.* **2019**, *16* (5), 2153–2161. https://doi.org/10.1021/acs.molpharmaceut.9b00106.

(82) Karamanos, T. K.; Jackson, M. P.; Calabrese, A. N.; Goodchild, S. C.; Cawood, E. E.; Thompson, G. S.; Kalverda, A. P.; Hewitt, E. W.; Radford, S. E. Structural Mapping of Oligomeric Intermediates in an Amyloid Assembly Pathway. *Elife* **2019**, *8*, 1–32. https://doi.org/10.7554/eLife.46574.001.

(83) Dear, A. J.; Meisl, G.; Šarić, A.; Michaels, T. C. T.; Kjaergaard, M.; Linse, S.; Knowles, T. P. J. Identification of On- and off-Pathway Oligomers in Amyloid Fibril Formation. *Chem. Sci.* **2020**, No. c. https://doi.org/10.1039/C9SC06501F.

(84) Cohen, S. I. A.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. From Macroscopic Measurements to Microscopic Mechanisms of Protein Aggregation. *J. Mol. Biol.* **2012**, *421* (2–3), 160–171. https://doi.org/10.1016/j.jmb.2012.02.031.

(85) Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An Analytical Solution

to the Kinetics of Breakable Filament Assembly. *Science (80-.).* **2009**, *326* (5959), 1533–1537. https://doi.org/10.1126/science.1178250.

(86) Brichtová, E. Studies on the Physical Stability of a C-Terminally Amidated Variant of GLP-1, University of Cambridge, 2019. https://doi.org/10.17863/CAM.43927.

(87) Ferguson, N.; Berriman, J.; Petrovich, M.; Sharpe, T. D.; Finch, J. T.; Fersht, A. R. Rapid Amyloid Fiber Formation from the Fast-Folding WW Domain FBP28. *Proc. Natl. Acad. Sci.* **2003**, *100* (17), 9814–9819. https://doi.org/10.1073/pnas.1333907100.

(88) Meisl, G.; Kirkegaard, J. B.; Arosio, P.; Michaels, T. C. T.; Vendruscolo, M.; Dobson, C. M.; Linse, S.; Knowles, T. P. J. Molecular Mechanisms of Protein Aggregation from Global Fitting of Kinetic Models. *Nat. Protoc.* **2016**, *11* (2), 252–272. https://doi.org/10.1038/nprot.2016.010.

(89) AmyloFit https://amylofit.com/amylofitmain/login/.

(90) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. *Annu. Rev. Biochem.* **2017**, *86* (1), 27–68. https://doi.org/10.1146/annurev-biochem-061516-045115.

(91) Mannini, B.; Mulvihill, E.; Sgromo, C.; Cascella, R.; Khodarahmi, R.; Ramazzotti, M.; Dobson, C. M.; Cecchi, C.; Chiti, F. Toxicity of Protein Oligomers Is Rationalized by a Function Combining Size and Surface Hydrophobicity. *ACS Chem. Biol.* **2014**, *9* (10), 2309–2317. https://doi.org/10.1021/cb500505m.

(92) Ahmed, M.; Davis, J.; Aucoin, D.; Sato, T.; Ahuja, S.; Aimoto, S.; Elliott, J. I.; Van Nostrand, W. E.; Smith, S. O. Structural Conversion of Neurotoxic Amyloid-B 1-42 Oligomers to Fibrils. *Nat. Struct. Mol. Biol.* **2010**, *17* (5), 561–567. https://doi.org/10.1038/nsmb.1799.

(93) Vivoli Vega, M.; Cascella, R.; Chen, S. W.; Fusco, G.; De Simone, A.; Dobson, C. M.; Cecchi, C.; Chiti, F. The Toxicity of Misfolded Protein Oligomers Is Independent of Their Secondary Structure. *ACS Chem. Biol.* **2019**, *14* (7), 1593–1600. https://doi.org/10.1021/acschembio.9b00324.

(94) Bolognesi, B.; Kumita, J. R.; Barros, T. P.; Esbjorner, E. K.; Luheshi, L. M.; Crowther, D. C.; Wilson, M. R.; Dobson, C. M.; Favrin, G.; Yerbury, J. J. ANS Binding Reveals Common Features of Cytotoxic Amyloid Species. *ACS Chem. Biol.* **2010**, *5* (8), 735–740. https://doi.org/10.1021/cb1001203.

(95) Musteikytė, G.; Jayaram, A. K.; Xu, C. K.; Vendruscolo, M.; Krainer, G.; Knowles, T. P. J. Interactions of α-Synuclein Oligomers with Lipid Membranes. *Biochim. Biophys. Acta - Biomembr.* **2021**, *1863* (4), 183536. https://doi.org/10.1016/j.bbamem.2020.183536.

(96) Fusco, G.; Chen, S. W.; Williamson, P. T. F.; Cascella, R.; Perni, M.; Jarvis, J. A.; Cecchi, C.; Vendruscolo, M.; Chiti, F.; Cremades, N.; et al. Structural Basis of Membrane Disruption and Cellular Toxicity by α-Synuclein Oligomers. *Science (80-.).* **2017**, *358* (6369), 1440–1443. https://doi.org/10.1126/science.aan6160.

(97) Niyangoda, C.; Barton, J.; Bushra, N.; Karunarathne, K.; Strauss, G.; Fakhre, F.; Koria, P.; Muschol, M. Origin, Toxicity and Characteristics of Two Amyloid Oligomer Polymorphs. *RSC Chem. Biol.* **2021**, *2* (6), 1631–1642. https://doi.org/10.1039/d1cb00081k.

(98) Paslawski, W.; Mysling, S.; Thomsen, K.; Jørgensen, T. J. D.; Otzen, D. E. Co-Existence of Two Different α-Synuclein Oligomers with Different Core Structures Determined by Hydrogen/Deuterium Exchange Mass Spectrometry. *Angew. Chemie Int. Ed.* **2014**, *53* (29), 7560–7563. https://doi.org/10.1002/anie.201400491.

(99) Breydo, L.; Uversky, V. N. Structural, Morphological, and Functional Diversity of Amyloid Oligomers. *FEBS Lett.* **2015**, *589* (19PartA), 2640–2648. https://doi.org/10.1016/j.febslet.2015.07.013.

(100) Dear, A. J.; Michaels, T. C. T.; Meisl, G.; Klenerman, D.; Wu, S.; Perrett, S.; Linse, S.; Dobson, C. M.; Knowles, T. P. J. Kinetic Diversity of Amyloid Oligomers. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117* (22), 28–31. https://doi.org/10.1073/pnas.1922267117.

(101) Hnath, B.; Dokholyan, N. V. Toxic SOD1 Trimers Are Off-Pathway in the Formation of Amyloid-like Fibrils in ALS. *Biophys. J.* **2022**, *121* (11), 2084–2095. https://doi.org/10.1016/j.bpj.2022.04.037.

(102) Nick, M.; Wu, Y.; Schmidt, N. W.; Prusiner, S. B.; Stöhr, J.; DeGrado, W. F. A Long-Lived Aβ Oligomer Resistant to Fibrillization. *Biopolymers* **2018**, *109* (8), e23096. https://doi.org/10.1002/bip.23096.

(103) Chen, S. W.; Drakulic, S.; Deas, E.; Ouberai, M.; Aprile, F. A.; Arranz, R.; Ness, S.; Roodveldt, C.; Guilliams, T.; De-Genst, E. J.; et al. Structural Characterization of Toxic Oligomers That Are Kinetically Trapped during α-Synuclein Fibril Formation. *Proc. Natl. Acad. Sci.* **2015**, *112* (16), E1994–E2003. https://doi.org/10.1073/pnas.1421204112.

(104) Cremades, N.; Cohen, S. I. A.; Deas, E.; Abramov, A. Y.; Chen, A. Y.; Orte, A.; Sandal, M.; Clarke, R. W.; Dunne, P.; Aprile, F. A.; et al. Direct Observation of the

Interconversion of Normal and Toxic Forms of α-Synuclein. *Cell* **2012**, *149* (5), 1048–1059. https://doi.org/10.1016/j.cell.2012.03.037.

(105) Dasari, A. K. R.; Hughes, R. M.; Wi, S.; Hung, I.; Gan, Z.; Kelly, J. W.; Lim, K. H. Transthyretin Aggregation Pathway toward the Formation of Distinct Cytotoxic Oligomers. *Sci. Rep.* **2019**, *9* (1), 1–10. https://doi.org/10.1038/s41598-018-37230-1.

(106) Sanders, A.; Jeremy Craven, C.; Higgins, L. D.; Giannini, S.; Conroy, M. J.; Hounslow, A. M.; Waltho, J. P.; Staniforth, R. A. Cystatin Forms a Tetramer through Structural Rearrangement of Domain-Swapped Dimers Prior to Amyloidogenesis. *J. Mol. Biol.* **2004**, *336* (1), 165–178. https://doi.org/10.1016/j.jmb.2003.12.011.

(107) Staniforth, R. A.; Giannini, S.; Higgins, L. D.; Conroy, M. J.; Hounslow, A. M.; Jerala, R.; Craven, C. J.; Waltho, J. P. Three-Dimensional Domain Swapping in the Folded and Molten-Globule States of Cystatins, an Amyloid-Forming Structural Superfamily. *EMBO J.* **2001**, *20* (17), 4774–4781. https://doi.org/10.1093/emboj/20.17.4774.

(108) Baskakov, I. V.; Legname, G.; Baldwin, M. A.; Prusiner, S. B.; Cohen, F. E. Pathway Complexity of Prion Protein Assembly into Amyloid. *J. Biol. Chem.* **2002**, *277* (24), 21140–21148. https://doi.org/10.1074/jbc.M111402200.

(109) Bitan, G.; Kirkitadze, M. D.; Lomakin, A.; Vollers, S. S.; Benedek, G. B.; Teplow, D. B. Amyloid -Protein (A) Assembly: A 40 and A 42 Oligomerize through Distinct Pathways. *Proc. Natl. Acad. Sci.* **2003**, *100* (1), 330–335. https://doi.org/10.1073/pnas.222681699.

(110) Kuo, Y.-M.; Webster, S.; Emmerling, M. R.; De Lima, N.; Roher, A. E. Irreversible Dimerization/Tetramerization and Post-Translational Modifications Inhibit Proteolytic Degradation of Aβ Peptides of Alzheimer's Disease. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1998**, *1406* (3), 291–298. https://doi.org/10.1016/S0925-4439(98)00014-3.

(111) Calabrese, M. F.; Miranker, A. D. Formation of a Stable Oligomer of β-2 Microglobulin Requires Only Transient Encounter with Cu(II). *J. Mol. Biol.* **2007**, *367* (1), 1–7. https://doi.org/10.1016/j.jmb.2006.12.034.

(112) Grelle, G.; Otto, A.; Lorenz, M.; Frank, R. F.; Wanker, E. E.; Bieschke, J. Black Tea Theaflavins Inhibit Formation of Toxic Amyloid-β and α-Synuclein Fibrils. *Biochemistry* **2011**, *50* (49), 10624–10636. https://doi.org/10.1021/bi2012383.

(113) Meng, X.; Munishkina, L. A.; Fink, A. L.; Uversky, V. N. Effects of Various Flavonoids on the α-Synuclein Fibrillation Process. *Parkinsons. Dis.* **2010**, *2010*, 1–16. https://doi.org/10.4061/2010/650794.

(114) Masuda, M.; Suzuki, N.; Taniguchi, S.; Oikawa, T.; Nonaka, T.; Iwatsubo, T.; Hisanaga, S.; Goedert, M.; Hasegawa, M. Small Molecule Inhibitors of α-Synuclein Filament Assembly. *Biochemistry* **2006**, *45* (19), 6085–6094. https://doi.org/10.1021/bi0600749.

(115) Ehrnhoefer, D. E.; Bieschke, J.; Boeddrich, A.; Herbst, M.; Masino, L.; Lurz, R.; Engemann, S.; Pastore, A.; Wanker, E. E. EGCG Redirects Amyloidogenic Polypeptides into Unstructured, off-Pathway Oligomers. *Nat. Struct. Mol. Biol.* **2008**, *15* (6), 558–566. https://doi.org/10.1038/nsmb.1437.

(116) Versluis, F.; Marsden, H. R.; Kros, A. Power Struggles in Peptide-Amphiphile Nanostructures. *Chem. Soc. Rev.* **2010**, *39* (9), 3434. https://doi.org/10.1039/b919446k.

(117) Cui, H.; Webber, M. J.; Stupp, S. I. Self-Assembly of Peptide Amphiphiles: From Molecules to Nanostructures to Biomaterials. *Biopolymers* **2010**, *94* (1), 1–18. https://doi.org/10.1002/bip.21328.

(118) Matson, J. B.; Zha, R. H.; Stupp, S. I. Peptide Self-Assembly for Crafting Functional Biological Materials. *Curr. Opin. Solid State Mater. Sci.* **2011**, *15* (6), 225–235. https://doi.org/10.1016/j.cossms.2011.08.001.

(119) Lin, B. F.; Marullo, R. S.; Robb, M. J.; Krogstad, D. V.; Antoni, P.; Hawker, C. J.; Campos, L. M.; Tirrell, M. V. De Novo Design of Bioactive Protein-Resembling Nanospheres via Dendrimer-Templated Peptide Amphiphile Assembly. *Nano Lett.* **2011**, *11* (9), 3946–3950. https://doi.org/10.1021/nl202220q.

(120) Tu, R. S.; Tirrell, M. Bottom-up Design of Biomimetic Assemblies. *Adv. Drug Deliv. Rev.* **2004**, *56* (11), 1537–1563. https://doi.org/10.1016/j.addr.2003.10.047.

(121) Hutchinson, J. A.; Burholt, S.; Hamley, I. W. Peptide Hormones and Lipopeptides: From Self-Assembly to Therapeutic Applications. *J. Pept. Sci.* **2017**, *23* (2), 82–94. https://doi.org/10.1002/psc.2954.

(122) Elgersma, R. C.; Meijneke, T.; de Jong, R.; Brouwer, A. J.; Posthuma, G.; Rijkers, D. T. S.; Liskamp, R. M. J. Synthesis and Structural Investigations of N-Alkylated β-Peptidosulfonamide–Peptide Hybrids of the Amyloidogenic Amylin(20–29) Sequence:

Implications of Supramolecular Folding for the Design of Peptide-Based Bionanomaterials. *Org. Biomol. Chem.* **2006**, *4* (19), 3587–3597. https://doi.org/10.1039/B606875H.

(123) Wolff, M.; Schüler, A.; Gast, K.; Seckler, R.; Evers, A.; Pfeiffer-Marek, S.; Kurz, M.; Nagel, N.; Haack, T.; Wagner, M.; et al. Self-Assembly of Exendin-4-Derived Dual Peptide Agonists Is Mediated by Acylation and Correlated to the Length of Conjugated Fatty Acyl Chains. *Mol. Pharm.* **2020**, *17* (3), 965–978. https://doi.org/10.1021/acs.molpharmaceut.9b01195.

(124) Hamley, I. W.; Kirkham, S.; Dehsorkhi, A.; Castelletto, V.; Reza, M.; Ruokolainen, J. Toll-like Receptor Agonist Lipopeptides Self-Assemble into Distinct Nanostructures. *Chem. Commun.* **2014**, *50* (100), 15948–15951. https://doi.org/10.1039/c4cc07511k.

(125) Hutchinson, J. A.; Hamley, I. W.; Edwards-Gayle, C. J. C.; Castelletto, V.; Piras, C.; Cramer, R.; Kowalczyk, R.; Seitsonen, J.; Ruokolainen, J.; Rambo, R. P. Melanin Production by Tyrosinase Activity on a Tyrosine-Rich Peptide Fragment and PH-Dependent Self-Assembly of Its Lipidated Analogue. *Org. Biomol. Chem.* **2019**, *17* (18), 4543–4553. https://doi.org/10.1039/c9ob00550a.

(126) Miravet, J. F.; Escuder, B.; Segarra-Maset, M. D.; Tena-Solsona, M.; Hamley, I. W.; Dehsorkhi, A.; Castelletto, V. Self-Assembly of a Peptide Amphiphile: Transition from Nanotape Fibrils to Micelles. *Soft Matter* **2013**, *9* (13), 3558–3564. https://doi.org/10.1039/c3sm27899a.

(127) Habila, N.; Kulkarni, K.; Lee, T. H.; Al-Garawi, Z. S.; Serpell, L. C.; Aguilar, M. I.; Del Borgo, M. P. Transition of Nano-Architectures Through Self-Assembly of Lipidated B3-Tripeptide Foldamers. *Front. Chem.* **2020**, *8* (March), 1–11. https://doi.org/10.3389/fchem.2020.00217.

(128) Yu, Y.-C.; Tirrell, M.; Fields, G. B. Minimal Lipidation Stabilizes Protein-Like Molecular Architecture. *J. Am. Chem. Soc.* **1998**, *120* (39), 9979–9987. https://doi.org/10.1021/ja981654z.

(129) Chu-Kung, A. F.; Bozzelli, K. N.; Lockwood, N. A.; Haseman, J. R.; Mayo, K. H.; Tirrell, M. V. Promotion of Peptide Antimicrobial Activity by Fatty Acid Conjugation. *Bioconjug. Chem.* **2004**, *15* (3), 530–535. https://doi.org/10.1021/bc0341573.

(130) Yu, Y.-C.; Berndt, P.; Tirrell, M.; Fields, G. B. Self-Assembling Amphiphiles for Construction of Protein Molecular Architecture. *J. Am. Chem. Soc.* **1996**, *118* (50), 12515–12520. https://doi.org/10.1021/ja9627656.

(131) Astbury, W. T.; Dickinson, S.; Bailey, K. The X-Ray Interpretation of Denaturation and the Structure of the Seed Globulins. *Biochem. J.* **1935**, *29* (10), 2351.

(132) Fändrich, M.; Meinhardt, J.; Grigorieff, N. Structural Polymorphism of Alzheimer Aβ and Other Amyloid Fibrils. *Prion* **2009**, *3* (2), 89–93. https://doi.org/10.4161/pri.3.2.8859.

(133) Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. . Common Core Structure of Amyloid Fibrils by Synchrotron X-Ray Diffraction 1 1Edited by F. E. Cohen. *J. Mol. Biol.* **1997**, *273* (3), 729–739. https://doi.org/10.1006/jmbi.1997.1348.

(134) Westermark, P.; Benson, M. D.; Buxbaum, J. N.; Cohen, A. S.; Frangione, B.; Ikeda, S.-I.; Masters, C. L.; Merlini, G.; Saraiva, M. J.; Sipe, J. D. A Primer of Amyloid Nomenclature. *Amyloid* **2007**, *14* (3), 179–183. https://doi.org/10.1080/13506120701460923.

(135) Eichner, T.; Radford, S. E. A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. *Mol. Cell* **2011**, *43* (1), 8–18. https://doi.org/10.1016/j.molcel.2011.05.012.

(136) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. Ø.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the Cross-β Spine of Amyloid-like Fibrils. *Nature* **2005**, *435* (7043), 773–778. https://doi.org/10.1038/nature03680.

(137) Yoo, J. M.; Lin, Y.; Heo, Y.; Lee, Y. Polymorphism in Alpha-Synuclein Oligomers and Its Implications in Toxicity under Disease Conditions. *Front. Mol. Biosci.* **2022**, *9* (August), 1–10. https://doi.org/10.3389/fmolb.2022.959425.

(138) Kurouski, D.; Dukor, R. K.; Lu, X.; Nafie, L. A.; Lednev, I. K. Normal and Reversed Supramolecular Chirality of Insulin Fibrils Probed by Vibrational Circular Dichroism at the Protofilament Level of Fibril Structure. *Biophys. J.* **2012**, *103* (3), 522–531. https://doi.org/10.1016/j.bpj.2012.04.042.

(139) Kurouski, D.; Lombardi, R. A.; Dukor, R. K.; Lednev, I. K.; Nafie, L. A. Direct Observation and PH Control of Reversed Supramolecular Chirality in Insulin Fibrils by Vibrational Circular Dichroism. *Chem. Commun.* **2010**, *46* (38), 7154. https://doi.org/10.1039/c0cc02423f.

(140) Ragonis-Bachar, P.; Landau, M. Functional and Pathological Amyloid Structures in the Eyes of 2020 Cryo-EM. *Curr. Opin. Struct. Biol.* **2021**, *68*, 184–193. https://doi.org/10.1016/j.sbi.2021.01.006.

(141) Fitzpatrick, A. W. P.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Crowther, R. A.; Ghetti, B.; Goedert, M.; Scheres, S. H. W. Cryo-EM Structures of Tau Filaments from Alzheimer's Disease. *Nature* **2017**, *547* (7662), 185–190. https://doi.org/10.1038/nature23002.

(142) Kollmer, M.; Close, W.; Funk, L.; Rasmussen, J.; Bsoul, A.; Schierhorn, A.; Schmidt, M.; Sigurdson, C. J.; Jucker, M.; Fändrich, M. Cryo-EM Structure and Polymorphism of Aβ Amyloid Fibrils Purified from Alzheimer's Brain Tissue. *Nat. Commun.* **2019**, *10* (1), 4760. https://doi.org/10.1038/s41467-019-12683-8.

(143) Radamaker, L.; Lin, Y.-H.; Annamalai, K.; Huhn, S.; Hegenbart, U.; Schönland, S. O.; Fritz, G.; Schmidt, M.; Fändrich, M. Cryo-EM Structure of a Light Chain-Derived Amyloid Fibril from a Patient with Systemic AL Amyloidosis. *Nat. Commun.* **2019**, *10* (1), 1103. https://doi.org/10.1038/s41467-019-09032-0.

(144) Röder, C.; Kupreichyk, T.; Gremer, L.; Schäfer, L. U.; Pothula, K. R.; Ravelli, R. B. G.; Willbold, D.; Hoyer, W.; Schröder, G. F. Cryo-EM Structure of Islet Amyloid Polypeptide Fibrils Reveals Similarities with Amyloid-β Fibrils. *Nat. Struct. Mol. Biol.* **2020**, *27* (7), 660–667. https://doi.org/10.1038/s41594-020-0442-4.

(145) Swuec, P.; Lavatelli, F.; Tasaki, M.; Paissoni, C.; Rognoni, P.; Maritan, M.; Brambilla, F.; Milani, P.; Mauri, P.; Camilloni, C.; et al. Cryo-EM Structure of Cardiac Amyloid Fibrils from an Immunoglobulin Light Chain AL Amyloidosis Patient. *Nat. Commun.* **2019**, *10* (1), 1269. https://doi.org/10.1038/s41467-019-09133-w.

(146) Iadanza, M. G.; Silvers, R.; Boardman, J.; Smith, H. I.; Karamanos, T. K.; Debelouchina, G. T.; Su, Y.; Griffin, R. G.; Ranson, N. A.; Radford, S. E. The Structure of a B2-Microglobulin Fibril Suggests a Molecular Basis for Its Amyloid Polymorphism. *Nat. Commun.* **2018**, *9* (1), 4517. https://doi.org/10.1038/s41467-018-06761-6.

(147) Liberta, F.; Loerch, S.; Rennegarbe, M.; Schierhorn, A.; Westermark, P.; Westermark, G. T.; Hazenberg, B. P. C.; Grigorieff, N.; Fändrich, M.; Schmidt, M. Cryo-EM Fibril Structures from Systemic AA Amyloidosis Reveal the Species Complementarity of Pathological Amyloids. *Nat. Commun.* **2019**, *10* (1), 1104. https://doi.org/10.1038/s41467-019-09033-z.

(148) Hervas, R.; Rau, M. J.; Park, Y.; Zhang, W.; Murzin, A. G.; Fitzpatrick, J. A. J.; Scheres, S. H. W.; Si, K. Cryo-EM Structure of a Neuronal Functional Amyloid Implicated

in Memory Persistence in Drosophila. *Science (80-.).* **2020**, *367* (6483), 1230–1234. https://doi.org/10.1126/science.aba3526.

(149) Lu, J.; Cao, Q.; Hughes, M. P.; Sawaya, M. R.; Boyer, D. R.; Cascio, D.; Eisenberg, D. S. CryoEM Structure of the Low-Complexity Domain of HnRNPA2 and Its Conversion to Pathogenic Amyloid. *Nat. Commun.* **2020**, *11* (1), 4090. https://doi.org/10.1038/s41467-020-17905-y.

(150) Cao, Q.; Boyer, D. R.; Sawaya, M. R.; Ge, P.; Eisenberg, D. S. Cryo-EM Structures of Four Polymorphic TDP-43 Amyloid Cores. *Nat. Struct. Mol. Biol.* **2019**, *26* (7), 619–627. https://doi.org/10.1038/s41594-019-0248-4.

(151) Fernandez-Escamilla, A.-M.; Rousseau, F.; Schymkowitz, J.; Serrano, L. Prediction of Sequence-Dependent and Mutational Effects on the Aggregation of Peptides and Proteins. *Nat. Biotechnol.* **2004**, *22* (10), 1302–1306. https://doi.org/10.1038/nbt1012.

(152) Conchillo-Solé, O.; de Groot, N. S.; Avilés, F. X.; Vendrell, J.; Daura, X.; Ventura, S. AGGRESCAN: A Server for the Prediction and Evaluation of "Hot Spots" of Aggregation in Polypeptides. *BMC Bioinformatics* **2007**, *8* (1), 65. https://doi.org/10.1186/1471-2105-8-65.

(153) Louros, N.; Orlando, G.; De Vleeschouwer, M.; Rousseau, F.; Schymkowitz, J. Structure-Based Machine-Guided Mapping of Amyloid Sequence Space Reveals Uncharted Sequence Clusters with Higher Solubilities. *Nat. Commun.* **2020**, *11* (1), 1–13. https://doi.org/10.1038/s41467-020-17207-3.

(154) Garbuzynskiy, S. O.; Lobanov, M. Y.; Galzitskaya, O. V. FoldAmyloid: A Method of Prediction of Amyloidogenic Regions from Protein Sequence. *Bioinformatics* **2010**, *26* (3), 326–332. https://doi.org/10.1093/bioinformatics/btp691.

(155) Maurer-Stroh, S.; Debulpaep, M.; Kuemmerer, N.; de la Paz, M. L.; Martins, I. C.; Reumers, J.; Morris, K. L.; Copland, A.; Serpell, L.; Serrano, L.; et al. Exploring the Sequence Determinants of Amyloid Structure Using Position-Specific Scoring Matrices. *Nat. Methods* **2010**, *7* (3), 237–242. https://doi.org/10.1038/nmeth.1432.

(156) Emily, M.; Talvas, A.; Delamarche, C. MetAmyl: A METa-Predictor for AMYLoid Proteins. *PLoS One* **2013**, *8* (11), e79722. https://doi.org/10.1371/journal.pone.0079722.

(157) Kuriata, A.; Iglesias, V.; Pujols, J.; Kurcinski, M.; Kmiecik, S.; Ventura, S. Aggrescan3D (A3D) 2.0: Prediction and Engineering of Protein Solubility. *Nucleic Acids Res.* **2019**, *47* (W1), W300–W307. https://doi.org/10.1093/nar/gkz321.

(158) Kim, C.; Choi, J.; Lee, S. J.; Welsh, W. J.; Yoon, S. NetCSSP: Web Application for Predicting Chameleon Sequences and Amyloid Fibril Formation. *Nucleic Acids Res.* **2009**, *37* (Web Server), W469–W473. https://doi.org/10.1093/nar/gkp351.

(159) Housmans, J. A. J.; Wu, G.; Schymkowitz, J.; Rousseau, F. A Guide to Studying Protein Aggregation. *FEBS J.* **2021**, 1–30. https://doi.org/10.1111/febs.16312.

(160) Mortensen, S. Extracellular Domain of GLP-1 Receptor in Complex with GLP-1 Variant Ala8Hcs/Thr11Hcs. **2017**. https://doi.org/10.2210/pdb5otu/pdb.

(161) Oddo, A.; Mortensen, S.; Thøgersen, H.; De Maria, L.; Hennen, S.; McGuire, J. N.; Kofoed, J.; Linderoth, L.; Reedtz-Runge, S. α-Helix or β-Turn? An Investigation into N-Terminally Constrained Analogues of Glucagon-like Peptide 1 (GLP-1) and Exendin-4. *Biochemistry* **2018**, *57* (28), 4148–4154. https://doi.org/10.1021/acs.biochem.8b00105.

(162) Tayeb-Fligelman, E.; Salinas, N.; Tabachnikov, O.; Landau, M. Staphylococcus Aureus PSMα3 Cross-α Fibril Polymorphism and Determinants of Cytotoxicity. *Structure* **2020**, *28* (3), 301-313.e6. https://doi.org/10.1016/j.str.2019.12.006.

(163) Salinas, N.; Colletier, J. P.; Moshe, A.; Landau, M. Extreme Amyloid Polymorphism in Staphylococcus Aureus Virulent PSMα Peptides. *Nat. Commun.* **2018**, *9* (1). https://doi.org/10.1038/s41467-018-05490-0.

(164) Tayeb-Fligelman, E.; Tabachnikov, O.; Moshe, A.; Goldshmidt-Tran, O.; Sawaya, M. R.; Coquelle, N.; Colletier, J. P.; Landau, M. The Cytotoxic Staphylococcus Aureus PSMα3 Reveals a Cross-α Amyloid-like Fibril. *Science (80-.).* **2017**, *355* (6327), 831–833. https://doi.org/10.1126/science.aaf4901.

(165) Salinas, N.; Tayeb-Fligelman, E.; Sammito, M. D.; Bloch, D.; Jelinek, R.; Noy, D.; Usón, I.; Landau, M. The Amphibian Antimicrobial Peptide Uperin 3.5 Is a Cross-α/Cross-β Chameleon Functional Amyloid. *Proc. Natl. Acad. Sci.* **2021**, *118* (3). https://doi.org/10.1073/pnas.2014442118.

(166) Landau, M. Mimicking Cross-α Amyloids. *Nat. Chem. Biol.* **2018**, *14* (9), 833–834. https://doi.org/10.1038/s41589-018-0118-0.

(167) Martins, P. M.; Navarro, S.; Silva, A.; Pinto, M. F.; Sárkány, Z.; Figueiredo, F.; Pereira, P. J. B.; Pinheiro, F.; Bednarikova, Z.; Burdukiewicz, M.; et al. MIRRAGGE – Minimum Information Required for Reproducible AGGregation Experiments. *Front. Mol. Neurosci.* **2020**, *13* (November), 1–18. https://doi.org/10.3389/fnmol.2020.582488.

(168) Ruggeri, F. S.; Vieweg, S.; Cendrowska, U.; Longo, G.; Chiki, A.; Lashuel, H. A.; Dietler, G. Nanoscale Studies Link Amyloid Maturity with Polyglutamine Diseases Onset. *Sci. Rep.* **2016**, *6* (1), 31155. https://doi.org/10.1038/srep31155.

(169) Adamcik, J.; Mezzenga, R. Study of Amyloid Fibrils via Atomic Force Microscopy. *Curr. Opin. Colloid Interface Sci.* **2012**, *17* (6), 369–376. https://doi.org/10.1016/j.cocis.2012.08.001.

(170) Jurado, R.; Adamcik, J.; Sánchez-Ferrer, A.; Bolisetty, S.; Mezzenga, R.; Gálvez, N. Understanding the Formation of Apoferritin Amyloid Fibrils. *Biomacromolecules* **2021**, *22* (5), 2057–2066. https://doi.org/10.1021/acs.biomac.1c00176.

(171) Fitzpatrick, A. W.; Saibil, H. R. Cryo-EM of Amyloid Fibrils and Cellular Aggregates. *Curr. Opin. Struct. Biol.* **2019**, *58*, 34–42. https://doi.org/10.1016/j.sbi.2019.05.003.

(172) Iadanza, M. G.; Jackson, M. P.; Hewitt, E. W.; Ranson, N. A.; Radford, S. E. A New Era for Understanding Amyloid Structures and Disease. *Nat. Rev. Mol. Cell Biol.* **2018**, *19* (12), 755–773. https://doi.org/10.1038/s41580-018-0060-8.

(173) Vilar, M.; Chou, H.-T.; Lührs, T.; Maji, S. K.; Riek-Loher, D.; Verel, R.; Manning, G.; Stahlberg, H.; Riek, R. The Fold of α-Synuclein Fibrils. *Proc. Natl. Acad. Sci.* **2008**, *105* (25), 8637–8642. https://doi.org/10.1073/pnas.0712179105.

(174) Wälti, M. A.; Ravotti, F.; Arai, H.; Glabe, C. G.; Wall, J. S.; Böckmann, A.; Güntert, P.; Meier, B. H.; Riek, R. Atomic-Resolution Structure of a Disease-Relevant Aβ(1–42) Amyloid Fibril. *Proc. Natl. Acad. Sci.* **2016**, *113* (34). https://doi.org/10.1073/pnas.1600749113.

(175) Rodriguez, J. A.; Ivanova, M. I.; Sawaya, M. R.; Cascio, D.; Reyes, F. E.; Shi, D.; Sangwan, S.; Guenther, E. L.; Johnson, L. M.; Zhang, M.; et al. Structure of the Toxic Core of α-Synuclein from Invisible Crystals. *Nature* **2015**, *525* (7570), 486–490. https://doi.org/10.1038/nature15368.

(176) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; et al. Atomic Structures of Amyloid Cross-β Spines Reveal Varied Steric Zippers. *Nature* **2007**, *447* (7143), 453–457. https://doi.org/10.1038/nature05695.

(177) Fitzpatrick, A. W.; Saibil, H. R. Cryo-EM of Amyloid Fibrils and Cellular Aggregates. *Curr. Opin. Struct. Biol.* **2019**, *58*, 34–42. https://doi.org/10.1016/j.sbi.2019.05.003.

(178) Kühlbrandt, W. The Resolution Revolution. *Science (80-.).* **2014**, *343* (6178), 1443–1444. https://doi.org/10.1126/science.1251652.

(179) Cristóvão, J. S.; Henriques, B. J.; Gomes, C. M. Biophysical and Spectroscopic Methods for Monitoring Protein Misfolding and Amyloid Aggregation; 2019; pp 3–18. https://doi.org/10.1007/978-1-4939-8820-4_1.

(180) Whitmore, L.; Wallace, B. A. Protein Secondary Structure Analyses from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers* **2008**, *89* (5), 392–400. https://doi.org/10.1002/bip.20853.

(181) Whitmore, L.; Wallace, B. A. DICHROWEB, an Online Server for Protein Secondary Structure Analyses from Circular Dichroism Spectroscopic Data. *Nucleic Acids Res.* **2004**, *32*, 668–673. https://doi.org/10.1093/nar/gkh371.

(182) Sreerama, N.; Woody, R. W. Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Anal. Biochem.* **2000**, *287* (2), 252–260. https://doi.org/10.1006/abio.2000.4880.

(183) Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y.; Goto, Y.; Réfrégiers, M.; Kardos, J. Accurate Secondary Structure Prediction and Fold Recognition for Circular Dichroism Spectroscopy. *Proc. Natl. Acad. Sci.* **2015**, *112* (24), 3095–3103. https://doi.org/10.1073/pnas.1500851112.

(184) Micsonai, A.; Wien, F.; Bulyáki, É.; Kun, J.; Moussong, É.; Lee, Y.-H.; Goto, Y.; Réfrégiers, M.; Kardos, J. BeStSel: A Web Server for Accurate Protein Secondary Structure Prediction and Fold Recognition from the Circular Dichroism Spectra. *Nucleic Acids Res.* **2018**, *46* (W1), W315–W322. https://doi.org/10.1093/nar/gky497.

(185) Micsonai, A.; Bulyáki, É.; Kardos, J. BeStSel: From Secondary Structure Analysis to Protein Fold Prediction by Circular Dichroism Spectroscopy; 2021; pp 175–189. https://doi.org/10.1007/978-1-0716-0892-0_11.

(186) Micsonai, A.; Moussong, É.; Murvai, N.; Tantos, Á.; Tőke, O.; Réfrégiers, M.; Wien, F.; Kardos, J. Disordered–Ordered Protein Binary Classification by Circular

Dichroism Spectroscopy. *Front. Mol. Biosci.* **2022**, *9*. https://doi.org/10.3389/fmolb.2022.863141.

(187) Micsonai, A.; Moussong, É.; Wien, F.; Boros, E.; Vadászi, H.; Murvai, N.; Lee, Y.-H.; Molnár, T.; Réfrégiers, M.; Goto, Y.; et al. BeStSel: Webserver for Secondary Structure and Fold Prediction for Protein CD Spectroscopy. *Nucleic Acids Res.* **2022**, *50* (W1), W90–W98. https://doi.org/10.1093/nar/gkac345.

(188) Manavalan, P.; Johnson, W. C. Variable Selection Method Improves the Prediction of Protein Secondary Structure from Circular Dichroism Spectra. *Anal. Biochem.* **1987**, *167* (1), 76–85. https://doi.org/10.1016/0003-2697(87)90135-7.

(189) Miles, A. J.; Ramalli, S. G.; Wallace, B. A. DichroWeb , a Website for Calculating Protein Secondary Structure from Circular Dichroism Spectroscopic Data. **2022**, No. June 2021, 37–46. https://doi.org/10.1002/pro.4153.

(190) Lees, J. G.; Miles, A. J.; Wien, F.; Wallace, B. A. A Reference Database for Circular Dichroism Spectroscopy Covering Fold and Secondary Structure Space. *Bioinformatics* **2006**, *22* (16), 1955–1962. https://doi.org/10.1093/bioinformatics/btl327.

(191) Kelly, S. M.; Jess, T. J.; Price, N. C. How to Study Proteins by Circular Dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1751* (2), 119–139. https://doi.org/10.1016/j.bbapap.2005.06.005.

(192) Andersson, D.; Carlsson, U.; Freskgård, P.-O. Contribution of Tryptophan Residues to the CD Spectrum of the Extracellular Domain of Human Tissue Factor. *Eur. J. Biochem.* **2001**, *268* (4), 1118–1128. https://doi.org/10.1046/j.1432-1327.2001.01981.x.

(193) Ozaki, Y.; Baranska, M.; Lednev, I. K.; Wood, B. R. Vibrational Spectroscopy in Protein Research: From Purified Proteins to Aggregates and Assemblies; Elsevier Science, 2020.

(194) *Peptide Self-Assembly*; Nilsson, B. L., Doran, T. M., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2018; Vol. 1777. https://doi.org/10.1007/978-1-4939-7811-3.

(195) Nafie, L. A. *Vibrational Optical Activity*; John Wiley & Sons, Ltd: Chichester, UK, 2011. https://doi.org/10.1002/9781119976516.

(196) Ma, S.; Cao, X.; Mak, M.; Sadik, A.; Walkner, C.; Freedman, T. B.; Lednev, I. K.; Dukor, R. K.; Nafie, L. A. Vibrational Circular Dichroism Shows Unusual Sensitivity to

Protein Fibril Formation and Development in Solution. *J. Am. Chem. Soc.* **2007**, *129* (41), 12364–12365. https://doi.org/10.1021/ja074188z.

(197) Pazderková, M.; Pazderka, T.; Shanmugasundaram, M.; Dukor, R. K.; Lednev, I. K.; Nafie, L. A. Origin of Enhanced VCD in Amyloid Fibril Spectra: Effect of Deuteriation and PH. *Chirality* **2017**, *29* (9), 469–475. https://doi.org/10.1002/chir.22722.

(198) Kurouski, D.; Dukor, R. K.; Lu, X.; Nafie, L. A.; Lednev, I. K. Normal and Reversed Supramolecular Chirality of Insulin Fibrils Probed by Vibrational Circular Dichroism at the Protofilament Level of Fibril Structure. *Biophys. J.* **2012**, *103* (3), 522–531. https://doi.org/10.1016/j.bpj.2012.04.042.

(199) Krupová, M.; Kessler, J.; Bour, P. Recent Trends in Chiroptical Spectroscopy: Theory and Applications of Vibrational Circular Dichroism and Raman Optical Activity. *Chempluschem* **2020**, 1–16. https://doi.org/10.1002/cplu.202000014.

(200) Ma, S.; Cao, X.; Mak, M.; Sadik, A.; Walkner, C.; Freedman, T. B.; Lednev, I. K.; Dukor, R. K.; Nafie, L. A. Vibrational Circular Dichroism Shows Unusual Sensitivity to Protein Fibril Formation and Development in Solution. *J. Am. Chem. Soc.* **2007**, *129* (41), 12364–12365. https://doi.org/10.1021/ja074188z.

(201) Measey, T. J.; Smith, K. B.; Decatur, S. M.; Zhao, L.; Yang, G.; Schweitzer-Stenner, R. Self-Aggregation of a Polyalanine Octamer Promoted by Its C-Terminal Tyrosine and Probed by a Strongly Enhanced Vibrational Circular Dichroism Signal. *J. Am. Chem. Soc.* **2009**, *131* (51), 18218–18219. https://doi.org/10.1021/ja908324m.

(202) Measey, T. J.; Schweitzer-Stenner, R. Vibrational Circular Dichroism as a Probe of Fibrillogenesis: The Origin of the Anomalous Intensity Enhancement of Amyloid-like Fibrils. *J. Am. Chem. Soc.* **2011**, *133* (4), 1066–1076. https://doi.org/10.1021/ja1089827.

(203) Keiderling, T. A. Structure of Condensed Phase Peptides : Insights from Vibrational Circular Dichroism and Raman Optical Activity Techniques. **2020**. https://doi.org/10.1021/acs.chemrev.9b00636.

(204) Krupová, M.; Kessler, J.; Bouř, P. Polymorphism of Amyloid Fibrils Induced by Catalytic Seeding: A Vibrational Circular Dichroism Study. *ChemPhysChem* **2021**, *22* (1), 83–91. https://doi.org/10.1002/cphc.202000797.

(205) Van de Vondel, E.; Baatsen, P.; Van Elzen, R.; Lambeir, A.-M.; Keiderling, T. A.; Herrebout, W. A.; Johannessen, C. Vibrational Circular Dichroism Sheds New Light on the Competitive Effects of Crowding and β-Synuclein on the Fibrillation Process of α-

Synuclein. *Biochemistry* **2018**, *57* (41), 5989–5995. https://doi.org/10.1021/acs.biochem.8b00780.

(206) Krupová, M.; Kessler, J.; Bouř, P. Recent Trends in Chiroptical Spectroscopy: Theory and Applications of Vibrational Circular Dichroism and Raman Optical Activity. *Chempluschem* **2020**, *85* (3), 561–575. https://doi.org/10.1002/cplu.202000014.

(207) Dusa, A.; Kaylor, J.; Edridge, S.; Bodner, N.; Hong, D.-P.; Fink, A. L. Characterization of Oligomers during α-Synuclein Aggregation Using Intrinsic Tryptophan Fluorescence. *Biochemistry* **2006**, *45* (8), 2752–2760. https://doi.org/10.1021/bi051426z.

(208) Gorinstein, S.; Goshev, I.; Moncheva, S.; Zemser, M.; Weisz, M.; Caspi, A.; Libman, I.; Lerner, H. T.; Trakhtenberg, S.; Martín-Belloso, O. Intrinsic Tryptophan Fluorescence of Human Serum Proteins and Related Conformational Changes. *J. Protein Chem.* **2000**, *19* (8), 637–642. https://doi.org/10.1023/A:1007192017291.

(209) Gade Malmos, K.; Blancas-Mejia, L. M.; Weber, B.; Buchner, J.; Ramirez-Alvarado, M.; Naiki, H.; Otzen, D. ThT 101: A Primer on the Use of Thioflavin T to Investigate Amyloid Formation. *Amyloid* **2017**, *24* (1), 1–16. https://doi.org/10.1080/13506129.2017.1304905.

(210) Naiki, H.; Higuchi, K.; Hosokawa, M.; Takeda, T. Fluorometric Determination of Amyloid Fibrils in Vitro Using the Fluorescent Dye, Thioflavine T. *Anal. Biochem.* **1989**, *177* (2), 244–249. https://doi.org/10.1016/0003-2697(89)90046-8.

(211) Biancalana, M.; Koide, S. Molecular Mechanism of Thioflavin-T Binding to Amyloid Fibrils. *Biochim. Biophys. Acta - Proteins Proteomics* **2010**, *1804* (7), 1405–1412. https://doi.org/10.1016/j.bbapap.2010.04.001.

(212) Yeasmin Khusbu, F.; Zhou, X.; Chen, H.; Ma, C.; Wang, K. Thioflavin T as a Fluorescence Probe for Biosensing Applications. *TrAC - Trends Anal. Chem.* **2018**, *109*, 1–18. https://doi.org/10.1016/j.trac.2018.09.013.

(213) De Ferrari, G. V.; Mallender, W. D.; Inestrosa, N. C.; Rosenberry, T. L. Thioflavin T Is a Fluorescent Probe of the Acetylcholinesterase Peripheral Site That Reveals Conformational Interactions between the Peripheral and Acylation Sites. *J. Biol. Chem.* **2001**, *276* (26), 23282–23287. https://doi.org/10.1074/jbc.M009596200.

(214) Groenning, M.; Olsen, L.; van de Weert, M.; Flink, J. M.; Frokjaer, S.; Jørgensen, F. S. Study on the Binding of Thioflavin T to β-Sheet-Rich and Non-β-Sheet Cavities. *J. Struct. Biol.* **2007**, *158* (3), 358–369. https://doi.org/10.1016/j.jsb.2006.12.010.

(215) Carrotta, R.; Bauer, R.; Waninge, R.; Rischel, C. Conformational Characterization of Oligomeric Intermediates and Aggregates in β-Lactoglobulin Heat Aggregation. *Protein Sci.* **2008**, *10* (7), 1312–1318. https://doi.org/10.1110/ps.42501.

(216) Grudzielanek, S.; Smirnovas, V.; Winter, R. Solvation-Assisted Pressure Tuning of Insulin Fibrillation: From Novel Aggregation Pathways to Biotechnological Applications. *J. Mol. Biol.* **2006**, *356* (2), 497–509. https://doi.org/10.1016/j.jmb.2005.11.075.

(217) Arad, E.; Green, H.; Jelinek, R.; Rapaport, H. Revisiting Thioflavin T (ThT) Fluorescence as a Marker of Protein Fibrillation – The Prominent Role of Electrostatic Interactions. *J. Colloid Interface Sci.* **2020**, *573*, 87–95. https://doi.org/10.1016/j.jcis.2020.03.075.

(218) Zapadka, K. L. Understanding the Mechanism of Aggregation and Amyloid Fibril Formation of Glucagon-like Peptide-1, University of Cambridge, 2016.

(219) Younan, N. D.; Viles, J. H. A Comparison of Three Fluorophores for the Detection of Amyloid Fibers and Prefibrillar Oligomeric Assemblies. ThT (Thioflavin T); ANS (1-Anilinonaphthalene-8-Sulfonic Acid); and BisANS (4,4′-Dianilino-1,1′-Binaphthyl-5,5′-Disulfonic Acid). *Biochemistry* **2015**, *54* (28), 4297–4306. https://doi.org/10.1021/acs.biochem.5b00309.

(220) Lebowitz, J.; Lewis, M. S.; Schuck, P. Modern Analytical Ultracentrifugation in Protein Science: A Tutorial Review. *Protein Sci.* **2009**, *11* (9), 2067–2079. https://doi.org/10.1110/ps.0207702.

(221) Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **2000**, *78* (3), 1606–1619. https://doi.org/10.1016/S0006-3495(00)76713-0.

(222) Brown, P. H.; Balbo, A.; Schuck, P. Characterizing Protein-Protein Interactions by Sedimentation Velocity Analytical Ultracentrifugation. *Curr. Protoc. Immunol.* **2008**, *81* (1). https://doi.org/10.1002/0471142735.im1815s81.

(223) Teller, D. C.; Swanson, E.; de Haën, C. [8] The Translational Friction Coefficient of Proteins; 1979; pp 104–124. https://doi.org/10.1016/0076-6879(79)61010-8.

(224) Lamm, O. Die Differentialgleichung Der Ultrazentrifugierung; Almqvist & Wiksell, 1929.

(225) Becher, F. J. Studies of the Physical Stability of GLP-1 & Chemically Modified Forms of GLP-1, University of Cambridge, 2020. https://doi.org/10.17863/CAM.78083.

(226) ExPASy: Bioinformatics Resource Portal expasy.org.

(227) Chen, Y.-H.; Yang, J. T.; Chau, K. H. Determination of the Helix and β Form of Proteins in Aqueous Solution by Circular Dichroism. *Biochemistry* **1974**, *13* (16), 3350–3359.

(228) Muñoz, V.; Serrano, L. Elucidating the Folding Problem of Helical Peptides Using Empirical Parameters. III>Temperature and PH Dependence. *J. Mol. Biol.* **1995**, *245* (3), 297–308. https://doi.org/10.1006/jmbi.1994.0024.

(229) Boreikaite, V.; Wicky, B. I. M.; Watt, I. N.; Clarke, J.; Walker, J. E. Extrinsic Conditions Influence the Self-Association and Structure of IF1, the Regulatory Protein of Mitochondrial ATP Synthase. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116* (21), 10354–10359. https://doi.org/10.1073/pnas.1903535116.

(230) Wales, D. J. GMIN: A Program for Finding Global Minima and Calculating Thermodynamic Properties. University of Cambridge, Cambridge, UK 2018.

(231) Wales, D. J. OPTIM - a Program for Optimising Geometries and Calculating Reaction Pathways. University of Cambridge, Cambridge, UK 2018.

(232) Wales, D. J. PATHSAMPLE: A Driver for OPTIM to Create Stationary Point Databases Using Discrete Path Sampling and Perform Kinetic Analysis. University of Cambridge, Cambridge, UK 2018.

(233) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106* (3), 765–784. https://doi.org/10.1021/ja00315a051.

(234) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7* (2), 230–252. https://doi.org/10.1002/jcc.540070216.

(235) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of

Molecules. *Comput. Phys. Commun.* **1995**, *91* (1–3), 1–41. https://doi.org/10.1016/0010-4655(95)00041-D.

(236) Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M. AMBER 12; University of California: San Francisco, 2012. 2010, pp 1–826.

(237) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

(238) Małolepsza, E.; Strodel, B.; Khalili, M.; Trygubenko, S.; Fejer, S. N.; Wales, D. J. Symmetrization of the AMBER and CHARMM Force Fields. *J. Comput. Chem.* **2010**, *31* (7), 1402–1409. https://doi.org/10.1002/jcc.21425.

(239) Małolepsza, E.; Strodel, B.; Khalili, M.; Trygubenko, S.; Fejer, S.; Carr, J. M.; Wales, D. J. Erratum: Symmetrization of the AMBER and CHARMM Force Fields [J. Comp. Chem. 31, 1402]. *J. Comput. Chem.* **2012**, *33* (27), 2209.

(240) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins Struct. Funct. Bioinforma.* **2006**, *65* (3), 712–725. https://doi.org/10.1002/prot.21123.

(241) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104* (15), 3712–3720. https://doi.org/10.1021/jp994072s.

(242) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. Application of a Pairwise Generalized Born Model to Proteins and Nucleic Acids: Inclusion of Salt Effects. *Theor. Chem. Acc.* **1999**, *101* (6), 426–434. https://doi.org/10.1007/s002140050460.

(243) Wales, D. J. Discrete Path Sampling. *Mol. Phys.* **2002**, *100* (20), 3285–3305. https://doi.org/10.1080/00268970210162691.

(244) Wales, D. J. Some Further Applications of Discrete Path Sampling to Cluster Isomerization. *Mol. Phys.* **2004**, *102* (9–10), 891–908. https://doi.org/10.1080/00268970410001703363.

(245) Noé, F.; Fischer, S. Transition Networks for Modeling the Kinetics of Conformational Change in Macromolecules. *Curr. Opin. Struct. Biol.* **2008**, *18* (2), 154–162. https://doi.org/https://doi.org/10.1016/j.sbi.2008.01.008.

(246) Wales, D. J. Energy Landscapes: Some New Horizons. *Curr. Opin. Struct. Biol.* **2010**, *20* (1), 3–10. https://doi.org/https://doi.org/10.1016/j.sbi.2009.12.011.

(247) Trygubenko, S. A.; Wales, D. J. A Doubly Nudged Elastic Band Method for Finding Transition States. *J. Chem. Phys.* **2004**, *120* (5), 2082–2094. https://doi.org/10.1063/1.1636455.

(248) Henkelman, G.; Jónsson, H. Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points. *J. Chem. Phys.* **2000**, *113* (22), 9978–9985. https://doi.org/10.1063/1.1323224.

(249) Sheppard, D.; Terrell, R.; Henkelman, G. Optimization Methods for Finding Minimum Energy Paths. *J. Chem. Phys.* **2008**, *128* (13), 134106. https://doi.org/10.1063/1.2841941.

(250) Munro, L. J.; Wales, D. J. Defect Migration in Crystalline Silicon. *Phys. Rev. B* **1999**, *59* (6), 3969–3980. https://doi.org/10.1103/PhysRevB.59.3969.

(251) Nocedal, J. Updating Quasi-Newton Matrices with Limited Storage. *Math. Comput.* **1980**, *35* (151), 773–773. https://doi.org/10.1090/S0025-5718-1980-0572855-7.

(252) Liu, D. C.; Nocedal, J. On the Limited Memory BFGS Method for Large Scale Optimization. *Math. Program.* **1989**, *45* (1–3), 503–528. https://doi.org/10.1007/BF01589116.

(253) Carr, J. M.; Trygubenko, S. A.; Wales, D. J. Finding Pathways between Distant Local Minima. *J. Chem. Phys.* **2005**, *122* (23), 234903. https://doi.org/10.1063/1.1931587.

(254) Strodel, B.; Whittleston, C. S.; Wales, D. J. Thermodynamics and Kinetics of Aggregation for the GNNQQNY Peptide. *J. Am. Chem. Soc.* **2007**, *129* (51), 16005–16014. https://doi.org/10.1021/ja075346p.

(255) Röder, K.; Wales, D. J. Energy Landscapes for the Aggregation of Aβ17-42. *J. Am. Chem. Soc.* **2018**, *140* (11), 4018–4027. https://doi.org/10.1021/jacs.7b12896.

(256) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095. https://doi.org/10.1021/ct400341p.

(257) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577–2637. https://doi.org/10.1002/bip.360221211.

(258) Wales, D. *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*; Cambridge University Press: Cambridge, 2004. https://doi.org/DOI: 10.1017/CBO9780511721724.

(259) PyMOL Molecular Graphics System. Schrödinger, L.L.C.: Cambridge, MA, 2013.

(260) Poon, S.; Birkett, N.; Fowler, S.; Luisi, B.; Dobson, C.; Zurdo, J. Amyloidogenicity and Aggregate Cytotoxicity of Human Glucagon-Like Peptide-1 (HGLP-1). *Protein Pept. Lett.* **2009**, *16* (12), 1548–1556. https://doi.org/10.2174/092986609789839232.

(261) Jha, N. N.; Anoop, A.; Ranganathan, S.; Mohite, G. M.; Padinhateeri, R.; Maji, S. K. Characterization of Amyloid Formation by Glucagon-Like Peptides: Role of Basic Residues in Heparin-Mediated Aggregation. *Biochemistry* **2013**, *52* (49), 8800–8810. https://doi.org/10.1021/bi401398k.

(262) Deva, T.; Lorenzen, N.; Vad, B. S.; Petersen, S. V; Thørgersen, I.; Enghild, J. J.; Kristensen, T.; Otzen, D. E. Off-Pathway Aggregation Can Inhibit Fibrillation at High Protein Concentrations. *Biochim. Biophys. Acta - Proteins Proteomics* **2013**, *1834* (3), 677–687. https://doi.org/10.1016/j.bbapap.2012.12.020.

(263) Souillac, P. O.; Uversky, V. N.; Fink, A. L. Structural Transformations of Oligomeric Intermediates in the Fibrillation of the Immunoglobulin Light Chain LEN †. *Biochemistry* **2003**, *42* (26), 8094–8104. https://doi.org/10.1021/bi034652m.

(264) Raynal, B.; Lenormand, P.; Baron, B.; Hoos, S.; England, P. Quality Assessment and Optimization of Purified Protein Samples: Why and How? *Microb. Cell Fact.* **2010**, *13* (1), 1–10. https://doi.org/10.1186/s12934-014-0180-6.

(265) ROBINSON, A. B.; RUDD, C. J. Deamidation of Glutaminyl and Asparaginyl Residues in Peptides and Proteins. In *Current Topics in Cellular Regulation*; ACADEMIC PRESS, INC., 1974; Vol. 8, pp 247–295. https://doi.org/10.1016/B978-0-12-152808-9.50013-4.

(266) Robinson, N. E.; Robinson, Z. W.; Robinson, B. R.; Robinson, A. L.; Robinson, J. A.; Robinson, M. L.; Robinson, A. B. Structure-Dependent Nonenzymatic Deamidation of Glutaminyl and Asparaginyl Pentapeptides. *J. Pept. Res.* **2004**, *63* (5), 426–436. https://doi.org/10.1111/j.1399-3011.2004.00151.x.

(267) Tonie Wright, H.; Urry, D. W. Nonenzymatic Deamidation of Asparaginyl and Glutaminyl Residues in Protein. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26* (1), 1–52. https://doi.org/10.3109/10409239109081719.

(268) Fusco, G.; Chen, S. W.; Williamson, P. T. F.; Cascella, R.; Perni, M.; Jarvis, J. A.; Cecchi, C.; Vendruscolo, M.; Chiti, F.; Cremades, N.; et al. Structural Basis of Membrane Disruption and Cellular Toxicity by α-Synuclein Oligomers. *Science (80-.).* **2017**, *358* (6369), 1440–1443. https://doi.org/10.1126/science.aan6160.

(269) Wu, J. W.; Breydo, L.; Isas, J. M.; Lee, J.; Kuznetsov, Y. G.; Langen, R.; Glabe, C. Fibrillar Oligomers Nucleate the Oligomerization of Monomeric Amyloid β but Do Not Seed Fibril Formation. *J. Biol. Chem.* **2010**, *285* (9), 6071–6079. https://doi.org/10.1074/jbc.M109.069542.

(270) Vivoli Vega, M.; Cascella, R.; Chen, S. W.; Fusco, G.; De Simone, A.; Dobson, C. M.; Cecchi, C.; Chiti, F. The Toxicity of Misfolded Protein Oligomers Is Independent of Their Secondary Structure. *ACS Chem. Biol.* **2019**, *14* (7), 1593–1600. https://doi.org/10.1021/acschembio.9b00324.

(271) Turner, A. L.; Watson, M.; Wilkins, O. G.; Cato, L.; Travers, A.; Thomas, J. O.; Stott, K. Highly Disordered Histone H1−DNA Model Complexes and Their Condensates. *Proc. Natl. Acad. Sci.* **2018**, *115* (47), 11964–11969. https://doi.org/10.1073/pnas.1805943115.

(272) Borgia, A.; Borgia, M. B.; Bugge, K.; Kissling, V. M.; Heidarsson, P. O.; Fernandes, C. B.; Sottini, A.; Soranno, A.; Buholzer, K. J.; Nettels, D.; et al. Extreme Disorder in an Ultrahigh-Affinity Protein Complex. *Nature* **2018**, *555* (7694), 61–66. https://doi.org/10.1038/nature25762.

(273) Maya-Martinez, R.; Xu, Y.; Guthertz, N.; Walko, M.; Karamanos, T. K.; Sobott, F.; Breeze, A. L.; Radford, S. E. Dimers of D76N-B2-Microglobulin Display Potent Antiamyloid Aggregation Activity. *J. Biol. Chem.* **2022**, *298* (12), 102659. https://doi.org/10.1016/j.jbc.2022.102659.

(274) Joseph, J. A.; Röder, K.; Chakraborty, D.; Mantell, R. G.; Wales, D. J. Exploring Biomolecular Energy Landscapes. *Chem. Commun.* **2017**, *53* (52), 6974–6988. https://doi.org/10.1039/C7CC02413D.

(275) Becker, O. M.; Karplus, M. The Topology of Multidimensional Potential Energy Surfaces: Theory and Application to Peptide Structure and Kinetics. *J. Chem. Phys.* **1997**, *106* (4), 1495–1517. https://doi.org/10.1063/1.473299.

(276) Wales, D. J.; Miller, M. A.; Walsh, T. R. Archetypal Energy Landscapes. *Nature* **1998**, *394* (6695), 758–760. https://doi.org/10.1038/29487.

(277) Löhr, T.; Kohlhoff, K.; Heller, G. T.; Camilloni, C.; Vendruscolo, M. A Kinetic Ensemble of the Alzheimer's Aβ Peptide. *Nat. Comput. Sci.* **2021**, *1* (1), 71–78. https://doi.org/10.1038/s43588-020-00003-w.

(278) Chebaro, Y.; Ballard, A. J.; Chakraborty, D.; Wales, D. J. Intrinsically Disordered Energy Landscapes. *Sci. Rep.* **2015**, *5* (1), 10386. https://doi.org/10.1038/srep10386.

(279) Joseph, J. A.; Wales, D. J. Intrinsically Disordered Landscapes for Human CD4 Receptor Peptide. *J. Phys. Chem. B* **2018**, *122* (50), 11906–11921. https://doi.org/10.1021/acs.jpcb.8b08371.

(280) Wales, D. Energy Landscapes: Applications to Clusters, Biomolecules and Glasses; Cambridge University Press, 2003.

(281) Li, Z.; Scheraga, H. A. Monte Carlo-Minimization Approach to the Multiple-Minima Problem in Protein Folding. *Proc. Natl. Acad. Sci.* **1987**, *84* (19), 6611 LP – 6615. https://doi.org/10.1073/pnas.84.19.6611.

(282) Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116. https://doi.org/10.1021/jp970984n.

(283) Wales, D. J.; Scheraga, H. A. Global Optimization of Clusters, Crystals, and Biomolecules. *Science (80-.).* **1999**, *285* (5432), 1368–1372. https://doi.org/10.1126/science.285.5432.1368.

(284) Sutherland-Cash, K. H.; Wales, D. J.; Chakrabarti, D. Free Energy Basin-Hopping. *Chem. Phys. Lett.* **2015**, *625*, 1–4. https://doi.org/10.1016/j.cplett.2015.02.015.

(285) Wales, D. J. Discrete Path Sampling. *Mol. Phys.* **2002**, *100* (20), 3285–3305. https://doi.org/10.1080/00268970210162691.

(286) Röder, K.; Joseph, J. A.; Husic, B. E.; Wales, D. J. Energy Landscapes for Proteins: From Single Funnels to Multifunctional Systems. *Adv. Theory Simulations* **2019**, *2* (4), 1800175. https://doi.org/10.1002/adts.201800175.

(287) Röder, K.; Wales, D. J. Energy Landscapes for the Aggregation of Aβ17-42. *J. Am. Chem. Soc.* **2018**, *140* (11), 4018–4027. https://doi.org/10.1021/jacs.7b12896.

(288) Joseph, J. A.; Chakraborty, D.; Wales, D. J. Energy Landscape for Fold-Switching in Regulatory Protein RfaH. *J. Chem. Theory Comput.* **2019**, *15* (1), 731–742. https://doi.org/10.1021/acs.jctc.8b00912.

(289) Strodel, B.; Wales, D. J. Free Energy Surfaces from an Extended Harmonic Superposition Approach and Kinetics for Alanine Dipeptide. *Chem. Phys. Lett.* **2008**, *466* (4), 105–115. https://doi.org/https://doi.org/10.1016/j.cplett.2008.10.085.

(290) Reid, L. M.; Guzzetti, I.; Svensson, T.; Carlsson, A.-C.; Su, W.; Leek, T.; von Sydow, L.; Czechtizky, W.; Miljak, M.; Verma, C.; et al. How Well Does Molecular Simulation Reproduce Environment-Specific Conformations of the Intrinsically Disordered Peptides PLP, TP2 and ONEG? *Chem. Sci.* **2022**, *13* (7), 1957–1971. https://doi.org/10.1039/D1SC03496K.

(291) Robinson, M. K.; Monroe, J. I.; Shell, M. S. Are AMBER Force Fields and Implicit Solvation Models Additive? A Folding Study with a Balanced Peptide Test Set. *J. Chem. Theory Comput.* **2016**, *12* (11), 5631–5642. https://doi.org/10.1021/acs.jctc.6b00788.

(292) Duong, V. T.; Chen, Z.; Thapa, M. T.; Luo, R. Computational Studies of Intrinsically Disordered Proteins. *J. Phys. Chem. B* **2018**, *122* (46), 10455–10469. https://doi.org/10.1021/acs.jpcb.8b09029.

(293) Wales, D. J. Exploring Energy Landscapes. *Annu. Rev. Phys. Chem.* **2018**, *69* (1), 401–425. https://doi.org/10.1146/annurev-physchem-050317-021219.

(294) Erak, M.; Bellmann-Sickert, K.; Els-Heindl, S.; Beck-Sickinger, A. G. Peptide Chemistry Toolbox – Transforming Natural Peptides into Peptide Therapeutics. *Bioorganic Med. Chem.* **2018**, *26* (10), 2759–2765. https://doi.org/10.1016/j.bmc.2018.01.012.

(295) Kieffer, T. J.; McIntosh, C. H.; Pederson, R. A. Degradation of Glucose-Dependent Insulinotropic Polypeptide and Truncated Glucagon-like Peptide 1 in Vitro and in Vivo by Dipeptidyl Peptidase IV. *Endocrinology* **1995**, *136* (8), 3585–3596. https://doi.org/10.1210/endo.136.8.7628397.

(296) Vilsbøll, T.; Agersø, H.; Krarup, T.; Holst, J. J. Similar Elimination Rates of Glucagon-Like Peptide-1 in Obese Type 2 Diabetic Patients and Healthy Subjects. *J. Clin. Endocrinol. Metab.* **2003**, *88* (1), 220–224. https://doi.org/10.1210/jc.2002-021053.

(297) Walensky, L. D.; Bird, G. H. Hydrocarbon-Stapled Peptides: Principles, Practice, and Progress. *J. Med. Chem.* **2014**, *57* (15), 6275–6288. https://doi.org/10.1021/jm4011675.

(298) Gentilucci, L.; De Marco, R.; Cerisoli, L. Chemical Modifications Designed to Improve Peptide Stability: Incorporation of Non-Natural Amino Acids, Pseudo-Peptide Bonds, and Cyclization. *Curr. Pharm. Des.* **2010**, *16* (28), 3185–3203. https://doi.org/10.2174/138161210793292555.

(299) Amso, Z.; Kowalczyk, R.; Watson, M.; Park, Y.-E.; Callon, K. E.; Musson, D. S.; Cornish, J.; Brimble, M. A. Structure Activity Relationship Study on the Peptide Hormone Preptin, a Novel Bone-Anabolic Agent for the Treatment of Osteoporosis. *Org. Biomol. Chem.* **2016**, *14* (39), 9225–9238. https://doi.org/10.1039/C6OB01455K.

(300) Sleep, D.; Cameron, J.; Evans, L. R. Albumin as a Versatile Platform for Drug Half-Life Extension. *Biochim. Biophys. Acta - Gen. Subj.* **2013**, *1830* (12), 5526–5534. https://doi.org/10.1016/j.bbagen.2013.04.023.

(301) Strohl, W. R. Fusion Proteins for Half-Life Extension of Biologics as a Strategy to Make Biobetters. *BioDrugs* **2015**, *29* (4), 215–239. https://doi.org/10.1007/s40259-015-0133-6.

(302) Pasut, G.; Veronese, F. M. PEG Conjugates in Clinical Development or Use as Anticancer Agents: An Overview. *Adv. Drug Deliv. Rev.* **2009**, *61* (13), 1177–1188. https://doi.org/10.1016/j.addr.2009.02.010.

(303) Bech, E. M.; Pedersen, S. L.; Jensen, K. J. Chemical Strategies for Half-Life Extension of Biopharmaceuticals: Lipidation and Its Alternatives. *ACS Med. Chem. Lett.* **2018**, *9* (7), 577–580. https://doi.org/10.1021/acsmedchemlett.8b00226.

(304) Poulsen, C.; Pedersen, M. Ø.; Wahlund, P. O.; Sjölander, A.; Thomsen, J. K.; Conde-Frieboes, K. W.; Paulsson, J. F.; Wulff, B. S.; Østergaard, S. Rational Development of Stable PYY3–36 Peptide Y2 Receptor Agonists. *Pharm. Res.* **2021**, *38* (8), 1369–1385. https://doi.org/10.1007/s11095-021-03077-x.

(305) Knudsen, L. B. Liraglutide: The Therapeutic Promise from Animal Models. *Int. J. Clin. Pract.* **2010**, *64* (SUPPL. 167), 4–11. https://doi.org/10.1111/j.1742-1241.2010.02499.x.

(306) Home, P.; Kurtzhals, P. Insulin Detemir: From Concept to Clinical Experience. *Expert Opin. Pharmacother.* **2006**, *7* (3), 325–343. https://doi.org/10.1517/14656566.7.3.325.

(307) Helleberg, H.; Bjelke, M.; Damholt, B. B.; Pedersen, P. J.; Rasmussen, M. H. Absorption, Metabolism and Excretion of Once-Weekly Somapacitan, a Long-Acting Growth Hormone Derivative, after Single Subcutaneous Dosing in Human Subjects. *Eur. J. Pharm. Sci.* **2021**, *167*, 106030. https://doi.org/10.1016/j.ejps.2021.106030.

(308) Kowalczyk, R.; Harris, P. W. R.; Williams, G. M.; Yang, S. H.; Brimble, M. A. *Peptide Lipidation - A Synthetic Strategy to Afford Peptide Based Therapeutics*; 2017; Vol. 1030. https://doi.org/10.1007/978-3-319-66095-0_9.

(309) Susi, H.; Michael Byler, D. Protein Structure by Fourier Transform Infrared Spectroscopy: Second Derivative Spectra. *Biochem. Biophys. Res. Commun.* **1983**, *115* (1), 391–397. https://doi.org/10.1016/0006-291X(83)91016-1.

(310) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta - Bioenerg.* **2007**, *1767* (9), 1073–1101. https://doi.org/10.1016/j.bbabio.2007.06.004.

(311) *Protein Folding, Misfolding, and Disease*; Hill, A. F., Barnham, K. J., Bottomley, S. P., Cappai, R., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2011; Vol. 752. https://doi.org/10.1007/978-1-60327-223-0.

(312) Nakajima, A. Intensity Enhancement Induced by Solute-Solvent Interaction between Pyrene and Polar Solvents. *Spectrochim. Acta Part A Mol. Spectrosc.* **1982**, *38* (6), 693–695. https://doi.org/10.1016/0584-8539(82)80093-7.

(313) Kalyanasundaram, K.; Thomas, J. K. Environmental Effects on Vibronic Band Intensities in Pyrene Monomer Fluorescence and Their Application in Studies of Micellar Systems. *J. Am. Chem. Soc.* **1977**, *99* (7), 2039–2044. https://doi.org/10.1021/ja00449a004.

(314) Wang, Y.; Lomakin, A.; Kanai, S.; Alex, R.; Benedek, G. B. Transformation of Oligomers of Lipidated Peptide Induced by Change in PH. *Mol. Pharm.* **2015**, *12* (2), 411–419. https://doi.org/10.1021/mp500519s.

(315) Kirkham, S.; Castelletto, V.; Hamley, I. W.; Inoue, K.; Rambo, R.; Reza, M.; Ruokolainen, J. Self-Assembly of the Cyclic Lipopeptide Daptomycin: Spherical Micelle Formation Does Not Depend on the Presence of Calcium Chloride. *ChemPhysChem* **2016**, 2118–2122. https://doi.org/10.1002/cphc.201600308.

(316) Hutchinson, J. A.; Hamley, I. W.; Torras, J.; Alemán, C.; Seitsonen, J.; Ruokolainen, J. Self-Assembly of Lipopeptides Containing Short Peptide Fragments Derived from the Gastrointestinal Hormone PYY 3-36 : From Micelles to Amyloid Fibrils. *J. Phys. Chem. B* **2019**, *123* (3), 614–621. https://doi.org/10.1021/acs.jpcb.8b11097.

(317) Procházka, K.; Bednář, B.; Tuzar, Z.; KoČiřík, M. Size Exclusion Chromatography of Associating Systems. I. The Theoretical Model. *J. Liq. Chromatogr.* **1988**, *11* (11), 2221–2239. https://doi.org/10.1080/01483918808067195.

(318) Cherney, L. T.; Krylov, S. N. Slow-Equilibration Approximation in Kinetic Size Exclusion Chromatography. *Anal. Chem.* **2016**, *88* (7), 4063–4070. https://doi.org/10.1021/acs.analchem.6b00415.

(319) Muneeruddin, K.; Thomas, J. J.; Salinas, P. A.; Kaltashov, I. A. Characterization of Small Protein Aggregates and Oligomers Using Size Exclusion Chromatography with Online Detection by Native Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2014**, *86* (21), 10692–10699. https://doi.org/10.1021/ac502590h.

(320) Lin, J. S.; Bekale, L. A.; Molchanova, N.; Nielsen, J. E.; Wright, M.; Bacacao, B.; Diamond, G.; Jenssen, H.; Santa Maria, P. L.; Barron, A. E. Anti-Persister and Anti-Biofilm Activity of Self-Assembled Antimicrobial Peptoid Ellipsoidal Micelles. *ACS Infect. Dis.* **2022**. https://doi.org/10.1021/acsinfecdis.2c00288.

(321) Novo, M.; Freire, S.; Al-Soufi, W. Critical Aggregation Concentration for the Formation of Early Amyloid-β (1–42) Oligomers. *Sci. Rep.* **2018**, *8* (1), 1783. https://doi.org/10.1038/s41598-018-19961-3.

(322) Venanzi, M.; Savioli, M.; Cimino, R.; Gatto, E.; Palleschi, A.; Ripani, G.; Cicero, D.; Placidi, E.; Orvieto, F.; Bianchi, E. A Spectroscopic and Molecular Dynamics Study on the Aggregation Process of a Long-Acting Lipidated Therapeutic Peptide: The Case of Semaglutide. *Soft Matter* **2020**, *16* (44), 10122–10131. https://doi.org/10.1039/d0sm01011a.

(323) Zaman, M.; Khan, A. N.; Wahiduzzaman; Zakariya, S. M.; Khan, R. H. Protein Misfolding, Aggregation and Mechanism of Amyloid Cytotoxicity: An Overview and Therapeutic Strategies to Inhibit Aggregation. *Int. J. Biol. Macromol.* **2019**, *134*, 1022–1037. https://doi.org/10.1016/j.ijbiomac.2019.05.109.

(324) Kamada, A.; Rodriguez-Garcia, M.; Ruggeri, F. S.; Shen, Y.; Levin, A.; Knowles, T. P. J. Controlled Self-Assembly of Plant Proteins into High-Performance Multifunctional Nanostructured Films. *Nat. Commun.* **2021**, *12* (1), 3529. https://doi.org/10.1038/s41467-021-23813-6.

(325) Hartgerink, J. D.; Beniash, E.; Stupp, S. I. Peptide-Amphiphile Nanofibers: A Versatile Scaffold for the Preparation of Self-Assembling Materials. *Proc. Natl. Acad. Sci.* **2002**, *99* (8), 5133–5138. https://doi.org/10.1073/pnas.072699999.

(326) Ouberai, M. M.; Dos Santos, A. L. G.; Kinna, S.; Madalli, S.; Hornigold, D. C.; Baker, D.; Naylor, J.; Sheldrake, L.; Corkill, D. J.; Hood, J.; et al. Controlling the Bioactivity of a Peptide Hormone in Vivo by Reversible Self-Assembly. *Nat. Commun.* **2017**, *8* (1), 1026. https://doi.org/10.1038/s41467-017-01114-1.

(327) Zhou, N.; Gao, X.; Lv, Y.; Cheng, J.; Zhou, W.; Liu, K. Self-Assembled Nanostructures of Long-Acting GnRH Analogs Modified at Position 7. *J. Pept. Sci.* **2014**, *20* (11), 868–875. https://doi.org/10.1002/psc.2678.

(328) Kunitake, T.; Nakashima, N.; Morimitsu, K. Enhanced Circular Dichroism and Fluidity of Disk-like Aggregates of a Chiral, Single-Chain Amphiphile. *Chem. Lett.* **1980**, *9* (11), 1347–1350. https://doi.org/10.1246/cl.1980.1347.

(329) Shimizu, T.; Mori, M.; Minamikawa, H.; Hato, M. Enhanced Circular Dichroism of Self-Assembled Peptidic Amphiphiles. *Chem. Lett.* **1989**, *18* (8), 1341–1344. https://doi.org/10.1246/cl.1989.1341.

(330) Liu, J.; Su, H.; Meng, L.; Zhao, Y.; Deng, C.; Ng, J. C. Y.; Lu, P.; Faisal, M.; Lam, J. W. Y.; Huang, X.; et al. What Makes Efficient Circularly Polarised Luminescence in the Condensed Phase: Aggregation-Induced Circular Dichroism and Light Emission. *Chem. Sci.* **2012**, *3* (9), 2737–2747. https://doi.org/10.1039/c2sc20382k.

(331) Zhang, L.; Liang, K.; Dong, L.; Yang, P.; Li, Y.; Feng, X.; Zhi, J.; Shi, J.; Tong, B.; Dong, Y. Aggregation-Induced Emission Enhancement and Aggregation-Induced Circular Dichroism of Chiral Pentaphenylpyrrole Derivatives and Their Helical Self-Assembly. *New J. Chem.* **2017**, *41* (17), 8877–8884. https://doi.org/10.1039/c7nj00583k.

(332) Li, Z.; Hirst, J. D. Quantitative First Principles Calculations of Protein Circular Dichroism in the Near-Ultraviolet. *Chem. Sci.* **2017**, *8* (6), 4318–4333. https://doi.org/10.1039/C7SC00586E.

(333) Pescitelli, G.; Di Bari, L.; Berova, N. Application of Electronic Circular Dichroism in the Study of Supramolecular Systems. *Chem. Soc. Rev.* **2014**, *43* (15), 5211–5233. https://doi.org/10.1039/C4CS00104D.

(334) Gottarelli, G.; Lena, S.; Masiero, S.; Pieraccini, S.; Spada, G. P. The Use of Circular Dichroism Spectroscopy for Studying the Chiral Molecular Self-Assembly: An Overview. *Chirality* **2008**, *20* (3–4), 471–485. https://doi.org/10.1002/chir.20459.

(335) Arosio, P.; Knowles, T. P. J.; Linse, S. On the Lag Phase in Amyloid Fibril Formation. *Phys. Chem. Chem. Phys.* **2015**, *17* (12), 7606–7618. https://doi.org/10.1039/C4CP05563B.

(336) Younan, N. D.; Viles, J. H. A Comparison of Three Fluorophores for the Detection of Amyloid Fibers and Prefibrillar Oligomeric Assemblies. ThT (Thioflavin T); ANS (1-Anilinonaphthalene-8-Sulfonic Acid); and BisANS (4,4′-Dianilino-1,1′-Binaphthyl-5,5′-Disulfonic Acid). *Biochemistry* **2015**, *54* (28), 4297–4306. https://doi.org/10.1021/acs.biochem.5b00309.

(337) Bolognesi, B.; Kumita, J. R.; Barros, T. P.; Esbjorner, E. K.; Luheshi, L. M.; Crowther, D. C.; Wilson, M. R.; Dobson, C. M.; Favrin, G.; Yerbury, J. J. ANS Binding Reveals Common Features of Cytotoxic Amyloid Species. *ACS Chem. Biol.* **2010**, *5* (8), 735–740. https://doi.org/10.1021/cb1001203.

(338) Foderà, V.; Groenning, M.; Vetri, V.; Librizzi, F.; Spagnolo, S.; Cornett, C.; Olsen, L.; van de Weert, M.; Leone, M. Thioflavin T Hydroxylation at Basic PH and Its Effect on Amyloid Fibril Detection. *J. Phys. Chem. B* **2008**, *112* (47), 15174–15181. https://doi.org/10.1021/jp805560c.

(339) Edu, I.-A. Studies on the Physical Stability and Aggregation of GLP-1 Analogues, University of Cambridge, 2022.

(340) Michaels, T. C. T.; Šarić, A.; Habchi, J.; Chia, S.; Meisl, G.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. Chemical Kinetics for Bridging Molecular Mechanisms and Macroscopic Measurements of Amyloid Fibril Formation. *Annu. Rev. Phys. Chem.* **2018**, *69* (1), 273–298. https://doi.org/10.1146/annurev-physchem-050317-021322.

(341) Michaels, T. C. T.; Šarić, A.; Meisl, G.; Heller, G. T.; Curk, S.; Arosio, P.; Linse, S.; Dobson, C. M.; Vendruscolo, M.; Knowles, T. P. J. Thermodynamic and Kinetic Design Principles for Amyloid-Aggregation Inhibitors. *Proc. Natl. Acad. Sci.* **2020**, *117* (39), 24251–24257. https://doi.org/10.1073/pnas.2006684117.

(342) Tantakitti, F.; Boekhoven, J.; Wang, X.; Kazantsev, R. V.; Yu, T.; Li, J.; Zhuang, E.; Zandi, R.; Ortony, J. H.; Newcomb, C. J.; et al. Energy Landscapes and Functions of Supramolecular Systems. *Nat. Mater.* **2016**, *15* (4), 469–476. https://doi.org/10.1038/nmat4538.

(343) Hamley, I. W.; Castelletto, V. Self-Assembly of Peptide Bioconjugates: Selected Recent Research Highlights. *Bioconjug. Chem.* **2017**, *28* (3), 731–739. https://doi.org/10.1021/acs.bioconjchem.6b00284.

(344) Ghosh, A.; Buettner, C. J.; Manos, A. A.; Wallace, A. J.; Tweedle, M. F.; Goldberger, J. E. Probing Peptide Amphiphile Self-Assembly in Blood Serum. *Biomacromolecules* **2014**, *15* (12), 4488–4494. https://doi.org/10.1021/bm501311g.

(345) Lock, L. L.; Reyes, C. D.; Zhang, P.; Cui, H. Tuning Cellular Uptake of Molecular Probes by Rational Design of Their Assembly into Supramolecular Nanoprobes. *J. Am. Chem. Soc.* **2016**, *138* (10), 3533–3540. https://doi.org/10.1021/jacs.6b00073.

(346) Staby, A.; Steensgaard, D. B.; Haselmann, K. F.; Marino, J. S.; Bartholdy, C.; Videbæk, N.; Schelde, O.; Bosch-Traberg, H.; Spang, L. T.; Asgreen, D. J. Influence of Production Process and Scale on Quality of Polypeptide Drugs: A Case Study on GLP-1 Analogs. *Pharm. Res.* **2020**, *37* (7). https://doi.org/10.1007/s11095-020-02817-9.

(347) Gallo, M.; Vanni, D.; Esposito, S.; Alaimo, N.; Orvieto, F.; Rulli, F.; Missineo, A.; Caretti, F.; Bonelli, F.; Veneziano, M.; et al. Oligomerization, Albumin Binding and Catabolism of Therapeutic Peptides in the Subcutaneous Compartment: An Investigation on Lipidated GLP-1 Analogs. *J. Pharm. Biomed. Anal.* **2022**, *210*, 114566. https://doi.org/10.1016/j.jpba.2021.114566.

(348) Hamley, I. W. Self-Assembly of Amphiphilic Peptides. *Soft Matter* **2011**, *7* (9), 4122. https://doi.org/10.1039/c0sm01218a.

(349) Bleiholder, C.; Bowers, M. T. The Solution Assembly of Biological Molecules Using Ion Mobility Methods: From Amino Acids to Amyloid β-Protein. *Annu. Rev. Anal. Chem.* **2017**, *10* (1), 365–386. https://doi.org/10.1146/annurev-anchem-071114-040304.

(350) Miller, S. A.; Jeanne Dit Fouque, K.; Ridgeway, M. E.; Park, M. A.; Fernandez-Lima, F. Trapped Ion Mobility Spectrometry, Ultraviolet Photodissociation, and Time-of-Flight Mass Spectrometry for Gas-Phase Peptide Isobars/Isomers/Conformers Discrimination. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (7), 1267–1275. https://doi.org/10.1021/jasms.2c00091.

(351) Depraz Depland, A.; Stroganova, I.; Wootton, C. A.; Rijs, A. M. Developments in Trapped Ion Mobility Mass Spectrometry to Probe the Early Stages of Peptide Aggregation. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (2), 193–204. https://doi.org/10.1021/jasms.2c00253.

(352) Byrd, E. J.; Wilkinson, M.; Radford, S. E.; Sobott, F. Taking Charge: Metal Ions Accelerate Amyloid Aggregation in Sequence Variants of α-Synuclein. *J. Am. Soc. Mass Spectrom.* **2023**, *34* (3), 493–504. https://doi.org/10.1021/jasms.2c00379.

(353) Zheng, Y.; Cong, Y.; Schmidt, E. W.; Nair, S. K. Catalysts for the Enzymatic Lipidation of Peptides. *Acc. Chem. Res.* **2022**, *55* (9), 1313–1323. https://doi.org/10.1021/acs.accounts.2c00108.

(354) Castelletto, V.; Hamley, I. W.; Seitsonen, J.; Ruokolainen, J.; Harris, G.; Bellmann-Sickert, K.; Beck-Sickinger, A. G. Conformation and Aggregation of Selectively PEGylated and Lipidated Gastric Peptide Hormone Human PYY 3–36. *Biomacromolecules* **2018**, *19* (11), 4320–4332. https://doi.org/10.1021/acs.biomac.8b01209.

(355) Wolff, M.; Schüler, A.; Gast, K.; Seckler, R.; Evers, A.; Pfeiffer-Marek, S.; Kurz, M.; Nagel, N.; Haack, T.; Wagner, M.; et al. Self-Assembly of Exendin-4-Derived Dual Peptide Agonists Is Mediated by Acylation and Correlated to the Length of Conjugated Fatty Acyl Chains. *Mol. Pharm.* **2020**, *17* (3), 965–978. https://doi.org/10.1021/acs.molpharmaceut.9b01195.

(356) Zhang, G.; Li, J.; Gao, L.; Wang, T.; Quan, D. Morphology of Nanostructures and Their Long-Acting Properties in Vivo for a Novel Synthetic Peptide of Gonadotropin-Releasing Hormone Antagonist. *J. Pharm. Pharmacol.* **2014**, *66* (8), 1077–1081. https://doi.org/10.1111/jphp.12233.

(357) Thompson, M. J.; Sievers, S. A.; Karanicolas, J.; Ivanova, M. I.; Baker, D.; Eisenberg, D. The 3D Profile Method for Identifying Fibril-Forming Segments of Proteins. *Proc. Natl. Acad. Sci.* **2006**, *103* (11), 4074–4078. https://doi.org/10.1073/pnas.0511295103.

(358) Zilberzwige-Tal, S.; Gazit, E. Go with the Flow-Microfluidics Approaches for Amyloid Research. *Chem. - An Asian J.* **2018**, *13* (22), 3437–3447. https://doi.org/10.1002/asia.201801007.

(359) Arosio, P.; Müller, T.; Rajah, L.; Yates, E. V.; Aprile, F. A.; Zhang, Y.; Cohen, S. I. A.; White, D. A.; Herling, T. W.; De Genst, E. J.; et al. Microfluidic Diffusion Analysis of the Sizes and Interactions of Proteins under Native Solution Conditions. *ACS Nano* **2016**, *10* (1), 333–341. https://doi.org/10.1021/acsnano.5b04713.

(360) Drucker, D. J. Advances in Oral Peptide Therapeutics. *Nat. Rev. Drug Discov.* **2020**, *19* (4), 277–289. https://doi.org/10.1038/s41573-019-0053-0.

(361) Haddadzadegan, S.; Dorkoosh, F.; Bernkop-Schnürch, A. Oral Delivery of Therapeutic Peptides and Proteins: Technology Landscape of Lipid-Based Nanocarriers. *Adv. Drug Deliv. Rev.* **2022**, *182*, 114097. https://doi.org/10.1016/j.addr.2021.114097.