Synthetic modification of fusogenic coiled coil peptides
Crone, N.S.A.

Citation

Version: Publisher's Version
License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from: https://hdl.handle.net/1887/3247291

Note: To cite this publication please use the final published version (if applicable).
General introduction
1.1 Peptide and protein folding

Peptides and proteins are the most abundant macromolecules encountered in living systems, and play a role in all vital cellular processes. They are constructed from chains of amino acids, but differ in the length of the chain, with the term ‘peptide’ referring to chains of less than 50 residues and anything larger referred to as proteins. In order to reduce steric interference between atoms, maximise backbone hydrogen bonding interactions and facilitate favourable side-chain interactions, peptides and proteins can adopt complex three-dimensional conformations. Biological activity is often directly related to the adopted structure, with a large scientific effort dedicated to studying, predicting and controlling protein folding.

Analysis of known protein structures have shown two dominant secondary structures; the α-helix and the β-strand, with proteins divided into classes based on the abundance of these structural elements. The two structures can be defined by the different angles describing the three-dimensional orientation of the amino acid around its α-carbon (Cα, Figure 1.1A). Hydrogen bonding within the same chain is maximized for α-helices, with every amino acid forming a hydrogen bond to the amino acid 4 positions further in the sequence (i to i+4 hydrogen bonding). The peptide chain rotates around a central axis with 3.6 residues per helix turn, resulting in a 7-amino acid sequence (called a ‘heptad repeat’) describing two full rotations. In contrast, β-structures form hydrogen bonds which can be between strands of the same, or different protein chains. These intermolecular hydrogen bonds results in multiple β-strands combining to form a β-sheet. To allow hydrogen bonding in the same sequence, these β-strand domains are connected via a short turn, hairpin or other structural motif; or can bridge to separate β-domains in the same sequence spaced further apart. Protein self-assembly can also be driven by the formation of intermolecular β-sheets between β-structured domains in different protein chains.

The tendency of an amino acid sequence to fold as one of these two conformations depends on the order of polar (p) and hydrophobic/non-polar (n) amino acids in the sequence, in combination with the tendency of individual amino acids to adopt either structure. When polar and non-polar amino acids are directly alternated,
Figure 1.1: Schematic overview of peptide structure and folding. Amino acid chain showing the $\phi$ and $\psi$ angles used to define geometry around the amino acid $\alpha$-carbon (A). Arrangement of polar and non-polar amino acids in $\beta$-structured (left) and $\alpha$-helical (right) peptide sequences (B). White circles indicate polar side chains, black circles represent nonpolar side chains and grey circles represent amino acids that can be either polar or non-polar (X in the sequence layout). Note that both structures result in polar and non-polar amino acids placed adjacent to each other. Image adapted from Kohn and Hodghes.}

the chains tend to fold as $\beta$-structured domains, whereas $\alpha$-helices tend to follow a heptad repeat, with alternating non-polar residues every 3 and 4 amino acids (Figure 1.1B). These general patterns result in secondary structures where amino acid side chains with comparable interactions with their environment are placed adjacent. Reducing interaction between aliphatic side chains and the aqueous environment especially is one of the main entropic driving forces behind protein folding and self-assembly.

**Coiled coils**

The heptad repeat sequence of $\alpha$-helical peptides results in non-polar amino acids being adjacent on one side of the helix. This results in a ‘stripe’ of non-polar amino acids that runs along the length of the helix. If this effect is strongly pronounced, with one side of the helix strongly hydrophobic and the other strongly polar, it is called an amphipathic helix. This type of helix tends to interact with other hydrophobic peptide domains, lipid bilayers, or hydrophobic small-molecules in order to reduce the aqueous interactions of the aliphatic side chains. One method to achieve this is the self-assembly of two or more helices, with the hydrophobic amino acids in the centre, forming a so-called ‘hydrophobic core’ in between the helices. Because the $\alpha$-helix completes a full turn every 3.6 residues, bundles of
α-helices rotate around each other in a ‘supercoil’ structure, which reduces the effective residues per turn and allows for continuous contact of the hydrophobic amino acids. This self-assembled structure of multiple α-helices is referred to as a coiled coil,\(^{11}\) and is an abundant structural motif both within proteins and a driving force for protein oligomerization.\(^{12}\)

In coiled-coil structures, amino acid side chains in one helix occupy empty spaces between side chains of adjacent helices, a regime which is referred to as ‘knobs-in-holes packing’.\(^{13}\) Non-polar amino acids at positions \(a\) and \(d\) of the heptad repeat sequence, \(abcd\cdot efg\), form the hydrophobic interface (Figure 1.2A), and packing of these side chains affects coiled-coil stability, with the size and branching also affecting the oligomer state of the coiled coil.\(^{14}\) Switching amino acids at the \(a\) and \(d\) positions additionally determines if the coiled coil assembles in a parallel (N-termini on the same side) or anti-parallel (N-termini on opposing ends) orientation, as shown in Figure 1.2B. The positions \(e\) and \(g\), flanking the hydrophobic amino acids, tend to be charged amino acids that form electrostatic interactions between different helical strands, and thereby provide pairing specificity. Finally, the other three amino acid positions are typically hydrophilic as they are solvent exposed.\(^{15}\) These rules can be used to predict coiled-coil formation both in synthetic and

**Figure 1.2:** Helical wheel diagram of parallel dimeric (top left), anti-parallel dimeric (top right), parallel trimeric (bottom left) and parallel tetrameric (bottom right) coiled-coil peptides (A). Straight arrows indicate hydrophobic interactions, where curved arrows indicate electrostatic side-chain interactions. Packing of hydrophobic side chains according to the ‘knobs-in-holes’ configuration for parallel (top) and anti-parallel (bottom) coiled coils with up to 4 chains (B). Images adapted from Apostolovic et al.\(^{10}\)
natural proteins,\textsuperscript{16-18} although for pentameric and larger coiled-coils the preferred repeat sequence starts to differ.\textsuperscript{19} Additionally, more variation in the sequence can be allowed for longer coiled-coil sequences, which can be used for increasing binding selectivity, introducing functional groups or introduction of an active component. A typical example of this are the asparagine residues found at position ‘a’ of GCN4, which promote formation of dimeric over a tetrameric coiled-coil and prevents anti-parallel or out-of-register self-assembly.\textsuperscript{20}

In the last decades, coiled coils have been investigated as structural motifs in synthetic nanoparticles,\textsuperscript{21} therapeutics such as vaccines,\textsuperscript{22} hydrogels,\textsuperscript{23} cellular delivery systems,\textsuperscript{24} and synthetic biology in general.\textsuperscript{25} The high binding strength, well defined structure and self-assembly, and ease of preparation via solid-phase peptide synthesis (SPPS) provide several advantages and flexibility in their application.\textsuperscript{26} The use of coiled coil lipopeptides in synthetic membrane fusion has shown to be particularly effective,\textsuperscript{27} using the energy associated with coiled-coil formation to drive the fusion process, and is the main focus of the work presented in this thesis.

1.2 Membrane fusion and SNARE proteins

Membrane fusion describes the process in which two separate (lipid) membranes combine into one continuous membrane. Fusion of membranes results in mixing of encapsulated content, when discussing fusion between vesicles, or release of content from a vesicle across a membrane. Fusion of biological membranes is a fundamental process for all living systems, ranging from continuous processes such as nutrient uptake and digestion by lysosomes to cell fusion during reproduction and the viral infection of host cells.\textsuperscript{28} Due to the importance of these processes, an understanding of the membrane fusion mechanism is of direct scientific interest. For example, understanding protein-membrane interactions should allow for the development of novel therapeutics against viral infections from HIV and novel variants of SARS-CoV-2.\textsuperscript{29, 30} Besides control over natural fusion processes, development of artificial fusion systems has shown potential in the fields of drug delivery, sensing, and synthetic biology.\textsuperscript{31}

Combining two separate lipid bilayers into a single continuous membrane requires rearrangement of the bilayers through (non-bilayer) metastable intermediates. These intermediate states are characterized by high curvature, are energetically
unfavourable, and therefore do not occur spontaneously. In Nature, membrane fusion proteins are thus employed by living systems to drive this process by lowering the energy barrier of the intermediates, and to control the instance and location where membrane fusion occurs.\(^{32-34}\) The basic steps of bilayer fusion are shown in Figure 1.3 and start with two membranes that are in close proximity, followed by closer contact called ‘docking’ and initiation of fusion. Rearrangement of the lipids leads to a hemifusion diaphragm, where both a double and a single bilayer are present, followed by the opening of a fusion pore and finally expansion of the pore until the membranes are fully fused. The hemifusion intermediate is deemed essential to all fusion processes, and has been shown to exist as a temporally stable state for multiple fusion systems.\(^{35, 36}\) Initiation of fusion has long been theorized to occur through a fusion stalk intermediate,\(^{34}\) with discussion as to the exact mechanism still ongoing.\(^{37-39}\)

One of the best understood biological fusion processes is that involved in the release of neurotransmitters for neuronal signal transduction. Fusion of neuronal vesicles with the axonal membrane is stimulated by Ca\(^{2+}\) influx in the cell, and
results in fast release of neurotransmitters in the synaptic cleft. The process is driven by the assembly of SNARE (soluble NSF attachment protein receptor) proteins on both membranes. A functional SNARE complex is formed by the assembly of a four-helix coiled-coil complex between the two lipid membranes, and is proposed as the energetic driving force behind fusion.\textsuperscript{40-42} The SNARE proteins are separated into v-SNARE or t-SNARE groups, depending on their occurrence on vesicle or target membranes. For neuronal vesicle fusion, synaptobrevin functions as the v-SNARE, whilst the other three helices of the complex originate from Syntaxin 1 and two copies of SNAP25 as t-SNARE proteins. The combination of different SNARE proteins in the formation of a SNARE complex is part of the mechanism that allows for selectivity of fusion of different cellular membranes.\textsuperscript{43} A long-standing question has been how many SNARE complexes are required to successfully fuse two membranes, with some evidence showing that a single SNARE complex might even be sufficient.\textsuperscript{44} However, more experimental work points towards multiple SNARE complexes in a single fusion interface,\textsuperscript{45} with theoretical studies showing the speed of fusion is entropically dependent on the number of complexes.\textsuperscript{46}

Although SNARE proteins provide the driving force behind fusion, a large number of other proteins are involved to regulate where and when fusion occurs. In the example of vesicle fusion in neuronal synapses, synaptotagmin-1 binds to the SNARE complex before fusion and couples the influx of Ca\textsuperscript{2+} to activation of the fusion process. Other regulators of SNARE-mediated fusion are members of the SM protein family (Sec1/Munc18) working together with complexins and Munc13s, Rab3, and HOPS proteins. The complex interplay between all these fusion regulators is still under investigation and several reviews discussing this matter have been published.\textsuperscript{47-51} Although these proteins are called regulators of the fusion process, many have been hypothesized or shown to have an active contribution to SNARE-mediated fusion.\textsuperscript{47} Active recycling of the SNARE complex is performed by the NSF protein, which binds to the fused SNARE complex in combination with multiple SNAPs and actively disassembles the proteins driven by ATP hyrolysis.\textsuperscript{49,52} After the complex has been deconstructed, the v-SNARE is recycled while the three t-SNARE proteins remain on the target membrane until the process is restarted. Together, these regulatory systems are highly efficient in their control over SNARE-mediated fusion. They facilitate the extremely fast release of neurotransmitters (signal transfer along nerve fibres can reach speeds of up to 120 m/s)\textsuperscript{53}, but prevent
the energetically costly release of neurotransmitters when no trigger signal is present.

1.3 Synthetic mimics of biological membrane fusion

Study of SNARE-mediated fusion and the complex systems in place to regulate fusion activity have progressed in the last decades, but remain difficult because of the number of proteins involved and the multiple roles that they can perform. Furthermore, reconstitution of the fusion machinery in model systems may lead to different results than observed in the cellular process. Because of these reasons, and the potential applications of fusion systems discussed earlier, synthetic model systems have been prepared as mimics of SNARE-mediated fusion. The ultimate goal in developing such model systems is the same fusion efficiency and selectivity as observed in the native system, with low leakage of contents, which indicates membrane instability rather than fusion. Minimalism in design allows for deconvolution of the contributing factors to the fusion process, but a multi-component strategy is essential to obtain the required selectivity, with a minimum of one component in each membrane. Synthetic fusion systems exist that do rely on the action of a single entity, such as the pore-forming GALA peptides, but these result in the random mixing of vesicles instead of controlled mixing of different entities. Liposomal fusion model systems can in principle be based on any strong and selective interaction, and have been demonstrated using complementary strands of single strand DNA or PNA, the hydrogen bonding interaction between cyanuric acid and melamime, the reversible covalent linkages between boronic acids and cis-diols, and binding of vancomycin to the D-Ala-D-Ala peptide sequence.

Coiled coil peptides as fusion mimics

The most investigated model system for membrane fusion has been developed in the Kros group and uses coiled-coil peptides (Figure 1.4). It consists of two peptides named ‘E’; (EIAALEK)_n, and ‘K’; (KIAALKE)_n, after the dominant charged amino acid in their respective sequences. These peptides are designed to form a tight-binding heterodimeric coiled-coil, as a synthetic alternative to the tetrameric coiled-coil formed by SNARE proteins. A lipid anchor connected to the peptide through a polyethylene glycol (PEG) spacer allows for their integration into lipid membranes. Incorporating these peptides in separate lipid membranes results in their fusion upon mixing. Because of the simplicity of the system, it allows
Figure 1.4: Schematic representation of membrane fusion mediated by coiled-coil peptides, as mimics of the biological fusion process. Comparison between the SNARE fusion complex (left) and the coiled-coil fusion system (right) (A), overview of the structure of the coiled coil lipopeptides consisting of peptide E or K connected to a lipid via a PEG spacer, and how they are incorporated in lipid vesicles (B), and how this fusion process can be controlled via the incorporation of a photolabile PEG lipid which shields the peptides from interaction (C). Images are adapted from Marsden et al. and Kong et al.

synthetic derivatization to determine how the structure is related to the fusion activity.

The E/K fusion system initially used phospholipids as the membrane anchor, but a study comparing different lipids revealed cholesterol as the optimal lipid anchor, possibly because it is easier for cholesterol-peptides to disperse through the lipid membrane. Increasing the coiled-coil length, which is directly related to the binding strength, results in increased fusion with the direction of the coiled coil (N-terminal versus C-terminal anchoring) not having a significant effect. Changing the length of the PEG spacer also altered fusion efficiency, with an optimum of around 12 polyethylene glycol repeats. Biophysical investigations of the coiled-coil fusion system have mainly focussed on the disparate roles of the two peptides during the fusion process. Peptide K was shown to interact with lipid membranes, a phenomenon which originates from its structure as a positively charged amphipathic helix. This membrane interaction was theorized as one of the driving forces behind the system's fusogenicity. Further investigation showed that peptide K, both by itself and when connected to the lipid membrane, can result in altered membrane curvature, viscosity and hydration. These interactions might lower the energy required for reorganization of lipid membranes in the fusion intermediates described earlier, and thereby result in higher fusion efficiency.
Because the E/K coiled-coil is designed *de novo*, its’ activity is orthogonal to normal biological processes, and its’ liposomal formulation is therefore of interest as a drug delivery system. Cellular studies using the E/K fusion system have shown successful delivery of liposomal doxorubicin,\textsuperscript{76} and showed promising results in zebrafish xenografts.\textsuperscript{77} In other studies using giant unilamellar vesicles (GUVs) as a model system, liposomal fusion was used to effectively introduce the hydrophobic bis-(thiourea)decalin anion transporter.\textsuperscript{78} Besides liposomes, the coiled-coil system could be used for delivery of mesoporous silica nanoparticles and nucleic acid lipid particles *in vitro*.\textsuperscript{79, 80} The strong interactions between the E/K peptides have also resulted in other applications such as the immobilization of vesicles on a patterned surface substrate,\textsuperscript{81} studying cellular exchange by forcing close proximity of cellular membranes through coiled-coil modification,\textsuperscript{82} and recently the sorting of cells with magnetic beads bearing a coiled-coil functionalized dextran polymer.\textsuperscript{83} The incorporation of lipids containing a long (M\textsubscript{n} = 2000 g/mol) PEG chain proved to block the peptides from fusing lipid vesicles, by preventing coiled-coil interactions in the vesicle docking phase.\textsuperscript{84} This effect was used to introduce more selectivity in the fusion process, by combining the E/K coiled-coil with a photocleavable PEG lipid to halt any fusion progress until the PEG was cleaved from the vesicle with UV light (Figure 1.4C).\textsuperscript{67} This mechanism allowed for photoactive doxorubicin delivery in a zebrafish cancer model,\textsuperscript{85} showing the potential of combining a fusion system with active components. This dePEGylation strategy could be seen as a mimic of the natural regulatory proteins involved in the control of SNARE protein activity, halting liposomal fusion until a trigger is applied. However, there is no selectivity in the interactions between PEG and the fusogenic coiled-coil peptides, resulting in the requirement for a large excess of PEG to shield the entire liposomal surface. Furthermore, a lack of reversibility in the dePEGylation prevents precise control over coiled-coil interactions, and limit its potential applications. Therefore, development of novel (photo)activation strategies for coiled-coil peptides is still of direct interest.

Together, all of this work shows not only the flexibility of the coiled-coil peptides, but also their compatibility with biological systems. Improving our understanding of the interactions of peptide E and K in membrane fusion directly leads to better drug delivery systems, but alterations to the coiled-coil that allow for increased binding, biocompatibility or introduction of other functionalities are also of general interest outside the field of membrane fusion.
1.4 Aims and scope of this thesis

The work presented in this thesis aims to understand the underlying mechanisms that drive coiled-coil based membrane fusion, using a synthetic investigation strategy to determine relationships between peptide structure, interactions with its binding partner and lipid membranes, and fusion of liposomal membranes. The second part of this thesis aims to introduce a light-active component in the peptide structure to gain spatiotemporal control over peptide folding and activity.

In Chapter 2 of this thesis the effect of lipid anchor attachment location in the E/K coiled-coil fusion mechanism is investigated. The cholesterol-PEG moiety used as the membrane anchor is repositioned from the peptide terminus to the center of the peptide sequence, mimicking side-chain palmitoylation common in the natural SNARE fusion system. The new variants showed lower fusion efficiency compared to the unmodified peptides, and liposomes prepared with the new E variant were shown to be unstable and aggregate, which was attributed to the homomeric interactions of peptide E on opposing membranes. The formation of weak homomeric coiled-coils of peptide E was previously assumed to have no effect on the fusion process, however here we showed how this stabilized the peptide before the docking stage, allowing fusion to occur in a leakage-free manner.

In order to separate the effects of coiled-coil and membrane interactions on fusion efficiency, a peptide stapling strategy is used in Chapter 3 to investigate the interactions of peptide K₃. Cysteines were introduced in the peptide backbone, spaced i to i+4 to span a single helical turn, and crosslinked to generate structural variants of peptide K. These stapled peptides were analysed for changes in folding, coiled-coil binding and membrane affinity, with stapling increasing coiled-coil binding through a pre-organization mechanism. Fusion assays showed a large increase in content mixing directly related to increased coiled-coil binding strength. The results in this chapter disprove the hypothesis that membrane interactions of peptide K are beneficial for fusion, and support the idea that fusion efficiency is directly related to coiled-coil binding strength.

Chapter 4 investigates coiled-coil peptide stapling using azobenzene crosslinkers as light-switchable alternatives to the rigid crosslinkers shown to affect coiled-coil formation in Chapter 3. These azobenzene crosslinkers can switch between two isomers upon light irradiation, and when connected to the peptide at two positions,
this change in morphology allows for photocontrol of peptide folding. A single azobenzene crosslinker was found ineffective for controlling peptide structure in a four-heptad coiled coil, but reducing the length of the coiled coil to three heptads resulted in a larger separation between the two isomers, and shows the potential of this strategy to control coiled-coil self-assembly using light.

Because the preparation of azobenzene-crosslinked peptides in Chapter 4 was synthetically challenging, a different strategy for coiled-coil photocontrol was investigated in Chapter 5, based on the incorporation of amino acids bearing an azobenzene moiety. Two azobenzene derivatives of phenylalanine were prepared, based on previous literature procedures, as well as a novel azobenzene derivative of phenylglycine. Substitution of a single isoleucine in the hydrophobic core of peptide K by these amino acids yielded peptides that could be switched between two states using UV light. The phenylglycine based photoswitch created the largest difference upon isomerization, and the mechanism of action was related to a switch in polarity of the diazene bond upon isomerization. Molecular dynamics (MD) simulations revealed the diazene of the phenylglycine based photoswitch to be positioned in the coiled-coil hydrophobic core, and demonstrated how the methylene group in phenylalanine-based photoswitches increased reorganization after isomerization. Together, these findings demonstrate the incorporation of azobenzene-based amino acids as an effective, novel method for coiled-coil photocontrol.

Disruption of hydrophobic domains, by amino acids containing an azobenzene moiety, was hypothesized as a general mechanism for peptide photocontrol. The viability of this strategy in β-structured peptides is investigated in Chapter 6 of this thesis. A photoswitchable amino acid was introduced in different positions of a peptide known to self-assemble to form β-sheet-based fibres, and the effect of isomerization on structure and self-assembly was studied. Prepared peptides showed β-sheet fibre formation, with different oligomerization and self-assembly kinetics depending on the location of the photoswitch. Peptide folding and the critical aggregation concentration could be altered through light isomerization, demonstrating the application of this amino acid in the photocontrol of β-sheet peptides and confirming our hypothesis. To test if azobenzene isomerization can be used to control peptide activity, histidine residues were introduced and the peptides were evaluated for organocatalytic activity, demonstrating an activity which was dependent on the state of the photoswitch.
1.5 References


Chapter 1


