Self-assembly of designed coiled coil peptides studied by small-angle X-ray scattering and analytical ultracentrifugation

Malik, Leila; Nygaard, Jesper; Christensen, Niels Johan; Streicher, Werner W.; Thulstrup, Peter W.; Arleth, Lise; Jensen, Knud J.

Published in:
Journal of Peptide Science

Link to article, DOI:
10.1002/psc.2497

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Self-assembly of designed coiled coil peptides studied by small-angle X-ray scattering and analytical ultracentrifugation

Leila Malik, a Jesper Nygaard, a,b Niels J. Christensen, a Werner W. Streicher, c Peter W. Thulstrup, a Lise Arleth a and Knud J. Jensen a*

α-Helical coiled coil structures, which are noncovalently associated heptad repeat peptide sequences, are ubiquitous in nature. Similar amphipathic repeat sequences have also been found in helix-containing proteins and have played a central role in de novo design of proteins. In addition, they are promising tools for the construction of nanomaterials. Small-angle X-ray scattering (SAXS) has emerged as a new biophysical technique for elucidation of protein topology. Here, we describe a systematic study of the self-assembly of a small ensemble of coiled coil sequences using SAXS and analytical ultracentrifugation (AUC), which was correlated with molecular dynamics simulations. Our results show that even minor sequence changes have an effect on the folding topology and the self-assembly and that these differences can be observed by a combination of AUC, SAXS, and circular dichroism spectroscopy. A small difference in these methods was observed, as SAXS for one peptide and revealed the presence of a population of longer aggregates, which was not observed by AUC. Copyright © 2013 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: coiled coil; SAXS; AUC; MD

Introduction

The α-helical coiled coil is a ubiquitous protein motif that exists in 5–10% of all protein sequences [1]. Coiled coil proteins are involved in a wide variety of functions, for example, gene regulation [2], cell division [3], and membrane fusion [4]. One of the main characteristics of a coiled coil is the simplicity of the sequence; they consist of a motif that repeats itself every seven residues, (abcdefg)n. These heptad repeats have alternating hydrophobic and hydrophilic residues, where the folding is driven by a hydrophobic collapse of the amphiphilic α-helices. In coiled coils, two or more helices wrap around each other in a left-handed or right-handed helical twist. These helices can adopt different topologies, as they can assemble in parallel, antiparallel, or in a mixed orientation, originally proposed by Crick [5], and later, this repeating pattern of hydrophobic residues was observed in rabbit skeletal α-tropomyosin [6,7]. Here, it was also found that other factors such as electrostatic interactions could further stabilize the coiled coil [6]. On the basis of the aforementioned criteria, the first design of peptide sequences, as inspired by α-tropomyosin, was synthesized by Lau et al. [8] and their TM29 (n = 4) was shown to be a dimer in solution.

Three-stranded coiled coils are found among others in many fusogenic proteins including influenza hemagglutinin [9,10], HIV [11], and Maloney murine leukemia virus [12]. One of the first high-resolution structures of trimeric coiled coil was a mutant of the GCN4 leucine zipper [13,14]. Another high resolution structure was the designed sequence CoilSer, which was the second generation structure based on TM29 [15]. This sequence was initially designed to form a double stranded coiled coil, but the crystal structure revealed a triple-stranded coiled coil with an antiparallel orientation [15]. One of the causes for the antiparallel orientation was a steric effect from the bulky indole side chain in Trp-2 in the hydrophobic core. CoilSer was found to exist in solution in a noncooperative monomer–dimer–trimer equilibrium [16]. The high-resolution structure findings of CoilSer lead to the design of CoilVaLd. In this sequence, the charged residues Glu and Lys in the e and g positions were preserved, and the Trp-2 residue and the Leu positions at 9, 16, and 23 were replaced by Val [17–19]. The solution structure analysis of CoilVaLd also revealed a cooperative monomer–dimer–trimer equilibrium, but apparently with the dimer state being a highly unstable intermediate [19]. Later, the crystal structure of CoilVaLd revealed a parallel three-helix bundle [17].

Coiled coils have previously been studied extensively by circular dichroism (CD), analytical ultracentrifugation (AUC), and X-ray crystallography [20]. Over the last decade, small-angle X-ray scattering (SAXS) has developed into a powerful biophysical tool. We have previously described the use of SAXS to study the topology of de novo designed proteins in solution [21]. Considering the importance of coiled coil structures, i.e. noncovalent assemblies of amphipatic peptides, we performed a systematic investigation of...
coiled coiling studied by SAXS and correlated this with AUC. This will also provide a ‘backdrop’ for the further use of SAXS in de novo design of proteins using these sequences. In addition, coiled coils with various sequence additions were studied, which are relevant for later use in de novo design.

Mutter et al. reported the concept of template-assembled synthetic proteins [22,23]. Our group has developed a class of de novo proteins termed carboproteins [24]. In previous work, we designed a carboprotein with a 4-helix bundle topology. Interestingly, our small angle X-ray scattering data clearly revealed that it instead formed a 3-helix bundle carboprotein [21], although the peptide sequence used had been reported to favor 4-helix bundles [25].

Here, we present the design and synthesis of seven different coiled coil motifs and analyze the differences in their folding topology using SAXS and AUC.

Materials and Methods

Experimental Section

All organic solvents and N\textsuperscript{9}-9-fluorenylmethyloxycarbonyl (Fmoc) amino acids were obtained from Iris Biotech GmbH. Tentagel Rink amide resin was obtained from Rapp Polymere GmbH. Preparative and analytical HPLC was performed on a Dionex Ultimate 3000 with Chromeloen 6.80SP3 software. Peptides were purified by preparative RP-HPLC on an FEF 300 Å C4 column (5 μm, 20 × 250 mm), applying a flow of 10.0 ml/min with a linear gradient of increasing amount of buffer (10% B to 100% B) over 37 min (buffer A: 0.1% trifluoroacetic acid in H\textsubscript{2}O; buffer B: 0.1% TFA in CH\textsubscript{3}CN). Mass spectroscopy was performed on an ESI-MS mass spectrometer (MSQ Plus, Thermo). Peptides were analyzed on a Phenomenex Gemini 110 Å C18 column (3 μm, 4.6 × 50 mm) or on a Phenomenex Gemini 110 Å C4 column (3 μm, 4.6 × 50 mm), applying a flow of 1.0 ml/min with a linear gradient with increasing amount of buffer B over 10–20 min (buffer A: 0.1% formic acid in H\textsubscript{2}O; buffer B: 0.1% formic acid in CH\textsubscript{3}CN).

Peptide Synthesis

The peptides were synthesized by automated peptide synthesis on a Syro II peptide synthesizer (Biotage AB) using standard solid-phase peptide synthesis on a Tentagel Rink amide resin with Fmoc for protection of N-9-amino groups. Side-chain protection groups were tert-butyloxycarbonyl (Boc) for (Lys), N\texttextsuperscript{-9}-methylmethan-aminium hexafluorophosphate (3 equiv), 1-hydroxybenzotriazole (3.6 equiv), 1-hydroxy-7-dimethylformamide (DMF) for 1 h, and triethylsilane (TES)-H\textsubscript{2}O (95:2.5:2.5) for 2 h at RT. The TFA solutions were precipitated with diethylether to yield the crude compounds.

CoilSer 1: Analytical RP-HPLC: purity 95%. ESI-MS (m/z) 1674.8 [M + 2H]+, 1116.8 [M + 3H]+, 837.4 [M + 4H]+ (MW\textsubscript{calc} = 3346.9).

CoilSer-G-desE 2: Analytical RP-HPLC: purity 97%. ESI-MS (m/z) 1090.1 [M + 3H]+, 817.5 [M + 4H]+, 654.2 [M + 5H]+ (MW\textsubscript{calc} = 3265.8).

CoilV\textsubscript{Ld} 3: Analytical RP-HPLC: purity 95%. ESI-MS (m/z) 1609.7 [M + 2H]+, 1073.4 [M + 3H]+, 805.4 [M + 4H]+ (MW\textsubscript{calc} = 3217.7).

CoilV\textsubscript{Ld}-YG 4: Analytical RP-HPLC: purity 95%. ESI-MS (m/z) 1143.3 [M + 3H]+, 858.0 [M + 4H]+, 686.6 [M + 5H]+ (MW\textsubscript{calc} = 3428.9).

CoilV\textsubscript{Ld}-3-Y 5: Analytical RP-HPLC: purity 99%. ESI-MS (m/z) 1251.2 [M + 2H]+, 834.7 [M + 3H]+, 626.2 [M + 4H]+ (MW\textsubscript{calc} = 2501.8).

CoilV\textsubscript{Ld}-3-Y-Ahx 7: Analytical RP-HPLC: purity 98%. ESI-MS (m/z) 1279.5 [M + 2H]+, 853.3 [M + 3H]+, 639.9 [M + 4H]+ (MW\textsubscript{calc} = 2556.7).

Circular Dichroism Spectroscopy

All peptides were dissolved in 50 mM acetate buffer, pH 5.5. Concentrations were determined by UV absorbance using a molecular absorption coefficient, ε\textsubscript{280} of 1490 M cm/Tyr, except for CoilV\textsubscript{Ld} 3 where the concentration was determined by amino acid analysis. Far-UV CD data were recorded on a JASCO J815 calibrated with ammonium d-10-camphorsulfonate. All spectra were recorded at room temperature using a 0.01 cm cell path length and a peptide concentration of about 1 mg/ml. The resulting Δε is based on the molar concentration of peptide bond. CD temperature scans were performed between 0 and 100 °C in a 1-mm quartz cuvette, by preparing the solution at room temperature and subsequently lowering the temperature in a thermostated cell holder in the CD instrument. After achievement of a constant ellipticity at 0 °C, the temperature was raised to 110 °C with a ramp rate of 1 °C/min to achieve complete thermal denaturation.

Analytical Ultracentrifugation

All peptides were dissolved in 50 mM acetate buffer pH 5.5 and extensively dialyzed against the same buffer. Sedimentation velocity and equilibrium experiments were performed on a Beckman XL-1 analytical centrifuge at 25 °C where the sedimentation was measured at an absorbance wavelength of 280 nm. The sedimentation velocity experiments were performed at 50 000 rpm, and the data was analyzed using a c(S) model implemented by SEDFIT [26]. The sedimentation equilibrium experiments were performed at 25 000, 30 000, and 35 000 rpm and the data fitted to the following equation:

\[
c_r = c_0 \exp \left[ \frac{M_w r - r_0^2}{2R T} \left( 1 - \left( \frac{\tilde{\nu}}{\tilde{\nu}_0} \right)^2 \right) \right] + E
\]

where \(c_r\) is the peptide concentration at radius \(r\), \(c_0\) is the concentration of monomeric peptide at \(r_0\), \(\tilde{\nu}\) is the angular velocity, \(R\) is the gas constant equal to 8.314 × 10\(^7\) erg mol/K, \(T\) is the temperature in degrees in Kelvin, \(M_w\) is the molecular mass, \(\tilde{\nu}\) is the partial specific volume of the peptide calculated using the additivity scheme described in Makhatazde et al. [27], \(\rho\) is the density of the solvent, and \(E\) is the baseline offset. All fits were carried out using nonlinear regression software (NLREG) and in-house written scripts.
Small-angle X-ray Scattering Measurements and Data Processing

Small-angle X-ray scattering measurements were performed on ID14-3 at ESRF, Grenoble, France, and at the JJ X-ray SAXS laboratory source at the University of Copenhagen, Faculty of Science, using standard procedures. The momentum transfer \( q = 4\pi \sin \theta / \lambda \), \( 2\theta \) is the scattering angle and \( \lambda \) is the X-ray wavelength (0.931 Å for ID14-3 and 1.54 Å for the home source). All SAXS measurements were performed at 24 ± 1 °C. All peptides were dissolved in 50 mM acetate buffer, pH 5.5 with a concentration of 2.5, 5, and 10 mg/ml. The scattering intensities of buffer backgrounds were measured both before and after the sample, and the averaged background scattering was subtracted from the scattering of the sample according to standard procedures. Furthermore, the data were background subtracted and converted into absolute scale units (cm²/cm³) by using water and BSA (5 mg/ml) as external standards. In the process of preliminary data analysis, the obtained scattering intensities, \( I(q) \), were converted into a direct space representation in terms of the pair distance distribution function, \( p(r) \), by means of indirect Fourier transformation (IFT) using the IFT method based on Bayesian statistics that has been built into BioXTAS RAW [28]. For proteins, which have an almost uniform electron density, the \( p(r) \) function can be considered as a histogram of inter-particle distances [29]. Thus, both the \( I(q) \) and \( p(r) \) plots contain information about the particle shape and size. Molecular mass determinations of the peptides were performed according to Mylonas et al. [29]. A partial specific volume of 0.73 cm³/g was used for the peptide.

Analysis software

The models were implemented using an in-house Fortran program containing a least squares fitting routine that allowed for fitting the mathematical models to the experimental data. CRYSOL was used to determine the numerical scattering intensities corresponding to the peptide monomeric, dimeric, and trimeric states.

Molecular Dynamics Simulations

The initial 3D model of CoilVaLd-3-Y 5 trimer was built with MOD-ELLER9v7 [30] using the crystal structure of CoilVaLd (PDB ID: 1COI) as template [17]. Because the C-terminal part of CoilVaLd deviates from \( z \)-helicity, the sequence alignment between CoilVaLd-3 and 1COI was performed starting from the N-terminal. Symmetry restraints were used in the modeling procedure to ensure C3 symmetry of the triple-helical CoilVaLd-3, 5, model. The initial 3D model of the CoilVaLd-3-Y dimer was created by manual deletion of a single \( z \)-helix in the CoilVaLd-3-Y, 5, trimer model. MD simulations were performed with Desmond 3.0 [31]. Using the Desmond System Builder, the dimeric and trimeric CoilVaLd models were each immersed in a cubic box of TIP3P water providing a minimum layer of 15 Å of water on each side of the peptide. Prior to MD simulations, a steepest descent minimization to a gradient of 1 kcal/mol Å was carried out.

This was followed by the default presimulation protocol employed in Desmond consisting of (i) minimization with restraints on solute, (ii) unrestrained minimization, (iii) Berendsen [32] NVT (constant number of particles (N), volume (V), and temperature (T)) simulation, \( T = 10^{K} \), small time steps, restraints on heavy solute atoms, (iv) Berendsen NPT (constant number of particles (N), pressure (P), and temperature (T)) simulation, \( T = 10^{K} \), restraints on solute heavy atoms, (v) Berendsen NPT simulation with restraints on heavy solute atoms, (vi) unrestrained Berendsen NPT simulation. Following the relaxation protocol, a 20 ns NPT simulation was carried out for the CoilVaLd dimer and trimer. The temperature was regulated with the Nose–Hoover chain thermostat [33] with a relaxation time of 1.0 ps. The pressure was regulated with the Martyna–Tobias–Klein barostat [34] with isotropic coupling and a relaxation time of 2.0 ps. The RESP integrator [35] was employed with bonded, near, and far time steps of 2.0 fs, 2.0 fs, and 6.0 fs, respectively. MD trajectories were saved to disk at 20 ps intervals. The OPLS_2005 force field [36] was employed for minimizations and simulations. A 9 Å cut-off was used for non-bonded interactions and the smooth particle mesh Ewald [37] method with a tolerance of \( 10^{-9} \) for long-range Coulomb interactions. MD trajectories were analyzed using the VMD program [38] and in-house Tcl scripts. Mean radius of gyration and intraprotein maximum atom–atom distance were calculated using 100 equally spaced snapshots from the MD trajectories. Images of molecular models were generated with VMD.

Results

Design and Synthesis

The aim was to use SAXS and AUC to study coiled coils, which would favor a three-stranded helical formation and could be used in subsequent de novo design of proteins. Well-known coiled coil sequences, as well as derived structures, were evaluated. We included CoilSer, 1, and a modified sequence, 2, as well as CoilVaLd, 3, and four modified sequences, 4–7 (Table 1). The

### Table 1. Synthesized peptide sequences and CD

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Sequences</th>
<th>( \alpha )-helicity/%b</th>
<th>( \theta_{222}/\theta_{208}$c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CoilSer</td>
<td>Ac-E WEALEKK LAALESK LQALEKK LEALEHG-NH₂</td>
<td>88</td>
<td>1.04</td>
</tr>
<tr>
<td>2</td>
<td>CoilSer-G-desE</td>
<td>Ac- EWALEKK LAALESK LQALEKK LEALEKG G-NH₂</td>
<td>86</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>CoilVaLd-3</td>
<td>Ac-E WEALEKK VAALESK VQALEKK VEALEHG-NH₂</td>
<td>69</td>
<td>1.01</td>
</tr>
<tr>
<td>4</td>
<td>CoilVaLd-3-Y</td>
<td>Ac-Y WEALEKK VAALESK VQALEKK VEALEKG G-NH₂</td>
<td>67</td>
<td>1.02</td>
</tr>
<tr>
<td>5</td>
<td>CoilVaLd-3-Y</td>
<td>Ac-Y VAALESK VQALEKK VEALEKG GG-NH₂</td>
<td>41</td>
<td>0.85</td>
</tr>
<tr>
<td>6</td>
<td>CoilVaLd-3-Y-Ahx</td>
<td>Ac-Y VAALESK VQALEKK VEALEK-Ahx G-NH₂</td>
<td>28</td>
<td>0.81</td>
</tr>
</tbody>
</table>

*Ahx: 6-amino-hexanoic acid.

*bCD was measured at 25 °C.

*\( \alpha \)-helicity calculated according to Chen et al [41]. The ratio describes the coiled coil formation ability.
modifications were selected also with a view to the use of these sequences in subsequent design of artificial proteins. Peptide 4 was derived from 3 by replacing His in the last heptad with Lys in order to facilitate the synthesis and make it more compatible with an eventual C-terminal modification as the aldehyde. Furthermore, an extension of the N-terminal and C-terminal with Tyr and Gly, respectively, was performed. Peptides 5–7 were truncated to comprise just three heptad repeats, indicated by a ‘3’ in their names. We expected that this substitution of His with Lys would have no effect on the tertiary structure, as it is a conservative change at the f position. Some of the coiled coil sequences were extended at the C-terminal to study the effect of these extensions, with the view to later tether to a carbohydrate template, resulting in a so-called carboprotein [24]. Moreover, all the CoIV_L_d sequences had an additional N-terminal Tyr to allow concentration determination by UV-absorption (Table 1) [39]. As a control, we designed CoII-L_d-G-2, which lacks a Glu in the first heptad. To evaluate if the length contributed to the degree of self-assembly into three-stranded helical coiled coils, the first heptad of CoIV_d was eliminated (Table 1, CoIV_L3-Y 5). Small changes in the peptide sequences can potentially cause radical changes in coiled coil topology [40]. Two sequences were designed, for eventual use in carboproteins, with the addition of only one amino acid residue. One contained an additional Gly at the N-terminal, whereas the other included 6-aminohexanoic acid (Table 1, CoIV_L3-YG 6, CoIV_L3-Y-Ahx 7, respectively). These analogs were synthesized to evaluate if a longer linker would induce any change to the self-assembly of the coiled coils. Sequences 1–6 were subjected to the PairCoil algorithm for prediction of coiled coil propensities (Supporting Information, Table S6), which showed a good correlation with CD results (vide infra).

The peptide amides were synthesized on a Syro II peptide synthesizer (Biotage AB) on a Tentagel Rink amide resin. After completion of the chain elongation, the peptides were N-acetylated. Cleavage from the resin was achieved using TFA-TES-H2O (95:2.5:2.5) for 2 h at room temperature. Each of the peptides was obtained in ≥95% purity after HPLC, and each gave an LC-MS spectrum consistent with the expected molecular weight.

Biophysical Studies

CD spectroscopy

The CD spectra of all peptides obtained at 25 °C in 50 mM acetate buffer pH 5.5 are shown in Figure 1(A) and (B). The seven peptides all exhibit a high degree of α-helicity with negative maxima at 222 and 208 nm. The helical content was calculated based on the mean residue ellipticity accordingly to Chen et al. (Table 1) [41]. The θ_222/θ_208 nm ratio ≥ 1 for coiled coil and ≤ 0.86 for isolated helices was also calculated (Tables 1 and S6) [8,42]. The helical contents for CoIV_d 3 and CoIV_L3-YG 4 were 69% and 67%, respectively. The θ_222/θ_208 ratios were 1.01 and 1.02, which were consistent with the existence of stable α-helical coiled coils in solution [8,42]. However, when the length was shortened with one heptad, the helicity was only in the range of 28–41%, and the θ_222/θ_208
ratios were less than 1 (CoilVaLd-3-Y, CoilVaLd-3-YG, and CoilVaLd-3-Y-Ahx, 5–7). CoilSer 1 and CoilSer-G-desE 2 remarkably had the same degree of helicity according to far-UV CD data (Figure 1(A)), but as the θ222/θ208 ratios for CoilSer 1 was 1.04, the ratio was slightly less for CoilSer-G-desE 2 0.98. This indicates that the double mutation in CoilSer-G-desE 2 afforded a lowering in the helical content.

A temperature scan at 0–110 °C using CD spectroscopy at 222 nm revealed that no transition to a fully unfolded state was observed up to 90 °C for CoilSer 1, CoilSer-G-desE 2, and CoilVaLd-YG 4, i.e. the folding of these peptides was very thermostable (data not shown). The thermal unfolding was found to be reversible upon lowering of the temperature. The shorter peptides (CoilVaLd-3-Y, CoilVaLd-3-YG, and CoilVaLd-3-Y-Ahx, 5–7) were much less stable with a Tm around 33 °C (data not shown), and also, here, the thermal unfolding was reversible. The thermodynamic stability of CoilVaLd-3-Y 5 was studied at with different guanidinium chloride (GdnCl) concentrations, to destabilize the native state and monitored at 222 nm (Figure 1(C)). The absolute ellipticity values at 0 °C decreased with increasing GdnCl, because of a decrease in the fraction of the native state of the peptide. The denaturation curves were all sigmoidal with a single inflection point; the transition profiles were consistent with a two-state equilibrium model. The CoilVaLd-3-Y, 5, had a Tm of 33 °C in the absence of GdnCl. At 4 M, GdnCl, θ222 approached zero, reaching a fully denatured state.

On the basis of the stability studies, we speculated that there might be a concentration-dependent folding process. Therefore, a CD concentration series of CoilVaLd-3-Y 5 was measured. It clearly showed that with increasing concentration, there was a concurrent rise in the degree of folding (Figure 1(D)). The θ222/θ208 ratio range of 0.83 (40 μM) to 0.98 (800 μM), indicated the formation of stabilized assemblies at higher concentration.

**Analytical ultracentrifugation**

Analytical ultracentrifugation sedimentation velocity experiments were performed on the different peptides to determine the distribution of the different oligomers present (Figure 2). Each peptide generally had a single peak showing that there was only one predominant species present at those concentrations (50–400 μM); there were no significant higher-order species, i.e. aggregates, present (Figure 2). The sediment coefficient (S_{w,20}) was estimated, and the values were in the expected range for small peptides (Table 2). Sedimentation equilibrium AUC was used to examine the oligomeric properties of the peptides and to determine the molecular weights of the species present. Each sample was measured at two concentrations (60 and 200 μM), and data sets were obtained at three rotor speeds. All peptides were analyzed except CoilSer 1 and CoilVaLd 3, which already had been reported to fit to single species with molecular masses of trimers (Table 2). The long sequences CoilSer-G-desE 2 and CoilVaLd-YG 4 both fitted best with a single species with a molecular size of 10.5 and 10.2 kDa, respectively. This also correlates very well with the molecular size of a trimer (Table 2). The sediment equilibrium analysis fit of CoilVaLd-YG 4 is shown in Figure 3. The ratio for the shorter peptides (CoilVaLd-3-Y, CoilVaLd-3-YG, and CoilVaLd-3-Y-Ahx, 5–7) corresponded to somewhere between a dimer and trimer. All attempts to fit the data set for CoilVaLd-3-Y 5 to a single species resulted in a poor fit to the experimental data. When fitted with more variables such as mixtures of oligomeric states, better fits were obtained especially when including a monomer/dimer/trimer equilibrium (data not shown).

**Small-angle X-ray scattering**

X-ray scattering data provides structural information on peptide self-assembly in solution, including information on the topology. Initially, scattering data were collected at three different concentrations for all peptides (2.5, 5, and 10 mg/ml). No concentration-dependent effects were observed at these concentrations (see supporting information). The recorded scattering profiles and the indirect Fourier transformed data are shown in Figure 4. The maximum diameter (D_{max}) of the molecular assemblies, the radius of gyration R_g and the molecular weight were determined (Table 2). The scattering data for CoilSer 1 and CoilSer-G-desE 2

### Table 2. Result from AUC and parameters from indirect Fourier transformation of SAXS data (20 °C)

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>S_{w,20}</th>
<th>M_{w,AUC} (kDa)</th>
<th>M_{w,IND} (kDa) (monomer)</th>
<th>Ratio</th>
<th>D_{max} (Å)</th>
<th>R_g (Å)</th>
<th>M_{w,SAXS} (kDa)</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CoilSer^a</td>
<td>n.d</td>
<td>10.2[16]</td>
<td>3.3</td>
<td>3.1</td>
<td>48.3 ± 0.6</td>
<td>15.1 ± 0.02</td>
<td>10.0</td>
<td>3.0</td>
</tr>
<tr>
<td>2</td>
<td>CoilSer-G-desE</td>
<td>1.33</td>
<td>10.1 (9.9)</td>
<td>3.3</td>
<td>3.1</td>
<td>88.4 ± 1.0</td>
<td>15.2 ± 0.2</td>
<td>10.2</td>
<td>3.1</td>
</tr>
<tr>
<td>3</td>
<td>CoilVaLd-3-Y</td>
<td>n.d</td>
<td>10.1[19]</td>
<td>3.2</td>
<td>3.1</td>
<td>42.0 ± 0.04</td>
<td>14.3 ± 0.02</td>
<td>10.1</td>
<td>3.1</td>
</tr>
<tr>
<td>4</td>
<td>CoilVaLd-3-YG</td>
<td>1.41</td>
<td>10.5 (10.2)</td>
<td>3.4</td>
<td>3.1</td>
<td>65.1 ± 0.4</td>
<td>17.4 ± 0.06</td>
<td>10.1</td>
<td>3.0</td>
</tr>
<tr>
<td>5</td>
<td>CoilVaLd-3-Y</td>
<td>0.90</td>
<td>5.6 (6.1)</td>
<td>2.5</td>
<td>2.2</td>
<td>49.0 ± 0.08</td>
<td>13.8 ± 0.01</td>
<td>6.0</td>
<td>2.4</td>
</tr>
<tr>
<td>6</td>
<td>CoilVaLd-3-YG</td>
<td>0.88</td>
<td>6.1 (6.1)</td>
<td>2.6</td>
<td>2.3</td>
<td>47.0 ± 0.07</td>
<td>13.4 ± 0.02</td>
<td>5.9</td>
<td>2.3</td>
</tr>
<tr>
<td>7</td>
<td>CoilVaLd-3-Y-Ahx</td>
<td>0.94</td>
<td>6.3 (6.4)</td>
<td>2.6</td>
<td>2.2</td>
<td>46.0 ± 0.06</td>
<td>13.5 ± 0.02</td>
<td>5.6</td>
<td>2.2</td>
</tr>
</tbody>
</table>

^aValues in parentheses are the sediment velocity molecular weights.
were recorded (Figure 4(A)). The corresponding pair distance distribution functions \((p(r))-function\) for CoilSer-G-desE \(2\) changed slightly compared with the known CoilSer \(1\) (Figure 4(B)) giving rise to a somewhat higher \(D_{\text{max}}\). The anomalous \(D_{\text{max}}\) of 88.4 for \(2\) is most likely caused by the presence of a small population of oligomers/aggregates. They both had an apparent molecular weight corresponding to a trimer (Table 2). CoilSer \(1\) and CoilSer-G-desE \(2\) both consist of four heptads with a length of 29 amino acids, but again, the latter has a His-to-Lys mutation and an additional C-terminal Gly. The major difference was the deletion of Glu in the first heptad. The effect of these small changes can clearly be seen in the \(p(r))-function were CoilSer \(1\) has a more compact structure with a smaller \(D_{\text{max}}\). CoilSer \(1\) and CoilVaLd \(3\) both consist of four heptads with a length of 29 amino acids, but again, the latter has a His-to-Lys mutation and an additional C-terminal Gly. The major difference was the deletion of Glu in the first heptad. The effect of these small changes can clearly be seen in the \(p(r))-function were CoilSer \(1\) has a more compact structure with a smaller \(D_{\text{max}}\). They both have nearly bell-shaped \(p(r))-functions with a \(D_{\text{max}}\) of 48 and 42 Å. In both cases, a molecular weight corresponding to a trimer (Table 2) was observed.

The \(p(r))-function for the four analogs of CoilVaLd (Figure 4(C) and (D)) have the same bell-shaped structure as CoilVaLd \(3\) itself, but they all have significantly higher \(D_{\text{max}}\) values (Table 2). One analog that is particularly different from the original CoilVaLd was the CoilVaLd-YG \(4\). It has the same number of heptad repeats as CoilVaLd. In CoilVaLd-YG \(4\), a further Gly was added to the C-terminal along with Tyr at the N-terminal. This analog with a \(D_{\text{max}}\) of 65 Å and \(R_g\) of 17.9 Å contained a small population of some larger assemblies that was seen from the \(p(r))-function.

**Figure 3.** Sedimentation equilibrium analysis. Measurements were made using 180 μM of CoilVaLd-YG. Constant used for the data analysis are \(\eta_{\text{bar}} = 0.761 \text{ ml/g}\); and \(\rho = 1.0017 \text{ g/ml}\). The data are shown to fit a single species. Residuals to the fit are shown in the lower panel.

**Figure 4.** (A) SAXS data of CoilSer \(1\) and CoilSer-G-desE \(2\). (B) The corresponding pair distance distribution \((p(r))-function\) for CoilSer and CoilSer-G-desE. (C) SAXS data of for CoilVaLd and its four analogs. (D) the corresponding pair distance distribution \((p(r))-function\) for CoilVaLd and its four analogs. All sample with a concentration of ~5 mg/ml in 50 mM acetate buffer pH 5.5.
The major component of CoilVaLd-YG 4 is a compact structure with a smaller \( D_{\text{max}} \). The obtained molecular weight of CoilVaLd-YG 4 is as expected for a full trimer (10 kDa). Finally, the \( p(r) \)-function of CoilVaLd-Y, CoilVaLd-3-Y, and CoilVaLd-3-Y-Ahx 5–7 were very similar. The combination of a low average \( M_w \) and a large \( D_{\text{max}} \) can only be explained in terms of a mixture of oligomeric states (Figure 4(D)).

**Molecular dynamics simulations**

A 20 ns MD simulation was performed on CoilVaLd-3-Y 5 as a dimer and a trimer, to understand the dynamics involved in the coiled coil interaction and to provide plausible atomistic models for the simulation of SAXS data. The MD simulations presented here do not as such yield a measure of the relative thermodynamic stability of the dimeric and trimeric association states of the CoilVaLd-3-Y 5 but rather provide visual and semi-quantitative information on their packing and dynamics. This insight is useful for subsequent peptide and protein design. The initial 3D models of CoilVaLd-3-Y 5 in dimeric and trimeric configuration are shown in Figures 5(A), (C) and 6(A), (C), respectively. Root-mean-square deviation plots for backbone atoms (Supporting Information, Figure S2) show that the middle part (residues 2 to 19) of each chain in both the dimer and trimer is structurally stable throughout the simulation, whereas the N-terminal and particularly C-terminal residues undergo significant movement. Backbone root mean square fluctuations (Supporting Information, Figure S2) quantify the dynamics: small (~0.8 Å) thermal fluctuations were found for residues 2–19 for all chains in both the dimer and trimer and larger fluctuations for the first N-terminal residue and the last four C-terminal residues in both simulations. The terminal fluctuations were increased in the trimer relative to the dimer, particularly at the C-terminal (average root mean square fluctuations across chains and last four residues for the dimer = 2.5 Å and for the trimer = 3.2 Å). This indicates a weaker association between the chains in the trimer.

The simulation averaged interchain distances between Val \( \alpha \)-carbons (residues 2, 9, and 16) and Leu \( \alpha \)-carbons (residues 5, 12, and 19) (Supporting Information, Table S2 - S5) confirmed that the two chains in the dimer approached each other more closely (average \( C_{\alpha} \) distance for Val = 5.9 Å and Leu = 6.3 Å) than any two chains in the trimer (average \( C_{\alpha} \) distance for Val = 7.4 Å and Leu = 8.3 Å).

Furthermore, monitoring secondary structure in the simulations showed that the chains in the trimer had a lower number of residues in \( \alpha \)-helical conformation relative to the chains in the dimer (Supporting Information, Figure S3–S7). On average, ~84% and ~76% of the residues are found in \( \alpha \)-helical form for the dimer and trimer simulations, respectively.

The statistics for root-mean-square deviation, interchain distances, and secondary structure conservation are congruent with visual inspection of the first and last structure snapshots from the dimer and trimer simulations. In the dimer at 20 ns (Figure 5(B)), the Val and Leu side chains were packed in similar fashion as in the starting structure (Figure 5(A)). This was also reflected in a modest change in the interhelical angle from 22.3° in the initial structure (Figure 5(C)) to 16.4° in the structure at 20 ns (Figure 5(D)). In contrast, the packing of Val and Leu side chains for the trimer at 20 ns (Figure 6(B)) clearly changed relative to the starting point for the simulation (Figure 6(A)). This is accompanied by a significant decrease in the interhelical angles (by up to 16°, see Supporting Information, Table S6) as evident from the side view of the starting structure (Figure 6(C)) and the structure at 20 ns (Figure 6(D)). Altogether, these findings indicate that the CoilVaLd-3-Y 5 trimer is destabilized relative to the dimer in the simulations. This is in qualitative agreement with our experimental findings from SAXS (*vide infra*) that indicate that the dominating species in vitro are monomers and dimers. From the MD simulations, a mean radius of gyration and maximum distance was found for the dimer \( (R_g = 11.7 \pm 0.2 \, \text{Å}, \, D_{\text{max}} = 40.8 \pm 2.1 \, \text{Å}) \) and trimer \( (R_g = 12.6 \pm 0.2, \, D_{\text{max}} = 44.5 \pm 1.7 \, \text{Å}) \). The dimer thus exhibited a shorter average maximum distance \( D_{\text{max}} \) than the trimer. This appears to be related to the tighter packing of residues observed in the MD simulation of the dimer.

**Figure 5.** Molecular dynamics: structures of peptide amide 5 as a parallel dimer. (A) Initial structure front view and (B) after 20 ns MD. (C) Initial structure side view and (D) after 20 ns MD. Val and Leu are color coded as red and green, respectively. \( \alpha \)-Helices and coils are shown in cylinder and tube representation, respectively.

**Figure 6.** Molecular dynamics: structures of the peptide amide 5 as a parallel trimer. (A) Initial structure front view and (B) after 20 ns MD. (C) Initial structure top view and (D) after 20 ns MD. Val and Leu are color coded as red and green, respectively. \( \alpha \)-Helices and coils are shown in cylinder and tube representation, respectively.
The scattering intensity from the computer models of CoilVaLd-3-Y 5 dimer and trimer were calculated using CRYSOI [43] and compared with the experimental SAXS data for CoilVaLd-3-Y 5.

First, the single species of monomer, dimer, and trimer were fitted using an in-house developed Fortran fitting routine. This afforded relatively good but not optimal fits. The best results were obtained for a mixture of monomer and dimer at a ratio of 90 ± 1% monomer and 10 ± 1% dimer. The fit was good with a relatively good chi square value $r^2 = 1.75$ (Figure 7).

**Discussion**

Our studies aimed at investigating self-assembly of a small ensemble of coiled coil peptides by SAXS and compared this with AUC. This would also be important for subsequent use of these sequences in de novo design of proteins. The original peptide design was to study the topological effect of small changes in well-known 3-stranded coiled coils, such as CoilSer 1 and CoilVaLd 3. CoilSer 1 is known to form an antiparallel triple stranded coiled coil, whereas CoilVaLd 3 is known to form a parallel triple stranded coiled coil. Here, we made one analog of CoilSer 1 with a switch of two amino acid in the first heptad (CoilSer-G-desE 2) of the N-terminal and compared the AUC and SAXS data with CoilSer 1. Furthermore, we synthesized four analogs of CoilVaLd 3 where His was substituted with Lys, and a Gly was added (CoilVaLd-YG 4). From CoilVaLd-YG 4, we deleted the first heptad from the N-terminal (CoilVaLd-3-Y 5), and on the basis of this new analog, we made two new analogs with small amino acid residue changes in the C-terminal end of CoilVaLd (CoilVaLd-3-YG and CoilVaLd-3-Y-Ahx, 6-7). These modifications of the trimeric coiled coil motifs, CoilSer 1 and CoilVaLd 3, lead us to several intriguing results.

First, we compared CoilSer 1 and CoilSer-G-desE 2. They both have the same number of heptad repeats but with minor deletion changes in the amino acid arrangement; they both fit well with a single species (trimer) by AUC and SAXS, and according to CD spectroscopy, they have a very high $\alpha$-helical content. Here, we can therefore conclude that the deletion of amino acids in the N-terminal end did not alter the oligomerization state of the CoilSer analog (CoilSer-G-desE 2).

The present data analysis revealed unequivocally that the truncated version of trimeric CoilVaLd 3, i.e. CoilVaLd-3-Y 5, resulted in a disruption of the 3-stranded coiled coil. The sequence with four heptads (CoilVaLd-3-YG 4), according to both AUC and SAXS data, confirmed the formation of a trimeric coiled coil, which had a molecular weight corresponding to a trimer. The CD spectrum (Figure 1B) indicates a very high helical content, and the ratio $\Theta_{222}/\Theta_{208}$ supported a coiled-coil tertiary structure. This is in agreement with literature data for CoilVaLd [16]. This meant that the substitution of His with Lys, as predicted, had no dramatic effect on the overall folding topology of 4 compared with the original CoilVaLd sequence 3 at pH 5.5.

The minimum length required for the formation of stable coiled coils has been reported to be in the range of 21–23 residues [44–46]. Thus, the sequence (CoilVaLd-3-Y 5) could be long enough to fold into well-defined tertiary structures. Interestingly, we observed that the deletion of the N-terminal heptad repeat altered the state of the oligomerization of CoilVaLd 3, according to both AUC and SAXS. For CoilVaLd-3-Y 5, the molecular weight both by AUC and SAXS did not correspond to a full trimer (ratio 2.2 and 2.4, respectively) and did not fit with a single species component by either AUC or SAXS. Neither did any of the other two analogs of CoilVaLd-3-YG and CoilVaLd-3-Y-Ahx, 6–7. The experimental absence of trimers of short CoilVaLd analogs is consistent with MD results. As detailed in the results section, the MD simulations of the CoilVaLd-3-Y 5 dimer showed decreased terminal backbone fluctuations, smaller average distances between backbone atoms in neighboring helices and a higher degree of $\alpha$-helicity compared with the trimer. Thus, given similar conditions, the MD simulations suggest that the dimer is more tightly associated than the trimer. This is also consistent with the CD data (Figure 1A) that showed that all three short versions of CoilVaLd possess a lower degree of $\alpha$-helicity, and the ratio $\Theta_{222}/\Theta_{208}$ indicated no occurrence of stable $\alpha$-helical coiled coils. It has been reported that deletion of heptad repeats can alter the oligomerization state of coiled coil [47]. Of course, on the basis of these studies, we cannot rule out that the deletion of a heptad from the other end could have a different effect, as the repeats are not fully identical. The thermal stability of CoilVaLd-3-Y 5 ($T_m=33$ C) was considerably reduced compared with CoilVaLd-YG 4 ($T_m=90$ C). This was probably due to the tighter packing of the hydrophobic residues in a trimeric coiled coil of CoilVaLd-YG 4 compared with the assembly of 3-heptad CoilVaLd-3-Y 5, where the selective formation of a three-helical coiled coil was not observed either by AUC or SAXS. $D_{max}$ for CoilVaLd-YG and CoilSer can be calculated to $\sim 43$ Å assuming 3.6 residues per turn and a helical repeat unit of 5.4 Å. This correlates well with the observed $D_{max}$ (42 and 48 Å) for both peptides by SAXS. Furthermore, the major components in CoilSer-G-desE 2 and CoilVaLd-YG 4 have a $D_{max}$ in that range. The experimental SAXS data for the shorter peptides (CoilVaLd-3-Y, CoilVaLd-3-YG and CoilVaLd-3-Y-Ahx, 5-7) all look very similar, and in general, adding a Gly to CoilVaLd-3-Y 5 did not significantly change the SAXS data as seen in the $p(r)$-function. The topology of CoilVaLd-3-YG 6 did not change compared with the original design (CoilVaLd-3-Y 5). Thus, the addition of a further amino acid did not alter the properties much. Instead, a long chain was added (6-amino hexanoic acid) to provide CoilVaLd-3-Y-Ahx 7, which had very similar properties compared with the original design (CoilVaLd-3-Y 5).

Figure 7. Representative model fit of retrieved MD simulation of CoilVaLd-3-Y. Experimental SAXS data of CoilVaLd-3-Y 5. Fit to a mixture of 90 ± 0.66% monomer and 10 ± 0.33% dimer (solid black line).
We further studied the short peptide (CoilV4Ld-3-Y)5. Both in the SAXS and AUC data, we observed molecular masses that did not correspond to a trimer. None of the attempts to fit the AUC data of 5 to a single species resulted in a satisfactory data fit. Different mixtures were also fitted, but none of them gave an optimal fit. The best fit was obtained when all components in the mixture (monomer/dimer/trimer) were present in the model. These findings also correspond to the SAXS data analysis using the MD simulated structures for dimer and trimer. Here, the best fit was obtained using a monomer/dimer mixture. The fit did not become significantly better if a monomer/dimer/trimer mixture was used, but the error bars of the fit parameters would become somewhat higher (data not shown).

Rewardingly, AUC and SAXS provided similar results regarding the aggregational state of these coiled coils, where sequences with four heptads lead to relatively stable and well-defined secondary and quaternary structure, whereas sequences with only three heptads provided a more heterogeneous ensemble of dimers and trimers. The pair-distance distribution information, which is relatively easy to obtain by SAXS, revealed that peptide 2, which is the slightly mutated version of 1, behaved mostly like 1 but also included a small population of oligomers with a larger Dmax. In contrast, AUC did not show the presence of larger aggregates for 2 thus revealing some slight differences provided by these two methods.

Conclusion

We studied a small ensemble of designed coiled coil peptides and analyzed their self-assembly, by CD spectroscopy, AUC, and SAXS. Our SAXS study corroborates previous findings on two known de novo designed coiled coil peptides (CoilSer and CoilVdLd). Minor changes such as a His-to-Lys mutation and short extensions still provided trimeric structures of four-heptad sequences. Our study confirmed that there is a balance between two-stranded and three-stranded coiled coils, as demonstrated by reducing the length of CoilV4La with one heptad repeat from the N-terminal end to a three heptad. Interestingly, this revealed that instead of forming a three-stranded coiled coil, we observed multiple aggregational states from a single sequence. SAXS and AUC provided comparable results in this study. However, for 2, the presence of population of larger aggregates was shown by SAXS and not AUC. This study also provides the background for future protein design studies, where it will be interesting to see, whether shorter three-heptad sequences can form stable assemblies when covalently linked.

Acknowledgements

A NABiIT grant from DSF is gratefully acknowledged. Beamtime at the beamline ID14-3 at ESRF Grenoble is gratefully acknowledged. We thank Dr. A. Round for the excellent support received at the beamline.

References


