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).
Azobenzene-based amino acids for photocontrol of coiled coil peptides

This chapter is in preparation as an original research paper: Niek S.A. Crone, Niek v. Hilten, Alex v.d. Ham, Aimee L. Boyle, Alexander Kros, 2021.
Abstract

Coiled coil peptides have been under investigation as binding motifs in synthetic biology and biomaterials, with active control over their self-assembly allowing for a wider range of applications. To introduce photocontrol in coiled-coil assembly, three azobenzene-containing amino acids were prepared and incorporated into the hydrophobic core of the heterodimeric coiled coil formed between peptide ‘K₃’, (KIAALKE)₃ and peptide ‘E₃’ (EIAALEK)₃. Two amino acids (APhe1 and APhe2) were based on phenylalanine, differing in the presence of a carboxylic acid group, and have previously shown to allow photomodulation of protein activity. When incorporated in peptide K₃, binding to peptide E₃GY changed upon trans to cis isomerization, with the two variants differing in the most folded state. The third azobenzene-containing amino acid, APgly, is a novel design based on phenylglycine, and showed comparable absorption bands and isomerization to the phenylalanine based amino acids. When APgly was incorporated in the coiled-coil, a 4.7-fold decrease in folding constant was observed upon trans to cis isomerization, the largest difference for all three amino acids. Removal of the methylene group was theorized to position the diazene closer to the hydrophobic amino acids and reduce the possible rotations of the amino acid, with molecular dynamics simulations supporting these hypotheses. Combined, these results show APgly as a novel photoswitchable amino acid for the photocontrol of coiled-coil assembly. The introduction of APgly in existing coiled-coil based biomaterials should therefore allow for control over activity related to coiled-coil formation.
5.1 Introduction

In Chapter 4, the photocontrol of coiled-coil peptide assembly was investigated via intramolecular cyclization using an azobenzene-based crosslinker. This cyclization-based strategy affects peptide structure by providing an energetic constraint on the length of the crosslinked region, and by design does not interact with the amino acid side chains taking part in coiled-coil formation. Because the crosslinking strategy proved to be a major bottleneck in the preparation of photoactive peptides, alternative methods were sought. Therefore, in this chapter, the incorporation of azobenzene-based amino acids into the hydrophobic positions of coiled-coil peptides is investigated as a novel method for coiled-coil photocontrol.

Single amino acid modifications in the hydrophobic domain of coiled-coil forming peptides have been investigated both to aid understanding of coiled-coil assembly, and to introduce selectivity or functionality. For example, the group of Hodges investigated substitution at the ‘a’ position of a cysteine-crosslinked homotrimeric system, and, in a separate study, the role of the ‘d’ position in a crosslinked homodimeric peptide. The two positions showed a different preference for amino acids, with the general trend being that hydrophobic amino acids stabilize, and polar or charged amino acids destabilize, coiled-coil formation relative to alanine. When comparing aromatic amino acids, phenylalanine and tyrosine are well tolerated but the large double ring system of tryptophan does not provide higher stability compared to alanine. Acharya et al. performed single amino acid substitutions in the basic leucine zipper protein VBP, and observed changes in the Gibbs free energy of binding of up to 9.2 kcal/mol by substituting opposing amino acids in the ‘a’ position. Although not all possible amino acids were investigated, the same trend was observed as previously shown by the group of Hodges, and the effect remained when both chains were substituted. The role of asparagine in the hydrophobic core has also been investigated, because it is a common feature in natural coiled-coil peptides, and is useful to control peptide oligomerization although at a cost of overall stability of the coiled coil. To introduce a response to an external trigger, redox switching of methionine residues in the hydrophobic core has been investigated and results in disruption of the coiled coil, and substitution of hydrophobic amino acids for histidine residues leads to self-assembly which is dependent on pH or metal ion coordination in a trimeric coiled-coil system. This previous work on hydrophobic core modifications suggests that
the incorporation of azobenzene-based amino acids, which change structure and polarity via cis/trans isomerization through light illumination (see Chapter 4), should allow for control over coiled-coil assembly.

Phenylazo-phenylalanine (APhe1, Scheme 5.1) was one of the first published amino acids containing an azobenzene moiety, and has been incorporated into synthetic peptides and into protein structures via genetic code expansion. To date, APhe1 and structural derivatives have been used to control enzyme dimerization necessary for catalytic activity, the DNA binding strength of the CAP transcription factor by destabilization of its cAMP binding site, incorporation into superfolder green fluorescent protein, and allosterically reducing the protein-ligand binding strength in chemiluminescent luciferase or subunit interaction in imidazole glycerol phosphate synthase. Derivatives of APhe1 containing reactive sites for the generation of intramolecular crosslinked proteins have also been studied for control over helical folding, similar to the crosslinking strategy investigated in the previous chapter. Since its incorporation into peptides and proteins has been demonstrated to be straightforward, the potential ability of Aphe1 derivatives to control coiled-coil assembly would provide a novel method for the introduction of photoswitching into a wide variety of natural and synthetic coiled-coil systems already under investigation.

To test whether azobenzene-based amino acids can be used to control coiled-coil folding, peptides incorporating three structurally diverse non-natural amino acids were prepared by solid phase peptides synthesis (SPPS). Two of the amino acids were derived from L-phenylalanine, namely Aphe1 and Aphe2, and these differ in

Scheme 5.1: Hypothesized interactions of azobenzene in the coiled-coil hydrophobic core (A), and the three azobenzene based amino acids investigated in this chapter for the control over coiled-coil self-assembly (B).
the presence of a carboxylic acid group \textit{para} to the diazene group for \textbf{APhe2} (Scheme 5.1B). The carboxylic acid present in \textbf{APhe2} is theorized to introduce electrostatic repulsion with the opposing glutamic acid residue, destabilizing the coiled coil when the azobenzene is in the \textit{trans} conformation. Switching electrostatic interactions has been investigated previously via enzymatic serine phosphorylation,\textsuperscript{23} and allowed both stronger and weaker interactions after phosphorylation depending on the design.\textsuperscript{24} Besides azobenzene derivatives of \textit{l}-phenylalanine, the azobenzene derivative of \textit{l}-phenylglycine was prepared (\textbf{APgly}, Scheme 5.1B), with the expectation that this amino acid would result in a better incorporation in the hydrophobic core compared to \textbf{APhe1} due to the absence of a methylene group. The three azobenzene-based amino acids were incorporated into a dimeric coiled-coil which is used in our group as a SNARE protein mimic.\textsuperscript{25} This coiled-coil system consists of peptide ‘\textit{K}_3’, (KIAALKE)\textsubscript{3} and peptide ‘\textit{E}_3’ (EIAALEK)\textsubscript{3}, which self-assemble into a dimeric parallel coiled-coil with high binding affinity.\textsuperscript{26} After incorporation in peptide \textit{K}_3, \textbf{APgly} showed the largest difference in coiled-coil binding upon \textit{trans} to \textit{cis} isomerization, with a 4.65 fold decrease in the folding constant. Molecular dynamics (MD) simulations of \textbf{APhe1} and \textbf{APgly} in the coiled coil show more rearrangement of the \textbf{APhe1} azobenzene after isomerization, and support \textbf{APgly} as a better switch for coiled-coil photocontrol.

### 5.2 Results and Discussion

\textit{Peptides containing azobenzene derivatives of phenylalanine.}

Photoswitchable amino acids based on phenylalanine were prepared following literature procedures (Scheme 5.2),\textsuperscript{14} with the 9-fluorenylmethoxy-carbonyl (Fmoc) protecting group on the \textalpha-amine to facilitate use in SPPS. Fmoc-4-nitro-\textit{l}-phenylalanine, 1, was prepared from \textit{l}-phenylalanine via nitration and Fmoc-protection, or was purchased commercially. Subsequently the nitro group of 1 was reduced to the respective aniline with Zn powder and ammonium chloride, followed by a Mills reaction with nitrosobenzene to yield N\textalpha-fmoc-4-(phenylazo)-\textit{l}-phenylalanine, 2. The same procedure was followed for the preparation of protected \textbf{APhe2}, utilizing \textit{tert}-butyl-4-nitrosobenzoate in the Mills reaction resulting in product 3. The tert-butyl ester is used to prevent side reactions during SPPS, and is removed by the acidic peptide cleavage conditions to yield the carboxylic acid. Derivatives of peptide \textit{K}_3 (sequence shown in Scheme 5.2) were prepared via SPPS, with the isoleucine at position 9 replaced with \textbf{APhe1} or \textbf{APhe2}. 

121
Scheme 5.2: Reaction scheme for the preparation of azobenzene containing amino acids based on phenylalanine (top) and amino acid sequences of peptides prepared in this chapter (bottom). AzoAA refers to APhe1, APhe2 or APgly. Reagents and conditions: (i) HNO₃, H₂SO₄; (ii) Fmoc-chloride, NaHCO₃; (iii) Zn powder, NH₄Cl, EtOH; (iv) AcOH.

The UV absorption spectra of the two peptides were very similar (Figure 5.1A), with a slight redshift for the absorption bands of APhe2 (335 nm, $\varepsilon = 9,480 \text{ M}^{-1} \text{ cm}^{-1}$) compared to APhe1 (327 nm, $\varepsilon = 10,020 \text{ M}^{-1} \text{ cm}^{-1}$). Both amino acids could be isomerized to the cis conformation with 340 nm light, as observed by the disappearance over time of the strong absorption peak (Figure S5.1). Cis to trans isomerization from the 340 nm photostationary state (PSS) could be achieved efficiently with 435 nm light (Figure 5.1B). After plotting the change in intensity of the absorption maximum in the dark-adapted state, a larger difference after 340 nm irradiation is observed for APhe1, indicating this amino acid is isomerized to the cis conformation more completely by 340 nm light, or the cis conformer of APhe2 has a stronger absorption band (Figure 5.1C). Back isomerization from cis to trans with 435 nm light is slightly more efficient for APhe2, resulting in 98% of the absorption maximum observed in the dark state, compared to 92% for APhe1.

The structure of the K₃ peptides and coiled coils was studied via CD spectroscopy to determine the effect of phenylalanine-based photoswitches on peptide folding. Peptide E₃GY was used as binding partner, since the tyrosine can be used to accurately determine peptide concentration. Both K₃-APhe1 and K₃-APhe2 peptides were mostly unstructured by themselves, and able to form coiled coils in...
combination with E₃GY (Figure 5.2), with only marginal differences in structure between the dark adapted and 340 nm irradiated states. Because no obvious differences in structure were observed after irradiation, thermal denaturation experiments were performed to determine whether the binding affinity was dependent on sample irradiation. For this, CD melting curves measured at different concentrations were fitted to a coiled-coil binding model (Figure S5.2 and Figure S5.3). The best fit was achieved for a dimeric coiled coil over other oligomeric states, and important parameters from this fitting are shown in Table 5.1.

**Figure 5.1:** UV spectra of peptide K₃ with position 9 changed for APhe1 or APhe2, showing spectra for the dark-adapted peptide and after trans to cis isomerization with 340 nm light (A). Spectra after cis to trans isomerization using 435 nm light (B), and an overview of the absorbance at 335 nm for the dark adapted, 340 nm PSS and 435 nm PSS (C). [Peptide] = 50 µM in PBS buffer.

**Figure 5.2:** CD spectra of peptide K₃ with the 9th amino acid changed for APhe1 or APhe2, showing both the dark adapted and 340 nm irradiated states, as well as spectra with and without binding partner E₃GY. All samples were measured in 10 mM phosphate containing 2 mM NaCl, with [peptide] = 50 µM.
Photoisomerization of K₃-APhe₁ using 340 nm light resulted in an increase in the unfolding constant Kᵤ, from 1.09 to 1.41 µM, showing reduced coiled-coil binding in the irradiated state. The opposite was observed for K₃-APhe₂, where irradiation yielded a reduction in Kᵤ from 4.58 to 3.07 µM, revealing the 340 nm irradiated state to contain the most stable coiled-coil. The carboxylic acid group in APhe₂ is intended to increase electrostatic repulsion with glutamic acid residues of E₃ in the trans conformation, resulting in the cis conformation as the most stable coiled-coil. Contrary to our expectations, the cis isomer of K₃-APhe₁ was not the most stable state, showing that the bulk of the trans azobenzene moiety is better accommodated in the hydrophobic positions of the coiled-coil, than the increased polarity of the diazene in the cis conformation. In summary, incorporation of APhe₁ and APhe₂ in the coiled-coil hydrophobic core was successful, but the design of these amino acids needs to be improved since the difference in coiled-coil formation between the dark state and 340 nm PSS was small.

Given the small differences observed for the peptides modified with azo derivatives of phenylalanine, a different amino acid structure was designed. Removal of the methylene group from APhe₁ results in a phenylglycine derivative that was theorized to position the diazene closer to the side chain of hydrophobic amino acids in the ‘a’ and ‘d’ positions, and increase the effect of photoisomerization. Amino acid 4-(azophenyl)phenylglycine (APgly, Scheme 5.1) was chosen to test this theory due to its structural simplicity.
Synthesis and peptide incorporation of 4-(Azophenyl)phenylglycine

Synthesis of N-fmoc-4-(phenylazo)-l-phenylglycine (Fmoc-APgly, 10, Scheme 5.3) follows the same general route as the synthesis of azobenzene derivatives of phenylalanine; preparation of the para-substituted aniline followed by a Mills reaction with nitrosobenzene to yield the azobenzene. Preparation of the aniline requires an alternative method, since nitration of phenylglycine does not yield the para-nitrated product, and preparation of the 4-nitro variant via a Strecker synthesis.

Scheme 5.3: Synthetic scheme for the preparation of Fmoc-APgly from 4-hydroxyphenylglycine. Reagents and conditions: (i) MeOH, SOCl₂, 0 °C to RT, 95%; (ii) Boc₂O, TEA, MeCN, 93%; (iii) Tf₂O, TEA, DCM, -20 °C to RT, 98% (88% from 5); (iv) Benzophenone imine, rac-BINAP, Pd(OAc)₂, Cs₂CO₃, PhMe, 100 °C, 65%; (v) NH₄HCO₂, 10% Pd/C, MeOH, 60 °C; (vi) Nitrosobenzene, AcOH, 74% over 2 steps; (vii) LiOH, H₂O/THF; (viii) TFA, DCM; (ix) Fmoc-chloride, NaHCO₃, H₂O/THF, 63% over 3 steps.
synthesis is not effective. Therefore, amination of the respective aryl triflate (6) was deemed the most effective route, since the 4-hydroxy precursor is readily available. This route was tried and ultimately effective in the amination of 6 in sufficient quantities, from which 10 could be prepared following the procedure described below.

Methylation of 4-hydroxy-\(\text{L}\)-phenylglycine to form ester 4 was followed by amine protection with Boc anhydride to yield the double protected amino acid 5, followed by installation of the aryl triflate to form intermediate 6 in high yields. Initially the use of zinc trimethylsilylamide as an ammonia equivalent was attempted in the palladium-catalysed amination of 6, but the use of benzophenone imine proved much more effective, producing intermediate 7 in 65% yield. The imine was then reduced to yield aniline 8, followed directly by the Mills reaction to form azobenzene 9. Ester hydrolysis and Boc deprotection yielded APgly, which was Fmoc-protected to yield product 10. The final three steps could be performed in good (63%) overall yield, however attempts to simplify the procedure in the form of a one-pot procedure resulted in a sharp reduction in product yield. The overall yield of 10 from 4-hydroxy-\(\text{L}\)-phenylglycine is good (23% over 9 steps), yet the atom efficiency could be improved as most steps concern protecting group manipulation. Since only mildly basic conditions are used for the conversion of 5 to 9, for the larger scale synthesis of 10 the amination of Fmoc-protected phenylglycine-4-triflate would be worth investigating, as it would eliminate 4 synthetic steps.

As azobenzene photoswitches based on phenylglycine have not been reported, the photoswitching behaviour of Fmoc-APgly (10) was analysed, as displayed in Figure 5.3. High-pressure liquid chromatography (HPLC) separation of 10 dissolved in MeCN and kept under dark conditions, showed predominantly the trans (94.2 ± 0.2%) isomer, which switches to predominantly cis (91.3 ± 0.6%) when irradiated with 340 nm light (Figure 5.3C). The largest proportion of trans to cis isomers achieved via photoisomerization was after irradiation with 385 nm light (83.2 ± 0.2% trans), although higher wavelengths showed very similar distributions. No difference was observed in the proportions of isomers when 10 was directly illuminated with 435 nm light, or first isomerized to be cis-dominant with 340 nm light followed by illumination with 435 nm light, as is expected for the PSS. Relaxation after irradiation with 340 nm light was tracked over time (Figure 5.3B), and showed proportional decline of the cis isomer and increase in trans isomer,
Figure 5.3: HPLC traces of Fmoc-APgly, showing the chromatogram of the compound after irradiation with different wavelengths of light (A), thermal relaxation back from cis to trans (B), and the percentage of the trans and cis isomers in solution at the different photostationary states (C). All illuminations and relaxation studies were performed in MeCN at 20 °C.

Figure 5.4: UV-Vis (left) and CD (right) spectra of peptide K3-APGly during photoswitching. Spectra are recorded with samples that were dark adapted, or irradiated with light until no changes could be observed. Spectra were recorded at 20 °C in PBS with [peptide] = 50 µM.
with a half-life of 1690 minutes (≈ 28 h), comparable to reported half-lives for APhe1 (1600 min) and APhe2 (2100 min).\textsuperscript{14} Incorporation of 10 in K\textsubscript{3} was achieved using normal Fmoc-SPPS methods. UV-Vis Spectra of K\textsubscript{3}-APgly in PBS showed a strong absorbance band at 325 nm (Figure 5.4, $\varepsilon = 13,700$ M$^{-1}$ cm$^{-1}$) that disappeared upon irradiation with 340 nm light, demonstrating trans to cis isomerization, accompanied by an increase in the absorbance at 428 nm. Irradiation of the sample with 435 nm light resulted in 92% of the dark absorbance at 325 nm, showing effective cis to trans isomerization. The observed absorbance values in UV-Vis are in agreement with the isomeric ratios observed with HPLC (Figure 5.3A), demonstrating that APgly incorporation in peptides does not affect isomerization. CD spectra of K\textsubscript{3}-APgly showed a partially folded helix in the dark, that becomes less structured when irradiated with 340 nm light. Addition of binding partner E\textsubscript{3}GY led to a well-folded coiled-coil, that also showed a reduction in signal upon isomerization. The difference in folding between dark adapted and irradiated samples is much larger for peptide K\textsubscript{3} containing APgly, than it was for either of the phenylalanine derivatives, showing that the azobenzene is better positioned for coiled-coil photocontrol in the APgly derivative. The effect of isomerization on coiled-coil binding strength was again determined using thermal denaturation titration experiments (Figure S5.4), with fitting parameters shown in Table 5.2 and Table S5.2. The coiled-coil formed by K\textsubscript{3}-APgly and E\textsubscript{3}GY was determined to have a $K_\text{u}$ of 0.72 µM in the dark, which increased to 3.37 µM after photoisomerization. This equates to a 4.65-fold reduction in binding affinity for K\textsubscript{3}-APgly by isomerization of the azobenzene to the cis dominant state.

Table 5.2: Fit results of CD thermal unfolding curves from K\textsubscript{3}-APgly as a coiled-coil with E\textsubscript{3}GY. A complete list of fitting parameters and confidence intervals can be found in Table S5.2.

<table>
<thead>
<tr>
<th>Coiled-coil system</th>
<th>K\textsubscript{3}-APgly dark</th>
<th>K\textsubscript{3}-APgly 340nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta H^a$ (kJ mol$^{-1}$)</td>
<td>285.2</td>
<td>236</td>
</tr>
<tr>
<td>$T^b$ (°C)</td>
<td>113.0</td>
<td>110.5</td>
</tr>
<tr>
<td>$\Delta C_p$ (kJ mol$^{-1}$ k$^{-1}$)</td>
<td>2.80</td>
<td>2.15</td>
</tr>
<tr>
<td>$K_f$ (M$^{-1}$)$^c$</td>
<td>1.38 x 10$^6$</td>
<td>2.97 x 10$^5$</td>
</tr>
<tr>
<td>$K_u$ (µM)$^c$</td>
<td>0.72</td>
<td>3.37</td>
</tr>
<tr>
<td>Dark/340nm PSS</td>
<td>4.65</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ $\Delta H^o$ and $T^*$ are the enthalpy and the temperature where $\Delta G = 0$ and $K_u = K_f = 1$. $^b$ $\Delta C_p$ is the change in heat capacity upon unfolding. $^c$ Binding model at 20 (°C).
The difference between the dark and 340 nm adapted state is larger for $K_3$-APgly than was observed for either of the peptides that incorporated the phenylalanine derivatives, and the overall highest folding constant was also observed for this peptide. When comparing APgly isomerization to amino acid substitution at position ‘a’ studied by the group of Hodges, the change in binding energy is similar to substitution of alanine by a polar (Ser) or basic (Lys) side chain.

For the potential application of $K_3$-APgly in an active (fusion) system, stability of the photoswitch to repeated light exposures (photocycling) is an important characteristic. A solution of $K_3$-APgly was illuminated repeatedly first with 340 nm, followed by 435 nm light, and the switching monitored via the UV absorbance band at 325 nm (Figure 5.5 and Figure S5.5). The azobenzene consistently yielded the same absorption maxima during 6 cycles, demonstrating cycling of photostationary states with minimal deviation. Finally, the peptide was irradiated to the 340 nm PSS and allowed to thermally relax, showing 65% recovery in 69 hours. This thermal relaxation is 1.6 times slower ($t_{1/2} = 46h$) than observed for 10 in MeCN, which can be explained by the increased dielectric constant of the solvent.

**Molecular simulations**

In the previous experiments we have shown that incorporation of an azobenzene amino acid in the hydrophobic core of the $K_3/E_3$ coiled-coil yields more effective

![Figure 5.5: Photocycling of $K_3$-APgly between the dark adapted, 340 nm and 435 nm PSS (left), thermal relaxation of the peptide after thermal cycling (middle) and distribution of the absorbance in the different states during cycling (right). Absorbance at 325nm is plotted (full spectra are displayed in Figure S5.5), measurements were performed at 20 °C in PBS with [Peptide] = 50 µM.](image)
photoswitching when it is derived from phenylglycine then from phenylalanine. To understand how one methylene group affects azobenzene interactions, molecular models of $K_3$-$\text{ APgly}$ and $K_3$-$\text{ APhe1}$ were prepared. These models were based on the NMR structure of the $K_3/E_3$ coiled-coil,\textsuperscript{32} but with the Ile residue at position 9 replaced by a geometry optimized model of the synthetic amino acids in the $\text{cis}$ or $\text{trans}$ conformations. 500 ns MD simulations of these initial helical models were performed, with all 4 peptides showing a mostly stable helical coiled-coil structure during the simulations (Figure S5.6), although some unfolding at the termini was observed.

Snapshots from these MD simulations, displaying the azobenzene-amino acid and all side chains of peptide $E_3$ within 5 Å of that residue are shown in Figure 5.6. For both peptides, the azobenzene is shown to be close to the hydrophobic amino acids (Leu5, Ile9, Leu12) in the ‘a’ and ‘d’ positions, as well as adjacent amino acids (Ala4, Glu8) in the ‘g’ and ‘c’ positions. Both azobenzene amino acids extend outside of hydrophobic core, therefore shielding of the azobenzene outer phenyl group (as viewed from the backbone, and referred to as ‘Ring 2’) from the surrounding water by amino-acid side chains will have an impact on coiled-coil stability. The snapshot of $\text{cis} K_3$-$\text{ APgly}$ places the diazo group in the same position as in the $\text{trans}$ conformation, with ring 2 rotated towards peptide $K$. The opposite is observed in the snapshots of $K_3$-$\text{ APhe1}$, where the diazene group is in a different position relative to $K_3$ and ring 2 is still able to interact with amino acids in peptide $E_3$.

Because snapshots are not representative of the entire simulations, the average change in distance between the azobenzene and the amino acids of $E_3$ over all simulations was determined (Figure 5.7 and Figure S5.8). Distance changes to the diazene, ring 1 and ring 2 were plotted separately, since different interactions are expected for these groups. Both the diazene and ring 2 of $\text{ APhe1}$ show a negative distance change for amino acids close to the C-terminus, indicating closer positioning to those groups in the $\text{cis}$ conformation, which changes to positive values when moving through the sequence. A normalized change in distance $\leq -0.2$ or $\geq 0.2$ was deemed ‘significant’ in order to compare the different groups and photoswitches. More significant distance changes are observed for ring 2, which is expected since it is further away from the isomerized diazene, and therefore any rotation is amplified. All amino acids in the ‘a’, ‘c’, ‘d’ and ‘g’ positions of $E_3$ show significant changes in distance to $\text{ APhe1}$, with the other three positions also showing some significant changes. The apparent crossover point observed in the
Figure 5.6: Snapshot from MD simulations of the coiled coil between peptide K₃ (blue) and E₃ (green), with peptide K₃ containing the photoswitchable amino acid APhe1 or APgly in the trans or cis conformation. Peptides backbones are shown as a cartoon, with the photoswitch and all amino acid side chains of peptide E₃ within 5 Å displayed as sticks. Representative snapshots were chosen based on a similar RMSD from an ideal helix. Complete images are shown in Figure S5.7.

distance change graphs, combined with significant changes in distance upon APhe1 isomerization for nearly all amino acids suggests reorganization after isomerization, where the azobenzene can move to a different conformation to accommodate the cis isomer. This is supported by the distance change for ring 1 (Figure S5.8), which shows the same general distribution of distance change throughout the sequence.

Peptide K₃-APgly shows the largest change in distance for Leu side chains in the ‘d’ position of the heptad repeat sequence of E₃, and to the Glu side chain at position8. There are more distinct changes in distance to side chains in the ‘a’ and ‘c’ positions,
Figure 5.7: Normalized change in distance upon photoisomerization between the non-hydrogen atoms in peptide E₃ and the azobenzene diazene or ring 2 (left) during MD simulations, and helical wheel diagrams of peptide E₃ indicating the positions of the amino acids with significant changes in distance (right). Change in distance upon trans to cis isomerization of the azobenzene is normalized to the trans simulations, with positive values indicating more distance in the cis simulations. Distances are averaged over 3 simulations of 500 ns. Dotted lines indicate an arbitrary cut-off for significance of 0.2, with amino acids of interest marked light blue if their value falls below this line. Amino acids in the helical wheel diagram are coloured to indicate significant differences to the diazo group (blue), ring 2 (red), or both (purple), with amino acids of interest which fall below the cut-off also coloured light blue.

but these fall outside of the significance threshold that we use for comparison and have been marked with a lighter colour in Figure 5.7. Isomerization to cis is expected to result in repulsion of the diazene by hydrophobic side chains, which is
observed in the increased distance for Leu12 and Glu8, which are directly opposite the azobenzene in the coiled coil. After isomerization, ring 2 of APgly shows positioning close to amino acid side chains of K3 (Figure S5.9), which stabilizes the cis conformation. Overall, the photoswitch of K3-APgly shows changes in distance with the amino acids that are already closely positioned in the trans conformation, indicating the changes in distance can be attributed to the rotation around the diazene bond after isomerization.

If we generalize the structures in the MD snapshots with the changes in distance shown in Figure 5.7, the effect of isomerization on the coiled-coil folding constant observed in the previous section can be explained. Upon trans to cis isomerization, the diazene group becomes more polar resulting in unfavourable interactions if it is positioned in the hydrophobic core, and with both azobenzene amino acids, a general increase in distance to the diazene is observed. APhe1 has two bonds between Cα and the azobenzene moiety, which allows for more degrees of freedom to reposition the azobenzene after isomerization compared to APgly, which can only change position via a single rotation and deviation from the optimal bond angles. The extra rotational freedom in APhe1 results in a decreased effect of the diazene polarity on coiled-coil stability through repositioning of the azobenzene. This repositioning also allows ring 2 to keep interacting with (different) hydrophobic side chains from peptide E3, which is not possible for K3-APgly. Both of these factors decrease the differences in binding strength between the two isomeric states of APhe1, and explain why APgly shows the largest difference in coiled-coil binding upon isomerization.

5.3 Conclusions

Two azobenzene derivatives of phenylalanine, APhe1 and APhe2, were prepared and incorporated into the hydrophobic core of Peptide K3. Both peptides showed good photoswitching behaviour, but azobenzene isomerization had only a moderate effect on coiled-coil interactions with binding partner E3GY. The 340 nm irradiated (cis dominant) state of K3-APhe1 showed a 1.3-fold decrease in $K_f$ compared to the dark (trans dominant) state. Peptide K3-APhe2 showed the opposite behaviour, increasing the $K_f$ by 1.5-fold, which can be attributed to the repulsion between the glutamic acid side chains and the carboxylic acid group of APhe2. A novel azobenzene amino acid, APgly, was synthesized based on L-phenylglycine and showed effective photoswitching between the cis and trans states as the Fmoc-protected amino acid in MeCN. When incorporated in peptide
K₃, APgly also showed stability to photocycling under aqueous conditions, and exhibited a 1.6-fold slower relaxation from the 340 nm PSS. CD experiments demonstrated a difference in folding between the dark and irradiated states, and thermal melting experiments revealed a coiled-coil K₃ that was reduced 4.65-fold after isomerization with 340 nm light. Combined with the highest overall Kᵢ in the dark state, APgly is the superior amino acid for coiled-coil photoswitching compared to both phenylalanine derivatives.

To better understand the interactions of the azobenzene amino acid in the coiled coil, MD simulations of K₃-APhe1 and K₃-APgly with binding partner E₃ were performed. These simulations showed more distance changes between the cis and trans isomers of APhe1 than was observed for APgly, indicative of a rearrangement of the amino acid after isomerization. Positioning of the APgly diazene group close to the centre of the hydrophobic coiled-coil interface, combined with less rearrangement of the azobenzene after isomerization makes APgly the more effective photoswitch for functional coiled-coil motifs and assemblies.

5.4 Methods

Fmoc-protected amino acids and fmoc chloride were purchased from Novabiochem (Amsterdam, The Netherlands). Acetic anhydride (Ac₂O), acetonitrile (MeCN), dimethylformamide (DMF), piperidine, pyridine, NaHCO₃, trifluoroacetic acid (TFA) and tetrahydrofuran (THF) were purchased from Biosolve (Valkenswaard, The Netherlands). Oxyma pure was purchased from Carl Roth (Karlsruhe, Germany). Acetic acid (AcOH), ammonia, ammonium chloride, ammonium formate, benzophenone imine, 2,2′-Bis(diphenylphosphino)-1,1′-binaphthalene (rac-BINAP), 1,2-bis(2-mercapto-ethoxy)-ethane (EDDT), cesium carbonate, di-tert-butyl decarbonate (Boc₂O), fmoc-4-nitro-L-phenylalanine, 4-hydroxy-L-phenylglycine, lithium hydroxide, nitrosobenzene, N,N′-diisopropylcarbodiimide (DIC), oxone, palladium acetate, palladium on carbon (Pd/C, 10%), L-phenylalanine, tert-butyl 4-aminobenzoate, thionyl chloride, triethylamine (TEA), trifluoromethanesulfonic anhydride (Tf₂O) and zinc powder were purchased from Sigma Aldrich (Zwijndrecht, The Netherlands). Chloroform, dichloromethane (DCM), diethyl ether (Et₂O), ethyl acetate (EtOAc), ethanol (EtOH), methanol (MeOH), pentane, petroleum ether (PE), sodium sulfate and toluene were supplied by Honeywell (Meppel, The Netherlands). All reagents were used as purchased. Ultrapure water was purified using a Milli-Q™ purification system from Millipore (Amsterdam, The Netherlands).
**Peptide synthesis** was performed via Fmoc-based SPPS, on a CEM Liberty Blue microwave-accelerated peptide synthesizer. Peptides were prepared on a 0.1 mmol scale using Tentagel S RAM resin (0.22 mmol/g). 5 equivalents each of amino acid, Oxyma pure and DIC, the latter two of which were coupling reagents were heated at 90 °C for 4 minutes to facilitate coupling. Deprotection was achieved with 20% piperidine in DMF heated to 90 °C for 1 minute. Between deprotection and peptide coupling 3 DMF washes were performed, with a single washing step between the coupling and deprotection steps. Azobenzene amino-acids 2, 3 and 10 were coupled manually, using 2.5 equivalents of amino acid, 2.5 equivalents of HATU and 5 equivalents DIPEA in DMF for 2-3 hours. After the last Fmoc deprotection, the N-terminus was acylated using 5 mL/mmol each of Ac₂O and pyridine in DMF, for 5 minutes. The resin was washed 3 times with DMF, MeOH and DCM followed by air-drying. Cleavage of peptide was achieved performed with TFA (8 mL) containing 2.5% water and 2.5% TIS for 1 hour, followed by precipitation of the product in Et₂O. The product was collected via centrifugation (4000 rpm, 10 minutes), the organic layer removed and the product resuspended in water for direct purification or lyophilization.

Peptides were purified using reverse-phase HPLC on a Shimadzu system consisting of two KC-20AR pumps and an SPD-20A or SPD-M20A detector fitted with a 21.2x150 mm Phenomenex Kinetix Evo C18 column. A nonlinear gradient from 10-90% MeCN in water was used, with 0.1% TFA as the ion-pair reagent, at a flow rate of 12 ml/min. Collected fractions were checked via analytical HPLC, pooled and lyophilized twice to yield the dry products. MS characterization of purified peptides can be found in Table S5.4.

**Analytical methods**

**LC-MS analysis** was performed on a Thermo Scientific TSQ quantum access MAX mass detector connected to an Ultimate 3000 liquid chromatography system fitted with a 50x4.6 mm Phenomenex Gemini 3 μm C18 column. LC-MS spectra were recorded using a linear gradient of 10%-90% MeCN in H₂O + 0.1% TFA.

**Analytical HPLC** was performed using a Shimadzu Prominence-i LC-2030C 3D system fitted with a 4.6x50 mm Phenomenex Kinetix Evo C18 column. The isomeric ratio of 10 was quantified using a linear gradient of 10-90% MeCN in water containing 0.1% TFA as buffer. For each measurement, 2 μL of a 200 mM solution of 10 was injected, and measurements were repeated 3 times for accuracy.
**UV-Vis** spectra were measured on an Agilent Cary-300 spectrophotometer fitted with an Agilent temperature controller. Spectra were measured at 20 °C in a 1 cm quartz low-volume cuvette, using a scanning speed of 200 nm/min. Spectra were baseline corrected using a blank measurement with the same solvent used for sample preparation.

**CD** spectra were recorded on a Jasco J-815 CD spectrometer fitted with a Peltier temperature controller. Spectra were recorded in a 2 mm quartz cuvette at 20 °C using either PBS or low-salt buffer (2 mM NaCl, 10 mM phosphate) at pH 7.4. Spectra were recorded between 190 and 280 nm with 1 nm intervals, at a scan rate of 100 nm/min with 5 subsequent spectra averaged to minimize noise. The mean residue molar ellipticity (θ, deg cm² dmol.res⁻¹), was calculated using equation 1:

\[
[\theta] = \left(\frac{100 \times [\theta]_{obs}}{c \times n \times l}\right) 
\]

Where \([\theta]_{obs}\) represents the observed ellipticity in mdeg, \(c\) represents the peptide concentration in mM, \(n\) the number of peptide bonds and \(l\) the path length of the cuvette in cm. Thermal melting curves were generated by recording θ₂₂₂ₙₙ between 5 and 90 °C, at a speed of 1 °C/min. If irradiated samples were used, samples were reilluminated every 30 minutes. Melting curves were measured at 4 different concentrations and fitted using the Fitdis software package.²⁷

**Sample illumination** at 340 nm was achieved using a Thorlabs M340F3 Fiber-coupled LED powered by a T-Cube driver at 1000 mA. For all other wavelengths, high-power single chip LEDs were purchased from Roithner Laser (Vienna, Austria) from the H2A1 series. LEDs from Roithner Laser were mounted on an aluminum back plate for heat dissipation, and powered at 350 mA current using a driver built in-house. For illumination the LED was placed parallel to the side of the cuvette at a distance of 5 mm, and centered to the width and height of the sample.

**Molecular simulations**

**Geometry Optimization:** Input geometries of the amino acids for the MD simulations were obtained by geometry optimization performed using the Amsterdam Density Functional (ADF2019.302)³³⁻³⁵ software package, using the BLYP functional with D3(BJ) dispersion correction,³⁶,³⁷ and a TZ2P basis set, which is of triple-ζ quality for all atoms and has been improved by two sets of polarization functions. The accuracies of the fit scheme (Zlm fit) and the integration grid (Becke
grid) were set to VERYGOOD. All structures were verified to be at the stationary point by the absence of negative frequencies.

**Force field parametrization:** A new atom type NX was introduced to the AMBER-96 force field\textsuperscript{38} for the diazo nitrogen atoms in APgly and APhe1. Bond lengths, angles and dihedral angles for this new atom type were derived from DFT calculated optimized geometries, and are displayed in table S5.3. Force constants from similar existing atom type combinations were adapted. For the production runs, a high force constant (1000 kJ mol\(^{-1}\) rad\(^{-2}\)) was used to restrain the cis/trans dihedral angle.

**System setup:** The NMR structure of the E\textsubscript{3}/K\textsubscript{3} coiled-coil heterodimer was retrieved from the protein data bank (PDB, ID: 1U0I).\textsuperscript{32} Ile9 of peptide K was mutared manually to APgly or APhe1 and solvated with TIP3P water\textsuperscript{39} in a 5x5x5 nm\(^3\) simulation box. After steepest descent energy minimization, the system was equilibrated for 500 ps with and, subsequently, without position restraints on the protein atoms. For each system, three independent production runs of 500 ns were performed with pseudorandom initial velocities.

**Simulation details:** All molecular dynamics (MD) simulations were performed with GROMACS 2019.3.\textsuperscript{40} A 1 fs time step was used. Constant temperature (300K, \(\tau_T = 0.1\) ps) and pressure (1 bar, \(\tau_P = 2\) ps) were maintained by the velocity rescaling thermostat\textsuperscript{41} and the Berendsen barostat,\textsuperscript{42} respectively. The system’s compressibility was set to 4.5\(\cdot\)10\(^{-5}\) bar\(^{-1}\). Neighbour lists were recalculated every 100 steps with a cut-off of 1 nm using the Verlet cut-off scheme.\textsuperscript{43} Particle-mesh Ewald (PME)\textsuperscript{44} electrostatics (0.11 nm grid) and Van der Waals interactions are shifted such that they switch off at the cut-off distance (1 nm).

*Organic Synthesis*

**Protocol A, general method for the reduction of nitrophenylalanine:** Fmoc-4-nitro-L-phenylalanine was dissolved in absolute EtOH (150 mL/g), combined with ammonium chloride (5 eq.) and Zn dust (4 eq.), and the reaction refluxed for 2 hours. After evaporation of solvent, the resulting solids were combined with EtOAc and excess 1 M HCl, the phases separated and the aqueous layer extracted twice with ethyl acetate. The two organic layers were combined, washed with deionized water, and dried with anhydrous Na\textsubscript{2}SO\textsubscript{4}. Solvent was evaporated under reduced pressure, yielding Fmoc-(4-amino)-L-phenylalanine that was used without further purification.
**4-tert-Butyl-nitrosobenzoate**: 4-Amino tert-butyl benzoate (765 mg, 3.96 mmol) was dissolved in 10 mL DCM and combined with Oxone (1263 mg, 8.29 mmol) in 13 mL H$_2$O, and the mixture was stirred vigorously under reflux for 20 hours. The layers were separated, and the aqueous layer extracted with an additional 20 mL DCM. The combined organic layers were washed sequentially with 10 mL 1M HCl, 10 mL half-saturated NaHCO$_3$ and 10 mL brine, dried over Na$_2$SO$_4$, and solvent was removed to yield the crude product. The product was purified via column chromatography (Et$_2$O/C$_5$H$_12$) to yield 220 mg of the product as bright green needles (1.06 mmol, 27%) that were used directly in the Mills reaction.

$^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 8.23 (d, $J = 8.7$ Hz, 2H), 7.91 (d, 2H), 1.62 (s, 9H).

$^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 164.78, 164.47, 137.28, 130.94, 130.65, 123.51, 120.47, 82.53, 28.25.

**N-Fmoc-4-nitro-l-phenylalanine (1):** To a solution of l-phenylalanine (4.96 g, 30 mmol) dissolved in H$_2$SO$_4$ (95%, 22.5 mL) and cooled over ice was added 4.2 mL nitrating solution (prepared by mixing 2.8 mL of 60% HNO$_3$ and 2.2 mL 95% H$_2$SO$_4$ on ice) dropwise over 3.5 hours. The solution was neutralized with ammonia solution (25%) added dropwise until pH 6 was achieved and precipitate observed. The precipitate was collected by filtration and washed with 5 x 30 mL water. After drying, 3.5 g (55%) of 4-nitrophenylalanine was collected as an off-white solid. This intermediate (1.05 g, 5.00 mmol) was dissolved in a mixture of 20 mL 0.5 M Na$_2$CO$_3$, 13 mL acetone, 13 mL deionized water and 49.6 mg dodecyl sulfate and cooled over ice. A solution of Fmoc-chloride (1.32 g, 5.1 mmol) in 10 mL acetone was added dropwise to the reaction, and it was left to stir overnight at room temperature. The reaction was quenched by dilution into 300 mL cold water and addition of 5 mL, 2 M HCl. The precipitate was filtered and redissolved in 100 mL, 0.5 M Na$_2$CO$_3$. The mixture was allowed to stir at 70 °C for 1 hour, after which the white precipitate was filtered out and the filtrate was collected and acidified to pH < 3 with 2 M HCl. The precipitate was collected and dried to yield 1.5 g of a white solid (69%, 38% over two steps). $^1$H NMR (400 MHz, DMSO-d6) 8.15 (d, 2H), 7.89 (d, $J = 7.6$ Hz, 2H), 7.80 (d, $J = 8.7$ Hz, 1H), 7.62 (dd, $J = 7.5$, 2.7 Hz, 2H), 7.56 (d, $J = 8.5$ Hz, 2H), 7.41 (t, $J = 7.4$ Hz, 2H), 7.30 (q, $J = 7.4$ Hz, 2H), 4.32 – 4.13 (m, 4H), 3.25 (dd, $J = 13.8$, 4.4 Hz, 1H), 3.02 (dd, $J = 13.8$, 10.8 Hz, 1H). LC-MS RT = 7.57 min, m/z= 178.47 (calc Fm$^+$ = 179.09), 454.61 (calc [1 + Na$^+$]$^+$ = 455.12), 470.00 (calc [1 + K$^+$]$^+$ = 471.10).

**Preparation of N-fmoc-(4-phenylazo)-l-phenylalanine, APhe1 (2):** Fmoc-4-nitrophenylalanine (1.34 g, 3.11 mmol) was reduced according to protocol A (see above), resulting in 1.04 g crude Fmoc-4-aminophenylalanine. The intermediate
was dissolved in 60 mL AcOH, before nitrosobenzene (350 mg, 3.27 mmol, 1.25 eq.)
dissolved in 5 mL AcOH was added and the solution left for 20 hours. The product
was precipitated by addition of the reaction mixture to 250 mL H$_2$O, and collected
via filtration. The crude product was redissolved in EtOAc, dried with Na$_2$SO$_4$,
filtered and solvent evaporated. The solids were purified via column chromatography (dry
loading, DCM/MeOH + 0.5% AcOH) and removal of solvents yielded 478 mg (0.97 mmol, 32% over 2 steps) of pure product as a brown powder.

$^1$H NMR (500 MHz, DMSO-d$_6$) 7.89 - 7.85 (m, 4H), 7.82 (d, J = 8.3 Hz, 3H), 7.66 - 7.54 (m,5H), 7.50 (d, J = 8.4 Hz, 2H), 7.38 (q, J = 7.9 Hz, 2H), 7.28 (dt, J = 14.9, 7.4 Hz, 2H),
4.30 - 4.24 (m,1H), 4.21 (d, J = 6.1 Hz, 2H), 4.19 - 4.13 (m, 1H), 3.21 (dd, J = 13.8, 4.4 Hz, 1H), 2.99 (dd, J = 13.8,10.7 Hz, 1H).

$^{13}$C NMR (126 MHz, DMSO) 173.14, 155.96, 151.99, 150.65, 143.77, 143.71, 142.06, 140.69,140.68, 131.39, 130.24, 129.48, 127.61, 127.59, 127.05, 125.25, 125.19, 122.47, 120.10, 65.61, 55.22, 46.57, 36.32.

**Preparation of N-fmoc-L-(4-(4'-tert-Butoxycarbonyl)phenylazo)phenylalanine, (3):**
Fmoc-(4-nitro)-L-phenylalanine (2.02 mmol, 874 mg) was reduced according to
general protocol A to yield 732 mg of Fmoc-(4-amino)-L-phenylalanine. The solids
were dissolved by heating in 70 mL AcOH, and after it had returned to room
temperature, 560 mg tert-Butyl-4-nitrosobenzoate was added (2.71 mmol)
dissolved in 5 mL AcOH, and the reaction stirred for 20 h. Another 220 mg (1.06 mmol) of the
nitrosobenzene was added to drive the reaction to completion, and
the reaction stirred for another 20 h. The product was precipitated in 250 mL H$_2$O,
collected via filtration, redissolved in acetone and the solvent was removed via rotary
evaporation. The crude product was purified via column chromatography (PE/Et$_2$O + 1% AcOH)
and solvent removed to yield 675 mg (1.14 mmol, 56% over two steps) of the pure product as a red semi-crystalline material.

$^1$H NMR (400 MHz, CDCl$_3$): δ 8.12 (d, J = 8.5 Hz, 2H), 7.89 (dd, J = 11.5, 8.3 Hz, 4H), 7.76 (d, J = 7.6 Hz, 2H), 7.55 (d, J = 7.2 Hz, 2H), 7.39 (t, J = 7.4 Hz, 2H), 7.33 – 7.28 (m, 4H), 5.30 (d, J = 8.1 Hz, 1H), 4.83 – 4.71 (m, 1H), 4.51 – 4.47 (m, 1H), 4.39 (dd, J = 10.7, 6.8 Hz, 1H), 4.21 (t, J = 6.9 Hz, 1H), 3.26 (ddd, J = 45.1, 13.9, 5.8 Hz, 2H), 1.63 (s, 9H).

$^{13}$C NMR (101 MHz, CDCl$_3$): δ 177.31, 165.40, 155.83, 154.93, 151.86, 143.84, 143.70, 141.47, 139.65, 133.84, 130.56, 130.38, 127.90, 127.23, 125.17, 125.11, 123.52, 122.63, 120.16, 81.71, 67.20, 54.58, 47.25, 37.82, 28.33.

**LC-MS RT = 8.79 (cis), 9.81 (trans) min, m/z= 178.47 (calc Fm$^+$ = 179.09), 513.31 (calc [2 + Na$^+$]$^+$ = 514.17).**
4-Hydroxy-L-phenylglycine methyl ester (4): Thionyl chloride (13.0 mL, 184 mmol, 9 eq.) was added dropwise to a suspension of 4-hydroxy-L-phenylglycine (3.39 g, 20.3 mmol) in 100 mL dry MeOH temperature controlled at 20 °C, over the course of 30 minutes. The clear solution was stirred overnight, and all liquid was evaporated under reduced pressure. The resulting oil was mixed with 50 mL Et₂O to yield an off-white precipitate, which was filtered and washed with 2x 50 mL Et₂O. The residue was dried under vacuum to yield 4.33 g of 4-hydroxy-phenylglycine methyl ester hydrochloride (19.9 mmol, 98%).

\[ ^{1}H \text{ NMR} (400 MHz, DMSO): \delta \text{ 9.98 (s, 1H), 8.98 (s, 3H), 7.28 (d, } J = 8.6 \text{ Hz, 2H), 6.84 (d, } J = 8.6 \text{ Hz, 2H), 5.09 (s, 1H), 3.69 (s, 3H), 3.38 (s, 1H). } \]

\[ ^{13}C \text{ NMR} (101 MHz, DMSO): \delta \text{ 169.26, 158.60, 129.64, 122.61, 115.70, 54.90, 53.02. } \]

N-Boc-4-hydroxy-L-phenylglycine methyl ester (5): A solution was prepared of 4-hydroxy-phenylglycine methyl ester (4, 1.11 g of the HCl salt, 5.11 mmol) in dry MeCN (70 mL) containing TEA (0.75 mL 5.39 mmol, 1.06 eq). Boc₂O (1.30 g, 5.96 mmol, 1.17 eq) was dissolved in 20 mL dry MeCN and added dropwise to the phenylglycine solution, and the solution stirred overnight. The solvent was evaporated and remaining solids redissolved in 50 mL DCM, which was subsequently washed with 25 mL 1M H₃PO₄, 50 mL H₂O and 25 mL brine. The organic layer was dried with Na₂SO₄ and the solvent evaporated. The crude product was purified via filtration through a plug of silica (eluent: EtOAc) and the solvent was removed to yield 1.33 g (4.71 mmol, 93%) of product as a white solid.

\[ ^{1}H \text{ NMR} (400 MHz, DMSO): \delta \text{ 9.50 (s, 1H), 7.62 (d, } J = 7.9 \text{ Hz, 1H), 7.20 – 7.13 (m, 2H), 6.75 – 6.65 (m, 2H), 5.05 (d, } J = 7.9 \text{ Hz, 1H), 3.59 (s, 3H), 1.38 (s, 9H). } \]

\[ ^{13}C \text{ NMR} (101 MHz, DMSO): \delta \text{ 171.91, 157.25, 155.20, 129.10, 126.64, 115.19, 78.42, 57.10, 52.00, 28.19. } \]

N-Boc-4-(trifluoromethanesulfonate)-L-phenylglycine methyl ester (6): N-boc-4-hydroxy-L-phenylglycine methyl ester (1.41 g, 5.02 mmol) was dissolved in 25 mL DCM and combined with 1 mL (7.5 mmol, 1.5 eq.) TEA. The reaction was cooled to -20 °C and Tf₂O (0.83 mL, 4.92 mmol, 0.98 eq.) was added in portions. After 30 minutes, the reaction was allowed to heat to room temperature and left to stir for 4 hours. The reaction was diluted with another 20 mL DCM and washed with 20 mL 0.5M HCl, 20 mL H₂O and 20 mL brine. The organic layer was dried over Na₂SO₄, the solvent was removed and the product purified over a silica column (Et₂O/C₅H₁₂) to yield 1.825 g of a clear oil (4.42 mmol, 98% yield relative to Tf₂O, 88% relative to 5). The oil could be turned into a solid via dissolution in pentane and removal of the
solvent. \textsuperscript{1}H NMR (500 MHz, CDCl\textsubscript{3}): \(\delta\) 7.47 (d, \(J = 8.7\) Hz, 2H), 7.30 – 7.23 (m, 2H), 5.77 – 5.61 (m, 1H), 5.37 (d, \(J = 7.2\) Hz, 1H), 3.75 (s, 3H), 1.43 (s, 8H). \textsuperscript{13}C NMR (126 MHz, CDCl\textsubscript{3}): \(\delta\) 170.93, 154.76, 149.45, 137.93, 129.17, 121.92, 118.25 (d, \(J_{CF} = 161\) Hz), 80.72, 56.89, 53.21, 28.40, 27.80. \textsuperscript{19}F NMR (471 MHz, CDCl\textsubscript{3}): \(\delta\) -72.86.

**N-Boc-4-((diphenylmethylene)-amino)-\(\text{\textalpha}\)-phenylglycine methyl ester (7):** Cs\textsubscript{2}CO\textsubscript{3} (460 mg, 1.4 mmol, 1.4 eq), N-Boc-[(diphenylmethyl)-amino]-\(\text{\textalpha}\)-phenylglycine methyl ester (6, 413 mg, 1.03 mmol) and benzophenone imine (210 \(\mu\)L, 1.25 mmol, 1.21 eq) were placed in an oven-dried shlenck reaction vessel, and the flask was placed under a nitrogen atmosphere. Pd(OAc)\textsubscript{2} (15.5 mg, 0.07 mmol, 7\%) and rac-BINAP (95 mg, 0.15 mmol, 15\%) were combined in a separate vial, placed under nitrogen, and combined with 4 mL toluene and 2.5 \(\mu\)L H\textsubscript{2}O (0.14 mmol, 14\%) and heated to 85 \(^\circ\)C. After 2 minutes the color of the catalyst solution had changed to bright red, transferred to the reaction flask and the reaction stirred at 100 \(^\circ\)C for 20h. After cooling to RT, the liquid phase of the reaction was transferred to a separatory funnel, the solids washed 3x with 10 mL Et\textsubscript{2}O, and this liquid also combined in the funnel. The organic layers were extracted with 2x 20 mL H\textsubscript{2}O, dried with Na\textsubscript{2}SO\textsubscript{4} and the solvent removed. The product purified over silica column (Et\textsubscript{2}O/C\textsubscript{5}H\textsubscript{12}) and dried under high vacuum to yield 296 mg of a light-yellow solid (0.66 mmol, 65\% yield). \textsuperscript{1}H NMR (500 MHz, CDCl\textsubscript{3}): \(\delta\) 7.72 (d, \(J = 7.0\) Hz, 2H), 7.46 (t, \(J = 7.3\) Hz, 1H), 7.39 (dd, \(J = 8.3, 6.8\) Hz, 2H), 7.25 (d, \(J = 8.1\) Hz, 4H), 7.14 – 7.07 (m, 4H), 6.69 (d, \(J = 8.4\) Hz, 2H), 5.39 (d, \(J = 7.5\) Hz, 1H), 5.19 (d, \(J = 7.5\) Hz, 1H), 3.68 (s, 3H), 1.43 (s, 8H). \textsuperscript{13}C NMR (126 MHz, CDCl\textsubscript{3}): \(\delta\) 171.96, 168.66, 154.95, 151.54, 139.62, 136.07, 131.34, 130.97, 129.61, 129.47, 128.83, 128.34, 128.09, 127.56, 121.53, 80.21, 57.33, 52.65, 28.42.

**N-Boc-4-(4-phenylazo)-\(\text{\textalpha}\)-phenylglycine methyl ester (8):** Ammonium formate (1.1 g, 17.5 mmol, 15.8 eq.) and 10\% Pd on carbon (50\% water by weight, 125 mg, 5.8\%) were combined with N-Boc-4-((diphenylmethylene)amino)-\(\text{\textalpha}\)-phenylglycine methyl ester (7, 496 mg, 1.11 mmol) under a nitrogen atmosphere. 4 mL MeOH was added and the reaction heated to 60 \(^\circ\)C for 2h. The reaction was diluted with 10 mL DCM and filtered over celite, then washed with another 20 mL DCM. The organic layers were washed with 2x 30 mL H\textsubscript{2}O, 1 x 10 mL brine, and dried over Na\textsubscript{2}SO\textsubscript{4}. The solvent was removed to yield 400 mg of an off-white solid. This was redissolved in 5 mL AcOH, and 167 mg (1.56 mmol, 1.4 eq) nitrosobenzene in 3 mL AcOH was added and the reaction stirred for 6 days. 50 mL DCM was added and the organic layer washed with 50 mL H\textsubscript{2}O, 50 mL 3\% NH\textsubscript{3} and 50 mL brine. The organic layer was dried
over dried over Na₂SO₄ and the solvent removed. The crude product was purified via column chromatography (Et₂O/C₅H₁₂) to yield 303 mg of an orange solid (0.82 mmol, 74%). ¹H NMR (400 MHz, CDCl₃): δ 7.94 – 7.88 (m, 4H), 7.52 (m, 5H), 5.68 (d, J = 7.3 Hz, 1H), 5.41 (d, J = 7.3 Hz, 1H), 3.74 (s, 3H), 1.44 (s, 8H). ¹³C NMR (101 MHz, CDCl₃): δ 171.29, 152.73, 152.66, 139.84, 131.33, 129.25, 128.05, 123.45, 123.05, 80.51, 57.48, 53.06, 28.45.

**N-fmoc-4-(phenylazo)-L-phenylglycine, Fmoc-APhe (10):** To N-Boc-4-(phenylazo)-L-phenylglycine methyl ester (9, 311 mg, 0.84 mmol) dissolved in 6 mL THF was added 1M LiOH (1.26 mL, 1.5 eq.) and the reaction stirred for 90 minutes. After cooling the reaction was combined with 50 mL DCM and 50 mL 0.5 M HCl. The layers were separated, and the organic layer washed with 50 mL H₂O, 50 mL brine and dried over Na₂SO₄. The solvent was evaporated and the intermediate was redissolved in 5 mL DCM and 5 mL TFA was added. The reaction was stirred for 2 hours, after which the intermediate was precipitated in 100 mL cold Et₂O: C₅H₁₂ and collected via centrifugation at 4000 rpm for 10 minutes. The precipitate was dried under a stream of air for 5 minutes, then suspended in 150 mL 25% H₂O/THF. NaHCO₃ (1.5 g) was added to neutralize the solution and Fmoc-chloride (258 mg, 1 mmol, 1.19 eq.) dissolved in 6 mL THF was added to the reaction dropwise. The solution was stirred overnight, and the next day all THF was evaporated. The reaction was combined with 60 mL DCM and 100 mL 1M HCl and separated. The Aqueous layer was extracted with 40 mL DCM, organic layers combined and washed with 100 mL H₂O, 50 mL brine and dried over Na₂SO₄. The solvent was removed and the remaining solids purified via column chromatography (DCM/MeOH with 0.5 % AcOH) to yield 254 mg of product as a bright orange powder (0.3 mmol, 63% over 3 steps). ¹H NMR (600 MHz, DMSO): δ 13.11 (s, 1H), 8.44 – 8.31 (m, 1H), 7.93 – 7.87 (m, 6H), 7.77 (d, J = 7.5 Hz, 2H), 7.66 (d, J = 8.5 Hz, 2H), 7.64 – 7.55 (m, 3H), 7.44 – 7.39 (m, 2H), 7.32 (dt, J = 11.7, 7.5 Hz, 2H), 5.32 (d, J = 8.2 Hz, 1H), 4.34 – 4.22 (m, 3H). ¹³C NMR (151 MHz, DMSO): δ 171.68, 155.91, 151.92, 151.50, 143.84, 143.77, 140.74, 140.66, 131.71, 129.55, 128.99, 127.70, 127.11, 125.42, 125.42, 122.64, 120.15, 66.01, 57.78, 46.62. **LC-MS** RT = 6.65 (cis), 8.80 (trans) min, m/z= 178.54 (calc Fm⁺ = 179.09), 238.71 (calc [10 + 2H⁺]²⁺ = 239.59), 477.28 (calc [10 + H⁺]⁺ = 478.17), 499.46 (calc [10 + Na⁺]⁺ = 500.16).
5.5 References


Supporting information for Chapter 5

**Figure S5.1:** UV-Vis spectra of peptide K₃ containing either APhe1 or APhe2 at different time intervals during illumination with 340 nm light. Measurements were performed at 20 °C in PBS with [peptide] = 50 µM. Arrows indicate direction of change after illumination.
Figure S5.2: Fitting of CD melting curves (top) for peptide K3-APhe1 with E3GY in the dark adapted (left) and 340 nm irradiated (right) states performed using the Fitdis software program. Residuals (middle) and derivatives (bottom) of the best fitting (dimeric) model are shown. Measurements were performed in PBS.
Figure S5.3: Fitting of CD melting curves (top) for peptide K₃-APhe2 with E₃GY in the dark adapted (left) and 340 nm irradiated (right) states performed by the Fitdis software program. Residuals (middle) and derivatives (bottom) of the best fitting (dimeric) model are shown. Measurements were performed in PBS.
Table S5.1: Fit results of thermal unfolding curves for K₃ peptides containing azophenylalanine amino acids with binding partner E₃GY, in either the dark adapted or 340 nm irradiated state.

<table>
<thead>
<tr>
<th>Coiled-coil system</th>
<th>K₃-APhe1 Dark</th>
<th>K₃-APhe1 340 nm</th>
<th>K₃-APhe2 dark</th>
<th>K₃-APhe2 340 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH° (kJ mol⁻¹)</td>
<td>224.7 ± 12.4</td>
<td>261.4 ± 7.4</td>
<td>175.0 ± 12.1</td>
<td>259.0 ± 9.7</td>
</tr>
<tr>
<td>T° (°C)</td>
<td>123 ± 3.1</td>
<td>114.1 ± 1.5</td>
<td>142.5 ± 5.4</td>
<td>102.1 ± 2.1</td>
</tr>
<tr>
<td>ΔCₚ (kJ mol⁻¹ K⁻¹)</td>
<td>1.69 ± 0.22</td>
<td>2.46 ± 0.15</td>
<td>1.07 ± 0.17</td>
<td>2.65 ±0.22</td>
</tr>
<tr>
<td>θₜ (deg cm² dmol⁻¹)</td>
<td>-22,067 ± 413</td>
<td>-23,009 ± 358</td>
<td>-21,790 ± 548</td>
<td>-21,112 ± 447</td>
</tr>
<tr>
<td>mₜ</td>
<td>61.3 ± 16.7</td>
<td>59.9 ± 14.03</td>
<td>87.2 ± 27.1</td>
<td>99.1 ± 20.6</td>
</tr>
<tr>
<td>θᵤ (deg cm² dmol⁻¹)</td>
<td>-3,566 ± 1062</td>
<td>-2,830 ± 876</td>
<td>-2,986 ± 887</td>
<td>-4,078 ± 545</td>
</tr>
<tr>
<td>mᵤ</td>
<td>-7.4 ± 13.0</td>
<td>-15.0 ± 10.7</td>
<td>-5.0 ± 11.0</td>
<td>0.0 ± 7.0</td>
</tr>
<tr>
<td>ΔG₂₀ (kJ mol⁻¹)</td>
<td>33.5</td>
<td>32.8</td>
<td>29.97</td>
<td>30.94</td>
</tr>
<tr>
<td>ΔS₂₀ (J mol⁻¹ K⁻¹)</td>
<td>57</td>
<td>-8.95</td>
<td>46.31</td>
<td>36.7</td>
</tr>
<tr>
<td>ΔH₂₀ (kJ mol⁻¹)</td>
<td>50.23</td>
<td>30.21</td>
<td>43.54</td>
<td>41.70</td>
</tr>
<tr>
<td>Kₙ (M⁻¹, 20 °C)</td>
<td>9.19 10⁵</td>
<td>7.09 10⁵</td>
<td>2.18 10⁵</td>
<td>3.26 10⁵</td>
</tr>
<tr>
<td>Kᵤ (µM, 20 °C)</td>
<td>1.09</td>
<td>1.41</td>
<td>4.58</td>
<td>3.07</td>
</tr>
<tr>
<td>Dark/340nm PSS</td>
<td>1.30</td>
<td></td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.997</td>
<td>0.998</td>
<td>0.996</td>
<td>0.997</td>
</tr>
<tr>
<td>RMSE</td>
<td>300.1</td>
<td>248.2</td>
<td>298.9</td>
<td>248.5</td>
</tr>
</tbody>
</table>

¹ ΔH° and T° are the enthalpy and the temperature where ΔG = 0 and Kᵤ = Kₙ = 1. ² ΔCₚ is the change in heat capacity upon unfolding. ³ Binding model at 20 (°C).
Figure S5.4: Fitting of CD melting curves (top) for peptide K₃-APgly with E₃GY in the dark adapted (left) and 340 nm irradiated (right) states performed by the Fitdis software program. Residuals (middle) and derivatives (bottom) of the best (dimeric) fitting model are shown. Measurements were performed in PBS.
Table S5.2: Fit results of thermal unfolding curves for K₃ peptides containing azophenylglycine with coiled-coil binding partner E₃GY, in either the dark adapted or 340 nm irradiated state

<table>
<thead>
<tr>
<th>Coiled-coil system</th>
<th>K₃-APgly dark</th>
<th>K₃-APgly 340 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆H°a (kJ mol⁻¹)</td>
<td>285.2 ± 6.5</td>
<td>236 ± 7.1</td>
</tr>
<tr>
<td>T°b (°C)</td>
<td>113 ± 1.2</td>
<td>110.5 ± 1.9</td>
</tr>
<tr>
<td>∆C_p (kJ mol⁻¹ k⁻¹)</td>
<td>2.8 ± 0.13</td>
<td>2.15 ± 0.14</td>
</tr>
<tr>
<td>θ_F (deg cm² dmol⁻¹)</td>
<td>-26877 ± 353</td>
<td>-27513 ± 520</td>
</tr>
<tr>
<td>m_F</td>
<td>112 ± 12.6</td>
<td>119 ± 23.9</td>
</tr>
<tr>
<td>θ_u (deg cm² dmol⁻¹)</td>
<td>0 ± 1153</td>
<td>-1424 ± 687</td>
</tr>
<tr>
<td>m_u</td>
<td>-46.6 ± 13.9</td>
<td>-28 ± 8.7</td>
</tr>
<tr>
<td>∆G_20c (kJ mol⁻¹)</td>
<td>34.5</td>
<td>30.7</td>
</tr>
<tr>
<td>∆S_20c (J mol⁻¹ k⁻¹)</td>
<td>-33</td>
<td>37.9</td>
</tr>
<tr>
<td>∆H_20c (kJ mol⁻¹)</td>
<td>24.8</td>
<td>41.8</td>
</tr>
<tr>
<td>K_f (M⁻¹, 20 °C)</td>
<td>1.38 10⁶</td>
<td>2.97 10⁵</td>
</tr>
<tr>
<td>K_u (µM, 20 °C)</td>
<td>0.72</td>
<td>3.37</td>
</tr>
<tr>
<td>Dark/340nm PSS</td>
<td>4.65</td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td>RMSE</td>
<td>274.3</td>
<td>288.7</td>
</tr>
</tbody>
</table>

a ∆H° and T° are the enthalpy and the temperature where ∆G = 0 and K_u = K_f = 1. b ∆C_p is the change in heat capacity upon unfolding. c Binding model at 20 °C.
Table S5.3: Bond lengths, angles and dihedral angles for new atom types required for the integration of APgly (top left) and APhe1 (top right) into the AMBER-96 force field for molecular dynamics simulations of azobenzene-amino acids. Bond lengths and angles as computed at BLYP-D3(BJ)/TZ2P.

<table>
<thead>
<tr>
<th>Bond type</th>
<th>Residue</th>
<th>Bond length (nm)</th>
<th>Force constant (kJ mol(^{-1}) nm(^{-1}))</th>
<th>Adapted from</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>CA-NX</td>
<td>0.14398</td>
<td>357313.6</td>
<td>CA-NA</td>
</tr>
<tr>
<td>**</td>
<td>NX-NX</td>
<td>0.14354</td>
<td>357313.6</td>
<td>CA-NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Angle type</th>
<th>Residue</th>
<th>Angle (degrees)</th>
<th>Force constant (kJ mol(^{-1}) rad(^{-1}))</th>
<th>Adapted from</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>CA-CT-C</td>
<td>109.139</td>
<td>527.184</td>
<td>CA-CT-CT</td>
</tr>
<tr>
<td>b</td>
<td>CA-CT-N</td>
<td>110.762</td>
<td>527.184</td>
<td>C-CT-N</td>
</tr>
<tr>
<td>c</td>
<td>CA-CT-H1</td>
<td>108.439</td>
<td>418.400</td>
<td>CA-CT-HC</td>
</tr>
<tr>
<td>d</td>
<td>CA-CA-NX</td>
<td>119.225</td>
<td>585.760</td>
<td>NA-CN-NC</td>
</tr>
<tr>
<td>e</td>
<td>CA-CA-NX</td>
<td>121.626</td>
<td>585.760</td>
<td>NA-CN-NC</td>
</tr>
<tr>
<td>f</td>
<td>CA-CA-NX</td>
<td>121.615</td>
<td>585.760</td>
<td>NA-CN-NC</td>
</tr>
<tr>
<td>g</td>
<td>CA-NX-NX</td>
<td>115.088</td>
<td>585.760</td>
<td>NA-CN-NC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dihedral type</th>
<th>Residue</th>
<th>Dihedral angle (degrees)</th>
<th>Force constant (kJ mol(^{-1}) rad(^{-2}))</th>
<th>Adapted from</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA-CA-CA-NX</td>
<td>All</td>
<td>180.0 (improper)</td>
<td>4.60240</td>
<td>CA-CA-CA-NT</td>
</tr>
<tr>
<td>CA-NX-NX-CA</td>
<td>Cis</td>
<td>186.5 (proper)</td>
<td>6.27600(^{+})</td>
<td>X-CA-NA-X</td>
</tr>
<tr>
<td>CA-NX-NX-CA</td>
<td>Trans</td>
<td>180.1 (proper)</td>
<td>6.27600(^{+})</td>
<td>X-CA-NA-X</td>
</tr>
<tr>
<td>CA-CA-NX-NX</td>
<td>Cis</td>
<td>227.4 (proper)</td>
<td>6.27600</td>
<td>X-CA-NA-X</td>
</tr>
<tr>
<td>CA-CA-NX-NX</td>
<td>Trans</td>
<td>180.1 (proper)</td>
<td>6.27600</td>
<td>X-CA-NA-X</td>
</tr>
</tbody>
</table>

\(^{+}\)Set to 1000 kJ mol\(^{-1}\) rad\(^{-2}\) for production runs
Figure S5.5: UV-Vis spectra of peptide K₃-APgly cycling between different photostationary states (left) and relaxing from the 340 nm PSS (right), which were used to prepare Figure 5.5. Spectra after illumination with 340 nm light are indicated with dotted lines (left). Graphs are numbered by time after start of the experiment. Measurements were performed at 20 °C in PBS with [peptide] = 50 µM.

Figure S5.6: Percentage of amino acids classified as helical (based on GROMACS’ default criteria for psi/phi dihedral angles and H-bond distances) over 500 ns of simulation time, for all amino acids in peptide E (left) and peptide K (right), averaged over 3 independent MD simulations. Standard deviations are plotted as dotted lines.
Figure S5.7: Snapshots from MD simulations of the coiled coil between peptide K$_3$ (blue) and E$_3$ (green), with peptide K$_3$ containing the photoswitchable amino acid APhe1 or APgly at position 9 in the peptide sequence, in the trans or cis conformation. Peptides backbones are shown as a cartoon, with the photoswitch and all amino acid side chains of peptide E$_3$ within 5 Å displayed as sticks.
Figure S5.8: Normalized change in distance upon photoisomerization between the non-hydrogen atoms in peptide E₃ and the azobenzene ring 1 (closest to the peptide backbone). Change in distance upon trans to cis isomerization of the azobenzene is normalized to the trans conformation, with positive values indicating more distance in the cis conformation. Distances are averaged over 3 simulations of 500 ns. Dotted lines indicate an arbitrary cut-off for significance of 0.2.
Figure S5.9: Snapshots from MD simulations of the coiled coil between E₃ (green) and K₃ (blue) containing APhe1 or APgly at position 9 in the peptide sequence, in either the cis or trans conformation. Photoswitchable amino acids in peptide K₃ have been coloured orange for clarity. Peptides backbones are shown as a cartoon, with the photoswitch and all amino acid side chains within 5 Å displayed as sticks.
**LC-MS of purified peptides**

**Table S5.4:** Overview of the calculated masses of all peptides used in this project, and the masses found by LCMS.

<table>
<thead>
<tr>
<th>Peptide name</th>
<th>Calculated mass (Da)</th>
<th>Measured mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_3GY$</td>
<td>$[M + 2H^+]^2+$ 1272.69</td>
<td>1271.24</td>
</tr>
<tr>
<td></td>
<td>$[M + H^+]^+$       2544.39</td>
<td>2545.88</td>
</tr>
<tr>
<td>$K_3$-P9-APhe1</td>
<td>$[M + 2H^+]^2+$ 1230.24</td>
<td>1228.81</td>
</tr>
<tr>
<td>$K_3$-P9-APhe2</td>
<td>$[M + 2H^+]^2+$ 1252.24</td>
<td>1250.85</td>
</tr>
<tr>
<td>$K_3$-P9-APgly</td>
<td>$[M + 2H^+]^2+$ 1223.23</td>
<td>1222.23</td>
</tr>
</tbody>
</table>